TECHNISCHE UNIVERSITÄT MÜNCHEN

Institut für Astronomische und Physikalische Geodäsie

Geometrical Theory
of
Satellite Orbits and Gravity Field

Vollständiger Abdruck
der von der Ingenieurfakultät Bau Geo Umwelt
der Technischen Universität München
zur Erlangung des akademischen Grades
eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigten Dissertation.

Drazen Svehla

Vorsitzender: Prof. Dr.-Ing. U. Stilla
Prüfer der Dissertation: 1. Prof. Dr.-Ing., Dr. h.c. R. Rummel (i.R.)
                            2. Prof. Dr.phil.nat. M. Rothacher, ETH Zürich

Abstract

The framework of this thesis consists of the three gravity field missions CHAMP, GRACE and GOCE in LEO orbit, the launch of the first Galileo satellites and the Space-Time Explorer mission (STE-QUEST) in the ESA Cosmic Vision Programme, jointly proposed by the timing community involved in the ACES mission on the International Space Station. The satellite missions CHAMP, GRACE and GOCE equipped with geodetic GPS receivers in the LEO orbit, have initiated a new era of space geodesy and accurate static and temporal gravity field observations from space based on precise orbit determination (POD) using GPS. The Space-Time Explorer mission covers space geodesy and relativistic geodesy as science objectives and aims to combine the terrestrial and celestial reference frame determination and to unify the reference frames for positioning, time and gravity. This thesis presents major results and achievements obtained with these space geodesy missions over the last 15 years. The major part of this thesis covers work done with Prof. M. Rothacher at TU München and ETH Zürich in the context of the LEO Working Group on Precise Orbit Determination of IAG and IGS, ESA mission GOCE, ESA Topical Team on ACES Geodesy and several Working Groups of the IGS. All developments in the Bernese GNSS Software were used for the orbit determination of the GOCE mission (PI Prof. R. Rummel) and the Formosat-3/COSMIC mission.

In all these space geodesy missions, precise orbit determination of satellites and determination of terrestrial reference frame parameters of the Earth represent the fundamental framework of all space geodesy activities. In this thesis, pioneering work has been done on the estimation of purely geometrical (i.e. kinematic) orbits of LEO satellites that has triggered the worldwide development of new approaches in gravity field determination, opened up new fields of application and significantly changed the way we think about the gravity field of the Earth from the point of view of satellite dynamics. This thesis not only presents pioneering work on the high-precision kinematic and reduced-dynamic orbit determination of LEO and GNSS satellites, and the sub-millimeter relative positioning between the two GRACE satellites flying in formation in LEO orbit, but also demonstrates the use of GPS measurements from LEO satellites in the determination of terrestrial reference frame parameters, and provides fundamental studies on the geometrical approach for other space geodesy techniques, such as the sub-millimeter double-difference SLR, Lunar Laser Ranging and their combination with the global GNSS solutions. The use of stable clocks on board Galileo satellites offered an extension of the kinematic POD approach from LEO to GNSS satellites by using Galileo clocks to map kinematically radial orbit errors. This has led to the development of new approaches in the modeling of solar radiation pressure and satellite thermal re-radiation. Several linear combinations were developed for the processing of multi-GNSS data and the integer nature of the ionosphere-free ambiguities is shown by means of the integer ambiguity algebra for the resolution of carrier-phase ambiguities. Several different strategies for the ambiguity resolution are presented including the track-to-track ambiguity resolution demonstrated with GPS data from the GRACE mission in LEO orbit. By introducing so-called "absolute" code biases, a consistent definition of carrier-phase ambiguities has been developed with satellite clock parameters and differential code biases that are estimated without ionosphere information. In the field of satellite orbit dynamics, it was demonstrated that the concept of geometrical rotations of spherical harmonics can be applied to the gravity

field modeling and subsequently to the orbit representation. In addition, geometrical rotations offer a direct representation of the spherical harmonics and their calculation to ultra-high degree and order, considering that a rotation about the polar axis is equivalent to the geometrical rotation of spherical harmonics about an equatorial axis. In this thesis, fundamental work on frequency transfer using GPS has been performed and a new approach consisting of the estimation of so-called phase-clock parameters for GNSS was introduced and tested. This demonstrated the feasibility of one-way frequency transfer between ground and space to support the geodetic applications of optical clocks that now provide relative frequency stability at the level of 10^{-18}. At the end of the thesis, the focus is on relativistic geodesy, related to ACES and STE-QUEST missions, covered by the work done on this thesis over several years. This new field of space geodesy is described, as it is a new field opened up by the capabilities of the new generation of optical atomic clocks.

As part of this thesis, three major developments in the Bernese GNSS Software were performed, including the implementation of kinematic and reduced-dynamic orbit determination of LEO satellites using zero- and double-difference GPS and SLR measurements and the combination with the GPS constellation in the determination of terrestrial reference frame parameters. This work also includes the processing of the GPS baseline with ambiguity resolution between the two GRACE satellites and the combination with GRACE K-band measurements. The second major development is related to the multi-GNSS data processing, in particular the implementation of Galileo and Beidou data processing and the combination with all other GNSS systems. The third major development is the double-difference SLR approach for GNSS with double-differences over time (free of SLR range biases) and the implementation of lunar laser ranging data processing in the barycentric and geocentric frame, including the estimation of the lunar orbit and all reference frame parameters (for GNSS).

As part of this work on the Ph.D. thesis, several conference sessions were organized, including the organization of an ESA conference with more than 100 participants at TU München, in the context of the ESA Topical Team on Geodesy, that triggered several activities described in this thesis. This work in the field of space geodesy was supported by the ESA GOCE mission, several developments of the Bernese GNSS Software and the ESA Topical Team on Geodesy of the ACES mission. This contributed to several ESA missions and mission proposals such as STE-QUEST (reference frames of the Earth), ACES, ASTROD-1, GPS reflectometry/altimetry on the International Space Station – three of them were selected.

Zusammenfassung

Den Rahmen dieser Arbeit bilden die drei Gravitationsfeldmissionen CHAMP, GRACE und GOCE als tieffliegende Satelliten (LEO), der Start der ersten Galileo-Satelliten und die Space-Time Explorer-Mission (STE-QUEST) im ESA Cosmic Vision Program, die gemeinsam mit der Timing-Community, die auch an der ACES-Mission auf der Internationalen Raumstation beteiligt ist, vorgeschlagen wurde. Die Satellitenmissionen CHAMP, GRACE und GOCE, die mit geodätischen GPS-Empfängern ausgestattet sind, haben eine neue Ära der Satellitengeodäsie und der genauen Vermessung des statischen und zeitvariablen Gravitationsfeldes aus dem Weltraum auf der Basis präziser Bahnbestimmung (POD) mit GPS initiiert. Die Space-Time Explorer-Mission deckt die Satellitengeodäsie und die relativistische Geodäsie als wissenschaftliche Ziele ab und zielt darauf ab, die Bestimmung des terrestrischen und zälestischen Referenzrahmens zu kombinieren und die Referenzrahmen für Positionierung, Zeit und Schwerkraft zu vereinheitlichen. Diese Arbeit präsentiert die Ergebnisse, die im Rahmen dieser Doktorarbeit mit den Satellitenmissionen in den letzten 15 Jahren erzielt wurden. Der Großteil dieser Arbeit beschäftigt sich mit der Arbeit an der TU München und der ETH Zürich mit Prof. M. Rothacher im Rahmen der LEO-POD-Arbeitsgruppe der IAG und IGS, der GOCE Mission, des ESA Topical Teams für ACES Geodäsie und mehreren Arbeitsgruppen des IGS. Alle Entwicklungen in der Berner GNSS Software wurden für die Bahnbestimmung der GOCE-Mission (PI Prof. R. Rummel) und der Formosat-3/COSMIC-Mission eingesetzt.

In alle diesen Satellitenmissionen stellt die präzise Bahnbestimmung von Satelliten und die Bestimmung von Parametern des terrestrischen Referenzrahmens der Erde das Fundament aller geodätischen Aktivitäten dar. In dieser Arbeit wurde Pionierarbeit bei der Schätzung von rein geometrischen (d.h. kinematischen) Bahnen von LEO-Satelliten geleistet, die eine weltweite Entwicklung neuer Ansätze in der Gravitationsfeldbestimmung ausgelöst, neue Anwendungsfelder erschlossen und die Art und Weise deutlich verändert haben, wie wir das Gravitationsfeld der Erde aus der Sicht der Satellitendynamik betrachten. Diese Doktorarbeit ist nicht nur eine Arbeit zur hochpräzisen kinematischen und reduziert-dynamischen Bahnbestimmung von LEO- und GNSS-Satelliten und der Sub-Millimeter-Relativpositionierung zwischen den beiden GRACE-Satelliten, die in einer Formation in einer LEO-Umlaufbahn fliegen, sondern zeigt auch die Verwendung von GPS-Messungen von LEO-Satelliten für die Bestimmung von Parametern des terrestrischen Referenzrahmens auf und liefert grundlegende Untersuchungen zu geometrischen Ansätzen im anderen geodätischen Raumverfahren wie dem Bilden von Submillimeter-Doppeldifferenzen bei SLR und beim Lunar Laser Ranging und deren Kombination mit den globalen GNSS-Lösungen. Die Verwendung von stabilen Uhren an Bord der Galileo-Satelliten bot eine Erweiterung des kinematischen POD-Ansatzes für LEO mit GNSS-Satelliten an, welche die Galileo-Uhren nutzt, um radiale Fehler der kinematischen Bahnen zu kartieren, was wiederum zur Entwicklung neuer Ansätze bei der Modellierung des solaren Strahldrucks und der thermischen Rückstrahlung führte. Für die Verarbeitung von Multi-GNSS-Daten wurden mehrere Linearkombinationen entwickelt, und die ganzzahlige Natur der ionosphärenfreien Mehrdeutigkeiten wird mit der ganzzahligen Mehrdeutigkeitsalgebra zur Auflösung von Trägerphasen-Mehrdeutigkeiten aufgezeigt. Mehrere verschiedene Strategien für die Mehrdeutigkeitsauflösung

werden vorgestellt, einschließlich der Mehrdeutigkeitslösung zwischen aufeinanderfolgenden Satelliten-durchgängen, die mit GPS-Daten von der GRACE-Mission in der LEO-Umlaufbahn demonstriert wird. Durch die Einführung sogenannter "absoluter" Code-Biases wurde eine konsistente Definition von Trägerphasen-Mehrdeutigkeiten mit Satelliten-Uhrparametern und Differential-Codebiases entwickelt, die ohne Ionosphären-information geschätzt werden. Auf dem Gebiet der Satelliten-Bahndynamik wurde gezeigt, dass das Konzept der geometrischen Rotationen der sphärisch-harmonischen Flächenfunktionen auf die Gravitationsfeldmodel-lierung und anschließend auf die Bahndarstellung angewendet werden kann. Darüber hinaus bieten geometrische Rotationen eine direkte Darstellung der sphärischen Oberschwingungen und deren Berechnung bis zu höchsten Entwicklungsgraden und -ordnungen, wenn man bedenkt, dass eine Rotation um die Polachse der geometrischen Rotation der sphärischen Harmonischen um eine Äquatorachse entspricht. In dieser Arbeit wurde eine grundlegende Studie zur Frequenzübertragung mit GPS durchgeführt und ein neuer Ansatz, der in der Schätzung der sogenannten Phasenuhrparameter für GNSS besteht, wurde eingeführt und getestet, was die Machbarkeit eines Einweg-Frequenztransfers zwischen Boden und Weltraum zur Unterstützung geo-dätischer Anwendungen von optischen Uhren demonstriert, mit einer relativen Frequenzstabilität im Bereich von 10^{-18}. Der letzte Teil der Arbeit konzentriert sich auf die relativistische Geodäsie, die durch diese Arbeit über mehrere Jahre hinweg wesentlich mitgestaltet wurde, und das neue Anwendungsfeld der Satelliten-geodäsie entstanden ist das durch die neue Generation optischer Atomuhren wird beschrieben.

Im Rahmen dieser Arbeit wurden drei wesentliche Entwicklungen in der Bernese GNSS Software durch-geführt, darunter die Implementierung der kinematischen und reduziert-dynamischen Bahnbestimmung von LEO-Satelliten mittels Null- und Doppeldifferenz-GPS und SLR-Messungen sowie die Kombination mit der GPS-Konstellation für die Bestimmung von terrestrischen Referenzsystemparametern. Diese Arbeit beinhaltet auch die Verarbeitung der GPS-Basislinie zwischen den beiden GRACE-Satelliten inclusive Mehrdeu-tigkeitsauflösung und die Kombination mit GRACE K-Band-Messungen. Die zweite große Entwicklung bezieht sich auf die Multi-GNSS-Datenverarbeitung, insbesondere die Implementierung der Galileo- und Beidou-Datenverarbeitung und die Kombination mit allen anderen GNSS-Systemen. Die dritte wesentliche Entwick-lung betrifft den Doppeldifferenz-SLR-Ansatz und die Implementierung der Lunar-Laser-Datenverarbeitung im barizentrischen und geozentrischen Bezugsrahmen einschließlich der Schätzung der Mondbahn und aller Parameter des Referenzrahmens.

Als Teil dieser Arbeit wurden mehrere Konferenzsession organisiert, darunter die Organisation einer ESA-Konferenz mit mehr als 100 Teilnehmern an der TU München im Rahmen des ESA Topical Team on Geodesy, das mehrere in dieser Arbeit beschriebene Aktivitäten auslöste. Diese Arbeit auf dem Gebiet der Weltraumge-odäsie wurde von der ESA GOCE Mission, dem ESA Topical Team für Geodäsie mit der ACES Mission und mehreren Entwicklungen der Bernese GNSS Software unterstützt. Dies trug zu mehreren ESA-Missionen und Missionsvorschlägen wie STE-QUEST (Bezugsrahmen der Erde), ACES, ASTROD-1, GPS-Reflektometry/Altimetrie auf der Internationalen Raumstation bei - drei davon wurden ausgewählt.

Acknowledgement

My deep gratitude goes to Prof. M. Rothacher for his support of our 90 conference talks given all over the world and at various institutions, and his kind review of this comprehensive thesis improving the style and the content. This includes endless discussions on the Bernese GNSS Software and his support at ETH Zurich. There are no words to express gratitude for all his time he invested, and to properly acknowledge it. Special thanks to Prof. R. Rummel for my involvement in the ESA GOCE mission for many years and support via the DAAD research scholarship at TU München. I would like to acknowledge his reading of this thesis and providing valuable remarks and advices. To colleagues in ESA, L. Cacciapuoti for the long cooperation on the ACES and the STE-QUEST missions and the joint work on the ESA Topical Team on Geodesy, and Rune Floberghagen and Roger Haagmans for their support and the time we together spent on the GOCE mission. Last, but not least, to Prof. H. Moritz, who showed the way, making all this possible. In the end, to Rory Bridson for the English proofreading of the thesis and to all friends at TU München and friends in the international community of expats around the European Patent Office in Munich.

Table of Contents

1. The First Geometric POD of LEO Satellites – a Piece of History

T he very first precise geometric (i.e., kinematic) orbit determination of a LEO satellite was presented in (Švehla and Rothacher 2002), where for the first time double-difference ambiguity resolution was demonstrated using the CHAMP satellite in LEO orbit and the ground IGS network. In (Švehla and Rothacher 2003a), (Švehla and Rothacher 2003b) and later in (Švehla and Rothacher 2005a) and (Švehla and Rothacher 2005b) geometric precise orbit determination (POD) was demonstrated to cm-level accuracy and presented as an established technique and as very attractive for gravity field determination. Here we give a chronological overview of the development of the method.

1.1 Introduction

In (Švehla and Rothacher 2005a) and (Švehla and Rothacher 2005b), kinematic (or geometric) precise orbit determination of Low Earth Orbiting (LEO) satellites was introduced as a new method of precise orbit determination of LEO satellites where the main application is in gravity field determination. The first geometric orbits of the CHAMP satellite were presented in (Švehla and Rothacher 2002). Later, in (Švehla and Rothacher 2004a) kinematic and reduced-dynamic POD were shown for a period of two years using CHAMP data. Kinematic or geometric POD can be considered as the third fundamental POD approach, along with dynamic and reduced-dynamic POD:

- Dynamic POD: (Kaula 1966), (Beutler 1977)
- Reduced-Dynamic POD: (Colombo 1986), (Yunck et al. 1994)
- Geometric POD: (Švehla and Rothacher 2003b)

An intermediate, or fourth basic approach to POD, is the reduced-kinematic POD, where the orbit kinematics (geometry) is reduced to a dynamic orbit by estimating normal kinematic points along an a priori dynamic orbit and making use of relative constraints between kinematic positions (Švehla and Rothacher 2003b). However, we applied the reduced-kinematic POD approach only for GPS and Galileo satellites. In the reduced-dynamic POD approach, the orbit dynamics is reduced by making use of geometrical information, i.e., estimating velocity pulses along a dynamic orbit (Yunck et al. 1993) or estimating empirical accelerations (Colombo 1986). The dynamic POD approach is based on numerical integration of the equation of motion, see e.g., (Beutler 1977). This numerical integration can be avoided in certain applications, and in the case of analytical POD the equation of motion is modeled as an analytical representation, see e.g., (Kaula 1966).

A considerable number of groups have been using our CHAMP kinematic positions to estimate Earth gravity field coefficients and to validate dynamic orbits and orbit models. Using the CHAMP kinematic positions together with the corresponding variance–covariance information, gravity field coefficients can be estimated geometrically by making use of the energy balance approach or the boundary value method rather than the classical numerical integration schemes, see e.g., (Gerlach et al. 2003a, 2003b), (Wermuth et al. 2004),

(Földváry et al. 2005) at TU München, (Mayer-Gürr et al. 2005) at TU Bonn now at TU Graz, (Reubelt et al. 2006), (Sneeuw et al. 2003, 2005) at TU Stuttgart and (Ditmar et al. 2006) at TU Delft, and (Fengler et al. 2004) from Prof. Freeden's Group at TU Kaiserslautern and (Schmidt et al. 2005) from DGFI (Deutsches Geodätisches Forschungsinstitut). With the GRACE and GOCE missions, kinematic orbits continued to be used world-wide and a number of groups have been reporting gravity field recovery based on GRACE and GOCE kinematic orbits, e.g., (Mayer-Gürr et al. 2010), (Jäggi et al. 2011), (Pail et al. 2010), (Pail et al. 2011), (Baur and Grafarend 2006), (Rummel et al. 2011). The validation of gravity field models computed in such a way showed that LEO kinematic positions contain high-resolution gravity field information. In combination with gravity gradients from the GOCE gradiometer in very low Earth orbit ($255\,\mathrm{km}$ altitude), kinematic orbits allow mapping of the gravity field of the Earth from space with the highest resolution reported so far. Kinematic positions with the corresponding variance–covariance information are a very attractive interface between the raw GPS data and gravity field models or other valuable information that can be derived from satellite orbits, e.g., air densities, thermospheric winds or orbit force model improvements. In this way, the groups that use kinematic positions do not have to undertake the laborious tasks of processing and analyzing the GPS observations and determining the reference frame.

In regard to kinematic POD for ESA mission GOCE we refer to (Bock et al. 2011), (Visser et al. 2007, 2009) (Bock et al. 2014). Several other groups reported calculation of kinematic orbits for gravity field determination e.g., (Zehentner and Mayer-Gürr 2015) for the GRACE mission and (Hwang et al. 2009, 2010), (Tseng et al 2012) using similar approach for the Formosat-3/COSMIC mission. (Zehentner and Mayer-Gürr 2015) demonstrated an approach that avoids ionosphere-free linear combination by estimating an additional bias per GPS satellite every epoch in order to remove remaining systematic effects in carrier-phase measurements. Among the aforementioned geometric gravity models, (Baur et al. 2013) identifies and compares 5 fundamental approaches in gravity field determination based on kinematic orbits:

- Short-Arc Approach: TU Graz
- Celestial Mechanics Approach: AIUB/University of Bern
- Averaged Acceleration Approach: DEOS/TU Delft
- Point-wise Acceleration Approach: University of Stuttgart/Austrian Acad. of Sciences
- Energy Balance Approach: TU München, TU Graz.

Time-variable gravity field determination using a CHAMP kinematic orbit was recently demonstrated in (Baur 2013), showing that the ice mass loss over Greenland is in line with the findings from GRACE data and the trend estimates differ by only 10%. This opens up the possibility of using kinematic orbits to bridge the gap between GRACE and the GRACE follow-on mission, making use of the GPS receivers on the 3 satellites of the ESA mission Swarm for mapping the time-variable gravity field of the Earth.

1.2 Geometric and Dynamic Equation of Motion

The theory of relativity is the frame of reference for satellite orbit determination and includes corrections to the Newtonian equations of motion, so-called post-Newtonian approximation of general relativity, often denoted as Parameterized Post-Newtonian Formalism or PPN-formalism. For the near-Earth orbiting satellites, the geocentric reference frame is used, whereas for planetary missions in the Solar System, a barycentric reference frame is more appropriate. A geocentric reference frame is more suitable for the orbit determination of Earth-orbiting satellites because the gravitational effects of the Moon, the Sun and other planets can be described solely as tidal forces, while the relativistic acceleration corrections to the equations of motion are very small. For satellites in Earth orbit, in the post-Newtonian approximation of general relativity, the main general relativistic effects are caused by the gravity field of the Earth and its rotation. The flat three-dimensional Euclidian space is used to model geometry and to dynamically integrate the satellite orbit. A geocentric terrestrial reference frame is created using space geodesy techniques such as GNSS, SLR, VLBI and DORIS.

The latest versions of this reference frame, e.g., ITRF2005 or ITRF2008 use terrestrial time defined on the geoid as the reference.

The geometric equation of kinematic motion of a satellite can be defined as

$$\vec{r}_{\oplus}^{satellite} := \vec{r}_{\oplus}^{frame} + \Delta\vec{r}_{frame}^{satellite} \tag{1.1}$$

where $\vec{r}_{\oplus}^{frame}$ defines the reference frame (coordinate system defined by station coordinates, GNSS satellite orbits and clock parameters, etc.) and $\Delta\vec{r}_{frame}^{satellite}$ is the relative geometric vector of the satellite $\vec{r}_{\oplus}^{satellite}$ w.r.t. to that reference frame, i.e., the vector between a GNSS and a LEO satellite, or a vector between a GNSS satellite and a ground station. Since the geometric equation of motion (1.1) does not include the dynamics of the satellite, it is kinematic in its nature. This is why geometric orbits of satellites are also often called kinematic orbits.

The dynamic equation of motion can be written as

$$\ddot{\vec{r}}_{\oplus}^{satellite} = \nabla V_{\oplus}^{satellite} + \ddot{\vec{r}}_{temp} + \ddot{\vec{r}}_{relativity} + \ddot{\vec{r}}_{non-gravitational} \tag{1.2}$$

where $\nabla V_{\oplus}^{satellite}$ is the gravitational acceleration, what one could call the dynamic reference frame, $\ddot{\vec{r}}_{temp}$ denotes temporal variations of the gravitational field (tides, etc.) and $\ddot{\vec{r}}_{relativity}$ represents relativistic corrections, what one could call the relativistic frame. The last term in (1.2) denotes non-gravitational contributions to the equation of motion, such as solar radiation, Earth albedo, aerodynamic drag, etc.

1.3 LEO GPS Observation Equation

The observation equation for LEO zero-difference POD using carrier-phase measurements for the frequency i between a LEO receiver and a GPS satellite s can be written as follows (in units of length)

$$\begin{aligned} L_{LEO,i}^s = \rho_{LEO}^s &+ c(\delta t_{LEO} + \delta t_{sys,i}) - c(\delta t^s + \delta t^{sys,i}) + \\ &+ \delta\rho_{ion,i} + \delta\rho_{rel} + \delta\rho_{mul,i} + \delta\rho_{pco,i} + \delta\rho_{pcv,i} + \\ &+ \lambda_i \cdot N_{LEO,i}^s + \varepsilon_i \end{aligned} \tag{1.3}$$

$L_{LEO,i}^s$ LEO zero-difference phase measurement,

ρ_{LEO}^s geometric distance,

c speed of light in vacuum $c = 299792458$ m/s ,

$\delta t_{LEO}, \delta t^s$ LEO and GPS satellite clock corrections,

$\delta t_{sys,i}, \delta t^{sys,i}$ LEO and GPS satellite system delays (cable, electronics, etc.),

$\delta\rho_{ion,i}$ ionospheric delay,

$\delta\rho_{rel}$ periodic relativistic correction and Shapiro correction,

$\delta\rho_{mul,i}$ multipath, scattering, bending effects,

$\delta\rho_{pco,i}$ LEO phase center offset,

$\delta\rho_{pcv,i}$ LEO phase center variations,

λ_i wavelength of the GPS signal (L_1 or L_2),

$N_{LEO,i}^s$ zero-difference phase ambiguity,

ε_i phase noise (L_1 or L_2)

For more on this subject see (Švehla and Rothacher 2005a). One can immediately recognize the well-known observation equation for a ground station, with one exception: in the LEO case there is no tropospheric delay to be taken into account. In order to eliminate ionospheric delays, the ionosphere-free L_3 linear combination (LC) can be formed between the LEO phase measurements $L_{LEO,1}^s$ and $L_{LEO,2}^s$ on carrier frequencies f_1 and f_2, respectively

$$L_{LEO,3}^s = \frac{f_1^2}{f_1^2 - f_2^2} L_{LEO,1}^s - \frac{f_2^2}{f_1^2 - f_2^2} L_{LEO,2}^s \tag{1.4}$$

In this case the LEO zero-difference observation equation can be written as follows

$$\begin{aligned} L_{LEO,3}^s = {} & \rho_{LEO}^s + c \cdot \delta t_{LEO,clk,3} - c \cdot \delta t^{s,clk,3} + \delta\rho_{rel} + \delta\rho_{mul,3} + \delta\rho_{pco,3} + \delta\rho_{pcv,3} + \\ & + B_{LEO,3}^s + \varepsilon_3 \end{aligned} \tag{1.5}$$

where $\delta t_{LEO,clk,3}$ denotes the ionosphere-free LEO clock parameter consisting of the real LEO clock value δt_{LEO} and the system delays $\delta t_{sys,1}$ and $\delta t_{sys,2}$ on both frequencies:

$$\delta t_{LEO,clk,3} = \delta t_{LEO} + \frac{f_1^2}{f_1^2 - f_2^2} \delta t_{sys,1} - \frac{f_2^2}{f_1^2 - f_2^2} \delta t_{sys,2} \tag{1.6}$$

In the same way the ionosphere-free GPS clock parameter can be defined as

$$\delta t^{s,CLK,3} = \delta t^s + \frac{f_1^2}{f_1^2 - f_2^2} \delta t^{sys,1} - \frac{f_2^2}{f_1^2 - f_2^2} \delta t^{sys,2} \tag{1.7}$$

$\delta\rho_{mul,3}$, $\delta\rho_{pco,3}$ and $\delta\rho_{pcv,3}$ denote multipath effects, phase center offset and phase center variations for the ionosphere-free linear combination, respectively. The zero-difference ionosphere-free ambiguity (phase bias) is denoted by $B_{LEO,3}^s$, for more details we refer to (Švehla and Rothacher 2005a).

The ionosphere-free observation equation for the LEO zero-difference code measurements can be written in the same way except that the LEO phase ambiguity parameter $B_{LEO,3}^s$ is not included and the first order ionosphere effect has an opposite sign. GPS satellite and LEO system delays are different for P_1 and P_2 code measurements. By convention, the ionosphere-free LC is said to have no Differential Code Bias (DCB), i.e., system delays are included in the GPS satellite and the receiver clocks, respectively, see (Schaer 1999).

The observation equation for POD based on double differences can be written by forming double-differences between the LEO and a ground station and between GPS satellites k and s:

$$L_{grd,LEO,3}^{s,k} = (L_{LEO,3}^k - L_{grd,3}^k) - (L_{LEO,3}^s - L_{grd,3}^s) \tag{1.8}$$

In this way we can form baselines between all ground IGS stations and the LEO satellite. It is very important to note that, by using double-differences between LEO and ground station, the absolute tropospheric delay for the ground station can be estimated and isolated.

As soon as we involve the GPS ground network (e.g., the IGS network), the troposphere zenith delays and station coordinates have to be considered. In our POD approach, weekly IGS solutions for station coordinates, computed at the CODE Analysis Center, and corresponding troposphere zenith delays and troposphere gradients are introduced and kept fixed. (For more about IGS products see (Dow et al. 2005).) For the CODE IGS products we refer to the ftp site ftp://ftp.unibe.ch/aiub/. In order to have full consistency between IGS products and the software used, we used GPS satellite orbits, ground station coordinates and troposphere parameters from the IGS Reprocessing Project (Steigenberger et al. 2006) run at TU München. With regard to the IGS and the quality of the IGS products, we refer to (Hugentobler 2012) and to the GGOS

Coventions to (Hugentobler et al. 2012). For more information on the GGOS Project of IAG (Global Geodetic Observing System) and the combination of space geodesy techniques in the generation of the terrestrial reference frame of the Earth, we refer to (Rothacher et al. 2002), (Rothacher et al. 2004) and (Rummel et al. 2000). For the latest generations of the international terrestrial reference frame we refer to e.g., (Altamimi et al. 2011).

1.4 Zero-, Double- and Triple-Difference POD Approaches

In the field of kinematic POD with spaceborne GPS receivers, three main approaches can be distinguished from the point of view of differencing: zero-difference (ZD), double-difference (DD) and triple-difference (TD), (see Figure 1.1), for more on this subject see (Švehla and Rothacher 2002), (Švehla and Rothacher 2005a).

The ZD approach, in contrast to the DD and TD approaches, only relies on the GPS observations of the LEO and avoids the use of the ground IGS network. This is, at the same time the weakness of this solution, because high-rate satellite GPS clocks are a prerequisite for this method of determining the position of the spaceborne GPS receiver. One has to use the data from the IGS network to estimate a very high number of GPS clock parameters first, and then use these to compute a kinematic orbit, which means that errors in the GPS satellite clocks propagate directly into the LEO orbit positions. A high level of correlation exists between clock parameters, zero-difference ambiguities and epoch-wise satellite positions. Thus we can say that the quality of ZD kinematic orbit determination greatly depends on the accuracy of GPS orbit data, which is itself strongly correlated to that of the GPS satellite clocks.

A very efficient alternative zero-difference approach, followed at the Astronomical Institute, University of Berne, avoids setting up zero-difference ambiguity parameters by forming differences between phase observations of consecutive epochs, (see (Bock et al. 2003)).

By forming double-differences, i.e., baselines between the stations of the IGS network and the LEO, all GPS/LEO satellite clock parameters can be eliminated. The present accuracy of the GPS orbits provided by the IGS, which is in the range of $1 - 2 \, \text{cm}$, is sufficient, according to the rule of thumb given by (Bauersima 1983), for there to be no significant impact on the double-difference solutions. The great advantage of the DD approach is the possibility of fixing ambiguities to integer values and thus of improving the accuracy of LEO POD.

By forming triple differences (differences of DD in time), ambiguities are eliminated, thus allowing very efficient processing algorithms to be employed. The drawback of this approach is the increase of the observation noise and the need for efficient methods to correctly account for the correlations between epochs.

All three aforementioned approaches make direct or indirect use of the IGS network. In the ZD case a global solution is needed to supply GPS satellite orbit and clock information for the subsequent kinematic POD using ZD. Similarly, in the DD and TD cases a global solution can be used to obtain highly accurate IGS site coordinates, the corresponding troposphere zenith delays and GPS satellite orbits. All of these parameters can be held fixed in both DD and TD POD. From the point of view of accuracy, IGS thus plays a major role in providing the framework for LEO POD by GPS.

In all three cases the effect of the ionosphere can be eliminated by forming the ionosphere-free linear combination. The remaining effect of multipath can be reduced to a great extent by elevation-dependent weighting of the GPS observations of the ground network as well as the spaceborne GPS receiver. Last but not least, the LEO antenna phase center position (offset and phase center variations) has to be exactly known in the satellite-fixed system and we need an accurate model of the attitude of the spacecraft (e.g., from quaternions provided by star trackers - with or without combining them with onboard accelerometer data).
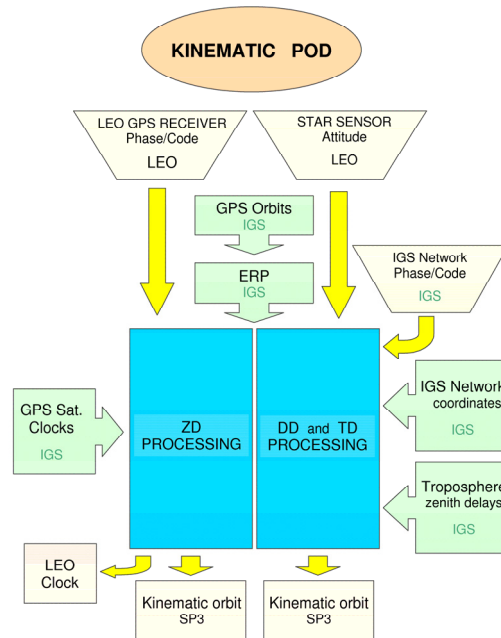
Figure 1.1 Zero- and double- difference approach in kinematic POD, (Švehla and Rothacher 2002).

1.5 Zero-Difference Approach

In the zero-difference kinematic POD for each epoch three LEO coordinates have to be estimated together with one LEO clock parameter. Zero-difference ambiguities are the only parameters in the adjustment procedure that are not epoch-specific. Figure 1.2 shows the normal equation matrix for zero-difference kinematic POD over eleven epochs. On the main diagonal we can easily recognize 3×3 blocks of epoch-wise kinematic LEO coordinates, 11 epoch-wise LEO clock parameters and, in the lower right corner, 6 zero-difference ionosphere-free ambiguities. We easily see the correlations between zero-difference ambiguities and epoch-wise parameters. All zero-difference approaches rely on the availability of highly accurate GPS satellite orbits and
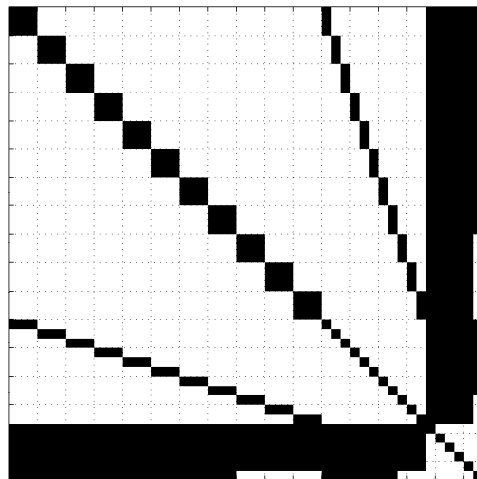
Figure 1.2 Normal equation matrix for zero-difference kinematic POD (11 epochs only). On the main diagonal: 3 by 3 blocks of epoch-wise kinematic coordinates, 11 epoch-wise LEO clock parameters, and in the lower right corner, 6 zero-difference ambiguity parameters, (Švehla and Rothacher 2005a).

clocks. They should be provided with the same sampling as used for the LEO kinematic POD. For the highest accuracy, GPS satellite clocks can be linearly interpolated only for sampling below 30 s . Linear interpolation of 5 min GPS satellite clocks is not recommended for high-precision applications. It is very important that GPS satellite orbits and clocks are consistent with each other because of the high correlations. If highly accurate GPS satellite orbits and clocks are available, this method is very simple and reliable because it does not involve the immense task of processing the ground IGS network. More about the zero-difference approach can be found in e.g., (Švehla and Rothacher 2002) or (Švehla and Rothacher 2005a). An alternative zero-difference approach based on forming differences between phase observations of consecutive epochs and avoiding zero-difference ambiguity parameters, may be found in (Bock et al. 2003).

The normal equations in the least-squares adjustment can be written in the form

$$A^t P A x = A^t P l \tag{1.9}$$

with the design matrix A , containing in our case partial derivatives of the observation equation (1.5), the weight matrix P of the observations, the vector of the unknown parameters x and the vector l containing the so-called observed-minus-computed values. If we denote in (1.9) the normal equation matrix as $N = A^t P A$ and $b = A^t P l$, the normal equations (1.9) can be written as

$$N x = b \tag{1.10}$$

The normal equation matrix for the kinematic POD can be considered as a block diagonal, see also Figure 1.3, thus we can separate ambiguities x_1 from epoch-wise parameters x_2

$$\begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{1.11}$$

where ambiguities are estimated first

$$\left(N_{11} - N_{12} N_{22}^{-1} N_{21} \right) x_1 = b_1 - N_{12} N_{22}^{-1} b_2 \tag{1.12}$$

and epoch-wise parameters are determined by a re-substitution of the estimated ambiguity parameters

$$x_2 = N_{22}^{-1} \left(b_2 - N_{21} \right) x_1 \tag{1.13}$$

In order to derive the variance-covariance matrix Q_{xx} of the estimated epoch-wise parameters we start with
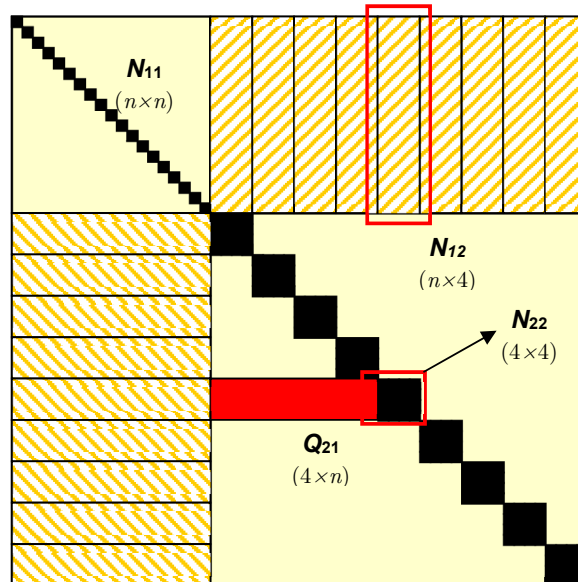
Figure 1.3 An elegant algorithm to calculate variance-covariances between several epochs.

$$Q_{xx}N = I \tag{1.14}$$

and obtain the variance-covariance matrix for the single epoch

$$\underset{(4\times4)}{Q_{22}} = \underset{(4\times4)}{N_{22}^{-1}} + N_{22}^{-1}N_{21} \underset{(n\times n)}{Q_{11}} N_{12}N_{22}^{-1} \tag{1.15}$$

Considering (1.14) and making use of the Shur-Frobenius relations for block-matrices, we derive the variance-covariance matrix of epoch-wise kinematic parameters $\underset{(4\times n)}{Q_{21}}$ over several epochs n

$$Q_{12} = -Q_{11}N_{12}N_{22}^{-1} \tag{1.16}$$

$$\underset{(4\times n)}{Q_{21}} = -\underset{(4\times4)}{N_{22}^{-1}} N_{12}N_{22}^{-1} \tag{1.17}$$

Figure 1.3 graphically shows the matrix $Q_{21(4\times n)}$ containing variance-covariance information of n kinematic epochs as used for the GOCE mission, where the $Q_{21(4\times n)}$ matrix is provided as an official product of the GOCE mission, accompanying the GOCE kinematic orbit positions. Figure 1.4 shows the first kinematic orbits of the CHAMP and GRACE satellites with a sampling of 30 s over one day against the reduced-dynamic orbit. One can see that the kinematic positions in the radial direction are more affected by noise than those in the along-track and cross-track directions, and the along-track differences show a clear once-per-rev. pattern. The variations of the kinematic positions are in the order of $1-2$ cm .

Figure 1.5 shows typical correlations of LEO kinematic positions, with correlation length ≈ 22 min and Figure 1.6 shows the first continuous CHAMP kinematic orbit with cm-level accuracy. The reduced-dynamic orbit model used in our approach for LEO satellites is based on the dynamic model originally developed at
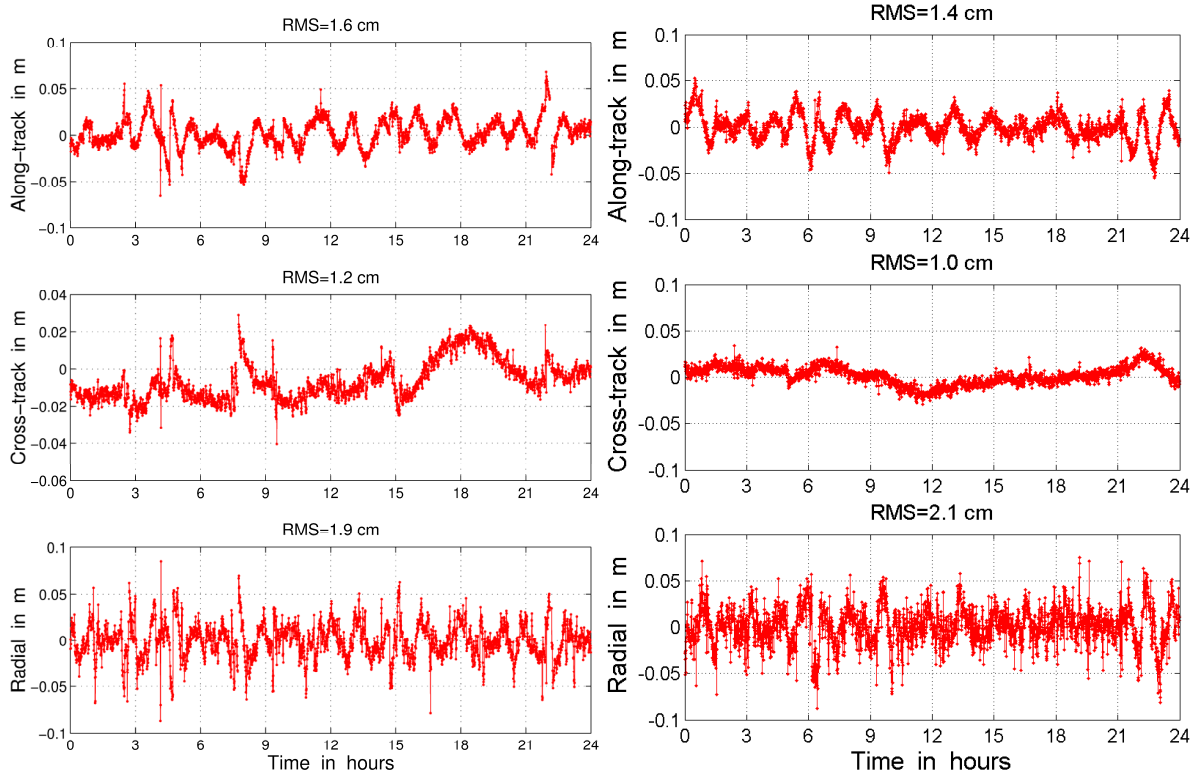
Figure 1.4. First kinematic orbit of CHAMP, day 200/2002 (left) and GRACE-A, day 200/2003 (right) against the reduced-dynamic orbit.
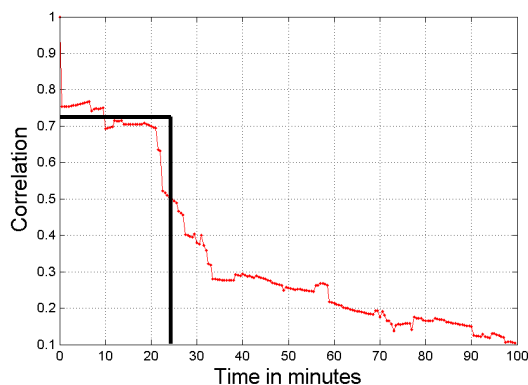
Figure 1.5. Typical correlations of CHAMP kinematic positions indicating significant white noise of the epoch-wise kinematic positions. Correlation length of approx. 22 min is similar in size to the typical observation time of carrier-phase ambiguities.

the CODE IGS Analysis Center for GPS satellites and here subsequently adapted for use in LEO satellite POD (Švehla and Rothacher 2002). The adaptation of this software involved, among other things, the development of an independent orbit modeling chain in the Bernese software including kinematic and reduced-dynamic orbit parameterization, and pre-processing of the data. First results with GPS measurements from the CHAMP satellite showed that frequent estimation of pseudo-stochastic pulses (small velocity changes) is a very efficient approach to modeling the orbit dynamics of a satellite at low orbit altitude. For the orbits of the CHAMP, GRACE and GOCE satellites, pseudo-stochastic pulses are set-up every 6 min in the numerical integration. Later (Jäggi et al. 2006) introduced the estimation of pseudo-stochastic accelerations estimated as constant parameters. However, comparing the accuracy of kinematic and reduced-dynamic orbits for GOCE (Bock et al. 2007, 2011), (Visser et al. 2007, 2009) with the CHAMP and GRACE results in Figure 1.4, one can see that over the last 10 years the LEO orbit accuracy has not been significantly improved. Figure 1.7 shows the daily RMS of GRACE kinematic orbits estimated for the first 4 months of GPS data provided to the GRACE Science Team.
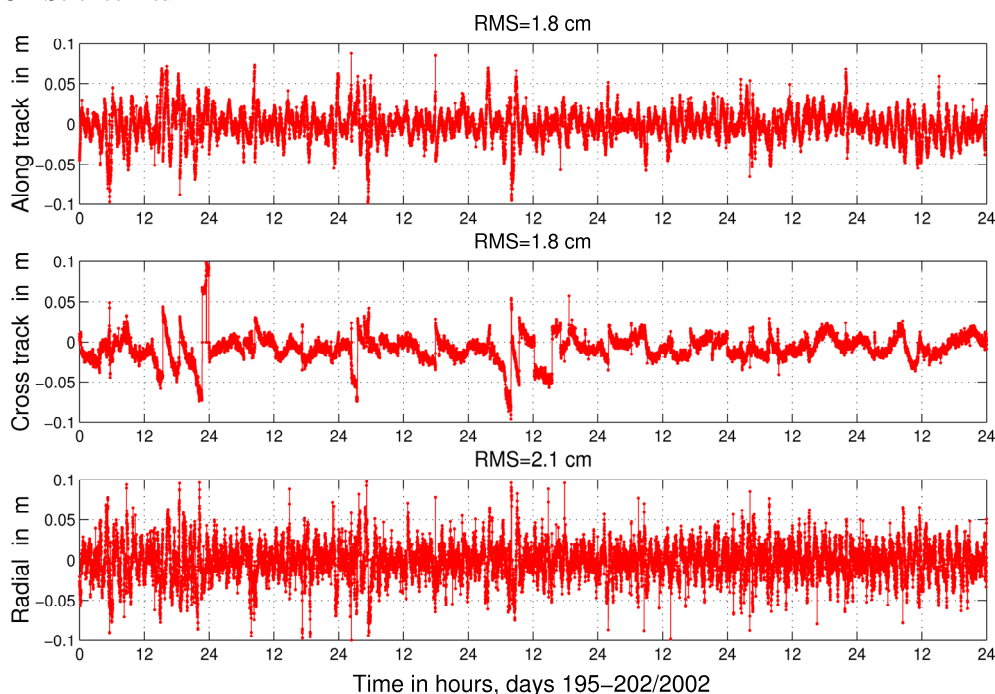
Figure 1.6. Differences between CHAMP kinematic and red.-dynamic orbit, week 1175/2002.
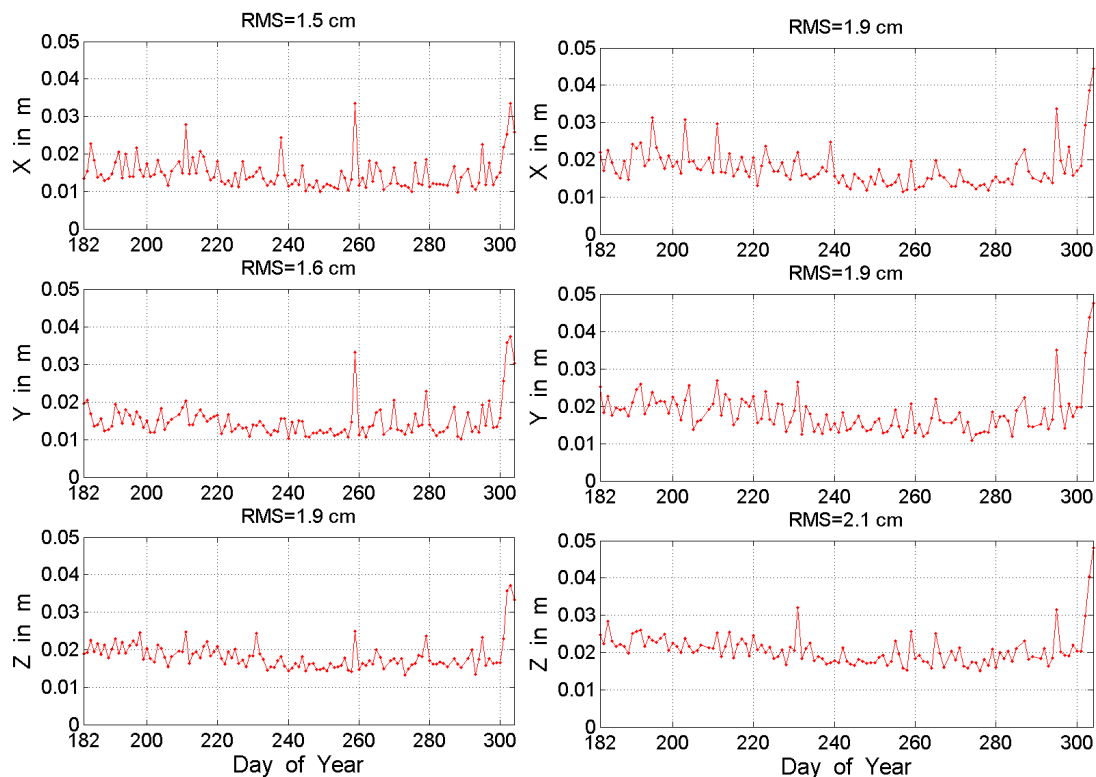
Figure 1.7. Daily RMS of kinematic orbits of the GRACE-A and GRACE-B satellite vs. the reduced-dynamic orbit for the first 4 months (days 182-303/2003) of GPS data provided to the GRACE Science Team.

1.5.1 GPS Receiver Clock and Kinematic POD

In the case of the CHAMP satellite, the estimated clock corrections of the internal GPS receiver clock used to time-tag carrier-phase and code measurements w.r.t. GPS time are in the order of $0.1\,\mu$s. In the case of the GRACE mission, GPS measurements can be synchronized to GPS time very accurately in post-processing due to the onboard K-band ranging system and this synchronization is carried out at the level of the accuracy of the P-code measurements. In the case of the CHAMP or GRACE missions, the estimated GPS receiver clock corrections are very small and do not create any problems concerning the interpolation of GPS clocks and the computation of the correct distance between GPS and LEO satellites. If we consider the LEO orbit velocity to be below $10\,$km/s, including the perigee velocity of a satellite in a highly elliptic orbit, a synchronization error of GPS measurements in the order of $0.1\,\mu$s will lead to an error of $1\,$mm. Therefore, any double-differences between a LEO satellite and the ground network can easily be formed and this synchronization error can easily be taken into account when forming zero- or double-differences. However, if the onboard navigation solution is not used for the steering of the GPS receiver clock, as is the case with the GOCE mission, the internal GPS receiver clock will slowly drift w.r.t. GPS time and the GPS measurements will be taken anywhere in the integer second interval, since the sampling interval of GOCE GPS measurements is $1\,$s. In this case, the orbit changes significantly from the nominal integer second position and this needs to be properly accounted for. If GPS measurements are taken anywhere between the integer seconds of receiver time, it is very difficult to form double- or triple-differences with the ground IGS network, since clock steering is used for all GNSS receivers in the IGS network. For more, see (Švehla and Rothacher 2002) and (Švehla and Rothacher 2003a). The BlackJack GPS receiver and derivatives of this device onboard several LEO missions use the calculated clock offset from the navigation solution to adjust the onboard GPS receiver clock to

GPS time. The BlackJack receiver clock is based on a voltage-controlled quartz oscillator and the frequency of oscillation is controlled, so that the drift is nearly zero. Navigation time solutions are used for clock steering only when at least 5 satellites are being tracked and a valid navigation solution can be calculated. Therefore, the GPS receiver clock drifts away from GPS time only during epochs without a valid navigation solution. The receiver also generates a 0.1 PPS timing pulse on both timing ports. This timing pulse is coincident with the receiver clock 10 second epoch and is used to provide a time source for the spacecraft and scientific instruments. Note that the BlackJack GPS receiver operates without knowledge of the Anti-Spoofing (AS) encryption code. More about the ICESat BlackJack receiver can be found in, e.g., (Williams et al. 2002).

1.5.2 Validation of Kinematic Positions with SLR

That we are not just talking about orbit consistency, but also orbit accuracy, can be seen in Figure 1.8, where SLR residuals are shown for the same kinematic and reduced-dynamic orbits as displayed in Figure 1.6. SLR residuals were calculated as the difference between the SLR measurements (corrected for signal propagation effects) minus the distance between the SLR station and the GPS-derived orbit position. For the validation of dynamic orbits, LEO positions were calculated directly from the dynamic orbit represented by the high order polynomial in the integration step. The offset between CHAMP center of mass and SLR retro-reflector was applied using the attitude provided in the form of quaternions. In the case of kinematic orbits, the only difference is that kinematic positions are given with a sampling of 30 s and an interpolation procedure is required in order to obtain positions at the epochs of the SLR normal points. A linear interpolation was used to obtain kinematic positions along an a priori dynamic orbit. We noticed that the SLR validation of kinematic orbits is more difficult and the necessary interpolation may easily increase the RMS. Another alternative would be to form SLR normal points exactly at the epochs where kinematic positions are defined, but in this case raw SLR data would have to be processed, and these are not readily available from all SLR stations.
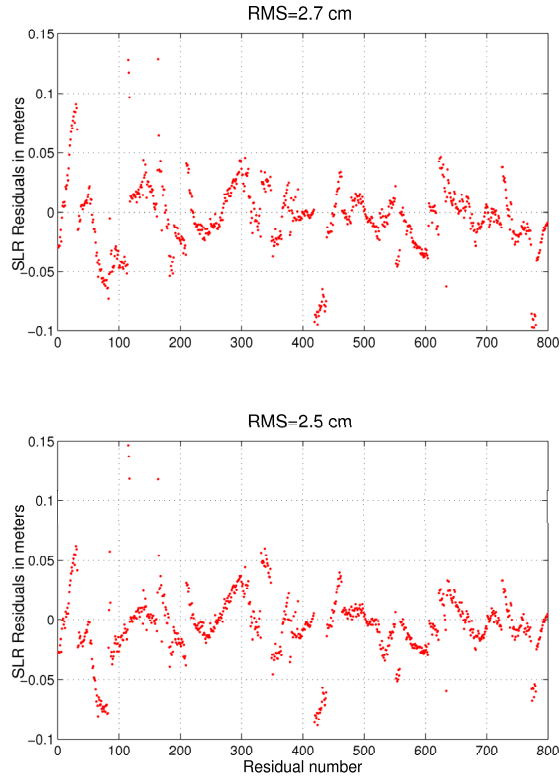
Figure 1.8 SLR residuals for CHAMP kinematic (top) and reduced-dynamic orbits (bottom) for GPS week 1175/2002 (days 195-201/2002). All SLR residuals were used in the analysis; elevation cut-off 10°.

Tropospheric delays for SLR measurements were modeled using the Marini-Murray model and standard corrections like ocean loading (GOT00.2), Shapiro relativistic effect and station velocities were applied. All SLR stations and SLR measurements were used in this validation (elevation cut-off $10°$). The RMS of the CHAMP kinematic and reduced-dynamic orbits is about 2.5 cm (days 195-201/2002). It is interesting to note that the SLR residuals show a similar pattern for both kinematic and reduced-dynamic orbits and that no significant bias can be identified in the SLR residuals. Table 4.3 summarizes the daily RMS of the SLR residuals for our CHAMP orbits based on four different POD approaches, namely kinematic and reduced-dynamic orbits based on zero- and double-differences. One can see that CHAMP orbits are of similar quality for both a purely kinematic and a reduced-dynamic approach. This also holds for CHAMP orbits calculated using either zero- or double-differences. Slightly better orbit quality, i.e., 2.56 cm is obtained with kinematic orbits based on double-differences.

Day	Zero-diff. Dynamic	Zero-diff. Kinematic	Double-diff. Dynamic	Double-diff. Kinematic
195	4.02	4.17	3.22	2.66
196	2.90	2.93	3.19	3.03
197	3.40	3.11	3.29	2.90
198	2.07	2.07	1.99	1.34
199	1.94	1.66	1.91	1.70
200	1.43	1.45	1.69	1.83
201	3.59	4.65	4.32	5.00
202	2.03	2.08	1.93	2.05
Mean	2.67	2.77	2.69	**2.56**

Table 1.1 Daily RMS of SLR residuals in cm for CHAMP kinematic and reduced-dynamic orbits based on zero- and double-differences (days 195-202/2002).

1.6 Double-Difference Approach

In comparison to the zero-difference kinematic POD approach, the double-difference approach requires simultaneous processing of the GPS ground network and the LEO GPS measurements. All possible baselines between the LEO and the ground IGS network are formed and processed together. For each epoch three kinematic LEO coordinates are estimated, together with the double-difference ambiguity parameters. By forming double-differences, all GPS satellite clocks are eliminated and there is thus no need for highly accurate GPS satellite clocks calculated from the ground GPS network, see (Švehla and Rothacher 2002).

The disadvantage of the double-difference kinematic approach is the very large number of observations and ambiguity parameters originating from the IGS network. The noise of the double-difference observable is twice as high as that of the zero-difference observable, but all clock parameters are eliminated and, what is most important, ambiguity resolution can be performed using double-differences. This advantage of ambiguity resolution, together with different ambiguity resolution strategies, will be discussed later in this thesis.

1.7 Triple-Difference Approach

By forming triple-differences (differences of double-differences in time), ambiguities are eliminated, which allows very efficient processing algorithms to be employed. The drawback of this approach is the increase of the observation noise and efficient methods are needed to correctly account for the correlations between epochs. More about the triple-difference approach can be found in (Ijssel et al. 2003) and in (Byun, 2003].

1.8 Parameter Space in Geometric and Dynamic POD

Table 1.2 shows the parameter statistics for zero- and double-difference kinematic and dynamic POD with real LEO GPS data over one day. We immediately notice the very large number of phase observations stemming from the approx. 100 IGS ground stations selected. This, together with the rapidly changing geometry, is also the reason why a great number of double-difference ambiguities are involved. Compared to dynamic parameterization, kinematic POD has many more epoch-wise parameters. Table 1.3 shows the treatment of parameters while forming the normal equation system. In order to speed up computation, epoch-wise parameters (LEO clocks and kinematic positions) are always pre-eliminated epoch-by-epoch. At the end, only the normal equation matrix consisting of parameters that are not epoch-specific remains. This is then inverted, and by back substitution, epoch-wise parameters are obtained epoch-by-epoch, see (1.13). In the double-difference kinematic case, if more than 100 ground IGS stations are used it is more efficient to pre-eliminate double-difference ambiguities using (1.12) and invert the normal equation with kinematic parameters first.

Solution	Zero-diff. Dynamic	Zero-diff. Kinematic	Double-diff. Dynamic	Double-diff. Kinematic
Ambiguities	450	450	13200	13200
Orbit Parameters	300	-	300	-
Kinematic Coordinates	-	8640	-	8640
LEO Clocks	2880	2880	-	-
Total Number	3630	11700	13500	21840
Number of Observations	18400	18400	340000	340000

Table 1.2 Parameter and observation statistics for zero- and double-difference kinematic and dynamic POD.

Solution	Zero-diff. Dynamic	Zero-diff. Kinematic	Double-diff. Dynamic	Double-diff. Kinematic
Ambiguities	pre-eliminated	estimated	pre-eliminated	pre-eliminated
Orbit Parameters	estimated	-	estimated	-
Kinematic Coordinates	-	pre-eliminated	-	estimated
LEO Clocks	pre-eliminated	pre-eliminated	-	-

Table 1.3 NEQ parameters in the zero- and double-difference kinematic and dynamic POD.

1.9 Ambiguity Resolution

The potential to resolve phase ambiguities and thus to achieve higher levels of LEO orbit accuracy is one of the main advantages of the double-difference technique. Ambiguity resolution is certainly the most challenging aspect of double-difference POD. Here we consider two major approaches to ambiguity resolution. The first is based on phase observations only, without making use of P code measurements, and is known as the QIF strategy (Quasi-Ionosphere-Free). The second strategy (wide-lane/narrow-lane) is based on wide-lane ambiguity resolution using the Melbourne-Wübbena linear combination and subsequent resolution of the narrow-lane ambiguities using the ionosphere-free linear combination of the phase observables. More about LEO ambiguity resolution can be found in (Švehla and Rothacher 2002) and (Švehla and Rothacher 2003a).

QIF Ambiguity Resolution was developed at the CODE Analysis Center for large-area permanent networks. The QIF strategy enables L_1 and L_2 ambiguities to be resolved in one step, in which the phase observations on L_1 and L_2 are processed together and epoch- and satellite-specific ionospheric parameters

are set up. These stochastic ionospheric parameters are slightly constrained and pre-eliminated epoch-wise. The QIF strategy can cope with larger ionospheric errors than the phase-based wide-lane method, i.e., with errors up to approximately two wide-lane cycles. In order to increase the percentage of ambiguities fixed by QIF, global ionosphere maps may be used. For LEO satellites, orbiting the Earth within the ionosphere, the total electron content (TEC) has to be reduced to account for only the free electrons above the LEO orbit. This can, e.g., be performed by an appropriate integration of the alpha-Chapman layer. We found that the fast-changing ionosphere (due to the high LEO velocity) and the difficulty in computing the vertical TEC (e.g., given by IGS ionosphere maps based on a single layer model) for the altitude of the LEO, are the reasons why the QIF approach is still problematic when used for LEO ambiguity resolution, and, therefore, it will not be discussed further. For more on this subject we refer to (Švehla and Rothacher 2005a, 2005b).

1.9.1 Melbourne-Wübbena Ambiguity Resolution

In order to completely avoid ionosphere effects in ambiguity resolution, the Melbourne-Wübbena (MW) linear combination of phase and code observations is used to first resolve the wide-lane ambiguities. At the double-difference level the observation equation of the MW linear combination may be written as

$$N_{iLEO,5}^{kl} = \frac{1}{\lambda_5}\left(\frac{f_1}{f_1 - f_2} L_{iLEO,1}^{kl} - \frac{f_2}{f_1 - f_2} L_{iLEO,2}^{kl} - \frac{f_1}{f_1 + f_2} P_{iLEO,1}^{kl} - \frac{f_2}{f_1 - f_2} P_{iLEO,2}^{kl} \right) \tag{1.18}$$

where $N_{iLEO,5}^{kl}$ denotes the wide-lane double-difference ambiguity, with wavelength $\lambda_5 \approx 86\,\text{cm}$, of the baseline from station i to the LEO satellite with the GPS satellites k and l; $L_{iLEO,j}^{kl}$ and $P_{iLEO,j}^{kl}$ are the phase and P code double difference observations on both frequencies. Observation equation (1.18) is free of geometry, clock parameters, ionosphere and troposphere delays and contains only the wide-lane ambiguity and possible effects of multipath. It is, therefore, independent of the baseline length involved and ambiguity resolution may be performed baseline by baseline. To resolve the wide-lane ambiguities an iterative approach (bootstrapping) is used, where, after a first float solution, ambiguities are sorted according to best RMS and iteratively resolved starting with the best determined ambiguities. In order to ensure that ambiguities are correctly resolved, a double-difference ambiguity is only set to an integer value if exactly one integer lies within the three RMS confidence interval of the real-valued ambiguity estimate. For practical reasons, two additional criteria are used to define the pull-in region of the integer bootstrapping: (1) if the RMS of a float ambiguity is smaller than a user-specified minimum value, this minimum value will be used to define the confidence interval. This is necessary, because often the formal RMS of an ambiguity is too small and obviously resolvable ambiguities will remain unresolved; (2) if the RMS of a float ambiguity is larger than a user-specified maximum value, the ambiguity will not be resolved.
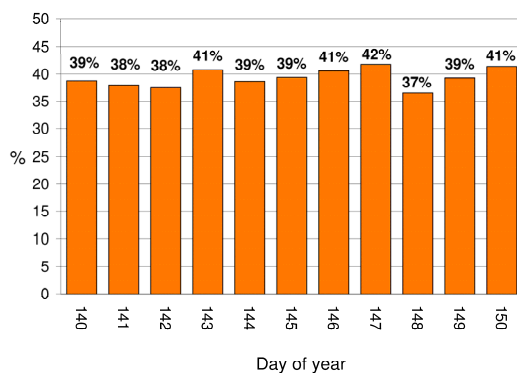
Figure 1.9 Percentage of resolved wide-lane ambiguities using the Melbourne-Wübbena linear combination.

Discussions on integer bootstrapping can also be found in (Teunissen 2001), where the decorrelation of ambiguities by Z-transformations is recommended in order to improve the success rate of the bootstrapping method. For more on this subject see (Švehla and Rothacher 2003a), (Švehla and Rothacher 2005a).

For the sake of completeness, Figure 1.9 shows the percentage of resolved wide-lane ambiguities using the Melbourne-Wübbena linear combination over 11 days (140-150/2001). However, not all GNSS receivers provide P code measurements on both the L_1 and L_2 frequencies. According to (Ray 2002), there are three main classes of GPS receivers within the IGS network, namely: 1) cross-correlators that observe C_1 and $P_2^* = C_1 + (P_2 - P_1)$ (e.g., Rogue SNR-x, AOA ICS-4000Z, Trimble 4000, and Trimble 4700); 2) Y-codeless, non-cross-correlators that observe P_1 and P_2 (e.g., Ashtech Z-XII3, AOA SNR-12 ACT, and AOA Benchmark ACT); 3) C_1, Y-codeless, non-cross-correlators that apparently function in a similar way to other modern Y-codeless receivers, but report C_1 (instead of P_1) and P_2 (Trimble 5700, Leica CRS1000, and Leica SR9600). For those receivers that do not provide P code on both frequencies, the use of differential code biases $P_1 - C_1$ for the GPS satellites, available from the CODE IGS AC, considerably improves wide-lane ambiguity resolution.

1.9.2 Narrow-lane Ambiguity Resolution

If wide-lane ambiguities have successfully been resolved, the ionosphere-free linear combination of the L_1 and L_2 phase observations can be used to resolve the corresponding narrow-lane ambiguities. The ionosphere-free linear combination may be written in the form

$$L_{iLEO,3}^{kl} = \rho_{iLEO,3}^{kl} + B_{iLEO,3}^{kl} \tag{1.19}$$

where the first term denotes double-difference geometrical distance and the second term the ionosphere-free ambiguity bias. Note that other terms such as tropospheric refraction delay, multipath and noise are not explicitly shown and higher-order ionospheric terms have been ignored. The ionosphere-free bias can be written as

$$B_{iLEO,3}^{kl} = \frac{f_1^2}{f_1^2 - f_2^2} \lambda_1 N_{iLEO,1}^{kl} - \frac{f_2^2}{f_1^2 - f_2^2} \lambda_2 N_{iLEO,2}^{kl} \tag{1.20}$$

where $\lambda_j\,(j = 1, 2)$ denote the wavelength of L_1 and L_2 and $N_{iLEO,j}^{kl}$ the corresponding double-difference ambiguity. By introducing the known wide-lane ambiguity

$$N_{iLEO,5}^{kl} = N_{iLEO,1}^{kl} - N_{iLEO,2}^{kl} \tag{1.21}$$

into (1.20) we obtain

$$B_{iLEO,3}^{kl} = \frac{cf_2}{f_1^2 - f_2^2} N_{iLEO,5}^{kl} + \frac{c}{f_1 + f_2} N_{iLEO,1}^{kl} \tag{1.22}$$

where the first term contains the resolved wide-lane ambiguity and the second term is known as narrow-lane ambiguity. In this way the ionosphere refraction has been eliminated and using only phase observations the remaining ambiguity $N_{iLEO,1}^{kl}$ can be resolved with the same algorithms as used for wide-lane ambiguities. In comparison to wide-lane ambiguities, all baselines have to be processed simultaneously to obtain the best possible kinematic orbit by accounting for the correct correlations between the baselines and thus obtaining the best possible bootstrapping results. Due to the short wavelength (11.6 cm) of the narrow-lane ambiguities, all biases stemming from the orbits of the GPS satellites and tropospheric refraction have to be modeled very

carefully. Tropospheric biases can be corrected for by using tropospheric zenith delays (and gradients). The impact of errors in the IGS Final Orbits for GPS satellites on double-differences with LEO satellites, given their current level of accuracy, is negligible. The station coordinates of the ground network should be consistent with the GPS satellite orbits. For more on this subject we refer to (Švehla and Rothacher 2005a, 2005b).

1.9.3 The Impact of Narrow-Lane Ambiguity Resolution and Tracking Geometry on Ground GPS Double-Differences with LEO Satellites

GPS phase observations for the CHAMP satellite were simulated with a white noise of 1 mm using the same physical and mathematical models as those used in the processing of real data. The white noise applied to the carrier-phase of IGS stations was 1 mm and no other error sources were simulated (no systematic effects). Simulation was carried out with a higher cut-off angle of 15°, with the maximum number of tracked GPS satellites set to 8, and with 105 stations of the IGS ground network. Figure 1.10 shows the kinematic orbit positions obtained with an ambiguity float/fixed solution against the true orbit used in the simulation. It is interesting to note the systematic excursions of up to a few centimeters in the float solution caused only by the observation noise, low number of tracked GPS satellites and probably also by the high correlation between ambiguities and kinematic coordinates. The large deviations at about 0.75 and 2.45 hours are the result of a small number of satellites tracked around these epochs. Figure 1.10 (right) shows the kinematic orbit with fixed ambiguities after narrow-lane bootstrapping with 98% of the narrow-lane ambiguities resolved. A systematic once-per-rev. pattern in the kinematic orbit with float ambiguities is clearly visible in Figure 1.10 (left) and is completely eliminated after ambiguity resolution in Figure 1.10 (right), producing kinematic orbit determination to an accuracy of less than one centimeter. This analysis shows that, if the kinematic orbit is estimated using double-differences from the IGS network, ambiguity resolution needs to be performed due to the very large number of ambiguities introduced by that network.
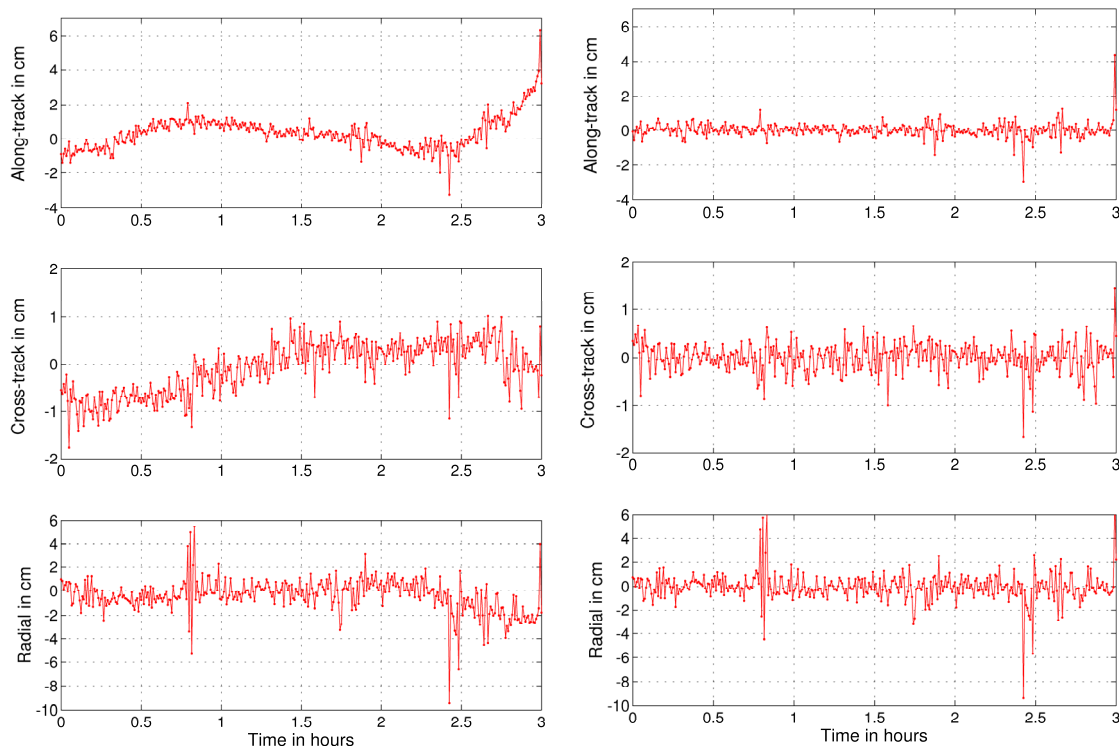
Figure 1.10 Kinematic orbit with float (left) and fixed ambiguities after narrow-lane bootstrapping (right) based on simulated data with high cut-off angle of 15° and max. number of tracked GPS satellites set to 8.

Comparing the two sets of results in Figure 1.10, we may expect orbit changes in the order of a few centimeters when fixing the double-difference ambiguities in kinematic POD based on double-differences. It is interesting to note that, analogous to the height component for ground stations, the radial kinematic component is less accurately determined by a factor of about 2-3.

We would like to point out that for POD of the present geodetic missions such as CHAMP, JASON and GRACE, GPS measurements over all elevations $0° - 90°$ are used. In the early days of CHAMP, GPS measurements were collected even below the antenna horizon (down to -15°) throughout the satellite constellation, but, due to their poor quality, all measurements below 0° elevation were rejected in the pre-processing stage and were not used in POD. Later on, the CHAMP BlackJack software was upgraded and GPS satellites below the antenna horizon were no longer tracked. The main part of the LEO GPS data is at elevations of $5° - 20°$ and, therefore, an elevation cut-off angle of 0° is strongly recommended for any satellite mission that requires orbits with high accuracy. The usage of a cut-off angle above 0°, e.g., above 15°, is very disadvantageous and may lead to gaps in kinematic POD as shown in Figure 1.10 . It is important to note that weighting of the phase measurements as a function of elevation is not necessary in the POD of CHAMP and GRACE, which means that the phase measurements over the entire elevation range from 0° to 90° are of similar quality. Elevation-dependent weighting is still required for ground GPS applications due to multipath and troposphere effects. For more on this subject see (Švehla and Rothacher 2003a), (Švehla and Rothacher 2005a).

1.9.4 Narrow-lane Kinematic and Reduced-Dynamic Bootstrapping

Using the ionosphere-free linear combination of the carrier-phase measurements and the resolved wide-lane ambiguities, an iterative resolution of the narrow-lane ambiguities (bootstrapping) can be performed. Two main methods were studied to perform the narrow-lane ambiguity resolution with LEO data. In the kinematic bootstrapping epoch-wise coordinates are pre-eliminated in order to reduce the size of the normal-equation matrix. The first solution is a float solution where the ambiguities are real numbers. Then the best estimated ambiguities are set to integer numbers, the normal equation system is updated and re-inverted and the whole procedure is repeated. More about this type of bootstrapping and the criteria applied for ambiguity fixing can be found in (Švehla and Rothacher 2002). The same procedure can also be used when estimating dynamic orbit parameters. We then speak of dynamic bootstrapping, see (Švehla and Rothacher 2005a, 2005b).

The reduced-dynamic orbit model used in our approach is based on the dynamic model originally developed at the CODE Analysis Center for GPS orbits and here subsequently used for LEO satellites (Švehla and Rothacher 2002) making use of the estimation of pseudo-stochastic pulses (small velocity changes). For the CHAMP and GRACE orbits, stochastic pulses are set-up every 6 minutes. Bootstrapping with this reduced-dynamic parameterization can be used as an independent check for the ambiguity resolution based on kinematic bootstrapping. When comparing the double-difference ambiguities obtained from the kinematic and the reduced-dynamic bootstrapping no discrepancies were found.

Baseline-wise ambiguity resolution could, in principle, be applied for kinematic as well as for dynamic orbits, but highly accurate a priori orbits have to be available in that case. The orbits are then fixed in the baseline by baseline ambiguity resolution. The drawback of this method is that the criteria to fix the ambiguities have to be very restrictive in order to ensure that ambiguities are correctly resolved. In principle, baseline-wise ambiguity resolution can be performed iteratively: after the first baseline-wise ambiguity resolution step a new orbit is computed making use of the fixed ambiguities and a new iteration of the baseline-wise ambiguity resolution is performed with the updated orbits. Our experience with baseline-wise ambiguity resolution shows that highly accurate dynamic orbit models are a prerequisite for this method. More details about Melbourne-Wübbena wide-laning with narrow-lane bootstrapping may be found in (Švehla and Rothacher 2002).

Ambiguity resolution was performed in the double-difference case for GPS week 1175/2002. Using the Melbourne-Wübbena linear combination, about 59% of the wide-lane ambiguities could be resolved. These wide-lane ambiguities were introduced in the next step to resolve the narrow-lane ambiguities. Epoch-wise

coordinates were pre-eliminated from the NEQ system in kinematic, and orbital parameters in reduced-dynamic POD, leaving ambiguities as remaining parameters for bootstrapping. The overall percentage of resolved narrow-lane ambiguities was 27% of all ambiguities or 59% of the ambiguities for which the wide-lane ambiguities were successfully resolved with the Melbourne-Wübbena approach. Comparing kinematic and reduced-dynamic bootstrapping, no discrepancies were found in the fixed ambiguities. Due to the large number of ambiguity parameters (5000 per day), bootstrapping is very time-consuming and requires about 100 inversions of the 1-day NEQ for both approaches.

Figure 1.11 shows the impact of ambiguity resolution on reduced-dynamic orbits based on double-differences. Ambiguity resolution changes the determined orbit by $1 - 2\,\mathrm{cm}$.
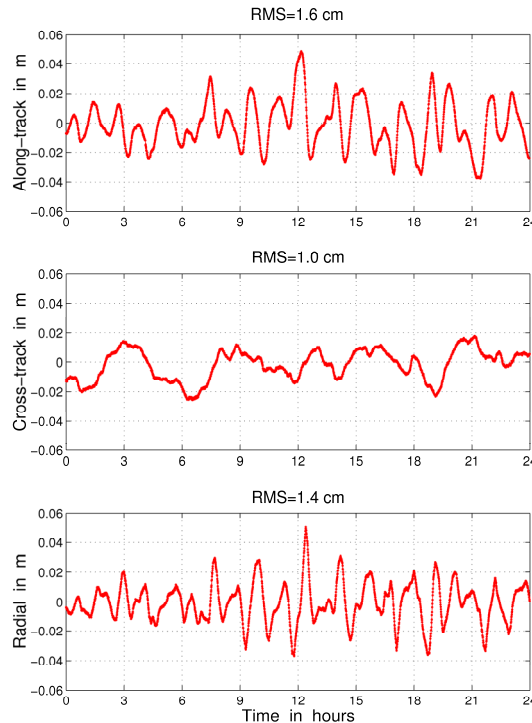
Figure 1.11 Impact of ambiguity resolution: difference between reduced-dynamic orbit with float and fixed ambiguities, day 200/2002.

1.10 Differential Code Biases and Kinematic POD

The term differential code biases (DCB) denotes biases in the tracking of different code observables, e.g., between P_1 and P_2 or C/A and P_1 code that can be individually assigned to each GPS satellite, as well as, to each GPS receiver. When estimating the GPS satellite clock corrections from the global IGS network using the ionosphere-free linear combination, the differential code biases are included in the clock correction. Phase ″iono-free″ GPS clocks are then consistent with the corresponding GPS orbits. Unfortunately, this is not the case for the inter-channel biases, and therefore this effect has to be correctly and very accurately calibrated. When performing ambiguity resolution based on the Melbourne-Wübbena linear combination, the quality of P code measurements has to be very high and GPS satellite differential code biases should be taken into account. In the case of the DCBs of a GPS receiver, they should be constant over time, and as small and as independent of the environment (e.g., temperature) as possible. DCBs play a role in kinematic POD only if

ambiguity resolution based on the Melbourne-Wübbena linear combination is performed, or if ionospheric delays are estimated.

2.Reference Frame From the Combination of a LEO Satellite with GPS Constellation and Ground Network of GPS Stations

I n this section we demonstrate the combination of a LEO satellite with the satellites of the GPS constellation and the ground networks of space geodesy techniques (GPS, SLR, DORIS) in the generation of reference frame parameters. We show clear improvements in terrestrial reference frame parameters after the combination of the GPS constellation in MEO with spaceborne GPS, DORIS and SLR measurements from the Jason-2 satellite in LEO orbit, including station coordinates, tropospheric zenith delays, Earth rotation parameters, geocenter coordinates and GPS satellite orbit and high-rate clock parameters. We analyze the impact of the LEO data on the terrestrial reference frame parameters and possible improvements they could bring. (See also (Svehla et al. 2010b).) This is a continuation of the work performed with the GPS data from the Jason-1 satellite, where the strong impact of the LEO data on the global parameters has already been demonstrated by means of simulated GPS measurements and variance-covariance analysis (Švehla and Rothacher 2006a).

Terrestrial reference frames are usually defined by a set of station coordinates that are estimated over a long period of time using a combination of different space geodesy techniques. However, in the case of Precise Point Positioning (PPP) of a GPS receiver on the ground or kinematic or dynamic POD of LEO satellites using GPS, reference stations on the ground are not directly used to estimate the orbit of a LEO satellite or coordinates of a GPS receiver on the ground. The PPP of a ground station or POD of LEO satellites is based on an intermediate reference frame defined by the GPS satellite orbits and epoch-wise estimates of GPS satellite clocks. Any error in the GPS satellite orbits and clocks, or in this intermediate space-based reference frame (that is highly temporal in nature), will map directly into the LEO kinematic/dynamic orbit and gravity field determination (CHAMP, GRACE, GOCE), altimetry results (Jason-2, Sentinel-3, etc.) or coordinates of a ground station. Therefore, an instantaneous terrestrial reference frame can be defined as a frame created by the epoch-wise solution of GNSS orbit and clock parameters supported by other space geodesy techniques such as SLR, DORIS and VLBI. In the next section we introduce the concept of phase clocks in order to consistently bridge the gap between ground-based and space-based terrestrial frames and show how a terrestrial frame can be transferred to the LEO orbit avoiding biases associated with the code GPS measurements.

At the end we give an insight into the generation of an instantaneous reference frame from different GPS frame solutions (e.g., provided by IGS ACs) by means of least-squares collocation using a so-called intermediate reference sphere in LEO or GNSS orbit. The use of a simple weighted average, which is often used in the combination of GNSS solutions from different IGS ACs without taking into account correlations in time (and space) of each individual solution, will always introduce systematic effects that are not equally distributed over an imaginary sphere at the GNSS orbit height.

2.1 General Remarks on the Combination of a LEO Satellite with the GPS Constellation for Reference Frame Determination

In (Švehla and Rothacher 2006a) and in various publications before we demonstrated the strong impact of LEO data (from GRACE-A&B and Jason-1 satellites) on reference frame parameters, indicating that altimetry satellites are the best candidates for such a combination. However, due to the onboard multipath and the performance of the Jason-1 GPS receiver, those results were not based on real GPS measurements, but rather on simulations. In the case of GRACE-A&B satellites in a lower LEO orbit, we noticed a strong impact of the gravity field used in the LEO POD on the combined reference frame solution.

The quality of the instantaneous reference frame defined by the GPS satellites will more strongly affect LEO satellites in very low orbit (such as GOCE) than satellites in a high LEO orbit (such as Jason-2). This is because the orbit of the Jason-2 satellite requires a rather modest number of orbital parameters comparable to the parameterization of the GPS satellite orbits. Furthermore, in terms of non-gravitational forces, satellites in a high LEO orbit are mainly affected by solar radiation pressure, whereas satellites in a very low LEO orbit are, besides solar radiation, mainly affected by air-drag. Satellites in higher LEO orbits are very good candidates for the combination of space geodesy techniques. With the Jason-2 satellite, all GPS satellites in the GPS constellation can be connected in only ≈ 1.5 hours, and all ground SLR and DORIS stations within the same timeframe. One can imagine the Jason-2 satellite as a station with well-defined ties between different space geodesy techniques collocated on the same satellite, flying below the constellation of GPS satellites and above the ground networks of the different space geodesy techniques (GPS, SLR, DORIS, VLBI).

Thus we can draw the conclusion that altimetry satellites in higher LEO orbits with an onboard GPS, DORIS and SLR are very good candidates for the combination of space geodesy techniques, since the orbit parameterization is very similar to GPS satellites and the orbit is also mainly affected by solar radiation pressure.

2.2 Terrestrial Frame Parameters from the Combination of a LEO Satellite with the GPS Constellation

Here we used GPS, SLR and DORIS measurements from the Jason-2 satellite during the CONT'08 Campaign (10.8.-31.8.2008) and combined them with GPS measurements from about 150 stations of the global IGS ground network and estimated typical reference frame parameters, such as GPS orbits and clocks, station coordinates, Earth rotation parameters, troposphere zenith delays and geocenter coordinates. In essence, we generated typical IGS-type daily solutions and added DORIS and SLR measurements from the Jason-2 satellite on the observational level. As a priori datum definition we used the station coordinates of GPS, DORIS and SLR stations in ITRF2005 and a no-net-rotation condition for GPS and DORIS stations. The scale was mainly defined by SLR measurements to Jason-2 and the coordinates of ground ILRS stations (high constraints). Absolute phase center variations from the robot calibration (Montenbruck et al. 2009) were used for the GPS antenna on board the Jason-2 satellite. In order to prevent the remaining systematic effects of the Jasson-2 antenna phase center offset propagating into the geocenter z-coordinate, we estimated the phase center offset for the Jason-2 GPS antenna in the up direction. Figure 2.1 shows the impact of GPS, DORIS and SLR measurements on Jason-2 POD as well as on the orbit determination of all satellites in the GPS constellation. This solution was based on ambiguity resolution for GPS measurements from the ground IGS network. For the orbits of GPS satellites, the effect is in the order of $12-16$ mm RMS. This is a significant effect, considering that the current accuracy of GPS satellite orbit determination is at a similar level. For LEO orbit, the main effect is in the along-track direction (three times higher than for the radial direction). However, the radial orbit component is changed by an RMS of about 5 mm.
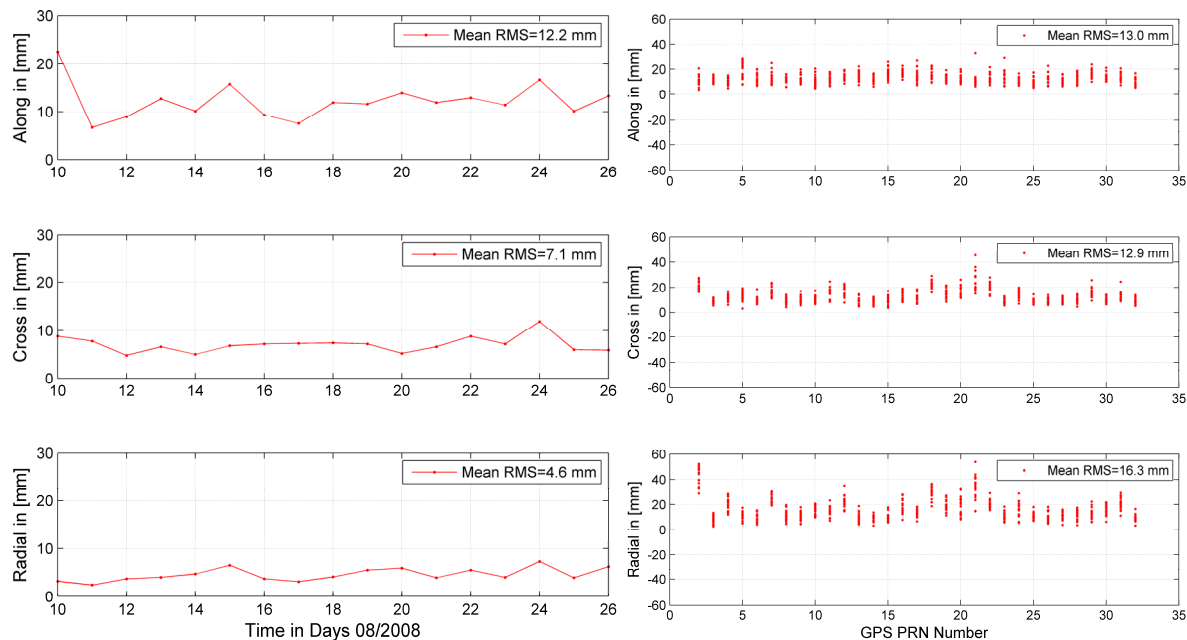
Figure 2.1 Impact of the combined GPS, DORIS and SLR measurements on the Jason-2 orbit (left) and the orbits of the GPS constellation (right). Ambiguity resolution was performed for GPS measurements from the ground network. For LEO, the main effect is in the along-track direction, for GPS all components are affected by $12 - 16$ mm RMS. Notice that for the LEO, the RMS in the radial component is in the order of 5 mm (significant for altimetry satellites that typically have an accuracy of the radial orbit component at that level).
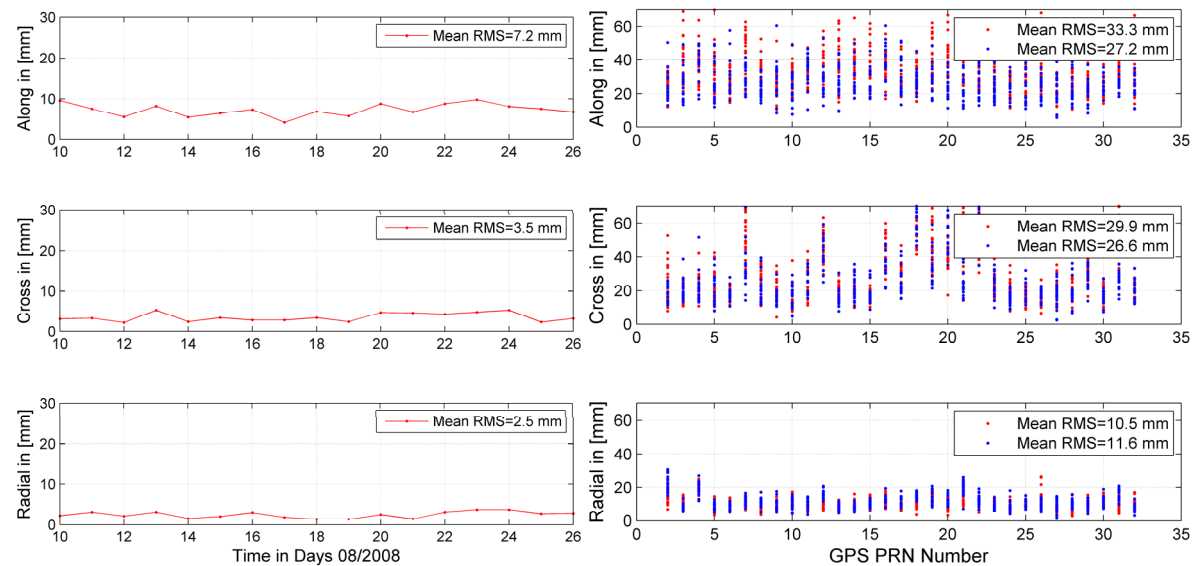
Figure 2.2 Impact of the ambiguity resolution carried out for GPS measurements from the ground IGS network on the combined Jason-2 orbit (with GPS, DORIS and SLR data from Jason-2) (left) and on the determined orbits of the GPS constellation (right). Blue/red dots show orbit solution with/without Jason-2 data in the combination respectively, vs. ambiguity-fixed solution. For both the LEO and GPS satellites, the main improvement is in the along-track direction. Notice that for LEO, ambiguity resolution improves the along-track orbit component by a factor of two compared to other components. Inclusion of GPS data from just one LEO in the combination has a similar effect to ambiguity resolution for the entire IGS network in the GPS-only case.

23

This is a significant effect for all altimetry satellites as the typical consistency of the radial orbit component between different solutions (e.g., JPL. CNES, ESOC, GFZ) is about $5-8\,\mathrm{mm}$, see e.g., (Flohrer et al. 2011).

Figure 2.2 shows the impact of ambiguity resolution carried out for the ground IGS network when GPS, DORIS and SLR measurements from the Jason-2 satellite are combined with measurements from the GPS constellation for the POD of Jason-2 and GPS satellites. One can see that ambiguity resolution improves the along-track orbit component of the Jason-2 satellite by a factor of 2 or even 3 compared with the radial orbit component. For the orbits of the GPS satellites, the effect of ambiguity resolution is surprisingly less visible. However, this is what is to be expected, since inclusion of GPS measurements from the Jason-2 satellite decorrelates all GPS orbit parameters, i.e., the LEO satellite connects all GPS satellites during just one orbit revolution of typically about 1.5 hours. Figure 2.2 (right) shows that inclusion of GPS measurements from just one satellite in higher LEO orbit has a similar effect to carrier-phase ambiguity resolution of the GPS measurements for the entire IGS network (the ambiguity-fixed solution was used as a reference).

More and more altimetry satellites are now carrying GPS receivers as well as DORIS and SLR. It is expected that, in future, GNSS receivers will track all GNSS systems as well as receive DORIS signals. With Jason-2 we clearly demonstrated that LEO data can be included in the generation of reference frame parameters and that there is a good reason to do so.

2.2.1 Geocenter Estimates from the Combination of a LEO Satellite with GPS Constellation

Table 2.1 shows the Helmert transformation of weekly station coordinates (after stacking of daily normal equations) against the ITRF2005. One can immediately notice a very large systematic translation of the geocenter by about $-5.8\,\mathrm{cm}$ in the Z-direction that is very much uniform for all three weeks of the CONT'08 Campaign and also very uniform in the daily solutions. The other six parameters of the Helmert transformation are significantly smaller. The reason for such a large effect is most likely the SLR frame, since the orbits of the Jason-2 satellite do not indicate that there is any bias introduced by GPS data processing (e.g., phase center definition). It is also interesting that weekly geocenter estimates drift slightly over the three weeks at a rate of about $2\,\mathrm{mm/yr}$ to $5\,\mathrm{mm/yr}$. This is consistent with the variation of the SLR origin values as given in (Pavlis 2012). However, Table 2.1 shows very smooth SLR origin variations estimated using only one LEO satellite, as depicted in Figure 2.3. Unfortunately, the data set of the CONT'08 campaign is limited to just three weeks to reliably extrapolate those values over a longer time span. Our combination of space geodesy techniques from a LEO satellite, GPS constellations and ground GPS/SLR/DORIS networks clearly demonstrates improvements of the combined solutions and the presence of biases in the ITRF2005 reference frame.

Weekly Geocenter Estimates (CONT'08 Campaign)

Week 1	Week 2	Week 3
dx = -0.83 mm	dx = -1.78 mm	dx = -1.72 mm
dy = -0.94 mm	dy = -1.67 mm	dy = -1.22 mm
dz = -5.90 mm	**dz = -5.75 mm**	**dz = -5.60 mm**
rx = 0.021 mas	rx = 0.067 mas	rx = 0.059 mas
ry = 0.052 mas	ry = 0.055 mas	ry = -0.011 mas
rz = -0.051 mas	rz = -0.077 mas	rz = -0.051 mas
scale = 0.13 ppb	scale = 0.14 ppb	scale = 0.16 ppb

Table 2.1 Helmert transformation of weekly coordinates solution (after stacking daily NEQs) from the combined GPS/Jason-2 constellation (GPS, DORIS, SLR) − CONT'08 Campaign. Notice a large systematic translation of about $-5.8\,\mathrm{mm}$ in the geocenter z-coordinate that is very much uniform for all three weeks.
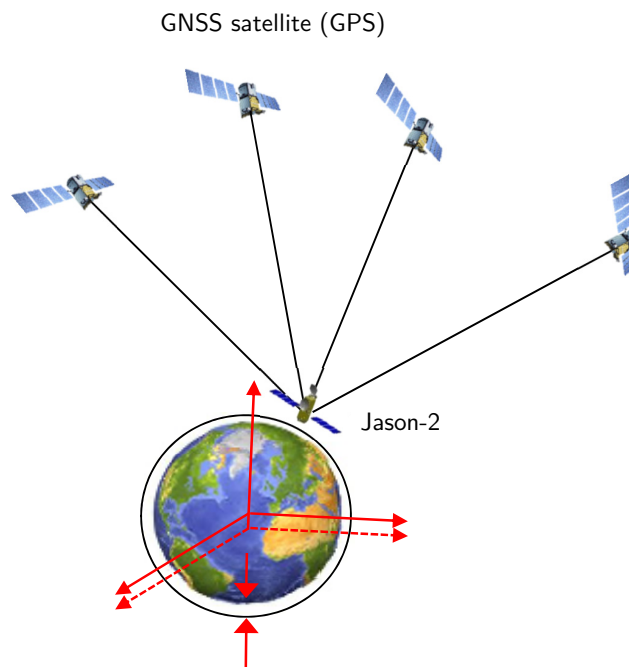
Figure 2.3 Geometrical representation of geocenter estimation from the combination of a Jason-2 satellite with the GPS constellation.

2.2.2 SLR Network Effect

Figure 2.3 and Figure 2.4 show a possible explanation for the Z-offset in the estimated geocenter – the fact that the majority of SLR stations are located in the northern hemisphere. Any common range bias and typically the high weight of SLR measurements will bias the determined orbit of the Jason-2 satellite towards the Northern hemisphere. Any frame translation in a west-east direction will average out. However, due to the uneven distribution of SLR stations this is not the case with the Z-direction.

Let us now see if there is any similar offset in the GPS satellite orbits after combination. Figure 2.5 shows the translation and scale of the GPS constellation after the combination with Jason-2 data (GPS, DORIS, SLR). One can immediately see a large offset of -5.4 mm in the geocenter z-coordinate that is very similar to the Z-offset of -5.8 mm in Table 2.1. Thus, both Jason-2 and GPS orbits are shifted by the same amount in the Z-direction. In addition, very interestingly, Figure 2.5 (right) shows that the combination of GPS measurements from a ground network observing the GPS constellation and GPS data from a LEO satellite
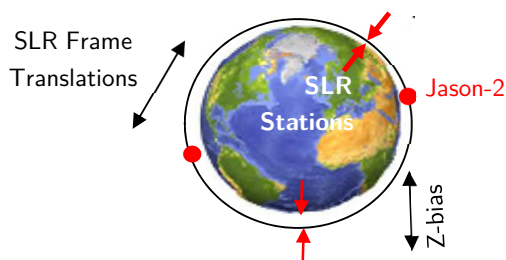
Figure 2.4 Possible explanation for the Z-offset in the geocenter – the fact that the majority of SLR stations are in the Northern hemisphere. Any frame translation in a west-east direction will average out.
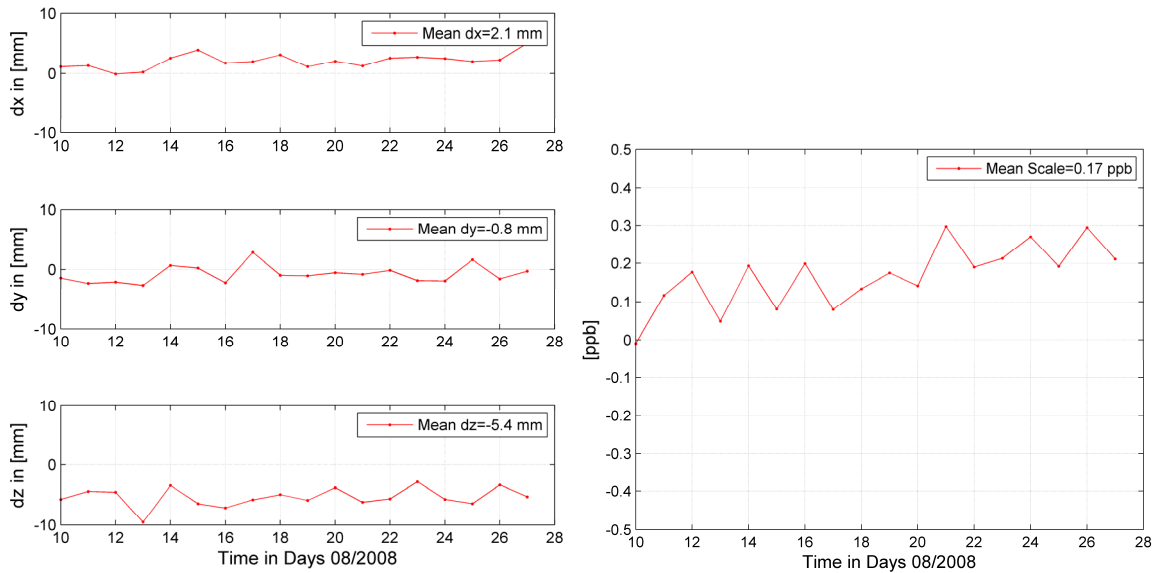
Figure 2.5 Translation (left) and scale (right) of the orbits of the GPS constellation after combination with Jason-2 data (GPS, DORIS, SLR). Notice a large systematic translation of $-5.4\,\text{mm}$ in the z-coordinate that is very similar to the geocenter Z-offset of $-5.8\,\text{mm}$ in Table 2.1. The scale difference after combination indicates that the Jason-2 data reduces the SLR bias in GPS orbits (GPS scale) by $5\,\text{mm}$ ($0.2\,\text{ppb}$).

reduces the SLR bias in GPS orbits by $5\,\text{mm}$. From this, we can conclude that there must be some residual SLR bias in GPS satellite orbits in the order of about $-5\,\text{mm}$. This effect is mapped into LEO orbits (mainly in along-track and radial) and due to the typically high weight of SLR measurements and the majority of SLR stations being in the Northern hemisphere, this then shifts the entire GPS/LEO frame in the Z-direction. Thus, there is a bias between SLR and GPS frames that could be removed to a great extent by estimating LEO antenna phase center offset.

2.2.3 Earth Rotation Parameters from the Combination of a LEO Satellite with GPS Constellation

Figure 2.6 shows the effect on X-Pole and Y-Pole coordinates of a combination of GPS, DORIS and SLR data from the Jason-2 satellite in high LEO orbit with the satellites of GPS constellation and ground IGS/IDS/ILRS networks. One can see a bias of the order of $0.4\,\text{mas}$ in both X-Pole and Y-Pole and the effect is within $0.15\,\text{mas}$ peak-to-peak over a period of three weeks of the CONT'08 Campaign. This bias of $0.4\,\text{mas}$ gives about $1.2\,\text{mm}$ at the Earth's surface or about $5.1\,\text{mm}$ at GPS orbit altitude. However, those are daily solutions, without any stacking of normal equations over a longer period of time. Figure 2.7 shows corresponding length-of-day estimated for the same period.

Since combining LEO space geodesy measurements (GPS, DORIS, SLR) with the GPS constellation gives the main orbit effect in the along-track direction (for both LEO and GPS satellites, see Figure 2.3), it is expected that about 16 LEO orbit revolutions per day could "see" the sub-daily parameters in Earth rotation. However for this, it is expected that the ambiguity resolution for LEO GPS measurements would need to be performed. Later in this thesis, we introduce the concept of track-to-track ambiguities, where, by connecting 16 LEO ambiguity parameters per GPS satellite over one day, one could obtain only one ambiguity per GPS satellite. Thus, by reducing the number of ambiguity parameters, it is expected that the LEO GPS data will significantly contribute to the estimation of daily and sub-daily Earth rotation parameters.
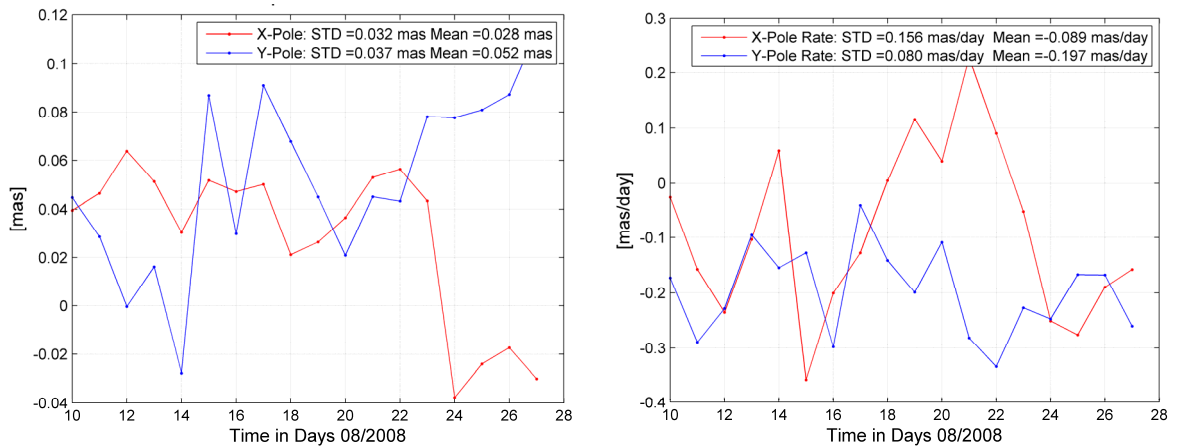
Figure 2.6 Impact of the combination of Jason-2 data (GPS, DORIS, SLR) with GPS constellation and ground IGS/ILRS/IDS networks on pole coordinates (left) and rates in pole coordinates (right). CONT'08 Campaign.
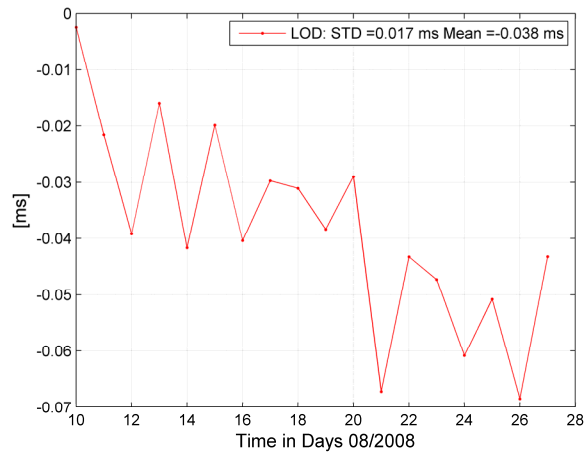
Figure 2.7 Impact of combination of Jason-2 data (GPS, DORIS, SLR) with GPS constellation on length-of-day (LOD). CONT'08 Campaign.

2.3 An Instantaneous Reference Sphere – A Proposal for the GNSS Orbit Combination and Terrestrial Frame Realization by Means of Least-Squares Collocation

An instantaneous terrestrial reference frame can be defined as a frame realized by the epoch-wise solution of GNSS orbit and clock parameters supported by other space geodesy techniques such as SLR, DORIS and VLBI. It is typically formed by 7-8 GPS satellites in the field of view of a ground station or LEO satellite. In the case of orbit determination of LEO satellites or Precise Point Positioning (PPP) of a ground GPS receiver, we use an intermediate space-based terrestrial reference frame given by GNSS orbit and clock parameters and not by station coordinates on the ground. Terrestrial reference frames are usually defined by a set of station coordinates that are estimated over a long period of time using a combination of different space geodesy techniques such as GNSS, SLR, DORIS and VLBI. This intermediate instantaneous space-based reference frame is temporal in nature and any error in, e.g., GNSS satellite clock parameters will map directly into the LEO kinematic/dynamic orbit, gravity field determination (CHAMP, GRACE, GOCE), altimetry results (Jason-2, Sentinel-3, etc.) or PPP coordinates of a ground GPS receiver. In the next section, we introduce the concept of phase clocks (carrier-phase estimation of GNSS clock parameters) in order to consistently bridge the gap between ground-based and space-based terrestrial frames and thus develop a bias-free means of transferring a terrestrial frame to LEO orbit considering typical systematic effects and biases associated with the GPS code measurements. One could assume that GNSS orbit solutions provided by different IGS ACs or IGS Final GNSS orbits themselves are a significantly better tool for generating phase clocks for GNSS satellites and defining this intermediate space-based terrestrial frame. By definition, IGS Final GNSS orbits are the best in terms of RMS compared to any other solution. However, colored noise introduced by a combination of different orbit solutions directly maps into the LEO kinematic/dynamic orbits, gravity field determination and altimetry results. The same is true of high-rate IGS clock parameters for GNSS satellites that are combined as a weighted average of different solutions, however without taking into account any correlation in time between the subsequent epochs of the individual solution. Compared to precise point positioning, for a series of applications in geosciences, IGS Final Orbits/Clocks are not always the best option.

Therefore, from the point of view of least-squares, least-squares collocation is an alternative and promising approach for the combination of different IGS orbit solutions and for the realization of the intermediate instantaneous space-based reference frame. Rather than using a weighted average between different GNSS solutions every epoch (as is done now), one could have a different covariance function for each individual solution that would correctly model noise and correlations between epochs over time. As a result, least-squares collocation would provide an unbiased estimate (zero-mean). Typically, in least-squares collocation one splits the noise from the signal associated with the homogeneous and isotropic covariance function to obtain the best estimate of parameters for a given set of observables. In this way, one could filter out and smooth spatial and temporal systematic effects in each individual solution.

Following (Moritz 1980) the observation equation in least-squares collocation can be written as

$$l = Ax + s + n \tag{2.1}$$

where x is the vector of estimated parameters, A is sensitivity or design matrix and l is the vector of observations, often denoted as "observed-minus-computed" ($-l$). In the case of least-squares collocation, the vector of errors is split into two parts: in addition to the measuring errors n ("noise") we have the "signal" s. The noise n is a random (stochastic) quantity with a probability distribution with the mathematical expectation denoted here by E. The signal s is not a stochastic quantity in the same sense as noise, i.e., repeated observation of the same quantity give different noise values, but the values for signal s remain the same. Thus, expectation $E\{s\} = s$ and $E\{n\} = 0$. If we now introduce an operator M that denotes a homogeneous and

isotropic average over the sphere, rather than an expectation in a probabilistic sense, we may write $M\{s\} = 0$ and $M\{n\} = n$. This leads us to the following condition of the least-squares collocation

$$s^t C_{ss}^{-1} s + n^t C_{nn}^{-1} n = min \tag{2.2}$$

with the estimated unknown parameter vector

$$\widehat{x} = (A^T \overline{C}^{-1} A)^{-1} A^T \overline{C}^{-1} l \tag{2.3}$$

and estimated values of the signal (predicted and/or filtered)

$$\widehat{s} = C_{\widehat{s}s} \overline{C}^{-1} (l - A \widehat{x}) \tag{2.4}$$

where

$$\overline{C} = C_{ss} + C_{nn} \tag{2.5}$$

where C_{nn} is the noise covariance matrix, C_{ss} the signal covariance matrix and $C_{\widehat{s}s}$ contains covariances between a new and the given signal points. (For more about least-squares collocation and determination of empirical covariance functions we refer to (Moritz 1980)). The reason why least-squares collocation can offer realization of an instantaneous reference frame that will provide homogeneous and isotropic positioning, is that the empirical covariance function of the signal is determined only as a function of distance (or time), i.e., an angle between two points on the reference sphere. Therefore, collocation can map the remaining residual signal in the combination of space geodesy techniques or in the generation of an instantaneous reference frame in a theoretically correct way. The use of a simple weighted average between different solutions that is often used in the combination of GNSS solutions from different IGS ACs without taking into account the correlations in time (and space) of each individual solution, will always introduce systematic effects that are not equally distributed over a reference sphere at the GNSS orbit altitude. Figure 2.8 shows an instantaneous reference sphere at the GNSS or LEO orbit altitude that one could use to model residual systematic effects in each individual GNSS solution.

GNSS satellites (GPS)

Instantaneouss Reference Sphere
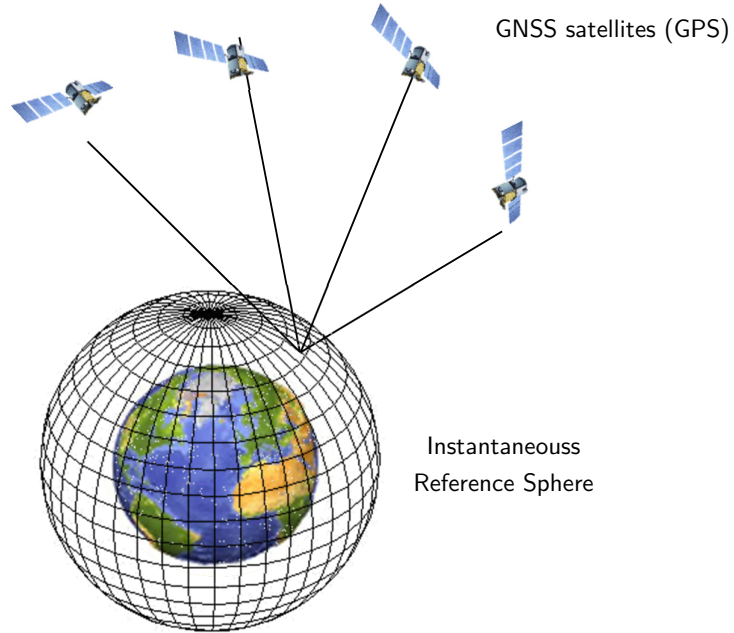
Figure 2.8 Reference frame realization by least-squares collocation on the reference sphere placed at the altitude of the LEO/GNSS orbit, offering optimal combination and variance-covariance properties.

In addition, for a given reference sphere at LEO orbit, one could construct a global grid of epoch-wise positioning solutions (epoch by epoch), and by generating temporal maps on that reference sphere one could monitor geographically correlated errors of the instantaneous reference frame realization based on GNSS orbits and clock parameters provided by different IGS Analysis Centers. This would be analogous to temporal gravity field maps modeled by spherical harmonics.

In the next step, the orbits of different LEO satellites could be mapped onto that reference sphere in LEO orbit and a combined instantaneous terrestrial reference frame based on a GNSS constellation (and LEO data) could be generated using least-squares collocation with parameters. This would be the spatial approach to the generation of the terrestrial frame using GNSS (and LEO) data. Another, straightforward, classical approach is to combine the LEO with a GNSS constellation and ground IGS/ILRS/IDS networks at the conventional normal equation or observation level. The advantage of the spatial combination strategy is the possibility of obtaining a reference frame that will give homogeneous and isotropic positioning results over the entire reference sphere at the LEO orbit altitude, irrespective of the location and direction (azimuth) on the sphere. Here, homogeneous positioning is defined as positioning that provides the same consistency or spatial correlation anywhere on the reference sphere and isotropic means over all azimuths, for any given point on the reference sphere. By definition, correlation functions of the instantaneous reference sphere include all information already contained in the normal equations of the individual IGS solutions or frame solutions of the space geodesy techniques, the difference is only that a spatial dimension is introduced in the combination or generation of the terrestrial reference frame by least-squares.

Least-squares collocation is thus a very good candidate for providing an alternative to conventional approaches in the combination of individual reference frame solutions (e.g., by IGS) or in the generation of terrestrial reference frames (e.g., by IERS) in order to provide globally homogeneous and isotropic positioning results.

3. Geometrical Model of the Earth's Geocenter Based on Temporal Gravity Field Maps

I n this section we derive the rate in geocenter z-coordinate from the secular rate of low-degree odd coefficients ("pear-shaped") over the last 10 years (GRACE RL05) and compare it with results from the global GPS and SLR solutions, tide-gauge records over the last 100 years and the limited data set of geocenter z-coordinates estimated from the combined orbit determination for the Jason-2 satellite and the GPS constellation. This confirms the initial assumption that the asymmetrical mean sea lever rise between the Northern and the Southern hemispheres is reflected in the rate of asymmetric surface spherical harmonics ("pear-shaped"). Following (Cazenave and Llovel 2010), satellite altimetry observations suggest that the mean sea level has been rising faster over the Southern than over the Northern Hemisphere, whereas recently (Wöppelmann et al. 2014) using selected tide-gauges measurements corrected with the glacial isostatic adjustment (GIA) and GPS velocities report the opposite sign, i.e., the mean sea level rise of 2.0 ± 0.2 mm/yr for the Northern hemisphere and 1.1 ± 0.2 mm/yr for the Southern hemisphere. Based on the 10 years of GRACE gravity field models (GRACE RL05), we can draw the conclusion that difference in the mean sea level rise between the Northern and the Southern hemispheres is reflected in the rate of the z-coordinate of the geocenter confirming the assertion (sign) of (Cazenave and Llovel 2010), i.e., that the mean sea level has been rising faster over the Southern than over the Northern hemisphere.

3.1 Interhemispheric Temperature Asymmetry and Ocean Mass Flux Between the Northern and Southern Hemispheres

The very first reaction when presenting the weekly z-coordinates of the geocenter in Table 2.1 arising from a combined orbit determination of Jason-2 and the GPS constellation of satellites (Svehla et al. 2010b), was that the source of the constant geocenter offset was in the inhomogeneous distribution of the SLR network between Northern and Southern hemispheres, (Pavlis, priv. com.). However, a closer look at Table 2.1 reveals a rate that is very much constant from week to week and, when extrapolated to the entire year, gives a rate of about 2 mm/yr to 4 mm/yr . However, this extrapolation is based on the very limited data set of the CONT'08 Campaign (about one month only). We should bear in mind that the relative dynamics of a LEO satellite and the GPS constellation is a new, unique tool, since in this case the orbit of the Jason-2 satellite is tied to the GPS constellation of satellites and not directly to the reference frame realized by the ground network, as is the case with DORIS satellites. From this point of view, the sensitivity of relative dynamics between a LEO satellite and GPS constellation is a completely new tool in the research of the system Earth and the estimation of the annual amplitude in the z-coordinate of the geocenter and secular rates. Any secular rate in the z-coordinate of the geocenter would indicate a secular rate in the mean sea lever rise between the

Northern and the Southern hemispheres. This analysis leads us to another idea, namely that rate in the z-coordinate of the geocenter could also be derived from the gravity maps provided by the GRACE mission.

(Cazenave and Llovel 2010) quantify the role of the thermal expansion of the oceans, land ice mass loss, and land water–storage change in the global sea-level rise measured by radar altimetry. Thermal expansion of the oceans and melting of the polar ice-sheets are the two main contributors to sea-level rise in general. Approximately one-third of the sea-level rise has been attributed to thermal expansion and two-thirds to the melting of the polar ice-sheets and mountain glaciers, (Cazenave and Llovel 2010). However, since 2003 acceleration in glacier melting and ice mass loss from the ice sheets has increased this to 80% (Cazenave and Llovel 2010), see also (Cazenave et al. 2009). The sea level variations due to anomalies in temperature and salinity, or so-called steric variations, are associated with the density or the volume of the water column.

Recently, (Friedman et al. 2013) showed that global warming is faster in the Northern hemisphere than in the Southern hemisphere, with some of the most rapid warming rates located in the Arctic regions of the Earth, where sea and land ice is rapidly thinning and shrinking faster than in Antarctica. (Friedman et al. 2013) introduce the so-called interhemispheric temperature asymmetry (ITA) as an emerging indicator of global climate change and report that the observed annual mean ITA (Northern minus Southern) has varied within an 0.8°C range over the last 100 years and has featured a significant positive trend since 1980. (Friedman et al. 2013) attribute this increase to the uneven spatial impacts of greenhouse forcing, which result in amplified warming in the Arctic and northern landmasses. This is largely because the Northern hemisphere has less ocean and more land than the Southern hemisphere, and oceans warm relatively slowly, (Friedman et al. 2013). Another consequence of the Northern hemisphere becoming warmer is the tendency tropical rainfall to extend northward. This means a northward extension of the wet season in sub-Saharan Africa and South America (Amazon) and an increase in extremes in the monsoon weather systems in Asia, see (Friedman et al. 2013). At the same time, (Luderer et al. 2013) point to global ocean currents as another factor confirming asymmetrical warming between Northern and Southern hemisphere. Global currents, such as the Gulf Stream, transport heat from the Southern hemisphere and into the Northern hemisphere, primarily to the North Pacific and North Atlantic.

All this implies that any asymmetric mass flux between the Northern and Southern hemispheres should be reflected in the z-coordinate of the geocenter. Following (Cazenave and Llovel 2010) satellite altimetry observations suggest that the mean sea level has been rising faster over the Southern than over the Northern Hemisphere. On the other hand, most altimetry satellites are placed at an inclination of approx. 66°, thus mainly mapping sea level rise in the mid-latitudes and equatorial waters. Although the altimetry orbits are symmetrical w.r.t. the equator, this is not the case for the amount water in the oceans, i.e., the Northern hemisphere has less ocean than the Southern hemisphere. Therefore, altimetry satellites measure sea level rise mainly in the southern waters and not globally.

Thus we can draw the conclusion that symmetries in the mass flux between the Northern and Southern hemisphere should be reflected in the gravity field maps from the GRACE mission and potentially also in the geocenter z-coordinates derived from the combined orbit determination of altimetry and GPS satellites.

3.2 The Geocenter Rate from Pear-Shaped Zonal Spherical Harmonics

Following (Heiskanen and Moritz 1967), degree one gravity field coefficients C_{10}, C_{11} and S_{11} are directly related to the center of mass coordinates (x, y, z) as the origin of the coordinate system by

$$C_{10} = \frac{z}{R_\oplus} \qquad C_{11} = \frac{x}{R_\oplus} \qquad S_{11} = \frac{y}{R_\oplus} \qquad (3.1)$$

Therefore, the rate in the translation of the geocenter z-coordinate could be related to the first degree gravity field coefficient C_{10} by

$$\frac{dz}{dt} = R_{\oplus} \frac{dC_{10}}{dt} \tag{3.2}$$

with $R_{\oplus}$ being the Earth's semi-major axis. Unfortunately, gravity field maps from the GRACE mission delivered on a routine basis by JPL, CSR and GFZ do not provide degree one harmonics of the spherical harmonic expansion of the Earth's gravity field. This would be a direct measure of this effect. However, any mass flux between the Northern and the Southern hemisphere should be reflected in the asymmetrical surface spherical harmonics (that are not symmetrical with the equator), such as odd zonal degree harmonics that depend only on geographical latitude. They are of odd degree and are asymmetric w.r.t. the equator, so-called "pear-shaped".

Let us now write Earth's gravitational potential V in the form of a spherical harmonic expansion as a function of the geocentric coordinates (r, θ, λ)

$$V(r, \theta, \lambda) = \frac{GM}{r} \sum_{n=0}^{\infty} \sum_{m=0}^{n} \frac{R_{\oplus}^n}{r^n} P_{nm}(\cos\theta) \Big(C_{nm} \cos m\lambda + S_{nm} \sin m\lambda\Big) \tag{3.3}$$

with the un-normalized spherical harmonic coefficients C_{nm} and S_{nm} and the $P_{nm}(\cos\theta)$ denoting the un-normalized associated Legendre polynomials of degree n and order m. Let us now write surface spherical harmonic $Y_{nm}(\theta, \lambda)$ in (3.3) in the complex form

$$Y_{nm}(\theta, \lambda) = P_{nm}(\cos\theta)e^{im\lambda} \tag{3.4}$$

that gives three forms of spherical harmonics: zonal, tesseral and sectorial harmonics

$$
\begin{array}{ccc}
\text{zonal } (m=0) & \text{tesseral } (m \neq n) & \text{sectorial } (m = n) \\[4pt]
 & \cos(m\lambda)P_{nm}(\cos\theta) & \cos(n\lambda)P_{nn}(\cos\theta) \\[2pt]
P_{nm}(\cos\theta) & \sin(m\lambda)P_{nm}(\cos\theta) & \sin(n\lambda)P_{nn}(\cos\theta)
\end{array} \tag{3.5}
$$

Considering that temporal variation of a spherical harmonic should be equivalent to the temporal variation of spherical harmonic coefficient itself, one can write the following relation for the translation along the z-direction for zonal harmonic with $m = 0$

$$\frac{dY_{n0}(\cos\theta)}{dt} = \frac{dY_{n0}(\cos\theta)}{dz}\frac{dz}{dt} = \frac{dC_{n0}}{dt} \qquad \rightarrow \qquad \frac{dz}{dt} = \sqrt{2n+1}\,\frac{d\overline{C}_{n0}}{dt}\left(\frac{dY_{n0}(\cos\theta_{mean})}{dz}\right)^{-1} \tag{3.6}$$

since for the vertically oriented zonal surface spherical harmonic we have $dY_{n0}(\cos\theta)/dt = dC_{n0}/dt$. This can also be seen if we scale surface spherical harmonic $Y_{nm}(\theta, \lambda)$ in (3.4) by spherical harmonic coefficients written in the complex form $K_{nm} = C_{nm} + S_{nm}i$

$$K_{nm}Y_{nm}(\theta, \lambda) = C_{nm}P_{nm}(\cos\theta)\cos m\lambda + S_{nm}P_{nm}(\cos\theta)\sin m\lambda \tag{3.7}$$

that for $m = 0$ gives

$$K_{n0}Y_{n0}(\theta) = C_{n0}P_{n0}(\cos\theta)\cos 0\lambda \tag{3.8}$$

The $\sqrt{2n+1}$ in (3.6) stands for the normalization factor. The mean derivative $dY_{n0}(\cos\theta_{mean})/dz$ can be calculated from the mean value theorem for integrals in the following way

$$\frac{dY_{n0}(\cos\theta_{mean})}{dz} = \frac{1}{2R_{\oplus}}\int_{-R_{\oplus}}^{R_{\oplus}} \frac{dY_{n0}(\cos\theta)}{dz}dz \tag{3.9}$$

Figure 3.1 Normalized odd zonal degree coefficients ("pear-shaped") from GRACE monthly gravity fields, RL05. One can clearly see an annual period and a very strong rate in all odd zonal degree coefficients up to degree $\bar{C}_{7,0}$. For higher degrees this rate is smaller and lost in noise. This consistent rate in the first odd zonal coefficients and the 10 times higher amplitude of $\bar{C}_{30}$ (compared to other odd coefficients) confirms the initial assumption that ocean mass flux between the Northern and the Southern hemispheres should be reflected in the asymmetric surface spherical harmonics and the observed rate of the z-coordinate of the geocenter from our combined reference frame solution with Jason-2 and GPS satellites.

Since $\cos\theta = z/R_\oplus$, for the first three odd spherical harmonics ("pear-shaped") we can derive

$$
\begin{aligned}
Y_{30}(\cos\theta) &= P_{30}(\cos\theta)\cos 0\lambda = \frac{5}{2}\frac{z^3}{R_\oplus^3} - \frac{3}{2}\frac{z}{R_\oplus} \\
Y_{50}(\cos\theta) &= P_{50}(\cos\theta)\cos 0\lambda = \frac{63}{8}\frac{z^5}{R_\oplus^5} - \frac{70}{8}\frac{z^3}{R_\oplus^3} + \frac{15}{8}\frac{z}{R_\oplus} \\
Y_{70}(\cos\theta) &= P_{70}(\cos\theta)\cos 0\lambda = \frac{429}{16}\frac{z^7}{R_\oplus^7} - \frac{693}{16}\frac{z^5}{R_\oplus^5} + \frac{315}{16}\frac{z^3}{R_\oplus^3} - \frac{35}{16}\frac{z}{R_\oplus}
\end{aligned}
\tag{3.10}
$$

In the next step we approximate annual periodic variations of odd zonal degree coefficients from monthly gravity field maps in Figure 3.1 with amplitudes A_c and A_s taking into account secular rates $\dot{\bar{C}}_{n0}$ and $\dot{\bar{S}}_{n0}$ relative to nominal epoch $t_0 = 2003.0$. Therefore, the adjusted model of temporal gravity field coefficients for $\bar{C}_{n0}(t)$ and $\bar{S}_{n0}(t)$ as a function of time is finally

$$
\begin{aligned}
\bar{C}_{n0}(t) &= \bar{C}_{n0}(t_0) + \dot{\bar{C}}_{n0}t + A_c\cos(\omega t + A_{0c}) \\
\bar{S}_{n0}(t) &= \bar{S}_{n0}(t_0) + \dot{\bar{S}}_{n0}t + A_s\cos(\omega t + A_{0s})
\end{aligned}
\tag{3.11}
$$

with

$$\bar{C}_{30} = (0.78 \pm 0.03) \cdot 10^{-10} \cos(\omega t)$$
$$\bar{C}_{50} = (1.04 \pm 0.01) \cdot 10^{-10} \cos(\omega t) \qquad (3.12)$$
$$\bar{C}_{70} = (0.29 \pm 0.01) \cdot 10^{-10} \cos(\omega t)$$

where time t is measured in days and ω denotes annual frequency $\omega = 2\pi \, / \, 365.25$.

$$\frac{dz}{dt}(\bar{C}_{30}) = R_{\oplus}\sqrt{7} \times (0.78 \cdot 10^{-10})\cos(\omega t) = 1.3 \text{ mm} \times \cos(\omega t) \quad \pm 0.05 \text{ mm}$$

$$\frac{dz}{dt}(\bar{C}_{50}) = R_{\oplus}\sqrt{11} \times (1.04 \cdot 10^{-10})\cos(\omega t) = 2.2 \text{ mm} \times \cos(\omega t) \quad \pm 0.03 \text{ mm} \qquad (3.13)$$

$$\frac{dz}{dt}(\bar{C}_{70}) = R_{\oplus}\sqrt{15} \times (0.29 \cdot 10^{-10})\cos(\omega t) = 0.7 \text{ mm} \times \cos(\omega t) \quad \pm 0.03 \text{ mm}$$

Considering that the pear-shaped term $\bar{C}_{30}$ is 10 times larger than other low-order odd coefficients, for the period of the CONT'08 Campaign (August 2008) we obtain $dz/dt(\bar{C}_{30}) = -0.1 \text{ mm/2 weeks}$, whereas for the geocenter rate from LEO/GPS combination we obtain $dz/dt = -0.15 \text{ mm/2 weeks}$. Note that the data set of the CONT'08 Campaign (August 2008) is limited to three weeks only. Our annual amplitude of 2.4 mm is consistent with the $\approx 4 \text{ mm}$ annual amplitude of SLR z-origin values as given for the same period in (Pavlis 2012) using 30-day boxcar smoothing of SLR geocenter values (but significantly noisier). This value is also consistent with the annual amplitude of the mean see level variations of about 2.5 mm due to geocenter variations as reported in (Pavlis 2012).

Let us now look at the secular rate of the odd zonal degree coefficients in Figure 3.1 based on GRACE monthly gravity fields. For

$$\dot{\bar{C}}_{30} = (-0.60 \pm 0.07) \cdot 10^{-10}/10 \text{ yr}$$
$$\dot{\bar{C}}_{50} = (-0.62 \pm 0.03) \cdot 10^{-10}/10 \text{ yr} \qquad (3.14)$$
$$\dot{\bar{C}}_{70} = (-0.31 \pm 0.03) \cdot 10^{-10}/10 \text{ yr}$$

we obtain the following geocenter translation rate

$$\frac{dz}{dt}(\bar{C}_{30}) = R_{\oplus}\sqrt{7} \times (-0.60 \cdot 10^{-10}) \, / \, 10 \text{ yr} = -1.0 \text{ mm/10 yr} \quad \pm 0.13 \text{ mm/10 yr}$$

$$\frac{dz}{dt}(\bar{C}_{50}) = R_{\oplus}\sqrt{11} \times (-0.62 \cdot 10^{-10}) \, / \, 10 \text{ yr} = -1.3 \text{ mm/10 yr} \quad \pm 0.07 \text{ mm/10 yr} \qquad (3.15)$$

$$\frac{dz}{dt}(\bar{C}_{70}) = R_{\oplus}\sqrt{15} \times (-0.31 \cdot 10^{-10}) \, / \, 10 \text{ yr} = -0.8 \text{ mm/10 yr} \quad \pm 0.07 \text{ mm/10 yr}$$

In all three cases we see a rate of the geocenter z-coordinate of -2.4 mm/10 yr, consistent with the rate of 0.232 mm/ yr in the mean sea level due to the SLR geocenter rate in ITRF2005, as reported in (Pavlis 2012). Therefore, we can draw the conclusion that the sea level rise of about 3.1 mm/yr can be explained by the asymmetric water mass flux in the oceans between Northern and Southern hemispheres giving a rate of -2.4 mm/10 yr for the z-coordinate of geocenter.

It is very interesting that geocenter translation rates derived from temporal gravity field maps are very close to the secular reference frame translation rates derived from the tide-gauge records, and from GPS and SLR solutions. During the review of this thesis, (Santamaría-Gómez et al. 2014) reported a secular rate in the z-coordinate of the geocenter of 0.29 mm/yr as derived from GPS and 0.18 mm/yr from SLR data over the last 20 years against the ITRF2008 frame. As expected, the tide gauge records over the last 100 years gave a higher value of 0.9 mm/yr (corrected for GIA, see (Santamaría-Gómez et al. 2014)). This smaller value is in line with the anticipated hemispherical pattern of sea level rise over the last several decades, indicating an increased rate of rise in the mean sea level compared to the longer 100 year period (by 50%).

Origin rate in z-axis	GPS last 20 years	SLR last 20 years	Tide-gauges last 100 years	GRACE Gravity Maps last 10 years
(Santamaría-Gómez et al. 2014)	0.29 mm/yr	0.18 mm/yr	0.90 mm/yr	-
This work, 2013	-	-	-	-1.0 mm/10yr ± 0.13 mm

Annual Amplitude in z-axis	SLR last 20 years	GRACE Gravity Maps last 10 years
(Pavlis 2012)	2.5 mm	-
This work, 2013	-	1.3 mm ± 0.05 mm

Table 3.1 Reported rate in the z-coordinate of the geocenter from the global GPS and SLR solutions for the period of the last 20 years compared to tide-gauges records (GIA corrected) over the last 100 years.

Here, for the first time, we have derived a similar reference frame rate from the analysis of the GRACE gravity field maps. The tide-gauge records over the last 100 years gave, as expected, a higher value compared to the values derived from the orbit determination of GPS and SLR satellites and GRACE gravity field models. The reference frame rate obtained from GRACE gravity field maps is very similar to that from GPS and SLR solutions with, however, an opposite sign. Bottom part of the table shows annual amplitude in the z-coordinate of the geocenter as derived from the GRACE gravity models over the last 10 years and shows a very good agreement with the amplitude derived from the SLR reference frame solution.

3.3 Rate in the Even-Degree Zonal Spherical Harmonics as a Measure of Sea Level Rise and Intrinsic Scale of the Reference Frame

The central term of the gravity field $\overline{C}_{00}$ of the spherical harmonic expansion defines the mean gravitational potential of the Earth. In the case of homogeneous sea level rise over all oceans it is expected that only zonal surface spherical harmonics will be affected since they are symmetrical w.r.t. the equator. The mean gravitational potential as well as the shape of the oblate ellipsoid will not be changed under this assumption. Thus, one could expect a scale type effect that will be reflected in a change of the mean sphere in the expansion of the Earth's gravity field in terms of spherical harmonics. The derivative of the radius $R_{\oplus}$ of the mean sphere of the spherical harmonic expansion can be calculated from the mean value theorem for integrals in the following way

$$\frac{dY_{n0}(\cos\theta_{mean})}{dR_{\oplus}} = \frac{2}{R_{\oplus}} \int_0^{R_{\oplus}} \frac{dY_{n0}(\cos\theta)}{dR_{\oplus}} dz \tag{3.16}$$

Since for the first odd zonal spherical harmonics

$$Y_{20}(\cos\theta) = P_{20}(\cos\theta)\cos 0\lambda = \frac{3}{2}\frac{z^2}{R_\oplus^2} - \frac{1}{2}$$

$$Y_{40}(\cos\theta) = P_{60}(\cos\theta)\cos 0\lambda = \frac{35}{8}\frac{z^4}{R_\oplus^4} - \frac{30}{8}\frac{z^2}{R_\oplus^2} + \frac{3}{8} \tag{3.17}$$

$$Y_{60}(\cos\theta) = P_{60}(\cos\theta)\cos 0\lambda = \frac{231}{16}\frac{z^6}{R_\oplus^6} - \frac{315}{16}\frac{z^4}{R_\oplus^4} + \frac{105}{16}\frac{z^2}{R_\oplus^2} - \frac{5}{16}$$

we obtain the rate in the scale of the geometrical frame that defines expansion of Earth's gravitational field in terms of spherical harmonics

$$\frac{dR_\oplus}{dt}(\bar{C}_{20}) = R_\oplus\sqrt{5}\times(-1.7\cdot10^{-10})\,/\,10\text{ yr} = -2.4\text{ mm/10 yr}$$

$$\frac{dR_\oplus}{dt}(\bar{C}_{40}) = R_\oplus\sqrt{9}\times(-1.0\cdot10^{-10})\,/\,10\text{ yr} = -1.9\text{ mm/10 yr} \tag{3.18}$$

$$\frac{dR_\oplus}{dt}(\bar{C}_{60}) = R_\oplus\sqrt{13}\times(-0.9\cdot10^{-10})\,/\,10\text{ yr} = -1.7\text{ mm/10 yr}$$

or the relative rate in the scale of -0.5 ppb/10 yr. The scale of the conventional terrestrial frame is defined by the scale of the station coordinates of the ground networks of space geodesy techniques fixed to the continental Earth's crust. Here we show that spherical harmonics also contain an intrinsic scale and one can use temporal gravity field maps to monitor its variations over time. This scale does not influence the mean gravitational potential nor the shape of the oblate ellipsoid, but rather defines the scale of the background geometrical reference frame that defines the expansion of spherical harmonics. Eq. (3.18) shows that this geometrical scale can be monitored by temporal gravity field maps. Since the radius of the mean sphere approximates the global mean see level, a constant rise of the mean sea level will be reflected in the rate of the estimated even degree zonal gravity field coefficients or equivalently in the radius of the mean sphere used in the expansion of spherical harmonics.

3.4 Is There a Secular Rate in the Gravitational Constant?

Is there any secular rate in the gravitational constant? Can we see it from the GRACE gravity field maps? Over the last 10 years, there have been many discussions arguing that the gravitational constant, as well as other constants of the Standard Model used in physics do experience variations over time, most likely in terms of secular rates. Since the gravitational constant is not estimated as a parameter of the GRACE mission it will be reflected in the estimated low-order gravity field coefficients, i.e., the rate in the gravitational constant G is equivalent to the rate of the first few low-degree harmonics. Therefore, we may establish the following relation

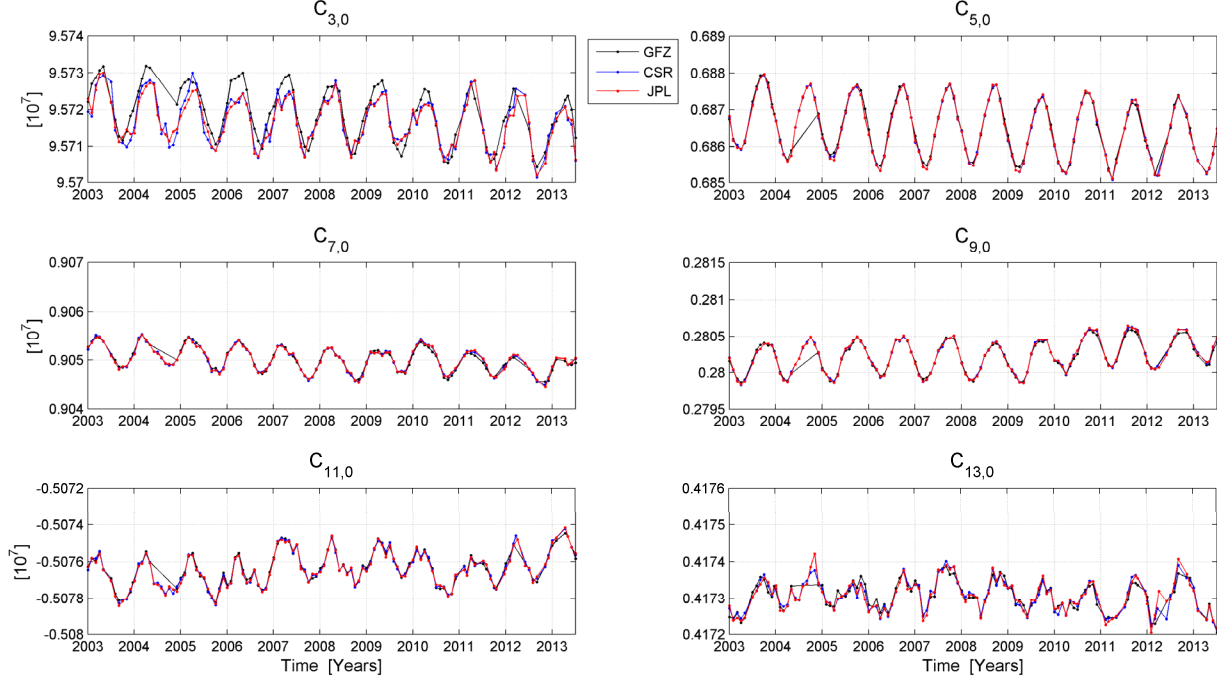$$\frac{dG}{dt} \sim \frac{dC_{n0}}{dt} \sim \frac{dS_{n0}}{dt} \tag{3.19}$$
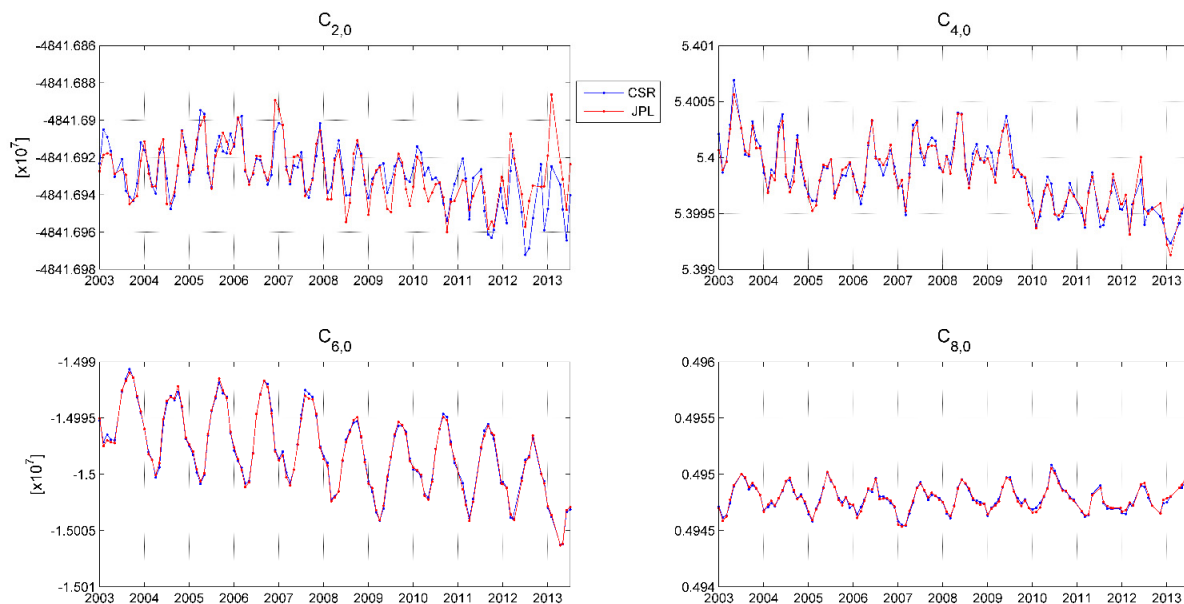
Figure 3.2 Normalized even zonal degree coefficients from GRACE monthly gravity fields, RL05. One can clearly see an annual period and a very strong rate in all even zonal degree coefficients up to degree $\bar{C}_{6,0}$ of the order of $d\bar{C}_{n0}/dt = -0.5 \cdot 10^{-10}/10$ yr . For $\bar{C}_{2,0}$ the secular rate is higher by about one order of magnitude.

Figure 3.1 shows a very uniform rate in the first few low-degree odd coefficients of the Earth's gravitational field as provided by the GRACE missions over the last 10 years. One can see this uniform rate especially for the pear-shaped coefficients $\bar{C}_{30}$ and $\bar{C}_{50}$, c.f. $d\bar{C}_{30}/dt \approx d\bar{C}_{50}/dt \approx -1 \cdot 10^{-10}/10$ yr . Figure 3.2 confirms a similar rate also for the first gravity field coefficients of the even-degree in the order of $d\bar{C}_{n0}/dt = -0.5 \cdot 10^{-10}/10$ yr . For $\bar{C}_{2,0}$, the secular rate is larger by about one order of magnitude. Recently, Kasevich's Group at Stanford has reported the measurement of the gravitational constant using a gravity gradiometer based on atomic interferometry, see (Fixler et al. 2007). The gradiometer measures the differential acceleration of two samples of laser-cooled Cs atoms. They reported a standard error of the estimated gravitational constant of $\pm 0.027 \times 10^{-11}$ m^{3}kg^{-1}s^{-2} and a systematic error of $\pm 0.021 \times 10^{-11}$ m^{3}kg^{-1}s^{-2} . The standard deviation of this measurement is about two to three orders of magnitude lower than that of the rates of low-degree gravity field coefficients one can see in Figure 3.1 and Figure 3.2.

On the other hand, we do have a possible geophysical explanation for the temporal rates in the low degree coefficients, especially for $\bar{C}_{2,0}$ and the "pear-shaped" term $\bar{C}_{3,0}$ that are also visible in the reference frame parameters. Therefore, we may draw the conclusion that a secular rate in the gravitational constant is not yet visible in the GRACE gravity field maps and, if it exists, is most likely hidden in the geophysical signal and aliasing effects. One should also bear in mind that the gravitational force is extremely weak when compared to other fundamental forces, i.e., 39 orders of magnitude weaker than the electromagnetic force. This can easily be shown by using the mass of the proton and the electron and Newton's universal law of gravitation to calculate the gravitational force and then comparing that with the electromagnetic force calculated between the same two particles. Here we did not consider any changes in the total mass of the Earth M and the exchange of material between the Earth's atmosphere and outer space.

4. First Phase Clocks and Frequency Transfer

In (Švehla and Rothacher 2004a), (Švehla and Rothacher 2004b), (Švehla and Rothacher 2005a), and in (Švehla and Rothacher 2006b) it was demonstrated for the first time that clock parameters for GPS satellites and ground stations can be estimated solely from the carrier-phase GPS measurements. These also allow frequency transfer with a very high level of accuracy of a few parts in 10^{-16} (≈ 25 ps/day in terms of linear time rate). The main motivation for the development of the phase clock approach is to avoid the colored systematic noise that is introduced by using code, or smoothed code GPS measurements and other possible biases in the official GNSS clock parameters provided by IGS. On the other hand, phase clocks completely absorb the GPS radial orbit error and are fully consistent with the LEO carrier-phase measurements when determining kinematic or reduced-dynamic LEO orbits, since in both cases carrier-phase ambiguities are estimated. Phase measurements from a GPS ground network of about 40-50 stations tracking about 30 GPS satellites in MEO orbit form a closed, internally connected system, in which the phase information of one clock can be related to that of any other GPS satellite or a ground station clock in the network, even on the antipodal side of the world. This opens up the possibility of high-precision positioning and especially intercontinental non-common view frequency transfer of utmost accuracy. We may say, phase clocks are the optimal way to compare phase information between ground station clocks and/or LEO/GNSS satellites. Later on in this thesis, we introduce the concept of track-to-track ambiguities to optimally fix carrier-phase ambiguities to their integer values.

Later, phase clocks were also studied in (Dach et al. 2005), (Bauch et al. 2006), (Dach et al. 2006) and in (Matsakis et al. 2006) over longer periods of time and have been compared to other time/frequency comparison techniques. Ambiguity resolution with phase clocks was demonstrated for the first time in (Švehla and Rothacher 2006a) and later on in (Mercier and Laurichesse 2007), (Delporte et al. 2007), (Delporte et al. 2008). Starting with GPS Week 1449, JPL started providing additional information on clock time bias and drift relative to the reference clock in the IGS network in their IGS reports, see (Desai 2007). In their IGS reports, as a reference clock JPL uses exclusively IGS station USN3 (US Naval Observatory), or in some cases AMC2 (Colorado Springs). Besides CNES, all IGS Analysis Centers provide satellite clock parameters calculated using carrier-phase and pseudo-range measurements in order to support both time and frequency transfer at the same time. Thus, IGS clock parameters are more applicable to PPP (Precise Point Positioning) than to frequency transfer. This section describes the estimation of phase clocks and their application in frequency transfer and precise point positioning.

4.1 The Concept of Phase Clocks

Phase clocks are biased clock parameters preserving the highly accurate relative epoch-to-epoch information of carrier-phase measurements. When carrier-phase is connected over all ground stations and all GPS satellites, any time bias and drift in the selected reference clock biases all other clocks in the network by exactly the same amount. Ground stations do not have to be connected to a stable frequency standard such as H-maser

or a clock assembly in a timing lab. There is a minimum number of about 40 ground stations needed to form a connected system with continuous carrier-phase information between all GPS satellites and ground stations, see Figure 4.1. We demonstrated for the first time that the frequency transfer between the best timing labs in the IGS network is possible with a precision of below 25 ps/day (few parts in 10^{-16}/day) (Švehla and Rothacher 2004a), (Švehla and Rothacher 2004b), (Švehla and Rothacher 2005a), and in (Švehla and Rothacher 2006b). Using carrier-phase data only, the impact of the pseudorange noise and accompanying systematic effects can be avoided. Code measurements are needed only to pre-synchronize all receiver clocks at the level
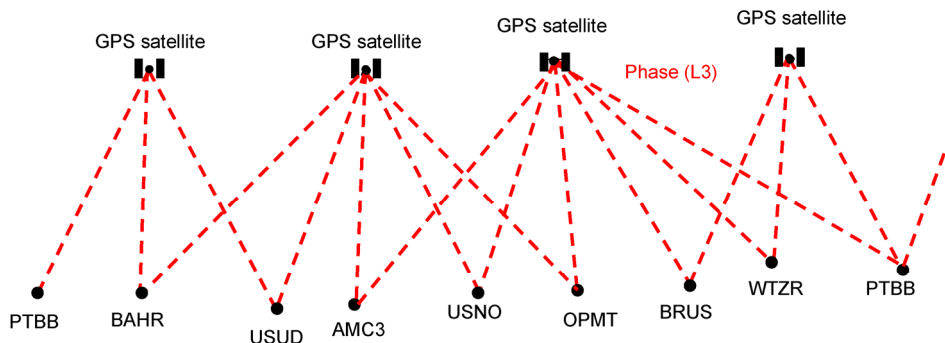
Figure 4.1 The concept of phase clocks. Phase information is connected and transferred between all ground stations and all GPS satellites.

of about $1\,\mu s$. Due to the low accuracy of code measurements, phase clocks can be aligned in an absolute time frame to about $1\,ns$. Increasing the number of ground stations increases the overall number of stations that contribute to the clock parameters of one GPS satellite at a given time. In this way, local effects such as multipath and other station-specific environmental errors are averaged out over a number of ground stations providing extremely precise and consistent phase information for the GPS satellites. GPS satellites are placed at high altitude in the MEO orbit and any radial orbit error can be fully represented by the estimated satellite clock parameters. This is the reason why GPS orbit errors do not propagate into frequency transfer between ground stations or LEO satellites, and allow for extremely accurate precise point positioning and orbit determination of LEO satellites. We may say, phase clocks are the optimal way to compare clock information between ground GPS stations and/or LEO satellites. Usually, one well-performing H-maser in the IGS network or timing lab is selected as the reference clock in the system, and any epoch-specific bias in the ensemble of such ground/space phase clocks will be removed when differences between different stations are formed. All common errors between ground stations will be removed as well, such as common troposphere and tidal errors. This opens up the possibility of extracting extremely accurate frequency information on two ground clocks a great distance apart or to study the frequency stability of clocks on board GPS or LEO satellites However, in the case of the clock parameters of GPS (and LEO) satellites, orbit errors will propagate into the estimated clock parameters, but due to the nature of satellite orbits, these orbit errors will average out over one or several orbit revolutions. Thus very accurate frequency offsets (time rate) can be calculated between ground-to-space or ground-to-ground despite the orbit determination errors of GPS (and LEO) satellites. This is especially true for the orbit errors of GNSS satellites that typically have once-per-revolution pattern.

4.2 Estimation of Phase Clocks

Figure 4.2 shows the ground network of IGS stations used for the calculation of phase clocks. It is a network of about 40-50 ground stations uniformly distributed over the globe. All stations are part of the IGS network used in the IGS Reprocessing Project running at TU München. In order to maintain consistency with the software, station coordinates, GPS satellite orbits and Earth rotation parameters were kindly provided by the IGS Reprocessing Project (Steigenberger et al. 2006). The disadvantage of these products is that many timing
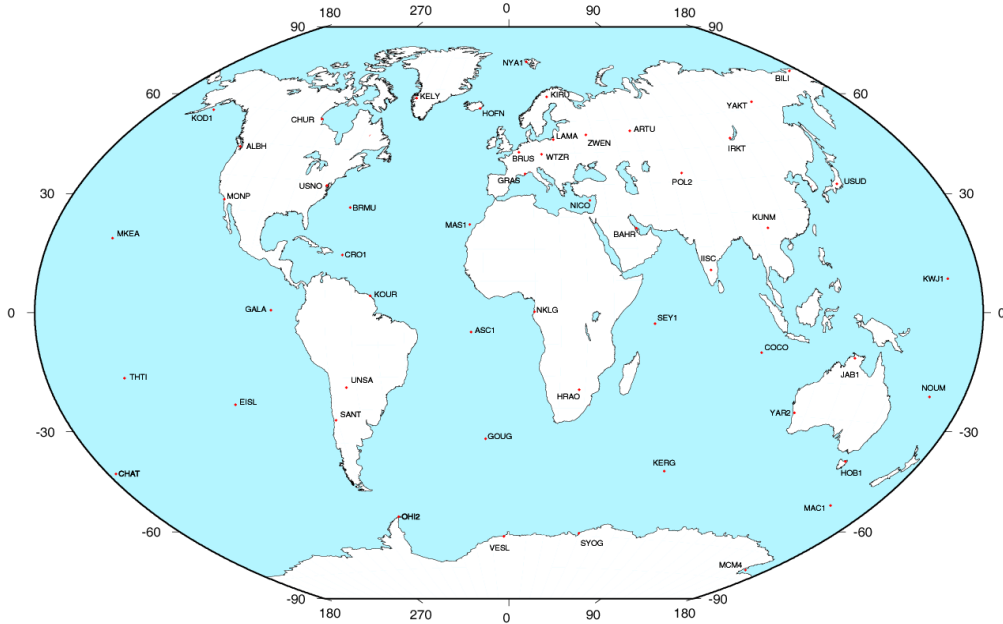
Figure 4.2 Ground GPS stations used for determination of phase GPS satellite clocks.

labs within the IGS network are not included in the IGS Reprocessing Project, mainly due to poor station monumentation or simply a lack of continuous tracking over many years. Figure 4.3 shows the procedure used to estimate phase clocks for GPS satellites and ground stations. In the first step, broadcast GPS satellite clock information is used to align the selected ground reference clock to GPS time. As the reference clock, the most stable H-maser is selected, such as the one available from the timing lab US Naval Observatory (USNO) or a geodetic IGS station connected to a local H-maser, e.g., Wettzell in Germany. In this alignment step time bias and drift are estimated for the reference H-maser using smoothed ionosphere-free pseudorange measurements. In this calculation, all broadcast GPS satellite clock parameters are held fixed. For all other ground GPS stations, an a priori clock synchronization to broadcast GPS time is performed by estimating one clock parameter every epoch using ionosphere-free code measurements. This step is required since GPS measurements are given in the GPS receiver time, which could differ from GPS time by up to a millisecond. Screening of code and phase measurements is based on the Melbourne-Wübbena linear combination. In the next step, the parameters of the aligned reference clock from the previous step are held fixed and all other GPS satellite/station clock parameters are estimated with a resolution of 5 minutes using smoothed-code ionosphere-free measurements. Once the first solution for GPS satellite clocks is available, it is used to pre-process carrier-phase measurements, i.e., to detect cycle-slips and outliers. Once the phase data have been screened, the clock estimation is repeated for all GPS satellites and ground stations without using any pseudorange measurements. The clock solution in this step is calculated with a resolution of 30 s . This procedure is repeated in order to further screen the phase data. For a 1-day arc, GPS satellite/ground station clocks can be estimated with a sampling of 30 s using the full normal equation system consisting of phase ambiguities and GPS satellite/receiver clocks as parameters only. With 45 ground stations we may easily expect up to 5000 ambiguities and this can easily be handled on a standard Linux PC system. The NEQ matrix contains only phase ambiguities (up to 5000) since all GPS satellite and ground clock parameters are pre-eliminated every epoch. Once the normal equation system is inverted, phase ambiguities are back-substituted and a normal equation matrix is set-up and inverted every epoch containing only clock parameters for about 30 GPS satellites and 40-50 ground stations.

By calculating high-rate GPS satellite phase clocks and CHAMP kinematic and reduced-dynamic orbits for a period of 2 years (Švehla and Rothacher 2004a), we demonstrated that such an approach can easily be
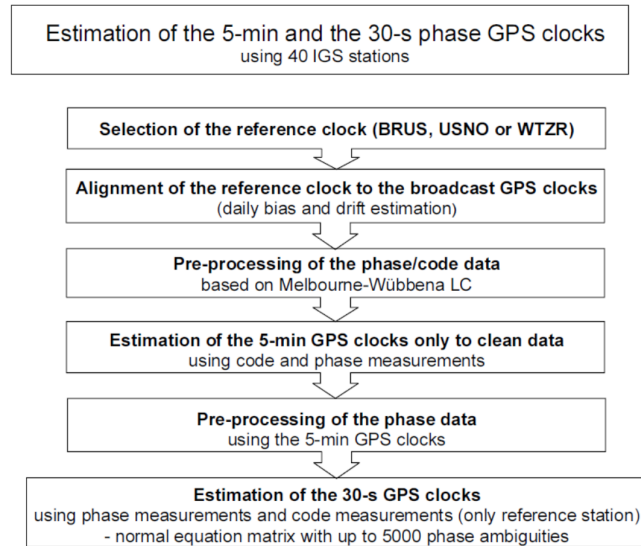
Figure 4.3 Overview of phase clocks calculation.

performed on a standard PC with 1 GB of RAM. The high-rate 30 s phase clock solution was based on about 40-50 ground IGS stations and one ground hydrogen maser as a fixed clock reference.

4.3 Frequency Transfer Based on Phase Clocks

In (Švehla and Rothacher 2004a), (Švehla and Rothacher 2004b), (Švehla and Rothacher 2005a), and in (Švehla and Rothacher 2006b) it was demonstrated for the first time that clock parameters estimated for ground stations allow frequency transfer with few parts in 10^{-16} (≈ 25 ps/day in terms of linear time rate). Later on those results were repeated by (Dach et al. 2005), (Bauch et al. 2006), (Dach et al. 2006) and (Matsakis et al. 2006). Figure 4.4 shows the differences in phase clocks between AMC2 and USNO for a period of one day. After removing a linear time drift (top figure) we obtained a residual clock noise with a standard deviation of ≈ 25 ps over one day (bottom figure) with results below 10 ps for the best days.
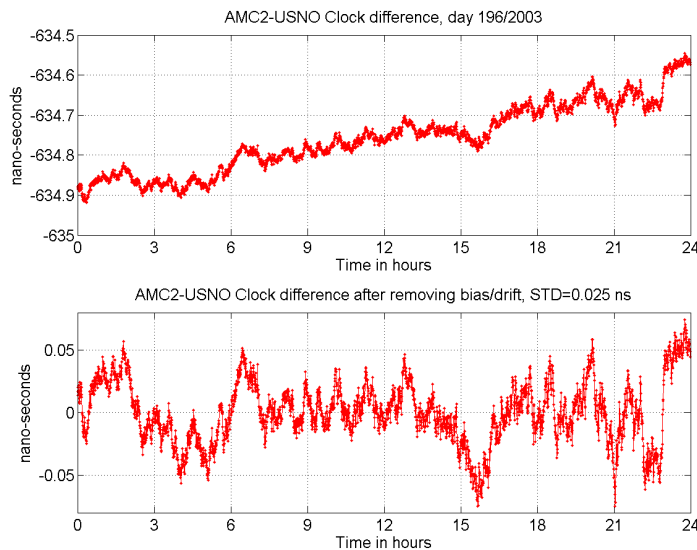
Figure 4.4 First high-precision frequency transfer using GPS with $\mathrm{STD} = 25$ ps, day 196/2003.
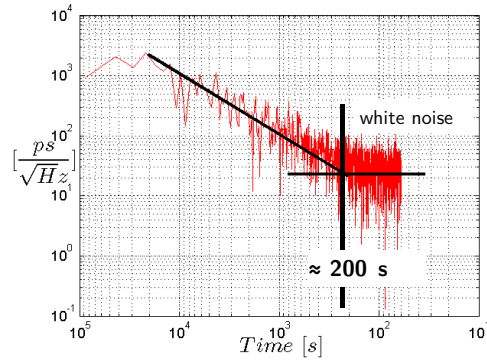
Figure 4.5 Power spectral density of phase clock differences between AMC2 and USNO, 196/2003.

Considering the short-term stability of H-masers in the IGS network, the residual systematic pattern in Figure 4.4 (bottom) is most likely mainly due to the modeling of troposphere delay, i.e., estimation of the troposphere zenith delays that, in this case, were estimated every hour as a piece-wise constant function. Figure 4.5 shows the power spectral density of the residual phase clock parameters between AMC2 and USNO given in Figure 4.4 (bottom). Comparing Figure 4.4 (bottom) and Figure 4.5, one can see that phase clocks show white noise up to 200 s, whereas flicker noise from 200 s - 24 h. It should be noted that the residual clock parameters in Figure 4.4 (bottom) include noise of the H-maser at both AMC2 and USNO as well as effects from the GPS data, including residual troposphere effects, signal multipath, station coordinates (residual atmospheric effects and tides), antenna phase center variations, antenna cable delays, in addition to GPS receiver effects (e.g., front-end). Here, the role of GPS orbit errors is significantly reduced since any radial orbit error is compensated for by the GPS satellite clock parameters being averaged over many ground stations. Thus the difference between the phase clocks of two separate timing labs is free of any common biases (to a great extent), including an overall common time offset. This is true for all estimated phase clocks, which only give relative time information between all clocks in the network.

GPS satellite clock parameters are not as smooth as ground receiver clock parameters derived from an external frequency such as a H-maser. Figure 4.6 shows the phase clocks of the GPS satellite PRN24. Compared to AMC2 or USNO, the noise level is about 120 ps after removing a low-order polynomial (quadratic term). The remaining clock residuals show a periodic pattern with a period of about 6 h (see Figure 4.6.) that can be explained by the periodic relativistic correction due to the J_2 gravity field coefficient and a variable semi-major axis, following the model presented in (Kouba 2004). The green line in Figure 4.6 shows the remaining periodic relativistic correction (Kouba 2004), and the black line represents a periodic signal with a period of 6 hours fitted to the phase clock parameters. One can see very good agreement between the model and the phase clocks of GPS PRN24. It should be noted that in GPS data processing only relativistic satellite clock correction due to the eccentricity of the GPS satellite orbit has been applied.
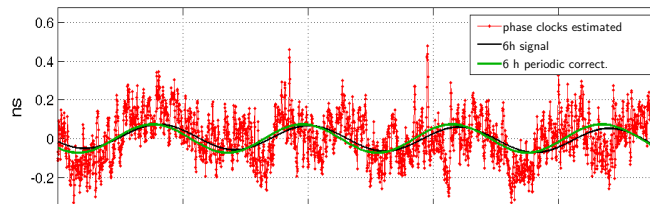
Figure 4.6 Phase clock parameters of GPS PRN 24 over a period of 24 h. The green line shows remaining periodic relativistic correction from (Kouba 2004), and the black line represents a periodic signal with a period of 6 hours fitted to the phase clock parameters.

4.4 Inter-Frequency and Inter-Channel Biases

Inter-frequency biases can be considered as a delay on the L_2 frequency measurements with respect to the measurements on the L_1 frequency. They are caused by hardware delays in the L_1 and L_2 signal paths and are mostly temperature-dependent. Inter-channel biases are differences in the signal path between the different receiver channels that track the GPS satellites. This effect is very difficult to estimate as it is receiver-specific and has not yet been estimated in global IGS processing. Inter-channel biases can be determined by calibration procedures and, when correctly applied, should not present a problem in the processing of GPS data. Inter-frequency biases can be eliminated when one clock parameter is estimated every epoch. This is not the case for inter-channel biases, whose constant parts can be eliminated by estimating phase ambiguities. Calibration for inter-channel biases can in fact be performed in the GPS receiver on the ground or in space, tracking the same GPS satellite on all channels.
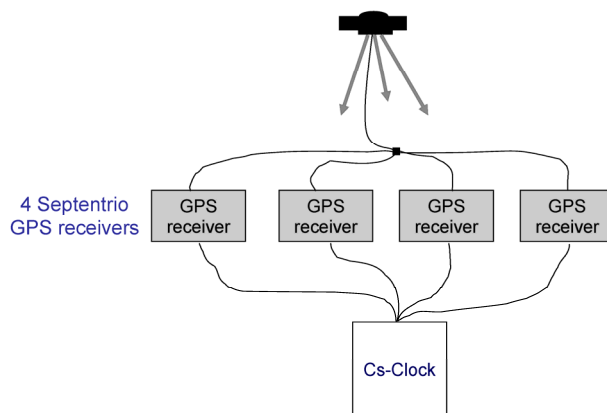
Figure 4.7 Test set-up with four Septentrio PolaRx2 GPS receivers.

In order to assess potential GPS receiver errors, we connected four Septentrio PolaRx2 GPS receivers to the same GPS antenna and external frequency, see Figure 4.7. This experiment was carried out in cooperation with the Institute of Navigation and Communication at DLR (Švehla et al. 2006a). Figure 4.8 shows very large variations on L_2 carrier-phase for those four 4 receivers denoted as UTC1, UTC2, GRX1 and GRX2. When forming single-differences between those four GPS receivers, w.r.t. the receiver denoted as UTC1, all signal propagation and receiver-specific effects should be removed, and phase noise is then the remaining effect. It is interesting that L_1 carrier-phase is not affected by the apparent clock variation that one can clearly see on L_2 data. Although all four GPS receivers are identical, residual multipath mitigation effects between the same receivers could also play a role. Following (Petit, priv. com.), such an effect could also be caused by the antenna cable splitters. A similar effect can be seen in Figure 4.9, comparing the carrier-phase from GPS PRN5 against PRN30 tracked by two GPS receivers, UTC1 and UTC2. Figure 4.9 indicates that the carrier-phase on L_2 shows some form of inter-channel or inter-satellite phase variations. Following (Simsky, priv. com.) these effects are most likely caused by the GPS receiver front-end.

Figure 4.10 shows code and carrier-phase measurements from four Septentrio GPS receivers connected to the same GPS antenna and the same external frequency standard over 10 days, (day 160-170/2006). For the code measurements, one can clearly see antenna cable delays of up to 10 m between different receivers and for the carrier-phase measurements a dominant periodic effect, most likely caused by the residual multipath effect (between receivers of the same type) or receiver front-ends, especially on the second GPS frequency (Simsky, priv. com.).
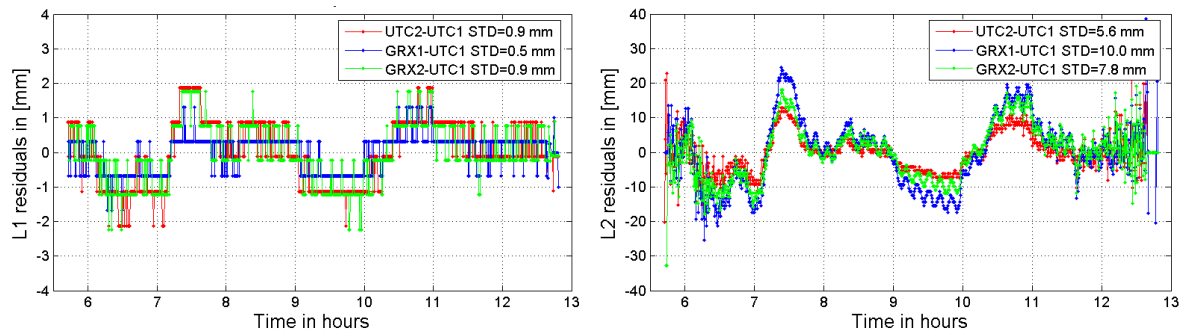
Figure 4.8 L_1 and L_2 carrier-phase of the GPS satellite PRN30 from four Septentrio receivers connected to the same antenna and an external frequency, (day 160/2006). Large variations on the L_2 carrier-phase are most likely due to delays in the receiver front-end (Simsky, priv. com.).
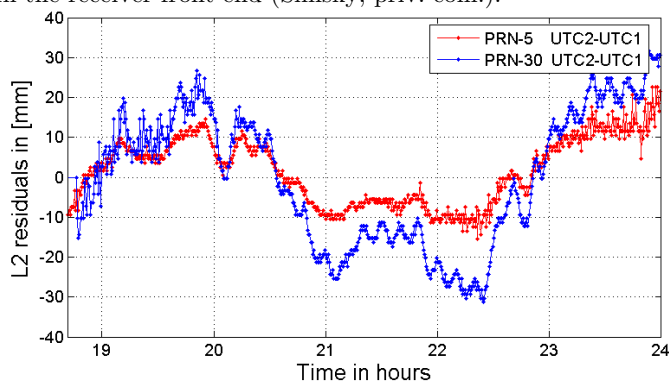
Figure 4.9 L_2 carrier-phase variations between two GPS satellites (PRN30 and PRN5) tracked by two identical GPS receivers connected to the same antenna and the same external frequency, (day 160/2006).
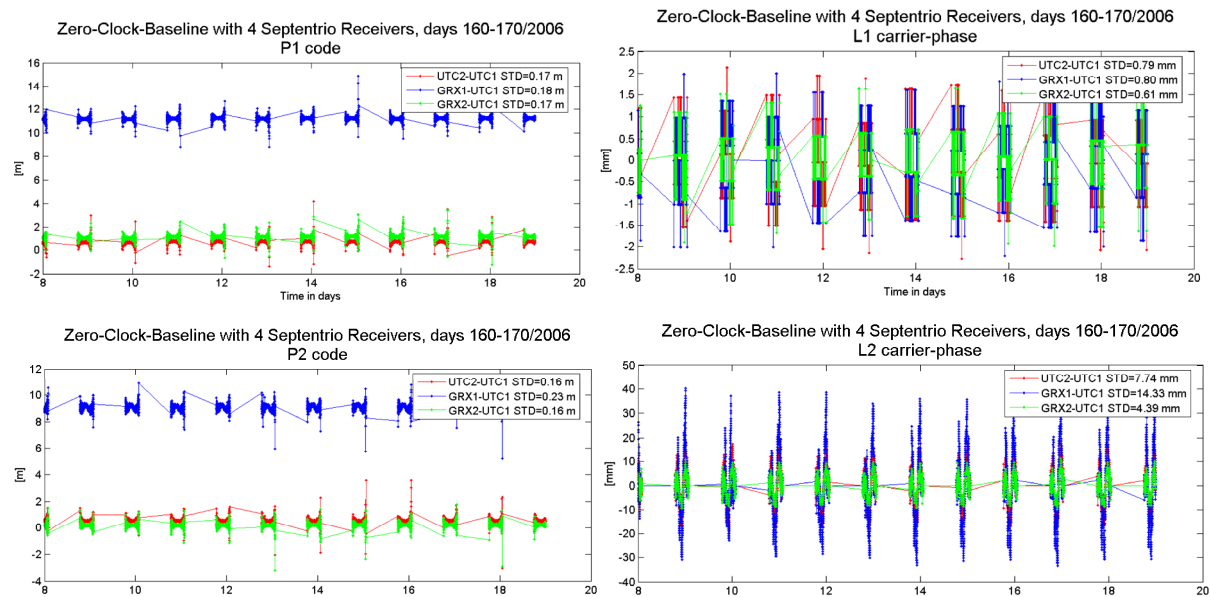
Figure 4.10 P_1 and P_2 code (left) and carrier-phase L_1 and L_2 (right) from 4 Septentrio GPS receivers connected to the same antenna and an external frequency standard over 10 days, (day 160-170/2006). One can clearly see antenna cable delays in code measurements between different receivers and periodic effects in carrier-phase (GPS satellite PRN30).

45

5. First Geometric POD of GPS and Galileo Satellites

W̌e have already estimated purely geometric orbits of several LEO satellites, and now one may ask how accurately a GPS satellite orbit can be estimated purely geometrically, i.e., kinematically. The main problem is that GNSS satellites are high above the Earth and positioning geometry is not as good as for satellites in LEO orbit. This section deals with the first estimation of one GPS satellite fully geometrically. New Galileo satellites are equipped with H-masers and in this case the satellite clock can be modeled very efficiently using a linear model over one day. We present here the first Galileo orbits estimated geometrically using a linear model for the H-maser on board the GIOVE-B satellite. The current accuracy of geometric GPS orbits is approximately $15\,\text{cm}$, whereas this improves to several centimeters in the case of Galileo. On the other hand, with Galileo, ambiguity resolution on the zero-difference level will be significantly improved, thus once the phase ambiguities are fixed, it is assumed that it will be feasible to estimate GNSS orbits fully geometrically with an accuracy comparable to dynamic orbits. For more on geometric POD of GNSS satellites see (Švehla and Rothacher 2005a).

5.1 The First Geometric Positioning of a GPS Satellite

The basic idea is to fix the coordinates of the IGS GPS points on the ground and to estimate three coordinates of the center-of-mass of the GPS satellite every epoch using zero- or double-difference phase measurements. The main difference to kinematic positioning of a ground station or a LEO satellite is that, due to the very high altitude, the GPS satellites "see" all ground stations within a very small range of nadir angles. A GPS antenna placed on a LEO satellite or located on the ground can receive signals from the GPS satellites at elevations ranging from $0°$ to $90°$. In contrast, the maximum nadir angle of a signal transmitted from a GPS satellite to a LEO satellite or ground station is about $14° - 15°$, see Figure 5.1. This angle is six times smaller than the maximum zenith angle of a LEO or ground GPS antenna and thus, the position of the ground stations in the local orbital system of the GPS satellite varies very little with time.

In the case of a LEO or a ground GPS station the kinematic positions are computed at the measurement epoch, which is the same for all GPS satellites tracked. This is not the case for the kinematic positioning of GPS satellites where, due to the GPS receiver clock correction and the light-travel time correction, different ground GPS stations "see" the GPS satellite at different places along its orbit for nominally the same observation epoch.

Due to the instability of the GPS receiver clock, the GPS measurements are not taken exactly at the integer second in GPS time. Steering of the GPS receiver clock on the ground or on the LEO satellite can be performed using the receiver's navigation solution based solely on the code measurements and broadcast GPS orbits and clocks. In the case of the Blackjack GPS receiver onboard the CHAMP satellite, the clock steering is performed to a precision of $0.1\,\mu s$. Nevertheless, for some ground GPS receivers (IGS network) the clock
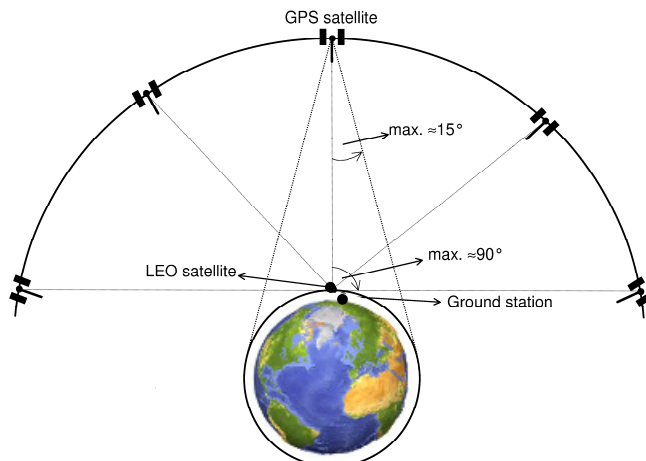
Figure 5.1 Geometry for LEO and GPS satellites and GPS station on the ground.

correction w.r.t. GPS time may vary by up to $1\,\mathrm{ms}$. In order to correct for this GPS receiver effect, aiming at an accuracy for the GPS orbit of $\Delta x = 1\,\mathrm{cm}$ and assuming a GPS receiver clock correction of $\Delta t = 1\,\mathrm{ms}$ the velocity of the GPS satellite has to be known with only a very low level of accuracy, about $\Delta v = \Delta x / \Delta t = 10\,\mathrm{m/s}$. The velocity of the GPS satellite is required to a higher level of accuracy, however, to correctly apply light-travel time and periodic relativistic corrections. For the GPS satellites, the light-travel time correction Δ_{LTT} and the periodic relativistic correction Δ_{PRC}, see (Ashby 2003), are given as

$$\Delta_{LTT} = -d\frac{\vec{n}_0\vec{v}_S}{c} \tag{5.1}$$

$$\Delta_{PRC} = 2\frac{\vec{r}_S\vec{v}_S}{c}, \tag{5.2}$$

where d and $\vec{n}_0$ denote distance and unit vector between GPS satellite and ground stations, respectively, $\vec{r}_S$ and $\vec{v}_S$ are the geocentric position and velocity vector of the GPS satellite, and c is the speed of light in vacuum. One can easily see that the periodic relativistic correction is satellite-specific and, therefore, is canceled out when forming double-differences or can be absorbed by the GPS satellite clock parameters when using zero-differences. Following (5.1), to compute the light-travel time to an accuracy of $1\,\mathrm{cm}$ (in terms of length), the velocity of the GPS satellite should be known to an accuracy of $\Delta v \approx 0.12\,\mathrm{m/s}$. Since the requirements imposed on the velocity in the computation of the light-travel time correction are not so demanding, the orbits of the GPS satellites can indeed be determined geometrically. Nevertheless, an approximate GPS orbit has to be available, and in principle could be computed solely based on smoothed code measurements. Figure 5.2 shows differences between a kinematic and dynamic orbit (assumed to be more accurate) for the GPS satellite PRN 20 and Figure 5.3 the corresponding a posteriori RMS values of the kinematic positions. Both types of orbit were determined using the same IGS stations, troposphere parameters, station coordinates and Earth rotation parameters and the only difference is in the estimated orbital parameters. Dynamic GPS orbits were modeled by six Keplerian elements, nine solar radiation pressure parameters and one pseudo-stochastic pulse for the one day arc, whereas three kinematic coordinates were estimated for PRN 20 (the parameters of the other satellites were held fixed) every epoch (i.e., every $30\,\mathrm{s}$). In both cases, the ambiguities were held fixed at their integer values. One can easily see that the accuracy of the estimated kinematic positions is in the order of $10-20\,\mathrm{cm}$. Replacing the kinematic parameterization by polynomials over a few $10\,\mathrm{min}$ intervals would considerably improve the "kinematic" GPS orbits. We should bear in mind that the dynamic GPS orbit is usually represented by a polynomial of degree 12 for each step of $1\,\mathrm{h}$ in the
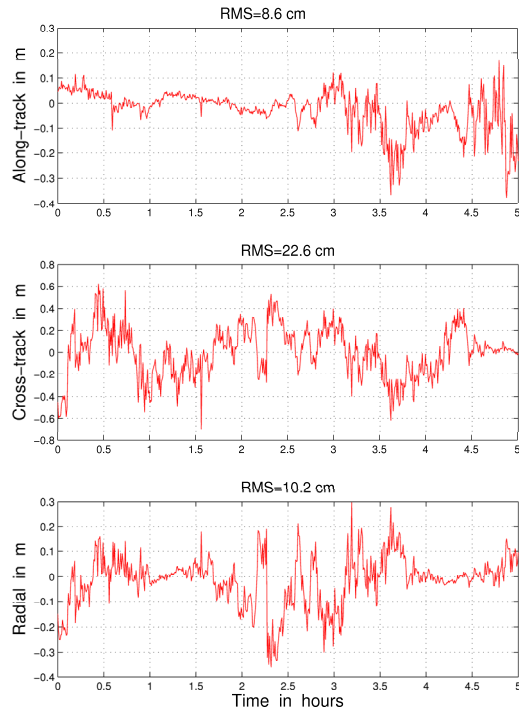
Figure 5.2 Differences between the kinematic and dynamic orbit for GPS satellite PRN 20, day 200/2002.

numerical integration method in the Bernese GPS Software. The rather large variations between kinematic and dynamic GPS positions in Figure 5.2 and the periodic behavior in the corresponding formal precision displayed in Figure 5.3 are certainly due to the weak and slowly changing geometry of ground stations as seen from the GPS satellite.
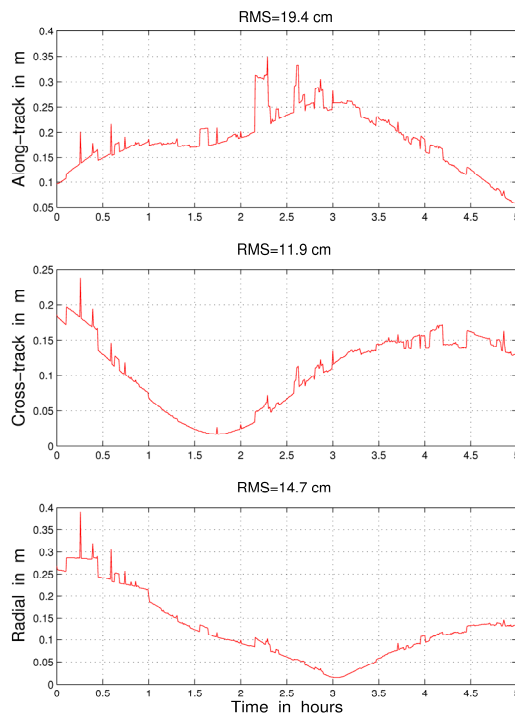
Figure 5.3 A posteriori RMS of the kinematic orbit for GPS satellite PRN 20, day 200/2002.

49
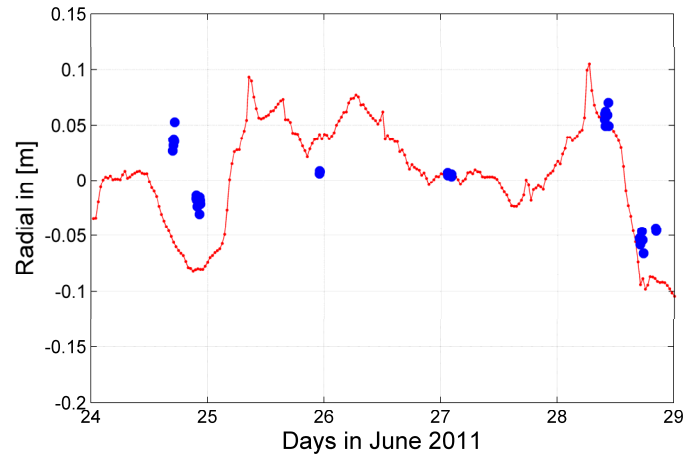
Figure 5.4 Kinematic orbit of the GIOVE-B satellite estimated using a linear model for the onboard H-maser against the SLR normal points (blue).

Figure 5.4 shows the kinematic orbit of the GIOVE-B satellite based on a linear clock model of the onboard passive H-maser. Normal points of the kinematic positions of the GIOVE-B satellite were estimated every 30 min over a period of six days using two-frequency carrier-phase measurements with a sampling of 30 s and without performing any ambiguity resolution at the zero-difference level. GPS orbits and satellite clocks, as well as station coordinates and troposphere parameters were held fixed in the estimation. Blue dots show SLR residuals of the estimated kinematic orbit giving an agreement of about 1-2 cm RMS with the kinematic positions. One can see that the SLR validation closely matches the shape of the estimated kinematic positions against the dynamic orbit for the entire period of time. Compared to Figure 5.2, the kinematic orbit of the GIOVE-B satellite based on a linear clock model is smooth and considerably more stable than the kinematic orbit of the GPS satellite.

6. Kinematics of IGS Stations

For comparison with the kinematic POD of LEO and GPS satellites, a ground GPS baseline from Greenbelt (GODE, US) to Algonquin Park (ALGO, Canada) with a length of 777 km was processed kinematically for a period of one day. The coordinates of one station of the baseline were kept fixed (GODE) and a set of three coordinates was estimated every 30 s for the second station ALGO using carrier-phase data only.

6.1 Ground Double-Difference GPS Baseline in IGS Network

Figure 6.1 shows the kinematic positions of the station ALGO against the ″true″ static coordinates estimated in the global IGS network solution. Ambiguities were resolved using the Melbourne-Wübbena linear combination and narrow-lane bootstrapping. One can see that an accuracy of $0.5 - 1 \, \text{cm}$ in horizontal position
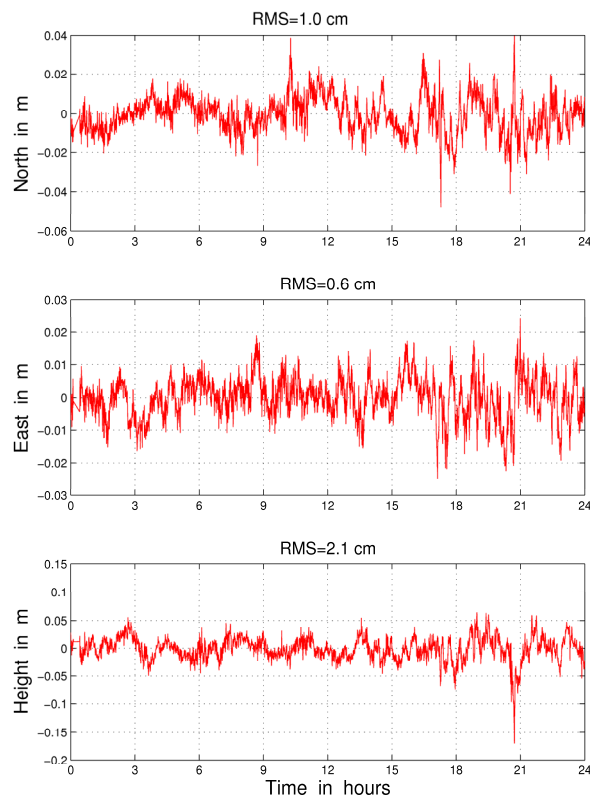
Figure 6.1 Kinematic estimation of the ground IGS point ALGO with respect to the fixed IGS station GODE. Ambiguity-resolved baseline with a length of 777 km, day 200 in the year 2002.

and $2\,\mathrm{cm}$ in height can easily be achieved. Similar results can be obtained, if troposphere parameters are taken from the global IGS solution or estimated every $1\,\mathrm{h}$. Other GPS baselines in the IGS network with lengths of up to $1000\,\mathrm{km}$ show similar results. Using the rule of thumb given by (Bauersima 1983),

$$\Delta l = \frac{l}{20\,000\ \mathrm{km}} \Delta\rho \tag{6.1}$$

with the GPS orbit error of, e.g., $\Delta\rho = 1\,\mathrm{cm}$ RMS and with a baseline length of $l = 1000\,\mathrm{km}$, one can expect an effect in the station coordinates in the order of $\Delta l = 0.5\,\mathrm{mm}$ RMS. For a baseline length of $10\,000\,\mathrm{km}$ (LEO) one can expect about $5\,\mathrm{mm}$ RMS. Therefore, for the ground GPS applications, a GPS orbit accuracy of $1\,\mathrm{cm}$ allows the cm-kinematic positioning for double-difference baselines up to $5000 - 10000\,\mathrm{km}$. From this analysis it follows that station multipath along with the troposphere delay errors (wet part), are probably the main sources of error in ground GPS positioning based on double-differences.

7. Reduced-Kinematic POD

Here we present the results of reduced-kinematic POD, as introduced and published in (Švehla and Rothacher 2005a). Reduced-kinematic POD can be defined as the fourth fundamental approach in precise orbit determination, along with kinematic, reduced-dynamic and dynamic POD. The main difference between reduced-kinematic and reduced-dynamic orbit determination is that in the reduced-kinematic POD the constrained normal equations are set up for the epoch-wise kinematic positions (with epoch-wise clock parameters), whereas in the reduced-dynamic approach, dynamic parameters (such as initial Keplerian state vector, aerodynamic drag coefficients, empirical accelerations, etc.) and/or some pseudo-stochastic parameters are determined. Thus, in the case of reduced-kinematic POD, degrees of freedom are reduced towards a dynamic orbit, whereas in the reduced-dynamic orbit, the dynamics of the orbit is reduced towards a kinematic orbit. Due to the relative or absolute constraints that are used in the reduced-kinematic POD, we did not use nor develop this approach further for LEO satellites. We merely present typical results for the sake of completeness.

7.1 Reduced-Kinematic POD of LEO Satellites

Compared to dynamic orbits, the main disadvantage of kinematic orbits is the presence of ″jumps″ between consecutive positions that occur when, e.g., small numbers of GPS satellites are tracked or when phase breaks occur. Although these ″jumps″ from epoch to epoch are fully reflected in the variance–covariance information, they can be clearly seen in Figure 7.1, where CHAMP kinematic positions are plotted against the dynamic orbit. Typical spikes in kinematic positions, and accordingly in the variance–covariances, can be seen around 1.1, 1.3, 2.5 and 4.1 h and phase breaks can be identified for the isolated arc from 4.1 to 4.6 h .

Compared to kinematic orbits, dynamic orbits are very smooth, i.e., high frequency noise is not visible from epoch to epoch due to the integration of the equation of motion. In order to reduce the size of the small jumps in kinematic position, constraints can be applied from epoch to epoch to the kinematic position differences w.r.t. corresponding differences in the a priori dynamic orbit. In this case, we may speak of ″reduced-kinematic″ orbit determination, where the kinematic degrees of freedom are reduced by constraints to the dynamic orbit. It can be shown that the a priori dynamic LEO orbit used for constraining can be of very low accuracy, e.g., defined by only 15 orbital parameters per day and estimated by means of code measurements only. The size of the relative constraints applied in the computation of reduced-kinematic orbits in Figure 7.1 was 5 mm between 30 s consecutive epochs. Using the reduced-kinematic approach, one can obtain very smooth kinematic orbits where spikes in the kinematic positions are removed or considerably reduced. This is illustrated in Figure 7.1, where kinematic and reduced-kinematic orbits are shown w.r.t. the best reduced-dynamic orbit. Although the stochastic process achieved by relative constraints is a random walk, the trajectory does not drift away from the a priori dynamic orbit. Depending on the strength of the constraints between consecutive epochs, the estimated reduced-kinematic orbit will be closer either to the dynamic or the kinematic orbit.
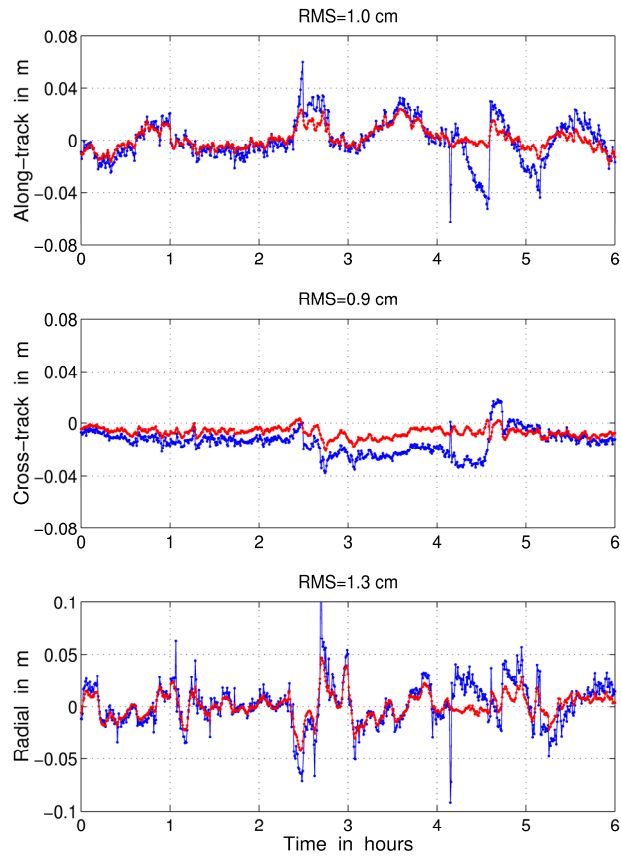
Figure 7.1 Kinematic (blue) and reduced-kinematic orbit (red) for the CHAMP satellite using relative constraints, day 200/2003.
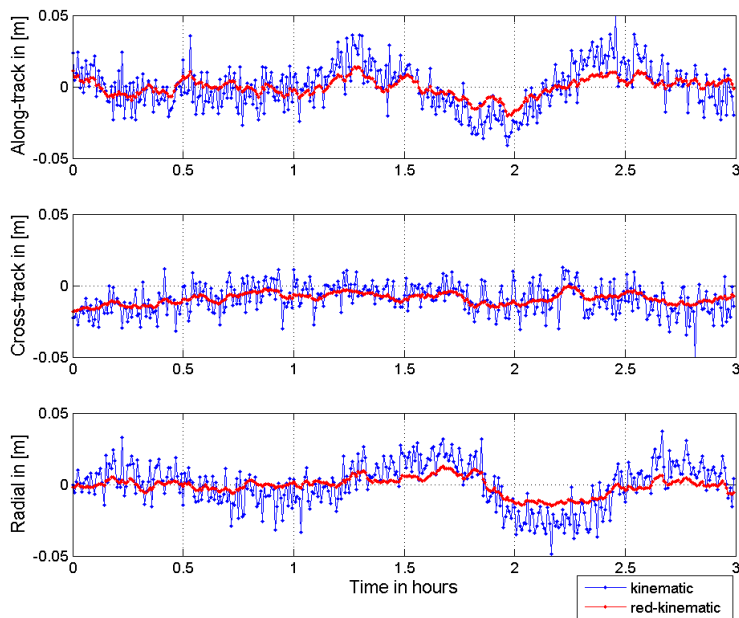
Figure 7.2 Reduced-kinematic orbit of the CHAMP satellite based on simulated phase measurements. Noise of simulated phase observables is $\sigma(L_1) = \sigma(L_2) = 5$ mm , data rate is 30 s .
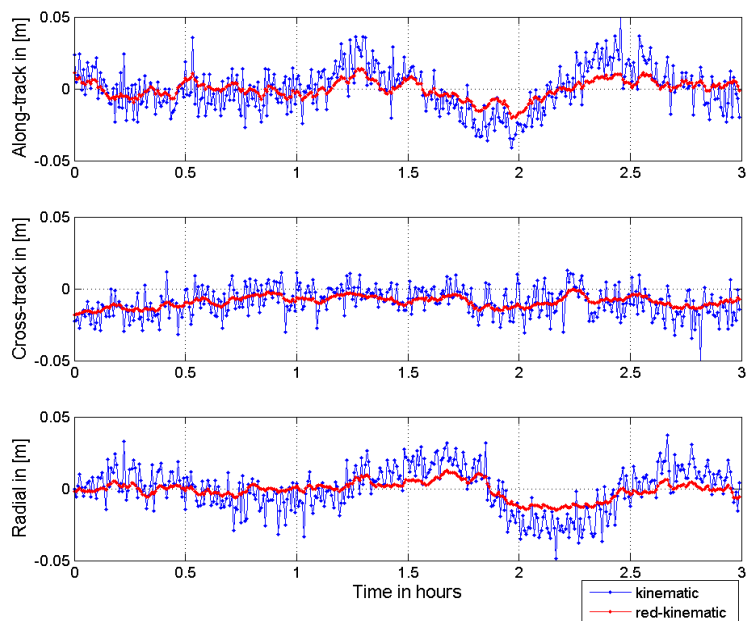
Figure 7.3 Reduced-kinematic orbit for the CHAMP satellite based on simulated phase measurements. Noise of simulated phase observables is $\sigma(L_1) = \sigma(L_2) = 5$ mm , data rate is 10 s .

The main difference between reduced-kinematic and reduced-dynamic orbit determination is that in the re-duced-kinematic POD the constrained normal equations are set up for the epoch-wise kinematic positions (with epoch-wise clocks), whereas in the reduced-dynamic approach dynamic parameters (such as initial Kep-lerian state vector, air-drag coefficients, empirical accelerations, etc.) and/or some pseudo-stochastic parameters are determined. The reduced-kinematic method improves the overall characteristics of the purely kinematic POD by a considerable reduction of spikes and jumps. Therefore, reduced-kinematic POD can be used for LEO applications that require a very smooth trajectory such as radio-occultation. Since the a priori dynamic orbit used in reduced-kinematic POD does not have to be of high accuracy and can be very easily computed, the reduced-kinematic positions will not rely significantly on an a priori gravity field, but will allow, e.g., better velocity computation for the energy balance approach of gravity field determination. However, there will still be a residual dependency on the a priori information. Figure 7.2 and Figure 7.3 show the reduced-kinematic orbit of the CHAMP satellite based on simulated phase measurements with a noise of 5 mm and data rates of 10 s and 30 s , respectively. Computation is performed based on the inversion of the
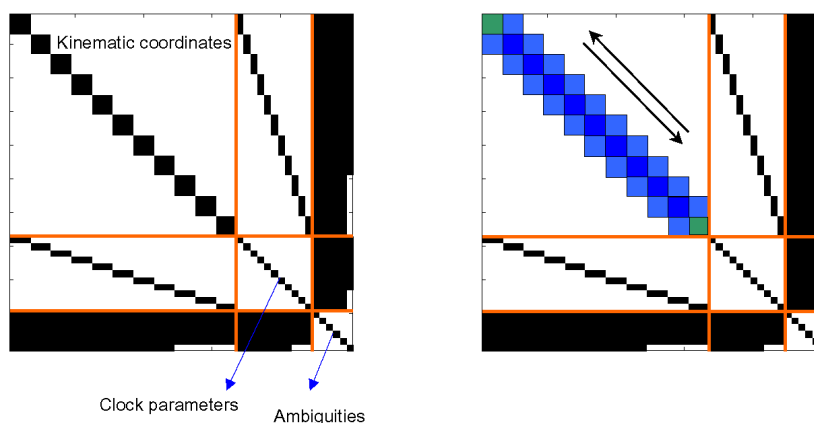
Figure 7.4 Normal equation matrix for kinematic (left) and block tridiagonal normal equation matrix for reduced-kinematic POD (right).

55

full normal equation matrix for a period of 3 hours. In the case of the GOCE mission, GPS measurements are provided every second and therefore kinematic and reduced-kinematic orbits have to be provided with a sampling interval of 1 s .

When relative constraints are set up between epochs, as depicted in Figure 7.4 (right), the normal equation matrix of the kinematic positions is no longer block diagonal, but rather tridiagonal. In kinematic POD, a very efficient parameter pre-elimination scheme is used, where in the first step kinematic positions and clock parameters are pre-eliminated to ambiguities. In the second step, after inversion of the reduced normal equation system (ambiguities only), ambiguities are back-substituted and epoch-wise 4×4 blocks are inverted providing kinematic positions every epoch. Ambiguities can be pre-eliminated to the epoch-wise parameters, but the block tridiagonal/diagonal property of the normal equation matrix is lost in that case, since the normal-equation matrix becomes fully populated.

In order to improve numerical stability and reduce execution time, we studied several algorithms to invert tridiagonal matrices with very large numbers of parameters (about 350 000 per day) as well as various algorithms for inverting sparse matrices. If such an algorithm is to be integrated into the official software, it has to be compatible with all other processing and parameter estimation methods. In particular, it should be compatible with the existing bookkeeping of ambiguity parameters and epoch-wise clock parameters.

After reviewing the software and algorithm design, the conclusion was drawn that reduced-kinematic POD requires a re-design. The main arguments for this decision were the amount of time required to perform the computations and, even more importantly, the significant biases introduced into the reduced-kinematic orbit by the relative constraints applied over a long orbit arc. Small, but significant biases in the cross-track components can be seen in Figure 7.1, Figure 7.2, Figure 7.3.

In the present design, all epochs are constrained, and therefore, relative constraints between the first epoch pair affect the solution of the last epoch pair within the same run. This is similar to applying a small absolute constraint to all epochs. In order to overcome this problem, the current strategy for the reduced-kinematic orbit was changed from relative constraining over the entire arc to a band-limited form of relative constraints preserving the local properties of the orbit. Another solution would be to represent the reduced-kinematic orbit by normal points, estimated every, e.g., few epochs. Computing a normal point over several kinematic positions would have a similar effect to setting up relative constraints between epochs. However, in the case of normal point estimation, the original sampling of the kinematic orbit is lost.
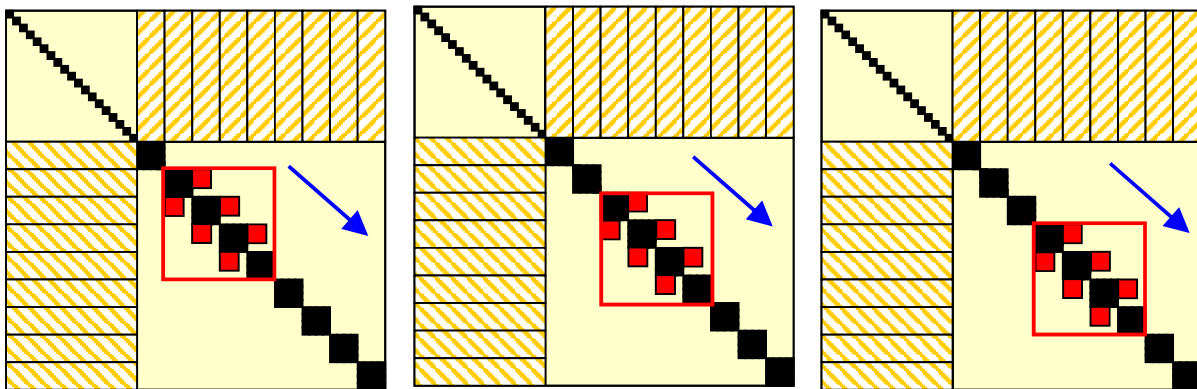
Figure 7.5 Successive use of relative constraints (3×3 red blocks) in the NEQ between coordinates of the consecutive epochs set-up over the three epochs (left, middle, right figure). In the first sub-matrix on the main diagonal (upper left), one can see the ambiguity parameters (1×1 black blocks) and in the second sub-matrix on the main diagonal (lower-right) the 4×4 black blocks (three coordinates and a receiver clock) of parameters set-up every epoch. The yellow color shows the empty fields in the NEQ.

In order to preserve the local properties of the orbit and avoid long-periodic biases introduced by setting up relative constraints over the entire orbital arc, we do not make use of the fully populated cofactor matrix, but rather select elements only over the specified band of epochs, see Figure 7.5. In fact, the same window over a selected number of epochs used to calculate the cofactor matrix Q_{xx} for the kinematic positions can be used to determine the reduced-kinematic orbit. Therefore, kinematic and reduced-kinematic orbit can be determined in the same processing run. However, we never calculated reduced-kinematic orbits for LEO satellite missions, merely tested the algorithm.

Since the reduced-kinematic orbits can be obtained by a parameter transformation of the kinematic orbit (linear combination), the computation of the matrix C with constraints can be extended using dynamic information over a short interval of time. In this way standard numerical integration could, in fact, be avoided, since the local properties of the orbit are preserved over a very short period of time. In this way one can talk of reduced-dynamic POD with local properties. Figure 7.6 (left) was calculated by setting up relative constraints over different bands of consecutive epochs (smoothing window over a number of epochs) and by varying the size of relative constraints. Figure 7.6 (right) shows the equivalence between the reduced-kinematic orbit (smoothing window of 30 s with 1 s sampling) and the highly-reduced-dynamic orbit with a significant number of empirical parameters estimated. Both orbits exhibit a similar power spectrum density (PSD).

To summarize, we have shown that the reduced-kinematic orbit is a very simple representation of the kinematic orbit preserving the local properties of the orbit and reducing the high-frequency noise in kinematic positions between consecutive epochs. However, since only kinematic orbits are used for gravity field determination for the GOCE, GRACE and CHAMP missions, the reduced-kinematic orbit determination strategy has not been developed further.
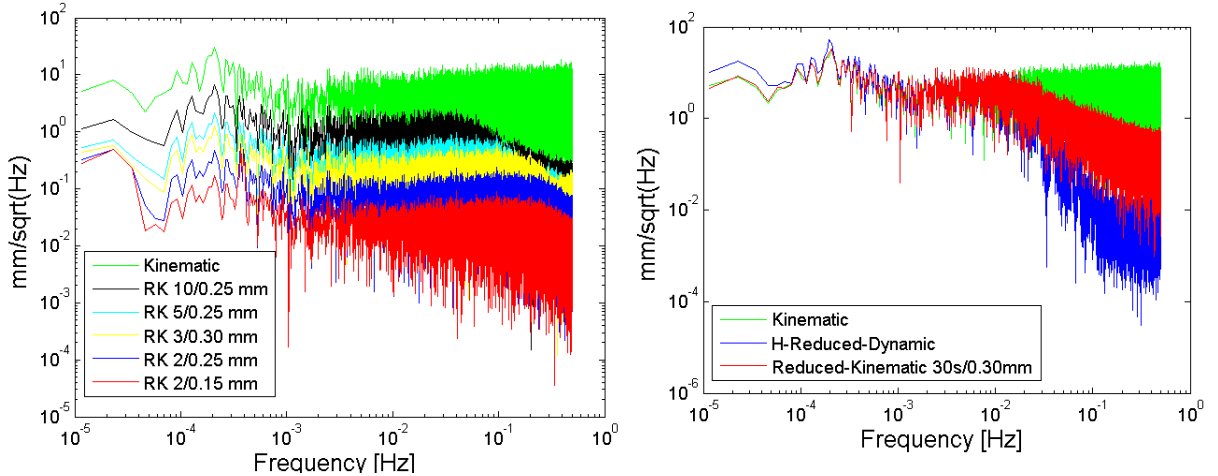
Figure 7.6 PSD of the radial reduced-kinematic positions (denoted as RK) by varying the smoothing window over a number of epochs and the size of the relative constraints (number of epochs/relative constraint). The figure to the right shows the equivalence between the reduced-kinematic orbit (sampling 1 s , constraints over 30 s) and the highly-reduced-dynamic orbit (H-Reduced-Dynamic) with a significant number of estimated empirical parameters. Both orbits show similar PSD.

7.2 Constraints in the Reduced-Kinematic POD

Although we did not use constraints in the kinematic POD, here we provide more information how constraining could be performed for the coordinates of a ground station or the reduced kinematic orbit in the case of a LEO satellite.

Let us write the normal equation system, with the normal equation matrix N and the vector of unknown kinematic positions x estimated along an a priori orbit in the geocentric Cartesian coordinates

$$Nx = n \tag{7.1}$$

constructed from the observation equation associated with the design matrix A and the vector of the reduced observations $-l$, typically termed "observed minus computed"

$$v = Ax - l \tag{7.2}$$

where the vector of errors is denoted with v and associated with the weight matrix P. The normal equation matrix is then

$$N = A^T P A, \qquad n = A^T P l \tag{7.3}$$

Typically, the vector of unknowns x is given in geocentric Cartesian coordinates. If we require unknowns in the local station coordinate system (north, east, up) or in the local orbital frame (along, cross, radial), we need to transform our geocentric unknowns $x = [\delta x, \delta y, \delta z]$, first to the geographic coordinates $\widehat{x} = [\delta\varphi, \delta\lambda, \delta h]$ and then to apply any additional rotation necessary to achieve a local station coordinate system (north, east, up) or a local orbital LEO coordinate system (along, cross, radial)

$$x = R\widehat{x} \qquad \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = R \cdot \widehat{x} = R \cdot \begin{bmatrix} d\varphi \\ d\lambda \\ dh \end{bmatrix} \tag{7.4}$$

with matrix R (partial derivatives) and the constraint (weight) matrix C

$$R = \begin{bmatrix} -\sin\varphi\cos\lambda & -\cos\varphi\sin\lambda & \cos\varphi\cos\lambda \\ -\sin\varphi\sin\lambda & \cos\varphi\cos\lambda & \cos\varphi\sin\lambda \\ \cos\varphi & 0 & \sin\varphi \end{bmatrix} \qquad C = \begin{bmatrix} \dfrac{m_0^2}{m_0^2(\varphi)} & & 0 \\ & \dfrac{m_0^2}{m_0^2(\lambda)} & \\ 0 & & \dfrac{m_0^2}{m_0^2(h)} \end{bmatrix} \tag{7.5}$$

with $m_0(\varphi)$, $m_0(\lambda)$ and $m_0(h)$ denoting the noise level or the constraints of the estimated parameters where $m_0 = \sqrt{v^T P v / (n-u)}$ with the number of measurements n and number of unknowns u. After substitution of (7.5) into (7.1) we derive

$$(R^T N R + C) \cdot \widehat{x} = R^T n = \widehat{n} \tag{7.6}$$

where $\widehat{n} = R^T A^T P l$. The absolute constraints could be represented by using the vector of unknowns as the pseudo-observations in the least squares adjustment with an identity matrix as the design matrix A.

If the relative constraints are set up between consecutive epochs, the approach is similar to that for absolute constraints, with the difference that the design matrix is no longer an identity matrix for the estimated coordinates of the consecutive epochs x_{i-1} and x_i,

$$A = \begin{bmatrix} 1 & -1 & & 0 \\ & 1 & -1 & \\ & & ... & \\ 0 & & 1 & -1 \end{bmatrix} \tag{7.7}$$

By introducing the matrix $\widehat{C}$ containing the relative constraints between consecutive epochs x_{i-1} and x_i (see (7.5)), we obtain

$$
\begin{aligned}
(\widehat{C} + N)\overline{x} &= n \\
N^{-1}\widehat{C}\overline{x} + \overline{x} &= N^{-1}n \\
N^{-1}\widehat{C}\overline{x} + \overline{x} &= x \\
(N^{-1}\widehat{C} + I)\overline{x} &= x \\
\overline{x} &= (N^{-1}\widehat{C} + I)^{-1}x \\
\overline{x} &= (Q_{xx}\widehat{C} + I)^{-1}x \\
\overline{x} &= \overline{Q}_{xx}x
\end{aligned}
\tag{7.8}
$$

with $\overline{Q}_{xx} = (Q_{xx}C + I)^{-1}$. As shown in (7.8), the reduced-kinematic orbit $\overline{x}$ can be obtained by a parameter transformation of the original kinematic orbit x (linear combination). As an alternative to (7.8), reduced-kinematic positions can be obtained by calculating a "small correction" to the existing kinematic orbit, e.g.,

$$
\begin{aligned}
\overline{x} &= x - N^{-1}\widehat{C}\overline{x} \\
\overline{x} &= x - Q_{xx}\widehat{C}\overline{x}
\end{aligned}
\tag{7.9}
$$

8. First GPS Baseline in Space – the GRACE Mission

In (Švehla and Rothacher 2004c) it was reported for the first time that the orbit vector between the two GRACE satellites equipped with GPS in the LEO orbit can be estimated with mm-level accuracy. This level of accuracy was achieved after performing ambiguity resolution for the GPS double-difference baseline and independently confirmed by the K-band measurements between the two GRACE satellites. Here we present the results of this GPS baseline in space.

8.1 Formation Flying Using GPS

Distributed space systems employ two or more spacecraft which act in a coordinated way to achieve the common mission objective. The architecture of such distributed systems can be based on rendezvous and docking scenarios with two spacecraft in close vicinity, formation flying with two or more spacecraft with a separation of a few tens of meters to a few 100 kilometers, constellations with several spacecraft distributed on a global scale or swarms with a multitude of spacecraft, each with limited functionality (Gill 2006). Following this very precise definition, the GRACE mission is a typical example of formation flying and the US/Taiwanese COSMIC mission a constellation of six satellites in LEO orbit, whereas the ESA Swarm mission is the first swarm in LEO orbit.

Let us now see what accuracy might be achievable for the inter-satellite baseline between the two GRACE satellites using a kinematic approach. In order to do this, phase zero-difference measurements were simulated for both GRACE satellites, assuming the noise level and the number of GPS satellites tracked to be similar to CHAMP (only a noise of 1.1 mm was considered, with multipath included in this noise). A typical noise value for the a posteriori RMS of the phase zero-differences in CHAMP kinematic POD is about $1.5 - 2.0$ mm or $1.2 - 1.4$ mm when using double-differences. Whereas zero-differences are mainly affected by the GPS satellite orbit/clock errors, double-differences primarily reflect ground station specific errors such as troposphere, multipath, etc. Therefore, the noise level of 1.1 mm adopted for the GRACE simulation might be considered rather pessimistic, bearing in mind that for the short GRACE baseline (about 220 km) the effect of GPS orbit errors should only be about 0.2 mm, tropospheric refraction is non-existent and multipath is expected to be very small. Figure 8.1 shows the GRACE kinematic baseline results with float, Figure 8.2 those with fixed ambiguities. In both cases, the GRACE-B positions were held fixed to the a priori orbit and GRACE-A positions were estimated kinematically. Comparing these two figures, one can clearly see that ambiguity resolution de-correlates kinematic coordinates and ambiguities and changes the colored noise present in the kinematic positions of the float solution into white noise. A decrease of the a posteriori RMS from 5 to 3 mm for the along-track component can also be noticed. Ambiguity resolution was performed as explained in section 1.9 (Melbourne-Wübbena wide-laning, narrow-lane bootstrapping) and all ambiguities were correctly resolved. GRACE GPS
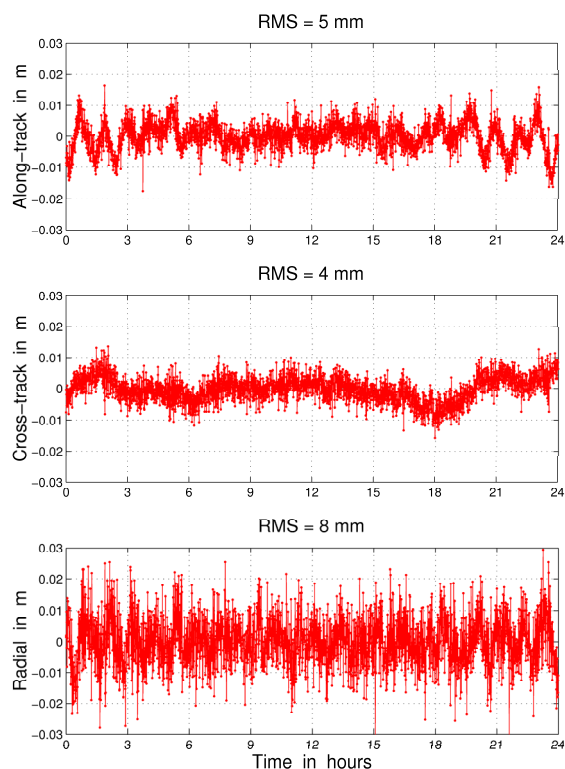
Figure 8.1 Kinematic positions of GRACE-A w.r.t. GRACE-B from simulated data with float ambiguities compared to the true baseline. Note the colored noise, reflecting correlations between positions and ambiguities.

Figure 8.2 Kinematic positions of GRACE-A w.r.t. GRACE-B from simulated data with resolved ambiguities compared to the true baseline. Note the white noise in the kinematic positions and the reduction of the a posteriori RMS from 5 to 3 mm .

data are a very nice playground to study, for the first time, an inter-satellite baseline with the unique possibility to validate the results with the much more accurate measurements of the K-band link.

8.2 GRACE GPS Baseline

The orbits of both the GRACE A and GRACE B satellites can be determined independently of each other using either zero-difference point positioning or double-difference baselines formed from IGS GPS stations to the GRACE satellites. In both cases, the GRACE satellites are treated as two independent satellites similar to CHAMP and their orbits are estimated independently. An alternative approach consists of a combined zero- and double-difference POD, where one LEO satellite is determined absolutely using zero-differences, and the other satellite is determined relatively to the reference satellite by forming a very accurate inter-satellite GPS baseline. In order to validate such a spaceborne double-difference GPS baseline, KBR measurements were used. The KBR observable is the biased distance between the two GRACE satellites measured to an accuracy of a few micrometers. Figure 8.4 shows the KBR residuals for the GRACE GPS baseline with fixed ambiguities and Figure 8.3 shows the KBR residuals for the reduced-dynamic orbits of the two GRACE satellites estimated separately using zero-differences. The KBR and accelerometer data were not used in the orbit determination. These two figures show that for LEO satellites flying in formation (e.g., the two GRACE satellites) the optimum POD strategy is to estimate the orbit or position of one satellite absolutely and those of the other satellites in the formation relatively by forming GPS baselines in space to the reference satellite.
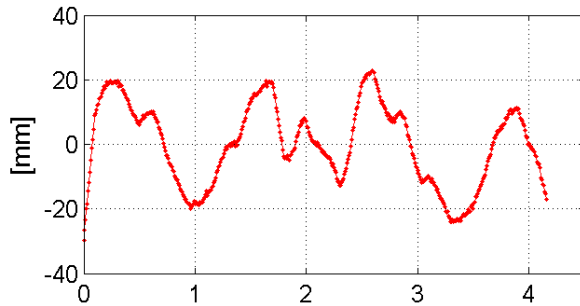
Figure 8.3 KBR residuals of the reduced-dynamic orbits (over 4 hours) of GRACE-A and GRACE-B satellites estimated independently of each other using zero-difference carrier-phase measurements, $\text{RMS} = 12.6$ mm . A clear once-per-orbit pattern can be recognized. GPS Day 200/2003.
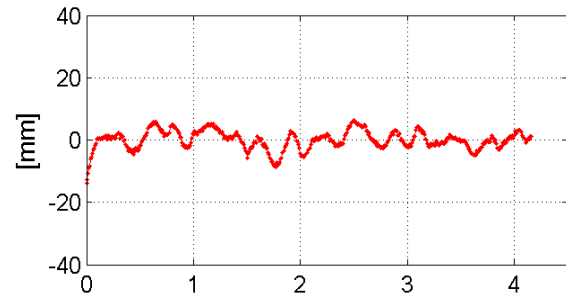
Figure 8.4 KBR residuals of the orbits of the GRACE-A and GRACE-B satellites (over 4 hours) estimated using the GRACE GPS baseline with fixed ambiguities, $\text{RMS} = 2.8$ mm . Peaks in the KBR residuals show the epochs where pseudo-stochastic pulses were introduced. GPS Day 200/2003.

In this case, the relative orbit information between the LEO satellites can be estimated to a level of $1-3$ mm (see Figure 8.4) compared to the $10-15$ mm in the case where all satellites are estimated independently from each other, e.g., using zero-difference GPS measurements, as in Figure 8.3. More about LEO formation flying and the GRACE GPS baseline can be found in (Švehla and Rothacher 2004c). Figure 8.5 shows the number of double-difference ambiguities and the percentage of resolved wide-lane ambiguities using the Melbourne-Wübbena linear combination for the two GRACE satellites for a period of four months, days 182-303/2003. Figure 8.6 shows the percentage of the resolved narrow-lane ambiguities using bootstrapping of the normal equation matrix (NEQ) with dynamic orbit parameters (left) and kinematic positions (right).
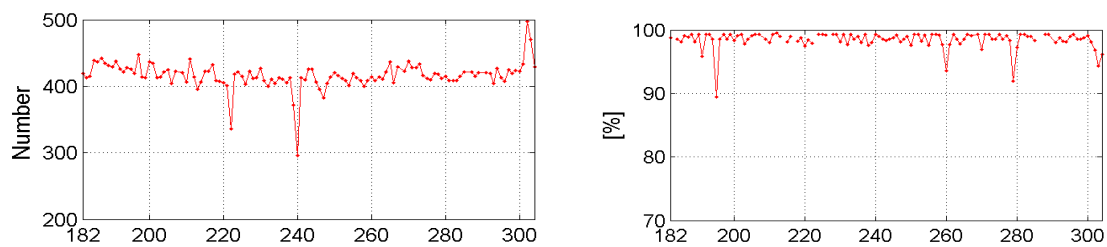
Figure 8.5 Left: Total number of ambiguities per day, mean= 416/day. GPS days 182-303/2003. Right: Resolved wide-lane ambiguities (Melbourne-Wübbena), mean= 98.4%.
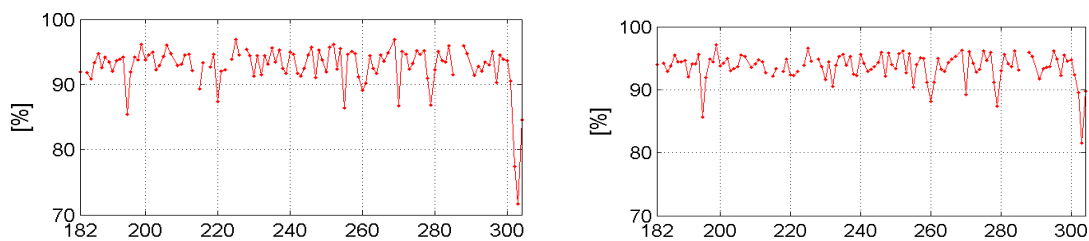
Figure 8.6 Resolved narrow-lane ambiguities using dynamic NEQ bootstrapping (left), mean= 92.8%. and kinematic NEQ bootstrapping (right), mean= 93.6%. GRACE data set 182-303/2003.

a) red.-dyn. baseline: float ambig. RMS = 11.6 mm

b) red.-dyn. baseline: fixed ambig. RMS = 2.8 mm

c) kinematic baseline: float ambig. RMS = 11.7 mm

d) kinematic baseline: fixed ambig. RMS = 6.1 mm
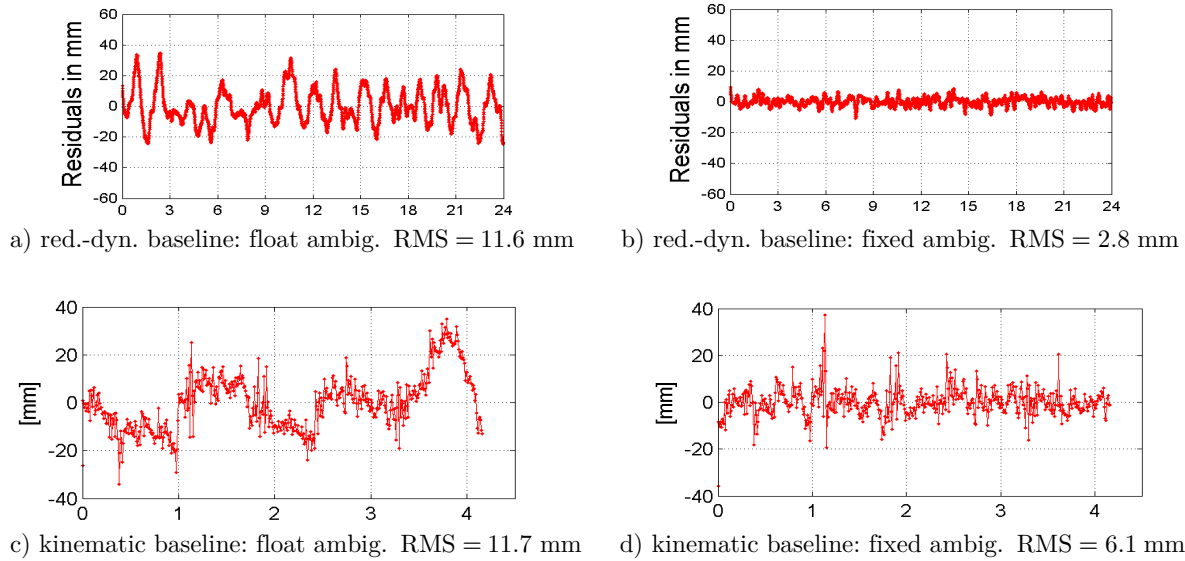
Figure 8.7 Kinematic and reduced-dynamic GPS baselines between the two GRACE satellites, estimated using double-differences with float ambiguities (left) and double-differences with fixed ambiguities (right) compared to KBR measurements, day 200/2003.

The impact of the ambiguity resolution on the kinematic and reduced-dynamic GPS baseline is shown in Figure 8.7. One can see that ambiguity resolution improves the relative orbit accuracy by about one order of magnitude in the case of a reduced-dynamic orbit, whereas in the case of a kinematic parameterization this improvement is about a factor of two. A clear, once-per-orbit pattern can be recognized in both the reduced-dynamic and the kinematic double-difference baselines with float ambiguities that is removed after performing the ambiguity resolution. A closer look at the reduced-dynamic baseline with fixed ambiguities in Figure 8.7. b), and especially in Figure 8.4, reveals a very strong systematic pattern in the KBR residuals, indicating epochs where pseudo-stochastic pulses were set up in the reduced-dynamic orbit parameterization (every 6 min in this case).
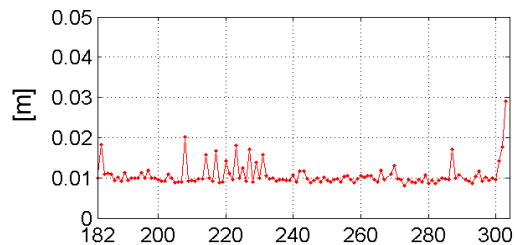
Figure 8.8 Daily RMS of the differences between the kinematic and reduced-dynamic baseline results, days 182-303/2003. After ambiguity resolution, kinematic and reduced-dynamic baseline results agree to within 1 cm . Compared to zero-difference GRACE orbits this is an improvement in accuracy by a factor of two.

64

8.3 Along-Track Sub-mm Kinematic Orbit Determination with GRACE – Combination of GPS and K-Band Measurements

Let us now see what happens when GPS measurements from the two GRACE satellites are combined with inter-satellite K-band measurements in kinematic and reduced-dynamic POD. For this, we first estimated the orbit of the GRACE-B satellite using zero-differences and in the second step we estimated the position of the GRACE-A satellite kinematically from the orbit of the GRACE-B satellite.

In this relative orbit determination, K-band measurements are combined with GPS measurements with fixed narrow-lane ambiguities. Figure 8.9. shows the differences between kinematic and reduced-dynamic orbits after fixing double-difference ambiguities and combining GPS with GRACE K-band measurements of μm-precision. One can see that differences are in the order of a few millimeters in the along-track and cross-track directions and up to one centimeter in the radial direction. Interestingly, the difference in the along-track orbit direction is not zero. This is what one would expect in kinematic POD, when combining GPS double-difference measurements with K-band measurements of very high weight. The most likely explanation for this fact is that the combined GPS/K-band reduced-dynamic baseline is limited by the level of accuracy of the dynamic orbit models used, as shown in Figure 8.9 in the along-track direction. The reasons for this lie, most likely, in the accuracy of the dynamic orbit modeling, in the orbit parameterization and the numerical integration (gravity field used). Thus, the accuracy of the reduced-dynamic baseline between two GRACE satellites can, in our case, be determined with an RMS of the order of 0.7 mm .
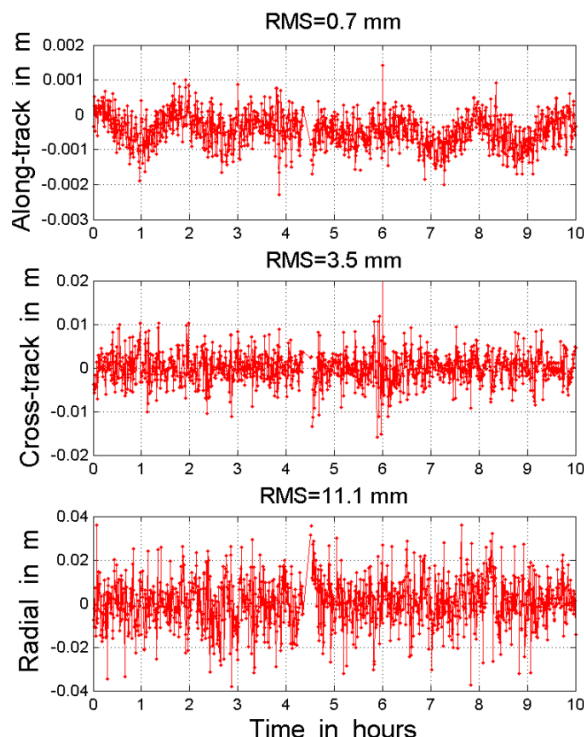
Figure 8.9 Kinematic GPS baseline in space with fixed ambiguities combined with the K-band measurements (sub-mm accuracy only in along-track orbit direction) between two GRACE satellites separated in the same orbital plane by about 200 km , day 300/2003.

9. Geometrical Modeling of the Ionosphere and the Troposphere with LEO Orbit

In this section, we first briefly describe the mathematical and physical background of the first and second order ionosphere effects on LEO GPS measurements and then give a geometrical interpretation of the second order ionosphere effect for one-way and two-way LEO tracking observables. We discuss systematic effects resulting from higher order ionosphere effects on LEO orbit determination and then on gravity field and altimetry results. We show that, when the IGS TEC maps are compared to the TEC observed along the CHAMP orbit (merely by applying a constant bias) during the solar maximum, the agreement is excellent and is at the level of about 1 TECU or below. We show how to calculate the fractional TEC below or above the LEO orbit, taking into account the Sun's position w.r.t. LEO orbit. We show that the fractional TEC for LEO orbit can be calculated exactly from the Chapman function, by transforming the Chapman function into the "error function" $erf(x)$, encountered when integrating the normal distribution in statistics. This allows a direct combination of LEO and ground IGS TEC maps. After that, we present a novel remove-restore approach in the combination of LEO and ground-based TEC measurements by means of least-squares collocation. The same approach could be applied to augment final and real-time IGS TEC maps. It is proposed to model the ionospheric TEC (by combining LEO and ground GNSS measurements) as a spherically-layered electron density distribution in three main Chapman layers, i.e., E, F1 and F2 with an additional layer for the plasmaspheric density above the ionosphere, using GOCE (above the E-layer), GRACE (above the F1 layer) and Jason-2 (above the F2-layer and below the plasmasphere). In the second part, we discuss tropospheric effects on the propagation of microwave and optical measurements and show the influence of tropospheric effects on the kinematic and reduced-dynamic POD of LEO satellites. We show that there is an effect of the tropospheric modeling on the estimated low-degree zonal gravity field coefficients based on LEO orbits. At the end, we propose a way forward in modeling ground-specific high-resolution tropospheric delays for all space geodesy techniques, making use of the high-performing clocks on board the new GNSS satellites and the more than 35 GNSS satellites in the field of view of a ground station, given that all four GNSS constellations will be deployed in a few years from now. For that, ground-specific tropospheric and ionospheric delays could be modeled making use of the rotation of spherical harmonics in order to account for temporal variations w.r.t. a fixed frame. Rotations of spherical harmonic coefficients provide continuous TEC information.

9.1 Ionospheric Refraction and LEO

The ionosphere is a dispersive and anisotropic medium for radio waves. The first-order ionospheric group delay (or phase advance) for microwave signals is in the order of $1-50\,\mathrm{m}$. For the GPS carrier phase and code observables, the observation equation with higher-order ionosphere effects can be derived as

$$L_i = \rho + N_i \lambda_i - \frac{c_2}{f_i^2} - \frac{c_3}{f_i^3} - \frac{c_4}{f_i^4} - \frac{c_5}{f_i^5} \tag{9.1}$$

$$P_i = \rho + \frac{c_2}{f_i^2} + 2\frac{c_3}{f_i^3} + 3\frac{c_4}{f_i^4} + 4\frac{c_5}{f_i^5} \tag{9.2}$$

where the index i refers to the GPS frequency f_i, P_i and L_i are the code and carrier phase measurements respectively, λ_i is the corresponding wavelength, and N_i is the integer ambiguity. The geometry part, denoted by ρ, includes the geometrical distance and the clock corrections, as well as other effects, including the phase wind-up, and the Shapiro and light-travel time corrections. From (9.1) and (9.2) the first-order ionosphere correction (it appears with factor $1/f$ for the carrier-phase in cycles) causes a group delay (code measurements) and phase advance (carrier phase measurements). The c_2, c_3, c_4 and c_5 are the coefficients of the first-, second-, third- and fourth-order ionosphere effects respectively. They approximate the phase refractive index n_{ph}

$$n_{ph} = 1 + \frac{c_2}{f_i^2} + \frac{c_3}{f_i^3} + \frac{c_4}{f_i^4} + \frac{c_5}{f_i^5} + \dots \tag{9.3}$$

Making use of the Rayleigh equation, the group refractive index n_{gr} can easily be derived from the phase refractive index n_{ph}

$$n_{gr} = n_{ph} + f\frac{dn_{ph}}{df} \tag{9.4}$$

and up to the fourth-order, for the group refractive index we can derive

$$n_{gr} = 1 - \frac{c_2}{f_i^2} - 2\frac{c_3}{f_i^3} - 3\frac{c_4}{f_i^4} - 4\frac{c_5}{f_i^5} + \dots \tag{9.5}$$

Since the velocity of carrier waves v_{ph} and the group velocity v_{gr} is given by

$$v_{ph} = \frac{c}{n_{ph}}, \qquad v_{gr} = \frac{c}{n_{gr}} \tag{9.6}$$

making use of the approximation $(1 + \varepsilon)^{-1} = 1 - \varepsilon$ for (9.3) and (9.5) in (9.6), we can derive higher order ionosphere effects in (9.1) and (9.2).

The coefficient c_2 of the first order ionosphere effect is typically given as $c_2 = -40.3 \cdot TEC$ and measured in $[\mathrm{Hz}^2]$, where TEC stands for the total electron content along the line of sight . The first order ionosphere free effect can be eliminated by forming the so-called ionosphere-free linear combination, denoted in some cases by L_3. For the ionosphere-free linear combination of the carrier-phase this is given by

$$L_3 = \frac{f_1^2}{f_1^2 - f_2^2} L_1 - \frac{f_2^2}{f_1^2 - f_2^2} L_2 . \tag{9.7}$$

This very nice formula can easily be derived by multiplying the original carrier phase measurements given in cycles with α_1 and α_2, and introducing the condition that the first-order ionosphere effect is eliminated by forming the linear combination

$$\alpha_1 \frac{c_2}{c \cdot f_1} + \alpha_2 \frac{c_2}{c \cdot f_2} = 0 \tag{9.8}$$

Setting $\alpha_1 = 1$, we obtain

$$\alpha_2 = -\frac{f_2}{f_1}.$$ (9.9)

The second-order ionosphere effect is caused by the Faraday rotation effect induced by the Earth's magnetic field and depends on the direction of signal propagation, (see, e.g., (Kedar et al. 2003)). The second-order ionosphere correction in (9.1) and (9.2) can be calculated by means of

$$s = \int f_g f_p^2 \cos\theta_B dL = 7527 \cdot c \int N B_0 \cos\theta_B dL$$ (9.10)

as originally given in (Kedar et al. 2003), where f_g is gyro frequency ($\sim 0.59\,\mathrm{MHz}$) and f_p is the plasma frequency integrated along the line of sight and c is the speed of light in vacuum. For more details see (Kedar et al. 2003). (9.10) is related to the coefficient of the second-order ionosphere effect in (9.1) and (9.2) by $s = 2c_3$. The integral part of (9.10) includes the integration of the total electron content TEC

$$TEC = \int N dL$$ (9.11)

along the line of sight, multiplied by the strength B_0 of the magnetic field vector $\vec{B}_0$ projected in the direction of signal propagation $\vec{k}$. Considering the definition of the scalar product of two vectors spanning the angle θ_B, (9.10) can further be written as

$$s = 7527 \cdot c \int (\vec{B}_0 \vec{k}) N dL$$ (9.12)

as originally given in (Kedar et al. 2003). A simple magnetic dipole model of the Earth's magnetic field was recommended in (Kedar et al. 2003), along with a single layer model for the ionosphere. For a ground station with magnetic latitude λ_m, colatitude θ_m and a satellite with elevation E_m and azimuth A_m (measured clockwise from the magnetic pole), the magnetic colatitude θ'_m of the sub-ionospheric point, where the signal propagation direction intersects the ionosphere layer is to the first order (Kedar et al. 2003)

$$\theta'_m = \theta_m - \frac{H}{R_E \sin E_m} \cos A_m \cos E_m$$ (9.13)

R_E denotes the Earth's radius ($R_E = 6370\,\mathrm{km}$) and H is the reference height of the ionosphere single layer model ($H = 400\,\mathrm{km}$). The scalar product of the magnetic field vector $\vec{B}_0$ and the signal propagation unit vector $\vec{k}$ reads as

$$\vec{B}_0 \vec{k} = B_g \left(\frac{R_E}{r_m}\right)^3 \left(\sin\theta'_m \cos E_m \cos A_m - 2\cos\theta'_m \sin E_m\right)$$ (9.14)

with radius $r_m = R_E + H$ and the amplitude of the equatorial magnetic field at the Earth's surface B_g ($\sim 3.12 \cdot 10^{-5}\,\mathrm{T}$). Finally, (Kedar et al. 2003) defines the second-order ionospheric group delay ΔIg_i in meters for the GPS signal wavelength λ_i as

$$\Delta Ig_i = 2.61 \cdot 10^{-18} \lambda_i^3 \left(\frac{R_E}{r_m}\right)^3 \left(\sin\theta'_m \cos E_m \cos A_m - 2\cos\theta'_m \sin E_m\right) \cdot TEC \quad [m]$$ (9.15)

From (9.1), the phase delay ΔIp_i (advance) is then

$$\Delta Ip_i = -\frac{1}{2}\Delta Ig_i$$ (9.16)

69

In fact, different TEC values should be used for the first- and the second-order ionosphere correction, as the Faraday rotation effect is due to electrons below $2000\,\text{km}$, but the effect would be very small. For more details see (Davies 1990). There are other higher order ionosphere effects that also include the additional bending of the signal, but they will not be discussed here.

9.2 Geometric Interpretation of the Second Order Ionosphere Effect for One-Way LEO and Two-Way LEO Observables

When the signal direction vector $\vec{k}$ is parallel to the magnetic field vector $\vec{B}_0$, the phase signal is delayed. The opposite is true as well, when the vector $\vec{k}$ is anti-parallel to $\vec{B}_0$. In both cases, the true position is shifted accordingly. By considering the geometry of the Earth's magnetic field lines, (see Figure 9.1) and the inclination of GPS satellite orbits, one can draw the conclusion that the second-order ionosphere effect mainly occurs at lower elevations (mid-latitudes). This means that the effect is close to zero towards the zenith, when the satellite signal from the zenith direction is orthogonal to the lines of the Earth's magnetic field. The effect is also highly dependent on the azimuth angle. As a rule of thumb, the apparent distance from GPS stations in the Northern hemisphere is shortened compared to that from stations in the Southern hemisphere. Therefore, stations appear further north than they really are, especially at higher latitudes. The same happens with the determination of a polar LEO orbit in the along-track orbit component, (see Figure 9.1), i.e., the determined orbit is translated within the geocentric frame.

The second order ionosphere effect changes the scale of the observables and therefore the scale of the corresponding GPS solutions, reference frame parameters, GPS baseline, ground network or a determined LEO orbit. Looking at Figure 9.1, one can see that the effect is strongly geographically correlated following the Earth's magnetic field profile. Therefore, in the case of a polar LEO orbit, this could lead to significant long-periodic errors in the determined orbit and shifts in the geocenter of that orbit. In the case of all POD approaches, we may expect the orbit to be systematically translated in the reference frame along the lines of the Earth's magnetic field, see Figure 9.1. This is significant for a very low GOCE orbit with the entire ionosphere above that orbit, c.f. Figure 9.2. On the other hand, altimetry satellites are typically aligned away from a polar inclination, and hence from a magnetic field axis, and will thus experience a different systematic distortion and offset of the orbit. Nevertheless, altimetry satellites are typically placed above the Chapman layer and therefore above the main part of the ionosphere, thus the overall effect will be significantly smaller than for a low LEO orbit. However, there is still an ionosphere effect stemming from the plasmasphere above the $1000\,\text{km}$ orbit altitude that affects GNSS measurements from altimetry missions above that altitude.

Compared to classical one way measurements, the advantage of using two-way measurements for, e.g., frequency transfer lies in the possibility of removing all geometrical and signal propagation effects. In this way, a frequency between two ground clocks can be compared directly without parameter estimation. However, the only propagation effect that is not eliminated in two-way measurements is the second-order ionosphere effect. The reason for that is that Faraday rotation depends on the signal propagation direction. In the case of receiving and sending a signal from the same ground station to LEO, HEO or an interplanetary orbit, the second-order ionosphere effect is compounded, i.e., doubled. Therefore, the only way forward in designing a
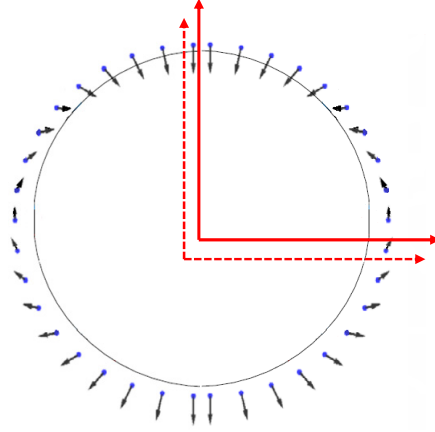
Figure 9.1 Profile of the Earth's magnetic field along the polar orbit. Arrows show direction and strength of the field and the corresponding distortion and offset of the orbit (e.g., along-track).

high-performance metrology ground-to-space link is to go towards higher X-band or Ka-frequencies where first- and higher-order ionosphere effects decrease rapidly with frequency.

GNSS frequencies are in the $1.2 - 1.5$ GHz range, compared to S-band at ≈ 2.248 GHz (microwave link for the ACES mission). The second order ionosphere effect for S-band is smaller by a factor of $\approx 3 - 4$ than that for the L_2 GPS frequency. In the zenith direction, there can easily be an effect of 1 cm ≈ 30 ps, whereas close to the horizon, the effect is multiplied by a factor of ≈ 10. During the solar maximum, the ionospheric TEC value can reach up to 200 TECU. TEC maps are provided by the IGS for the zenith direction, thus towards the horizon the effect is increased by $1/\cos$ (zenith angle), or one can use a multiplication factor of $6 - 12$ for elevation angles in the range of $5° - 10°$. Of all space geodesy techniques, only SLR is free from ionosphere effects.

9.3 Ionosphere Effect at LEO Altitude

Here we look at the possibility of using global TEC maps provided on a regular basis by the IGS to calculate the fractional TEC above or below a LEO orbit. First, a few words about ionospheric modeling using the single layer model we have referred to. The IGS provides Global Ionosphere Maps (GIM) on a daily basis with a time resolution of two hours, (see e.g., (Dow et al. 2005)). These maps are generated using estimates from the ground IGS network and contain the total electron content between the Earth's surface and the GPS orbit height.

Figure 9.2 shows the vertical profile of electron density often called the Chapman layer. The Chapman function provides a simple model of the ion production rate as a function of altitude h and the zenith angle χ with respect to the Sun (Davies 1990)

$$q(h, \chi) = q_0 e^{(1 - z - \sec \chi e^{-z})} \tag{9.17}$$

The scaled altitude for the altitude h and the reference height h_0 (when the Sun is at its zenith, $\chi = 0$) reads as

$$z = \frac{h - h_0}{\Delta h} \tag{9.18}$$

with Δh denoting the scale height (typically $h_0 = 450$ km, $\Delta h = 100$ km). h_0 is the reference height of maximum ion production when the Sun is at its zenith.
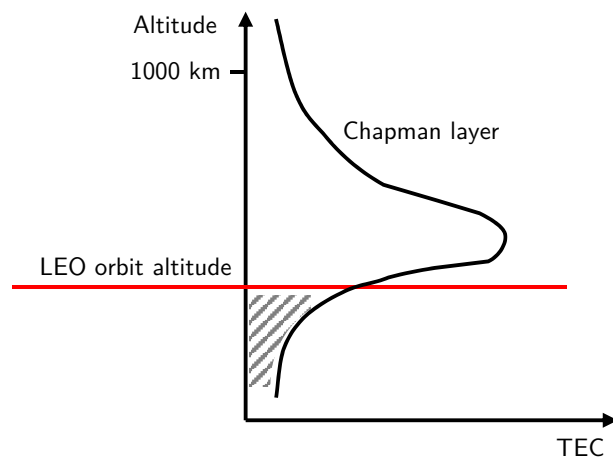
Figure 9.2 LEO orbit altitude and vertical electron density distribution.

For all other zenith angles the height of maximum ion production is given by

$$h_{\max} = h_0 + \Delta h \ln \frac{1}{\cos \chi} \,. \tag{9.19}$$

The ion production rate q_0 is given by

$$q_0 = \frac{\phi(\infty) \cdot \eta}{\Delta h \cdot \mathrm{e}} \tag{9.20}$$

where $\phi(\infty)$ denotes the solar flux density outside the atmosphere (photons per square meter), η is the number of ion pairs produced per photon and ″ e ″ is the base of natural logarithms. The electron density distribution corresponding to the Chapman function in (9.17) is called the Chapman layer and is given by

$$N_e(z, \chi) = N_{e,0} \mathrm{e}^{\frac{1}{2}\left(1 - z - \sec \chi \mathrm{e}^{-z}\right)} \tag{9.21}$$

with α denoting the mean recombination coefficient for molecular ions and $N_{e,0}$ is the electron density at $z = 0$

$$N_{e,0} = \sqrt{\frac{q_0}{\alpha}} \tag{9.22}$$

and the maximum electron density is given by

$$N_{e,\max}(\chi) = N_{e,0} \sqrt{\cos \chi} \,. \tag{9.23}$$

Figure 9.3 shows the ionosphere profile from CHAMP GPS measurements given in terms of the first-order ionosphere delay for P_1 code in the zenith direction along the CHAMP orbit during the last ″normal″ solar maximum. We see that the effect of the ionosphere is significantly reduced for LEO orbits above 400 km . On the other hand, if ionosphere maps provided by IGS are corrected for the LEO altitude (applying a constant bias), we see that the agreement with observed TEC values from CHAMP is excellent and is at the level of about 1 TECU or below. Larger deviations can only be expected when a satellite is passing the equatorial anomaly.

In (Montenbruck and Gill 2002), the following model is given to calculate the fractional TEC above a LEO orbit. The coefficient α is given as a scaling factor and reads as
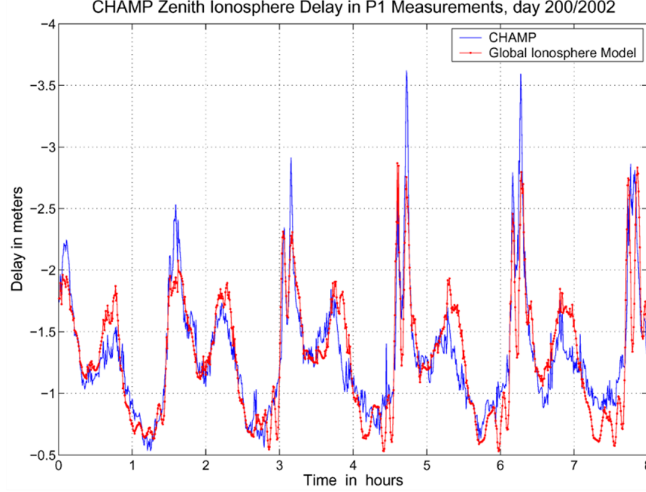
Figure 9.3 Ionosphere profile estimated using CHAMP P_1 and P_2 code measurements during solar maximum as a delay on P_1 in the zenith direction along the CHAMP orbit (red), in comparison with the global iono-maps estimated by IGS and corrected for the LEO altitude applying a constant bias (blue). Agreement with the IGS TEC maps is at the level of 1 TECU or below. Larger variations are due to the equatorial anomaly.

$$\alpha = \frac{e - e^{1 - e^{-z_{IP}}}}{e - e^{1 - e^{h_0/H}}} \tag{9.24}$$

Figure 9.4 shows the ionosphere profile along the GRACE orbit in terms of the geometry-free linear combination P_4 with and without applying the fractional TEC model (9.24) from (Montenbruck and Gill 2002). One can see that agreement is not as good as in Figure 9.3. Therefore, it is proposed to refine this fractional TEC model including the zenith angle χ with respect to the Sun $f\big(\text{Sun position}\big)$, e.g., including $\cos\chi$

$$\alpha = \frac{e - e^{1 - e^{-z_{IP}}}}{e - e^{1 - e^{h_0/H}}} \; f\big(\text{Sun position}\big) \tag{9.25}$$

A closer look at the Chapman function (9.21) that describes the shape of the Chapman layer (vertical TEC profile of ionosphere), shows that correct calculation of the fractional TEC above or below the LEO orbit altitude involves solution of the following integral

$$\int_{z_{LEO}}^{\infty} e^{\frac{1}{2}(1 - z - \sec\chi\, e^{-z})} dz \tag{9.26}$$

We can show that (9.26) can be reduced to the "error function" $erf(x)$ well known in statistics

$$\int_{z_{LEO}}^{\infty} e^{\frac{1}{2}(1 - z - \sec\chi\, e^{-z})} dz = \sqrt{e\pi\cos\chi} \cdot erf\left(\frac{e^{-\frac{z}{2}}}{\sqrt{\cos\chi}}\right) = \sqrt{e\pi\cos\chi} \cdot erf(\mathrm{x}) \tag{9.27}$$

The "error function" in integrating the normal distribution is given as, e.g., (Bronstein and Semendjajew 1996)

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \left(x - \frac{x^3}{3} + \frac{x^5}{10} - \cdots \right) = \frac{2}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k \, x^{2k+1}}{k!\,(2k+1)} \tag{9.28}$$

We see that the fractional TEC along the LEO orbit can be calculated exactly from the Chapman function.
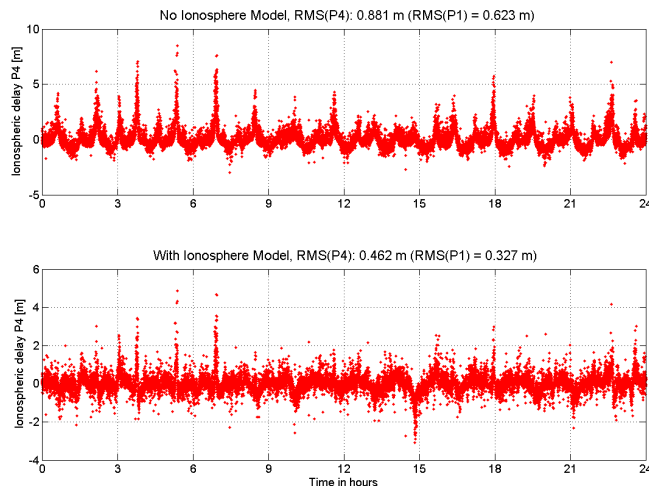
Figure 9.4 Ionosphere profile (GRACE-B) estimated using P_1 and P_2 code in terms of the geometry-free linear combination P_4, with and without applying the fractional TEC model (9.24) (agreement is not as good as in Figure 9.3).

9.4 Proposal for A Novel Remove-Restore Approach for Ionosphere/Plasmasphere Modelling with LEO Satellites Based on Least-Squares Collocation and Four Chapman Layers

Global ionosphere TEC maps provided by IGS have considerable spatial and temporal deficiencies due to the irregular distribution of the ground IGS stations (e.g., low density over oceans, polar caps and in the Southern hemisphere in general). This is especially true for the IGS real-time network, considering the recent attempts by IGS to provide ionosphere maps in real-time. In the combination of LEO and ground TEC measurements for the generation of final IGS TEC maps or the augmentation of real-time TEC maps, one could use the remove-restore approach. This would be similar to the remove-restore approach used in geoid determination by least-squares collocation. For both the final and real-time IGS maps, observed TEC obtained from ground IGS receivers and/or along the LEO orbit (for the final IGS maps only) is "removed" or reduced by employing a background ionosphere model, such as IRI2010 or NeQuick-2. In the second step, the derived LEO and ground-based TEC residuals are then modeled and properly combined using least-squares collocation. In the third step, the ionosphere model is "restored" to the reduced and combined TEC measurements. This proposed remove-restore method would augment the real-time IGS TEC maps with the background ionosphere model and combine LEO and ground IGS measurements for the final TEC maps. Such a remove-restore approach could especially improve spatial and temporal resolution of TEC maps in the regions where ground based or space based TEC observations are insufficient, e.g., oceans, polar caps or the Southern hemisphere in general. Compared to ground TEC measurements, a LEO orbit is typically placed within the ionosphere, thus the main challenge in the combination of LEO and ground TEC measurements is how to correctly account for the fractional TEC of the ionosphere below the orbit altitude, (see Figure 9.2). However, homogeneous and isotropic covariance functions used in least-squares collocation are designed to clearly distinguish between signal and noise in the data combination and filter out geographically correlated errors allowing the consistent combination of LEO and ground TEC measurements over the entire sphere. Here homogeneous means that statistical properties of the combination are preserved uniquely over the entire sphere and isotropic means over all azimuths. This is typically achieved by the design of the covariance function. For more on least-squares collocation we refer to (Moritz 1980).

Compared to all other LEO satellites equipped with a GPS receiver, GOCE with an altitude of 240 km is placed in a very low LEO orbit below the Chapman height (about 450 km altitude) and thus is a good candidate for studying improvements in the IGS global ionosphere maps, combining total electron content derived from the GOCE orbit and the ground TEC measurements. The GOCE satellite performs 16 revolutions per day around the Earth and thus it is expected that the future ionosphere products provided by IGS will be based on a combination of ground- and space-based LEO GNSS measurements including in addition about 100 GNSS satellites of the GPS, GLONASS, Galileo and Beidou constellations.

Inclusion of GOCE, GRACE and Jason-2 data can considerably improve modeling of the layered structure of the ionosphere, considering that GOCE TEC measurements at 240 km altitude represent almost the complete effect of the ionosphere, whereas TEC measurements taken by Jason-2 above the 1300 km altitude are mainly driven by the plasmasphere. Therefore, we propose to model the ionosphere as a spherically-layered electron density distribution in three main Chapman layers, i.e., E, F1 and F2 and an additional layer for the plasmasphere density above the ionosphere. The GOCE orbit is located above the E layer and below the F1 and F2 layers, the GRACE orbit is above the F1 layer and below the F2 layer, whereas Jason-2 is above the E, F1 and F2 layers, just above the ionosphere, where the plasmasphere starts. Therefore, there is great potential in combining these three missions with ground IGS measurements in constructing a layered model of the ionosphere.

At the moment, GPS measurements provided by the GOCE GPS receiver are solely used for kinematic and reduced-dynamic precise orbit determination. Here we are proposing applications of the GOCE GPS data in other scientific disciplines. In particular, applications to enhance IGS products and to study potential application of the LEO GPS measurements for the Space Weather segment section of ESA's Space Situational Awareness (SSA) Programme. The SSA Programme is based on the following three areas: 1) Space Surveillance and Tracking; 2) Near Earth Objects; 3) Space Weather. Ionosphere monitoring is one of the components of the Space Weather section of ESA's SSA Programme and GOCE GPS data could help to answer the question of how GPS measurements from the LEO satellites could improve the temporal and spatial resolution of the global ionosphere models. The GOCE satellite is an excellent candidate for such a study, since the error in TEC reduction from the very low GOCE orbit to the location of the ground TEC measurements is not as significant as that in TEC reduction from other LEO missions.

Typically, TEC maps provided by the CODE IGS AC are calculated in terms of a single-layer model represented by a spherical harmonic expansion in a frame defined by the axis of the geomagnetic pole and the geomagnetic equator w.r.t. the position of the Sun, (see (Schaer 1999)). Thus, instead of using geographic latitude, geomagnetic latitude is used, calculated for the intersection point of the line of sight with the single layer (ionospheric pierce point). Instead of geographic longitude, the Sun-fixed longitude of the ionospheric pierce point is used w.r.t. the longitude of the Sun. This rotation from an Earth-system, where TEC measurements and coordinates of ground stations are given, to the Sun-fixed geomagnetic coordinate system can be performed at the level of spherical harmonic coefficients by a rotation about the polar axis by an angle α

$$\begin{bmatrix} S_{n_{\max}m}(\alpha) \\ \dots \\ S_{mm}(\alpha) \\ C_{mm}(\alpha) \\ \dots \\ C_{n_{\max}m}(\alpha) \end{bmatrix} := \begin{bmatrix} S_{n_{\max}m} \\ \dots \\ S_{mm} \\ C_{mm} \\ \dots \\ C_{n_{\max}m} \end{bmatrix} \cos m\alpha + \begin{bmatrix} C_{n_{\max}m} \\ \dots \\ C_{mm} \\ -S_{mm} \\ \dots \\ -S_{n_{\max}m} \end{bmatrix} \sin m\alpha \tag{9.29}$$

where the coordinates of the Geomagnetic Pole $\left(x_{mp}(t), y_{mp}(t)\right)$ are related to the transformed coefficients as

$$\begin{bmatrix} \overline{C}_{21}(t) \\ \overline{S}_{21}(t) \end{bmatrix} := \begin{bmatrix} \sqrt{3}\overline{C}_{20} - \overline{C}_{22} & \overline{S}_{22} \\ -\overline{S}_{22} & -\sqrt{3}\overline{C}_{20} - \overline{C}_{22} \end{bmatrix} \begin{bmatrix} x_{mp}(t) \\ y_{mp}(t) \end{bmatrix} \tag{9.30}$$

In this case, global TEC mapping could be performed with station coordinates in an Earth-fixed terrestrial frame. For more on rotation of spherical harmonic coefficients, see Section 25. In the same way, rotations of spherical harmonic coefficients could be directly introduced as parameters of TEC maps. Therefore, instead of calculating a set of spherical harmonic coefficients every, e.g., 2 hours, one could calculate a set of rotations $\alpha = \alpha(t)$ for initial spherical harmonic coefficients given for a period of one day. In that case, one could produce a continuous transition of temporal ionosphere maps over one day or longer. That is not the case now, where every TEC map is calculated separately and there is no smooth transition between those maps. To our knowledge, only CODE Analysis Center uses constraints between the TEC maps, but one still needs to use an interpolation method to obtain the TEC value between the two TEC maps.

9.5 Tropospheric Refraction and Low-Order Zonal Gravity Field Coefficients from LEO Orbits. Is There a Connection?

A LEO orbit is located high above the troposphere and therefore only ionospheric effects are relevant in determining a given LEO orbit. However, since GPS satellite orbits, and especially GPS satellite clock parameters, are estimated by means of the ground GPS network, the troposphere has an indirect impact on LEO orbit determination and subsequently on the estimated gravity field and altimetry results. We have noticed that with the kinematic orbits of CHAMP and the low-order zonal gravity field coefficients. It was reported for the first time by (Mayer-Gürr et al. 2006) that some of the solutions of CHAMP kinematic orbits show very significant differences in low-order zonal gravity field coefficients, namely J_2, J_4 and J_6. The error was above the error-bars one would expect and was significant compared to the first GRACE gravity models. Over several years we were very puzzled as to what the root cause was and the background effect. Looking at the affected low order zonal gravity field coefficients, they define the shape of the Earth's gravity field, i.e., the flattening of the rotational ellipsoid and zonal effects along the parallels (e.g., at mid latitudes). The only effect that is similar is that of troposphere gradients, where a strong north-south component follows the shape of the troposphere (flattening at the poles). As with the Earth's gravity field, the troposphere also flattens at the poles, having a typical maximum height of $18-20$ km above the equator and $8-9$ km above the poles. In the calculation of phase clocks for GPS satellites we did not take into account tropospheric gradients, although they have a strong effect on GPS signals at $10°$ elevation and below. It was assumed that processing GPS measurements from the IGS network above $10°$ elevation and estimating tropospheric zenith delays as piece-wise linear functions every hour, would be sufficient to properly model the effect of the troposphere on ground GPS measurements. However, we did not take into account any data below $10°$ elevation or tropospheric gradients, which basically model the troposphere mapping function as a function of azimuth. This unisotropical effect caused by this chosen model of tropospheric refraction affected GPS satellite clock parameters and subsequently LEO kinematic POD. In the case of reduced-dynamic orbits or gravity field modeling based on dynamic orbits, the effect is coupled with the once-per-rev. empirical accelerations that are typically estimated in dynamic POD (and could partially remove it), but not in kinematic POD.

9.6 An Overview of Tropospheric Effects on Microwave and Optical Measurements

This overview is fully based on the existing literature, see e.g., IERS Conventions (Petit and Luzum 2010) and given here for the sake of completeness, thus readers familiar with the topic may wish to forego this

summary. The atmosphere is a layer of gases surrounding the Earth that is held in place by the Earth's gravity field. The Earth's atmosphere has several layers that differ in properties such as temperature, pressure and composition that extend from the troposphere (the lowest layer up to some 10 km), to the stratosphere, mesosphere, thermosphere, up to the exosphere that includes ionosphere and plasmasphere.

Atmospheric refraction is the main accuracy-limiting factor in all microwave space-based geodetic techniques such as GPS, DORIS, VLBI and satellite altimetry. This is also true, to a great extent, for optical space-based geodetic techniques, such as SLR, that are also influenced by range biases. Moreover, for kinematic POD, tropospheric refraction and ground station multipath are the main sources of error in determining GPS satellite clock parameters and consequently LEO kinematic orbit. Therefore, we decided to give here an overview of the state of the art in the modeling of tropospheric refraction and to propose improvements.

The troposphere is non-dispersive for radio signals with frequencies up to 40 GHz. Due to the refractive index and its variation within the troposphere, microwave signals are delayed. The same is true for a laser pulse transmitted and received by a SLR telescope. Typically, the total delay of the radio signal is divided into "hydrostatic" and "wet" components. The hydrostatic delay is caused by the refractivity of the dry gases in the troposphere and by the non-dipole component of water vapor refractivity. The main part, (about 90%) of the total delay, is caused by the hydrostatic component and can be very accurately predicted using surface pressure data. The dipole component of water vapor refractivity is responsible for the wet delay and amounts to about 10% of the total delay. This corresponds to $5 - 40$ cm (max.) for very humid conditions. A mapping function is used to transform the zenith tropospheric delay to the elevation of each observation. In recent years, the so-called Niell Mapping Function (NMF) has become the standard for the processing of microwave measurements. It is based on one year of radiosonde profiles, primarily from the Northern hemisphere (Niell 1996). In order to improve accuracy, it was recommended that troposphere mapping functions based on data from numerical weather models (NWM), such as ECMWF (European Centre for Medium-Range Weather Forecasts) be used. They provide the spatial distribution of refractivity throughout the troposphere with high temporal resolution. Today, these mapping functions (e.g., Vienna Mapping Function - VMF1 (Boehm et al. 2006b) or IMF (Niell 2001)) are available as time series of coefficients with a resolution of six hours ((Boehm et al. 2006a)). As an alternative, if NWM-based mapping functions are not available for a particular station or period of time, the global mapping function (GMF) can be used without introducing systematic biases (in the coordinate time series), see (Boehm et al. 2006a). The GMF is a compatible empirical representation of the more complex NWM-based mapping functions, the differences being mainly in short-term precision. The GMF provides better precision than the NMF and smaller height biases with respect to VMF1 (Boehm et al. 2006a). VMF1 is currently the mapping function providing globally the most accurate and reliable geodetic results. However, systematic station height changes of up to 10 mm occur when changing from NMF to VMF1 (Boehm et al. 2006a).

Traditionally, the correction of the tropospheric delay at optical wavelengths has been performed using the formulation of (Marini and Murray 1973), a model developed for the 0.6943 μm wavelength (McCarthy and Petit 2004). The model formulated in (Hulley and Pavlis 2007) is now the standard zenith delay model, the so-called M-P model, for modeling the refraction of SLR measurements and is valid for a wide spectrum of wavelengths ($355 - 1064$ nm) with sub-mm accuracies. The accompanying mapping functions (FCULa and FCULb) published in (Mendes et al. 2002) showed a 2-year average RMS (model minus ray tracing through radiosonde data) of approximately 7 mm at 10° elevation (Hulley and Pavlis, 2007). However, these are models based on an unrealistic spherically symmetric atmosphere neglecting contributions from horizontal refractivity gradients around the SLR tracking sites. (Hulley and Pavlis, 2007] addressed the contribution of horizontal refractivity gradients to the computation of the total tropospheric delay for SLR measurements by direct ray tracing through three-dimensional atmospheric fields generated using AIRS and NCEP data. AIRS stands for the Atmospheric Infra-Red Sounder instrument on NASA's AQUA Earth Observing System (EOS) platform. They calculated horizontal gradient delays at any selected azimuth and elevation angle for 10 of the most prolific, globally distributed ILRS stations during 2004 and 2005. They showed that AIRS North-South

(NS) and East-West (EW) gradients have annual means of between 1 and 4 mm in absolute magnitude at 10° elevation. The NS component had larger standard deviations ranging from 6 to 12 mm, while the standard deviations of the EW component were between 5 and 9 mm at all the stations analyzed. Maximum NS gradient delays of up to 50 mm were found at Yarragadee (Australia) and Herstmonceux (UK) at 10° elevation. They found that the largest variations occur as a result of seasonal and diurnal changes. Stations situated in mountainous regions, such as McDonald and Monument Peak, had larger horizontal pressure gradients, while stations in close proximity to large bodies of water (for example, Yarragadee) had larger horizontal temperature gradients. No significant non-hydrostatic (wet) gradients were found, with maximum wet delays only reaching a few tenths of a millimeter during the summer at Greenbelt. They found that the gradient delays decreased by a factor of 3 from 10° to 20° elevation and were at sub-mm levels at higher elevation angles. The NS and EW gradients varied primarily by station location and time of year. Gradient variations in the NS and EW directions increased from winter to summer at Yarragadee and Monument Peak and from summer to winter at Herstmonceux and Zimmerwald. By using uncertainties in the most recent AIRS validation results, they were able to estimate error variations in the gradient delay results. They found monthly RMS differences (original minus simulated data) of less than 5 mm for an elevation angle of 10° at Herstmonceux and Yarragadee. Actual day-to-day variations in the gradients were larger and ranged from 7 to 14 mm. The effects of replacing the M-P delay model by ray-tracing results in order to calculate the total tropospheric correction (including gradients) resulted in a reduction in the variance of the SLR observation residuals for LAGEOS 1 and 2 of 25–43% for NCEP and 10–30% for AIRS during 2004 and 2005. They concluded that NCEP had much larger biases than AIRS at most stations, and an optimum solution would need to be developed (e.g., using ECMWF) in order to extract the best results for future corrections, see (Hulley and Pavlis, 2007).

Compared to data relying on microwave technology, the two main advantages of SLR measurements are, firstly, that they are free from first- and higher-order ionospheric effects, and, secondly, that water vapor delays can easily be modeled. The signal delay due to refraction by the water vapor in the atmosphere is significantly different in the optical and in the microwave band. The ratio is about 67:1, meaning that a typical "wet component" in the zenith direction of about $5 - 40$ cm for the microwave band (GPS) corresponds to a delay of about $1 - 6$ mm for SLR observations. Since the effect is relatively small, about 80% of the delay can be modeled by using surface pressure, temperature and humidity measured at the station.

Atmospheric water vapor is the dominant greenhouse gas in the Earth's atmosphere, and quantifying the feedback of water vapor in global warming is therefore of paramount importance, (Bengtsson et al. 2003). The lack of detailed knowledge of the hydrological cycle is thus a major factor limiting a better understanding of the Earth's climate system. The inaccuracy is substantial and concerns practically all aspects of the hydrological cycle (Bengtsson et al. 2003). Recently, GNSS-based measurements have offered new and promising possibilities. The global IGS network and dense regional GNSS networks have been developed around the world, and these provide highly temporal and spatial information (e.g., up to 20 km) about the integrated atmospheric water vapor; vertical profiling using the GPS radio-occultation technique is similarly taking place, using satellites in LEO orbit. Tropospheric zenith delays are estimated on a regular basis using regional GPS networks and the global IGS ground network. These are then used to assimilate and constrain numerical weather models, (see e.g., (Guerova et al. 2006)). However, GPS networks provide total zenith delay, and the water vapor information is extracted using models.

9.7 The Way Forward in High-Resolution Modeling of Tropospheric Delays for all Space Geodesy Techniques

From the overview of tropospheric modeling approaches given above, a number of possible improvements spring to mind. Use of water vapor radiometers is the way forward, although these instruments have not yet found operational application in space geodesy. This is due to their inability to consistently deliver tropospheric delays in all directions in the field of view and especially in all weather conditions (e.g., rain). On the other hand, the use of numerical weather models can improve spatial and temporal resolution of the background troposphere model (e.g., for troposphere mapping or ray-tracing). However, the state-of-the art numerical weather models still have a temporal resolution of several hours and cannot represent the effect of the troposphere to a spatial resolution of below some $20\,\mathrm{km}$. We know that the water vapor content can change significantly within about $30\,\mathrm{min}$ and over several kilometers in terms of spatial resolution.

From Section 15.2 we will learn that in just a few years from now one can expect the operation of four complete GNSS constellations (GPS, GLONASS, Galileo, and Beidou) providing more than 35 GNSS satellites in the field of view 10° above the horizon. This opens up the possibility of modeling station-specific tropospheric delays in terms of temporal spherical harmonics or spherical grids. In this way, fine structures in the station troposphere and multipath could be mapped at the same time on the reference sphere placed around the ground station. This will be very similar to estimating a PCV map for the specific location of a ground station. Such an approach will be feasible, since more and more GNSS satellites are equipped with high-performing satellite clocks that allow modeling of GNSS clock parameters with a simple linear model over a one-day period, (see (Svehla 2010a) or Section 20). The same trend can be seen in the inclusion of H-masers in the ground IGS network. Therefore, it is to be expected that Galileo will require modeling of tropospheric delays to a significantly higher resolution in order to fully benefit from the short- and long-term stability of the on-board H-maser. To understand why, one just has to consider that at each epoch about 8 different ground stations contribute to the estimation of a single GNSS clock parameter. Thus, the residual tropospheric effect is averaged over those 8 different stations every epoch and the noise is much higher than carrier-phase noise or noise from the ″instability″ of the Galileo H-maser, (see, e.g., Section 18). In addition, the tracking geometry changes slightly from epoch to epoch and new ground stations enter this averaging process typically at very low elevations. For validation of the H-maser on board GIOVE-B using SLR measurements see (Svehla 2010a) and for the first Galileo FOC satellites Section 20. However, one should always consider correlations of any additional parameters with the station coordinates, if they are estimated in the same processing run, especially the station height.

Once, high-resolution troposphere maps are being provided by ground GNSS stations, other space geodesy techniques, such as SLR, VLBI and DORIS could use those maps to accurately account for very small changes in tropospheric delays. In almost all cases, GNSS receivers are co-located with all other space geodesy sensors in very close proximity, so such an approach is already feasible. Combination with space geodesy techniques could also bring an added value. On the other hand, rotation of spherical harmonics can very efficiently account for any temporal variations in modeling of station-specific tropospheric delays. In the scope of this thesis we have developed a new technique for the rotation of spherical harmonics that can be used for the modeling of temporal variations represented by spherical harmonics (e.g., gravity field, ionosphere maps, troposphere maps).

10. Aerodynamics in Low LEO: A Novel Approach to Modeling Air Density Based on IGS TEC Maps

Here we present some theoretical aspects of the modeling of aerodynamic acceleration in the precise orbit determination of a LEO satellite. We have included this section because of the great importance of the role that aerodynamic drag plays in all gravity field missions, as they are typically placed in a very low LEO orbit. Thus, here we look at the geometrical properties of this effect. We show that the accuracy of the velocity in the calculation of the aerodynamic drag for a LEO satellite, in particular the velocity of thermospheric horizontal winds, is as important as the atmospheric density. We then give a geographical representation of the models used to calculate atmospheric density and thermospheric horizontal winds, with an emphasis on the GOCE (Sun-synchronous) orbit, and compare this with the orbits of altimetry satellites in high LEO. In addition, we present the prospects of investigating atmospheric density and thermospheric winds using the GOCE mission at $220 - 250$ km altitude. Models of neutral horizontal winds show that thermospheric winds mainly occur around the geomagnetic poles where they are driven by the perturbations in the geomagnetic field. The highest thermospheric wind velocities may be expected along the dawn-dusk regions, and from that point of view, the GOCE orbit is the perfect candidate to provide unique information on the neutral horizontal winds in the lower thermosphere. Section 10.3 of this thesis triggered an ESA study that demonstrated the retrieval of thermospheric wind parameters from GOCE data. At the end of this section, we demonstrate a novel approach to calculating and predicting air density in the thermosphere based on the global TEC maps provided by IGS. This approach could be used to predict solar activity in an alternative way, independent of the number of Sun spots or the solar flux index at a wavelength of 10.7 cm ($F_{10.7}$). We also show that information on the ionization of the thermospheric part of the ionosphere, as provided in IGS TEC maps, can be used to calculate the LEO mission duration (as was done for GOCE). This opens up new applications for the global IGS TEC maps in monitoring air density in the thermosphere, including spatial and temporal variations. In addition, we show that variations in air density driven by variations in solar activity (heating) are empirically proportional to the ionization of the ionosphere. Thermospheric density and TEC can be related by an empirical linear model as shown here.

10.1 Aerodynamic Drag

Aerodynamic drag is the most significant non-gravitational force acting on a satellite in low LEO orbit. Atmospheric density decreases exponentially with increasing orbit altitude and, as a result, aerodynamic drag becomes negligible at the outer boundary of the thermosphere (≈ 1000 km). Due to the energy dissipation caused by air resistance, natural orbital motion below 120 km orbit altitude cannot be sustained in the Earth's atmosphere and so is followed by orbital re-entry. The ESA mission GOCE, in Sun-synchronous orbit, uses a dedicated electric ion propulsion system to counteract aerodynamic drag and to maintain the satellite orbit at $220 - 250$ km altitude. Thus, the duration of the GOCE mission is limited by the capacity of the 40-kg tank

of xenon on board, as xenon is used for propulsion. Neutral xenon atoms are converted into fast-moving ions by an electric discharge generated by the satellite's photo-voltaic panels. The ions are then ejected aft of the satellite giving a very smooth thrust of $1-20$ mN, depending on the measured drag in the along-track direction. On the orbit determination of the International Space Station, see (Shum et al. 2008, 2009).

The aerodynamic acceleration of the satellite due to air drag reads as, e.g., (Montenbruck and Gill 2000)

$$\ddot{\vec{r}} = -\frac{1}{2} c_D \frac{A}{m} \rho \dot{r}_r \dot{\vec{r}}_r \tag{10.1}$$

with ρ denoting the air density, c_D the empirical drag coefficient, A/m is the so-called form factor or the aerodynamic reference cross-section with satellite mass m and satellite velocity $\dot{\vec{r}}_r$ relative to the atmosphere (assuming that the atmosphere co-rotates with the Earth). Air drag acceleration can easily be derived by considering the linear momentum of a small mass element of a column of the atmosphere that hits the satellite's cross-sectional area. (For more details see, e.g., (Montenbruck and Gill 2000)). This is the reason why the acceleration of the satellite due to air drag is directly proportional to the square of the relative velocity. The relative velocity or free-stream flow velocity, as a function of the satellite velocity $\dot{\vec{r}}$, is

$$\dot{\vec{r}}_r = \dot{\vec{r}} - \vec{\omega}_\oplus \times \vec{r}_\oplus - \dot{\vec{r}}_{HW} \tag{10.2}$$

with $\vec{\omega}_\oplus$ denoting the Earth's angular velocity vector and $\vec{r}_\oplus$ is the satellite position in the Earth-fixed frame. In (10.2), we have included, in addition, the thermospheric horizontal wind velocity denoted as $\dot{\vec{r}}_{HW}$. The second term in (10.2) assumes that the entire atmosphere co-rotates with the Earth and the third term models more closely the real dynamics of the atmosphere, making use of the model for horizontal neutral winds in the upper thermosphere.

In the case of a more refined model, the satellite surface can be considered as an array of finite elements, where the surface element A_k has a corresponding drag coefficient $C_D(k)$. By introducing the normal vector for each surface element $\vec{A}_k$, with the length set to the actual surface element area, we can derive the refined model for the aerodynamic acceleration of the satellite due to aerodynamic drag, giving

$$\ddot{\vec{r}} = -\frac{1}{2} \frac{\rho}{m} \dot{\vec{r}}_r \sum_k C_D(k) \vec{A}_k \dot{\vec{r}}_r \tag{10.3}$$

Accuracy of the satellite velocity relative to the atmosphere is limited by the complex atmosphere dynamics modeled by the horizontal wind models, see Figure 10.2 and Figure 10.3. Estimation of the absolute velocity of the atmosphere is at least five orders of magnitude less accurate than determination of the actual satellite velocity (≈ 0.01 mm/s). Aerodynamic drag modeling is mainly limited by the accuracy of the models for atmospheric density and neutral thermospheric wind velocity as well as by the drag coefficients that describe the interaction of the atmosphere's constituents with the satellite surface. These limitations can be reduced by empirical orbit modeling, i.e., by estimating frequent air drag coefficients and other empirical parameters.

Comparing the kinematic and reduced-dynamic orbits of the CHAMP satellite it was demonstrated that aerodynamic-drag could only have a smooth effect on POD with very long periodicity. It can therefore easily be removed by estimating empirical parameters (pseudo-stochastic pulses) allowing cm- level orbit accuracy to be achieved (Švehla and Rothacher 2002). However, in comparison with the kinematic CHAMP orbit, the remaining systematic errors in the along-track of the reduced-dynamic orbit can easily be identified in the polar regions (Švehla and Rothacher 2002), (Švehla and Rothacher 2005b). These are regions where the dynamics of the atmosphere is very complex and larger errors in the thermosphere wind/density models can be expected. (Bruinsma et al. 2003) compared methods to model acceleration for the CHAMP satellite and showed that the level of geomagnetic activity is highly correlated with the atmospheric drag model error, and that the largest errors occur around the geomagnetic Poles.

The size of the drag coefficient depends, firstly, on the flow conditions which are characterized by the Mach, Reynolds and Knudsen numbers, and, secondly, on the scattering mechanisms taking place at the satellite surface, such as specular, elastic and diffuse reflections. For LEO orbits, the satellite is in the free molecular flow regime, which means that the incident flow is undisturbed by the satellite moving through it, i.e., particles re-emitted from the surface of the satellite do not interfere with the incident flow. A typical drag coefficient for LEO free molecular flow that one can find in the relevant literature, is in the order of $2 - 2.3$. As the orbit altitude decreases, air density increases exponentially and the satellite moves from a free molecular flow regime into intermolecular collision flow and finally into continuum flow. A typical value for the aerodynamic drag coefficient in this transitional flow regime (below 200 km) is about 1.0, however, and the increased air density causes orbital re-entry of the satellite.

10.2 Geographical Representation of Atmosphere Density and Thermospheric Horizontal Wind Models

Thermospheric density models play an important role in POD, orbit predictions, orbital station keeping maneuvers, ground-track maintenance, collision risk analysis and orbit reentry predictions. In order to model aerodynamic drag, we employed the NRLMSIS-00 atmosphere density model (Picone et al. 2002) along with the thermospheric horizontal wind model HWM93 (Hedin et al., 1996). NRLMSIS-00 is the recent major upgrade of the MSISE-90 model of the thermosphere (Picone et al. 2002). The MSISE-90 model is a revision of the MSIS-86 empirical model (Hedin, 1987) of the lower thermosphere extended into the mesosphere and lower atmosphere taking into account data derived from space shuttle flights and from incoherent scatter radar (Hedin, 1991). Compared to MSISE-90, the NRLMSIS-00 model is based on the following data: (1) total mass density from satellite accelerometers and from orbit determination (including the Jacchia and Barlier data sets), (2) temperature from incoherent scatter radar covering the years 1981–1997, and (3) molecular oxygen number density, from solar ultraviolet occultation aboard the Solar Maximum Mission (Picone et al. 2002). A new component, "anomalous oxygen", allows for appreciable O^+ and hot atomic oxygen contributions to the total mass density at high altitudes and applies primarily to drag estimation above 500 km (Picone et al. 2002). The same paper reports a large O^+ contribution to the total mass density when there is a combination of summer, low solar activity, high latitude, and high altitude. Under these conditions, except when there is very little solar activity, the Jacchia-70 model shows a significantly higher total mass density than does MSISE-90. However, under the corresponding winter conditions, the MSIS-class models represent a noticeable improvement relative to Jacchia-70 over a wide range of solar fluxes. Considering the two regimes together, NRLMSISE-00 achieves an improvement over both, MSISE-90 and Jacchia-70, by incorporating advantages from both (Picone et al. 2002).

Figure 10.1 shows the air density for a sphere placed at 250, 500, 700 and 1000 km above the Equator at 12 UT. The solar flux F10.7 was set to 150 and the Ap indices to 4 as approx. values for day 200/2003. The maximum density at 250 km altitude occurs two hours after the local noon around the geomagnetic equator, whereas for higher altitudes this maximum is shifted to the south-east. From Figure 10.1 one can draw the conclusion that the air density at 250 km is about one order of magnitude higher than that at an altitude of 400 km, three orders of magnitude higher than that at 700 km and about four orders of magnitude higher than that at 1000 km. In a Sun-synchronous orbit the satellite is not exposed to maximum atmospheric density. For the GOCE satellite placed in a Sun-synchronous orbit at an orbit altitude of 240 km, the main density perturbation is avoided. However, residual perturbations can be expected around the geomagnetic poles.

The thermospheric horizontal wind model HWM93 (Hedin et al., 1996) is a revision of the previous HWM90 model (Hedin et al., 1991) for the lower thermosphere and extended into the mesosphere, stratosphere and lower atmosphere to provide a single analytic model for calculating zonal and meridional wind profiles representative of the climatological average for various geophysical conditions (Hedin et al., 1996). Gradient winds from CIRA-86, plus rocket soundings, incoherent scatter radar, MF radar, and meteor radar provided
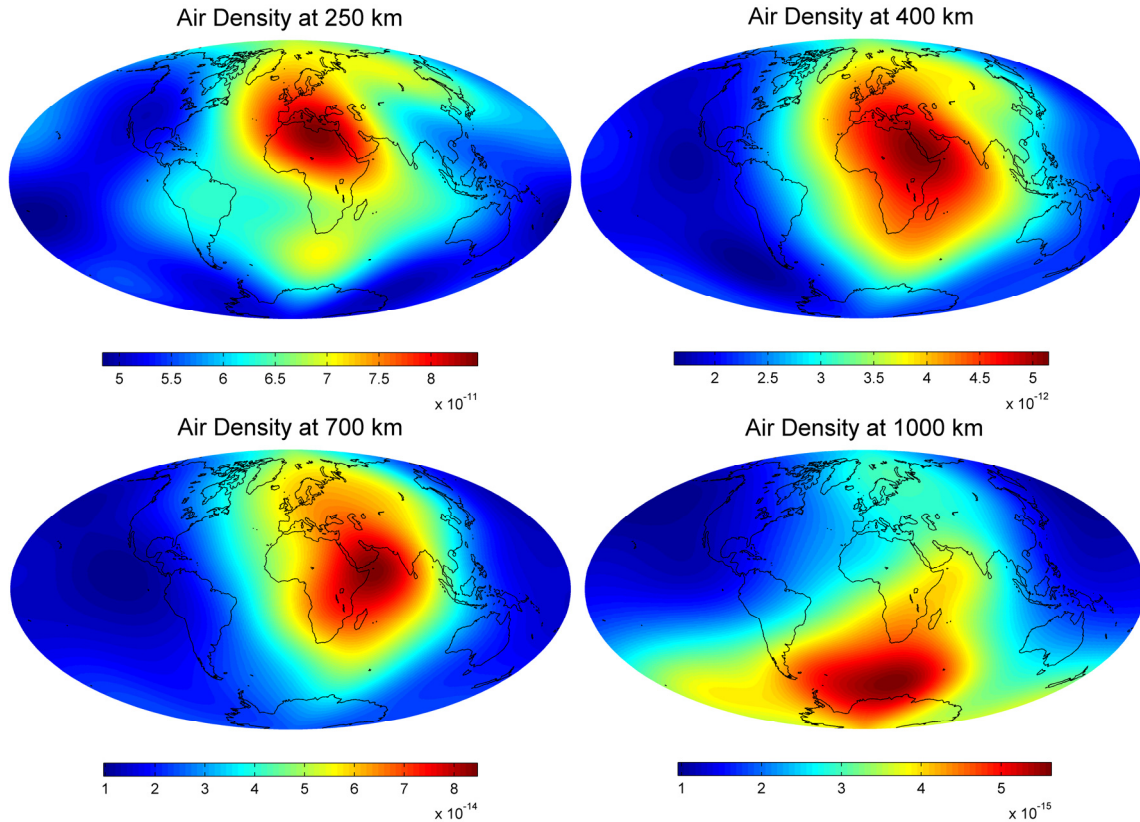
Figure 10.1 Air density in kg/m^3 based on the NRLMSISE-00 model for a sphere placed at 250, 500, 700 and 1000 km above the equator at 12 UT (F10.7 solar index was set to 150 (during last solar maximum) and Ap index to 4, day 200/2003). Comparing top and bottom figures, one can see that the density of the atmosphere is much higher below an altitude of 400 km and that for these altitudes, air density shows a geographical distribution similar to global TEC maps provided by the IGS.

the data base and were supplemented by previous models. Low-order vector spherical harmonics and Fourier series were used to describe the major variations in the atmosphere including factors such as latitude, annual, semiannual and local time (tides), and longitude, with a cubic spline interpolation in altitude (Hedin et al., 1996). The MSIS models are based on the so-called Bates-Walker temperature profile -- a function of geopotential height for the upper thermosphere and an inverse polynomial in geopotential height for the lower thermosphere. Exospheric temperature and other atmospheric quantities are expressed as functions of geographical and solar/magnetic parameters. The temperature profiles allow for exact integration of the hydrostatic equation for a constant mass to determine the density profile based on a density specified at 120 km as a function of geographic and solar/magnetic parameters (Hedin et al., 1996).

Although the agreement between various data sources was reported to be good, systematic differences were reported, particularly near the mesopause. RMS differences between data and the model values are of the order of 15 m/s in the mesosphere and 10 m/s in the stratosphere for zonal winds, and 10 m/s and 5 m/s, respectively, for meridional winds. (For more detail see (Hedin et al., 1996)) The output of the model are zonal and meridional wind components for altitudes from 0 km to 2000 km. Velocities of up to 1 km/s can be reached across the poles at altitudes of 300 km. In the vertical direction, the mean wind velocity is generally less than 1 cm/s and can be neglected for all applications.

Figure 10.2 shows the total horizontal thermospheric wind velocity in m/s based on the HWM93 model at 250, 500, 700 and 1000 km altitude at 12 UT. The same solar and geomagnetic parameters were used as in the computation of atmospheric density. The neutral horizontal wind model shows that thermospheric winds
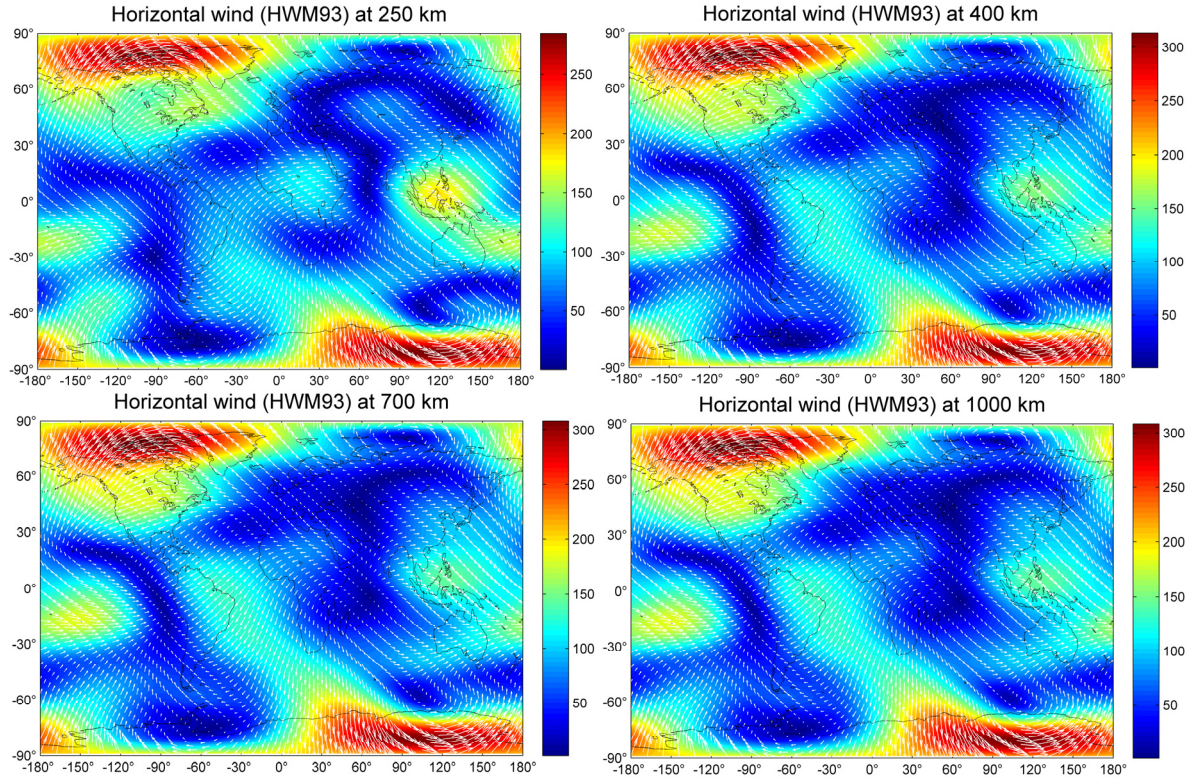
Figure 10.2 Total horizontal air velocity in m/s based on the HWM93 model at 250, 500, 700 and 1000 km altitude at 12 UT ($F_{10.7}$ index set to 150 and Ap indices to 4, approx. values for day 200/2003).

mainly occur around the geomagnetic poles, where they are caused by the perturbations in the geomagnetic field. The highest wind velocities may be expected along the dawn-dusk regions. At low latitudes, more stable (accurate) and moderate velocities are to be found and unlike with atmospheric density and the ionosphere, no correlation with the sub-solar point can be observed (as is the case with IGS TEC maps).

The thermosphere is the "LEO layer" of the Earth's atmosphere above the mesosphere and below the exosphere, where ultraviolet radiation causes ionization and the creation of the ionosphere. The exosphere is the uppermost layer of the atmosphere (roughly above 1000 km) and is sometimes used synonymously with outer space, since there is no clear boundary between the two. In the exosphere, a molecule can escape into space or can be pulled back to Earth by gravity with almost no probability of colliding with another molecule.

Figure 10.3 shows the neutral atmospheric density and horizontal velocity at 1300 km altitude (altimetry satellites such as T/P, JASON-1/2). One can see that atmospheric density is lower by a factor of 5 compared to an altitude of 1000 km, but horizontal winds show a very similar behavior to those in the lower thermosphere at altitudes of 250 or 400 km. Figure 10.3 confirms again that atmospheric winds are driven mainly by perturbations in the magnetic field and that atmospheric density is driven by the solar flux at a wavelength of 10.7 cm ($F_{10.7}$). Maximum air density occurs about 2 hours after the local noon and is placed close to the South Magnetic Pole. Both models for thermospheric density and models for thermospheric winds, are mainly driven by the solar flux index $F_{10.7}$ as an input and the mean solar flux over the previous three 27-day rotations of the Sun. Due to the interaction between the solar wind and the Earth's magnetic field, the geomagnetic field is perturbed and related variations in atmospheric density can be expected. Variations in the Earth's magnetic field are globally represented by the so-called (three-hourly) planetary geomagnetic index and its daily mean, often denoted as Ap.
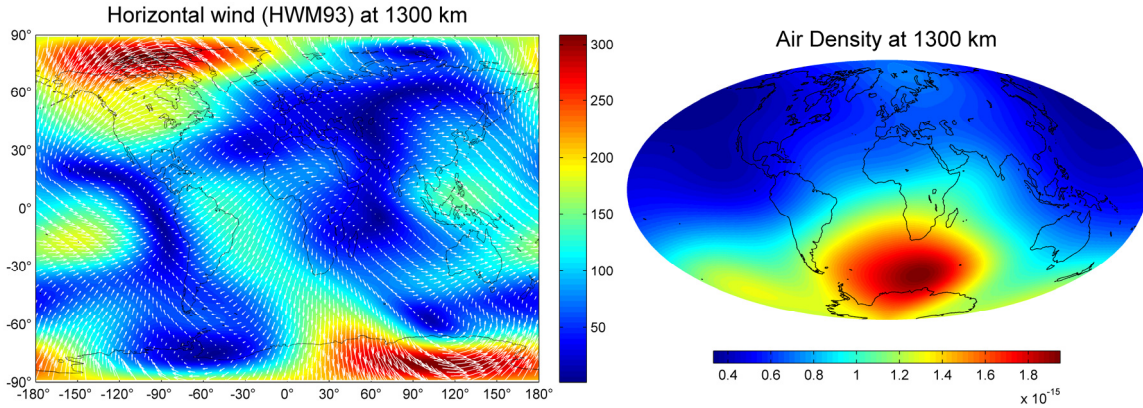
Figure 10.3 Total horizontal atmosphere velocity in m/s and air density in kg/m³ based on the HWM93 model at 1300 km altitude at 12 UT ($F_{10.7}$ index set to 150 and Ap indices to 4, approx. for 200/2003).

Figure 10.4 shows the mean observed solar radio flux at a wavelength of 10.7 cm over more than 60 years. One can clearly recognize the 11-year solar cycle. The same periods may be identified in the ionosphere maps provided by IGS and in the atmospheric density models that use the solar flux index as an input. The Sun emits radio energy that is driven by the layers high in the Sun's chromosphere and low in its corona, and the rate at which that energy is emitted changes in unison with the number of spot groups on the disk. By looking at the number of Sun spot groups on the Sun's disk we can identify the 27-day Sun rotation period. This rotation period can also be seen in the variations of the solar flux as reflected in the Total Electron Content shown on the IGS ionosphere maps or in the atmospheric density.

The solar flux density at 2.8 GHz corresponds to a wavelength of 10.7 cm and has been recorded routinely by radio telescopes. Figure 10.4 shows observed monthly means of the solar flux recorded since 1947 by the radio telescope near Ottawa and starting with June 1991, from Penticton, in Canada. The observed time series contain fluctuations that arise from the variations in the Sun-Earth distance over one year. Absolute solar fluxes are corrected and referred to the mean Sun-Earth distance. In addition, they are multiplied by 0.90 to compensate for uncertainties in the antenna gain and in waves reflected from the ground (NOAA 2009).
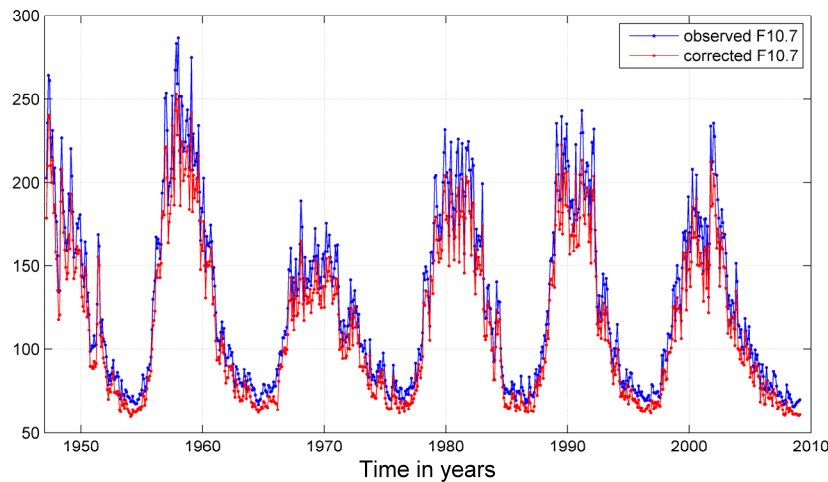
Figure 10.4 Observed and absolute (corrected to the mean Sun-Earth distance) solar flux at 10.7 cm wavelength (data source NRCAN).

10.3 Probing the Thermospheric Density and Thermospheric Horizontal Winds Using the GOCE Mission

During the writing of this thesis, this section triggered a dedicated ESA Study that demonstrated the use of GOCE data in examining of thermospheric horizontal winds.

The common mode of the GOCE accelerometers contains the signal of the non-gravitational forces acting on the satellite. However, acceleration in the along-track direction is counteracted by the electric ion-propulsion system. This, in turn, is controlled by the measurements from the accelerometers. Hence, they measure the near-zero drag acceleration in a closed loop. Therefore, thermospheric density can be derived mainly from the force that is applied by the ion-propulsion system on the satellite. Since the drag-free system is acting only in the along-track direction, GOCE accelerometers should be able to provide information on the horizontal crosswind velocity (in cross-track direction) since the ion-propulsion system does not counteract the effect of these on the satellite.

High-quality accelerometer measurements from the CHAMP and the two GRACE satellites in LEO orbit have shown that existing state-of-the-art thermospheric density and horizontal wind models such as JB2006 (Bowman et al. 2008), JB2008 (Bowman et al. 2008a), NRLMSIS-00 (Picone et al. 2002) and HWM93 (Hedin et al. 1996) contain systematic errors and their use in precise orbit determination has to be heavily supported by the estimation of empirical orbital parameters (pseudo-stochastic pulses, i.e., empirical velocities), see e.g., (Švehla and Rothacher 2005a). However, at the same time, air density provided by those models can easily be calibrated against the accelerometer measurements, providing very good predicted variations of the air density along the orbit. Thermospheric density models and solar radiation pressure at higher LEO altitudes are the main source of error in the precise determination and prediction of the orbits of LEO satellites, considering the high spatial and temporal resolution of the Earth's gravity field models available after the GOCE and GRACE mission.

Due to its Sun-synchronous, very low LEO orbit (only $220 - 250$ km altitude), GOCE could provide a new insight into non-gravitational forces acting on LEO satellites at altitudes of $220 - 250$ km . This is particularly true for forces related to air-density and horizontal winds in the lower thermosphere, but is also the case for those connected with other effects such as solar radiation and albedo. For instance, GOCE is the first LEO mission with highly sensitive accelerometers in a dawn-dusk Sun-synchronous orbit and could provide for the first time, a spectral characterization of solar radiation pressure, which, in the case of GOCE, acts approximately orthogonally to the aerodynamic drag.

Before the satellite gravity missions such as CHAMP, GRACE and GOCE equipped with highly sensitive accelerometers, launched over the last 10 years, there was very little high quality data available on thermospheric density and thermospheric winds. However, at times of low solar activity, and especially at the higher altitude of the GRACE and forthcoming Swarm satellites, the determination of thermospheric cross-winds is likely to remain much more challenging. Firstly, because of the reduced aerodynamic effect under those conditions and errors in the data calibration, and secondly, due to solar radiation pressure that is a more dominant effect at those altitudes.

On the other hand, GOCE could offer unprecedented information on air density and neutral horizontal winds in the thermosphere at very low LEO altitudes never investigated before. This region of the thermosphere is of special interest for research involving the orbital re-entry analysis of space objects as well as calibration of the air density models to be used at higher LEO altitudes in the upper thermosphere. For the GOCE satellite in a Sun-synchronous orbit, and for dawn-dusk orbits in general, the local mean solar time of passage for equatorial longitudes is around sunrise or sunset, so that the satellite rides the terminator between day and night. In that position the aerodynamic drag along the GOCE orbit is not significantly perturbed by the Sun, as is the case for a Sun-synchronous orbit placed at the noon-midnight position. This could help in calibrating thermospheric air density models at $220 - 250$ km altitude, which could then be used as reference for higher altitudes, where density is considerably lower. Such a strategy is also used in the e.g., MSIS-type

87

models, where temperature profiles allow for the exact integration of the hydrostatic equation for a constant mass to determine the density profile based on a density specified at 120 km as a function of geographic and solar/magnetic parameters.

On the other hand, neutral horizontal wind models show that thermospheric horizontal winds mainly occur around the geomagnetic poles, where they are driven by the perturbations in the geomagnetic field. The highest thermospheric wind velocities may be expected in the dawn-dusk regions, and from that point of view, the GOCE orbit is a perfect candidate for providing, for the first time, information on neutral horizontal winds in the lower thermosphere.

For more on the dedicated ESA study triggered by this section that demonstrated for the first time the use of GOCE data on thermospheric winds, see (Doornbos et al. 2012) and (Peterseim et al. 2011).

10.4 A Novel Approach to Modeling Thermospheric Air Density Using Ionosphere TEC Maps

Can we make use of the global TEC maps, regularly provided by the IGS, to improve the thermospheric density models used in the orbit determination of LEO satellites? Can we use IGS TEC maps to predict solar activity and from that the duration of a LEO mission? The current solar cycle (Solar Cycle 24) is extremely mild, and thus the GOCE mission in very low LEO orbit has now two additional mission phases. This clearly indicates that the atmospheric density in the thermosphere is lower than predicted. A similar effect can be seen in the TEC maps provided by the IGS, i.e., due to a lower level of solar activity, there are fewer free electrons in the ionosphere, as measured by GNSS receivers in the global IGS network.

Figure 10.5 shows the daily Sun spot number over the last 150 years using data from the National Geophysical Data Center (NOAA). Both this set of data and the solar flux index F10.7, clearly show that Solar Cycle 24 is the mildest for the last 150 years and up to 50% milder than the other solar cycles.

The same can be seen in Figure 10.6, showing global mean TEC values calculated using the IGS TEC maps (CODE AC) over the last two solar cycles. Since CODE IGS AC uses a spherical harmonic expansion to generate the global TEC maps, we plotted the central term C_{00} of the spherical harmonic expansion that shows that Solar Cycle 24 (with the maximum in 2012-2013) is the mildest for the last 150 years and up to 50% milder than other cycles.
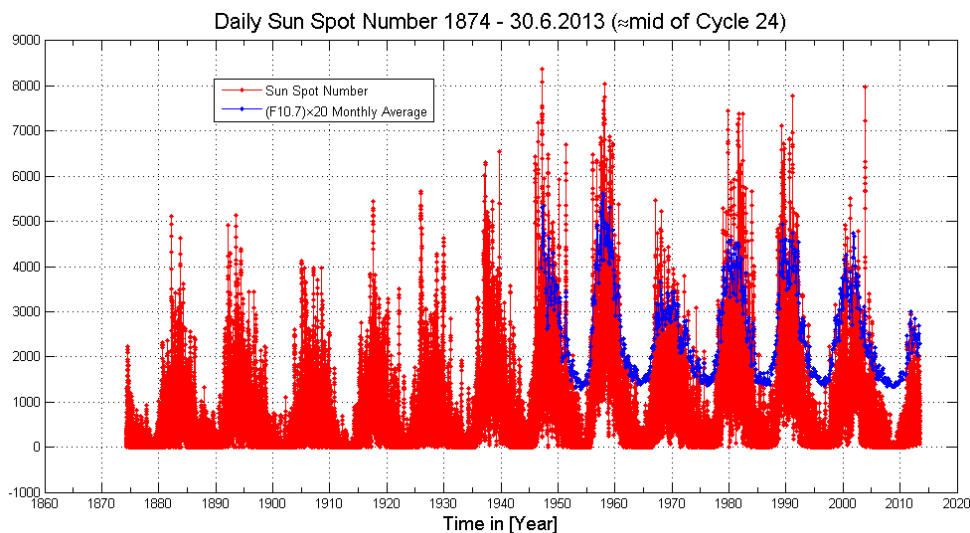
Figure 10.5 Daily Sun Spot Number (in red) from the year 1874 to 30.6.2013 ($\approx$mid Cycle 24) against the monthly mean of the Solar flux index F10.7 (in blue) scaled by a factor of 20. Both sets of data represent the mean over the entire sphere placed at the Chapman height of 450 km.
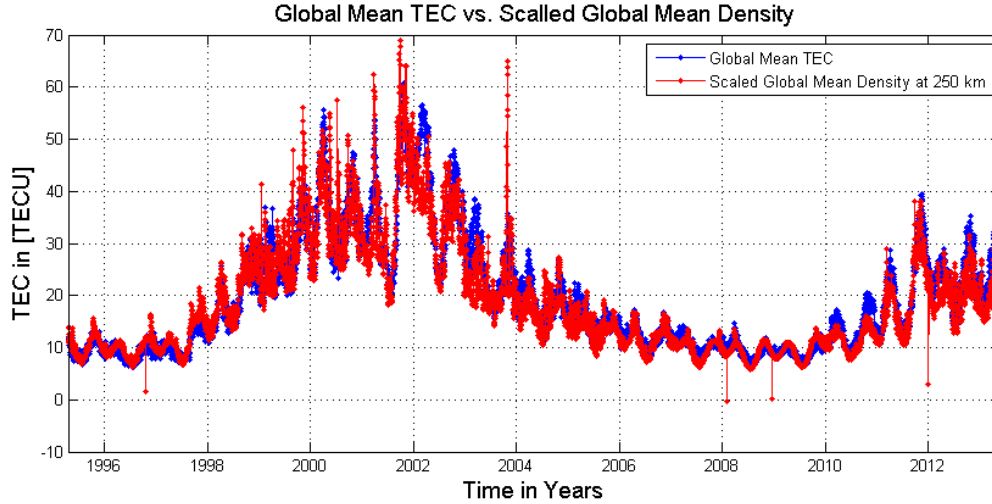
Figure 10.6 Daily global mean TEC based on IGS TEC maps (CODE AC) against daily global mean density in $[\text{kg/m}^3]$ at 250 km altitude calculated using the NRLMSISE-00 model (for a $5° \times 5°$ grid every 6 hours) based on the linear model (10.5) over the last 20 years. The agreement between the two different physical quantities is excellent and the relation can be modeled by a simple linear model with an accuracy of a few TECU over two solar cycles. One can clearly see the annual and the 27-day Sun rotation period in both time series as well as the maxima of Solar Cycle 23 and Solar Cycle 24 around the years 2002 and 2013 respectively.

The question one can now ask is, ″Can we see the same effect in atmospheric density?″. To answer this we calculated daily global mean density using the NRLMSISE00 model for a global grid $5° \times 5°$ at 250 km altitude every 6 hours. Figure 10.6 shows the daily global mean density in $[\text{kg/m}^3]$ scaled by a constant factor over the last 20 years. The agreement between the two different physical quantities is astonishing. A linear model was fitted by least-squares for a period of 20 years, covering the last two solar cycles. As one can see from Figure 10.6, the relation between TEC and thermospheric density can be modeled by a simple linear model with an accuracy of a few TECU over two solar cycles. One can clearly see the annual and the 27-day Sun rotation period in both time series, as well as the maxima of Solar Cycle 23 and Solar Cycle 24 around the years 2002 and 2013, respectively.

Over shorter time scales, e.g., half a solar cycle as shown in Figure 10.7, we see that the agreement between the global mean TEC and the mean thermospheric density is even better, at a level of $1 - 2$ TECU over the last 7 years. The NRLMSISE-00 model was used with the solar index $F_{10.7}$ and the geomagnetic Ap index from the National Geophysical Data Center (NOAA). The calculation using the NRLMSISE-00 model is very sensitive to the solar index $F_{10.7}$, whereas the 3-hourly Ap indices provide only short-term sub-daily data. Looking at those time series, given for the last 20 years, one could also ask the question, ″How stable are the differential code biases (DCBs) over those 20 years?″. DCBs define the absolute datum for IGS TEC maps, and the estimation of global ionosphere maps is used as the reference to determine them.

To calculate the mean daily TEC based on the mean thermospheric density at 250 km altitude for the period of the two solar cycles displayed in Figure 10.6, we used the following linear model

$$\overline{\text{TEC}} = a \times \bar{\rho} + b \tag{10.4}$$

with coefficients a and b, and the mean density $\bar{\rho}$. After least-squares adjustment (fit to IGS TEC maps) we obtained

$$\overline{\text{TEC}} = 5.0 \cdot 10^{11} \times \bar{\rho}_{250\,\text{km}} - 7.4 \tag{10.5}$$
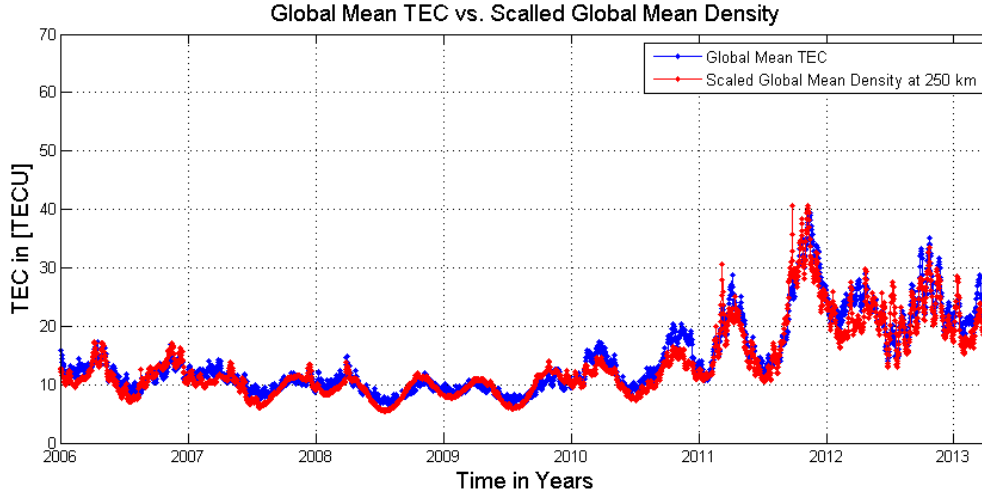
Figure 10.7 Daily global mean TEC based on IGS TEC maps against daily global mean density in $[\text{kg/m}^3]$ at 250 km altitude (NRLMSISE-00 model, $5° \times 5°$ grid every 6 hours) based on the linear model (10.5) over the last 20 years. Agreement between the two different physical quantities is to a level of $1-2$ TECU over the last 7 years. One can clearly see the annual and the 27-day Sun rotation period in both time series.

where $\overline{\text{TEC}}$ stands for daily global mean of the TEC in [TECU], $\overline{\rho}_{250\,\text{km}}$ is the mean air density given in $[\text{kg/m}^3]$ calculated using the NRLMSISE-00 model for a $5° \times 5°$ grid every 6 hours at 250 km altitude. Figure 10.6 was calculated using $\text{Ap} = 4$, since the use of the 3-hourly Ap indices increases only the high-frequency part.

From (10.5) it follows that ionization in the ionosphere is directly proportional to air density, i.e., a greater density of the thermosphere due to a higher level of solar activity (heating) is accompanied by proportionally more free electrons in the ionosphere. The linear model of fractional thermospheric density at 250 km altitude is similar to the fractional TEC at LEO altitude, both fractional quantities can be modeled using a simple linear model (10.5). In Section 9.3, we showed with GPS measurements from the CHAMP satellite that integration of the Chapman function, i.e., fractional TEC above LEO orbit altitude, can be calculated using a bias applied to ground TEC values.

Making use of the liner model (10.5), we can combine ground TEC or fractional LEO TEC measurements with thermospheric density at a given altitude. This could be used to indirectly predict solar activity in order to calculate LEO mission duration (as was done for the GOCE mission) and opens up new applications of the global IGS TEC maps in monitoring air density in the thermosphere.

A similar linear model for thermospheric density was derived for an altitude of 500 km

$$\overline{\text{TEC}} = 252.0 \cdot 10^{11} \times \overline{\rho}_{500\,\text{km}} + 8.6 \tag{10.6}$$

(see also Figure 10.8). Although, compared to (10.5), the orbit altitude was increased by a factor of 2, the scaling factor in (10.6) increased by a factor of 50. Figure 10.9 (left) shows a geographical map of air density at 400 km altitude scaled to the TEC values by a linear model (scale and offset) at 12 UT, while the figure on the right shows TEC values as provided by IGS (CODE IGS AC). One can see that the overall agreement is very good and in both cases the maximum value occurs at about 14 h local time, two hours after the Sun has passed the meridian of that geographical location.
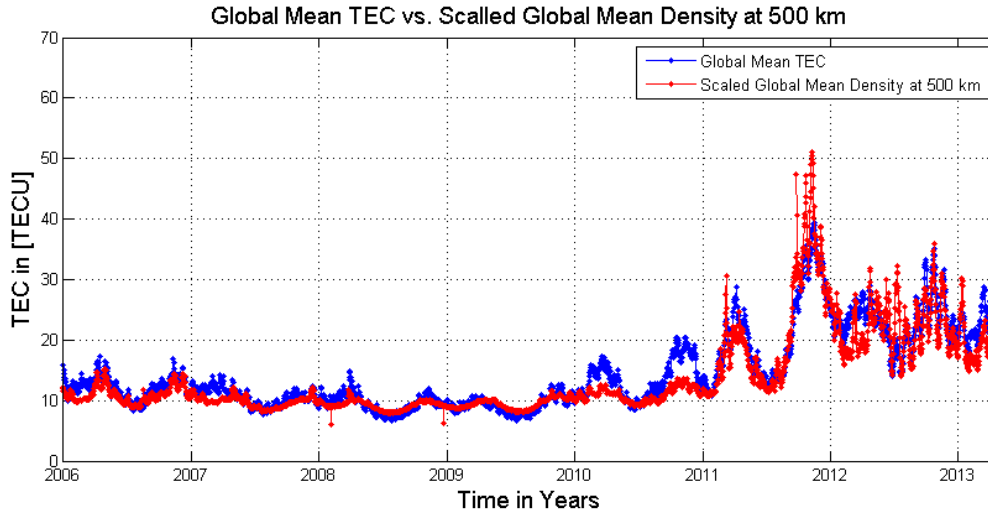
Figure 10.8 Daily global mean TEC based on IGS TEC maps against daily global mean density in $[\text{kg/m}^3]$ at 500 km altitude calculated using NRLMSISE-00 model (for a $5° \times 5°$ grid every 6 hours) scaled by a constant factor over the last 20 years. Agreement between the two different quantities is to a level of $1 - 2$ TECU over the last 7 years. One can clearly see the annual and the 27-day Sun rotation period in both time series.

What is the mechanism that relates density of the thermosphere to ionization in the ionosphere? When the Sun is more active it emits more high-energy radiation, i.e., X-ray and extreme UV radiation (XUV) that is almost completely absorbed in the thermosphere. This radiation creates ionospheric layers and increases the temperature at those altitudes. Due to this high-energy radiation, the thermosphere becomes hotter and so expands. Expansion of the thermosphere moves lower levels of the thermosphere with higher density to higher altitudes. This, in turn, increases the aerodynamic drag on satellites at those altitudes. In the auroral regions additional heating of the thermosphere can be caused by the solar wind interacting with the magnetosphere. At the same time, this high-energy radiation from the Sun in the form of high-energy photons tears electrons away from gas molecules creating ions at the same thermospheric altitudes (ionosphere). This is described by the Chapman function (9.17) that gives the ion production rate as a function of height for the entire ionosphere. Thus we have two mechanisms that work in parallel at similar altitudes, i.e., ionization of the ionosphere and heating of the electrically neutral thermosphere.
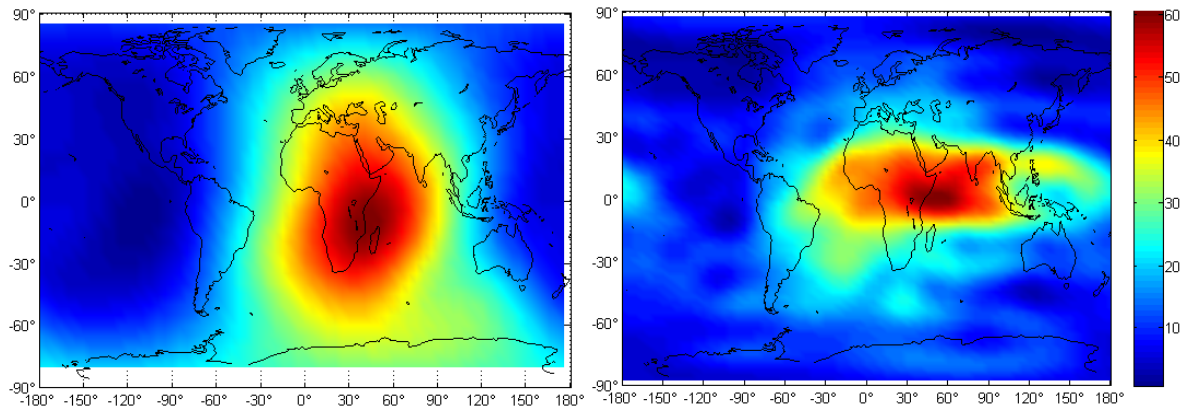
Figure 10.9 Air density at 400 km altitude scaled to TEC map (left) vs. global TEC (right) in [TECU] during the current solar maximum (day 55/2013, 12 UT). Air density was calculated using the NRLMSISE-00 model (for a $5° \times 5°$ grid) and scaled to the TEC values by a linear model (scale and offset). One can see that in both cases maximum values occurs at about 14 h local time.

91

Analogous to the Chapman function (9.17) that defines the vertical profile of the ionosphere, thermospheric temperature is given by the so-called Bates profile, (Bates 1959)

$$T = \ T_\infty - (T_\infty - T_0) \ e^{-s(z\text{-}z_0)} \tag{10.7}$$

with reference temperature $T_0 = 355 \ \text{K}$ given at $z_0 = 120 \ \text{km}$. The exospheric temperature T_∞ is directly related to solar activity as a function of the solar index $F_{10.7}$ by the following empirical formula

$$T_\infty = 500 + 3.4 F_0 \tag{10.8}$$

with the Covington index F_0 having a typical range of $70 - 250$ over one complete solar cycle. The shape of the Bates profile is given by the empirical parameter s that typically decreases with T_∞. Once the temperature profile of the thermosphere (10.7) is given, one can calculate the corresponding pressure profile and from that the thermospheric density profile. Taking into account the ideal gas law and integration of the hydrostatic equation, the simplest form of the density profile as a function of temperature T and altitude h reads as

$$\rho = \rho_0 e^{-h/H_0} \tag{10.9}$$

with

$$H_0 = \frac{\text{R}}{\mu g} T \tag{10.10}$$

with R denoting the universal gas constant, g is the gravity at altitude h, μ represents the molecular weight of the atmospheric constituents and ρ_0 in (10.9) is the atmospheric density at the reference height.

10.5 The Remove-Restore Approach to Modeling the Density of the Thermosphere

The previous subsection clearly shows that there is a high correlation between the density of the thermosphere and the total electron content in the ionosphere and that there is a similar physical mechanism governing both. The next step would be to improve the temporal and spatial resolution and accuracy of the thermospheric models. One possible approach is to look at the existing information on the geomagnetic indices and solar radio flux measurements that drive input parameters for the thermospheric models. Is there an alternative?

Here we propose studying the empirical coupling between thermospheric density and ionospheric total electron content. We intend to use data provided with a high degree of spatial and temporal resolution by the IGS. The idea is to study temporal and spatial correlations between global TEC maps and air density. It is known that both effects are highly correlated with, and driven by the solar radio flux index $F_{10.7}$. Monitoring of the ionosphere is performed by IGS providing global maps of the vertical TEC every 2 hours. Our proposed approach could be based on the temporal and spatial correlations between variations in the TEC at the GOCE altitude, against the air density provided by the models and the air densities derived from the GOCE accelerometer. One way to carry out such an approach is to use a standard remove-restore technique with the thermospheric model as a background model. It can be shown that ionosphere models such as IRI2007 or NeQuick can be used as background models in the very sparse real-time IGS network to improve spatial and temporal resolution of the real-time/predicted TEC maps. In a similar way, coupling and correlation between ionospheric charging and thermospheric heating could be studied, where the TEC information is used as a precursor for density variation. The GOCE in-situ density measurements could be used as a reference in this modeling. The quality of orbit prediction would be the first criterion in assessing the performance of such empirically derived density models. External validation can be carried out by independent comparison with density estimates from other missions, e.g., the TIMED mission (NASA) with an orbit inclination of 74°.

10.6 Sustainable Mapping of the Earth's Gravity Field at Very Low LEO Altitudes of 195–205 km and Below

It is expected that GOCE follow-on missions will be single satellite missions to monitor both, the static and temporal gravity field of the Earth, based on atomic interferometry. These missions will need to fly at very low orbit altitude in order to map the static and temporal gravity field of the Earth with very high degree and order in terms of spherical harmonic expansion, most likely in the range of $195 - 205$ km orbit altitude with drag levels at 15 mN and above. At the GOCE orbit altitude of 192 km in a Sun-synchronous orbit, the reported measured atmospheric drag level of GOCE accelerometers was 24 mN on average, with peaks up to 35 mN and an average natural orbit decay of 4 km/day. At the GOCE orbit altitude of about 224 km (last mission phase) the drag level is nearly halved in size and is at the level of 8 mN compared to the orbit altitude of 205 km. Since the GOCE mission demonstrated for the first time that ion propulsion is a viable technique for maintaining a satellite at low LEO orbit for a period of nearly 5 years, new generations of gravity missions will push the borders of ion propulsion even further, in terms both of duration and of lower orbit altitude. However, maintenance of an extremely low LEO orbit is always limited by the onboard fuel capacity and depends on the air density at those altitudes, i.e., solar activity. It is expected that future propulsion systems will need to be able to maintain the orbit altitude for about 10 years (depending on solar activity), at orbit altitudes $195 - 205$ km with a natural orbit decay of $2 - 3$ km/day at those altitudes. This appears even more attractive if the very low level of solar activity in the current solar cycles continues into the future solar cycles (as highly expected). At the GOCE orbit altitude of 190 km, orbit decay was already 4 km/day and doubled at the orbit altitude of 170 km to 8 km/day. At the 160 km orbit altitude, GOCE orbit decay was 13 km/day with an average air-drag of around 90 mN. At an orbit altitude of 147 km, less than 18 hours before re-entry, GOCE was dropping at a rate of more than 1 km/hour with an average drag level of about 165 mN. Interestingly, the temperatures of payloads and GOCE subsystems close to the front of the satellite increased by only about 13°C from those of the altitude of about 160 km the day before, as reported by the GOCE mission operations team in ESA/ESOC. For the sake of completeness, at the orbit altitude of 122 km the orbit decay was about 2.7 km/hour.

Going to lower orbit altitudes, an additional lift force could be gained by the increased density levels and optimizing the angle of attack. In the case of drag, the surface force is parallel to the air flow direction, whereas the lift force is the component of the total aerodynamic force perpendicular to the oncoming flow direction. When the angle of attack α (typically in the order of several degrees) is optimized for the platform area A one can obtain a lift coefficient $C_L(\alpha)$ that will give maximum lift acceleration $a_l(\alpha)$ for a given angle of attack

$$a_l(\alpha) = \frac{1}{2}\rho v^2 \frac{C_L(\alpha) \cdot A}{m} \tag{10.11}$$

From (10.11) we see that lift acceleration is proportional to air density ρ and to the square of the relative velocity v. Similar to the cross-section ratio for air drag, the ratio A/m could be called platform ratio. Thus, for future gravity field missions flying at 200 km orbit altitude and below it is expected that significant lift could be generated by increasing the platform ratio and optimizing the angle of attack to gain the maximum lift coefficient $C_L(\alpha)$. Air planes typically maintain an optimized angle of attack by using the onboard computer to ensure that air flow generates maximum lift at all times. A similar optimization could be performed in astronautics for satellites in low LEO orbit. At the Karman line, the LEO orbit cannot be sustained any longer and the lift force is equal to gravitation g

$$a_l(\alpha) = \frac{1}{2}\rho v^2 \frac{C_L(\alpha) \cdot A}{m} = g(h) \qquad h \approx 100 \text{ km}, \quad \text{Karman line} \qquad (10.12)$$

GOCE was the first satellite that re-entered the Earth's atmosphere with a drag-free system active prior to orbit re-entry, and was the first uncontrolled ESA re-entry in 25 years. Although the onboard fuel was spent, the net effect was that the re-entry of the GOCE satellite took place at very low angle of attack w.r.t. the Earth's atmosphere, i.e., the so-called Karman line at ≈ 100 km altitude. This is due to the drag-free mode that was active at very low orbit altitudes, much below the nominal orbit altitude when the mission was planned some 15 years ago. For ATV and the Shuttle missions, an orbit maneuver is usually needed to achieve the correct angle of attack for safe orbit re-entry or in order to burn up the satellite in the atmosphere (ATV). However, for a drag-free satellite with an uncontrolled re-entry, the angle of attack is close to zero with lower relative velocity, thus the re-entry will take longer and there is a high probability that many parts of the satellite will survive thermal effects. For the GOCE mission, it was estimated that the proof mass could survive the satellite re-entry. However, after maintaining the GOCE satellite at a significantly lower orbit altitude than that planned some 15 years ago, it is expected that more parts of the satellite survived re-entry and impacted on landing. Typically for all satellite missions, parts with high melting temperatures, such as fuel tanks made of stainless steel or titanium could survive orbit re-entry. As showed in (Hansen 1987), the heat load experienced by a satellite re-entering the atmosphere is inversely proportional to the air-drag coefficient, i.e., the greater the air-drag, the lower the heat load. Higher air-drag or cross-section area acts in a similar way to an air-bag by keeping hot gases away from direct contact with the satellite, the heat energy moves around the satellite and dissipates in the atmosphere. Thus, with a low angle of re-entry the air drag will be maximal and with relatively lower velocity (entering slowly), there is a high probability that such a satellite could re-enter the Earth's atmosphere and impact on landing with many parts. The GOCE gradiometer itself is protected by a carbon-carbon structure that has a very high melting point. This poses the question of whether, with some additional thermal protection on the port side of the satellite (GOCE shadow side) and flying a high-drag altitude profile, one could land the main part of the satellite on the ground. Some early predictions from 15 years ago claimed that 25% of the GOCE satellite (250 kg) will survive re-entry. Thanks to its aerodynamic shape, it is expected that the GOCE spacecraft could maintain the nominal attitude by the atmospheric drag forces alone, flying like a "needle" in the Earth's atmosphere.

New generation mini shuttle missions and other re-entry space vehicles or sub-orbital flights capabilities show that it will be possible to land the payload after the mission is over and to re-launch the same system. This could be a sustainable option for the core satellite missions that require decades of continuous Earth monitoring with a significant number of satellites at extremely low LEO orbits equipped with propulsion systems. The ion propulsion could be supported by the new generation of combustion engines such as the rotational detonation engine that could both maintain a satellite orbit at very low altitude for a very long time.

11. GPS Single-Frequency: From First cm-POD to Single Frequency GNSS-RO/R

I n this section we introduce what we call "Positive Code-Phase" linear combination or the LP linear combination (phase and code added) to eliminate the first-order ionosphere effect and estimate LEO orbits using single-frequency GPS measurements, (see (Švehla and Rothacher 2003a), (Švehla and Rothacher 2005b)). We do not smooth code measurements with the linear model as proposed by the GRAPHIC (Group and Phase Ionospheric Calibration) linear combination in (Yunck 1993; Gold et al. 1994; Muellerschoen et al. 2004). We show that in the case of the GRACE-B satellite it is possible to estimate LEO orbits to an accuracy of $2-3$ cm RMS (1.3 cm radial) using single-frequency GPS measurements only, (see also (Svehla et al. 2010a)). This is similar to the orbit accuracy of $1-2$ cm one can typically achieve with dual-frequency carrier-phase measurements. This is possible due to the very low noise level of the code measurements from the GRACE-B satellite and recent gravity field models from the GRACE and GOCE missions that provide very accurate gravity field coefficients up to degree and order 120 allowing an orbit parameterization with a very modest number of empirical parameters. In addition, thanks to the excellent precision of the real-time GPS satellite clock parameters provided by the IGS, we show that this cm-orbit accuracy can be achieved even in real-time. Subsequently, we introduce an estimation of the group delay pattern of GNSS satellite antennae based on the LP linear combination. We show that the LP linear combination can be used to estimate single-code group delay variations (GDV) for GNSS satellite antennae at the single-frequency level and present the first GDV pattern based on GPS measurements from the GRACE-B satellite. The GDV pattern based on LP linear combination is related to a single code observable and not to an ionosphere-free linear combination, a strong advantage in the presence of multi-GNSS data. After that, we present the concept of using single-frequency GPS radio-occultations (RO) as a very promising alternative to standard GPS-RO based on dual-frequency measurements. The advantage of this approach is that carrier and code measurements on the same GPS frequency follow the same path in the ionosphere. This is not the case for the bended carrier-phase GPS-RO measurements on different GPS frequencies that can reach a vertical separation of up to 500 m in some cases. Since the antenna used for GPS-RO is typically a high-gain antenna, the noise level of the code measurements is very low and, with an additional smoothing, this approach could be used for GPS-RO with SBAS satellites in GEO. The same approach could also be applied to GNSS reflectometry (GNSS-R).

11.1 Positive Code-Phase Linear Combination

Following (Švehla and Rothacher 2003b), a simplified version of the observation equation for the phase $L_{LEO,i}^s$ and code $P_{LEO,i}^s$ observations (GPS frequency i, distance between LEO satellite and GPS satellite s) is given as

$$L^s_{LEO,i} = \rho^s_{LEO} + \lambda_i N^s_{LEO,i} + I^s_{LEO,i} + c\delta t_{LEO} - c\delta t^s + \varepsilon(L_i)$$
$$P^s_{LEO,i} = \rho^s_{LEO} \qquad\qquad - I^s_{LEO,i} + c\delta t_{LEO} - c\delta t^s + \varepsilon(P_i)$$

(11.1)

where ρ^s_{LEO} denotes the geometry term of the distance between the LEO and the GPS satellite s, $N^s_{LEO,i}$ is the zero-difference phase ambiguity with wavelength λ_i, $I^s_{LEO,i}$ is the first order ionospheric correction, δt_{LEO} and δt^s are the LEO and GPS satellite clock values and $\varepsilon(L_i)$ and $\varepsilon(P_i)$ denote carrier-phase and code noise, respectively. The LP linear combination ("Positive Code-Phase") of phase and code measurements is then defined as (Švehla and Rothacher 2003a), (Švehla and Rothacher 2005b)

$$LP^s_{LEO,i} := \frac{1}{2}\Big(P^s_{LEO,i} + L^s_{LEO,i}\Big)$$

(11.2)

Since the first-order ionosphere effect has opposite signs for phase and code observables, it can be eliminated by adding code and carrier-phase measurements together and the ionosphere-free linear combination is then

$$LP^s_{LEO,i} = \rho^s_{LEO} + \frac{1}{2}\lambda_i N^s_{LEO,i} + c\delta t_{LEO} - c\delta t^s + \varepsilon(LP_i)$$

(11.3)

Any bias in the GPS satellite clocks or bias in the code measurements is absorbed by the estimated carrier-phase ambiguities. The wavelength of the LP linear combination is half that of the original wavelength λ_i and the noise $\varepsilon(LP_i)$ is half that of the original code.

Figure 11.1 shows the first reduced-dynamic orbit of the CHAMP satellite based on the LP linear combination of the L_1 and P_1 measurements, day 200/2002. The accuracy level is about 10 cm, when compared against the best reduced-dynamic orbit estimated using dual-frequency carrier-phase measurements. However,
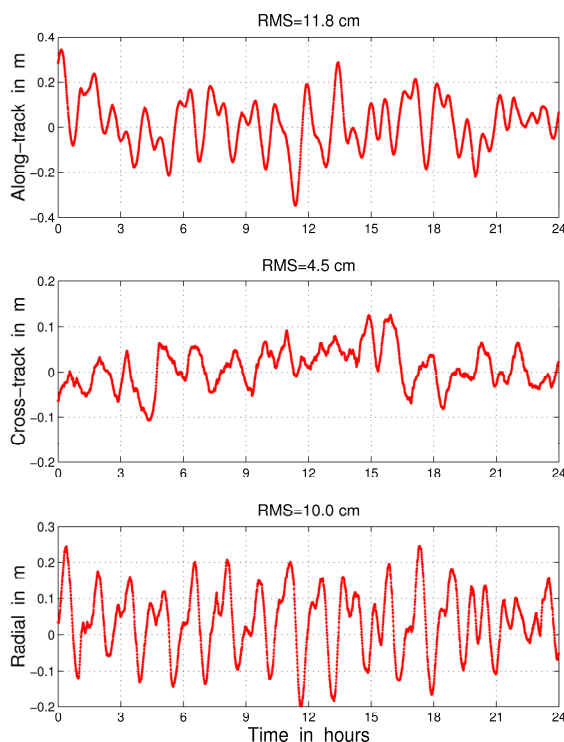
Figure 11.1 The first CHAMP reduced-dynamic orbit estimated using the LP linear combination of the L_1 and P_1 measurements, day 200/2002 (Švehla and Rothacher 2005b) based on the EIGEN-1 gravity field model and IGS orbit/clock quality in 2002. In comparison, the GRACE-B orbit can be estimated with an accuracy of $2-3$ cm RMS using single-frequency fata and gravity models from the GRACE mission, (see Figure 11.2).

it should be noted that this accuracy level is mainly driven by the GPS satellite orbit and clock quality available from IGS in 2002 and the very first CHAMP gravity models, such as EIGEN-1. In order to compensate for orbit modeling deficiencies, empirical parameters need to be estimated. The orbit results in Figure 11.1 are based on the frequent estimation of so-called pseudo-stochastic parameters (empirical velocity pulses) that, in this particular case, were estimated every 6 minutes. One can expect that this is correlated with the carrier-phase ambiguities that are estimated per tracking pass ($15-20$ min), and due to the noise level of the LP linear combination the resulting orbit is not better than about 10 cm RMS. We will see in the next subsection that, when the duration of the empirical parameterization is increased to about $1-2$ hours, and when making use of the GRACE gravity field models and the IGS orbit/high-rate clock parameters, the orbit quality improves to about $2-3$ cm RMS.

Apparently, in comparison to the LP linear combination in (11.2), a similar linear combination was introduced by (Yunck 1993), (Gold et al. 1994) and (Bertiger and Wu 1996) for C/A code measurements, where it was called GRAPHIC (Group and Phase Ionospheric Calibration) linear combination. Although developed independently, the GRAPHIC linear combination was re-discovered in 2002 and used for the first CHAMP data using more accurate P code measurements, as presented at the CHAMP Workshop in Potsdam in 2003 (Švehla and Rothacher 2005b). However, the GRAPHIC linear combination is based on the smoothed code measurements, see (Muellerschoen et al. 2004), where a linear or quadratic smoothing operator $\langle \cdot \rangle$ is employed on the difference between the code and the carrier phase measurements

$$LP_{LEO,i}^{s} := L_{LEO,i}^{s} + \frac{1}{2}\left\langle P_{LEO,i}^{s} - L_{LEO,i}^{s} \right\rangle \tag{11.4}$$

For more on GRAPHIC see (Muellerschoen et al. 2004), where a linear fit was used to smooth code measurement in (11.4).

11.2 The 1-cm Single-Frequency Orbit in a Radial Direction Based on Real-Time GPS Satellite Clocks

The LP linear combination not only reduces the noise level of code measurements by about 50%, in addition, the noise level is also averaged over the tracking pass (typically $15-20$ min) and over all tracked GPS satellites every epoch by estimating one phase ambiguity per tracking pass and receiver clock parameters every epoch. Since the precision of the CHAMP ionosphere-free observables based on C/A and P code measurements is about 48 cm (from the kinematic POD), we expect the precision of the code measurements to be about $15-16$ cm . This leads to a noise level of the LP_1 observable of about $5-8$ cm for CHAMP, whereas for GRACE-B the noise level is halved

$$\varepsilon(LP_i) \approx \frac{1}{2}\varepsilon(P_i) \approx 5-8 \text{ cm} \quad \text{CHAMP}$$
$$\varepsilon(LP_i) \approx \frac{1}{2}\varepsilon(P_i) \approx 2-4 \text{ cm} \quad \text{GRACE-B} \tag{11.5}$$

Galileo and future GNSS will introduce wide-band signals that will enable a low code noise in the cm-range to be achieved. The Galileo E5 wide-band signal (nominal bandwidth of 51.15 MHz) and AltBOC modulation will offer a code noise at the cm-level. However, this is not the case for its subcarriers E5a and E5b.

Figure 11.2. shows SLR residuals of the GRACE-B reduced-dynamic orbit estimated using the LP linear combination, while Figure 11.3 shows daily RMS errors in the along-track, cross-track and radial directions against the JPL orbit estimated by means of dual-frequency carrier-phase. One can see that the radial orbit

component can be determined down to $14\,\text{mm}$ RMS using the *LP* linear combination. Typical RMS of the single orbit component is $26\,\text{mm}$ and is similar to the $25.5\,\text{mm}$ RMS of the SLR residuals, (see Figure 11.2).

It should be noted that the GRACE orbits are based on the GRACE gravity field models (Tapley et al. 2005) and the IGS orbit/high-rate clock parameters. The GRACE gravity field models allow the orbit to be modelled dynamically with a relatively modest number of empirical parameters, e.g., velocity pulses every $1-2$ hours. Thus there is a weaker correlation with the frequent carrier-phase ambiguities that are estimated per tracking pass (typically $15-20\,\text{min}$ in duration). At this point, it is interesting to note the noise in the estimated carrier-phase ambiguities (see Figure 11.4). The noise level of the *LP* residuals is in the order of $2\,\text{cm}$ compared to the wavelength that is of the order of $10\,\text{cm}$.
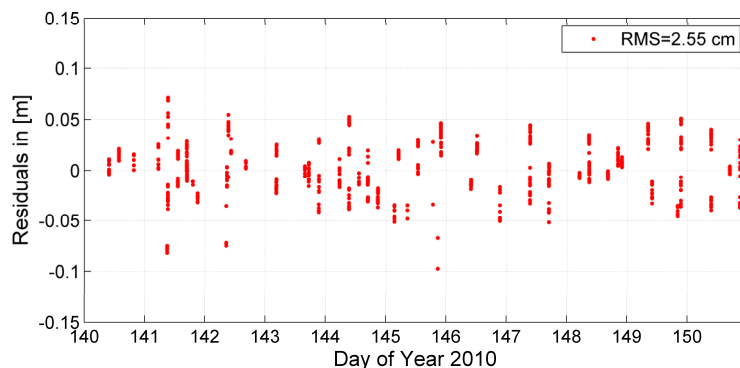
Figure 11.2 SLR residuals of the GRACE-B orbit based on the ″Positive Code-Phase″ or the *LP* linear combination using Final IGS orbit and clock products and the GRACE gravity field models (days 140-150/2010).
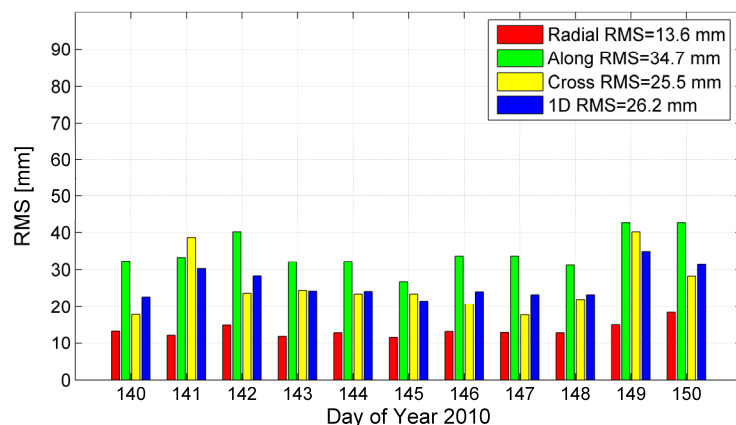
Figure 11.3 GRACE-B orbit based on the IGS Final GPS orbit and clock products against the GRACE-B orbit provided by JPL (GRACE Level 2 Product).

Observed Noise in the LP Ambiguities
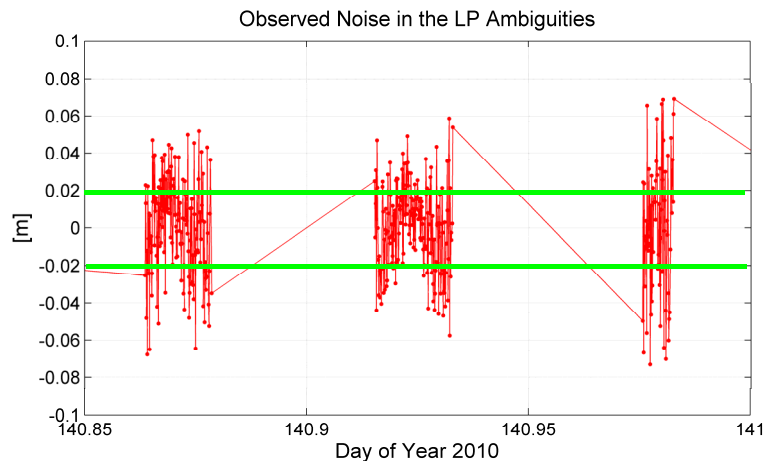
Figure 11.4 Observed noise in the estimated carrier-phase ambiguities using the (Positive Code-Phase) linear combination of the L_1 and C/A code measurements.
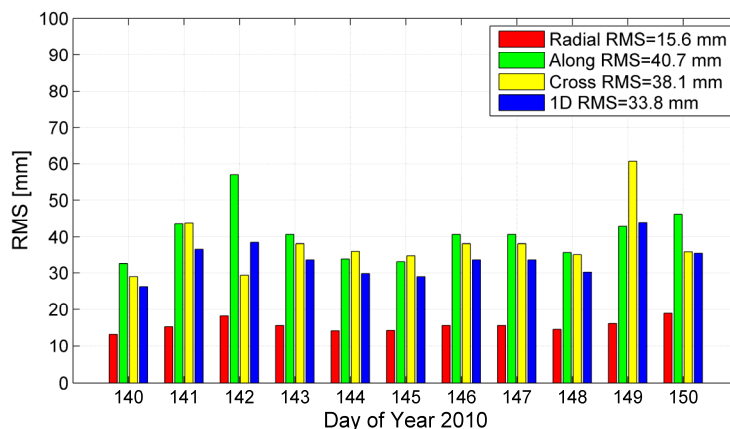
Figure 11.5 Single-frequency GRACE-B orbit based on real-time GPS orbit and clock products, versus the GRACE-B orbit from JPL.

This clearly opens doors to fix track-to-track carrier-phase ambiguities on L_1 (see Section 21). It should be noted that biases in the code measurements, which are common to all tracking passes, are eliminated by forming track-to-track ambiguities.

Figure 11.5 shows the GRACE-B orbit based on real-time IGS clock products (latency 10 s). One can see that the orbit quality is only slightly reduced when real-time GPS satellite clocks are used, i.e., from 26 mm to 33.8 mm as a typical RMS for all three orbit components. Again, the radial orbit component is the most accurate (15.6 mm RMS).

11.3 Estimation of GPS Satellite Group Delay Patterns Using the LP Linear Combination

Figure 11.6 shows the LP residuals from the reduced-dynamic orbit determination of the GRACE-B satellite as a function of GPS satellite nadir angle. One can see that residuals, when plotted in the GPS satellite frame, are strongly nadir dependent, as is to be expected, when elevation dependency of the residuals is observed for the receiving GPS antenna on the ground. This test confirms that the group delay patterns for C/A code are

99
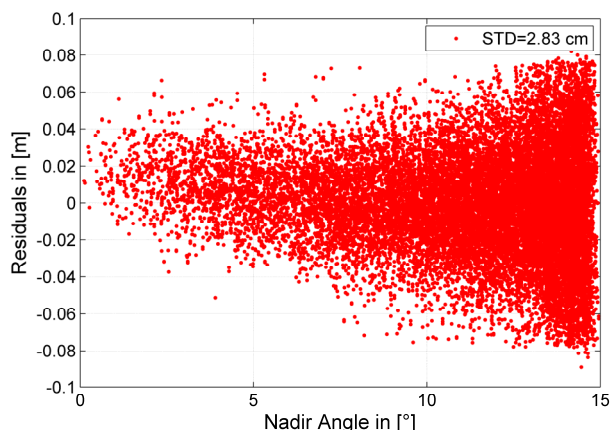
Figure 11.6 GRACE-B residuals (SVN49) from the reduced-dynamic orbit estimated using the LP linear combination of L_1 and C/A code as a function of GPS satellite nadir angle (day 150/2010).
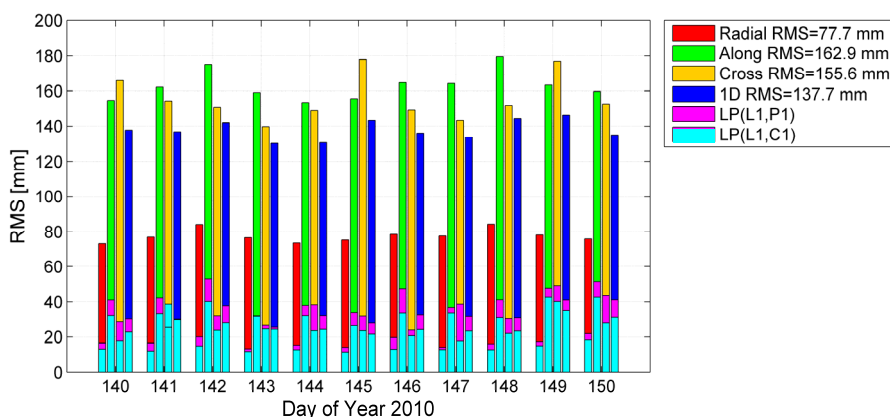
Figure 11.7 Single-frequency GRACE-B orbit based on P_2 and L_2 measurements compared with P_1 and C/A code measurements. One can see a significant degradation of the P_2 based orbit compared to other code measurements, most likely due to the P_2 group-delay patterns of GPS satellites.

flat (within about $6-8$ cm peak-to-peak) and that both the choke-ring antenna on board the GRACE-B and the GPS satellite transmitter (SVN49) have similar characteristics in terms of group delay variations.

Let us now see if the same POD performance can be achieved when other code observables are used, namely P_1, P_2 and L_2 carrier-phase. Figure 11.7 shows the daily RMS of the orbit estimated using all alternative code observables against the orbit based on dual-frequency carrier-phase used as a reference. One can clearly see the significant degradation of the orbit based on P_2 code that could be explained by higher variations in the group-delay variation (GDV) patterns on P_2. A significantly smaller effect can be seen in the orbit based on P_1 code. Figure 11.7 shows that the LP linear combination could be used to estimate single-code group delay patterns of GNSS satellites. The GDV patterns estimated based on the LP linear combination are related to the single code observables and not to an ionosphere-free linear combination, an advantage in the presence of multi-GNSS data.

Figure 11.8 shows the GDV pattern on P_2 for the GPS satellite GPS-08 based on code measurements from the GRACE-B satellite and the LP linear combination. One can see a strong nadir dependency as well as variations with azimuth. This GDV pattern was estimated based on the choke-ring antenna on the GRACE-

B satellite that has a very low multipath environment in LEO orbit. This is the lowest code noise < 10 cm RMS, reported for a GPS receiver.
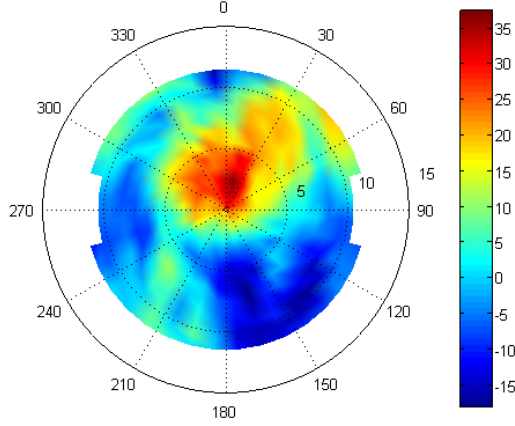
Figure 11.8 Preliminary map of the group delay pattern on P_2 in [cm] for the GPS satellite GPS-08 in the satellite-specific reference frame based on P_2 code measurements from the GRACE-B satellite and the LP linear combination, day 214/2008. Typically, Block IIR-M satellites show large group delay variations, (Svehla et al. 2010a).

11.4 "Negative Code-Phase" Linear Combination: A Geometrically Correct Ionosphere-Free Linear Combination for GNSS-Radio-Occultations

The first-order ionospheric effects can be eliminated by adding code and phase measurements together. Let us now see what happens when those two types of observables are subtracted from each other. In this case we obtain what we call the LM linear combination (phase minus) or "Negative Code-Phase" linear combination, defined as follows

$$LM_{LEO,i}^s := \frac{1}{2}\Big(P_{LEO,i}^s - L_{LEO,i}^s\Big) \tag{11.6}$$

from which we can derive the observation equation of the geometry-free linear combination

$$LM_{LEO,i}^s = -\frac{1}{2}\lambda_i N_{LEO,i}^s - I_{LEO,i}^s + \varepsilon(LM_i) \tag{11.7}$$

As with the LP linear combination, the wavelength of the LM linear combination is half that of the original wavelength, λ_i, and the noise $\varepsilon(LM_i)$ is half the code noise

$$\varepsilon(LM_i) \approx \frac{1}{2}\varepsilon(P_i) \approx 5 - 8 \text{ cm} \quad \text{CHAMP}$$
$$\varepsilon(LM_i) \approx \frac{1}{2}\varepsilon(P_i) \approx 2 - 4 \text{ cm} \quad \text{GRACE-B} \tag{11.8}$$

From (11.7) it follows that the ionospheric slant delay between the LEO and the GPS satellite can be calculated as

$$I_{LEO,i}^s = -LM_{LEO,i}^s - \frac{1}{2}\lambda_i N_{LEO,i}^s + \varepsilon(LM_i) \tag{11.9}$$

Eq. (11.9) is biased by an unknown carrier-phase ambiguity $N^s_{LEO,i}$ that could be back-substituted from the orbit determination procedure based on the LP (Positive Code-Phase) linear combination. When GPS measurements are provided at a high sampling rate an additional averaging or smoothing of (11.9) can be employed. The corrected, ionosphere-free carrier-phase measurement $\bar{L}^s_{LEO,i}$ is then

$$\bar{L}^s_{LEO,i} = L^s_{LEO,i} + \left\langle LM^s_{LEO,i} + \frac{1}{2}\lambda_i N^s_{LEO,i} + \varepsilon(LM_i) \right\rangle \tag{11.10}$$

where $\langle \cdot \rangle$ denotes the smoothing or averaging operator. With an increased sampling rate, one could average code noise and even form normal points at a sampling rate below the GPS-RO signal. In addition, the code noise could be reduced by the GPS-RO antenna with high-gain (phased-array, etc.).

For GPS radio-occultations, the first derivative of (11.10) is actually needed. Thus, as a smoothing operator one could employ a simple polynomial. The first derivative of the fitted polynomial can be used directly as an input for the inversion of GPS-RO data. If the multipath level on board the LEO satellite is low, the single-frequency approach described above could provide an alternative GPS-RO observable with very low noise. GPS-RO with 10-15 GEO satellites could provide added value to the standard GPS-RO approach with GPS satellites in MEO. Typically, satellites such as EGNOS, and WAAS provide single-frequency carrier-phase and code measurements that are collected by the ground GPS receivers, but hardly used for any application.

It is very important to mention that the L_1 carrier-phase and the C/A code follow the same path in the ionosphere, even in the case of extremely bent GPS-RO signals. This is not the case with the GPS-RO carrier-phase measurements on two GPS frequencies, since it is well known that the vertical separation between the paths of L_1 and L_2 signals in the GPS-RO profile can reach up to 500 m (Axel von Engeln, priv. com.). In the case of GPS-RO carrier-phase measurements, such significant bending leads to difficulties in forming the ionosphere-free linear combination in order to completely remove the first-order ionosphere effect. The error created when forming such a dual-frequency ionosphere-free GPS-RO observable can easily be above the noise level of the alternative single-frequency approach. Therefore, the single-frequency approach described above is an attractive alternative to the standard GPS-RO strategy, especially in the light of possible future GNSS signals, considering other applications in GNSS radio-occultation and GNSS reflectometry, providing code measurements at frequencies outside the conventional L-band and under different tracking conditions.

11.5 Pre-processing and Synchronization of Single-Frequency GPS Data

A disadvantage of the LP linear combination lies in the data pre-processing, since in the case of single-frequency GPS receivers, pre-processing has to be performed without the second GPS frequency. The pre-processing approach, as implemented in the Bernese GNSS software for undifferenced dual-frequency carrier-phase measurements, is based on the estimation of position differences and one clock parameter between subsequent epochs, (see (Švehla and Rothacher 2003b)). Considering the relatively high sampling rate of carrier-phase measurements compared to the changes in the ionospheric TEC, it can be shown that phase cycle-slips can be reliably detected by looking at the differences between successive epochs. Thus the same algorithm used to pre-process carrier-phase measurements could be used to pre-process single-frequency measurements. When dual-frequency data are processed as single-frequency, the pre-processing can be performed at the dual-frequency level using the ionosphere-free linear combination and the aforementioned algorithm used in the Bernese GNSS software.

When GPS measurements are provided at high sampling rates (e.g., 30 s), the following linear combination could be employed utilizing epoch-wise differencing between consecutive epochs t_{k+1} and t_k

$$L_{LEO,i}^{s}(t_{k}^{k+1}) = L_{LEO,i}^{s}(t_{k+1}) - L_{LEO,i}^{s}(t_{k}) = \left[\rho_{LEO}^{s}(t_{k+1}) - \rho_{LEO}^{s}(t_{k})\right]$$
$$+ \left[I_{LEO,i}^{s}(t_{k+1}) - I_{LEO,i}^{s}(t_{k})\right] + \left[c\delta t_{LEO}(t_{k+1}) - c\delta t_{LEO}(t_{k})\right] \tag{11.11}$$

$$P_{LEO,i}^{s}(t_{k}^{k+1}) = P_{LEO,i}^{s}(t_{k+1}) - P_{LEO,i}^{s}(t_{k}) = \left[\rho_{LEO}^{s}(t_{k+1}) - \rho_{LEO}^{s}(t_{k})\right]$$
$$- \left[I_{LEO,i}^{s}(t_{k+1}) - I_{LEO,i}^{s}(t_{k})\right] + \left[c\delta t_{LEO}(t_{k+1}) - c\delta t_{LEO}(t_{k})\right] \tag{11.12}$$

Large clock variations from epoch-to-epoch can further be detected and eliminated by forming differences between two GPS satellites s and r tracked at the same epoch

$$L_{LEO,i}^{s}(t_{k}^{k+1}) - L_{LEO,i}^{r}(t_{k}^{k+1}) = \left[\rho_{LEO}^{s}(t_{k}^{k+1}) - \rho_{LEO}^{r}(t_{k}^{k+1})\right] + \left[I_{LEO,i}^{s}(t_{k}^{k+1}) - I_{LEO,i}^{r}(t_{k}^{k+1})\right] \tag{11.13}$$

$$P_{LEO,i}^{s}(t_{k}^{k+1}) - P_{LEO,i}^{r}(t_{k}^{k+1}) = \left[\rho_{LEO}^{s}(t_{k}^{k+1}) - \rho_{LEO}^{r}(t_{k}^{k+1})\right] - \left[I_{LEO,i}^{s}(t_{k}^{k+1}) - I_{LEO,i}^{r}(t_{k}^{k+1})\right] \tag{11.14}$$

The advantage of this alternative pre-processing algorithm, is that the variation of the ionosphere effect from epoch to epoch is smooth and small enough to detect phase breaks between epochs. This approach could be combined with the estimation of kinematic differences between successive epochs along an a priori reduced-dynamic orbit as described above. The a priori LEO orbit needed for this algorithm is obtained by making use of the single- or dual-frequency code measurements and a relatively small number of orbit parameters (e.g., 6 Keplerian parameters and 9 empirical accelerations per day). Variations in the ionospheric delay from epoch to epoch and relative orbit errors between subsequent epochs are small enough to limit epoch-wise kinematic orbit differences. In this parameter estimation, single-frequency L_1 phase measurements are used between two consecutive epochs and four parameters are estimated, including three kinematic position differences and one clock parameter between two consecutive epochs.

In the case of dual-frequency GPS data, the ionosphere-free linear combination of code measurements is used to obtain a priori LEO positions and to approximately synchronize LEO measurements to GPS time. For single-frequency GPS data, receiver clock synchronization of raw phase and code measurements can only be performed by means of the single-frequency code measurements fully affected by ionosphere effects. The use of IGS ionosphere maps corrected for the LEO altitude could be used to further improve this procedure.

In the case of dual-frequency GPS receivers, the synchronization of the GPS receiver time to the GPS time scale is limited by the noise of code measurements and the a priori orbit errors. For a maximum orbit error of e.g., $1\,\mathrm{cm}$ and a velocity of the LEO satellite of about $7.7\,\mathrm{km/s}$, synchronization could be carried out with an accuracy below $1.3\,\mu\mathrm{s}$, if code measurements of similar quality were available

$$\frac{0.01\,\mathrm{m}}{7700\,\mathrm{m/s}} \approx 1.3\,\mu\mathrm{s} \tag{11.15}$$

This corresponds to about $400\,\mathrm{m}$ in terms of the code error. If we now consider total electron content (TEC) above the LEO satellite to be very extreme, reaching about 200 TECU (TEC Units, $1\,\mathrm{TECU} = 10^{16}$ electrons per m^2) in the vertical direction (VTEC), the maximum ionospheric error reads as

$$I_{LEO,1} = \frac{1}{\sin E} \frac{40.3}{f_1^2} \mathrm{VTEC} \tag{11.16}$$

which at an elevation of $E = 10°$ is about $330\,\mathrm{m}$. We see that such an error is below the synchronization error of $400\,\mathrm{m}$. Generally speaking, even during the solar maximum, and under very extreme ionospheric conditions, the a priori synchronization of the receiver clock can be performed with sufficient accuracy without using any a priori ionosphere model.

12. Absolute Code Biases Based on the Ambiguity-Free Linear Combination – DCBs without TEC

$\mathbf{A}$ bsolute code biases and associated DCBs determined using absolute code biases are called ″absolute″ because they do not require TEC information to estimate them and are defined against the IGS Clock Convention (″P_3 clocks″).

Differential code biases (DCBs) are typically determined by co-estimating the first-order ionosphere effect using the geometry-free linear combination of code measurements from two different GNSS frequencies. We develop ambiguity-free linear combinations based on the dual- or triple-frequency GPS carrier-phase and code measurements on only one GPS frequency. In this way, we can estimate code biases on a single GPS frequency. Since the datum of the GPS satellite clock corrections is defined by the ionosphere-free linear combination of the P-code measurements on L_1 and L_2 we can estimate these single-frequency code biases as ″absolute biases″ using the geometry-free approach. Our ambiguity-free linear combination removes single-frequency ambiguities, but it requires the estimation of one wide-lane ambiguity with a very long wavelength, a wavelength that is significantly greater than the size of the code biases. In addition, by forming single-differences between two GNSS satellites using measurements from one station, one can separate satellite-based from station-based code biases. We show the relationship between the code biases and the narrow-lane biases in the Melbourne-Wübbena linear combination and DCBs. The same approach is extended to other multi-GNSS code observables.

Absolute code biases defined for single-frequency observables can be used to combine carrier-phase and code measurements consistently in a multi-GNSS environment and to define carrier-phase ambiguities and ionospheric effects in an ″absolute sense″. Absolute code biases can provide a datum for estimated global ionosphere maps and for all calibration of multi-GNSS code measurements (e.g., group delays). We show here absolute code bias in P_1 and C_5 code GPS measurements on L_1 and L_5 carrier-phases and present calibration of ¼-ambiguities associated with L_5. We discuss absolute code biases in the light of the S-curve bias and group delay variation maps for GNSS satellites. We show how, by introducing absolute code biases, we can consistently define a datum for GNSS satellite clock parameters and ionosphere maps in a multi-frequency GNSS environment. Galileo and future GNSS will introduce wide-band signals that will lead to low code noise (in the cm-range). Specifically, the Galileo E5 wide-band signal (nominal bandwidth of 51.15 MHz) and the AltBOC modulation will offer code noise at cm-level. The same approach could be applied to Galileo using wide-band signals as reference signals to determined absolute code biases.

12.1 Definition of Absolute Code Biases in the Light of Multi-GNSS Data

In the case of the positive code-phase linear combination, any bias in the GPS satellite clock parameters or any satellite/receiver code biases are absorbed by the estimated carrier-phase ambiguities. By definition, GPS satellite clock parameters provided by the IGS are based on the standard ionosphere-free linear combination of P code measurements on both GPS carrier-frequencies (L_1 and L_2). The use of any other code observable (e.g., C/A, $L2C$ code) or linear combination (e.g., Melbourne-Wübbena), requires a consistent handling of the code biases. At the moment, only relative or so-called differential code biases (DCBs) are used by the IGS for GPS satellites, relating two code observables at a given time and fulfilling the zero-mean condition over all GPS satellites in the constellation. By forming the negative code-phase linear combination, one can measure the first-order ionosphere effect. However, even by knowing the absolute values for carrier-phase ambiguities, we will not be able to define a datum for ionosphere measurements in an absolute sense. This is because DCBs are always defined between two different code observables and the absolute single-frequency biases have not yet been considered by the IGS. Therefore, in the light of multi-frequency GNSS there is a need to introduce absolute code biases, defined separately for each code observable relative to the corresponding carrier-phase on the same frequency.

The LP linear combination ("Positive Code-Phase") of phase L_i and code P_i measurements on the carrier-frequency i is defined as (Švehla and Rothacher 2003a), (Švehla and Rothacher 2005b) (for more information see Section 11)

$$LP_i \; := \frac{1}{2}(P_i \; + L_i) \tag{12.1}$$

The observation equation for the LP linear combination including the absolute code bias AB_i is then

$$LP_i \; = \rho + \frac{1}{2}\lambda_i N_i \; + c\delta t - c\delta t^s + \frac{1}{2}AB_i + \varepsilon(LP_i) \tag{12.2}$$

with the geometry term ρ, and the satellite and receiver clock parameters $c\delta t$ and $c\delta t^s$. The wavelength of the LP linear combination is half that of the original wavelength λ_i and the noise $\varepsilon(LP_i)$ is half that of the original code noise. If we look at the difference between any two associated code and carrier-phase measurements on the carrier frequencies i, j and k we obtain the following two possibilities

$$
\begin{aligned}
LP_i \; - LP_j \; &= \frac{1}{2}(\lambda_i - \lambda_j)N_i + \frac{1}{2}\lambda_j N_{W(i,j)} + \frac{1}{2}(AB_i - AB_j) + \varepsilon(LP_i - LP_j) \\
LP_i \; - LP_k \; &= \frac{1}{2}(\lambda_i - \lambda_k)N_i + \frac{1}{2}\lambda_k N_{W(i,k)} + \frac{1}{2}(AB_i - AB_k) + \varepsilon(LP_i - LP_k)
\end{aligned}
\tag{12.3}
$$

where $N_{W(i,j)} = N_i - N_j$ and where the third term represents the relative differential code bias between two frequencies

$$DCB_{i,j} \; := AB_i \; - AB_j \tag{12.4}$$

From the reference absolute bias, e.g., on the first frequency AB_i, we can estimate any other absolute bias

$$AB_j \; = AB_i \; + DCB_{i,j} \tag{12.5}$$

12.2 Absolute Code Biases Based and the Ambiguity-Free Linear Combination

Let us first define the ambiguity-free linear combination AF with only one code observable at a given time. For this, we make use of the LP linear combination (12.2) and the ionosphere-free linear combination L_3 of two carrier-phase measurements L_1 and L_2

$$AF_1 := \kappa_1^{af} L_3 + \kappa_2^{af} LP_1 \qquad (12.6)$$

(12.6) only contains absolute code bias on the P_1 code measurement, see (12.2). The geometry-preserving condition for the multiplication factors κ_1^{af} and κ_2^{af} is then as follows

$$\kappa_1^{af} + \kappa_2^{af} := 1 \qquad (12.7)$$

For the first time, we introduce here an ambiguity-free condition that for ambiguity N_1 on L_1 carrier-phase is defined as

$$\kappa_1^{af} \lambda_N + \kappa_2^{af} \frac{\lambda_1}{2} := 0 \qquad (12.8)$$

where $\lambda_N = c / (f_1 + f_2)$ denotes the narrow-lane wavelength of narrow-lane ambiguity in the ionosphere-free linear combination L_3 with the two GPS frequencies f_1 and f_2. The $\lambda_1/2$ is the wavelength of the L_1 ambiguity in the LP_1 linear combination. The basic idea of the ambiguity-free condition (12.8) is to eliminate the N_1 ambiguity that appears both in the ionosphere-free L_3 and the LP_1 linear combination

$$\begin{aligned} L_3 &= \frac{f_1^2}{f_1^2 - f_2^2} L_1 - \frac{f_2^2}{f_1^2 - f_2^2} L_2 = \rho + \lambda_N N_1 + \frac{1}{2}(\lambda_W - \lambda_N)N_W + c\delta t - c\delta t^s \\ LP_1 &= \frac{1}{2}(L_1 + P_1) = \rho + \frac{1}{2}\lambda_1 N_1 + \frac{1}{2}AB_1 + c\delta t - c\delta t^s \end{aligned} \qquad (12.9)$$

where AB_1 is the absolute code bias on P_1. In a similar way, in order to obtain the AB_2, the absolute code bias on P_2, we need to eliminate the N_1 ambiguity that appears both in the ionosphere-free L_3 and the LP_2 linear combination

$$\begin{aligned} L_3 &= \frac{f_1^2}{f_1^2 - f_2^2} L_1 - \frac{f_2^2}{f_1^2 - f_2^2} L_2 = \rho + \lambda_N N_1 + \frac{1}{2}(\lambda_W - \lambda_N)N_W + c\delta t - c\delta t^s \\ LP_2 &= \frac{1}{2}(L_2 + P_2) = \rho + \frac{1}{2}\lambda_2 N_1 - \frac{1}{2}\lambda_2 N_W + \frac{1}{2}AB_2 + c\delta t - c\delta t^s \end{aligned} \qquad (12.10)$$

For the AB_5, the absolute code bias on code measurements on L_5 carrier-phase, denoted here as C_5, we use the following two linear combinations

$$\begin{aligned} L_3 &= \frac{f_1^2}{f_1^2 - f_2^2} L_1 - \frac{f_2^2}{f_1^2 - f_2^2} L_2 = \rho + \lambda_N N_1 + \frac{1}{2}(\lambda_W - \lambda_N)N_W + c\delta t - c\delta t^s \\ LP_5 &= \frac{1}{2}(L_5 + C_5) = \rho + \frac{1}{2}\lambda_5 N_1 - \frac{1}{2}\lambda_5 N_{W(1,5)} + \frac{1}{2}AB_5 + c\delta t - c\delta t^s \end{aligned} \qquad (12.11)$$

where $N_{W(1,5)} = N_1 - N_5$. The ambiguity-free condition (12.8) is fulfilled as long as the wide-lane ambiguities $N_W = N_1 - N_2$ are fixed in the ionosphere-free linear combination, i.e., all ambiguities are aligned to each

other so that $N_1 = N_2$, using e.g., the Melbourne-Wübbena linear combination. After solving (12.8) and (12.7) for the multiplication factors of the ambiguity-free linear combination in (12.6) we obtain

$$\kappa_1^{af} = -\frac{f_1 + f_2}{f_1 - f_2}, \qquad \kappa_2^{af} = \frac{2f_1}{f_1 - f_2} \tag{12.12}$$

or ambiguity-free linear combination

$$AF_1 = \rho - \frac{c \cdot f_2}{(f_1 - f_2)^2} N_W + \frac{f_1}{f_1 - f_2} AB_1 + c\delta t - c\delta t^s \tag{12.13}$$

where the wide-lane ambiguity is associated with a long wavelength of

$$-\frac{c \cdot f_2}{(f_1 - f_2)^2} \approx -3.04 \,\text{m} \tag{12.14}$$

We can also write the ambiguity-free linear combination (12.6) for other code measurements and frequencies. For $LP_2 = (L_2 + P_2)/2$ we obtain

$$AF_2 := \kappa_{1(2)}^{af} L_3 + \kappa_{2(2)}^{af} LP_2 \tag{12.15}$$

with

$$\kappa_{1(2)}^{af} = \frac{f_1 + f_2}{f_1 - f_2}, \qquad \kappa_{2(2)}^{af} = -\frac{2f_2}{f_1 - f_2} \tag{12.16}$$

or

$$AF_2 = \rho + \frac{c \cdot f_1}{(f_1 - f_2)^2} N_W - \frac{f_2}{f_1 - f_2} AB_2 + c\delta t - c\delta t^s \tag{12.17}$$

where the wide-lane ambiguity is associated with a long wavelength of

$$\frac{c \cdot f_1}{(f_1 - f_2)^2} \approx 3.90 \,\text{m} \tag{12.18}$$

and for $LP_5 = (L_5 + C_5)/2$ and ionosphere-free linear combination L_3

$$AF_5 := \kappa_{1(5)}^{af} L_3 + \kappa_{2(5)}^{af} LP_5 \tag{12.19}$$

with

$$\kappa_{1(5)}^{af} = -\frac{f_1 + f_2}{2f_5 - f_1 - f_2}, \qquad \kappa_{2(5)}^{af} = \frac{2f_5}{2f_5 - f_1 - f_2} \tag{12.20}$$

or

$$AF_5 = \rho - \frac{c}{2f_5 - f_1 - f_2} \cdot \frac{f_2}{f_1 - f_2} N_W - \frac{c}{2f_5 - f_1 - f_2} N_{W(1,5)} + \frac{f_5}{2f_5 - f_1 - f_2} AB_5 + c\delta t - c\delta t^s \tag{12.21}$$

According to the IGS convention, GNSS satellite clock parameters are defined by the ionosphere-free linear combination L_3 of the two carrier-phase measurements L_1 and L_2, and the ionosphere-free linear combination P_3 of the two code measurements P_1 and P_2. The use of any other code observable requires the introduction of differential code biases. Any bias in those two code observables will move into a clock parameter, or, in other words, by convention ionosphere-free linear combination P_3 does not contain any code bias, thus one can define code biases in an absolute way. This also means that any absolute bias in P_1 or P_2 would need to be defined in terms of the P_3 observable. Therefore, in the next step we define the geometry-free form of the ambiguity-free linear combination (12.6)

$$\overline{AF}_1 := AF_1 - P_3 = \kappa_1^{af} L_3 + \kappa_2^{af} LP_1 - P_3 = \frac{\kappa_2^{af}}{2} AB_1 \tag{12.22}$$

From (12.22) we may calculate the absolute bias AB_1 on the P_1 code measurements using the absolute bias linear combination defined as

$$AB_1 := \frac{2}{\kappa_2^{af}}(AF_1 - P_3) = \frac{f_1 - f_2}{f_1}(AF_1 - P_3) \tag{12.23}$$

or

$$AB_1 - \frac{c \cdot f_2}{f_1(f_1 - f_2)} N_W = AB_1 - \frac{f_2}{f_1} \lambda_W N_W = \frac{2}{\kappa_2^{af}}(AF_1 - P_3) = \frac{f_1 - f_2}{f_1}(AF_1 - P_3) \tag{12.24}$$

where the wide-lane ambiguity N_W is associated with the wavelength of

$$-\frac{c \cdot f_2}{f_1(f_1 - f_2)} = -\frac{f_2}{f_1} \lambda_W \approx -0.67 \text{ m} \tag{12.25}$$

and the wide-lane wavelength $\lambda_W = c/(f_1 - f_2)$. Assuming that $\sigma(P_1) \approx \sigma(P_2)$, for the noise level of the estimated absolute bias AB_1, we may write

$$\sigma(AB_1) \approx \frac{f_2}{f_1} \cdot \sigma(P_1) \approx 0.78 \cdot \sigma(P_1) \tag{12.26}$$

For code observables on the second GPS frequency we may write

$$AB_2 := \frac{2}{\kappa_{2(2)}^{af}}(AF_2 - P_3) = -\frac{f_1 - f_2}{f_2}(AF_2 - P_3) \tag{12.27}$$

or

$$AB_2 - \frac{c \cdot f_1}{f_2(f_1 - f_2)} N_W = AB_2 - \frac{f_1}{f_2} \lambda_W N_W = \frac{2}{\kappa_{2(2)}^{af}}(AF_2 - P_3) = -\frac{f_1 - f_2}{f_2}(AF_2 - P_3) \tag{12.28}$$

where the wide-lane ambiguity N_W is associated with the wavelength of

$$-\frac{c \cdot f_1}{f_2(f_1 - f_2)} = -\frac{f_1}{f_2} \lambda_W \approx -1.11 \text{ m} \tag{12.29}$$

Assuming that $\sigma(P_1) \approx \sigma(P_2)$, for the noise level of the estimated absolute bias AB_2 we get

$$\sigma(AB_2) \approx \frac{f_2}{f_1} \cdot \sigma(P_1) \approx 1.28 \cdot \sigma(P_1) \tag{12.30}$$

and for the third GPS frequency

$$AB_5 := \frac{2}{\kappa_{2(5)}^{af}}(AF_5 - P_3) = \frac{2f_5 - f_1 - f_2}{f_5}(AF_5 - P_3) \tag{12.31}$$

or

$$AB_5 - \frac{c \cdot f_2}{f_5(f_1 - f_2)}N_W - \frac{c}{f_5}N_{W(1,5)} = \frac{2}{\kappa_{2(5)}^{af}}(AF_5 - P_3) = \frac{2f_5 - f_1 - f_2}{f_5}(AF_5 - P_3) \tag{12.32}$$

or

$$AB_5 - \frac{f_2}{f_5}\lambda_W N_W - \lambda_5 N_{W(1,5)} = \frac{2}{\kappa_{2(5)}^{af}}(AF_5 - P_3) = \frac{2f_5 - f_1 - f_2}{f_5}(AF_5 - P_3) \tag{12.33}$$

where the wide-lane ambiguity N_W is associated with the wavelength of

$$-\frac{c \cdot f_2}{f_5(f_1 - f_2)} = -\frac{f_2}{f_5}\lambda_W \approx -0.90 \text{ m} \tag{12.34}$$

and assuming that $\sigma(P_1) \approx \sigma(P_2) \approx \sigma(C_5)$, with the noise level

$$\sigma(AB_5) \approx \frac{f_1 + f_2 - f_5}{f_5} \cdot \sigma(P_1) \approx 1.38 \cdot \sigma(P_1) \tag{12.35}$$

Let us now remove wide-lane ambiguity in (12.24) and (12.28) with the following ambiguity-free condition

$$\kappa_1^w \frac{f_2}{f_1} + \kappa_2^w \frac{f_1}{f_2} := 0, \qquad \kappa_1^w + \kappa_2^w := 1 \tag{12.36}$$

from which we can derive the following multiplication factors κ_1^w and κ_2^w

$$\kappa_1^w = \frac{f_1^2}{f_1^2 - f_2^2} = \kappa_1, \qquad \kappa_1^w = -\frac{f_2^2}{f_1^2 - f_2^2}\kappa_2 \tag{12.37}$$

which are equal to the multiplication factors of the ionosphere-free linear combinations of κ_1 and κ_2. Since the ionosphere-free linear combination of P_1 and P_2 code observables is by convention free of biases we obtain

$$\kappa_1^w AB_1 + \kappa_2^w AB_2 = \kappa_1 AB_1 + \kappa_2 AB_2 = 0 \tag{12.38}$$

from which it follows that the relationship between absolute code biases and differential code bias DCB_{P_1,P_2}

$$AB_2 = \frac{f_1^2}{f_2^2}AB_1 \quad \rightarrow \quad AB_1 = -\frac{f_2^2}{f_1^2 - f_2^2}(AB_1 - AB_2) = -\frac{f_2^2}{f_1^2 - f_2^2}DCB_{P_1,P_2} \tag{12.39}$$

If we now subtract (12.24) and (12.28) we derive

$$AB_1 - AB_2 + \frac{f_1^2 - f_2^2}{f_1 f_2}\lambda_W N_W = DCB_{P_1,P_2} + \frac{f_1^2 - f_2^2}{f_1 f_2}\lambda_W N_W = \frac{f_1 - f_2}{f_1}\overline{AF_1} + \frac{f_1 - f_2}{f_2}\overline{AF_2} \tag{12.40}$$

Let us now estimate wide-lane ambiguity N_W in (12.24), (12.28) and (12.33). For this we use the following two ionosphere-free linear combinations (omitting the receiver/satellite clock parameters)

$$L_3 = \rho + \lambda_N N_1 + \frac{1}{2}\left(\lambda_W - \lambda_N\right) N_W$$

$$L_3^{2,5} = \rho + \lambda_{N(2,5)} N_1 - \lambda_{N(2,5)} N_W + \frac{1}{2}\left(\lambda_{W(2,5)} - \lambda_{N(2,5)}\right) N_{W(2,5)}$$

(12.41)

In order to eliminate the N_1 ambiguity we use the following ambiguity-free and geometry-free condition

$$\kappa_1^{af*}\lambda_N + \kappa_2^{af*}\lambda_{N(2,5)} := 0, \qquad \kappa_1^{af*} + \kappa_2^{af*} = 1$$

(12.42)

from which we obtain the multiplication factors

$$\kappa_1^{af*} = \frac{-\lambda_{N(2,5)}}{\lambda_N - \lambda_{N(2,5)}} = \frac{f_1 + f_2}{f_1 - f_5} \approx 7.02 \qquad \kappa_2^{af*} = \frac{\lambda_N}{\lambda_N - \lambda_{N(2,5)}} = -\frac{f_2 + f_5}{f_1 - f_5} \approx -6.02$$

(12.43)

For more on this linear combination we refer to Section 22. Finally, for this ambiguity-free linear combination we derive

$$L_3^{af*} := \kappa_1^{af*} L_3^{1,2} + \kappa_2^{af*} L_3^{2,5}$$

$$= \rho + \left[\frac{\kappa_1^{af*}}{2}\left(\lambda_W - \lambda_N\right) - \kappa_2^{af*}\lambda_{N(2,5)}\right] N_W + \frac{\kappa_2^{af*}}{2}\left(\lambda_{W(2,5)} - \lambda_{N(2,5)}\right) N_{W(2,5)}$$

(12.44)

with the wavelengths of wide-lane ambiguities that are relatively very long, i.e.,

$$\lambda_W^{af*} = \frac{\kappa_1^{af*}}{2}\left(\lambda_W - \lambda_N\right) - \kappa_2^{af*}\lambda_{N(2,5)} \quad \approx 3.40 \text{ m}$$

$$\lambda_{W(2,5)}^{af*} = \frac{\kappa_2^{af*}}{2}\left(\lambda_{W(2,5)} - \lambda_{N(2,5)}\right) \qquad \approx -17.28 \text{ m}$$

(12.45)

In a similar way, we can eliminate wide-lane ambiguity by combining (12.13) with (12.44)

$$-\frac{c \cdot f_2}{(f_1 - f_2)^2}\kappa_1^{af**} + \kappa_2^{af**}\lambda_W^{af*} := 0, \qquad \kappa_1^{af**} + \kappa_2^{af**} = 1$$

(12.46)

with

$$\kappa_1^{af**} \approx 0.53, \quad \kappa_2^{af**} \approx 0.47$$

(12.47)

and (12.17) with (12.44), using the following ambiguity-free and geometry-free condition

$$\frac{c \cdot f_1}{(f_1 - f_2)^2}\kappa_1^{af***} + \kappa_2^{af***}\lambda_W^{af*} := 0, \qquad \kappa_1^{af***} + \kappa_2^{af***} = 1$$

(12.48)

with

$$\kappa_1^{af***} = -6.80, \quad \kappa_2^{af**} = 7.80$$

(12.49)

After removing the geometry term by subtracting P_3 we obtain

$$AF_1^* = \kappa_1^{af**} AF_1 + \kappa_2^{af**} L_3^{af*} - P_3 = \kappa_1^{af**}\frac{f_1}{f_1 - f_2} AB_1 + \kappa_2^{af**}\lambda_{W(2,5)}^{af*} N_{W(2,5)}$$

(12.50)

and

$$AF_2^* = \kappa_1^{af***} AF_2 + \kappa_2^{af***} L_3^{af*} - P_3 = \kappa_1^{af***} \frac{-f_2}{f_1 - f_2} AB_1 + \kappa_2^{af***} \lambda_{W(2,5)}^{af*} N_{W(2,5)} \tag{12.51}$$

that give

$$AB_1 = \frac{f_1 - f_2}{\kappa_1^{af**} f_1} \left[\left(\kappa_1^{af**} AF_1 + \kappa_2^{af**} L_3^{af*} - P_3 \right) - \kappa_2^{af**} \lambda_{W(2,5)}^{af*} N_{W(2,5)} \right] \tag{12.52}$$

$$AB_2 = -\frac{f_1 - f_2}{\kappa_1^{af***} f_2} \left[\left(\kappa_1^{af***} AF_2 + \kappa_2^{af***} L_3^{af*} - P_3 \right) - \kappa_2^{af***} \lambda_{W(2,5)}^{af*} N_{W(2,5)} \right] \tag{12.53}$$

and finally

$$AB_1 = \frac{f_1 - f_2}{\kappa_1^{af**} f_1} AF_1^* + \lambda^{AF^{**}} N_{W(2,5)} \tag{12.54}$$

$$AB_2 = -\frac{f_1 - f_2}{\kappa_1^{af***} f_2} AF_2^* + \lambda^{AF^{***}} N_{W(2,5)} \tag{12.55}$$

with the wavelengths $\lambda^{AF^{**}}$ and $\lambda^{AF^{***}}$ of the wide-lane ambiguity $N_{W(2,5)}$

$$\lambda^{AF^{**}} \approx 3.41 \text{ m} \tag{12.56}$$

$$\lambda^{AF^{***}} \approx 5.62 \text{ m} \tag{12.57}$$

Assuming that $\sigma(P_1) \approx \sigma(P_2)$, the noise levels are

$$\sigma(AB_1) = \sqrt{0.06^2 \sigma^2(P_1) + 0.65^2 \sigma^2(P_2)} \approx 0.65 \cdot \sigma(P_2) \approx 0.65 \cdot \sigma(P_1) \tag{12.58}$$

$$\sigma(AB_2) \approx \sqrt{1.06^2 \sigma^2(P_1) + 0.11^2 \sigma^2(P_2)} \approx 1.07 \cdot \sigma(P_1) \tag{12.59}$$

It is interesting to note that the noise level of AB_1 is mainly driven by $\sigma(P_2)$, whereas the noise level of AB_2 by $\sigma(P_1)$. Closer look at (12.56) and (12.55), in addition to noise level (12.59) confirms the scaling factor $AB_2 = f_1^2 / f_2^2 AB_1$ in (12.39).

Another approach to estimate AB_1 is to subtract (12.17) and (12.44)

$$AF_2 - L_3^{af*} = \left(\frac{c \cdot f_1}{(f_1 - f_2)^2} - \lambda_W^{af*} \right) N_W - \lambda_{W(2,5)}^{af*} N_{W(2,5)} - \frac{f_2}{f_1 - f_2} AB_2 \tag{12.60}$$

Inserting $AB_2 = f_1^2 / f_2^2 AB_1$ we derive

$$AB_1 = -\frac{f_2(f_1 - f_2)}{f_1^2} \left[AF_2 - L_3^{af*} + \left(\frac{c \cdot f_2}{f_1(f_1 - f_2)} - \lambda_W^{af*} \right) N_W + \lambda_{W(2,5)}^{af*} N_{W(2,5)} \right] \tag{12.61}$$

12.3 Absolute Code Biases and Melbourne-Wübbena Linear Combination

Since our ambiguity-free linear combination (12.22) is geometry-free, there must be a direct relation to the wide-lane biases in the Melbourne-Wübbena linear combination

$$MW_{1,2} = \lambda_W N_W = L_W - P_N = \frac{f_1}{f_1 - f_2} L_1 - \frac{f_2}{f_1 - f_2} L_2 - (\frac{f_1}{f_1 + f_2} P_1 + \frac{f_2}{f_1 + f_2} P_2) \tag{12.62}$$

that is also both geometry-free and ionosphere-free. In (12.62) L_W and P_N represent the wide-lane and narrow-lane linear combinations of the carrier-phase and code measurements respectively, while N_W is the wide-lane ambiguity and $\lambda_W = c / (f_1 - f_2)$ the wide-lane wavelength. In contrast to the ionosphere-free linear combination P_3, the Melbourne-Wübbena linear combination is in general, by convention, not free of biases. Therefore, in (12.62) we need to introduce the narrow-lane bias denoted as δ_N

$$MW_{1,2} = L_W - P_N = \lambda_W N_W + \delta_N \tag{12.63}$$

We will see later that the wide-lane biases can be removed by forming track-to-track ambiguities (pass-to-pass ambiguities) between consecutive tracking passes at the zero-difference level which can then be removed by forming double-differences. However, this is not the case for the ambiguity resolution of wide-lane ambiguities using zero-difference measurements. By estimating absolute code bias, wide-lane biases can be adequately dealt with for all GNSS code observables in a multi-frequency GNSS environment. It can be shown that the following relation exists between our geometry-free form (12.22) of the ambiguity-free linear combination $\overline{AF}$ and the Melbourne-Wübbena linear combination (12.62)

$$MW_{1,2} = -\frac{f_1 - f_2}{f_2} \overline{AF}_1 \tag{12.64}$$

$$MW_{1,2} = \frac{f_1 - f_2}{f_1} \overline{AF}_2 \tag{12.65}$$

From (12.64) and (12.65) it follows that absolute code biases can be calculated directly from the narrow-lane biases δ_N

$$AB_1(P_1) = AB_1 = -\frac{f_2}{f_1} \delta_N \tag{12.66}$$

$$AB_2(P_2) = AB_2 = -\frac{f_1}{f_2} \delta_N \tag{12.67}$$

Since GNSS satellite clock data provided by the IGS refer, by convention, to the ionosphere-free linear combination P_3, following our expression for the absolute code biases (12.66) and (12.67) we may write

$$AB_2(P_2) = \frac{f_1^2}{f_2^2} AB_1(P_1) \tag{12.68}$$

Therefore, we can calculate differential-code bias directly from the narrow-lane bias

$$DCB_{P_1,P_2} = AB_{P_1} - AB_{P_2} = -\frac{f_2}{f_1} \delta_N + \frac{f_1}{f_2} \delta_N \tag{12.69}$$

from which we may derive the following relation between the differential-code bias and the narrow-lane bias

$$DCB_{P_1,P_2} = \frac{f_1^2 - f_2^2}{f_1 f_2} \delta_N \tag{12.70}$$

This also means that for a given differential-code bias, one can calculate the narrow-lane bias

$$\delta_N = \frac{f_1 f_2}{f_1^2 - f_2^2} DCB_{P_1,P_2} \tag{12.71}$$

113

and for the absolute code biases we finally obtain

$$AB_1(P_1) = AB_1 = -\frac{f_2^2}{f_1^2 - f_2^2} DCB_{P_1,P_2} \tag{12.72}$$

$$AB_2(P_2) = AB_2 = -\frac{f_1^2}{f_1^2 - f_2^2} DCB_{P_1,P_2} \tag{12.73}$$

Similar relations can be derived for the Melbourne- Wübbena linear combination $MW_{1,5}$ for code and carrier-phase measurements on f_1 and f_5 frequencies as

$$MW_{1,5} = -\frac{f_1 - f_5}{f_5} \overline{AF}_{5(1,5)} \tag{12.74}$$

where $\overline{AF}_5^*$ is the ambiguity-free linear combination defined relative to the ionosphere-free code measurements $P_{3(1,5)}$

$$\overline{AF}_5^* := AF_5 - P_{3(1,5)} = \kappa_1^{af} L_{3(1,5)} + \kappa_2^{af} LP_5 - P_{3(1,5)} = \frac{\kappa_2^{af}}{2} AB_1 \tag{12.75}$$

$$AB_{1(1,5)}(P_1) = AB_{1(1,5)} = -\frac{f_5}{f_1} \delta_{N(1,5)} \tag{12.76}$$

As mentioned above, Galileo and future GNSS will introduce wide-band signals with low code noise (in the cm-range). The Galileo E5 wide-band signal (nominal bandwidth 51.15 MHz) and the AltBOC modulation will offer code noise at cm-level. However, this is not the case for its subcarriers E5a and E5b. Therefore, the same approach could be applied to Galileo using the E5 signal as a reference observable to derive absolute code biases.

12.4 Estimation of DCBs and Absolute Code Biases

In order to demonstrate this new approach, Figure 12.1 shows the absolute code biases on P_1 for satellites of the GPS constellation from the station ZIMJ over a period of 11 days, using dual- and triple-frequency GPS measurements. The triple-frequency ambiguity-free linear combination offers a very long wavelength of 3.41 m (12.56). Figure 12.1 show that the noise level over 11 days is $\sigma = \pm 0.065$ m without any elevation-dependent weighting and wind-up effect applied. An additional effect, the apparent clock variations, was reported for the third GPS frequency f_5, see (Montenbruck et al. 2012) that could affect the code measurements on the third GPS frequency. For comparison, the noise of the dual-frequency code biases is $\sigma = \pm 0.027$ m, as can be seen
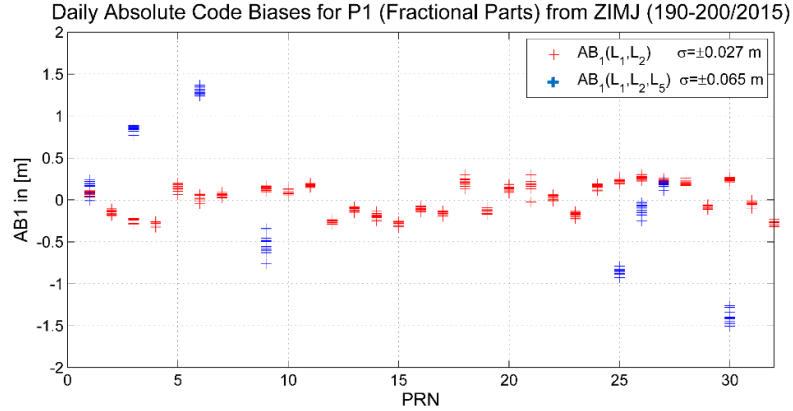
Figure 12.1 Daily estimates of P_1 absolute code biases (11 days) for GPS constellation from ZIMJ station. One can see a very low noise of $\sigma = \pm 0.027$ m for the two-frequency solution and for the triple-frequency solution with wavelengths of 0.67 m and 3.41 m respectively.
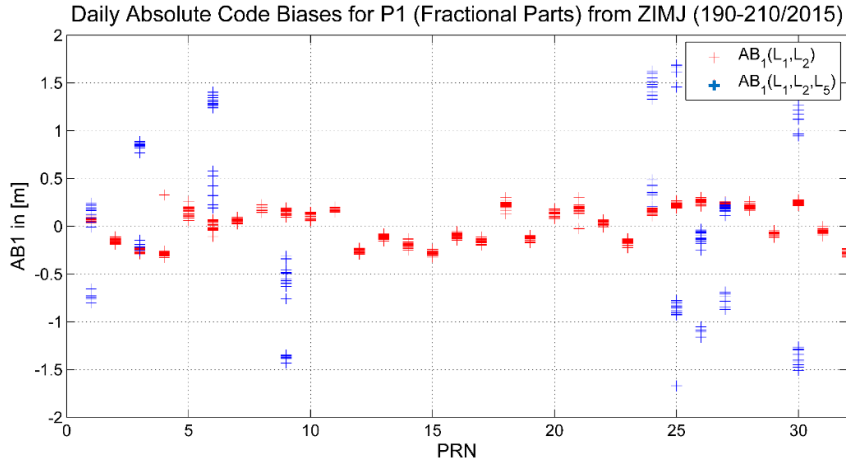
Figure 12.2 Daily estimates of absolute code biases (21 days) for GPS constellation from ZIMJ station (Javad). Over a longer period of time, one can see ¼ carrier-phase ambiguities due to f_5 that give a code bias of about 0.85 cm.

in Figure 12.1. It is very interesting to note that the estimated code ambiguities are very stable and show similar fractional parts over those 11 days. Hoverer, when longer data sets are processed, such as the 21 days in Figure 12.2, one can see that triple-frequency solution shows the ¼ carrier-phase ambiguities typically associated with the L_5 carrier-phase measurements and Javad GPS receiver (Javad TRE_G3TH Delta3.4.9). One quarter of the 3.41-m-wavelength gives a code bias of about 0.85 cm. From this, we can draw the conclusion that resolution of code biases could also be used to detect the ¼-ambiguities associated with carrier-phase measurements.

About 10 satellites in the GPS constellation currently transmit on three, rather than two, GPS frequencies. This opens the opportunity of comparing the estimation of absolute code biases using the two GPS frequency $AB_1(L_1, L_2)$ with the triple-frequency solution $AB_1(L_1, L_2, L_5)$ of higher wavelength. Figure 12.3 shows the resolution of dual-frequency solution against the triple-frequency solution.
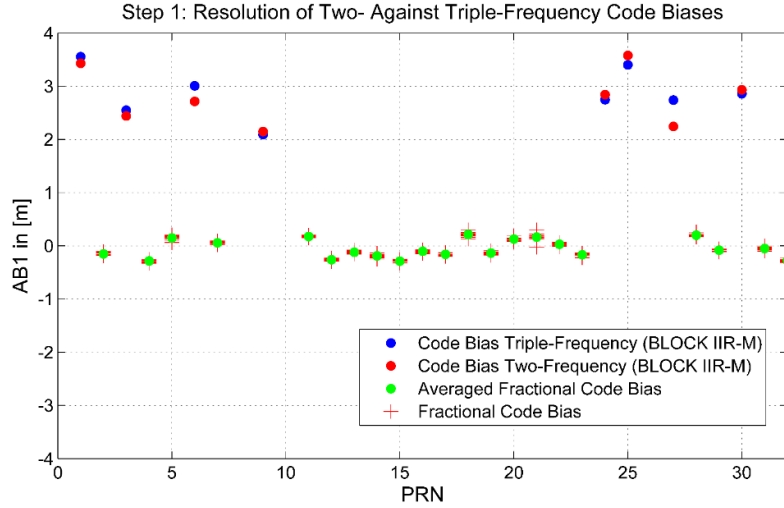
Figure 12.3 Step 2: Resolution of absolute code biases $_A B_1(L_1, L_2)$ with carrier-phase L_5 benefitting from the large wavelength of 3.41 m. Code biases with triple-frequency linear combination are used as a reference for the two-frequency data. We can see that for one satellite, out of 10 satellites in the GPS constellation, the wavelength was fixed incorrectly by 0.67 m. Elevation dependent weighting and wind-up effect were not used.
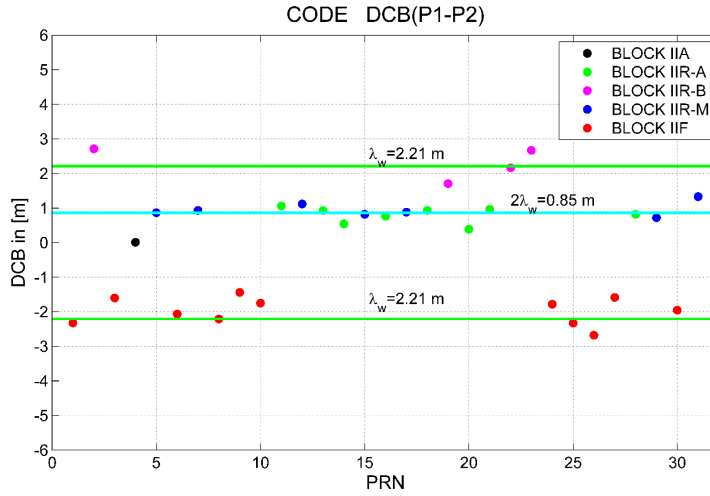
Figure 12.4 Step 3: Estimation of DCBs based on absolute code biases $_A B_1(L_1, L_2)$ and $AB_1(L_1, L_2, L_5)$. Note that DCBs between GPS satellites within the same GPS BLOCK are very small.

The advantage of such a "fixed" dual-frequency solution is the lower noise level (about 50 % lower than that of the triple-frequency solution, see Figure 12.1). For the remaining satellites in the GPS constellation with dual-frequency GPS measurements only, we averaged the fractional code biases in Figure 12.1.

In step 3, from the estimated absolute code biases in Figure 12.3 we derived differential-code biases (DCBs) using (12.72). Due to the multiplication factor in (12.72), the wavelength of 3.41 m is reduced to 2.21 m. and the wavelength of 0.67 m to 0.43 m for the dual-frequency code biases. Figure 12.4 shows that the estimated DCBs are very close to each other within the same GPS BLOCK. This is more visible in Figure 12.5 where the mean DCB is calculated for every GPS BLOCK and subtracted from the individual DCB value for every GPS satellite. From Figure 12.5, one can draw the conclusion that estimated DCB values are within a wavelength of 0.43 m for all satellites in the GPS constellation. A closer look at Figure 12.4 shows that values for all GPS BLOCK IIF satellites with the third GPS frequency are centered at about -2.21 m.
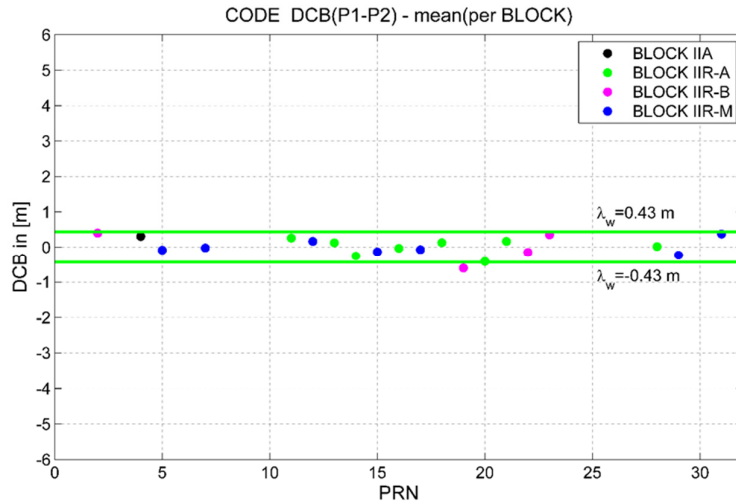
Figure 12.5 Step 3: DCBs between GPS BLOCKs are small: GPS BLOCK II-A/IIR-A/IIR-B/IIR-M.
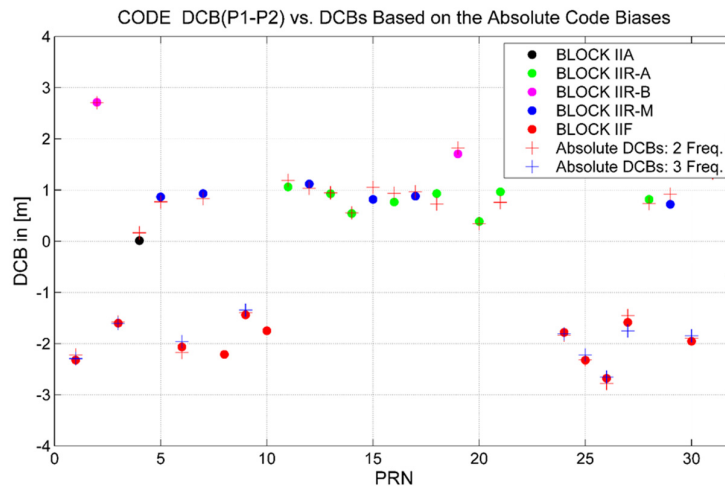
Figure 12.6 DCBs from the absolute code biases in comparison with the DCB available from the CODE Analysis Center. One can see a good overall agreement for dual- and triple-frequency solutions.

Figure 12.6 shows DCBs from Figure 12.5, estimated by making use of the absolute code biases, in comparison with the DCBs provided by the CODE Analysis Center. The difference is also displayed in Figure 12.7, where a mean DCB value (per GPS BLOCK) is subtracted from the single satellite DCB solution. It should be noted that CODE DCBs are based on the two zero-mean conditions, separately applied for the DCBs of GPS satellites and ground receivers, constraining in this way the DCBs for all satellites and receivers to the zero value. In the same way as the GPS satellite clock parameters provided by IGS are referenced to a reference clock in the ground IGS network, our estimates of DCBs values are solely based on the ZIMJ ground station. However, the overall agreement with the DCB values provided by the CODE AC is very good for such a limited data set of only 11 days. From this we can draw the conclusion that the approach presented for deriving absolute code biases offers relatively low noise and a resolution of code biases that can be used in the next step in calculating DCB values. Generally speaking, one can identify three applications of the approach associated with the dual- and triple-frequency GPS data:

- estimation of wide-lane ambiguities (two- and triple-frequency)
- detection of ¼-ambiguities
- resolution of absolute code biases and DCBs.

117
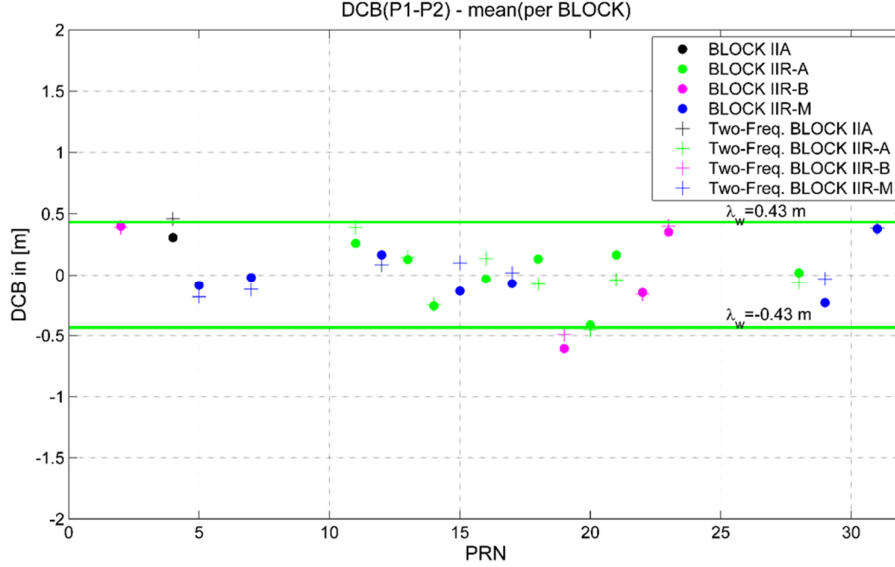
Figure 12.7 CODE DCBs vs. DCBs based on the absolute code biases (mean DCB removed).

12.5 Consistent Datum Definition for GNSS Clock Parameters and Ionosphere Maps

So far we have used the ambiguity-free linear combination (12.22) to derive a geometry-free definition of absolute code biases in terms of the ionosphere-free linear combination of P_1 and P_2 code GPS measurements. However, in the multi-frequency GNSS environment, it would be interesting to estimate GNSS clock parameters based on all carrier-phase measurements, since all forms of error in carrier-phase and code measurements would be averaged over a range of different frequency bands and signal modulations. This is particularly true for the GPS "apparent clock variations" in the case of carrier-phase on a different frequency and different multipath sensitivity in the case of code measurements on several frequencies. It was reported in (Montenbruck et al. 2012) that small delays can be noticed between carrier-phase on different GNSS frequencies. For Galileo, such a thermal delay or internal multipath delay will be negligible. However, ionosphere-free carrier-phase values obtained by averaging measurements over several frequencies will always be more accurate than a single ionosphere-free carrier-phase measurement. In order to be consistent with the IGS convention for GNSS clock parameters, we can always use absolute code biases from different GNSS observables to transform GNSS clock parameters into the two-code measurements used by IGS per convention. Based on (12.68), absolute code biases for the ionosphere-free linear combination can be written as

$$AB_3(P_1, P_2) := \frac{f_1^2}{f_1^2 - f_2^2} AB_1(P_1) - \frac{f_2^2}{f_1^2 - f_2^2} \cdot \frac{f_1^2}{f_2^2} AB_1(P_1) := 0 \qquad (12.77)$$

that gives a zero bias for $AB_3(P_1, P_2) := 0$, as expected. For any other two-frequency ionosphere-free linear combination, the ionosphere-free bias $AB_3(P_i, P_j)$ is

$$AB_3(P_i, P_j) := \frac{f_i^2}{f_i^2 - f_j^2} AB_i(P_i) - \frac{f_j^2}{f_i^2 - f_j^2} AB_j(P_j) \qquad (12.78)$$

Therefore, we may use the zero-bias condition (12.77) to estimate a frequency-independent GNSS clock parameter that is consistent with the IGS convention for GNSS satellite clock parameters. Relative code biases between different observables on the same GNSS frequency can be measured directly.

For a permanent GNSS network, such as that of IGS, track-to-track ambiguities (integer ambiguities between consecutive tracking passes) can be estimated or fixed to their integer values over longer periods of time. After resolving track-to-track ambiguities for a given station and a given GNSS satellite, there remains only one carrier-phase ambiguity to be estimated over several days, weeks or even months. Since absolute code biases are removed by forming track-to-track ambiguities, we can establish a consistent datum to define GNSS satellite clock parameters and absolute code biases. The only possible ambiguity that could arise in this datum definition is the size of the wide-lane ambiguity in (12.23). The size of the wide-lane ambiguity for GPS is around $86 \, \text{cm}$ or $5.86 \, \text{cm}$ for the super-wide-lane, and so can easily be detected by a ground H-maser or GNSS satellite clock. From this, we may draw the conclusion that the ambiguity-free linear combination can be used to define a geometry-free datum for GNSS satellite clock parameters that is consistent with the IGS definition of clock parameters and at the same time is absolute in nature. For the evolution of ground TEC maps over longer periods of time, such a datum definition will provide long-term TEC stability and at the same time can be used to give consistency to any time/frequency transfer over long periods. Moreover, absolute code biases provide a framework for combining all multi-GNSS observables.

12.6 S-Curve Bias and Group Delay Variations

There are several other justifications for the use of absolute code biases in the multi-frequency GNSS, such as code-carrier coherency, S-curve bias and related code-offset delay variations. Code measurements need to be coherent with the carrier-phase for both the satellite and the ground receiver. Any synchronization offset between code and carrier-phase will introduce an additional carrier-phase bias. Such a bias should be constant and satellite-specific. It has been noticed that some GPS receivers in the IGS network sporadically experience so-called ″ms-jumps″, i.e., the code measurements do not have the same receiver time as the carrier-phase measurements. However, for GPS satellites any delay can be considered as constant bias and is driven by the analog technology of the satellite electronics causing different delays between code and carrier-phase. These delays can be measured and monitored by phasemeters on board the GNSS satellite. They can also be partially reduced in a relative sense by applying information on group delay variations provided in the navigation message.

The S-curve bias is an effect that can be measured by a GPS receiver connected to a high-gain antenna (e.g., the size of a VLBI antenna). The very large size of such a GPS antenna reduces the thermal noise of the tracked GPS signal and offers code measurements with sub-cm precision. If the GPS signal is sampled with such an antenna in the open-loop mode or if several GPS receivers are connected to this antenna with different correlator spacing (narrow-band correlator to wide-bands with long integration time), one can observe a bias as a function of the correlator spacing. This S-curve bias effect can also be seen if the group delay pattern of the GPS satellite transmitter is estimated using GPS receivers with different correlator spacing, since results will be biased to each other. This is also one of the reasons why so far no reliable maps of group delay variations for GPS satellites have been provided by the IGS, as GPS receivers based on different correlator spacing will give slightly different results.

One can show that when different GPS satellites are compared, S-curve bias effects can easily reach $1.5 - 2.5$ ns for C/A and P code with differences in the order of 0.5 ns between different GPS satellites. The related code-offset delay variations could easily reach $0.5 - 1.0$ ns between different GPS satellites. These variations in the absolute code biases between different GPS satellites and receivers play a very important role in the definition of the IGS time scale, especially on a day-to-day basis. It is well known that the daily solutions of the GPS satellite clock parameters are biased by about 1 ns when the common clock parameter at day-boundary is compared.

The best way forward to calibrate code measurements for different GNSS receiver classes and to estimate group delay variation maps for GNSS satellites is to use a high-gain ground GNSS antenna. Such an antenna should be as large as possible, i.e. at least as the VLBI antenna, in order to reduce thermal noise and should

be connected to several different GNSS receivers with different correlator spacing (narrow- to wide-bands with longer integration time).

13. LEO Near-Field Multipath and Antenna Effects

In an internal technical note (Švehla and Rothacher 2004d), it was suggested to the GOCE Project Office in ESA that a study be conducted on the effect of the near-field multipath on a POD antenna due to the structural environment of the GOCE satellite itself. The idea was that by performing an absolute calibration of the GOCE antenna, with and without a mock-up (solar panel wing), the near-field multipath effect could be described as the difference between the two estimated PCV maps. In the case of near-field multipath, the total antenna PCV correction can be defined as the sum of the nominal antenna PCV map and the antenna map resulting from the near-field multipath. This section studies multipath effects originating from the satellite environment and the impact of GPS antenna calibration on orbit determination of LEO satellites. It is shown that near-field multipath has a very strong effect on the kinematic POD of a LEO satellite using carrier-phase measurements. At the end of this section, a near-field multipath calibration method is proposed and then discussed for GNSS satellites.

13.1 Near-Field Multipath Onboard LEO Satellite

Multipath is one of the main factors limiting the positioning accuracy of GNSS. For carrier-phase measurements, its theoretical maximum is a quarter of the carrier wavelength, or about 4.8 cm for L-band frequencies. For pseudo-range measurements, the situation is significantly worse and the theoretical maximum effect is half the code chip length, i.e., for P code measurements it is about 15 m and for C/A code up to 150 m when the reflected/direct signal amplitude ratio is 1. Several authors have reported measured multipath on P code pseudo-ranges between 1.3 m, in a benign environment, and 4 to 5 m in a highly reflective environment, (for more see (Langley 1998)). For C/A code, values that are up to one order of magnitude larger may be expected. Generally speaking, multipath can be to a great extent mitigated by the GPS receiver's multipath mitigation techniques and by a choke-ring antenna.

There is a principal difference between the multipath of a ground GPS antenna and that of a GPS antenna on board a satellite. A LEO antenna is less affected by the far-field reflectors and the main multipath driver on board a LEO satellite is the near-field antenna environment and solar panels. Other payloads in the vicinity can also interfere with the POD antenna, e.g., a radio-occultation antenna placed close to the POD antenna, as is the case with the CHAMP satellite (cross-talk). By tuning tracking loops and optimizing multipath mitigation techniques within the spaceborne GPS receiver, multipath can be minimized to a large extent. In the case of geodetic satellite missions with the highest POD requirements, it is important to ensure that there is a flat surface and a clear horizon surrounding the GPS antenna. The possible impact of the satellite surface rims on the GPS signal, as in the case of the CHAMP satellite, was pointed out by (Isler, priv. com.) and confirmed using GPS data as shown in Figure 13.6 later in this section. In some cases, the GPS receiver itself can be a multipath generator. With the four Septentrio Polarex GPS receivers connected

to the same GPS antenna and external Cs-frequency standard, we noticed that single-differences formed between different receivers show significant differences of up to 1.5 cm on the second GPS frequency. The first explanation offered for this multipath-like effect was that it was due to the front-end of the receiver (Simsky, priv. com.). However, the signal splitters of the antenna cable has also been put forward as a possible cause of this effect (Petit, priv. com.). That a GPS receiver itself can generate a multipath-like effect was also reported for the IGOR GPS receiver (Montenbruck, priv. com.) on the TerraSAR mission.

It is known that multipath interference induced by reflecting objects in the very close vicinity of GPS antennae (e.g., surfaces of pillars) as well as antenna imaging effects and diffraction cause near-field effects on the GPS signal received, (see e.g., (Elósegui et al. 1995)). (Elósegui et al. 1995) reported that the part of the GPS signal scattered from the surface of a pillar on which a GPS antenna is mounted interferes with the direct signal. The error depends on the elevation angle of the satellite, varies slowly with elevation angle and time and is not necessarily eliminated by changing the antenna configuration and/or lengthening the baselines. It introduces systematic errors at the centimetre-level in the estimates of all parameters including site coordinates and residual tropospheric propagation delays, see (Elósegui et al. 1995). Although imaging and true multipath are similar phenomena and are often simply called multipath, they are frequently distinguished from scattering (Langley 1998). Multipath effects, when averaged over a longer time period, will be considerably reduced for ground static positioning or reduced-dynamic POD of LEO satellites. However, this is not true for imaging effects, which, by definition, leave biases in the measurements, since the reflecting object generates an image of the antenna and the effective antenna is a combination of the nominal antenna and its image. Imaging effects for LEO satellites can easily be demonstrated in the case of the Jason-1 satellite, where the GPS antenna is tilted towards the satellite body by about 45° and, together with the reflecting satellite surface, generates a new antenna pattern. Looking at the calibrated Jason-1 antenna map, see e.g., (Haines et al. 2004), one can easily recognize the satellite's structure in the derived antenna PCV maps.

The far-field multipath effects caused by reflecting objects located further away from the GPS antenna tend to be much weaker compared to the signals reflected in the vicinity of the antenna and can be to a greater extent reduced by the antenna design and receiver mitigation methods. Despite the difference in amplitude, the near-field and far-field multipath also have different periodic behavior. By means of signal processing techniques, a GNSS receiver can mitigate the effect of multipath when the multipath distance (the difference between the direct path and the indirect path) is more than about 10 m. In cases where the antenna is mounted on a satellite or a boom, the multipath distance is much shorter than 10 m and the multipath cannot be mitigated significantly. A number of different receiver-tracking techniques have been developed to mitigate multipath, e.g., using narrow correlators or using multiple-correlator channels to estimate multipath. For more about multipath mitigation techniques, especially those based on receiver-internal approaches such as the narrow correlation technique, double delta correlator implementation and Early/Late Slope (ELS) techniques we refer to (Irsigler 2008). The author also discusses several other mitigation approaches, such as those based on arrays of closely spaced antennae. It is demonstrated that carrier smoothing does not ensure efficient multipath mitigation in any situation (Irsigler 2008). In the same source, a new multipath monitoring approach is proposed based on multi-correlator observations. It allows instantaneous detection of multipath signals and can be used to detect very weak multipath-affected observations (Irsigler 2008).

Near-field multipath, caused by the satellite structure in the vicinity of the antenna or the satellite underneath the antenna, can be mitigated to a large extent by choke-ring ground planes. A choke-ring ground plane consists of several concentric thin rings around the antenna element in the center. The principal disadvantage of the choke-ring design is that the radial spacing of the rings is related to the wavelength of the GNSS signal and therefore the choke-ring can be tuned only for one frequency at a time. This is the reason why a choke-ring antenna in a conical form has been developed for different GNSS frequencies (Leica AR25 r.4), although often incorrectly termed a ″pyramid″.

In (Švehla and Rothacher 2004d), it was suggested to the GOCE Project Office in ESA that a study be conducted on the near-field multipath caused by the structural environment of a GOCE satellite. We proposed calibrating the GOCE GPS antenna using the robot absolute calibration method, with and without the GOCE

mock-up, i.e., the solar panel (wing) as shown in Figure 13.1. The antenna calibration was then performed by ESA (the GOCE Project Office). With the calibration set up depicted in Figure 13.1 it is possible to calibrate phase center variations in an absolute manner by tilting and rotating an antenna. The errors from sources such as the ionosphere or troposphere, or satellite orbit/station errors are eliminated by using an additional nearby reference station. The idea was to perform antenna calibration with and without a mock-up and to describe the near-field multipath effect as the difference between the two estimated antenna PCV maps. In the case of near-field multipath, the total antenna PCV correction δ_{PCV} can be defined as the sum of the nominal antenna PCV map δ_{PCV}^{nom} and the antenna map originating from the near-field multipath δ_{NFM}

$$\delta_{PCV} \coloneqq \delta_{PCV}^{nom} + \delta_{NFM} \tag{13.1}$$

It is assumed that near-field multipath can be represented in much the same way as the nominal antenna map using an elevation/azimuth grid or a spherical harmonic representation. However, due to the shape and structure of the reflecting antenna environment, large gradients can be expected, especially for the L_2 frequency and therefore, the near-field multipath map should be provided with sufficient resolution. Figure 13.2 confirms that larger gradients can be expected for the L_2 frequency. In the case of satellites with movable parts, such as solar panels on board COSMIC satellites in LEO orbit, near field multipath cannot be adequately represented by only one antenna map, but requires a function of time or Sun position in the antenna reference frame. In this case, several antenna maps could be used to model near-field multipath based on representative cases of the antenna environment.

Later on in (Wübbena et al. 2006), the near-field multipath was studied for different configurations of a ground GPS antenna, in particular an antenna mounted on a pillar with different antenna heights from 7 cm up to 27 cm, as well as an antenna installation mounted on a standard tripod (height of 175 cm). A significant low-frequency effect even at high elevations was reported, especially in the first case. In particular, a systematic bias, predominantly in the height component, was reported, which does not average out over long observation time periods and increases with lowering antenna height. Although only elevation-dependent effects have been shown, there are also azimuth-dependent influences for asymmetric configurations. The same paper (Wübbena et al. 2006) includes report analyses for the two common geodetic set-ups using Dorne Margoline choke-ring antenna with a tribrach on a round (diameter 20 cm) and a quadratic pillar (sides of 30 cm). The influence of the near-field has a magnitude of up to 7.5 mm at low elevations and even 5 mm at $10°$ elevation. For some azimuthal regions at the horizon the effect was even larger.

In (Lau and Cross 2007) it is developed a new ray-tracing approach for carrier-phase multipath modeling. It takes into account the relative positions of the receiving antenna and reflectors, relative permittivity of the reflecting surfaces, the correlator spacing of the receiver, the RCP gain pattern of the receiving antenna and the phase center offset and variation map of the receiving antenna. Sensitivity tests with the model showed that the accuracy of the predicted multipath errors was highly dependent on precise knowledge of the relative antenna-reflector geometry. For instance, errors of up to 1 cm in their relative height can cause errors in the modelled multipath from reflectors below the antenna of up to 1 cm. It was shown that an error of up to 10% in the assumed permittivity of the reflector would not have a noticeable effect on the modeled multipath (Lau and Cross 2007).

Figure 13.1 shows the proposed set-up for the near-field multipath calibration of the GOCE antenna using a robot (Švehla and Rothacher 2004d). One can recognize the GOCE helix antenna mounted on the solar panel (mock-up). The original idea was to perform absolute calibration both with and without the mock-up, and compare the resulting antenna maps. The difference between the two respective antenna maps is a measure of the constant near-field multipath environment.

Figure 13.1 GOCE GPS antenna set-up for near-field multipath calibration using a GOCE mock-up as originally proposed to ESA (GOCE Project Office) in (Švehla and Rothacher 2004d). (credit ESA)

The impact of the GOCE solar panel wing (mock-up), depicted in Figure 13.1, on the L_1 and L_2 antenna phase-center is shown in Figure 13.2, based on the PCV maps provided by the GOCE Project Office in ESA. One can clearly see an effect due to the satellite mock-up and a significantly increased effect for the L_2 frequency of about -3 mm at 45° zenith angle and 10 mm at 90° (close to antenna horizon).

By forming the ionosphere-free linear combination of the original L_1 and L_2 phase patterns, it can be determined that the phase-center variations are increased by up to 1 cm when the wing is present, see Figure 13.3. This is an increase by at least a factor of three compared to the accuracy of the original L_1 and L_2 GPS carrier phase measurements. Reference directions (0° azimuth) for all GOCE PCVs is the along-satellite axis (flight direction).
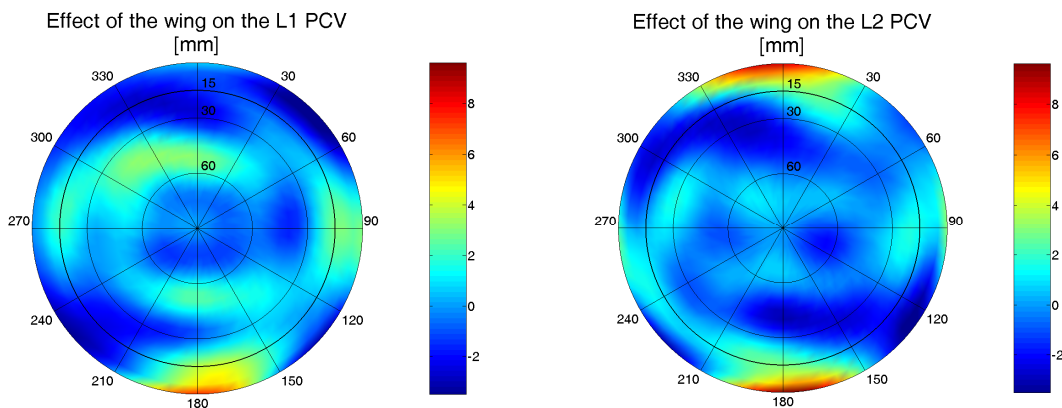
Figure 13.2 The GOCE near-field multipath. Impact of the GOCE solar panel (Figure 13.1) on the GPS antenna phase-center pattern. Phase-center variations are at the level of a few millimeters.

Effect of the wing on the iono-free PCV
[mm]
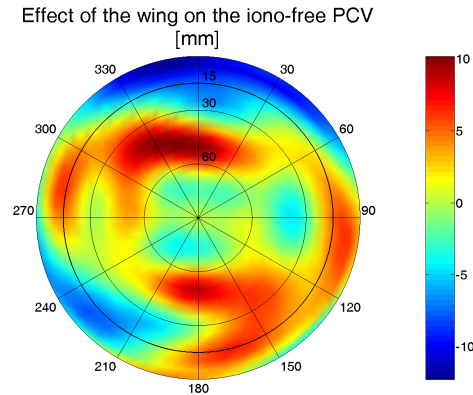
Figure 13.3 GOCE antenna phase-center variations for the ionophere-free linear combination.

This study shows that the antenna calibration results depend heavily on the antenna environment during the calibration. Therefore, it is necessary to calibrate the antenna together with the satellite mock-up, i.e., using a set-up similar to the real satellite. For highly accurate LEO POD, the use of a choke-ring antenna to suppress the near-field multipath is required, i.e., a "conical design" for all GNSS frequencies.

13.2 Impact of the Near-Field Multipath on GOCE Kinematic POD

In order to study the effect of the near-field multipath on the GOCE kinematic POD, GPS phase measurements were simulated for the GOCE antenna with and without PCV maps characterizing the near-field multipath on board the satellite. The simulation was carried out with a cut-off angle of $0°$ for the nominal case and also with $15°$, given the limitations of the GOCE GPS receiver that only locks the signals above $12°$ elevation. From Figure 13.4, one can see variations in the kinematic positions in the order of several centimeters for a single orbit component ($1-3$ cm RMS), or $3-5$ cm RMS for the 3D orbit error. To a large extent this effect can be smoothed out by employing reduced-dynamic POD. The long-periodic structure is clearly visible.

It was assumed that elevation-dependent weighting for phase measurements would down-weight the effect of near-field multipath on the GOCE antenna. Figure 13.5 shows the results when elevation-dependent weighting was employed. The differences are even greater ($3-5$ cm 3D RMS), due to the fact that the entire antenna map is affected, and that the conventional elevation-dependent weighting is not optimal in this case. It should be noted that the antenna phase-center offset was not corrected in the case when the antenna PCV map was used with a cut-off angle of $15°$ elevation. However, such an offset would mainly produce bias in the radial direction, since the kinematic coordinates were estimated every epoch. The presence of a radial bias is not visible in Figure 13.4.

125
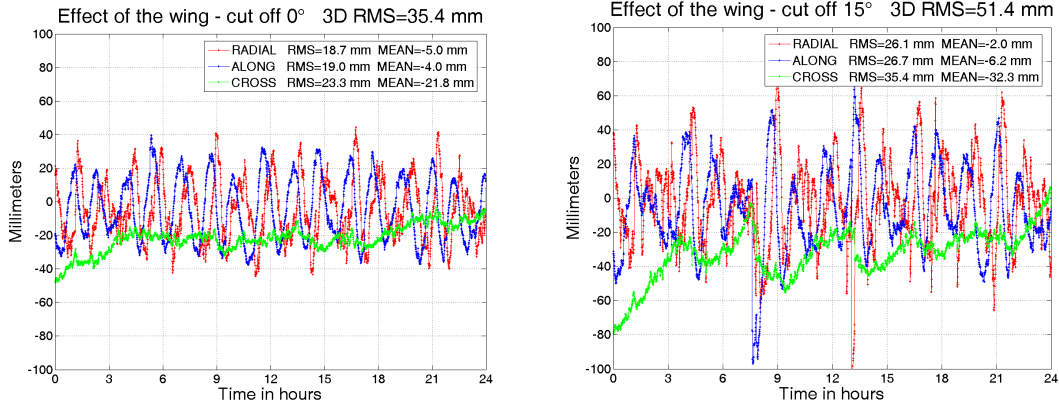
Figure 13.4 The effect of the near-field multipath on the GOCE kinematic POD for 2 different cut-off angles.
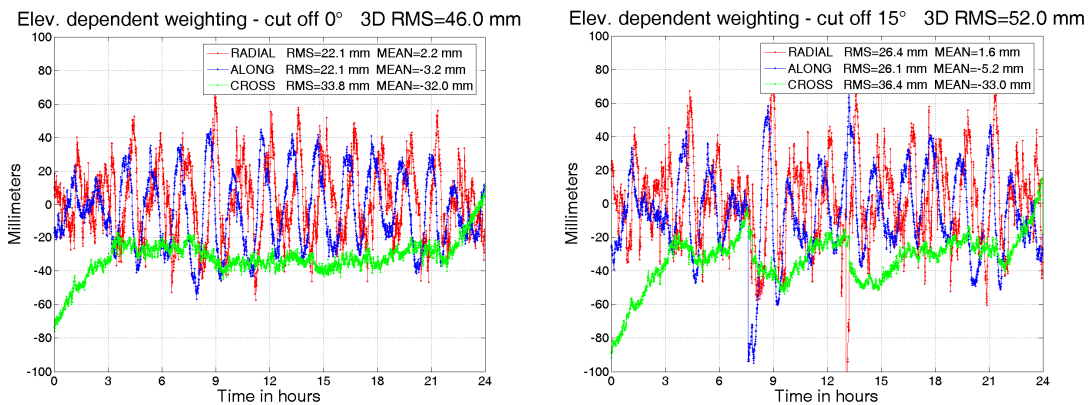
Figure 13.5 The effect of the near-field multipath on the GOCE kinematic POD for 2 different antenna cut-off angles when using elevation-dependent weighting.

13.3 CHAMP Near-Field Multipath

Figure 13.6 shows the CHAMP ionosphere-free code residuals vs. azimuth, after precise orbit determination (POD). The orbit was estimated with carrier-phase measurements and to obtain residuals of code measurements, the CHAMP reduced-dynamic orbit was fixed and clock parameters were estimated every epoch using only ionosphere-free P code measurements. Considering the very low noise level of these code measurements and the orbit quality of several centimeters, it can be assumed that the code residuals obtained are only affected by noise and multipath effects. One can clearly see multipath originating from the aft side of the satellite, (Švehla and Rothacher 2003c). Two particular directions can be clearly identified: at azimuth 135° and at azimuth 225° (Švehla and Rothacher 2003c) and this effect is related to the shape of the CHAMP satellite structure (sharp edges on the aft side), see Figure 13.7. The satellite holding mechanism, visible in Figure 13.7 in the same direction, was also identified as a potential candidate to explain those two particular directions. Multipath at azimuth 180° is mainly driven by the receiving radio-occultation antenna, see also (Montenbruck and Kroes 2003).
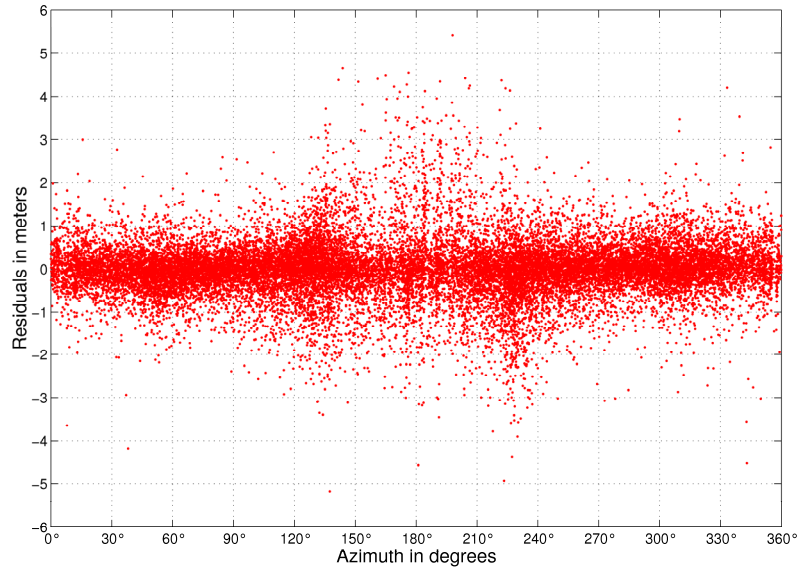
Figure 13.6 CHAMP ionosphere-free code residuals vs. azimuth for day 200/2002 clearly show multipath originating from two particular directions: at azimuth 135° and 225° (due to satellite structure - edges, see Figure 13.7). Multipath at azimuth 180° is due to radio-occultation antenna (cross-talk with POD antenna).

13.4 CHAMP/GRACE GPS Antenna

The POD antenna on board the CHAMP, GRACE and TerraSAR/Tandem missions is a space qualified GPS patch antenna (Sensor Systems S67-1575-14 model), see Figure 13.7. This model was selected for the CHAMP mission because of the slightly better performance of this antenna on the L_2 frequency and its better thermal behavior (Grunwaldt, priv. com.). The helix radio occultation antenna on board the CHAMP satellite is the JPL design with a gain of up to 9 dB along boresight and a half cone (at 3 dB) of 45°. For radio-occultation satellites such as COSMIC, radio-occultation antennae are built up from stacked patch arrays of 4×1 elements, which should have a gain of up to 11.5 dB, but exhibit a very narrow gain characteristic (Grunwaldt, priv. com.). For more information on the POD of the Formosat-3/COSMIC mission we refer to (Hwang et al. 2009), (Tseng et al. 2012).

Figure 13.7 shows the main CHAMP POD antenna together with the choke ring, pointing in the zenith direction, and the helix antenna, directed aft. The lightweight choke ring developed by the GeoForschungsZentrum (GFZ) Postdam is used to suppress multipath and to achieve high-grade code and carrier-phase measurements. In Figure 13.7, one can also see a small POD antenna without a choke ring close to the helix antenna and also directed aft. This is a spare POD antenna and has never been used. The GRACE mission uses BlackJack GPS receivers and GPS antennae of the same type as CHAMP. There are two omni-directional POD antennae (one primary in the zenith direction and one backup directed aft) and one high-gain helix antenna with a 45° field of view directed aft. The aft-pointing POD antenna serves as a redundant source for orbit determination in case of a failure of the zenith antenna.
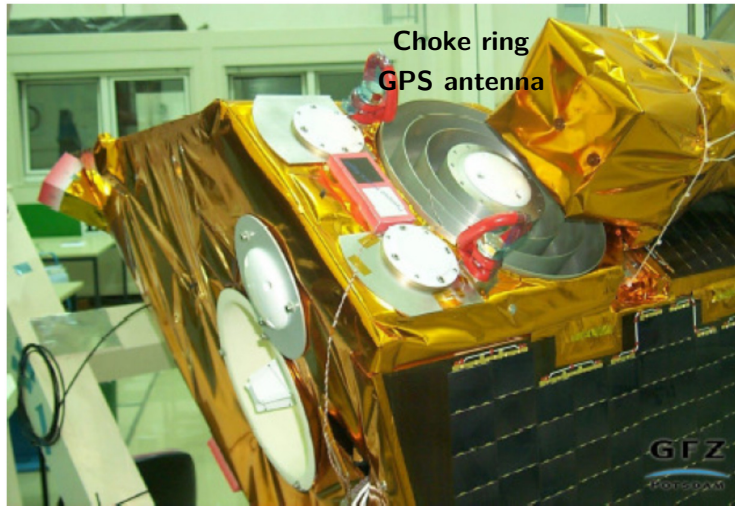
Figure 13.7 CHAMP POD antenna with the choke ring (zenith) and radio-occultation antenna (aft). Next to the radio-occultation (helix) antenna is the spare POD antenna without choke ring (aft). (credit GFZ)

13.5 Antenna Calibration on Board CHAMP, GRACE and JASON Satellites

Two types of antenna calibration can be performed: relative and absolute. Relative antenna calibration is based on the Dorne Margolin T choke-ring antenna as the reference antenna for all other GPS antennae. It can be performed on a very short baseline using a series of GPS measurements. However, the absolute antenna calibration, carried out independently with a robot and in a separate procedure in an anechoic chamber, showed in both instances that the Dorne Margolin T antenna phase center varies with elevation and azimuth. Starting with the GPS week 1400, IGS has included absolute phase-center offsets and patterns for all ground GPS antennae and GPS satellites in its routine processing of global IGS data, (Schmid and Rothacher 2003), (Gendt 2006).

In the case of the Jason-1 mission, several attempts have been made to calibrate the GPS antenna using dynamic POD and to estimate the GPS antenna parameters together with the orbit parameters. Due to the high orbit altitude, it is possible to perform fully dynamic POD for the Jason-1 satellite by estimating a relatively small number of orbit parameters. Hence in-orbit antenna calibration is feasible. At the same time, highly accurate and fully independent Jason-1 orbits based on SLR and DORIS are available for comparison. More about Jason-1 POD and the related GPS antenna problem on the Jason-1 satellite can be found in (Luthcke et al. 2003), (Haines et al. 2004) and (Flohrer et al. 2011).

Figure 13.8 shows the CHAMP absolute phase-center variation over all elevations estimated using ionosphere-free carrier phase measurements together with all orbital parameters, see (Švehla and Rothacher 2004a). The elevation-dependent PCV was estimated in bins of 5° and 10° and both series of parameters show very close agreement (about ± 0.15 mm). The estimation of azimuth-dependent phase-center variations revealed the high correlation between the phase center variations in the along-track direction, the along-track orbit component and the pseudo-stochastic pulses. The azimuth-dependent pattern can be estimated using the higher accuracy in the cross-track direction. Elevation-dependent weighting is not used for the CHAMP carrier-phase, although in ground GPS applications it is mostly applied. Recently, the GPS antenna on board CHAMP and GRACE has been absolutely calibrated by the robot calibration system developed by the Institut für Erdmessung (IfE) and Geo++ in Hannover, see (Montenbruck et al. 2009). Comparing Figure 13.8 and Figure 13.9 one can see a close overall agreement between the CHAMP elevation-dependent antenna patterns estimated in-flight using GPS carrier-phase and the ground calibration using the robot. The opposite sign is

due to the different sign conventions used, namely the sign of the PCVs provided by IGS is opposite to that of the IfE. In both cases, we see an effect from -10 mm to 10 mm with a maximum value at $50°$ zenith angle. Deviations can only be seen in the zenith direction and are most likely due to the slightly different mean phase-center offset.
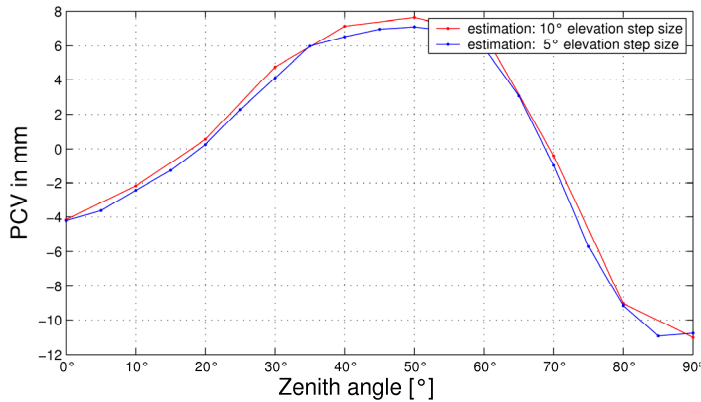
Figure 13.8 CHAMP absolute phase-center pattern based on one week of measurements, days 195-201/2002, (Švehla and Rothacher 2004a). The opposite sign w.r.t. robot calibration is due to the different sign conventions used, namely the sign of the PCVs provided by IGS is opposite to that of the robot calibration from IfE.
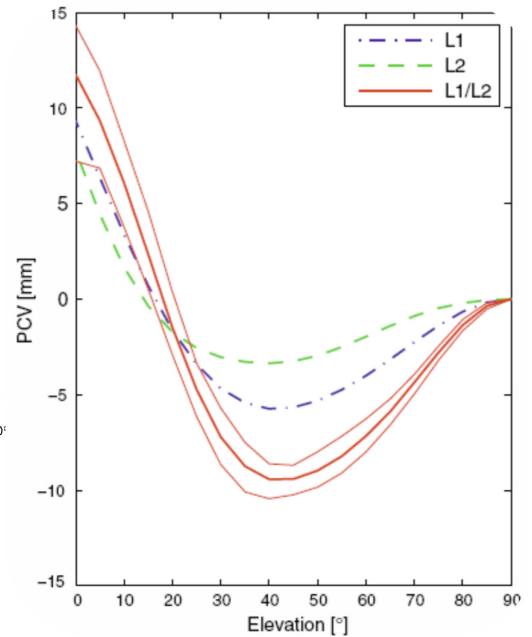
Figure 13.9 CHAMP elevation-dependent PCV after absolute antenna calibration using a robot, (Montenbruck et al. 2009), for L_1 and L_2 and for iono-free linear combination "L1/L2". Besides the mean value (bold) upper/lower bounds with 1σ (thin lines) over all azimuth angles are shown.

13.6 The Ray-Tracing Technique for Multipath Maps of GNSS and LEO Satellites

Antenna phase-center maps for GPS satellites were estimated using a least-squares adjustment and GPS measurements from the ground IGS network, applying absolute antenna PCV maps for the ground GPS antennae from robot calibrations, see (Schmid and Rothacher 2003). GPS satellite antenna PCV parameters were estimated together with all other reference frame parameters, including GPS satellite orbits, troposphere parameters and station coordinates. As a result, the estimated GPS satellite antenna PCVs contain residual effects due to high correlations with other reference frame parameters and poor geometry, given the high altitude of GPS satellites and the relatively small GPS satellite antenna aperture angle of about $28°$. Those maps were estimated and then averaged over a long period of time and do not necessarily represent the GPS satellite PCV affected by near-field multipath stemming from the rotation of large GPS solar panels. Therefore it would be interesting to calibrate and derive near-field multipath maps using an alternative technique (e.g., as a function of solar beta angle and argument of latitude of the satellite relative to the Sun's position in the orbital frame). Although calibration on a robot or in an anechoic chamber provides similar results for the receiving ground antennae, using such a technique for calibration of GNSS transmitting antennae together with a rotating solar panel would be a more demanding and challenging task. Perhaps this could still be done for new GNSS satellites.

An alternative method is the ray-tracing technique based on the antenna electro-magnetic characterization and coupling between the antenna environment and the phase-center variations. Figure 13.10 shows the GOCE antenna PCV profile (along the satellite axis) for the L_2 frequency. This PCV profile was estimated using a model of the complete GOCE satellite. The green line shows the nominal L_2 PCV, whereas the blue line is the PCV profile considering the GOCE mock-up, the same as that used for the robot calibration shown in Figure 13.2. The red line shows the near-field effects stemming from the complete satellite structure. The blue line shows a good overall agreement with Figure 13.2, where the entire PCV map is depicted. One can clearly recognize an effect of about -3 mm at 45° zenith angle and 10 mm at 90° (horizon).
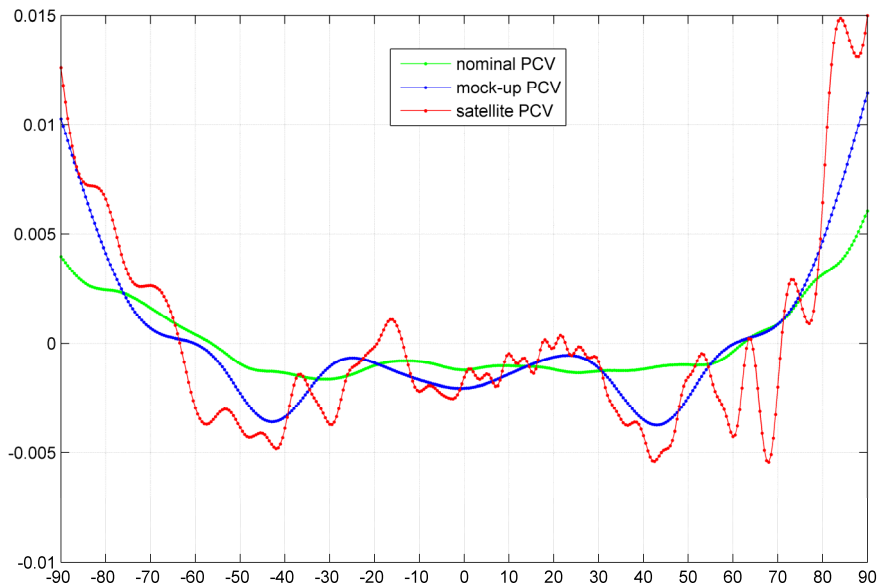
Figure 13.10 GOCE L_2 PCV profile (along the satellite axis) in [mm] vs. zenith angle in [°] estimated using the ray-tracing technique. The blue line is the PCV estimated using a mock-up and shows very close agreement with Figure 13.2, namely an effect of -3 mm at 45° and 10 mm at 90° (data: ESA GOCE Project Office).

It is interesting to note that the complete model of the GOCE satellite gives a very variable signal over all elevations. A high-resolution GOCE PCV map is difficult to derive using a POD technique that is independent of the gravity field, since a very small step size is needed to pick up all PCV variations, especially close to the antenna horizon and in the direction of the solar wing. This very good overall agreement between robot calibration and the ray-tracing technique opens up the possibility of calibrating GPS antennae on any LEO satellite, or even calibrating near-field multipath on GNSS satellites.

In the case of the GNSS or the six COSMIC satellites, antenna phase center maps could be provided for several characteristic orientations of the satellite solar panels. In this way, temporal PCV maps could be established for GNSS and LEO satellites to more accurately model the near-field multipath (including multipath, imaging and scattering effects) generated by the on-board near-field antenna environment and large solar panels. The same approach could be used for the ground or spaceborne LEO GNSS antennae and transmitting antennae on board GNSS satellites.

13.7 Multipath Linear Combination

Here we derive a multipath linear combination that can be used to assess and monitor multipath in GPS measurements, as was performed for the GOCE mission, but which can also be easily extended to any other GNSS observable. Similar linear combination, without derivation, can be found in (Estey and Meertens 1999).

Let us look again at the ionosphere-free linear combination L_3 of the carrier-phase measurements L_1 and L_2 converted from cycles to meters

$$L_3 = \frac{1}{f_1^2 - f_2^2}\left(f_1^2 L_1 - f_2^2 L_2\right) = \rho + \lambda_N N_1 + \frac{1}{2}(\lambda_W - \lambda_N)N_W = \rho + B_3 \tag{13.2}$$

with the geometry term ρ and the ionosphere-free bias denoted as B_3

$$B_3 = \lambda_N N_1 + \frac{1}{2}(\lambda_W - \lambda_N)N_W \tag{13.3}$$

with the narrow-lane wavelength λ_N, the wide-lane wavelength λ_W and the corresponding wide-lane ambiguity N_W. The ionosphere effects can also be removed by forming the so-called LP linear combination by adding carrier-phase and code measurements on the same GPS frequency

$$LP = \frac{L_1 + P_1}{2} \tag{13.4}$$

Bearing in mind that multipath and the noise of the carrier-phase measurements can be neglected compared to multipath M_{P1} and noise of the code measurements ε_{P1}, the following relation between the LP linear combination and the ionosphe-free linear combination L_3 can be written

$$\begin{aligned} 2LP = L_1 + P_1 &= 2\rho + \lambda_1 N_1 + M_{P1} + \varepsilon_{P1} \\ = L_1 + P_1 &= 2(L_3 - B_3) + \lambda_1 N_1 + M_{P1} + \varepsilon_{P1} \\ &= 2L_3 + B_{P1} + M_{P1} + \varepsilon_{P1} \end{aligned} \tag{13.5}$$

where λ_1 is the wavelength of the L_1 carrier-frequency with the corresponding integer ambiguity N_1. M_{P_1} denotes the multipath and ε_{P_1} the noise of P_1 code measurements. B_{P1} can be considered as a bias in the LP linear combination

$$B_{P1} = -2B_3 + \lambda_1 N_1 = -2(\lambda_N - \frac{\lambda_1}{2})N_1 - (\lambda_W - \lambda_N)N_W \tag{13.6}$$

The following expressions can be written and used for the evaluation of multipath on C/A, P_1 and P_2 code measurements, respectively:

$$M_{CA} + \varepsilon_{CA} = L_1 + C/A - \frac{2}{f_1^2 - f_2^2}\left(f_1^2 L_1 - f_2^2 L_2\right) - B_{P1}$$

$$M_{P1} + \varepsilon_{P1} = L_1 + P_1 - \frac{2}{f_1^2 - f_2^2}\left(f_1^2 L_1 - f_2^2 L_2\right) - B_{P1} \tag{13.7}$$

$$M_{P2} + \varepsilon_{P2} = L_2 + P_2 - \frac{2}{f_1^2 - f_2^2}\left(f_1^2 L_1 - f_2^2 L_2\right) - B_{P2}$$

The float ambiguity should be constant throughout each pass and can be removed by calculating the mean. The B_{P2} denotes a bias in the LP linear combination for the second GPS frequency

$$B_{P2} = -2B_3 + \lambda_2 N_2 = -2(\lambda_N - \frac{\lambda_2}{2})N_1 - (\lambda_W - \lambda_N + \lambda_2)N_W \tag{13.8}$$

14. Probing the Flyby Anomaly Using Kinematic POD – Exotic Applications of Kinematic POD

T he idea presented here is to use the GPS receiver for the comparison of kinematic and dynamic orbits of an interplanetary mission during Earth flyby, e.g., BepiColombo, Juno. Purely geometrical orbits can be estimated to an accuracy of $1\,\text{cm}$ RMS using GPS carrier-phase measurements, whereas dynamic orbits will be affected by any potential flyby anomaly effect on the spacecraft while it is in Earth flyby.

The flyby anomaly is an unexpected increase in the spacecraft velocity or orbital energy during Earth flyby. This anomaly has been observed in Doppler measurements by a number of ground ESA/NASA stations operating in S- and X-band, for more details see e.g., (Anderson et al. 2008). The orbit velocity increase is in the order of $7 - 13\,\text{mm/s}$ and it has not been reported for all swingbys (Morley, priv. com.).

The minimum altitude for a flyby is in the order of $500 - 2500\,\text{km}$, which means that an Earth flyby could be observed using a GPS receiver over several hours (up to altitudes of e.g., $10\,000\,\text{km}$). In the case of kinematic POD, the velocity of the satellite can be estimated geometrically to an accuracy well below $0.05\,\text{mm/s}$ (Švehla and Földváry 2006). Using a GNSS receiver on a future interplanetary mission during Earth flyby we will be able to monitor the flyby anomaly geometrically and compare the results with dynamic orbits. An additional SLR retro-reflector would enable ground laser stations to monitor the flyby anomaly.

In 2009 we proposed mounting a GPS receiver on the BepiColombo spacecraft. However, after approaching the BepiColombo Project Office it was deemed to be too late to include one in the payload. Nevertheless, the concept is well worth further test, since kinematic POD can assess the flyby anomaly effect far more accurately than Doppler S-band/X-band tracking from the ground. Figure 14.1 shows the predicted ground track of the Juno satellite in interplanetary orbit at Earth swingby on 9.10.2013.
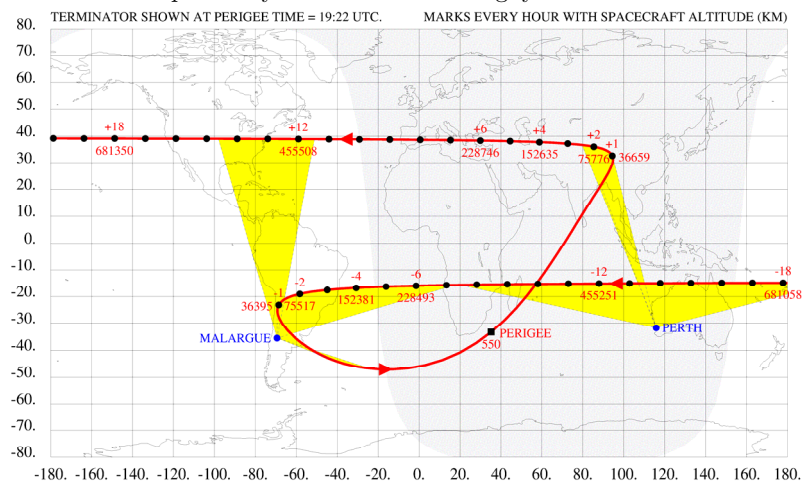
Figure 14.1 Predicted ground track of the Juno satellite at Earth swingby on 9.10.2013. Every hour is marked in red with an orbit altitude in [km]. The yellow areas show tracking visibility from the two ground stations.

15. Galileo-2: A Highly Accurate Dynamical GEO Reference Frame to Complement the TRF

In (Švehla 2007a), (Švehla et al. 2008) and in (Švehla 2008a) we presented a novel design for the GNSS system called here Galielo-2 based on recent developments in optical clocks, frequency combs and time/frequency comparison technology. We demonstrated a concept of a navigation system in MEO based on master clocks in the GEO orbit and two-way optical/microwave links to transfer their stable frequency to the navigation satellites in MEO orbit (either from the ground or via GEO). In this way, the use of H-masers and Cs- or Rb-clocks in the GNSS satellites can be avoided and frequency combs could be used to generate the desired navigation radio (and optical) signal in the MEO orbit. The development of "Ultra"-USO, e.g., for the STE-QUEST mission with a frequency stability in the order of 10^{-15} at $1\,\mathrm{s}$ is sufficient to meet the required GNSS clock stability over a longer period of time (e.g., one day), and thus one could separate precise orbit determination of GNSS satellites from estimation of GNSS clock parameters. GNSS clock frequency can be steered either from the ground or from the GEO orbit making use of the two-way metrology links. For this, master clocks in GEO do not need to be of the highest accuracy, they could be optical clocks or the latest Rb-clocks with high short-term stability. However, the assembly of several GEO clocks equipped with optical/microwave links for frequency transfer will meet the needs of the timing community for clock comparison in the generation of the global TAI/UTC time scale. Thus, the idea of Galileo-2 is twofold: Firstly, to combine positioning and timing systems under one umbrella, and, secondly, to enable new applications in geosciences.

Generally speaking, a highly accurate dynamic reference frame in the GEO orbit would, in future, have the potential in terms of accuracy to provide an alternative to, and to complement, the terrestrial reference frame of the Earth. Drag-free and ranging technology as developed for the LISA mission provide very strong arguments in this direction. A GEO reference frame could provide the basis for a real-time positioning/timing facility for all GNSS Earth-based applications, from LEO to GEO orbit and beyond towards lunar orbits. Intersatellite ranging between such (drag-free) GEO satellites could be obtained to a very high level of accuracy, e.g., sub-micrometer – several orders of magnitude higher than the accuracy of a terrestrial reference frame. Considering the orbit-redshift equivalence principle we introduce in Section 29 (a symmetry between the error in orbit position and velocity such that these cancel or compensate each other out in generating the net redshift effect), an orbit in space (GEO) offers the best environment to define and realize the frequency standard and define the SI second using an atomic clock. A far more reliable method than using the geoid and the surface of the Earth. This is mainly due to the fact that cold atoms in the clock can be observed for a long time in space (weightlessness) and are not limited by free-fall as they are on Earth. This typically gains an additional 3-4 orders of magnitude in sensitivity. Therefore, in future, GEO orbit could offer the best place to define the datum for time and so support positioning on Earth. The terrestrial reference frame of the Earth is, by definition, tied to the ground network of station coordinates on the Earth's crust. Thus the proposed realization using GEO orbit is an extended and complementary realization of the terrestrial frame which aims to achieve higher accuracy and precision and to obtain synergy with time realization.

15.1 Galileo and Beidou – Paving the Way Towards the new GNSS Science?

Can we design a navigation system that can meet the requirements of both navigation and geosciences at the same time? Is there a place for improving GPS, for something better than introducing a third navigation frequency? Can we enhance GNSS so that it can be used for novel applications in geosciences? What new developments and potential will Galileo and Beidou bring? Can we use satellites in the GEO orbit for real-time positioning of GNSS satellites, just as GNSS is used for to the POD of satellites in the LEO orbit?

If we look back some 30 years, GPS was primarily designed to meet US military requirements for navigation. However, over the years, GPS has become much more than just a navigation system. It is now a driving force in geodesy, with applications ranging from precise geodetic positioning, geodynamics and timing, to meteorology and remote sensing techniques, such as radio-occultation and reflectometry/scatterometry. Today GNSS receivers are readily available and GNSS has become an indispensable part of the infrastructure in every aspect of human activity.

Block I	11	1 CS	2 Rb
Block II	9	2 CS	2 Rb
Block IIa	19	2 CS	2 Rb
Block IIR	13		3 Rb
Block IIR-M	6		3 Rb
Total :	58	67 CS	135 Rb

Figure 15.1 Number of cesium and rubidium atomic clocks on board 58 GPS satellites since the launch of the first GPS satellite on 22 February 1978.

Every GNSS satellite carries several atomic clocks, and as a result GNSS satellites are very expensive, bulky, heavy and the entire constellation needs to be maintained by launching new satellites every 5-10 years. Atomic clocks on board GNSS satellites have demanding requirements in terms of power and payload and hence all GNSS satellites are equipped with large solar panels and with three-axis attitude stabilization. The atomic clocks placed on board such satellites are probably the most crucial single element in achieving a high-performance GNSS (Hein et al. 2007). A more accurate and stable frequency inside the GNSS satellite means a reduction of uncertainty in clock prediction and hence improved real-time positioning and greater integrity of information.

Figure 15.1 shows the number of atomic clocks put into space since the launch of the first GPS satellite in early 1978. One can see that of the 58 GPS satellites launched to date, only 30 are still active and over the last 30 years altogether 202 atomic clocks have been launched into space. Altogether 170-180 atomic clocks are orbiting the planet Earth on board the decommissioned GNSS satellites without sending any signal towards the Earth. Although the lifetime of GPS satellites is higher compared to those of GLONASS, satellites of both GNSS systems have to be decommissioned in orbit and their lifetime is limited by the lifetime of the on-board batteries and the radiation environment in MEO orbit. If we now take 30 years of GPS and 200 clocks per single GNSS system and multiply this by 4 GNSS systems in the near future (GPS, GLONASS, Galileo and Beidou, not including Indian and Japanese regional navigation systems IRNSS and QZSS), we end up with about 1000 atomic clocks in space in just 20 years from now. Is there an alternative? It is generally considered that clocks in timing labs will always be more accurate than even the most modern optical clocks developed for space. However, this is not completely true since it is well known that weightlessness in space offers orders of magnitude improvements in clock performance compared to ground clocks. Therefore, it would seem logical to develop a number of very high-quality clocks and put them into GEO orbit. Their frequency could always be controlled by the ultra-accurate optical ground clocks. Their main purpose would be to compare optical ground clocks defining TAI/UTC and to distribute this frequency standard to GNSS satellites in MEO orbit.

In order to ensure integrity, 3-5 master clocks would be required in GEO orbit. As in the case of GNSS satellites, additional onboard clocks would meet all redundancy and integrity requirements. Frequency dissemination in space between GEO and MEO is easier given the atmospheric conditions close to the Earth's surface and can either be performed optically or in the microwave domain.

Despite high quality onboard clocks, GNSS are not designed for time/frequency transfer. Unlike ground Cs-fountains in TAI labs that provide frequency with a stability of below one part in 10^{-16} or optical clocks with a stability of less than one part in $10^{-17}/3\,\mathrm{h}$, GNSS systems cannot meet the demanding requirements of time/frequency transfer for TAI/UTC. One can draw the general conclusion that positioning and timing are two separate worlds and both communities are using their own global timing and positioning systems. The main problem is that positioning is based on one-way systems and time/frequency transfer requires a two-way system. Why not combine them and benefit mutually? In the case of a two-way system, such as TWSTFT, a signal is sent in both directions and, by differencing, the first order Doppler effect and all geometry and propagation delays are removed. There are still residual higher-order ionosphere terms present in the two-way microwave measurements, but due to the very high frequency used they are very small (they sum up for the uplink and downlink). In the case of optical two-way measurements, all propagation effects are eliminated, and atmospheric turbulence is the main source of error.

Currently, there is no operational system available which can compare on a global scale the best ground optical clocks that have already demonstrated an accuracy of two parts in 10^{-18}, (Nicholson et al. 2015). In the very near future, there will be a gap in performance between the TAI clocks and the satellite-based time/frequency comparison systems. In fact, the best ground optical clocks have reached such a level of accuracy that it is already now feasible to measure dynamic heights (geopotential numbers) using terrestrial clocks, but there is no satellite system available to compare ground clocks with sufficient accuracy. GNSS receivers measure geometric heights above the ellipsoid, whereas physical height systems use the equipotential surface of the reference geopotential, called geoid, as a datum. Therefore, a two-way link on a GNSS satellite would allow the unification of the timing and positioning systems, and hence the unification of geometrical and gravitational positioning (gravity potential).

Compared to carrier-phase and pseudo-range measurement of the present and the forthcoming one-way GNSS constellations, a two-way system would provide geometry-free transfer of clock frequency. This would allow geometry-free steering of the GNSS satellite clock frequency. In the case of Galileo H-maser, we already see that the satellite clock can be modelled with just two linear parameters per day (time drift and bias) providing a standard deviation of remaining residual clock parameters at the cm-level. By introducing frequency steering of the Galileo satellite clock, one could predict satellite clock over a longer time period and thus separate orbit from determination of clock parameters. In the case of pseudo-ranges, or, generally speaking, observables of all traditional one-way GNSS systems, receiver and satellite clock parameters need to be estimated or removed every epoch. That clock parameters cannot be separated from the propagation effects in the orbit determination or parameter estimation is the main disadvantage of the one-way GNSS systems. Even with the three or four Galileo frequencies we cannot estimate absolute TEC and calibrate all biases at the mm-level in order to "measure carrier-phase ambiguities". In the processing of one-way GNSS data, we fix something that we call "phase ambiguity" by estimating a very large number of other global parameters such as station coordinates, tropospheric zenith delays and gradients, Earth rotation and satellite orbit parameters. This is the case with zero- and double-difference carrier-phase measurements. Even with the three-carrier ambiguity resolution strategies, these problems still remain to a great extent, and propagation effects need to be separated from the integer phase ambiguities. There is always a trade-off between the ambiguity space and the parameter space (all global GNSS parameters including ambiguities). In zero-difference GNSS applications, carrier-phase ambiguities are additional "nuisance" parameters that need to be estimated and they constrain the capacity to reduce the influence of systematic errors of GPS orbit/clock products, and tropospheric, multipath and other effects. The absolute ionospheric effects and tropospheric delays cannot be separated entirely from GNSS satellite/receiver clock parameters. The bottom line is that modern and future GNSS systems must be a combination of one-way and two-way systems.

In the orbit determination for GNSS satellites we estimate typically 9 empirical solar radiation pressure (SRP) parameters per daily orbit arc and the SRP effects propagate into geocenter results, EOPs, and the orbits of altimetry and gravity field missions that require the highest accuracy. On the other hand, GNSS satellites are placed at a very high altitude above the Earth with a very small antenna aperture angle of about 14° half angle, and due to this "bad geometry", orbit errors such as residual solar radiation pressure propagate into all global GNSS parameters (EOPs, geocenter, etc.).

Therefore, one could generate a dynamic reference frame in the GEO orbit consisting of several GEO satellites, similar to the three drag-free satellites of the LISA mission in a triangular constellation. In this way, one could extend and complement the classical definition of the terrestrial reference frame based on a network of ground stations and thus tied to the Earth's crust. Intersatellite ranging between those GEO satellites could lead to a very high level of accuracy, several orders of magnitude higher than that of a terrestrial reference frame. However, since the conventional terrestrial frame is by definition tied to the ground stations on the Earth's crust, one would still need a tie with such a dynamic system in space. Thus, one could talk about complementarity between the space-based and the ground-based reference frames, where space-based frames provide higher accuracy and stability.

One could make use of frequency combs as a metrology system between GEO satellites as well as a generator of microwave/optical frequencies for the navigation signals. Frequency combs were proposed for the various ESA missions that use formation flying and high-accuracy long-distance metrology (Holzwarth et al. 2008). In the latter, femtosecond-based laser systems are combined with incoherent time-of-flight absolute distance measurement capabilities over long distances using coherent high-resolution interferometric methods. Such optical systems provide sub-micrometer resolution in an absolute measurement of nearly arbitrary distances (Holzwarth et al. 2008). On the other hand, the GEO orbit is high enough above the Earth and the microwave downlink transmitter of the two-way system can be tracked by VLBI antennas in S- and X-band and up to Ka-band. This is the same frequency band used by VLBI to observe quasars (VLBI2010). Thus the use of VLBI in combination with a satellite in a higher orbit (e.g., GEO), would open up new possibilities in combining the terrestrial reference frame and the, at the moment, fully independent VLBI celestial reference frame based on extragalactic radio sources (quasars).

15.2 Geometrical Properties of Positioning with Four GNSS – Homogeneous and Isotropic Positioning with Galileo

What improvements will Galileo and Beidou bring to global positioning? We carried out a simulation of IGS-type processing with four GNSS. For that purpose the Bernese GPS software was adapted for GNSS, i.e., Galileo/Beidou data processing, within the scope of a project with Astrium and GFZ. The first results in processing GIOVE-A data with this new version of Bernese multi-GNSS software were presented in (Švehla and Heinze 2007). The simulation covered a period of one day and included 31 satellites of the GPS constellation (day 62/2007), 24 GLONASS satellites (8 satellites were added to simulate the complete GLONASS constellation), 30 Galileo satellites and 30 Beidou satellites. This gives 115 GNSS satellites in total, the number one can expect to be in Earth orbit in the near future, see Figure 15.2 and Figure 15.3. For the IGS network we considered a grid of $15° \times 15°$ which covers about 200 ground stations. Carrier-phase measurements were simulated with a white noise of $3\,\text{mm}$.
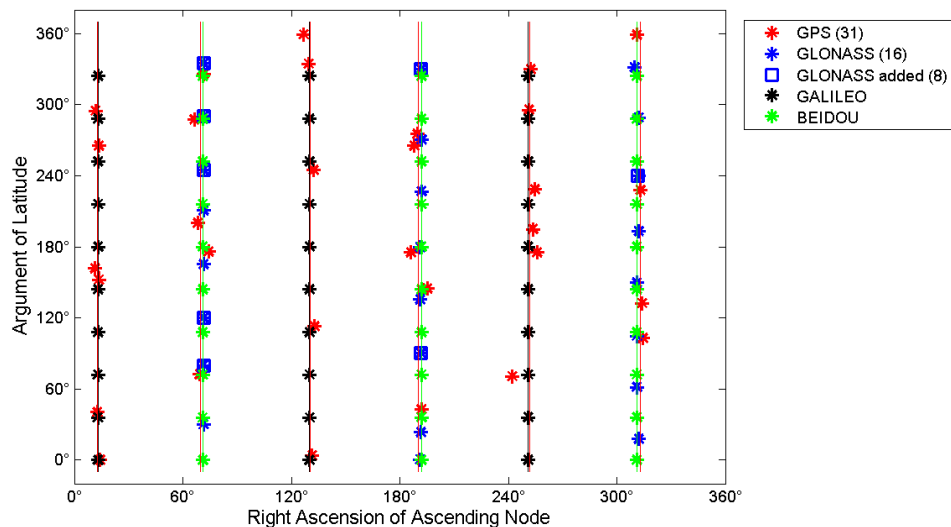
Figure 15.2 GNSS satellites and orbital planes used in the simulation (day 62/2007).

For Galileo and Beidou we considered the L_1 and $E5a$ frequencies, which leads to an increase in the noise level of the ionosphere-free linear combination in the order of ≈ 2.588 (w.r.t. the noise of L_1), compared to ≈ 2.978 in the case of GPS and GLONASS. For the ground stations, the so-called elevation-dependent weighting was used to model noise as a function of elevation. For GPS and GLONASS we used IGS Final Orbits for day 62/2007, see Figure 15.2, whereas Galileo and Beidou were simulated in a Walker constellation based on the constellation parameters available in 2007.

Figure 15.4 shows the number of visible GNSS satellites using an antenna cut-off angle of $10°$. One can see that, whereas there are about 9 GPS satellites in the field of view at present, in future we can expect about 16 GNSS satellites by considering the complete GLONASS constellation in addition to GPS, and 26 GNSS satellites with the addition of both the GLONASS and Galileo constellations. The additional Beidou satellites increase the number of visible GNSS satellites to 35. From Figure 15.4 one can clearly see that the number of visible GNSS satellites varies strongly with geographical latitude, with the highest number of GNSS satellites visible in polar regions and along the equator.
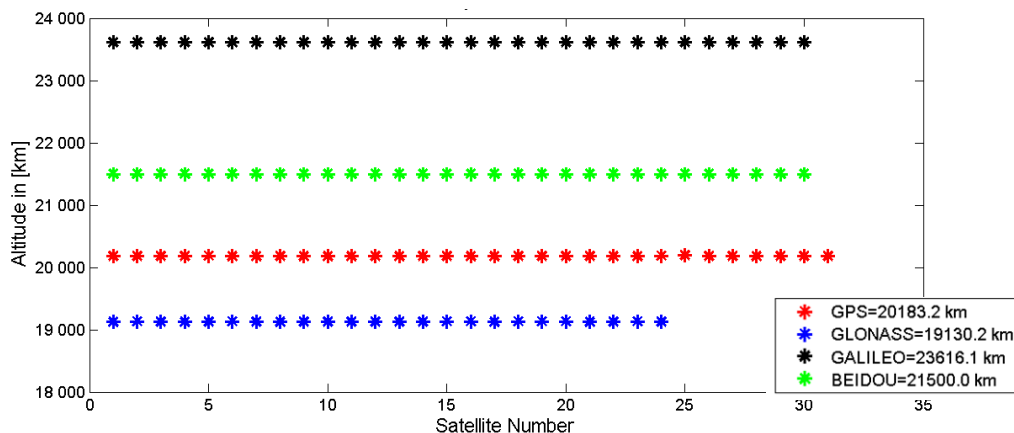
Figure 15.3 Orbit altitude of GNSS satellites used in the simulation.
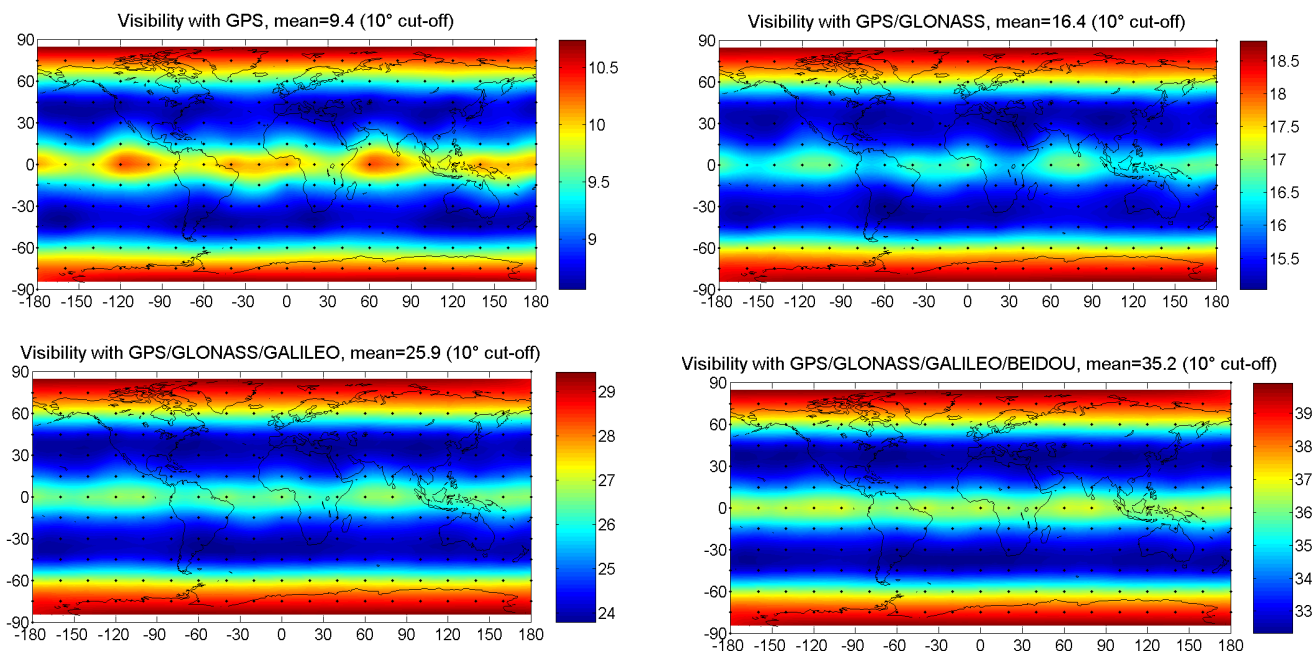
Figure 15.4 Mean number of visible GNSS satellites over one day with all four GNSS (10° cut-off angle).

Figure 15.5 shows the error ellipses of the horizontal position based on one-day PPP solutions with all four GNSS (10° cut-off angle). We calculated the Helmert error ellipse (central ellipse) M

$$A^2 + B^2 = M^2 = 2R^2 \tag{15.1}$$

with the semi-major axis A and the semi-minor axis B. The Helmert ellipse is often called the central ellipse because it is the smallest in size and can be described by a circle of radius R with a probability of 0.39. For the sake of completeness we would like to mention that an ellipse with a probability of 0.63 in the interval $\pm 1\sigma$ has semi-major axes $A\sqrt{2}$ and $B\sqrt{2}$.

Based on the central Helmert ellipse of radius R, one can calculate an improvement factor by adding each individual GNSS constellation separately. One can see an improvement in the central ellipse by a factor of 1.51 when adding GLONASS and by 2.22 when adding in addition Galileo and Beidou. More homogeneous and isotropic positioning with Galileo can clearly be seen from Figure 15.5, since by increasing the number of GNSS satellites, error ellipses become smaller and more circular, i.e., the dominant east-west orientation is reduced.

Isotropic positioning in our definition here means that the error ellipses are circular in shape, i.e., the accuracy of the estimated horizontal station coordinates is the same at all azimuth angles. Homogeneous positioning in our definition here refers to a mean accuracy of station coordinates that is similar or equal over all geographical longitudes and latitudes, i.e., irrespective of the location of the station. It is interesting to note the east-west orientation of error ellipses also in the polar regions. Figure 15.6 shows the formal errors of the station height based on PPP over one day with four GNSS. One can clearly see that the highest accuracy can be expected in mid-latitudes. Around the polar regions, despite the highest number of visible GNSS satellites, the accuracy of the estimated station heights is lowest, simply due to the satellite geometry and the low elevations of GNSS satellites tracked.
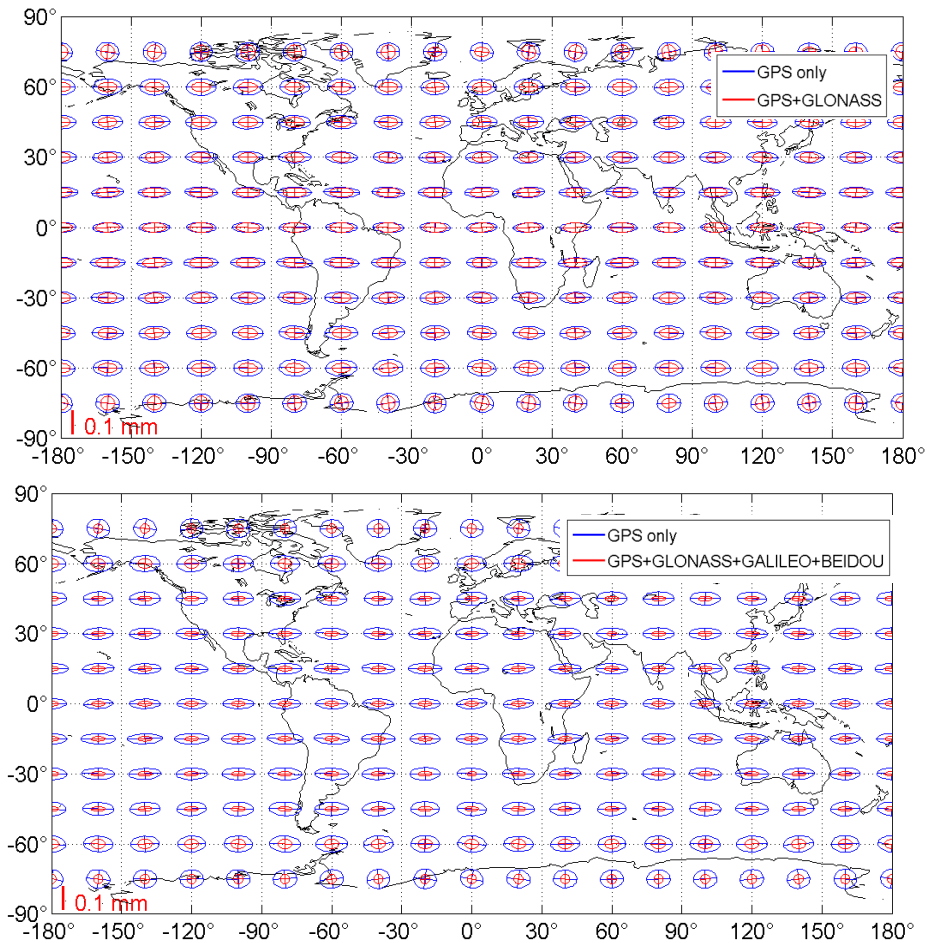
Figure 15.5 Error ellipses of the horizontal position based on one-day PPP solutions with all four GNSS (10° cut-off angle). More homogeneous and isotropic positioning with Galileo. Improvement in the central Helmert error ellipse by a factor of 1.51 when adding GLONASS and 2.22 when adding in addition Galileo and Beidou.
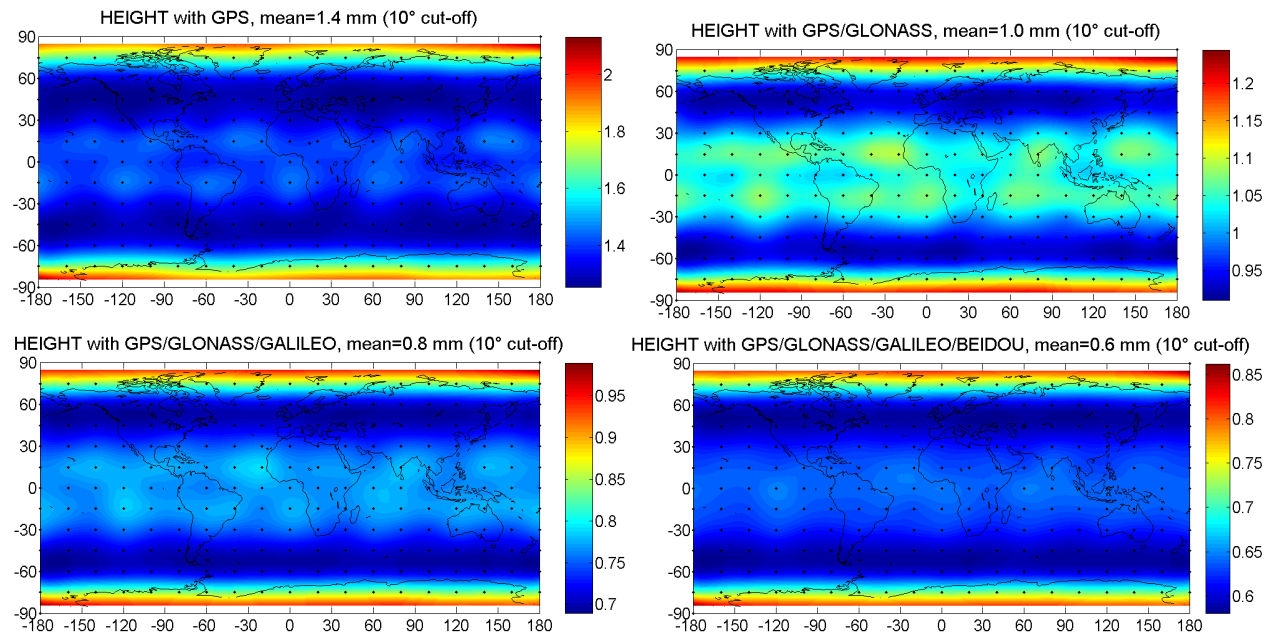
Figure 15.6 Formal error of station heights based on one day PPP with all four GNSS (10° cut-off antenna angle). Improvements by a factor of 1.8 are obtained by adding Galileo and 2.4 with all 4 GNSS.
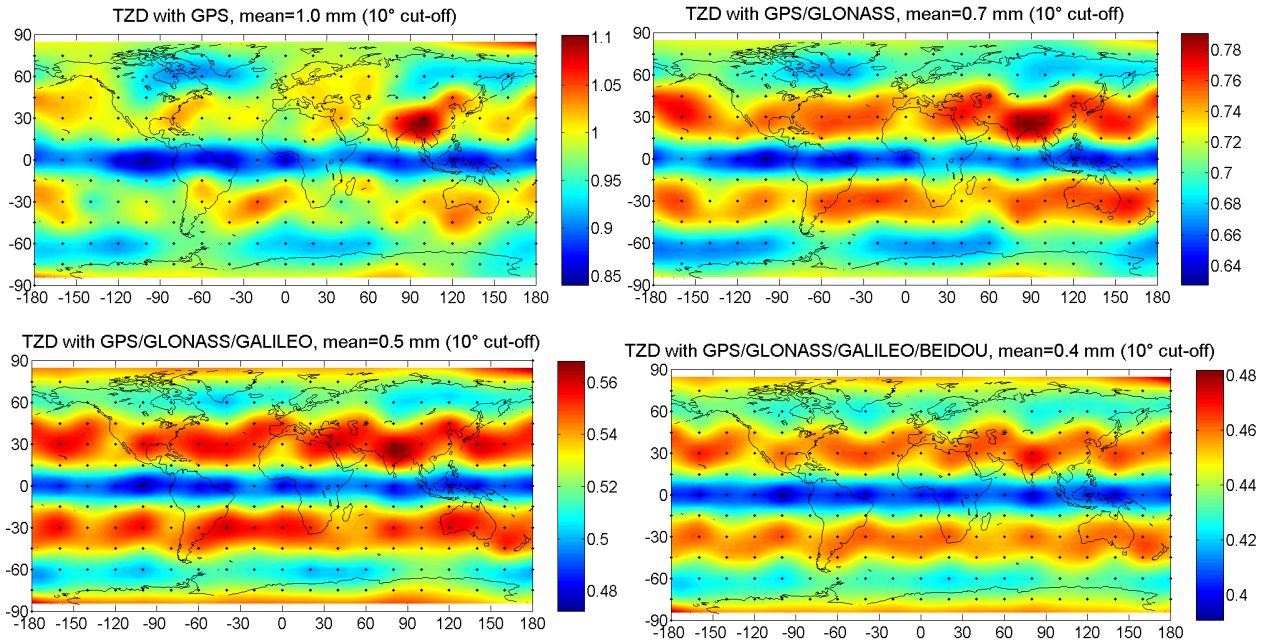
Figure 15.7 Formal errors of troposphere zenith delays with all four GNSS (estimated every two hours, 10° cut-off angle). Periodic patterns are visible at mid-latitudes, most likely due to the six orbital planes used.

By adding Galileo to complete the GPS and GLONASS constellations one can reduce the formal errors by a factor of 1.8 of the estimated station heights and by a factor of 2.4 by using all 4 GNSS. Figure 15.7 shows the formal errors of tropospheric zenith delays (TZDs) estimated every two hours using all four GNSS. The improvement compared to the GPS-only scenario due to the addition of the GLONASS and Galileo constellations amounts to a factor of 2 in terms of formal errors of the estimated tropospheric zenith delays.

Comparing tropospheric zenith delays in Figure 15.7 with station heights in Figure 15.6, one can clearly see that station heights can be estimated most accurately in the mid-latitudes, whereas tropospheric zenith delays are estimated most accurately around the equator. This must be due to correlations between station heights, tropospheric zenith delays and station clocks as well as observation geometry. A closer look at Figure 15.7, reveals very strong periodic patterns at mid-latitudes, most likely related to the six orbital planes.

15.3 Can we Improve GPS Satellite Orbits With Galileo?

Can we improve the orbit determination of GPS satellites with Galileo? The answer is ″yes″. Galileo measurements contribute to common parameters estimated together with GPS measurements, such as tropospheric zenith delays and station coordinates, EOPs and GNSS receiver clock parameters. We have extended our GPS simulation by adding the Galileo constellation and Figure 15.8 shows the typical improvements in the orbit of one GPS satellite against orbit estimation based only on the GPS constellation. The effect is in the order of $1-2$ cm . The estimation was based on zero-difference carrier-phase measurements and the orbit parameterization is exactly the same as that used at the CODE IGS AC for the one-day orbit arc. The only difference is that the combined GPS/Galileo processing was based on zero-difference measurements without ambiguity resolution.
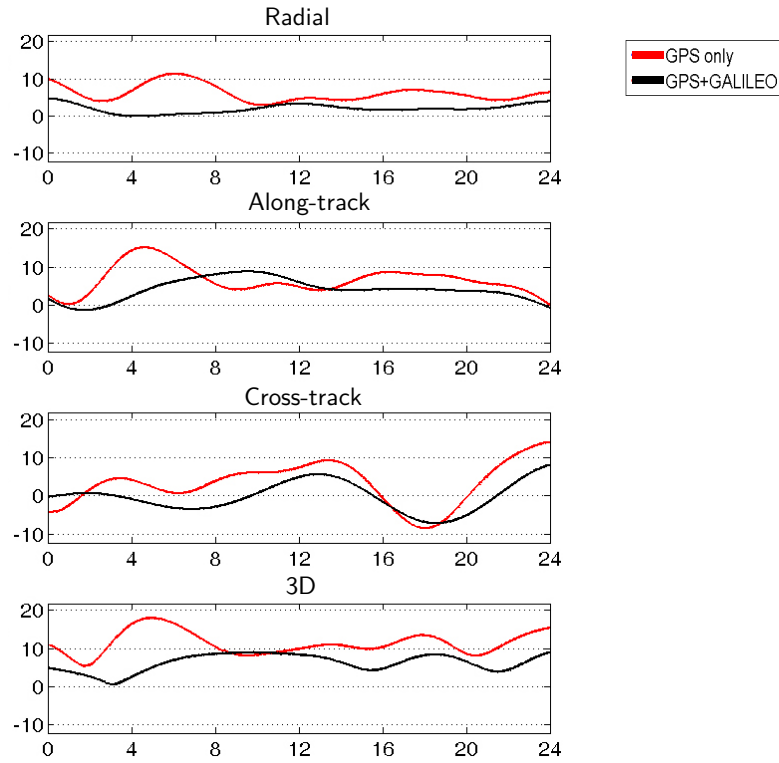
Figure 15.8 Typical differences in GPS satellite orbits by adding measurements from 30 Galileo satellites. For the simulation of the combined processing of the GPS and Galileo constellations, see (Švehla and Heinze 2007).

15.4 Orbit Determination of GNSS Satellites From GEO

How accurately can one estimate the orbit of a GEO satellite? Can we generate a highly accurate space-based reference frame in GEO orbit and combine such a geometric/dynamic frame with the conventional terrestrial and celestial frame? With just three to five satellites in GEO orbit one could cover, in terms of visibility, the entire Earth, (see Figure 15.9) and continuously measure range or range-rate between the reference GEO
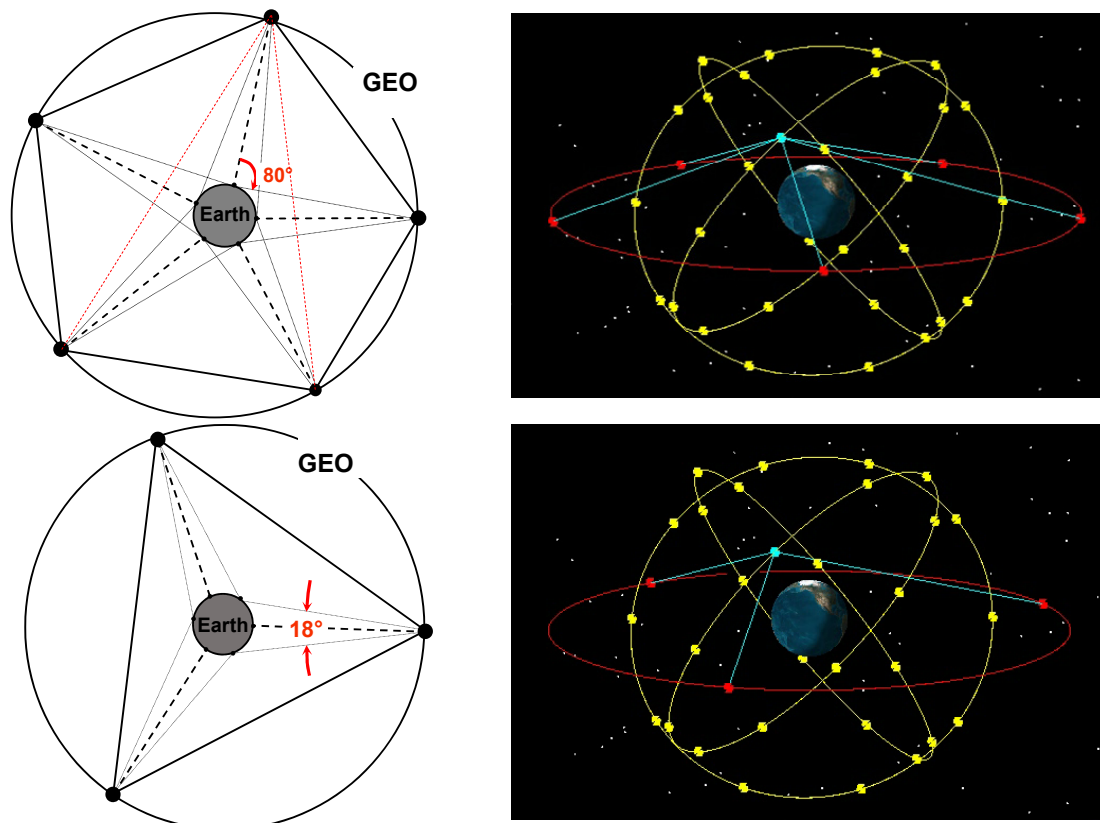
Figure 15.9 Intermediate GEO Reference Frame based on 3-5 reference satellites connected with intersatellite metrology links and closely tied to the ground terrestrial frame and GNSS satellites.

satellites with very high, e.g., sub-micrometer accuracy. Any additional reference "station" in GEO orbit will be in the field of view of all other reference "stations" in the GEO orbit, as shown in Figure 15.9. No matter how accurately one could determine the GEO orbit from the ground, any reduced-dynamic or dynamic POD approach will constrain the relative GEO orbit information to highly accurate range (or range-rate) measurements. This is also demonstrated with simulated data later in the text in more detail, see Figure 15.10. The relative orbit information between reference "stations" in the GEO orbit will be several orders of magnitude more accurate than the relative information between the satellites and the ground geodetic stations of space geodesy techniques such as GPS, VLBI, SLR and DORIS. This can be argued based on the intersatellite link between two GRACE satellites that uses K-band measurements with a noise level below $10\ \mu m$. The relative radial and along-track orbit information will be of the highest accuracy, whereas out-of-plane orbit information will strongly depend on the accuracy of the ground-to-space link and the ability to orient this orbital frame in space. The gravity field of the Earth, like e.g., J_2 coefficient, will provide additional constraints to the accuracy of the cross-track orbit direction. Additional ranging from GEO to any satellite in MEO or LEO orbit will provide a space-based reference frame of utmost accuracy in all directions.
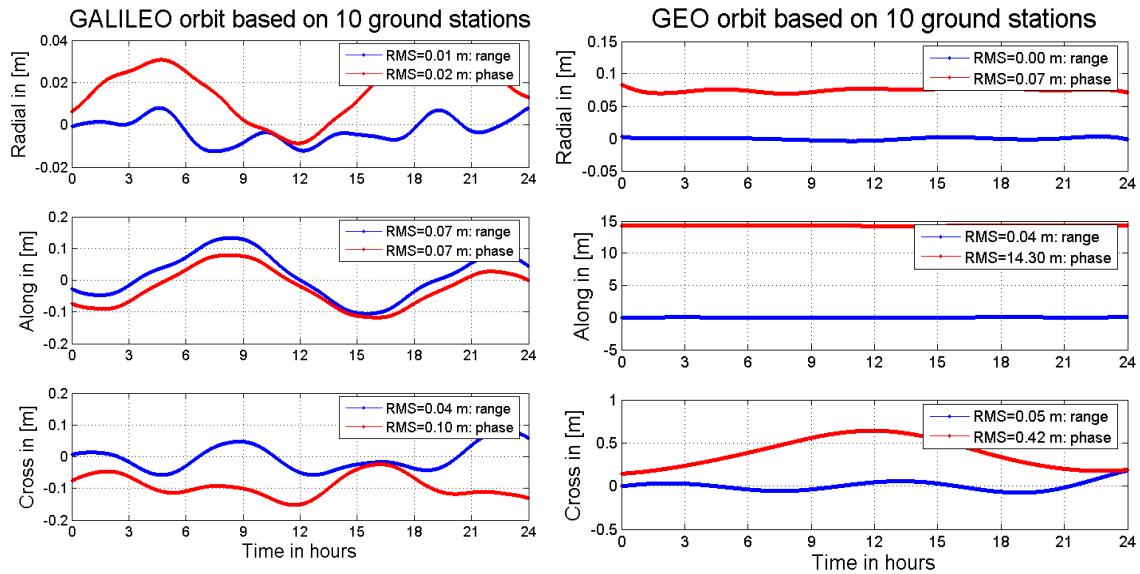
Figure 15.10 Accuracy of Galileo orbit (MEO) based on tracking from 10 ground stations (left) and accuracy of GEO orbit based on 10 ground stations (right).

This GEO reference frame needs to be tied to the ground to complement the Earth terrestrial reference frame defined by the global network of ground stations fixed to the Earth's crust.

Compared to pseudo-range and carrier-phase measurements from the GPS or Galileo system, SLR measurements provide ranges that do not require the estimation of clock and ambiguity parameters in orbit determination. This is the reason why simulations show that using such range measurements, orbits of GEO satellites could be determined with an accuracy of a few centimeters based on only 10 ground stations, see Figure 15.10. At the moment, not all SLR stations can be used for tracking GNSS satellites and thus also not for GEO satellites. It should be noted that such level of accuracy of orbit determination cannot be obtained with one-way measurements, such as carrier phase or pseudo-range measurements. GNSS based one-way measurements require the estimation of additional parameters, such as phase ambiguities and clock parameters,
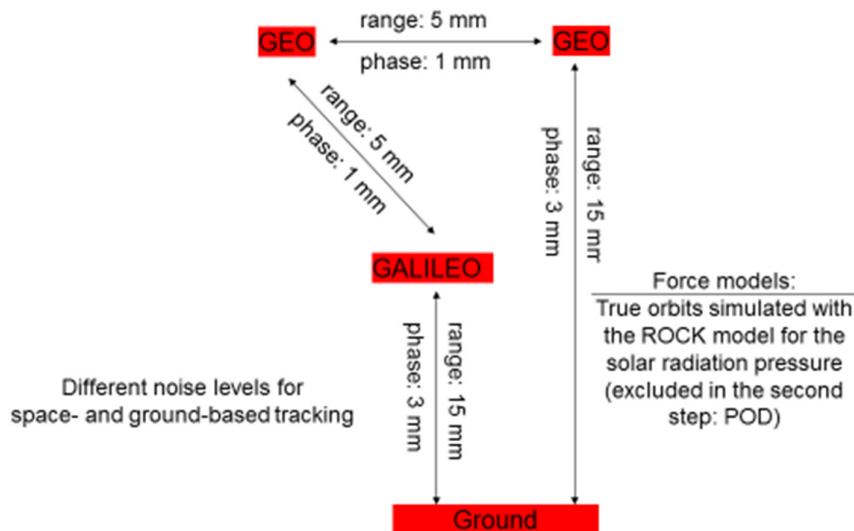
Figure 15.11 Noise level used in the simulation for range and carrier-phase of the navigation concept with 3-5 GEO satellites, Galileo in MEO and 10 ground stations. The ROCK solar radiation pressure model was applied a priori in the simulation. Typically, 15 orbital parameters were estimated for daily GNSS orbits.
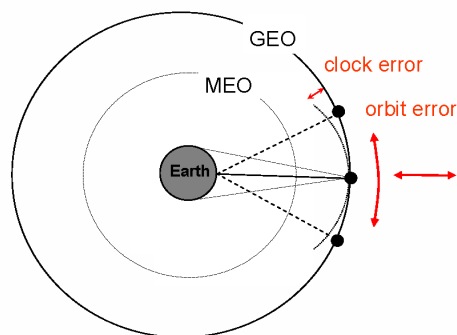
Figure 15.12 Geometry of the GEO and MEO orbit sensitivity, distinguishing the orbit error in the along-track and the radial orbit directions influenced by the clock error.

that are highly correlated with the GEO orbit parameters and, in this case, cannot be separated from orbit parameters with sufficient accuracy. This can be clearly seen in Figure 15.10 (right). The RMS of 14.3 m in the along-track direction for the orbit solution based on carrier-phase measurements compares with 0.04 m based on ranges. The simulation was carried out using 10 globally distributed SLR stations, assuming noise in the range measurements of $\sigma(range) = 15$ mm and in the carrier-phase $\sigma(phase) = 3$ mm. For the noise level applied in the simulation for range and carrier-phase measurements of the navigation concept with 3-5 GEO satellites, we refer to Figure 15.11. Since the gravity field can be assumed as error-free for the GEO orbit altitudes, the main source of error remains solar radiation pressure. It should be noted that GEO orbit determination will be sensitive to the polar and equatorial flattening of the Earth's gravity field as well as lunar and solar gravitational forces, and they will be driving factors in the dynamic orientation of the orbital plane in the inertial frame. One should also consider resonances in the GEO orbit and periodic maneuvers. However, since all 3-5 satellites are affected by resonances in longitude in a similar way, it is expected that the entire GEO constellation could also drift as a whole over a longer period of time. For more on resonances of GEO satellites see (Hugentobler et al. 1999).

Solar radiation pressure remains the main source of error in the realization of the GEO reference frame. In the simulation, we employed the ROCK model developed for GPS satellites as implemented in the Bernese GPS Software v5.1, see e.g., (Rothacher and Mervart 1996). The GEO orbit determination was performed using orbit parameterizations similar to GPS satellites, i.e., an arc length of only 24 hours and the standard 9 solar radiation pressure parameters in the Bernese GPS Software v5.1. Since we did not make use of longer arcs, in reality one could expect significantly better results, especially when all GEO satellites are combined together with cross-link measurements and the solar radiation pressure parameters are estimated simultaneously for all GEO satellites. Looking at Figure 15.10, one can observe that the orbit quality in the radial direction is well below the noise level of the simulated measurements, i.e., it is heavily constrained by the gravity field of the Earth. It has been shown, (Thaller et al. 2010), that for good ground ILRS stations, the noise floor of SLR measurements to GPS satellite G06 is at the level of 13 mm. Performing the same simulation for GPS satellites and using the same 10 ground stations reveals that GEO orbits can, in fact, be estimated with better accuracy than MEO orbits, see Figure 15.10 (left). This is especially true for the radial and along-track orbit components, whose accuracy is better by a factor of ≈ 2 compared to the GPS orbit. This paradox in our simulation can be explained by the fact that the radial orbit error for GEO is smaller by at least a factor of 2 than that for GPS, and thus the along-track component of the GEO orbit can indeed be better estimated. However, this is only true if highly accurate range measurements are available that heavily constrain the radial orbit component.

Figure 15.12 graphically explains why the along-track orbit component of a GEO orbit is estimated with very low accuracy when using carrier-phase measurements, or any observable that requires the estimation of GEO satellite clock parameters. True range measurements such as SLR can provide enough information to accurately constrain in-plane orbit rotation. It should be noted that MEO orbit can only be observed by the

same ground station for several hours, whereas GEO orbit allows continuous tracking from the terrestrial reference frame. Other alternatives for determining an accurate along-track GEO orbit component include differential VLBI or GNSS double-differences (against the GNSS constellation and the GEO satellite). However, ambiguity resolution would play an important role in this case and should be performed using a geometry-free method. One could also assume that significant information will come from the highly accurate intersatellite ranging between GEO satellites, since sub-micrometer level accuracy could be achieved in the free space in GEO orbit.

Generally speaking, a GEO dynamic reference frame has the potential to provide an alternative realization of the frame and complement and extend realization of the conventional terrestrial reference frame of the Earth. In this way, both the celestial and the terrestrial reference frame of the Earth could be combined with a GEO dynamic reference frame at the same time. Drag-free and ranging technology as developed for the LISA mission support this contention. A GEO reference frame could not only provide a basis for the real-time positioning/timing facility for GNSS Earth-based applications, but could also be used for positioning and time/frequency dissemination in the very populated GEO belt. MEO (GNSS), LEO, and satellites in GEO orbit could make use of this reference frame in GEO orbit, e.g., for real-time orbit determination and time/frequency dissemination. Figure 15.13 shows the accuracy of a Galileo orbit based on tracking from five GEO satellites. One can see that all three orbit components can be estimated with a similar level of accuracy. If tracking from several GEO satellites in the equatorial plane is available to satellites in MEO, one should expect that one orbit component could be determined with less accuracy. However, considering that GNSS orbit is determined with only 15 parameters, one can see from Figure 15.13 that all orbit components for a GNSS satellite can be determined with a similar level of accuracy.
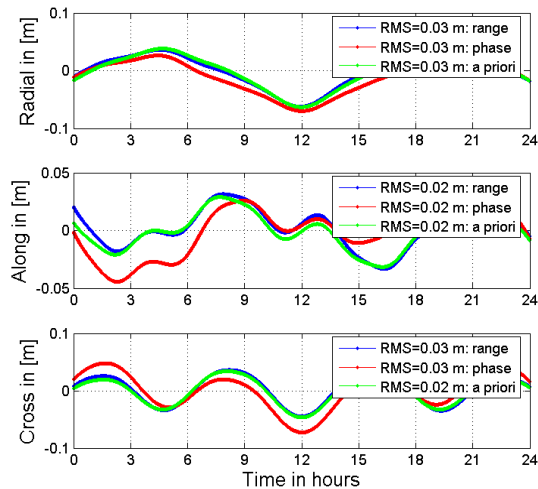
Figure 15.13 Accuracy of a Galileo orbit (MEO) based on tracking from five GEO satellites. Considering that GNSS orbit is determined with only 15 parameters, one can see that all orbit components can be determined with a similar level of accuracy. The along-track component is slightly more accurate, considering equatorially placed GEO satellites.

16. The GPS Transponder Concept – Towards One-way and Two-way GNSS Frequency Transfer

I n this section we discuss alternative, geometry-free approaches for positioning and time/frequency transfer using one-way and two-way measurements. The transmitter and receiver clock parameters can be separated or removed from the tracking geometry by using two-way measurements or introducing one-way measurements into the geometry-free linear combination. Clocks on board GNSS have become so stable that it makes interesting to steer their frequency using a geometry-free approach as demonstrated here. Galileo satellite clock parameters can be modelled using just two parameters per day (time drift and offset) with the remaining residual clock parameters showing the standard deviation at the level of 15 mm , see Section 18. Therefore, frequency steering of the satellite clock could be performed far more infrequently, (e.g., once a day) using the two-way frequency transfer approach. This could also bring to the separation of the prediction of GNSS satellite clock parameters (based on frequency steering) from the orbit prediction. We also discuss an application of the one-way frequency transfer approach based on geometry-free linear combination between two satellites (e.g., between GNSS satellites in MEO or with GEO). On the development of the two-way microwave metrology links for atomic clocks of the ACES mission we refer to (Cacciapuoti and Salomon 2009).

In addition to providing a two-way frequency transfer capability for GNSS, one could also consider the GPS-transponder concept, where a GNSS signals is tracked by a LEO GBSS receiver and then re-transmitted by the LEO satellite to a ground station (e.g., on a slightly shifted frequency). This opens up the possibility of separating tracking geometry from clock information when using a one-way approach for positioning, similar to the geometry-free two-way approach. One could also consider combining the standard one-way GPS positioning with the one-way frequency transfer. Observables in the one-way frequency transfer based on geometry-free linear combination would then be free of propagation effects, such as the effects of the ionosphere and the troposphere. The one-way approach based on geometry-free linear combination would also eliminate errors due to tropospheric effects and atmospheric turbulence in the case of optical measurements, and tropospheric effects and first and higher-order ionospheric corrections in the case of microwave measurements.

We also discuss the geometrical mapping of GNSS constellations with VLBI against extragalactic radio sources in the GPS-transponder configuration. At the end of this section, we discuss the idea of a similar two-way approach constructed using VLBI to observe both LAGEOS and passive laser retro-reflectors on the Moon in a bi-static radar configuration.

16.1 Principles of the One-Way and Two-Way Tracking

One-way tracking involves one signal transmitter with a stable frequency reference and one receiver, whereas in the case of two-way tracking an additional transponder is used. Such a configuration can be implemented for both radio and optical frequency bands anywhere within the Solar System. Transponders in navigation/data communication in space (e.g., interplanetary satellites) typically operate by sending the received

radio signal back to the transmitter, only with amplification of the received signal and shifting the signal from the uplink to the downlink frequency in order to avoid signal interference. Thus transponders work as a frequency translator, using an onboard ultra-stable oscillator (USO) and a frequency mixer to convert the frequency of the received incoming Doppler-shifted signal to the frequency required for the transmitted downlink signal. An onboard satellite receiver uses a phase-locked loop to lock the uplink carrier and to generate a reference signal coherent with that uplink carrier. Similar to GNSS, this reference signal is used to demodulate the ranging signal (ranging tones) received on the uplink carrier. As with GNSS, this ranging signal is again phase-modulated onto the downlink carrier that is shifted in frequency and coherent with the uplink carrier (reference signal). Thus, the frequency transmitted by the satellite is a Doppler-shifted replica of the uplink frequency. Typically, for the Deep-Space Network (DSN) for example, the downlink carrier frequency is higher by a factor of 880/749 in X-band and 3344/749 in Ka-band for an X-band uplink. The station that generated and transmitted the uplink signal receives the downlink signal and uses a PLL (Phase Locked Loop) to generate a reference signal coherent with the received signal. The round trip two-way transit time is determined by comparing the received range code with a model of the transmitted range code on the uplink. The same ground frequency standard is used to generate ranging codes consisting of a sequence of sinusoidal tones. In addition to range measurements, two-way Doppler measurements are derived by comparing the received reference signal with the same ground frequency reference used to generate the uplink carrier. The Doppler cycle counter measures the phase change of the Doppler tone (frequency difference) during a given count time, thus providing a measure of the range change over a given time interval.

The state of the art of technology in two-way interplanetary tracking is the radio-science instrument developed for the BepiColombo mission based on a Ka/Ka-band digital transponder enabling a high phase coherence between uplink and downlink carriers and supporting a wideband ranging tone. (For more details see (Iess et al. 2009).) The BepiColombo wideband ranging system is designed for an end-to-end accuracy of 20 cm using integration times of a few seconds based on the simultaneous transmission and reception of multiple frequencies in X- and Ka-bands with two-way range-rate measurements accurate to $3 \mu \mathrm{m/s}$, (Iess et al. 2009). In the case of the SELENE mission, differential same-beam VLBI interferometry has been successfully demonstrated between the two Lunar orbiters tracked by the same ground VLBI antenna, and further differenced between two ground stations. Differential same-beam interferometry provides extremely accurate relative position measurements in the plane-of-the-sky, thus complementing the line-of-sight information one can obtain from the two-way Doppler and range measurements. It was reported in (Goossens et al. 2010) that the differential phase delay obtained in this way on the X-band signal can be estimated to within 1 ps (0.3 mm). In the case of S-band data, obtained with wider beamwidth compared to X-band, differential phase delay was determined with an error of a few picoseconds (roughly 1 mm) for narrow separation angles of the spacecraft, and about 10 ps (3 mm) for wider angles. These accuracies include effects of the ionosphere and atmosphere, (Goossens et al. 2010). The advantage of the differential same-beam VLBI measurements lies in the differencing out of common errors over a very narrow beamwidth angle. However, if the differential measurement is performed on only a single frequency, the total phase delay is biased by an integer ambiguity. To overcome the cycle ambiguity problem in the same-beam VLBI interferometry and to increase the accuracy of the SELENE measurement, a multi-frequency method was used, with three carriers in the S-band (2212, 2218 and 2287 MHz) and one in the X-band (8456 MHz), (Goossens et al. 2010).

Following (Border and Kursinski 1991), the internationally allocated frequency bands for uplink/downlink used in the communication/navigation of interplanetary missions are given in Table 16.1. In the case of a very long round trip transit time, e.g., a distant interplanetary mission, when the downlink signal reaches Earth, the satellite might no longer be in the field of view of the ground station which transmitted the uplink signal. Thus a second ground station is required to receive the downlink signal. Such tracking is referred to as "three-way tracking". For example, for the distances to Neptune, the round-trip light travel time is more than 8 hours. Similar scenarios may arise with two satellites (e.g., the SELENE mission in the lunar orbit) and one ground station, where one can even identify four-way tracking. In all these cases, high stability of the onboard

Frequency Band	Uplink Frequency [MHz]	Downlink Frequency [MHz]
S	2110 – 2120	2290 - 2300
X	7145 – 7190	8400 - 8450
Ka	34 200 - 34 700	31 800 - 32 300

Table 16.1 Internationally allocated frequency bands used for navigation/communication of interplanetary missions (DSN) (Border and Kursinski 1991).

(ultra-) stable oscillator is essential and, typically, additional parameters need to be taken into account in orbit determination to model the onboard frequency offset. Thus, any instability or inaccuracy of the onboard frequency reference translates directly into an error $\Delta\dot{\rho}$ in range rate

$$\Delta\dot{\rho} = c\frac{\Delta f}{f} \ . \tag{16.1}$$

Assuming the frequency instability over a tracking pass to be in the order of $\Delta f / f = 10^{-14}$, we have an error in the range rate in the order of $3\,\mu m/s$. For comparison, the typical accuracy of the radial velocity of GPS satellites is in the order of $5-10\,\mu m/s$ (based on orbit solutions provided by the IGS Analysis Centers), whereas in the case of GOCE, in very low LEO, the velocity can be determined with an accuracy in the order of $15-25\,\mu m/s$ for all three components.

The state-of-the-art two-way approach was developed for the ACES mission in LEO orbit, see (Cacciapuoti and Salomon 2009) making use of the small ground and spaceborne antenna. Therefore, there are enough arguments to consider the two-way approach for future GNSS. This is especially true considering that clocks on board Galileo have become so stable that they can easily be steered from the ground using the two-way frequency transfer approach. This could be performed very infrequently, e.g., once a day, using ground clocks in the UTC/TAI network that have several orders of magnitude better accuracy and stability than the Galileo clocks. Such an approach could even be extended by using master clock(s) in the GEO orbit. Frequency steering for GNSS is a very interesting new technique for GNSS, considering that Galileo satellite clock parameters can be modelled with just two parameters per day (time drift and offset). See Section 18, for more details on the Galileo clock performance, where we showed that remaining residuals for Galileo clocks (after removing the linear model) have a standard deviation at the level of 15 mm. This could also bring to the separation of GNSS satellite clock parameters (frequency steering) from the orbit prediction.

16.2 Geometry and Propagation Constraints from LEO to Interplanetary Distances

The two-way approach can be used to transfer frequency between two clocks free of any geometry effects, since these are removed by differentiating downlink and uplink measurements. The observation equation for the downlink $L_1^2(t_2)$ from e.g., a satellite to a ground station and uplink $L_2^1(t_2)$ carrier-phase measurements in the two-way form can be given as

$$L_1^2(t_2) = \rho_1^2 + \lambda_1^2 N_1^2 - b^{sat}(t_1) + b_{rec}(t_2)$$
$$L_2^1(t_2) = \rho_2^1 + \lambda_2^1 N_2^1 + b^{sat}(t_2) - b_{rec}(t_1) \tag{16.2}$$

where $L_1^2(t_2)$ is tracked by the ground receiver at the reception time t_2. The $b_{rec}(t_2)$ denotes the receiver clock error at the reception epoch t_2 and the satellite clock error $b^{sat}(t_1)$ is referred to the transmitting epoch t_1. The term ρ_1^2 includes all geometry terms for the downlink between epochs t_1 and t_2. One can say that

the first equation given for the downlink in (16.2) is the same as for the GNSS one-way measurements. The uplink carrier-phase measurements $L_2^1(t_2)$ can be performed on the satellite, or referred to the ground receiver if the satellite is used as a transponder of the uplink signal and the signal is sent by the ground station. If we assume that the satellite measures carrier-phase $L_2^1(t_2)$ from the uplink, the observation equation is given as the second equation in (16.2). For the GNSS orbit altitude, we can model light-travel time for the uplink and the downlink using the line-of sight velocity of the GNSS satellites relative to the ground station, similar to the geometry between GNSS and a ground station or a LEO satellite in space, see Section 2. This means that the geometry terms for downlink ρ_1^2 and uplink ρ_2^1 are nearly equal and can be removed with the sufficient accuracy. Making a difference of (16.2) we derive

$$\frac{1}{2}\left(L_1^2 - L_2^1\right) = \frac{1}{2}\left(\lambda_1^2 N_1^2 - \lambda_2^1 N_2^1\right) - \frac{1}{2}\left(b^{sat}(t_1) + b^{sat}(t_2)\right) + \frac{1}{2}\left(b_{rec}(t_2) + b_{rec}(t_1)\right) \tag{16.3}$$

From (16.3) we can see that our observation model is still biased by the carrier-phase ambiguities for downlink $\lambda_1^2 N_1^2$ and uplink $\lambda_2^1 N_2^1$ that are typically given at different frequency. To remove ambiguity parts, one can make use of differencing over time. Frequency difference between a ground clock and a space clock from time t_A and t_B

$$\left[\frac{1}{2}\left(L_1^2 - L_2^1\right)\right]\Bigg|_{t_B} - \left[\frac{1}{2}\left(L_1^2 - L_2^1\right)\right]\Bigg|_{t_A} = \left[-\frac{1}{2}\left(b^{sat}(t_1) + b^{sat}(t_2)\right) + \frac{1}{2}\left(b_{rec}(t_2) + b_{rec}(t_1)\right)\right]\Bigg|_{t_B} -$$

$$\left[-\frac{1}{2}\left(b^{sat}(t_1) + b^{sat}(t_2)\right) + \frac{1}{2}\left(b_{rec}(t_2) + b_{rec}(t_1)\right)\right]\Bigg|_{t_A} \tag{16.4}$$

$$= f_{rec} - f^{sat}$$

gives the frequency difference $f^{sat} - f_{rec}$ that can be written

$$\frac{d}{dt}\left[\frac{1}{2}\left(L_1^2 - L_2^1\right)\right] = \Delta f = f_{rec} - f^{sat} \tag{16.5}$$

If the satellite is used as a transponder of the uplink signal sent from a ground station, carrier-phase measurements can be performed separately for uplink and downlink by the ground receiver. Carrier-phase measurements can also be performed between an uplink and a downlink signal by the ground station. In case the ground station is used as a transponder, carrier-phase measurements can be performed by the satellite. Again, a geometry-free frequency offset is determined between a ground station clock and a satellite clock. Such measurements can be used to steer very accurately the onboard frequency of a GNSS satellite. We do not consider an error budget in full detail here. However, visibility time of a GNSS satellite from a ground station is typically several hours (e.g., 6 hours), compared to a very short observation time, limited to about 5 min, for the ACES mission in LEO orbit. This gives a lot of confidence for future GNSS considering that the ACES two-way link with satellite clocks showing two orders of magnitude better performance compared to Galileo clocks is a guarantee of such an approach.

The question is what are the limitations of the two-way approach? Considering that there is a light-travel time between a satellite and a ground station Δt, one could distinguish a Δ-configuration when differencing is referred to the common epoch on a satellite or V-configuration when differencing between downlink and uplink measurements is done for an epoch referred to the a ground station, see Figure 16.1. The propagation path in the atmosphere for Δ-configuration and V-configuration is slightly different, and could be a source of error. This is especially true for the ground-to-LEO or ground-to-GNSS clock comparison in Δ-configuration with a large point-ahead angle between the ground station and the LEO satellite, or for a HEO orbit with a very long light-travel time. In both cases, atmospheric turbulence (with a spectrum up to some 1000 Hz) can
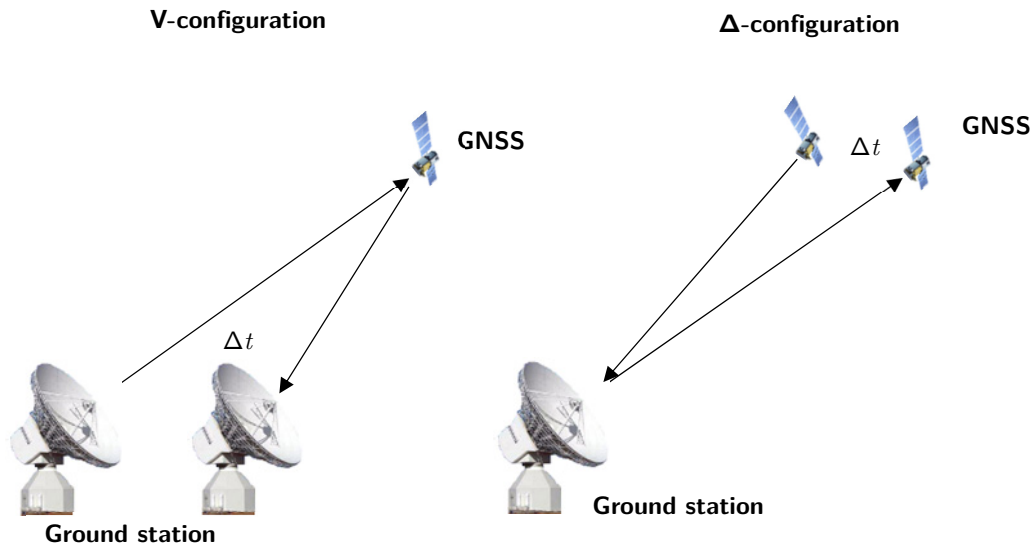
Figure 16.1 The Δ-configuration and V-configuration of the two-way approach for uplink and downlink measurements. Due to the light-travel time Δt, signal propagates slightly differently for the atmosphere conditions in the Δ-configuration, whereas satellite moves during the light-travel time in the V-configuration.

generate an effect in the optical or microwave phase that is not eliminated by forming differences between carrier-phase measurements of the two symmetric paths. Considering the very short wavelength of the optical frequencies used, in the case of optical measurements, this asymmetry could even prevent coherent tracking of the optical signal, i.e., coherent optical carrier with GHz-modulation. In the case of a clock on an interplanetary satellite, the light-travel time could easily reach 30 min (2×8.3 min/AU) and during that period of time the dry/wet part of the atmospheric delay could significantly change (not only due to atmospheric turbulence). Because of Earth's rotation, the point-ahead angle changes by $\approx 2 \times 2.1°$/AU, and is about $10°$ at the distance to, e.g., Mars ($a = 1.5$ AU). In the case of the ionosphere this asymmetry introduces different bending angles between two counter-propagating waves and the ionospheric/plasmaspheric effects are generally different for the two waves. This complicates the removal of the first order ionosphere-effect by using the ionosphere-free linear combination. In the case of higher order effects of the ionosphere, these are not eliminated by forming differences or ionosphere-free linear combination, but effects on uplink and downlink sum up. Ionospheric/plasmaspheric effects can be reduced by making use of the higher microwave bands such as Ka-band or higher, where the second and higher order effects are insignificant. In the case of the LISA mission (The Laser Interferometer Space Antenna) with an armlength of 5 Mkm, or eLISA (Extended LISA) with an armlength of 1 Mkm, an additional constraint is precise pointing or alignment for the optical telescope. Even ILRS stations with very good ground stabilization very often report difficulties in directing SLR telescopes at GNSS satellites. Tracking over lunar distances is feasible only for a few ILRS stations. Considering the previous example with the Mars distance, the SLR telescope should be re-aligned by about $10°$ between an uplink and a downlink.

All these geometry and propagation constraints would be eliminated if one designed a one-way metrology link, because in that case the signals would propagate along the same path through the atmosphere or interplanetary plasmasphere.

16.3 The One-way Geometry-Free Approach to Frequency Transfer

Let us imagine that a satellite clock is transmitting to the receiver the same signal twice, i.e., with frequency f^{sat}, and a frequency $f^{sat,2}$ shifted by an offset Δf, with $f^{sat,2} = f^{sat} + \Delta f$, see Figure 16.2. Such a scenario is typical for GNSS satellites, for time and frequency transfer, and for some interplanetary missions. In addition to the reference clock error b^{sat} associated with the frequency f^{sat}, any use of a transponder or frequency multiplication, such as the case of GNSS, will introduce an additional time delay error $b^{sat,2}$ in the generation of the frequency offset Δf. The observation equation for the carrier-phase measurements denoted here as $L_1(b^{sat})$ and $L_2(b^{sat}, b^{sat,2})$ tracked by the receiver is

$$L_1(b^{sat}) := \rho + \lambda_1 N_1(b^{sat}) - b^{sat} + b_{rec} + I_1$$

$$L_2(b^{sat}, b^{sat,2}) := \rho + \lambda_2 N_2(b^{sat}, b^{sat,2}) - b^{sat} - b^{sat,2} + b_{rec} + (\frac{f^{sat}}{f^{sat,2}})^2 I_1$$

(16.6)

where b_{rec} is the receiver error on the ground with geometry term ρ, ambiguity terms $\lambda_1 N_1$ and $\lambda_2 N_2$ and the first-order ionosphere-effect I_1. The frequency offset can be small enough to guarantee that there is no interference between the two signals. The geometry-free linear combination $L_4(b^{sat}, b^{sat,2})$ is then

$$L_4(b^{sat}, b^{sat,2}) = L_1(b^{sat}) - L_1(b^{sat}, b^{sat,2})$$

$$= b^{sat,2} + \lambda_1 N_1(b^{sat}) - \lambda_2 N_2(b^{sat}, b^{sat,2}) + \left[1 - (\frac{f^{sat}}{f^{sat,2}})^2\right] I_1$$

(16.7)

with the last term denoting the differential first-order ionosphere effect ΔI_1. If we now make a difference of (16.7) over a time interval t, the ambiguity parameter will be removed

$$\Delta L_4(b^{sat}, b^{sat,2}) = \Delta\left[L_1(b^{sat}) - L_1(b^{sat}, b^{sat,2})\right]_0^t = b^{sat,2}(t) - b^{sat,2} + \Delta I_1(t) - \Delta I_1$$

$$= \int_0^t \left(\Delta f + \Delta I_1\right) \cdot dt$$

(16.8)

By increasing the frequency, the first order ionosphere-effect reduces by $1/f^2$. We will show later in this section that differential ionospheric effects ΔI_1 between two frequencies are proportional to $1/f^3$ and thus
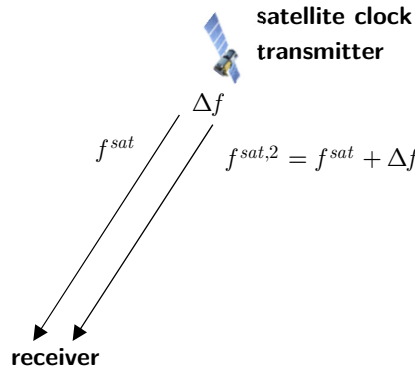
Figure 16.2 General concept of the one-way approach to transfer frequency offset Δf. When received by the receiver, geometry and propagation effects can be removed by using geometry-free linear combination in the time domain. In this way, frequency offset Δf can be directly measured by the ground receiver. The first order ionosphere-effect reduces by $1/f^2$ and is significantly smaller in size when differenced in time.

are significantly smaller in size when differenced in the time domain. Thus, they can either be neglected or removed using an a priori ionosphere-model. The use of two frequencies in X-band or in the Ka-band will decrease the first order ionosphere effect by a factor of at least 100 compared to the GNSS frequencies in the L-band. The high-order ionosphere-effects that are proportional to $1/f^3$, $1/f^4$ or $1/f^5$ are reduced even faster by increasing the frequency. When considering such a concept in space, using a space-based receiver and a space-based transmitter, frequency steering between satellites or a GEO satellite could, in principle, be performed using the one-way approach. GNSS satellites are high above the ionosphere and plasmasphere effects are significantly reduced.

The clock parameter $b^{sat,2}(t)$ in (16.8) is associated with the frequency $f^{sat,2}$ and measured against the frequency of the receiver f_{rec}

$$b^{sat,2}(t) = \int\limits_0^t \Delta f \cdot dt \tag{16.9}$$

In this way, we can measure a frequency offset Δf of the satellite clock relative to the same frequency offset generated by the receiver. The geometry and propagation effects can be removed by using the geometry-free linear combination in the time domain (16.8). This concept could be realized with optical and microwave measurements and is basically free of all propagation effects. If we assume that typical LEO orbit velocity is known with a standard deviation of 0.01 mm/s (over a daily period), the error in the first order Doppler effect will give a relative frequency offset in the reference frequency to an accuracy of $3 \cdot 10^{-14}$ for a single station and the effect will be reduced by using single-differences with two stations as a function of nadir angle α_{nadir}

one station   two stations
$$\frac{\Delta f}{f} = \frac{0.01\,{}^{mm}\!/_{s}}{c} \approx 3 \cdot 10^{-14} \qquad \frac{\Delta f}{f} = \frac{0.01\,{}^{mm}\!/_{s}}{c}\Big[1 - \cos(\alpha_{nadir})\Big] \approx 4 \cdot 10^{-16}\,/\,10° \tag{16.10}$$

Since the determined radial orbit velocity is more accurate for GNSS and averages out for a typical GNSS orbit, this approach is very interesting for the application of tri-carrier GPS measurements (GPS BLOCK-IIF). In the case of common-view single-differencing the effect is significantly reduced by about two orders of magnitude, offering the possibility of achieving a level of accuracy for the relative frequency comparison in the order of 10^{-18} over several hours of averaging. Currently, there are about 12 GPS BLOCK-IIF satellites in the GPS constellation.

16.3.1 Differential Atmospheric Effects in Optical and Microwave Bands

Here we look at the differential ionospheric and tropospheric effects on the one-way signal on the two frequencies f_1^* and f_2^* close to one another in the optical band and separately, two frequencies f_1^r and f_2^r in the microwave frequency band.

Let us first look at the differential ionospheric effect on the two frequencies close to one another in the GNSS L-band. The derivative of the first order ionosphere-effect I in the zenith direction gives

$$dI = \frac{2 \cdot k_{TEC}}{f^3} TEC \cdot df = -2\frac{df}{f} I \tag{16.11}$$

with typically used value for $k_{TEC} = 40.3$ in the first-order ionosphere-effect. With a differential microwave frequency df at the MHz-level, the first order effect is reduced by about three orders of magnitude, thus it can be neglected or easily eliminated to below the mm-level by the simple Klobuchar-grade ionosphere models available from the broadcast navigation message. Higher-order ionospheric effects are completely eliminated

as are the tropospheric effects. At higher frequencies, such as Ka-band, the first order ionosphere-effect is further reduced by a factor of about 30 compared to the f_1 GPS frequency. For the differential atmospheric effect on optical frequencies, we make use of the Marini-Murray model, IERS Conventions 2003 (McCarthy and Petit 2004). The range correction due to the Marini-Murray is

$$\Delta R = \frac{f(\lambda)}{f(\phi, \mathrm{H})} \cdot \frac{A + B}{\sin E + \dfrac{B/(A+B)}{\sin E + 0.01}} \tag{16.12}$$

with elevation of satellite E and A and B given in (McCarthy and Petit 2004). The laser site function is denoted by $f(\phi, \mathrm{H})$ and the laser frequency parameter $f(\lambda)$ is

$$f(\lambda) = k_1 + \frac{k_2}{\lambda^2} + \frac{k_3}{\lambda^4} \tag{16.13}$$

given for the wavelength λ in micrometers. For a ruby laser $f(\lambda) = 1$. For the constants $k_1 = 0.9650$ and $k_2 = 0.0164$ and $k_3 = 0.000228$ we refer to (McCarthy and Petit 2004). From (16.13) it follows that the differential in the range correction (meters) is

$$d\Delta R \approx -\Delta R \frac{0.0334}{\lambda^3} d\lambda \tag{16.14}$$

which for the infrared wavelength of 1064 nm gives

$$d\Delta R \approx -0.0311 \Delta R \cdot d\lambda \tag{16.15}$$

A difference in the wavelengths of 1 nm gives about 30 μm per 1 m of range correction for a wavelength of 1064 nm. The accuracy of the troposphere models used for SLR is below one millimeter, (see IERS Conventions 2010 (Petit and Luzum 2010)). Therefore, by utilizing an a priori troposphere model, e.g., (Mendes and Pavlis 2004), the accuracy of our differential troposphere model (16.15) can be significantly increased.

16.3.2 A Concept for an Interferometric Metrology Link

Let us now look at the case where a satellite is transmitting a carrier-wave on two frequencies f_1^* and f_2^* in the optical or near-infrared spectrum separated by the beat frequency Δf^* in the microwave domain. This frequency separation could be chosen to be in the microwave band of GNSS frequencies, e.g.,

$$\Delta f^* = f_2^* - f_1^* = 154 \cdot f_0 = f_1, \tag{16.16}$$

where f_0 denotes the fundamental GPS frequency $f_0 = 10.23$ MHz and f_1 is the GPS frequency. Taking into account only the first order Doppler effect, the frequencies of the signal received on the ground are

$$f_{R,1}^* = \left(1 - \frac{\dot{\rho}}{c}\right) f_1^*, \qquad f_{R,2}^* = \left(1 - \frac{\dot{\rho}}{c}\right) f_2^* \tag{16.17}$$

where $\dot{\rho}$ denotes the line-of-sight range rate with the Doppler shifted beat frequency Δf_R^*

$$\Delta f_R^* = f_{R,2}^* - f_{R,1}^* = \left(1 - \frac{\dot{\rho}}{c}\right) \Delta f^* = \left(1 - \frac{\dot{\rho}}{c}\right) f_1 \tag{16.18}$$

Instead of tracking each optical frequency separately, we combine them in order to generate the beat frequency Δf_R^* using optical heterodyning

$$\sin\left(2\pi f_{R,1}^* \, t\right)\sin\left(2\pi f_{R,2}^* \, t\right) = \frac{1}{2}\cos\left[2\pi\left(f_{R,2}^* - f_{R,1}^*\right)t\right] - \frac{1}{2}\cos\left[2\pi\left(f_{R,2}^* + f_{R,1}^*\right)t\right]. \tag{16.19}$$

Heterodyning is a radio/optical signal processing technique in which two reference frequencies are linearly combined or mixed in order to create two new frequencies (differencing/summation). From the trigonometric relation (16.19), we see that the multiplication of two carrier waves generates two new signals. Applied to our case, two frequencies close to one another in the visible part of the spectrum $f_{R,1}^*$ and $f_{R,2}^*$ illuminate the photo-detector in the receiver and the oscillating electrical signal corresponds to the difference between their frequencies, i.e., the beat frequency Δf_R^*. In our case, this beat frequency corresponds to the Doppler-shifted GPS frequency f_1 in the L-band. In the next step, the generated beat signal with frequency Δf_R^* is compared against the reference signal from the ground frequency reference in order to generate carrier-phase measurements. Using this approach, we can obtain very precise measurements of phase and frequency differences between two optical signals. Optical heterodyne detection is used for many applications, such as coherent Doppler LIDAR measurements that are capable of detecting very weak light scattered in the atmosphere or monitoring wind speeds in the atmosphere with a high degree of accuracy. One can find many applications in high-accuracy optical frequency measurements, including frequency combs.

Considering the relative velocity of the GPS satellite in (16.18) for the range rate of, e.g., $\dot{\rho} = 4\ \text{km/s}$, the beat frequency Δf_R^* is in the order of 21 kHz and it generates the same carrier-phase signal as the GPS frequency f_1. The received beat frequency Δf_R^* is free of the first and higher-order ionospheric-effects, and the influence of atmospheric turbulence as well as the dry/wet part of the tropospheric delay is basically eliminated. Another approach would be to modulate the GPS frequency f_1 onto the optical carrier with frequency f_1^* (e.g., using phase modulation) and use the optical carrier as an "atmosphere tunnel", since optical frequencies are not affected by ionospheric effects and the dry/wet part of the troposphere can be modeled at the sub-mm-level, see IERS Conventions 2010 (Petit and Luzum 2010). In that case, carrier-phase measurements would need to be performed using a modulated "GPS carrier" on the optical carrier comparing it against the reference frequency used by the receiver. Such a concept would be feasible with only one optical frequency.

16.4 The GPS One-way Approach to Frequency Transfer

Typically, a GPS satellite uses the fundamental frequency $f_0 = 10.23\ \text{MHz}$ to generate by multiplication $f_1 = 154 \cdot f_0$, $f_2 = 120 \cdot f_0$ and $f_5 = 115 \cdot f_0$. Let us imagine a frequency offset Δf associated with, e.g., $f_5^* = (115 \cdot f_0 + \Delta f)$. Such an error will generate an additional clock error $b_5^{sat}(t)$ accumulated over time t

$$b_5^{sat}(t) = \frac{c}{f_5}\int_0^t \Delta f \cdot dt = c\int_0^t \frac{\Delta f}{f_5} \cdot dt = \frac{c}{115}\int_0^t \frac{\Delta f}{f_0} \cdot dt \tag{16.20}$$

given in meters and using the speed of light c. In addition to the satellite clock error b^{sat} associated with an error in reference GPS frequency df_0 and corresponding receiver clock error b_{rec} estimated every epoch, the observation equation for carrier-phase L_1 and L_2 with this newly modified L_5^* observable using f_5^* is

$$
\begin{aligned}
L_1 &= \rho + \lambda_1 N_1 - b^{sat} + b_{rec} + I_1 \\
L_2 &= \rho + \lambda_2 N_2 - b^{sat} + b_{rec} + \frac{f_1^2}{f_2^2} I_1 \\
L_5^* &= \rho + \lambda_5^* N_5^* - b^{sat} - b_5^{sat}(t) + b_{rec} + \frac{f_1^2}{f_5^{*2}} I_1
\end{aligned} \tag{16.21}
$$

with geometry term ρ, ambiguities N_1 and N_2, and wavelengths λ_1 and λ_2. For the new frequency f_5^*, we have an integer ambiguity N_5^* and wavelength λ_5^*. The first order ionosphere-effect on f_1 is denoted by I_1. If we now form the ionosphere-free linear combination relative to L_1 we obtain

$$
\begin{aligned}
L_3(L_1, L_2) &= \rho + \frac{c}{f_1 + f_2} N_1 + \frac{f_2 \cdot c}{f_1^2 - f_2^2}(N_1 - N_2) - b^{sat} + b_{rec} \\
L_3(L_1, L_5^*) &= \rho + \frac{c}{f_1 + f_5^*} N_1 + \frac{f_5 \cdot c}{f_1^2 - f_5^{*2}}(N_1 - N_5^*) + \frac{f_5^{*2}}{f_1^2 - f_5^{*2}} b_5^{sat} - b^{sat} + b_{rec}
\end{aligned}
\tag{16.22}
$$

The satellite clock error $b^{sat}(t)$ accumulated over time t for ionosphere-free linear combination $L_3(L_1, L_2)$ can be derived from a clock error df_0 in the fundamental GPS frequency f_0 as

$$
b^{sat}(t) = \frac{f_1 \cdot c}{f_1^2 - f_2^2} \int_0^t 154 \cdot df_0 \cdot dt - \frac{f_2 \cdot c}{f_1^2 - f_2^2} \int_0^t 120 \cdot df_0 \cdot dt = c \cdot \int_0^t \frac{df_0}{f_0} \cdot dt
\tag{16.23}
$$

where df_0 / f_0 denotes the relative frequency error of the fundamental clock frequency f_0. The same satellite clock is defined using the ionosphere-free linear combination on f_1 and f_5

$$
b^{sat}(t) = \frac{f_1 \cdot c}{f_1^2 - f_5^2} \int_0^t 154 \cdot df_0 \cdot dt - \frac{f_5 \cdot c}{f_1^2 - f_5^2} \int_0^t 115 \cdot df_0 \cdot dt = c \cdot \int_0^t \frac{df_0}{f_0} \cdot dt
\tag{16.24}
$$

Following the IGS convention, GPS satellite clock parameters are defined using the ionosphere-free linear combination $L_3(L_1, L_2)$ of L_1 and L_2 measured on P-code. From, (16.24) one can see that the same clock parameter is defined by f_1 and f_5 as long as $f_5^* = f_0 = 115 \cdot f_0$ and $\Delta f = 0$.

If we now subtract the two equations in (16.22), we derive ΔL_3

$$
\begin{aligned}
\Delta L_3 = L_3(L_1, L_5^*) - L_3(L_1, L_2) &= \frac{f_5^{*2}}{f_1^2 - f_5^{*2}} b_5^{sat} + \left(\lambda_{N(1,2)} - \lambda_{N(1,5)}^*\right) N_1 \\
&+ \frac{1}{2}\left(\lambda_{W(1,2)}^* - \lambda_{N(1,2)}\right) N_{W(1,2)} - \frac{1}{2}\left(\lambda_{W(1,5)}^* - \lambda_{N(1,5)}\right) N_{W(1,5)}^*
\end{aligned}
\tag{16.25}
$$

with the narrow-lane $\lambda_{N(1,2)} = c / (f_1 + f_2)$ and $\lambda_{N(1,5)}^* = c / (f_1 + f_5^*)$, and the wide-lane wavelengths $\lambda_{W(1,2)} = c / (f_1 - f_2)$ and $\lambda_{W(1,5)}^* = c / (f_1 - f_5^*)$, and ambiguities $N_{W(1,2)} = N_1 - N_2$ and $N_{W(1,5)}^* = N_1 - N_5^*$

The ambiguity part in (16.25) can be removed by making a difference in time t, measuring directly frequency offset Δf using the geometry-free linear combination $\Delta L_3(t)$ differenced over time

$$
\Delta L_3(t) = \frac{f_5^{*2}}{f_1^2 - f_5^{*2}} b_5^{sat}(t) = \frac{f_5^{*2}}{f_1^2 - f_5^{*2}} \cdot c \cdot \int_0^t \frac{\Delta f}{f_5^*} \cdot dt
\tag{16.26}
$$

considering that $b_5^{sat}(t)$ is an accumulated time error over t. The size of the term $(f_5^2 / f_1^2 + f_5^2) \approx 1.26$ is modest. Eq. (16.26) shows that one can transfer frequency offset Δf or relative frequency $\Delta f / f_5^*$ from space to ground in a geometry-free way, as in the case with the one-way approach with two frequencies.

16.5 The GPS Transponder Concept – Towards "Geometry-Free" Positioning

As demonstrated earlier in this section, using a one-way or two-way approach it is possible to eliminate frequency offset between receiver and transmitter from equations used in orbit determination.

Quasars can be observed over the entire observable electromagnetic spectrum, including radio, infrared, optical, ultraviolet, X-ray and even gamma rays. A selected frequency band could be observed by a phased-array antenna on board a satellite. The same approach could be applied to a GNSS and a LEO satellite, (Figure 16.3). This tracking could be performed in open-loop, as in the case of VLBI, or using a frequency comb to measure the spectrum of the received signal. The ideal solution would be to perform carrier-phase measurements, as in the GNSS-LEO case. At the same time, the signal tracked by the antenna array could be transmitted to Earth or measurements taken could be downloaded to a ground station using an (optical) communication link or retransmitted towards the Earth ("GPS transponder" concept) in the case of GNSS-LEO configuration. If the same quasar is tracked in the vicinity of the satellite (in the line-of-sight), one could directly measure the Doppler shift, i.e., the line-of-sight velocity of the satellite, by comparing the same signal from satellite and/or quasar. This measurement is geometry-free, since the frequency of the clock onboard the satellite could be measured from the ground using either one-way or two-way frequency measurements. This measurement of the line-of-sight velocity is not only geometry-free, but is also free of any propagation effects. The more than 3000 radio sources listed in ICRF2 are sufficient to enable any satellite to carry out such measurement in the light-of-sight direction. For the ground tracking, one would need to use VLBI or phased-array antennae. Due to good multipath mitigation capabilities and the low noise of GNSS observables, it is believed that phased-array antennae will find an application in permanent GNSS networks such as IGS.

This type of geometry-free measurement could support a pulsar-based time scale. According to (Hobbs et al. 2012), there are about 30 quasars that in terms of timing stability provide an alternative to TAI, as demonstrated in (Hobbs et al. 2012). With the proposed geometry-free approach, one could transfer a pulsar time scale from the satellite to the ground, eliminating atmospheric effects with the one-way or two-way approach. STE-QUEST has the potential to be the first mission to demonstrate the geometry-free one-way approach for positioning (making use of the existing onboard payload), see Section 27. The same approach could be applied to deep space missions carrying metrology links.
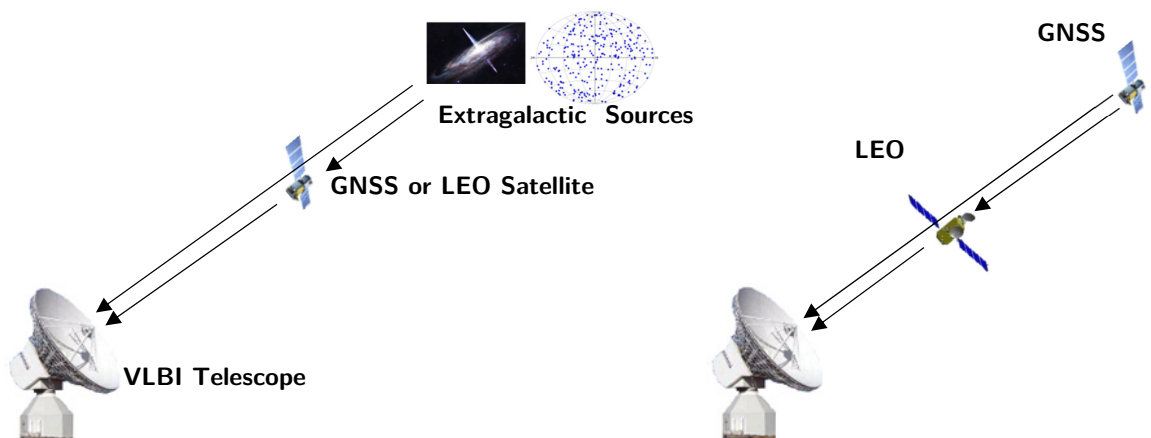
Figure 16.3 Concept of one-way geometry-free positioning: with a GNSS satellite and a quasar (left) and a GNSS satellite and a LEO or a GEO satellite (right) "GPS transponder" concept. Since time could be eliminated from orbit determination, the orbit itself can be observed w.r.t. another satellite or object. A navigation signal needs to be received and re-transmitted from the target satellite relative to the known object, e.g., the GNSS satellite or quasar.

16.6 Geometrical Mapping of a GNSS Constellation Against Extragalactic Radio Sources

Geometrical mapping of a GNSS constellation against extragalactic radio sources (quasars) can be realized by observing quasars at the approximate location of GNSS satellites. This is similar to the Delta-DOR approach used in the tracking of interplanetary satellites, where an open-loop receiver samples VLBI signals in the S-/X- and/or Ka-bands. To track GNSS satellites, one would also need to sample GNSS signals in the L-band, and correlate them on a correlator. Compared to GPS, there would be an advantage for Galileo, due to the wide range of different modulations on several frequencies and the higher bandwidth one could obtain from the Galileo signals. A second approach would be to observe GNSS carrier-phase and code measurements using a GNSS receiver connected to a VLBI antenna. Due to the size of the typical VLBI antenna dish and the pointing of the antenna, the thermal noise of the received signal would be significantly less than that experienced with the standard choke-ring omni-directional antenna currently used by the IGS.

Figure 16.4 shows ionosphere-free code against ionosphere-free carrier-phase measurements from the GNSS receiver connected to a 25 m antenna dish. Translated to the original single frequency observable, the noise of the code measurements is at 6 mm precision. Such a low code noise significantly simplifies ambiguity resolution of the carrier-phase measurements. This opens up the possibility of using the differential same-beam VLBI interferometry approach to track GNSS constellations. In this technique, two ground-based VLBI stations track the same two close-by GNSS satellites within the beam width of the VLBI antennae. Differential same-beam interferometry provides very accurate relative positioning measurements in the plane-of-the-sky (same plane), thus complementing the line-of-sight information one can obtain from the one-way GNSS carrier-phase and code measurements. The main advantage of differential same-beam VLBI lies in the differencing out of common errors over a narrow beam-width angle. However, if the differential measurement is performed on only a single frequency, the total phase delay is biased by an integer ambiguity, thus ambiguity resolution is required (with very accurate code measurements in that case). A VLBI session would always need to be scheduled in such a way that several GNSS satellites are visible in the same beam-width in the vicinity of a selected quasar seen from two different VLBI stations. Differential same-beam interferometry between GNSS satellites in close proximity to one another is, by its very nature, a double-difference approach.

First attempts have already been made to observe GNSS satellites using VLBI (Kodet et al. 2013) by observing GLONASS satellites from the Wettzell and Onsala VLBI stations with an open-loop receiver. The receiver of the Wettzell 20 m VLBI antenna has been modified to measure the GNSS L_1 signal without changing the local ties (Kodet et al. 2013). It is very important to mention that with the VLBI technique it is possible to determine GNSS satellite orbits without using any other GNSS or SLR measurements. The noise level of the positions of 3414 S-/X-band radio sources listed in ICRF2 (295 defining sources) is in the order of ≈ 40 μas with an axis stability of ≈ 10 μas (Gordon et al. 2010). Translated to GNSS altitudes, this gives a position precision of 5 mm RMS for these 3414 S-/X-band radio sources. Very small steerable antennae will be required to observe the orbits of GNSS satellites against quasars using higher frequencies (Ka-band or W-band). Such a configuration could be improved by making use of a fixed phased-array antennae with beam forming on receive. The phased-array technique would improve the gain of the antenna and it could track all in-view GNSS satellites and refer them to the common measurement epoch. Table 16.2 lists the advantages and disadvantages of VLBI at higher frequencies. First attempts to define a celestial frame at 32 GHz have already been made, mainly driven by the radio-science objectives of the most recent space missions. It is expected that the next realization of the ICRS will include radio-sources observed at higher frequencies, in at least the Ka-band (32 GHz).
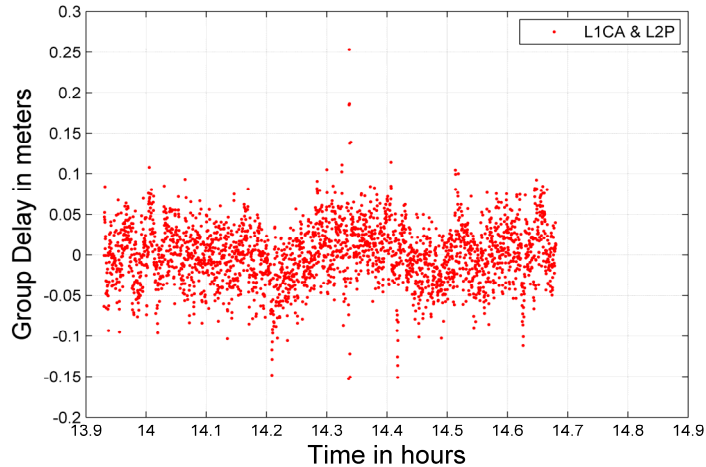
Figure 16.4 GPS signal observed with an L-band receiver connected to a 25 m antenna. Difference between ionosphere-free code (C/A on L_1) and ionosphere-free carrier-phase measurements (Svehla et al. 2010a) converted to the noise level of the original single-frequency measurements. The code measurements have a precision of about 6 mm.

Advantages

- Main drivers are new space missions (Mars Reconnaissance Orbiter, BepiColombo, JUICE, Netlander…) requiring higher telemetry rates, radio-science, improved deep-space navigation (gravity field), lower ionosphere/solar plasma-effects, etc.
- Higher telemetry data rates in deep space
- Onboard RF systems are smaller (antenna) and lighter
- Avoidance of RF interference in S-band
- Ionospheric & solar plasma effects decreased by $16 - 100$ times at 32 GHz/90 GHz compared to 8 GHz
- Observations possible closer to the Sun/galactic center
- Very compact sources (spatial distribution of flux) that give more stability in position over time
- Compared to ICRF2 frame defined in S/X-band, positions in Ka-band are closer to optical positions (GAIA)

Disadvantages

- More weather-sensitive (close to the 22 GHz water vapor line)
- Antenna pointing requirements 4-10 times higher at 32 GHz/90 GHz than at 8 GHz (beam forming technique)
- In order to increase sensitivity, sampling rate needs to be $4 - 10$ times higher at 32 GHz/90 GHz compared to 8 GHz
- Currently no celestial frame in the W-band, first realization of celestial frame at 32 GHz

Table 16.2 Advantages and disadvantages of VLBI at higher frequencies (Ka/W-band) that could allow the use of smaller antennae for combined GNSS/VLBI tracking using a phased-array antenna design.

16.7 Can LAGEOS or Lunar Retro-Reflectors be Observed by VLBI?

Radars are used to detect and track objects in space with metre-grade ranging, (Joint Space Operations Center (JSpOC) and NORAD in the USA and TIRA (Tracking and Imaging Radar) of the Fraunhofer Institute in Europe). JSpOC tracks more than 16 000 objects and uses infrared sensors to detect the re-entry of satellites.

If the ground radar sends microwave signals towards a LAGEOS satellite, this signal will be reflected and scattered by the surface of the satellite into all directions and a tiny part of the wave's energy will be directed towards the ground VLBI antennae, see Figure 16.5.

How the microwaves transmitted by the ground radar scatter on the surface of the satellite depends on their wavelength and the shape of the satellite. If the wavelength of the microwave signal is smaller than the size of the satellite, the wave will be reflected in a specular way similar to light. However, due to diffraction, divergence of the reflected signal will allow tracking the same signal by the VLBI antennae at different locations. It is assumed that such a wide-band microwave signal could be tracked by the open-loop receiver or similar techniques used for tracking extragalactic sources and inter-planetary satellites.

Making use of such a bi-static VLBI concept, the LAGEOS orbit could be determined by SLR and microwave VLBI and tied against the positions of extragalactic radio sources. The same principle could be applied to lunar laser retro-reflectors. So-called persistent scatterers, as they are known in SAR interferometry, are objects that reflect radar well, e.g., metallic structures, buildings etc. If the radar is directed towards the laser retro-reflector on the Moon, the diffraction pattern from the laser retro-reflectors will be different to that from the surrounding lunar surface. Thus one could correlate VLBI signals observed by several VLBI antennae on Earth. Potentially, this could be extended to all 5 retro-reflectors on the Moon in order to monitor lunar orientation (librations). The proposed bi-static concept could potentially open up new applications of VLBI in combining geometric and dynamic frames. Here we only outline the idea and perform no simulations.

Another possible approach is to use a principle of photoconductive antennae, where a passive detector (e.g., on lunar/Mars surface), after being illuminated by SLR or a ground radar, transmits a wide-band microwave signal observed by several VLBI antennae on Earth. This could be called planetary VLBI.

LAGEOS or lunar laser retro-reflector

VLBI antenna

Radar or wide-band microwave signal

VLBI antenna
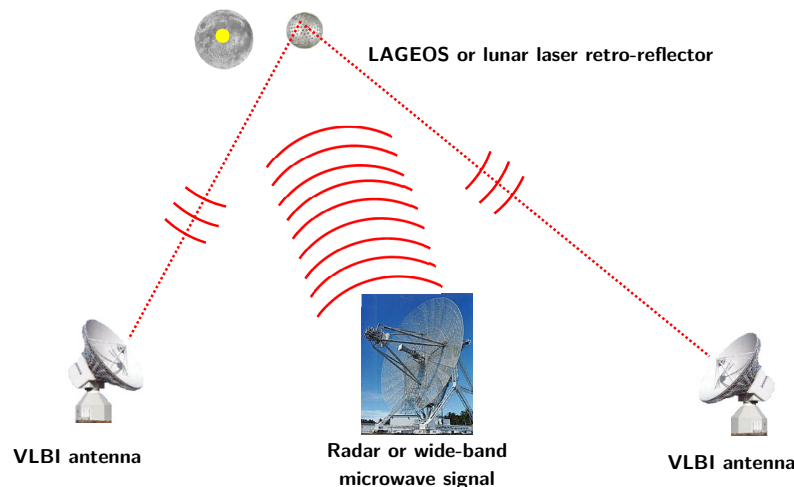
Figure 16.5 A possible bi-static concept of VLBI with a LAGEOS satellite or lunar laser retro-reflectors. A radar microwave signal is transmitted towards the LAGEOS satellite and after reflection/scattering by the satellite, is tracked by the VLBI radio-telescopes on Earth. In the case of laser retro-reflectors on the Moon, the diffraction pattern from the retro-reflectors will be different to that from the surrounding surface, thus one could possibly correlate wide-band signals received by different VLBI antennae. Another approach is to use the principle of photoconductivity, where a passive detector (e.g., on lunar/Mars surface), after being illuminated by SLR or a ground-based radar, transmits a wide-band signal tracked by several VLBI antennae.

17. The SLR/LLR Double-Difference Baseline

Here we present a novel SLR double-difference approach with GNSS satellites. It is shown how forming double-differences of SLR measurements between Herstmonceux (HERL) and Graz (GRZL) ILRS stations and two Galileo satellites removes common SLR biases: i.e., ILRS station range biases and common retro-reflector effects. By using the orbits of GNSS satellites from IGS as fixed in the parameter estimation, the double-difference SLR approach offers a bias-free estimation of relative coordinates with the mm-accuracy between two ILRS stations (SLR baseline) that are separated by about 5000 km. In this way, we obtain SLR observables of utmost precision and accuracy at sub-millimeter level with the standard deviation $\sigma = 0.5 - 1.0$ mm . We show that after differencing the remaining noise in the SLR measurements nicely averages out, leading to estimation of station coordinates, local ties between different space geodesy techniques and precise comparison of optical/microwave tropospheric effects. Considering that relative station coordinates between ILRS stations can be estimated in a similar way between collocated GNSS stations using the GNSS double-differences, the SLR approach allows direct estimation of local ties between SLR and GNSS ground stations. We extend the common-view SLR and make double-differences over time by considering the different observation times for all SLR measurements between all SLR stations. SLR range biases and small biases between SLR sessions are removed. The scale is preserved when double-differencing SLR and free of range biases (at mm-level), making this approach very attractive to combine ILRS network with IGS network in the global GNSS solution. We show that LLR offers estimation of UT0 and with differential SLR the global GNSS can estimate a complete terrestrial frame. For the un-differenced SLR we refer to (Pearlman et al. 2002).

When a LEO satellite is observed by two SLR stations quasi-simultaneously with a GNSS satellite, one can calculate the "vertical SLR baseline" (vector) between the GNSS and the LAGEOS (LEO) satellite as well as the "vertical SLR range" (GNSS-LEO range) derived from geometry. This provides radial orbit information that can be used for altimetry and gravity field missions as well as reference frame satellites. At the end we extend the double-difference approach to other space geodesy techniques such as lunar laser ranging, VLBI and DORIS and discuss estimation of local ties and global reference frame parameters. We also derive a relationship between a possible bias in LAGEOS center of mass correction and radial bias in GNSS orbits. At the end we extend the concept of SLR double-differencing to lunar laser ranging (LLR) and present first results for the LLR double-difference baseline. We succeeded in processing LLR measurements to Apollo and Luna retro-reflectors on the Moon, and, in a similar way, have processed SLR measurements to GPS satellites considering only the geocentric frame in order to model the uplink and downlink for lunar laser ranges.

17.1 SLR Double-Differences – Over Time and Common-view

Double-differences are widely used in the processing of GPS measurements, forming so-called GPS baselines, or vectors between ground GPS stations. In the case of common-view double-difference SLR, the approach is very much the same, we need SLR ranges or SLR normal points given at the same (common) epoch t from
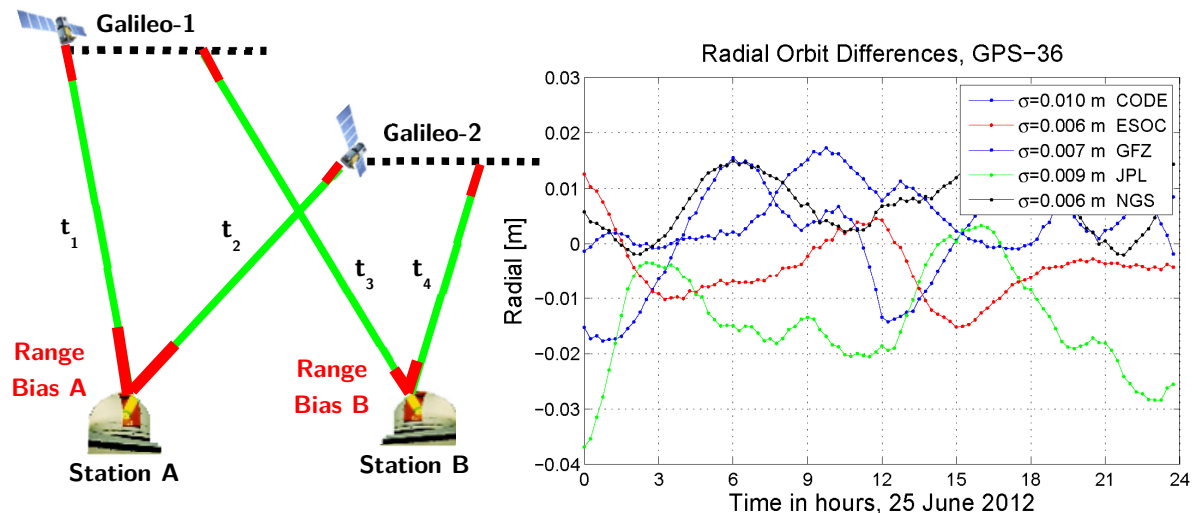
Figure 17.1 Figure on the left shows the general case, i.e. the SLR double-differences over time (green) with four different observation epochs observing two Galileo satellites from two ILRS stations with SLR range biases (red). Figure on the right shows radial orbit differences between different IGS solutions and the IGS Final Orbit. The high level of orbit precision for GPS satellites enables interpolation of SLR normal points to the common epoch to form common-view SLR double-differences. Range- and satellite-biases are removed.

two stations (one station as reference), see Figure 17.1 (left). In that case, the SLR single-difference $SD^1_{A,B}$ between the ranges d of stations A and B to a satellite $''1''$ for a given common epoch t can be defined as

$$SD^1_{A,B}(t) := d^1_B(t) - d^1_A(t) \tag{17.1}$$

If we observe quasi-simultaneously a second satellite $''2''$ from both stations, similar to GPS, we can define the common-view SLR double-difference or the common-view SLR baseline as

$$DD^{1,2}_{A,B}(t) := \left[d^2_B(t) - d^2_A(t) \right] - \left[d^1_B(t) - d^1_A(t) \right] \tag{17.2}$$

as originally proposed in (Svehla et al. 2012) and later extended with more measurements and discussed in more detail in (Svehla et al. 2013a; Svehla et al. 2014; Svehla et al. 2015b). From (17.1), we see that by forming SLR single-differences between two ILRS stations and a common GNSS satellite, common orbit errors are removed as well as common retro-reflector effects. According to (17.2), common-view SLR double-differences remove range-biases and station-specific effects such as common troposphere effects between the two ILRS stations and the same two GNSS satellites. In general case, that is more appropriate when ILRS network is processed with IGS network, we can define the SLR double-differences over time, considering different observation time for all ILRS stations. SLR range biases and small biases between SLR sessions are removed.

In the case of GPS, measurements are taken at integer seconds of receiver time that is synchronized to the global GPS time scale, and the navigation solution is calculated internally by the GPS receiver. Hence, GPS double-differences can be formed between any two stations in the world that have two GPS satellites in common-view. The velocity of the observed satellite drives the synchronization accuracy required to form common-view double-differences. A synchronization error of 0.1 μs will lead to an orbit error of 0.4 mm in the case of GNSS satellites or 0.8 mm in the case of LEOs. In order to form common-view SLR double-differences with an accuracy better than 0.4 mm RMS, SLR measurements between two stations need to be synchronized (e.g., to GPS time) with an accuracy of about 50 ns RMS that corresponds to a GNSS orbit-induced error of 0.2 mm. The typical accuracy of a GPS receiver clock parameter estimated by the navigation solution in a GPS receiver is in the order of 10 ns RMS. The required level of synchronization for the SLR double-difference approach is already provided by GPS and available at ILRS stations.

The easiest way to form common-view double-differences of SLR measurements between two ILRS stations is to generate SLR normal points at the common epoch for both stations. Since this is currently not done, SLR normal points need to be interpolated using the epoch of one of the stations in the pair as a reference. Figure 17.1 shows the orbit differences in the radial direction for the GPS-36 satellite between different IGS Analysis Centres and the IGS Final Orbit. One can see that for the best orbit solutions, the radial orbit error is always under 1 cm, which corresponds to a standard deviation of about $\sigma = 3$ mm . If we look at the first derivative, we have a slope in the radial orbit error in the order of up to 1 cm/3 h (vs. IGS Final Orbit). Therefore, if we assume the interleaving time between the SLR observations of two GNSS satellites to be, e.g., 10 min, we have a systematic error of about 0.5 mm. However, it should be noted that any mean in this interpolation is removed by single-differencing to the same satellite, thus a standard deviation below 0.1 mm is more realistic considering also that IGS Final Orbits should be more accurate than any of the individual orbit solutions. For some ILRS stations (e.g., Herstmonceux and Graz), the interleaving time between GNSS satellites can be reduced to 30-60 s, thus several GNSS satellites could be observed simultaneously in the same session. This analysis shows that SLR single-differences do not remove the interpolation error of SLR normal points. However, the use of the precise orbit keeps this orbit error below the precision of the SLR normal points even for longer interleaving time intervals.

SLR range biases are not eliminated by forming single-differences between two stations, thus the single-difference to another satellite in common-view is needed. When orbit and range biases are removed by double-differencing, SLR with sub-millimeter precision is feasible, and is mainly limited by the station-internal noise. The use of zero-signature retro-reflectors, kHz-ranging systems for GNSS arrays, and the use of enhanced troposphere modeling have produced a degree of precision in SLR that is heading towards the sub-millimetre level. Both GNSS satellites need to be observed quasi-simultaneously (within some e.g., 10-30 min) so that SLR residuals from two stations can be interpolated to the common epoch. Although double-differencing increases the noise level by a factor of 2 w.r.t. the original SLR observable, all session-based systematic effects are at much higher levels and are removed, thus paving the way for sub-millimeter SLR.

We call this approach geometrical because, for the separation of ground SLR stations up to a distance of about 1000-5000 km, GNSS orbit errors of 1 cm RMS do not have a significant impact on the SLR double-difference baseline, or they are significantly reduced to below 1.7 mm. This topic is further discussed in this section. Therefore, relative station coordinates can be determined using double-difference SLR without further improving the orbits of the target satellites that need to be at higher altitudes (GNSS). In this way, this approach is similar to geometrical VLBI, where relative station coordinates are estimated. In our view, when LAGEOS and Etalon satellites are observed by SLR for reference frame realization, any orbit error or deficiency in the orbit modeling, such as e.g., solar radiation pressure and other effects, propagates directly into the estimated station coordinates. In addition, SLR measurements are very sparse in nature: the orbits of SLR satellites used for reference frame realization are not observed continuously, as is the case with GNSS. In this way, the quality of the satellite orbit determination that is based on SLR measurements has a significant impact on reference frame realization and the averaging process is essential for the estimation of high-precision station coordinates over a long period of time. In our view, all these deficiencies of the classical SLR approach can be avoided by making use of double-differences with satellites in high Earth orbit, e.g., GNSS. In this case, ground stations are within half of the max. nadir angle, i.e., $12° - 14°$, as seen from the GNSS satellite.

Differencing of SLR measurements was considered back in the 80ies, where the use of simultaneously (from two stations) observed range differences to LAGEOS satellites was investigated. For more details see (Pavlis 1985) or (Dedes and Mueller 1989). However, the LAGEOS orbit is too low for the common-view double-differencing that is used in this paper with GNSS satellites.

17.2 Biases in SLR Measurements

In order to assess the size of the biases in the SLR range measurements, we compared SLR residuals of the Galileo E11 satellite from different ILRS stations against Galileo residual clock parameters calculated by removing a daily time offset and time drift from the estimated clock parameters. Due to the high altitude of the Galileo satellite orbits, any radial orbit error is compensated by the estimated clock parameter in the orbit determination. Therefore, when a linear model is removed from the estimated Galileo clock parameters, residual clock parameters map radial orbit errors along the orbit with an opposite sign relative to the SLR residuals. Figure 17.2 shows a very good agreement between SLR residuals and residual clock parameters for the Galileo E11 satellite. We chose a period of 30 days (95-125/2013) with high Sun elevation angle above the orbital plane ($\beta = 60° - 67°$) in order to avoid a large impact of solar radiation pressure effects in the radial orbit direction. Galileo E11 clock parameters were corrected only for the periodic relativistic correction due to J_2 gravity field coefficient, following (Kouba 2004). The standard deviation of the calculated residual clock parameters is 20.7 mm, whereas SLR residuals show a higher standard deviation of 25.3 mm. For this analysis we used the Galileo orbit/clock solution submitted to MGEX Campaign of IGS by the Astronomical Institute of the University of Bern (AIUB). For more on the MGEX Campaign we refer to (Steigenberger et al. 2014). We have carried out very realistic simulations of Galileo H-masers based on ground test results, and it can be shown that the standard deviation of simulated Galileo residual clock parameters is at the level of 15.5 mm for a period of 24 h and about 7 mm for a period of half the orbit revolution. In this simulation we also considered all onboard environmental effects such as variations due to temperature and magnetic field along the Galileo orbit, for more information see (Svehla et al. 2015a) and (Svehla et al. 2016). Thus the standard deviation of the corresponding Galileo radial orbit error should be at the level of about 14 mm for the selected period of 30 days. That is about a factor of 2 smaller than the standard deviation of the SLR residuals of 25.3 mm. From this, we can draw a conclusion that space/ground local ties as well as biases in some of the SLR ranges and in the reference frame (e.g., geocenter), prevent the maximum exploitation of SLR normal points
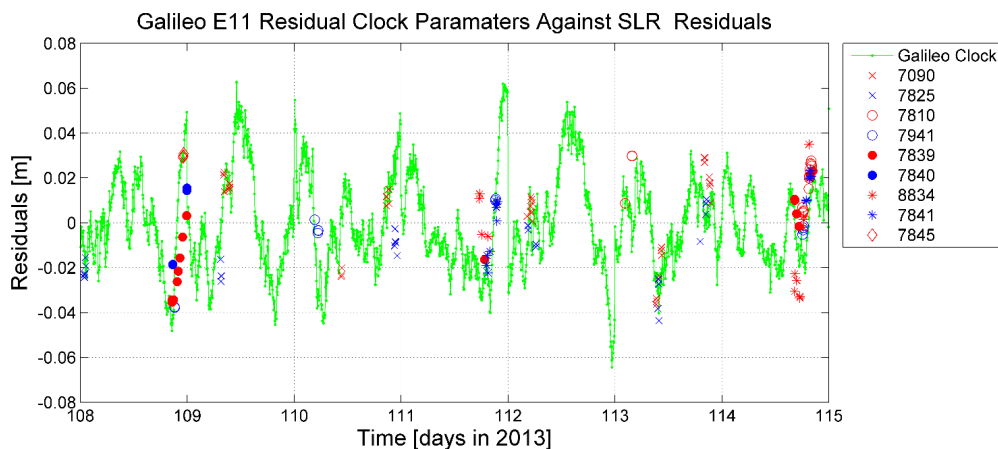
Figure 17.2 Residual clock parameters of Galileo E11 satellite against SLR residuals for a 7-day subset of the 30-day analysis period. From the Galileo E11 clock parameters (MGEX-AIUB) a daily time drift and time offset was removed. The standard deviation of residual clock parameters is 20.7 mm for days 95-125/2013, whereas noise contribution of the Galileo H-maser is about 15.5 mm over a 24 h period (based on simulated data of Galileo H-masers using ground test results). This leads to radial Galileo orbit error at the level of ≈ 14 mm. In comparison with clock parameters, SLR residuals show a higher standard deviation of 25.3 mm and this factor of ≈ 2 is most likely due to space/ground local ties, biases in some of the SLR ranges and in the reference frame (e.g., geocenter). The SLR residuals were calculated using the Bernese GNSS Software at TU München and provided by AIUB, but this should introduce no inconsistencies.
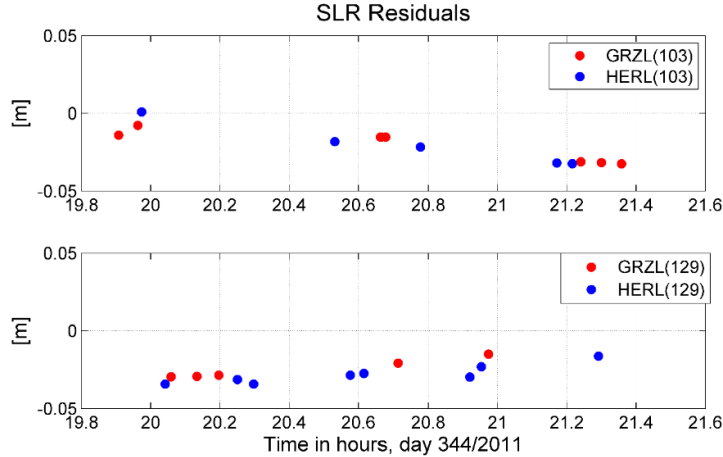
Figure 17.3 SLR residuals of the GLONASS 103 (top) and 129 (bottom) satellites from the Graz (GRZL) and Herstmonceux (HERL) ILRS station. SLR residuals based on the orbit solution from CODE IGS Center.

that show standard deviations at much lower noise levels, going down to some 0.2 mm, which is actually the limit for the best SLR ground stations. On the other hand, orbit predictions operationally provided for the first Galileo satellites are currently not of high accuracy compared to those for other GNSS satellites, thus they are not easy targets for the ground ILRS stations. As a consequence, any change in the ranging gate at ground stations will also result in session-specific SLR range biases. Figure 17.3 shows SLR measurements from Herstmonceux and Graz ILRS stations taken to the GLONASS 103 and GLONASS 129 satellites. One can see clear common orbit errors of the order of 1-5 cm in the SLR residuals from both stations. The second interesting feature is the long-periodic systematic effects spread over several tracking passes observed by both stations at the mm-level. This effect is either caused by the orbit dynamics, satellite reflector signature, troposphere modeling or time-varying station effects, e.g., unmodelled tidal effects or atmosphere loading. The third interesting feature to note is the small relative range biases, in the order of about 3.4 mm between the two stations. These that are similar in size (except for the first normal point), but there is a clear difference between the consecutive tracking passes measured at the two stations. For LAGEOS satellites, used for reference frame determination, there exists a 7 mm difference between the CoM corrections to be applied to stations GRZL and HERL, as recommended by the ILRS and applied by the analysis centres (Otsubo and Appleby 2003). It is interesting to note that SLR residuals to Galileo E11 in Figure 17.2 also show similar SLR range bias between the two stations with the same sign. Tentatively, we suggest that loading effects of a few millimeters may contribute, and these will be investigated in the future.

The question remains as to whether there are any signature effects due to the SLR array or to the variable angle of incidence. The SLR arrays are flat, thus the only systematic effect introduced into the range measurement will be via the angle of incidence, (Otsubo et al. 2001). However, GRZL and HERL ILRS stations only receive single photons (due to 0.4 mJ pulses). With single photons, the mean reflection point is very close to the center of the SLR array, and it will remain there, regardless of the angle of incidence. Hence there is no systematic range error from the ″array signature″. Small variations, as induced by variations in the far field diffraction pattern due to non-perfect prisms, are at the mm-level, and should not appear here. Variations in the angle of incidence will only have an effects on the RMS of the measurements: min. at 90° angle of incidence, and max. at lower elevations/angles of incidence.

The first normal point to GLONASS 103 in Figure 17.2 has a slightly different range bias, thus differencing could be used in the SLR data pre-processing for screening and calibrating SLR normal points. This is more visible for the Galileo satellites in Figure 17.4, where the second tracking pass (after 24 h) shows a small bias for both satellites compared to the tracking pass 24 h before. Orbits for both Galileo satellites in Figure 17.4 were generated as two independent daily 24-h arcs for both days and SLR measurements were most likely taken during the same session. Thus, apparently we could have two independent tracking passes.
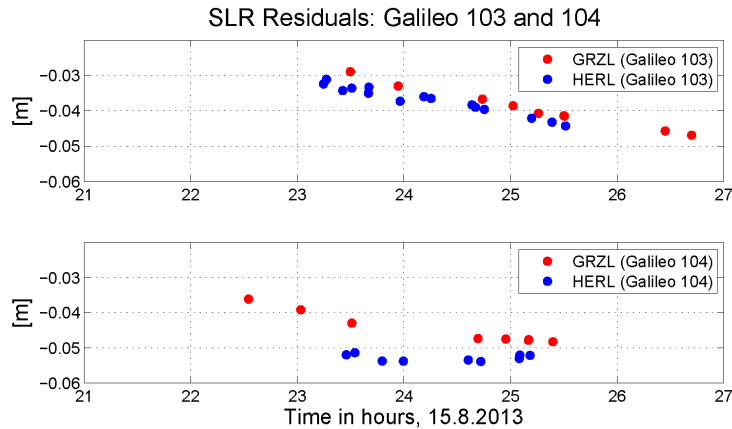
Figure 17.4 SLR residuals of Galileo 103 (top) and 104 (bottom) satellites from Graz (GRZL) and Herstmonceux (HERL). SLR residuals based on the orbit solution from MGEX IGS Campaign (AIUB).

17.3 The First SLR Double-Difference Baseline and the Local Tie

In order to form SLR normal points at common epochs for the Galileo 103 and Galileo 104 satellites in Figure 17.4, a linear model (first order polynomial) was fitted to the normal points of the GRZL station separately for both tracking passes. In this way, the SLR normal points of the GRZL station were interpolated to epochs of the normal points of the HERL station, separately for each satellite and tracking pass.

Figure 17.5 shows single-difference SLR measurements for both Galileo satellites. One can clearly see that residuals are grouped for each tracking pass, whereas differences between the two satellites within the tracking pass are very small (mm-level). Single differences cannot remove station-specific range biases, and this explains why SLR differences to both Galileo satellites show the same bias. This bias is removed by forming double-differences in Figure 17.6. However, here we did not use a linear model (first order polynomial) to interpolate SLR normal points, we merely calculated a mean SLR bias for the single-differences of the Galileo 104 satellite. At this level, residuals shows random nature and it is difficult to model any trend using a linear model.
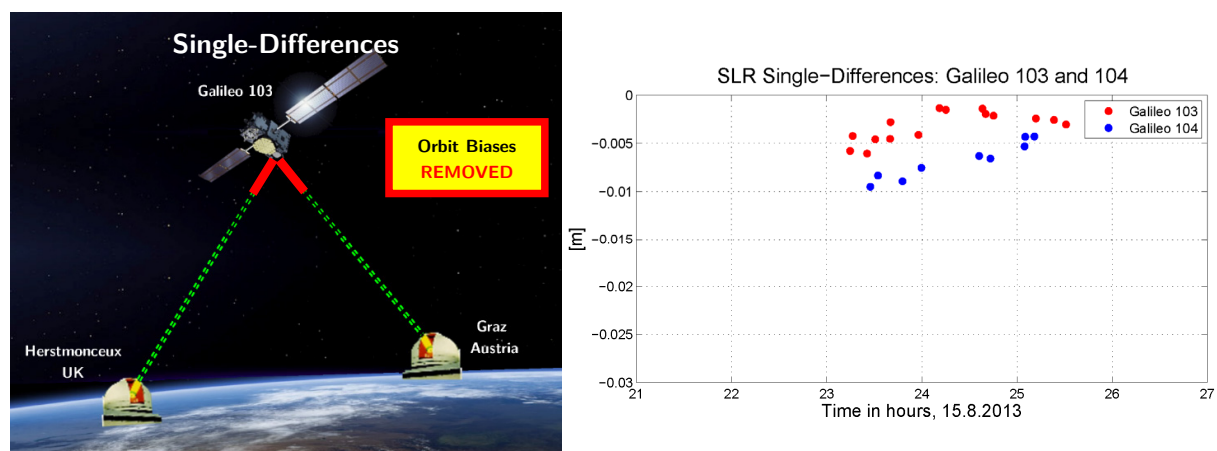
Figure 17.5 Concept of SLR single-differences (left) and the first SLR single-differences (right) to the Galileo 103 and Galileo 104 satellites using SLR measurements from HERL and GRZL stations. Orbit errors in the original SLR measurements are removed, since the single-difference residuals are very similar for both Galileo satellites. The remaining biases reflect range biases between the two stations.
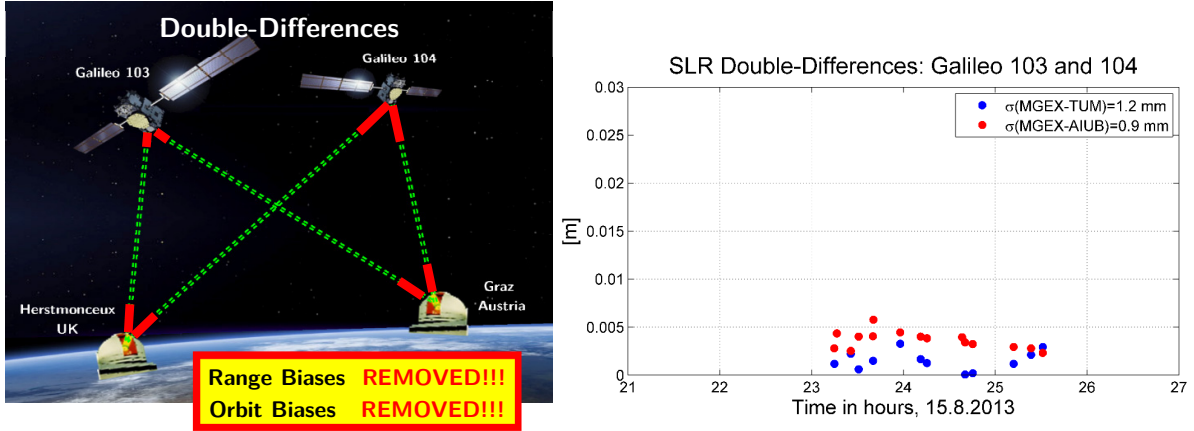
Figure 17.6 Concept of SLR double-difference (left) and the first SLR double-differences (right) between the Galileo 103 and Galileo 104 satellites using SLR measurements from HERL and GRZL. Orbit and range biases in the original SLR measurements are removed, giving a SLR baseline with a standard deviation of only 1.2 mm for MGEX-TUM orbits and 0.9 mm for MGEX-AIUB orbits. Results were confirmed with more data from different days and different ITRS (Svehla et al. 2015b).

Least-squares prediction with an empirical covariance function would probably be more suitable for interpolating single-difference normal points. After subtracting the single-differences of the Galileo 103 and Galileo 104 satellites in Figure 17.5 we obtained the double-difference SLR residuals shown in Figure 17.6. Figure 17.6 nicely shows that all orbit and SLR range biases are removed by forming double-differences of SLR measurements. The standard deviation of double-difference residuals is $\sigma = 1.2$ mm for MGEX-TUM orbits and $\sigma = 0.9$ mm for MGEX-AIUB orbits. Since by forming double-differences the noise is increased by a factor of 2, the noise level of the original SLR normal points is about $\sigma = 0.5$ mm. A small bias in the remaining SLR double-differences for both orbits in Figure 17.6 indicates remaining systematic effects that were not removed by differencing. However, the scale of the SLR measurements is preserved by differencing and should be free of biases in the case of double-differences. Table 17.1 shows estimated coordinates of the HERL station w.r.t. the ITRF2008 terrestrial reference frame using the SLR double-difference baselines from GRZL and HERL stations. The a posteriori sigma of unit weight from the least-squares adjustment is $\sigma_0 = \pm 0.7$ mm. Figure 17.6 is the first demonstration of sub-millimeter differential ranging from Earth to space, to the Galileo satellites with an orbit altitude of 23 222 km, see Figure 17.7. In our view, the SLR double-difference approach

Estimated coordinates of HERL in (mm) using SLR double-differences from GRZL	
MGEX-TUM Orbits	MGEX-AIUB Orbits
$\sigma_0 = \pm 1.2$ mm	$\sigma_0 = \pm 0.7$ mm
N=1.2±2.5	N=4.2±7.2
E=2.1±0.8	E=8.1±2.4

Given Local Tie   GRAZ(SLR-GPS): ΔX=2.5580   ΔY=-8.5160   ΔZ=1.3210
Baseline   GRAZ-HERS (SLR-GPS): ΔX=-9.0427   ΔY=-1.7543   ΔZ=2.6192
Estimated Local Tie HERS(SLR-GPS): ΔX= -6.4847   ΔY=-10.2703 ΔZ=3.9402
ITRF2014 Local Tie HERS(SLR-GPS): ΔX= -6.4868   ΔY=-10.2700 ΔZ=3.9487

Difference Local Tie (Measured − ITRF2014): **N=-0.007   E=0.0005**

Table 17.1 Left: estimated horizontal coordinates (N-North, E-East) of the HERL station w.r.t. ITRF2008 terrestrial reference frame using the SLR double-difference baselines from GRZL based on only 15 double-difference normal points. Both solutions, based on the MGEX-TUM and MGEX-AIUB orbits for Galileo 103 and Galileo 104 provide similar results with an accuracy of several millimeters. If all three local coordinates are estimated, accuracy is at the cm-level based on only 15 normal points and two GNSS satellites observed over 2 h. Right: estimated local tie in ITRF2014 between SLR and GPS at HERL station based on collected double-difference SLR and GPS baseline between GRAZ and HERL. Local tie at GRAZ is from ITRF2014.
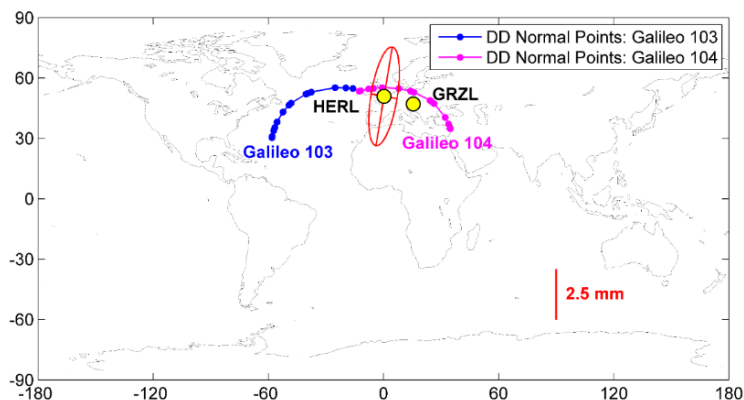
Figure 17.7 Common-view observation of Galileo 103 and Galileo 104 satellites from GRZL and HERL ground ILRS stations. Blue and magenta dots denote 15 SLR normal points on the ground tracks of the Galileo 103 and Galileo 104 used to form the first SLR double-difference baseline. The error-ellipse (red) refers to estimated HERL coordinates with semi-major axes $m_A = 2.5$ mm and $m_B = 0.7$ mm (based on MGEX-TUM orbits).

allows for a precision and accuracy significantly better than any other space geodesy technique (GNSS, VLBI, DORIS, or classical SLR). With sub-millimeter precision and accuracy, this approach recommends itself for a suite of novel applications in geodesy and terrestrial reference frame realization, especially considering effects that could be monitored between SLR stations, such as tidal effects and atmosphere loading. Making use of long SLR baselines, the double-difference SLR approach offers bias-free estimation of all terrestrial reference frame parameters. SLR double-differences are similar to GPS double-differences with fixed carrier-phase ambiguities. However, SLR is much more precise and accurate than GPS considering multipath, antenna phase center effects and other signal propagation effects, such as troposphere and higher-order ionosphere effects. When a pair of GNSS satellites is observed simultaneously using both microwave (GNSS/VLBI) and SLR techniques, one could use this configuration to estimate very accurately local ties by comparing (or subtracting) GNSS and SLR double-difference baselines, see Figure 17.8, showing that there is only one local tie between IGS and ILR networks and the same approach for local ties could be extended to VLBI and DORIS.
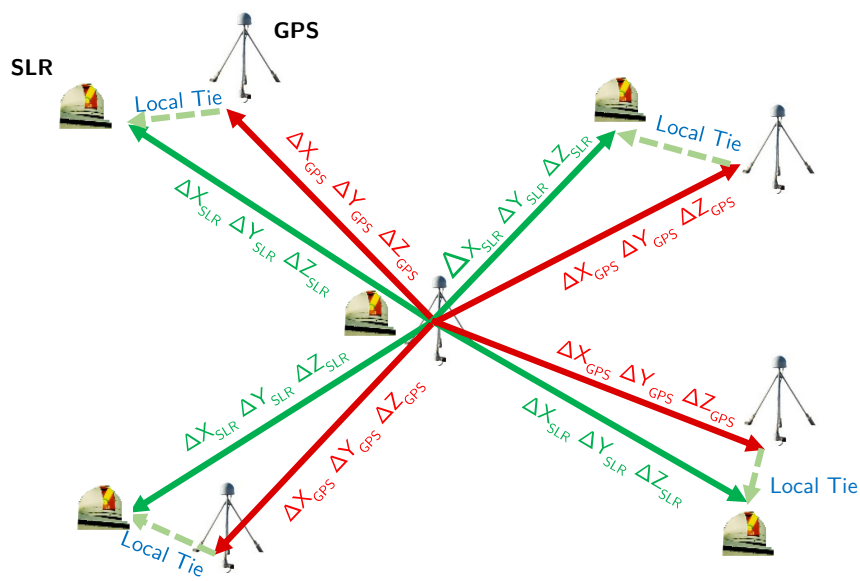
Figure 17.8 An efficient approach to estimate local ties between SLR and GPS ground stations by estimating baselines (relative coordinates) using double-difference SLR and collocated double-difference GPS relative to the reference station. Therefore, there is only one local tie between the global IGS and ILRS networks.

17.4 Sensitivity Analysis of SLR Double-Differences

Let us now try to estimate station coordinates using simulated measurements for an SLR baseline. Table 17.2 shows the ZIML station coordinates estimated relative to WETL based on simulated SLR double-differences with normal points every 5, 10 and 15 min. Before differencing, SLR measurements were simulated with an RMS of 2.2 mm for GPS and 6.0 mm for GLONASS satellites. This RMS corresponds to $\sigma = 3.2$ mm of epoch-wise differences of normal points for GPS-36 and 8.5 mm for epoch-wise differences for GLONASS R07, taken over a period of 7 years from SLR station GRZL, (Thaller et al. 2011). Table 17.2 shows that with just two SLR double-difference passes (based on three GNSS satellites) one can estimate station coordinates at the mm-level. The noise level is a factor of 2-3 higher for GLONASS. When all satellites of the GPS or GLONASS constellations are taken into account over a period of one day (last column in Table 17.2), the precision of the station coordinates is within the sub-millimeter level, assuming white noise only. However, in our case the noise level of the original SLR measurements of $\sigma = 0.5$ mm is 3-4 times smaller. This tells us that with the geometrical SLR double-difference approach station coordinates could be estimated with millimeter precision and accuracy for all three coordinates, as we showed for the first time in (Svehla et al. 2012).

An error in the order of 4-6 cm RMS was introduced to the GNSS orbits for the processing of the simulated SLR baseline. The effect on station coordinates was negligible over such a short SLR baseline, confirming the "rule of thumb" in Eq. (17.3). Eq. (17.3) relates the station vector component error $\delta\rho_{xyz}$ (scale) with an orbit error δr multiplied by the baseline length l and normalized by the orbit altitude R, and is identical to the "rule of thumb" given by (Bauersima 1983) for GNSS

$$\delta\rho_{xyz} = \frac{l}{R}\delta r \tag{17.3}$$

Considering that GNSS orbits can be estimated with an accuracy of about 1 cm RMS, one can see that for baselines of 1000-5000 km the impact of orbit errors on station coordinates is in the order of only 1.7 mm, whereas for a baseline of 1000 km the effect is only 0.3 mm and for a baseline of 500 km only 0.2 mm

$$\begin{aligned}
\delta r(\text{GNSS}) &= 1 \text{ cm} \\
l = 500 \text{ km} &\rightarrow \delta\rho_{xyz} = 0.2 \text{ mm} \\
l = 1000 \text{ km} &\rightarrow \delta\rho_{xyz} = 0.4 \text{ mm} \\
l = 5000 \text{ km} &\rightarrow \delta\rho_{xyz} = 2.2 \text{ mm}
\end{aligned} \tag{17.4}$$

Simulation shows that with just a few double-difference passes one can estimate station coordinates at the mm-level or even at the sub-mm level, whereas for longer SLR baselines it is suggested that IGS Final Orbits are used in order to reduce the impact of the GNSS orbit on the estimated station coordinates.

The SLR double-difference approach is similar to the GPS double-difference approach. Since the ionosphere-free linear combination is used to process GPS measurements, the noise of the calculated GPS double-differences is increased by a factor of 3, in addition to the factor of two resulting from forming double-differences. Compared to this factor of 6 in the increase in noise of the original GPS measurements provided by a geodetic GPS receiver, in the case of SLR, double-differencing increases the noise of SLR normal points by a factor of 2, but the size of range biases in the SLR measurements is significantly higher than the noise of SLR normal points. In addition, range biases and satellite orbit error are removed or significantly reduced in the case of double-differencing. Thus, the SLR double-differences are significantly more accurate than the original SLR measurements.

This is why one could claim that the double-difference SLR approach has the potential to offer a level of precision and accuracy that is significantly better than any other space geodesy technique (GNSS, VLBI, DORIS, or classical SLR). As with GPS, with very long SLR baselines all terrestrial reference frame parameters could be estimated, including station coordinates, geocenter and Earth rotation parameters (ERPs), since following Eq. (17.3) the effect of the orbit error will be linearly scaled in the estimated station coordinates.

Two Double-Differences with 3 GNSS Satellites GPS/GLONASS			Full GNSS Constellation
ZIML Coordinates [mm]	Normal Point every 5 min	Normal Point every 10 min	Normal Point every 15 min	Normal Point every 10 min
Up	-1.4/-3.7	5.4/14.6	-5.7/-15.6	-0.1/-0.3
North	0.3/0.7	-0.7/-2.0	0.1/0.3	0.0/0.0
East	0.2/0.5	0.1/0.2	0.0/-0.1	0.0/0.0

Table 17.2 Sensitivity analysis of simulated SLR double-differences for **GPS**/GLONASS: ZIML station coordinates estimated relative to WETL based on only two SLR tracking passes with three GNSS satellites (left columns) and the full GNSS constellation for **GPS**/GLONASS (last column), day 293/2012.

However, the estimation of ERPs and the geocenter will most likely require the modelling of the satellite orbits, or relative dynamics between the two GNSS satellites involved in double-differencing. Galileo satellites could also be treated as geometrical targets where geocenter and ERP errors are common to all observed satellites and also mapped geometrically in the radial direction by the onboard Galileo H-maser. This issue on combination of Galileo clock information and SLR, is outside the scope here and will be addressed in future work.

17.5 How to Observe Four GNSS Constellations with SLR

Figure 17.9 shows the first common-view SLR ranging to the Galileo constellation from three ILRS stations that was used to form the first SLR double-difference baseline. The complete Galileo and Beidou constellations as well as GLONASS and future GPS satellites equipped with SLR retro-reflectors will provide about 35 SLR targets above 10° elevation. With three GNSS constellations this global mean number of SLR targets over all latitudes and longitudes is about 26. Given that the SLR double-difference approach may allow precision and accuracy that is much better than any other space geodesy technique, it is assumed that SLR telescopes will be improved in the future, enabling wide-angle SLR ranging, see Figure 17.10. Beam steering within the optical telescope has been developed in optical communication providing wide-angle tracking in space (up to some 120° without a loss in energy). With a telescope that does not move during one SLR session, very accurate pointing could be achieved, allowing very fast tracking of all common-view GNSS satellites including all LEO and reference frame SLR satellites.
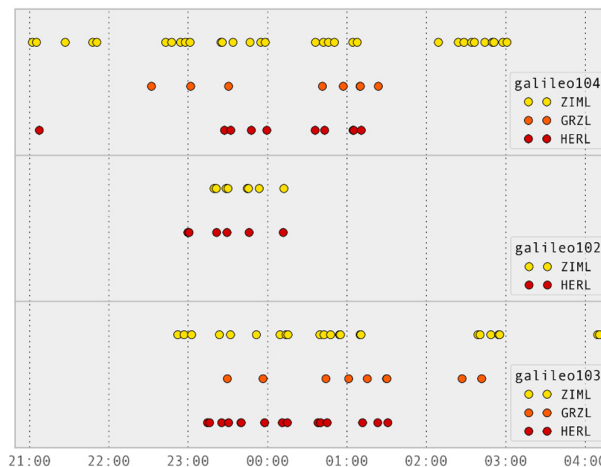
Figure 17.9 First common-view SLR observation of Galileo constellation on Aug 15, 2013.
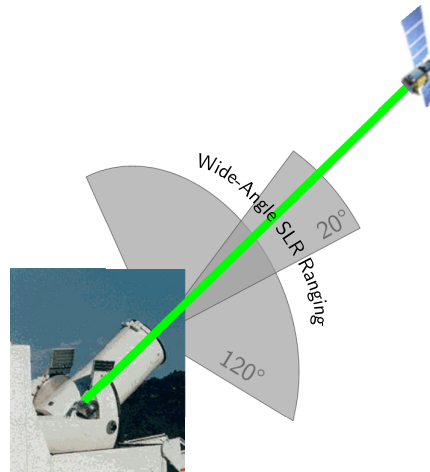
Figure 17.10 Proposed wide-angle SLR ranging to several GNSS with a telescope using beam steering (over e.g., 10° - 60° angle) and does not move during one session.

17.6 Vertical SLR Double-Difference Baseline and Vertical SLR Range Between GNSS and LEO Satellites

For altimetry and gravity field missions, the radial component is the most important orbit component as it is the direction of the main gravity gradient (Rummel et al. 2011) and the direction in which the range to the sea surface topography is measured by satellite altimetry. In addition, we have SLR range biases for measurements to LEO satellites that cannot be directly assessed due to the very low orbit altitude, i.e., no common-view to a LEO satellite (including LAGEOS) from two stations in, e.g., the US and Europe. If we observe a quasi-simultaneously a LEO and a GNSS satellite from two SLR ground stations A and B, see Figure 17.11, one can define the vertical SLR double-difference baseline $DD_{GNSS,LEO}^{AB}$ defined for a common epoch t as
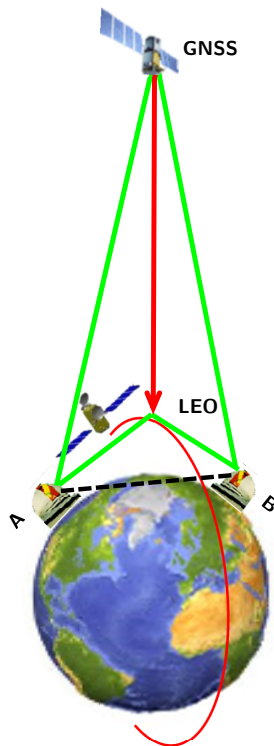
Figure 17.11 Quasi-simultaneous SLR tracking of a LEO and a GNSS satellite from two ground SLR stations (in green). Vertical double-difference SLR baseline (red) between a GNSS and a LEO satellite w.r.t. two ground SLR stations.

$$DD_{GNSS,LEO}^{AB}(t) := \left[d_{LEO}^{B}(t) - d_{GNSS}^{B}(t) \right] - \left[d_{LEO}^{A}(t) - d_{GNSS}^{A}(t) \right] \qquad (17.5)$$

with SLR ranges to the LEO d_{LEO}^{A} and d_{LEO}^{B}, and to the GNSS satellite d_{GNSS}^{A}, d_{GNSS}^{B}. In this case, the orbit of a LEO satellite can be defined w.r.t. the GNSS satellite. We call this baseline "vertical", because GNSS and the LEO satellite are observed by SLR at different elevations, the GNSS-LEO baseline itself is always "vertical" or "radial". The advantage of the "vertical SLR baseline" lies in the elimination of SLR range biases and radial GNSS orbit errors, since both stations are within a small angular separation as seen from the GNSS satellite. It is assumed that the station range biases are independent of the satellite altitude, which is not necessarily true for LEO and GNSS (e.g., when the time delay measurement system has non-linear errors). Since SLR measurements to the LEO and the GNSS satellite need to refer to the same observation epoch t, only the LEO satellite needs to be observed simultaneously from both stations. We have seen in the beginning of this section that SLR residuals of GNSS satellites can be interpolated very accurately to the common epoch over an interval of e.g., 10-30 min. In the case of a ground twin-SLR telescope, the LEO and the GNSS satellite could even be observed with the same laser pulse generated for both telescopes in the twin-configuration. The same approach could be applied to the two LAGEOS reference frame satellites in combination with GNSS.

Taking, in addition, a fixed distance between the two SLR stations in Figure 17.11, we can calculate another completely independent observable, what we call the "vertical SLR range" between a GNSS and a LEO satellite. For this, GNSS needs to be observed in approx. the radial direction as seen from a satellite in the lower orbit (e.g., LAGEOS, JASON-2). The "vertical SLR range" and residuals will refer to the radial orbit direction in that case and can be calculated from the geometry of the two "observed" triangles A-LEO-B and A-GNSS-B, even without any LEO/GNSS orbit information.

17.7 Double-Difference Approach in Space Geodesy: SLR/GNSS/VLBI

Figure 17.12 shows the double-difference concept of space geodesy. Figure 17.12 (left) depicts different ways to form SLR double-differences based on satellites in different orbits, such as lunar, MEO and LAGEOS orbit. In all cases SLR satellites are observed quasi-simultaneously against the background GNSS constellation. By forming SLR double-differences, one can combine, with reduced SLR biases, the orbits of GNSS satellites with the ETALON and LAGEOS satellites used for definition of the terrestrial reference frame, as well as Lunar Laser Ranging (LLR). One could also form double-differences between two retro-reflectors on the Moon, considering that the baseline/altitude ratio in (17.3) approaches zero in that case. In a few years from now, when the Galileo and Beidou constellations have been deployed, together with GLONASS we will have three GNSS constellations completely equipped with SLR arrays - more than 70 GNSS satellites in space with SLR reflectors. Currently, only GPS-36 is equipped with SLR reflectors, but future GPS satellites will carry new generation SLR arrays. Figure 17.12 (right) depicts SLR, GNSS and VLBI double-differences with GNSS satellites. In the case of SLR measurements, double-differences can be used to geometrically map SLR reference frame satellites against GNSS constellations, whereas VLBI double-differences can be used to geometrically map the GNSS constellations against the VLBI quasars (extragalactic sources) that define ICRF-2 (International Celestial Reference Frame).

When both GNSS satellites are observed simultaneously using both the microwave (GNSS/VLBI) and SLR techniques, one could use this configuration to estimate very accurately local ties by comparing (or subtracting) GNSS and SLR double-difference measurements (17.6). In (17.6), ρ_{AB}^{jk} denotes the geometry term and $\delta\rho_{AB}^{jk}$ represents tropospheric effects. Thus (17.6) can be used for very precise comparisons of troposphere models and mapping functions between the optical and microwave domains, as well as local ties between different space geodesy techniques

$$DD_{AB}^{jk}(\text{GNSS}) = \rho_{AB}^{jk} + \delta\rho_{AB}^{jk}(\text{TZD}_{microwave})$$

$$DD_{AB}^{jk}(\text{SLR}) = \rho_{AB}^{jk} + \delta\rho_{AB}^{jk}(\text{TZD}_{optical}) + \text{local tie}_{\text{SLR}} \qquad (17.6)$$

$$DD_{AB}^{jk}(\text{VLBI}) = \rho_{AB}^{jk} + \delta\rho_{AB}^{jk}(\text{TZD}_{microwave}) + \text{local tie}_{\text{VLBI}}$$

Following the "Bauersima rule of thumb" (Bauersima 1983), we see that in all three cases (GNSS, SLR and VLBI double-difference baseline), we do not need very accurate GNSS satellite orbits to estimate station coordinates. In all these cases, GNSS satellites could be considered as geometrical targets on the celestial sphere, i.e., similar to quasars in VLBI. From this point of view, the double-difference concept of space geodesy, as outlined in this section, is very much a geometrical technique by its nature, similar to VLBI.

However, in the case of very long double-difference baselines, the estimated vector will be affected by an additional rotation of the GNSS reference frame (common to all baselines in the network). This will not be the case, if this baseline is composed of two shorter baselines, e.g., by adding one or more stations in between. From this we can draw the conclusion that orbits of GNSS satellites could be mapped against the celestial frame (e.g., using the Delta-DOR approach), and thus GNSS satellites could serve geometrically as "moving quasars" on the celestial sphere. Observing these geometrical targets with SLR, GNSS or VLBI double-difference approaches we could estimate all other parameters, not only station coordinates, but also parameters such as Earth rotation/orientation and geocenter coordinates. Since Earth orientation and rotation can be considered as dynamic in nature, especially regarding the parameters used to interpolate normal points to the common epoch, the double-difference concept of space geodesy as outlined in this section is a viable method for observing and combining the geometry and dynamics of reference frames.
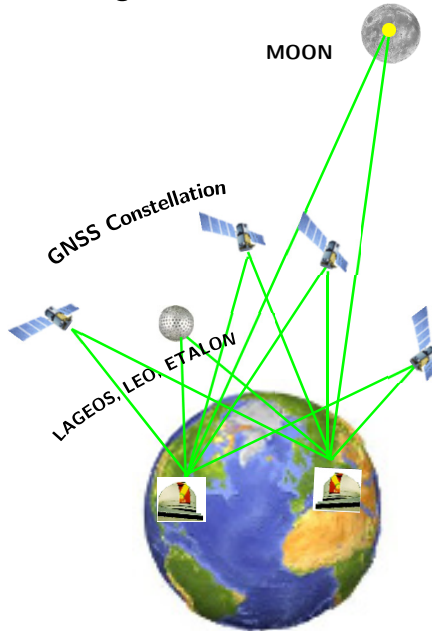
a) SLR Double-Differences against GNSS Constellation   b) VLBI/ SLR/GNSS Double-Differences
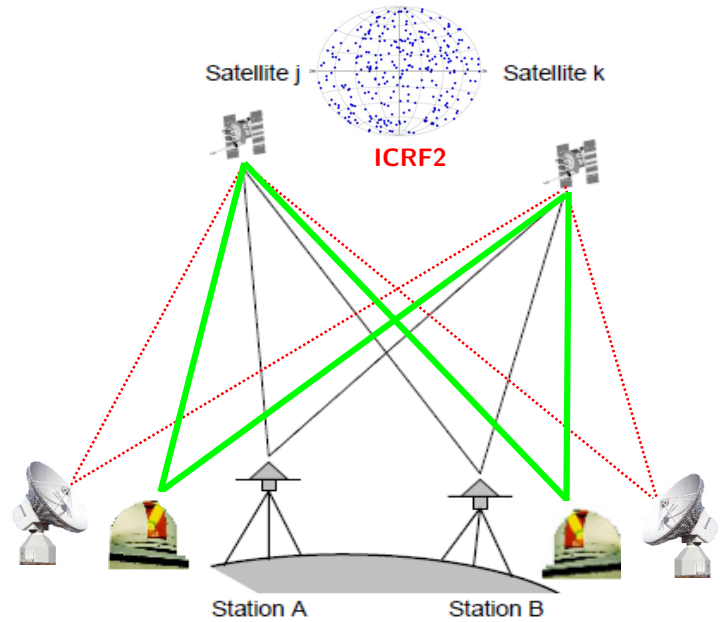
Figure 17.12 Double-difference concept of space geodesy. SLR observation of reference frame satellites (ETALON, LAGEOS) and Moon, quasi-simultaneously with (against) background GNSS constellation (left). SLR, GNSS and VLBI double-differences with GNSS satellites (right). In the case of SLR, an a priori orbit is used whereas in the case of VLBI, the new generation of GNSS clocks will allow interpolation of VLBI measurements to a common epoch (e.g., Galileo Passive H-Maser).

So far we have not referred to DORIS, the fourth space geodesy technique. Tracking of DORIS and GNSS is very similar and we are now seeing the first attempts to upgrade geodetic space GNSS receivers with DORIS tracking, i.e., as proposed for the STE-QUEST mission in highly elliptical orbit for terrestrial and celestial reference frame determination. For this part of the STE-QUEST mission see (Svehla et al. 2013b). DORIS has the potential to complement GNSS with a nadir pointing antenna at higher altitudes. By making use of the phased-array antenna design and beam forming on receive it should be possible to increase the gain of the DORIS antenna and hopefully achieve DORIS tracking at higher altitudes.

17.8 Global Solution with Double-Difference SLR Approach

In order to simulate global SLR baselines, we have chosen 4 globally distributed ILRS stations, see Figure 17.13, with one short SLR baseline between GRZL and HERL, (Svehla et al. 2014). We simulated double-difference SLR measurements as normal points (NPT) every 10 min with two common GPS satellite (denoted by PRN numbers in Figure 17.13) between two ILRS stations in common-view. We used simulation for common-view due to simplicity. The noise level used in the simulation was ± 1 mm for SLR measurements from GODL, GMSL, HARL, and ± 0.5 mm for those from HERL and GRZL. In the next step, for a period of 17 days we estimated daily solutions for global parameters including X- and Y-pole and rates, length-of-day (LOD) and the geocenter coordinates in the Z-direction. The X-pole and Y-pole coordinates with rates and LOD parameters were estimated against the C04 values. In this estimation, we used two independent solutions, keeping the orbits of the GPS satellites fixed and simulating the orbit error using the daily orbit difference between orbits provided by the two IGS AC centers (CODE and ESOC), see Figure 17.14, Figure 17.15 and Figure 17.16. Solution with fixed GPS orbits shows the sensitivity of the measurements to the estimated parameters. When orbit error is introduced for GPS satellites, one can see that by forming double-differences with SLR measurements, the long SLR baselines are still affected by the orbit error and one would need to estimate orbit parameters, or to combine DD-SLR with GPS measurements. Simulation shows that by introducing an orbit error, the noise of all estimated parameters is about 2-3 times higher compared to difference between CODE and ESOC ACs. This is based on only 4 ground ILRS stations.
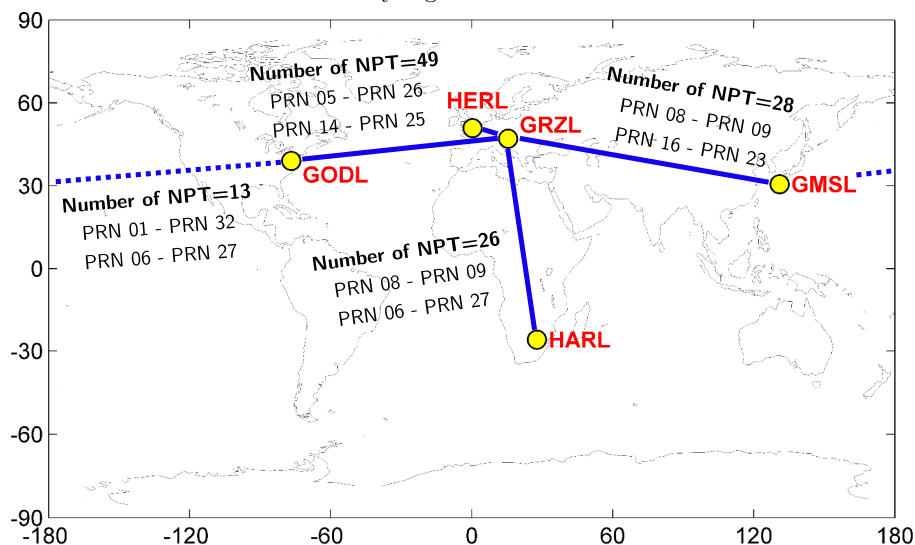
Figure 17.13 Simulation of the global double-difference (DD) SLR approach with 4 ground ILRS stations (GRZL, HARL, GSL and GODL). To form the baseline we 2 GPS satellites observed simultaneously from two ILRS stations in common-view (blue). The ″Number of NPT″ shows the daily number of DD-normal points.

a) Without orbit error
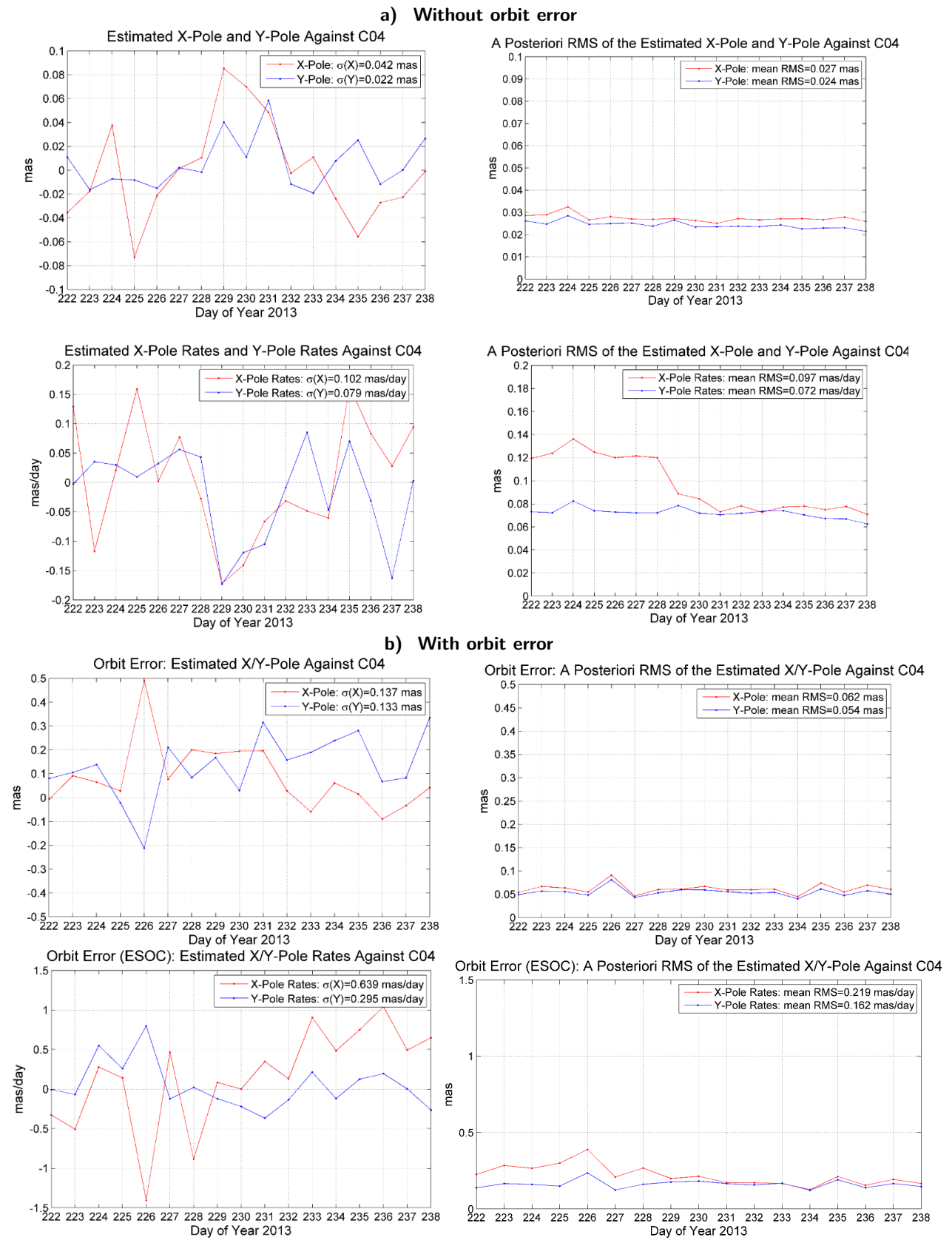
b) With orbit error

Figure 17.14 The X-Pole and Y-pole coordinates and rates estimated with a posteriori RMS values against the C04 values using only 4 global ILRS stations. The solution a) refers to the fixed orbits of GPS satellites, whereas in b) the orbit error was simulated using the daily orbit difference between two IGS AC centers.
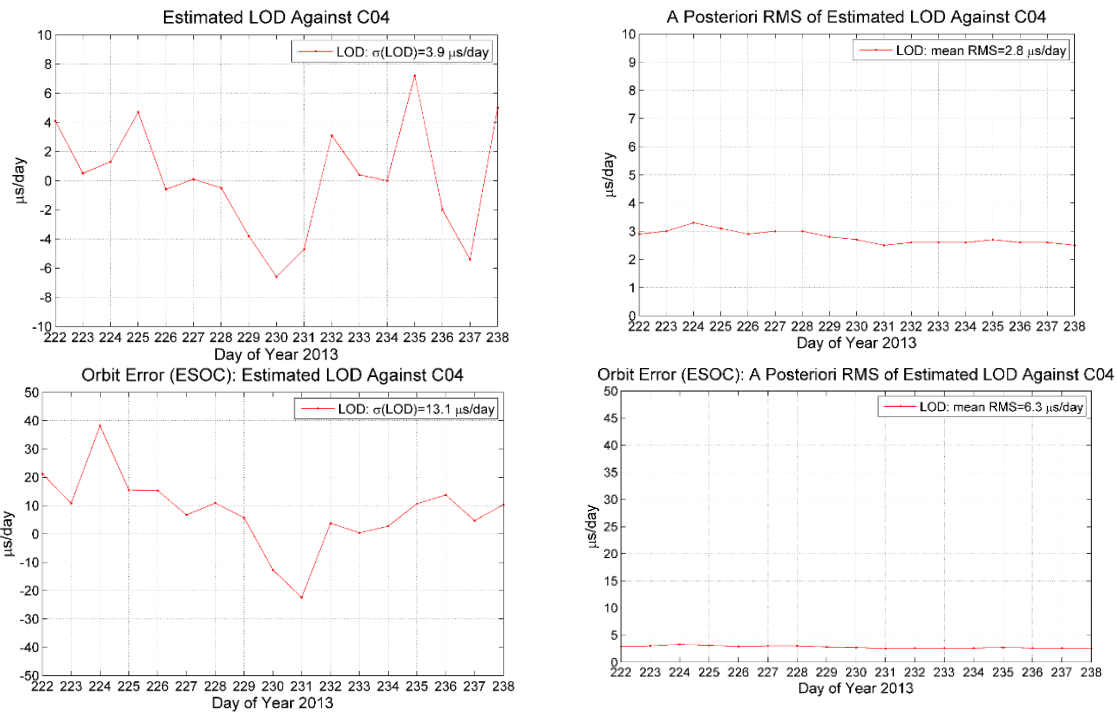
Figure 17.15 The estimated length-of-day (LOD) with a posteriori RMS values against the C04 values using only 4 global ILRS stations. The top figures refer to the fixed orbits of GPS satellites, whereas in the bottom figures an orbit error is simulated as the daily orbit difference between orbits provided by the two IGS AC.
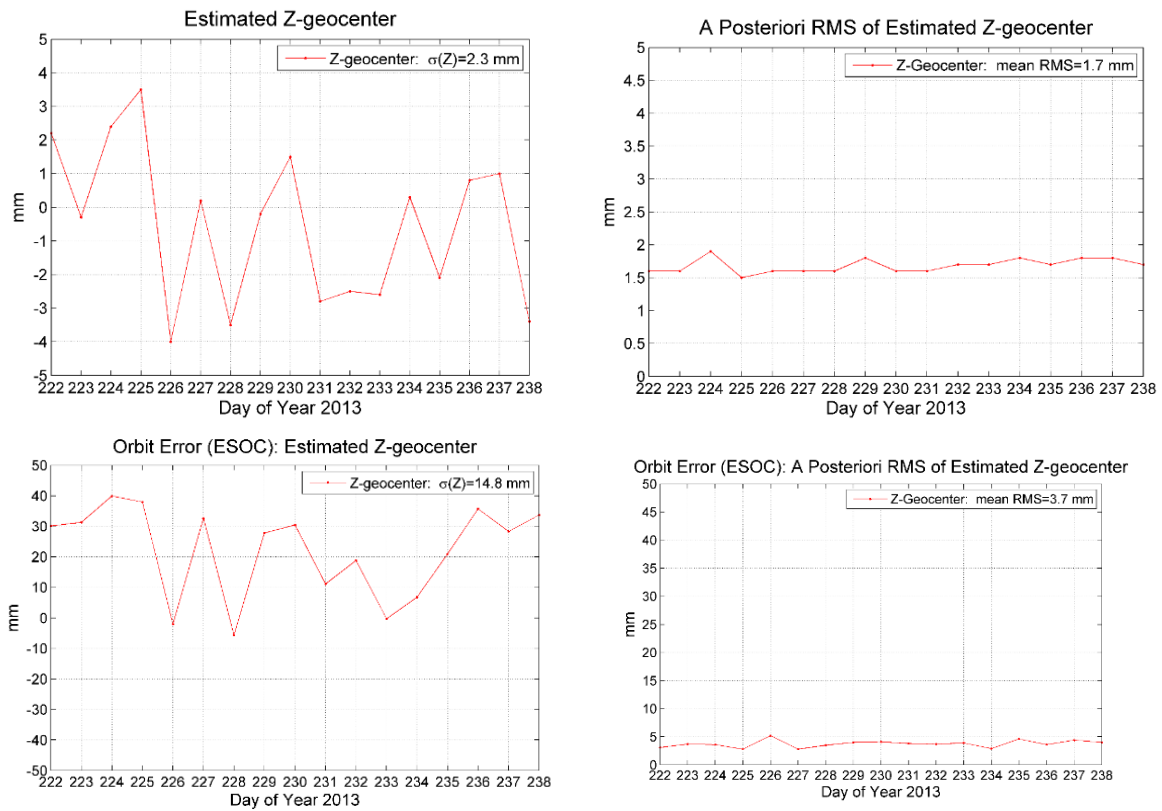
Figure 17.16 The geocenter Z-coordinate with a posteriori RMS values estimated with fixed orbits of GPS satellites and with a daily orbit difference between two IGS AC centers using only 4 global ILRS stations.

17.9 Relationship Between Bias in LAGEOS Center of Mass Correction and Radial Bias in Orbits of GNSS Satellites

Let us now see what would be the impact of the bias of e.g., 7 mm in the LAGEOS center of mass correction on the GNSS orbits. The existence of a potential small bias is indicated in the single-difference SLR measurements between HERL and GRZL in Figure 17.3. As we mentioned before, HERL employs strictly single-photon ranging to all satellites with a center of mass correction of 245 mm, whereas GRZL uses "leading edge" post-processing with a center of mass correction of 252 mm. The SLR frame bias of 7 mm in the radial orbit of LAGEOS satellites (reflected as the bias in the center of mass correction) will give a scale error or radial bias in the orbits of GNSS satellites, since the scale of the GNSS frame is typically taken from the SLR frame. From Kepler's third law written in the form $n^2 = GM / a^3$, we can derive the following relation for the semi-major axis a_{LAGEOS} of LAGEOS and GNSS satellites a_{GNSS} respectively

$$\Delta a_{GNSS} = \frac{a_{GNSS}}{a_{LAGEOS}} \Delta a_{LAGEOS} \quad \rightarrow \quad \Delta a_{GPS} \approx 2.2 \cdot \Delta a_{LAGEOS}, \quad \Delta a_{Galileo} \approx 2.4 \cdot \Delta a_{LAGEOS} \quad (17.7)$$

A bias in the semi-major axis of LAGEOS satellites (a bias in the center of mass correction) of $\Delta a_{LAGEOS} = -7$ mm would give a radial orbit bias $\Delta a_{Galileo} \approx -17$ mm in the orbits of Galileo satellites and $\Delta a_{GPS} \approx -15$ mm in the orbits of GPS and GLONASS satellites, see Table 17.3

LAGEOS	Galileo	GPS	GLONASS
−7 mm	−17.1 mm	−15.4 mm	−14.8 mm
−4.1 mm	−10 mm	−9 mm	−8.7 mm

Table 17.3 Radial bias in the orbits of GNSS satellites calculated as a function of the bias in the center of mass value of LAGEOS satellites. One can see that the Galileo radial orbit bias of −10 mm corresponds to a bias of −4.1 mm in the SLR measurements to LAGEOS that determine scale of the GNSS terrestrial frame.

17.10 Lunar Laser Ranging Double-Differences and Estimation of UT0

We have processed undifferenced and double-difference lunar laser ranging (LLR) measurements to Luna and Apollo retro-reflectors on the Moon in a similar way we are processing SLR measurements to GPS satellites, see (Svehla et al. 2015). We made use of the latest lunar libration models and DE430 ephemerides given in the Solar system barycentric frame and modeled uplink and downlink LLR ranges in the geocentric frame as one-way measurements, like the SLR to GPS satellites. We estimated all orbital parameters including UT0.

For the lunar orbit, we implemented the latest DE430 ephemerides given in the barycentric frame and described in (Williams et al. 2009). The same model provides physical librations of the Moon and coordinates of the two Luna and three Apollo lunar laser retro-reflectors. The DE430 model includes solid-body tides of the Moon in the form of permanent tidal displacements separately for each retroreflector array.

Following (Williams et al. 2009), the LLR retro-reflector principal axis coordinates were determined during the solution of DE430 ephemerides. These coordinates are rotated from the LLR principal axis frame (PA) to the lunar mean Earth/mean rotation axis frame (MER) by

$$\vec{p} = R_z(67.573'')R_y(78.580'')R_x(0.285'') \cdot \vec{m} \tag{17.8}$$

- • DE430 Ephemerides
- • Frame aligned to the International Celestial Reference Frame v.2.0
- • Solar System barycentric frame
- • TDB used as the Solar System barycentric coordinate time
  - TCB-TCG (IERS2010 Conventions)
- • Lunar librations (DE430)
- • Lunar reflector coordinates (DE430)
- • Principal axes and mean Earth/mean rotation axes
- • Constant tidal displacements from the Earth and the Sun (DE430)
- • Different force modeling for Moon (compared to GNSS/LEOs)
- • **Shapiro effect:**
  1. **Sun gravitational field: 7.5 m**
  2. Earth gravitational field: **0.04 m**
  3. Moon gravitational field: $<$**1 mm**
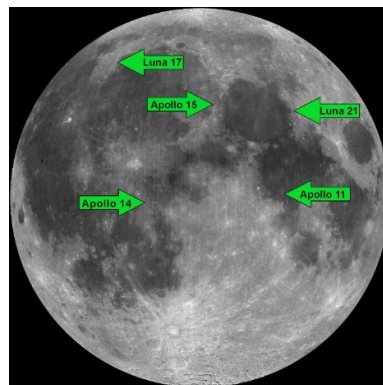
Table 17.4 Processing standard for the LLR measurements following the IERS Conventions 2010 [Petit and Luzum, 2010] and (Williams et al. 2009). Figure on the right shows position of the 5 Lunar retro-reflectors.

where $\vec{m}$ is the vector from the Lunar center of mass to a surface point in the mean Earth/mean rotation axis frame (MER) and $\vec{p}$ is the same vector in the principal axis (PA) frame. Such a transformation needs to be performed for the lunar libration rotation matrix provided by the DE430 ephemerides. For a description of DE430 models, we refer to (Williams et al. 2009), see Table 17.4. The lunar solid tides (constant tidal displacements due to Earth and Sun) are applied to the coordinates of the lunar retro-reflectors given by the DE430 ephemerides, (Williams, 2013).

It is interesting to note that the size of the Shapiro effect (Petit and Luzum 2010) in Table 17.4 for LLR measurements is in the order of 7.5 m for the Sun gravitational field and only 4 cm for the gravitational field of the Earth. For the analysis of LLR data in the geocentric frame, we used the following formulation in the IERS Conventions 2010 (Petit and Luzum 2010),

$$\vec{r}_b = \left(\vec{r} - \frac{U}{c^2}\right) - \frac{1}{2}\left(\frac{\vec{V} \cdot \vec{r}}{c^2}\right)\vec{V} \tag{17.9}$$

that provides transformation of the vector $\vec{r}$, a geocentric position vector expressed in the GCRS (Geocentric Celestial Reference System), to $\vec{r}_b$, the vector expressed in the BCRS (Barycentric Celestial Reference System). U is the gravitational potential at the geocenter (excluding the Earth's mass) and $\vec{V}$ is the barycentric velocity of the Earth. The geocentric and barycentric systems are chosen so that the geocentric space coordinates (position vector $\vec{r}_{TT}$) are consistent with terrestrial time (TT) and that the barycentric space coordinates are TDB-compatible (position vector $\vec{r}_{TDB}$ from DE430 ephemerides) (Petit and Luzum 2010). The transformation of $\vec{r}_{TT}$ to $\vec{r}_{TDB}$ is then given by

$$\vec{r}_{TDB} = \vec{r}_{TT}\left(1 - \frac{U}{c^2} - L_C\right) - \frac{1}{2}\left(\frac{\vec{V} \cdot \vec{r}_{TT}}{c^2}\right)\vec{V} \tag{17.10}$$

with the conversion factor L_C given in the IERS Conventions 2010, (Petit and Luzum 2010). The difference between TCB and TCG time scales (TCB-TCG) is calculated at the geocenter, using the approximation of the time ephemeris TE405. The IERS subroutine HF2002.f provided by the IERS Conventions approximates TE405 time ephemeris (including the trend) with an error of 0.453 ns (RMS) over the years 1600-2200, (Petit and Luzum 2010). For other time transformations: TCG-TT, TDB-TCB, TDB-TT we refer to IERS Conventions 2010, (Petit and Luzum 2010). The novelty is that we processed LLR measurements in the geocentric frame in a similar way we process SLR measurements in the geocentric frame for GPS satellites. For this, we calculate a light-travel time for LLR measurements from a ground ILRS station to
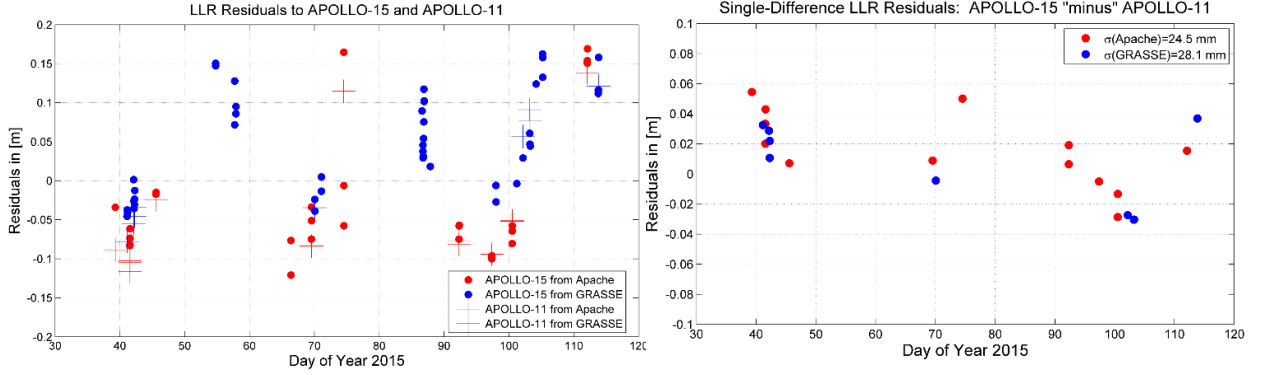
Figure 17.17 Undifferenced and single-difference LLR residuals to Apollo-11 and Apollo-15 lunar retro-reflectors from ground stations GRASE and Apache Point Observatory. Single-differencing is performed by making use of the residuals from the nearest epoch. Lunar ephemerides were fixed to DE430, along with all other parameters, and station coordinates were fixed to SLR2008. The cm-accuracy is achieved by single-differencing.

a lunar retro-reflector, and evaluated lunar libration at the epoch when LLR photons sent by a ground LLR station arrive at the lunar retro-reflectors. It was noticed that the physical lunar librations change significantly during the light-travel time. This separation between the epoch of lunar librations and the epoch of the Lunar orbit, enables to model uplink and downlink lunar laser ranges in the geocentric frame as the one-way measurements (similar to SLR measurements to GPS satellites). SLR measurements for GPS satellites are typically calculated at the reception epoch, when the reflected SLR signal arrives at the ILRS station. The same occurs with LLR, with the difference that the lunar orbit is provided by the DE430 ephemerides and given in TDB time. We calculate lunar librations from DE430 ephemerides at the reflection point (reception time minus one-way light-travel time) and took into account the velocity of the Earth $\vec{v}_{Earth}$ in the barycentric frame. The one-way light-travel equation for a distance d between a ground receiver and a satellite, for GPS and SLR measurements, is given by (17.11) in the equatorial true system of date. The same equation (17.11) can also be used for lunar laser ranging, taking into account the velocity of the Earth, $\vec{v}_{Earth}$:

$$GPS : \overline{d} = d \cdot \left(1 - \frac{\Delta \vec{x}_{rec}^{\,sat} \cdot \vec{v}^{\,sat}}{d \cdot c}\right), \qquad d = \left|\Delta \vec{x}_{rec}^{\,sat}\right|$$

$$SLR : \overline{d} = \frac{d}{2} \cdot \left[\left(1 - \frac{\Delta \vec{x}_{rec}^{\,sat} \cdot \vec{v}^{\,sat}}{c \cdot d}\right) + \left(1 - \frac{\Delta \vec{x}_{rec}^{\,sat} (\vec{v}^{\,sat} - 2 \cdot \vec{v}_{rec})}{c \cdot d}\right)\right] \qquad (17.11)$$

$$LLR : \overline{d} = \frac{d}{2} \cdot \left[\left(1 - \frac{\Delta \vec{x}_{rec}^{\,sat} \cdot \vec{v}^{\,sat}}{c \cdot d}\right) + \left(1 - \frac{\Delta \vec{x}_{rec}^{\,sat} (\vec{v}^{\,sat} - 2 \cdot \vec{v}_{rec})}{c \cdot d}\right)\right] - \frac{\Delta \vec{x}_{rec}^{\,sat} \cdot \vec{v}^{\,sat}}{c^2} \frac{\Delta \vec{x}_{rec}^{\,sat} \cdot \vec{v}_{Earth}}{d}$$

where $\Delta \vec{x}_{rec}^{\,sat}$ denotes to the station-satellite vector and $\vec{v}^{\,sat}$ and $\vec{v}_{rec}$ are the satellite and receiver velocity.

Figure 17.17 shows undifferenced and single-difference LLR residuals to Apollo-11 and Apollo-15 lunar retro-reflectors for a period of 90 days. All parameters were kept fixed, including lunar ephemerides, and station coordinates were in the SLR2008 frame. One can see that the accuracy of DE430 ephemerides and the ranging model is at the level of several centimeters, whereas single-difference residuals show significantly smaller scatter, with a standard deviation of about $\sigma = \pm 2.5$ cm. In the next step, we formed double-difference LLR measurements between two lunar retro-reflectors and two LLR stations, see Figure 17.7. Since, by forming double-differences of LLR measurements, all range biases are removed and orbit errors are significantly reduced (the lunar orbit is much further away than GPS orbits), one can consider the double-difference LLR as an ″orbit-free″ and ″bias-free″ differential approach. This is the reason why the noise level of residuals is reduced significantly in Figure 17.17, and for double-difference Apache LLR measurements achieved a noise level of $\sigma = \pm 7.5$ mm (one-way) and those from GRASSE a level of $\sigma = \pm 7.1$ mm.
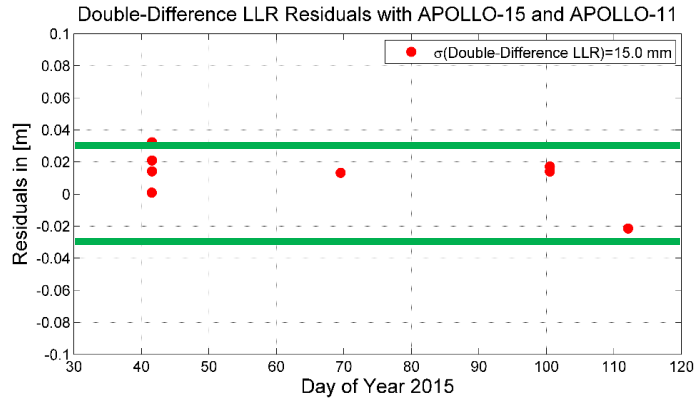
Figure 17.18 Double-difference LLR residuals to Apollo-11 and Apollo-15 from GRASE and Apache Point Observatory. Apache LLR measurements (one-way) show noise $\sigma = \pm 7.5$ mm (mean -3.6 cm), compared to GRASSE $\sigma = \pm 7.1$ mm (mean 4.6 cm). Differencing performed with residuals at the nearest epoch.

Estimation of 6 Keplerian parameters for the lunar orbit over 90 days improves the RMS of LLR residuals by $\times 2$ (from RMS=8.4 cm to RMS=4.8 cm). These residuals are further improved by estimating UT0 (every 10 days) to an RMS of about 3.5 cm, see Figure 17.19. LLR residuals plotted relative to Sun position in the lunar orbital plane in Figure 17.19 show a distinct pattern around 90° and 270° relative argument of latitude. This indicates that remaining modelling errors could also be associated with errors in the Earth orbit around the Sun. UT0 results are similar when additional empirical parameters are estimated in Figure 17.20.
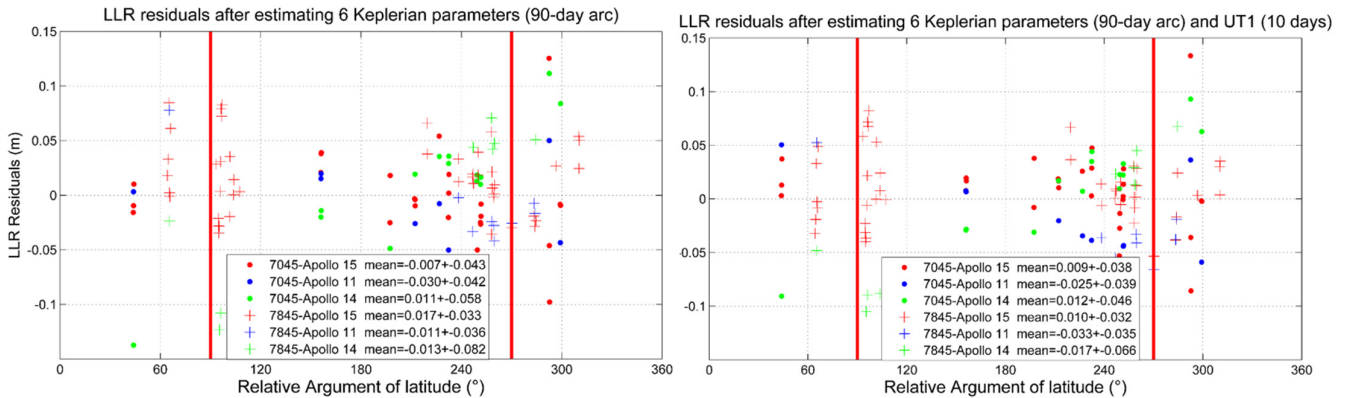
Figure 17.19 LLR residuals over 90 days after estimating 6 Keplerian parameters (left) and UT0 (right).
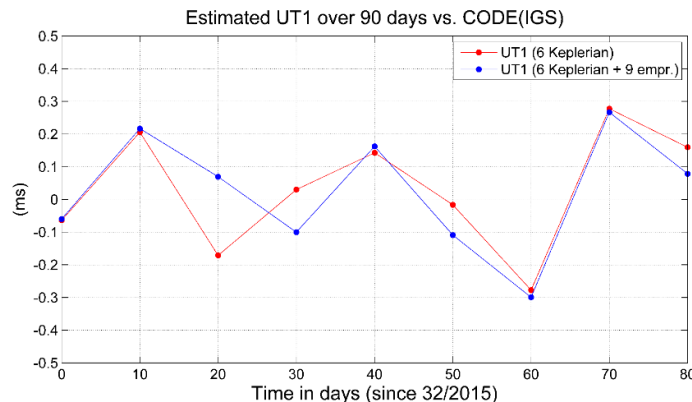
Figure 17.20 Estimation of UT0 over 90 days with an additional 9 empirical parameters (vs. CODE EOPs).

18. Noise Model of the Galileo "mm-Clock"

Galileo is the first GNSS system equipped with a highly stable H-maser. In this section we conduct a comprehensive analysis of the performance of the Passive H-Maser (PHM) used as a primary clock on board Galileo navigation satellites. PHM ground test results are compared to the clock parameters estimated from the MGEX data. The time evolution of the relativistic effects arising from the J_2 term of Earth's gravity field, as well as Sun and Moon gravitational potential have been calculated and taken into account. In addition, an analysis has been performed of the space environment (temperature and magnetic field variations) and the corresponding perturbations on the timing signal evolution .

Based on available ground test results, we derived relevant noise processes for the Galileo onboard passive maser, including the white frequency noise at the level of 5.9×10^{-13} – defining the short- to medium-term performance, and the flicker frequency noise of 7.9×10^{-16} – defining the clock long-term behavior. The white phase noise of 9.8×10^{-13} plays a role only for very short integration times (up to about 10 s), whereas a relatively low frequency drift of $< 1 \times 10^{-15}$/day plays a role only for measurement times longer than a few days.

Galileo clock parameters simulated according to the noise processes above show a residual standard deviation of $\sigma = 15.5$ mm, when time offset and time drift (linear model) are removed at 24 h intervals from the simulated epoch-wise Galileo clock parameters over 10 days. This standard deviation is reduced to $\sigma = 11.2$ mm, when the linear model is removed every 14 h (orbit period), going down to $\sigma = 2.7$ mm after time offset and time drift removal at 1 h intervals. For more see (Svehla, Cacciapuoti, Rothacher 2015, 2016, 2017), (Svehla et al. 2017).

The simulated data where then compared to the real in-orbit data. The Galileo clock solution from AIUB submitted to the MGEX Campaign of IGS shows a standard deviation of residual clock parameters at the level of $\sigma = 20.7$ mm, whereas SLR residuals show a higher standard deviation of $\sigma = 25.3$ mm. From this, one can derive a standard deviation of the radial orbit error to a level of $\sigma \approx 13 - 14$ mm. This factor of about 2 in precision between Galileo clock and SLR is most likely due to space/ground local ties, biases in some of the SLR ranges and in the reference frame (e.g., geocenter). We analyzed a period of 30 days 95-125/2013 of MGEX data with a high Sun elevation angle ($> 60^{\circ}$) above the Galileo satellite orbit plane in order to decouple orbit errors from the clock noise in the estimated Galileo clock parameters. In this case, the orbit errors originating from the modelling of solar radiation pressure are very modest compared to the rest of the draconic year of about 357 days for Galileo orbits. Similar results were obtained for Sun elevation angles $< -60°$ and four Galileo IOV satellites.

The main perturbation affecting the Galileo clock parameters for the analyzed period with high Sun elevation ($> 60^{\circ}$) is the periodic relativistic effect due to the J_2 gravity field coefficient that contributes an amplitude of about 18 mm at twice the orbital frequency. Accumulated time along the Galileo orbit due to the gravitational potential of Sun and Moon after removing daily time offset and time drift shows distinct, twice per revolution effects below 0.4 mm for the Sun potential and 1 mm for the Moon potential. Environmental effects, such as variations in temperature and magnetic field, were integrated along the orbit, but did

not have a significant impact on the Galileo residual clock parameters. The maximum effect due to the magnetic field is below 0.8 mm while temperature perturbations are well below 1×10^{-15} .

This analysis clearly shows that the onboard Galileo passive maser is stable enough to map for the first time radial perturbations continuously along the orbit. This is also confirmed by the close agreement with SLR residuals.

Estimated GNSS satellite clock parameters completely absorb variations in radial orbit error along the orbit. As a result, one can talk about an equivalence between the Galileo clock and SLR residuals, such that the Galileo clock can be considered as providing "continuous SLR" measurements along the orbit.

In summary, based on the simulated and real Galileo clock data, as well as the independent SLR measurements, the Galileo primary clock offers a wide spectrum of new applications, such as:

- geometrical mapping of the orbit perturbations along the orbit;
- clock modeling with only two linear parameters (time offset and time drift) or with a low-degree polynomial for a period up to one day, considerably reducing the number of estimated parameters in the orbit determination;
- primary clock on future LEO missions (e.g., DORIS on altimetry missions, or gravity missions) and for one-way ranging on interplanetary missions;
- mapping of troposphere slant delays between Galileo and a ground H-maser of similar stability.

18.1 An Overview of Galileo Clocks

The first two satellites of the European navigation system Galileo were launched on 21 October 2011, followed by the launch of two additional satellites on 12 October 2012 (ESA Portal 2014). The first four Galileo satellites are part of the Galileo In-Orbit Validation (IOV) Phase and contribute to the full constellation of 30 Galileo satellites. The Galileo navigation payload consists of two Passive Hydrogen Masers, two Rubidium Atomic Frequency Standards (RAFS) serving as backup, the Clock Monitoring and Control Unit (CMCU), the navigation signal generator unit, the L-band antenna for transmission of the navigation signal, the C-band antenna for uplink signal detection, the two S-band antennae for telemetry and telecommands, and the search and rescue antenna (ESA Portal 2014). The first Satellite Laser Ranging (SLR) to the retro-reflector arrays of the first two Galileo IOV satellites, denoted as Galileo-101 and Galileo-102 by ILRS (International SLR Service) and Galileo E11 and Galileo E12 by IGS (International GNSS Service), was carried out on 27 and 29 November 2011, respectively, using a near-infrared laser beam, (Svehla and Navarro-Reyes 2011).

The development of on-board clocks was initiated by ESA in the late nineties and resulted in the validation and qualification of two technologies. The Rubidium Atomic Frequency Standard is a microwave clock based on a vapour-cell with buffer gas operated on the double optical-microwave resonance of rubidium atoms. The clock, very compact and with low power consumption, has a fractional frequency stability better than $5 \times 10^{-12} \tau^{-1/2}$ over one day of integration time (Waller et al. 2009). The Passive Hydrogen Maser is based on the stimulated emission of microwave radiation on the hyperfine transition of the hydrogen ground state. Its fractional frequency stability is about 5 times better than that of RAFS (Waller et al. 2009). An overview of the Galileo clocks and their specifications can be found in (Rochat et al. 2012) and (Waller et al. 2009).

In the light of the new Galileo and BeiDou global navigation satellite systems (GNSS), as well as regional navigation and augmentation systems such as the Japanese Quasi-Zenith Satellite System (QZSS) and the Indian Regional Navigation Satellite System (IRNSS), the IGS initiated the Multi-GNSS EXperiment (MGEX). The goal of MGEX is the data collection and analysis of all available GNSS (Montenbruck 2013). The MGEX Tracking Network currently consists of about 90 active tracking stations contributed by about 25 different institutions (Steigenberger et al. 2014). The general consistency of the MGEX orbit products for Galileo is slightly better than one decimeter (Steigenberger et al. 2014). This rather rough orbit quality limits evaluation of the Galileo clock performance, since any orbit error will also be reflected in the estimated Galileo clock parameters. Perturbations which strongly depend on the satellite orbit have been recently observed in

the analysis of MGEX Galileo clock solutions (Steigenberger et al. 2014). These measurements also confirm earlier results reported in (Waller et al. 2009) or (Rochat et al. 2012) and clearly indicate that clock performance evaluation is heavily biased by orbit errors. Improvements in the quality of Galileo IOV orbit determination were reported recently in (Montenbruck et al. 2014) by employing an empirical a priori solar radiation pressure model that reduces the overall standard deviation of SLR residuals from 8-10 cm to 5-7 cm for all four Galileo IOV satellites (Montenbruck et al. 2014).

Modelling of the solar radiation pressure (SRP) based on the CODE SRP model (Beutler et al. 1994) introduces an error in the orbit as a function of the Sun elevation angle β above the orbital plane and the satellite argument of latitude (u) relative to the Sun's position in the orbital plane, see Section 19. At lower elevations there will be an additional effect on the orbit due to orbit eclipses. The same is true, if an empirical a priori SRP model is used, as this improves the overall accuracy of the orbit, but also introduces an additional signal at different orbit frequencies.

This has led to the development of a completely different approach to assess the quality of the Galileo PHM clock. Our proposed method for evaluating Galileo clock performance is based on two distinct facts. We decouple orbit and clock error by analyzing estimated clock parameters at high Sun elevation above the orbital plane $-60^{\circ} > \beta > 60^{\circ}$, where orbit quality is increased by a factor of 5-8 compared to low Sun elevations. In addition, we introduce what we call an equivalence between orbit error and clock error (see Figure 18.1), and use the SLR measurements in direct comparison with the Galileo epoch-wise clock parameters. This leads us to the first geometrical mapping of GNSS orbit perturbations. Power Spectral Density, Allan deviation and other metrics of the simulated and estimated Galileo clock parameters corrected by all known relativistic and environmental effects are then analyzed.
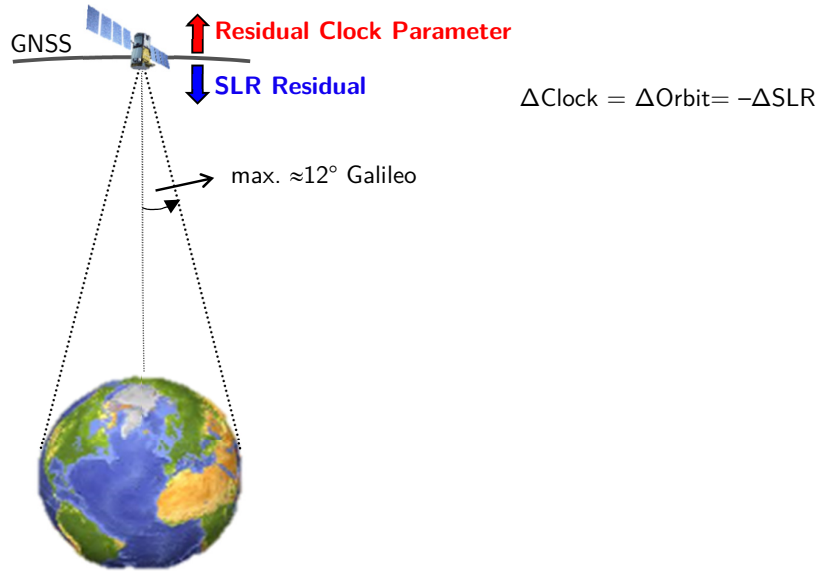
Figure 18.1 Equivalence between the radial orbit error and the residual clock parameters. Any radial orbit error (Δ Orbit) is compensated by the estimated clock parameter (Δ Clock) that corresponds to a negative SLR residual ($-\Delta$ SLR). Since the majority of ground stations are visible from a Galileo orbit at a nadir angle of $\approx 10^{\circ}$, a slant error of, e.g., 10 mm will give an error of only -1.5 mm, when projected in a radial direction.

18.2 First Geometrical Mapping of GNSS Orbit Perturbations

For satellites at GNSS orbit altitudes, any radial orbit error is directly mapped into the estimated clock parameters resulting in an opposite sign between the SLR residuals and the residual clock parameters (see Figure 18.1). SLR residuals are calculated as "observed-minus-computed", i.e., as the laser-measured ranges minus the ranges calculated from the solved-for satellite orbit. In this case, it is possible to establish an equivalence between the radial orbit error and the residual clock parameter. Therefore, if the clock onboard a GNSS satellite is stable enough, it can be used to map orbit perturbations along the satellite orbit. That was the original idea to assess the quality of the Galileo primary clock: to compare epoch-wise estimated Galileo clock parameters with the SLR measurements.

Figure 18.2 shows residual GIOVE-B clock parameters over a period of 4 days after subtracting a daily time offset and time drift from the clock parameters estimated epoch-wise every 30 s. One can clearly see a distinct pattern in the orbital period, highly correlated with the SLR residuals (plotted with an opposite sign) used only for the external orbit validation (dark blue). By adjusting just two linear parameters (time offset and drift over a one-day period) to the estimated Galileo clock parameters, the passive H-maser can be modelled with cm-accuracy, mapping the radial error continuously along the orbit with an excellent agreement with SLR measurements. We call this approach geometrical, as the stable Galileo clock measurements are equivalent to "continuous" SLR at every given GNSS epoch. The SLR residuals have a RMS of 5.4 cm.

This RMS value is significantly higher than the differences between the SLR residuals and GIOVE-B clock residuals, as one can see in Figure 18.2, indicating that any potential use of SLR measurements in the dynamic orbit determination is irrelevant to this approach.

Figure 18.3 shows residual clock parameters for the Galileo E11 satellite against SLR residuals from different ground ILRS stations. We used the Galileo clock solution from AIUB submitted to the MGEX Campaign of the IGS, days 95-125/2013.
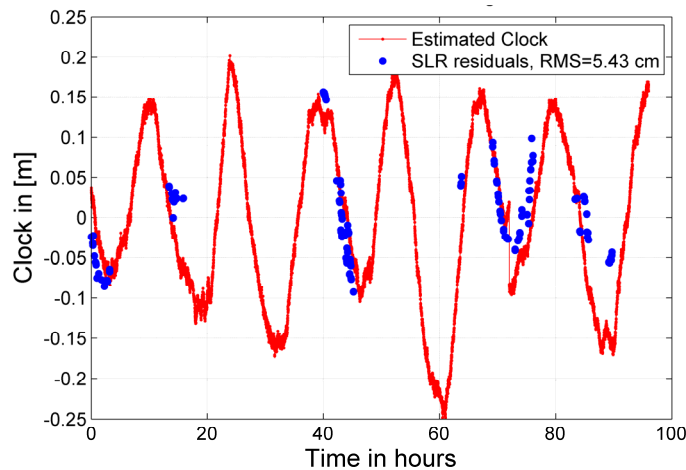
Figure 18.2 Residual GIOVE-B clock parameters after subtracting a daily time offset and drift from the satellite clock parameters estimated every 30 s (starting with day 250 in 2009). Dark blue dots represent SLR residuals (with an opposite sign) used only for orbit validation, showing that the passive H-maser on board GIOVE-B can be used to geometrically map orbit errors with remaining clock variations at the cm-level. This figure shows the first use of a stable GNSS clock in precise orbit determination/validation, (Svehla 2010a).
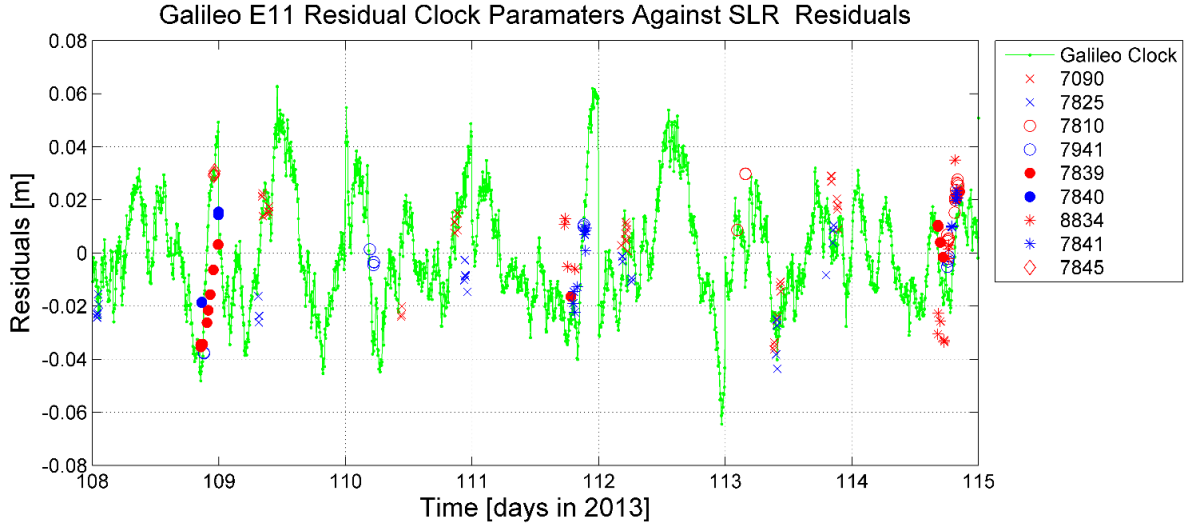
Figure 18.3 Residual clock parameters of the Galileo E11 satellite against SLR residuals for a 7-day subset of the 30-day analysis period. From the Galileo E11 clock parameters (MGEX solution from AIUB) a daily time drift and time offset were removed. Remaining residual clock parameters map radial orbit errors with an opposite sign relative to the SLR residuals calculated as "observed-minus-computed". The standard deviation of residual clock parameters is $\sigma = 20.7$ mm, for days 95–125/2013, whereas the noise contribution of the Galileo H-maser is about $\sigma = 15.5$ mm over a 24 h period (see Figure 18.6). This gives the radial Galileo orbit error with $\sigma \approx 13 - 14$ mm. In comparison with clock, SLR residuals show a higher standard deviation of $\sigma = 25.3$ mm (mean is -4.9 cm) and this factor of about 2 is most likely due to space/ground local ties, biases in some of the SLR ranges and in the reference frame (e.g., geocenter). The SLR residuals were provided by and compared with the Bernese GNSS Software v.5.3 at AIUB.

When a daily time drift and time offset is removed from the calculated satellite clock parameters, the remaining residual clock parameters map radial orbit errors with an opposite sign relative to the SLR residuals at the sub-cm level. We selected a period with high Sun elevation above the orbital plane $< -60°$ and $> 60°$ to significantly reduce the distinct periodic perturbation observed in Figure 18.2 at low Sun elevations. Only a periodic relativistic correction (Kouba 2004) due to the J_2 gravity field coefficient was applied to estimated Galileo E11 clock parameters. It is interesting to note that the standard deviation of residual clock parameters is $\sigma = 20.7$ mm, whereas SLR residuals show a higher standard deviation of $\sigma = 25.3$ mm. We will see later in the text from the end-to-end simulation of the Galileo clock that the noise contribution of the PHM is about $\sigma = 15.5$ mm over a 24 h period (see e.g., Figure 18.6). From this, one can derive a standard deviation of the radial Galileo orbit error of $\sigma \approx 13 - 14$ mm. This factor of about 2 in precision between Galileo clock parameters and SLR measurements is most likely due to space/ground local ties, biases in some of the SLR ranges and in the reference frame (e.g., geocenter).

This analysis confirms that the Galileo PHM can be used as "continuous SLR" along the orbit. However, the Galileo clock only maps the radial orbit error, whereas SLR maps, in addition, the contribution of the along-track and cross-track error. Since we did not account for the noise contribution of the global ground network in the standard deviation of residual clock parameters of $\sigma = 20.7$ mm, we can assume that the standard deviation of the radial orbit error is $\sigma \approx 10$ mm.

18.3 Noise Model of the Galileo H-maser

In order to evaluate the quality of the measured orbit errors in the radial direction, i.e., the stability of the residual Galileo clock parameters, we analyzed the performance of the Galileo onboard H-maser and evaluated possible environmental effects along the orbit.

The overlapping Allan variance corresponding to one of the best stability curves achieved during PHM performance tests on the ground (P. Rochat, private communication and (Wang et al. 2013) as available from SpectraTime) was considered as a reference. A model function, including all the relevant noise processes, was fitted to the data points:

$$f(\tau) = \frac{A^2}{\tau^2} + \frac{B^2}{\tau} + C^2 + D^2\tau + E^2\tau^2 \tag{18.1}$$

where the coefficients A, B, C, D and E are the fit coefficients. Only the first three coefficients in (18.1) were considered here: white phase noise, white frequency noise and the flicker frequency noise, respectively.

The relevant noise processes for the passive maser include the white frequency noise, defining the short to medium-term performance, and the flicker frequency noise, defining the clock long-term behavior. The white phase noise only plays a role for very short integration times (up to about 10 s) and becomes irrelevant for our analysis. Both, experimental data and the fitting function are shown in Figure 18.4. In the next step, the fit results are used in the Stable32 software (Riley 2014) to generate a time series of simulated clock data covering the same time span of 10 days that is available for the MGEX space clock parameter data. A frequency drift of 1×10^{-15}/day, as measured during the flight model tests on Galileo passive masers, was also added to the model function. The drift considered here is an upper estimate, which anyhow plays a role only for measurement times longer than a few days. Simulated clock data were generated at a sampling rate of 30 s, according to the model function parameters (Allan deviation at $\tau = 1$ s) listed below:

- White phase noise: 9.8×10^{-13}
- White frequency noise: 5.9×10^{-13}
- Flicker frequency noise: 7.9×10^{-16}
- Frequency drift: 1.2×10^{-20} /s^2

Figure 18.5 shows the resulting Allan deviation compared with the Allan deviation of real Galileo residual clock parameters from the MGEX Campaign. The higher noise observed in the MGEX data at about 7 hours (half the orbit period) will be discussed later in detail.
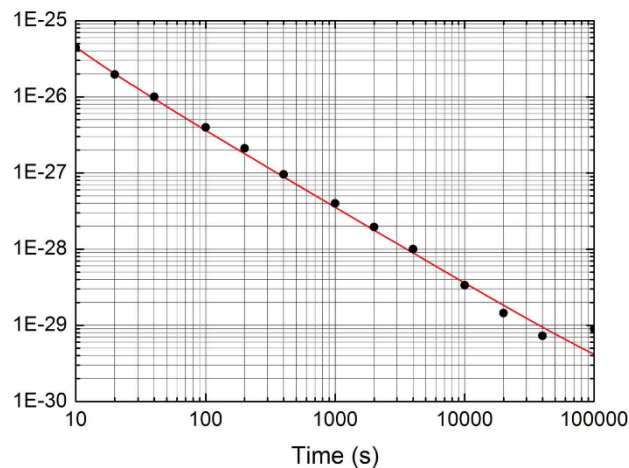
Figure 18.4 Overlapping Allan variance (black) of one of the best performing Galileo passive masers characterized during performance tests on the ground (P. Rochat, private comm.; see also (Wang et al. 2013) as available from SpectraTime) and best fit of the data points including all the relevant noise processes (red).
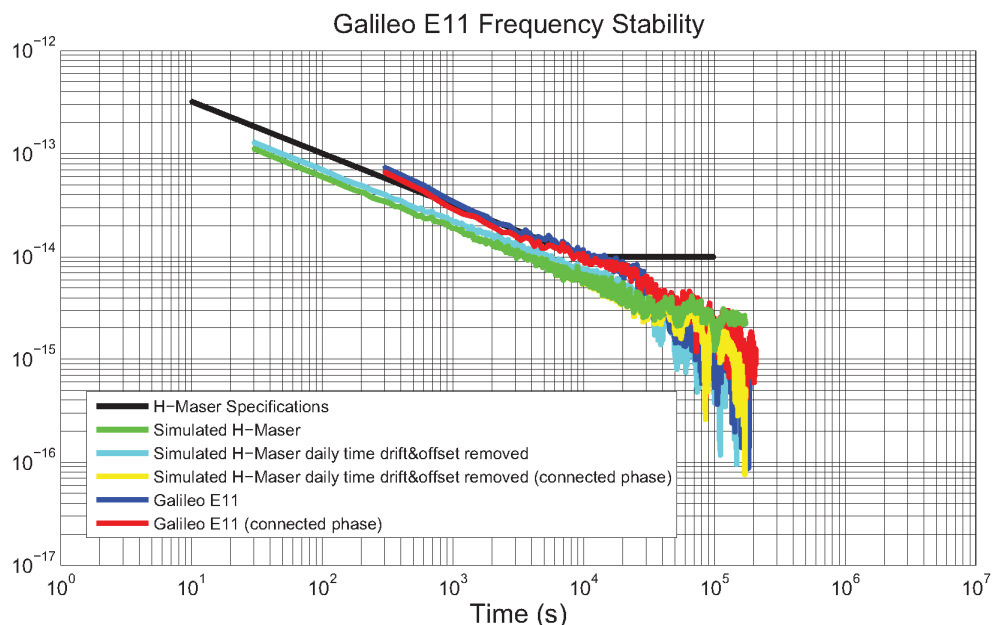
Figure 18.5 Allan deviation of the simulated clock parameters (red) and ground Galileo test results (green) against PHM specifications (cyan). In blue is the Allan deviation of the real residual clock parameters after time drift and bias removal every 24 h from the MGEX clock solution (AIUB) of IGS (days 96–106/2013). The "connected phase" denotes residual clock parameters connected at day boundaries and therefore showing better short term stability.

Simulated Galileo PHM data were then used to estimate the standard deviation of the clock error and compare it to the results obtained from MGEX data after applying the same processing algorithms. Simulated Galileo residual clock parameters show a standard deviation of $\sigma =15.5$ mm, when time offset and time drift (linear model) are removed at 24 h intervals. The corresponding data are shown in Figure 18.6. For comparison, it is interesting to note the qualitative agreement between Figure 18.3 and Figure 18.6 in terms of peak-to-peak variations and noise behavior. The standard deviation is reduced to $\sigma =11.2$ mm, when a linear model is removed every 14 h, down to $\sigma = 2.7$ mm after time offset and drift removal at 1 h intervals.
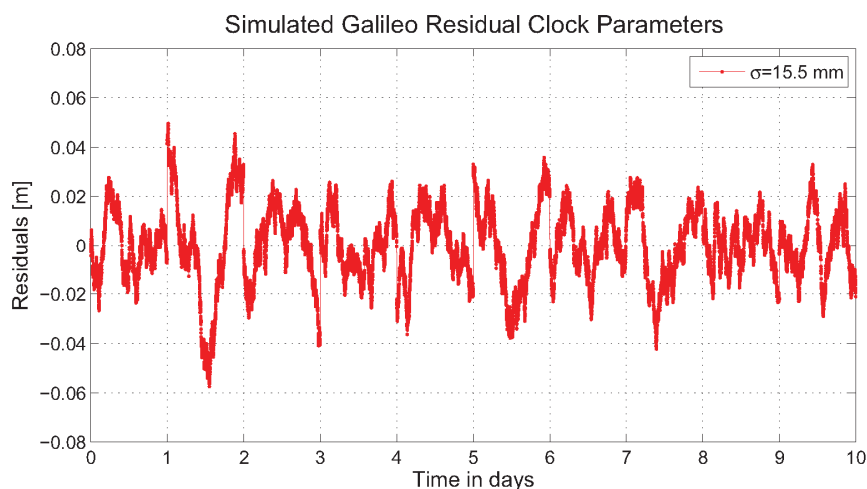
Figure 18.6 Simulated Galileo residual clock parameters over a period of 10 days. Only time offset and time drift were removed from the Galileo clock parameters every 24 h. Simulated residual clock parameters show a very good agreement with the real Galileo clocks in Figure 18.3, where the periodic relativistic effect of J_2 was removed, and Figure 18.7. The noise introduced by the GNSS network has not been considered.

The results of our analysis are shown in Table 18.1. With a polynomial of higher degree, mm-accuracy can be reached (Table 18.1). This is in line with Figure 18.3 which clearly shows consistency between clock and SLR residuals at the sub-cm level, when time offset and time drift is removed at 24 h intervals. However, for short intervals in Table 18.1, it is difficult to fully confirm values with real MGEX data, due to the short-term noise in the estimated clock parameters of the limited MGEX network, see Table 18.2. This is reflected in higher ADEV values in Figure 18.5 for MGEX satellite clock parameters for integration times up to several hours.

Table 18.3 shows standard deviation differences between MGEX clock parameters and simulated clock parameters based on the ground test results in Table 18.1. By forming such a differences, Table 18.3 reflects the noise of the ground data processing in the estimated MGEX Galileo clock parameters. It is interesting for a linear model to note that values at 12-h, 24-h and at the orbit period interval, are very similar, indicating that there is no significant signal at the orbit period and is most likely represented by flicker frequency noise after about 8-10 h (see also Allan deviations in Figure 18.5). This noise figure is most likely related to the significant orbit error represented by the orbit period being in the order of 7 h.

Simulated Galileo H-Maser (σ in mm)								
Degree	0.2 h	0.25 h	0.5 h	1.0 h	1.5 h	6 h	12 h	14 h	24 h
1	1.2	1.4	2.0	2.7	3.4	6.8	9.3	11.2	15.5
2	1.0	1.1	1.5	2.2	2.7	5.7	7.7	8.8	10.3
3	0.8	0.9	1.3	1.9	2.3	4.7	6.5	7.8	9.8
4	0.7	0.8	1.2	1.7	2.1	4.3	5.8	6.6	8.7
5	0.8	0.9	1.1	1.5	1.9	3.8	5.2	5.6	7.8

Table 18.1 Standard deviation of the simulated clock parameters for Galileo passive H-maser over a period of 10 days, after removing a polynomial of degree 1-5 over time intervals from 0.2 h to 24 h. Accuracy at the mm-level can be achieved by using the low-degree polynomial, significantly reducing the number of estimated clock parameters.

MGEX (AIUB) Clock Parameters (σ in mm)								
Degree	0.2 h	0.25 h	0.5 h	1.0 h	1.5 h	6 h	12 h	14 h	24 h
1	-	1.8	2.9	4.2	5.3	10.9	16.2	18.3	20.2
2	-	-	2.3	3.7	4.2	8.3	12.5	14.1	17.8
3	-	-	1.8	3.1	3.8	7.1	10.4	12.4	16.9
4	-	-	1.3	2.8	3.4	6.4	9.3	10.4	12.9
5	-	-	-	2.8	3.2	5.6	8.5	9.6	11.9

Table 18.2 Standard deviation of the MGEX clock parameters (AIUB) for Galileo passive H-maser over a period of 10 days (96-106/2013), after removing a polynomial of degree 1-5 over time intervals from 0.2 h to 24 h. Missing values are due to the lower sampling of MGEX clock parameters, given every 300 s.

Difference: MGEX (AIUB) - Simulated, (σ in mm)								
Degree	0.2 h	0.25 h	0.5 h	1.0 h	1.5 h	6 h	12 h	14 h	24 h
1	-	1.1	2.1	3.2	4.0	8.5	13.3	14.5	13.1
2	-	-	1.7	2.9	3.2	6.1	9.8	11.0	14.4
3	-	-	1.2	2.4	3.0	5.3	8.1	9.7	13.8
4	-	-	0.5	2.2	2.7	4.7	7.3	8.1	9.6
5	-	-	-	2.3	2.6	4.2	6.7	7.9	9.0

Table 18.3 Standard deviation difference between MGEX clock parameters (AIUB) in Table 18.2 and simulated clock parameters based on ground test data in Table 18.1 for Galileo passive H-maser over a period of 10 days (96-106/2013), after removing a polynomial of degree 1-5 over time intervals from 0.2 h to 24 h.

18.4 Relativistic Effects of Earth's Oblateness and Gravitational Fields of the Sun and Moon on the Galileo Clock Parameters

Following (Petit and Luzum 2010), the proper time τ of a clock with the coordinate position $\mathbf{x(t)}$ in the Geocentric Celestial Reference System (GCRS) moving with the coordinate velocity $\mathbf{v} = \mathbf{dx} / dt$, where t is Geocenteric Coordinate Time (TCG), is

$$\frac{d\tau}{dt} = 1 - \frac{1}{c^2}\Big[\frac{\mathbf{v}^2}{2} + U_E(\mathbf{x}) + V(\mathbf{X}) - V(\mathbf{X}_E) - x^i\partial_i V(\mathbf{X}_E)\Big] \qquad (18.2)$$

where c is the speed of light and U_E the gravitational potential of the Earth at the clock position $\mathbf{x}$ in the geocentric frame. V denotes the sum of the gravitational potential of the Sun and the Moon calculated at a location $\mathbf{X}$ in barycentric coordinates of the Solar system, separately for the Earth's center of mass $\mathbf{X}_E$ and the clock location $\mathbf{X}$. GNSS satellite clock parameters provided by IGS only include conventional periodic relativistic correction due to satellite orbit eccentricity. Considering only the central term of the Earth's gravity field (Kouba 2004), the U_E term in (18.2),

$$\Delta t_{per} = -\frac{2}{c^2}\sqrt{\mathrm{GM}a} \cdot e\sin E \qquad (18.3)$$

where a, e and E are the osculating semi-major axis, the eccentricity and the eccentric (angular) anomaly of the GNSS satellite orbit and GM is the geocentric gravitational constant. This periodic effect, with the orbit frequency mainly depends on the orbit eccentricity, i.e., special and general relativity effects due to satellite height and velocity variations from the mean values along the orbit. Considering the very small orbit eccentricity of $e = 0.0002$, the amplitude of this effect is only about -0.15 m for Galileo E11 (similar to all four IOV satellites), and it is about one order of magnitude higher for the constellation of GPS satellites that typically have higher orbit eccentricities by at least one order of magnitude. An alternative, but more convenient formulation of (18.3) applied directly in GNSS software packages is $\Delta t_{per} = -2\mathbf{r} \cdot \mathbf{v} / c^2$, where $\mathbf{r}$ and $\mathbf{v}$ denote for the satellite position and velocity vectors, respectively (Kouba 2004).

The periodic relativistic effect due to the J_2 gravity field coefficient was calculated using the following expression (Kouba 2004)

$$\Delta t(J_2)_{per} = -\frac{3}{2}\frac{a_E^{\,2}}{a^2c^2}J_2\sqrt{\mathrm{GM}a} \cdot \sin^2 i\sin 2u \qquad (18.4)$$

where a is the semi-major axis of the orbit, i the orbit inclination, u the argument of latitude and a_E the semi-major axis of the Earth's ellipsoid. This effect is due to special and general relativistic effects of the elliptical orbit perturbed by the Earth's oblateness, reflected in the J_2 coefficient (dynamic flattening) of the Earth's gravity field. Periodic effects of other low-degree zonal gravity field coefficients are negligible in our case. An additional time drift due to the J_2 coefficient in (Kouba 2004) is not considered here, since residual clock parameters are calculated by removing time offset and time drift of the satellite clock parameters estimated against the reference H-maser on the ground.

Figure 18.7 shows Galileo residual clock parameters (MGEX) at high Sun elevations from $60^°$ to $65^°$ together with the calculated J_2 contribution. The periodic relativistic correction (18.3) was added to the Galileo residual clock parameters in Figure 18.7 after multiplication by the speed of light c in a vacuum. The standard deviation of the calculated residual clock parameters is reduced from 2.5 cm to 2.1 cm. The amplitude of the periodic effect (18.4) for the orbit of the Galileo E11 satellite is about 18 mm. Figure 18.8 shows the
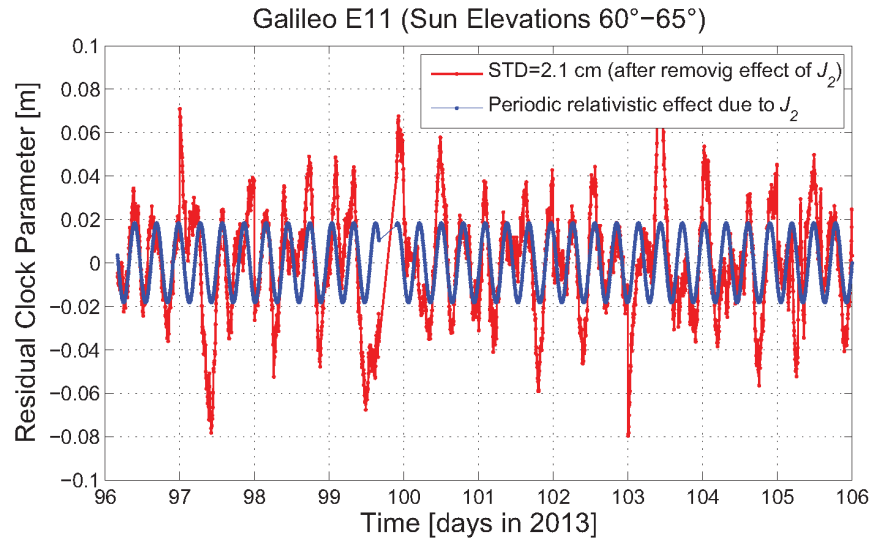
Figure 18.7 Galileo E11 residual clock parameters at high Sun elevations from $60°$ to $65°$ (clock solution from MGEX/AIUB). After removing the periodic relativistic effect due to the J_2 gravity field coefficient, remaining residual clock parameters show a standard deviation (STD) of 2.1 cm.

power spectral density of the Galileo E11 residual clock parameters before and after applying the correction for the periodic relativistic effect due to the J_2 gravity field coefficient. The peak originally present at twice the orbital frequency is removed after accounting for the J_2 perturbation.
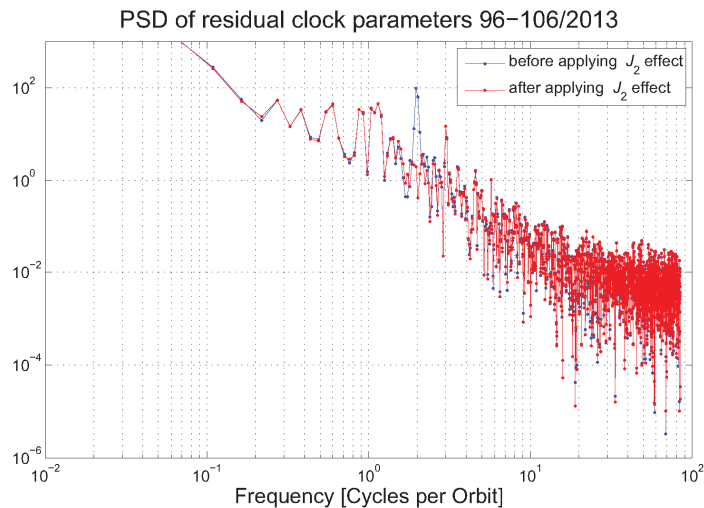
Figure 18.8 Power Spectral Density (PSD) of the Galileo E11 residual clock parameters (unit of length) at high Sun elevations from $60°$ to $65°$ before (blue) and after removing (red) the periodic relativistic effect due to the J_2 gravity field coefficient. The effect at 3 cycles per orbit revolution is still to be understood. Clock solution from MGEX/AIUB, days 96–106 in 2013.
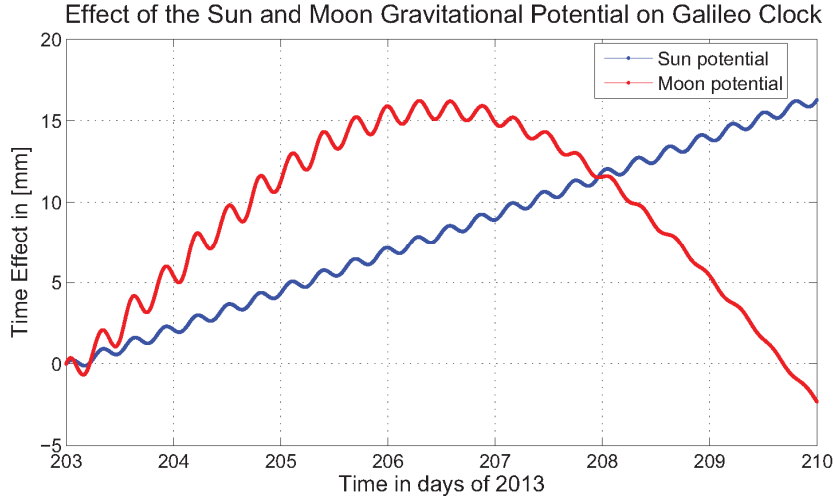
Figure 18.9 Accumulated time along the orbit of the Galileo E11 satellite due to the gravitational potential of Sun and Moon for a selected period of 6 days with low Sun elevation $4^\circ > \beta > -2^\circ$ above the orbital plane. When the Sun is in the Galileo orbital plane, the satellite orbit spans about 2×29600 km per orbit revolution in the Sun's gravitational field (max.-min. distance to the Sun). The maximum is reached for a Moon elevation of 28.5°.

The CMCU onboard the Galileo satellite can be used to adjust the constant frequency offset of the clock due to the effects of the general and special theories of relativity (Svehla 2010a) arising from the orbit altitude, see e.g., (Kouba 2004). Since this frequency adjustment could be performed in small finite steps, the absolute frequency of the Galileo clock is ambiguous by a constant step of the CMCU unit (Svehla 2010a). Therefore, here we are not considering the absolute frequency of the Galileo primary onboard clock (i.e., the time drift).

Since the first two terms in the brackets of (18.2) consider the periodic relativistic corrections (18.4) and (18.3) in the Earth's gravitational field, the accumulated time due to gravitational potential of the Sun and the Moon in (18.2) along the Galileo orbit was evaluated by the following expression based on (Wolf and Petit 1995)

$$V(\mathbf{X}_A) - V(\mathbf{X}_E) - x_A^i \partial_i V(\mathbf{X}_E) = \sum_{A \neq E} GM_A \Big[\frac{1}{r_{AP}} - \frac{1}{r_{AE}} + \frac{\mathbf{x}_{AE}\mathbf{x}_{EP}}{r_{AE}^3} \Big] \tag{18.5}$$

and displayed in Figure 18.9. A summation was carried out with the subscript A denoting Sun and Moon, and the subscript E is Earth. r is the modulus of the corresponding vector x to satellite P in the barycentric frame. Figure 18.9 shows that the net relativistic effect due to the Sun's and the Moon's gravitational potential is very small and, after removing the daily time offset and drift (see Figure 18.10), it reduces to 0.4 mm for the Sun and 0.8 mm for the Moon. We selected a period of 6 days with low Sun elevation in order to have a maximum extension of the Galileo satellite orbit in the Sun's gravitational field of about 2×29600 km over one orbit revolution. The larger oscillations for the Moon gravitational potential for the first few days in Figure 18.10 are due to the low elevation of the Moon above the satellite orbital plane.
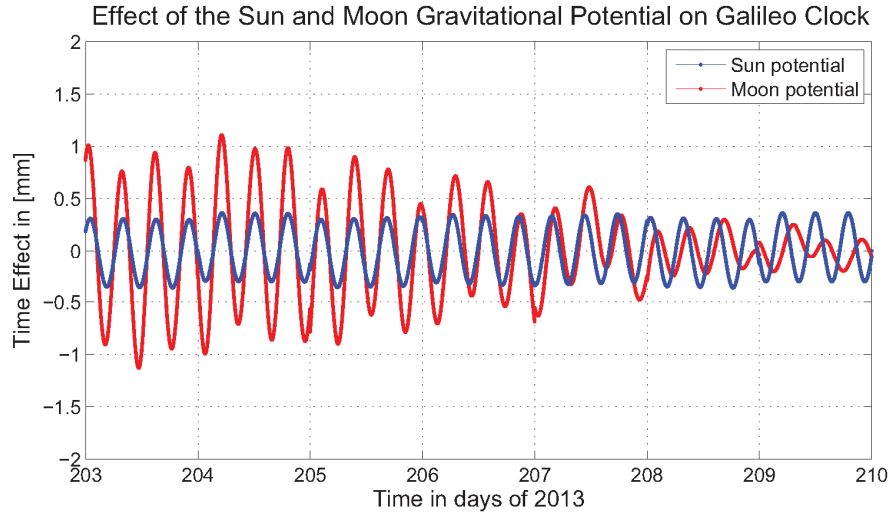
Figure 18.10 Accumulated time along the orbit of the Galileo E11 satellite due to the gravitational potential of Sun and Moon after removing daily time offset and drift. One can clearly see a distinct twice per revolution effect for the Sun potential. After removing daily time offset and time drift, the remaining effect on the residual clock parameters is below 0.4 mm for the Sun and up to 1 mm for the Moon potential.

18.5 Environmental Effects on the Galileo Clock Parameters

In this section, we discuss the impact of the in-orbit environment (magnetic field and temperature variations) on the Galileo clock performance. Magnetic field perturbations can be estimated by using the magnetic sensitivity coefficient of $< 3 \times 10^{-13}$ /G (one gauss equals 10^{-4} tesla) in fractional frequency, as measured during ground tests (Boving et al. 2009) and (Rochat et al. 2012). Magnetic filed variations along the Galileo orbit were calculated by using the International Geomagnetic Reference Field (IGRF) model (International Association of Geomagnetism and Aeronomy et al. 2010) in the direction of the satellite X, Y and Z axes. The time accumulated along the orbit was obtained by integrating the fractional frequency variations due to the ambient magnetic field (see Figure 18.11). Considering that the magnetic field is in the order of 300–550 nT along the Galileo orbit (days 100–116 in 2013), the contribution of magnetic perturbations to the estimated residual clock parameters is in the order of several millimeters. However, assuming the orientation of the Galileo maser cavity along the satellite X-axis (that never faces the Sun), the maximum effect of the magnetic field is below 0.8 mm, see Figure 18.11. When applied as a correction, the standard deviation of the residual clock parameters in Figure 18.7 was improved by only 0.1 mm. We can therefore conclude that the impact of magnetic field variations on the Galileo clock parameters is very small and negligible. In addition, shielding of the satellite further reduces their effect. However, this would not be the case if the same clock were placed in a LEO orbit, where the magnetic field strength is higher by two orders of magnitude.

Unfortunately, no public data is available on the in-orbit temperature at the clock reference point, therefore not much can be said about thermal perturbations. (Boving et al. 2009) reported a thermal sensitivity coefficient of the Galileo H-maser as measured on the ground of $\leq 2 \times 10^{-14}$ /°C. The cavity temperature of the Galileo H-maser is stabilized by a two-stage thermal control and an additional electronic Automatic Cavity Tuning (ACT) system is used to optimize the cavity frequency pulling effect caused by the residual thermal drift (Mattioni et al. 2002). From (Mattioni et al. 2002), one can see that, for platform temperature variations of 5 °C, the cavity thermal control stabilizes the temperature within 3 m°C.
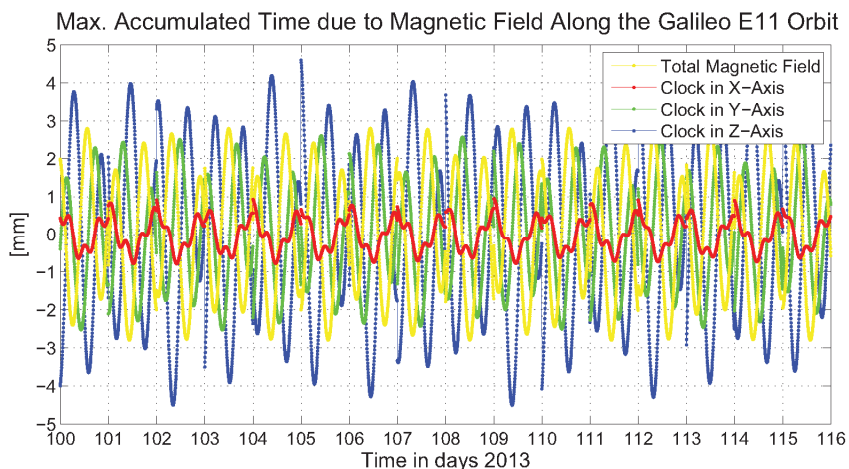
Figure 18.11 Maximum accumulated time in (mm) along the orbit of Galileo E11 satellite assuming a maximum magnetic sensitivity coefficient of 3×10^{-13} /Gauss. The magnetic field along the X, Y and Z satellite axes was calculated using the IGRF model giving a magnetic field variation of 300-550 nT along the orbit. Assuming the orientation of the H-maser cavity along the X satellite axis (never faces Sun), the maximum effect of the magnetic field on residual clock parameters is below 0.8 mm. Along the Z axis, the effect is about three times higher. Daily bias and drift were removed.

Temperature variations at the PHM reference point on-board the Galileo satellite are expected to have two different periods, the orbital period and the period of the stabilization loop. An analysis of the Allan deviation and PSD curves in Figure 18.7 and Figure 18.8, reveals no perturbation at the orbital period nor for periods shorter than 1000 s that could be attributed to temperature effects. This allows us to conclude that the temperature stability at the PHM reference point is at the level of a few tenths of a Kelvin.

19. Model of Solar Radiation Pressure and Thermal Re-Radiation

T he non-gravitational force solar radiation pressure is the main source of error in the precise orbit determination of GNSS satellites. All deficiencies in the modeling of solar radiation pressure map into estimated terrestrial reference frame parameters as well as into derived gravity field coefficients and altimetry results when LEO orbits are determined using GPS. Here we introduce a new approach to geometrically map radial orbit perturbations of GNSS satellites, in particular due to solar radiation pressure along the orbit, using high-performing clocks on board the first Galileo satellites. We have seen in Section 18 that only a linear model (time offset and time drift) need be removed from the estimated Galileo clock parameters and the remaining clock residuals will map all radial orbit perturbations along the orbit. Agreement between SLR residuals and clock residuals is at the cm-level RMS for an orbit arc of 24 h . Looking at the clock parameters determined along one orbit revolution over a period of one year, we show that the so-called SLR bias in Galileo and GPS orbits can be represented by a translation of the determined orbit in the orbital plane away from the Sun. This orbit translation is due to thermal re-radiation and does not account for the Sun's elevation above the orbital plane in the parameterization of the estimated solar radiation pressure parameters. SLR ranging to GNSS satellites takes place typically at night, e.g., between 6 p.m. and 6 a.m. local time, when the Sun is in opposition to the satellite. Therefore, SLR mostly observes that part of the GNSS orbit with a radial orbit error that is mapped as an artificial bias into the SLR observables. The Galileo clocks clearly show an orbit translation for all Sun elevations: the radial orbit error is negative when the Sun is in conjunction (orbit noon) and positive when the Sun is in opposition (orbit midnight). The magnitude of this SLR bias depends on the accuracy of the determined orbit and should rather be called "GNSS orbit bias" instead of "SLR bias". All LEO satellites, such as CHAMP, GRACE and JASON-1/2, need an adjustment of the radial antenna phase center offset. When LEO satellite orbits are estimated using GPS, this GPS orbit bias is mapped into the antenna phase center. GNSS orbit translation away from the Sun in the orbital plane not only propagate into the estimated LEO orbits, but also into derived gravity field and altimetry products. The mapping of orbit perturbations using an onboard GNSS clock is a new technique to monitor orbit perturbations along the orbit and was successfully applied in the modeling of solar radiation pressure. We show that the CODE solar radiation pressure parameterization lacks the dependency on the Sun's elevation above the orbital plane, i.e., the elongation angle (rotation of solar arrays), especially at low Sun elevations (eclipses). Sun elongation angle is used in the so-called T30 model (ROCK) that includes thermal re-radiation. A preliminary version of a solar radiation pressure model for the first five Galileo and the GPS-36 satellite is based on the orbit/clock solution of 2×180 days of the MGEX Campaign. We show that, in addition, Galileo clocks map the Yarkowsky effect along the orbit, i.e., a small time lag between the Sun's illumination of the satellite and its thermal re-radiation. We present the first geometrical mapping of the anisotropic thermal emission of absorbed sunlight of an illuminated satellite.

19.1 Galileo Clock Parameters and the SLR Bias in GNSS Orbits
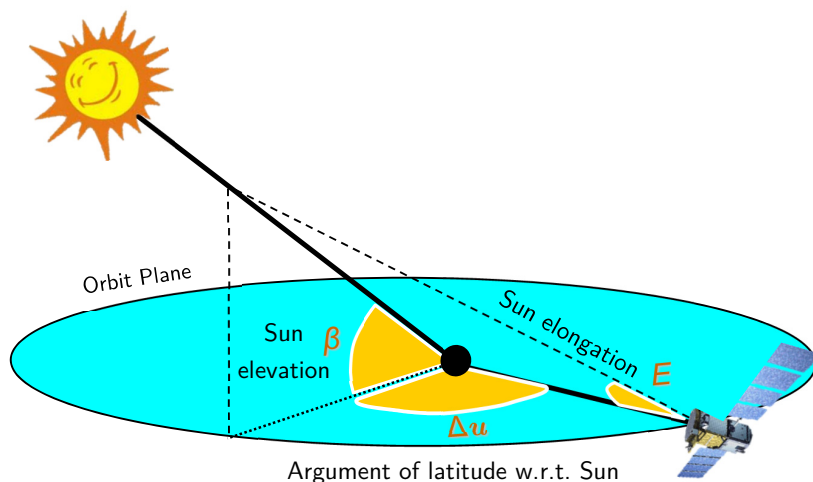
Figure 19.1 Definition of a Sun-fixed orbit coordinate system. The β angle denotes the elevation of the Sun above the orbital plane, Δu is the argument of latitude w.r.t. the argument of latitude of the Sun and E denotes Sun elongation angle.

In Section 18 we demonstrated that the estimated epoch-wise Galileo clock parameters can be used to map radial orbit error continuously along the Galileo orbit. That was confirmed by an external validation with SLR measurements. Based on this analysis of Galileo clock parameters, it was reported in (Svehla et al. 2013c) that modelling solar radiation pressure (SRP) based on the CODE SRP model (Beutler et al. 1994) will always introduce an error in the orbit modeling as a function of the Sun elongation angle E. The Sun elongation angle E is the angle at which the satellite "sees" the Sun and the geocenter and can be determined from spherical geometry, see Figure 19.1,

$$\cos E = -\cos \beta \cos \Delta u = -\cos \beta \cos(u - u_\odot) \tag{19.1}$$

as the function of the Sun elevation angle β above the orbital plane and satellite argument of latitude u relative to Sun position in the orbital plane of the satellite $u_\odot$. The Sun elongation angle is also the angle defining the orientation of the solar array with respect to the satellite body. The same Sun elongation angle is explicitly used in the so-called T30 model (ROCK) an a priori SRP model for GPS satellites that includes thermal re-radiation (Fliegel and Gallini 1996) and an empirical SRP model from JPL (Bar-Sever and Kuang 2004). In these two models amplitudes are typically given for the following harmonics: E, $3E$, $5E$ in the GPS satellite Z-direction with an additional $2E$ and $7E$ in the X-component. However, as reported in (Svehla et al. 2013c), Galileo clock parameters clearly show only the first harmonic E in the Galileo orbit over all Sun elevations that can be modelled as $A \cos \beta \cos \Delta u = -A \cos E$, where A denotes amplitude. Recently, (Montenbruck et al. 2014) reported an a priori SRP model for Galileo satellites that is very similar in parameterization to (Svehla et al. 2013c). The same approach was also recently applied in the parameterizations of solar radiation pressure for GPS and GLONASS (Arnold et al. 2014), reporting estimation of harmonic amplitudes of the elongation angle that go up to $3E$ and $4E$.

Let us now analyze Galileo residual clock parameters for all Sun elevations above the orbital plane. We use the Sun-fixed orbital frame, as defined in Figure 19.1 with Sun elevation angle β above the orbital plane and argument of latitude of the Galileo satellite relative to the Sun position in the orbital plane Δu. Figure 19.2 shows Galileo E11 residual clock parameters for rising and setting Sun (ascending and descending Sun elevations) based on MGEX orbit/clock solutions from AIUB and GFZ Potsdam. As expected, one can see a very close agreement between different MGEX solutions. Figure 19.2 shows that residual Galileo clock parameters are centered at an argument of latitude of $\Delta u = 180°$ relative to the Sun and the magnitude decreases with increasing Sun elevation. The maximum effect is when Sun and satellite are in opposition $\Delta u = 180°$, and the minimum at $\Delta u = 0°$ when they are in conjunction. The same effects can be seen in the MGEX orbits available from TU München in Figure 19.3.

Since the minimum and maximum are reached at Sun/satellite conjunction (orbit noon) and opposition (orbit midnight) for all Sun elevation points towards translation of the calculated orbit away from the Sun in the Sun-fixed orbital frame, i.e., the radial orbit error is positive when the Sun is in opposition and negative when Sun is in conjunction, see Figure 19.3 and Figure 19.4.
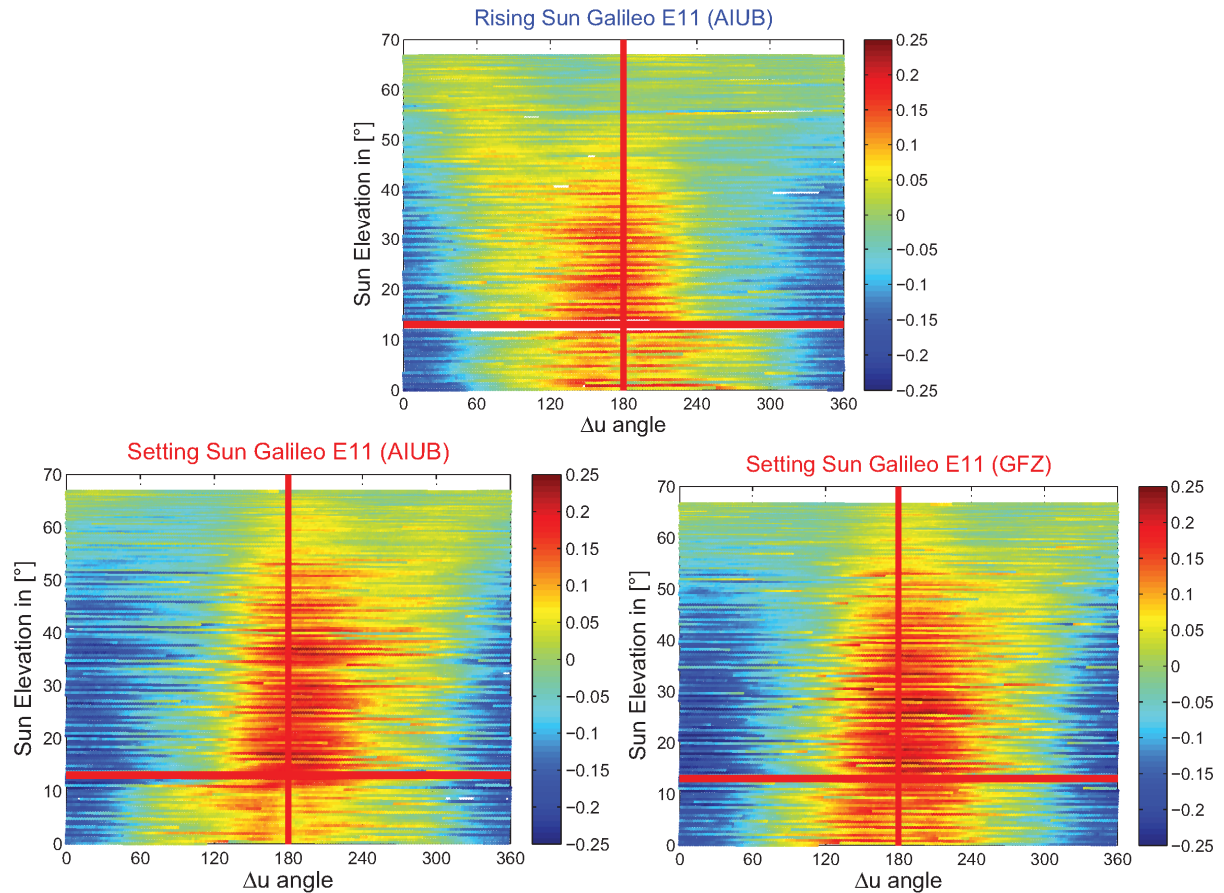
Figure 19.2 Galileo E11 residual clock parameters in Sun-fixed orbital frame for rising (top) and setting Sun elevations (bottom) against the argument of latitude relative to the Sun argument of latitude. Figures are based on the MGEX clock solutions from AIUB and the bottom-right figure on the MGEX solution from GFZ Potsdam. One can see a very close agreement between different MGEX solutions and asymmetry in argument of latitude between rising and setting Sun elevations. Max. effect is at $\Delta u = 180°$ (vertical red line), when the Sun and the satellite are in opposition and min. effect at $\Delta u = 0°$ when they are in conjunction. The horizontal red lines show Sun elevations between $-12° < \beta < 12°$ (satellite passing eclipses).
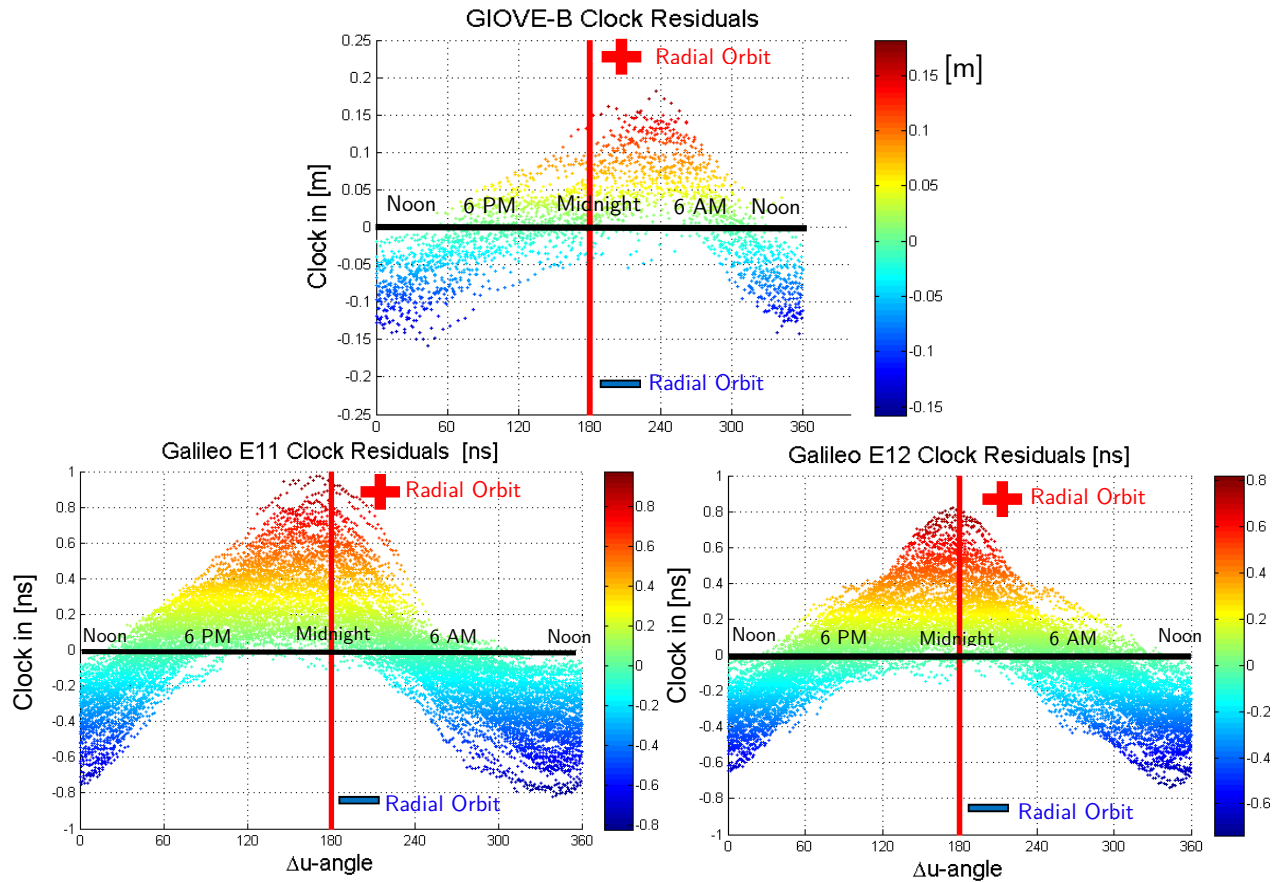
Figure 19.3 Clock residuals of GIOVE-B, Galileo E11 and Galileo E12 satellites against the argument of latitude of the satellite relative to the Sun argument of latitude. Max. effect is at $\Delta u = 180°$, when the Sun and the satellite are in opposition and at $\Delta u = 0°$, when they are in conjunction. Note also a slight asymmetry for the GIOVE-B satellite clock parameters that follows the high Sun elevation. Based on 2×180 days (2012/2013) of data from the MGEX Campaign of IGS (Galileo Clock Solution from TU München).

Estimated clock parameters for all three Galileo satellites show a periodic effect (cosine function) highly correlated with the argument of latitude relative to the position of the Sun. The maximum effect is reached when Sun and satellite are in opposition $\Delta u = 180°$, and at $\Delta u = 0°$, when they are in conjunction. Due to the fact that SLR ranging to GNSS satellites takes place typically at night, e.g., between 6 p.m. and 6 a.m. local time when the Sun is in opposition to the satellite, SLR measurements observe only one part of the GNSS orbit, including radial orbit error that leads to an artificial negative bias in SLR measurements, see Figure 19.4. The Galileo clocks clearly show this orbit translation for all Sun elevations.
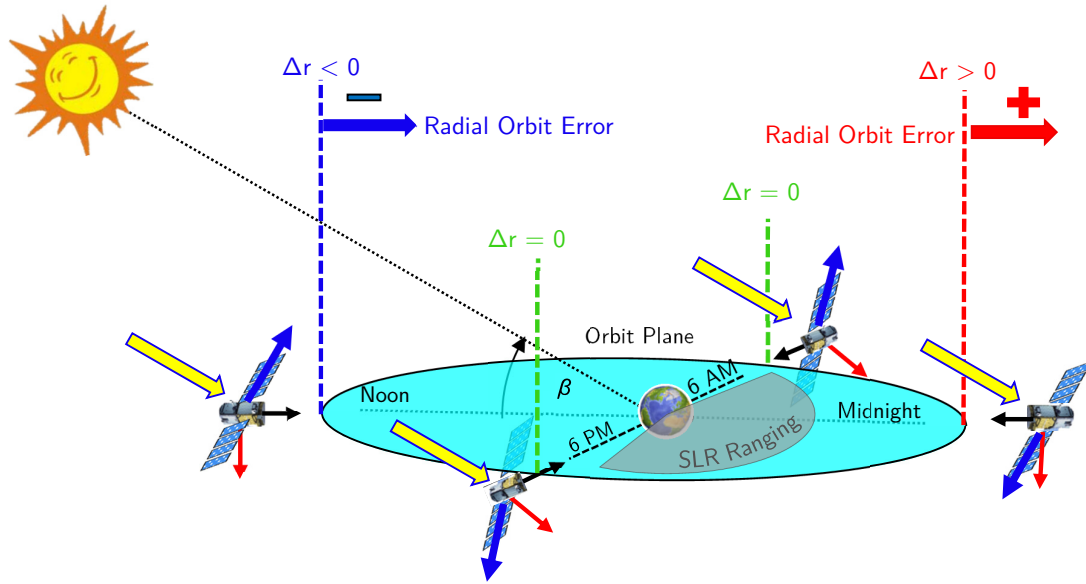
Figure 19.4 Translation of the GNSS orbit away from the Sun in the orbit plane, as mapped by the Galileo clock. Night-time SLR ranging (depicted in grey) covers mainly that part of the orbit with positive radial orbit errors $\Delta r > 0$. This explains why the SLR bias should be called "GNSS orbit bias" instead of "SLR bias". Based on information provided on the ILRS homepage, the blue/black arrows depict +Y and +Z-axes of the attitude yaw steering such that the +Y axis has an opposite sign to the +Y axis of GPS II/IIA, i.e., the +X (red arrow) spacecraft panel is maintained away from the Sun. This is the same as for GPS IIR satellites.

for all Sun elevations: the radial orbit error is positive, when the Sun is in conjunction (orbit noon) and negative when the Sun is in opposition (orbit midnight), see also Figure 19.5. The magnitude of this artificial negative SLR bias depends on the orbit quality and, therefore, should rather be called "GNSS orbit bias", instead of "SLR bias". For example, early Galileo orbits were showing a bias of -10 cm that dropped to some -6 cm, when orbits improved by a factor of 2. When LEO satellite orbits are estimated using GNSS, this orbit bias could be reflected as phase center offset in the radial direction, and this could be the reason why all LEO satellite missions need an adjustment of the antenna phase center in the radial direction. Moreover, the GNSS orbit translation in the Earth-Sun direction in the orbital plane directly maps into the estimated LEO orbits and subsequently into derived gravity field or altimetry products in the case of gravity and altimetry missions.

In order to model the periodic effect in Figure 19.2 and Figure 19.3, as a first approximation we may use the cosine function of the satellite argument of latitude relative to the Sun position in the orbital frame Δu.
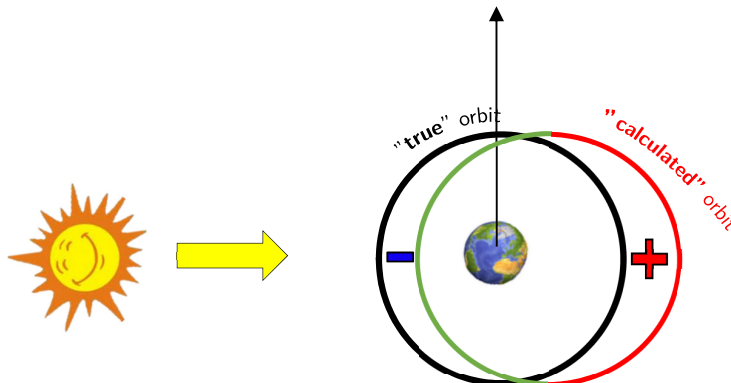
Figure 19.5 True and calculated orbit as revealed by the Galileo clock parameters.

In addition, by making use of the Sun elevation above the orbital plane β, the satellite radial orbit error Δr along the orbit can be approximated by

$$\Delta r = A \cdot \cos\beta \cos\Delta u \qquad (19.2)$$

In the case of GIOVE-B and the first four Galileo satellites, the amplitude A is in the order of $A \approx 20$ cm and depends also on the orbit quality. Let us now introduce the elongation angle E from (19.1) at which the satellite "sees" the Sun and the geocenter, see Figure 19.1. Hence, as a first approximation we can introduce an empirical model δ_{clk} for the Galileo residual clock parameters as

$$\delta_{clk} = \Delta r = -A \cdot \cos E \qquad (19.3)$$

Figure 19.6 shows the first approximation model of the Galileo residual clock parameters parameterized by the elongation angle in (19.3). Since (19.3) gives the circular pattern in Figure 19.6 that can also be seen with real Galileo data in Figure 19.2, we have given this effect the name "eye-effect". Figure 19.6 also shows a similar pattern for the model of GPS radial error due to solar radiation pressure modelling deficiencies for GPS 06 (Svehla et al. 2011). Compared to Galileo, one can see the very modest amplitude of about 10 cm due to the Galileo orbit quality at that time.

Eq. (19.3) and Figure 19.6 clearly point towards a translation of the calculated orbit away from the Sun in the Sun-fixed orbital frame, i.e., the radial orbit error is positive at orbit midnight, when Sun is in opposition $E = 180°$, and negative at orbit noon $E = 0°$. Such an orbit translation will introduce an orbit bias and subsequently an SLR bias when GNSS orbit is observed by nigh-time SLR ranging between e.g., 6 pm and 6 am that corresponds to the interval of about $\Delta u = 90° - 270°$.

If we now plot residual clock parameters as a function of elongation angle, we obtain Figure 19.7, showing that Galileo residual clock parameters (radial error) closely follow the Sun elongation angle, i.e., the orientation of the solar array w.r.t. to the satellite body. This is also confirmed by SLR residuals plotted with a negative sign in Figure 19.7 (right). The use of Sun elongation angle in (19.3) clearly points towards deficiencies in the modeling of solar radiation pressure. Sun elongation angle is explicitly used in the so-called T30 model (ROCK) that includes thermal re-radiation (Fliegel and Gallini 1996) and in an a priori empirical model of solar radiation pressure for GPS satellites used at JPL (Bar-Sever and Kuang 2004).

The clock estimates for all five Galileo satellites clearly show this orbit translation for all Sun elevations: the radial orbit error is positive when the Sun is in conjunction (orbit noon) and negative when the Sun is in opposition (orbit midnight). This is fully in line with (Urschl et al. 2007) that reported for the first time an eye-type pattern in the SLR residuals of the two GPS and GLONASS satellites equipped with SLR retro-reflectors, indicating negative SLR residuals with a maximum effect when Sun and satellite are in opposition.
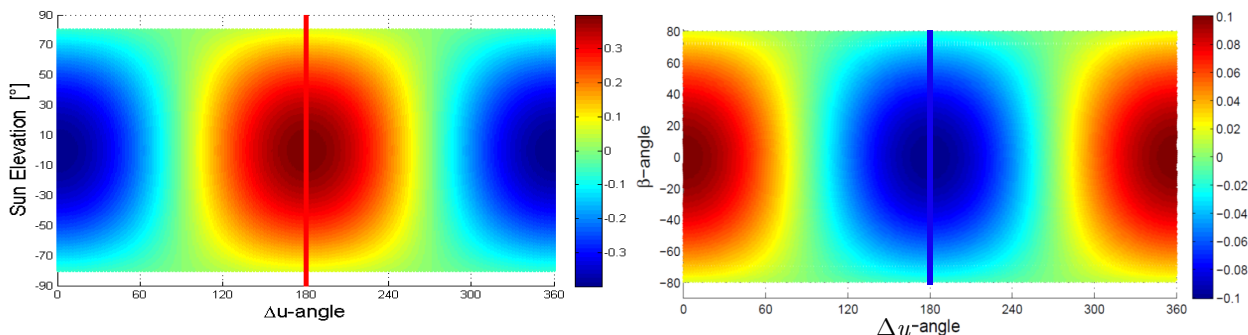
Figure 19.6 The first model of the Galileo clock residuals (left) in [m] using elongation angle and a model for SLR residuals for GPS 06 (right) (Svehla et al. 2011). For GPS, one can see the very modest amplitude of about 10 cm and change of the sign for SLR. The higher amplitude of the effect for Galileo is due to the orbit quality available for Galileo satellites from TU München at that time.
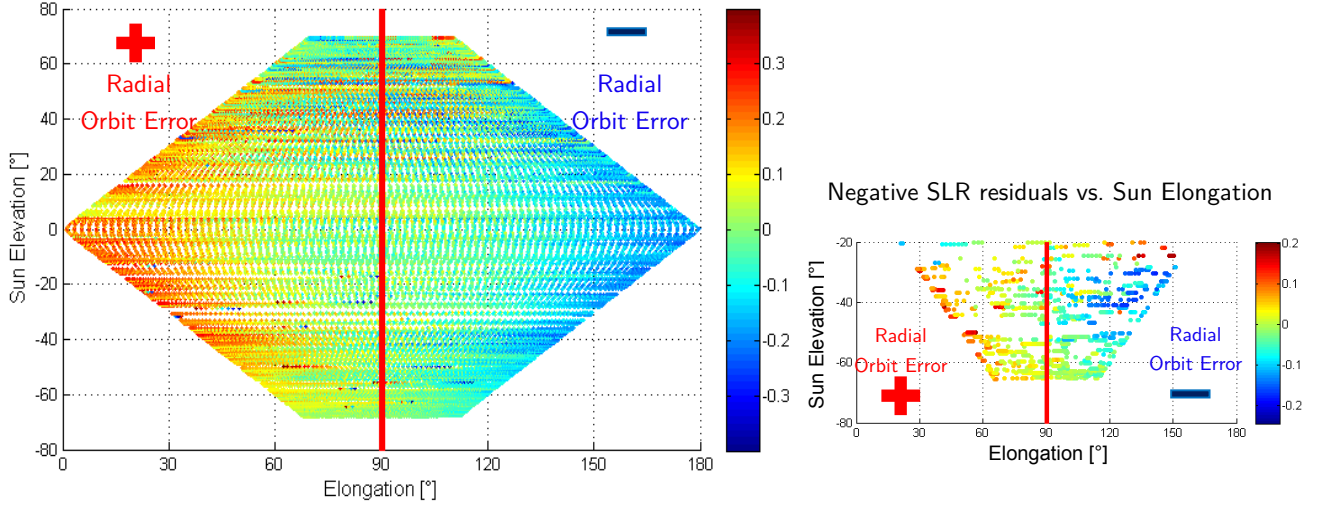
Figure 19.7 Residual clock parameters (left) and SLR residuals (right) in [m] against Sun elongation. Galileo E11 clock residuals follow the Sun elongation angle, i.e., the orientation of the solar array, as confirmed by independent SLR residuals (right) given with an opposite sign. Residual clock parameters are based on the Galileo clock solutions from TU München submitted to the MGEX Campaign of IGS.

Compared to SLR, Galileo clock parameters map the radial orbit error along the entire orbit, including when the Sun and satellite are in conjunction $\Delta u = 0°$. This is a strong argument to claim the orbit translation, not only for Galileo, but also for GPS and GLONASS. Due to the fact that SLR ranging to GNSS takes place typically at night, e.g., between 6 pm and 6 am local time, when the Sun is mainly in opposition to the satellite, SLR observes mainly that side of the GNSS orbit with a negative radial orbit error that is mapped as an artificial bias into the SLR ranges. The magnitude of this artificial negative SLR bias depends on the orbit quality and, therefore, should rather be called "GNSS orbit bias", instead of "SLR bias". For example, early Galileo orbits were showing a bias of 10 cm that dropped to some 5 cm when orbits improved by a factor of 2. This can also be seen in Figure 19.2 and Figure 19.3 where MGEX orbits from AIUB show a smaller amplitude compared to orbit/clock solutions from TU München. Partially, this orbit translation is also affected by albedo effects, although the net albedo effects tend to move the orbit in an opposite direction (towards the Sun) compared to our case. Related to albedo see (Ziebart et al. 2007), (Rodriguez-Solano et al. 2012). When LEO orbits are estimated using GNSS, any GNS orbit translation maps into the estimated LEO orbits.

Figure 19.8 shows a histogram of SLR residuals as a function of satellite argument of latitude Δu relative to the position of the Sun in the orbital plane. SLR residuals refer to two periods of about 50 days (days 69-131/2013 and 300/2013-52/2014) for AIUB orbit solutions showing a mean SLR bias of -6.5 cm. One can see that SLR measurements are not spread uniformly along the orbit and for Galileo the majority of SLR measurements are taken around midnight, whereas fewer SLR measurements are available for when Galileo satellites are closer to the Sun. Therefore, the mean SLR bias (orbit bias) $\overline{\delta}_{SLR}$ can be decomposed into one part due to an orbit modelling, accounting for mismodeling of e.g., solar radiation pressure $\overline{\delta}_{SLRmodel}$ by using e.g., (19.3) and a constant part along the orbit $\overline{\delta}_{const}$ generated by e.g., the antenna trust effect or constant part of the Earth's albedo.

$$\overline{\delta}_{SLR} = \overline{\delta}_{SLRmodel} + \overline{\delta}_{const} \tag{19.4}$$

If we now calculate the weighted average of SLR residuals in Figure 19.8, making use of (19.3)

$$\overline{\delta}_{SLR} = \frac{\sum n_i \cdot A \cos E_i}{\sum n_i} + \overline{\delta}_{const} \approx -4.1 \text{ cm} - 2.4 \text{ cm} = -6.5 \text{ cm} \tag{19.5}$$
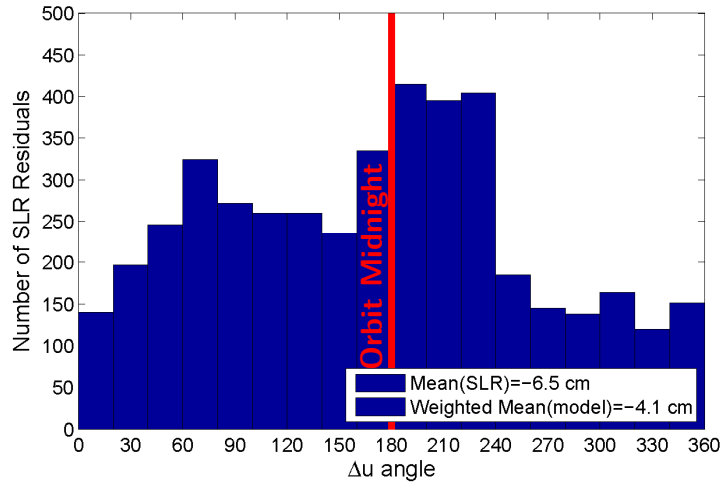
Figure 19.8 Histogram of SLR residuals based on the MGEX solution from AIUB (days 69-131/2013 and 300/2013-52/2014). One can see that the Galileo E11 orbit is observed when Sun and satellite are in opposition.

as a functional model and weighting by the number of measurements n_i in histogram bins shown in Figure 19.8. The weighted SLR bias (orbit bias) is $\overline{\delta}_{SLR} = -6.5\ \text{cm}$, giving an estimated constant SLR bias of $-2.4\ \text{cm}$. We will see in the next section that the Earth's planetary radiation contributes approx. $-14.6\ \text{mm}$ to the constant bias in the radial direction. For GPS we estimated this value to be about $-6.3\ \text{mm}$ for GPS Block-IIR and $-7.7\ \text{mm}$ for GPS Block-IIF. Considering the transmitted power of Galileo IOV satellites, our estimate of the Galileo trust effect is in the order of $-9\ \text{mm}$ to $-11\ \text{mm}$. These values are in line with (Ziebart et al. 2007) that also reported a constant effect of the Earth's albedo in the radial orbit error of GPS satellites at the cm-level and an antenna trust effect of $-5\ \text{mm}$ for GPS Block-IIF satellites.

We conclude this section by validating the derived empirical Galileo clock model in (19.3) with SLR measurements over all Sun elevations. Figure 19.9 shows SLR residuals (with an opposite sign) in the Sun-fixed orbital frame for two periods of about 50 days in 2013 and 2014 with rising Sun elevation. One can see very close agreement with the residual clock parameters displayed in Figure 19.2. Distinct asymmetry for rising and setting Sun elevations vs. orbit noon and midnight is consistent for both Galileo clock residuals and SLR.
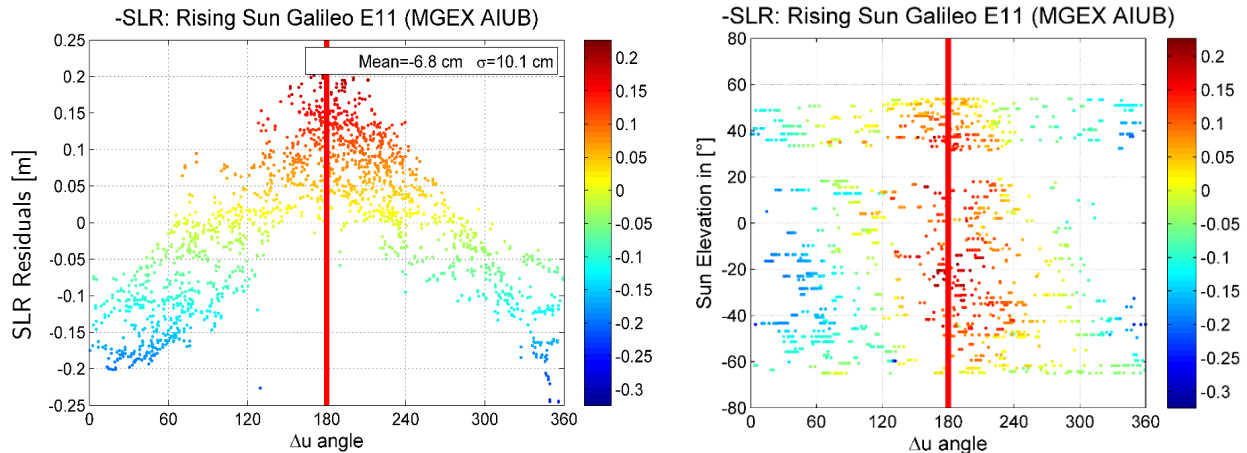
Figure 19.9 Galileo E11 SLR residuals (with negative sign) for rising Sun elevations based on the MGEX clock solutions from AIUB (days 69-131/2013 and 300/2013-52/2014). The figure on the left and on the right show negative SLR residuals relative to the satellite argument of latitude (relative to Sun position in the orbit frame). Notice a small asymmetry of residuals vs. orbit noon and midnight for SLR and Galileo clock residuals.

19.2 A Model of Solar Radiation Pressure Based on Galileo Clock Parameters and Circular Perturbations

Eq. (19.2) can be written as the radial perturbation equation in the form $A \cos(u + u_0)$, see (19.23), which is the general solution of the radial harmonic oscillator. Thus, we may use the following circular model to approximate the associated perturbations:

$$\vec{r} = \vec{c}_1 \cos nt + \vec{c}_2 \sin nt, \qquad \vec{c}_1 \perp \vec{c}_2, \ \ \left\| \vec{c}_1 \right\| = \left\| \vec{c}_2 \right\| = r \tag{19.6}$$

with two orthogonal vectors $\vec{c}_1$ and $\vec{c}_2$, the mean motion $n = 2\pi / P$ of the satellite and the orbit period $P \approx 14 \,\mathrm{h}$ for Galileo. The second time derivative is then

$$\ddot{\vec{r}} = -n^2 \left(\vec{c}_1 \cos nt + \vec{c}_2 \sin nt \right) = -n^2 \vec{r} \tag{19.7}$$

that gives circular radial orbit perturbation $\Delta \ddot{r} = -n^2 \cdot \Delta r$ assuming constant mean motion n. We may approximate $\Delta \ddot{r}$ with the radial component of acceleration due to mismodeled solar radiation pressure, as observed by clock residuals in (19.2). After substituting with (19.2), we obtain

$$\Delta \ddot{r} = -n^2 \cdot A \cdot \cos \beta \cos \Delta u \tag{19.8}$$

By introducing the elongation angle E (19.1), (see Figure 19.1), and after substitution into (19.8), we obtain the circular perturbation of the modelled clock residuals

$$\Delta \ddot{r} = -A \cdot \left(\frac{2\pi}{P} \right)^2 \cdot \cos \beta \cos \Delta u = A \left(\frac{2\pi}{P} \right)^2 \cos E \tag{19.9}$$

We now note that $-\cos \beta \cos \Delta u$ is the projection of the Sun unit vector $\vec{s}_\odot$ onto the Z-axis in the satellite body frame pointing radially inwards towards the geocenter. For all three components of the Sun unit vector

$$
\begin{aligned}
s_{\odot x} &= -\cos \beta \sin \Delta u \\
s_{\odot y} &= -\sin \beta \\
s_{\odot z} &= -\cos \beta \cos \Delta u
\end{aligned}
\tag{19.10}
$$

In an analogous way, similar to ROCK-type models where only X- and Z- directions are considered (Fliegel and Gallini 1996), we may define an orthogonal effect in the X-direction $\vec{x} = -(\vec{s}_\odot \times \vec{z}) \times \vec{z}$. As a result we may thus propose a perturbation model for both components in the satellite frame parameterized with two amplitudes A_x and A_z as follows

$$
\begin{aligned}
\Delta \ddot{r}_x &= A_x \cdot \left(\frac{2\pi}{P} \right)^2 \cos \beta \sin \Delta u = -A_x \left(\frac{2\pi}{P} \right)^2 \sin E \\
\Delta \ddot{r}_z &= A_z \cdot \left(\frac{2\pi}{P} \right)^2 \cos \beta \cos \Delta u = -A_z \left(\frac{2\pi}{P} \right)^2 \cos E
\end{aligned}
\tag{19.11}
$$

The amplitudes in (19.11) can be determined from the estimated clock parameters or estimated as parameters in the global GNSS solution. We typically remove daily time offset and drift from the Galileo clock parameters, thus (19.11) is a good approximation for the residual SRP acting along the satellite orbit (radial offset and drift removed). Eq. (19.2) or the form $\delta_{clk} = \Delta r = -A \cdot \cos E$, is a general solution of the radial harmonic oscillator and (19.11) is a good approximation that gives an order of magnitude of the total effect. The Galileo clock amplitude of $\approx 20 \,\mathrm{cm}$ shows a very close agreement in Figure 19.10 with the solution of Hill equations

(Colombo 1986). Small terms due to orbit velocity in Hill equations (Hill-Clohessy-Wiltshire equations) are not modelled and will affect the radial orbit (radial linear model removed), (Clohessy and Wiltshire 1960).

19.3 Thermal Re-Radiation Acceleration and Thermal Inertia of the Satellite

A satellite illuminated by the Sun experiences acceleration due to the absorption and reflection of photons on the exposed surface areas. This effect is commonly known as solar radiation pressure (SRP) and is dependent on the optical properties of the satellite surfaces. Solar radiation pressure is driven by the solar radiation intensity J_s that for a given distance d from Sun can be calculated as, see e.g., (Fortescue et al. 2011)

$$J_s = \frac{P}{4\pi d^2} \tag{19.12}$$

where P is the total power output from the Sun, or the solar flux 3.856×10^{26} W. At the Earth's mean distance from the Sun (1 AU) it is approx. 1371 ± 5 W/m^2 and often referred to as the Solar Constant. Since the satellite acceleration induced by solar radiation is proportional to the projected area exposed to the Sun (here denoted as A_c) and inversely proportional to the total mass m of the satellite, the SRP acceleration in satellite-Sun direction $\vec{e}_\odot$ is

$$\ddot{\vec{r}}_{SRP} = -C_{SRP} \frac{J_s}{c} \frac{A_c}{m} \vec{e}_\odot \tag{19.13}$$

where c is the speed of light in a vacuum and C_{SRP} the Solar radiation pressure coefficient describing optical properties of the satellite surface. Let us now define a normal to the surface $\vec{n}$ with an angle θ defined as $\cos\theta = \vec{n}^T \cdot \vec{e}_\odot$. We introduce the optical properties on the satellite surface by defining absorptivity, specular and diffusive reflectivity

- the specular reflectivity: $-2\rho_s \cdot \cos^2\theta \vec{n}$

- the diffusive reflectivity: $-\rho_d \cdot \cos\theta \cdot \vec{e}_\odot - \frac{2}{3}\rho_d \cdot \cos\theta \cdot \vec{n}$

- the absorptivity: $-\alpha \cdot \cos\theta \cdot \vec{e}_\odot$

with specular, diffusive and absorptivity coefficient $\rho_s + \rho_d + \alpha = 1$. From this we can derive an equation for the solar radiation pressure acceleration

$$\ddot{\vec{r}}_{SRP} = -\frac{J_s}{c} \frac{A_c}{m} \cos\theta \left[(1 - \rho_s)\vec{e}_\odot + 2(\rho_s \cos\theta + \frac{1}{3}\rho_d)\vec{n} \right] \tag{19.14}$$

In the case of solar arrays oriented towards the Sun

$$\ddot{\vec{r}}_{SRP} = -\frac{J_s}{c} \frac{A_c}{m} \left[1 + \rho_s + \frac{2}{3}\rho_d \right]\vec{e}_\odot = -C_{SRP} \frac{J_s}{c} \frac{A}{m} \vec{e}_\odot \tag{19.15}$$

where $C_{SRP} = 1 + \rho_s + \frac{2}{3}\rho_d$. A similar expression can be found in (Milani et al. 1987).

Solar radiation acceleration is typically estimated as part of orbit determination by utilizing the widely used CODE SRP model (Beutler et al. 1994). As a function of argument of latitude u

$$D(u) = D_0 + D_c \cos(u) + D_s \sin(u)$$
$$Y(u) = Y_0 + Y_c \cos(u) + Y_s \sin(u) \qquad (19.16)$$
$$B(u) = B_0 + B_c \cos(u) + B_s \sin(u)$$

The CODE SRP model (19.16) defines estimated empirical acceleration in the satellite-Sun direction $D(u)$, along the solar panel axis $Y(u)$, and $B(u)$ completes the orthogonal triad. Typically, the CODE 5-parameter version is used where of the nine empirical parameters in (19.16) only the direct accelerations D_0, Y_0, B_0 are estimated, along with two periodic components B_c and B_s. The remaining four amplitudes D_c, D_s, Y_c and Y_s in (19.16) are either not estimated or constrained in the orbit determination. It is neither well known nor available in the relevant literature, but due to variable satellite-Sun distance d along the orbital plane, all nine SRP parameters in (19.16) are scaled to one Astronomical Unit (1 AU), making use of the scaling factor $(1 \text{ AU} / d)^2$.

The SRP acceleration is induced by incident solar radiation due to the exchange of momentum with the satellite surface depending on how much power is absorbed or reflected either diffusely or specularly by the satellite surface. This exchange of momentum depends also on the nature of the Sun radiation. Since a satellite is not a black body, it absorbs only a fraction of the incident Sun energy (absorptance α). The actual temperature T of the satellite surface will cause infrared re-radiation emission at thermal infrared wavelengths generating thermal re-radiation intensity according to Stefan-Boltzmann's law

$$J_{radiated} = \varepsilon \cdot \sigma \cdot T^4 \qquad (19.17)$$

where ε denotes the emittance and σ the Stefan-Boltzman constant $5.67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4}$, see (Fortescue et al. 2011). With the effective area of the satellite for absorbing A_α and for emitting A_ε, with no internal heat dissipation, the equilibrium temperature T is achieved when absorbed thermal flux q_α and emitted thermal flux q_ε are equal, $q_\alpha = q_\varepsilon$

$$A_\alpha \cdot \alpha \cdot J_s = A_\varepsilon \cdot \varepsilon \cdot \sigma \cdot T^4 \qquad (19.18)$$

For a given ratio between absorptance and emittance α / ε which mainly depends on the surface color, one can calculate the equilibrium temperature T at the exposed satellite surface.

According to the ESA News of 11.7.2013, each of the solar arrays in the pair on board a Galileo satellite is 1×5 m in size and consist of more than 2500 state-of-the-art gallium arsenide (GaAs) solar cells. This type of solar cells is also used on GPS Block-IIF satellites, see Table 19.1 and other ESA satellite missions, such as Rosetta. Typical values for absorptance and emittance for GaAs solar cells can be found in the relevant literature, e.g., (Fortescue et al. 2011), and are $\alpha = 0.88$ and $\varepsilon = 0.80$. For the ratio between absorptance and emittance for the Galileo solar arrays this gives $\alpha / \varepsilon = 1.10$. For the black paint that is typically used for the satellite body one obtains 1.16. At the distance of 1 AU for Galileo solar arrays this gives an equilibrium temperature of $T = 339.60$ K or $T = 66.45\,^\circ$C. This is based on the assumption that $A_\varepsilon / A_\alpha = 2$, as a first approximation, it was assumed that both the front and the rear side of the Solar array are radiating equally. When the Sun is in the orbital plane, the max. difference in temperature along the orbit (between orbit noon and orbit midnight) is only $0.07\,^\circ C$.

Since the satellite acceleration due to thermal re-radiation is proportional to the area of the radiating satellite body surface A_ε and inversely proportional to the total mass m of the satellite, the final expression for thermal re-radiation acceleration in the Sun-satellite direction can be derived from the emitted thermal flux Q_ε

$$\ddot{r}_t = -\frac{2}{3}\frac{Q_\epsilon}{c}\frac{A_\epsilon}{m} = -\frac{2}{3}C_{ther}\epsilon_f T_f^4 \qquad (19.19)$$

where T_f denotes the equilibrium temperature at the satellite body surface considering only Lambertian diffuse reflectivity and neglected specular reflectivity. Thus, the factor 2/3 in (19.19) comes from Lambert's cosine law integrated over the whole hemisphere. Following (Rievers et al. 2009), if the radiating surface is an ideal radiator, the radiation pattern is hemispheric and the distribution of intensity over the hemisphere can be expressed by Lambert's cosine law. We define the thermal coefficient C_{ther} in the following way

$$C_{ther} = \frac{A_\epsilon}{m} \frac{\sigma}{c} \qquad [\mathrm{m} / (\mathrm{s}^2 \mathrm{K}^4)] \tag{19.20}$$

In the case of solar panels we need to account for the thermal re-radiation from both sides of the solar panels, i.e., a difference between emitted thermal flux from the front and the rear side of the solar panel $Q_{\epsilon f} - Q_{\epsilon r}$

$$
\begin{aligned}
\ddot{r}_{ther} &= \ddot{r}_{ther}^{front} + \ddot{r}_{ther}^{rear} \\
&= -\frac{Q_{\epsilon f} - Q_{\epsilon r}}{c} \frac{A}{m} = -C_{ther}(\epsilon_f T_f^4 - \epsilon_r T_r^4)
\end{aligned}
\tag{19.21}
$$

where T_r denotes the temperature of the rear side of the solar panel. Following (Fortescue et al. 2011), absorptance of a satellite surface illuminated by solar radiation that has a peak intensity at about 0.45 μm in the optical part of the spectrum has a corresponding emittance of a surface radiating in the infrared region with peak intensity at about 10 μm in the infrared spectrum.

Since the heat flow through the typical honeycomb core structure of the solar arrays occurs by conduction only, the thermal emission properties of solar arrays are practically unaffected by outgassing and radiation of the heat flux through the cavities within the solar array core. Radiation is the main mode of heat transfer in a vacuum and thus in space. Therefore, we may calculate the temperature difference between the front and the rear side of the solar array ΔT due to the heat flow from the warmer front panel to the colder rear panel knowing the conductive heat flow rate Q_c that is equal to the absorbed thermal flux $A_\alpha \cdot \alpha \cdot J_s$ in (19.18)

$$Q_c = \frac{\lambda A_c}{l} \Delta T = h_c(T_f - T_r) \tag{19.22}$$

where h_c is the thermal conductance as a function of cross-sectional area A_c, l the conductive path length (approx. thickness of the solar array) and λ the thermal conductivity. Since the GaAs solar cells are also used on other ESA missions, such as Rosetta, we assumed that the inner core of the Galileo solar arrays consist of a thin honeycomb structure made of aluminum (Al), whereas the external front and rear solar array surfaces are made of Carbon Fiber Reinforced Plastic (CFRP), with the front surface being covered by the GaAs solar cells. Typical thermal conductivity for Al-honeycombs as $\lambda = 109 + 0.245 \cdot (T_a - 273.15)$ where T_a is the mean temperature $T_a = (T_f - T_r) / 2$. Emissivity of CFRP surface have a strong temperature dependency that is empirically given as $\epsilon = 0.312 + 0.003288 \cdot T - 0.00000533 \cdot T^2$.

Figure 19.10 shows estimated Galileo clock parameters after removing linear clock model (time offset and drift) against the effect of the thermal re-radiation acceleration of the satellite body in the radial direction. For this calculation we used analytical Hill equations for the radial orbit direction given in (Colombo 1986) perturbed by the analytical effect of thermal re-radiation from this section. For the calculated radial perturbation we removed offset and drift in order to be comparable to the Galileo clock parameter. Figure 19.10 shows very close agreement between both analytical effects without any parameter estimation. For the small asymmetry with the orbit non-midnight direction we used thermal inertia of 4.7 min.

Solar radiation pressure for orbits of GNSS satellites is mainly driven by the large solar panels. Since these are relatively thin, the main component of the thermal re-radiation of Solar panels at infrared wavelengths act in the opposite direction to that of the solar radiation pressure. Considering that the same
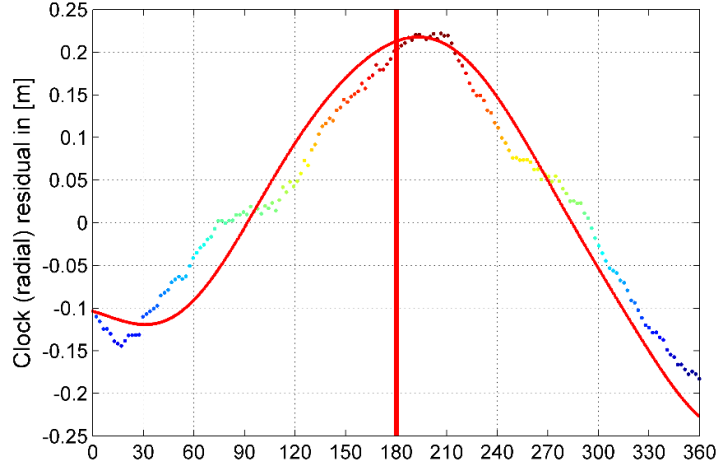
Figure 19.10 Estimated Galileo clock parameters (linear model removed) (dotted line) against effect of the thermal re-radiation (red) in radial direction calculated using analytical orbit Hill equation, (Colombo 1986). For the thermal inertia, a small asymmetry with the orbit non-midnight direction, we used a value of 4.7 min. Note the size of the amplitude ≈ 20 cm in the radial direction that is similar to our simple model $A \cdot \cos E$.

area of solar panels is illuminated by the Sun along the orbit, both solar radiation and related thermal re-radiation of solar panels generate a net force along the GNSS orbit that is removed by the estimated CODE 5-parameter model. This is not the case with the thermal re-radiation of the satellite body that when heated by the Sun generates a re-radiation force acting in the same direction as the solar radiation pressure, but with a delay needed to heat the surface. This is so called thermal inertia or Yarkovsky effect, often associated with the orbital dynamics of asteroids, see (Chesley et al. 2003). Thus, once illuminated by the Sun, the satellite surface will warm up after some delay and stay warmer even after pointing to the Sun. This afternoon side of the satellite is hotter and thus will generate thermal re-radiation acceleration that is away from the Sun-satellite direction and not co-linear with the SRP acceleration. Due to the size of the satellite body, the solar radiation pressure is significantly smaller compared to the thermal re-radiation for the satellite body. Because of the time lag, the net effect due to thermal re-radiation is not collinear with the direction of solar radiation pressure and we see an asymmetric effect when comparing rising and setting Sun elevations for orbit noon and midnight. This thermal inertia of the satellite or the Yarkovsky effect, can be confirmed with Galileo clock residuals and SLR residuals plotted against the satellite argument of latitude in all figures in this section, see e.g., Figure 19.2 or Figure 19.7.

The Yarkovsky effect was first claimed for the asteroid 6489 Golevka tracked by the Arecibo radio telescope in 1991–2003. The asteroid drifted 15 km from its predicted position over 12 years, (see Science paper (Chesley et al. 2003)). An illuminated object, or a Solar array and a satellite body in our case, takes some time to become warm when illuminated and to cool down when this illumination stops. Recently (Turyshev et al. 2012) have claimed that the anomalous acceleration of the Pioneer 10 and 11 (Pioneer anomaly) is due to the recoil force associated with an anisotropic emission of thermal radiation from these vehicles.

In (Lucchesi et al. 2004), a part of the total Yarkovsky effect is analyzed for the LAGEOS-2 satellite (called the Yarkovsky–Schach effect) that is modulated only during the eclipse passages through the Earth's shadow. For satellites that are rapidly spinning, such as LAGEOS-2, one can assume a latitudinal distribution of temperature across the satellite surface, and therefore, the thermal re-radiation acceleration is directed along the satellite spin axis. Due to the absence of solar radiation in the eclipse passages, and associated change in the surface temperature, the finite thermal inertia of the spinning satellite produces a small change in the thermal re-radiation acceleration along the spin axis. This gives rise to a non-null along-track acceleration along the orbit revolution and associated long-term effects in the satellite semimajor axis (Lucchesi et al. 2004). (Rubincam 1987) discusses a similar thermal inertia effect for the rapidly spinning LAGEOS-2 satellites

due to the Earth's infrared radiation, causing a net force along the direction of the spin axis. This effect is often called the Earth-Yarkovsky or Rubicam effect, see e.g., (Lucchesi et al. 2004).

However, GNSS satellites or typical gravity or altimetry missions in the polar Earth orbits, do not rapidly spin as does the LAGEOS-2 satellite. This is a significant factor, as one cannot assume a latitudinal distribution of temperature across the satellite surface and easily distinguish between the "cold" and the "hot hemisphere" for a spherical approximation of satellite surface. In this case, thermal re-radiation acceleration is fixed in the inertial space, relative to Sun direction. For nadir- pointing satellites typical rotation is associated with one orbital period. For GNSS satellites one should also consider yaw steering along the nadir direction with typical oscillations from β to $180^{\circ} - \beta$ outside the fixed yaw-steering regime when Sun elevation is close to zero.

In the general case of a spinning satellite both a spin component and an equatorial component of the acceleration are present. The recoil acceleration for a spinning satellite is generated by the imbalance of the temperature distribution across the satellite surface and directed along the satellite spin axis, away from the colder pole. As soon as the spinning satellite is in full sunlight, i.e., in the absence of eclipses, the along-track acceleration at a given point of the orbit is compensated by an equal and opposite acceleration at the opposite point of the orbit, giving a resultant null acceleration over one orbital revolution.

Since all GNSS satellites are pointing towards the Earth, there will always be a component of thermal re-radiation in the radial orbit direction as a function of relative Sun argument of latitude $\cos \Delta u$ that is not removed by the estimated CODE 5-parameter model. The estimated CODE 5-parameter model removes only solar radiation/re-radiation pressure of the solar panels constantly oriented towards the Sun. Since $\cos(0^{\circ}) = -\cos(180^{\circ})$, we get the maximum effect of the thermal re-radiation of the satellite body in the radial direction when Sun and satellite are in opposition $\Delta u = 180^{\circ}$, and the minimum at $\Delta u = 0^{\circ}$ when they are in conjunction. Thus, the net effect translates the orbit away from the Sun. Satellite payloads also generate heat within the satellite and radiators placed on the satellite surface channel this heat outside the satellite. However, they are typically placed symmetrically to each other along the Y-axis (Solar panel axis) of the satellite. Therefore the net thermal effect of the internal heat dissipation is zero and with appropriate thermal design should not have a significant effect on the satellite orbit.

Figure 19.11 graphically depicts the Yarkovsky effect on satellite orbit around the Earth. Radiation from the Sun heats the satellite body on the nearest side to the Sun (orbit noon). The net effect in the along-track direction accelerates the satellite in Sun-satellite opposition and slows it down in Sun-satellite conjunction. This can be geometrically measured in the radial direction by the Galileo H-maser. One can distinguish the Yarkovsky effect at orbit period in Figure 19.11 and at draconic period between rising and setting Sun elevations in Figure 19.12. The GNSS draconic year is the repeat period of the GNSS constellation w.r.t. Sun which is approximately 351 days for Galileo and 357 days for Galileo.

For rising and setting Sun elevations, Galileo radial orbit error or residual clock parameters can be approximated by

$$
\begin{aligned}
\Delta_{clk} = \Delta r_t = A_t \cos \beta \cos(u - u_{\odot}) & \qquad \text{Rising Sun} \\
\Delta_{clk} = \Delta r_t = A_t \cos \beta \cos(u - u_{\odot} - 180^{\circ}) & \ \ \text{Setting Sun}
\end{aligned}
\tag{19.23}
$$

where $u_{\odot}$ denotes the argument of latitude of the Sun's ascending node on the satellite orbit plane. By introducing a time lag for the thermal inertia α, the clock model Δ_{clk} is then given for rising and setting Sun elevations

$$
\begin{aligned}
\Delta_{clk} = \Delta r_t = A_t \cos \beta \cos(u - u_{\odot}) & \qquad \text{Rising Sun} \\
\Delta_{clk} = \Delta r_t = -A_t \cos \beta \cos(u - u_{\odot} + \alpha) & \quad \text{Setting Sun}
\end{aligned}
\tag{19.24}
$$
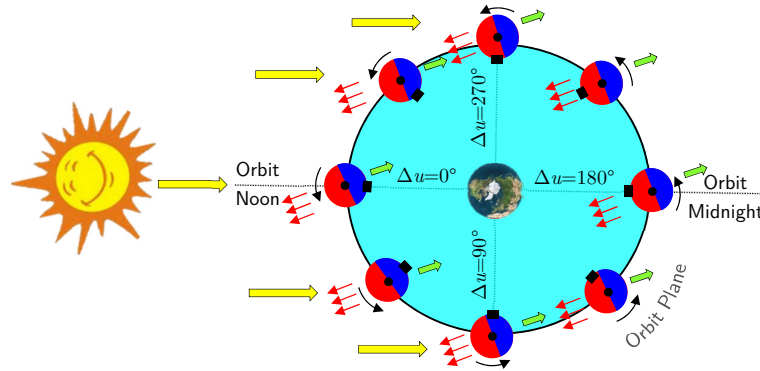
Figure 19.11 Yarkovsky effect on a spherical, nadir-pointing satellite in a prograde orbit around the Earth in the Sun-fixed orbital coordinate system. Due to thermal inertia, the maximum of the surface temperature (red) and subsequently its thermal radiation acceleration (green arrows) is displaced from the Sun-satellite direction. The hotter side of the satellite (red) is the afternoon side (past the orbit noon) that re-radiates most of the absorbed solar radiation (red arrows). As long as the satellite is in sunlight, the effect will result in zero net acceleration over one orbital revolution, since the projection of thermal acceleration in the along-track orbit direction at any given point along the orbit will be compensated by equal and opposite acceleration at the antipodal point of the orbit. When Sun and satellite are in opposition, $\Delta u = 180°$, additional along-track acceleration increases the satellite velocity, whereas it is compensated by an equal and opposite accelerations at the orbit Sun/satellite conjunction $\Delta u = 0°$, where it is opposite to the along-track velocity. Thus, the net effect along the orbit will result in translation of the orbit in the Sun-satellite direction away from the Sun.
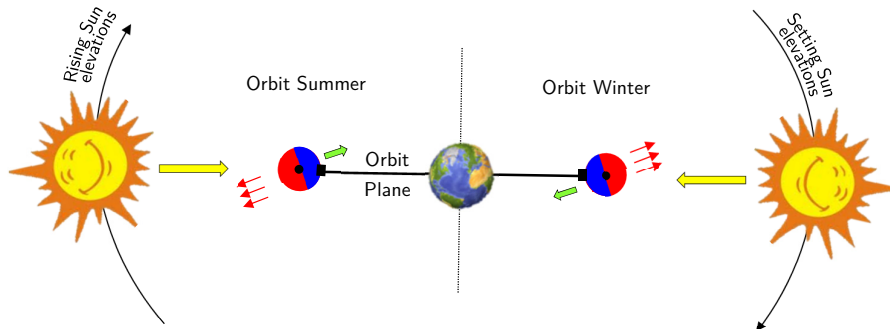
Figure 19.12 Yarkovsky effect on a spherical nadir-pointing satellite in orbit around the Earth due to rising and setting Sun elevations over one draconic period (357 days for Galileo). Due to thermal inertia, the maximum of the surface temperature (red) and subsequently its thermal radiation acceleration (green arrow) is displaced from the Sun-satellite direction. This seasonal Yarkovsky effect between orbit summer and orbit winter is equivalent to the Yarkovsky effect with orbit revolution between orbit noon and orbit midnight. The hotter side of the satellite (red) is the summer side of the orbit (rising Sun elevations) that re-radiates the most of the thermal radiation (red arrows). As long as the satellite is in sunlight, the effect will result with null acceleration over one Sun draconic period. Projection of thermal acceleration to the radial orbit direction at any given orbit noon along the summer orbit will be compensated by an equal and opposite acceleration in the antipodal point of the winter orbit. When Sun and satellite are in opposition $\Delta u = 180°$, additional radial acceleration increases the satellite velocity whereas it is compensated by an equal and opposite accelerations in the orbit Sun/satellite conjunction $\Delta u = 0°$, where it is opposite to the along-track velocity. Thus, the net effect along the orbit will result in the orbit rotation along the orbital plane direction.

19.4 Planetary radiation of the Earth

The Earth and other planetary bodies in the Solar System have non-zero temperature. Therefore, in addition to thermal flux due to solar radiation intensity J_s given by (19.12) there is also the planetary radiation of the Earth to be considered. This has a wavelengths in the infrared spectrum between 2 and $50\ \mu\mathrm{m}$, exhibiting peak intensity around $10\ \mu\mathrm{m}$, and is generated by the whole cross-sectional area of the Earth. Intensity of planetary radiation J_p is a function of orbit altitude R_{orbit} and is given by

$$J_p = 237 \cdot \frac{R_p}{R_{orbit}} \tag{19.25}$$

where R_p is the radius of the effective radiating surface, and in the case of the Earth can be approximated by the mean Earth's radius. For the Galileo orbit altitude one can find $J_p = 11.0037$ and for GPS $J_p = 13.7292$. This corresponds to about 0.8% of the solar intensity for the Galileo orbit and 1.0% for the GPS orbit at 1 AU from (19.12). Estimated radial orbit bias is given in Table 19.1 for Galileo and GPS satellites calculated using satellite properties available from http://www.gps.gov and http://www.gsa.europa.eu/galileo/programme.

From Table 19.1 one can see that the orbit bias $\Delta r = -14.6$ mm of calculated Earth's radiation for Galileo satellites is in a very good agreement with the mean bias in SLR residuals that is in the order of -2.4 cm. The remaining bias of -9.4 mm is close to the estimated antenna trust effect, see previous section.

For solar arrays, the effect of the Earth's radiation is strongly dependent on the cross-sectional area of the solar arrays in the nadir direction and the orientation of the solar arrays. This relationship can be modelled using the elongation angle E

$$\Delta r = A_s \cdot \cos E \tag{19.26}$$

In order to calculate the amplitude A_s one also needs to take into account the emittance of the rear side of the solar panel.

	Galileo	GPS Block-IIR	GPS Block-IIRM	GPS Block-IIF	GPS III
Mass	696.815 kg	1127 kg	1127 kg	1465 kg	2161 kg
Solar array	2×5×1.58 m²	13.4 m²	13.4 m²	13.4 m²	13.4 m²
Nadir surface	(3.02-0.18)×1.58 m	1.57×2.21 m	1.57×2.21 m	2.49×2.24 m	2.49×2.24 m
Earth's thermal radiation for Solar array and satellite body					
Absorptance α (nadir)	(0.95)	(0.95)	(0.95)	(0.95)	(0.95)
Emittance ε (nadir)	(0.82)	(0.82)	(0.82)	(0.82)	(0.82)
Δr (nadir surface)	14.6 mm	6.3 mm	6.3 mm	7.7 mm	7.9 mm
Absorptance α (nadir)	(0.88) GaAs cells	(0.75) silicon cells	(0.75) silicon cells	(0.88) GaAs cells	(0.88) GaAs cells
Emittance ε (nadir)	(0.80) GaAs cells	(0.82) silicon cells	(0.82) silicon cells	(0.80) GaAs cells	(0.80) GaAs cells
Δr (Solar array) [mm]	$23.8 \cdot \cos E$	$9.5 \cdot \cos E$	$9.5 \cdot \cos E$	$8.6 \cdot \cos E$	$5.8 \cdot \cos E$

Table 19.1 Effect of planetary thermal radiation on Galileo and GPS satellite orbits in the radial direction. Assumed values used in the calculation are in brackets.

19.5 Galileo Clock Parameters and Attitude

According to the description of Galileo satellite parameters provided on the ILRS homepage, GIOVE-A, GIOVE-B and the Galileo satellites follow the yaw steering law. The satellite body +Z axis points continuously to nadir (as in GPS), and a rotation performed around the Z axis maintains the satellite +Y axis perpendicular to the Sun. The +X spacecraft panel is maintained away from the Sun. From the information provided on the ILRS homepage it follows that the +Y axis has the opposite sign to the +Y axis of GPS Block II/IIA satellites, i.e., the axis definition for Galileo is the same as for GPS Block IIR satellites, (see IGSMail#16353 for a description of Block IIR satellites). The Galileo clock residuals and Figure 19.3 show that the accumulated carrier-phase due to the antenna wind-up is similar to that for a GPS orbit, indicating that orientation of the yaw steering for Galileo is the same as for GPS. Since the clock residuals nicely match the SLR residuals, we may draw the conclusion that the Galileo wind-up effect was correctly calculated and that the assumptions used in the Galileo axis definition and attitude law are correct. In addition, the ILRS homepage states, "As with GIOVE-A, it is foreseen that the theoretical attitude will not be achieved at times where the beta angle is small, due to limitations in the reaction wheels and yaw measurement (Sun co-linearity)". This is similar to GPS Block IIR satellites. According to IGSMail#1653, it was reported that for low Sun elevations, $-1.6° < \beta < 1.6°$, GPS IIR satellites switch from yaw steering to a fixed yaw mode. This transition happens at orbit dusk and in this mode the yaw angle is fixed, i.e., the X and Z axes are in the orbital plane, while the +X points approximately in the direction of the velocity vector (axes definition for GPS Block IIR), (see IGSMail#1653). In the case of Galileo, the yaw steering algorithm was presented in (Gonzalez 2010), where it was indicated that the beta angle, at which yaw steering is switched to the fixed yaw mode is below 2°. Figure 19.13 shows the clock residuals of the Galileo E11 satellite during fixed yaw steering (Sun elevation $\beta = 0°$). Since the antenna wind-up effect was calculated for nominal yaw steering, one can clearly see a jump at $\Delta u = 180°$ and a slightly smaller jump at $\Delta u = 0°$. This indicates that during fixed yaw steering (Sun elevation $\beta = 0°$), the satellite rotates by 180° about the Z axis at $\Delta u = 180°$, i.e., it makes a turn in the yaw angle from 0° to 180° over an interval in the argument of latitude of about 15°. A rotation by 180° in the yaw angle corresponds to the wind-up effect of half of the narrow-lane wavelength, and this is mapped into estimated satellite clock parameters. Figure 19.13 indicates that this yaw rotation turn also takes place at $\Delta u = 0°$, in the opposite direction and is less visible.
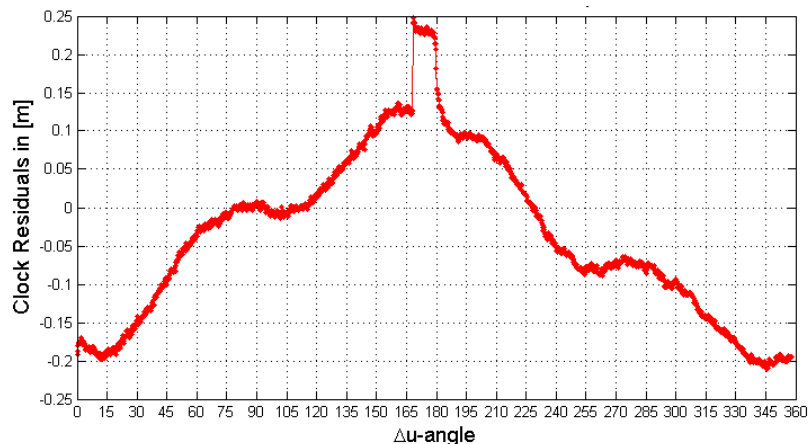
Figure 19.13 Galileo E11 clock residuals against the argument of latitude relative to the Sun's position. One can see a clear jump at $\Delta u = 180°$ and a slightly smaller jump at $\Delta u = 0°$ of about half a narrow-lane wavelength, indicating that during fixed yaw steering (Sun elevation $\beta = 0°$) the satellite rotates approx. 180° about the Z axis, i.e., a turn in the yaw angle from 0° to 180° over an argument of latitude of $\approx 15°$. The antenna wind-up was based on nominal yaw steering. A rotation of 180° in yaw corresponds to half a narrow-lane wavelength.

Calculation of the antenna wind-up effect was based on nominal yaw steering. Figure 19.13 shows the clock residuals of the Galileo E11 against the Sun's elevation above the orbital plane and the argument of latitude of the satellite relative to the Sun's position. One can see that at low Sun elevations the clock residuals experience higher variations, which are most likely due to eclipses. For GNSS, eclipse periods take place when $-14° < \beta < 14°$. For Galileo, the eclipse interval is slightly narrower, i.e., $-12° < \beta < 12°$ due to the higher orbit altitude. The angle of $12°$ is the angle of the Earth's radius as seen from the Galileo orbit altitude.

19.6 Comparison with a Thermal Re-Radiation Model for GPS Satellites at Low Sun Elevations

By inserting the mass of the GPS PRN06 (975 kg) and the model of solar radiation pressure (19.11) we obtain for the force due to solar radiation pressure

$$f = 0.10 \cdot \left(\frac{2\pi}{P}\right)^2 \cdot \cos E \cdot 975 = 0.21 \cdot \cos E \qquad [10^{-5}\text{N}] \tag{19.27}$$

in units of $[10^{-5}\text{N}]$. Comparison of (19.27) with the T30 thermal re-radiation model of ROCK-type Solar radiation pressure models, reveals parameterization with the elongation angle E that is similar to our model (19.11). Following (Fliegel and Gallini 1996), for BLOCK IIR GPS satellites, the T30 model including thermal re-radiation in the X and Z directions of a satellite body-fixed system is

$$\begin{aligned} f_Z &= -11.3 \cos E + 0.1 \cos 3E + 0.2 \cos 5E \\ f_X &= -11.0 \sin E - 0.2 \sin 3E + 0.2 \sin 5E \end{aligned} \tag{19.28}$$

in units of $[10^{-5}\text{N}]$, as a function of the elongation angle only. Explanation for the frequencies $3E$ and $5E$ is not given in (Fliegel and Gallini 1996). Similar parameterization to the T30 thermal re-radiation model was presented in (Bar-Sever and Kuang 2004). Note that the Z direction for GNSS satellites is a negative radial direction, hence the change in sign compared to our model (19.27). The 5 or 9 standard CODE solar radiation pressure parameters are not suited to absorbing an effect that varies significantly with the Sun $\beta -$ angle. A variation in Sun elevation by one degree will generate an additional acceleration at the $\approx 10^{-9}$ m/s^2 level that can explain the effect in (19.11). This is why GPS and Galileo orbits are at their most accurate levels when the Sun is high above the orbital plane. We can draw the conclusion that 9 CODE solar radiation pressure parameters should be used in addition to our thermal re-radiation model, or the parameters of the T30 model in (19.28) should be estimated empirically in addition to the CODE Solar radiation model. Considering the single term in (19.11), one could estimate empirically two additional frequencies such as

$$\Delta \ddot{r} = A \left(2\pi/P\right)^2 \cos E + A_3 \left(6\pi/P\right)^2 \cos 3E + A_5 \left(10\pi/P\right)^2 \cos 5E \tag{19.29}$$

in order to properly model the "side lobes" at lower Sun elevations, (see Figure 19.14). At low Sun elevations, the amplitude of the twice-per-rev. frequency is more visible, due to the high Sun beta angle in Figure 19.14.
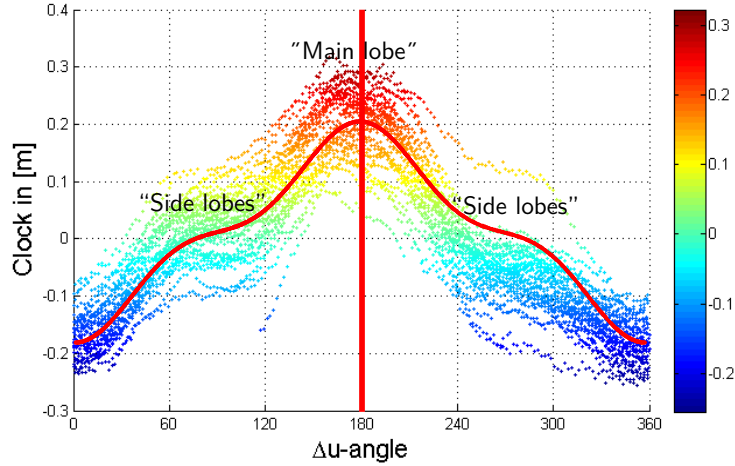
Figure 19.14 Clock parameters of the Galileo E11 satellite at low Sun elevations $<12°$. At these Sun elevations, the amplitudes of the ″side lobes″ due to the $3E$ and $5E$ frequency are more visible.

19.7 Solar Wind Pressure and its Symmetry with Solar Radiation Pressure

Solar wind pressure has not been considered so far in precise orbit determination. However, with increasing orbit accuracy, this effect is becoming more interesting. Here we derive a theoretical model of solar wind pressure and discuss its application in precise orbit determination.

With several groups performing POD of SLR satellites, (Ciufolini et al. 2012) reports that the recently launched LARES satellite shows the smallest deviations from a geodesic motion of any artificial satellite, i.e., its residual mean acceleration away from a geodesic motion is less than 0.5×10^{-12} m/s^2 after modelling non-gravitational perturbations. When talking about orbit modeling at the $10^{-12} - 10^{-13}$ m/s^2 level, the effect of solar wind pressure becomes far more interesting, not only for LARES, but also for GPS and Galileo satellites with very large cross-section-to-mass ratios (form factor) and long orbit arcs.

Analogous to solar radiation pressure due to the photon flux from the Sun that propagates at the speed of light, we may consider, in addition, pressure due to solar wind that propagates at velocities $v_p = 400 - 800$ km/s. According to (Feldman et al. 2005), solar wind has two components: slow solar wind with a velocity of about 400 km/s and a composition similar to the Sun's corona; and fast solar wind with a typical velocity of 750 km/s and whose composition nearly matches that of the Sun's photosphere. The slow solar wind is twice as dense as the fast solar wind. Solar wind is believed to originate very close to the Solar surface, but since it is accelerated significantly above the solar surface, its velocity cannot be correlated with remote observations to trace its origin, (Feldman et al. 2005). Sun particles travelling at a velocity of $400 - 750$ km/s reach the Earth after about $2.2 - 4.4$ days from an apparent direction that is $2.2° - 4.4°$ away from the Earth-Sun direction. Satellites such as Ulysses (ESA) or ACE (NASA) at the L1 Lagrangian point (about 1 Mkm away from the Earth towards the Sun) measure the speed of the solar wind and the number of protons per cm^3. Taking into account the mass of a proton $m_p = 1.672621777(74) \times 10^{-27}$ kg and the number of protons n_p per cm^3, we obtain the pressure of the solar wind as a function of solar wind velocity v_p in [km/s] and proton density n_p (number of protons per cm^3) given in nPa

$$P_\odot = m_p \cdot n_p \cdot v_p^2 \tag{19.30}$$

or

$$P_{\odot} = 1.6726 \cdot 10^{-7} \cdot n_p \cdot v_p^2 \quad [n\mathrm{Nm}^{-2}] \tag{19.31}$$

where $\frac{1}{2} m_p v_p^2$ is the kinetic energy of a single particle. Introducing the effective cross-sectional area A divided by the mass of the satellite m, or the form factor of the satellite A/m, we can obtain the acceleration of the satellite due to the force exerted by the solar wind

$$a_p = -2 \cdot m_p \cdot \frac{A}{m} \cdot n_p \cdot v_p^2 \tag{19.32}$$

or

$$a_p = -2 \cdot 1.6726 \cdot 10^{-16} \cdot \frac{A}{m} \cdot n_p \cdot v_p^2 \quad [\mathrm{nm/s^2}] \tag{19.33}$$

where the factor of 2 arises when there is pure absorption. We may introduce the solar wind pressure coefficient C_w similar to that for solar radiation. The acceleration of a satellite due to solar wind can then be defined as

$$a_p := -m_p \cdot C_w \cdot \frac{A}{m} \cdot n_p \cdot v_p^2 \tag{19.34}$$

or

$$a_p := -1.6726 \cdot 10^{-16} \cdot C_w \cdot \frac{A}{m} \cdot n_p \cdot v_p^2 \tag{19.35}$$

As an example, for a wind speed of $400 \ \mathrm{km/s}$, $n_p = 20 \ \mathrm{protons/cm^3}$, and pure absorption $C_w = 2$ we obtain

$$a_p = -2.7 \cdot 10^{-12} \cdot \frac{A}{m} \quad [\mathrm{m/s^2}] \tag{19.36}$$

For the form factor in the order of $A/m = 0.02$ for GNSS satellites, we get an effect in the order of about $a_p = -0.5 \cdot 10^{-13} \ \mathrm{m/s^2}$. One should bear in mind that the effect of solar wind pressure is very systematic, i.e., it does not average out and, in our example, is about $4.4°$ away from the Sun's direction.

20. Track-to-Track Ambiguity Resolution for Zero-Differences – Integer Phase Clocks

In this section we introduce a novel approach for GNSS ambiguity resolution at the zero-difference level, what we call Track-to-Track (T2T) ambiguity resolution. The T2T approach is based on the resolution of wide-lane and narrow-lane ambiguities between consecutive satellite tracking passes, what we call track-to-track or pass-to-pass ambiguities. To fix T2T ambiguities to their integer values, GNSS measurements from only a single GNSS receiver are used without forming any double-differences or similar combinations between different GNSS receivers. Thus, the T2T approach is especially appropriate for LEO applications, to connect very short tracking passes (typically $15 - 20$ min) that introduce a very large number of zero(double)-difference ambiguities, and for ground networks, where the ambiguities of a single GNSS satellite can be connected over a longer period of time (e.g., one week). This opens up a new application for T2T ambiguities to monitor stability and to define code biases and GNSS clock parameters over a long period of time. In this section, we demonstrate the T2T ambiguity resolution approach using LEO and ground GPS measurements. We show that LEO T2T ambiguity resolution leads to an optimal combination of LEO and ground GPS measurements and thus opens doors to form a network of LEO satellites in space for the determination of combined GNSS/LEO terrestrial reference frame parameters. This is possible thanks to the connected LEO ambiguities over all tracking passes (about 16 ambiguities per day per GPS satellite). Hence double-differences between a LEO satellite and ground stations are connected, reducing the number of zero-difference or double-difference ambiguities with the ground IGS network by nearly 95%.

The same Track-to-Track (T2T) ambiguity resolution approach based on carrier-phase measurements could be applied to double-differences. Biases in the double-differences that are common and repeat from one GPS tracking pass to another tracking pass (e.g., multipath effects, orbit errors, etc.) will be removed when forming differences of double-difference ambiguities between consecutive tracking passes. This is particularly true for the narrow-lane ambiguities where the reduction of common systematic effects between tracking passes will significantly improve ambiguity resolution. In this way reducing the effects like near-field multipath and orbit errors, that repeat in a similar way from the track to the track.

20.1 Direct Resolution of T2T Wide-Lane and Narrow-Lane Ambiguities at the Zero-Difference Level

Wide-lane ambiguities can easily be fixed at the double-difference level using the Melbourne-Wübbena linear combination with a very high success rate close to 100%. Thanks to improved receiver-tracking and multipath mitigation techniques and better antenna design, the relatively low noise of the pseudo-range measurements guarantees very robust resolution of wide-lane ambiguities. In the light of the forthcoming Galileo navigation system offering a wide range of different pseudo-range observables with relatively low noise, the
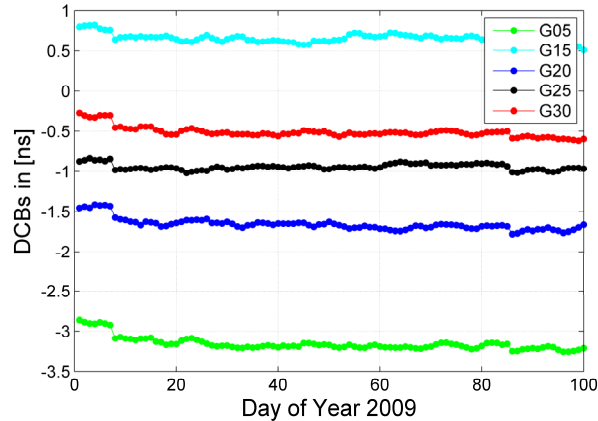
Figure 20.1 Stability of the differential code biases (P1-P2) provided by IGS. Note jumps of up to one narrow-lane ambiguity and differences larger than one wide-lane ambiguity.

success rate in fixing wide-lane ambiguities will follow this trend. Compared to double-differences, wide-lane ambiguities at the zero-difference level are affected by additional satellite and receiver code biases that need to be correctly modelled. One of the main problems stems from the convention used to define satellite and receiver differential code biases (DCBs). By convention, satellite and receiver DCBs are defined as a zero mean over all GPS satellites and over all ground receivers, respectively. This convention is inappropriate for the resolution of wide-lane ambiguities at the zero-difference level, since after applying the DCBs, the Melbourne-Wübbena linear combination will always be affected by an additional wide-lane bias. Figure 20.1 shows satellite DCBs for a period of about three months. One can clearly see jumps in the time series of up to one narrow-lane ambiguity, and differences between different GPS satellites larger than the wavelength of the wide-lane ambiguity. DCBs are typically estimated using global ionosphere maps, and any change in the number of satellites in the GPS constellation or tracking problems of a single GPS satellite, have an impact on the DCBs of all GPS satellites.

When resolution of wide-lane and narrow-lane ambiguities is performed with the DCBs depicted in Figure 20.1, the percentage of the resolved wide-lane ambiguities at the zero-difference level is only about $20 - 30\%$. This considerably limits the ambiguity resolution of subsequent narrow-lane ambiguities to $20 - 30\%$ or less. Narrow-lane ambiguity resolution is directly limited by the success rate in fixing wide-lane ambiguities and
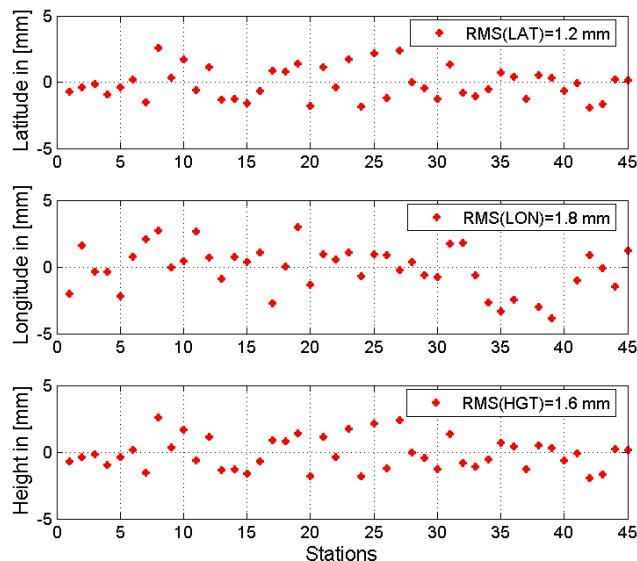
Figure 20.2 Impact of direct zero-difference ambiguity resolution on station coordinates with 45 ground stations (float minus ambiguity fixed solution), day 200/2003.

218
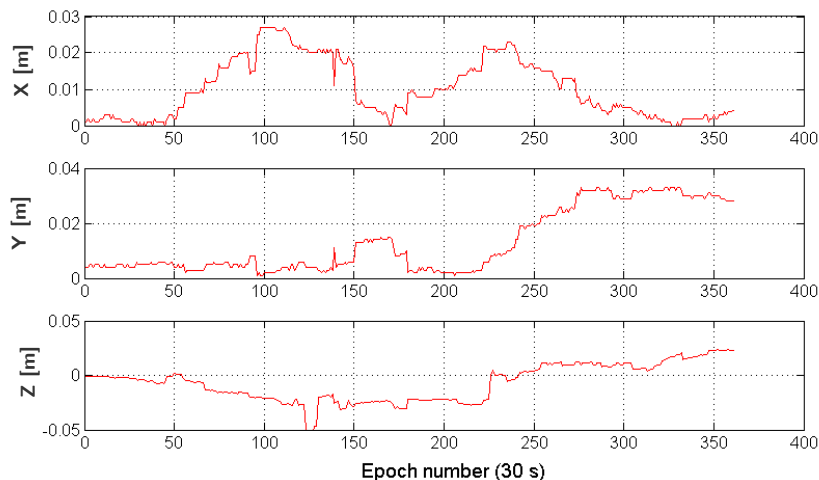
Figure 20.3 Impact of direct zero-difference ambiguity resolution on GRACE-A determined orbit (float minus ambiguity fixed solution).

can only be equal to or lower than the number of successfully resolved wide-lane ambiguities. Wide-lane ambiguities align ambiguities on both GPS frequencies. Figure 20.2 shows the influence of the resolved zero-difference narrow-lane ambiguities on the station coordinates using the phase clock approach, i.e., carrier-phase measurements only. For this test, a global network of about 45 ground stations has been processed for a period of one day, estimating all relevant global parameters, such as station coordinates, GPS satellite orbits, troposphere parameters and satellite and receiver high-rate clock parameters. The effect of the ambiguity resolution on station coordinates is relatively small, and this can easily be explained by the low number of successfully resolved narrow-lane ambiguities (about 22%), limited by the number of resolved wide-lane ambiguities. From Figure 20.2 one can see that the main effect of the direct ambiguity resolution on station coordinates is in the East-West component, whereas the North-South component is less affected. A similar effect, in terms of Cartesian coordinates, can be seen in Figure 20.3, where the impact of direct ambiguity resolution is shown in the case of a LEO orbit. The effect is in the order of 1 cm RMS. Figure 20.3 shows the GRACE-A orbit calculated for a period of 3 hours. Figure 20.4 shows the impact of the direct resolution of narrow-lane ambiguities on the GPS satellite orbits. Although the reduction of 20 − 30% in the overall number of narrow-lane ambiguities is relatively small, the difference in GPS satellite orbits between the zero- and the double-difference solution with fixed ambiguities is in the order of 1 − 3 cm RMS.
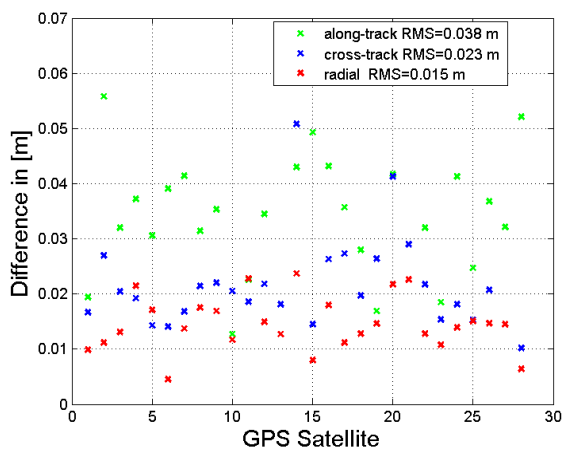
Figure 20.4 GPS satellite orbits based on phase clocks with a limited number of fixed narrow-lane ambiguities (direct approach with about 22% of fixed narrow-lane ambiguities) in comparison to double-difference orbits with fixed ambiguities (close to 100%).
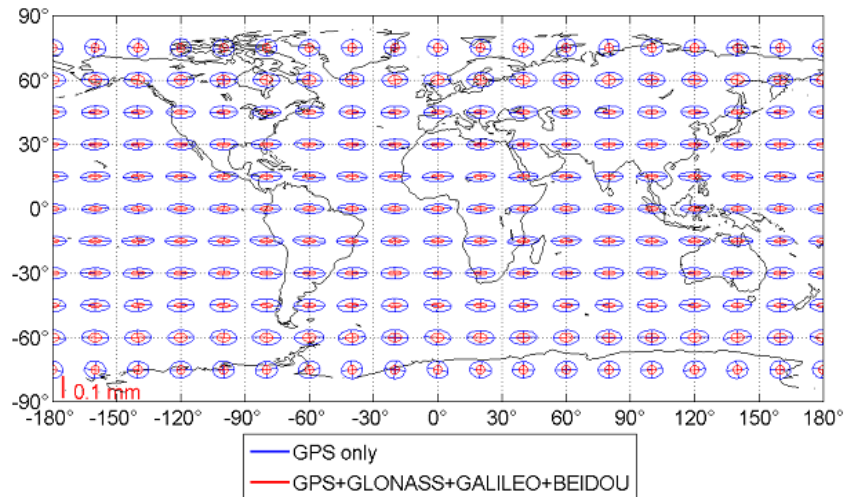
Figure 20.5 East-west effect of float ambiguities on the error ellipses in the PPP solution (24 h) based on simulated carrier-phase measurements for all four GNSS (day 3.3.2007). On can notice reduced noise and more isotropic errors when measurements for all four GNSS are included.

A typical geographically correlated East-West error structure can be seen in the PPP results shown in Figure 20.5 based on simulated carrier-phase measurements with float ambiguities of four GNSS systems (E1/E5 used for Galileo/Compass). One can notice homogeneous and more isotropic positioning and an improvement by a factor of 2.2 in the Helmert error ellipse radius, compared to GPS-only results. Figure 20.6 shows the effects of float ambiguities in the troposphere zenith delays estimated as a piece-wise constant function every hour for an evenly distributed global network of ground stations. Carrier-phase measurements were simulated for the network of ground stations over one day with noise $\sigma = 1$ mm and sampling interval of 30 s .
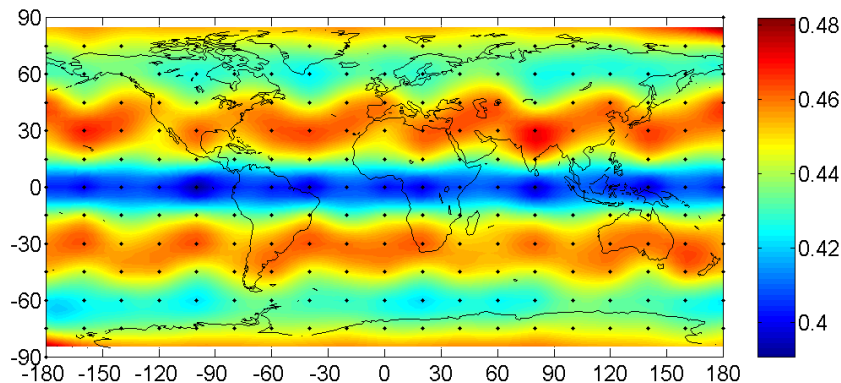
Figure 20.6 Effect of the float ambiguities in the troposphere zenith delays in [mm] based on PPP with simulated GPS constellation (day 3.3.2007). Noise of the estimated TRP parameters is reduced by a factor of 2.5 compared to the solution with four GNSS (E1/E5 used for Galileo/Compass). The 6 simulated orbital GPS planes can easily be recognized as geographically correlated errors. Black dots shows an evenly distributed global network where carrier-phase measurements were simulated over a period of one day.
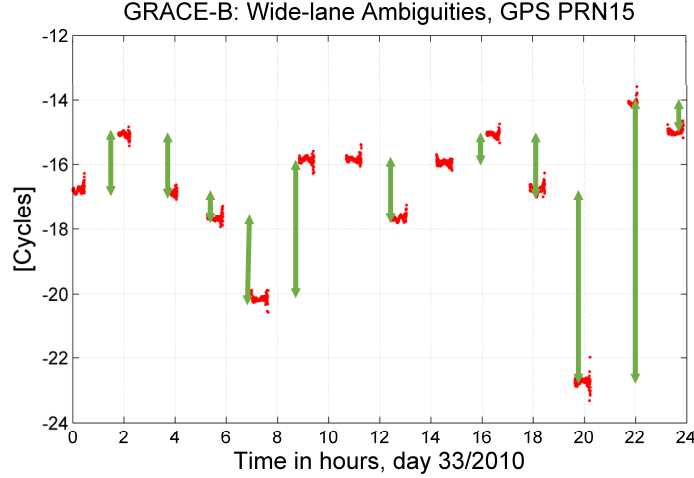
Figure 20.7 Schematic view (in green) of fixing ambiguities between consecutive passes to the same GPS satellite (from GRACE-B GPS receiver) over a period of one day – called here track-to-track ambiguities (T2T). One can see wide-lane ambiguities (in red) every $30\,\mathrm{s}$ as estimated using Melbourne-Wübbena linear combination, affected by the same wide-lane bias for all tracking passes. The integer property of wide-lane ambiguities is preserved by forming differences between consecutive tracking passes (T2T ambiguities).

20.2 Track-to-Track Ambiguity Resolution of Wide-Lane Ambiguities

Over the last couple of years, several ambiguity resolution approaches have been under development at the zero-difference level. The ambiguity resolution approach followed by the IGS Analysis Centre at CNES is based on a very frequent estimation of biases in the Melbourne-Wübbena linear combination, leading to a very high success rate of almost 100%, (Laurichesse and Mercier 2007). However, the very frequent estimation of calibration biases might introduce additional nuisance parameters in the least-squares adjustment, leading to aliasing effects in all other GPS parameters. In the ambiguity resolution approach proposed by (Ge et al. 2007), a network of ground receivers is required to estimate so-called uncalibrated phase delays in the GPS measurements. However, in the case of GPS measurements from LEO satellites, very short tracking passes in LEO GPS measurements (typically $15-20\,\mathrm{min}$) introduce a large number of double-difference ambiguities with the stations of the global ground network. Thus, an ambiguity resolution approach needs to be developed for zero-differences that overcomes both problems, i.e., it does not require a ground network in order to resolve ambiguities from a single GPS receiver, and estimation of biases should be limited and preferably avoided.

In order to avoid estimation of the satellite and receiver code biases (b^{sat}, b_{rec}) in the least-squares approach and possible aliasing effects on other GPS parameters, we show that it is possible to remove those biases between subsequent tracking passes. Let us first write the Melbourne-Wübbena linear combination L_{MW} between two consecutive tracking passes $i, i+1$ to the same GPS satellite

$$
\begin{aligned}
L_{MW}(L_1, L_2, P_1, P_2)_i &\coloneqq \lambda_W N_W^i + b^{sat} + b_{rec} \\
L_{MW}(L_1, L_2, P_1, P_2)_{i+1} &\coloneqq \lambda_W N_W^{i+1} + b^{sat} + b_{rec}
\end{aligned}
\tag{20.1}
$$

with $\lambda_W N_W$ denoting the wide-lane ambiguity. We use the standard definition of the Melbourne-Wübbena linear combination as the difference between the wide-lane linear combination L_W of carrier-phase measurements (L_1, L_2) and the narrow-lane linear combination P_N of code measurements (P_1, P_2). In addition, we

calculate the mean of all measurements j related to one tracking pass with the number of epochs denoted here as n_e

$$L_{MW}(L_1, L_2, P_1, P_2) := \frac{1}{n_e} \sum_{j=1}^{n_e} \left[L_W(L_1, L_2) - P_N(P_1, P_2) \right]_j \tag{20.2}$$

Since the noise of code GPS measurements is typically dependent on the zenith angle, the weighted mean Melbourne-Wübbena linear combination over one tracking pass is finally defined as

$$L_{MW}(L_1, L_2, P_1, P_2) := \frac{\sum_{j=1}^{n_e} p_j \cdot \left[L_W(L_1, L_2) - P_N(P_1, P_2) \right]_j}{\sum_{j=1}^{n_e} p_i}, \qquad p_j = \cos^2(z_j) \tag{20.3}$$

with the elevation-dependent weighting p_j. Wide-lane and narrow-lane observables are then

$$
\begin{aligned}
L_W(L_1, L_2) &= \kappa_{W1} L_1 + \kappa_{W2} L_2 \\
P_N(P_1, P_2) &= \kappa_{N1} P_1 + \kappa_{N2} P_2
\end{aligned}
\qquad
\begin{cases}
\kappa_{W1} = \dfrac{f_1}{f_1 - f_2}, \ \kappa_{W2} = -\dfrac{f_2}{f_1 - f_2} \\[2mm]
\kappa_{N1} = \dfrac{f_1}{f_1 + f_2}, \ \kappa_{N2} = \dfrac{f_2}{f_1 + f_2}
\end{cases}
\tag{20.4}
$$

By differentiating weighted mean Melbourne-Wübbena linear combinations between consecutive tracking passes, we define the track-to-track (T2T) wide-lane ambiguity ΔN_W^i as

$$\lambda_W \Delta N_W^i := \lambda_W N_W^{i+1} - \lambda_W N_W^i \tag{20.5}$$

defined as the bias-free integer wide-lane ambiguity between Melbourne-Wübbena linear combinations of consecutive tracking passes of the same GPS satellite

$$\lambda_W \Delta N_W^i = L_{MW}(L_1, L_2, P_1, P_2)_{i+1} - L_{MW}(L_1, L_2, P_1, P_2)_i \tag{20.6}$$

assuming that the satellite and receiver code biases (b^{sat}, b_{rec}) are constant between the consecutive tracking passes i and $i+1$. Considering that the duration of data gaps between consecutive tracking passes is about $6-12$ hours and less than 30 min in the case of a LEO GPS receiver, we will show with real GPS data that satellite and receiver biases are stable enough over this period of time and are almost completely reduced by forming the difference (20.6). It should be noted that T2T wide-lane ambiguities can be fixed to their integer values without knowing any geometry, even in real-time, by the GPS receiver.

The cumulative integer wide-lane ambiguity N_W^i of the tracking pass i is defined then as the sum of all T2T wide-lane ambiguities ΔN_W^{k-1}, added to the initial wide-lane ambiguity N_W^1

$$\lambda_W N_W^i := \lambda_W N_W^1 + \lambda_W \sum_{k=2}^{i} \Delta N_W^{k-1} \tag{20.7}$$

The initial or reference wide-lane ambiguity N_W^1 is affected by satellite and receiver biases b^{sat} and b_{rec}

$$B_{rec}^{sat} := b^{sat} + b_{rec} = \frac{1}{n_t} \sum_{n_t} \left[L_{MW}(L_1, L_2, P_1, P_2)_i - \lambda_W N_W^i \right] \tag{20.8}$$

where the bias B_{rec}^{sat} is the fractional difference to the nearest wide-lane integer, common to all tracking passes. In the light of this approach, one could consider defining these receiver and satellite code biases as "absolute DCBs", to associate them with the term "relative DCBs" used by IGS. The absolute DCBs in (20.8) should
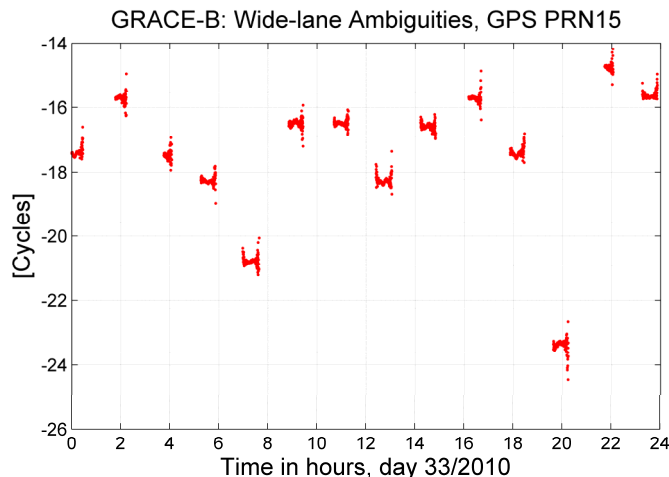
Figure 20.8 Float wide-lane ambiguities estimated using the Melbourne-Wübbena linear combination based on C/A and P_2 code measurements

allow for the "absolute" resolution of wide-lane ambiguities at the zero-difference level. One way to define "absolute DCBs" is to consider them to be zero in the ionosphere-free linear combination of code measurements P_1 and P_2. This is in line with the IGS convention for the estimated GPS satellite clock parameters that by definition are not affected by the ionosphere-free DCBs. In this way, one could talk about the "absolute DCBs", keeping in mind that by forming T2T ambiguities all systematic effects between consecutive tracking passes are removed, and considering that the bias B_{rec}^{sat} requires an "absolute" integer number of wide-lane cycles. In order to demonstrate the T2T ambiguity resolution approach, Figure 20.8 and Figure 20.9 show float wide-lane ambiguities for the GRACE-B satellite estimated using the Melbourne-Wübbena linear combination. Figure 20.8 is based on C/A and P_2 code measurements, whereas Figure 20.9 on P_1 and P_2 code. One can see that the variation between consecutive tracks or tracking passes can be up to several cycles of wide-lane ambiguity in size. To demonstrate the robustness of the approach, typical modeling effects such as satellite and receiver antenna phase-center variations and offsets for different carrier-phase frequencies, as well as the antenna wind-up effect, were not applied. Only raw GPS measurements were used to form the Melbourne-Wübbena linear combination without any screening or data pre-processing. Elevation-dependent weighting was not applied, and for each track the mean Melbourne-Wübbena linear combination was calculated using (20.2). Typically, at the beginning and at the end of every tracking pass one can expect greater noise in the code measurements that could be dealt with by using elevation-dependent weighting, c.f. (20.3). It should be noted that the noise of the Melbourne-Wübbena linear combination is in the order of 70% of the original noise floor for the GPS code measurements.

Comparing Figure 20.9 with Figure 20.8 based on C/A and P_2 code measurements (Melbourne-Wübbena linear combination), one can clearly see a constant bias over all tracking passes. This bias is more visible, when wide-lane ambiguities from Figure 20.9 and Figure 20.8 are rounded to the nearest integer value, as
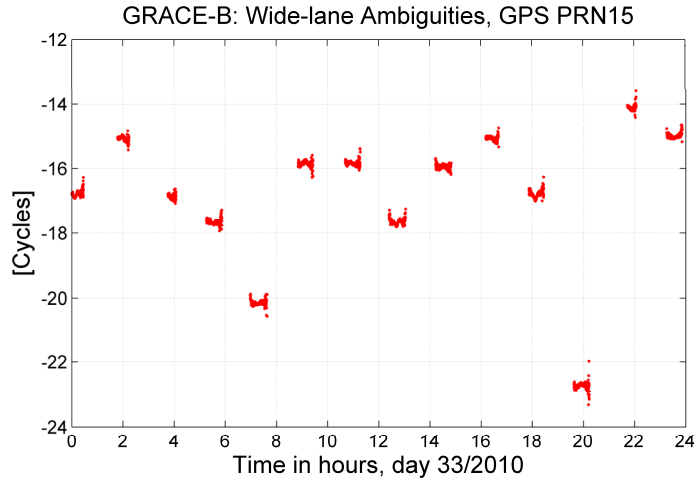
Figure 20.9 Float wide-lane ambiguities estimated using the Melbourne-Wübbena linear combination based on P_1 and P_2 code measurements.

shown in Figure 20.11 and Figure 20.10, respectively. One can see that the fractional parts of wide-lane ambiguities show a clear bias for all tracking passes of about -0.2 cycles for P_1 and P_2 code measurements, and about -0.4 cycles for C/A and P_2 code measurements. These biases are the reason why direct resolution of wide-lane ambiguities at the zero-difference level has a very low success rate, although the noise of the code measurements is sufficiently low to fix wide-lane ambiguities reliably. Ambiguity resolution at the zero-difference level without properly considering these biases could give misleading results. There are two approaches possible: either to estimate wide-lane biases as parameters or to remove them by forming T2T ambiguities. If the wide-lane biases are estimated as parameters, one should count on additional correlations with ambiguity parameters in the least-squares.

The two outliers in Figure 20.11 are due to rounding to the nearest integer, since the bias is very close to 0.5 cycles. Variations in the estimated wide-lame ambiguities between successive tracking passes are within 0.1 cycles, or significantly less in the case of C/A code measurements. This depends on the DCBs applied to the code measurements of the GPS satellite and a GPS receiver. In this particular case, we did not apply any DCB between C/A and P_1 code data. When differences are formed between consecutives passes, such a bias is removed. If the bias in carrier-phase tracking in the GPS receiver is randomly initiated for every
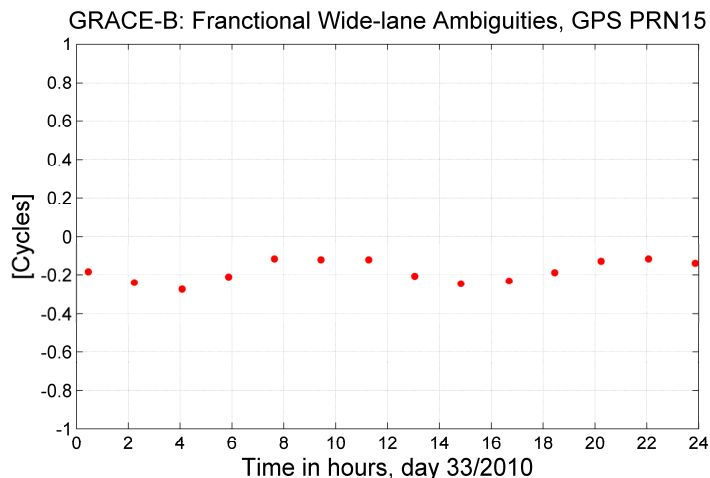
Figure 20.10 Fractional parts of the float wide-lane ambiguities from the nearest integer values, based on P_1 and P_2 code measurements.

224
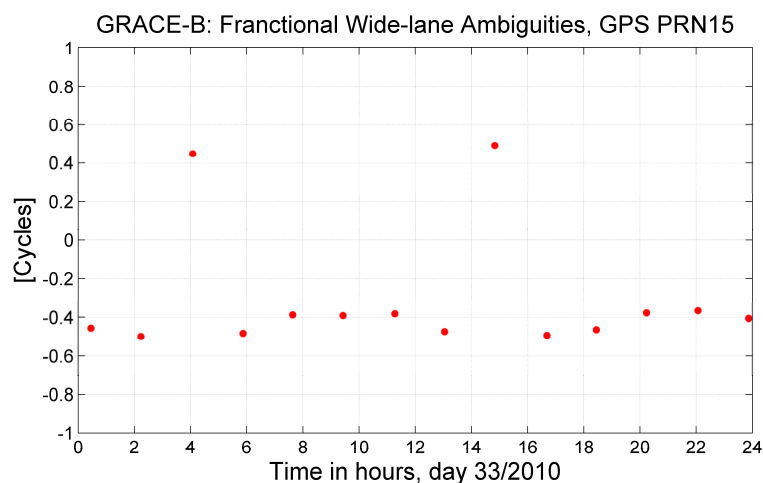
Figure 20.11 Fractional parts of float wide-lane ambiguities from the nearest integer values, based on C/A and P_2 code measurements. The two outliers (≈ 4 h and 14 h) are due to rounding to the nearest integer.

tracking pass, the common bias in T2T ambiguities would experience a random property, but this is not visible. Figure 20.13 shows residuals after fixing T2T ambiguities, or differences between mean float wide-lane ambiguities along consecutive tracking passes. One can see that the common bias is removed between consecutive tracking passes and remaining residuals are below 0.1 cycles. Figure 20.13 clearly shows that T2T wide lane ambiguity resolution can be performed with a very high success rate very close to 100%. Figure 20.12 shows the same T2T ambiguities, but based on C/A and P_2 code measurements.

In a similar way, we show in Figure 20.14 mean wide-lane ambiguities for the ground station ALGO and all GPS satellites tracked, for a period of one day. Again, one can see a clear bias between consecutive float wide-lane ambiguities. Typically, for one day of ground GPS measurements, one can expect $2-3$ tracking passes with $2-3$ wide-lane ambiguities to the same GPS satellite. After forming differences between consecutive passes, common biases are eliminated for all GPS satellites and fractional parts of the T2T ambiguities are below 0.1 cycles, see Figure 20.15. This shows again that track-to-track differences remove common biases and the remaining float ambiguity can be fixed with a success rate close to 100%. Thanks to the very long observation time, noise in the code measurements is averaged over a period of $4-6$ hours
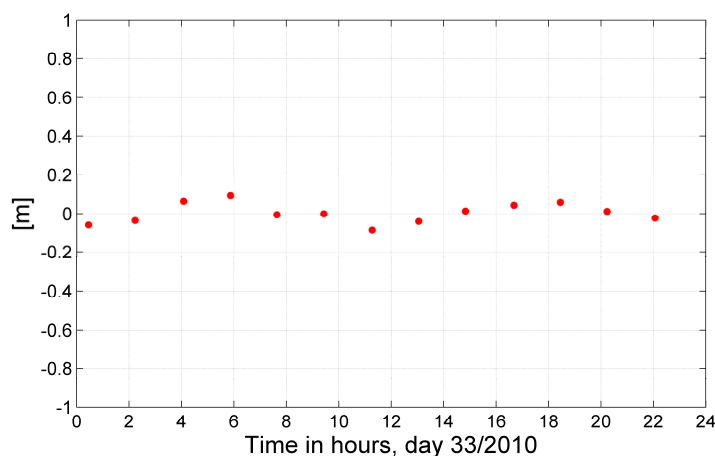
Figure 20.12 Residuals in wide-lane ambiguities after fixing track-to-track ambiguities to their integer values ($\mathrm{RMS} = 4.8\,\mathrm{cm}$, antenna wind up, PCVs and other similar effects not applied). This solution is based on C/A and P_2 code measurements (GPS PRN 15).
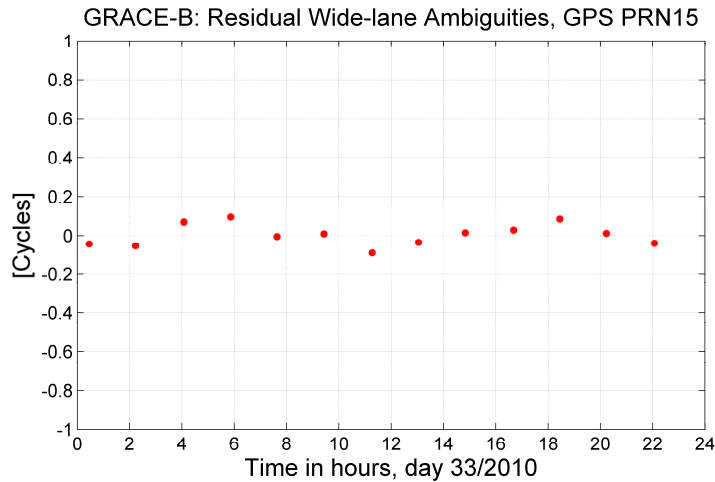
Figure 20.13 Residuals in wide-lane ambiguities after fixing T2T ambiguities ($\text{RMS} = 5.2$ cm , antenna wind up, PCVs and other similar effects not applied). This solution is based on P_1/P_2 code measurements.

leading to very small errors in the fractional parts of the wide-lane ambiguities. However, if the tracking pass is very short (LEO or a ground station), the ambiguity resolution could be critical, and therefore, GPS data should be properly combined between consecutive days (day boundaries). Typically, the beginning and the end of a tracking pass show higher noise and multipath effects compared to the middle. Elevation-dependent weighting could give misleading results if only a short fraction of a tracking pass (close to a day boundary) is processed. However, Figure 20.15 shows that even in this case, when one could expect higher noise for the tracking passes close to day boundaries (see in Figure 20.14 cases with GPS satellites with three passes per day), estimated fractional parts are below 0.2 cycles. In this calculation we used standard processing models for carrier-phase and code measurements, such as satellite and station PCO/PCVs, antenna wind-up effect and elevation-dependent weighting. Other geometrical effects such as light-travel time corrections, relativistic corrections, tidal effects in the station coordinates, etc. do not play any role in the Melbourne-Wübbena linear combination.
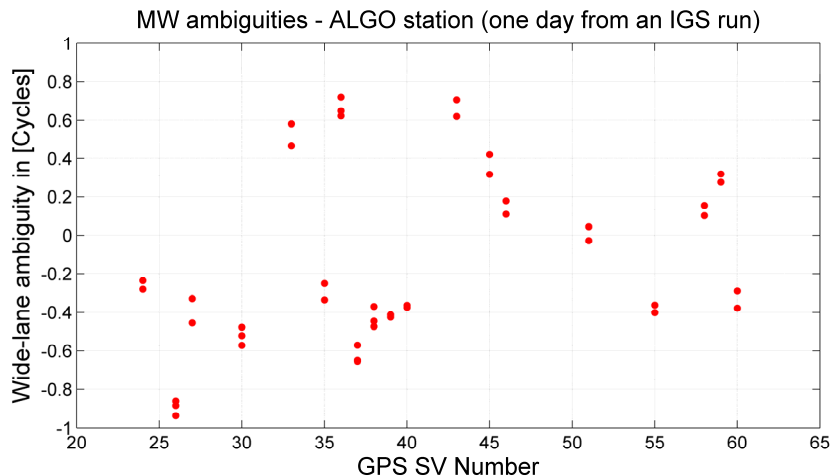
Figure 20.14 Float wide-lane ambiguities for the ground station ALGO over one day. Typically, $2 - 3$ wide-lane ambiguities have to be set up per satellite for a one-day arc. Note the different common biases between GPS satellites and the very similar magnitude of the float wide-lane ambiguities for the same GPS satellite.
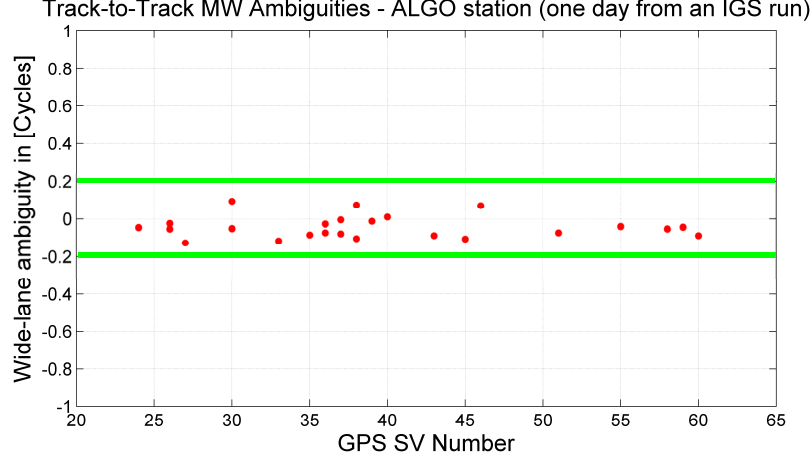
Figure 20.15 The fractional parts of the track-to-track wide-lane ambiguities over one day are well below 0.2 cycles and can be reliably fixed to their integer values (green lines). Please note that the wide-lane ambiguities were processed for a period of one day, thus, a short fraction of a tracking pass (close to a day boundary) could in principle produce a **f**ractional part with higher error.

20.3 Track-to-Track Ambiguity Resolution of Narrow-Lane Ambiguities

We first write the ionosphere-free linear combination L_3 between consecutive tracking passes $i, i+1$

$$L_3(L_1, L_2)_i := \rho^i + \lambda_N N_1^i + \frac{1}{2}(\lambda_W - \lambda_N)N_W^i - \delta_{sat}^i + \delta_{rec}^i - b^{sat} + b_{rec}$$
$$L_3(L_1, L_2)_{i+1} := \rho^{i+1} + \lambda_N N_1^{i+1} + \frac{1}{2}(\lambda_W - \lambda_N)N_W^{i+1} - \delta_{sat}^{i+1} + \delta_{rec}^{i+1} - b^{sat} + b_{rec}$$

(20.9)

Compared to the Melbourne-Wübbena linear combination, additional terms are involved, namely geometry ρ and satellite and receiver clock parameters (δ_{sat}, δ_{rec}). Thus the easiest way to form track-to-track differences is to extract narrow-lane ambiguities from the parameter estimation procedure based on float ambiguities and then to form T2T ambiguities. If GPS satellite orbits and high-rate GPS satellite clock parameters are available, one would need to estimate station coordinates or the LEO orbit together with GPS receiver clock parameters, and subsequently form T2T ambiguities. In the float solution, it is a prerequisite that GPS satellite clock parameters are continuous, i.e., clock parameters of successive tracking passes are connected. Typically, GPS satellite clock parameters estimated using the phase clock approach, or a combination of code and phase measurements with down-weighted code measurements, show excellent stability between GPS tracking passes. They are, however, biased in the absolute sense by b^{sat}. This is especially the case with phase clocks estimated using only carrier-phase measurements. A global ground network of about 45 stations is sufficient to estimate GPS satellite clocks that do not experience discontinuities, and can thus be used to connect consecutive tracking passes. However, they are biased in an absolute sense. This common bias can be removed by forming differences between consecutive tracking passes, defining the track-to-track narrow-lane ambiguity ΔN_N^i

$$\lambda_N \Delta N_1^i := \lambda_N N_1^{i+1} - \lambda_N N_1^i$$

(20.10)

as the bias-free, integer narrow-lane ambiguity between ionosphere-free linear combinations of consecutive tracking passes to the same GPS satellite

$$\lambda_N \Delta N_1^i := L_3(L_1, L_2)_{i+1} - L_3(L_1, L_2)_i - (\rho^{i+1} - \rho^i) - (\delta_{sat}^{i+1} - \delta_{sat}^i) - (\delta_{rec}^{i+1} - \delta_{rec}^i).$$

(20.11)

From (20.11), we see that satellite and receiver code biases are completely removed. Any bias in the GPS satellite clock parameters is removed by forming differences between consecutive float ambiguities. There are only a small remaining effects on the carrier-phase that are not constant between consecutive passes, e.g., due to the geometry terms (troposphere, GPS orbits), multipath, receiver front-end, etc.

The cumulative narrow-lane ambiguity N_1^i is then defined as the sum of all consecutive narrow-lane ambiguities ΔN_1^{k-1} added to the initial narrow-lane ambiguity denoted as N_N^1

$$\lambda_N N_1^i := \lambda_N N_1^1 + \lambda_N \sum_{k=2}^{i} \Delta N_1^{k-1} \tag{20.12}$$

Compared to the ambiguity resolution of T2T wide-lane ambiguities, the additional geometry and satellite/receiver clock parameters need to be modeled to an accuracy of $1-2$ cm RMS. This is required in order to obtain a noise level of the estimated fractional T2T ambiguities in the order of $10-20\%$ of the narrow-lane wavelength of ≈ 10.7 cm .

Figure 20.16 shows fractional (residual) T2T narrow-lane ambiguities of the IGS station ALGO for all GPS satellites over a period of one day. One can see that the noise level of the estimated T2T ambiguities is < 0.3 cycle , and thus most T2T ambiguities can be fixed to their integer numbers. In this solution, GPS satellite orbits and high-rate clock parameters were estimated, together with the ground station coordinates, Earth's rotation and troposphere parameters, using all state-of-the-art modeling and processing standards for GPS measurements.

In order to align the integer T2T narrow-lane ambiguities, an additional satellite/receiver bias needs to be estimated together with the initial narrow-lane ambiguity. This could be solved by estimating the initial narrow-lane ambiguity as a float integer, or in a similar way to wide-lane ambiguities, by estimating a common fractional part in the narrow-lane ambiguities over all tracking passes. For LEO measurements, T2T ambiguity resolution can be performed first (for the very short-duration LEO ambiguities), and in the second step, the common bias can be removed using ground-to-LEO baselines. In this way, T2T ambiguities are estimated together with the ground-to-ground (long duration) and ground-to-space phase ambiguities. T2T ambiguity resolution could be considered as the optimal method for combining LEO and ground GPS measurements. In this way, the LEO can serve as a "flying station" connecting carrier-phase ambiguities for all ground stations in only 90 min (LEO orbit period).

By forming double-differences, biases in the initial narrow-lane ambiguities are removed. If those biases are not stable over time, double-difference ambiguities cannot be fixed to their integer values. For all GPS receivers in the IGS ground network it is well-known that double-differences remove all common biases between a GPS receiver and a GPS satellite.

It is important to mention that, if the bias in the tracking of carrier-phase in the GPS receiver were randomly initiated for every tracking pass, the common bias in the narrow-lane T2T ambiguities would experience a random property from one tracking pass to the next, but this is not visible in the data.
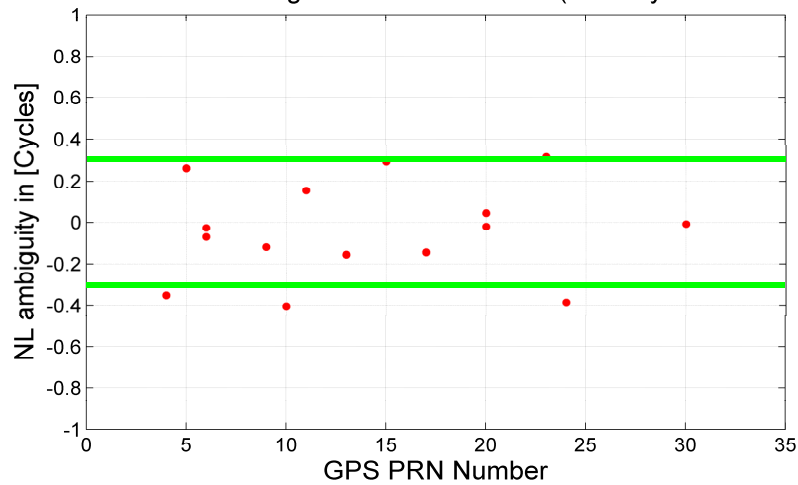
Figure 20.16 Fractional parts in track-to-track narrow-lane ambiguities for a period of one day. All residuals are within an interval of ± 3 cm.

20.4 L1-L1A Track-to-Track Ambiguities

For GPS measurements from the GRACE mission, two types of carrier-phase measurements are available on the first GPS frequency f_1: carrier-phase measurements from C/A, and measurements from the P_1 code tracking. Figure 20.17 (left) shows the differences between L_1 and L_{1A} carrier-phase measurements, abbreviated to "L1-L1A" float ambiguities. One can see that there is a common bias in all ambiguities of about one wide-lane ambiguity and the differences between consecutive tracks are in the order of one wavelength λ_1. Once track-to-track measurements are formed, the common bias is removed and the integer nature of the track-to-track ambiguities can be clearly seen in Figure 20.17 (right). After rounding the track-to-track L1-L1A ambiguities to their integer values, the remaining phase residuals are in the order of 0.29 mm RMS, as shown in Figure 20.18. This value corresponds to the typical noise floor of carrier-phase measurements. It is important to note that Figure 20.18 does not show any systematic effects in the carrier-phase measurements between consecutive tracking passes. Figure 20.17 (left) shows that larger differences between L_1 and L_{1A} phase measurements can be expected at the beginning and end of the tracking pass and making use of the elevation-dependent weighting, the residuals in Figure 20.18 might be even smaller.
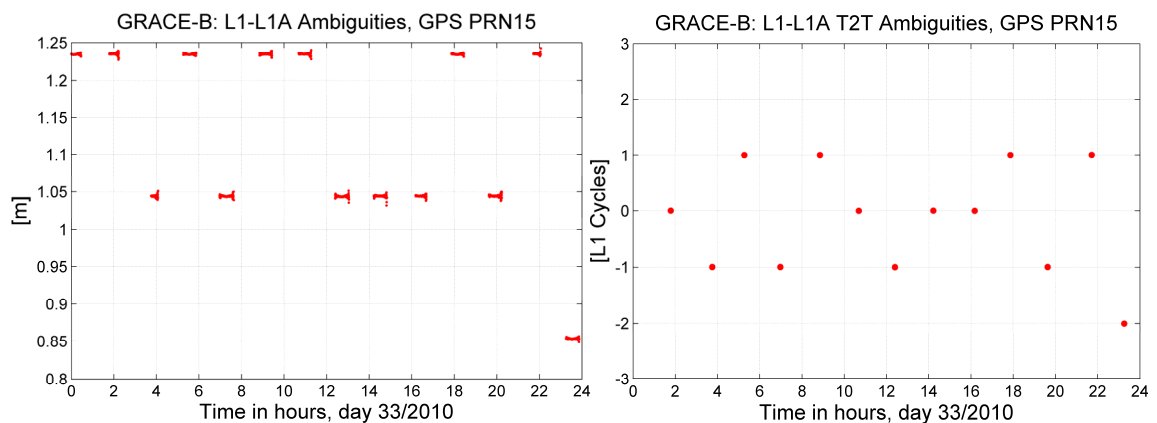
Figure 20.17 Differences in L1-L1A phase measurements (left) and track-to-track ambiguities (right).
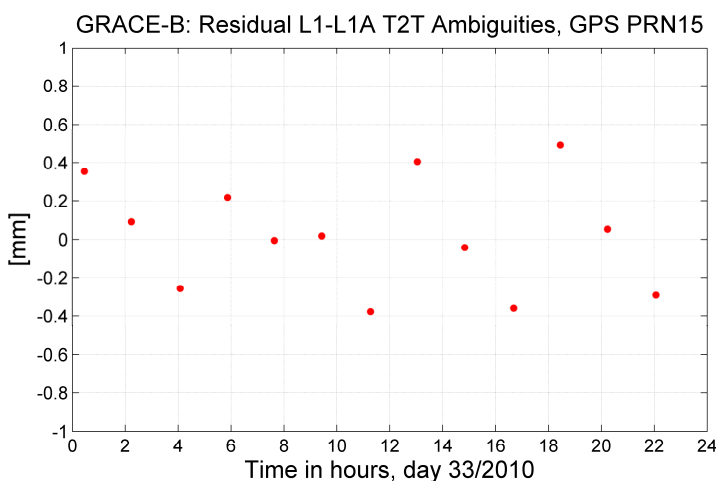
Figure 20.18 Residuals in the L1-L1A T2T ambiguities after rounding to their integer values.

20.5 Using Stable Satellite Clocks for Track-to-Track Ambiguity Resolution

Let us now see what satellite clock stability would be needed to predict clock parameters reliably over the data gaps caused by tracking geometry, if one had a linear combination with a sufficiently high wavelength. One could then estimate integer track-to-track ambiguity as simple carrier-phase cycle-slip.

As long as phase clocks are estimated without code measurements and GPS phase measurements remain connected for all satellites and ground stations, we may select one reference H-maser and relate all epoch-wise ground and station clock parameters in the IGS network to this reference clock. In that sense, phase clocks estimated with float ambiguities are a closed system and can be used for the ambiguity resolution of track-to-track narrow-lane ambiguities. The absolute bias in the phase clocks of the same GPS satellite is removed by forming track-to-track differences. Or one could use IGS Final clock solutions and relate carrier-phase between consecutive passes. On the other hand, one could use stable clocks in the IGS network and attempt to treat consecutive narrow-lane ambiguities as cycle-slips. Considering that there are about 70 H-masers and other atomic clocks in the IGS network, ambiguity resolution of track-to-track narrow-lane ambiguities could be considered as cycle-slip fixing.

Let us now see what level of clock stability would be required to reliably predict or estimate the receiver and satellite clock terms in (20.11) between two successive tracking passes, i.e., over a period of 6–12 hours. For an H-maser, given the Allan deviation $\mathrm{ADEV}(\tau)$ over an integration time $\tau = 1\,\mathrm{s}$, we can calculate the Allan deviation for an integration time τ using

$$\mathrm{ADEV}(\tau) = \mathrm{ADEV}(1\,s) / \sqrt{\tau} \tag{20.13}$$

The time deviation of the receiver or satellite clock over a time interval τ, for a given (modified) Allan deviation (MDEV) is then

$$\sigma\left(\delta_{clk}(\tau)\right) = \frac{\tau}{\sqrt{3}} \cdot \mathrm{MDEV}(\tau) \tag{20.14}$$

or in simple terms, TDEV is equal to MDEV whose slope is normalized by √3, (Riley 2014). The time Allan variance is equal to the standard variance of the time deviations for white phase modulated noise. It is particularly useful for measuring the stability of a time distribution network (Riley 2014).

In Section 18 on the performance of the Galileo passive H-Maser (PHM) based on ground data, we derived white phase noise in the order of 9.8×10^{-13}, white frequency noise of 5.9×10^{-13}, flicker frequency noise of 7.9×10^{-16} and a very small frequency drift of 1.2×10^{-20} /s. Figure 18.5 shows the Galileo PHM clock model based on ground data with the linear model removed (time offset and time drift over a period of time τ) in comparison with simple TDEV without the linear model removed. One can see that removal of time drift and time offset significantly improves performance of the Galileo PHM, especially for flicker frequency and white frequency noise. Flicker frequency is the dominant error source only after Galileo orbit period (14 h). Note that frequency drift is very small. This confirms that the Galileo PHM clock is stable enough to maintain carrier-phase over data gaps to the same ground station and can be used for T2T ambiguity resolution.

$\sigma\left(\delta_{clk}(\tau)\right)$	$\sigma\left(\tau = 0.5\,\text{h}\right)$	$\sigma\left(\tau = 1\,\text{h}\right)$	$\sigma\left(\tau = 7\,\text{h}\right)$	$\sigma\left(\tau = 14\,\text{h}\right)$
Galileo PHM all frequency noises (linear model removed)	2 mm	2.7 mm	6.8 mm (6 h)	11.2 mm
1×10^{-12} (Galileo PHM) (linear model included (20.14))	7 mm	10 mm	27 mm	39 mm
1×10^{-13} (H-maser)	0.7 mm	1.0 mm	2.7 mm	3.9 mm
1×10^{-16} (optical clock)	0.0007 mm	0.001 mm	0.0027 mm	0.0039 mm

Table 20.1 Allan deviation of an H-maser for Galileo and a highly stable H-maser in the IGS network in comparison with an optical clock in terms of time standard deviation over an interval of 0.5 to 14 hours.

Table 20.1 shows that a highly stable H-Maser in the IGS network can predict and keep phase between two consecutive tracking passes of the same GPS satellite up to 6–12 hours. It should be noted that Allan variance actually gives the accuracy of the linear time drift, i.e., the accuracy of the slope defined by two parameters (time offset and drift), and therefore, the estimated GPS receiver clock parameters are considerably more stable than depicted in Table 20.1. In a similar way, any gap in the Galileo satellite clock could be preserved over a period of about 0.5–1 hour. However, if the Galileo clock is modeled using linear bias and drift over a period of one day, the estimated results are sufficiently stable to resolve the T2T ambiguities. The last line in Table 20.1 refers to an optical clock and represents the state-of-the-art in clock performance.

This analysis shows that T2T narrow-lane ambiguities can be considered as cycle-slips, and stable clocks in the IGS network could be used to correct them between subsequent tracking passes. This statement is true, as long as 40–50 well performing H-masers in the IGS network can be modeled with just two linear clock parameters per day. An additional geometry term, including station coordinates and troposphere parameters can be estimated with sufficient accuracy and its impact could be considered smaller than the clock contribution.

On the other hand, the best IGS Final clocks for GPS and GLONASS satellites (estimated epoch-wise satellite clock parameters) show standard deviation of about 15 ps (a typical comparison between the best GPS/GLONASS clock solutions and the IGS Final Product in 2011). This corresponds to a standard deviation of about 4.5 mm or roughly 3.4 mm above the noise floor of the ionosphere-free linear combination. The 3 mm noise floor of the ionosphere-free linear combination corresponds to a phase noise of about 1 mm on L_1 and on L_2. It is expected that the noise floor of 15 mm will be improved in the near future to $5 - 10$ ps ($2 - 3$ mm), especially when additional Galileo satellites become available. Galileo satellites can reduce noise in the estimated epoch-wise clock parameters of GPS and GLONASS satellites by using common ground station clock parameters. This is what one can see when processing GLONASS data together with GPS − a clear improvement in the estimated GPS satellite clock parameters. This analysis shows that estimated satellite clock parameters are of sufficient accuracy to be used for T2T ambiguity resolution, i.e., to bridge the gap and fix the cycle-slip ambiguity between two consecutive tracking passes. Since the GPS satellite orbit can be predicted very accurately, assuming Galileo satellite clock stability one could estimate T2T ambiguities as

cycle-slips. In the following sections, we will develop a Galileo/GPS three-carrier linear combination with a wavelength sufficient to fix ambiguities to their integer values by treating them as simple cycle-slips.

20.6 Towards the LEO Network in Space and Combined LEO/GNSS Frame Parameters Based on the Cumulative Track-to-Track Ambiguities

For one LEO satellite and the constellation of 30 GPS satellites, one can expect about 450 zero-difference ambiguities for a period of one day. Connecting the carrier-phase between consecutive tracking passes, the T2T ambiguity resolution leads to a reduction in the overall number of LEO ambiguities of about 95%. After fixing ambiguities between consecutive tracking passes (16 LEO revolutions), we end up with only one cumulative or core float zero-difference ambiguity per GPS satellite (and LEO) for the period of one day or longer (i.e., 30 ambiguities in total for all GPS constellation a total of about 30 GPS satellites).

If we now look at a constellation of several LEO satellites, or just two LEO satellites flying in formation (as with the GRACE-A/B mission), after T2T ambiguity resolution we need to fix only one arc-specific float ambiguity, i.e., one core float ambiguity per GPS satellite and one LEO satellite for the entire arc. Once carrier-phase between the LEO satellite and one GPS satellite is connected for the entire arc (about 16 orbits per day), one can form baselines between the LEO satellites and/or between the LEO satellite and the ground stations. Typically, for the ground-to-LEO GPS baselines with about 100 ground stations, we have about 5000 double-difference ambiguities for the period of one day. Following the proposed T2T approach, this total number of ambiguities can be reduced to one double-difference ambiguity per LEO satellite and ground station. It should be noted that the duration of LEO passes is very short (typically $15-20$ min) and when GPS baselines are formed between LEO satellites or a ground station and a LEO satellite, the observation time of one double-difference ambiguity is significantly shorter than that for the original zero-difference ambiguity. However, after fixing T2T ambiguities, the observation time of one double-difference ambiguity is increased to the entire duration of the arc (e.g., 24 h or even one week).

It should be noted that LEO orbits can be estimated with an accuracy of $1-2$ cm without any ambiguity resolution. Thus, the float orbit solution could be used and T2T ambiguities could be fixed to their integer values. After T2T ambiguity resolution, an efficient combination of LEO and ground GPS measurements is feasible, leading to a combined LEO-to-LEO or ground-to-LEO reference frame solution of utmost accuracy.

When the global IGS network is processed at the zero-difference level by estimating all GNSS terrestrial frame parameters and zero-difference ambiguities are fixed using, e.g., the "GFZ approach", the additional constraints at the normal equation level for T2T ambiguities improves the overall ambiguity resolution by a total of about 30-40%. However, the best results are obtained if core T2T ambiguities are fixed first and the carrier-phase is connected for all tracking passes. This reduces the number of all narrow-lane ambiguities by about 95%.

21. Integer Ambiguity Algebra

I n this section we develop integer ambiguity algebra, a mathematical approach to handle integer ambiguities between different GNSS frequencies and introduce what we call the ambiguity-free linear combination. We first show the vector form of the wide-lane ambiguity for multi-frequency GNSS and then develop integer ambiguity algebra and show in detail the integer property of the ionosphere-free ambiguity for GPS and Galileo. We show that any GNSS ionosphere-free linear combination can be represented by an integer ambiguity without resolving wide-lane ambiguity. This opens up the possibility of forming an integer ambiguity of arbitrary wavelength, when combined with narrow-lane ambiguity. We introduce an elegant way to resolve wide-/narrow-lane ambiguities using the ambiguity-free linear combination that is consistent with what we term absolute code biases. The advantage of this approach is the consistent resolution of wide-lane ambiguities and calibration of wide-lane biases in an absolute sense, since the same ambiguity-free linear combination can be used to estimate absolute code biases, (see section on absolute code biases). Code biases can be defined in an absolute sense if one uses the IGS convention for estimated clock parameters that the net effect of code biases is zero for the ionosphere-free linear combination of P-code measurements, or so-called P_3-clocks. They are still limited by the full number of wide-lane ambiguities that can be defined separately for two- and three-carriers with a wavelength of 0.67 m and 3.41 m respectively. Since absolute code biases are determined against the ionosphere-free P-code, we obtain a consistent framework for ambiguity resolution for all four GNSS. Then, by using integer ambiguity algebra, we develop three-carrier wide-/narrow-lane linear combinations for GPS/Galileo and show how to use this approach for ambiguity resolution and retrieval of ionospheric effects. We show that a three-carrier-type Melbourne-Wübbena linear combination can be derived by means of ambiguity algebra.

21.1 Code-Ambiguity Linear Combination

Due to receiver tracking difficulties, e.g., due to missing broadcast orbits, code measurements can also be biased by integer ambiguities, or what we call "code ambiguities". A typical example occurred with the GIOVE-A/B and early Galileo data, when ground receivers were tracking the Galileo signal without knowing satellite's position (the broadcast navigation message not being transmitted). In this case, a GNSS receiver could not resolve the full number of code chip lengths from the receiver to the satellite, as noted for the first time in (Svehla et al. 2008). It was reported by the Galileo Project Office at ESA/ESTEC that (Svehla et al. 2008) was the first solution of this problem [*F. Gonzalez, priv. com.*]. This was already available to the Galileo Project Office (in 2007). Table 21.1 shows code ambiguities applied to different code observables and the clear clock bias once the code ambiguities are applied. Another aspect of tracking that could cause code integer biases is incorrect locking in the tracking loops. Figure 21.1 shows the so-called 10-m jumps in the Galileo PRS (Public Regulated Service) modulation code (C1A-C1B) residuals from GIOVE-A station GIEN, due to incorrect tracking lock to the nearest asymmetric side-peak (on the right), spaced at $1/12$ chip length ≈ 9.77 m .

Code	Raw Pseudorange [km]	Chip Length [msec]	Code Ambiguity [sec]	Corrected Pseudorange [km]
C1A	399801	100	-12	40050
C1B	-49888	4	75	40050
C1C	-49888	100	3	40050
C5I	1149282	20	-185	40050
C5Q	1149282	100	-37	40050
C7Q	-49888	100	3	40050
C8Q	1149282	100	-37	40050

Table 21.1 Ambiguities in pseudorange measurements (broadcast navigation message not being transmitted) in the early GIOVE-A data (GNOR, day 70/2007). The last column shows corrected pseudoranges after applying code ambiguities with chip length (third column) that we found to match the original data. One can see a common clock bias in the corrected pseudoranges.

The PRS code chip length corresponds to about ≈ 117.2 m and if we consider a correlation profile with 12 chips (as shown in Figure 21.1) the offset to the nearest side-peak is ≈ 9.77 m , (Svehla et al. 2008). Similar code ambiguities caused by the tracking loops in the receivers could be seen in the early GPS measurements from the CHAMP mission in 2001, typically at low elevations.

The general form of the code-ambiguity linear combination for measurement on the frequency pair $(1, q)$ to resolve code and phase ambiguities to their integer numbers can be defined as a difference of wide-lane phase $L_{W(1,q)}$ and narrow-lane code $P_{N(1,q)}$

$$
\begin{aligned}
L_{CAMB}(L_1, L_q, P_1, P_q) &:= L_{W(1,q)} - P_{N(1,q)} \\
&= \lambda_{W(1,q)} N_{W(1,q)} - \left(\kappa_{N(1)}^{1-q} \Lambda_{P1} + \kappa_{N(2)}^{1-q} \Lambda_{Pq} \right) + b^{sat} + b_{rec}
\end{aligned}
\tag{21.1}
$$

where Λ_{P1} and Λ_{Pq} denote code ambiguities on both tracking frequencies and $N_{W(1,q)}$ the wide-lane carrier-phase ambiguity with wavelength $\lambda_{W(1,q)}$. The satellite and receiver code ambiguities are denoted as b^{sat} and b_{rec}. Furthermore, the code ambiguity Λ_{Pq}, can be defined as the sum of an integer number (Π_{CL}) of code chip lengths (Λ_{CL}) between receiver and satellite, and what we call the integer side-peak offset ambiguity (Π_{SP}) with length Λ_{SP}

$$
\Lambda_{Pq} := \Lambda_{CL} \Pi_{CL} + \Lambda_{SP} \Pi_{SP}, \qquad \Pi_{SP} = \pm 1
\tag{21.2}
$$

Typically, the integer side-peak offset ambiguity $\Pi_{SP} = \pm 1$. It should be noted that by forming T2T ambiguities, the code ambiguity Λ_{CL} will be completely removed for all tracking passes, enabling resolution of T2T wide-lane ambiguities. However, this is not the case with the integer side-peak offset ambiguity $\Lambda_{CP} \Pi_{CL}$ that can change from track to track, as shown in Figure 21.1. That additional effect is due to incorrect lock to the nearest asymmetric side-peak of the correlation profile in the receiver tracking loop. As long as broadcast ephemerides are transmitted from a GNSS satellite and the receiver knows the approximate position of the GNSS satellite, code ambiguities can be directly fixed by the receiver. However, in some extreme tracking situations, especially at low elevations, with a poor S/N ratio, or in a strong multipath environment, the receiver can in addition incorrectly lock the signal to the nearest side-peak in the correlation profile. This incorrect lock could be detected in the early measurements from the GIOVE-A satellite and the very first GPS measurements from the CHAMP satellite in LEO orbit. Code ambiguities were present in GNSS measurements from GIOVE-B as well as early Galileo satellites, or in all cases where broadcast navigation messages were not being transmitted by the GNSS satellite.
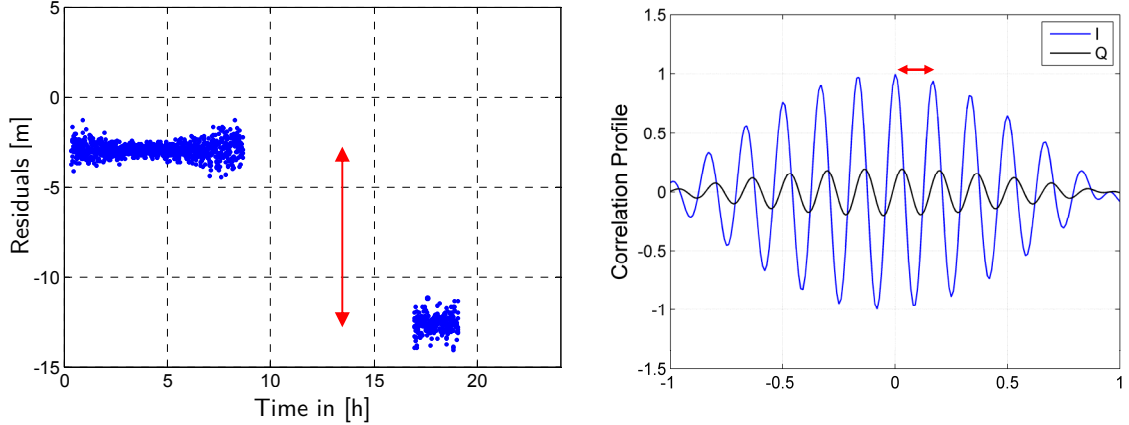
Figure 21.1 The so-called 10-m jumps (left) in the GIOVE-A PRS code residuals (C1A-C1B) due to incorrect lock to the nearest asymmetric side-peak in the receiver tracking loop (right) spaced at $1/12$ chip length of ≈ 9.77 m (GIEN).

21.2 Ambiguity Resolution based on a Symmetric Geometry-Free Form of the Ionosphere-Free Linear Combination

In the case of three- or multi-frequency GNSS measurements, receiver and satellite code biases will always be present in the estimated wide-lane and narrow-lane ambiguities at the zero-difference level. T2T ambiguity resolution can be used to remove these biases in the zero-difference GNSS measurements (by forming difference between satellite tracks) and reliably fix ambiguities to their integer values. Absolute code biases can be estimated using ambiguity-free linear combination (see section on absolute code biases). This enables an absolute datum for the remaining float ambiguities to be obtained after the T2T ambiguity resolution.

For any combination of two-frequency GNSS measurements, the geometry-free form of the ionosphere-free linear combination L_3 (comprising only the ambiguity part) can be written as

$$2\lambda_N N_1 + (\lambda_W - \lambda_N)N_W := (L_N - P_W) + (L_W - P_N) := 2(L_3 - P_3) \tag{21.3}$$

with P_N and P_W denoting the narrow-lane and wide-lane linear combination of code measurements with wavelength λ_W and λ_N of the wide-lane N_W and narrow-lane ambiguities. The main drawback of (21.3) is the very high noise of the wide-lane linear combination of code measurements. This noise level can be reduced by the symmetric form of the ionosphere-free linear combination (21.3), with the negative wide-lane ambiguity

$$\begin{aligned}2\lambda_N N_1 - (\lambda_W - \lambda_N)N_W &:= (L_N - P_W) - (L_W - P_N) \\ &:= -(\kappa_{W1} + \kappa_{N1})\big[(L_1 - L_2) + (P_1 - P_2)\big]\end{aligned} \tag{21.4}$$

where κ_{W1} and κ_{N1} are the wide-lane and narrow-lane multiplication factors for the first GPS frequency. In both cases, the noise of the linear combination is too high to reliably fix the narrow-lane ambiguity $\lambda_N N_1$ and thus an additional transformation is needed to increase the wavelength of the ambiguity with respect to the noise of the code measurements.

Galileo and future GNSS will introduce a wide-band signal that will lead to very low code noise (in the cm-range). The Galileo E5 signal with a wide-band signal (nominal bandwidth of 51.15 MHz) and AltBOC modulation will offer a code noise at the cm-level, enabling reliable ambiguity resolution of the narrow-lane, or, generally speaking, original carrier-phase ambiguities. The multipath level is expected to be in the order of several centimeters for the worst-case environment. An alternative to a broadband signal is to use a high-gain antenna with a very large antenna size. Since the thermal noise is significantly reduced by an increase in
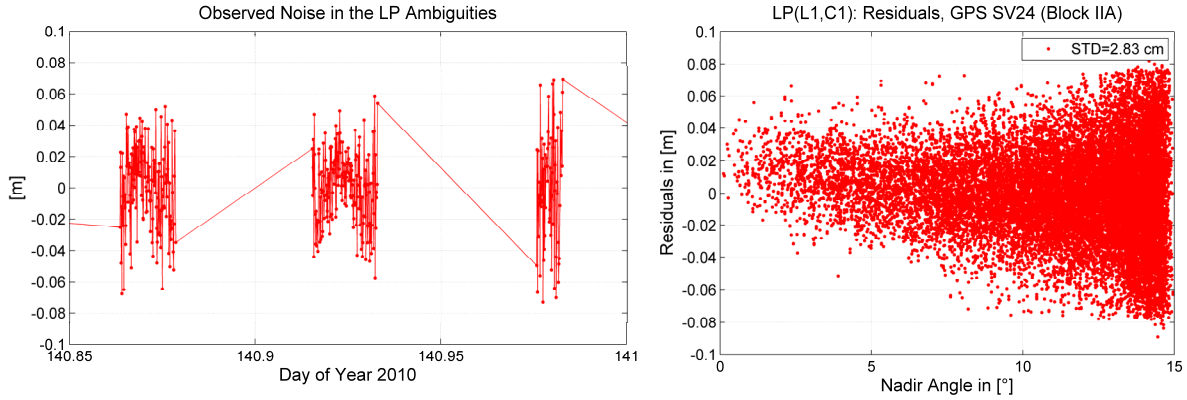
Figure 21.2 Ambiguity resolution using a symmetric geometry-free form of the ionosphere-free linear combination (21.4). The figure on the left shows the noise level in fixing the T2T N_1 ambiguities in terms of single-frequency LP linear combination ("graphic data" of L_1 and C/A code). GPS measurements from the GRACE-B satellite.

the antenna size, in both cases we can get code measurements with noise at the cm-level or even sub-cm precision with very large antennae (VLBI). This offers direct resolution of the narrow-lane ambiguities using a geometry-free form of the ionosphere-free linear combination (21.3) or its symmetric counterpart (21.4).

A closer look at the symmetric geometry-free form of the ionosphere-free linear combination (21.4) reveals differences in the LP linear combination (mean sum of code and phase) on both GPS frequencies that are scaled by constant wide-lane κ_{W1} and narrow-lane κ_{N1} multiplication factors. By forming the LP linear combination, the first order ionosphere effect is removed and the code noise is reduced by 50%. One of the best code tracking performances can be seen in the case of the GRACE-B mission, with code noise at a level of $5-6$ cm . Figure 21.2 shows residuals in the original T2T narrow-lane N_1 ambiguities after forming the LP linear combination. Figure 21.2 confirms that with a low code noise, the symmetric geometry-free form of the iono-free linear combination (21.4) can be used for ambiguity resolution for two-frequency GNSS measurements.

21.3 General Geometry-Free Form of the Ionosphere-free Linear Combination

Let us now find the general form of the linear combination of carrier-phase and/or pseudo-range measurements that fulfils both the ionosphere-free and the geometry-free condition at the same time. In addition, such a linear combination should be applicable to two-frequency as well as multi-frequency GNSS measurements. The general form of the linear combination $L_{LC}^{(1,2,...,q)}$ of q carrier-phase observables can be defined as

$$L_{LC}^{(1,2,...,q)}(L_1, L_2, ..., L_q) := \alpha_1 \frac{f_1}{f_{LC}} L_1 + \alpha_2 \frac{f_2}{f_{LC}} L_2 + ... + \alpha_q \frac{f_q}{f_{LC}} L_q \tag{21.5}$$

where $f_{LC}^{(1,2,...,q)}$ denotes the frequency of the linear combination

$$f_{LC}^{(1,2,...,q)} := \alpha_1 f_1 + \alpha_2 f_2 + ... + \alpha_q f_q \qquad \alpha_i \in R \tag{21.6}$$

In the case of narrow-lane and wide-lane type linear combinations

$$\alpha_i \in \left\{ -1, 1 \right\} \tag{21.7}$$

Introducing the multiplication factor κ_i

$$\kappa_i := \alpha_i \frac{f_i}{f_{LC}} \tag{21.8}$$

we finally obtain the general form of the linear combination for multi-carrier GNSS measurements as

$$L_{LC}^{(1,2,...,q)}(L_1, L_2, ..., L_q) := \kappa_1 L_1 + \kappa_2 L_2 + ... + \kappa_q L_q \tag{21.9}$$

The basic condition one can formulate in forming any linear combination is the geometry-free condition that removes the geometrical terms from the observation equation. The geometry-free condition is fulfilled if the sum of all multiplication factors κ_i, of n observables, is equal to zero

$$\kappa_1 + \kappa_2 + ... + \kappa_n := 0 \tag{21.10}$$

The geometry-free condition guarantees that ambiguities are estimated solely by means of measurements. In a similar way, the geometry-preserving condition is given when the sum of the coefficients is equal to one

$$\kappa_1 + \kappa_2 + ... + \kappa_n := 1 \tag{21.11}$$

In order to remove the first-order ionosphere effect, we need to formulate an ionosphere-free condition that could easily be derived by setting the sum of the first-order ionosphere effects I_i equal to zero for each observable

$$\kappa_1 I_1 + \kappa_2 I_2 + ... + \kappa_n I_n := 0 \tag{21.12}$$

or in the final form

$$\kappa_1 + \kappa_2 \frac{f_1^2}{f_2^2} + ... + \kappa_n \frac{f_1^2}{f_n^2} := 0 \tag{21.13}$$

It is assumed that higher-order ionosphere effects can be pre-computed with sufficient accuracy, and, considering their size, will not have any significant impact on ambiguity resolution. When multiplication factors fulfill the ionosphere-free condition, we may define the ambiguity linear combination

$$\kappa_1 \lambda_1 N_1 + \kappa_2 \lambda_2 N_2 + ... + \kappa_q \lambda_q N_q := \lambda N \tag{21.14}$$

with the ionosphere-free ambiguity λN. In this section we will show in more detail that the ionosphere-free ambiguity term can be represented by an integer N and an ionosphere-free wavelength λ, as with any other carrier-phase observable. In the general case of the ambiguity linear combination (21.14)

$$q \leq n \tag{21.15}$$

if carrier-phase measurements are combined with pseudorange measurements. In order to preserve the integer nature of the ionosphere-free ambiguity, we need to formulate an additional, what we call, the integer ambiguity condition. One way to derive the integer ambiguity condition is to introduce into (21.14) the normalized wavelength $\bar{\lambda}_i$, defined as

$$\bar{\lambda}_i := \frac{\lambda_i}{\lambda} \tag{21.16}$$

and the normalized ambiguity linear combination

$$\kappa_1 \bar{\lambda}_1 N_1 + \kappa_2 \bar{\lambda}_2 N_2 + ... + \kappa_q \bar{\lambda}_q N_q := N \tag{21.17}$$

which gives the integer ambiguity equation defined as

$$N := i_1 N_1 + i_2 N_2 + i_3 N_3 + ... + i_q N_q \qquad i_1, i_2, i_3, ... i_q \in Z \tag{21.18}$$

with

$$i_1 = \kappa_1 \overline{\lambda}_1, \quad i_2 = \kappa_2 \overline{\lambda}_2, \quad ..., \quad i_q = \kappa_q \overline{\lambda}_q \tag{21.19}$$

From (21.6) it follows that in the case of wide-lane and narrow-lane type linear combinations the following integer ambiguity equation can be defined

$$N := \alpha_1 N_1 + \alpha_2 N_2 + ... + \alpha_q N_q \qquad \alpha_i \in Z \tag{21.20}$$

Another integer ambiguity condition for four-frequency Galileo measurements can be found in (Ji et al. 2007) and for code-phase linear combinations in (Henkel 2008).

An elegant way to find the integer multiplication factors in (21.18) is to make use of the wide-lane ambiguities that can be resolved directly at the zero-difference level. Wide-lane ambiguities align the phase ambiguities between two different carrier-phase observables and for a particular frequency pair can be reliably determined to their integer values using the Melbourne-Wübbena linear combination. Wide-lane and narrow-lane ambiguities relative to the reference ambiguity N_1 can then be defined as

$$\begin{aligned} N_W &:= N_1 - N_2 \quad ... \quad N_{W(1,q)} := N_1 - N_q \\ N_N &:= N_1 + N_2 \quad ... \quad N_{N(1,q)} := N_1 + N_q \end{aligned} \tag{21.21}$$

and after substitution into the ambiguity linear combination (21.14), we obtain the expression for the narrow-lane-wide-lane ambiguity linear combination of the generalized ionosphere-free ambiguity

$$(\kappa_1 \lambda_1 + \kappa_2 \lambda_2 + ... + \kappa_q \lambda_q) N_1 - (\kappa_2 \lambda_2 N_W + \kappa_3 \lambda_3 N_{W(1,3)} + ... + \kappa_q \lambda_q N_{W(1,q)}) = \lambda N \tag{21.22}$$

or, in short

$$\lambda N = N_1 \sum_{i=1}^{q} \kappa_i \lambda_i - \sum_{i=2}^{q} \kappa_i \lambda_i N_{W(1,i)} \tag{21.23}$$

Substituting (21.8) for $\alpha_i \in Z$ (wide-lane/narrow-lane type linear combinations) into (21.23) we obtain

$$N = N_1 \sum_{i=1}^{q} \alpha_i - \sum_{i=2}^{q} \alpha_i N_{W(1,i)} \tag{21.24}$$

Since

$$N_{W(1,q)} = N_W + N_{W(2,q)} \tag{21.25}$$

we finally obtain

$$N = N_1 \sum_{i=1}^{q} \alpha_i - N_W \sum_{i=2}^{q} \alpha_i - \sum_{i=3}^{q} \alpha_i N_{W(2,i)} \tag{21.26}$$

The simplest form of (21.22) can be written as

$$\begin{aligned} &\frac{1}{q}\left(\lambda_N + \lambda_{N(1,3)} + ... + \lambda_{N(1,q)}\right) N_1 + \\ &+ \frac{1}{2q}\left[\left(\lambda_W - \lambda_N\right) N_W + \left(\lambda_{W(1,3)} - \lambda_{N(1,3)}\right) N_{W(1,3)} + ... + \left(\lambda_{W(1,q)} - \lambda_{N(1,q)}\right) N_{W(1,q)}\right] = \lambda N \end{aligned} \tag{21.27}$$

or

$$\lambda N := N_1 \frac{1}{q} \sum_{i=2}^{q} \lambda_{N(1,i)} + \frac{1}{2q} \sum_{i=2}^{q} \left(\lambda_{W(1,i)} - \lambda_{N(1,i)} \right) N_{W(1,i)} \qquad (21.28)$$

that reduces to

$$\lambda N := N_1 \frac{1}{q} \sum_{i=2}^{q} \lambda_{N(1,i)} + \frac{1}{2q} \sum_{i=2}^{q} \left(\lambda_{W(1,i)} - \lambda_{N(1,i)} \right) N_W + \frac{1}{2q} \sum_{i=3}^{q} \left(\lambda_{W(1,i)} - \lambda_{N(1,i)} \right) N_{W(2,i)} \qquad (21.29)$$

Eq. (21.28) combines all possible carrier-phase measurements in a multi-frequency GNSS environment, reducing the noise level by $\sqrt{q}$, and thus is equivalent to processing all measurements without forming any linear combinations and estimating one ionosphere-free slant delay per epoch and satellite. The advantage of estimating an additional ionosphere-free slant delays is in the absorption of one common multipath effect per epoch and satellite. However, such an epoch-wise bias could also be estimated on the level of ionosphere-free linear combinations. Nevertheless, if precise point positioning is based on estimated clock parameters using either of these two approaches, the results will be consistent in both cases if carried out in a consistent manner. This is especially important considering that two-frequency ionosphere-free linear combinations will be standard for all Galileo services, as is the case for GPS and all four GNSS. In the case of precise point positioning, an additional epoch-wise bias can always be estimated to average out common systematic effects, such as multipath and front-end effects of the receiver. The estimation of an epoch-wise bias per satellite and receiver was first performed in (Schaer 1999) in the case of two-frequency GPS measurements, where this parameter was called the SIP or the stochastic ionosphere parameter.

In the case of carrier-phase measurements from two GPS frequencies L_1 and L_2, the general form (21.28) reduces to the well-known expression for the ionosphere-free bias that is actually a float ambiguity

$$\lambda_N N_1 + \frac{1}{2} (\lambda_W - \lambda_N) N_W = \lambda_3 N_3 \qquad (21.30)$$

typically denoted as N_3 and the associated wavelength as λ_3. We will see later that N_3 is an integer ambiguity with the specific wavelength λ_3. For the integer properties of the ionosphere-free ambiguity, we refer to Section 21.5. The ionosphere-free ambiguity in (21.30) is directly related to the general form of the ionosphere- and geometry-free linear combination for two-frequency GPS measurements.

21.4 Triangular Form of Wide-Lane Ambiguities

Considering all possible dual-frequency pairs of multi-frequency GNSS measurements, wide-lane ambiguities can be resolved using the Melbourne-Wübbena linear combination. If the frequencies in such a pair are very close to each other, e.g., L_2 and L_5, the resulting wide-lane ambiguity will have a so-called super-wavelength (≈ 5.86 m), about an order of magnitude larger than the original wide-lane wavelength between the L_1 and L_2 carrier-phase observables. However, the noise floor of such a super wide-lane linear combination will be increased by a factor of about 33. Nevertheless, it will still be, by a factor of about 2, the wide-lane ambiguity to be best determined of the three wide-lane linear combinations. Thus, in order to use such a super wide-lane ambiguity as an additional constraint in the estimation of wide-lane ambiguities, we introduce the vector form of the three-carrier wide-lane ambiguities as depicted in Figure 21.3
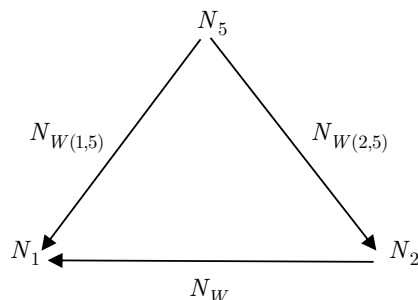
Figure 21.3 Triangular form of the three-carrier wide-lane ambiguities. The three pairs of wide-lane ambiguities are fully linearly dependent, but only two pairs can be estimated independently.

$$
\begin{aligned}
N_W &= N_1 - N_2 \\
N_{W(1,5)} &= N_1 - N_5 \\
N_{W(2,5)} &= N_2 - N_5
\end{aligned}
\tag{21.31}
$$

from which it follows the triangular form of the wide-lane ambiguity

$$
N_{W(2,5)} = N_{W(1,5)} - N_W
\tag{21.32}
$$

that can be used to additionally constrain the resolution of the other two wide-lane ambiguities. From (21.32) we see that for all frequency pairs, the wide-lane ambiguities can be reliably fixed to their integer values and be used to align carrier-phase ambiguities between different frequencies. However, all three pairs of ambiguities are fully linearly dependent and one can estimate only two pairs independently.

A similar geometry-free approach can also be applied when different measurements on the same frequency are available. For instance, in the case of two-frequency GPS measurements from the GRACE-B satellite, we can form two different narrow-lane and two wide-lane ambiguities between the $L_1 - L_{1A}$ phase measurements on the first frequency and the second GPS frequency. It is assumed that the common ambiguities on the same frequency can easily be fixed between $L_1 - L_{1A}$ phase measurements, as demonstrated in Section 20.4. Thus with the reference ambiguity denoted as N_1 on P_1, we can write the following transformed ambiguity equation

$$
\lambda N = N_1 \frac{1}{2n} \sum_1^n \lambda_{N(1,i)} + \frac{1}{2n} \sum_1^n (\lambda_{W(1,i)} - \lambda_{N(1,i)}) N_{W(1,i)}
\tag{21.33}
$$

From (21.33) we see that, with this technique related to a parameter transformation, we can combine all independent ionosphere-free linear combinations, transforming all narrow-lane ambiguities into the ionosphere-free linear combination with the common N_1 ambiguity, see also (21.29). Thus in the case of multi-frequency GNSS measurements, the number of parameters is the same as for two-frequency GPS using L_3. However, the noise level can be decreased by about $\sqrt{2}$ by adding the second ionosphere-free linear combination $L_{3(1,5)}$ in the case of the third GPS frequency, or by adding the ionosphere-free linear combination based on L_{1A} carrier-phase $L_{3(1A,5)}$.

21.5 Ambiguity-Free Linear Combinations – Geometry-Free Ambiguity Resolution of Wide-Lane and Narrow-Lane Ambiguities

Here we introduce an elegant way to resolve wide-lane and narrow-lane ambiguities with the ambiguity-free linear combination that is consistent with the code biases (see section on the absolute code biases). Let us write the ionosphere-free linear combination for L_1 and L_2, and, in addition, for L_2 and L_5 carrier-phase measurements

$$
\begin{aligned}
L_3^{1,2} &= \rho + \lambda_N N_1 + \frac{1}{2}\big(\lambda_W - \lambda_N\big) N_W \\
L_3^{2,5} &= \rho + \lambda_{N(2,5)} N_1 - \lambda_{N(2,5)} N_W + \frac{1}{2}\big(\lambda_{W(2,5)} - \lambda_{N(2,5)}\big) N_{W(2,5)}
\end{aligned}
\tag{21.34}
$$

We now define the following ambiguity-free condition

$$
\kappa_1^{af*}\lambda_N + \kappa_2^{af*}\lambda_{N(2,5)} := 0
\tag{21.35}
$$

satisfying the geometry condition

$$
\kappa_1^{af*} + \kappa_2^{af*} = 1
\tag{21.36}
$$

from which the following expression to calculate ambiguity-free multiplication factors that are of very moderate magnitude results:

$$
\kappa_1^{af*} = \frac{-\lambda_{N(2,5)}}{\lambda_N - \lambda_{N(2,5)}} = \frac{f_1 + f_2}{f_1 - f_5} \approx 7.02
\qquad
\kappa_2^{af*} = \frac{\lambda_N}{\lambda_N - \lambda_{N(2,5)}} = -\frac{f_2 + f_5}{f_1 - f_5} \approx -6.02
\tag{21.37}
$$

Applying the ambiguity-free condition to the ionosphere-free linear combination, we obtain the following narrow-lane ambiguity-free linear combination

$$
\begin{aligned}
L_3^{af*} &:= \kappa_1^{af*} L_3^{1,2} + \kappa_2^{af*} L_3^{2,5} \\
&= \rho + \left[\frac{\kappa_1^{af*}}{2}\big(\lambda_W - \lambda_N\big) - \kappa_2^{af*}\lambda_{N(2,5)}\right] N_W + \frac{\kappa_2^{af*}}{2}\big(\lambda_{W(2,5)} - \lambda_{N(2,5)}\big) N_{W(2,5)}
\end{aligned}
\tag{21.38}
$$

with the following wide-lane wavelengths of considerable magnitude

$$
\begin{aligned}
\lambda_W^{af*} &= \frac{\kappa_1^{af*}}{2}\big(\lambda_W - \lambda_N\big) - \kappa_2^{af*}\lambda_{N(2,5)} &\approx 3.40 \text{ m} \\
\lambda_{W(2,5)}^{af*} &= \frac{\kappa_2^{af*}}{2}\big(\lambda_{W(2,5)} - \lambda_{N(2,5)}\big) &\approx -17.28 \text{ m}
\end{aligned}
\tag{21.39}
$$

Let us now repeat the procedure with the (L_1, L_2) and the (L_1, L_5) combination of the ionosphere-free linear combination

$$
\begin{aligned}
L_3^{1,2} &= \rho + \lambda_N N_1 + \frac{1}{2}\big(\lambda_W - \lambda_N\big) N_W \\
L_3^{1,5} &= \rho + \lambda_{N(1,5)} N_1 + \frac{1}{2}\big(\lambda_{W(1,5)} - \lambda_{N(1,5)}\big) N_{W(1,5)}
\end{aligned}
\tag{21.40}
$$

We thus obtain the ambiguity-free multiplication factors

$$\kappa_1^{af^{**}} = \frac{-\lambda_{N(1,5)}}{\lambda_N - \lambda_{N(1,5)}} = \frac{f_1 + f_2}{f_2 - f_5} \approx 54.8 \qquad \kappa_2^{af^{**}} = \frac{\lambda_N}{\lambda_N - \lambda_{N(1,5)}} = -\frac{f_1 + f_5}{f_2 - f_5} \approx -53.8 \qquad (21.41)$$

and the following narrow-lane ambiguity-free linear combination

$$
\begin{aligned}
L_3^{af^{**}} &:= \kappa_1^{af^{**}} L_3^{1,2} + \kappa_2^{af^{**}} L_3^{1,5} \\
&= \rho + \frac{\kappa_1^{af^{**}}}{2}\Big(\lambda_W - \lambda_N\Big)N_W + \frac{\kappa_2^{af^{**}}}{2}\Big(\lambda_{W(1,5)} - \lambda_{N(1,5)}\Big)N_{W(1,5)}
\end{aligned}
\tag{21.42}
$$

with the following wide-lane wavelengths

$$
\begin{aligned}
\lambda_W^{af^{**}} &= \frac{\kappa_1^{af^{**}}}{2}\Big(\lambda_W - \lambda_N\Big) &&\approx 20.69 \text{ m} \\
\lambda_{W(1,5)}^{af^{**}} &= \frac{\kappa_2^{af^{**}}}{2}\Big(\lambda_{W(1,5)} - \lambda_{N(1,5)}\Big) &&\approx -17.28 \text{ m}
\end{aligned}
\tag{21.43}
$$

Repeating the procedure with the third combination of ionosphere-free linear combination

$$
\begin{aligned}
L_3^{1,5} &= \rho + \lambda_{N(1,5)}N_1 + \frac{1}{2}\Big(\lambda_{W(1,5)} - \lambda_{N(1,5)}\Big)N_{W(1,5)} \\
L_3^{2,5} &= \rho + \lambda_{N(2,5)}N_1 - \lambda_{N(2,5)}N_W + \frac{1}{2}\Big(\lambda_{W(2,5)} - \lambda_{N(2,5)}\Big)N_{W(2,5)}
\end{aligned}
\tag{21.44}
$$

we obtain the ambiguity-free multiplication factors

$$\kappa_1^{af^{***}} = \frac{-\lambda_{N(2,5)}}{\lambda_{N(1,5)} - \lambda_{N(2,5)}} = \frac{f_1 + f_5}{f_1 - f_2} \approx 7.91, \qquad \kappa_2^{af^{***}} = \frac{\lambda_{N(1,5)}}{\lambda_{N(1,5)} - \lambda_{N(2,5)}} = -\frac{f_2 + f_5}{f_1 - f_2} \approx -6.91 \quad (21.45)$$

and the following narrow-lane ambiguity-free linear combination

$$
\begin{aligned}
L_3^{af^{***}} &:= \kappa_1^{af^{***}} L_3^{1,5} + \kappa_2^{af^{***}} L_3^{2,5} = \\
&= \rho - \kappa_2^{af^{***}}\lambda_{N(2,5)}N_W + \frac{\kappa_1^{af^{***}}}{2}\Big(\lambda_{W(1,5)} - \lambda_{N(1,5)}\Big)N_{W(1,5)} + \frac{\kappa_2^{af^{***}}}{2}\Big(\lambda_{W(2,5)} - \lambda_{N(2,5)}\Big)N_{W(2,5)}
\end{aligned}
\tag{21.46}
$$

where $\kappa_1^{af^{***}}\lambda_{N(1,5)}N_1 + \kappa_2^{af^{***}}\lambda_{N(2,5)}N_1 = 0$ due to the ambiguity-free condition for N_1, i.e., the multiplication factors (21.45). For the wide-lane wavelengths in (21.46) we obtain

$$
\begin{aligned}
\lambda_W^{af^{***}} &= -\kappa_2^{af^{***}}\lambda_{N(2,5)} &&\approx -0.86 \text{ m} \\
\lambda_{W(1,5)}^{af^{***}} &= \frac{\kappa_1^{af^{***}}}{2}\Big(\lambda_{W(1,5)} - \lambda_{N(1,5)}\Big) &&\approx 2.54 \text{ m} \\
\lambda_{W(2,5)}^{af^{***}} &= \frac{\kappa_2^{af^{***}}}{2}\Big(\lambda_{W(2,5)} - \lambda_{N(2,5)}\Big) &&\approx -19.82 \text{ m}
\end{aligned}
\tag{21.47}
$$

It can be shown that the differences of the two linear combinations (21.38) and (21.42) is equal to zero

$$L_3^{af^*} - L_3^{af^{**}} = 0 = \lambda_W^+ N_W + \lambda_{W(1,5)}^+ N_{W(1,5)} + \lambda_{W(2,5)}^+ N_{W(2,5)} \tag{21.48}$$

where

$$\lambda_W^+ = \frac{\kappa_1^{af*} - \kappa_1^{af**}}{2}\left(\lambda_W - \lambda_N\right) - \kappa_2^{af*}\lambda_{N(2,5)} \quad \approx -17.28 \text{ m}$$

$$\lambda_{W(1,5)}^+ = -\frac{\kappa_2^{af**}}{2}\left(\lambda_{W(1,5)} - \lambda_{N(1,5)}\right) \quad\quad \approx \ \ 17.28 \text{ m} \quad\quad (21.49)$$

$$\lambda_{W(2,5)}^+ = \frac{\kappa_2^{af*}}{2}\left(\lambda_{W(2,5)} - \lambda_{N(2,5)}\right) \quad\quad \approx -17.28 \text{ m}$$

or

$$L_3^{af*} - L_3^{af**} = \left[\frac{\kappa_1^{af*}}{2}\left(\lambda_W - \lambda_N\right) - \kappa_2^{af*}\lambda_{N(2,5)}\right]N_W + \frac{\kappa_2^{af*}}{2}\left(\lambda_{W(2,5)} - \lambda_{N(2,5)}\right)N_{W(2,5)} -$$
$$-\frac{\kappa_1^{af**}}{2}\left(\lambda_W - \lambda_N\right)N_W - \frac{\kappa_2^{af**}}{2}\left(\lambda_{W(1,5)} - \lambda_{N(1,5)}\right)N_{W(1,5)} \quad\quad (21.50)$$

that can be reduced to

$$L_3^{af*} - L_3^{af**} = \left[\frac{\kappa_1^{af*} - \kappa_1^{af**}}{2}\left(\lambda_W - \lambda_N\right) - \kappa_2^{af*}\lambda_{N(2,5)}\right]N_W +$$
$$+\frac{\kappa_2^{af*}}{2}\left(\lambda_{W(2,5)} - \lambda_{N(2,5)}\right)N_{W(2,5)} - \frac{\kappa_2^{af**}}{2}\left(\lambda_{W(1,5)} - \lambda_{N(1,5)}\right)N_{W(1,5)} \quad\quad (21.51)$$

Thus, there is a way to resolve the wide-lane ambiguities and obtain ambiguity-free linear combination considering only wide-lane ambiguities from the GPS carrier-phase measurements on the three frequencies.

Another approach to remove geometry in these linear combinations is to form ambiguity-free linear combinations of single code measurements. This could be very interesting for future wide-band GNSS signals, such as PRS code on Galileo E6 that offers cm-level precision. For this, we make use of the LP linear combination $LP = (L + P)/2$ and the ionosphere-free linear combination L_3 of two carrier-phase measurements L_1 and L_2

$$AF_1 := \kappa_1^{af}L_3 + \kappa_2^{af}LP_1 \quad\quad (21.52)$$

that contains only an absolute code bias from P_1 code measurements. The geometry-preserving condition for multiplication factors κ_1^{af} and κ_2^{af} is then as follows

$$\kappa_1^{af} + \kappa_2^{af} := 1 \qu\quad (21.53)$$

For the first time, we are introducing here an ambiguity-free condition (a condition to remove an ambiguity that is common to a pair of linear combinations) that for ambiguity N_1 on L_1 carrier-phase is defined as

$$\kappa_1^{af}\lambda_N + \kappa_2^{af}\frac{\lambda_1}{2} := 0 \qu\quad (21.54)$$

where λ_N denotes the narrow-lane wavelength of the narrow-lane ambiguity in ionosphere-free linear combination L_3 and $\lambda_1/2$ is the wavelength of the L_1 ambiguity in the LP_1 linear combination. The ambiguity-free condition is fulfilled as long as wide-lane ambiguities are fixed, i.e., all ambiguities are aligned to each other $N_1 = N_2 = N_5$, using e.g., Melbourne-Wübbena linear combination. After solving (21.53) and (21.54), for the multiplication factors of ambiguity-free linear combination in (21.52) we obtain

$$\kappa_1^{af} = -\frac{f_1 + f_2}{f_1 - f_2}, \quad\quad \kappa_2^{af} = \frac{2f_1}{f_1 - f_2} \qu\quad (21.55)$$

We can also write ambiguity-free linear combination (21.52) for other code measurements and frequencies. For $LP_2 = (L_2 + P_2)/2$ we obtain

$$AF_2 := \kappa_{1(2)}^{af} L_3 + \kappa_{2(2)}^{af} LP_2 \tag{21.56}$$

with

$$\kappa_{1(2)}^{af} = \frac{f_1 + f_2}{f_1 - f_2}, \qquad \kappa_{2(2)}^{af} = -\frac{2f_2}{f_1 - f_2} \tag{21.57}$$

and for $LP_5 = (L_5 + C_5)/2$

$$AF_5 := \kappa_{1(5)}^{af} L_3 + \kappa_{2(5)}^{af} LP_5 \tag{21.58}$$

with

$$\kappa_{1(5)}^{af} = -\frac{f_1 + f_2}{2f_5 - f_1 - f_2}, \qquad \kappa_{2(5)}^{af} = \frac{2f_5}{2f_5 - f_1 - f_2} \tag{21.59}$$

Accordingly, for the Galileo E6 signal we introduce $LP_6 = (L_6 + E_6)/2$ that divides E_6 by 2 and reduces the code noise of the original E6 frequency by 50%

$$AF_6 := \kappa_{1(6)}^{af} L_{3(1,6)} + \kappa_{2(6)}^{af} LP_6 \tag{21.60}$$

with

$$\kappa_{1(E6)}^{af} = \frac{f_1 + f_6}{f_1 - f_6}, \qquad \kappa_{2(E6)}^{af} = -\frac{2f_6}{f_1 - f_6} \tag{21.61}$$

Such an ambiguity-free linear combination could offer a noise level of $10-15$ cm and could be used for the resolution of wide-lane and narrow-lane ambiguities by making differences to e.g., (21.38) and removing the geometry term. Once singe-frequency ambiguity is resolved on the wideband GNSS signal, e.g., E_6, all ambiguities are resolved, since wide-lane ambiguities can be fixed to their integer values. This is also true for the absolute code biases that could be used as a reference for the absolute calibration of code measurements and the resolution of wide-lane ambiguities; see the section on absolute code biases and calibration of code measurements. The advantage of this approach is that it offers consistent resolution of wide-lane ambiguities and calibration of wide-lane biases in an absolute sense, since the same ambiguity-free linear combination is used to estimate single-frequency absolute code biases. Since absolute code biases are determined against the ionosphere-free P-code observable (IGS conventions for clock parameters) we can establish a consistent framework to process observables that could be extended to all four-GNSS.

21.6 Integer Ambiguity Algebra and the Integer Property of the Ionosphere-Free Linear Combination

The ionosphere-free ambiguity of L_1 and L_2 carrier-phase measurements is a real number and can be decomposed into the sum of a narrow-lane and a wide-lane ambiguity. Here we show that it is possible to express an ionosphere-free linear combination as a function of a single integer ionosphere-free ambiguity without resolving the wide-lane ambiguity.

We start with the ionosphere-free linear combination and introduce the integer ionosphere-free ambiguity N_3 with the wavelength λ_3 as follows:

$$\kappa_1 \lambda_1 N_1 + \kappa_2 \lambda_2 N_2 = \lambda_3 N_3, \qquad N_3 \in Z \tag{21.62}$$

with corresponding multiplication factors κ_1 and κ_2 defined as

$$\kappa_1 = \frac{f_1^2}{f_1^2 - f_2^2}, \quad \kappa_2 = -\frac{f_2^2}{f_1^2 - f_2^2} \tag{21.63}$$

These can easily be derived from the ionosphere-free (21.13) and geometry-preserving condition (21.11)

$$\kappa_1 + \kappa_2 \frac{f_1^2}{f_2^2} = 0, \quad \kappa_1 + \kappa_2 = 1 \tag{21.64}$$

It is interesting to note that the ionosphere-free multipliers can be represented as the product of the wide-lane and the narrow-lane multipliers on the first and the second frequency

$$\kappa_1 = \kappa_{W(1)} \cdot \kappa_{N(1)}, \quad \kappa_2 = \kappa_{W(2)} \cdot \kappa_{N(2)} \tag{21.65}$$

with

$$\kappa_{N(1)} = \frac{f_1}{f_1 + f_2}, \quad \kappa_{N(2)} = \frac{f_2}{f_1 + f_2} \tag{21.66}$$

$$\kappa_{W(1)} = \frac{f_1}{f_1 - f_2}, \quad \kappa_{W(2)} = -\frac{f_2}{f_1 - f_2} \tag{21.67}$$

Let us now substitute (21.63) into (21.62) to obtain

$$\frac{f_1}{f_1^2 - f_2^2} N_1 - \frac{f_2}{f_1^2 - f_2^2} N_2 = \frac{1}{f_3} N_3 \tag{21.68}$$

that in the case of aligned ambiguities $N_1 = N_2$, after wide-lane ambiguity resolution reduces to

$$\frac{1}{f_1 + f_2} N_1 = \frac{1}{f_3} N_3 \tag{21.69}$$

$$f_3 = f_1 + f_2 = (154 + 120) \cdot f_0 \qquad \rightarrow \qquad \lambda_3 = \lambda_N \tag{21.70}$$

Let us now write (21.62) in the following form

$$N_1 + \frac{\kappa_2 \lambda_2}{\kappa_1 \lambda_1} N_2 = \frac{\lambda_3}{\kappa_1 \lambda_1} N_3 \tag{21.71}$$

and since

$$\frac{\kappa_2 \lambda_2}{\kappa_1 \lambda_1} = -\frac{f_2}{f_1} = -\frac{120 \cdot f_0}{154 \cdot f_0} = -\frac{60}{77} \tag{21.72}$$

it follows that

$$f_1 N_1 - f_2 N_2 = \frac{f_1^2 - f_2^2}{f_3} N_3 \tag{21.73}$$

Eq. (21.73) will remain unchanged if the two GPS frequencies $154 \cdot f_0$ and $120 \cdot f_0$ are divided by the fundamental GPS frequency $f_0 = 10.23$ MHz. The same is true for (21.62). Therefore, we may normalize all GPS

frequencies with the fundamental GPS frequency f_0 and consider f_1 and f_2 to be integers of 154 and 120 respectively.

One of the integer ambiguity solutions of (21.73) is

$$f_3 = (f_1^2 - f_2^2) / f_0 = (f_1 - f_2)(f_1 + f_2) / f_0 = f_N f_W \cdot f_0 \tag{21.74}$$

or finally

$$f_3 := f_N f_W \cdot f_0 \tag{21.75}$$

with the normalized wide-lane frequency $f_W = (f_1 - f_2)/f_0$ and the normalized narrow-lane frequency $f_N = (f_1 + f_2)/f_0$. When ambiguities $N_1 = N_2$ are aligned, e.g., after wide-lane ambiguity resolution, we may write

$$f_3 = f_1^2 - f_2^2 \quad \rightarrow \quad f_3 N_1 = (f_1 + f_2) N_3 \tag{21.76}$$

Thus

$$f_W N_1 = N_3 \tag{21.77}$$

giving a direct relationship between the narrow-lane wavelength λ_N and the ionosphere-free wavelength λ_3

$$\lambda_3 = \frac{\lambda_N}{f_W \cdot f_0} \approx 3.14 \text{ mm} \tag{21.78}$$

Since f_W is an even number in the case of the two GPS frequencies, we may further write

$$\lambda_3 = \frac{\lambda_N}{\frac{f_W}{2} \cdot f_0} = 2 \frac{\lambda_N}{f_W \cdot f_0} \approx 6.29 \text{ mm} \tag{21.79}$$

The same expression can be developed following (21.68) that for two GPS frequencies gives

$$77 N_1 - 60 N_2 = 17 N_1 + 60 N_W = 77 \frac{\lambda_3}{\kappa_1 \lambda_1} N_3 \tag{21.80}$$

If N_1 and N_2 are integers, the ionosphere-free ambiguity N_3 will be an integer when the following condition is met

$$77 \frac{\lambda_3}{\kappa_1 \lambda_1} = \pm 1 \tag{21.81}$$

Finally, the integer equation of the ionosphere-free ambiguity N_3 of the ionosphere-free linear combination can be defined as

$$N_3 := 77 N_1 - 60 N_2 = 17 N_1 + 60 N_W = 137 N_1 - 60 N_N \tag{21.82}$$

N_W and N_N denote the wide-lane and narrow-lane ambiguity respectively and λ_3 is the ionosphere-free wavelength

$$77 \frac{\lambda_3}{\kappa_1 \lambda_1} = 1 \quad \rightarrow \quad \lambda_3 = \frac{\kappa_1 \lambda_1}{77} \quad \rightarrow \quad \lambda_3 = \frac{1}{4658} \cdot \frac{c}{f_0} \approx 6.29 \text{ mm} \tag{21.83}$$

with the fundamental GPS frequency $f_0 = 10.23$ MHz and the speed of light c. The ionosphere-free linear combination with the integer ionosphere-free ambiguity N_3 is then defined as

$$L_3 = \kappa_1 L_1 + \kappa_2 L_2 := \rho + \lambda_3 N_3 \tag{21.84}$$

with the geometry term denoted by ρ. Introducing

$$\frac{n}{m} = \frac{\kappa_1 \lambda_1}{\kappa_2 \lambda_2} = -\frac{f_1}{f_2}, \qquad m, n \in Z \tag{21.85}$$

we obtain the general form of the integer ambiguity equation (21.82)

$$n N_1 + m N_2 = N_3 \tag{21.86}$$

with

$$\lambda_3 = \frac{\kappa_1 \lambda_1}{n} \tag{21.87}$$

The final form of the equation for the wavelength of the ionosphere-free ambiguity λ_3 with an integer ionosphere-free ambiguity N_3 is

$$\lambda_3 := \frac{\lambda_1 \lambda_2}{n \lambda_2 + m \lambda_1} \tag{21.88}$$

or in terms of frequencies

$$\lambda_3 := \frac{c}{n f_1 + m f_2} \tag{21.89}$$

Let us now define a new, transformed ionosphere-free integer ambiguity in the following way

$$\bar{N}_3 := 77 N_3 \tag{21.90}$$

and after substitution in (21.81) we obtain the following new solution for the wavelength λ_3 denoted as $\bar{\lambda}_3$

$$\frac{\bar{\lambda}_3}{\kappa_1 \lambda_1} = 1 \quad \rightarrow \quad \bar{\lambda}_3 := \kappa_1 \lambda_1 \approx 48.44 \text{ cm} \tag{21.91}$$

that is considerably longer than the original ionosphere-free wavelength. From (21.90) and (21.91) we obtain the following ambiguity equation

$$17 N_1 + 60 N_W = 77 N_3 \tag{21.92}$$

showing that in an arbitrary case our new transformed ambiguity $\bar{N}_3$ is not an integer, but rather a float ambiguity. However, the integer condition is fulfilled in the special case

$$\frac{N_1}{N_W} = 1 \quad \rightarrow \quad \bar{N}_3 \in Z \tag{21.93}$$

From the integer equation of the ionosphere-free ambiguity (21.82) and from the transformed ambiguity equation (21.92) we see that adding one narrow-lane ambiguity to our integer equation (21.82) will modify the wide-lane ambiguity to $(N_1 + 1 - N_2)$ and the ionosphere-free linear combination by the wavelength $\bar{\lambda}_3 \approx 48.44$ cm (21.91). Therefore, instead of aligning initial ambiguities $N_1 = N_2$ first, by applying a wide-lane ambiguity, one can first determine the narrow-lane ambiguity N_1 with a relatively long wavelength of $\bar{\lambda}_3 \approx 48.44$ cm. In the second step, the wide-lane ambiguity can be applied, aligning the initial ambiguities $N_1 = N_2$.

From this, we can draw the conclusion that one can add an arbitrary number of integer wide-lane ambiguities to iono-free linear combination, as long as the single-frequency ambiguity N_1 or N_2 is estimated. This also means that there is a mechanism to form iono-free linear combination with an arbitrary wavelength. Let us now find the simplest solution when $N_3 = 0$, i.e., when iono-free integer ambiguity is fixed. From the integer ambiguity equation (21.82) we obtain

$$17N_1 + 60N_W = N_3 = 0 \tag{21.94}$$

and

$$N_1 = -\frac{60}{17}N_W \qquad N_1 \in Z, \quad N_W \in Z \tag{21.95}$$

Therefore, after aligning carrier-phase measurements on both frequencies by wide-lane ambiguity resolution, one can add an arbitrary number of wide-lane ambiguities N_W under the condition

$$N_W = 17 \cdot k \qquad k \in Z \tag{21.96}$$

that gives the following solution for the single-frequency ambiguity N_1

$$N_1 = -60 \cdot k \tag{21.97}$$

This means that adding a number of wide-lane ambiguities $N_W = 17k$ to iono-free linear combination of GPS carrier-phase measurements is equivalent to adding single-frequency ambiguities $N_1 = -60k$, since the total number of iono-free ambiguities in (21.94) will not change.

21.7 Integer Ambiguity Algebra for Narrow-lane and Wide-lane Ambiguities

It can easily be shown that narrow-lane and wide-lane ambiguities with wavelengths

$$\begin{aligned}
\lambda_N &= \frac{c}{f_1 + f_2} = \frac{1}{154 + 120} \cdot \frac{c}{f_0} \quad \rightarrow \quad \lambda_N = \frac{1}{274} \cdot \frac{c}{f_0} \approx 10.70 \text{ cm} \\
\lambda_W &= \frac{c}{f_1 - f_2} = \frac{1}{154 - 120} \cdot \frac{c}{f_0} \quad \rightarrow \quad \lambda_W = \frac{1}{34} \cdot \frac{c}{f_0} \approx 86.19 \text{ cm}
\end{aligned} \tag{21.98}$$

have direct integer properties

$$\frac{\lambda_N}{\lambda_3} = \frac{4658}{274} = 17 \quad \rightarrow \quad \lambda_N = 17 \cdot \lambda_3 \qquad \frac{\lambda_W}{\lambda_3} = \frac{4658}{34} = 137 \quad \rightarrow \quad \lambda_W = 137 \cdot \lambda_3 \tag{21.99}$$

satisfying the following integer relations with wide-lane ambiguities

$$\frac{\kappa_1 \lambda_1 + \kappa_2 \lambda_2}{\lambda_3} N_1 - \frac{\kappa_2 \lambda_2}{\lambda_3} N_W = N_3 \quad \rightarrow \quad \lambda_{3(W)} = \lambda_3, \quad \frac{n}{m} = \frac{17}{60} \tag{21.100}$$

that gives

$$(77 - 60)N_1 + 60N_W = N_3 \tag{21.101}$$

or

$$17N_1 + 60N_W = N_3 \tag{21.102}$$

Once the wide-lane ambiguity is introduced in (21.102) one can form an integer ambiguity of arbitrary wavelength considering that the ambiguity-free linear combination provides absolute (pseudo)-range with an accuracy of about 20 cm RMS. Thus, we obtain in that case

$$17N_1 := N_3 \tag{21.103}$$

that could be used as a "ruler" in the ambiguity space when forming a wavelength of an integer ambiguity. With narrow-lane type ambiguity $\lambda_N N_1$ in the ionosphere-free linear combination we obtain

$$\lambda_N N_1 + \lambda_{NW} N_W = \lambda_3 N_3 \tag{21.104}$$

denoting the intermediate wavelength λ_{NW}, where

$$\lambda_{NW} = \frac{1}{2}(\lambda_W - \lambda_N) = \frac{f_2}{f_1^2 - f_2^2} c \tag{21.105}$$

For the narrow-lane ambiguities it follows that

$$\frac{\kappa_1 \lambda_1 - \kappa_2 \lambda_2}{\lambda_3} N_1 + \frac{\kappa_2 \lambda_2}{\lambda_3} N_N = N_3 \quad \rightarrow \quad \frac{n}{m} = -\frac{137}{60} \tag{21.106}$$

and considering $(77 + 60)N_1 - 60N_N = N_3$ we obtain

$$137N_1 - 60N_N = N_3 \tag{21.107}$$

with a wide-lane type ambiguity $\lambda_W N_1$ in the ionosphere-free linear combination

$$\lambda_W N_1 - \lambda_{NW} N_N = \lambda_3 N_3 \tag{21.108}$$

The sum of narrow-lane and wide-lane ambiguity is then

$$L_N + L_W = 2L_3 \quad \rightarrow \quad \frac{n}{m} = \frac{\lambda_N}{\lambda_W} = \frac{17}{137} \tag{21.109}$$

After substituting (21.98) we derive

$$17N_N + 137N_W = 2N_3 \tag{21.110}$$

21.8 Integer Ambiguity Algebra for the Third GPS Frequency

For other combinations of two fundamental GPS frequencies including L_5 we have

$$\frac{n}{m} = -\frac{f_1}{f_5} = -\frac{154}{115} \qquad \begin{cases} f_1 = 154 \cdot f_0 \\ f_5 = 115 \cdot f_0 \end{cases} \tag{21.111}$$

$$\frac{n}{m} = -\frac{f_2}{f_5} = -\frac{120}{115} = -\frac{24}{23} \quad \begin{cases} f_2 = 120 \cdot f_0 \\ f_5 = 115 \cdot f_0 \end{cases} \tag{21.112}$$

that in the first and second cases gives the following ionosphere-free wavelength

$$\lambda_{3(1,5)} = \frac{1}{10491} \cdot \frac{c}{f_0} \approx 2.8 \text{ mm} \tag{21.113}$$

Similarly, for the second and third GPS frequencies we obtain

$$\lambda_{3(2,5)} = \frac{1}{235} \cdot \frac{c}{f_0} = \frac{c}{f_2 + f_5} = \lambda_{N(2,5)} \approx 12.5 \text{ cm} \tag{21.114}$$

In the second case, the wavelength of the ionosphere-free ambiguity is equal to the narrow-lane ambiguity. In a similar way, the following ambiguity equation can be obtained for the first GPS frequency pair

$$39N_1 + 115N_{W(1,5)} = N_{3(1,5)} \tag{21.115}$$

$$269N_1 - 115N_{N(1,5)} = N_{3(1,5)} \tag{21.116}$$

and the second GPS frequency pair

$$N_2 + 23N_{W(2,5)} = N_{3(2,5)} \tag{21.117}$$

$$47N_2 - 23N_{N(2,5)} = N_{3(2,5)} \tag{21.118}$$

After substituting $N_W = N_1 - N_2$, (21.117) reduces to

$$N_1 - N_W + 23N_{W(2,5)} = N_{3(2,5)} \tag{21.119}$$

From (21.117) it follows that in the case of L_2 and L_5 phase measurements, wide-lane ambiguities can be represented as multiples of the narrow-lane ambiguity

$$\frac{\lambda_{W(2,5)}}{\lambda_{N(2,5)}} = 47 \tag{21.120}$$

with the super wide-lane wavelength

$$\lambda_{W(2,5)} \approx 5.86 \text{ m} \tag{21.121}$$

From (21.117) it follows that for the ionosphere-free linear combination based on L_2 and L_5 phase measurements there is no need to solve wide-lane ambiguities before solving narrow-lane ambiguities, since the iono-free integer has the same wavelength as the narrow-lane ambiguity

$$\begin{aligned}
L_{3(2,5)} &= \rho + \lambda_{N(2,5)}N_2 + \frac{1}{2}(\lambda_{W(2,5)} - \lambda_{N(2,5)})N_{W(2,5)} \\
&= \rho + \lambda_{N(2,5)}(N_2 + 23N_{W(2,5)}) \\
&= \rho + \lambda_{N(2,5)}N_{3(2,5)}
\end{aligned} \tag{21.122}$$

21.9 Integer Ambiguity Algebra for Galileo Ambiguities

In a similar way we can apply integer ambiguity algebra to Galileo measurements. In the case of Galileo frequencies, the longest wavelength can be obtained by combining L_{5a} and L_6 observables

$$\frac{n}{m} = -\frac{f_6}{f_{5a}} = -\frac{125}{115} = -\frac{25}{23} \qquad \begin{cases} f_6 = 125 \cdot f_0 \\ f_{5a} = 115 \cdot f_0 \end{cases} \tag{21.123}$$

from which follows the ionosphere-free wavelength, which is half the narrow-lane wavelength $\lambda_{N(6,5a)}$

$$\lambda_{3(6,5a)} = \frac{1}{480} \cdot \frac{c}{f_0} = \frac{c}{2(f_6 + f_{5a})} = \frac{1}{2}\lambda_{N(6,5a)} \quad \rightarrow \quad \lambda_{3(6,5a)} = 6.1 \text{ cm} \tag{21.124}$$

The corresponding ionosphere-free frequency $f_{3(6,5a)}$ is then defined as

$$f_{3(6,5a)} := 480 \cdot f_0 = 2\left(f_6 + f_{5a}\right) = 2f_{N(6,5a)} \quad \rightarrow \quad \lambda_{3(6,5a)} = \frac{1}{2}\lambda_{N(6,5a)} \tag{21.125}$$

with the corresponding integer ambiguity equation

$$25N_6 - 23N_{5a} = N_{3(6,5a)} \tag{21.126}$$

Inserting the wide-lane $N_{W(6,5a)}$ and the narrow-lane $N_{N(6,5a)}$ ambiguity, we finally obtain

$$2N_6 + 23N_{W(6,5a)} = N_{3(6,5a)} \tag{21.127}$$

$$48N_6 - 23N_{N(6,5a)} = N_{3(6,5a)} \tag{21.128}$$

Let us now define the frequencies of the new wide-lane and narrow-lane linear combination, respectively

$$f_W^{(1,6,3)} := f_1 + f_{5a} - \frac{1}{2}f_{3(6,5a)}$$
$$f_N^{(1,6,3)} := f_1 - f_{5a} + \frac{1}{2}f_{3(6,5a)} \tag{21.129}$$

as linear combinations of L_1, L_6 and the ionosphere-free linear combination $L_{3(6,5a)}$. The frequencies of the new wide-lane and narrow-lane linear combinations are essentially equal to the frequencies of the wide-lane and narrow-lane linear combinations of the original L_1 and L_2 measurements

$$f_W^{(1,6,3)} = f_1 - f_6 = f_{W(1,6)}$$
$$f_N^{(1,6,3)} = f_1 + f_6 = f_{N(1,6)} \tag{21.130}$$

With the wide-lane and narrow-lane wavelengths defined as

$$\lambda_W^{(1,6,3)} := \frac{c}{f_1 + f_{5a} - 0.5f_{3(6,5a)}} = \frac{c}{f_1 - f_6} = \frac{c}{f_{W(1,6)}} = \frac{1}{29} \cdot \frac{c}{f_0} = \lambda_{W(1,6)} \approx 101.1 \text{ cm}$$
$$\lambda_N^{(1,6,3)} := \frac{c}{f_1 - f_{5a} + 0.5f_{3(6,5a)}} = \frac{c}{f_1 + f_6} = \frac{c}{f_{N(1,6)}} = \frac{1}{279} \cdot \frac{c}{f_0} = \lambda_{N(1,6)} \approx 10.5 \text{ cm} \tag{21.131}$$

the three-carrier wide-lane ambiguity is then

$$N_W^{(1,6,3)} = N_1 + N_{5a} - \frac{1}{2}N_{3(6,5a)} = N_1 + N_{5a} - \frac{1}{2}(25N_6 - 23N_{5a})$$
$$= N_1 + N_{5a} - N_6 - \frac{23}{2}(N_6 - N_{5a}) \tag{21.132}$$
$$= N_1 - \frac{25}{2}N_{W(6,5a)}$$

In its final form the three-carrier wide-lane ambiguity is defined as

$$2N_W^{(1,6,3)} := 2N_1 - 25N_{W(6,5a)} \tag{21.133}$$

In a similar way, we can derive the three-carrier narrow-lane ambiguity

$$N_N^{(1,6,3)} = N_1 - N_{5a} + \frac{1}{2}N_{3(6,5a)} = N_1 - N_{5a} + \frac{1}{2}(25N_6 - 23N_{5a})$$
$$= N_1 - N_{5a} + N_{6a} + \frac{23}{2}(N_6 - N_{5a}) \tag{21.134}$$
$$= N_1 + \frac{25}{2}N_{W(6,5a)}$$

In its final form, the three-carrier narrow-lane ambiguity is defined as

$$2N_N^{(1,6,3)} := 2N_1 + 25N_{W(6,5a)} \tag{21.135}$$

Eqs. (21.133) and (21.135) show that with integer ambiguity algebra it is possible to express the wide-lane and the narrow-lane linear combinations of the original L_1 and L_6 measurements as a linear combination of the wide-lane ambiguity between L_6 and L_{5a} measurements. However, the noise floor of the wide-lane linear combination $L_{W(6,5a)}$ needed to derive the super wide-lane ambiguity $N_{W(6,5a)}$ is higher by a factor of 17 than that of the original measurements L_1 on the reference Galileo frequency

$$\sigma(L_{W(6,5a)}) \approx 17 \cdot \sigma(L_1) \tag{21.136}$$

Therefore, in order to obtain the super wide-lane ambiguity $N_{W(6,5a)}$, we propose to make use of the vector form of the wide-lane ambiguity (21.32)

$$\vec{N}_{W(6,5a)} := \vec{N}_{W(1,5a)} - \vec{N}_{W(1,6)} \tag{21.137}$$

since the wide-lane ambiguities $N_{W(1,5a)}$ and $N_{W(1,6)}$ can be fixed to their integer values using the Melbourne-Wübbena linear combination.

Let us now derive a mathematical model of the three-carrier wide-lane and narrow-lane linear combinations and prove that the final form of the three-carrier linear combinations can be reduced to the wide-lane and narrow-lane linear combinations of two frequencies. By means of (21.129) the three-carrier wide-lane linear combination is

$$L_W^{(1,5a,3)}(L_1, L_{5a}, L_{N(6,5a)}) = \frac{f_1}{f_1 + f_{5a} - 0.5f_{3(6,5a)}} L_1 + \frac{f_{5a}}{f_1 + f_{5a} - 0.5f_{3(6,5a)}} L_{5a}$$
$$- \frac{0.5f_{3(6,5a)}}{f_1 + f_{5a} - 0.5f_{3(6,5a)}} L_{N(6,5a)} \tag{21.138}$$

and in its final form can be defined as

$$L_W^{(1,5a,3)}(L_1, L_{5a}, L_{N(6,5a)}) := \frac{f_1}{f_1 - f_6} L_1 + \frac{f_{5a}}{f_1 - f_6} L_{5a} - \frac{f_6 + f_{5a}}{f_1 - f_6} L_{N(6,5a)}$$
$$:= L_{W(1,6)} \tag{21.139}$$

In a similar way, by means of (21.151) and (21.153) we obtain the three-carrier narrow-lane linear combination

$$L_N^{(1,5a,3)}(L_1, L_{5a}, L_{N(6,5a)}) = \frac{f_1}{f_1 - f_{5a} + 0.5f_{3(6,5a)}} L_1 - \frac{f_{5a}}{f_1 - f_{5a} + 0.5f_{3(6,5a)}} L_{5a}$$
$$+ \frac{0.5f_{3(6,5a)}}{f_1 - f_{5a} + 0.5f_{3(6,5a)}} L_{N(6,5a)} \tag{21.140}$$

leading to the final form of the three-carrier narrow-lane linear combination

$$L_N^{(1,5a,3)}(L_1, L_{5a}, L_{N(6,5a)}) := \frac{f_1}{f_1 + f_6} L_1 - \frac{f_{5a}}{f_1 + f_6} L_{5a} + \frac{f_6 + f_{5a}}{f_1 + f_6} L_{N(6,5a)}$$
$$:= L_{N(1,6)} \tag{21.141}$$

As expected, the three-carrier wide-lane and narrow-lane linear combinations (21.139) and (21.141) have been reduced to the wide-lane and narrow-lane linear combination of the original L_1 and L_6 measurements. The mathematical model of the three-carrier wide-lane linear combination is then defined as

$$L_W^{(1,5a,3)}\left(L_1,L_6\right) := \rho + \lambda_{W(1,6)}N_1 - 25\frac{\lambda_{W(1,6)}}{2}N_{W(6,5a)} - \frac{f_1}{f_6}I_1 \tag{21.142}$$

and for the three-carrier narrow-lane linear combination

$$L_N^{(1,5a,3)}\left(L_1,L_6\right) := \rho + \lambda_{N(1,6)}N_1 + 25\frac{\lambda_{N(1,6)}}{2}N_{W(6,5a)} + \frac{f_1}{f_6}I_1 \tag{21.143}$$

The code version of the three-carrier narrow-lane linear combination (21.141) is then

$$P_N^{(1,5a,3)}\left(P_1,P_6\right) := \rho - \frac{f_1}{f_6}I_1 \tag{21.144}$$

21.10 Exotic Three-Carrier Wide-Lane and Narrow-Lane Combinations

In Section 21.5, we developed ambiguity integer algebra that can easily be extended to any GNSS frequency and observable. Let us now form the ionosphere-free $L_{3(2,5)}$, wide-lane $L_{W(2,5)}$ and narrow-lane $L_{N(2,5)}$ linear combinations of L_2 and L_5 phase measurements

$$\begin{aligned}
L_{3(2,5)} &= \rho + \lambda_{3(2,5)}N_{3(2,5)} \\
L_{W(2,5)} &= \rho + \lambda_{W(2,5)}N_{W(2,5)} - \frac{f_1^2}{f_2f_5}I_1 \\
L_{N(2,5)} &= \rho + \lambda_{N(2,5)}N_{N(2,5)} + \frac{f_1^2}{f_2f_5}I_1
\end{aligned} \tag{21.145}$$

with the ionosphere-free ambiguity $N_{3(2,5)}$, defined by the integer equation (21.117) in Section 21.5

$$N_2 + 23N_{W(2,5)} = N_{3(2,5)} \quad \rightarrow \quad \lambda_{3(2,5)} = \frac{1}{235}\cdot\frac{c}{f_0} = \frac{c}{f_{3(2,5)}} \approx 12.5 \text{ cm} \tag{21.146}$$

The corresponding ionosphere-free frequency $f_{3(2,5)}$ is then equal to the narrow-lane frequency $f_{N(2,5)}$

$$f_{3(2,5)} := 235\cdot f_0 = f_2 + f_5 = f_{N(2,5)} \quad \rightarrow \quad \lambda_{3(2,5)} = \lambda_{N(2,5)} \tag{21.147}$$

From (21.147) we can see that in the case of L_2 and L_5 phase measurements the ionosphere-free and the narrow-lane linear combinations have exactly the same frequency and wavelength. Following (21.117), the ionosphere-free integer ambiguity can be defined as

$$N_{3(2,5)} := 24N_2 - 23N_5 \tag{21.148}$$

Following (21.5), the general form of the linear combination $L_{LC}^{(1,2,5)}$ of three carrier-phase observables is

$$L_{LC}^{(1,2,5)}(L_1,L_2,L_5) := \alpha_1\frac{f_1}{f_{LC}}L_1 + \alpha_2\frac{f_2}{f_{LC}}L_2 + \alpha_5\frac{f_5}{f_{LC}}L_5 \tag{21.149}$$

with the frequency of the linear combination f_{LC} defined as

$$f_{LC}^{(1,2,5)} := \alpha_1 f_1 + \alpha_2 f_2 + \alpha_5 f_5 \qquad \alpha_i \in R \tag{21.150}$$

Let us now define the frequency of the three-carrier wide-lane and narrow-lane linear combination in the following way

$$f_W^{(1,5,3)} := f_1 + f_5 - f_{3(2,5)}$$
$$f_N^{(1,5,3)} := f_1 - f_5 + f_{3(2,5)}$$

(21.151)

i.e., as a linear combination of L_1, L_5 and the ionosphere-free linear combination $L_{3(2,5)}$. The frequency of the new wide-lane and narrow-lane linear combinations is essentially equal to the frequency of the wide-lane and narrow-lane linear combinations of the original L_1 and L_2 measurements

$$f_W^{(1,5,3)} = f_1 - f_2 = f_W \cdot f_0 = 34 \cdot f_0$$
$$f_N^{(1,5,3)} = f_1 + f_2 = f_N \cdot f_0 = 274 \cdot f_0$$

(21.152)

With fundamental GPS frequency $f_0 = 10.23$ MHz, the wide-lane and narrow-lane wavelengths can be defined as

$$\lambda_W^{(1,5,3)} := \frac{c}{f_1 + f_5 - f_{3(2,5)}} = \frac{c}{f_1 - f_2} = \frac{c}{f_W \cdot f_0} = \frac{1}{34} \cdot \frac{c}{f_0} = \lambda_W \approx 86.2 \text{ cm}$$
$$\lambda_N^{(1,5,3)} := \frac{c}{f_1 - f_5 + f_{3(2,5)}} = \frac{c}{f_1 + f_2} = \frac{c}{f_N \cdot f_0} = \frac{1}{274} \cdot \frac{c}{f_0} = \lambda_N \approx 10.7 \text{ cm}$$

(21.153)

By means of (21.117), the three-carrier wide-lane ambiguity is then

$$\begin{aligned} N_W^{(1,5,3)} &= N_1 + N_5 - N_{3(2,5)} = N_1 + N_5 - (24N_2 - 23N_5) \\ &= N_1 + N_5 - N_2 - 23(N_2 - N_5) \\ &= N_5 + N_W - 23N_{W(2,5)} \\ &= N_2 + N_W - 24N_{W(2,5)} \\ &= N_1 - 24N_{W(2,5)} \end{aligned}$$

(21.154)

In its final form, the three-carrier wide-lane ambiguity is defined as

$$N_W^{(1,5,3)} := N_1 - 24N_{W(2,5)}$$

(21.155)

or by adding $N_2 - N_2$ to (21.155)

$$N_W^{(1,5,3)} := N_2 + N_W - 24N_{W(2,5)}$$

(21.156)

In a similar way, we can derive the three-carrier narrow-lane ambiguity

$$\begin{aligned} N_N^{(1,5,3)} &= N_1 - N_5 + N_{3(2,5)} = N_1 - N_5 + (24N_2 - 23N_5) \\ &= -N_5 + N_1 + N_2 + 23(N_2 - N_5) \\ &= -N_5 + N_N + 23N_{W(2,5)} \end{aligned}$$

(21.157)

In its final form, the three-carrier narrow-lane ambiguity is defined as

$$N_N^{(1,5,3)} := N_1 + 24N_{W(2,5)}$$

(21.158)

Eqs. (21.155) and (21.158) show that with integer ambiguity algebra it is possible to express the wide-lane and narrow-lane linear combinations of the original L_1 and L_2 measurements as a linear combination of the wide-lane ambiguity between L_1 and L_5 measurements. However, the noise floor of the wide-lane linear combination $L_{W(2,5)}$ is higher by a factor of about 33 than that of the original measurements L_1 on the first GPS frequency

$$\sigma(L_{W(2,5)}) \approx 33 \cdot \sigma(L_1) \tag{21.159}$$

Therefore, in order to obtain the super wide-lane ambiguity $N_{W(2,5)}$, we propose to make use of the vector form of the wide-lane ambiguity (21.32)

$$\vec{N}_{W(2,5)} := \vec{N}_{W(1,5)} - \vec{N}_W \tag{21.160}$$

since the wide-lane ambiguities $N_{W(1,5)}$ and N_W can be fixed to their integer values using Melbourne-Wübbena linear combination.

Let us now derive a mathematical model of the three-carrier wide-lane and narrow-lane linear combinations defined by (21.151) and prove that the final form of the three-carrier linear combinations is reduced to the wide-lane and narrow-lane linear combinations of two frequencies. The three-carrier wide-lane linear combination is then

$$L_W^{(1,5,3)}(L_1, L_5, L_{3(2,5)}) = \frac{f_1}{f_1 + f_5 - f_{3(2,5)}} L_1 + \frac{f_5}{f_1 + f_5 - f_{3(2,5)}} L_5 -$$
$$- \frac{f_{3(2,5)}}{f_1 + f_5 - f_{3(2,5)}} \left(L_{N(2,5)} + \Delta_{N(2,5)} \right) \tag{21.161}$$

with $\Delta_{N(2,5)}$ denoting

$$\Delta_{N(2,5)} = -\lambda_{3(2,5)} N_5 + 23\lambda_{3(2,5)} N_{W(2,5)} - \frac{f_1^2}{f_2 f_5} I_1 \tag{21.162}$$

since from (21.145) and (21.146) we have

$$L_{3(2,5)} = L_{N(2,5)} - \lambda_{3(2,5)} N_5 + 23\lambda_{3(2,5)} N_{W(2,5)} - \frac{f_1^2}{f_2 f_5} I_1$$
$$= L_{N(2,5)} + \Delta_{N(2,5)} \tag{21.163}$$

In its final form, (21.161) can be defined as

$$L_W^{(1,5,3)}(L_1, L_5, L_{3(2,5)}) := \frac{f_1}{f_1 - f_2} L_1 + \frac{f_5}{f_1 - f_2} L_5 - \frac{f_2 + f_5}{f_1 - f_2} \left(L_{N(2,5)} + \Delta_{N(2,5)} \right)$$
$$:= L_W - \frac{f_2 + f_5}{f_1 - f_2} \Delta_{N(2,5)} \tag{21.164}$$

that reduces to

$$L_W^{(1,5,3)}(L_1, L_5, L_{3(2,5)}) = \rho + \lambda_W (N_1 - 24 N_{W(2,5)}) + \frac{f_1 + f_5}{f_1 - f_2} \frac{f_1}{f_5} I_1 \tag{21.165}$$

and after substitution of (21.155), we finally obtain the (super) ionosphere linear combination or the three-carrier wide-lane linear combination

$$L_W^{(1,5,3)}(L_1, L_5, L_{3(2,5)}) := \rho + \lambda_W N_W^{(1,5,3)} + \frac{f_1 + f_5}{f_1 - f_2} \frac{f_1}{f_5} I_1 \tag{21.166}$$

Since in (21.166) the first order ionosphere effect is multiplied by the very large factor

$$I(L_W^{(1,5,3)}) = \frac{f_1 + f_5}{f_1 - f_2} \frac{f_1}{f_5} I_1 \approx 10.6 \cdot I_1 \tag{21.167}$$

it can be used to derive the first order ionosphere effect. It can be shown that the multiplication coefficients in (21.161) follow the following properties

$$\frac{f_1}{f_1 + f_5 - f_{3(2,5)}} = \frac{f_1}{f_1 - f_2}$$

$$\frac{f_5}{f_1 + f_5 - f_{3(2,5)}} = \frac{f_5}{f_1 - f_2} \tag{21.168}$$

$$-\frac{f_{3(2,5)}}{f_1 + f_5 - f_{3(2,5)}} = -\frac{f_{3(2,5)}}{f_1 - f_2}$$

and so in its final form we obtain the three-carrier wide-lane linear combination

$$L_W^{(1,5,3)}(L_1, L_5, L_{3(2,5)}) := \frac{f_1}{f_1 - f_2} L_1 + \frac{f_5}{f_1 - f_2} L_5 - \frac{f_{3(2,5)}}{f_1 - f_2} L_{3(2,5)} \tag{21.169}$$

In a similar way, by means of (21.151) and (21.153) we obtain the three-carrier narrow-lane linear combination

$$L_N^{(1,5,3)}(L_1, L_5, L_{3(2,5)}) = \frac{f_1}{f_1 - f_5 + f_{3(2,5)}} L_1 - \frac{f_5}{f_1 - f_5 + f_{3(2,5)}} L_5 +$$

$$+ \frac{f_{3(2,5)}}{f_1 - f_5 + f_{3(2,5)}} \left(L_{N(2,5)} + \Delta_{N(2,5)} \right) \tag{21.170}$$

leading to the final form of the three-carrier narrow-lane linear combination

$$L_N^{(1,5,3)}(L_1, L_5, L_{N(2,5)}) := \frac{f_1}{f_1 + f_2} L_1 - \frac{f_5}{f_1 + f_2} L_5 + \frac{f_2 + f_5}{f_1 + f_2} \left(L_{N(2,5)} + \Delta_{N(2,5)} \right)$$

$$:= L_N + \frac{f_2 + f_5}{f_1 + f_2} \Delta_{N(2,5)} \tag{21.171}$$

that reduces to

$$L_N^{(1,5,3)}(L_1, L_5, L_{3(2,5)}) := \rho + \lambda_N (N_1 + 24 N_{W(2,5)}) - \frac{f_1 - f_5}{f_1 + f_2} \frac{f_1}{f_5} I_1 \tag{21.172}$$

and after substitution of (21.158) we finally obtain

$$L_N^{(1,5,3)}(L_1, L_5, L_{3(2,5)}) := \rho + \lambda_N N_{3(2,5)} - \frac{f_1 - f_5}{f_1 + f_2} \frac{f_1}{f_5} I_1 \tag{21.173}$$

or

$$L_N^{(1,5,3)}(L_1, L_5, L_{3(2,5)}) := \rho + \lambda_N N_N^{(1,5,3)} - \frac{f_1 - f_5}{f_1 + f_2} \frac{f_1}{f_5} I_1 \tag{21.174}$$

Note that in (21.173) the first order ionosphere effect is multiplied by the very small factor

$$I(L_N^{(1,5,3)}) = -\frac{f_1 - f_5}{f_1 + f_2} \frac{f_1}{f_5} I_1 \approx -0.19 \cdot I_1 \tag{21.175}$$

As expected, the three-carrier wide-lane and narrow-lane linear combinations (21.164) and (21.171) have been reduced to the wide-lane and narrow-lane type linear combination of the original L_1 and L_2 measurements.

It can be shown that multiplication coefficients in (21.170) follow the following properties

$$\frac{f_1}{f_1 - f_5 + f_{3(2,5)}} = \frac{f_1}{f_1 + f_2}$$
$$-\frac{f_5}{f_1 - f_5 + f_{3(2,5)}} = -\frac{f_5}{f_1 + f_2} \qquad (21.176)$$
$$\frac{f_{3(2,5)}}{f_1 - f_5 + f_{3(2,5)}} = \frac{f_{3(2,5)}}{f_1 + f_2}$$

and in its final form we derive the three-carrier narrow-lane linear combination

$$L_N^{(1,5,3)}(L_1, L_5, L_{3(2,5)}) := \frac{f_1}{f_1 + f_2} L_1 - \frac{f_5}{f_1 + f_2} L_5 + \frac{f_{3(2,5)}}{f_1 + f_2} L_{3(2,5)} \qquad (21.177)$$

with the following property

$$\frac{f_1}{f_1 + f_2} \lambda_1 - \frac{f_5}{f_1 + f_2} \lambda_5 = 0 \qquad (21.178)$$

Thus, only the third multiplication factor in (21.177) effectively contributes to ambiguity resolution when the original zero-difference ambiguities are aligned using wide-lane ambiguities.

The mathematical model of the three-carrier wide-lane linear combination we will use for ambiguity resolution is defined as

$$L_W^{(1,2,5)}(L_1, L_2) := \rho + \lambda_W N_1 - 24\lambda_W N_{W(2,5)} + \frac{f_1 + f_5}{f_1 - f_2} \frac{f_1}{f_5} I_1 \qquad (21.179)$$

and for the three-carrier narrow-lane linear combination

$$L_N^{(1,2,5)}(L_1, L_2) := \rho + \lambda_N N_1 + 24\lambda_N N_{W(2,5)} - \frac{f_1 - f_5}{f_1 + f_2} \frac{f_1}{f_5} I_1 \qquad (21.180)$$

The code version of the three-carrier wide-lane linear combination (21.179) is then

$$P_W^{(1,2,5)} := P_W - \frac{f_2 + f_5}{f_1 - f_2} \frac{f_1^2}{f_2 f_5} I_1 \qquad (21.181)$$

and in its final form

$$P_W^{(1,2,5)} := \rho - \frac{f_1 + f_5}{f_1 - f_2} \frac{f_1}{f_5} I_1 \qquad (21.182)$$

The code version of the three-carrier narrow-lane linear combination (21.180) is then

$$P_N^{(1,2,5)} := P_N + \frac{f_2 + f_5}{f_1 + f_2} \frac{f_1^2}{f_2 f_5} I_1 \qquad (21.183)$$

and in its final form

$$P_N^{(1,2,5)} := \rho + \frac{f_1 - f_5}{f_1 + f_2} \frac{f_1}{f_5} I_1 \qquad (21.184)$$

21.11 Three-Carrier Type Melbourne-Wübenna Linear Combination

By subtracting the three-carrier wide-lane linear combination of phase measurements (21.179) from the three-carrier narrow-lane linear combination of code measurements (21.183) we derive the three-carrier Melbourne-Wübbena linear combination

$$MW^{(1,2,3)}(L_1, L_5, L_{3(2,5)}, P_1, P_5, P_{3(2,5)}) := L_W^{(1,2,3)}(L_1, L_5, L_{3(2,5)}) - P_3^{(1,2,5)}(P_1, P_5, P_{3(2,5)}) \tag{21.185}$$

that can be reduced to

$$MW^{(1,2,5)}(L_1, L_5, L_{N(2,5)}, P_1, P_5, P_{N(2,5)}) := MW = L_W - P_N = \lambda_W N_W \tag{21.186}$$

with Melbourne-Wübbena linear combination MW. The mathematical model of the three-carrier Melbourne-Wübbena linear combination is defined as

$$MW^{(1,2,5)} := \lambda_W N_1 - 24\lambda_W N_{W(2,5)} \tag{21.187}$$

with an acceptable noise floor, mainly driven by the code measurements

$$\sigma(MW^{(1,2)}) = \sigma(MW^{(1,2,5)}) \approx \sigma(MW^{(1,2,5)}) \approx 0.71 \cdot \sigma(P_1) \tag{21.188}$$

From (21.187) we see that the three-carrier Melbourne-Wübbena linear combination can be used to fix the reference ambiguity N_1 to an integer value. However, the reference ambiguity N_1 can also be estimated using L_1 and L_5 phase measurements, i.e., using the following Melbourne-Wübbena linear combination

$$MW^{(1,5)} := \lambda_{W(1,5)}\left(N_W + N_{W(2,5)}\right) = \lambda_{W(1,5)}N_1 - 23\lambda_{W(1,5)}N_{W(2,5)} \tag{21.189}$$

By adding (21.187) and (21.189) we obtain the following observation equation for wide-lane ambiguities

$$MW^{(1,5)} + MW = (\lambda_W + \lambda_{W(1,5)})N_1 - (24\lambda_W + 23\lambda_{W(1,5)})N_{W(2,5)} \tag{21.190}$$

Although the noise level of (21.190) is increased by a factor of about $\sqrt{2}$ in this way, the wavelength of the reference ambiguity $\lambda_{W(1,2,5)}$ is approximately doubled in size and defined as

$$\lambda_{W(1,2,5)} := \lambda_W + \lambda_{W(1,5)} = \frac{2f_1 - f_2 - f_5}{(f_1 - f_2)(f_1 - f_5)}c \approx 0.86 + 0.75 \approx 1.61 \text{ m} \tag{21.191}$$

22. Earth Orientation Quaternion

I n (Švehla 2006), it was proposed for the first time to represent Earth orientation and rotation by means of an Earth Orientation Quaternion (EOQ). Quaternions are a very practical way to represent the Earth's orientation parameters (EOPs), because the transformation between the terrestrial and the inertial system can be performed without calculating rotation matrices. Most importantly, the use of EOPs stored in the form of a quaternion avoids the use of the latest models and standards available from the IERS Conventions, as in the case of the EOP/ERP parameters provided by IGS and IERS. In this way, information about the Earth's rotation/orientation is straightforward and the transformation can be performed much in the same way as for satellite attitude. This idea that was originally presented in (Švehla 2006), was included in the recommendations of the Workshop on Precise Orbit Determination for the future ESA Earth observation missions, held at ESTEC/ESA in 2007 (Švehla 2007c). Following this recommendation, the ESA Core Mission GOCE provides Earth Orientation Quaternions as a separate product accompanying the kinematic and reduced-dynamic orbit. The sampling rate of Earth Orientation Quaternions, as provided in the scope of the GOCE mission.

The four Euler symmetric parameters written in the form of a quaternion are a minimal set of parameters for defining non-singular mapping to the corresponding rotation matrix. Besides their symmetrical properties, modeling finite rotations using quaternions has many advantages compared to using Euler angles since any interpolation or integration can be performed on the sphere, preserving the orthonormality of the rotation transformation (Švehla 2006).

Hamilton or quaternion algebra avoids the use of a rotation matrix and any sequence of successive rotations can be represented very elegantly by the quaternion operator. This also holds for the derivatives of the successive rotations and the treatment of the kinematic equation of rotation. We show how to interpolate and extrapolate the Earth orientation quaternions preserving the orthonormality of the transformation. We introduce a transition quaternion derived from the kinematic equation of rotation.

In the field of numerical solutions of ordinary differential equations, geometric integration is defined as a numerical method that preserves the geometric properties of the exact flow of a differential equation. Therefore, when talking about integrating quaternions on the sphere and preserving orthonormality of the rotation transformation at the same time, we are actually talking about using geometric integration.

22.1 Kinematic Equation of Earth's Rotation in Terms of Quaternions

The kinematic rotation of a planet such as the Earth, or a satellite in the orbital plane, can be defined as a rotation irrespective of the forces that govern that rotation. Kinematic rotation describes rotation of a body, e.g., (Operation and Wertz 1978), and can be given by a set of first-order differential equations specifying the time evolution of the rotation parameters. Modeling rotation is, in essence, modeling an instantaneous angular velocity vector. Space geodesy techniques, such as VLBI, measure the geometric rotation and orientation of the Earth. Much in the same way, star trackers placed on a satellite take images of stars to provide orientation.

The Earth Orientation Quaternion defines a rotation between a terrestrial reference frame, such as ITRF, and the inertial, quasi-inertial, true system of date, or a celestial reference frame, here donoted as ICRF. The quaternion q is defined in terms of Euler symmetric parameters $\{q_1, q_2, q_3, q_0\}$ defined as e.g., (Hamilton 1853)

$$
\begin{aligned}
q &= q_0 + iq_1 + jq_2 + kq_3 \\
q_1 &= e_1 \sin\frac{\Phi}{2} \\
q_2 &= e_2 \sin\frac{\Phi}{2} \\
q_3 &= e_3 \sin\frac{\Phi}{2} \\
q_0 &= \cos\frac{\Phi}{2}
\end{aligned}
\tag{22.1}
$$

where $\{e_1, e_2, e_3\}$ are the components of the Euler axis and Φ the corresponding rotation angle. The quantity q_0 is the real or scalar part of the quaternion and $iq_1 + jq_2 + kq_3$ is the imaginary or vector part. $\{i, j, k\}$ are the hyperimaginary numbers satisfying the conditions

$$
\begin{aligned}
i^2 &= j^2 = k^2 = -1 \\
ij &= -ji = k \\
jk &= -kj = i \\
ki &= -ik = j
\end{aligned}
\tag{22.2}
$$

For more detail on the definition of quaternions and geometric algebra see the original paper (Hamilton 1853). When working with quaternions available from satellite missions, one needs to take into account the scalar term q_0 in (22.1) that can be provided either as the first or the last element.

The relationship between quaternions and the rotation matrix can be derived from the so-called "Euler/axis-angle" representation of the rotation. Following (Operation and Wertz 1978), the direction cosine matrix A is in this case given by

$$
A = \begin{vmatrix}
\cos\Phi + e_1^2(1 - \cos\Phi) & e_1 e_2(1 - \cos\Phi) + e_3\sin\Phi & e_1 e_3(1 - \cos\Phi) - e_2\sin\Phi \\
e_1 e_2(1 - \cos\Phi) - e_3\sin\Phi & \cos\Phi + e_2^2(1 - \cos\Phi) & e_2 e_3(1 - \cos\Phi) + e_1\sin\Phi \\
e_1 e_3(1 - \cos\Phi) + e_2\sin\Phi & e_2 e_3(1 - \cos\Phi) - e_1\sin\Phi & \cos\Phi + e_3^2(1 - \cos\Phi)
\end{vmatrix}
\tag{22.3}
$$

From there, the direction cosine matrix expressed in terms of the Euler symmetric parameters, or in our case the rotation matrix R from the terrestrial into the inertial reference frame, is defined as

$$
R = \begin{bmatrix}
q_1^2 - q_2^2 - q_3^2 + q_0^2 & 2(q_1 q_2 + q_3 q_0) & 2(q_1 q_3 - q_2 q_0) \\
2(q_1 q_2 - q_3 q_0) & -q_1^2 + q_2^2 - q_3^2 + q_0^2 & 2(q_2 q_3 + q_1 q_0) \\
2(q_1 q_3 + q_2 q_0) & 2(q_2 q_3 - q_1 q_0) & -q_1^2 - q_2^2 + q_3^2 + q_0^2
\end{bmatrix}
\tag{22.4}
$$

For a position vector X_{ITRF} and a velocity vector $\dot{X}_{\mathrm{ITRF}}$ given in the Earth-fixed reference frame, the transformation into the quasi-inertial reference frame can be calculated as follows

$$
\begin{aligned}
X_{\mathrm{ICRF}} &= R\ X_{\mathrm{ITRF}} \\
\dot{X}_{\mathrm{ICRF}} &= R\ \dot{X}_{\mathrm{ITRF}} + \dot{R}\ X_{\mathrm{ITRF}}
\end{aligned}
\tag{22.5}
$$

Expressing the first derivative of the rotation matrix by means of the skew-symmetric matrix Ω we have

$$
\dot{X}_{\mathrm{ICRF}} = R\ \dot{X}_{\mathrm{ITRF}} + \Omega_{3\times3} R\ X_{\mathrm{ITRF}}
\tag{22.6}
$$

where the skew-symmetric matrix $\Omega_{3\times3}$ can be defined by means of the angular velocity vector $\vec{\omega} = \{\omega_1, \omega_2, \omega_3\}$ as follows

$$\Omega_{3\times3} := \begin{bmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{bmatrix} \approx \begin{bmatrix} 0 & \omega_3 & 0 \\ -\omega_3 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\omega_3 = -7292115.1567 \cdot 10^{-11} \text{ rad/s}$$

(22.7)

and can, in some cases, be approximated with sufficient accuracy by a rotation about only one axis. Using quaternions, the calculation of the rotation matrix can be avoided and the rotation can be replaced by the quaternion multiplication with (4×1) vectors X_{ITRF} and X_{ICRF} (the fist value is zero)

$$X_{\text{ICRF}} = q^* \cdot X_{\text{ITRF}} \cdot q$$

(22.8)

where q^* denotes the conjugate or inverse quaternion q^* defined as

$$q^* := q_0 - iq_1 - jq_2 - kq_3$$

(22.9)

The multiplication of two quaternions, q and q', can be written as

$$q'' = q' \cdot q$$

$$\begin{bmatrix} q_0'' \\ q_1'' \\ q_2'' \\ q_3'' \end{bmatrix} = \begin{bmatrix} q_0' & -q_1' & -q_2' & -q_3' \\ q_1' & q_0' & -q_3' & q_2' \\ q_2' & q_3' & q_0' & -q_1' \\ q_3' & -q_2' & q_1' & q_0' \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}$$

(22.10)

For the GOCE mission, the Earth Orientation Quaternions are provided for every integer second t_0 of GPS time (terrestrial time). To obtain quaternion information for the actual epoch time t_{epo}, the kinematic equation of rotation may be used to propagate quaternion information between the two nearest integer seconds $q_{4\times1}(t_0)$ and $q_{4\times1}(t_1)$. Following (Operation and Wertz 1978), the time derivative of a quaternion reads as

$$q(t + \Delta t) \approx \left[I + \frac{1}{2}\Omega\Delta t \right] q(t)$$

(22.11)

where I denotes the (4×4) identity matrix and Ω is the skew-symmetric (4×4) matrix defined as

$$\Omega := \begin{bmatrix} 0 & \omega_3 & -\omega_2 & \omega_1 \\ -\omega_3 & 0 & \omega_1 & \omega_2 \\ \omega_2 & -\omega_1 & 0 & \omega_3 \\ -\omega_1 & -\omega_2 & -\omega_3 & 0 \end{bmatrix}$$

(22.12)

Finally, the first time derivative of a quaternion, or kinematic equation of rotation reads as

$$\frac{dq}{dt} = \lim_{\Delta t \to 0} \frac{q(t + \Delta t) - q(t)}{\Delta t} = \frac{1}{2}\Omega q$$

(22.13)

Assuming the angular vector to be constant between two epochs, by integrating (22.13), we can obtain a closed solution for the kinematic equation of rotation, see e.g., (Operation and Wertz 1978)

$$q(t) = e^{\frac{\Omega t}{2}} \cdot q(t_0)$$

(22.14)

22.2 Transition Quaternion

Since any sequence of successive rotations can be represented very elegantly by the quaternion multiplication operator as given in (22.10), we introduce and define the transition quaternion q_s in the following way

$$q(t_1) = q_s q(t_0)$$

$$q_s := q(t_1) \frac{q^*(t_0)}{\left\| q(t_0) \right\|^2} = e^{\frac{1}{2} \Omega t} \tag{22.15}$$

where q^* is the conjugate or inverse quaternion defined in (22.9) with the norm of a quaternion given as

$$\left\| q \right\| = \sqrt{q q^*} \tag{22.16}$$

Eq. (22.15) allows the calculation of a transition quaternion between two consecutive epochs. Let us now see how the transition quaternion can be calculated for an intermediate epoch. Using the expansion of the exponential function as given in (Bronstein and Semendjajew 1996) we obtain

$$e^{\frac{\Omega t}{2}} = \sum_{n=0}^{\infty} \frac{\left(\dfrac{\Omega t}{2}\right)^{2n}}{n!} \tag{22.17}$$

which can be written as

$$e^{\frac{\Omega t}{2}} = \sum_{n=0}^{\infty} \left[\frac{\left(\dfrac{\Omega t}{2}\right)^{2n}}{(2n)!} + \frac{\left(\dfrac{\Omega t}{2}\right)^{2n+1}}{(2n+1)!} \right]. \tag{22.18}$$

And since

$$\Omega^{2n} = (-1)^n \omega^{2n} I_{4\times 4} \tag{22.19}$$

with the identity matrix $I_{4\times 4}$, we obtain

$$e^{\frac{\Omega t}{2}} = I_{4\times 4} \sum_{n=0}^{\infty} \frac{(-1)^n \left(\dfrac{1}{2}\omega t\right)^{2n}}{(2n)!} + \Omega \omega^{-1} \sum_{n=0}^{\infty} \frac{(-1)^n \left(\dfrac{1}{2}\omega t\right)^{2n+1}}{(2n+1)!} \tag{22.20}$$

$$= I_{4\times 4} \cos\left(\frac{1}{2}\omega t\right) + \Omega \omega^{-1} \sin\left(\frac{1}{2}\omega t\right)$$

with the rotation rate ω, and finally

$$e^{\frac{\Omega t}{2}} = \begin{vmatrix} \cos\left(\dfrac{1}{2}\omega t\right) & \dfrac{\omega_3}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) & -\dfrac{\omega_2}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) & \dfrac{\omega_1}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) \\[2mm] -\dfrac{\omega_3}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) & \cos\left(\dfrac{1}{2}\omega t\right) & \dfrac{\omega_1}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) & \dfrac{\omega_2}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) \\[2mm] \dfrac{\omega_2}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) & -\dfrac{\omega_1}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) & \cos\left(\dfrac{1}{2}\omega t\right) & \dfrac{\omega_3}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) \\[2mm] -\dfrac{\omega_1}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) & -\dfrac{\omega_2}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) & -\dfrac{\omega_3}{\omega}\sin\left(\dfrac{1}{2}\omega t\right) & \cos\left(\dfrac{1}{2}\omega t\right) \end{vmatrix} \tag{22.21}$$

In the case of an extrapolation or interpolation, e.g., between two consecutive epochs, the transition quaternion q_s at epoch t_{epo} can be obtained in the following way

$$q_s(t_{epo}) = S_{4\times4}q_{4\times1}(t_0)$$

$$\Delta t = t_{epo} - t_0$$

$$S_{4\times4} = \begin{bmatrix} \cos\left(\frac{1}{2}\omega\Delta t\right) & \frac{\omega_3}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) & -\frac{\omega_2}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) & \frac{\omega_1}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) \\ -\frac{\omega_3}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) & \cos\left(\frac{1}{2}\omega\Delta t\right) & \frac{\omega_1}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) & \frac{\omega_2}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) \\ \frac{\omega_2}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) & -\frac{\omega_1}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) & \cos\left(\frac{1}{2}\omega\Delta t\right) & \frac{\omega_3}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) \\ -\frac{\omega_1}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) & -\frac{\omega_2}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) & -\frac{\omega_3}{\omega}\sin\left(\frac{1}{2}\omega\Delta t\right) & \cos\left(\frac{1}{2}\omega\Delta t\right) \end{bmatrix} \tag{22.22}$$

When the Earth Orientation Quaternions are provided with a sufficient sampling rate, (22.22) can be further approximated by

$$S_{4\times4} = \begin{bmatrix} \cos\left(\frac{1}{2}\omega\Delta t\right) & -\sin\left(\frac{1}{2}\omega\Delta t\right) & 0 & 0 \\ \sin\left(\frac{1}{2}\omega\Delta t\right) & \cos\left(\frac{1}{2}\omega\Delta t\right) & 0 & 0 \\ 0 & 0 & \cos\left(\frac{1}{2}\omega\Delta t\right) & -\sin\left(\frac{1}{2}\omega\Delta t\right) \\ 0 & 0 & \sin\left(\frac{1}{2}\omega\Delta t\right) & \cos\left(\frac{1}{2}\omega\Delta t\right) \end{bmatrix} \tag{22.23}$$

with the rotation rate $\omega = -7292115.1567 \cdot 10^{-11}$ [rad/s]. Before interpolation or extrapolation of the quaternions given at two epochs, one first needs to check if there is any ambiguity in the quaternion between consecutive epochs, i.e., that the rotation is carried out in the correct direction.

Once the Earth Orientation Quaternion is known for the given epoch, the position vector X_{ITRF} is first written in the form of a quaternion (the scalar part, or the first value is zero). Finally, transformation from the Earth-fixed reference frame (ITRF) into ICRF can be calculated as follows

$$\begin{bmatrix} 0 \\ X_{\text{ICRF}} \\ Y_{\text{ICRF}} \\ Z_{\text{ICRF}} \end{bmatrix} = \begin{bmatrix} q_0 & q_1 & q_2 & q_3 \\ -q_1 & q_0 & q_3 & -q_2 \\ -q_2 & -q_3 & q_0 & q_1 \\ -q_3 & q_2 & -q_1 & q_0 \end{bmatrix} \begin{bmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & q_3 & -q_2 \\ q_2 & -q_3 & q_0 & q_1 \\ q_3 & q_2 & -q_1 & q_0 \end{bmatrix} \begin{bmatrix} 0 \\ X_{\text{ITRF}} \\ Y_{\text{ITRF}} \\ Z_{\text{ITRF}} \end{bmatrix} \tag{22.24}$$

23. A Geometrical Approach to Model Circular Rotations

H ere we introduce an elegant way to geometrically model the rotation of a rigid body in vector form. Typically, to perform a rotation in Euclidian space $\Re^3$ one uses rotation matrices based on a given sequence of Euler angles. Another approach is to use quaternions. A matrix exponent is often used to describe rotations in mathematical expressions and derivations, i.e., the exponential map from so(3) to SO(3). However, the nine elements of the rotation matrix are still exclusively used for calculating rotations in Euclidian space. The axis/angle representation in terms of quaternions and Rodrigues' rotation formula are alternative approaches. However, hidden geometrical properties, or the complexity of using quaternion algebra are the stumbling blocks that lead to the situation that rotation matrices are still almost exclusively used nowadays. Here we introduce the spherical orthodrome rotation that describes a rotation purely geometrically in a highly transparent way as an orthodrome, or a great arc on a sphere. The application of such transparent geometrical rotations in vector form has many advantages compared to any other rotation. Here we introduce spherical rotation and show basic geometrical properties, i.e., the use of vector algebra to very efficiently perform rotation of a vector in Euclidian space or to describe any orientation. Thus, this approach could be used to model Earth orientation and rotation as well as the attitude of a satellite. We also show that this geometrical rotation approach could be used in orbit modeling, since orbit perturbations can be represented by circular rotations with an axis of rotation very close to the main axis of the satellite orbit.

23.1 Vector Rotations: Spherical Rotation

Spherical rotation on the sphere, as introduced here, is based on the equation of a great circle on a sphere, called the orthodrome. More on the equations of orthodrome and loxodrome on a sphere, given in a very elegant orthogonal vector form, can be found in (Švehla 1995) and (Švehla 1996), two student theses (not *Diplom*). The first was awarded the Rector's Prize in 1995 and the second the same prize in 1996. In (Švehla 1995), the following equation of the orthodrome was elaborated in the light of differential geometry and various projections on a sphere based on two orthogonal vectors $\vec{c}_1$ and $\vec{c}_2$

$$\vec{r}(\alpha) = \vec{c}_1 \cos\alpha + \vec{c}_2 \sin\alpha \qquad \vec{c}_1 \perp \vec{c}_2, \quad \left|\vec{c}_1\right| = \left|\vec{c}_2\right| = \left|\vec{r}(\alpha)\right| \tag{23.1}$$

Thus, to describe a great arc on a sphere we need an orthogonal basis $\left\{\vec{c}_1, \vec{c}_2\right\}$ and an angle α. The normal to the orthodrome is then given by $\vec{\omega} = \vec{c}_1 \times \vec{c}_2$. Representation (23.1) can be extended and used to describe the rotation of any vector $\vec{r}$ around an axis of rotation $\vec{\omega}$ and the rotation rate $\omega = \left|\vec{\omega}\right|$, along the given arc of the orthodrome on the sphere.
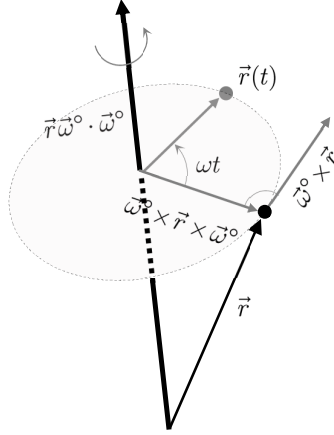
Figure 23.1 Spherical rotation - an elegant method of vector rotation, avoiding the use of rotation matrices, quaternions, etc.

We first consider a sphere of radius r, with the fixed rotation axis $\vec{\omega}$ and the rotation angle ωt defined by $\omega = |\vec{\omega}|$. In the second step we consider a plane defined by the normal that is collinear with the rotation axis $\vec{\omega}$ and intersects the sphere in a circle that describes the rotation of the vector $\vec{r}(t)$, see Figure 23.1. A rotation of the vector $\vec{r}(t)$ is then described uniquely by the following orthogonal basis

$$\left\{ \vec{\omega}^{\circ}, \quad \vec{\omega}^{\circ} \times \vec{r} \times \vec{\omega}^{\circ}, \quad \vec{\omega}^{\circ} \times \vec{r} \right\} \tag{23.2}$$

where unit vector $\vec{\omega}^{\circ}$ points along the rotation axis $\vec{\omega}$, $\vec{\omega}^{\circ} \times \vec{r} \times \vec{\omega}^{\circ}$ defines the direction in the meridian towards the vector $\vec{r}$ and $\vec{\omega}^{\circ} \times \vec{r}$ is collinear with the normal to the meridional plane defined by the vector $\vec{r}$. The spherical rotation of the vector $\vec{r}$ around the rotation axis $\vec{\omega}$ and the angle of rotation ωt is then defined as

$$\vec{r}(t) := \vec{r}\vec{\omega}^{\circ} \cdot \vec{\omega}^{\circ} + \vec{\omega}^{\circ} \times \vec{r} \times \vec{\omega}^{\circ} \cdot \cos \omega t + \vec{\omega}^{\circ} \times \vec{r} \cdot \sin \omega t \tag{23.3}$$

After including the scalar product $\vec{r}\vec{\omega}^{\circ} = r \cos \alpha_{r\omega}$ that is constant for all rotation angles and

$$\vec{a} = \left[\vec{r}^{\circ} \times \vec{\omega}^{\circ} \cdot \cos \omega t + \vec{r}^{\circ} \cdot \sin \omega t \right] \tag{23.4}$$

we obtain the second form of the spherical rotation

$$\vec{r}(t) := r \left[\vec{\omega}^{\circ} \cos \alpha_{r\omega} + \vec{\omega}^{\circ} \times \vec{a} \right] \tag{23.5}$$

The inverse rotation denoted here as $\vec{r}^{*}(t)$ is defined by the negative argument $-t$

$$\vec{r}^{*}(t) := \vec{r}(-t) = \vec{r}\vec{\omega}^{\circ} \cdot \vec{\omega}^{\circ} + \vec{\omega}^{\circ} \times \vec{r} \times \vec{\omega}^{\circ} \cdot \cos \omega t - \vec{\omega}^{\circ} \times \vec{r} \cdot \sin \omega t \tag{23.6}$$

thus we obtain the following property

$$\vec{r}(t) - \vec{r}^{*}(t) = 2 \cdot \vec{\omega}^{\circ} \times \vec{r} \cdot \sin \omega t \tag{23.7}$$

from where it follows

$$\vec{r}*(t) := \vec{r}(t) - 2 \cdot \vec{\omega}^{\circ} \times \vec{r} \cdot \sin \omega t \tag{23.8}$$

Keeping the radius constant, the first and second derivatives are

$$\frac{d\vec{r}(t)}{dt} = r\vec{\omega}^{\circ} \times \dot{\vec{a}}$$
$$= -\omega \cdot \vec{\omega}^{\circ} \times \vec{r} \times \vec{\omega}^{\circ} \cdot \sin\omega t + \omega \cdot \vec{\omega}^{\circ} \times \vec{r} \cdot \cos\omega t \qquad (23.9)$$
$$= \omega r \cdot \vec{\omega}^{\circ} \times \left[-\vec{r}^{\circ} \times \vec{\omega}^{\circ} \cdot \sin\omega t + \vec{r}^{\circ} \cdot \cos\omega t \right]$$

$$\frac{d^2\vec{r}(t)}{dt^2} = r\vec{\omega}^{\circ} \times \ddot{\vec{a}}$$
$$= -\omega^2 \cdot \vec{\omega}^{\circ} \times \vec{r} \times \vec{\omega}^{\circ} \cdot \cos\omega t - \omega^2 \cdot \vec{\omega}^{\circ} \times \vec{r} \cdot \sin\omega t \qquad (23.10)$$
$$= -\omega^2 r \cdot \vec{\omega}^{\circ} \times \left[\vec{r}^{\circ} \times \vec{\omega}^{\circ} \cdot \cos\omega t + \vec{r}^{\circ} \cdot \sin\omega t \right]$$
$$= -\omega^2 r \cdot \vec{\omega}^{\circ} \times \vec{a}$$

or

$$\frac{d^2\vec{r}(t)}{dt^2} = -\omega^2 \cdot \left(\vec{r}(t) - \vec{r}\vec{\omega}^{\circ} \cdot \vec{\omega}^{\circ} \right) \qquad (23.11)$$

The kinematic rotation of a planet such as the Earth, or a satellite, can be defined as a rotation irrespective of the forces that cause that rotation. Here the focus is on the model using uniform circular rotation. This kinematic rotation can be described by a set of first order differential equations specifying the time evolution of the rotation parameters.

23.2 Multipole Spherical Rotation

Let us now look at the case when the rotation axis $\vec{\omega}$ in is not fixed, but slowly rotating or precessing around a fixed axis $\vec{\omega}_1$. Generally speaking, we can add any number of additional frequencies and additional axes of rotation. For instance, in the case of Earth rotation, one could also add Chandler wobble, daily and annual terms, nutation due to tidal forces of the Moon and Sun, with the main period of 18.6 years, as well as precession. Thus, to add an additional rotation around an axis $\vec{\omega}_1^{\circ}$ by an angle $\omega_1 t$, we may write

$$\vec{\omega}^{\circ}(t) = \vec{\omega}_1^{\circ} \cdot \cos\alpha_{\omega\omega_1} + \vec{\omega}_1^{\circ} \times \left(\vec{\omega}^{\circ} \times \vec{\omega}_1^{\circ} \cdot \cos\omega_1 t + \vec{\omega}^{\circ} \cdot \sin\omega_1 t \right) \qquad (23.12)$$

or

$$\vec{\omega}^{\circ}(t) = \vec{\omega}_1^{\circ} \cos\alpha_{\omega\omega_1} + \vec{\omega}_1^{\circ} \times \vec{a}_{\omega_1} \qquad (23.13)$$

with

$$\cos\alpha_{\omega\omega_1} = \vec{\omega}^{\circ}\vec{\omega}_1^{\circ}$$
$$\vec{a}_{\omega_1} = \vec{\omega}^{\circ} \times \vec{\omega}_1^{\circ} \cdot \cos\omega_1 t + \vec{\omega}^{\circ} \cdot \sin\omega_1 t \qquad (23.14)$$

such a nested rotational spherical structure can be extended to any frequency argument $\vec{h}(t) := \vec{\omega}^{\circ} \cdot \cos\alpha_{h\omega} + \vec{\omega}^{\circ} \times \vec{a} \; n \cdot \omega_n t$ and rotation axis $\vec{\omega}_n^{\circ}$. The same model could be applied to the attitude of GNSS, or of LEO satellites, such as GOCE.

23.3 Transition Spherical Rotation

If two vectors are given on a sphere, the question is how to define the spherical rotation that directly connects them. This would be the same as the so-called second geodetic task on a sphere, i.e., given the positions of

two points on a unit sphere $\left\{\vec{r}_1^{\,\circ}(t_1), \vec{r}_2^{\,\circ}(t_2)\right\}$ we need to define the orthodrome or great arc between them. Following (23.1) we may define

$$\vec{c}_1 = \vec{r}_1^{\,\circ}(t_1) \tag{23.15}$$

from where it follows

$$\cos \sigma = \cos \omega(t_2 - t_1) = \vec{r}_1^{\,\circ}(t_1) \cdot \vec{r}_2^{\,\circ}(t_2)$$
$$\vec{c}_2 = \frac{\vec{r}_2^{\,\circ} - \vec{r}_1^{\,\circ} \cos \sigma}{\sin \sigma} \tag{23.16}$$

Finally, for the transition spherical rotation we obtain

$$\vec{r}^{\,\circ}(t) = \vec{c}_1(t_1) \cos \omega t + \vec{c}_2(t_1) \sin \omega t \tag{23.17}$$

with the first and second derivative

$$\frac{d\vec{r}^{\,\circ}(t)}{dt} = \omega\left[-\vec{c}_1(t_1)\sin \omega t + \vec{c}_2(t_1)\cos \omega t\right]$$
$$\frac{d^2\vec{r}^{\,\circ}(t)}{dt^2} = -\omega^2 \cdot \vec{r}^{\,\circ}(t) \tag{23.18}$$

In a similar way we obtain for the apsidal precession, i.e., the precession of the line of apsides $d\omega/dt$, around the unit momentum vector $\vec{h}^{\,\circ}(t)$

$$\vec{r}_{aps}^{\,\circ}(t) = \vec{r}_{an}^{\,\circ}(t)\cos\left(\omega_0 t + \dot{\omega}t\right) + \left[\vec{h}^{\,\circ}(t) \times \vec{r}_{an}^{\,\circ}(t)\right]\sin\left(\omega_0 t + \dot{\omega}t\right) \tag{23.19}$$

with an initial angular value ω_0. Since the precession of the orbital plane is uniquely determined by the normal of the orbital plane, one can directly model the angular momentum vector $\vec{h}$ by rotating it around the normal to the equatorial plane $\vec{\omega}_\Omega^{\,\circ}$ using the following orthogonal rotation

$$\vec{r}(t) = \vec{r}\vec{\omega}^{\,\circ} \cdot \vec{\omega}^{\,\circ} + \vec{\omega}^{\,\circ} \times \vec{r} \times \vec{\omega}^{\,\circ} \cdot \cos \omega t + \vec{\omega}^{\,\circ} \times \vec{r} \cdot \sin \omega t \tag{23.20}$$

After including the scalar product $\vec{h}\vec{\omega}_\Omega^{\,\circ} = \cos \alpha_{h\omega}$, that is constant for all rotation angles, we obtain

$$\vec{h}(t) = \vec{\omega}^{\,\circ} \cdot \cos \alpha_{h\omega} + \vec{\omega}^{\,\circ} \times \left(\vec{h} \times \vec{\omega}^{\,\circ} \cdot \cos \omega t + \vec{h} \cdot \sin \omega t\right) \tag{23.21}$$

that reduces to a very elegant orthogonal spherical rotation defined as

$$.\, \vec{h}(t) := \vec{\omega}^{\,\circ} \cdot \cos \alpha_{h\omega} + \vec{\omega}^{\,\circ} \times \vec{a} . \tag{23.22}$$

with vector $\vec{a}$ where

$$\vec{a} := \vec{h} \times \vec{\omega}^{\,\circ} \cdot \cos \omega t + \vec{h} \cdot \sin \omega t \tag{23.23}$$

Such an elegant method to geometrically rotate a vector around an axis for a given angle of rotation has not been reported so far in literature.

To calculate the first derivative of the angular momentum vector $\vec{h}(t)$ in (25.23), only the second term plays a role

$$\dot{\vec{h}}(t) := \vec{\omega}^{\,\circ} \times \dot{\vec{a}} = -\dot{\omega} \cdot \vec{\omega}^{\,\circ} \times \left[\vec{h} \times \vec{\omega}^{\,\circ} \cdot \sin \omega t - \vec{h} \cdot \cos \omega t\right] \tag{23.24}$$

If we would like to rotate the vector $\vec{r}$ by an angle θ around a rotation vector $\vec{\omega}_1$, or arbitrary number of rotation vectors $(\vec{\omega}_1, \vec{\omega}_2, \vec{\omega}_3, \ldots, \vec{\omega}_n)$ we can also use the following nested relations

$$\vec{r}°(\vec{\omega}_1) = \vec{r}° \cos\theta + \left(\vec{\omega}_1 \times \vec{r}°\right)\sin\theta$$

$$\vec{r}°(\vec{\omega}_1, \vec{\omega}_2) = \vec{r}°(\vec{\omega}_1)\cos\theta + \left(\vec{\omega}_2 \times \vec{r}°(\vec{\omega}_1)\right)\sin\theta$$

$$\vec{r}°(\vec{\omega}_1, \vec{\omega}_2, \vec{\omega}_3) = \vec{r}°(\vec{\omega}_1, \vec{\omega}_2)\cos\theta + \vec{\omega}_3 \times \vec{r}°(\vec{\omega}_1, \vec{\omega}_2)\sin\theta$$

$$\ldots \tag{23.25}$$

$$\vec{r}°(\vec{\omega}_1, \vec{\omega}_2, \vec{\omega}_3, \ldots, \vec{\omega}_n) = \vec{r}°(\vec{\omega}_1, \vec{\omega}_2, \ldots, \vec{\omega}_{n-1})\cos\theta + \vec{\omega}_n \times \vec{r}°(\vec{\omega}_1, \vec{\omega}_2, \ldots, \vec{\omega}_{n-1})\sin\theta$$

$$\left|\vec{\omega}_1 \times \vec{r}°\right| = \left|\vec{\omega}_2 \times \vec{r}°(\vec{\omega}_1)\right| = \ldots = \left|\vec{\omega}_n \times \vec{r}°(\vec{\omega}_1, \vec{\omega}_2, \ldots, \vec{\omega}_{n-1})\right| = 1$$

Generally speaking, an Earth-centered satellite orbit has a main axis of rotation that is precessing around an another axis that defines apsidal and nodal precession of the orbit. Typically, non-gravitational forces such as air-drag or solar radiation pressure have a clear orbit period signal. Therefore, all orbit perturbations can be described by multipole rotations with an axis of rotation close to the main axis of the orbit. The concept of circular motion and orbit representation will be discussed further in this thesis.

24. The Concept of Counter-Rotating Circular Orbits

$\mathbf{H}$ere we discuss the concept of bi-circular orbits and bi-circular orbit perturbations. It is shown how an elliptical orbit can be decomposed into two counter-rotating circular orbits. In this way, orbital dynamics can be approximated geometrically by circular orbits or circular rotations. Two counter-rotating orbits remove the variation of the orbit radius. Bi-circular orbit representation is essentially a linear combination of two harmonic oscillators with an opposite direction of rotation. In Section 19, we applied a simple harmonic oscillator to daily estimates of residual Galileo clock parameters. We just looked into the remaining amplitude in the clock parameters that measure the radial orbit error after removing a linear model (time offset and drift removed). Similar results to the circular representation of the effect where obtained when a solution of Hill equations in the radial direction (Colombo 1986) was plotted after removing a linear model (bias and drift) in the radial direction, see Section 19. The use of harmonic oscillators leads us also to the synergy or unification in modeling of orbital and rotational dynamics. We will show in the next section an interesting feature of circular orbit representation: that for a Keplerian orbit the velocity vector describes a circle. The velocity vector of the satellite in the presence of any point-like mass will rotate about that object along a circle with a constant radius. Thus an interesting application is in supporting numerical integration.

Another interesting feature of circular perturbations is in preserving the orthonormality of the rotation transformation, i.e., the geometrical properties of the orbit. The term orthonormality group denotes an orthogonal set of vectors that are normalized in terms of length. Most analytical orbit theories use a form of Keplerian motion as a reference and in numerical integration, typically, higher-order polynomials are used to approximate the orbit over an integration step. Here we use a combination of two uniform circular motions to represent the orbit in terms of orbit positions and in the next section we will see how to use a circular representation and its multipole expansion in modeling orbit velocity.

24.1 The Concept of Bi-Circular Orbits

The simplest orbit in celestial mechanics is the circular orbit. It can be represented as a special case of a Keplerian orbit which is the general solution of the two-body problem. In the case of circular orbit, geometry of the orbit is represented by a circle and Kepler's equation reduces to the equation of uniform mean motion. The vector of motion of a satellite in uniform circular rotation with a radius c, and a constant rotation rate n can be defined as

$$\vec{r}_{\odot} := c(\vec{C}_1 \cos(nt) + \vec{C}_2 \sin(nt)) \tag{24.1}$$

with the orthonormal vector basis $\vec{C}_1$ and $\vec{C}_2$. The mean motion is derived from Kepler's Third Law

$$n = \sqrt{\frac{GM}{c^3}} \tag{24.2}$$

with the geocentric constant GM. Denoting the prograde orthonormal rotation with $\vec{o}^+(n)$

$$\vec{o}^+(n) := \vec{C}_1 \cos(nt) + \vec{C}_2 \sin(nt) \tag{24.3}$$

the final kinematic equation of the prograde orthonormal rotation can be written as

$$\vec{r}_\odot := c\vec{o}^+(n) \tag{24.4}$$

Introducing the second derivative

$$\ddot{\vec{o}}^+(n) = -n^2 \vec{o}^+(n) \tag{24.5}$$

we then obtain the dynamic equation of the prograde orthonormal rotation

$$\ddot{\vec{r}}_\odot := c\ddot{\vec{o}}^+(n) \tag{24.6}$$

or

$$\ddot{\vec{r}}_\odot = -n^2 \vec{r}_\odot = -\frac{GM}{c^2} \vec{o}^+(n) \tag{24.7}$$

Rearranging both sides we can see that (24.7) is the differential equation of a simple harmonic motion

$$\ddot{\vec{r}}_\odot + n^2 \vec{r}_\odot = \vec{0} \tag{24.8}$$

with n as angular frequency of oscillation. Finally, introducing the mean motion we obtain

$$\ddot{\vec{r}}_\odot + \frac{GM}{c^3} \vec{r}_\odot = \vec{0} \tag{24.9}$$

which is the equation of motion resulting from a central gravity term.

The velocity and orbit (gravity) gradient can be obtained in a similar way starting with the first derivative of the prograde orthonormal vector basis

$$\dot{\vec{o}}^+(n) = n\left(-\vec{C}_1 \sin(nt) + \vec{C}_2 \cos(nt)\right) \tag{24.10}$$

and finally

$$\dot{\vec{r}}_\odot = c\dot{\vec{o}}^+ \tag{24.11}$$

or

$$\ddot{\dot{\vec{r}}}_\odot = -n^2 c\dot{\vec{o}}^+ = -n^2 \dot{\vec{r}}_\odot = -\frac{GM}{c^3} \dot{\vec{r}}_\odot = -\frac{GM}{c^2} \dot{\vec{o}}^+ \tag{24.12}$$

Let us now introduce the retrograde orthonormal circular motion $\vec{o}^-(n)$ that describes a circular motion in the opposite direction

$$\vec{o}^-(n) := c\vec{o}^+(-n) = c\vec{C}_1 \cos(nt) + c\vec{C}_2 \sin(-nt) \tag{24.13}$$

or

$$\vec{r}_\odot := c\vec{o}^-(n) \tag{24.14}$$

The dynamic equation of the retrograde orthonormal orbit follows as

$$\ddot{\vec{o}}^-(n) = -n^2 \vec{o}^-(n)$$

$$\ddot{\vec{r}}_\odot = -\frac{GM}{c^2}\vec{o}^-(n) \tag{24.15}$$

Velocity and orbit (gravity) gradient can be obtained in a similar way as

$$\dot{\vec{o}}^-(n) = n\left(-\vec{C}_1 \sin(nt) - \vec{C}_2 \cos(nt)\right) \tag{24.16}$$

$$\dot{\vec{r}}_\odot = c\dot{\vec{o}}^- \tag{24.17}$$

$$\dddot{\vec{r}}_\odot = -n^2 c\dot{\vec{o}}^- = -n^2\dot{\vec{r}}_\odot = -\frac{GM}{c^3}\dot{\vec{r}}_\odot = -\frac{GM}{c^2}\dot{\vec{o}}^- \tag{24.18}$$

So far, we have considered two simple circular orbits with counter-rotation. Let us now define a linear combination of a prograde and a retrograde circular orbit with radii of rotation c and d that rotate at the same rotation rate n

$$\vec{r} := c\vec{o}^+(n) + d\vec{o}^-(n) \tag{24.19}$$

Introducing the vector basis $\vec{C}_1$ and $\vec{C}_2$ into (24.19) we obtain the equation of an ellipse

$$\vec{r} = (c+d)\vec{C}_1 \cos(nt) + (c-d)\vec{C}_2 \sin(nt) \tag{24.20}$$

$$\vec{r} = a\vec{C}_1 \cos(nt) + b\vec{C}_2 \sin(nt) \tag{24.21}$$

with semi-major a and semi-minor axis b defined as

$$a := c + d$$
$$b := c - d \tag{24.22}$$

from which it follows

$$c := \frac{a+b}{2} \qquad d := \frac{a-b}{2} \tag{24.23}$$

The constant radius of rotation c of the first circular orbit is given as the mean between the semi-major and semi-minor axis whereas the radius of the second circular orbit is computed as half the difference between the ellipse axes. The magnitude of the resulting radius vector can be derived as follows

$$r^2 = \left|c\vec{o}^+ + d\vec{o}^-\right|^2 = \left|c\vec{o}^+\right|^2 + \left|d\vec{o}^-\right|^2 + 2cd\vec{o}^+\vec{o}^- \tag{24.24}$$

$$r^2 = c^2 + d^2 + 2cd\cos(2nt) \tag{24.25}$$

which is the same as

$$r^2 = c^2 + d^2 - 2cd\cos(\pi - 2nt) \tag{24.26}$$

confirming that the sum of two vectors satisfies the cosine law, since $(\pi - 2nt)$ is the supplement of the angle between them. Introducing (24.23) into (24.25) leads to the equation of an ellipse

$$r^2 = a^2 \cos^2(nt) + b^2 \sin^2(nt) \tag{24.27}$$

As a conclusion, we have demonstrated that an elliptical motion can be represented as a superposition of two counter-rotating circular orbits. The general solution of the ordinary differential equation of a simple harmonic motion (24.8) is an ellipse and it can be decomposed into a superposition of two circular motions (24.21) with opposite rotation. This is graphically shown in Figure 24.1. If we rotate the inner and outer circle by the same angle in opposite directions, point A represented by the vector $c\vec{C}_1$ on the outer circle rotates to point A^+
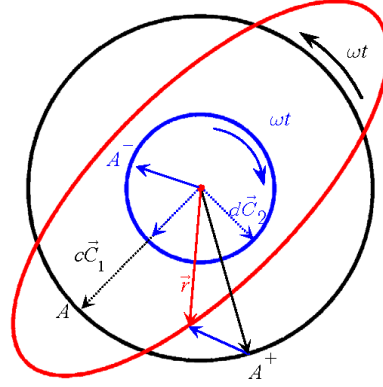
Figure 24.1 Elliptical motion as a superposition of two counter-rotating circular motions.

and the corresponding point on the inner circle is rotated together with the vector $d\vec{C}_2$ to the point A^-. The resulting vector $\vec{r}$ on the ellipse is the superposition of these two vectors. The eccentricity of the ellipse is then defined as

$$e := \frac{\sqrt{a^2 - b^2}}{a} = \frac{2\sqrt{cd}}{c + d}. \tag{24.28}$$

Combining (24.22) with (24.28), we can write the radius of the second circular orbit as a function of the orbit eccentricity and the radius of rotation of the first circular orbit

$$d := \frac{1 - \sqrt{1 - e^2}}{1 + \sqrt{1 - e^2}} c \tag{24.29}$$

Once the eccentricity has been derived, the equation of Kepler's ellipse centered at one of the focii is

$$\vec{r} = c\vec{o}^+(n) + d\vec{o}^-(n) - 2\sqrt{cd}\vec{C}_1 \tag{24.30}$$

or by introducing the eccentricity vector as

$$\vec{e} := e\vec{C}_1 \tag{24.31}$$

the kinematic equation of motion in terms of a uniform prograde and retrograde rotation, is given as

$$\vec{r} = c\vec{o}^+(n) + d\vec{o}^-(n) - c\vec{e} - d\vec{e} \tag{24.32}$$

From (24.32) we see that for every circular motion we have one translation to decompose Kepler's ellipse. The orthonormal vector basis $\vec{C}_1$ and $\vec{C}_2$ is defined such that $\vec{C}_1$ points towards the orbit pericenter and $\vec{C}_2$ is a perpendicular coplanar vector pointing in the direction of the satellite velocity at the pericenter. Figure 24.2 shows how the Kepler orbit can easily be oriented and for $c = d$ one obtains a linear motion represented by circular rotations.

In the next section we will see how to use circular representation and its multipole expansion in modeling orbit velocity. We will discuss in more detail the property that the velocity vector of the satellite in the presence of any point-like mass will rotate about that object along a circle with a constant radius. Therefore, the potential application of this model is in supporting numerical integration over long integration arcs, e.g., reference frame satellites, interplanetary orbits, etc.
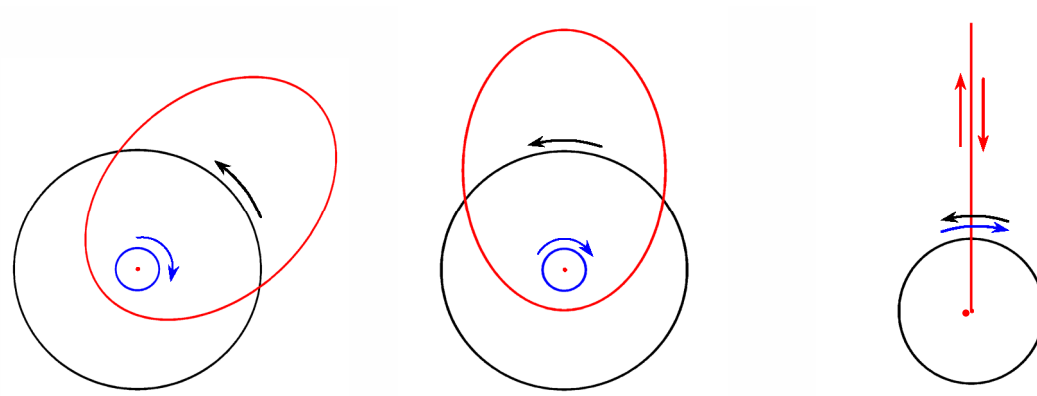
Figure 24.2 Elliptical orbit (left and middle) and linear motion (right) as a result of a coplanar counter-rotating motion. Ellipse: $c = 0.85a$, $d = 0.15a$ (left and middle), $c = d = 0.5a \rightarrow e = 1$ linear motion (right).

25. The Circular Kinematic and Dynamic Equation of a Satellite Orbit

Here we discuss the kinematic equation of a satellite orbit based on a circular representation of the velocity vectors of a Kepler orbit, otherwise known as the two-body problem in celestial mechanics. The velocity vector for Keplerian orbit describes a circle, i.e., we show that the velocity vector of the satellite in the presence of any point-like mass will rotate about that object along a circle with a constant radius. Thus, an interesting advantage of using circular perturbations is that this method preserves the orthonormality of the rotational transformation, i.e., the geometrical properties of the orbit. We show that the proposed circular model could be applied to kinematic as well as dynamic modeling of the orbit and rotation of a rigid body (satellite, Earth, etc.). In the case of circular perturbations, the radius of rotation is preserved, as is also the case with rotation of a rigid body (satellite, planet, etc.). At the end of this section, we discuss the proposed model in the light of geometrical integration, a special kind of integration that preserves the properties of the orbit, i.e., the exact flow of differential equations or Hamiltonian systems that govern satellite motion and rotation. In the light of circular perturbations we extend Newton's theorem of revolving orbits that defines a special central force as one that is changing the angular speed of the orbit by some constant factor, while the radial motion remains unaffected.

25.1 The Circular Kinematic and Dynamic Equation of Orbit

Let us first write the dynamic equation of satellite motion given by the radius-vector $\vec{r}$ including the central gravity term $GM \, / \, r^2$

$$\ddot{\vec{r}} = \frac{d\dot{\vec{r}}}{dt} = -\frac{GM}{r^2}\,\vec{r}^{\,\circ} \tag{25.1}$$

with the unit vector defined as $\vec{r}^{\,\circ}$. The associated constant angular momentum $\vec{h}$ of the orbit (given per unit mass, i.e., or the specific angular momentum) can be written as

$$\vec{h} = \vec{r} \times \dot{\vec{r}} \tag{25.2}$$

and considering the areal velocity over an angle θ, i.e., the area of the ellipse swept over a given period

$$\vec{h} = 2\,\frac{\text{area}}{\text{period}} = r^2\,\frac{d\theta}{dt}\,\vec{h}^{\,\circ} = 2\,\frac{ab\pi}{2\pi \, / \, \sqrt{\dfrac{GM}{a^3}}}\,\vec{h}^{\,\circ} = \sqrt{GMa(1-e^2)}\,\vec{h}^{\,\circ} = \sqrt{\frac{GMb^2}{a}}\,\vec{h}^{\,\circ} \tag{25.3}$$

with the ellipse semi-major axis a and semi-major axis b and eccentricity e, and the unit vector $\vec{h}^{\,\circ}$

$$h = \sqrt{GMa(1-e^2)} = \text{const} \tag{25.4}$$

279

Including *semi-latus rectum* p of the Kepler orbit

$$p = \frac{b^2}{a} = a(1 - e^2) = \text{const} \tag{25.5}$$

for the specific angular momentum we obtain

$$\vec{h} = \sqrt{GMp}\,\vec{h}^{\circ} \quad \rightarrow \quad h = \sqrt{GMp} \quad \rightarrow \quad GM = \frac{h^2}{p} \tag{25.6}$$

and from (25.3) we then have

$$\frac{dt}{d\theta} = \frac{r^2}{h} = \frac{r^2}{\sqrt{GMa(1 - e^2)}} = \frac{r^2}{\sqrt{GMp}} \tag{25.7}$$

By inserting (25.7) into the dynamic equation of satellite motion (25.1), we obtain

$$\frac{d\dot{\vec{r}}}{d\theta} = -\frac{GM}{h}\vec{r}^{\circ} = -\frac{GM}{\sqrt{GMa(1 - e^2)}}\vec{r}^{\circ} = -\sqrt{\frac{GM}{p}}\vec{r}^{\circ} = -\frac{\sqrt{GMa}}{b}\vec{r}^{\circ} \quad \rightarrow \quad \frac{d\dot{\vec{r}}}{d\theta} = -\frac{GM}{h}\vec{r}^{\circ} \tag{25.8}$$

Finally, considering (25.6), we obtain the derivative

$$\frac{d\dot{\vec{r}}}{d\theta} = -\frac{h}{p}\vec{r}^{\circ} \tag{25.9}$$

or

$$\frac{d\dot{\vec{r}}}{d\theta} = -\frac{h}{p}\vec{r}^{\circ} = -\frac{\left|\vec{r} \times \dot{\vec{r}}\right|}{p}\vec{r}^{\circ} \tag{25.10}$$

We see that the velocity vector describes a circle as a function of the true anomaly θ, i.e., the derivative of the velocity vector w.r.t. true anomaly is a circle with a constant radius h/p. If we now integrate (25.9) we obtain the equation of a circle centered at $\vec{k}$

$$\dot{\vec{r}} = -\frac{h}{p}\vec{r}_{\perp}^{\circ} + \vec{k} \qquad \vec{r}^{\circ} \cdot \vec{r}_{\perp}^{\circ} = 0, \quad \vec{r}^{\circ} \times \vec{r}_{\perp}^{\circ} = \vec{h}^{\circ} \tag{25.11}$$

where $\vec{k}$ is a constant velocity vector of integration and $\vec{r}_{\perp}^{\circ}$ is unit vector orthogonal to $\vec{r}^{\circ}$. It is interesting to note that the size of the circle in (25.10) does not depend on the orientation of the orbit, only on the shape of the orbit given by the *semi-latus rectum* p. In addition, there is one more interesting property: since $\vec{r}^{\circ}$ and $\vec{r}_{\perp}^{\circ}$ are two orthogonal vectors, the velocity $\dot{\vec{r}}$ and the orbit vector $\vec{r}$ are orthogonal under the following condition

$$\left(\dot{\vec{r}} - \vec{k}\right) \perp \vec{r} \tag{25.12}$$

Generally speaking, the specific angular momentum is not constant $\dot{h} \neq 0$, i.e., $\vec{h} = \vec{h}(t)$ and $p = p(t)$. Thus we obtain a torque exerted by the perturbing $\ddot{\vec{r}}$

$$\dot{\vec{h}}(t) = \frac{d}{dt}(\vec{r} \times \dot{\vec{r}}) = \dot{\vec{r}} \times \dot{\vec{r}} + \vec{r} \times \ddot{\vec{r}} = \vec{r} \times \ddot{\vec{r}} \tag{25.13}$$

Finally, combining (25.10) and (25.13), we obtain the equation of motion in the form defined as

$$\frac{d\dot{\vec{r}}}{d\theta} = -\frac{h(t)}{p(t)}\vec{r}^{\circ} = -\frac{\left|\vec{r} \times \dot{\vec{r}}\right|}{p(t)}\vec{r}^{\circ} \tag{25.14}$$

We call (25.14) the kinematic equation of motion or the kinematic form of the equation of motion, because the central gravity term that governs the motion does not appear in the equation. Kinematic equation is the terminology typically reserved for the description of the rotation of a body irrespective of the dynamics that govern that motion. We will see later in this section how a multipole representation could be used to model the general case, including all perturbations.

The importance of the kinematic equation of motion (25.14) is two-fold. Firstly, we see that the velocity vector of the satellite in the presence of any point-like mass (like central term of the gravity field) will rotate about that object along a circle with a constant radius GM/h or h/p. This "dynamic"-like constant radius GM/h or "kinematic"-like constant radius h/p is a constant in a Kepler orbit, analogous to the constant radius of a circular orbit. Thus one can model the orbit of a satellite in a way similar to the way we model the rotation of a rigid body, e.g., the attitude of a satellite or Earth rotation, making use of the specific angular momentum $\vec{h}$ and $\ddot{\vec{r}}$. We can see that in the case of a Kepler orbit, GM/h as well as h/p are both constants. Thus $d\dot{\vec{r}}/d\theta$ is a constant in a Kepler orbit, dependent only on the shape of the orbit, i.e., dependent only on the geocentric gravitational constant GM.

Eq. (25.14) can be integrated kinematically or dynamically with the initial state vector $\left\{\vec{r}_0, \dot{\vec{r}}_0\right\}$ defining the initial osculating Kepler orbit. This leads us to a special type of integration of ordinary differential equations that is often termed geometric integration, a numerical integration method that preserves the geometric properties of the exact flow of the differential equations. This means that the geometric properties of the orbit will be preserved even over a very long integration time, as well as if one were looking at the orbit at very short "microscopic" intervals. In this particular case, one can define energy to have conservative property in the geometric integration. This also opens up the possibility of separate numerical integration for the conservative and the non-conservative part of the orbit. This is not the case with polynomial representation of the orbit, as in the case of collocation approaches used in numerical integration. Geometric integration is very often considered in highly oscillatory mechanical systems as it preserves the properties of the Hamiltonian systems. Since geometrical integration is well known in literature, we do not give a specific reference. However, compared to Hamiltonian systems in celestial mechanics, where often generalized coordinates are used in terms of "momentum" and position, here we are using the geometrical properties of differential orbit velocity, that, according to (25.14), follow a simple circular motion (similar to a harmonic oscillator). Therefore, the second important aspect of (25.14) is that any satellite orbit can be represented by circular perturbations, i.e., geometrical rotations only. It is also astonishing that the kinematic "circular velocity equation" (25.14) is not a function of time at all, but depends only on the geometric angle θ, and h/p also has a purely geometrical representation, i.e., it does not explicitly depend on the gravity field. This is the reason why we call it a kinematic equation. Circular perturbations, e.g., in multipole expansion, offer a new way to represent and numerically integrate satellite orbits and are an alternative to the high-degree polynomials that are used in numerical integration at the moment. This is especially true for applications that require long integration time, as it is often the case in planetary geodesy, where the orbits of satellites and planets are integrated over long time periods, or gravity field missions for temporal gravity field variations. In the case of the general collocation methods often used in orbit integration, the polynomial model that approximates the orbit for each orbit component separately in the integration step reads as

$$r(t) = \sum_{i=0}^{q} (t - t_0)^i \cdot r_{0i} \tag{25.15}$$

with q denoting the degree of the polynomials and r_{0i} the coefficients that are estimated when fitting the second derivative of (25.15) to the acceleration field that governs the equation of motion and is calculated from models.

If we now consider the principle of moments from mechanics (Varignon's theorem), where the sum of the torques exerted by several forces (c.f. due to different harmonics in the spherical harmonic expansion) is equal to the torque of the resultant force, we can derive

$$\dot{\vec{h}} = \vec{r} \times \ddot{\vec{r}} = \vec{r} \times \ddot{\vec{r}}_2 + \vec{r} \times \ddot{\vec{r}}_3 + ... + \vec{r} \times \ddot{\vec{r}}_n \tag{25.16}$$

Thus, instead of integrating the acceleration field $\ddot{\vec{r}}$ along the orbit, one could use specific angular momentum (torque) $\dot{\vec{h}}$ since, in the case of a nearly circular orbit in the Earth's gravity field, angular momentum changes very slowly, c.f. precession of the orbital plane and apsidal line due to the J_2 coefficient of the Earth's gravity field. This makes angular momentum suitable for numerical integration of the orbit. We will see later in this section that to preserve the circular property of the orbit one could also make use of Newton's theorem of revolving orbits to account for perturbations in rotation, and the concept of bi-circular orbits to account for perturbations in a radial direction. Generally speaking, the circular property of the orbit can be preserved by integrating the velocity vector (25.10) making use of the linear Hill equations for constant acceleration along an orbit.

25.2 Orbit Representation Using Spherical Rotation

Let us now first see how, by introducing spherical rotation from the previous section and secular perturbations in orbital elements (Kaula 1966), one can easily represent a satellite orbit over a long period of time. Let us define the Kepler orbit by the specific angular momentum vector $\vec{h}$ and the line of nodes $\vec{r}_{an}$ (pointing towards the right ascension of the ascending nodes) and introduce precession of the orbital plane $d\Omega/dt$ due to the J_2 coefficient of the gravity field. Generally speaking, we may model the rotation of the vector $\vec{r}_{an}$ around the normal to the equatorial plane $\vec{\omega}_\Omega^\circ \approx \{0,0,1\}$ in a very elegant way by using the following orthogonal vector form, which we call spherical rotation

$$\vec{r}_{an}^\circ(t) = \vec{r}_{an}^\circ \cos \dot{\Omega} t + \left[\vec{\omega}_\Omega^\circ \times \vec{r}_{an}^\circ \right] \sin \dot{\Omega} t \tag{25.17}$$

In a similar way we can write for the apsidal precession, i.e., precession of the line of apsides $d\omega/dt$, around the momentum vector $\vec{h}^\circ(t)$

$$\vec{r}_{aps}^\circ(t) = \vec{r}_{an}^\circ(t) \cos \left(\omega_0 + \dot{\omega} t \right) + \left[\vec{h}^\circ(t) \times \vec{r}_{an}^\circ(t) \right] \sin \left(\omega_0 + \dot{\omega} t \right) \tag{25.18}$$

with the orbital plane defined uniquely by

$$\vec{h}^\circ(t) = \left| \vec{r}_{an}^\circ(t) \times \vec{r}_{aps}^\circ(t) \right| \tag{25.19}$$

Following (Kaula 1966), secular perturbations in the Keplerian elements due to the J_2 gravity field coefficients are given by

$$\frac{d\Omega}{dt} = \frac{3nC_{20}a_e^2}{2(1-e^2)^2 a^2} \cos i \qquad \frac{d\omega}{dt} = \frac{3nC_{20}a_e^2}{4(1-e^2)^2 a^2} [1 - 5\cos^2 i] \qquad \frac{dM}{dt} = n - \frac{3nC_{20}a_e^2}{4(1-e^2)^{3/2} a^2} [3\cos^2 i - 1]$$
$$\tag{25.20}$$

with $da/dt = de/dt = di/dt = 0$. Therefore, with just a few parameters, it is possible to model an orbit with an orthogonal vector basis over a long period of time. Since precession of the orbital plane is uniquely determined by the normal of the orbital plane, one can directly model specific angular momentum vector $\vec{h}$ by rotating it around the normal to the equatorial plane $\vec{\omega}_\Omega^\circ$ using the following orthogonal rotation

$$\vec{h}(t) = \vec{h}\vec{\omega}_\Omega^\circ \cdot \vec{\omega}_\Omega^\circ + \vec{\omega}_\Omega^\circ \times \vec{h} \times \vec{\omega}_\Omega^\circ \cdot \cos\dot{\Omega}t + \vec{\omega}_\Omega^\circ \times \vec{h} \cdot \sin\dot{\Omega}t \tag{25.21}$$

and after including the scalar product $\vec{h} \cdot \vec{\omega}_\Omega^\circ = \cos\alpha_{h\omega}$, that is constant for all rotation angles, we obtain

$$\vec{h}(t) = \vec{\omega}_\Omega^\circ \cdot \cos\alpha_{h\omega} + \vec{\omega}_\Omega^\circ \times \left(\vec{h} \times \vec{\omega}_\Omega^\circ \cdot \cos\dot{\Omega}t + \vec{h} \cdot \sin\dot{\Omega}t\right) \tag{25.22}$$

that reduces to a very elegant orthogonal spherical rotation given by

$$\vec{h}(t) = \vec{\omega}_\Omega^\circ \cdot \cos\alpha_{h\omega} + \vec{\omega}_\Omega^\circ \times \vec{a} \tag{25.23}$$

with the vector $\vec{a}$

$$\vec{a} = \vec{h} \times \vec{\omega}_\Omega^\circ \cdot \cos\dot{\Omega}t + \vec{h} \cdot \sin\dot{\Omega}t \tag{25.24}$$

To our knowledge, such an elegant way to geometrically rotate a vector about an axis for a given angle of rotation has never before been published.

To calculate the first derivative of the angular momentum vector $\vec{h}(t)$ in (25.23), only the second term plays a role

$$\dot{\vec{h}}(t) := \vec{\omega}_\Omega^\circ \times \dot{\vec{a}} = -\dot{\Omega} \cdot \vec{\omega}_\Omega^\circ \times \left[\vec{h} \times \vec{\omega}_\Omega^\circ \cdot \sin\dot{\Omega}t - \vec{h} \cdot \cos\dot{\Omega}t\right] \tag{25.25}$$

In the general form of spherical rotation (25.23), we can add different frequencies and additional axes of rotation. For instance, in the case of Earth rotation, in addition to precession we have nutation due to tidal forces of the Moon and Sun, with the main period of 18.6 years, the same as that of the precession of the Moon's orbital nodes. Thus, to add an additional rotation on top of (25.23) around an axis $\vec{\omega}_1^\circ$ by an angle $\omega_1 t$ we may write

$$\vec{\omega}_\Omega^\circ(t) = \vec{\omega}_1^\circ \cdot \cos\alpha_{\omega\omega_1} + \vec{\omega}_1^\circ \times \left(\vec{\omega}_\Omega^\circ \times \vec{\omega}_1^\circ \cdot \cos\omega_1 t + \vec{\omega}_\Omega^\circ \cdot \sin\omega_1 t\right) \tag{25.26}$$

Such a nested rotational structure can be extended to any frequency argument $n\omega_1 t$ and rotation axis $\vec{\omega}_n^\circ$.

$$\vec{\omega}_\Omega^\circ = \vec{\omega}_1^\circ \cos\omega_{12}t + \vec{\omega}_2^\circ \sin\omega_{12}t \qquad \vec{\omega}_1^\circ \perp \vec{\omega}_2^\circ \perp \vec{\omega}_\Omega^\circ \tag{25.27}$$

25.3 Multipole Circular Perturbations and Newton's Theorem of Revolving Orbits

To continue this discussion on orbit representation, let us now see if one can separate radial motion from angular motion. In Proposition XLIII and in Proposition XLIV of Newton's *Principia* (Newton 1687), it is stated

PROPOSITION XLIII. PROBLEM XXX.

It is required to make a body move in a trajectory that revolves about the centre of force in the same manner as another body in the same trajectory at rest.

PROPOSITION XLIV. THEOREM XIV.

The difference of the forces, by which two bodies may be made to move equally, one in a quiescent, the other in the same orbit revolving, is in a triplicate ratio of their common altitudes inversely.
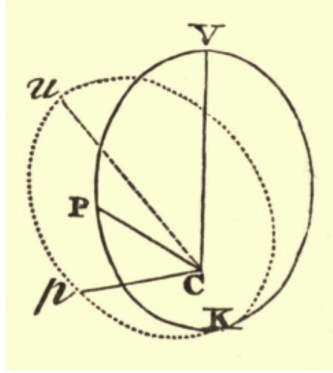
"It is required to make a body move in a trajectory that revolves about the centre of force in the same manner as another body in the same trajectory at rest." – Proposition XLIII

"The difference of the forces, by which two bodies may be made to move equally, one in a quiescent, the other in the same orbit revolving is in a triplicate ratio of their altitudes inversely. " – Proposition XLIV

Figure 25.1 Newton's theorem of revolving orbits, as published in *Philosophiæ Naturalis Principia Mathematica*, (Newton 1687) showing apsidal precession of the Kepler orbit.

Following Newton's theorem of revolving orbits in Figure 25.1, *Proposition XLIII* introduces apsidal precession under the special category of a central force. *Proposition XLIV* says that the difference of the central forces between those two orbits (perturbed and unperturbed *"at rest"*) varies inversely as the cube of their radial distances. Newton's theorem of revolving orbits defines a central force as one that is changing the angular speed of the orbit by some constant factor k, while the radial motion stays unaffected. Thus, for the true anomalies between those two Newtonian orbits we may write

$$\frac{d\theta_2}{dt} = k \frac{d\theta_1}{dt} \tag{25.28}$$

and for the corresponding specific angular momentums

$$h_2 = r^2 \frac{d\theta_2}{dt} = r^2 k \frac{d\theta_1}{dt} = k \cdot h_1 \tag{25.29}$$

If we now assume that the orbit is circular or nearly circular we may write the following Euler-Lagrange equation

$$\frac{d^2 r}{dt^2} - r \left(\frac{d\theta}{dt}\right)^2 = \frac{d^2 r}{dt^2} - \frac{h^2}{r^3} \tag{25.30}$$

separating radial $\dfrac{d^2 r}{dt^2}$ and rotational part $r \left(\dfrac{d\theta}{dt}\right)^2$. Considering (25.29), we obtain for the difference in radial acceleration (Newton 1687)

$$\Delta \ddot{r} = \ddot{r}(r_2) - \ddot{r}(r_1) = \frac{h_1^2}{r^3} - \frac{k^2 h_1^2}{r^3} = \frac{h_1^2}{r^3}\left(1 - k^2\right) \tag{25.31}$$

Thus, by considering an inverse-cube acceleration, the angular speed or angular momentum will be changed by a constant factor k. There will be no effect of the new radial acceleration in (25.31) if $k = \pm 1$. However, the total angular momentum in (25.28) depends on the sign of k. By setting $k = 2, 3, ..., n$, (25.31) can be used for multipole expansion of the orbit representation, where each angular frequency defined by k gives a different perturbation at a different orbit frequency, and so a different contribution to the final orbit.

In multipole representation, the kinematic equation of orbit (25.14) can be written as

$$\frac{d\dot{\vec{r}}}{d\theta}(k_i) := -\frac{k_i \cdot h}{p}\vec{r}^{\circ} = -k_i\frac{\left|\vec{r} \times \dot{\vec{r}}\right|}{p}\vec{r}^{\circ} = -k_i\frac{\left|\vec{r} \times \dot{\vec{r}}\right|}{a(1-e^2)}\vec{r}^{\circ} \tag{25.32}$$

Therefore, the integration of the individual frequencies will lead to an orbit velocity

$$\dot{\vec{r}} = -\int\frac{k \cdot h}{p}\vec{r}^{\circ}d\theta = -\sum_i\frac{k_i \cdot h}{p}\vec{r}_{\perp}^{\circ} + \vec{k}_0 \tag{25.33}$$

This model could be extended considering that each frequency k_i contributes with a coefficient c_i, which leads to

$$\dot{\vec{r}} = -\int\frac{k \cdot h}{p}\vec{r}^{\circ}d\theta = -\sum_i c_i\frac{k_i \cdot h}{p}\vec{r}_{\perp}^{\circ} + \vec{k}_0 \tag{25.34}$$

with an initial condition $\vec{k}_0$. Let us now see how to consider the general case of (25.32), when $k = k(\theta)$, thus $k/a(1-e^2)$ is no longer a constant radius in (25.32). For this we will introduce the concept of bi-circular orbits.

With the theorem of revolving orbits, (Newton 1687) introduced the concept of an inverse-cube central force in order to explain the apsidal precession of the Moon's orbit. Newton's theorem of revolving orbits defines a central force that increases the angular rate of the orbit by some constant factor k, while the radial motion remains unaffected

$$\frac{d\theta_2}{dt} = k\frac{d\theta_1}{dt} \tag{25.35}$$

which is the same as (25.28). This is a very important theorem that tells us that by adding inverse-cube type acceleration to any type of central force, the angular rate of the corresponding orbit will be changed by a constant factor, while the radial motion is the same for both orbits. The specific angular momentum for the second orbit is then again

$$h_2 = r^2\frac{d\theta_2}{dt} = r^2 k\frac{d\theta_1}{dt} = k \cdot h_1 \tag{25.36}$$

Let us consider a spherically symmetric gravitational potential, i.e., a potential that depends only on the radial distance, so that $V \sim GM \cdot a^n/r^{n+1}$. Making use of spherical harmonic rotation, the general form of the spherical harmonic representation of the gravity field of the Earth can be written as spherically symmetric gravitational potential $V \sim GM \cdot a^n/r^{n+1}$ considering that the rotation w.r.t. the initial state is purely a matter of datum definition. When acceleration is expressed in polar coordinates, the radial component is non-zero and we may write, see (25.30)

$$\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = \frac{d^2r}{dt^2} - \frac{h^2}{r^3} \tag{25.37}$$

For the two orbits sharing the same radial motion, we may write for the difference in radial acceleration

$$\Delta\ddot{r} = \ddot{r}(r_2) - \ddot{r}(r_1) = -\frac{k^2 h_1^2}{r^3} + \frac{h_1^2}{r^3} = \frac{h_1^2}{r^3}\left(1-k^2\right) \tag{25.38}$$

that gives the inverse-cube acceleration. Considering only the central term of the Earth's gravitational field $V = GM/r$, the specific angular momentum of the Keplerian orbit is $h_1 = \sqrt{GMa(1-e^2)}$, hence we obtain

$$\Delta\ddot{r} = \frac{GM}{r^2} \cdot \frac{a}{r}(1-e^2)\left(1-k^2\right) \tag{25.39}$$

285

an inverse-cube acceleration, whereas the specific angular momentum is changed by constant factor k. Since both initial and perturbed Keplerian orbits share the same radial motion, we may introduce the equation of the initial orbit $a(1-e^2)/r = (1 + e\cos\theta)$

$$\Delta\ddot{r} = \frac{GM}{r^2} \cdot (1 + e\cos\theta)\left(1 - k^2\right) \tag{25.40}$$

Thus the equation of motion of the perturbed orbit can be written in the form

$$\frac{d\dot{\vec{r}}}{dt} = -\frac{GM}{r^2}\vec{r}^{\circ} + \frac{GM}{r^2} \cdot (1 + e\cos\theta)\left(1 - k^2\right)\vec{r}^{\circ} \tag{25.41}$$

Assuming a circular orbit $e = 0$ we obtain

$$\Delta\ddot{r} = \frac{GM}{r^2} - \frac{GM}{r^2}k^2 \tag{25.42}$$

Thus the equation of motion of the perturbed orbit can be written in the form

$$\frac{d\dot{\vec{r}}}{dt} = -\frac{GM}{r^2}\vec{r}^{\circ} + \frac{GM}{r^2}\left(1 - k^2\right)\vec{r}^{\circ} = -k^2\frac{GM}{r^2}\vec{r}^{\circ} \tag{25.43}$$

We see that for $k = \pm 1$ we have the initial unperturbed Keplerian orbit. By setting $k = 2, 3, ..., n$, the angular rate or the specific angular momentum will be perturbed by a constant factor k, whereas the radial acceleration is changed by k^2. However, orbits will share the same radial motion, i.e., the radius of both central orbits will be the same $r = a$. This can easily be seen if (25.43) is multiplied by

$$\frac{dt}{d\theta} = \frac{r^2}{k \cdot h_1} \tag{25.44}$$

and we thus obtain the circular velocity perturbation

$$\frac{d\dot{\vec{r}}}{d\theta} = -k\frac{GM}{h_1}\vec{r}^{\circ} = -k \cdot R^*\vec{r}^{\circ} \qquad \Rightarrow \qquad \frac{d\dot{\vec{r}}}{k \cdot d\theta_1} = -\frac{GM}{h_1}\vec{r}^{\circ} = -R^*\vec{r}^{\circ} \tag{25.45}$$

where R^* is the constant radius of the circle $R^* = GM/h_1$. If we denote a as the radius or semi-major axis of the initial central orbit we see that there is a geometrical interpretation of the factor k, i.e., for central orbits we may write $k = a/r$. Thus for degree $n = 2$ of the spherical harmonic expansion we may write

$$k_2 = \frac{a}{r} \tag{25.46}$$

Eq. (25.31) can be used for a multipole expansion of the orbit representation, since each angular frequency defined by k defines a perturbation at the frequency that is a harmonic of the original orbit frequency. If $k > 1$, the added inverse-cube force is attractive, whereas it is repulsive when $-1 < k < 1$. When $k \approx 1$, both orbits are similar and the net effect is either apsidal precession, if k is slightly lower than 1, or regression, if k is slightly higher than 1.

Let us now derive the circular perturbations for Newton's inverse-cube acceleration (25.31) of the initial Keplerian orbit in the form

$$\frac{d\dot{\vec{r}}}{dt} = -\frac{GM}{r^2}\vec{r}^{\circ} + \frac{h_1^2}{r^3}\left(1 - k^2\right)\vec{r}^{\circ} \tag{25.47}$$

The specific angular momentum of the perturbed orbit is $h_2 = kr^2 \cdot d\theta_1/dt = kr^2 \cdot d\theta/dt$. Thus by multiplying (25.47) with

$$\frac{dt}{d\theta} = \frac{r^2}{k \cdot h_1} \tag{25.48}$$

we obtain

$$\frac{d\dot{\vec{r}}}{d\theta} = -\frac{GM}{kh_1}\vec{r}^\circ + \frac{h_1}{kr}\left(1 - k^2\right)\vec{r}^\circ \tag{25.49}$$

Considering that both orbits share the same radial motion of the Keplerial orbit in the form $r = a(1 - e^2)/(1 + e\cos\theta)$ and considering the specific angular momentum of the initial Keplerian orbit $h_1^2 = GMa(1 - e^2)$, we obtain

$$\frac{h_1}{r} = \frac{GM}{h_1}\left(1 + e\cos\theta\right) \tag{25.50}$$

In addition, by denoting $R^* = GM/h_1$, equation (25.49) reduces to

$$\frac{d\dot{\vec{r}}}{d\theta} = -R^*k\left[1 - \frac{1-k^2}{k^2}e\cos\theta\right]\vec{r}^\circ \tag{25.51}$$

Finally, by introducing the eccentricity of the perturbed orbit $e^* = \frac{1-k^2}{k^2}e$, we obtain

$$\frac{d\dot{\vec{r}}}{d\theta} = -k \cdot R^*\left(1 - e^*\cos\theta\right)\vec{r}^\circ \tag{25.52}$$

This is the equation of an ellipse as long as $\frac{1-k^2}{k^2}e < 1$, and can be written in bi-circular representation and be directly integrated without any numerical integration. For $k = 1$, (25.52) reduces to the equation of a circle for the Keplerian orbit.

Let us now see, what happens when k is not constant along the orbit and is dependent on the radial distance to the satellite, i.e., $k \sim 1/r$. In general form, the radial gravitational acceleration can be written as

$$\frac{d\dot{\vec{r}}}{dt} = -GM(n+1)\frac{C_n \cdot a_e^n}{r^{n+1}}R \cdot \vec{r}^\circ \tag{25.53}$$

where $R \cdot \vec{r}^\circ$ is the rotation of the unit-radius vector $\vec{r}^\circ$ of the orbit in the direction of gravitation. In the case of Earth-bounded orbits, those two vectors are nearly collinear. The a_e is the equatorial Earth radius used in the spherical harmonic expansion of the gravity field and $C_n = C_n(\vec{r}^\circ)$ can be related to the initial state vector $\vec{r}_\circ^\circ$ of the orbit

$$C_n = \Re_n \cdot C_n(\vec{r}_\circ^\circ) \tag{25.54}$$

where $\Re_n$ is the transformation matrix of the spherical harmonic coefficients defining rotation of the harmonic coefficients C_n given for a degree n. Therefore, in the general case $k_n = c_n/r^n$ we may write

$$\frac{d\theta_2}{dt} = k_n \frac{d\theta_1}{dt} = \frac{c_n}{r^n}\frac{d\theta_1}{dt} \tag{25.55}$$

with $c_n = c_n(\vec{r}^\circ)$ for a degree n. Considering again the Keplerian orbit in the form $r = a(1 - e^2)/(1 + e\cos\theta)$ for $k_1 = c_1/r$ we derive

$$\frac{d\dot{\vec{r}}}{d\theta} = -c_1 \cdot R^* \left(1 - e^* \cos\theta\right)\left(1 - e \cos\theta\right)\vec{r}^{\,\circ} \tag{25.56}$$

Thus, since adding an inverse-cube radial acceleration to the inverse-square acceleration corresponds to a potential $\sim 1/r^2$ (first degree in terms of spherical harmonics), we can derive

$$V_2 = -\frac{h_1^2}{2r^2}\left(1 - k^2\right) \tag{25.57}$$

or

$$V_2 = -\frac{GM}{r}\frac{a}{r} \cdot \frac{1 - e^2}{2}\left(1 - k^2\right) \tag{25.58}$$

Discussion in this section shows that there are very interesting alternative approaches for representing an orbit from the geometrical point of view and that Newton's theorem of revolving orbits, although not well known in the relevant literature, leads to very interesting geometrical properties.

26. A Geometrical Approach for the Computation and Rotation of Spherical Harmonics and Legendre Functions up to Ultra-High Degree and Order

In this section we introduce a new algorithm for the computation and rotation of spherical harmonics, Legendre polynomials and associated Legendre functions up to ultra-high degree and order. The algorithm is based on the geometric rotation of Legendre polynomials in Hilbert space. It is shown that Legendre polynomials can be calculated using geometrical rotations and can be treated as vectors in the Hilbert space leading to unitary Hermitian rotation matrices with geometric properties. We use the term geometrical rotations because although rotation itself is not governed by gravity and it can be used as a proxy to represent a gravity field geometrically. This novel method allows the calculation of spherical harmonics up to an arbitrary degree and order, i.e., up to degree and order 10^6 and beyond.

26.1 Basic Definitions

Following (Arfken et al. 1995), Legendre polynomials may appear in many different mathematical and physical solutions: 1) they may originate as solutions of the Legendre differential equation, 2) they may appear as a consequence of Rodrigues' formula, 3) they may be constructed as a consequence of the requirement for a complete, orthogonal set of functions (Gram-Schmidt orthogonalization), 4) we find them in gravity field modeling when representing a function in terms of spherical harmonics or in quantum mechanics to represent angular momentum eigenfunctions, 5) they may be generated by a generating function. The so-called Legendre differential equation is a second-order ordinary differential equation with two linearly independent solutions. The associated Legendre function of the first kind, often denoted as $P_{nm}(\cos\theta)$ of degree n and order m is a solution of the Legendre differential equation which is regular for all co-latitude angles θ. The associated Legendre function of the second kind, often denoted as $Q_{nm}(\cos\theta)$ is singular for $\theta \in \{0, \pi\}$. The complete solution of the Legendre differential equation is a linear combination of the associated Legendre functions of the first and second kind. In their famous textbook on physical geodesy (Heiskanen and Moritz 1967), one can find surface spherical harmonics as the angular portion of the solution to the Laplace equation in spherical coordinates, assuming that azimuthal symmetry is not present. This is the standard representation of the spherical harmonics used in geodesy, i.e., modeling the gravity field of the Earth and other planets.

The general method for calculating Legendre polynomials is by using a hypergeometric series (Abramowitz and Stegun 1965), (Koepf 1998) (see the given references for the description of arguments)

$$P_n(x) = {}_2F_1\left(-n, n+1, 1, \frac{1-x}{2}\right), \quad n = 0, 1, 2, \ldots \tag{26.1}$$

Hypergeometric functions are solutions to the hypergeometric differential equation, which is a general second-order ordinary differential equation. Many elementary functions, such as Bessel functions, elliptic integrals, error functions, gamma functions, and classical orthogonal polynomials are in fact special cases of hypergeometric functions. There are several alternative methods to evaluate Legendre functions, and the standard recursion formulae exclusively used to compute associated Legendre functions $P_{nm} = P_{nm}(\cos\theta)$ are as follows (e.g., (Hobson 1931)):

$$
\begin{aligned}
P_{mm} &= (2\mathrm{m}-1)\sin\theta P_{m-1,m-1} && \searrow \\
P_{m+1,m} &= (2m+1)\cos\theta P_{mm} && \downarrow \\
P_{nm} &= \frac{1}{n-m}\Big((2n-1)\cos\theta P_{n-1,m} - (n+m-1)P_{n-2,m}\Big) && \downarrow
\end{aligned}
\tag{26.2}
$$

where arrows symbolically show the direction of computation (recursions) over degree n and order m. This approach has the disadvantage that Legendre functions of a particular degree/order require the computation of previous degrees/orders in the recursion chain. Thus, the numerical errors accumulate with increasing degree and the absolute size of the functions may reach the critical size for representation on standard computation platforms. Usually, for a particular degree of expansion, the recursions start with associated Legendre functions of sectorial harmonics of the previous degree. In this case, order and degree are equal and Legendre functions reach extreme values that cannot be handled on standard computation platforms.

26.2 Admissible Underflow Co-Latitudes for the Computation of Associated Legendre Functions

Compared to associated Legendre functions, normalized or unnormalized Legendre polynomials up to ultra-high degree (i.e., 3000 or even up to 10 000 or higher) are very uniform in size and experience neither computational nor numerical problems. This will be demonstrated in the following sections.

In the relevant literature one can typically find three basic numerical aspects in computing spherical harmonics or Legendre functions of ultra-high degree: 1) the numerical efficiency of the algorithm, 2) the stability of the recurrence relations in the computation of the Legendre functions and 3) the underflow problem in recurrence relations. Recurrence relations are crucial in all three categories. The term underflow or arithmetic underflow (or floating point underflow) is a condition where the result of a computer program calculation is a number that has a smaller absolute value than the smallest value that computer can store in its memory. The underflow problem in recurrence relations can easily be seen in the asymptotic approximation of the normalized associated Legendre functions (Smith et al. 1981)

$$
\bar{P}_{nn} \sim \frac{1}{2}\left(\frac{n}{\pi}\right)^{1/4}(\sin\theta)^n, \quad n \to \infty, \quad \theta \text{ fixed},
\tag{26.3}
$$

where $\bar{P}_{nn}$ denotes the fully normalized associated Legendre functions of degree and order n. The standard recursion commonly used to compute $\bar{P}_{nm}$ cannot be initialized due to an underflow during the computation of $\bar{P}_{nn}$. Following (Wittwer et al. 2008) the maximum admissible degree n for a given function of the smallest non-zero positive and the largest non-zero negative normalized number ω that is storable for the given compiler and software is

$$
n_{\max} < \frac{2\lg 2\omega - \dfrac{1}{2}\lg\dfrac{2\lg 2\omega}{\pi\lg(\sin^2\theta)}}{\lg(\sin^2\theta)} \approx \frac{2\lg 2\omega}{\lg(\sin^2\theta)}.
\tag{26.4}
$$

According to the IEEE standard for binary floating-point arithmetics in double precision $\omega \approx \pm 2.225 \times 10^{-308}$. For instance, an underflow will occur in IEEE double precision for co-latitudes outside the interval from $21.7°$ to $158.3°$ if the maximum degree is 720 (Wittwer et al. 2008). One can draw the general conclusion that errors may occur for co-latitude angles close to $0°$ and $180°$. For instance, with the expansion up to degree 360, an underflow will occur for all co-latitudes below $8°$ and above its complementary co-latitude angle of $172°$. For an expansion up to degree 240 the underflow co-latitude angle is $3°$ and for degree 180, the underflow will occur for all latitudes below $1.13°$ and above its complementary co-latitude angle of $1.13°$. Considering that almost all geodetic LEO satellites are in polar orbits, i.e., with orbit inclinations close to $90°$, such effects will take place in the dynamic orbit modeling, especially for missions where gravity field determination requires high resolution. There are a number of proposed ways to extend the interval of admissible co-latitudes, such as (Wenzel 1998), where all the upward computations are scaled by a factor of 10^{200}.

In (Libbrecht 1985) and (Holmes and Featherstone 2002) a method was presented based on a recurrence relation for $\bar{P}_{nm} / \sin(m\theta)$ that eliminates the problematic $\sin(m\theta)$ term from the recursive algorithms and reintroduces it by employing Horner's scheme. However, in order to avoid an overflow during the recursions, a scale factor of 10^{280} is introduced. In (Jekeli et al. 2007) it is observed that Legendre functions for specific orders show a very strong attenuation w.r.t. the degree/order domain as a function of the degree and the co-latitude.

A closer look at asymptotic expressions for Legendre polynomials, e.g., given in (Press et al. 2007), reveals two particular cases that do not pose any numerical problems, i.e., for sectorials $m = n$ we obtain

$$\bar{P}_{nn} \sim \frac{1}{2}\left(\frac{n}{\pi}\right)^{1/4} (\sin\theta)^n \quad \Rightarrow \quad \begin{cases} \theta = \dfrac{\pi}{2}: \quad \bar{P}_{nn} \sim \dfrac{1}{2}\left(\dfrac{n}{\pi}\right)^{1/4} \\[3ex] \theta = 0: \quad \bar{P}_{nn} = 0 \end{cases} \tag{26.5}$$

and for Legendre polynomials (zonals $m = 0$)

$$P_n \sim \left(\frac{2}{n\pi}\right)^{1/2} \sin\left[(n+1)\frac{\pi}{2}\right] \quad \Rightarrow \quad \begin{cases} \theta = \dfrac{\pi}{2}: \quad P_n \sim \left(\dfrac{2}{n\pi}\right)^{1/2} \\[3ex] \theta = 0: \quad P_n = 1 \end{cases} \tag{26.6}$$

For an ultra-high degree and order, e.g., $n = 10^6$, we obtain

$$\theta = \frac{\pi}{2}: \qquad n = 10^6 \rightarrow \bar{P}_{nn}(\cos\frac{\pi}{2}) \approx 47.5 \tag{26.7}$$

$$\theta = \frac{\pi}{2}: \qquad n = 10^6 \rightarrow \bar{P}_n(\cos\frac{\pi}{2}) \approx 1.6 \tag{26.8}$$

showing that there are no numerical problems for the calculation of zonal and sectorial spherical harmonics at the equator and pole. This means, if a rotation of the spherical harmonics can be decomposed into several rotations and where that about the equatorial axis is limited to a rotation only from equator to pole, we can calculate spherical harmonics to any desired ultra-high degree and order. Or in other words, the algorithm to calculate associated Legendre functions could be based on pre-calculating associated Legendre functions at the equator (with recursions that are stable) and solely use an equivalent rotation along the equator to obtain associated Legendre functions at any location on the sphere.

26.3 Geometrical Rotation of Spherical Harmonics in Hyperspace

Spherical basis functions, e.g., spherical harmonics or wavelets, play a central role in modeling spatial and temporal processes in the system Earth. So far, to our knowledge, no usable algorithm has been published neither for rotation of spherical harmonics nor wavelet representations. The transformation of spherical harmonics under an arbitrary rotation of the coordinate system has been studied in the past and the earliest reference dates back to (Schmidt 1899). Most of the work in this field over the last 50 years has been based on (Wigner 1959) and (Edmonds 1960) and related to group theory in quantum mechanics. In geodesy, the rotation of spherical harmonics has been related to inclination functions and the analysis of the perturbations of satellite orbits. Inclination functions were introduced in [*Kaula*, 1961] and in [*Kaula*, 1966]. If we write

$$V = \sum_{n=0}^{\infty} \sum_{m=0}^{n} V_{nm} \tag{26.9}$$

$$V_{nm} = GM \frac{a_e^n}{r^{n+1}} \sum_{p=0}^{n} F_{nmp}(i) \left\{ \begin{bmatrix} C_{nm} \\ S_{nm} \end{bmatrix}_{n-\text{m odd}}^{n-m \text{ even}} \cos\left[(n-2p)u + m\Lambda\right] + \begin{bmatrix} S_{nm} \\ C_{nm} \end{bmatrix}_{n-\text{m odd}}^{n-m \text{ even}} \sin\left[(n-2p)u + m\Lambda\right] \right\}$$

$$u = \omega + \nu$$

$$\Lambda = \Omega - \theta$$

$$\tag{26.10}$$

expressing the gravitational potential V as a function of orbit inclination i, argument of latitude u (sum of eccentric anomaly ω and true anomaly ν), right ascension of the ascending node Ω and Greenwich Sidereal Time θ. The equatorial radius is denoted by a_e and GM is the geocentric gravitational constant. The corresponding inclination function $F_{nmp}(i)$ is

$$F_{nmp}(i) = \sum_t \frac{(2n-2t)!}{t!(l-t)!(n-m-2t)!2^{2n-2t}} \sin^{n-m-2t} i$$

$$\times \sum_{s=0}^{m} \binom{m}{s} \cos^s i \sum_c \binom{n-m-2t+s}{c} \binom{m-s}{p-t-c} (-1)^{c-k} \tag{26.11}$$

where k is the integer part of $(l-m)/2$ and t is summed from 0 to the lesser of p or k, and c is summed over all values making the binomial coefficients non-zero, see [*Kaula*, 1966]. This expansion is based on the particular form of the associated Legendre functions that can be found in (Hobson 1931)

$$P_{nm}(\sin\varphi) = \cos^m \varphi \sum_{t=0}^{k} T_{nmt} \sin^{n-m-2t} \varphi \tag{26.12}$$

and

$$T_{lmt} = \frac{(-1)^t (2l-2t)!}{2^l t!(l-t)!(l-m-2t)!} \tag{26.13}$$

Note that gravity models are provided with normalized coefficients and, therefore, the inclination function in (26.11) needs to be normalized in order to be consistent. At the moment two of the most stable and accurate algorithms to calculate inclination functions can be found in (Emeljanov and Kanter 1989) and (Gooding and Wagner 2008). Based on a re-parameterization of the potential using orbital elements (26.10) and in combination with the Lagrange Planetary Equations, (Kaula 1966) developed his famous first-order linear perturbation theory of satellite orbits. The main application of this theory is in very efficient error-assessment tools developed for satellite-to-satellite tracking ((Rosborough and Tapley 1987), (Casotto 1993)) and for satellite gradiometry (Sneeuw 2000). (Sneeuw 2000) applied Kaula's first-order theory to a Hill orbit and

showed how gravity field coefficients can easily be interpreted as the 2D Fourier spectrum of a function on a torus, reducing the gravity field inversion to a very simple block-diagonal normal equation matrix. (Goldstein 1984) was the first to introduce complex inclination functions and (Masters and Richards-Dinger 1998) proved to be about twice as efficient as (Goldstein 1984) and provide results which agree to one part in 10^{15} up to harmonic degree 256 (Masters and Richards-Dinger 1998).

Furthermore, the rotation of spherical harmonics has been used by (Balmino and Borderies 1978) to expand the gravitational potential in terms of harmonic coefficients relative to the axis of rotation of a rotating solid body. (Kleusberg 1980) derived an approximation of spherical harmonic rotation valid for small rotations that were used recently by (Kotsakis). Complete transformations of spherical harmonics, including translations and rotations were developed in (Giacaglia and Burša 1980) using Clebsch-Gordan coefficients. However, following (Goldstein 1984), due to the complexity of the general transformation formula and numerical instability in the propagation of the transformation coefficients, this has been used only for low degree ($<$10) expansions. Although (Goldstein 1984) presented the rotation of spherical harmonics with expansions up to degree 180, the mathematical apparatus is very complex, and numerically and computationally extensive. Generally speaking, all algorithms for spherical harmonic rotations are based on the recursions starting with Wigner matrices of degree one, or the actual rotation matrix of the coordinate frame rotation. The problem is that these recursions are instable in themselves, as frequently reported in the associated literature.

The gravitational potential in terms of real spherical harmonics reads as (Heiskanen and Moritz 1967)

$$V(r,\theta,\lambda) = \frac{GM}{r}\left[1 + \sum_{n=2}^{\infty}\left(\frac{a}{r}\right)^n \sum_{m=-n}^{n}\left(C_{nm}\cos m\lambda + S_{nm}\sin m\lambda\right)P_{nm}(\cos\theta)\right] \tag{26.14}$$

where C_{nm} and S_{nm} represent unnormalized spherical harmonic coefficients (SH). Typically, SH are normalized, employing the normalization function (27.28) in order to obtain the normalized associated Legendre functions

$$\bar{P}_{nm} = N_{nm}\cdot P_{nm} \tag{26.15}$$

and the normalized spherical harmonic coefficients $\bar{C}_{nm}$ and $\bar{S}_{nm}$ using

$$C_{nm} = N_{nm}\cdot\bar{C}_{nm}, \qquad S_{nm} = N_{nm}\cdot\bar{S}_{nm}. \tag{26.16}$$

By introducing

$$c_{nm} = \begin{cases} C_{nm}, & m \geq 0 \\ S_{n|m|}, & m < 0 \end{cases} \tag{26.17}$$

we obtain

$$V(r,\theta,\lambda) = \frac{GM}{r}\left[1 + \sum_{n=2}^{\infty}\left(\frac{a}{r}\right)^n \sum_{m=-n}^{n} c_{nm}Y_{nm}(\theta,\lambda)\right] \tag{26.18}$$

or the general case, assuming an arbitrary position of the center of gravitation with respect to the figure axis of the Earth

$$V(r,\theta,\lambda) = \frac{GM}{r}\sum_{n=0}^{\infty}\left(\frac{a}{r}\right)^n \sum_{m=-n}^{n} c_{nm}Y_{nm}(\theta,\lambda) \tag{26.19}$$

or shortened using the general form with the harmonic $V_{nm}(r,\theta,\lambda)$ similar to (26.9)

$$V(r,\theta,\lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} V_{nm}(r,\theta,\lambda) \,. \tag{26.20}$$

We can identify two scaling factors in (26.19), namely a geometrical scale a and dynamical scale GM that refer to the size of the central term V_{00} of the underlying gravitational field represented by a sphere with the radius a and the geocentric gravitational constant GM

$$V_{00}(a,\theta,\lambda) = \frac{GM}{a} \,. \tag{26.21}$$

The real spherical harmonic functions can be further written as

$$Y_{nm}(\theta,\lambda) = P_{n|m|}\big(\cos(\theta)\big) \begin{Bmatrix} \cos m\lambda, & m \geq 0 \\ \sin|m|\lambda, & m < 0 \end{Bmatrix} \,. \tag{26.22}$$

In the complex notation spherical harmonic functions $Y_{nm}(\theta,\lambda)$ read as

$$Y_{nm}(\theta,\lambda) = e^{im\lambda} P_{nm}\big(\cos(\theta)\big) \tag{26.23}$$

By the rotation of spherical harmonics we find a new set of spherical harmonic coefficients $\{\kappa_{nm}\}$ representing the rotated gravitational potential

$$V\big(\Re(\vec{r}_0)\big) = \frac{GM}{r_0} \left[1 + \sum_{n=2}^{\infty} \left(\frac{a}{r_0}\right)^n \sum_{m=-n}^{n} \kappa_{nm} Y_{nm}(\vec{r}_0) \right] \tag{26.24}$$

where $\Re$ denotes the rotation matrix applied to the initial position $\vec{r}_0$ and κ_{nm} a set of coefficients as a function of the rotated position

$$\kappa_{nm} = \kappa_{nm}\big(\Re(\vec{r}_0)\big) \,. \tag{26.25}$$

The gravitational potential for a specific degree n can be exactly represented by the rotated set of coefficients ν_n obtained from the rotation of the initial spherical harmonic coefficients Θ_n. This rotation can be carried out as a simple linear transformation for a specific degree n

$$\nu_n = R_n \Theta_n \tag{26.26}$$

considering all SH coefficients of the same degree n. In the general case, including SH coefficients of all degrees, rotation matrix $R_{(n,n)}$ is a block-sparse rotation matrix with a dimension $n \times n$

$$\nu = R_{(n,n)} \Theta \tag{26.27}$$

with blocks R_n on the main diagonal, see Figure 26.1. The Θ denotes a vector of spherical harmonic coefficients over all degrees. Rotation matrices in space with an arbitrary dimension or so-called Wigner D functions are the matrix representation of the rotation operators on the basis of spherical harmonics. The rotation matrix of the first degree, i.e., R_1 is a rotation matrix in the Euclidian space and in spherical harmonic space at the same time

$$\vec{r} = R_1 \vec{r}(t_0), \quad \nu_1 = R_1 \Theta_1 \quad \text{with } \Theta_1 = \big\{ S_{11}, C_{10}, C_{11} \big\} \tag{26.28}$$

A graphical representation of the structure of the rotation matrix of the spherical harmonic coefficients for every degree n can be seen in Figure 26.1.

$$
R_{(n,n)} = \begin{bmatrix}
\overset{C_{00}}{1} & \overset{S_{11}}{0} & \overset{C_{10}}{} & \overset{C_{11}}{} & & \overset{S_{22}}{} & \overset{S_{21}}{} & \overset{C_{20}}{} & \overset{C_{21}}{} & \overset{C_{22}}{} & \cdots \\
 & & R_1 & & & & & & & & \\
 & \times & & \times & & & & & & & \cdots \\
0 & & 1 & & & & 0 & & & & \cdots \\
 & \times & & \times & & & & & & & \cdots \\
 & & & & R_2 & & & & & & \\
 & & & \times & & & & \times & & & \cdots \\
 & & & & \times & & & \times & & & \cdots \\
 & & 0 & & & & 1 & & & & \cdots \\
 & & & \times & & & & \times & & & \cdots \\
 & & & \times & & & & \times & & & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots
\end{bmatrix}
= \begin{bmatrix}
1 & & & & \\
 & R_1 & & 0 & \\
 & & R_2 & & \\
 & 0 & & \ddots & \\
 & & & & R_n
\end{bmatrix}
$$

Figure 26.1 Structure of the spherical harmonic rotation matrix for a rotation about polar or equatorial axis. The symbol $\times$ denotes the populated elements of the rotation sub-matrices.

Apart from the rotation of surface spherical harmonics in (26.24), the general case in the transformation of spherical harmonics includes an arbitrary scale s defined as

$$
s = \frac{r}{r_0} \tag{26.29}
$$

Thus we have

$$
V\left(s \cdot \Re(\vec{r}_0)\right) = \frac{GM}{s \cdot r_0}\left[1 + \sum_{n=2}^{\infty}\left(\frac{a}{s \cdot r_0}\right)^n \sum_{m=-n}^{n} \kappa_{nm} Y_{nm}(\vec{r}_0)\right] \tag{26.30}
$$

where the original set of coefficients is re-scaled per-degree

$$
V\left(s \cdot \Re(\vec{r}_0)\right) = \frac{GM}{r}\left[1 + \sum_{n=2}^{\infty}\left(\frac{r_0}{r}\right)^n \sum_{m=-n}^{n} \hat{\kappa}_{nm} Y_{nm}(\vec{r}_0)\right] \tag{26.31}
$$

with

$$
\hat{\kappa}_{nm} = \left(\frac{a}{r_0}\right)^n \kappa_{nm} \tag{26.32}
$$

Generally speaking, such a scale can be defined per degree of the spherical harmonic expansion and is reflected in the degree variance that is invariant under rotation. In the case of GRACE gravity field maps, Figure 26.2 shows that the variation of the degree variance is very uniform in amplitude, i.e., the amplitude is $<5\permil$ for $n \leq 60$ and $<0.5\permil$ for $n < 30$. From this one can draw the conclusion that temporal gravity field maps can be parameterized by rotations of spherical harmonics, as depicted in Figure 26.3. One could model temporal gravity field variations between the gravity Field A and the gravity Field B (e.g., two successive monthly gravity fields) with continuous rotations of spherical harmonics about the polar axis. This is possible since the pole coordinates estimates from the coefficients $\bar{C}_{21}$ and $\bar{S}_{21}$ of the GRACE monthly gravity fields follow the
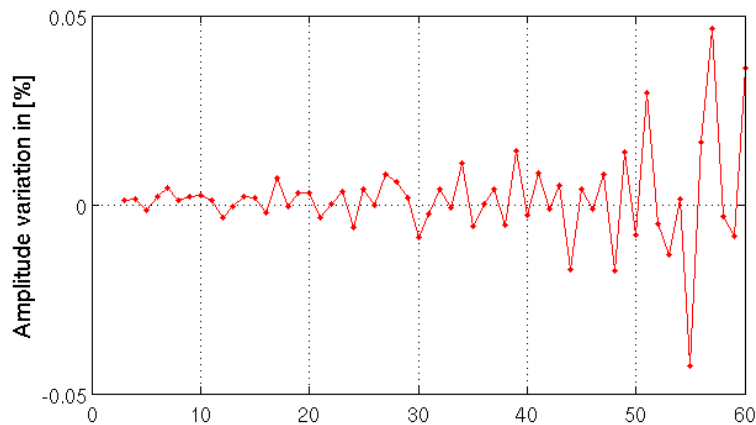
Figure 26.2 Variation of the degree variance (amplitude) vs. degree of spherical harmonics expansion from GRACE monthly gravity fields (RL4). One can see that degree variance is very uniform in amplitude $<5‰$ for $n \leq 60$ and $<0.5‰$ for $n < 30$.

conventional IERS pole. In this way the rotation of spherical harmonic coefficients provides a continuous parameterization of the temporal variations. Generally speaking, one could use a different rotation axis for each degree of SH expansion, however, the use of a polar axis is more efficient in this case. This topic and rotation about the polar axis is discussed later in this section. Coefficients of the GRACE gravity monthly fields follow the conventional IERS pole. Rotation provides continuous parameterization of spherical harmonic coefficients.
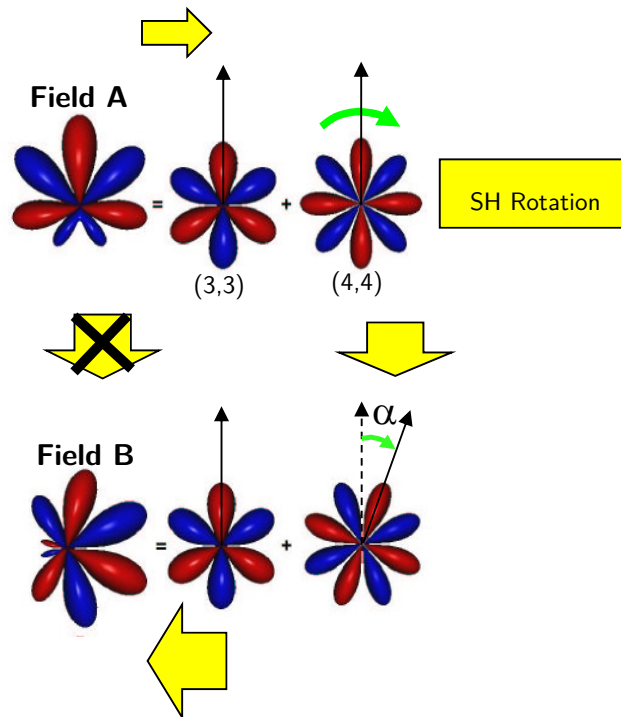
Figure 26.3 Schematic description of modeling temporal gravity field variations with rotation of spherical harmonics. Instead of having no physical connection between gravity Field A and gravity Field B (e.g., two successive monthly gravity fields) one could model temporal gravity field variations with continuous rotations of spherical harmonics about the polar axis. This is possible since pole coordinates estimates from $\bar{C}_{21}$ and $\bar{S}_{21}$ of the GRACE monthly gravity fields follow the conventional IERS pole.

26.3.1 Geometrical Rotation of Spherical Harmonics About the Polar Axis

A SH rotation about the polar axis is very simple and can be derived using simple trigonometric addition theorem. Let us introduce surface spherical harmonics of degree n

$$\boldsymbol{p}_n(\theta)^T \boldsymbol{\Lambda}_n(\lambda) := \sum_{m=-n}^{n} c_{nm} Y_{nm}(\theta, \lambda) \tag{26.33}$$

with $\boldsymbol{p}_n(\theta)$ denoting the vector of Legendre functions for a specific degree n and vector $\boldsymbol{\Lambda}_n(\lambda)$ given as

$$\boldsymbol{\Lambda}_n(\lambda) := \begin{bmatrix} S_{n,n} \sin n\lambda \\ \dots \\ S_{n,1} \sin \lambda \\ C_{n0} \\ C_{n,1} \cos \lambda \\ \dots \\ C_{n,n} \cos n\lambda \end{bmatrix} \tag{26.34}$$

thus obtaining the gravitational potential in the form

$$V(r, \theta, \lambda) = \frac{GM}{r} \left[1 + \sum_{n=2}^{\infty} \left(\frac{a}{r} \right)^n \boldsymbol{p}_n(\theta)^T \boldsymbol{\Lambda}_n(\lambda) \right] \tag{26.35}$$

If we now apply trigonometric addition theorem to (26.34) we derive

$$\boldsymbol{\Lambda}_n(\lambda + \alpha) := \begin{bmatrix} S_{n,n} \sin n\lambda \cos n\alpha \\ \dots \\ S_{n,1} \sin \lambda \cos \alpha \\ C_{n0} \\ C_{n,1} \cos \lambda \cos \alpha \\ \dots \\ C_{n,n} \cos n\lambda \cos n\alpha \end{bmatrix} + \begin{bmatrix} S_{n,n} \cos n\lambda \sin n\alpha \\ \dots \\ S_{n,1} \cos \lambda \sin \alpha \\ 0 \\ -C_{n,1} \sin \lambda \sin \alpha \\ \dots \\ -C_{n,n} \sin n\lambda \sin n\alpha \end{bmatrix} \tag{26.36}$$

Because of the symmetry in spherical harmonics, the associated Legendre function appears twice for the same order and hence the sine and cosine terms in (26.36) can be written in the following way

$$\boldsymbol{\Lambda}_n(\lambda + \alpha) := \begin{bmatrix} S_{n,n} \sin n\lambda \cos n\alpha \\ \dots \\ S_{n,1} \sin \lambda \cos \alpha \\ C_{n0} \cos 0\alpha \\ C_{n,1} \cos \lambda \cos \alpha \\ \dots \\ C_{n,n} \cos n\lambda \cos n\alpha \end{bmatrix} + \begin{bmatrix} -C_{n,n} \sin n\lambda \sin n\alpha \\ \dots \\ -C_{n,1} \sin \lambda \sin \alpha \\ C_{n0} \sin 0\alpha \\ S_{n,1} \cos \lambda \sin \alpha \\ \dots \\ S_{n,n} \cos n\lambda \sin n\alpha \end{bmatrix}. \tag{26.37}$$

Separating out the starting vector $\boldsymbol{\Lambda}_n(\lambda)$ in the form of a diagonal matrix

$$\mathbf{\Lambda}_n(\lambda + \alpha) = \begin{bmatrix} \sin n\lambda & & & & & & \\ & \cdots & & & & 0 & \\ & & \sin\lambda & & & & \\ & & & 1 & & & \\ & & & & \cos\lambda & & \\ & 0 & & & & \cdots & \\ & & & & & & \cos n\lambda \end{bmatrix} \widehat{\mathbf{\Lambda}}_n(\alpha). \tag{26.38}$$

Since $\sin n\lambda$ and $\cos n\lambda$ appear in both terms, (26.37) can be reduced to the argument α only

$$\widehat{\mathbf{\Lambda}}_n(\alpha) := \begin{bmatrix} S_{n,n}\cos n\alpha \\ \cdots \\ S_{n,1}\cos\alpha \\ C_{n0}\cos 0\alpha \\ C_{n,1}\cos\alpha \\ \cdots \\ C_{n,n}\cos n\alpha \end{bmatrix} + \begin{bmatrix} -C_{n,n}\sin n\alpha \\ \cdots \\ -C_{n,1}\sin\alpha \\ C_{n0}\sin 0\alpha \\ S_{n,1}\sin\alpha \\ \cdots \\ S_{n,n}\sin n\alpha \end{bmatrix} \tag{26.39}$$

As an example for the degree $n = 3$ we obtain

$$\widehat{\mathbf{\Lambda}}_3(\alpha) = \begin{bmatrix} S_{3,3}\cos 3\alpha \\ S_{3,2}\cos 2\alpha \\ S_{3,1}\cos\alpha \\ C_{30}\cos 0\alpha \\ C_{3,1}\cos\alpha \\ C_{3,2}\cos 2\alpha \\ C_{3,3}\cos 3\alpha \end{bmatrix} + \begin{bmatrix} -C_{3,3}\sin 3\alpha \\ -C_{3,2}\sin 2\alpha \\ -C_{3,1}\sin\alpha \\ C_{30}\sin 0\alpha \\ S_{3,1}\sin\alpha \\ S_{3,2}\sin 2\alpha \\ S_{3,3}\sin 3\alpha \end{bmatrix}. \tag{26.40}$$

We see that the rotation of spherical harmonic coefficients $\widehat{\Lambda}_n$ about the polar axis is very simple and, based on (26.37), rotated coefficients $\kappa_n(\alpha)$ can be defined using the orthonormal rotation matrix $\mathrm{K}_n(\alpha)$

$$\kappa_n(\alpha) := \begin{bmatrix} \cos n\alpha & & & & & & \sin n\alpha \\ & \ddots & & & & \cdot^{\cdot} & \\ & & \cos\alpha & & \sin\alpha & & \\ & & & 1 & & & \\ & & -\sin\alpha & & \cos\alpha & & \\ & \cdot^{\cdot} & & & & \ddots & \\ -\sin n\alpha & & & & & & \cos n\alpha \end{bmatrix}^T \begin{bmatrix} S_{nn} \\ \vdots \\ S_{n1} \\ C_{n0} \\ C_{n1} \\ \vdots \\ C_{nn} \end{bmatrix} \tag{26.41}$$

or shortened

$$\kappa_n(\alpha) := \mathrm{K}_n(\alpha)\widehat{\Lambda}_n \tag{26.42}$$

After performing QR decomposition of the matrix $\mathrm{K}_n = \mathrm{K}_n(\alpha)$, we obtain a new orthogonal matrix $\widetilde{\mathrm{K}}_n$ and the right triangular matrix $\widetilde{\mathrm{I}}_n$

$$\mathrm{K}_n := \overset{\sim}{\mathrm{K}}_n \overset{\sim}{\mathrm{I}}_n \tag{26.43}$$

$$\mathrm{K}_n(\alpha) := \begin{bmatrix} -\cos n\alpha & & & & & & & \sin n\alpha \\ & \ddots & & & & & \cdot^{\cdot^{\cdot}} & \\ & & -\cos\alpha & & \sin\alpha & & & \\ & & & 1 & & & & \\ & & \sin\alpha & & \cos\alpha & & & \\ & \cdot^{\cdot^{\cdot}} & & & & \ddots & & \\ \sin n\alpha & & & & & & & \cos n\alpha \end{bmatrix} \begin{bmatrix} -1 & & & & & \\ & \ddots & & & & \\ & & -1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \tag{26.44}$$

It can be shown that both matrices are Hermitian matrices. A Hermitian matrix is a square matrix with complex entries which is equal to its own conjugate transpose, i.e., the i-th row and j-th column is equal to the complex conjugate. However, in our case both matrices are real without complex or conjugate complex parts and it can be shown that they are at the same time unitary matrices. A matrix is unitary if and only if it has an inverse which is equal to its conjugate transpose, or as in our case, with all elements real numbers

$$\overset{\sim}{\mathrm{K}}_n \cdot \overset{\sim}{\mathrm{K}}_n = \mathrm{I}, \qquad \overset{\sim}{\mathrm{I}}_n \cdot \overset{\sim}{\mathrm{I}}_n = \mathrm{I} \tag{26.45}$$

the inverse of the matrix is equal to the original matrix. From isometry, it follows that all eigenvalues of a unitary matrix are complex numbers of absolute value 1, i.e., they lie on the unit circle centered at 0 in the complex plane. Or in other words, QR decomposition transforms our rotation matrix into an unitary matrix, a normal matrix with eigenvalues lying on the unit circle.

We see that QR decomposition of the rotation matrix K_n decomposes the matrix into two reflection matrices $\overset{\sim}{\mathrm{K}}_n, \overset{\sim}{\mathrm{I}}_n$. Rotation matrices have $\det(\mathrm{K}_n) = 1$ and to obtain $\det(\overset{\sim}{\mathrm{K}}_n) = \det(\overset{\sim}{\mathrm{I}}_n) = 1$, we simply change the sign of the central element of both matrices

$$\mathrm{K}_n(\alpha) := \begin{bmatrix} -\cos n\alpha & & & & & & & \sin n\alpha \\ & \ddots & & & & & \cdot^{\cdot^{\cdot}} & \\ & & -\cos\alpha & & \sin\alpha & & & \\ & & & -1 & & & & \\ & & \sin\alpha & & \cos\alpha & & & \\ & \cdot^{\cdot^{\cdot}} & & & & \ddots & & \\ \sin n\alpha & & & & & & & \cos n\alpha \end{bmatrix} \begin{bmatrix} -1 & & & & & \\ & \ddots & & & & \\ & & -1 & & & \\ & & & -1 & & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \tag{26.46}$$

which reduces to

$$\kappa_n(\alpha) := \begin{bmatrix} -\cos n\alpha & & & & & & & \sin n\alpha \\ & \ddots & & & & & \cdot^{\cdot^{\cdot}} & \\ & & -\cos\alpha & & \sin\alpha & & & \\ & & & -1 & & & & \\ & & \sin\alpha & & \cos\alpha & & & \\ & \cdot^{\cdot^{\cdot}} & & & & \ddots & & \\ \sin n\alpha & & & & & & & \cos n\alpha \end{bmatrix} \begin{bmatrix} -S_{nn} \\ \vdots \\ -S_{n1} \\ -C_{n0} \\ C_{n1} \\ \vdots \\ C_{nn} \end{bmatrix} \tag{26.47}$$

Comparing (26.47) with (26.41) we see that with QR decomposition it is possible to transform a rotation matrix into a reflection matrix.

Let us now perform a rotation only within the same order m

$$\kappa_m(\alpha) := \begin{bmatrix} \cos m\alpha & & & & & & \sin m\alpha \\ & \ddots & & & & \cdot^{\cdot^{\cdot}} & \\ & & \cos m\alpha & \sin m\alpha & & & \\ & & -\sin m\alpha & \cos m\alpha & & & \\ & \cdot^{\cdot^{\cdot}} & & & & \ddots & \\ -\sin m\alpha & & & & & & \cos m\alpha \end{bmatrix} \begin{bmatrix} S_{nm} \\ \vdots \\ S_{mm} \\ C_{mm} \\ \vdots \\ C_{nm} \end{bmatrix} \tag{26.48}$$

This leads us to the final expression for a spherical harmonic rotation about the polar axis for the particular order m

$$\begin{bmatrix} S_{n_{\max}m}(\alpha) \\ \dots \\ S_{mm}(\alpha) \\ C_{mm}(\alpha) \\ \dots \\ C_{n_{\max}m}(\alpha) \end{bmatrix} := \begin{bmatrix} S_{n_{\max}m} \\ \dots \\ S_{mm} \\ C_{mm} \\ \dots \\ C_{n_{\max}m} \end{bmatrix} \cos m\alpha + \begin{bmatrix} C_{n_{\max}m} \\ \dots \\ C_{mm} \\ -S_{mm} \\ \dots \\ -S_{n_{\max}m} \end{bmatrix} \sin m\alpha \tag{26.49}$$

with $n_{\max}$ denoting the maximum degree in the spherical harmonics expansion. (26.49) can be written in a very short form

$$\widehat{\Lambda}_m(\alpha) := \Lambda_m \cos m\alpha + \Lambda_m^* \sin m\alpha \tag{26.50}$$

where Λ_m^* corresponds to a form of conjugate transpose of the original vector Λ_m. Both vectors, Λ_m and Λ_m^* have a very nice orthonormal property: they are orthogonal and of the same length

$$\Lambda_m \perp \Lambda_m^* \qquad \left\|\Lambda_m\right\| = \left\|\Lambda_m^*\right\| \tag{26.51}$$

Special cases of (26.50) include for $\alpha = 0$

$$\widehat{\Lambda}_m(\alpha = 0) = \Lambda_m \tag{26.52}$$

and for $\alpha = \dfrac{\pi}{2}$

$$\begin{aligned} \widehat{\Lambda}_m\left(\alpha = \frac{\pi}{2}\right) &= (-1)^{(m-1)/2}\,\Lambda_m^* \qquad && \forall m = odd \\ \widehat{\Lambda}_m\left(\alpha = \frac{\pi}{2}\right) &= (-1)^{m/2}\,\Lambda_m \qquad && \forall m = even \end{aligned} \tag{26.53}$$

which corresponds to a rotation about the polar axis by $\pi/2$.

A schematic ordering of (26.49) for a fixed order m is shown in Figure 26.4, i.e., the specific order m is kept fixed and the degree index n runs over all degrees, up to $n_{\max}$.

$$
\begin{array}{ccccccccccc}
& & & & & C_{00} & & & & & \\
& & & & \boxed{S_{11}} & C_{10} & \boxed{C_{11}} & & & & \\
& & & S_{22} & S_{21} & C_{20} & C_{21} & C_{22} & & & \\
& & S_{33} & S_{32} & \downarrow & C_{30} & \downarrow & C_{32} & C_{33} & & \\
& & & & S_{31} & & C_{31} & & & & \\
& S_{44} & S_{43} & S_{42} & S_{41} & C_{40} & C_{41} & C_{42} & C_{43} & C_{44} & \\
S_{55} & S_{54} & S_{53} & S_{52} & \boxed{S_{51}} & C_{50} & \boxed{C_{51}} & C_{52} & C_{53} & C_{54} & C_{55}
\end{array}
$$

Figure 26.4 Ordering of spherical harmonic coefficients for a geometrical rotation about the polar axis.

26.3.2 Conventional Sequence for the Rotation of Spherical Harmonics About an Arbitrary Axis

Here we describe the conventional sequence in the rotation of spherical harmonics one can often find in literature, and in particular the case represented by an orbital plane, see e.g., (Sneeuw 2000). We first decompose the rotation matrix R_1 into the zxz orthogonal rotation sequence, i.e., represented by two rotations about the polar z axis (angles Λ and u) and one rotation about an arbitrary equatorial x axis (angle i). Furthermore, the rotation about the x axis is decomposed into the yzy rotation sequence, where the rotation about the y axis is represented by a pre- and post-rotation by $\pi/2$, whereas a rotation about the z axis by an angle i. The complete $zyzyz$ orthogonal rotation sequence for the x axis is then

$$
R_1 = R_z(u)R_y\left(-\frac{1}{2}\pi\right)R_z(-i)R_y\left(\frac{1}{2}\pi\right)R_z\left(\Lambda\right) \tag{26.54}
$$

In the case of an orbit represented by the Keplerian parameters, the x axis (Greenwich meridian) is, after the first rotation matrix with angle Λ, aligned with the line of nodes and pointing towards the ascending node of the orbit. The rotation about the x axis is decomposed into the yzy sequence and tilts the orbital plane by the inclination i from the equatorial plane. The last, fifth rotation about the rotated z axis, describes the rotation of the x axis in the orbital plane by the argument of latitude u. The yzy rotation sequence is decomposed into pre- and post-rotation of the x axis by $\pi/2$, meaning that rotation by the inclination angle is, in the end, carried out around the z axis and not around the x or y axis. The rotation of spherical harmonics about the x axis is in this way carried out about the z axis. The $zyzyz$ orthogonal rotation sequence allows us to limit the rotation about the y axis to merely a rotation between the equator and the Poles.

One can easily see a very nice advantage of spherical harmonic rotation by means of this approach: the transformation between equator and Pole or the tilt of the orbital plane from the equator by an inclination angle needs to be calculated only once and the values are valid for all points along an orbit. The final rotation matrix R_n, for the spherical harmonic coefficients of degree n is then

$$
R_n = R_{n(z)}(u)R_{n(y)}\left(-\frac{1}{2}\pi\right)R_{n(z)}(-i)R_{n(y)}\left(\frac{1}{2}\pi\right)R_{n(z)}\left(\Lambda\right) \tag{26.55}
$$

The main drawback of this approach is in the calculation of the Wigner matrices that is still challenging from the numerical point of view. Figure 26.5 shows a graphical representation of the spherical harmonic rotation of the second zonal harmonic tilted by $30°$ from the equatorial plane in the Euclidian space.

Figure 26.5 Graphical representation of the second zonal harmonic and it's rotated version tilted by $30°$ from the equatorial plane after rotation about an arbitrary equatorial axis.

26.4 A Fast Geometrical Approach to Calculate and Rotate Legendre Polynomials and their Derivatives to Ultra-High Degree Without Recurrence Relations

Following the Lecture Notes on Physical Geodesy (Rummel 2006), the spherical harmonics addition theorem or addition theorem for associated Legendre functions can be written as

$$P_n(\cos\gamma) = P_n(\cos\theta_1) \cdot P_n(\cos\theta_2) + 2\sum_{m=1}^{n} \frac{(n-m)!}{(n+m)!} \cdot P_n^m(\cos\theta_1) \cdot P_n^m(\cos\theta_2) \cdot \cos\left[m(\lambda_1 - \lambda_2)\right] \quad (26.56)$$

where the spherical distance γ between the two points on the sphere can be calculated from spherical coordinates θ and λ making use of the spherical law of cosine

$$\cos\gamma = \cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2 \cos(\lambda_1 - \lambda_2) \quad (26.57)$$

The question arises as to whether there is any geometrical representation of the addition theorem (26.56). Let us first introduce the associated Legendre functions of negative order by means of

$$P_n^{-m}(\cos\theta) = (-1)^m \frac{(n-m)!}{(n+m)!} \cdot P_n^m(\cos\theta) \quad (26.58)$$

Since

$$P_{nm}(\cos\theta) = (-1)^m P_n^m(\cos\theta) \quad (26.59)$$

the addition theorem for the associated Legendre functions can be written as

$$P_n(\cos\gamma) = P_n(\cos\theta_1) \cdot P_n(\cos\theta_2) + 2\sum_{m=1}^{n} P_{n(-m)}(\cos\theta_1) \cdot P_{nm}(\cos\theta_2) \cdot \cos\left[m(\lambda_1 - \lambda_2)\right] \quad (26.60)$$

If we now use the following vector form of the associated Legendre functions

$$\boldsymbol{p}_{nm}^{-}(\cos\theta_1) = \begin{Bmatrix} P_{n,-m}(\cos\theta_1) \\ \dots \\ P_{n0}(\cos\theta_1) \\ \dots \\ P_{n,-m}(\cos\theta_1) \end{Bmatrix} \qquad \boldsymbol{p}_{nm}(\cos\theta_2) = \begin{Bmatrix} P_{nm}(\cos\theta_2) \\ \dots \\ P_{n0}(\cos\theta_2) \\ \dots \\ P_{nm}(\cos\theta_2) \end{Bmatrix} \qquad (26.61)$$

we may write the addition theorem of the associated Legendre functions as a scalar product in Hilbert space as

$$P_n(\cos\gamma) = \boldsymbol{p}_{nm}^{-T}(\cos\theta_1)\cdot\boldsymbol{p}_{nm}(\cos\theta_2) \qquad (26.62)$$

where $\gamma = \theta_1 - \theta_2$. For an equatorial arc of the same length $\gamma = \theta_1 - \theta_2 = \theta = \lambda_1 - \lambda_2 = \Delta\lambda$, we obtain

$$P_n(\cos\theta) = P_n(\cos\Delta\lambda) = \boldsymbol{p}_{nm}^{-T}(0)\begin{bmatrix} \cos m\Delta\lambda & & & \\ & \dots & & \\ & & 1 & \\ & & & \dots & \\ & & & & \cos m\Delta\lambda \end{bmatrix}\boldsymbol{p}_{nm}(0) \qquad (26.63)$$

And by denoting the middle cos-matrix by $C(\Delta\lambda)$ we may write in shortened form

$$P_n(\cos\theta) = P_n(\cos\Delta\lambda) = \boldsymbol{p}_{nm}^{-T}(0)\cdot C(\Delta\lambda)\cdot\boldsymbol{p}_{nm}(0) \qquad (26.64)$$

Since $\boldsymbol{p}_{nm}(0)$ can easily be pre-calculated for all arguments, see, e.g., Figure 26.6, we may use (26.63) for a very elegant calculation of the Legendre polynomials and their derivatives. The first derivative is merely a function of λ measured along the equator

$$\frac{dP_n(\cos\theta)}{d\theta} = -P_{n1}(\cos\Delta\lambda) = \boldsymbol{p}_{nm}^{-T}(0)\begin{bmatrix} -m\sin m\Delta\lambda & & & \\ & \dots & & \\ & & 0 & \\ & & & \dots & \\ & & & & -m\sin m\Delta\lambda \end{bmatrix}\boldsymbol{p}_{nm}(0) \qquad (26.65)$$

as long as $\theta = \Delta\lambda$. The properties of the first-order associated Legendre functions $P_{n,1} = -dP_{n0}/d\theta$ will be derived in the next section. The second derivative is accordingly

$$\frac{d^2 P_n(\cos\theta)}{d\theta^2} = -\frac{dP_{n1}(\cos\Delta\lambda)}{d\theta} = \boldsymbol{p}_{nm}^{-T}(0)\begin{bmatrix} -m^2\cos m\Delta\lambda & & & \\ & \dots & & \\ & & 0 & \\ & & & \dots & \\ & & & & -m^2\cos m\Delta\lambda \end{bmatrix}\boldsymbol{p}_{nm}(0) \qquad (26.66)$$

or in shortened form

$$\frac{d^2 P_n(\cos\theta)}{d\theta^2} = -\frac{dP_{n1}(\cos\Delta\lambda)}{d\theta} = \boldsymbol{p}_{nm}^{-T}(0)\left(-m^2 C(\Delta\lambda)\right)\boldsymbol{p}_{nm}(0) \qquad (26.67)$$

where (26.64) is multiplied by $-m^2$ on the main diagonal.

In the theory of least squares adjustment, the matrix multiplication in (26.65) is known as the bilinear form, with two vectors $\boldsymbol{y}$ and $\boldsymbol{x}$ in a multiplication with a matrix A giving the scalar denoted here as u
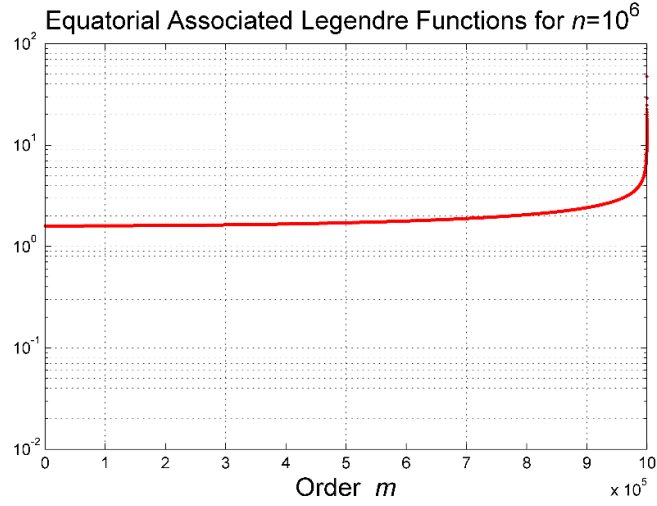
Figure 26.6 Equatorial fully normalized associated Legendre functions for $n = 10^6$. One can see that the magnitude is very moderate at the equator, thus a rotation can be performed along the equator and used for the calculation of associated Legendre functions at an arbitrary location.

such that

$$\boldsymbol{y}^T \cdot A \cdot \boldsymbol{x} = u \tag{26.68}$$

However, since in our case we may directly multiply the diagonal elements of matrix A by $\dfrac{(n-m)!}{(n+m)!}$, instead of multiplying them by $\boldsymbol{p}_{nm}$ in (26.58), the bilinear form (26.68) reduces to the so-called quadratic form

$$P_n(\cos\theta) = P_n(\cos\Delta\lambda) = \boldsymbol{p}_{nm}{}^T(0) \cdot A_n \cdot \boldsymbol{p}_{nm}(0) \tag{26.69}$$

or

$$\boldsymbol{x}^T \cdot A \cdot \boldsymbol{x} = u \tag{26.70}$$

where

$$A_n = A_n(\Delta\lambda) = \begin{bmatrix} \dfrac{(n-m)!}{(n+m)!}\cos m\Delta\lambda & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & \dfrac{(n-m)!}{(n+m)!}\cos m\Delta\lambda \end{bmatrix} \tag{26.71}$$

Therefore, with the spherical harmonics addition theorem we may very elegantly calculate the Legendre polynomials and their first and higher derivatives.

Let us now look at the full geometrical potential of this approach. Since any rotation along the meridional arc will give the same value of the Legendre polynomial for the equivalent rotation along the equator $\Re_n(\theta)$

$$P_n(\cos\theta) = \boldsymbol{p}_{nm}^-(0)^T \cdot \Re_n(\theta) \cdot \boldsymbol{p}_{nm}(0) \tag{26.72}$$

or

304

$$
P_n(\cos\theta) =
\begin{bmatrix}
P_{n,-m}(0) \\
\cdots \\
P_{n0}(0) \\
\cdots \\
P_{n,-m}(0)
\end{bmatrix}^T
\begin{bmatrix}
\cos m\theta & & & & \sin m\theta \\
& \cdots & & \cdots & \\
& & 1 & & \\
& \cdots & & \cdots & \\
-\sin m\theta & & & & \cos m\theta
\end{bmatrix}
\begin{bmatrix}
P_{nm}(0) \\
\cdots \\
P_{n0}(0) \\
\cdots \\
P_{nm}(0)
\end{bmatrix}
\tag{26.73}
$$

and we may rotate the first equatorial Legendre polynomial in (26.62) $\boldsymbol{p}_{nm}^{-}{}^{T}(0)$ by $\pi/2$ in the equatorial plane by the rotation matrix $\Re_n(\pi/2)$ and the second equatorial Legendre polynomial $\boldsymbol{p}_{nm}(0)$ by the geographical latitude $\varphi = \pi/2 - \theta$ using rotation matrix $\Re_n(\varphi)$. Both rotations in the equatorial plane will give the net rotation about the polar axis equal to the original zenith distance $\theta = \pi/2 - (\pi/2 - \theta) = \theta$. However, since the new rotation matrix $\Re_n(\pi/2 + \varphi)$ is a function of geographical latitude instead of zenith distance θ , this is equivalent to a rotation of the Legendre polynomial about an equatorial axis by $\pi/2$.

$$
P_n(\cos\theta) = \left[\Re_n(\pi/2)\boldsymbol{p}_{nm}^{-}(0)\right]^T \cdot \Re_n(\varphi)\boldsymbol{p}_{nm}(0)
\tag{26.74}
$$

$$
P_n(\cos\theta) =
\begin{bmatrix}
P_{n,-m}(0) \\
\cdots \\
P_{n0}(0) \\
\cdots \\
P_{n,-m}(0)
\end{bmatrix}^T
\begin{bmatrix}
\cos m\frac{\pi}{2} & & & & \sin m\frac{\pi}{2} \\
& \cdots & & \cdots & \\
& & 1 & & \\
& \cdots & & \cdots & \\
-\sin m\frac{\pi}{2} & & & & \cos m\frac{\pi}{2}
\end{bmatrix}^T
\begin{bmatrix}
\cos m\varphi & & & & \sin m\varphi \\
& \cdots & & \cdots & \\
& & 1 & & \\
& \cdots & & \cdots & \\
-\sin m\varphi & & & & \cos m\varphi
\end{bmatrix}
\begin{bmatrix}
P_{nm}(0) \\
\cdots \\
P_{n0}(0) \\
\cdots \\
P_{nm}(0)
\end{bmatrix}
$$
$$
\tag{26.75}
$$

that can be written as

$$
P_n(\cos\theta) = \boldsymbol{p}_{nm}^{-}(0)^T \cdot \Re_n(\pi/2)^T \Re_n(\varphi) \cdot \boldsymbol{p}_{nm}(0)
\tag{26.76}
$$

The matrix $\Re_n$ has the very nice property that it is a unitary matrix, i.e., its inverse is equal to its transpose

$$
\Re_n \cdot \Re_n^T = I
\tag{26.77}
$$

thus

$$
\Re_n(\alpha) \cdot \Re_n(-\alpha) = I
\tag{26.78}
$$

Since $\Re_n(\pi/2)^T = \Re_n(-\pi/2)$ we finally obtain

$$
P_n(\cos\theta) = \boldsymbol{p}_{nm}^{-}(0)^T \cdot \Re_n(-\pi/2)\Re_n(\varphi) \cdot \boldsymbol{p}_{nm}(0)
\tag{26.79}
$$

That is equivalent to a rotation by $\theta = -\pi/2 + \varphi = -\pi/2 + (\pi/2 - \theta) = \theta$

$$
P_n(\cos\theta) =
\begin{bmatrix}
P_{n,-m}(0) \\
\cdots \\
P_{n0}(0) \\
\cdots \\
P_{n,-m}(0)
\end{bmatrix}^T
\begin{bmatrix}
\cos m(\varphi - \frac{\pi}{2}) & & & & \sin m(\varphi - \frac{\pi}{2}) \\
& \cdots & & \cdots & \\
& & 1 & & \\
& \cdots & & \cdots & \\
-\sin m(\varphi - \frac{\pi}{2}) & & & & \cos m(\varphi - \frac{\pi}{2})
\end{bmatrix}
\begin{bmatrix}
P_{nm}(0) \\
\cdots \\
P_{n0}(0) \\
\cdots \\
P_{nm}(0)
\end{bmatrix}
\tag{26.80}
$$

If we now compare (26.80) with the rotation of the spherical harmonic coefficients about the polar axis (26.48) we can represent the geometrical rotation $\Re_n(-\pi/2)\Re_n(\varphi)\boldsymbol{p}_{nm}(0)$ in (26.80) as an orthogonal span or linear combination of two orthogonal vectors describing orthogonal geometrical rotation along the order m

$$
\begin{bmatrix} P_{n_{\max}m}(\varphi-\pi/2) \\ \cdots \\ P_{mm}(\varphi-\pi/2) \\ P_{mm}(\varphi-\pi/2) \\ \cdots \\ P_{n_{\max}m}(\varphi-\pi/2) \end{bmatrix} := \begin{bmatrix} P_{n_{\max}m}(0) \\ \cdots \\ P_{mm}(0) \\ P_{mm}(0) \\ \cdots \\ P_{n_{\max}m}(0) \end{bmatrix} \cos m\theta - \begin{bmatrix} P_{n_{\max}m}(0) \\ \cdots \\ P_{mm}(0) \\ -P_{mm}(0) \\ \cdots \\ -P_{n_{\max}m}(0) \end{bmatrix} \sin m\theta
\tag{26.81}
$$

Then (26.81) can be written in a very short orthogonal form similar to spherical harmonic coefficients (26.50)

$$
\widehat{\Lambda}_m\left(\varphi-\frac{\pi}{2}\right) := \Lambda_m \cos m\left(\varphi-\frac{\pi}{2}\right) + \Lambda_m^* \sin m\left(\varphi-\frac{\pi}{2}\right)
\tag{26.82}
$$

or by denoting $\alpha = \varphi - \pi/2$

$$
\widehat{\Lambda}_m(\alpha) := \Lambda_m \cos m\alpha + \Lambda_m^* \sin m\alpha
\tag{26.83}
$$

where Λ_m^* corresponds to a form of conjugate transpose of the original vector Λ_m with orthonormal property.

After geometrical rotation of the equatorial associated Legendre functions along the order m ("vertical rotation"), the "scalar product" along the same degree n ("horizontal") gives the Legendre polynomial of the rotation

$$
P_n(\cos\theta) = \begin{bmatrix} P_{n,-m}(0) \\ \cdots \\ P_{n0}(0) \\ \cdots \\ P_{n,-m}(0) \end{bmatrix}^T \begin{bmatrix} P_{nm}(\varphi-\pi/2) \\ \cdots \\ P_{n0}(\varphi-\pi/2) \\ \cdots \\ P_{nm}(\varphi-\pi/2) \end{bmatrix}
\tag{26.84}
$$

as depicted in Figure 26.7 in
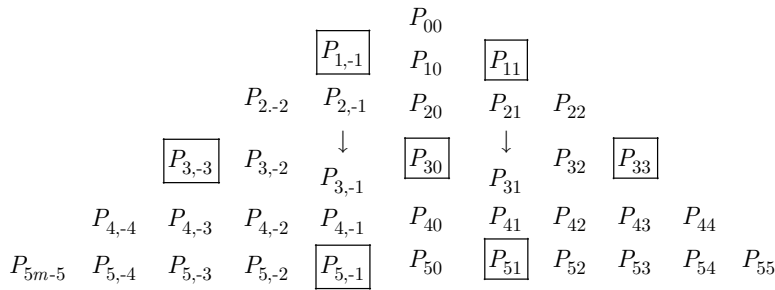
Figure 26.7 Schematic calculation of Legendre polynomial as a "scalar product" of rotated equatorial associated Legendre functions along the order m ("vertical rotation") with the one along the same degree n ("horizontal").

26.5 A Fast Geometrical Approach to Calculate First-Order and Sectorial Associated Legendre Functions

Let us now introduce the following recursion, which can be found in the excellent collection of recursion relations by (Ilk 1983)

$$2\frac{dP_{nm}}{d\theta} = (n+m)(n-m+1)P_{n,m-1} - P_{n,m+1} \tag{26.85}$$

using the associated Legendre function of negative order

$$P_{n,-m} = (-1)^m \frac{(n-m)!}{(n+m)!} P_{n,m} \tag{26.86}$$

for $m = -1$, we obtain

$$P_{n,-1} = -\frac{(n-1)!}{(n+1)!} P_{n,1} \tag{26.87}$$

and introducing (26.87) into (26.85)

$$2\frac{dP_{n0}}{d\theta} = -\left[n(n+1)\frac{(n-1)!}{(n+1)!} + 1 \right] P_{n,1} \tag{26.88}$$

since

$$\frac{(n-1)!}{(n+1)!} = \frac{1}{n(n+1)} \tag{26.89}$$

and finally the first-order associated Legendre function is

$$P_{n,1} = -\frac{dP_{n0}}{d\theta} \tag{26.90}$$

i.e., the associated Legendre function of the first-order can be calculated as the first derivative of the Legendre polynomial of the same degree. Let us now make use of the general definition of the associated Legendre functions based on Legendre polynomials, given by Rodriguez (see (Heiskanen and Moritz 1967)):

$$P_{nm}(\cos\theta) = \sin^m \theta \, \frac{d^m P_n(\cos\theta)}{d(\cos\theta)^m} \tag{26.91}$$

If we now take order $m = 1$ it follows the same property

$$P_{n1}(\cos\theta) = \sin\theta \frac{dP_n(\cos\theta)}{d(\cos\theta)} = \sin\theta \frac{dP_n(\cos\theta)}{d\theta}\frac{d\theta}{d(\cos\theta)} = -\frac{d}{d\theta}P_n(\cos\theta) \tag{26.92}$$

Finally, associated Legendre functions of order one $(m = 1)$ can be defined in terms of trigonometric expansion or rotations about the polar axis by longitude angle $\Delta\lambda = \theta$ as

$$P_{n,1}(\cos\theta) := -\frac{d}{d\theta}P_n(\cos\theta) = -P_{n1}(\cos\Delta\lambda) \tag{26.93}$$

This property (26.93) was presented for the first time in (Švehla 2008b) and later in (Svehla 2010b). A similar property can be used for sectorials $m = n$, as from (27.62) we obtain

$$2\frac{dP_{mm}}{d\theta} = (m+m)(m-m+1)P_{m,m-1} - P_{m,m+1} \tag{26.94}$$

and considering $P_{m,m+1} = 0$, we obtain

$$P_{m,m-1} = \frac{1}{m}\frac{dP_{mm}}{d\theta} \tag{26.95}$$

It is well known that sectorial associated Legendre functions P_{mm} can be calculated directly, e.g., (Hobson 1931) reads as

$$P_{mm} = \frac{(2m)!}{2^m \cdot m!}\sin^m \theta \tag{26.96}$$

or the Rodriguez formula one can also find in (Ilk 1983)

$$P_{mm} = (2m-1)!!\sin^m \theta \tag{26.97}$$

from which we derive

$$P_{m,m-1} = \frac{(2m)!}{2^m \cdot m!}\cos \theta \sin^{m-1} \theta \tag{26.98}$$

$$P_{mm} = (2m-1)!!\cos \theta \sin^{m-1} \theta \tag{26.99}$$

where !! is the double factoriel. For $n = m = 10$ the amplitude $(2m)!/(2^m \cdot m!)$ already reaches a very high value of $3.1983e+23$, thus such an approach is not an elegant method to calculate associated Legendre functions. In the next section we will see that there is a very nice symmetry between Legendre polynomials and sectorial Legendre functions.

Let us now try to express sectorial associated Legendre functions as a function of Legendre polynomials. Following (Abramowitz and Stegun 1965) we may write inverse relations such as

$$\cos \theta = P_1(\cos \theta)$$

$$\cos^4 \theta = \frac{1}{35}\Big[7P_0(\cos \theta) + 20P_2(\cos \theta) + 8P_4(\cos \theta)\Big]$$

$$\cos^2 \theta = \frac{1}{3}\Big[P_0(\cos \theta) + 2P_2(\cos \theta)\Big]$$

$$\cos^5 \theta = \frac{1}{63}\Big[27P_1(\cos \theta) + 28P_3(\cos \theta) + 8P_5(\cos \theta)\Big] \tag{26.100}$$

$$\cos^3 \theta = \frac{1}{5}\Big[3P_1(\cos \theta) + 2P_3(\cos \theta)\Big]$$

$$\cos^6 \theta = \frac{1}{231}\begin{array}{l}[33P_0(\cos \theta) + 110P_2(\cos \theta) + \\ +72P_4(\cos \theta) + 16P_6(\cos \theta)]\end{array}$$

Or, in general form

$$\cos^n \theta = \sum_{l=n,n-2,\cdots} \frac{(2l+1)n!}{2^{(n-l)/2}\left(\frac{1}{2}(n-l)\right)!(n+l+1)!!}P_l(\cos \theta) \tag{26.101}$$

from which we can derive sectorial associated Legendre functions as a function of Legendre polynomials by making use of $\cos \theta = \sin \varphi$,

$$P_{nn} = (2n-1)!! \sum_{l=n,n-2,\cdots} \frac{(2l+1)n!}{2^{(n-l)/2}\left(\frac{1}{2}(n-l)\right)!(n+l+1)!!}P_l(\cos \varphi) \tag{26.102}$$

$$P_{n,n-1} = \frac{1}{n}\frac{dP_{nn}}{d\theta} = \frac{(2n-1)!!}{n} \sum_{l=n,n-2,\cdots} \frac{(2l+1)n!}{2^{(n-l)/2}\left(\frac{1}{2}(n-l)\right)!(n+l+1)!!}\frac{dP_l(\cos \varphi)}{d\varphi} \tag{26.103}$$

which are directly related to the Legendre polynomials by rotation of equatorial Legendre polynomials about the polar axis. In the next section we will present an algorithm for the associated Legendre functions.

26.6 A Fast Geometrical Approach to Calculate Associated Legendre Functions to Ultra-High Degree and Order

Once Legendre polynomials are available, together with associated Legendre functions of the first order, one can use recursive relations to calculate the remaining associated Legendre functions. Recursions could also be used starting with the sectorial associated Legendre functions and the associated Legendre functions of the order $m = n - 1$ we derived in the previous section.

To calculate associated Legendre functions one could make direct use of the formula given by Ferrers, see, e.g., (Heiskanen and Moritz 1967)

$$P_{nm}(\cos\theta) = \sin^m\theta \frac{d^m P_n(\cos\theta)}{d(\cos\theta)^m} \tag{26.104}$$

In our case, Legendre polynomials can be represented by rotations, therefore, one would need to calculate high order derivatives of the geometrical rotations

$$P_{nm}(\cos\theta) = \sin^m\theta \frac{d^m P_n(\cos\theta)}{d(\cos\theta)^m} = \sin^m\theta \frac{d^m P_n(\cos\Delta\lambda)}{d(\cos\theta)^m} = \boldsymbol{p}_{nm}^{-T}(0) \cdot \left[\sin^m\theta \frac{d^m \mathrm{C}(\Delta\lambda)}{d(\cos\theta)^m}\right] \cdot \boldsymbol{p}_{nm}(0)$$

$$\tag{26.105}$$

It should be noted that calculation of higher order derivatives of geometrical rotations is more elegant than the calculation of higher order derivatives of Legendre polynomials. In our case we need to calculate the following terms on the main diagonal

$$\frac{d^m(\cos m\Delta\lambda)}{d(\cos\theta)^m} \tag{26.106}$$

considering that $\Delta\lambda = \theta$. To calculate higher derivatives of (26.106) we make use of the following recurrence relation

$$\cos n\theta = 2\cos\theta\cos(n-1)\theta - \cos(n-2)\theta \tag{26.107}$$

If we now take the first and second derivatives we obtain

$$\frac{d(\cos n\theta)}{d(\cos\theta)} = 2\cos(n-1)\theta + 2\cos\theta\frac{d\big(\cos(n-1)\theta\big)}{d(\cos\theta)} - \frac{d\big(\cos(n-2)\theta\big)}{d(\cos\theta)}$$

$$\frac{d^2(\cos n\theta)}{d(\cos\theta)^2} = 4\frac{d\big(\cos(n-1)\theta\big)}{d(\cos\theta)} + 2\cos\theta\frac{d^2\big(\cos(n-1)\theta\big)}{d(\cos\theta)^2} - \frac{d^2\big(\cos(n-2)\theta\big)}{d(\cos\theta)^2} \tag{26.108}$$

$$\frac{d^3(\cos n\theta)}{d(\cos\theta)^3} = 6\frac{d^2\big(\cos(n-1)\theta\big)}{d(\cos\theta)^2} + 2\cos\theta\frac{d^3\big(\cos(n-1)\theta\big)}{d(\cos\theta)^3} - \frac{d^3\big(\cos(n-2)\theta\big)}{d(\cos\theta)^3}$$

The general form of the k-th derivative of the recurrence relations (26.108) is thus

$$\frac{d^k(\cos n\theta)}{d(\cos\theta)^k} = (2k)\frac{d^{k-1}\big(\cos(n-1)\theta\big)}{d(\cos\theta)^k} + 2\cos\theta\frac{d^k\big(\cos(n-1)\theta\big)}{d(\cos\theta)^k} - \frac{d^k\big(\cos(n-2)\theta\big)}{d(\cos\theta)^k} \tag{26.109}$$

This approach allows the calculation of associated Legendre functions up to an ultra-high degree and order, see also Figure 26.8 where the algorithm is sketched in two steps.
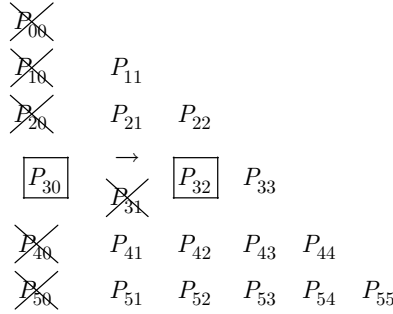
Figure 26.8 Algorithm sketch: Step 1: calculate Legendre polynomials, e.g., P_{30} using geometrical rotation along the equator. Step 2: derive associated Legendre functions, e.g P_{32}. The "$\times$" is the steps or calculation.

26.7 A Fast Geometrical Approach to Calculate Legendre Polynomials and Associated Legendre Functions at the Equator

With the addition theorem we can calculate the "scalar product" of two associated Legendre functions over the same degree. The question is whether we can calculate a similar "scalar product" over the same order.

Let us now imagine two points on a sphere with co-latitude θ_1 and θ_2. For these two points, we may write recurrence relations for the associated Legendre functions (running over the degree)

$$
\begin{aligned}
(2n+1)\cos\theta_1 P_{nm}(\cos\theta_1) &= (n-m+1)P_{n+1,m}(\cos\theta_1) + (n+m)P_{n-1,m}(\cos\theta_1) \\
(2n+1)\cos\theta_2 P_{nm}(\cos\theta_2) &= (n-m+1)P_{n+1,m}(\cos\theta_2) + (n+m)P_{n-1,m}(\cos\theta_2)
\end{aligned}
\tag{26.110}
$$

If we now multiply the upper equation by $P_{nm}(\cos\theta_2)$ and the lower by $P_{nm}(\cos\theta_1)$, and then subtract these two equations, we obtain

$$
\begin{aligned}
\big(\cos\theta_2 - \cos\theta_1\big)(2n+1)P_{nm}(\cos\theta_1)P_{nm}(\cos\theta_2) = (n-m+1)\cdot \\
\Big[P_{nm}(\cos\theta_1)P_{n+1,m}(\cos\theta_2) - P_{n+1,m}(\cos\theta_1)P_{nm}(\cos\theta_2) \Big] \\
-(n+m)\cdot \\
\Big[P_{n-1,m}(\cos\theta_1)P_{nm}(\cos\theta_2) - P_{nm}(\cos\theta_1)P_{n-1,m}(\cos\theta_2) \Big]
\end{aligned}
$$

$$\tag{26.111}$$

When added together up to degree n, for $m = 0$, we derive

$$
\begin{aligned}
\big(\cos\theta_2 - \cos\theta_1\big)\sum_{k=0}^{n}(2k+1)P_k(\cos\theta_1)P_k(\cos\theta_2) = (n+1)\cdot \\
\Big[P_n(\cos\theta_1)P_{n+1}(\cos\theta_2) - P_{n+1}(\cos\theta_1)P_n(\cos\theta_2) \Big]
\end{aligned}
\tag{26.112}
$$

If we now set $\theta_2 = 0$ and $\theta_1 = \theta$ for Legendre polynomials we obtain the very elegant expression

$$
P_{n+1}(\cos\theta) = P_n(\cos\theta) - \frac{1-\cos\theta}{n+1}\sum_{k=0}^{n}(2k+1)P_k(\cos\theta)
\tag{26.113}
$$

to calculate Legendre polynomials. For an equatorial point $\theta = \pi/2$ we obtain

$$
P_{n+1}(0) = P_n(0) - \frac{1}{n+1}\sum_{k=0}^{n}(2k+1)P_k(0)
\tag{26.114}
$$

In a similar way for associated Legendre functions we set $\theta_2 = \pi/2$ and $\theta_1 = \theta$ and obtain an elegant algorithm to calculate the "scalar product" between two associated Legendre functions of the same order

$$P_{n+1,m}(\cos\theta)P_{nm}(0) - P_{nm}(\cos\theta)P_{n+1,m}(0) = -\frac{1-\cos\theta}{(n-m+1)}\sum_{k=m}^{n}(2k+1)P_{nm}(\cos\theta)P_{nm}(0) \qquad (26.115)$$

If we now take into account the following expression from (Hobson 1931)

$$(2n+1)P_n(x) = \frac{dP_{n+1}(x)}{dx} - \frac{dP_{n-1}(x)}{dx} \qquad (26.116)$$

we obtain

$$P_{n+1}(x) - P_n(x) = -\frac{1-\cos\theta}{n+1}\sum_{k=0}^{n}\left[\frac{dP_{k+1}(x)}{dx} - \frac{dP_{k-1}(x)}{dx}\right] \qquad (26.117)$$

from where it follows

$$\frac{dP_{n+1}(\cos\theta)}{d\theta} = \frac{dP_n(\cos\theta)}{d\theta} - \frac{n+1}{1-\cos\theta}\sin\theta\left(P_{n+1}(\cos\theta) - P_n(\cos\theta)\right) \qquad (26.118)$$

Since we have already shown that $P_{n,1} = -dP_n(\cos\theta)/d\theta$ we can write

$$P_{n+1,1}(\cos\theta) - P_{n,1}(\cos\theta) = -\frac{n+1}{1-\cos\theta}\sin\theta\left(P_{n+1}(\cos\theta) - P_n(\cos\theta)\right) \qquad (26.119)$$

that for an equatorial point $\theta = \pi/2$ gives

$$P_{n+1,1}(0) - P_{n,1}(0) = -(n+1)\left(P_{n+1}(0) - P_n(0)\right) \qquad (26.120)$$

and reduces to

$$P_{n+1,1}(0) = (n+1)P_n(0) \qquad (26.121)$$

In the general case when $m \neq 0$ we obtain

$$P_{n+1,m}(0) = (n+m)P_{nm}(0) \qquad (26.122)$$

and for the derivative

$$\frac{dP_{n,m}(0)}{d\theta} = (n+m)P_{n-1,m}(0) \qquad (26.123)$$

In the general case we can derive

$$\frac{dP_{n+1}(\cos\theta)}{d\theta} = -P_{n,1}(\cos\theta) + \sin\theta\sum_{k=0}^{n}(2k+1)P_k(\cos\theta) \qquad (26.124)$$

$$\frac{dP_{n+1}(x)}{dx} - \frac{dP_n(x)}{dx} = -\sum_{k=0}^{n}(2k+1)P_k(\cos\theta) \qquad (26.125)$$

Since equatorial associated Legendre functions contain alternating zero values between consecutive orders we can make use of the following expression that is initiated with the Legendre polynomial

$$P_{n,m+1}(0) = -(n+m)(n-m+1)P_{n,m-1}(0) \qquad (26.126)$$

27. Trigonometric Representations of Legendre Functions

Although the trigonometric representation of associated Legendre functions has been considered in literature, here we give a new insight into the trigonometric reduction of Legendre polynomials. We show that Legnedre polynomials can be calculated up to an ultra-high degree, e.g., $n = 10^6$ and beyond without recursive relations and this can be used as a basis for the calculation of associated Legendre functions. The approach presented here was reported for the first time in (Švehla 2008b) and in (Svehla 2010b). In addition, we derive orthogonal geometrical forms of associated Legendre functions. However, in terms of performance, our geometrical approach based on the addition theorem of Legendre functions and geometrical rotations along the equator (previous section) is significantly more elegant.

27.1 A Slow Algorithm for the Computation of Legendre Polynomials Without Recursions Based on Trigonometric Series

Following (Heiskanen and Moritz 1967), the Legendre polynomials developed by means of recursion functions for the first low degrees are given as

$$
\begin{aligned}
P_0(t) &= 1 \\
P_1(t) &= t \\
P_2(t) &= \frac{3}{2}t^2 - \frac{1}{2} \\
P_3(t) &= \frac{5}{2}t^3 - \frac{3}{2}t \\
P_4(t) &= \frac{35}{8}t^4 - \frac{15}{4}t^2 + \frac{3}{8} \\
P_5(t) &= \frac{63}{8}t^5 - \frac{35}{4}t^3 + \frac{15}{8}t
\end{aligned}
\tag{27.1}
$$

with $t = \cos\theta$. (Heiskanen and Moritz 1967) continue with this development, expressing the powers of $\cos\theta$ in terms of the cosines of multiples of θ such as

$$
\cos^2\theta = \frac{1}{2}\cos 2\theta + \frac{1}{2} \qquad \cos^3\theta = \frac{1}{4}\cos 3\theta + \frac{3}{4}\cos\theta
\tag{27.2}
$$

and they obtain the following form of the Legendre polynomials in terms of trigonometric series

$$P_0(\cos\theta) = 1$$
$$P_1(\cos\theta) = \cos\theta$$
$$P_2(\cos\theta) = \frac{1}{4}\big(3\cos 2\theta + 1\big)$$
$$P_3(\cos\theta) = \frac{1}{8}\big(5\cos 3\theta + 3\cos\theta\big)$$
$$P_4(\cos\theta) = \frac{1}{64}\big(35\cos 4\theta + 20\cos 2\theta + 9\big)$$
$$P_5(\cos\theta) = \frac{1}{128}\big(63\cos 5\theta + 35\cos 3\theta + 30\cos\theta\big)$$

(27.3)

This particular form of the Legendre polynomials is interesting, because it is an alternative method of deriving Legendre polynomials compared to (27.1). Similar expressions for Legendre polynomials can be found in (Hobson 1931), providing two more degrees

$$P_6(\cos\theta) = \frac{1}{512}\big(231\cos 6\theta + 126\cos 4\theta + 105\cos 2\theta + 50\big)$$
$$P_7(\cos\theta) = \frac{1}{1024}\big(429\cos 7\theta + 231\cos 5\theta + 189\cos 3\theta + 175\cos\theta\big)$$

(27.4)

Unfortunately, (Heiskanen and Moritz 1967) and (Hobson 1931) developed only the first few terms of the trigonometric expansion and did not provide a general formula. Following (Hobson 1931), the first trigonometric expansion of spherical harmonics was originally given by Laplace and by Legendre by writing the distance l between two points with radius vectors r and r' using the complex domain

$$\big(1 - 2h\cos\theta + h^2\big)^{-\frac{1}{2}} = \big(1 - he^{i\theta}\big)^{-\frac{1}{2}}\big(1 - he^{-i\theta}\big)^{-\frac{1}{2}}$$

(27.5)

where the distance l and the ratio h are defined as

$$\frac{1}{l} = \frac{1}{r}\big(1 - 2h\cos\theta + h^2\big)^{-\frac{1}{2}}$$
$$h = \frac{r'}{r}$$

(27.6)

However, it was stated in (Hobson 1931) that binomial expansion is absolutely convergent and, therefore, the Cauchy product converges to the product of the sums. The Cauchy product c_n is the discrete convolution of two sequences a_n and b_n

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}$$

(27.7)

and therefore guarantees that our expansion is convergent. Now, we prove that (27.5), can be obtained by means of Euler's formula

$$e^{i\theta} = \cos(\theta) + i\sin(\theta)$$
$$e^{i\theta} + e^{-i\theta} = 2\cos\theta$$

(27.8)

and after insertion into the left-hand side of (27.5) we obtain

$$\big(1 - 2h\cos\theta + h^2\big)^{-\frac{1}{2}} = \big(1 - he^{i\theta} - he^{-i\theta} + e^0 h^2\big)^{-\frac{1}{2}} = \big(1 - he^{i\theta}\big)^{-\frac{1}{2}}\big(1 - he^{-i\theta}\big)^{-\frac{1}{2}}$$

(27.9)

In the next step, we use the famous expansion of the reciprocal distance between two points as defined in (27.5) in Legendre polynomials,

$$\left(1 - 2h\cos\theta + h^2\right)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} h^n P_n(\cos\theta) \tag{27.10}$$

Although this expansion was first reported by Legendre himself, it has a special place in Potential Theory because it is the starting point in developing the gravitational potential into a multipole expansion such as spherical harmonics. Let us now make use of the binomial theorem, but in the form generalized by Newton in terms of an indefinite series and by means of complex number z and the real exponent α, e.g., (Bronstein and Semendjajew 1996), to obtain

$$\frac{1}{(1-z)^{\alpha+1}} = \sum_{n=0}^{\infty} \binom{n+\alpha}{n} z^n \tag{27.11}$$

Applied to (27.5) for the first and second terms we obtain

$$\left(1 - he^{\theta i}\right)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} \binom{n-\frac{1}{2}}{n} h^n e^{n\theta i}, \quad \left(1 - he^{-\theta i}\right)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} \binom{n-\frac{1}{2}}{n} h^n e^{-n\theta i} \tag{27.12}$$

Using binomial coefficients in the form

$$\binom{z}{k} = \prod_{n=1}^{k} \frac{z-k+n}{n} \tag{27.13}$$

we finally derive

$$\left(1 - he^{\theta i}\right)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} \prod_{q=1}^{n} \frac{2q-1}{2q} h^n e^{n\theta i}, \quad \left(1 - he^{-\theta i}\right)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} \prod_{q=1}^{n} \frac{2q-1}{2q} h^n e^{-n\theta i} \tag{27.14}$$

or

$$\sum_{n=0}^{\infty} h^n P_n = \left(\sum_{n=0}^{\infty} \prod_{q=1}^{n} \frac{2q-1}{2q} h^n e^{n\theta i}\right)\left(\sum_{n=0}^{\infty} \prod_{q=1}^{n} \frac{2q-1}{2q} h^n e^{-n\theta i}\right)$$
$$\sum_{n=0}^{\infty} h^n P_n = \sum_{k=0}^{\infty} \sum_{s=0}^{\infty} \prod_{q=1}^{k} \frac{2q-1}{2q} \prod_{p=1}^{s} \frac{2p-1}{2p} h^{k+s} e^{(k-s)\theta i} \tag{27.15}$$

Inserting the following substitution

$$k + s = n \quad \rightarrow \quad k - s = n - 2s \tag{27.16}$$

we obtain

$$\sum_{n=0}^{\infty} h^n P_n = \sum_{n-s=0}^{\infty} h^n \prod_{q=1}^{n-s} \frac{2q-1}{2q} \sum_{s=0}^{\infty} \prod_{p=1}^{s} \frac{2p-1}{2p} e^{(n-2s)\theta i}$$
$$\sum_{n=0}^{\infty} h^n P_n = \sum_{n=0}^{\infty} h^n \prod_{q=1}^{n-s} \frac{2q-1}{2q} \sum_{s=0}^{n} \prod_{p=1}^{s} \frac{2p-1}{2p} e^{(n-2s)\theta i} \tag{27.17}$$

or the Legendre polynomials in the explicit form

$$P_n = \sum_{s=0}^{n} \prod_{q=1}^{n-s} \frac{2q-1}{2q} \prod_{p=1}^{s} \frac{2p-1}{2p} e^{(n-2s)\theta i} \tag{27.18}$$

A similar, but still slightly different approach can be found in (Sigl 1985) and (MacMillan 1930). Grouping complex conjugate terms, the Legendre polynomial for even degrees can be written as

$$P_n = 2 \sum_{s=0}^{2s<n} \prod_{q=1}^{n-s} \frac{2q-1}{2q} \prod_{p=1}^{s} \frac{2p-1}{2p} \left(e^{(n-2s)\theta i} + e^{-(n-2s)\theta i} \right) \tag{27.19}$$

from where it follows that the general expression to compute the Legendre polynomials of even degrees is

$$P_n(\cos\theta) := 2 \sum_{s=0}^{2s\leq n} \left\{ \prod_{q=1}^{n-s} \frac{2q-1}{2q} \prod_{p=1}^{s} \frac{2p-1}{2p} \right\} \cos(n-2s)\theta + \prod_{p=1}^{n/2} \frac{2p-1}{2p} \prod_{q=1}^{n/2} \frac{2q-1}{2q} \tag{27.20}$$

and for odd degrees

$$P_n(\cos\theta) := 2 \sum_{s=0}^{2s<n} \left\{ \prod_{q=1}^{n-s} \frac{2q-1}{2q} \prod_{p=1}^{s} \frac{2p-1}{2p} \right\} \cos(n-2s)\theta \tag{27.21}$$

In terms of trigonometric series, further simplifications lead to

$$P_n(\cos\theta) := \begin{cases} t_n + \displaystyle\sum_{k=0}^{n/2-1} A_{nk} \cos(n-2k)\theta & \text{for } n = even \\ \displaystyle\sum_{k=0}^{n/2} A_{nk} \cos(n-2k)\theta & \text{for } n = odd \end{cases} \tag{27.22}$$

with the amplitudes A_{nk} defined as

$$A_{nk} := 2 \prod_{q=1}^{n-k} \frac{2q-1}{2q} \cdot \prod_{p=1}^{k} \frac{2p-1}{2p} \qquad k \in \left[0, n/2\right\rangle \tag{27.23}$$

and the translation term t_n

$$t_n := \left(\prod_{q=1}^{n/2} \frac{2q-1}{2q} \right)^2 \tag{27.24}$$
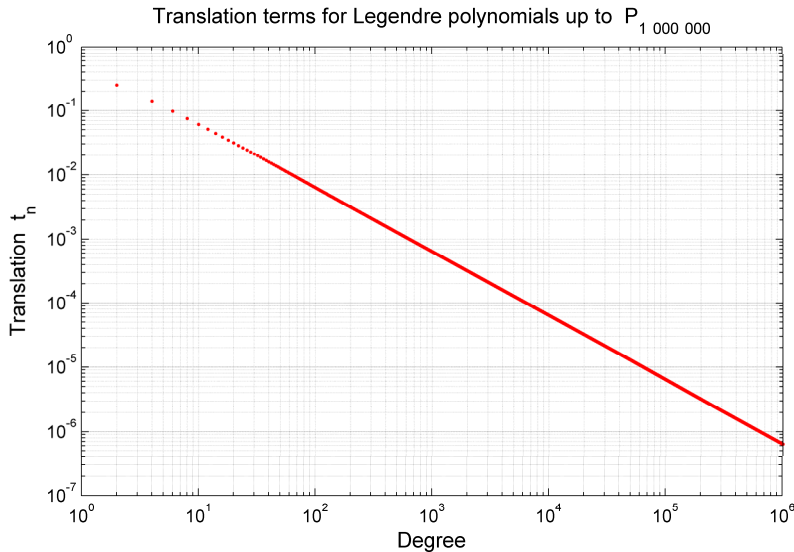
Figure 27.1 Translation terms for the Legendre polynomials up to degree $n = 1\,000\,000$.

316

From (27.22) we see that an efficient computation of Legendre polynomials reduces to the computation of amplitudes of a trigonometric series. They can be precomputed and are valid for all angular arguments. The translation term can be very efficiently computed by evaluating product by product in (27.24)

$$t_n := \left(\prod_{q=1}^{n/2} \frac{2q-1}{2q} \right)^2 = \left(\frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdot \frac{7}{8} \cdot \frac{9}{10} \cdot \ \cdots \ \frac{n-1}{n} \right)^2 \qquad (27.25)$$

In this case, extremely large nominators and denominators in (27.25) are avoided and every multiplication term has a magnitude very close to 1

$$\lim_{n \to \infty} \frac{n-1}{n} = 1. \qquad (27.26)$$

This algorithm allows stable and accurate calculation of translation terms to an ultra-high degree, e.g., $n = 1\,000\,000$. Figure 27.1 reveals that translation terms can be approximated by the following rule of thumb

$$t_n :\approx 0.6 \frac{1}{n} \qquad (27.27)$$

It is important to note that the calculated amplitudes and translation terms are moderate in size and are unnormalized.

Amplitudes A_{nk} in (27.23) calculated for the degree $n = 1\,000\,000$ are shown in Figure 27.2. Normalized Legendre polynomials $\overline{P}_n(\cos\theta)$ can easily be derived by multiplying the original Legendre polynomials by a normalization function (Heiskanen and Moritz 1967)

$$N_{nm} = \left[\frac{(n-m)!}{(n+m)!} (2n+1)(2-\delta_{0m}) \right]^{1/2} \qquad \overline{P}_{nm} = N_{nm} \cdot P_{nm} \qquad (27.28)$$

with Kronecker delta δ_{0m}

$$\delta_{nm} = \begin{cases} 1, & \text{if } n = m \\ 0, & \text{if } n \neq m \end{cases} \qquad (27.29)$$

For $m = 0$, the normalized Legendre polynomials avoiding the use of the Kronecker delta

$$\overline{P}_n(\cos\theta) = (2n+1)^{1/2} P_n(\cos\theta) \qquad (27.30)$$

and for $m \neq 0$

$$N_{nm} = \left[\frac{(n-m)!}{(n+m)!} (4n+2) \right]^{1/2} \qquad (27.31)$$

Normalization further increases the value of the Legendre polynomials and for degree $n = 1\,000\,000$ we have

$$\overline{P}_{1\,000\,000} \approx 1400 \cdot P_{1\,000\,000} \qquad (27.32)$$

leading to the absolute size of the smallest amplitude or translation term to be in the order of 10^{-3}. Moreover, amplitudes and translation terms need to be precomputed only once at the beginning and are valid for all further angular arguments.
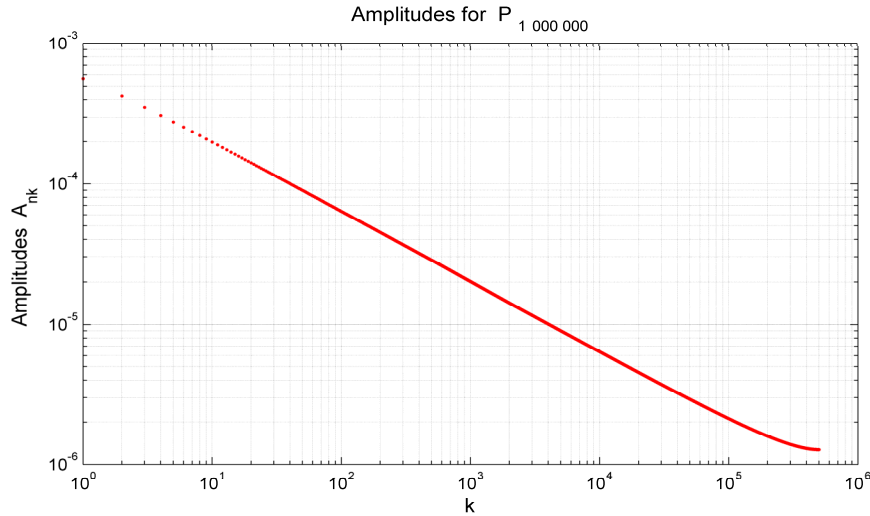
Figure 27.2 Amplitudes A_{nk} for the Legendre polynomial of degree $n = 1\,000\,000$.

Figure 27.3 and Figure 27.4 show amplitudes A_{nk} from (27.23) calculated for degree $n = 100$ and $n = 10\,000$ respectively. Computation of multiple cosine terms could be simplified by means of the following expression derived from Chebyshev polynomials avoiding additional sine terms used in the standard recursions for trigonometric series one can find in the literature

$$\cos k\theta = 2\cos\theta\cos(k-1)\theta - \cos(k-2)\theta \tag{27.33}$$

For the sake of completeness, multiple sine terms can be calculated by means of

$$\sin k\theta = 2\cos\theta\sin(k-1)\theta - \sin(k-2)\theta \tag{27.34}$$

However, for small increments of the angular component δ in a linear sequence $\theta = \theta_0 + k\delta$, $k = 0, 1, 2, \dots$, it is more elegant to use the following recurrence (Press et al. 2007)

$$\begin{aligned}
\cos(\theta + \delta) &= \cos\theta - \left[\alpha\cos\theta + \beta\sin\theta\right] \\
\sin(\theta + \delta) &= \sin\theta - \left[\alpha\sin\theta - \beta\cos\theta\right]
\end{aligned} \tag{27.35}$$
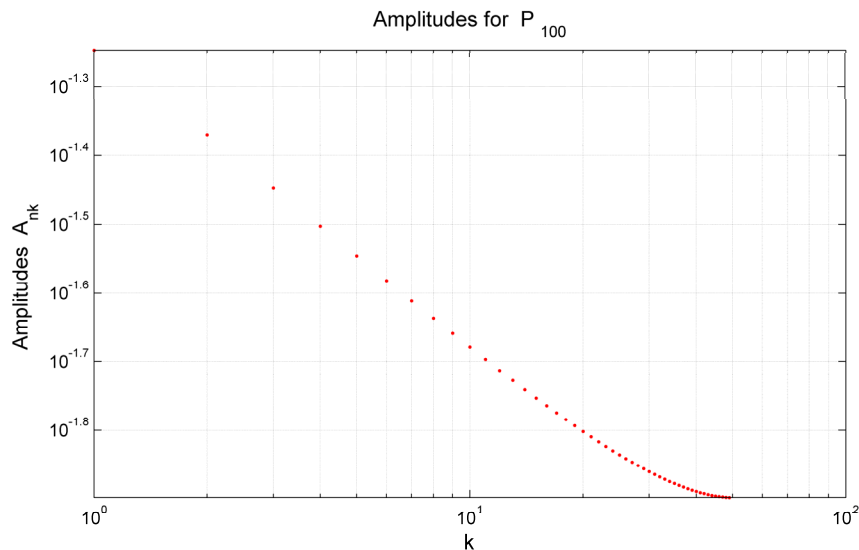
Figure 27.3 Amplitudes A_{nk} for the Legendre polynomial of degree $n = 100$.

318
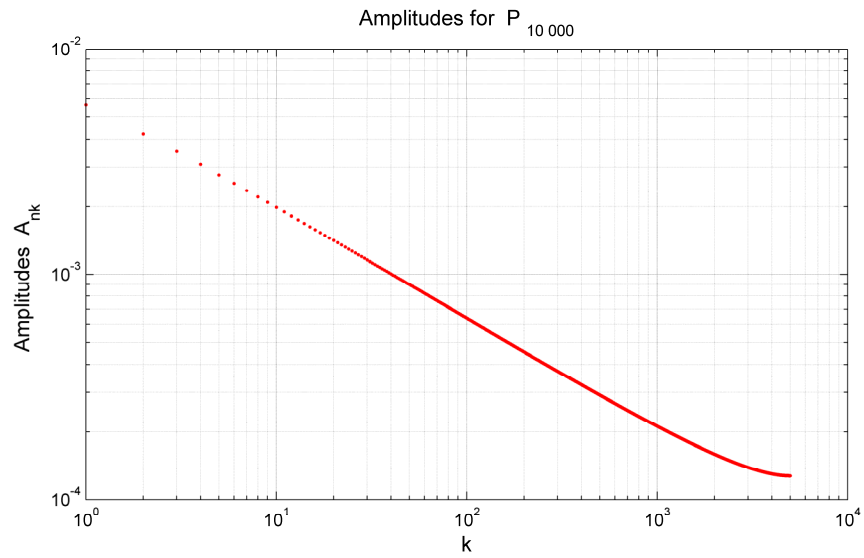
Figure 27.4 Amplitudes A_{nk} for the Legendre polynomial of degree $n = 10\,000$.

where α and β are the precomputed coefficients

$$\alpha = 2\sin^2\left(\frac{\delta}{2}\right), \qquad \beta = \sin\delta . \qquad (27.36)$$

In the case of (27.35), α and β do not lose significance if the incremental δ is small (Press et al. 2007).

It is very important to note that, compared to associated Legendre functions, normalized or unnormalized Legendre polynomials up to super ultra-high degrees (e.g., 1 000 000) are very uniform in size and thus do not lead to computational nor numerical problems. This can be clearly seen in Figure 27.5, where unnormalized Legendre polynomials were calculated up to degree $n = 10\,000$.
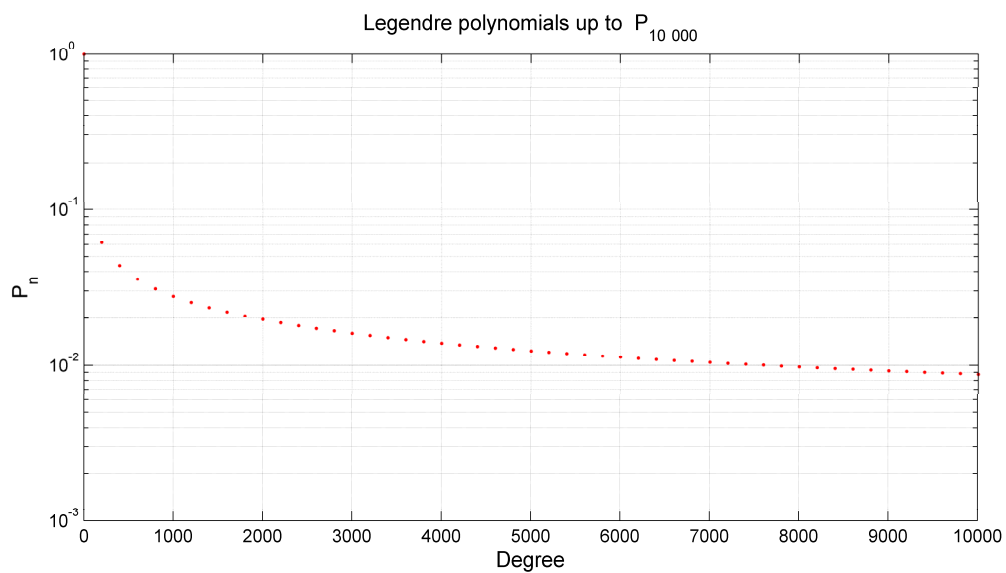
Figure 27.5 Legendre polynomials $P_n(\cos\theta)$ up to degree $n = 10\,000$ for $\theta = 70°$.

319

27.2 Multipole Derivatives of Legendre Polynomials Based on Trigonometric Series

The derivation of (27.22) w.r.t. θ leads to the following expressions for the even derivatives of Legendre polynomials

$$m = even$$

$$
\frac{d^m P_n(\cos\theta)}{d\theta^m} := (-1)^{m/2} \sum_{k=0}^{n/2-1} (n-2k)^m A_{nk} \cos(n-2k)\theta \qquad \begin{cases} n = even \\[2ex] n = odd \end{cases} \tag{27.37}
$$

$$
\frac{d^m P_n(\cos\theta)}{d\theta^m} := (-1)^{m/2} \sum_{k=0}^{n/2} (n-2k)^m A_{nk} \cos(n-2k)\theta
$$

and for the odd derivatives

$$m = odd$$

$$
\frac{d^m P_n(\cos\theta)}{d\theta^m} := (-1)^{(m+1)/2} \sum_{k=0}^{n/2-1} (n-2k)^m A_{nk} \sin(n-2k)\theta \qquad \begin{cases} n = even \\[2ex] n = odd \end{cases} \tag{27.38}
$$

$$
\frac{d^m P_n(\cos\theta)}{d\theta^m} := (-1)^{(m+1)/2} \sum_{k=0}^{n/2} (n-2k)^m A_{nk} \sin(n-2k)\theta
$$

The amplitudes A_{nk} are the same as in (27.23) for the computation of Legendre polynomials and can be pre-computed once. One can see that an algorithm can be based on the set of pre-computed amplitudes

$$
\begin{aligned}
A_{nk}{}^{(m)} &:= (-1)^{m/2}(n-2k)^m A_{nk} \\
A_{nk}{}^{(m)} &:= (-1)^{(m+1)/2}(n-2k)^m A_{nk}
\end{aligned} \qquad \begin{cases} m = even \\[2ex] m = odd \end{cases} \tag{27.39}
$$

Finally, the expressions to calculate the multi-derivatives of Legendre polynomials are thus

$$m = even$$

$$
\frac{d^m P_n(\cos\theta)}{d\theta^m} := \sum_{k=0}^{n/2-1} A_{nk}{}^{(m)} \cos(n-2k)\theta \qquad \begin{cases} n = even \\[2ex] n = odd \end{cases} \tag{27.40}
$$

$$
\frac{d^m P_n(\cos\theta)}{d\theta^m} := \sum_{k=0}^{n/2} A_{nk}{}^{(m)} \cos(n-2k)\theta
$$

$$m = odd$$

$$
\frac{d^m P_n(\cos\theta)}{d\theta^m} := \sum_{k=0}^{n/2-1} A_{nk}{}^{(m)} \sin(n-2k)\theta \qquad \begin{cases} n = even \\[2ex] n = odd \end{cases} \tag{27.41}
$$

$$
\frac{d^m P_n(\cos\theta)}{d\theta^m} := \sum_{k=0}^{n/2} A_{nk}{}^{(m)} \sin(n-2k)\theta
$$

In the computation of gravitational field quantities such as acceleration or gravity gradients we are often interested in particular in the first and second derivatives of Legendre polynomials. From (27.40) and (27.41) we obtain for the first derivative

$$
\frac{d P_n(\cos\theta)}{d\theta} := -\sum_{k=0}^{n/2-1} (n-2k) A_{nk} \sin(n-2k)\theta \qquad \begin{cases} n = even \\[2ex] n = odd \end{cases} \tag{27.42}
$$

$$
\frac{d P_n(\cos\theta)}{d\theta} := -\sum_{k=0}^{n/2} (n-2k) A_{nk} \sin(n-2k)\theta
$$

and for the second derivative

$$
\begin{cases}
\dfrac{d^2 P_n(\cos\theta)}{d\theta^2} := -\displaystyle\sum_{k=0}^{n/2-1} (n-2k)^2 A_{nk} \cos(n-2k)\theta & n = even \\[4mm]
\dfrac{d^2 P_n(\cos\theta)}{d\theta^2} := -\displaystyle\sum_{k=0}^{n/2} (n-2k)^2 A_{nk} \cos(n-2k)\theta & n = odd
\end{cases}
\tag{27.43}
$$

27.3 A Slow Algorithm for Direct Computation of Associated Legendre Functions Without Recursions

In the previous sections we learned that Legendre polynomials do not experience computational problems because amplitudes and translation terms do not diverge with increasing degree of expansion. This is the reason why efficient computation of associated Legendre functions can be based on the pre-computed amplitudes for Legendre polynomials.

To calculate associated Legendre functions one could make direct use of the formula given by Ferrers and the multiple derivatives of the Legendre polynomials derived in the previous section

$$
P_{nm}(\cos\theta) = \sin^m\theta \, \frac{d^m P_n(\cos\theta)}{d(\cos\theta)^m}
\tag{27.44}
$$

Generally speaking, Rodrigues' formula provides the means for producing a series of expressions by repeated differentiation of some other functions. A typical application is in the generation of a series of orthogonal polynomials, such as associated Legendre functions. The powers of the sine functions in (27.44) can be calculated as (e.g., (Bronstein and Semendjajew 1996))

$$
\sin^n\theta =
\begin{cases}
\dfrac{2}{2^n} \displaystyle\sum_{k=0}^{\frac{n-1}{2}} (-1)^{\frac{n-1}{2}-k} \binom{n}{k} \sin(n-2k)\theta & n = odd \\[6mm]
\dfrac{1}{2^n} \binom{n}{\frac{n}{2}} + \dfrac{2}{2^n} \displaystyle\sum_{k=0}^{\frac{n}{2}-1} (-1)^{\frac{n}{2}-k} \binom{n}{k} \cos(n-2k)\theta & n = even
\end{cases}
\tag{27.45}
$$

From (27.45) it follows that the expressions for the first few powers of the sine functions are

$$
\begin{aligned}
\sin^2\theta &= \frac{1}{2} - \frac{1}{2}\cos\theta \\[2mm]
\sin^3\theta &= \frac{3}{4}\sin\theta - \frac{1}{4}\sin 3\theta \\[2mm]
\sin^4\theta &= \frac{3}{8} - \frac{1}{2}\cos 2\theta + \frac{1}{8}\cos 4\theta \\[2mm]
\sin^5\theta &= \frac{5}{8}\sin\theta - \frac{5}{16}\sin 3\theta + \frac{1}{16}\sin 5\theta
\end{aligned}
\tag{27.46}
$$

Let us now introduce the vector form of the associated Legendre functions of degree n, combining the Legendre polynomial (here denoted as P_{n0}) and the associated Legendre functions in the vector form $\boldsymbol{p}_n(\theta)$

$$\boldsymbol{p}_n(\theta) := \begin{bmatrix} P_{n0} \\ P_{n1} \\ P_{n2} \\ \dots \\ P_{nn} \end{bmatrix} \tag{27.47}$$

Combining (27.46) with (27.44) we generate associated Legendre functions in vector form for the first few degrees

$$\boldsymbol{p}_0(\theta) = 1 \qquad \boldsymbol{p}_1(\theta) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \qquad \boldsymbol{p}_2(\theta) = \begin{bmatrix} \dfrac{1}{4} \\ 0 \\ \dfrac{3}{2} \end{bmatrix} + \begin{bmatrix} \dfrac{3}{4}\cos 2\theta \\ \dfrac{3}{2}\sin 2\theta \\ -\dfrac{3}{2}\cos 2\theta \end{bmatrix}$$

$$\boldsymbol{p}_3(\theta) = \begin{bmatrix} \dfrac{3}{8}\cos\theta \\ \dfrac{3}{8}\sin\theta \\ \dfrac{15}{4}\cos\theta \\ \dfrac{45}{4}\sin\theta \end{bmatrix} + \begin{bmatrix} \dfrac{5}{8}\cos 3\theta \\ \dfrac{15}{8}\sin 3\theta \\ -\dfrac{15}{4}\cos 3\theta \\ -\dfrac{15}{4}\sin 3\theta \end{bmatrix} \qquad \boldsymbol{p}_4(\theta) = \begin{bmatrix} \dfrac{9}{64} \\ 0 \\ \dfrac{45}{16} \\ 0 \\ \dfrac{315}{8} \end{bmatrix} + \begin{bmatrix} \dfrac{5}{16}\cos 2\theta \\ \dfrac{5}{8}\sin 2\theta \\ \dfrac{15}{4}\cos 2\theta \\ \dfrac{105}{4}\sin 2\theta \\ -\dfrac{105}{2}\cos 2\theta \end{bmatrix} + \begin{bmatrix} \dfrac{35}{64}\cos 4\theta \\ \dfrac{35}{16}\sin 4\theta \\ -\dfrac{105}{16}\cos 4\theta \\ -\dfrac{105}{8}\sin 4\theta \\ \dfrac{105}{8}\cos 4\theta \end{bmatrix} \tag{27.48}$$

based on the standard convention for the associated Legendre functions used in geodesy

$$P_{nm} = (-1)^m P_n^m \tag{27.49}$$

In vector form, (27.48) can be written as

$$\boldsymbol{p}_1(\theta) = \tilde{\boldsymbol{p}}_1(\theta), \quad \boldsymbol{p}_2(\theta) = \boldsymbol{t}_2 + \tilde{\boldsymbol{p}}_2(2\theta), \quad \boldsymbol{p}_3(\theta) = \tilde{\boldsymbol{p}}_3(\theta) + \tilde{\boldsymbol{p}}_3(3\theta) \tag{27.50}$$

with translation vectors $\boldsymbol{t}_2$, $\boldsymbol{t}_4$, etc., and the corresponding vectors with multiple angular arguments $\tilde{\boldsymbol{p}}_1(\theta)$, $\tilde{\boldsymbol{p}}_2(2\theta)$, $\tilde{\boldsymbol{p}}_3(\theta)$, $\tilde{\boldsymbol{p}}_3(3\theta)$, etc. Compared to the odd associated Legendre functions, even associated Legendre functions always contain additional translational vector terms $\boldsymbol{t}_n$ and, generally, associated Legendre functions of degree n can be written in the form

$$\boldsymbol{p}_n(\theta) = \begin{cases} \boldsymbol{t}_n + \displaystyle\sum_{k=0}^{n/2-1} \tilde{\boldsymbol{p}}_n(n-2k)\theta & \textit{for } n = \textit{even} \\ \displaystyle\sum_{k=0}^{n/2} \tilde{\boldsymbol{p}}_n(n-2k)\theta & \textit{for } n = \textit{odd} \end{cases} \tag{27.51}$$

We have already seen that the associated Legendre function of the first order is the first derivative of the Legendre polynomial of the same degree. Making use of the general definition of associated Legendre functions, given by Ferrers, it follows for the order $m = 1$

$$P_{n1}(\cos\theta) = \sin\theta \frac{dP_n(\cos\theta)}{d(\cos\theta)} = \sin\theta \frac{dP_n(\cos\theta)}{d\theta} \frac{d\theta}{d(\cos\theta)} = -\frac{d}{d\theta} P_n(\cos\theta) \tag{27.52}$$

This was presented for the first time in (Švehla 2008b) and later in (Svehla 2010b). We may now identify two cases for even and odd degrees of the trigonometric expansion

$$P_{n,1} := \begin{cases} \sum_{k=0}^{n/2-1} (n-2k)A_{nk} \sin(n-2k)\theta & \text{for } n = even \\ \sum_{k=0}^{n/2} (n-2k)A_{nk} \sin(n-2k)\theta & \text{for } n = odd \end{cases} \tag{27.53}$$

with the coefficients A_{nk} taken from the Legendre polynomials (27.22). Thus

$$A_{n1k} := (n-2k)A_{nk} \tag{27.54}$$

Expression (27.44) given by Ferrers can be transformed into a trigonometric series of multiple arguments in the following way

$$P_{nm}(\cos\theta) = \sin^m\theta \frac{d^m P_n(\cos\theta)}{d(\cos\theta)^m} = \sin^m\theta \frac{d^m P_n(\cos\theta)}{d\theta^m} \frac{d\theta^m}{d(\cos\theta)^m} \tag{27.55}$$

from where it follows

$$\frac{d(\cos\theta)^m}{d\theta^m} P_{nm}(\cos\theta) = \sin^m\theta \frac{d^m P_n(\cos\theta)}{d\theta^m} \tag{27.56}$$

with

$$\frac{d(\cos\theta)^m}{d\theta^m} = \begin{cases} (-1)^{\frac{m+1}{2}} \sin(\theta) & \text{for } m = odd \\ (-1)^{\frac{m}{2}} \cos(\theta) & \text{for } m = even \end{cases} \tag{27.57}$$

Associated Legendre functions can then be calculated for even orders with

$$P_{nm}(\cos\theta) = (-1)^{\frac{m+1}{2}} \sin^{m-1}\theta \frac{d^m P_n(\cos\theta)}{d\theta^m} \qquad m = even \tag{27.58}$$

and for odd orders with

$$P_{nm}(\cos\theta) = (-1)^{\frac{m}{2}} \sin^{m-1}\theta \frac{d^m P_n(\cos\theta)}{d\theta^m} \qquad m = odd \tag{27.59}$$

Making use of the first order derivative we finally obtain

$$P_{nm}(\cos\theta) = (-1)^{\frac{m+1}{2}} \sin^{m-1}\theta \frac{d^{m-1} P_{n1}(\cos\theta)}{d\theta^{m-1}} \qquad m = even \tag{27.60}$$

and for odd orders

$$P_{nm}(\cos\theta) = (-1)^{\frac{m}{2}} \sin^{m-1}\theta \frac{d^{m-1} P_{n1}(\cos\theta)}{d\theta^{m-1}} \qquad m = odd \tag{27.61}$$

Let us now make use of the following recursion

$$2\frac{dP_{nm}}{d\theta} = (n+m)(n-m+1)P_{n,m-1} - P_{n,m+1}. \tag{27.62}$$

$P_{n,m-1}, P_{n,m+1}$ are both either functions of only cosine or only sine functions. Since the first derivative in (27.62) removes the translation terms from the associated Legendre functions, we can define an expression for the translation term as

$$t_{n,m+1} := (n+m)(n-m+1)t_{n,m-1} \tag{27.63}$$

and since there is no translation term for the associated Legendre functions for all odd orders we have

$$t_{nm} := 0, \qquad \forall m = odd \tag{27.64}$$

Thus, the final expression to calculate translation terms for the associated Legendre functions is

$$t_{nm} := \begin{cases} t_n \prod_{i=1}^{m/2} (n+2i-1)(n-2i+2) & m = even \\ 0 & m = odd \end{cases} \tag{27.65}$$

A similar property can be used for sectorials $(m = n)$, and thus from (27.62) we obtain

$$P_{m,m-1} = \frac{1}{m}\frac{dP_{mm}}{d\theta} \tag{27.66}$$

It is well known that P_{mm} can be calculated directly and (Hobson 1931) reads as

$$P_{mm} = \frac{(2m)!}{2^m \cdot m!}\sin^m \theta \tag{27.67}$$

or

$$P_{mm} = (2m-1)!!\sin^m \theta \tag{27.68}$$

Hence, from (27.66) we obtain

$$P_{m,m-1} = \frac{(2m)!}{2^m \cdot m!}\cos\theta \sin^{m-1}\theta \tag{27.69}$$

or

$$P_{mm} = (2m-1)!!\cos\theta \sin^{m-1}\theta \tag{27.70}$$

Looking at (27.48) one can inductively observe that amplitudes between successive orders of the same degree also satisfy similar recursion relations, or (27.62) can be applied along the same order (see the next section) making use of the orthogonal properties within the same degree. Thus, for arbitrary order m and term k, amplitudes could, in principle, be calculated as a function of the corresponding amplitudes of the Legendre polynomials. However, considering the number of amplitudes and terms involved, a trigonometric algorithm is, generally speaking, not very elegant for practical use. On the other hand, compared to the Legendre polynomials, amplitudes for the associated Legendre functions are not uniform in size and can reach a magnitude likely to cause overflow numerical problems. Therefore, normalization has to be employed during the calculation

$$\overline{A}_{nmk} := N_{nm}A_{nmk} \tag{27.71}$$

with the same normalization function N_{nm} (27.28). The normalized translation terms are given accordingly

$$\overline{t}_{nm} := N_{nm}t_{nm} \tag{27.72}$$

The final expression to calculate fully normalized associated Legendre functions is thus

$$\overline{P}_{nm} = \begin{cases} \sum_{k=0}^{n/2-1} \overline{A}_{nmk} \cos(n-2k)\theta & m = even \\ \sum_{k=0}^{n/2-1} \overline{A}_{nmk} \sin(n-2k)\theta & m = odd \end{cases} , \quad \overline{P}_{nm} = \begin{cases} \overline{t}_{nm} + \sum_{k=0}^{n/2} \overline{A}_{nmk} \cos(n-2k)\theta & m = even \\ \sum_{k=0}^{n/2} \overline{A}_{nmk} \sin(n-2k)\theta & m = odd \end{cases} \tag{27.73}$$

over *n = odd* and *n = even*

or in vector form

$$\boldsymbol{p}_n(\theta) = \begin{cases} \boldsymbol{t}_n + \sum_{k=0}^{n/2-1} \tilde{\boldsymbol{p}}_n(n-2k)\theta & n = even \\ \sum_{k=0}^{n/2} \tilde{\boldsymbol{p}}_n(n-2k)\theta & n = odd \end{cases} \tag{27.74}$$

However, we should bear in mind that associated Legendre functions rapidly increase their absolute size by increasing order of the expansion (m) and thus, to avoid numerical problems, normalization is required. The algorithm is depicted in Figure 26.8.

27.4 Application of Downward and Upward Clenshaw's Recurrence Formula for the Calculation of Trigonometric Series

The question remains of how to calculate the sum of a long time series or trigonometric functions of very high order. Following (Press et al. 2007), Clenshaw's recurrence formula is an elegant way to calculate the sum of coefficients multiplied by a given function that obeys a recurrence formula. Let us write the sum of the series of coefficients c_k multiplied by a given function F_k

$$f(x) = \sum_{k=0}^{N} c_k F_k(x) \tag{27.75}$$

where F_k obeys a recurrence relation for a given $\alpha(n,x)$ and a given $\beta(n,x)$

$$F_{n+1}(x) = \alpha(n,x)F_n(x) + \beta(n,x)F_{n-1}(x) \tag{27.76}$$

The sum of the series (27.75) can then be calculated using Clenshaw's recurrence formula (Press et al. 2007)

$$f(x) = \beta(1,x)F_0(x)y_2 + F_1(x)y_1 + F_0(x)c_0 \tag{27.77}$$

$$y_k = \alpha(k,x)y_{k+1} + \beta(k+1,x)y_{k+2} + c_k, \quad k = N, N-1, ..., 1 \tag{27.78}$$

with the recurrence condition

$$y_{N+2} = y_{N+1} = 0 \tag{27.79}$$

where the sums y_k are calculated in downward order, with k decreasing.

Following (Press et al. 2007), if the functions F_k are small when k is large, and if the coefficients c_k are small when k is small, then the calculated sum can be dominated by small F_k's. In that case, the "remembered" coefficients in (27.78) will involve a delicate cancellation and there can be a significant loss of significance. The solution in such cases is to use an alternative Clenshaw's recurrence that incorporates coefficients c_k in an upward direction (Press et al. 2007)

$$y_k = \frac{1}{\beta(k+1,x)}[y_{k-2} - \alpha(k,x)y_{k-1} - c_k], \quad k = 0,1,...,N-1 \tag{27.80}$$

with the recurrence condition

$$y_{-2} = y_{-1} = 0 \tag{27.81}$$

and the final sum given by

$$f(x) = c_N F_N(x) - \beta(N,x)F_{N-1}(x)y_{N-1} - F_N(x)y_{N-2} \tag{27.82}$$

Following (Press et al. 2007), the rare case where (27.80) and (27.82) should be used instead of (27.77) and (27.78) can be detected automatically by testing whether the operands in the first sum in (27.77) are opposite in sign and nearly equal in magnitude. Other than in this special case, Clenshaw's recurrence is always stable, irrespective of whether the recurrence for the functions F_k is stable in the upward or downward direction or not (Press et al. 2007).

27.5 The Orthogonal Geometrical Form of Associated Legendre Functions in Terms of Trigonometric Series

Using addition trigonometric formulae we can derive the associated Legendre functions as a function of the angle θ rotated by an angle α. For the first few degrees we obtain, for $n = 1$

$$\boldsymbol{p}_1(\theta + \alpha) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}\cos\alpha + \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}\sin\alpha \tag{27.83}$$

for $n = 2$

$$\boldsymbol{p}_2(\theta + \alpha) = \begin{bmatrix} \dfrac{1}{4} \\ 0 \\ \dfrac{3}{2} \end{bmatrix} + \begin{bmatrix} \dfrac{3}{4}\cos 2\theta \\ \dfrac{3}{2}\sin 2\theta \\ -\dfrac{3}{2}\cos 2\theta \end{bmatrix}\cos 2\alpha + \begin{bmatrix} -\dfrac{3}{4}\sin 2\theta \\ \dfrac{3}{2}\cos 2\theta \\ \dfrac{3}{2}\sin 2\theta \end{bmatrix}\sin 2\alpha \tag{27.84}$$

or in general form, representing the translation term as an additional zero rotation $(\cos 0\alpha)$

$$\boldsymbol{p}_2(\theta + \alpha) = \begin{bmatrix} \dfrac{1}{4}\cos 0\theta \\ 0\cos 0\theta \\ \dfrac{3}{2}\cos 0\theta \end{bmatrix}\cos 0\alpha + \begin{bmatrix} -\dfrac{1}{4}\sin 0\theta \\ 0\sin 0\theta \\ -\dfrac{3}{2}\sin 0\theta \end{bmatrix}\sin 0\alpha + \begin{bmatrix} \dfrac{3}{4}\cos 2\theta \\ \dfrac{3}{2}\sin 2\theta \\ -\dfrac{3}{2}\cos 2\theta \end{bmatrix}\cos 2\alpha + \begin{bmatrix} -\dfrac{3}{4}\sin 2\theta \\ \dfrac{3}{2}\cos 2\theta \\ \dfrac{3}{2}\sin 2\theta \end{bmatrix}\sin 2\alpha \tag{27.85}$$

for $n = 3$

326

$$\boldsymbol{p}_3(\theta+\alpha) = \begin{bmatrix} \dfrac{3}{8}\cos\theta \\[2mm] \dfrac{3}{8}\sin\theta \\[2mm] \dfrac{15}{4}\cos\theta \\[2mm] \dfrac{45}{4}\sin\theta \end{bmatrix}\cos\alpha + \begin{bmatrix} -\dfrac{3}{8}\sin\theta \\[2mm] \dfrac{3}{8}\cos\theta \\[2mm] -\dfrac{15}{4}\sin\theta \\[2mm] \dfrac{45}{4}\cos\theta \end{bmatrix}\sin\alpha + \begin{bmatrix} \dfrac{5}{8}\cos3\theta \\[2mm] \dfrac{15}{8}\sin3\theta \\[2mm] -\dfrac{15}{4}\cos3\theta \\[2mm] -\dfrac{15}{4}\sin3\theta \end{bmatrix}\cos3\alpha + \begin{bmatrix} -\dfrac{5}{8}\sin3\theta \\[2mm] \dfrac{15}{8}\cos3\theta \\[2mm] \dfrac{15}{4}\sin3\theta \\[2mm] -\dfrac{15}{4}\cos3\theta \end{bmatrix}\sin3\alpha \qquad (27.86)$$

for $n = 4$

$$\boldsymbol{p}_4(\theta+\alpha) = \begin{bmatrix} \dfrac{9}{64} \\[2mm] 0 \\[2mm] \dfrac{45}{16} \\[2mm] 0 \\[2mm] \dfrac{315}{8} \end{bmatrix} + \begin{bmatrix} \dfrac{5}{16}\cos2\theta \\[2mm] \dfrac{5}{8}\sin2\theta \\[2mm] \dfrac{15}{4}\cos2\theta \\[2mm] \dfrac{105}{4}\sin2\theta \\[2mm] -\dfrac{105}{2}\cos2\theta \end{bmatrix}\cos2\alpha + \begin{bmatrix} -\dfrac{5}{16}\sin2\theta \\[2mm] \dfrac{5}{8}\cos2\theta \\[2mm] -\dfrac{15}{4}\sin2\theta \\[2mm] \dfrac{105}{4}\cos2\theta \\[2mm] \dfrac{105}{2}\sin2\theta \end{bmatrix}\sin2\alpha + \begin{bmatrix} \dfrac{35}{64}\cos4\theta \\[2mm] \dfrac{35}{16}\sin4\theta \\[2mm] -\dfrac{105}{16}\cos4\theta \\[2mm] -\dfrac{105}{8}\sin4\theta \\[2mm] \dfrac{105}{8}\cos4\theta \end{bmatrix}\cos4\alpha + \begin{bmatrix} -\dfrac{35}{64}\sin4\theta \\[2mm] \dfrac{35}{16}\cos4\theta \\[2mm] \dfrac{105}{16}\sin4\theta \\[2mm] -\dfrac{105}{8}\cos4\theta \\[2mm] -\dfrac{105}{8}\sin4\theta \end{bmatrix}\sin4\alpha$$

$$(27.87)$$

If we hold the angle θ fixed, we see that the associated Legendre functions can easily be represented by geometric rotations in hyperspace

$$\boldsymbol{p}_2(\theta+\alpha) = \boldsymbol{t}_2 + \left(\tilde{\boldsymbol{p}}_2(2\theta)\cos2\alpha + \frac{1}{2}\frac{d\tilde{\boldsymbol{p}}_2}{d\theta}(2\theta)\sin2\alpha \right) \qquad (27.88)$$

$$\boldsymbol{p}_3(\theta+\alpha) = \left(\tilde{\boldsymbol{p}}_3(\theta)\cos\alpha + \frac{d\tilde{\boldsymbol{p}}_3}{d\theta}(\theta)\sin\alpha \right) + \left(\tilde{\boldsymbol{p}}_3(3\theta)\cos3\alpha + \frac{1}{3}\frac{d\tilde{\boldsymbol{p}}_3}{d\theta}(3\theta)\sin3\alpha \right) \qquad (27.89)$$

Denoting

$$\tilde{\boldsymbol{p}}_3^*(k\theta) := \frac{1}{k}\frac{d\tilde{\boldsymbol{p}}_3}{d\theta}(k\theta) \qquad (27.90)$$

we can finally define

$$\tilde{\boldsymbol{p}}_n^+(k\alpha) := \tilde{\boldsymbol{p}}_n(k\theta)\cos k\alpha + \tilde{\boldsymbol{p}}_n^*(k\theta)\sin k\alpha \qquad (27.91)$$

For the second and third degree the associated Legendre functions are

$$\boldsymbol{p}_2(\theta+\alpha) = \boldsymbol{t}_2 + \tilde{\boldsymbol{p}}_2^+(2\alpha) \qquad (27.92)$$

$$\boldsymbol{p}_3(\theta+\alpha) = \tilde{\boldsymbol{p}}_3^+(\alpha) + \tilde{\boldsymbol{p}}_3^+(3\alpha) \qquad (27.93)$$

or generally, for even and odd degrees

$$\boldsymbol{p}_n(\theta+\alpha) = \boldsymbol{t}_n + \sum_{k=1}^{n/2}\tilde{\boldsymbol{p}}_n^+(2k\alpha) \quad n = \text{even} \qquad (27.94)$$

$$\boldsymbol{p}_n(\theta+\alpha) = \sum_{k=0}^{(n-1)/2}\tilde{\boldsymbol{p}}_n^+\big((2k+1)\alpha\big) \quad n = \text{odd} \qquad (27.95)$$

We see that for even degrees, associated Legendre functions, in addition to a rotational component, also contain a translation vector that can be represented as a zero rotation $\cos(0 \cdot \alpha)$.

Let us now make the following substitution

$$\theta + \alpha = \frac{\pi}{2} - \varphi \tag{27.96}$$

or

$$\theta = \frac{\pi}{2}, \quad \alpha = -\varphi . \tag{27.97}$$

We can thus produce a very simple representation of the associated Legendre functions in terms of latitude angle φ . For the first degree $n = 1$, we can derive the orthonormal vector span

$$\boldsymbol{p}_1(\theta) = \boldsymbol{p}_1\left(\frac{\pi}{2} - \varphi\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cos\varphi + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \sin\varphi \tag{27.98}$$

for $n = 2$, depicting the use of recurrence relation (27.62)

$$\boldsymbol{p}_2(\theta) = \boldsymbol{p}_2\left(\frac{\pi}{2} - \varphi\right) = \begin{bmatrix} \frac{1}{4} \\ 0 \\ \frac{3}{2} \end{bmatrix} + \begin{bmatrix} -\frac{3}{4} \\ 0 \\ \frac{3}{2} \end{bmatrix} \cos 2\varphi + \begin{bmatrix} 0 \\ \frac{3}{2} \\ 0 \end{bmatrix} \sin 2\varphi \tag{27.99}$$

for $n = 3$

$$\boldsymbol{p}_3(\theta) = \boldsymbol{p}_3\left(\frac{\pi}{2} - \varphi\right) = \begin{bmatrix} 0 \\ \frac{3}{8} \\ 0 \\ \frac{45}{4} \end{bmatrix} \cos\varphi + \begin{bmatrix} \frac{3}{8} \\ 0 \\ \frac{15}{4} \\ 0 \end{bmatrix} \sin\varphi + \begin{bmatrix} 0 \\ -\frac{15}{8} \\ 0 \\ \frac{15}{4} \end{bmatrix} \cos 3\varphi + \begin{bmatrix} -\frac{5}{8} \\ 0 \\ \frac{15}{4} \\ 0 \end{bmatrix} \sin 3\varphi \tag{27.100}$$

and for $n = 4$

$$\boldsymbol{p}_4(\theta) = \boldsymbol{p}_4\left(\frac{\pi}{2} - \varphi\right) = \begin{bmatrix} \frac{9}{64} \\ 0 \\ \frac{45}{16} \\ 0 \\ \frac{315}{8} \end{bmatrix} + \begin{bmatrix} -\frac{5}{16} \\ 0 \\ -\frac{15}{4} \\ 0 \\ \frac{105}{2} \end{bmatrix} \cos 2\varphi + \begin{bmatrix} 0 \\ -\frac{5}{8} \\ 0 \\ -\frac{105}{4} \\ 0 \end{bmatrix} \sin 2\varphi + \begin{bmatrix} \frac{35}{64} \\ 0 \\ -\frac{105}{16} \\ 0 \\ \frac{105}{8} \end{bmatrix} \cos 4\varphi + \begin{bmatrix} 0 \\ \frac{35}{16} \\ 0 \\ -\frac{105}{8} \\ 0 \end{bmatrix} \sin 4\varphi \tag{27.101}$$

A closer look at (27.99) and (27.100) reveals a rotation of the associated Legendre functions from the equator to an arbitrary point along the meridian

$$\boldsymbol{p}_2(\varphi=0) = \boldsymbol{p}_2\left(\frac{\pi}{2}\right) = \begin{bmatrix} \frac{1}{4} \\ 0 \\ \frac{3}{2} \end{bmatrix} + \begin{bmatrix} -\frac{3}{4} \\ 0 \\ \frac{3}{2} \end{bmatrix}\cos\left(2\cdot0\right) + \begin{bmatrix} 0 \\ \frac{3}{2} \\ 0 \end{bmatrix}\sin\left(2\cdot0\right) \tag{27.102}$$

and what is very important to note is that this rotation is orthogonal, i.e., all subsequent rotations in (27.98) , (27.99) and (27.100) consist of orthogonal vectors! In addition, (27.98) is nothing else but the equation of an orthodrome or a great circle on a sphere in Euclidian space. A closer look at all associated Legendre functions of even and odd degrees reveals that all subsequent rotations within the same degree are nothing else but orthogonal rotations, or orthogonal forms of associated Legendre functions. In order to make it clear for even degrees as well, we write the translation terms as a zero associated rotation. For $n=2$ we obtain

$$\begin{bmatrix} \frac{1}{4} \\ 0 \\ \frac{3}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ 0 \\ \frac{3}{2} \end{bmatrix}\cos 0 \cdot \varphi + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\sin 0 \cdot \varphi \tag{27.103}$$

The second vector in (27.103) can be arbitrary, considering that $\sin 0 \cdot \varphi = 0$. Accordingly, for the associated Legendre functions of the third and fourth degree we obtain

$$\boldsymbol{p}_3(\varphi=0) = \boldsymbol{p}_3\left(\frac{\pi}{2}\right) = \begin{bmatrix} 0 \\ \frac{3}{8} \\ 0 \\ \frac{45}{4} \end{bmatrix}\cos 0 + \begin{bmatrix} \frac{3}{8} \\ 0 \\ \frac{15}{4} \\ 0 \end{bmatrix}\sin 0 + \begin{bmatrix} 0 \\ -\frac{15}{8} \\ 0 \\ \frac{15}{4} \end{bmatrix}\cos\left(3\cdot0\right) + \begin{bmatrix} -\frac{5}{8} \\ 0 \\ \frac{15}{4} \\ 0 \end{bmatrix}\sin\left(3\cdot0\right) \tag{27.104}$$

$$\boldsymbol{p}_4(\varphi=0) = \begin{bmatrix} \frac{9}{64} \\ 0 \\ \frac{45}{16} \\ 0 \\ \frac{315}{8} \end{bmatrix} + \begin{bmatrix} -\frac{5}{16} \\ 0 \\ -\frac{15}{4} \\ 0 \\ \frac{105}{2} \end{bmatrix}\cos\left(2\cdot0\right) + \begin{bmatrix} 0 \\ -\frac{5}{8} \\ 0 \\ -\frac{105}{4} \\ 0 \end{bmatrix}\sin\left(2\cdot0\right) + \begin{bmatrix} \frac{35}{64} \\ 0 \\ -\frac{105}{16} \\ 0 \\ \frac{105}{8} \end{bmatrix}\cos\left(4\cdot0\right) + \begin{bmatrix} 0 \\ \frac{35}{16} \\ 0 \\ -\frac{105}{8} \\ 0 \end{bmatrix}\sin\left(4\cdot0\right) \tag{27.105}$$

In general, the orthogonal form of associated Legendre functions can be written as

$$\boldsymbol{p}_n(\theta) = \begin{cases} \displaystyle\sum_{l=0}^{n/2}\left(\boldsymbol{p}_{nl}\cos(n-2l)\varphi + \bar{\boldsymbol{p}}_{nl}\sin(n-2l)\varphi\right) & n = even \\ \displaystyle\sum_{l=0}^{int(n/2)}\left(\boldsymbol{p}_{nl}\cos(n-2l)\varphi + \bar{\boldsymbol{p}}_{nl}\sin(n-2l)\varphi\right) & n = odd \end{cases} \tag{27.106}$$

where

$$\boldsymbol{p}_{nl} \perp \bar{\boldsymbol{p}}_{nl}. \tag{27.107}$$

An additional interesting rotation can be obtained if the coordinate system is rotated about the y axis by $\pi/2$

$$\widehat{\varphi} = \varphi - \frac{\pi}{2} \tag{27.108}$$

For $\boldsymbol{p}_1(\theta)$ we then have

$$\boldsymbol{p}_1(\theta) = \boldsymbol{p}_1\left(\frac{\pi}{2} - \varphi\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cos\varphi + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \sin\varphi \tag{27.109}$$

and for $\boldsymbol{p}_1(\widehat{\varphi})$

$$\boldsymbol{p}_1(\widehat{\varphi}) = \boldsymbol{p}_1\left(\varphi - \frac{\pi}{2}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cos\widehat{\varphi} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} \sin\widehat{\varphi} \tag{27.110}$$

Accordingly, for $\boldsymbol{p}_2(\widehat{\varphi})$ we obtain

$$\boldsymbol{p}_2(\widehat{\varphi}) = \boldsymbol{p}_2\left(\varphi - \frac{\pi}{2}\right) = \begin{bmatrix} \frac{1}{4} \\ 0 \\ \frac{3}{2} \end{bmatrix} - \begin{bmatrix} -\frac{3}{4} \\ 0 \\ \frac{3}{2} \end{bmatrix} \cos 2\widehat{\varphi} - \begin{bmatrix} 0 \\ \frac{3}{2} \\ 0 \end{bmatrix} \sin 2\widehat{\varphi} \tag{27.111}$$

and for $\boldsymbol{p}_3(\widehat{\varphi})$

$$\boldsymbol{p}_3(\widehat{\varphi}) = \boldsymbol{p}_3\left(\varphi - \frac{\pi}{2}\right) = \begin{bmatrix} \frac{3}{8} \\ 0 \\ \frac{15}{4} \\ 0 \end{bmatrix} \cos\widehat{\varphi} - \begin{bmatrix} 0 \\ \frac{3}{8} \\ 0 \\ \frac{45}{4} \end{bmatrix} \sin\widehat{\varphi} - \begin{bmatrix} -\frac{5}{8} \\ 0 \\ \frac{15}{4} \\ 0 \end{bmatrix} \cos 3\widehat{\varphi} - \begin{bmatrix} 0 \\ -\frac{15}{8} \\ 0 \\ \frac{15}{4} \end{bmatrix} \sin 3\widehat{\varphi} \ . \tag{27.112}$$

27.6 Special Cases of Associated Legendre Functions: Pole and Equator

There are two particular cases of special interest in the calculation of associated Legendre functions, namely when $\theta = 0$ and $\theta = \pi/2$. These two cases are interesting because there are no numerical problems in the calculation of associated Legendre functions at the equator and poles . This means that if the rotation of spherical harmonics or associated Legendre functions can be decomposed into several rotations and where the one about the equatorial axis is limited to just rotation from pole to equator, we can calculate spherical harmonics to an arbitrary ultra-high degree and order. The same is true if we use the addition theorem in the calculation of associated Legendre functions using geometrical rotations of associated Legendre functions along the equator.

For a point at the pole, all associated Legendre functions are equal to 0 and all Legendre polynomials are equal to 1

$$\theta = 0 \quad \Rightarrow \quad \begin{cases} P_{nm} = 0, & \forall\, m \in \left\langle 0, n \right] \\[2mm] P_{n0} = 1. \end{cases} \tag{27.113}$$

For an equatorial point, i.e., $\theta = \pi/2$, we obtain for Legendre polynomials

$$\theta = \pi \,/\, 2 \quad \Rightarrow \quad \begin{cases} P_n = 0 & \forall \ n = odd \\[2em] P_n = t_n + \displaystyle\sum_{k=0}^{n/2-1} (-1)^{k+1} A_{nk} & \forall \ n = even \end{cases} \tag{27.114}$$

It can be shown that for a second case in (27.114) the Legendre polynomials can be calculated using the following expression

$$P_n\left(\theta = \pi\,/\,2\right) := \left(-1\right)^{n/2} t_n^{1/2} \qquad \forall n = even\,. \tag{27.115}$$

The final expression for the calculation of Legendre polynomials for an equatorial point is then

$$P_n\left(\theta = \pi\,/\,2\right) := \left(-1\right)^{n/2} \left[\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{5}{6}\cdot\frac{7}{8}\cdot\frac{9}{10}\cdot\ \cdots\ \frac{n-1}{n}\right] \qquad \forall n = even\,. \tag{27.116}$$

In a similar way, we can develop very fast expressions for the calculation of associated Legendre functions

$$\theta = \pi \,/\, 2 \quad \Rightarrow \quad \begin{cases} P_{nm} = 0 & n = even \ \& \ m = odd \\[2em] P_{nm} = 0 & n = odd \ \& \ m = even \end{cases} \tag{27.117}$$

and when $n = even$ and $m = even$

$$P_{nm}\left(\theta = \pi\,/\,2\right) := \left(-1\right)^{m/2} \frac{t_{nm}}{t_n^{1/2}} \qquad m \in \left[2, n\right], \quad \forall m = even \tag{27.118}$$

that can be reduced to

$$P_{nm}\left(\theta = \pi\,/\,2\right) := \left(-1\right)^{m/2} \frac{t_{nm}}{P_{n0}} \qquad m \in \left[2, n\right], \quad \forall m = even \tag{27.119}$$

where $P_{n0} = P_{n0}\left(\theta = \pi\,/\,2\right)$.

The final expression for the calculation of associated Legendre functions for an equatorial point, when $n = even$ and $m = even$ is

$$P_{nm}\left(\pi\,/\,2\right) := \left(-1\right)^{m/2} P_{n0}\left(\pi\,/\,2\right) \prod_{i=1}^{m/2} (n+2i-1)(n-2i+2) \quad m \in \left[2, n\right], \forall m = even \tag{27.120}$$

with $P_{n0}(\pi\,/\,2) = P_{n0}\left(\theta = \pi\,/\,2\right)$.

In a similar way, for the calculation of associated Legendre functions, when $n = odd$ and $m = odd$, we have

$$P_{nm}\left(\pi\,/\,2\right) = \sum_{k=1}^{(n+1)/2} \left(-1\right)^{k-1} A_{nmk} \qquad \forall \ n = odd \ \text{and} \ m = odd \tag{27.121}$$

Expression (27.121) could be further simplified given the following interesting property

$$P_{nm}\left(\pi\,/\,2\right) = (n+m-1)P_{n-1,m-1} \qquad \forall \ n = odd \ \& \ m = odd \tag{27.122}$$

where $(n+m-1)$ denotes odd numbers starting with the degree of the associated Legendre function

$$\boldsymbol{p}_0(\pi/2) = 1 \qquad \boldsymbol{p}_1(\pi/2) = \begin{bmatrix} 0 \\ 1P_{00} \end{bmatrix} \qquad \boldsymbol{p}_2(\pi/2) = \begin{bmatrix} -\dfrac{1}{2} \\ 0 \\ 3 \end{bmatrix} \qquad \boldsymbol{p}_3(\pi/2) = \begin{bmatrix} 0 \\ 3P_{20} \\ 0 \\ 5P_{22} \end{bmatrix}$$

(27.123)

$$\boldsymbol{p}_4(\pi/2) = \begin{bmatrix} \dfrac{3}{8} \\ 0 \\ -\dfrac{15}{2} \\ 0 \\ 105 \end{bmatrix} \qquad \boldsymbol{p}_5(\pi/2) = \begin{bmatrix} 0 \\ 5P_{40} \\ 0 \\ 7P_{42} \\ 0 \\ 9P_{44} \end{bmatrix}$$

Eq. (27.123) shows that associated Legendre functions for odd degrees can be calculated in a very fast way by multiplying the associated Legendre functions of the previous degree by even numbers starting with the current degree and shifted by one order.

Introducing

$$P_{n-1} = -\frac{(n-1)!}{(n+1)!} P_{n1} \tag{27.124}$$

a similar property can also be found for Legendre polynomials. As an example, we give the associated Legendre functions calculated for an equatorial point for the first six degrees

$$\boldsymbol{p}_0(\pi/2) = 1 \quad \boldsymbol{p}_1(\pi/2) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1P_{00} \end{bmatrix} \quad \boldsymbol{p}_2(\pi/2) = \begin{bmatrix} -\dfrac{1}{2} \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1P_{1-1} \\ 0 \\ 3P_{11} \end{bmatrix} \quad \boldsymbol{p}_3(\pi/2) = \begin{bmatrix} 0 \\ -\dfrac{3}{2} \\ 0 \\ 15 \end{bmatrix} = \begin{bmatrix} 0 \\ 3P_{20} \\ 0 \\ 5P_{22} \end{bmatrix}$$

$$\boldsymbol{p}_4(\pi/2) = \begin{bmatrix} \dfrac{3}{8} \\ 0 \\ -\dfrac{15}{2} \\ 0 \\ 105 \end{bmatrix} = \begin{bmatrix} 3P_{3-1} \\ 0 \\ 5P_{31} \\ 0 \\ 7P_{33} \end{bmatrix} \quad \boldsymbol{p}_5(\pi/2) = \begin{bmatrix} 0 \\ \dfrac{15}{8} \\ 0 \\ -\dfrac{105}{2} \\ 0 \\ 945 \end{bmatrix} = \begin{bmatrix} 0 \\ 5P_{40} \\ 0 \\ 7P_{42} \\ 0 \\ 9P_{44} \end{bmatrix} \quad \boldsymbol{p}_6(\pi/2) = \begin{bmatrix} \dfrac{-5}{16} \\ 0 \\ \dfrac{105}{8} \\ 0 \\ -\dfrac{945}{2} \\ 0 \\ 10395 \end{bmatrix} = \begin{bmatrix} 5P_{5-1} \\ 0 \\ 7P_{51} \\ 0 \\ 9P_{53} \\ 0 \\ 11P_{55} \end{bmatrix}$$

(27.125)

It is very important to note that associated Legendre functions for an equatorial point after normalization are very small. To show that we calculated all associated Legendre functions for a degree and order $n = 1\,000\,000$, see Figure 27.6. The zero values in Figure 27.6 are due to odd orders as defined in (27.114). In order to show that there are no numerical problems in the calculation of associated Legendre functions for the lower orders, we display values for the first $10\,000$ orders, see Figure 27.7. The maximum absolute size of the associated Legendre functions is in the sectorial functions $(n = m)$ reaching the maximum value of ≈ 47.5 for a degree and order $n = m = 1\,000\,000$.
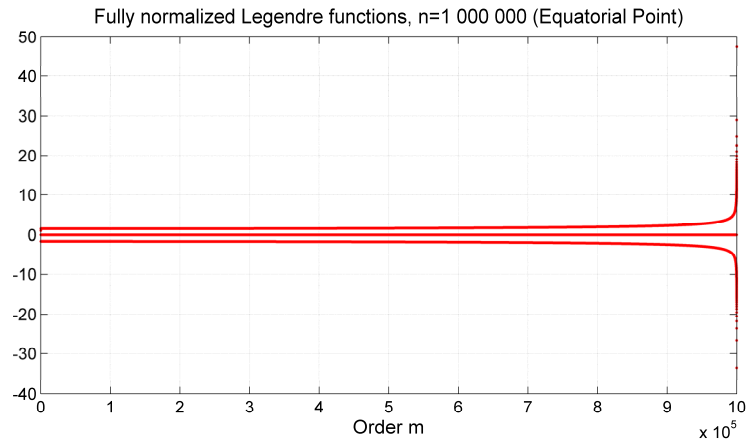
Figure 27.6 Fully normalized associated Legendre functions for an equatorial point for a degree $n = 1\,000\,000$.
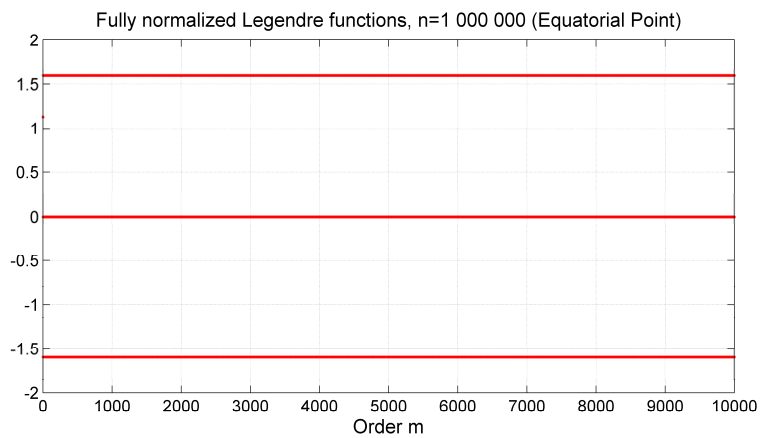
Figure 27.7 Fully normalized associated Legendre functions for an equatorial point for degree $n = 1\,000\,000$ (displayed only up to an order $m = 10\,000$).

28. Insight into the Earth's Interior from Geometrical Rotations in Temporal Gravity Field Maps and Earth's Rotation

To use the dynamics of GPS satellites to complement geometrical Earth rotation and orientation parameters from VLBI has always been a challenge. Geometric VLBI differs from other space geodesy techniques, such as GNSS, SLR and DORIS, in that it does not rely on satellite dynamics to estimate terrestrial reference frame parameters. Here we present a geometrical approach that combines dynamic and geometric variations in the Earth's rotation with temporal gravity field variations. We demonstrate that this novel approach provides a new insight into the Earth's interior, especially into processes and dynamics associated with the Earth's fluid outer core and the great earthquakes over the last 10 and 100 years. Firstly, we demonstrate that by combining two LAGEOS satellites in low MEO orbit we can remove errors in secular orbit perturbations stemming from low zonal harmonics (J_2) and give new insights into the Earth's rotation and nutation rates. Nutation rates were first estimated from GPS data including orbit determination of GPS satellites (Rothacher et al. 1999). Here we extend the theoretical model of nutation rates and show how, with the nodal separation close to 180° of the two LAGEOS satellites, common orbital errors in terms of nodal and apsidal orbit precession are eliminated. This approach based on celestial mechanics opens up the possibility of using satellite dynamics to determine rates of nutation and variations in length-of-day (LOD) very accurately and correlate them against the variations in the temporal gravity field (errors in J_2 are eliminated). This then leads us to the unexplained rate of variations in dynamical LOD estimated from GPS/LAGEOS (orbits driven by Earth's gravity) and from geometrical LOD from VLBI (external measure of Earth's orientation). We show how the rotation of spherical harmonics can explain this unresolved effect since rotation of the tri-axial Earth ellipsoid is the real physical phenomenon measured by gravity field missions as well as by SLR to LAGEOS satellites. We show how the geometrical rotation of spherical harmonics is equivalent to temporal gravity field variations and in the case of second degree harmonics is directly proportional to the rate of variations in LOD. This was presented for the first time in (Švehla 2008b). The conventional IERS mean pole model is in very good agreement with the terrestrial pole of the GRACE monthly gravity field models (derived from $\overline{C}_{21}$ and $\overline{S}_{21}$ gravity field coefficients). We show that temporal variations in the orientation of the tri-axial Earth ellipsoid (sectorials) are taking place along the equatorial plane, i.e., sharing the same axis of rotation within <0.02 arcsec w.r.t. the IERS mean pole model. This dynamic of the tri-axial Earth ellipsoid is very highly correlated with the major earthquakes over the last 10 (GRACE mission) and 100 years. Recently, (Holme and de Viron 2013) showed that variations in the Earth's rotation that occur with a 5.9-year cycle are probably related to motions within the Earth's fluid outer core (contemporaneous with geomagnetic jerks). Here we show that temporal gravity field variations in the second degree harmonics, represented by a rotation of the tri-axial Earth ellipsoid, most likely have the same or a similar origin. The idea to study the Earth's orientation is further extended with a highly elliptical orbit as proposed for the

Space-Time Explorer (STE-QUEST) mission in the ESA Cosmic Vision Programme. We discuss the potential of tracking the STE-QUEST satellite in a highly elliptical orbit with VLBI, especially during long apogee dwells, against extragalactic radio sources, thus, combining a geometrical celestial VLBI frame and a terrestrial reference frame. We show how a highly elliptical orbit can be considered as a sensor for Earth rotation, for low-order spherical harmonics coefficients and subsequently for the Earth's interior dynamics. A satellite dwells for a considerable period of time at the apogee of a highly elliptical orbit, thus it is a perfect target for VLBI to map satellite dynamics against the positions of extra-galactic radio sources. In LEO, a satellite can be observed only for a very short period of time with VLBI and other ground-based techniques. In addition, lunar third-body perturbations are very much uniform along the LEO orbit. Thus, in comparison with HEO, the LEO orbit precesses mainly due to the J_2 coefficient of the Earth's gravity field.

28.1 The Theoretical Basis of Length of Day Variations and Nutation Rates and Their Extension to First-Order Perturbation Theory

SLR is one of the major space geodetic techniques used to establish a link between geometry (reference frames), the Earth's rotation and the gravity field of the Earth. Besides providing scale for the reference frame, SLR gives direct information about the geocenter coordinates and this information is also included in the low-degree gravity field coefficients C_{10}, C_{11} and S_{11}. Moreover, SLR has provided information about the principle axes of inertia relating Earth's rotation to the orientation of the gravity field (C_{21}, S_{21}, S_{22}) and to the ITRF orientation (link to VLBI). SLR is the most accurate space geodesy technique in providing estimates of the J_2 coefficient (and its rate) of the gravity field of the Earth over a long period of time. Here we look at the J_2 coefficient of the Earth's gravity field and its relation to the orbital mechanics of the two LAGEOS satellites in order to estimate length of day variations of the Earth's rotation and nutation rates using SLR and the dynamics of the LAGEOS satellites.

Space geodetic techniques such as GNSS have been used to determine the length of day (or rates in $UT1 - UTC$) on a routine basis. The estimation of offsets in $UT1 - UTC$ and corrections to nutation models was, in the past, uniquely reserved to very long baseline interferometry (VLBI) and lunar laser ranging (LLR). Following (Rothacher et al. 1999), it was demonstrated for the first time that significant contributions to the estimation of nutation by GNSS are possible for periods below about 16 days. Since March 1994, daily nutation rates have been estimated at the Center for Orbit Determination in Europe (CODE) using the data collected by the global GPS network (Rothacher et al. 1999). There is no fundamental difference between the estimation of $UT1 - UTC$ rates (length of day) and nutation rates in obliquity and longitude from satellite data. However, GNSS is more sensitive to high frequencies because the dynamics of the satellites is involved. (Rothacher et al. 1999) reported the computation of corrections for a set of 34 nutation periods between 4 and 16 days. They reported that formal uncertainties of the estimated nutation coefficients in obliquity $\Delta\varepsilon$ and longitude $\Delta\psi \sin \varepsilon_0$ grow linearly with the period from several microarcseconds at periods of a few days to about 30 μas at periods of 16 days. In the case of the LAGEOS satellites, the estimation of length of day and nutation rates is significantly more challenging, given their very low orbit altitude and the resulting sensitivity to the gravity field of the Earth and its temporal variations.

Following (Rothacher et al. 1999), differential changes in $UT1 - UTC$, and nutation in obliquity $\Delta\varepsilon$ and longitude $\Delta\psi \sin \varepsilon_0$ as a function of differential changes in the right ascension of the ascending node Ω, the inclination i and the argument of latitude u_0 are as follows

28.1 The Theoretical Basis of Length of Day Variations and Nutation Rates and Their Extension to First-Order Perturbation Theory

$$\Delta\left(UT1 - UTC\right) = -\left(\Delta\Omega + \cos i \cdot \Delta u_0\right) / \rho$$
$$\delta\Delta\varepsilon = \cos\Omega \cdot \Delta i + \sin i \cdot \sin\Omega \cdot \Delta u_0 \qquad (28.1)$$
$$\delta\Delta\psi \cdot \sin\varepsilon_0 = -\sin\Omega \cdot \Delta i + \sin i \cdot \cos\Omega \cdot \Delta u_0$$

where length of day (LOD), in the absence of leap seconds is defined as

$$UT1\text{-}\dot{UTC} = -LOD \qquad (28.2)$$

The inverse relations read as

$$\Delta i = \cos\Omega \cdot \delta\Delta\varepsilon - \sin\Omega \cdot \left(\delta\Delta\psi \cdot \sin\varepsilon_0\right)$$
$$\Delta\Omega \cdot \tan i = -\sin\Omega \cdot \delta\Delta\varepsilon - \cos\Omega \cdot \left(\delta\Delta\psi \cdot \sin\varepsilon_0\right) - \tan i \cdot \rho \cdot \Delta(UT1 - UTC) \qquad (28.3)$$
$$\Delta u_0 \cdot \sin i = \sin\Omega \cdot \delta\Delta\varepsilon + \cos\Omega \cdot \left(\delta\Delta\psi \cdot \sin\varepsilon_0\right)$$

where ρ is the ratio of universal to sidereal time ($\rho \approx 1.0027379$). (28.1) and (28.3) show, how Earth rotation parameters and orbital elements are related to each other, and answers why offsets in nutation and $UT1 - UTC$ cannot be estimated together with orbital elements, based on the dynamic POD of GNSS or SLR satellites. To estimate offsets in nutation and $UT1 - UTC$ one needs VLBI (or LLR). Any offset in the orbital nodes is one-to-one related to the offset in $UT1 - UTC$, and nearly circular Earth-centered satellite orbits are not very sensitive to the absolute orientation of the nodes. However, secular variations of orbital elements are driven by the gravity field of the Earth and, in the case of J_2 perturbations (Earth approximated to an oblate spheroid), nodal and apsidal lines precess. For a polar orbit with inclination $i = 90°$, there is no nodal precession of the orbit due to the J_2 gravity field coefficient. Thus, the orbital elements are perturbed in terms of secular rates and one can use it to estimate rates in $UT1 - UTC$ and nutation. The first derivative of (28.1) gives a functional model relating rates in the Earth's rotation and secular rates in the orbital elements

$$\left(UT1\text{-}\dot{UTC}\right) = -\left(\dot{\Omega} + \cos i \cdot \dot{u}_0\right) / \rho$$
$$\Delta\dot{\varepsilon} = \cos\Omega \cdot \dot{i} + \sin i \cdot \sin\Omega \cdot \dot{u}_0 \qquad (28.4)$$
$$\Delta\dot{\psi} \cdot \sin\varepsilon_0 = -\sin\Omega \cdot \dot{i} + \sin i \cdot \cos\Omega \cdot \dot{u}_0$$

From Kaula's first order perturbation theory (Kaula 1966), the relationship between the J_2 gravity field coefficient (C_{20}) and orbital elements is given as

$$\frac{da}{dt} = \frac{de}{dt} = \frac{di}{dt} = 0$$
$$\frac{d\Omega}{dt} = \frac{3nC_{20}a_e^2}{2(1-e^2)^2 a^2}\cos i$$
$$\frac{d\omega}{dt} = \frac{3nC_{20}a_e^2}{4(1-e^2)^2 a^2}(1 - 5\cos^2 i) \qquad (28.5)$$
$$\frac{dM}{dt} = n - \frac{3nC_{20}a_e^2}{4(1-e^2)^{3/2} a^2}(3\cos^2 i - 1)$$

with the semi-major axis a, the eccentricity e, the argument of perigee ω, the mean anomaly M, the mean motion n and a_e is the equatorial radius of the Earth. (28.5) shows that LOD and rates in nutation can directly be related to the nodal and apsidal orbit precession. From (28.5), we see that the J_2 term does not introduce a secular rate in inclination, thus, following (28.4), errors in the nutation rates are affected by the errors in modeling apsidal precession and they increase with orbit inclination. On the other hand, nodal

337

precession ($\dot{\Omega}$) is directly related to the length of day. The Moon, Sun and planets introduce additional secular rates in nodal and apsidal orbit precession, as well as several relativistic effects such as frame dragging. The third-body perturbations in orbital elements are discussed in Section 28.8.2, hence for nodal and apsidal precession an additional rate due to the Moon, Sun and planets using Keplerian elements reads as

$$
\begin{aligned}
\frac{d\omega}{dt} &= \frac{3\mu' n'^2}{8n\sqrt{1-e^2}}\left[1 + \frac{3}{2}e'^2 + \frac{15}{8}e'^4\right]\left[(5\cos^2 i - 1 + e^2) + 5(1 - e^2 - \cos^2 i)\cos 2\omega\right] \\
\frac{d\Omega}{dt} &= \frac{3\mu' n'^2}{8n\sqrt{1-e^2}}\left[1 + \frac{3}{2}e'^2 + \frac{15}{8}e'^4\right]\left[5e^2\cos 2\omega - 3e^2 - 2\right]\cos i
\end{aligned}
\tag{28.6}
$$

where the orbit of the perturbing body is denoted with the eccentricity e' and the mean motion n'. The $\mu' = m'/(m_0 + m')$ denotes the mass ratio referring the mass of the perturbing body m' to the mass of the Earth m_0, see Section 28.8.2 for more details. Following the IERS Conventions 2010 (Petit and Luzum 2010), the relativistic correction to the acceleration $\Delta\ddot{\vec{r}}$ of an artificial Earth satellite considering the full post-Newtonian formulation, neglecting the Earth's oblateness and including both the effects of Lense-Thirring precession or frame-dragging (second term) and geodesic (de Sitter) precession (third term), is given as

$$
\begin{aligned}
\Delta\ddot{\vec{r}} = \frac{GM_E}{c^2 r^2}\Bigg\{&\left[\left[2(\beta + \gamma)\frac{GM_E}{r} - \gamma\dot{\vec{r}}\cdot\dot{\vec{r}}\right]\vec{r} + 2(1+\gamma)(\vec{r}\cdot\dot{\vec{r}})\dot{\vec{r}}\right] + \\
&+ (1+\gamma)\frac{GM_E}{c^2 r^3}\left[\frac{3}{r^2}(\vec{r}\times\dot{\vec{r}})(\vec{r}\cdot\vec{J}) + (\dot{\vec{r}}\times\vec{J})\right] + \\
&+ \left\{(1+\gamma)\left[\dot{\vec{R}}\times\left(\frac{-GM_S\vec{R}}{c^2 R^3}\right)\times\dot{\vec{r}}\right]\right\}\Bigg\}
\end{aligned}
\tag{28.7}
$$

with

c speed of light in vacuum

β,γ PPN parameters equal to 1 in general relativity

$\vec{r}$ position of the satellite with respect to the Earth

$\vec{R}$ is the position of the Earth with respect to the Sun,

$\vec{J}$ is the Earth's angular momentum per unit mass, $\left|\vec{J}\right| \cong 9.8\times 10^8 \ m^2/s$

GM_E and GM_S are the products of the gravitational constant and the mass of the Earth and Sun, respectively. The main difference between the de Sitter geodesic precession and the Lense–Thirring effect is that the de Sitter precession is due to the gravity of the central gravity field, whereas the Lense–Thirring precession is caused by the rotation of the central gravity field. These effects cause additional secular precessions of the orbital plane along the equator and the argument of perigee, similar to J_2. From (28.7), we see that Lense-Thirring precession is directly related to nodal orbit precession ($\dot{\Omega}$), since $\dot{\vec{r}}\times\vec{J}$ in the second term in (28.7) has an equatorial component symmetrical with the nodal line of the orbit. Recent results on the confirmation of frame dragging relativity based on orbital dynamics (Ciufolini and Pavlis 2004), heavily depend on the synergy of all three main components: low-degree zonal gravity field coefficients, LOD, nutation rates and Lense-Thirring precession. In addition, there are small relativistic propagation effects such as the so-called Shapiro effect, that, e.g., in the case of the CHAMP orbit, introduces an offset of about 6 mm in the radial orbit direction.

We see that Earth's rotation and orbital dynamics are highly coupled with the gravity field of the Earth and other planets, as well as relativistic effects. By lowering the orbit altitude, the estimation of the Earth's rotation parameters becomes more difficult due to the high-frequency signal from the gravity field. However,

recent gravity field missions offer gravity models of very high accuracy with high spatial and temporal resolution, beyond the sensitivity of dynamic POD. This opens up new possibilities for modeling Earth rotation as we will demonstrate here.

28.2 Removal of Aliasing Effects from the Low-Degree Spherical Harmonics Using Counter-Precessing Orbits in the Estimation of Length of Day Variations and Nutation Rates

Apart from temporal gravity field maps derived on a weekly and monthly basis by the GRACE mission, the first few low-degree harmonics, especially the zonal degree coefficients are of special interest that are not constant, but vary with time. Secular rates in the low-degree harmonics have been explained mainly by the postglacial rebound from the mantle (Peltier and Jiang 1996) and to some extent by sea level change due to the melting of the ice caps (James and Ivins 1997). The Earth's dynamic oblateness (J_2) had been decreasing, according to space geodetic observations over the past 30 years, until around 1998, when it switched quite suddenly to an increasing trend. This change in the global mass distribution measured by J_2 was first reported by (Cox and Chao 2002), reporting that this J_2 effect considerably overshadowed that of mantle rebound. This increase signifies a large change in global mass distribution. Using, the ECCO ocean circulation model, (Dickey et al. 2002) determined that the observed increase in J_2 is caused primarily by the surge in subpolar glacial melting and by mass shifts in the Southern Pacific and Indian Oceans.

Figure 28.1 shows the C_{20} coefficient from monthly gravity field maps provided by the GRACE mission, Release-05 (RL05). One can see a mainly decreasing trend over the last decade as well as a periodic annual term. In the case of the orbits of GNSS satellites, errors in J_2 and other low-degree zonal coefficients will introduce a systematic rate in estimated length of day and nutation rates, since the orbit inclination of all GNSS satellites is very similar, c.f. (28.4). These results will be biased by the same amount for all GNSS satellites. Therefore, we can consider two satellites with a nodal separation of 180°, with prograde and retrograde orbital motion. From (28.4) and (28.5) we see that, in the case of errors in J_2, the total orbit effect in terms of nodal/apsidal precession will cancel out for both satellites, since the two orbits precess in opposite directions (precession/regression of orbital planes).
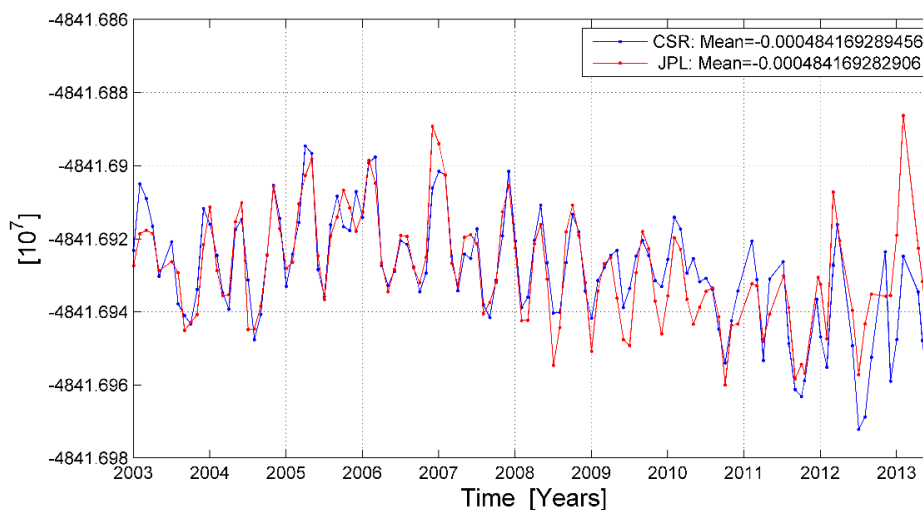
Figure 28.1 Gravity field coefficient C_{20} from GRACE monthly gravity maps (RL05) over the last 10 years.

Orbital Elements	LAGEOS-1	LAGEOS-2
Altitude	5894 km	5782 km
Eccentricity	0.0044	0.0138
Inclination	110°	53°
Right Ascension of Ascending Node	102°	-37°
Perigee	-8°	15°

Table 28.1 Estimated orbital elements of LAGEOS-1/2. Separation of $\approx 140°$ in ascending nodes and $\approx 60°$ in inclination mean they are almost ideal for removing the effects in low-degree zonal gravity field coefficients.

This can clearly be seen in Figure 28.2 that shows the offsets in $UT1 - UTC$ and nutation, drifting in opposite directions when using a single satellite. However, when data from both LAGEOS satellites are combined, errors in the orbit precession significantly cancel and LOD and nutation rates can be estimated. Table 28.1 shows the orbital elements of the LAGEOS satellites. One can see that LAGEOS orbits are separated by $\approx 140°$ in ascending nodes and $\approx 60°$ in inclination. These orbital characteristics mean that they are almost ideal for reducing the effects of errors in the low-degree zonal coefficients in the estimated LOD and nutation rates. There is also an additional effect due to correlations between the precession, LOD and nutation rates.

A closer look at Figure 28.2 reveals that the estimated offset in $UT1 - UTC$ of LAGEOS-2 drifts faster (by a factor of ≈ -2) in comparison to that of LAGEOS-1. A similar effect can also be noted in the nutation for both satellites. In order to explain this effect, we calculated secular perturbation in the longitude of the ascending node following the first order perturbation theory, (Kaula 1966). The average rates of precession for the LAGEOS-1 and regression for the LAGEOS-2 orbit node are then

$$\text{LAGEOS-1}: \quad \frac{d\Omega}{dt} = 0.345° \, / \, \text{day} \qquad \text{LAGEOS-2}: \quad \frac{d\Omega}{dt} = -0.627° \, / \, \text{day} \qquad (28.8)$$
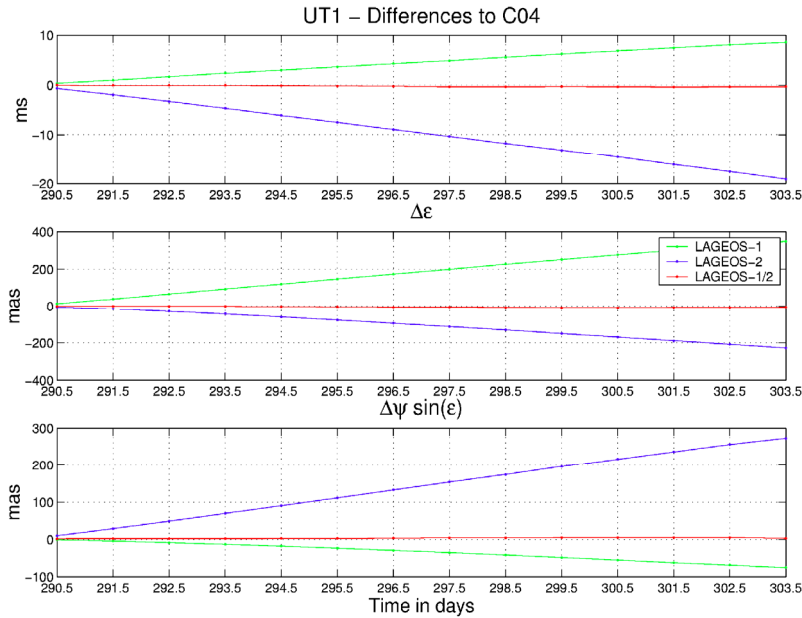
Figure 28.2 UT1 and nutation offsets estimated from dynamic POD of the LAGEOS satellites. Nutation rates drift in opposite directions when using single satellite data, clearly indicating errors in low-degree gravity field coefficients (J_2). However, when data from both LAGEOS satellites are combined, errors in the orbit precession/regression significantly cancel out. Therefore, length of day and nutation rates can be significantly estimated. Since an offset in UT1 and nutation cannot be estimated using LAGEOS dynamic POD, the initial value was set to C04. There is also an additional effect due to correlations between precession and LOD.

that differs by a ratio ≈ -2. However, the drift in UT1 in Figure 28.2 will correspond to ΔC_{20} that is two orders of magnitude higher than the annual amplitude variation of the C_{20} gravity field coefficient from GRACE gravity fields, see Figure 28.1. Following (28.5), we obtain an error of $\Delta C_{20} = 1.555 \times 10^{-8}$ in C_{20} when *LOD* or rate in UT1 from Figure 28.2 for LAGEOS-1 is back-substituted into (28.5). This indicates that C_{20} is most likely not the candidate to explain the effect of *LOD* in Figure 28.2.

What is the effect of non-gravitational forces? Can Solar radiation pressure and other non-gravitational effects be reduced or removed by the counter-precessing LAGEOS orbits? Gaussian perturbation equations for the perturbing accelerations in the radial R, transversal S and out of plane direction W, are given as

$$
\begin{aligned}
\dot{a} &= \frac{2}{n} S \\
\dot{e} &= \frac{1}{na}\big(\sin u \cdot R + 2\cos u \cdot S\big) \\
\dot{i} &= \frac{\cos u}{na} W \\
\dot{\Omega} &= \frac{\sin u}{na \sin i} W \\
\dot{u}_0 &= -\frac{2}{na} R + \frac{3n}{2a}\dot{a}\big(t - t_0\big) - \cos i \cdot \dot{\Omega}
\end{aligned}
\tag{28.9}
$$

with the mean motion $n = \sqrt{GM/a^3}$. Substituting (28.9) into (28.4) we derive expressions relating perturbing accelerations with *LOD* and nutation rates

$$
\begin{aligned}
\dot{\text{UT1-UTC}} &= \frac{2}{\rho na}\cos i R - \frac{3}{\rho a}(t-t_0)S - \frac{\sin i}{\rho na}\sin u W \\
\Delta\dot{\varepsilon} &= \left[-\frac{2}{na}R + \frac{3}{a}S\big(t-t_0\big)\right]\sin i \sin\Omega + \left[\frac{\cos\Omega}{na}\cos u - \cos i \frac{\sin\Omega}{na}\sin u\right]W \\
\Delta\dot{\psi}\sin\varepsilon_0 &= \left[-\frac{2}{na}R + \frac{3}{a}S\big(t-t_0\big)\right]\sin i \cos\Omega - \left[\frac{\sin\Omega}{na}\cos u + \cos i \frac{\cos\Omega}{na}\sin u\right]W
\end{aligned}
\tag{28.10}
$$

Following (28.10), we see that the length of day and the nutation rates may be separated into two components, namely in-orbital and out-of-orbital plane components. The out-of-plane component is dependent on the argument of latitude and thus is highly correlated with the empirical SRP parameters estimated as part of the POD. This correlation is clearly visible, decomposing the GPS SRP model (used also for LAGEOS) into four terms (28.11). The first term $\vec{a}_{CODE}$ in (28.11) is an a priori model, and the other three components in the first line are empirical accelerations estimated in the satellite-fixed coordinate frame (X, Y, D), see (Rothacher and Mervart 1996). a_{D0} denotes the direct solar radiation pressure and a_{Y0} the Y-bias

$$
\begin{aligned}
\vec{a}_{rpr} &= \vec{a}_{CODE} + D(u)\vec{e}_D + Y(u)\vec{e}_Y + X(u)\vec{e}_X \\
D(u) &= a_{D0} + a_{DC}\cos(u) + a_{DS}\sin(u) \\
Y(u) &= a_{Y0} + a_{YC}\cos(u) + a_{YS}\sin(u) \\
X(u) &= a_{X0} + a_{XC}\cos(u) + a_{XS}\sin(u)
\end{aligned}
\tag{28.11}
$$

We may draw the general conclusion that with counter-precessing orbits we may also reduce effects due to the non-gravitational forces (that act symmetrically on both satellites) such as solar radiation pressure, as well as effects predominately in the radial direction, such as albedo. Due to differing shadow/sunlight geometry, solar radiation pressure may significantly differ between two LAGEOS satellites, since the empirical acceleration model is estimated only during sunlight periods. However, non-gravitational forces are not a candidate to explain the large UT1 drift in Figure 28.2.

28.3 Length of Day Variations and Nutation Rates from Counter-Precessing LAGEOS-1 and LAGEOS-2 Orbits

In order to assess the sensitivity of the length of day variations and nutation rates based on the dynamic POD of the LAGEOS-1 and LAGEOS-2 satellites, we processed SLR data for the period of two weeks during the CONT02 campaign. LAGEOS SLR normal equations calculated here were later combined with the normal equations from GPS and VLBI solutions, see (Thaller 2009). The LAGEOS orbits were modeled with two weekly orbit arcs. SLR data were processed in the same way as for the orbit validation of LEO satellites. However, in order to ensure full consistency with GPS and up-to-date IGS-type data processing, the numerical integration of the orbits was carried out using the JGM-3 gravity field model (Tapley et al. 1996).

The solar radiation pressure model or the model of empirical accelerations estimated as part of the dynamic POD is given in (28.11). The only difference to (28.11) is that no a priori solar radiation pressure model was applied in the dynamic POD of the LAGEOS satellites. Empirical accelerations were estimated only during the sunlight period of the orbit and do not have any effect during the shadow passages.

Figure 28.3 shows daily pole coordinates (x, y) estimated using POD of the LAGEOS-1 and LAGEOS-2 satellites and their comparison from C04. One can note a very good agreement of 1 cm RMS (daily solutions). The solution in Figure 28.3 was obtained after stacking daily normal equations, and this is the reason why larger deviations to C04 are present only at the beginning and at the end of the combined orbit arc. The orbits of the LAGEOS-1 and LAGEOS-2 satellites were estimated as two weekly orbit arcs and after stacking of daily NEQs, the continuous 14-day orbit arc does not show any discontinuity at the nominal arc boundaries.

Figure 28.4 shows the estimated *LOD* and nutation rates in obliquity and longitude. One can clearly see the improvement in the solution where both LAGEOS satellites are included, since offsets in *LOD* and nutation rates are much reduced. When SLR data from the LAGEOS-1 and the LAGEOS-2 satellites are combined, errors due to orbit precession are much smaller, thus length of day and nutation rates can be estimated much better using dynamic POD. This is the reason why the single satellite solution in Figure 28.4 shows a very large offset in the estimated length of day variations and nutation rates. In Figure 28.2, errors in orbit precession are reflected as a clear drift in the estimated UT1 and nutation offsets.
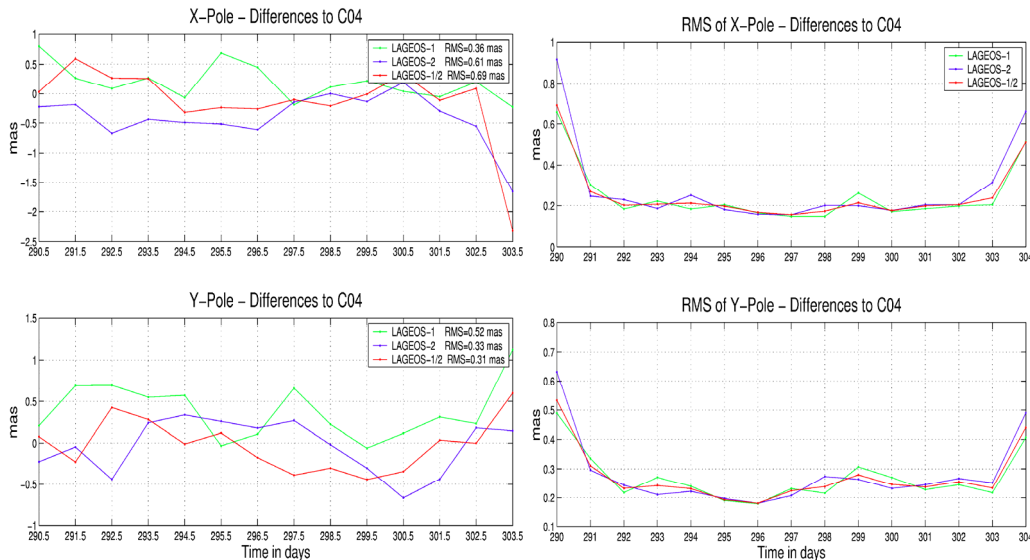
Figure 28.3 Pole coordinates from dynamic POD of LAGEOS-1 and LAGEOS-2 satellites. One can clearly see a larger RMS at the beginning and the end of the 14-day orbit arc. In this solution daily NEQs with the orbital parameters of LAGEOS-1 and LAGEOS-2 were stacked to obtain the continuous 14-day orbit arc.
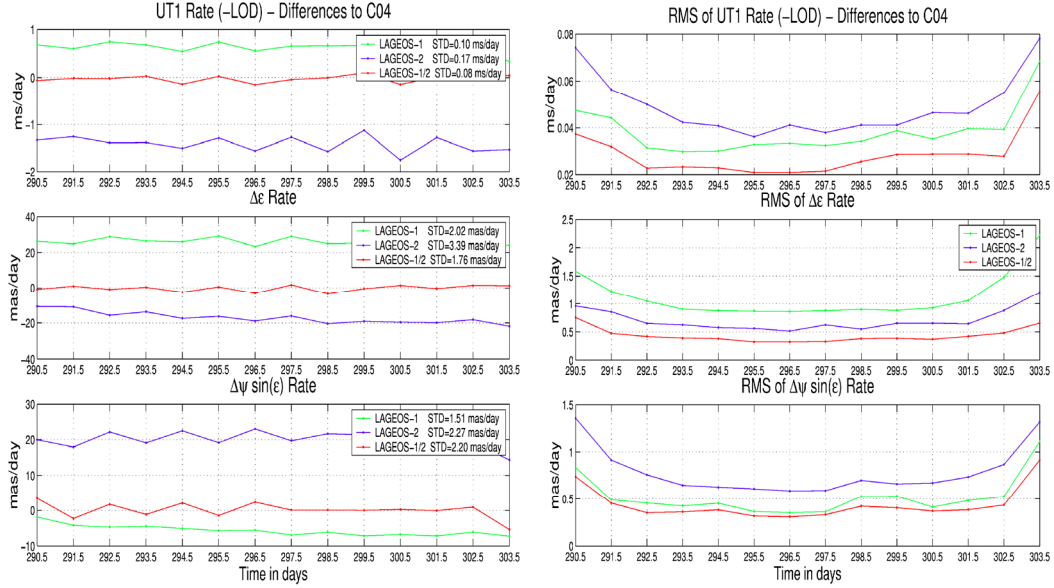
Figure 28.4 LOD variations and nutation rates from dynamic POD of LAGEOS satellites. One can see that when counter-precessing orbits of both LAGEOS satellites are used together, errors due to orbit precession are considerably reduced and LOD variations and nutation rates can be better estimated.

Figure 28.4 clearly shows a factor of ≈ -2 in the single satellite solution of the estimated length of day from LAGEOS-2 in comparison to results from the LAGEOS-1 satellite, see also Figure 28.2. A similar effect can also be noted in the nutation rates for both satellites. Following (28.8), the ratio between regression (backwards motion) and precession of the orbital node of the two LAGEOS satellites is

$$\frac{\dfrac{d\Omega}{dt}(\text{LAGEOS-2})}{\dfrac{d\Omega}{dt}(\text{LAGEOS-1})} = \frac{-0.627° / \text{day}}{0.345° / \text{day}} \approx -1.82 \tag{28.12}$$

that is fully in line with Figure 28.4 and Figure 28.2. Again, the size of the error $\Delta C_{20} = 1.555 \times 10^{-8}$ in the C_{20} gravity field coefficient (based on nodal precession $\dot{\Omega}$) cannot explain the *LOD* offset for LAGEOS-1 in Figure 28.4 since it is one to two orders of magnitude higher than the annual amplitude variation of C_{20} from GRACE temporal gravity fields, see Figure 28.1. Later in this section we will explain this effect in detail.

In Figure 28.5, the nutation rates estimated from the dynamic POD of the LAGEOS-1 and LAGEOS-2 satellites are compared to the solution from the dynamic POD of GPS satellites and to the geometrical solution from VLBI. Normal equations from the dynamic POD of LAGEOS-1 and LAGEOS-2 based on SLR measurements were calculated here and combined with the normal equations from the GPS and VLBI solutions (Thaller 2009). In Figure 28.5, one can see a good agreement for nutation in obliquity.

The agreement between the nutation rates estimated from the dynamic POD of LAGEOS and GPS satellites and the nutation from VLBI is good. However, this is just an apparent agreement since in the estimation of nutation offsets and UT1, external information from VLBI is needed. Typically, after two weeks of the CONT02 campaign dynamic UT1 will have accumulated a difference of about 0.75 ms compared to the geometric UT1 estimated from VLBI. This discrepancy for GPS between the rate of Earth's rotation, when determined using dynamics, and that determined using geometry, that is also consistent with the LAGEOS rate in Figure 28.4, will be discussed in the next section.
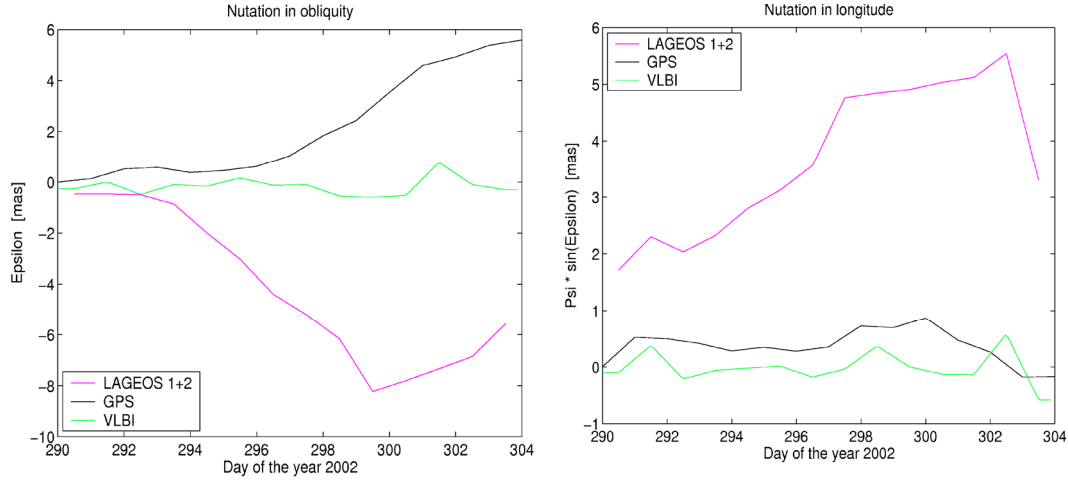
Figure 28.5 Nutation in obliquity/longitude from dynamic POD of LAGEOS and GPS satellites against VLBI (geometry). A good agreement for nutation in obliquity can be seen. Normal equations from the dynamic POD of LAGEOS-1 and LAGEOS-2 based on SLR measurements were calculated here and combined later with the normal equations from GPS and VLBI solutions (Thaller 2009).

28.4 Variations in the Orientation of the Earth's Tri-Axial Ellipsoid – LOD from LAGEOS/GPS and LOD from VLBI

The shape of the Earth is not a perfect oblate spheroid, since the equatorial ellipse is flattened by about 70 m (Indian Ocean). Earth's equatorial flattening depends mainly on the second-degree sectorial harmonics (J_{22}, S_{22}) as well as the even zonal harmonics of degrees $n = 4$, 6 and 8 (Burša et al. 1984). Following (Marchenko and Schwintzer 2003), the mean equatorial flattening based on JGM-3, EGM96, GRIM5-S1, and GRIM5-S1CH1 gravity models is $1/f_e = 91\ 436.6 \pm 0.3$. This value corresponds to a difference in the equatorial major axes of 69.8 m. Using JGM-3 (Tapley et al. 1996), one can calculate the longitude of the axis of the Earth's ellipsoid of inertia with respect to the principal moment of inertia, $\lambda_M = -14.9291° \pm 0.0004°$. This value is very similar to that of (Burša et al. 1984), that showed in the early 80's that the direction of the largest axis of the best-fitting tri-axial Earth's ellipsoid is practically identical to the direction of the axis of the Earth's ellipsoid of inertia. If we consider the sectorial surface spherical harmonics of second degree

$$\overline{C}_{22} \cdot \overline{P}_{22} \cos 2\lambda, \qquad \overline{S}_{22} \cdot \overline{P}_{22} \sin 2\lambda \tag{28.13}$$

we may calculate the orientation of the tri-axial Earth's ellipsoid, i.e., the orientation of the semi-major axis of the equatorial ellipse as

$$\tan 2\lambda_M = \frac{S_{22}}{J_{22}} = \frac{\overline{S}_{22}}{\overline{C}_{22}} \tag{28.14}$$

A similar development can be found in (Burša et al. 1984). The orientation of the semi-minor axis is then $\lambda_M + \pi/2$ and located in the Indian Ocean. The second degree zonal coefficient $\overline{C}_{20}$ defines the dynamic flattening of the meridional ellipse, whereas orthogonally, sectorial coefficients $\overline{C}_{22}, \overline{S}_{22}$ define the equatorial ellipse. The estimation of the full Earth's tensor of inertia and its eigenvalues from recent global gravity field solutions can be found in (Marchenko and Schwintzer 2003). From (28.14) we see that any temporal gravity field variations in the second-degree harmonics will introduce a rotation of the tri-axial Earth's ellipsoid.

Generally speaking, the second degree spherical harmonic (sectorial) will rotate about the polar axis and this rotation can be defined by the rotation matrix $\Re$

$$V\left(\Re^{-1}(\vec{r})\right) = \frac{GM}{r}\left[1 + \sum_{n=2}^{\infty}\left(\frac{a}{r}\right)^2 \sum_{m=-2}^{2} \kappa_{2m}(\alpha)\overline{Y}_{2m}(\vec{r})\right] \tag{28.15}$$

We can either rotate the orbit about the polar axis (or any other axis) in Euclidian space $\Re(\vec{r})$ to evaluate the gravitation potential, or rotate the gravity field coefficients in Hilbert space; the resulting gravitational potential V will be the same. From this, we can draw the conclusion that temporal variations in the gravity field itself can be explained by a rotation of the spherical harmonics. For this, the easiest way is to use rotation about the polar axis to represent such a temporal variation. For the second degree spherical harmonic, the rotation about the polar axis $\kappa_2(\alpha)$ can be defined using rotation matrix with $5{\times}5$ elements $\kappa_{2m}(\alpha)$ in (28.15)

$$\kappa_2(\alpha) := \begin{bmatrix} \cos 2\alpha & & & & \sin 2\alpha \\ & \cos\alpha & & \sin\alpha & \\ & & 1 & & \\ & -\sin\alpha & & \cos\alpha & \\ -\sin 2\alpha & & & & \cos 2\alpha \end{bmatrix}\begin{bmatrix} \overline{S}_{22} \\ \overline{S}_{21} \\ \overline{C}_{20} \\ \overline{C}_{21} \\ \overline{C}_{22} \end{bmatrix} \tag{28.16}$$

that is equivalent to a geometrical rotation of the terrestrial frame $\Re(\vec{r})$, or Euclidian space. Any temporal variation of the gravitational potential can be represented or is equivalent to a geometrical rotation in Euclidian space. Looking only at the first row in (28.16), we can easily derive (28.14). For this discussion on geometrical rotations and temporal gravity field variations we refer to Section 26.

If we now calculate the rotation of the second degree spherical harmonics around the polar axis based on temporal gravity maps provided by the GRACE mission (RL05), we obtain the values given in Figure 28.6 and in Table 28.2. In Figure 28.6 one can clearly see an annual period with an amplitude of about $2''$.

Following first order perturbation theory (Kaula 1966), the perturbation in the longitude of the ascending node $\Delta\Omega_{lmpq}$ given in Keplerian elements reads as

$$\Delta\Omega_{lmpq} = GMa_e^l \frac{\left(\partial F_{lmp}/\partial i\right)G_{lpq}\overline{S}_{lmpq}}{na^{l+3}(1-e^2)^{1/2}\sin i[(l-2p)\dot{\omega} + (l-2p+q)\dot{M} + m(\dot{\Omega}-\dot{\theta})]} \tag{28.17}$$

with the inclination functions F_{lmp} and the eccentricity functions G_{lpq}. For $l = m = 2$ and considering only secular perturbations in the ascending node $(p = 1, q = 0)$ we obtain
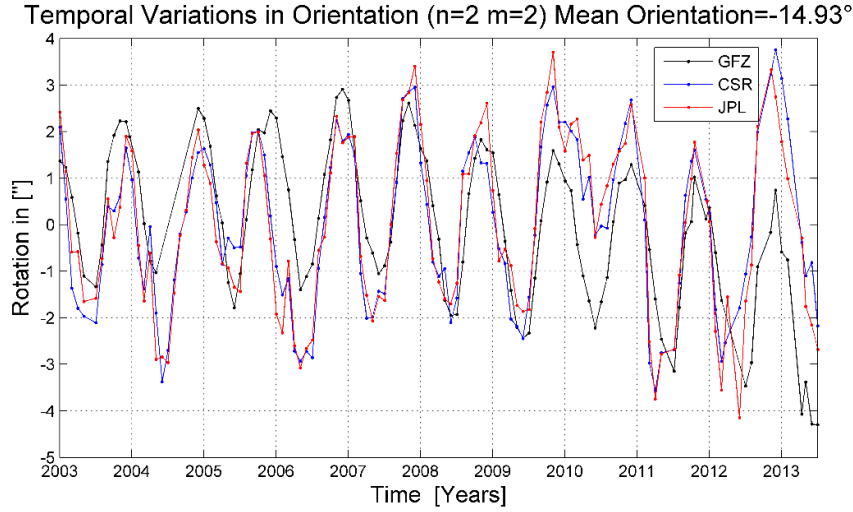
Figure 28.6 Temporal variations in the orientation of the tri-axial Earth's ellipsoid (from $\overline{C}_{22}$ and $\overline{S}_{22}$) in arcsec. over a period of the last 10 years based on GRACE monthly gravity fields (RL05) from JPL, CSR and GFZ. One can clearly see an annual period with an amplitude of about $2''$. This corresponds to the rotation of the equatorial ellipse with an amplitude of ≈ 60 m over a period of one year. The mean orientation of $-14.93°$ places the semi-minor axis of the equatorial ellipse of the Earth's tri-axis ellipsoid at longitude $75.07°$ over the Indian Ocean.

$$\Delta\Omega_{2210} = GMa_e^2 \frac{\left(\partial F_{221}/\partial i\right)G_{210}\overline{S}_{2210}}{2na^5(1-e^2)^{1/2}(\dot{\Omega}-\dot{\theta})\sin i} \tag{28.18}$$

Since $F_{221} = 3\sin^2 i/2$, $G_{210} = (1-e^2)^{-3/2}$, $S_{2210} = C_{22}\cos 2(\Omega-\theta) + S_{22}\sin 2(\Omega-\theta)$ and considering that $\overline{S}_{2210}$ is the integral of S_{2210} with respect to its argument, we obtain the daily perturbation of the longitude of the ascending node due to the equatorial ellipticity defined by C_{22} and S_{22}

$$\Delta\Omega_{2210} = GMa_e^2 \frac{3\cos i}{4na^5(1-e^2)^2(\dot{\Omega}-\dot{\theta})}\left[-C_{22}\sin 2(\Omega-\theta) + S_{22}\cos 2(\Omega-\theta)\right] \tag{28.19}$$

Inserting $\dot{\Omega} = -14.2°/\text{year}$ for the precession of orbital nodes of GPS satellites into (28.19) we finally obtain the perturbation of the longitude of the ascending node for the GPS orbits

$$\Delta\Omega_{2210} = -0.03686 \cdot \left[-C_{22}\sin 2(\Omega-\theta) + S_{22}\cos 2(\Omega-\theta)\right] \tag{28.20}$$

and for LAGEOS-I

$$\Delta\Omega_{2210} = 0.44280 \cdot \left[-C_{22}\sin 2(\Omega-\theta) + S_{22}\cos 2(\Omega-\theta)\right] \tag{28.21}$$

and LAGEOS-II

$$\Delta\Omega_{2210} = -0.80241 \cdot \left[-C_{22}\sin 2(\Omega-\theta) + S_{22}\cos 2(\Omega-\theta)\right] \tag{28.22}$$

From Figure 28.6 we can see the rotational rate of the tri-axial Earth's ellipsoid to be in the order of $4''/180$ days, which corresponds to the amplitude of $2''$ for the annual period. Coresponding perturbation in the longitude of the ascending nodes explains the *LOD* offsets, i.e., the drift in UT1 for LAGEOS-1 and LAGEOS-2 in Figure 28.4 and in Figure 28.2 is

$$\begin{aligned}
\Delta\Omega_{2210}(\text{LAGEOS}-1) &= \quad 0.44280 \cdot 4''/180 = 0.7 \text{ ms/day} &\rightarrow& \quad +9.2 \text{ ms/14 days} \\
\Delta\Omega_{2210}(\text{LAGEOS}-2) &= -0.80241 \cdot 4''/180 = 1.2 \text{ ms/day} &\rightarrow& \quad -16.6 \text{ ms/14 days} \\
\Delta\Omega_{2210}(\text{GPS}) &= -0.03686 \cdot 4''/180 = 0.05 \text{ ms/day} &\rightarrow& \quad +0.76 \text{ ms/14 days}
\end{aligned} \qquad (28.23)$$

with the ratio between LAGEOS-2 and LAGEOS-1

$$\frac{\Delta\Omega_{2210}(\text{LAGEOS}-2)}{\Delta\Omega_{2210}(\text{LAGEOS}-1)} \approx -1.8 \qquad (28.24)$$

and LAGEOS-2 and GPS

$$\frac{\Delta\Omega_{2210}(\text{LAGEOS}-2)}{\Delta\Omega_{2210}(\text{GPS})} \approx 21.8 \qquad (28.25)$$

If we compare the rate in LOD determined using data from GPS/LAGEOS with that determined by VLBI there is an unexplained difference (Thaller 2009) of about -0.7 ms per two weeks that at the Earth's equator gives a rotation of about -32 cm per two weeks, see Figure 28.7. Figure 28.7 shows the rate in dynamic LOD (GPS/LAGEOS) w.r.t. that in geometric LOD (VLBI) for a period of about two weeks. One can clearly see an accumulated difference of about 0.7 ms over a period of two weeks that corresponds to an arc length of about -31.5 cm on the Earth's equator. All three temporal gravity field maps provided by GFZ in Potsdam, CSR in Texas and JPL give an accumulated angular rotation rate of about -0.75 ms for a period of 2 weeks. That corresponds to a rotation (arc length) at the equator of about -35 cm.

Temporal variations in the second-degree harmonic can be determined by POD, however, they should be properly correlated with data from VLBI (fixed to the Earth's crust). This is especially true if the rate in LOD is estimated. Any misorientation between the ITRF realization and the temporal gravity field model (rotation about the polar axis) will result in an additional rotation of the orbit, i.e., the terrestrial frame for both VLBI and gravity needs to be the same. The current realization of the ITRF like ITRF2008 does not take into account temporal variations in the gravity field from a geometrical point of view. Temporal variations in the gravity field are taken into account in the integration of the equation of satellite motion, but there is also a geometrical effect, due to the rotation of the orbit, that is currently not considered. This is given in Figure 28.8 that graphically shows the simplified scheme for the equivalence between the rate in LOD, the orbit precession and the rotation of second degree spherical harmonics. When LOD is considered, typically the atmosphere angular momentum needs to be considered. The geometrical rotation of spherical harmonics is related to temporal gravity field variations and, in the case of the second degree harmonic is directly proportional to the rate in LOD. This was presented for the first time in (Švehla 2008b).

The question still remains: What is the additional rotation we see in the temporal gravity field of the second degree coefficients? Why is the terrestrial frame for temporal gravity maps not the same as for VLBI which is fixed with the station coordinates to the Earth's crust?

Solution	LOD (difference to C04)	Rotation of Spherical Harmonics second-degree harmonic
GPS vs. VLBI	-0.70 ms/(2 weeks) **≈-32 cm/(2 weeks)**	
		Mean over all monthly fields
LAGEOS-1 (RL05)		-0.765 ms (**-35 cm/(2 weeks**)
LAGEOS-2 (RL05)		-0.762 ms (-35 cm/(2 weeks)

Table 28.2 Very good agreement between the rate in (dynamic) LOD (GPS/LAGEOS) and (geometric) LOD (VLBI) ≈ -32 cm/2 weeks, and rotation of the second degree harmonic from the GRACE temporal gravity field maps (RL05 solution) ≈ -35 cm/2 weeks.
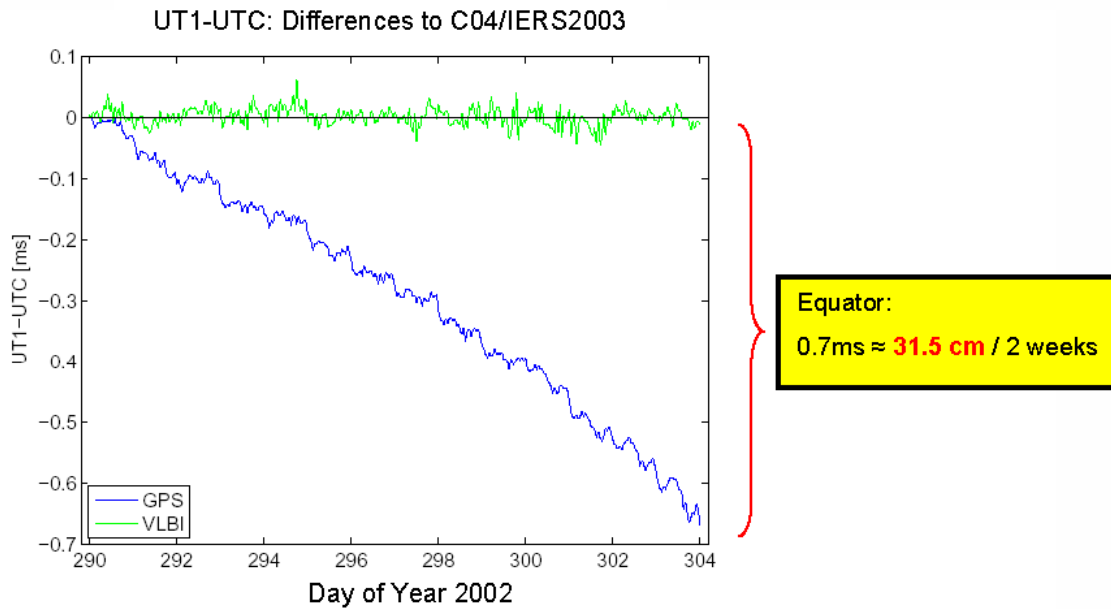
Figure 28.7 Rate in dynamic *LOD* (GPS) w.r.t. geometric *LOD* (VLBI). The background figure was provided by D. Thaller after stacking NEQs. The difference of 0.7 ms corresponds to an equatorial arc of ≈ 31.5 cm accumulated over a period of two weeks.
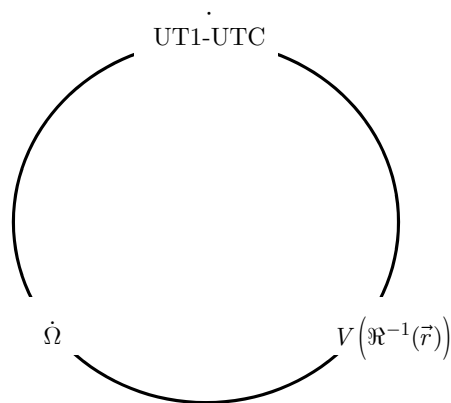
Figure 28.8 Simplified scheme for the equivalence between the rate in *LOD*, precession of orbital plane and rotation of the second degree spherical harmonics (sectorials).

28.4.1 The 6-Yearly Period in the Earth Core Orientation and GRACE Results

GRACE non-tidal high-frequency atmospheric and oceanic mass variation models are routinely generated at GFZ as part of the GRACE monthly gravity field determination. These so-called Atmosphere and Ocean De-aliasing Level-1B (AOD1B) products are added to the background static gravity model. The AOD1B product is a 6-hourly series of spherical harmonic coefficients up to degree and order 100, which is routinely provided to the GRACE Science Data System and the user community with only a few days delay. These products reflect spatio-temporal mass variations in the atmosphere and oceans deduced from an operational atmospheric model and corresponding ocean dynamics provided by an ocean model. The variability is derived by subtraction of a long-term mean of vertical integrated atmospheric mass distributions and a corresponding mean of ocean bottom pressure as simulated using the ocean model. The AOD1B de-aliasing product is applied in estimating monthly gravity field maps (GRACE RL05), thus those effects are properly covered in our analysis.

Recently, (Holme and de Viron 2013) showed that variations in the Earth's rotation that occur on a 5.9-year cycle are probably related to motions within the Earth's fluid outer core (contemporaneous with geomagnetic jerks). They looked at fluctuations in length of day over the last 50 years correcting measured LOD values for atmospheric and oceanic effects by general circulation models that account for most of the variations over yearly and shorter periods. They clearly show a decadal trend in the order of 3 ms /50 years and a constant 5.9-year periodic signal with an amplitude of 0.127 ms . Here we show that the temporal gravity field variations of the second degree harmonics, represented by the geometrical rotation of the tri-axial Earth ellipsoid within the figure of the Earth, most likely has the same origin, namely the Earth's fluid outer core. We came to this conclusion after showing that the rotation of the second degree harmonics (sectorials) can explain the rate in the difference between the dynamically determined LOD from GPS/LAGEOS data, and the geometrically determined LOD from VLBI. Dynamically determined LOD is driven by the gravity within the Earth's interior, whereas geometrically determined LOD from VLBI is purely a geometric measure of the rotation of the Earth's crust against extragalactic radio-sources. Any discrepancy between those two fundamentally different types of approach in measuring the variation of the Earth's rotation will be reflected in the temporal gravity variation of low-degree gravity field coefficients. The discrepancy between dynamic and geometric LOD is equivalent to the geometric rotation of low-degree spherical harmonics. Thus, there must be a difference between the rotation of the Earth's crust and that of the gravity-generating body within the Earth. Everything indicates that this effect has the rotation of the Earth's fluid outer core as its origin, since the magnitude of the missing rate in LOD could be compared with intradecadal periodic variations due to movements within the Earth's molten outer core reported in (Holme and de Viron 2013).

The variations in the Earth's rotation measured by LOD arise from external tidal torques, or from an exchange of angular momentum between the solid Earth and its fluid components. Over short timescales (annual or shorter) the non-tidal component is dominated by the atmosphere, with small contributions from the oceans and the hydrological system. Over decadal timescales, the dominant contribution is from angular momentum exchange between the solid mantle and the fluid outer core (Holme and de Viron 2013).

As a conclusion: The combination of geometric VLBI (Earth's crust frame) with the dynamic space geodesy techniques GNSS, SLR and DORIS in determining length-of-day variations and for comparison with angular temporal gravity field variations of low-degree gravity field coefficients is a promising technique to provide an insight into processes within the Earth's interior. This is especially interesting for the angular momentum exchange between the solid mantle and fluid outer core where geomagnetism is generated. The Swarm mission (ESA) will provide a new insight into the Earth's geomagnetism and any relation with the space geodesy data will be very interesting.

In the next few sections we will look at the geometry of the low-degree gravity field variations and thereafter, we will focus on the possibility of using highly-elliptical orbits for the combination of geometrically determined (VLBI) and dynamically determined reference frames (GNSS, SLR, DORIS).

28.5 Orientation of the Tri-Axial Ellipsoid Against the Conventional IERS Mean Pole Model

While the C_{20} gravity field coefficient of the second degree harmonic drives the dynamic flattening of the Earth, and C_{22} and S_{22} define equatorial flattening, the other two second degree coefficients C_{21} and S_{21} describe the position of the Earth's figure axis. They define a mean figure axis of the Earth (Moritz 1980) that corresponds to the mean pole positions consistent with the terrestrial reference frame. The figure axis should closely coincide with the observed positions of the rotation pole when averaged over a period of several years (Petit and Luzum 2010). Following the IERS Conventions 2010 (Petit and Luzum 2010), any difference between the averaged positions of the mean figure and the mean rotation pole would be due to long-period fluid motions in the atmosphere, oceans, or Earth's fluid core. The mean figure axis which coincides with the mean pole consistent with the terrestrial reference frame is given by the IERS conventional mean pole $\overline{x}_p(t)$ and $\overline{y}_p(t)$ (Petit and Luzum 2010)

$$
\begin{aligned}
\overline{C}_{21}(t) &= \sqrt{3}\overline{x}_p(t)\overline{C}_{20} - \overline{x}_p(t)\overline{C}_{22} + \overline{y}_p(t)\overline{S}_{22} \\
\overline{S}_{21}(t) &= -\sqrt{3}\overline{y}_p(t)\overline{C}_{20} - \overline{y}_p(t)\overline{C}_{22} - \overline{x}_p(t)\overline{S}_{22}
\end{aligned}
\tag{28.26}
$$

where $\overline{C}_{21}(t)$ and $\overline{S}_{21}(t)$ are calculated as a function of time from a given $\overline{x}_p(t)$ and $\overline{y}_p(t)$ and the gravity field coeffitient $\overline{C}_{20}$, $\overline{C}_{22}$ and $\overline{S}_{22}$. Let us now rearrange those expressions and here write them in terms of rotations

$$
\begin{bmatrix} \overline{C}_{21}(t) \\ \overline{S}_{21}(t) \end{bmatrix} = \begin{bmatrix} \sqrt{3}\overline{C}_{20} - \overline{C}_{22} & \overline{S}_{22} \\ -\overline{S}_{22} & -\sqrt{3}\overline{C}_{20} - \overline{C}_{22} \end{bmatrix} \begin{bmatrix} \overline{x}_p(t) \\ \overline{y}_p(t) \end{bmatrix}
\tag{28.27}
$$

obtaining an inverse relation referring the second degree gravity coefficients and the mean pole coordinates

$$
\begin{bmatrix} \overline{x}_p(t) \\ \overline{y}_p(t) \end{bmatrix} = \begin{bmatrix} \sqrt{3}\overline{C}_{20} - \overline{C}_{22} & \overline{S}_{22} \\ -\overline{S}_{22} & -\sqrt{3}\overline{C}_{20} - \overline{C}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \overline{C}_{21}(t) \\ \overline{S}_{21}(t) \end{bmatrix}
\tag{28.28}
$$
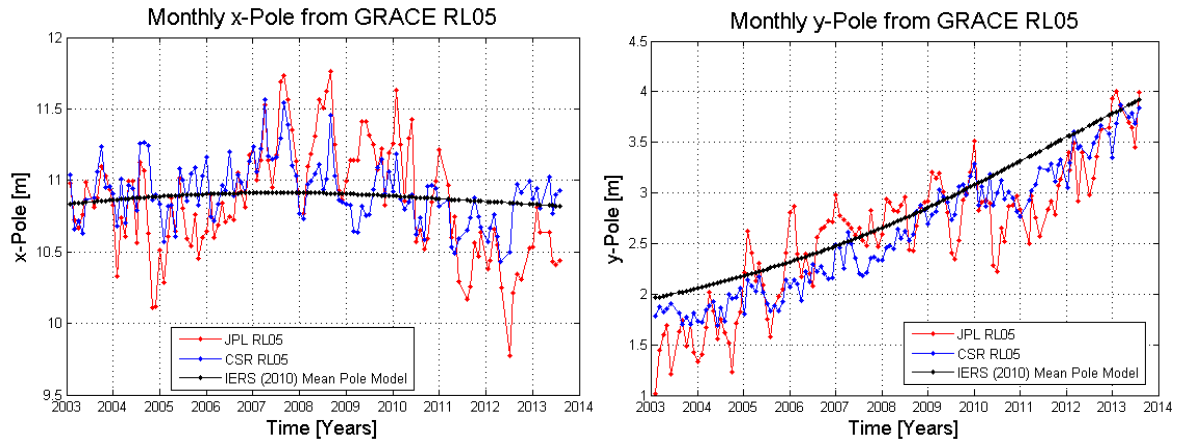
Figure 28.9 Pole coordinates derived from the second degree gravity field coefficients, showing that the conventional IERS mean pole model is in good agreement with the GRACE monthly gravity field models. This indicates that temporal variations in the orientation of the tri-axial Earth ellipsoid (sectorials) take place along the equatorial plane, i.e., sharing the same axis of rotation within <0.02 arcsec w.r.t. the IERS mean pole model.

plotted in Figure 28.9 based on the GRACE monthly gravity field maps (RL05). Figure 28.9 clearly shows that temporal variations in the orientation of the tri-axial Earth ellipsoid (sectorials) are taking place along the Equator, i.e., sharing the same mean axis of rotation within <0.02 arcsec w.r.t. IERS mean pole model.

28.6 Correlations in the Orientation of Earth's Tri-Axial Ellipsoid and the Major Earthquakes Over the Last 100 Years

Figure 28.10 shows the statistics for the number of earthquakes with magnitude >8.0 since the year 1900 and over the last 10 years covering the period of the GRACE mission. Earthquake data were obtained from the earthquake archive maintained by the US Geological Survey. One can see that great earthquakes take place mainly in the March-April and October-November intervals, matching the annual maximum and minimum of temporal variations in the C_{22} gravity field coefficient, i.e., oscillations in the orientation of the tri-axial Earth's ellipsoid, (see Figure 28.6). On the other hand, it is well known that an increased number of earthquakes around the Pacific plate leads to an increased number of earthquakes around the Adriatic-microplate, or generally, in the subduction zone between the European and African lithospheric plates. However, this has never been explained, although noted by many people. Any strong earthquake in the Pacific typically generates a number of smaller earthquakes in the Mediterranean region. Thus, oscillations in the orientation of the tri-axial ellipsoid point towards a missing link to explain the relationship between earthquakes in the Pacific and in the Mediterranean basins, both separated by about 180° in geographical longitude. Low-frequency oscillations in the gravity field within the Earth induce a tremendous load on the very thin Earth's crust twice a year, i.e., separated by about 6 months between minimum and maximum, see Figure 28.6. Recent earthquake activity east of Japan and the increased number of earthquakes in the Mediterranean region over the last few years (south Italy, south Croatia, Greece) together with the increased activity of the Etna volcano clearly indicate a correlation in this direction.
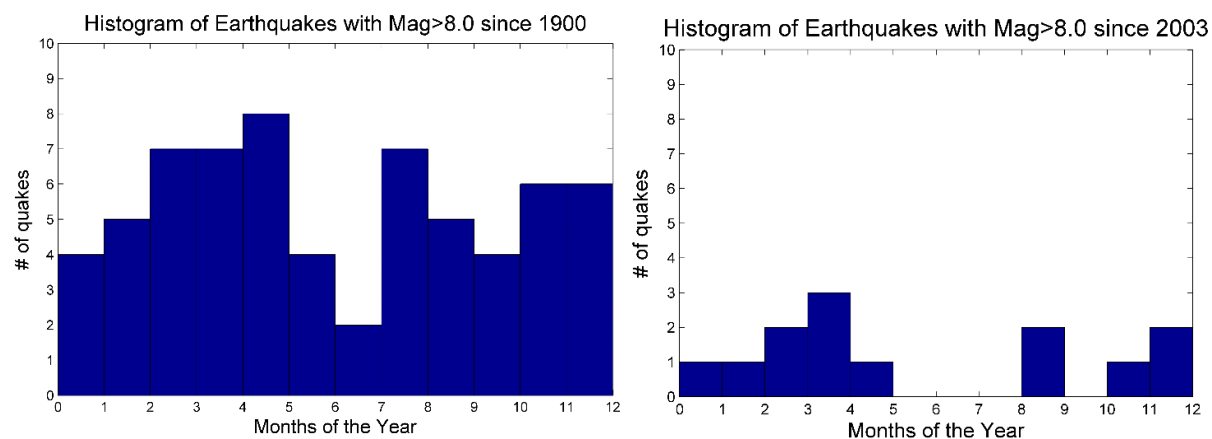
Figure 28.10 Histogram of the number of earthquakes with magnitude >8.0 since the year 1900 (left) and since 2003 – (GRACE mission) (right), (source USGS). One can see that great earthquakes most often occur in the March-April and October-November intervals, matching the annual maximum and minimum of temporal variations or oscillations in the $\overline{C}_{22}$ coefficient. Variations in orientation or oscillations in the orientation of the tri-axial ellipsoid indicate a link to explain the connection between earthquakes in the Pacific and in the Mediterranean basin. It is well known that an increased number of earthquakes around the Pacific plate leads to an increased number of earthquakes around the Adriatic-microplate, i.e., the subduction zone of the European/African lithospheric plates (south Italy, south Croatia, Greece), both separated by 180° in longitude.

28.7 Temporal Variations in the Orientation of the Tri-Axial Earth's Ellipsoid and Low-Degree Sectorial Harmonics

We have shown that variations in the orientation of the second degree zonal/sectorial harmonics take place around the Equator, sharing the same axis of rotation within 0.02 arcsec w.r.t. the IERS mean pole model. Thus, if we assume a representation of the spherical harmonics in terms of rotations

$$V\left(\Re^{-1}(\vec{r})\right) = \frac{GM}{r}\left[1 + \sum_{n=2}^{\infty}\left(\frac{a}{r}\right)^n \sum_{m=-n}^{n} \kappa_n(\lambda) \cdot \overline{Y}_{nm}(\vec{r})\right] \tag{28.29}$$

about the polar axis with longitude λ as the parameter

$$\kappa_n(\lambda) := \begin{bmatrix} \cos n\lambda & & & & \sin n\lambda \\ & \cdots & & \cdots & \\ & & 1 & & \\ & \cdots & & \cdots & \\ -\sin n\lambda & & & & \cos n\lambda \end{bmatrix}\begin{bmatrix} \overline{S}_{nn} \\ \cdots \\ \overline{C}_{n0} \\ \cdots \\ \overline{C}_{nn} \end{bmatrix} \tag{28.30}$$

we may calculate the orientation longitude of the sectorial harmonics along the Equator by means of

$$\tan n\lambda = \frac{\overline{S}_{nn}}{\overline{C}_{nn}} \tag{28.31}$$

Figure 28.11 and Figure 28.12 show time series of the orientation longitude along the Equator for even and odd low-degree sectorial harmonics based on the GRACE monthly gravity maps provided by JPL, CSR and GFZ.
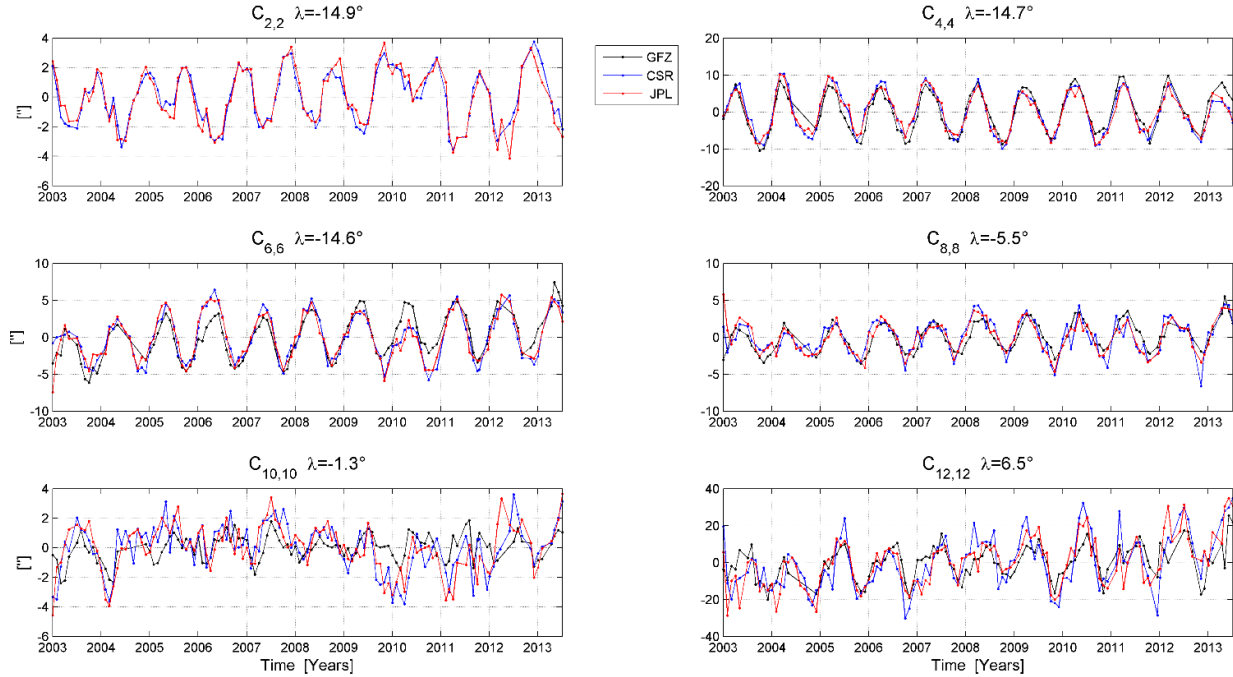
Figure 28.11 Orientation longitude of the even sectorial harmonics from GRACE monthly fields (RL05) calculated from $\overline{C}_{nn}$ and $\overline{S}_{nn}$ coefficients. One can clearly observe periodic variations up to degree/order 8 with an amplitude of several arcsec. Longitude of the orientation is very similar up to degree/order 6.

Figure 28.12 Orientation longitude of the odd sectorial harmonics from GRACE monthly gravity fields (RL05) calculated from $\overline{C}_{nn}$ and $\overline{S}_{nn}$ gravity field coefficients. One can clearly observe periodic variations up to degree/order 9 with an amplitude of several arcseconds.

Comparing both figures one can clearly see periodic variations up to degree 8 or 9 with an amplitude of the emporal variations of several arcseconds. In both cases, the longitude of orientation decreases with increased order of the spherical harmonics, explicitly given in Figure 28.13 up to degree 60. Figure 28.13 shows that the mean orientation of the low-degree sectorial harmonics goes up to 20° for the first 10 degrees and is in the order of several degrees for the max. degree/order. One can draw the conclusion that low-degree sectorial harmonics, especially below degree 12, show clear temporal variations in orientation in a very consistent way with similar amplitudes below 10 arcsec.
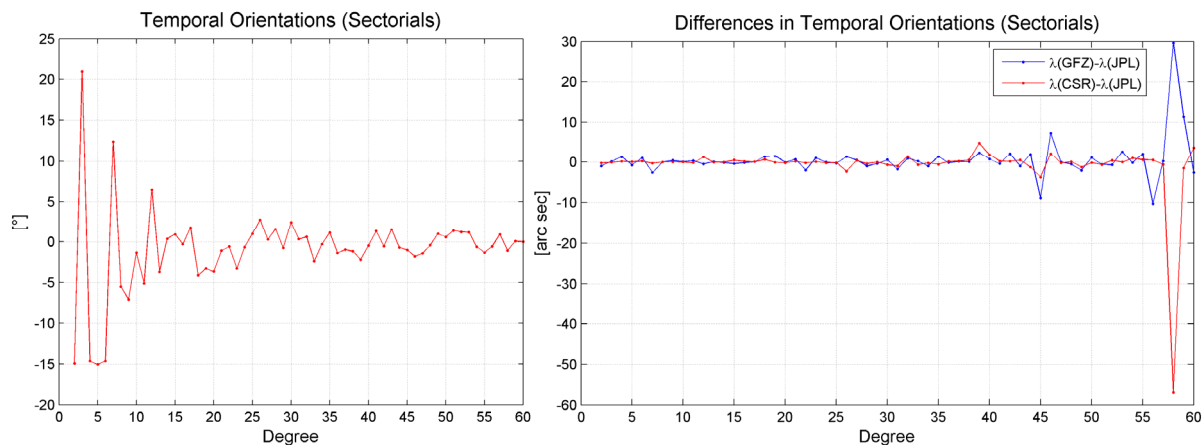
Figure 28.13 Mean orientation longitude (left) in [°] of sectorial harmonics up to degree 60 (GRACE RL05) over the last 10 years. From the figure on the left one can see that the main signal in the orientation of the sectorial harmonics is up to degree 12. The figure on the right shows that for the orientation of high-degree sectorial harmonics there are systematic effects at degree 58 for all three solutions, most likely caused by the filtering of stripping effects in the GRACE monthly gravity fields. Note that units (right) are in [arc sec].
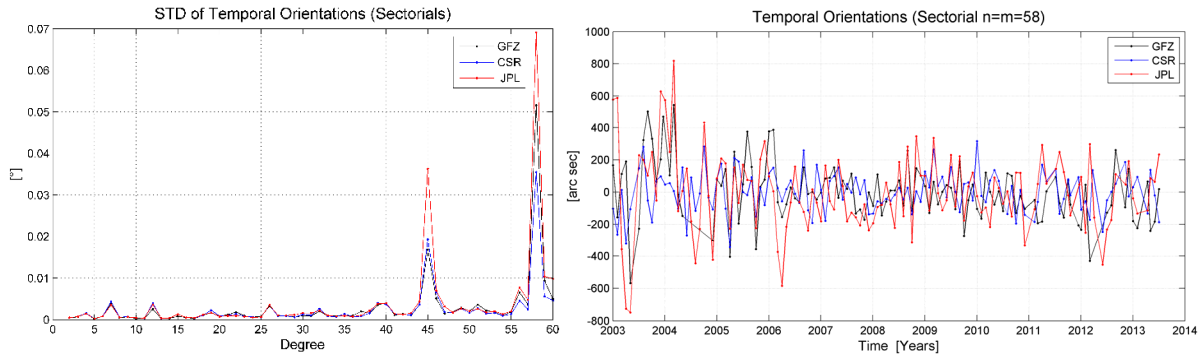
Figure 28.14 Standard deviation (left) of the orientation longitude of sectorial harmonics up to degree 60 (GRACE RL05) over the last 10 years. The figure on the left shows that sectorial harmonics of degree 45 and 58 show higher noise than the other degrees. A closer look at the sectorial harmonics of degree 58 (right) shows high noise in all three solutions provided by JPL, GFZ and CSR (GRACE RL05).

The differences between different solutions (JPL, CSR, GFZ) are well below the arcsec level. Figure 28.13 (right) shows that for the orientation of high-degree sectorial harmonics there are systematic effects at degree 58 for all three solutions, most likely caused by the filtering of stripping effects in the GRACE monthly gravity fields. This is clearly visible in Figure 28.14, where the standard deviation of the temporal variations in orientation is calculated for the sectorial harmonics up to degree 60, clearly identifying peaks at degree/order 45 and degree/order 58. Since GFZ provides gravity maps up to degree/order 90, we also identified degree/order 88 showing a very high standard deviation of temporal orientations. These anomalies are most likely associated with stripping effects in the GRACE gravity fields and associated filtering that is well reported by the GRACE mission.

Since the conventional IERS mean pole model is in very good agreement with the terrestrial pole of the GRACE monthly gravity field (derived from $\overline{C}_{21}$ and $\overline{S}_{21}$), temporal gravity field variations can be represented by simple rotation of spherical harmonic coefficients about the polar axis. Thus, for low-degree spherical harmonics one could establish a form of GRS80 system defining a normal gravity field of the Earth, including both static and temporal fields in terms of geometrical rotations.
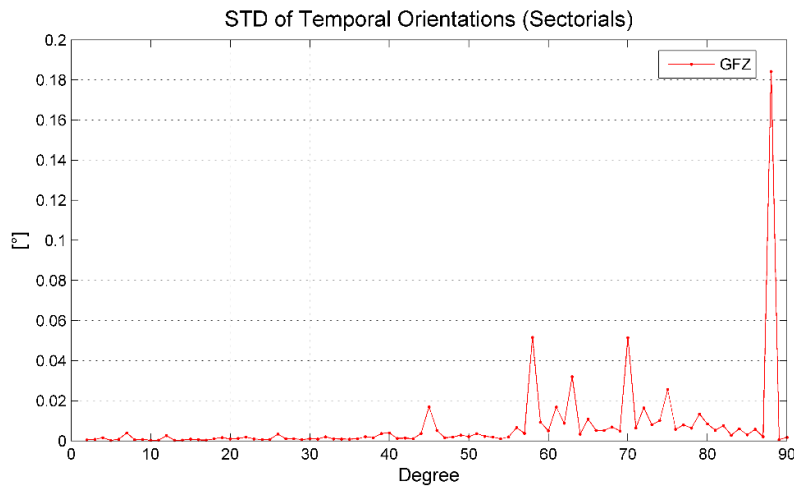
Figure 28.15 Standard deviation of the orientation longitude of sectorial harmonics up to degree 90 (GRACE RL05 from GFZ) over the last 10 yr. Higher values can be noticed for degree 45, 58, 63, 70 and especially for degree/order 88. High standard deviation at degree 88 is most likely related to stripping effects in the GRACE gravity fields and associated filtering. Above degree 60 standard deviation is significantly higher for all degrees.

This could improve the static gravity field derived from GRACE and GOCE, in order to accurately model annual signals in low-degree harmonics as a continuous function over a longer period of time. This is not the case now, where every monthly gravity field map is an independent solution.

28.8 The STE-QUEST Mission: Synergy of Terrestrial and Celestial Reference Frames with Low-Degree Gravity Field Terms Using a Highly Elliptical Orbit

To use highly elliptical orbits for terrestrial and celestial reference frame realizations was first proposed in (Svehla et al. 2013b). The **S**pace-**T**ime **E**xplorer and **QU**antum **E**quivalence Principle **S**pace **T**est (STE-QUEST) is a Medium Class fundamental physics mission pre-selected for the M3 slot of the ESA Cosmic Vision Programme to test Einstein's Equivalence Principle using atom interferometry and the General and Special Theory of Relativity based on atomic clocks and optical/microwave metrology links for time/frequency transfer. Had it been finally selected in 2014, the highly elliptical orbit of the STE-QUEST satellite can be used for terrestrial reference frame (TRF) realization by means of on board GNSS, SLR and VLBI radio source (microwave metrology link observed by VLBI antennae − compatible with VLBI2010). By upgrading the onboard GNSS receiver for DORIS tracking, the STE-QUEST mission will be similar to the GRASP mission proposal from JPL. However, the highly elliptical orbit of STE-QUEST provides advantages for terrestrial and celestial reference frame determination (e.g., ground/space VLBI tracking in apogee), compared to the GRASP mission proposal. VLBI measurements are very challenging for satellites at low LEO altitudes, whereas the STE-QUEST has a highly elliptical orbit where the satellite "dwells" in the apogee and can be observed for a long time against the quasars defining the celestial reference frame. The secondary scientific objectives of the STE-QUEST mission related to space geodesy are as follows, (Svehla et al. 2013b), (Svehla et al. 2014a)

- meet the GGOS (Global Geodetic Observing System) goals for a terrestrial reference frame of the Earth, i.e., 1 mm accuracy and 0.1 mm/yr stability
- implement the realization and unification of the terrestrial and celestial reference frames of the Earth
- improve the orbit accuracy of GNSS satellites (GPS, GLONASS, Galileo) by tracking orbits of GNSS constellations and SLR reference frame satellites against the STE-QUEST highly elliptical orbit and quasars defining the celestial frame (double-difference SLR/GNSS/VLBI approach)
- properly align the GAIA optical reference frame with the unified terrestrial and celestial reference frame and common optical/radio quasars observed at higher VLBI frequencies (that are closer to optical positions)
- determine the long-wavelength variability in the gravity field of the Earth, including central term and low-degree spherical harmonic coefficients that are either not observed or poorly observed by GRACE and GOCE gravity field missions (e.g., dynamic flattening of the Earth)
- significantly improve satellite altimetry (Jason-2, Sentinel-3) and tide gauge records of global mean sea level rise by using the highly accurate terrestrial reference frame from the STE-QUEST mission
- contribute to the monitoring of mass transport in polar regions (ice mass loss) by referencing altimetry (Cryosat, ICESat) and gravity data (GOCE and GRACE gravity missions) to the common terrestrial reference frame from the STE-QUEST mission
- contribute to the monitoring of the Earth's rotation and orientation parameters making use of the highly elliptical orbit of the STE-QUEST mission (UT1, LOD variations, etc.) and VLBI tracking from the ground
- provide a common time scale for all space geodesy techniques (GNSS, DORIS, VLBI and SLR)
- disseminate the terrestrial/celestial reference frame anywhere on Earth or in space (altimetry/gravity missions in LEO orbit, BepiColombo, etc.)

The geodetic scientific community is currently establishing a Global Geodetic Observing System (GGOS), (Rummel et al. 2000). Its objectives are the measurement of temporal changes of land, ice and ocean surfaces as well as the monitoring of mass transport processes in the Earth system and the early detection of natural hazards. Space geodesy and GGOS provide the foundation for most Earth observation and planetary missions, as well as for monitoring the Earth's geometry, gravity field and rotation, which are all related to mass

transport in the Earth system and the Earth system dynamics are related to climate change. Realizing the importance of the geodetic terrestrial reference frame and the contribution of geodesy to Earth observations, GEO (Group on Earth Observations – currently about 75 member countries) has included a specific task "Global geodetic reference frames" in its Work Plan, (GEO 2005). The main purpose of GEO is to facilitate the implementation of the Global Earth Observation System of Systems (GEOSS), see e.g., (GEO 2005).

Two of the most demanding requirements for the terrestrial reference frame are monitoring the water cycle at global to regional scales, and monitoring and modeling sea surface and ocean mass changes in order to detect global change signals in ocean currents, volume, mass, and sea level. In order to monitor all these processes in the system Earth, the background terrestrial reference frame should be accurate to a level of 1 mm RMS and be stable to a level of 0.1 mm/yr, (see e.g., (Gross et al. 2009)). Several decades of altimetry missions such as Topex/Poseidon, ERS-1/2, Jason-1/2, Envisat, Sentinel-3 and gravity field missions such as GRACE and GOCE have provided observations of the Earth system. However, the accuracy of the background terrestrial reference frame is far below that required to fully exploit the potential of all these missions. Global change processes have very long timescales and are therefore difficult to quantify. They thus require a reference frame of sufficient accuracy. On the other hand, a celestial reference frame with the Earth's rotation parameters is fundamentally important for the tracking of interplanetary satellites, navigation and planetary sciences (Rothacher et al. 2009). The highly elliptical STE-QUEST orbit and the unique suite of STE-QUEST instruments could demonstrate, for the first time, the unification of the terrestrial and celestial reference frames of the Earth and all space geodesy techniques used for its realization such as GNSS, VLBI, SLR and DORIS based on collocation. In this way, the STE-QUEST mission has immense potential to significantly improve the current accuracy of the conventional reference frames of the Earth and to meet the GGOS requirements for the terrestrial reference frame.

To use a highly elliptical orbit for terrestrial and celestial reference frame realization was first proposed in (Svehla et al. 2013b). A highly elliptical orbit is a sensor, not only for Earth rotation and orientation, but also for the estimation of low-degree spherical coefficients of the Earth's gravity field (e.g., Earth dynamic flattening) that are either not observed or poorly observed by the GRACE and GOCE gravity field missions. Geometrical mapping of the STE-QUEST orbit against extragalactic radio sources (quasars) can be realized by observing quasars at the approximate locations of the STE-QUEST satellite. This is similar to the Delta-DOR approach used in the tracking of interplanetary spacecraft. The STE-QUEST satellite dwells for a long time at the apogee of a highly elliptical orbit, thus it is a perfect target for a ground network of about 30 VLBI radio telescopes used to map the satellite orbit and the associated terrestrial frame against the positions of extra-galactic radio sources. In this way, we will be in a position, for the first time, to combine the geometrical celestial frame from VLBI and the dynamic terrestrial reference frame from GNSS constellations, SLR and DORIS satellites. In a similar way, using a double-difference SLR and GNSS approach, the orbits of GNSS satellites and SLR reference frame satellites can be dynamically mapped against the highly elliptical orbit of the STE-QUEST mission. From this point of view and with the unique suite of STE-QUEST instruments, the highly elliptical orbit of STE-QUEST is a good candidate for the combination of all space geodesy techniques (GNSS, VLBI, SLR and DORIS) and unification of the celestial and terrestrial frames.

The STE-QUEST orbit is designed to allow for long common-view frequency comparison between clocks located on different continents. Such measurements can be used to establish a global reference frame for time and the gravitational potential of the Earth. This reference frame could be used for the realization of TAI (International Atomic Time), as well as to support realization of the global height system. The first ground optical clocks achieved a frequency stability at the 10^{-18} level that corresponds to 1 cm in terms of geoid height. Temporal gravity field maps are provided routinely by the GRACE mission, however with significantly lower resolution compared to the static field. Thus, it will be very interesting to use the STE-QUEST mission to establish a unified terrestrial reference frame for positioning, time, and temporal gravity field of the Earth.

Since optically bright quasars (V<18 mag) that are covered by GAIA missions can also be observed by the ground network of VLBI radio-telescopes, the combined terrestrial and celestial reference frame from the

STE-QUEST mission could also be used to accurately align the optical GAIA reference frame. With a suite of different frequencies, the STE-QUEST metrology ground-to-space link allows differential VLBI observation of optically bright quasars at higher frequencies (Ka). Quasar positions in Ka-band are closer to optical positions (GAIA) than is the conventional celestial frame (ICRF2) defined by quasar positions based on S/X-band. We may draw the conclusion that a unified terrestrial and celestial reference frame from the STE-QUEST mission will serve Earth observation sciences, and planetary sciences at the same time (GAIA, BepiColombo).

We have unique combination of space geodesy techniques on board STE-QUEST, including GNSS, DORIS, SLR and ground and space VLBI. On top of this, all measurements are performed against the STE-QUEST clock that can be modeled with only two parameters over long period of time. On the other hand, optical and microwave metrology links developed for the STE-QUEST mission will provide range and Doppler measurements of extremely high accuracy. This opens doors for demonstration of the first geometry-free one-way positioning with the STE-QUEST mission, free of tropospheric and ionosperic effects, see Section 16. In addition, the optical metrology link offers the possibility of downloading a high volume of data from the STE-QUEST satellite, thus space VLBI signals could be sampled with the phased-array antenna at a very high sampling rate and at very high frequencies (even 1-2 orders higher than ground VLBI). In that case one could make use of relatively small phased-array antennae and cover wide frequency bands from S/X/Ka-band and higher (> 100 GHz). Due to the atmosphere, this cannot easily be performed from the ground.

There are about 30 millisecond quasars ("galactic clocks") that are providing time scale as precise as TAI, (Hobbs et al. 2012). Using signals from these 30 millisecond pulsars it is possible to derive an average pulsar time scale that has a stability better than that of atomic time, (Petit and Tavella 1996). Since the pulsar signal is observed by the ground radio-astronomical telescopes, the main limitation is due to atmospheric refraction and the use of the atmosphere radio window at relatively low microwave frequencies. An international project is underway to make use of pulsar timing for gravitational wave detection, see http://www.ipta4gw.org/. If signal from quasars are sampled at higher frequencies (against the STE-QUEST clock), these 30 millisecond quasars will be observed in the celestial reference frame at very high frequencies that, due to the atmosphere, cannot be achieved from the ground. Thus, STE-QUEST offers a very nice platform to combine space geodesy, timing and fundamental physics projects. Pulsar timing received top ranking in the "medium size" category for priorities from the Particle Astrophysics and Gravitational Panel of the Decadal Review of the National Academy in the USA.

28.8.1 Inverse Molniya Orbit and Latitude-Dwell Orbit for Highly Elliptical Orbits

Two orbit scenarios have been proposed, the first orbit design has been named the "inverse Molniya orbit" as shown in Figure 28.16 and the second orbit design is named the "latitude-dwell highly elliptical orbit", see Figure 28.17. The term Molniya comes from a series of Russian satellites called Molniya ("lightning") which have been using this type of orbit since the mid 1960s to keep mainly communication satellites over Russian territory (higher geographical latitudes) for a long time. A Molniya orbit is a highly elliptical orbit with the so-called "critical inclination" of 63.4°. The main characteristic of the orbit is the critical inclination that allows apogee-dwell, minimizing the secular effect of the J_2 perturbations on the argument of perigee.

This is very important for the STE-QUEST mission in order to make comparisons of the redshift in the orbit apogee and perigee over a long period of time. A special type of Molniya orbit has been designed named "inverse Molynia orbit", see Figure 28.16, in order to keep the ground track over major timing labs, and at the same time minimizing the perigee drift. Such an orbit scenario allows optimal non-common view clock comparison. The second orbit was designed to keep a satellite over higher latitudes over longer time periods in order to perform long common-view clock comparison. We named this type of orbit "latitude-dwell highly elliptical orbit" because the satellite appears to dwell above the Earth at the same latitude for a long period of time ($i = 30°$ or $i = 63.4°$ apogee $< 50\,000$ km altitude).
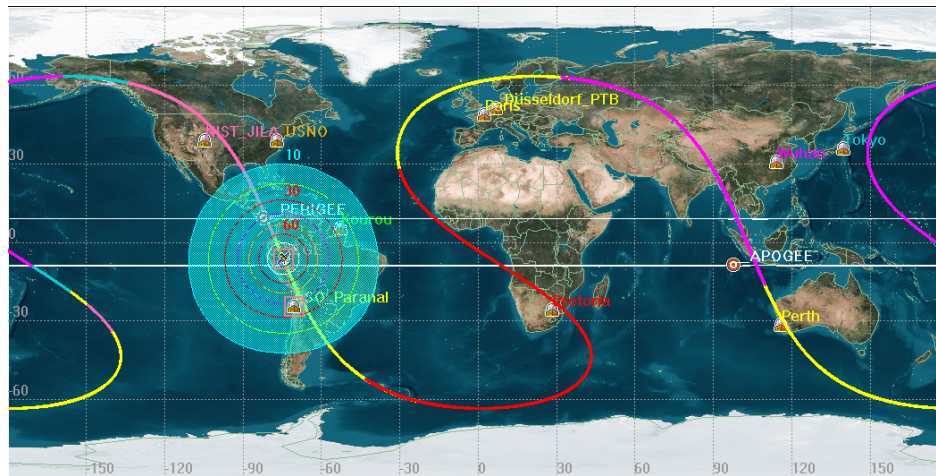
Figure 28.16 Inverse Molniya orbit, one of the proposed orbits for the STE-QUEST mission ($i = 63.4°$, apogee $\approx 38\,000$ km altitude). The satellite ground track passes over the major timing labs, allowing optimal non-common view clock comparison. The simulated period is two months, elevation angle every 10° (in cyan).

The latitude over which it dwells is determined by the orbit inclination and the number of "latitude-dwells" depends on the orbital period and sidereal day. Three "latitude dwells" were initially considered.

In the case of the inverse Molniya orbit (Scenario 1) typical contact times for perigee ground stations are $12 - 24$ min ($10°$ elevation cut-off), and for the apogee stations $4 - 8$ h . The common-view contact time is about 6 h in Europe, while between Europe and the USA it is ≈ 3.5 h . In the case of the "latitude-dwell highly elliptical orbit" (Scenario 2) the satellite passes over the same station every 48 h . Contact durations are $30 - 40$ min for the perigee ground stations and about $12 - 15$ h for the apogee stations. Scenario 2 provides about 13 h of common-view contact within Europe and about 12 h between Europe and the USA.

The selected STE-QUEST orbit scenario is the latitude-dwell highly elliptical orbit with a critical inclination $i = 63.4°$ and orbit period of 16 hours, with the repeated ground track after 3 days above the three selected ground timing labs in Boulder, Tokyo and Turin (three "latitude-dwells"). Orbital elements in the STE- QUEST proposal re-submitted to the ESA Cosmic Vission Call in 2015 are displayed in Table 28.3.
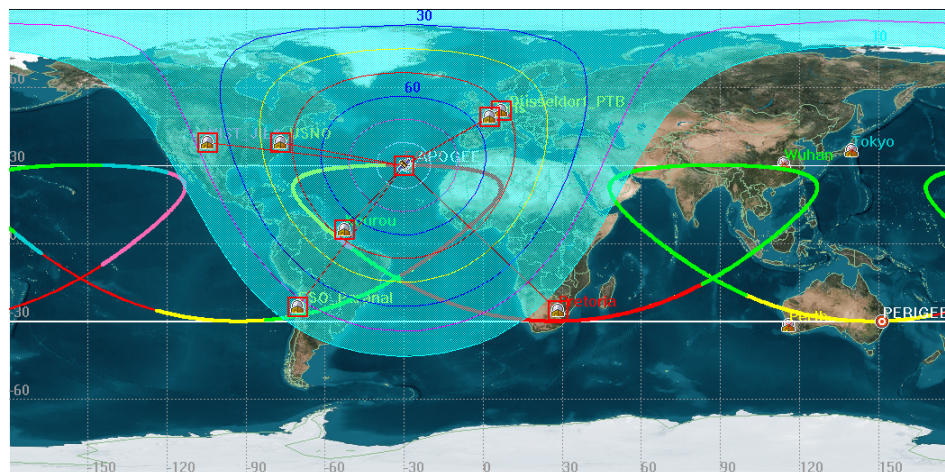
Figure 28.17 Latitude-dwell highly elliptical orbit of the STE-QUEST mission ($i = 30°$ or $i = 63.4°$ apogee $< 50\,000$ km altitude). The satellite maintains constant latitude for a long time, allowing long common-view clock comparisons and space/ground VLBI. The simulated period is two months, ground elevation angle steps are shown in cyan, given every 10°. The selected inclination for STE-QUEST is $i = 63.4°$.

Period	10.6 h
Semi-major axis	24 450 km
Eccentricity	0.636
Apogee Altitude	~33600 km
Perigee Altitude	~2500 km
Inclination	63.4° (argument of perigee 270°)
RAAN	265°

Table 28.3 The final orbit design in the STE-QUEST proposal re-submitted to the M4 Call of the ESA Cosmic Vision Programme in 2015.

28.8.2 Third-Body Perturbations and a Highly Elliptical Orbit

What is the impact of the dynamics of a highly-elliptical orbit on estimated reference frame parameters? Can rates in LOD and nutation be better estimated with a highly-elliptical orbit? Performing orbit simulations for the STE-QUEST mission over a longer period of time, it was noticed that highly elliptical orbits with their apogees very high above the Earth exhibit very significant third-body perturbations arising from the Sun and the Moon, especially in terms of secular variations. The main effect is in the longitude of the ascending node and the argument of perigee. These secular variations in the third-body perturbations can be explained by the "gyroscopic" precession of the highly elliptical orbit about the ecliptic pole. The effect of lunar perturbations on the orbits of spacecrafts around the Earth can be found in (Domingos et al. 2008). A semi-analytical and numerical study is presented of the perturbation caused by a third-body using a double averaged analytical model. Following (Domingos et al. 2008), the disturbing function R in terms of Legendre polynomial P_n is

$$R = \frac{\mu' G(m_0 + m')}{r'} \sum_{n=2}^{\infty} \left(\frac{r}{r'} \right)^n P_n \left(\cos(S) \right) \tag{28.32}$$

where m_0 and m' are the mass of the central body and of the perturbing body respectively, mass ratio $\mu' = m'/(m_0 + m')$, G is the gravitational constant, S is the central elongation between the perturbed body (spacecraft) with the radius vector r and the perturbing body (Moon or Sun) with the radius vector r'. After eliminating terms due to the short periodic motion, one can obtain the evolution of the mean orbital elements for a long-time period. An analytical model based on Lagrange's planetary equations for the third-body perturbations, for the case where the perturbing body is in an elliptical orbit, is according to (Domingos et al. 2008)

$$\frac{da}{dt} = 0$$

$$\frac{de}{dt} = \frac{15\mu' n'^2 e\sqrt{1-e^2}}{8n} \left[1 + \frac{3}{2} e'^2 + \frac{15}{8} e'^4 \right] \sin^2 i \sin 2\omega$$

$$\frac{di}{dt} = \frac{-15\mu' n'^2 e^2}{16n\sqrt{1-e^2}} \left[1 + \frac{3}{2} e'^2 + \frac{15}{8} e'^4 \right] \sin 2i \sin \omega$$

$$\frac{d\omega}{dt} = \frac{3\mu' n'^2}{8n\sqrt{1-e^2}} \left[1 + \frac{3}{2} e'^2 + \frac{15}{8} e'^4 \right] \left[(5\cos^2 i - 1 + e^2) + 5(1 - e^2 - \cos^2 i)\cos 2\omega \right] \tag{28.33}$$

$$\frac{d\Omega}{dt} = \frac{3\mu' n'^2}{8n\sqrt{1-e^2}} \left[1 + \frac{3}{2} e'^2 + \frac{15}{8} e'^4 \right] \left[5e^2 \cos 2\omega - 3e^2 - 2 \right] \cos i$$

$$\frac{dM}{dt} = \frac{-\mu' n'^2}{8n} \left[1 + \frac{3}{2} e'^2 + \frac{15}{8} e'^4 \right] \left[(3e^2 + 7)(3\cos^2 i - 1) + 15(1 + e^2)\sin^2 i \cos^2 \omega \right]$$

where $(a, e, i, \omega, \Omega, M)$ are 6 Keplerian parameters for the satellite orbit and (n', e') the corresponding Keplerian parameters of the orbit of the perturbing body. The model essentially depends on inclination i, eccentricity e and the argument of periapsis ω. It is interesting to note that the semi-major axis a is not perturbed at all and that the equation for the longitude of the ascending node Ω depends on the eccentricity, inclination and argument of periapsis, but does not influence their perturbations. Furthermore, one can draw the conclusion that in elliptic restricted three-body problems, the evolution over time of the orbital elements of the satellite depend on its initial state, on the eccentricity of the disturbing body, and on the mass ratio μ'. If the distance between the central body and the satellite increases, gravitation of the central body decreases with the square of this distance, so the perturbations of the third body become more important. The increase of the distance between the central body and the satellite may cause regions of stable orbits, quasiperiodic or chaotic orbits (Domingos et al. 2008). So, escape or collision of the satellite may occur as well. For the nearly circular orbits

$$\frac{da}{dt} = 0, \qquad \frac{de}{dt} = 0, \qquad \frac{di}{dt} = 0 \tag{28.34}$$

only perturbations in the longitude of the ascending node, argument of perigee, and mean anomaly should be considered. When $e = 0$ or $i = 0$ there are no perturbations of inclination or eccentricity, and the orbit remains circular and/or planar. These circular orbits with constant inclination appear due to the truncation of the expansion of the disturbing function and are not a physical phenomenon. In the real case (full restricted three-body problem), circular solutions with constant inclination do not exist (Domingos et al. 2008). One can see that eccentricity of the spacecraft increases with eccentricity e' of the perturbing body, which can be explained by the decrease of the minimum distance between the main bodies. Therefore, the perturbations are at a maximum when the secondary body is near the pericenter of its orbit. Interestingly, there is a decrease in the inclination when the eccentricity increases.

When reference frame parameters are determined using a highly elliptical orbit one needs to take into account secular precession and apsidal precession of both Earth flattening and third body perturbations

$$\frac{d\Omega}{dt}(\mu', e', \omega) = \frac{3\mu' n'^2}{8n\sqrt{1 - e^2}} \left[1 + \frac{3}{2}e'^2 + \frac{15}{8}e'^4\right] \left[5e^2 \cos 2\omega - 3e^2 - 2\right] \cos i$$

$$\frac{d\Omega}{dt}(C_{20}) = \frac{3n C_{20} a_e^2}{2(1 - \mathrm{e}^2)^2 a^2} \cos i \tag{28.35}$$

Therefore, differential changes in length of day LOD, obliquity and longitude as a function of differential changes in the right ascension of the ascending node due to lunar third body perturbations can be separated from the J_2 effect. They heavily depend on the Moon's position in the case of a highly elliptical orbit. In this way, a highly elliptical orbit becomes a sensor to monitor Earth rotation and orientation. In addition, the critical inclination of 64.3° due to the J_2 coefficient of the gravity field of the Earth is strongly affected by lunar secular third-body perturbation. With the symmetric negative term $(1 - 5\cos^2 i)$ that appears in both expressions, these two apsidal precessions tend to compensate each other

$$\frac{d\omega}{dt}(\mu', e', \omega) = \frac{3\mu' n'^2}{8n\sqrt{1 - e^2}} \left[1 + \frac{3}{2}e'^2 + \frac{15}{8}e'^4\right] \left[(5\cos^2 i - 1 + e^2) + 5(1 - e^2 - \cos^2 i) \cos 2\omega\right]$$

$$\frac{d\omega}{dt}(C_{20}) = \frac{3n C_{20} a_e^2}{4(1 - e^2)^2 a^2} (1 - 5\cos^2 i) \tag{28.36}$$

In that case, critical inclination is not constant, but in addition is a function of eccentricity e of a highly elliptical orbit and argument of perigee ω.

One can draw the general conclusion that a highly elliptical orbit is a sensor, not only for Earth rotation and orientation, but also for the estimation of low order spherical coefficients, especially the J_2 coefficient of the Earth's gravity field. In addition, the satellite dwells for a long time at the apogee of a highly elliptical orbit, and is thus a perfect target for VLBI. This can be used to map the satellite's dynamics against the positions of extra-galactic radio sources in order to combine a geometric celestial reference frame from VLBI and a dynamic terrestrial reference frame. From this point of view, a highly elliptical orbit is the best orbit for the combination of all space geodesy techniques and for unifying celestial and terrestrial reference frames. This is not the case with LEO orbit, where satellites can be observed with VLBI and other space geodesy techniques for only a very short period of time. For LEO orbit, lunar third-body perturbations are very much uniform along the orbit, thus the orbit precesses only due to the J_2 coefficient of the Earth's gravity field.

28.9 Two Equally Precessing LEO and HEO Orbits

How to design a gravity field mapping concept that could combine a reference frame mission and a gravity field mission at the same time, as well as the high-resolution mapping of the static and temporal gravity field of the Earth? Since low degree coefficients contain the terrestrial reference frame, such a mission should also combine gravity. We consider a scenario with two equally precessing LEO and HEO orbits $\dot{\Omega}_{LEO} = \dot{\Omega}_{HEO}$

$$\cos i_{HEO} = \frac{a_{HEO}^3 \sqrt{a_{HEO}}}{a_{LEO}^3 \sqrt{a_{LEO}}} \frac{(1 - e_{HEO}^2)^2}{(1 - e_{LEO}^2)^2} \cos i_{LEO} \tag{28.37}$$

and assuming both orbits to be circular

$$\cos i_{HEO} = \frac{a_{HEO}^3 \sqrt{a_{HEO}}}{a_{LEO}^3 \sqrt{a_{LEO}}} \cos i_{LEO} \tag{28.38}$$

For example a LEO orbit at $200\,\text{km}$ altitude and inclination $i_{LEO} = 89°$, will require a circular HEO orbit at an altitude of some $30\,000\,\text{km}$ and $i_{HEO} = 77°$. For a circular HEO orbit at $20\,000\,\text{km}$ altitude inclination will be $i_{HEO} = 88.5°$. Such a formation can be further optimized by introducing eccentricity into the HEO orbit and perturbations due to Sun and Moon.
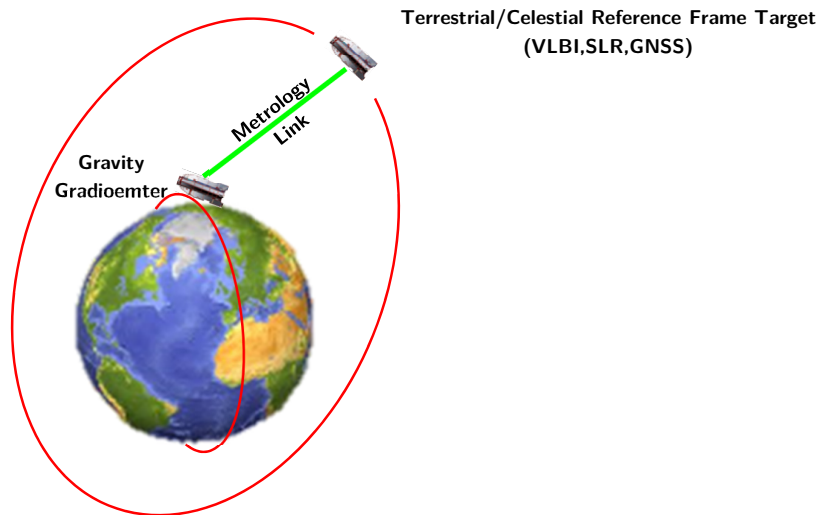
Figure 28.18 Gravity field mapping concept based on two equally precessing high-low orbits.

Both satellites could be launched together and the second satellite deployed in higher orbit using ion propulsion or a deploying vehicle. The satellite in the higher orbit will be an excellent target for VLBI, SLR and GNSS to combine all space geodesy techniques in combined realization of the terrestrial and celestial frame.

29. Geometrical Representation of Gravity

T o use atomic clocks for the *in situ* determination of differences in the gravitational potential of the Earth's gravity field was proposed for the first time by (Bjerhammar 1985) and (Vermeer 1983), taking up Einstein's postulation that two atomic clocks will tick at different rates due to different gravity potential values at different locations. However, this concept has not been demonstrated so far due to limitations in comparing clock frequency at $\leq 10^{-18}$ relative accuracy between two distant locations. Recently, a frequency transfer was demonstrated below 10^{-18} relative accuracy over a distance of ca. 920 km using an optical fibre (Predehl et al. 2012), with only one optical clock placed at one end of the optical fibre and a H-maser at the other end. In (Švehla and Rothacher 2005b) it was proposed to use atomic clocks in space to measure the gravitational potential along an orbit, to measure together with GNSS, both position and gravity in a purely geometrical way. Here we provide the physical background to relativistic geodesy that is not given in (Bjerhammar 1985) and, based on this, provide a geometrical representation of gravity and its relation to orbital motion and reference frames for time. We also show that in special cases, it is possible to measure absolute gravity potential values using quantum mechanics, which opens up new possibilities for the use of state-of-the-art optical clocks. Beyond the Standard Model in theoretical physics based on four fundamental forces, gravitation is still separated from the electromagnetic, strong nuclear, and weak nuclear interactions that are successfully related by the quantum field theory at the level of atomic, particle and high energy physics. On the other hand, general relativity brilliantly describes all observed phenomena related to gravitation in our Solar System and at galactic and cosmological scales. However, general relativity is fundamentally incomplete, because it does not include quantum effects. A unified theory relating all four known interactions will represent a step towards the unification of all fundamental forces of nature. Here we show that circular perturbations could provide an interesting representation between quantum mechanics and orbit mechanics. We try to establish an equivalence between the orbit mechanics based on circular perturbations and basic principles of quantum mechanics. We show that gravity at quantum level and at celestial level can be represented with the same property as light, i.e., gravity and light can be represented as oscillating at the equivalent rate and thus propagate at the same rate. In the essence of every orbit one could consider a wave represented by matter and time that could be modelled or represented by two geometrical rotations. We try to represent gravitational potential by two geometrical counter-rotations, with the rotation of spherical harmonic coefficients as generating functions. This dualistic concept is similar to the electromagnetic force where electricity and magnetism are elements of the same phenomenon orthogonal to each other. Following the general relativity, any form of energy that couples with spacetime creates differential geometrical forms that can describe gravity. Thus, gravitation can be considered purely as a geometrical property. However, our geometrical representation using two counter-oscillations (bi-circular orbits) can be considered as describing gravitation from the scalar point of view at the quantum as well as at the celestial level. Thus it gives geometrical and scalar properties of gravitation at the same time. This is similar to the concept of a magnetic field generated on top of an existing electric field, or similar to the concept of matter and antimatter in particle physics, where antimatter is described as material composed of antiparticles with the same mass as particles,

but with opposite charge (leptons, baryons). Following recent results from the Planck mission (Planck Collaboration et al. 2013), there is strong evidence that 26.8.% of the mass-energy of the Universe is made of non-baryonic dark matter particles, which should be described by the Standard Model.

29.1 Recent Theories of Gravity

To give an overview of the geometrical frame of orbit reference, we need to go back a few centuries and perhaps it would be best to start with Galileo Galileo back in the 17ᵗʰ century

When, therefore, I observe a stone initially at rest falling from an elevated position and continually acquiring new increments of speed, why should I not believe that such increases take place in a manner which is exceedingly simple and rather obvious to everybody? -- Galileo Galilei, 1638

Following Newton, space and time exist absolutely and independently from each other (Newton). Massive bodies such as the Earth and planets do not affect their existence. Later on, Einstein defined the speed of light as a constant (Einstein 1905) and stated that space and time exist in an absolute way, but are not independent from each other. He named this duality spacetime (Einstein 1916). Following Einstein's Theory of Relativity (General Relativity), matter in the form of massive rotating bodies such as the Earth and planets warp and twist spacetime in their vicinity. At the same time, spacetime governs the motion of matter. In Einstein's spacetime, gravity is a curvature of spacetime caused by massive bodies, e.g., Earth, whereas in Newton's space gravity is an instantaneous attraction force between two objects. Over the 350 years from the time of Newton to the time of Einstein, the concept of speed of gravity changed the value by a very high amount, from infinity to the speed of light we know today. Due to the curvature of spacetime, the orbit of a satellite around the Earth in Euclidian space will be perturbed. The existence of Einsteins's General Relativity has been confirmed by a number of different tests and phenomena, including the precession of the perihelion of the planet Mercury, light bending near a massive celestial object, Shapiro delay and tests with gravitational redshift using atomic clocks (Pound and Rebka 1959). The so-called Pound-Rebka experiment of 1959 started a new era of testing the General Theory of Relativity. The Vessot experiment of 1980 (Vessot et al. 1980) was the first test of gravitational redshift using a *space-borne hydrogen maser.* One test of Einsteins's warped spacetime currently underway is the Gravity Probe-B mission (GP-B). The idea is to observe the deflection of the spinning axis over time of a rotating gyroscope placed in Earth orbit, so-called geodetic precession or de Sitter precession. In the vicinity of the Earth, spacetime is not only warped. Due to the rotation of the Earth, spacetime is also twisted. The rotation of the Earth drags local spacetime causing the so-called Lense-Thirring or frame-dragging effect. In Newton's formulation, the rotation of a satellite about a spinning axis or an orbital axis of a Keplerian orbit around a central body will be fixed in space (central gravity term). As a result of frame-dragging, the orbit of a satellite in a polar orbit will precess along the equator and in the case of a gyro placed in a polar Earth orbit (Gravity Probe-B) the spin axis will be deflected mainly in the cross-track direction chasing the Earth's rotation. In the case of a polar orbit, both effects, the geodetic precession and the frame-dragging are perpendicular to each other and separable (Gravity Probe-B). A recent test of the Lense-Thirring effect, based on the latest gravity models from the CHAMP mission, is available in (Ciufolini and Pavlis 2004). An indirect measurement of the speed of gravity using Shapiro time delay was performed by (Kopeikin 2003) by means of VLBI measurements of the quasar QSO J0842+1835, while it was nearly aligned with the planet Jupiter. It confirmed that an arbitrarily moving gravitating body deflects light not instantaneously, but with a small retardation caused by the finite speed of gravity propagating from the body to the light ray. This test claims that gravitation propagates at the speed of light. Gravitational waves have been confirmed recently, see (LIGO Scientific Collaboration and Virgo Collaboration et al. 2016). General relativity predicts gravitational radiation, when the energy is transported by gravitational waves and radiated by celestial objects in motion. When matter accelerates, it emits gravitational waves. They can be described as a fluctuation in the curvature of spacetime that propagates outward from an object at the speed of light.

The ESA-NASA mission LISA (Laser Interferometer Space Antenna) is designed to confirm gravitational waves over very large distances in space. Three pairs of gold-platinum particles will be set at the apexes of a triangle with sides five million kilometres long. These pairs of free-falling particles, when hit by a gravitational wave, will undergo a tiny oscillating acceleration (relative to each other) measured by a laser interferometer.

Beyond the Standard Model in theoretical physics based on four fundamental forces, gravitation is still separated from the electromagnetic, strong nuclear, and weak nuclear interactions that are successfully related by the quantum field theory at the level of atomic, particle and high energy physics. On the other hand, general relativity brilliantly describes all observed phenomena related to gravitation in our Solar System and at galactic and cosmological scales. However, general relativity is fundamentally incomplete, because it does not include quantum effects. The state-of-the-art in the theory of gravitation is quantum gravitation, describing gravitation with the principles of quantum mechanics. A unified theory relating all four known interactions will represent a step towards the unification of all four fundamental forces of nature.

Geometrical model of two orthogonal counter-oscillations (bi-circular orbits) could be considered to represent an orbit as a wave and subsequently to describe gravitation in a dualistic way. For this wave nature of the orbit, we show that a similar constant to a Plank constant at the quantum level could be established representing the specific angular momentum of the orbit. This dualistic concept could consider gravitation as being similar to a electromagnetic force with a magnetic field orthogonal to an existing electric field, or similar to the concept of matter and antimatter in particle physics, introducing the concept of antiparticles with the same mass as particles, but with opposite charge.

29.2 The Physics Background to Relativistic Geodesy

General relativity, or generally speaking, metric theories of gravity, are based on the equivalence principle (EEP) stating that local effects of gravity are the same in both a static and in an accelerated reference frame. The EEP is based on three cornerstones: the weak equivalence principle, or universality of free fall (independency from the composition and structure of the body in free fall); local Lorentz invariance (independence from the velocity of the frame) and a local position invariance (independence from the position of the frame). Local position invariance says that the results of a non-gravitational experiment, e.g., the frequency of an oscillating system in a clock, are independent of the spatial and temporal coordinates of the experiment. Local position invariance can be tested by measuring gravitational redshift, (see e.g., (Vessot et al. 1980)). The ACES mission on board the ISS aims to demonstrate (in 2018) measurements of the redshift, and thus local position invariance in space, to an uncertainty level of 2×10^{-6} (Cacciapuoti and Salomon 2009). A completely new approach to measuring gravitational redshift was demonstrated by an experiment based on the quantum interference of atoms which showed the interference of matter waves (Muller et al. 2010). However, the distance scale of the experiment carried out in (Muller et al. 2010) varies from micrometer to millimeter compared to hundreds of km for LEO orbits, e.g., 400 km in the case of the altitude of the ACES mission.

In order to gain an insight into the gravitational redshift effect and put it into the context of a satellite orbit, we need first to introduce Einstein's famous mass-energy equivalence equation (Einstein 1905), where the total energy of a body is given by

$$E = m \cdot c^2 \tag{29.1}$$

with mass (rest mass) m, speed of light in vacuum $c = 299792458$ m/s . (29.1) states that energy and mass are linearly equivalent, the ratio being the square of the speed of light in vacuum. In other words, the energy content of a body can be measured through its mass. Following (Einstein 1905) and (Einstein 1916), the speed of light is the maximum velocity in the universe at which energy and matter can travel as postulated by special relativity. This is the velocity of all massless particles and that of the propagation of the associated fields of which they are the quanta (e.g., electromagnetic fields, including light) as well as the velocity of gravity, i.e., of gravitational waves, the velocity at which changes in a gravitational field propagate through

space. However, in Newtonian gravitation any change in the mass leads to an instantaneous adjustment of the gravitational field. Thus Newtonian gravitation has an infinite speed. This is used in the post-Newtonian parameterization (PPN) of gravity applied in precise orbit determination that correctly accounts for these effects in orbit mechanics by means of relativistic corrections to the equation of motion in Euclidian space (geodesic precession, Lense-Thirring effect).

The idea of using mass to measure energy leads us to the current activities in the re-definition of SI units, since the mass of the International Prototype of the Kilogram (made from platin-iridium and located at BIPM in Paris) is still the official mass unit of the SI system (since 1889). Following the current re-definition of SI units by BIPM, mass will be re-defined in the SI system by the equivalent energy of a photon via Plank's constant in quantum mechanics.

The basis of all frequency measurements is the Planck relation that gives the relationship between the energy of a photon and the frequency f of the associated electromagnetic wave

$$E = h \cdot f \tag{29.2}$$

where the Planck constant is given as $h = 6.6260695729 \cdot 10^{-34}$ J·s . The Planck constant is the quantum of action in quantum mechanics and gives the proportionality of the energy of a photon and the frequency of its associated electromagnetic wave. If we now introduce angular frequency $\omega = 2\pi f$ or the wavenumber $k = 2\pi/\lambda$ we derive

$$E = \frac{h}{2\pi} \cdot \omega = \hbar \cdot \omega = \hbar \cdot k \tag{29.3}$$

where $\hbar$ is the reduced Planck constant or Dirac constant. A photon is an elementary particle and the force carrier for the electromagnetic force. Photons exhibit wave-particle duality, having properties of both waves and particles. Thus, the photon is the quantum of light and all other forms of electromagnetic radiation. Photons are elementary particles emitted in many natural processes. During a molecular, atomic or nuclear transition to a lower energy level, photons are emitted with an energy level spanning the electromagnetic spectrum. Photons are massless particles, thus following the equivalence principle. They experience the same gravitational acceleration as other particles, because although they are massless they have relativistic mass. However, to understand this we need to introduce the relativistic mass of the particle and the concept of matter waves introduced by (de Broglie 1924). The de Broglie relation shows that the wavelength of its matter waves (de Broglie wavelength) is inversely proportional to the momentum of a particle, p

$$\lambda = \frac{h}{p} \tag{29.4}$$

whereas the frequency (of the matter waves) f is related to its kinetic energy E

$$f = \frac{E}{h} \tag{29.5}$$

Thus, the concept of matter waves or de Broglie waves, accurately reflects the wave-particle duality of matter and elementary particles such as photons or electrons.

Ther relativistic momentum in special relativity is given as

$$p = \gamma \cdot m_0 \cdot v \tag{29.6}$$

where the velocity v of the particle that has non-zero mass is always $v < c$. The particle's rest mass (w.r.t. relativistic mass in motion m) is denoted by m_0

$$m = \gamma \cdot m_0 \tag{29.7}$$

and where the Lorentz factor is $\gamma = 1 / \sqrt{1 - v^2 / c^2}$, (see (Moritz 1993)). In special relativity, the Lorentz factor relates time dilatation and space contraction between a moving frame and a frame at rest

$$
\begin{aligned}
\Delta t' &= \gamma \cdot \Delta t \\
\Delta x' &= \Delta x / \gamma
\end{aligned}
\tag{29.8}
$$

where Δt and Δx denote the time and the space intervals in the frame at rest and $\Delta t'$ and $\Delta x'$ those in the moving frame with proper time t' (Moritz 1993). In (29.6), v represents the particle's velocity (group velocity) that, in principle, could be different from the phase velocity (particle's frequency times wavelength) in a non-dispersive medium.

Combining special relativity and de Broglie's matter-waves relationship, particles with inertial mass, i.e., every quantum of energy E of matter (photon, electron, atom, etc.) have an inertial mass m_0 and an inertial energy $E_0 = m_0 \cdot c^2$ in a frame at rest, i.e.,

$$
h \cdot f_0 = m_0 \cdot c^2
\tag{29.9}
$$

see also (Solarić et al. 2012). Finally, following the de Broglie relation (29.4) and (29.5) we obtain, for the wavelength of a wave associated with an elementary particle (photon, electron, etc.)

$$
\lambda = \frac{h}{\gamma m_0 v} = \frac{h}{m_0 v} \sqrt{1 - \frac{v^2}{c^2}}
\tag{29.10}
$$

and for its frequency

$$
f = \frac{\gamma m_0 c^2}{h} = \frac{m_0 c^2}{h \sqrt{1 - \dfrac{v^2}{c^2}}}
\tag{29.11}
$$

If we now multiply (29.10) by (29.11) we derive the phase/group velocity of the matter waves

$$
v_p = v_g = f \cdot \lambda = \frac{c^2}{v} = \frac{E}{p} > c
\tag{29.12}
$$

As we know from special relativity, the velocity v of a particle that has non-zero mass is always $v < c$, thus the phase velocity of matter waves always exceeds c, but for a photon we have $v = c$ and the frequency of the matter waves reduces to

$$
f = \frac{m_0 c^2}{h}
\tag{29.13}
$$

Therefore, to calculate the fractional frequency shift Δf due to the velocity of the moving frame, for photons we can directly make use of the time dilation due to the Lorentz factor γ (29.8). Denoting the period of oscillation T in a moving frame with frequency f_B and frequency f_A in the frame at rest we have

$$
\frac{f_B - f_A}{f_A} = \frac{\Delta f}{f} = \frac{\frac{1}{\gamma T} - \frac{1}{T}}{\frac{1}{T}} = \frac{1 - \gamma}{\gamma}
\tag{29.14}
$$

Since the velocity of the satellite $v \ll c$, we may write, without any significant loss of accuracy,

$$
\gamma = \frac{1}{\sqrt{1 - v^2 / c^2}} \approx 1 + \frac{v^2}{2c^2}
\tag{29.15}
$$

And finally, the fractional frequency offset due to the velocity of the clock v is

$$\frac{f_B - f_A}{f_A} = \frac{\Delta f}{f} = -\frac{v^2}{2c^2} \tag{29.16}$$

From (29.16) we can draw the conclusion that, due to the velocity of the clock v on the rotating Earth or on board the satellite, the frequency of the clock in a moving frame will be lower than in the frame at rest, as stated by the second order Doppler effect in special theory of relativity. Thus, by increasing the satellite velocity we are proportionally decreasing the time rate of the clock.

Let us now see how to calculate the total energy along the orbit in celestial mechanics. In the case of the central gravitational field of the Earth we may write the gravitational potential in short form as $V = GM/r$ whereas the potential energy of the orbit needs to be written with a negative sign $E_{pot} = -GM \cdot m/r$, see e.g., (Montenbruck and Gill 2000). The total energy along the orbit can be then calculated as $E_{tot} = -GM \cdot m/r + mv^2/2 = -GM \cdot m/(2a)$ ($m << M$). According to this convention, the potential energy is higher with higher altitude, whereas the gravitational potential is higher towards the planetary body generating the potential. Considering this convention, the potential energy of a photon at a point A in the gravitational field is given by

$$E_A = -V_A \cdot m \tag{29.17}$$

and after changing location from point A to point B in the gravitational field, the differential change in the potential energy of the photon will be

$$\Delta E = E_B - E_A = -(V_B - V_A)m = -(V_B - V_A)\frac{h \cdot f}{c^2} \tag{29.18}$$

or relative to the original potential energy E_A at point A

$$\frac{\Delta E}{E_A} = \frac{E_B - E_A}{E_A} = -\frac{(V_B - V_A)}{c^2} = -\frac{\Delta V}{c^2} \tag{29.19}$$

However, if we change the energy of a photon by a small amount, we need to take into account the Plank relation (29.2) and the associated frequency of the photon will be changed

$$\Delta E = E_B - E_A = h \cdot f_B - h \cdot f_A \tag{29.20}$$

For a given frequency of the photon at point A we obtain

$$\frac{\Delta E}{E_A} = \frac{f_B - f_A}{f_A} = \frac{\Delta f}{f} \tag{29.21}$$

Finally, from (29.19) and (29.21), we obtain a relationship between the changes in the frequency of the photon Δf as a function of the change in the gravitational potential ΔV

$$\frac{\Delta f}{f} = \frac{f_B - f_A}{f_A} = -\frac{(V_B - V_A)}{c^2} = -\frac{\Delta V}{c^2} \tag{29.22}$$

Thus, by measuring the frequency offset between two points in the gravitational potential field it is possible to measure the differences in gravitational potential. In terms of quantum angular frequency $\omega = 2\pi f$, from (29.3) we have

$$\frac{\Delta \omega}{\omega} = \frac{\omega_B - \omega_A}{\omega_A} = -\frac{(V_B - V_A)}{c^2} = -\frac{\Delta V}{c^2} \tag{29.23}$$

Thus, the same photon will carry a different quantum of energy, depending on its location in the gravitational field and the clock rate will be higher towards the lower gravitational potential.

Although, for the last 50 years the electromagnetic oscillation which drives absorption in a cesium atom has been used by atomic clocks to define the SI second, recently (Lan et al. 2013) and (Debs et al. 2013) demonstrated for the first time a completely different approach to lock an atomic clock to the mass of the particle itself. They used momentum spectroscopy of an atom to stabilize the atomic clock. If we divide (29.9) by 2π we have

$$\hbar \cdot f_0 = \frac{1}{2\pi} m_0 \cdot c^2 \tag{29.24}$$

or the so-called the Compton frequency ω_C

$$\omega_C = \frac{1}{\hbar} m_0 \cdot c^2 \tag{29.25}$$

Thus, for a single particle of mass m_0 we may calculate the Compton frequency ω_C from (29.25) and use it as a reference frequency in the clock to enable high-precision mass measurements and the fundamental definition of the second and meter. We will see later that the Compton frequency can be very large in size, but a frequency comb can be used to overcome this problem.

If we now compare the Compton frequency ω_C in (29.25) with our expression for the quantum angular frequency $\Delta\omega / \omega$ along the orbit, we see that, indirectly, we can relate variations in the gravitational potential with the relativistic mass of the particle. For this, one would need to measure the Compton frequency ω_C along the satellite orbit.

The results of (Lan et al. 2013) and (Debs et al. 2013) have very important implications in metrology, fundamental physics and for the definition of the kilogram against the unit of time and length.

29.3 Is it Possible to Measure Absolute Gravitational Potential Using Optical Clocks?

It is clear that with optical clocks one could measure only relative gravitational potential between two points in the Earth's gravitational field. However, is there a special case, e.g., for a sufficiently high altitude, where one could also consider measuring absolute potential? Considering only the central term of the gravitational field of the Earth $V = GM/r$ (high Earth's orbits, interplanetary missions) we have

$$\frac{\Delta f}{f} = \frac{f_B - f_A}{f_A} = -\frac{(V_B - V_A)}{c^2} = -\frac{\Delta V}{c^2} = \frac{GM}{c^2} \cdot \frac{r_B - r_A}{r_A \cdot r_B} \tag{29.26}$$

From (29.26) we see that, due to the gravitational potential, the fractional frequency shift of the clock will be increased by increasing the orbit altitude, i.e., the clock rate will be higher further away from the planetary body generating the gravitational potential, as postulated by general relativity. From (29.26) we can determine absolute potential V_A or V_B considering only the central term of the gravity field of the Earth

$$\frac{\Delta f}{f} = \frac{f_B - f_A}{f_A} = \frac{V_A}{c^2} \cdot \frac{r_B - r_A}{r_B} = \frac{V_B}{c^2} \cdot \frac{r_B - r_A}{r_A} = \frac{V_B}{c^2} \cdot \frac{\Delta r}{r_A} \tag{29.27}$$

with $\Delta r = r_B - r_A$, from where it follows that the geometrical definition of the gravitational potential for the central term of the gravity field of the Earth is

$$V_A := \frac{f_B - f_A}{f_A} \cdot \frac{r_B}{r_B - r_A} c^2 = \frac{\Delta f}{\Delta r} \cdot \frac{r_B}{f} c^2 = \frac{\Delta f}{f} \cdot \frac{r_B}{\Delta r} c^2 \tag{29.28}$$

or for V_B

$$V_B := \frac{f_B - f_A}{r_B - r_A} \cdot \frac{r_A}{f_A} c^2 = \frac{\Delta f}{\Delta r} \cdot \frac{r_A}{f} \cdot c^2 = \frac{\Delta f}{f} \cdot \frac{r_A}{\Delta r} \cdot c^2 \tag{29.29}$$

Can we measure absolute gravitational potential on the Earth surface relative to the high orbit altitude, where gravitational potential is well known to resemble the properties of the central gravitational field? Thus, by measuring frequency offsets between two points in the gravitational potential field, it is possible to measure absolute gravitational potential, either at the first or at the second point considering only the central term of the gravity field of the Earth. For a ground station and a GNSS orbit altitude the ratio $r_A / (r_B - r_A) \approx -1.3$ and for a ground station and a GEO satellite $r_A / (r_B - r_A) \approx -1.2$. Therefore, from (29.28) and (29.29), we can draw the conclusion that absolute gravitational potential can indeed be measured with 1 cm accuracy in terms of geoid height on the ground or in LEO orbit using the reference optical clock at high altitude. For this, we would need to have two optical clocks with an absolute frequency accuracy of $\Delta f/f \approx 10^{-18}$ and to separate them over a large radial distance in the gravitational field, i.e., for a ground clock or LEO clock, the second clock would need to be placed at least at the GNSS altitude or in GEO. In that case, the second clock at very high orbit altitude will only be affected by the low-order gravity field coefficients that are well determined from the space gravity field missions. Typically, in numerical integration GPS orbit is sensitive to degree and order $n \leq 10$ of the Earth's gravity field. With a second optical clock in high Earth orbit, the relative gravitational potential difference between ground and space will be a high-frequency (absolute) gravitational potential on the Earth's surface measured by a ground optical clock relative to the optical clock in high Earth orbit.

If the gravitational field is represented by spherical harmonics up to degree n and order m

$$V(r,\theta,\lambda) = \frac{GM}{r} + \frac{GM}{r} \sum_{n=2}^{\infty} \left(\frac{a}{r}\right)^n \sum_{m=-n}^{n} c_{nm} Y_{nm}(\theta,\lambda) \tag{29.30}$$

one would need to measure differences in the gravitational potential over the entire sphere

$$\frac{\Delta f}{f} = -\frac{\Delta V}{c^2} = \frac{GM}{c^2} \cdot \frac{r_B - r_A}{r_A \cdot r_B} + \\ + \frac{GM}{c^2} \cdot \left[\frac{1}{r_A} \sum_{n=2}^{\infty} \left(\frac{a}{r_A}\right)^n \sum_{m=-n}^{n} c_{nm} Y_{nm}(\theta_A, \lambda_A) - \frac{1}{r_B} \sum_{n=2}^{\infty} \left(\frac{a}{r_B}\right)^n \sum_{m=-n}^{n} c_{nm} Y_{nm}(\theta_B, \lambda_B) \right] \tag{29.31}$$

in order obtain the global coverage necessary to estimate a set of spherical harmonic coefficients. In the particular case when we consider only the central term and the J_2 coefficient of the gravity field of the Earth, we may write

$$V = \frac{GM}{r} - J_2 \frac{GM}{r} \left(\frac{a_E}{r}\right)^2 \left(\frac{1}{4} - \frac{3}{4} \cos 2\phi\right) \tag{29.32}$$

$$\frac{\Delta f}{f} = -\frac{\Delta V}{c^2} = \frac{GM}{c^2} \cdot \frac{r_B - r_A}{r_A \cdot r_B} - J_2 \frac{GM}{r_A c^2} \left(\frac{a}{r_A}\right)^2 \left(\frac{1}{4} - \frac{3}{4} \cos 2\phi_A\right) + J_2 \frac{GM}{r_B c^2} \left(\frac{a}{r_B}\right)^2 \left(\frac{1}{4} - \frac{3}{4} \cos 2\phi_B\right) \tag{29.33}$$

At about $55°$ and $-55°$ geographical latitude, the gravitational potential due to the J_2 coefficient of the gravity field of the Earth is near zero. Therefore, at these geographical latitudes only the central term of the gravitational field, i.e., the GM constant, plays a major role with other (high-frequency) gravity content mainly coming from the surrounding topography. The majority of timing labs are placed at mid-latitudes, and the high-frequency part in the gravity field of the Earth can be taken from gravity missions such as GOCE and combined terrestrial/space gravity models. It should be very interesting to combine gravity information

from optical clock with space gravity missions, and to estimate the GM constant and low-order gravity field coefficients using data from optical clocks.

Figure 29.1 shows the two main approaches in the determination of gravitational potential using an optical clock. In the absolute approach, we measure a high resolution gravitational potential on the ground using an optical ground clock assuming that the gravitational potential of the satellite clock is known, and therefore we call this approach an absolute approach. The GNSS or a GEO orbit is high above the Earth and typical sensitivity to the Earth's gravity field is $n \le 10$ in terms of the spherical harmonic expansion used in the orbit numerical integration. We call this approach absolute, because we assume that the gravitational potential of a space optical clock is given (e.g., from the space gravity missions) and the absolute frequency of the space clock is determined by the clock itself. Any optical clock can be used by definition to determine the unit of time. Any error in the frequency realization of the satellite clock will give a bias in the two-way frequency transfer between a ground and a space clock. This satellite clock error or any additional two-way link error is removed or reduced, if a difference in the gravitational potential is determined relative to the first or the reference ground clock in a two-way frequency transfer (right). The main difference between the absolute and the relative approach is that relative approach requires a satellite clock with very high short-term stability during the light-travel time of the signal in the two-way approach, whereas in the absolute approach, an absolute realization of the frequency and gravitational potential is required for the satellite clock.
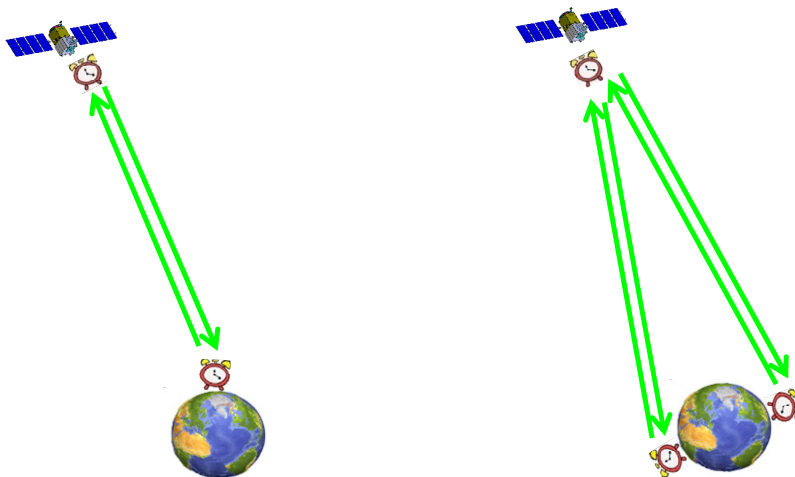
Figure 29.1 Measuring ground gravitational potential using an absolute approach (left), between a satellite and a ground clock assuming that the satellite clock is with a known gravitational potential (e.g., in GEO or GNSS orbit). Typical sensitivity of the GNSS or GEO orbit to the Earth's gravity is in terms of spherical harmonics degree/order $n \le 10$ (used in the numerical integration of the orbit). The figure on the right shows determination of the relative gravitational potential betweeen two ground clocks, assuming that the gravitational potential of one of the ground clocks is known. Green lines denote frequency transfer between a space clock and an optical clock on the ground.

29.4 Relativistic Orbit Determination

From (29.27) there follows a very useful property that could be uutilized in orbit determination: the relationship between the radial difference Δr or radial orbit perturbation in this case, and the relative frequency offset Δf in terms of the central gravitational field, given by

$$\frac{\Delta r}{r} = \frac{\Delta f}{f} \cdot \frac{c^2}{V_B} \tag{29.34}$$

For planetary missions, where the gravitational potential can be represented by only the central gravity term, one could use (29.34) in orbit determination, e.g., for the constellation of LISA satellites. There are several proposals to fly optical clocks in deep space, such as SAGAS (Wolf et al. 2009) and ASTROD (Braxmaier et al. 2012), where one could use such an approach.

Since the GOCE mission, the gravitational potential is known by at least a factor of ≈ 2 better in terms of degree and order of spherical harmonic expansion, compared to the degree and order of up to $100 - 120$ we typically need in the POD of LEO satellites using GPS. (29.34) opens the way for relativistic POD where the frequency variation of a clock along the orbit is used in the orbit determination. Knowing the gravitational potential and measuring the relative frequency offset along the orbit, one could derive the satellite position. The gravitational potential in terms of spherical harmonics is a function of position. Future GNSS satellites will fly optical clocks, so this concept is very interesting indeed. Environmental effects on the space clock, such as the magnetic field or the temperature variations along the satellite orbit are engineering issues and are not considered here. We have not seen significant influence of these effects on the Galileo H-maser. To obtain a LEO orbit determination with a 1 cm radial orbit accuracy, one would need an optical clock running in the Galileo satellites with a relative frequency stability of $\Delta f / f \approx 10^{-18}$. (29.34) could be considered as the geometry-free POD, since the relative frequency offset is typically measured using a two-way method that removes all geometry between GNSS and LEO or a ground station. The advantage of this approach is that there are no signal propagation effects, since the design of the two-way metrology link removes all propagation as well as geometrical terms. As already demonstrated with ground optical fiber measurements over 2000 kilometers, relative frequency offsets can be measured at the $\Delta f / f \approx 10^{-19}$ level over, e.g., 100 s , (Droste et al. 2013). On the other hand, rapid developments in optical clocks have enabled these to reach the $\Delta f / f \approx 10^{-18}$ level of stability already, and quantum wave interferometry has shown that the gravitational redshift can be measured to an accuracy orders of magnitude better than when using the present day optical clocks (Muller et al. 2010). In this type of positioning, the absolute velocity of the satellite could be estimated as a parameter.

Therefore, in the field of central gravitational potentials (e.g., planetary orbits), an interesting property can further be derived from (29.28) to measure energy E_A and position purely geometrically

$$-\frac{V_A \cdot f_A}{c^2} = -\frac{V_A \cdot m}{h} = \frac{E_A}{h} = f_A = -\frac{r_B}{r_B - r_A} \cdot \left(f_B - f_A \right) \tag{29.35}$$

knowing position and frequency offset, multiplied by the Planck constant

$$E_A = -\frac{r_B}{r_B - r_A} \cdot \left(f_B - f_A \right) \cdot h \tag{29.36}$$

that gives

$$\frac{f_B - f_A}{f_A} = -\frac{r_B - r_A}{r_B} \tag{29.37}$$

Thus, the geometrical definition of the relative frequency offset is

$$\frac{\Delta f}{f} = \frac{f_B - f_A}{f_A} = -\frac{E_B - E_A}{E_A} = \frac{m \left(V_B - V_A \right)}{m V_A} = -\frac{r_B - r_A}{r_B} \tag{29.38}$$

In a more general case, when the gravitational potential is given in terms of spherical harmonic expansion, the location can be calculated by the inverse relation, that for Legendre polynomials gives, e.g., (Abramowitz and Stegun 1965),

$$(\cos \gamma)^n = \sum_{l=n,n-2,\ldots} \frac{(2l+1)n!}{2^{(n-l)/2}\left[\frac{1}{2}(n-l)!\right](l+n+1)!!} P_n(\cos \gamma) \tag{29.39}$$

Since the inverse distance of the difference of any two vectors $\vec{r}$ and $\vec{s}$ in Euclidian space can be expanded into Legendre polynomials and since this is the basis for the spherical harmonic expansion of the gravitation potential

$$\frac{1}{|\vec{r} - \vec{s}|} = \frac{1}{r} \sum_{n=0}^{\infty} \left(\frac{s}{r}\right)^n P_n(\cos \gamma) \qquad \cos \gamma = \frac{\vec{r} \cdot \vec{s}}{r \cdot s} \tag{29.40}$$

following (Rummel 2006), the addition theorem of spherical harmonics is the addition theorem for the associated Legendre functions and can be written as

$$P_n(\cos \gamma) = P_n(\cos \theta_1) \cdot P_n(\cos \theta_2) + 2 \sum_{m=1}^{n} \frac{(n-m)!}{(n+m)!} \cdot P_n^m(\cos \theta_1) \cdot P_n^m(\cos \theta_2) \cdot \cos\left[m(\lambda_1 - \lambda_2)\right] \tag{29.41}$$

For the spherical harmonic expansion the vector $\vec{s}$ can be held constant and scaled to the radius of the reference sphere $|\vec{s}| = a$. The surface harmonics $P_n^m(\cos \theta_1) \cos m\lambda_1$ can be represented by spherical harmonic coefficients. The spherical distance γ between the two vectors $\vec{r}$ and $\vec{s}$ in Euclidian space can also be calculated from spherical coordinates θ and λ making use of the spherical law of cosine

$$\cos \gamma = \cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2 \cos(\lambda_1 - \lambda_2) \tag{29.42}$$

By introducing the rotation of the associated Legendre functions along the equator we obtain

$$P_n(\cos \gamma) = \boldsymbol{p}_n^-(0)^T \cdot \Re_n(\gamma) \cdot \boldsymbol{p}_n(0) \tag{29.43}$$

and finally

$$\frac{1}{|\vec{r} - \vec{s}|} = \frac{1}{r} \sum_{n=0}^{\infty} \left(\frac{s}{r}\right)^n \boldsymbol{p}_n^-(0)^T \cdot \Re_n(\gamma) \cdot \boldsymbol{p}_n(0) \tag{29.44}$$

Thus, the inverse spherical harmonics problem could either be solved iteratively based on an approximate location of the relative potential measurements or be derived by spherical harmonics, or spherical harmonic rotations.

29.5 A Satellite Orbit as a Wave

Analogous to the Planck constant in quantum mechanics (29.2), we may define a similar constant h_K for the Keplerian orbit

$$E = -\frac{GM}{2a} = h_K \cdot f_\oplus = h_K \cdot \frac{n}{2\pi} = h_K \cdot \frac{1}{2\pi} \sqrt{\frac{GM}{a^3}} \tag{29.45}$$

relating the total energy of the orbit $E = -GM/2a$ per unit mass to the orbit frequency $f_\oplus = n/2\pi$.
From which it follows

$$h_K := -\pi \frac{GM}{a} \cdot \sqrt{\frac{a^3}{GM}} = -\pi\sqrt{GMa} \tag{29.46}$$

375

that is nothing else than the specific angular momentum of the corresponding circular orbit multiplied by π

$$h_K := -\pi \cdot \sqrt{GMa} = \pi \cdot h_\oplus \tag{29.47}$$

The specific angular momentum of the corresponding circular orbit is thus

$$h_\oplus = -\frac{E}{\pi \cdot f_\oplus} \tag{29.48}$$

By analogy to (29.3) denoting $\hbar_K = \dfrac{h_K}{2\pi}$ we can write

$$\hbar_K := \frac{h_K}{2\pi} = -\frac{1}{2}\sqrt{GMa} = \frac{1}{2}h_\oplus \tag{29.49}$$

This is a very interesting relationship showing that an orbit can be represented or considered as a wave, much in the same way as photons can be considered as waves in quantum mechanics. For comparison, see the de Brougli relation (29.4). At the same time, the position of the photon can be measured, revealing its particle property. The same is true for a satellite and so it could be extended to satellite orbits. Since Keplerian orbit can be decomposed into bi-circular orbits using two circular orbits rotating into opposite direction, one could extend this analogy and state that gravity can be represented with geometrical properties in its origin. In the same way as the Planck relation describes the wave nature at the quantum level, we may write by analogy for the Keplerian orbit

$$f_\oplus = \frac{E}{h_K} = \frac{E}{\pi \cdot h_\oplus} \tag{29.50}$$

Thus, in essence, every orbit could be considered or imagined as a wave represented by matter and time that could be modeled by geometrical rotations. The use of bi-circular orbits representing the Keplerian orbit opens the door towards the dualistic concept of gravitation similar to the electromagnetic force where electricity and magnetism are elements of the same phenomenon and are orthogonal to each other.

In Section 26 we have seen that the rotation of spherical harmonic coefficients about the polar axis is equivalent to the rotation of equatorial associated Legendre functions or to the counter-rotation of the equatorial associated Legendre functions about the polar axis. The same analogy can be applied to orbit using bi-circular orbits, and one can draw the conclusion that the rotation of spherical harmonic coefficients could be used as a generating function, either to represent an orbit or the gravitation itself.

Due to the gravity field, modeled in terms of spherical harmonics or a multipole representation of bi-circular orbits, each quantum of energy ΔE is represented by the corresponding specific angular momentum defined again in the multipole representation by circular perturbations. In terms of semi-major axis and specific angular momentum h_A and h_B, (29.38) can be written as

$$\frac{\Delta E}{E_A} = \frac{E_B - E_A}{E_A} = -\frac{V_B - V_A}{V_A} = -\frac{\Delta V}{V_A} = -\frac{\dfrac{GM}{a_B} - \dfrac{GM}{a_A}}{\dfrac{GM}{a_A}} = \frac{a_B - a_A}{a_B} = \frac{h_B^2 - h_A^2}{h_B^2} \tag{29.51}$$

From this mathematical analogy one could consider that gravitation both at the quantum level and at the celestial level has similar properties to light or any other radiation.

An interesting extension of this approach is the conservation of angular momentum, at both quantum and celestial levels. One can see this every day in our planetary system, in particular in the transfer of angular momentum between Earth and Moon, considering that the total angular momentum is conserved. This can be measured by lunar laser ranging and reflected as an increase in the Earth-Moon distance by about 3.4

cm/yr, i.e., an increase in the radius of the Moon's orbit that is subsequently reflected in the angular momentum of the Moon's orbit. This exchange of angular momentum between Earth and Moon is driven by the tidal torque exerted by the Moon on the Earth and results in a slowing down of the Earth's rotation rate and thus the angular momentum of the Earth is decreased. A similar exchange of momentum can be identified at the quantum level between different energy levels.

Following general relativity, any form of energy that interacts with spacetime creates differential geometrical forms that can describe gravity. Thus gravitation can be considered as purely geometrical property. However, our geometrical model of two counter-oscillations could represent gravitation from the point of view of a scalar field at the quantum as well as at the celestial level. Thus, it gives geometrical and scalar properties of gravitation at the same time. This is similar to the concept of electromagnetic force where electricity and magnetism are elements of the same phenomenon or the concept of matter and antimatter in particle physics, where antimatter is described as material composed of antiparticles with the same mass as particles, but with opposite charge (leptons, baryons). Following recent results from the Planck mission (Planck Collaboration et al. 2013), there is strong evidence that 26.8.% of the mass-energy of the Universe is made of non-baryonic dark mater particles, which should be described by the Standard Model. Thus, our geometrical circular model of scalar gravitation is an interesting model that could be considered as describing gravitation at the quantum as well as at the celestial level.

30. Geometrical Representation of Gravity Field Determination

Geometrical or kinematic orbit determination, demonstrated for the first time using GPS on board the CHAMP satellite (Švehla and Rothacher 2003b), was the basis for the retrieval of the very first determination of the gravitational field of the Earth making use of the energy balance approach, see (Gerlach et al. 2003). By means of numerical differentiation, the geometric positions of the CHAMP satellite were used to determine geometrical velocities along the orbit, and making use of the energy integral, the very first geometrical gravity model of the Earth was developed. One advantage of gravity field determination based on the energy balance approach is that we can work directly with the gravitational potential as a scalar field instead of having to integrate the equation of motion. In the case of the GOCE mission, a gravity gradiometer maps gravity gradients along the orbit, (Rummel et al. 2011). Geometrical positions determined using GPS are used to position the gravity gradient measurements within the terrestrial reference frame and to estimate low-order gravity field coefficients. Here we present gravity field determination using kinematic orbits, and in addition, introduce a concept of gravity field determination based on gravitational redshift and atom interferometry. The possibility of determining kinematic orbits of LEO satellites has triggered the development of new approaches in gravity field determination, opened up new fields and significantly changed the way we think about the gravity field of the Earth, not only from the point of view of satellite dynamics and numerical integration. One of the most important applications of the metric theories of gravity, such as the General Theory of Relativity, is that a clock moved further away from the source of the gravitational potential will run faster, thus one can measure perturbations in the gravitational potential along an orbit by measuring variations in the optical clock frequency. Very soon mechanical test masses used to observe gravity from space will be replaced by atoms and test particles at quantum level. One advantage of quantum mechanics compared to the classical post-Newtonian framework we use in geodesy is that atoms can be used to directly measure not only the acceleration of motion, but, in addition, also relative frequency offsets, i.e., gravitational redshift. A gravity gradiometer could be constructed based on atom interferometry and this is most likely the next step in the determination of the Earth's gravity field. On the other hand, the redshift effect for matter waves is by orders of magnitude higher in frequency than the frequencies of standard microwave and optical clocks. The Compton frequency ω_C of matter waves is very high since it includes the rest mass energy multiplied by c^2, e.g., for cesium one obtains $\omega_C/2\pi = 3.2 \times 10^{25}$ Hz. This is significantly higher than the frequency used to measure time and to define the SI second using cesium atomic clocks. Considering that an orbit error is consistent with an error in the orbit velocity, the net redshift effect for the clock determined from the satellite position is compensated by the second order Doppler effect calculated from the satellite velocity. In size, the net effect on the total redshift effect is smaller and satellite orbit in terms of radial position is required with less accuracy compared to the accuracy of the static position for a ground clock placed on the Earth. A smaller variation in frequency can be measured at higher matter wave frequencies or by an atom gradiometer concept. This symmetry principle could be used to map gravity fields from space and

in the construction of an atom gradiometer. Here we discuss the question of how the new relativistic technique based on optical clocks and atom interferometers, in general, can contribute to global, regional and local gravity field determination and the realization of a global height system. We show that there are applications for this new technique in reference frame realization for positioning, time and temporal gravity determination and how this new geometric technique could unify all three fundamental reference frames in geodesy. The principle of error compensation in the calculation of the redshift effect, considering an orbit error in satellite position and the error in the second order Doppler effect calculated from the satellite velocity, has been discussed in the timing community. This is one of the main arguments, why an orbit in space (GEO) offers the best environment to define and establish the standard of frequency and define the SI second using an atomic clock, far better than using the geoid and the surface of the Earth. The main argument is, however, that cold atoms can be observed for a long time in space and are not limited by the free-fall on Earth, gaining an additional 3-4 orders of magnitude in sensitivity for atomic clocks. Thus a GEO or a GNSS orbit could offer the best place to define the datum for time on Earth and be used in supporting definition of the fundamental reference frames in geodesy.

30.1 General Aspects of the Relativistic Gravity Field Determination with Optical Clocks and Atom Interferometers

In (Švehla and Rothacher 2005b) it was proposed for the first time to use atomic clocks in space to measure gravitational potential along an orbit, while simultaneously using GPS to measure both the orbit and gravity with a purely kinematic or geometrical technique. Although relative frequency stability along an orbit and the sensitivity of such measurements at a high sampling rate is very demanding, this approach has considerable potential for use in future gravity field missions. The main reason to believe this is the fact that relative measurements of clock frequency variations along an orbit can be performed at a much higher level of precision than that of absolute clock measurements. For atomic clocks, the clock stability is always higher than the absolute clock accuracy. This also stands for the short averaging time (e.g., 1 s), and from that point of view, optical atomic clocks have the potential to be used to measure gravitational potential along an orbit.

The clock stability is a measure of how much the frequency is changing over a specified time interval, whereas accuracy is a measure of how much the frequency is offset from the absolute frequency that defines the SI unit of time, i.e., 1 s. This means that clock accuracy is a measure of how well the clock produces an exact signal frequency in terms of the SI second. Clock accuracy has typically two meanings and is often associated with the word uncertainty, because the atomic clock itself defines the accuracy. Uncertainty is the measure of how well the frequency can be assessed, i.e., how well the clock standard represents the natural frequency of atomic transition.

In terms of stability, the latest optical clocks at NIST in the US have reached a stability of 10^{-18} over several hours of averaging, while the accuracy is $8.6 \cdot 10^{-18}$ (Chou et al. 2010). The sstability of optical clocks far surpasses that of all other types of clocks. So-called optical lattice clocks are designed for high stability, with the latest reported results being $1.8 \cdot 10^{-18}$ in $20\,000$ s (Ludlow et al. 2013). Optical clocks have reached the accuracy level of 10^{-18}, as reported by several groups, and recently, e.g., (Nicholson et al. 2015), reports a total uncertainty of the JILA Sr clock to be $2.1 \cdot 10^{-18}$ in fractional frequency units.

Although optical clocks provide measurements of the gravitational potential, they are still insufficiently developed for space applications. Therefore, the atom interferometers have the potential to replace GOCE-type electrostatic accelerometers on future gravity field missions. It is assumed that the GOCE follow-on gradiometer will actually be an atom interferometer that will measure gravity gradient in three orthogonal directions (or just cross-track) using atoms as test masses. Atom interferometers have the potential to provide gravity gradients with long-term stability and allow the measurement of temporal gravity field changes at very low LEO altitudes. Formation-flying concepts at low LEO altitudes of 250 km and below are difficult to

realize due to the dynamics of the satellites (attitude dynamics, aerodynamic drag, etc.). Thus, the concept of a single satellite at 200 km altitude and below (with a new-generation propulsion system) equipped with a gradiometer based on atomic interferometry is the future for gravity field determination. Atom interferometry offers accuracy and long-term stability in the measurement of acceleration, and is therefore the way forward in measuring temporal gravity field variations of the Earth at very low LEO altitudes.

Regarding the non-gravitational forces and satellite dynamics, for future gravity field missions, one would either need to build a completely drag-free GRACE-type satellite capable of formation flying at low LEO altitudes or use gravity gradiometry based on atomic interferometry to directly eliminate non-gravitational effects. Even if some form of pendulum formation-flying concept is employed, the main gravity signal that is measured by the GRACE-concept will remain in the along-track direction, and partially in the cross-track direction. However, there will be no tracking system of sufficient accuracy to model satellite dynamics accurately enough in all three directions. Even with four GNSS systems, the accuracy of LEO orbit determination will not improve significantly. Thus, highly accurate along-track orbit dynamics will be affected by deficiencies in the modeling of the other two orbit components. From that point of view, a combination of pendulum formation-flying and gravity gradiometry could help, but if gravity gradiometry can improve the pendulum formation-flying concept with long-term stability of gravity gradients, there is enough justification to fly a single gradiometer at significantly lower LEO altitudes and measure both static and temporal gravity field variations at a very high resolution. Considering that the main gravity signal is in the radial direction, gravity gradiometry with a single satellite is a very robust concept that can also be applied to all planetary gravity fields.

Atom interferometry theoretically allows accelerometer resolution of up to 7×10^{-20} g (quantum limit), considering that all systematics at quantum level are currently known at the 10^{-16} g level (Kasevich 2013). Atom interferometry at 5×10^{-13} g has been demonstrated by (Kasevich 2013) opening up the possibility of using the same approach between two atoms separated by a constant baseline on the same satellite, or even between two satellites flying in formation. Measurements of relative acceleration (gravity gradients) is always more accurate than that of absolute acceleration itself. Generally speaking, one could say that for any relative measurement. As a result, it was reported that the accuracy level of 10^{-15} g that was achieved could be reached after only about 15 s of averaging (Kasevich 2013).

In atom interferometry atoms can act as clocks, measuring the light travel time across the baseline and at the same time they are the test masses. Atom interferometers with an arm of $1 - 1.5$ m , just in the orbit cross-track direction, could offer data on gravity gradients of extremely high sensitivity and accuracy. The cross-track direction is more favorable, since there is no direct rotational component in the cross-track, i.e., the measurement is orthogonal to the satellite rotation in the orbital plane, allowing a very high level of performance. Another advantage of a gradiometer based on atom interferometry is that it can provide data on angular rates to a similarly high level of precision, accuracy and stability at both, low and high frequencies. In the case of an atom interferometer, the frequency can be generated by a H-maser, thus there is no need for an optical clock in this case. A differential atom interferometer is under development for the ESA STE-QUEST mission, (Rasel 2013), to test the equivalence principle using two different isotopes of Rubidium.

What about optical clocks and redshift in comparison to the performance of atom interferometers? As a direct consequence of Einstein's general theory of relativity, a source of radiation in the gravitational potential V_B appears shifted in frequency to an observer in a different gravitational potential V_A by an amount $\Delta f/f = -\Delta V/c^2$, where $\Delta V = V_B - V_A$ is the difference between the gravitational potential at the position of the source, B, and the gravitational potential at the position of the observer, A. If we approximate the gravitational potential by only the central term GM/r (with r denoting the radial geocentric distance, GM the geocentric gravitational constant, and c the speed of light in vacuum), following (Švehla and Rothacher

2005b) we obtain the following expression for the fractional frequency shift as a function of gravitational potential perturbations ΔV along the orbit

$$\frac{\Delta f}{f} = \frac{\Delta V}{c^2} \approx \frac{GM}{c^2 r^2} \Delta r \qquad (30.1)$$

Thus by measuring the frequency variations Δf of the clock along the orbit, one can directly map the gravitational potential ΔV along the orbit. Since the internationally agreed definition of the SI second is based on atomic time, atomic clock frequency f is absolute in nature. Local environmental effects, like magnetic field, temperature, etc. can be engineered and compensated for along the orbit. However, the clock frequency is given generally in the local environment. The stability of the frequency, which is needed here, is typically one order of magnitude better than the absolute accuracy. Let us assume two extremely well performing clocks in space that are stable at a level of 10^{-18} over, e.g., 15 s (corresponding to ≈ 100 km in the orbit of the clocks). If the gravitational frequency shift between these two clocks can be measured with a similar accuracy, we will be able to directly measure differences in the gravitational potential that correspond to a change Δr in the equipotential surface of ≈ 1 cm over 100 km. Change of the gravitation field over this period of time could be significant, however, the clock will measure total energy along the orbit that is constant for the Keplerian orbit. The non-gravitational forces are typically integrated along the orbit. Since kinematic positions can already be determined with an accuracy of $1-2$ cm RMS, and the relative orbit accuracy between successive epochs (with smoothing) or between two satellites < 1 mm RMS, the positions (geometry) of the pair of clocks is well-enough known to support such measurements of the gravity potential difference. (30.1) does not include the second-order Doppler effect and other effects. Any differential change Δv of the velocity of the satellite v will modify the relative frequency shift by

$$\frac{\Delta f}{f} = -\frac{v}{c^2} \Delta v \qquad (30.2)$$

If one assumes a velocity error in the order of $\Delta v = 0.01$ mm/s it gives $\Delta f/f = -v/c^2 \cdot \Delta v \approx -0.9 \cdot 10^{-19}$ in terms of relative fractional frequency offset. Therefore, to measure gravitational potential one would need to know the velocity with a similar level of accuracy. We will show later in this section that orbit velocity error and orbit position error contributing in the total red-shift effect compensate each other, thus the net effect on the measured fractional frequency offset $\Delta f/f$ is significantly smaller than the same position error of a static clock on the ground. For circular orbits driven only by the central gravity term, there is a constant positive fractional frequency offset acting towards a higher orbit altitude. Thus, by raising the orbit altitude fractional frequency offset is increased, since the total energy along the orbit is $E_{tot} = -GMm/(2a)$.

Frequency comparison in space is significantly simpler over large distances than when using a space-to-ground metrology link, since there is no atmosphere and atmospheric turbulence to affect the space-to-space link. Therefore, such a concept could fly at significantly lower orbit altitudes, either as low-low formation flying or high-low (e.g., with STE-QUEST in highly elliptical orbit). It is proposed to fly metrology links for the timing community in the second-generation Galileo satellites (e.g., on board one Galileo satellite per orbital plane) (Švehla 2008c) or in GEO orbit. Optical and microwave metrology links are under development for the STE-QUEST mission and for ACES. Therefore, in $10-15$ years from now one could expect the timing community to have an infrastructure in place to compare ground optical clocks at the 10^{-18} level. There are already optical communication terminals in GEO orbit that can be used in single-difference mode to compare clock frequencies between two locations. In space, such a frequency comparison will be more accurate, due to the absence of atmospheric turbulence.

New developments in gravity sensors for space will most likely trigger a new synergy of observables and try to combine gravity signals with other information. One can envisage that a nadir altimeter, or rather wide-swath GNSS altimetry, will further extend the concept of a gravity mission based on a single satellite with a

high-performing gradiometer based on atom interferometry. GNSS reflectometry (GNSS-R) has the potential
to evolve into the next generation of altimetry – wide-swath altimetry based on the reflected signals from the
more than 100 GNSS satellites that will be available in just a few years from now. That was the reason for
proposing to use the International Space Station to demonstrate GNSS altimetry from LEO orbit for the first
time, (Svehla 2008). Most likely, decades of altimetry missions will be combined with decades of gravity field
missions. There is a high probability that gravity missions will be collocated with altimeters. This will provide
a complete measurement of mass transport within the system Earth, since altimetry directly measures *in situ*
the geometry of the ocean surface beneath the satellite orbit that is driving the temporal gravity field, while
gradiometers measure its gravitational signal. Such a combined observation of altimetry and gravity has not
yet been performed with a single satellite. Looking into the future, most of the current and forthcoming
planetary missions, such as BepiColombo, Mars Express, ExoMars, Juice, are, or are going to be equipped
with a laser altimeter to measure the topography of the planet in question (Mars, Mercury, Jupiter, etc.).
However, information on gravity is lacking in all these missions. In combination with altimetry, gravity gra-
diometry will provide an insight into the planetary interior (Bouguer anomalies, Moho, density, ice, water,
etc.). Some attempts have already been made or proposed in this direction for the BepiColombo mission,
combining laser altimetry with planetary gravity field determination based on orbit tracking from Earth.
However, the resolution of such gravity field determination is not comparable to the resolution of the topog-
raphy provided by planetary laser altimetry. Atom interferometry offers new types of gradiometers that will
be suitable for future planetary missions and will allow the combination of planetary altimetry and planetary
gradiometry on one satellite. The concept of a single satellite is much more suitable for planetary gravity field
mapping than intersatellite tracking, as is the case with the GRAIL mission. The main reason for this is that
it is easier to operate a single satellite and reach lower orbit altitude than to operate two satellites flying in
formation.

30.2 The Energy Balance Approach for Gravity Field Determination –
Using Kinematic Orbits or the Onboard Optical Clock

The CHAMP, GRACE and GOCE missions are based on measuring Newtonian quantities of the Earth's
gravity field. GOCE directly observes gravity gradients of the gravitational field, whereas GRACE measures
changes in the inter-satellite distance between two satellites caused by the satellites' dynamics driven by the
gravitational acceleration of the Earth. The comparison with reduced-dynamic orbits and the external valida-
tion with SLR show, that, due to the nature of the phase observable, changes in kinematic position are very
smooth from epoch to epoch and thus geometrically map the satellite orbit with a high resolution. As a
consequence, high-frequency gravity signals may be extracted from these positions. An elegant way to derive
gravity field coefficients from kinematic positions is to use the energy conservation law which may be written
along the satellite orbit in an inertial frame as

$$V = \frac{1}{2}\left(\frac{d\vec{x}}{dt}\right)^2 - \int_{\vec{x}} \vec{a}_t \cdot d\vec{x} - \int_{\vec{x}} \vec{a}_{non} \cdot d\vec{x} - C \tag{30.3}$$

with the gravitational potential V, the acceleration $\vec{a}_t$ due to the time-varying part of the gravity field, e.g.,
tides, and the non-gravitational forces denoted as $\vec{a}_{non}$. The total energy constant is denoted as C. A similar
formulation in the Earth-fixed frame can be found in (Gerlach et al. 2003) where we originally published this
approach. A big step in improving the accuracy of this approach was the discovery that tides are not con-
servative and thus it was proposed to numerically integrate them together with the non-gravitational
accelerations. An advantage of gravity field determination based on the energy integral is that one can work
directly with the gravity potential as a scalar field instead of having to integrate the equation of motion.

Whereas $\vec{a}_t$ can be obtained from models, the non-gravitational accelerations $\vec{a}_{non}$ are measured by the onboard accelerometers. The kinetic energy of the satellite can be calculated using velocities derived from kinematic positions by numerical differentiation procedures. LEO kinematic positions are typically given at a sampling rate of $30\,\text{s}$, which means that the spatial resolution of the estimated gravity field is limited to about $200\,\text{km}$ and that much care has to be taken when deriving kinematic velocities. This is the reason that for the GOCE mission, kinematic positions are provided at a sampling rate of $1\,\text{s}$. By changing to a higher sampling rate, numerical differentiation becomes more accurate and a higher spatial resolution is possible, see (Švehla and Földváry 2006). In the light of forthcoming accelerometers based on atomic interferometry, the integration of non-gravitational acceleration will be straightforward, since there will be no need for a very frequent estimation of biases and scaling factors as is the case with CHAMP and GRACE accelerometers. However, those parameters will need to be determined.

An interesting approach for gravity field determination can be realized if an onboard optical clock is collocated with the onboard GPS receiver. In this case, an onboard optical clock measures the total energy along the satellite orbit including the gravitational potential as well as the second order Doppler effect based on (30.3). This sum also includes accumulated parts due to integration of the non-gravitational accelerations. The onboard GPS receiver provides satellite position and velocity along the orbit. However, for this approach a short-term clock stability of e.g., $< 10^{-18}$ (or better at the averaging time of e.g., 1 s) would be needed in order to achieve sufficient sensitivity of the clock to the gravity field. Since a clock measures the total energy of the orbit, an averaging could be employed for measured $\Delta f / f$ over a given time window, e.g., 100 sec, shifted every 1 s in time. Making use of the averaging window, clock stability could be increased by a factor of 10-100 in the observation equation. Considering that clock stability improves with averaging time, the energy balance approach as presented in (30.3) would benefit from measurement of the total energy along the orbit and gravitational potential. The calculated total energy along the orbit could be averaged for a given time window using state-of-the art gravity models. For this, frequency measurements Δf of the optical clock along the orbit would need to be performed relative to the reference epoch or a given constant energy level of the clock with the gravitational potential V_0 and velocity $d\vec{x}_0 / dt$

$$\frac{\Delta f}{f} = -\frac{1}{c^2}\left[V - \frac{1}{2}\left(\frac{d\vec{x}}{dt}\right)^2 + \int_{\vec{x}} \vec{a}_t \cdot d\vec{x} + \int_{\vec{x}} \vec{a}_{non} \cdot d\vec{x} - V_0 + \frac{1}{2}\left(\frac{d\vec{x}_0}{dt}\right)^2\right] \tag{30.4}$$

In all other cases a relative frequency measurements would need to be performed, either between two satellites or between relative positions on the same satellite. The tidal effects $\vec{a}_t$ in (30.4), needs to be integrated along the orbit. The sensitivity of (30.4) is significantly limited by the orbit-redshift equivalence principle, discussed in more detail later in this section, where an error in the satellite velocity (used to generate second-order Doppler effect in (30.4)) is compensated for by an error in the orbit position when the net redshift effect is calculated.

The energy balance approach based on kinematic orbits determined using GPS which we published in (Gerlach et al. 2003), suffered significantly from the integration of accelerometer measurements $\vec{a}_{non}$ from the CHAMP satellite (measuring the non-gravitational effects such as air-drag etc.). That was primarily due to the high sensitivity to temperature variations along the orbit that required frequent estimation of accelerometer calibration parameters. Atom accelerometers and optical clocks could provide very interesting new approaches to retrieve the Earth's gravity field from space. Since we are not interested in the absolute accuracy of an optical clock, fractional frequency stability of $< 10^{-18}$ (at the averaging time of e.g., 100 s) would appear feasible considering the level of development of ground optical clocks in 2016. Considering the fast development of optical clocks over the last few years, (Nicholson et al. 2015) reports a total uncertainty of the JILA Sr clock to 2.1×10^{-18} in fractional frequency units reached by improving the atom's thermal environment and the atomic response to room-temperature blackbody radiation.

30.3 The Orbit-Redshift Equivalence Principle

In this section we look more closely into the relativistic effect for an optical clock in space. This section demonstrates that for a nearly circular orbit, a clock shifted radially for a small distance on the satellite will not experience any differential frequency offset, i.e., the clock will "tick" at the same or similar rate. Increasing velocity by a small "delta-v" Δv in (30.2) will give rise to Δr in (30.1), but the net fractional frequency offset after adding (30.2) to (30.1) will be close to zero. Hence, the net effect on total energy will be close to zero and thus the apparent fractional frequency offset will stay constant. This symmetry could be called the "orbit-redshift equivalence principle", since any error in the orbit position is equivalent to an error in the orbit velocity (second-order Doppler effect) that appears as equivalence in the net redshift effect. We analytically show that property for a circular orbit, and, by rigorous integration of the Schwarzschild metric, for the orbit of the CHAMP satellite at 408 km altitude (Švehla et al. 2006b).

30.3.1 Differential Gravitational Redshift and Radial Orbit Error

General relativity predicts that a clock further away from the Earth, i.e., away from the center of the gravity field, runs faster than a clock closer to the Earth. The effect is proportional to the gravitational potential due to the Earth (and other celestial bodies). In relativity, the geopotential is defined by convention as having a negative value, approaching zero as a particle moves towards infinity away from an attracting body. However, confusion often arises since in geodesy the sign convention for the geopotential is the opposite to that used in physics. In geodesy, all potential is positive, so that a higher potential would generally be closer to the Earth. In the geodetic convention, all geopotential is positive, thus potential energy along the orbit is negative.

Therefore, for a clock shifted radially by an offset d we may write the fractional frequency shift

$$\frac{\Delta f}{f} = -\frac{1}{c^2}\left[\frac{GM}{r+d} - \frac{GM}{r}\right] \tag{30.5}$$

or finally

$$\frac{\Delta f}{f} = \frac{GM}{c^2}\left[\frac{d}{r(r+d)}\right] \tag{30.6}$$

For a LEO orbit at $r = (6371 + 408)$ km and shifted by $d = -10$ m, the clock will run more slowly by

$$\frac{\Delta f}{f} \approx -9.65 \cdot 10^{-16}. \tag{30.7}$$

30.3.2 Differential Special Relativity and Radial Orbit Errors

The so-called second order Doppler shift of special relativity states that a standard clock runs slower if it moves faster, relative to a clock at rest with the observer. If we now consider an optical clocks shifted in the radial direction on the same satellite it must have the same mean motion ω. The velocity for a satellite in circular orbit is

$$v = \omega \cdot r \tag{30.8}$$

Since for both points the mean motion is the same, it must hold

$$v_{true} = \omega \cdot r \quad > \quad v_{quasi} = \omega \cdot (r - 10 \text{ m}) \tag{30.9}$$

This explains why the point shifted in the radial direction towards the Earth has a lower velocity, although from Kepler's Third Law the opposite would be the case. If these two points are not connected, they will describe two different orbits. Let us derive an expression for the special relativity part of the frequency offset against the true orbit with velocity v_{true}

$$\frac{\Delta f}{f} = \frac{1}{c^2}\left(-\frac{v_{quasi}^2}{2} + \frac{v_{true}^2}{2}\right) \tag{30.10}$$

introducing circular velocity $v = \omega r = r\sqrt{GM/r^3}$ for the true and the shifted orbit we obtain

$$\frac{\Delta f}{f} = \frac{GM}{2c^2}\left(-\frac{\left(r+d\right)^2}{r^3} + \frac{1}{r}\right) \tag{30.11}$$

that reduces to the fractional frequency shift

$$\frac{\Delta f}{f} = -\frac{GM}{2c^2}\left(\frac{2r+d}{r^3}\right)\cdot d \tag{30.12}$$

For our case $r = (6371 + 408)$ km and $d = -10\,\text{m}$ we obtain

$$\frac{\Delta f}{f} \approx 9.65\cdot 10^{-16} \tag{30.13}$$

If one compares (30.13) with gravitational redshift (30.7), one can see that the magnitude is the same, only the sign is opposite. Hence, the total effect is zero. This is the reason why for small clock displacements in nearly circular orbits we do not need highly accurate orbits to calculate accurate time and relativistic corrections along the orbit.

For comparison, in the case of a ground optical clock there is a difference between the potential due to gravitation and that due to gravity. The former arises from the presence of attracting masses only, the latter contains, in addition, the centripetal potential due to the Earth's rotation. The rotation of the Earth, therefore, gives rise to a centripetal potential that also changes the clock's frequency. For altitudes of 350 km and 450 km we have amplitudes of

$$\frac{\Delta f}{f}(10\text{ m at }350\text{ km}) \approx 4.91\cdot 10^{-16} \quad \frac{\Delta f}{f}(10\text{ m at }450\text{ km}) \approx 4.77\cdot 10^{-16} \tag{30.14}$$

For an offset of 1 m we obtain

$$\frac{\Delta f}{f}(1\text{ m at }350\text{ km}) \approx 4.91\cdot 10^{-17} \quad \frac{\Delta f}{f}(1\text{ m at }450\text{ km}) \approx 4.77\cdot 10^{-17} \tag{30.15}$$

If we increase the offset d to 50 m

$$\frac{\Delta f}{f}(50\text{ m at }350\text{ km}) \approx 2.45\cdot 10^{-15} \quad \frac{\Delta f}{f}(50\text{ m at }450\text{ km}) \approx 2.38\cdot 10^{-15} \tag{30.16}$$

30.3.3 Integration of the Schwarzschild Metric Along a LEO Orbit

To prove that general and special relativity compensate for each other for an orbit error in position and velocity, we derive an expression for the complete effect (30.6) and (30.12)

$$\frac{\Delta f}{f} = \frac{GM}{c^2}\left(\frac{d}{r(r+d)}\right) - \frac{GM}{2c^2}\left(\frac{2r+d}{r^3}\right) \cdot d \tag{30.17}$$

or finally, the fractional frequency shift is

$$\frac{\Delta f}{f} = -\frac{GM}{2c^2}\frac{d^2}{r^3}\left(\frac{3r+d}{r+d}\right) \tag{30.18}$$

For $r = (6371 + 408)$ km and $d = -10$ m we obtain, as expected

$$\frac{\Delta f}{f} \approx -2.13 \cdot 10^{-21} \approx 0 \tag{30.19}$$

A radial shift of a clock on board a satellite in a circular orbit does not induce any relativistic effect. That was reported for the first time in (Švehla et al. 2006b).

In (Švehla 2007b), the Schwarzschild metric given in e.g., (Petit and Luzum 2010), was integrated along the CHAMP orbit using velocity information from the POD and gravity field evaluated up to degree and order 120 in the calculation of the gravitational redshift. Two CHAMP orbits, with differences in the order of $RMS = 3.5$ cm, were used to evaluate the impact of orbit errors on the calculated proper time (based on total energy along the orbit), see Figure 30.1. One can see that the gravitational redshift, as well as the second-order Doppler effect is $< 10^{-17}$, fully in line with the orbit errors < 10 cm. Integrated proper time along the orbit shows that orbit errors are periodic in nature, thus errors in proper time do not accumulate as a random walk, but average out over time. Figure 30.1 also shows that the net effect of the frequency offset due to general and special relativity is more than one order of magnitude smaller than the individual contributions. To evaluate the net relativistic effect in more detail, an offset of 1 m was added to the radial antenna coordinate in the satellite-fixed reference frame. Thus, the center of mass of the satellite was assumed to be 1 m below its true center of mass and hence the acceleration from the models was calculated for the wrong place. When a reduced-dynamic orbit is determined, additional parameters are estimated (pseudo-stochastic pulses) in order to minimize the effects of unmodeled forces acting on the satellite (stemming from, e.g., air-drag and solar radiation). Figure 30.2 shows that 1 m in the radial orbit direction corresponds to a gravitational shift of 10^{-16} along the orbit and is symmetric to the contribution due to special relativity. The total frequency shift is not zero, but is 1-2 orders of magnitude smaller, i.e., $3 \cdot 10^{-18}$ with a clear once-per-rev pattern, Figure 30.2. This is due to empirical parameters that do not completely absorb the errors in the non-gravitational forces, thus Figure 30.2 (right) indicates that the total energy along the orbit is not constant.
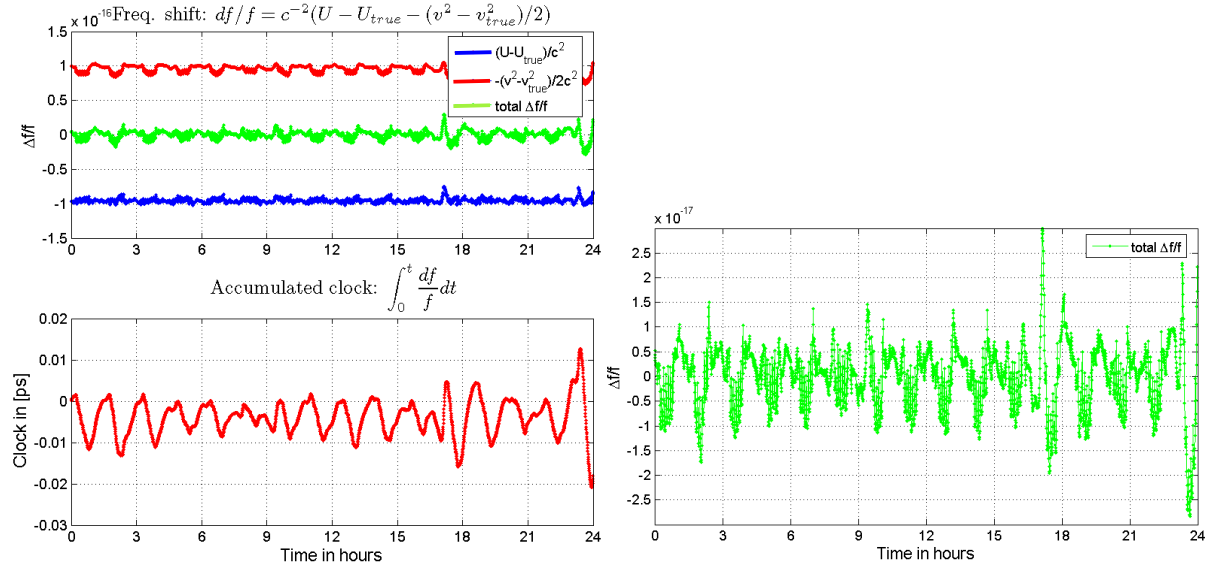
Figure 30.1 Fractional frequency shift and integrated proper time for a radial orbit bias of 1 m (left). The radial bias of 1 m gives a constant gravitational frequency shift of 10^{-16} along the LEO orbit, but the total frequency offset is $1-2$ orders of magnitude smaller, $\text{STD} = 3 \cdot 10^{-18}$ (right). The gravitational redshift is of a similar size to the frequency shift due to special relativity (left top).
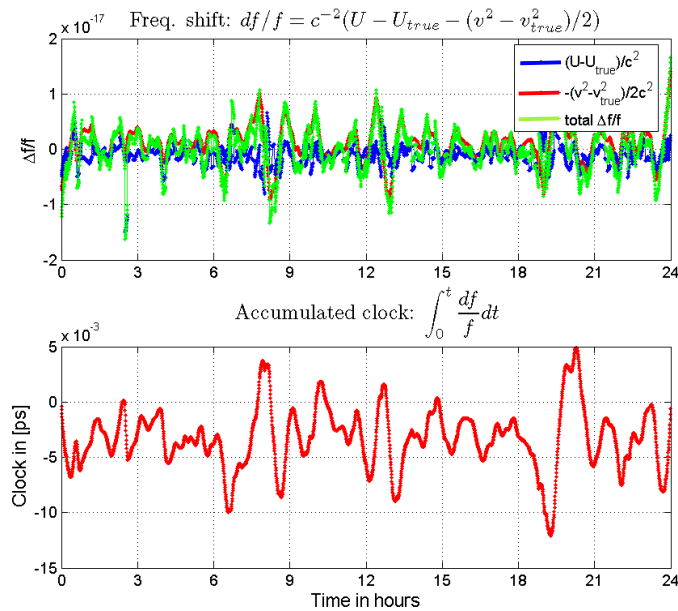
Figure 30.2 Typical fractional frequency shift between two orbits that differ by $\text{RMS} = 3.5$ cm (top) due to the general (red) and special (blue) theory of relativity along a LEO orbit, and integrated proper time along the orbit (bottom). Gravitational redshift as well as second-order Doppler effect are $< 10^{-17}$, in line with the orbit errors < 10 cm

30.3.4 Proper and Terrestrial Time Along an Orbit

Let us now see how to calculate terrestrial time along the orbit, i.e., how to relate the onboard time on the satellite to the time scale on the ground. Instead of measuring the fractional frequency offset along the orbit, one could base this technique on time comparison. However, frequency offset measurements are still the most precise measurements performed in any field of science.

For the non-rotating Earth, relativistic time in the vicinity of the geocenter is called Geocentric Coordinated Time (TCG). The rate of a moving clock outside the Earth and running with proper time T is related to TCG, at the 10^{-18} precision level, by the general and special relativity transformations (Petit 1998)

$$\frac{d\mathrm{T}}{d\mathrm{TCG}} = 1 - \frac{V(x,y,z) + \Delta V(x,y,z)}{c^2} - \frac{v^2}{2c^2} \tag{30.20}$$

with gravitational potential V and tidal potential ΔV (due to the Moon, Sun, and planets) given at the clock position (x, y, z). The clock velocity is denoted as v and given in the inertial geocentric coordinate system. Since the gravitational potential and velocity are always positive, by integrating (30.20), a clock with the time scale T runs slower than a clock aligned to TCG.

$$\mathrm{T} < \mathrm{TCG} \tag{30.21}$$

TCG is the coordinate time to be used for positioning and geophysical studies in a conventional terrestrial frame, such as ITRF. However, TCG cannot be directly observed and so-called Terrestrial Time (TT) is used as a proxy. TT is defined at the geoid and is related to TCG by the following conventional formula [*Petit*, 1998]

$$\frac{d\mathrm{TT}}{d\mathrm{TCG}} = 1 - \frac{W_0}{c^2} \tag{30.22}$$

where W_0 denotes the gravity potential at the geoid and is defined as

$$W_0 = V(x,y,z) + \Delta V(x,y,z) + \omega^2 \frac{(x^2 + y^2)}{2} \tag{30.23}$$

where we have the sum of the gravitational potential (first term), tidal potential (second term) and the third term is the centripetal potential resulting from the Earth rotation rate denoted by ω. Since W_0 is constant on the surface of the geoid, TT, by convention, differs from TCG by a constant time rate. By integrating relation (30.22), one can see that a clock synchronized to TT runs slower than a clock aligned to TCG

$$\mathrm{TT} < \mathrm{TCG} \tag{30.24}$$

In practice, International Atomic Time (TAI) serves as the realization of TT. TAI is a weighted average of a large number of atomic clocks (T_i), and using relativistic theory referenced on the geoid (Petit 1998)

$$\frac{d\mathrm{T}_i}{d\mathrm{TT}} = 1 - \frac{W_i + \Delta W_i}{c^2} + \frac{W_0}{c^2} \tag{30.25}$$

where ΔW_i, is the corresponding tidal potential. To calculate the gravity potential W_i, at the surface of the Earth requires a precise knowledge of the height above the geoid, or the geopotential difference $W_i - W_0$. The gravitational potential is always positive and by definition zero at infinity. This is why a standard clock on the surface of the Earth (above the geoid) runs faster than a clock on the geoid. This can easily be confirmed by integrating (30.25)

$$\mathrm{T}_i > \mathrm{TT}. \tag{30.26}$$

Finally, combining (30.20) and (30.22) we can derive the time-rate transformation for a moving clock in space running with proper time T on board a satellite,

$$\frac{d\mathrm{T}}{d\mathrm{TT}} = 1 - \frac{V(x,y,z) + \Delta V(x,y,z)}{c^2} + \frac{W_0}{c^2} - \frac{v^2}{2c^2} \tag{30.27}$$

which gives the required relativistic time-rate transformation between the time T and the conventional terrestrial time TT at the 10^{-18} precision level (Petit 1998). Furthermore, the TT in (30.27) can be replaced by GPS time, another atomic time which is freely and easily accessible using GNSS, assuming that both have the same nominal time scale (clock rate). Note that TT, TAI, and UTC (Universal Coordinated Time) all have nominally the same time scale.

Let us now assume two clocks in a satellite, running with proper times T_1 and T_2, respectively. By means of (30.27) we obtain as a time difference

$$\frac{d\mathrm{T}_1}{d\mathrm{TT}} - \frac{d\mathrm{T}_2}{d\mathrm{TT}} = \frac{V(x,y,z)_2 + \Delta V(x,y,z)_2}{c^2} - \frac{V(x,y,z)_1 + \Delta V(x,y,z)_1}{c^2} + \frac{v_2^2}{2c^2} - \frac{v_1^2}{2c^2} \tag{30.28}$$

Assuming that TT can be replaced by the GPS time scale, denoted here as t, and leaving out terms due to the tidal potential, we obtain after integration

$$\mathrm{T}_1 - \mathrm{T}_2 = \int_0^t \left[\frac{V(x,y,z)_2 - V(x,y,z)_1}{c^2} + \frac{v_2^2}{2c^2} - \frac{v_1^2}{2c^2} \right] dt . \tag{30.29}$$

Assuming that both clocks are shifted radially and placed on the same satellite in a circular orbit we have

$$\mathrm{T}_1 - \mathrm{T}_2 = \frac{1}{c^2} \int_0^t \left(\frac{GM}{r_2} - \frac{GM}{r_1} + \frac{v_2^2}{2} - \frac{v_1^2}{2} \right) dt \tag{30.30}$$

from which it follows

$$\frac{\mathrm{T}_1 - \mathrm{T}_2}{t} = \frac{1}{c^2} \left[\frac{GM}{r_2} - \frac{GM}{r_1} + \frac{v_2^2}{2} - \frac{v_1^2}{2} \right] \tag{30.31}$$

or finally, for the relative frequency offset

$$\frac{\Delta f}{f} = \frac{\mathrm{T}_1 - \mathrm{T}_2}{t} = \frac{1}{c^2} \left(\frac{GM}{r_2} - \frac{GM}{r_1} + \frac{v_2^2}{2} - \frac{v_1^2}{2} \right) \tag{30.32}$$

Considering geopotential, it follows from (30.32), that when clock T_1 runs faster, its frequency will be positively shifted relative to clock T_2. If we assume that clock T_1 is above clock T_2,

$$r_1 > r_2 \tag{30.33}$$

and due to a smaller gravitational potential, clock T_1 runs faster. This is in line with the statement in general relativity that a standard clock further away from the attracting body runs faster, see ((Švehla 2007b), (Pavlis and Weiss 2000), (Pavlis and Weiss 2003)).

30.3.5 Hamiltonian and Fractional Frequency Offset Along an Orbit

In order to further examine the role of non-gravitation accelerations in the calculation of fractional frequency offset along an orbit, we calculate the Hamiltonian H^*, known in mechanics, that represents the total energy along the orbit

$$H^* = \frac{1}{2}v^2 - V = -\int \frac{\partial V}{\partial t} dt + const. \tag{30.34}$$

Introducing Terrestrial Time TT (time scale on the geoid) we may write for the proper time T

$$\frac{d\mathrm{T}}{d\mathrm{TT}} = 1 - \frac{H^*}{c^2} + \frac{W_0}{c^2} \,. \tag{30.35}$$

where W_0 denotes gravitational potential on the geoid (including the centripetal part). Fractional frequency offset along the orbit over an accumulated time denoted here as t, can then be written as

$$\frac{\Delta f}{f} = \frac{\mathrm{T} - \mathrm{TT}}{t} = -\frac{1}{t \cdot c^2} \int_t H^* dt + \frac{W_0}{c^2} = -\frac{H^*}{c^2} + \frac{W_0}{c^2} \tag{30.36}$$

Including non-gravitational forces $\vec{a}$ (air-drag, solar radiation pressure, albedo, acceleration induced by attitude maneuvers and gradients due to an unknown satellite center of mass) and integrating them along the orbit we obtain

$$H^* = \frac{1}{2}v^2 - V + \int_t \vec{a} \cdot \vec{v} \, dt = -\int \frac{\partial V}{\partial t} dt + \int_t \vec{a} \cdot \vec{v} \, dt + const. \tag{30.37}$$

With regard to the integration of non-gravitational forces along an orbit and determination of the Earth's gravity field based on the energy conservation of the kinematic orbits of LEO satellites, we refer to (Gerlach et al. 2003). Due to non-gravitational forces, total energy is not conserved along the orbit and we may separate the non-conservative from the conservative Hamiltonian leading to the remaining non-gravitational constant

$$H^* = \frac{1}{2}v^2 - V + H^*_{non-con} + const. \tag{30.38}$$

and after integration over the time period t, we derive

$$\frac{\Delta f}{f} = \frac{\mathrm{T} - \mathrm{TT}}{t} = \frac{W_0}{c^2} + \frac{1}{t \cdot c^2} \int_t V dt - \frac{1}{t \cdot c^2} \int_t v^2 dt - \frac{1}{t \cdot c^2} \int_t H^*_{non-con} dt - \frac{const.}{c^2} \tag{30.39}$$

and finally for the fractional frequency offset we obtain

$$\frac{\Delta f}{f} = \frac{\mathrm{T} - \mathrm{TT}}{t} = \frac{W_0}{c^2} - \frac{H^*_{con}}{c^2} - \frac{1}{t \cdot c^2} \int_t H^*_{non-con} dt - \frac{const.}{c^2} \tag{30.40}$$

One can see that the separation of the fractional frequency offset due to the gravitational potential from the total effect requires the removal of the non-gravitational part and the integration of the non-gravitational acceleration along the orbit.

30.3.6 Relativistic Scale of Satellite Orbits in the Terrestrial Reference Frame

What is the impact of a different reference time on the scale of the terrestrial frame and ultimately on the satellite orbit? Starting with ITRF00, all versions of ITRF, such as ITRF05 or ITRF08, have used terrestrial time TT as a datum, whereas in earlier realizations, such as ITRF97, the time scale was referred to TCG. This can be verified by the release information of each ITRF. Looking at (30.22), one can see that replacing TCG by TT will not only change the time scale, but also the scale of the terrestrial frame. Since the speed of light is constant in all realizations of terrestrial systems (ITRF) and independent of the time convention in use, for a distance s measured in ITRF05, it must hold that

$$\frac{s_{\text{ITRF}}(\text{TT})}{d\text{TT}} = \frac{s_{\text{ITRF}}(\text{TCG})}{d\text{TCG}} \,. \tag{30.41}$$

From which, the scale of the terrestrial system can be obtained as

$$\frac{s_{\text{ITRF}}(\text{TT})}{s_{\text{ITRF}}(\text{TCG})} = \frac{d\text{TT}}{d\text{TCG}} = 1 - \frac{W_0}{c^2} \tag{30.42}$$

Based on this, we see that all distances measured in a terrestrial system with TT (such as ITRF05, ITRF08 or ITRF10) will be shorter than comparable distances measured in a system with a TCG time scale (ITRF97)

$$s_{\text{ITRF}}(\text{TT}) = s_{\text{ITRF}}(\text{TCG}) - \frac{W_0}{c^2} s_{\text{ITRF}}(\text{TCG}) \tag{30.43}$$

by about

$$\frac{W_0}{c^2} \approx 0.697 \text{ ppb} \tag{30.44}$$

or ≈ 0.7 mm / 1000 km. In the case of GNSS orbits, the orbit is shifted radially by about -18.5 mm, or in the case of a LEO orbit at 400 km altitude, the radial error is about -4.7 mm, when TT is used as a reference.

30.4 A Method to Measure Gravitational Gradient and Gravitational Redshift from the Interference of Matter Waves – Quantum Gravity Gradiometer

Atom interferometers based on atoms and light can measure acceleration and rotation to a very high level of precision. We may use atoms as drag-free test masses and utilize the wave-like or particle-like nature of atoms to perform interferometric measurements of the effect of gravitation on the atoms. In this way, we may probe relative gravitation between two locations in close proximity (gravity gradiometry) or to probe relativistic properties (general and special relativity). The redshift effect for matter waves is orders of magnitude higher than for the frequencies of standard microwave and optical clocks. The Compton frequency ω_C is very high since it includes the rest mass energy multiplied by c^2, e.g., for cesium one has $\omega_C/2\pi = 3.2 \times 10^{25}$ Hz. This is significantly higher than the frequency used to measure time and to define the SI Second by means of cesium atomic clocks. Here we will address both approaches and discuss, how to use matter-wave interference for a gravitational gradiometer in space.

If one utilizes the wave-like nature of atoms, atom interferometry could be performed in a similar manner to laser interferometry. Since photons carry momentum, when an atom is illuminated by a laser light it absorbs and emits a photon changing its momentum. Thus, by using a sequence of laser pulses $\pi/2$, π, $-\pi/2$ (phase), we can change the momentum and can obtain wave splitting and mirroring effects (as with a classical interferometer), see (Muller et al. 2010). The first laser pulse $\pi/2$ has a similar effect to wave beam splitting and sets the atom in a superposition (giving the momentum $k_1 \cdot \hbar$) of the ground and excited states. While the excited state of the atom changes its momentum due to the photon absorption, the ground state remains unchanged, thus accomplishing the atom wave beam splitting. For a description of the quantum gravity gradiometer based on the Mach-Zehnder interferometer at JPL see (Yu et al. 2006). The second laser pulse π akts like a mirror in redirecting the atom wave, giving the momentum $(k_1 + k_2) \cdot \hbar$ and emitting a photon $k_2 \cdot \hbar$. Thus, after the sequence of three laser pulses, in the absence of gravitation or angular acceleration,

the two paths of the interferometer will be identical. However, if the atom experiences an acceleration during this time a phase difference $\Delta\phi$ will be accumulated

$$\Delta\phi = kgT^2 \tag{30.45}$$

that is proportional to the effective laser wave number k (frequency), acceleration (gravitation) g and where T is the interrogation time, i.e., the time between the light pulses. In the case of a quantum gravitational gradiometer, the interferometric measurement needs to be performed at two locations in close proximity in order to remove non-gravitational acceleration (e.g., air-drag) common to both locations. Thus, we may define two types of measurements: differential mode measurements

$$\begin{aligned}\Delta\phi_2 - \Delta\phi_1 &= k(g_2 + a)T^2 - k(g_1 + a)T^2 \\ &= k(g_2 - g_1)T^2\end{aligned} \tag{30.46}$$

and common mode measurements

$$\begin{aligned}\frac{1}{2}(\Delta\phi_2 + \Delta\phi_1) &= \frac{1}{2}\Big(k(g_2 + a)T^2 + k(g_1 + a)T^2\Big) \\ &= \frac{1}{2}k(g_2 + g_1)T^2 + kaT^2\end{aligned} \tag{30.47}$$

where a stands for the common non-gravitational acceleration (e.g., air-drag). Other effects such as angular accelerations due to the rotation of the instrument are the same as for classical GOCE gradiometry, see (Rummel et al. 2011). It is important to note from (30.45) and (30.46) that the time interval T at the ground is limited to a fraction of a second due to the full amount of gravity that is practically unlimited in space (weightlessness). From (30.45) and (30.46) we can also see that the sensitivity of the derived gravitational gradients increases by $1/T^2$ compared to $1/T$ which we typically have if the gravitational potential is derived using an atomic clock. Thus, atom interferometry is a very good candidate for future gravity field missions.

Let us now see how to measure the gravitational redshift along the orbit with only one satellite and to derive gravity and gravity gradients. A paper in the journal *Nature* (Muller et al. 2010), reported a ground laboratory experiment based on quantum interference of atoms that enables much more precise measurements of gravitational redshift, yielding an accuracy of 7×10^{-9}. It stated an improvement by a factor of $10\,000$ compared to the ACES mission goals, where it is anticipated that the gravitational redshift can be tested to a precision of 2 ppm. Interestingly, the same paper reports that it should be possible to improve the reported accuracy $10 - 100$ fold by more precise mapping of the local gravity gradient. Let us try to look at this experiment the other way round, with the main interest being to extract the gravity signal itself.

(Muller et al. 2010) triggered considerable discussion and a number of follow-up papers in the fundamental physics community, because it shows that gravitational redshift can be measured at a distance scale of micrometers to millimeters rather than the thousands of kilometers we typically have with GNSS, STE-QUEST or ACES missions in LEO. We repeat again here that the Compton frequency ω_C is very high since it includes the rest mass energy multiplied by c^2, e.g., for cesium one has $\omega_C/2\pi = 3.2 \times 10^{25}$ Hz. That is significantly higher than the frequency used to define the SI Second by means of cesium atomic clocks, or optical clocks in the near future. We have already demonstrated a symmetry principle that any error in the orbit position is suppressed by an error in orbit velocity by generating net redshift effect. Therefore, under differential conditions the net redshift effect cancels or is very small also for high matter-wave frequencies, but it can be accurately measured. This makes interesting to discuss this principle in mapping of gravity field from space.

The basic idea is similar to that of a Mach-Zehnder interferometer where the phases of two waves travelling on two different trajectories with a very small separation 0.1 mm are compared. A single wave is split by a laser pulse and after travelling on a different trajectory it is superimposed on the original wave by an

additional laser pulse that re-adjusts its trajectory. On both trajectories, quantum mechanics describes the atom as a de Broglie matter wave. As they arise from the same wave, their oscillations are initially in phase, but, travelling along different paths, their phases will be slightly shifted due to differences in the gravitational potential and this can be measured. Thus it would be very interesting to apply such a principle for a GOCE-type gradiometer length of 50 cm.

Time measured by a clock moving in curved spacetime is given in general relativity by

$$\tau = \int d\tau = \int (-g_{\mu\nu} dx^\mu dx^\nu)^{1/2} \tag{30.48}$$

with $g_{\mu\nu}$ describing the Schwarzschild metric (4×4 matrix) of space-time geometry at the location x^μ of the clock including the gravitational redshift and the special relativistic contribution due to the velocity of the clock. (30.48) can be calculated more easily by numerically integrating the differential frequency offset along the orbit, including the part due to potential and kinetic energy. Thus, if we introduce the total energy or Hamiltonian of the particle H^* (including the rest mass energy) with relativistic mass m along the orbit, we obtain for the accumulated phase for each matter wave

$$\Delta\varphi_{free} = \frac{1}{\hbar} \int H^* d\tau = \frac{1}{\hbar} \int mc^2 d\tau = \int \omega_C d\tau \tag{30.49}$$

with the Compton frequency $\omega_C = mc^2/\hbar$ we have already derived in (29.25). Although the velocity v of the satellite is relatively low compared to the speed of light c, it is considered by using the Lorentz factor $\gamma = 1/\sqrt{1 - v^2/c^2}$ in (29.7). However, here we are dealing only with the differential effects between two trajectories. On the other hand, frequency ω_C is very high, since it includes the rest mass energy multiplied by c^2, but the phase difference (30.49) can be measured (Muller et al. 2010). The accumulated phase $\Delta\varphi_{free}$ contains the relativistic contribution $\Delta\varphi_{redshift}$

$$\Delta\varphi_{free} = \Delta\varphi_{redshift} + \Delta\varphi_{time} + \Delta\varphi_{light} \tag{30.50}$$

and $\Delta\varphi_{time}$ is an additional phase due to time dilation considering special relativity. For the differential change in the trajectory, we will demonstrate later that $\Delta\varphi_{redshift} = -\Delta\varphi_{time}$, i.e., any orbit error or differential change in the trajectory is compensated for by the special relativistic term. Since both waves do not propagate along the same trajectory, the accumulated phase due to redshift $\Delta\varphi_{redshift}$ over the time interval T is

$$\Delta\varphi_{redshift} = \int_0^T \omega_C \frac{\Delta V}{c^2} dt \tag{30.51}$$

or assuming constant gravitation g (acceleration) during the laser pulse interaction

$$\Delta\varphi_{redshift} = \omega_C \int_0^T \frac{g \cdot z}{c^2} dt \tag{30.52}$$

where the differential trajectory is denoted as z. The third term in (30.50) $\Delta\varphi_{light}$ is the phase accumulation due to the interaction of the laser pulse with the wave, see (Muller et al. 2010). Since the laser pulse is used to split the original wave over the interval T and later to mirror the trajectory for superposition on the original wave, each laser pulse introduces an additional momentum $\Delta E = \hbar(k_1 + k_2 + k_3)$, where k_i denotes the wavenumber (angular frequency of the laser light), see (29.3). Assuming constant local gravity g during the interval T (Muller et al. 2010) we again obtain (30.45)

$$\Delta\varphi_{light} = k \cdot g T^2 \tag{30.53}$$

In (Muller et al. 2010) the pulse separation time was $T = 160\,\text{ms}$ and the peak separation of the trajectories $0.12\,\text{mm}$. Since differential velocity between the two waves is only due to the laser pulse we may write $\Delta\varphi_{light} = -\Delta\varphi_{time}$. Therefore, the measured phase after superposition of the two matter waves is driven only by the gravitational redshift, i.e., the interferometer phase is equal to the redshift phase $\Delta\varphi = \Delta\varphi_{redshift}$. Knowing the frequency of the laser light and measured phase, one can indirectly calculate gravity g.

By choosing different trajectories, i.e., three orthogonal directions, one could measure $\Delta\varphi_{redshift}$ in all three directions along the orbit. This would provide measurements of differential gravitational redshift and gravitational gradients in all three directions. Measured phase on one satellite could even be transferred to another satellite flying in formation, similar to very long baseline interferometry. In the case of a gravitational gradiometer, the measurement would be relative, thus a gravitational gradient could be used to remove the common mode accelerations (e.g., air-drag). The relationship between the orbit position and the velocity error in generating the error in the net redshift effect plays a key role, since the separated interferometer locations over the gradiometer baseline will experience a similar net frequency offset, although redshift and angular accelerations will be different. Thus, differential mode measurements could in principle be zero in this closed loop. The second innovation we are introducing here is the fact that the Compton frequency is extremely high and for cesium one obtains $\omega_C/2\pi = 3.2 \times 10^{25}\,\text{Hz}$. Therefore, for an accuracy of the gravitational potential measurements in the range of 10^{-19} we obtain from (29.23)

$$\frac{\Delta\omega}{\omega_C} = 10^{-19} \quad \rightarrow \quad \frac{\Delta\omega}{2\pi} = 10^{-19} \cdot 3.2 \times 10^{25}\,\text{Hz} = 3.2 \times 10^6\,\text{Hz} \tag{30.54}$$

$3.2\,\text{MHz}$ and it can be measured. If we now consider a LEO orbit, an eccentricity of approx. $10\,\text{km}$ gives variations in the gravitational potential and Compton frequency in the order of $1\,\text{THz}$, and this could be measured by an optical clock. A clock referenced to the Compton frequency using an optical frequency comb has been demonstrated in (Lan et al. 2013).

Although, both approaches outlined here and proposed for spaceborne gravitational gradiometry appear at first sight very similar, atoms move slower than light, thus atom interferometers are candidates for achieving greater inertial sensitivity than their optical counterparts based on quantum matter interference. However, a ground-based experiment performed by (Muller et al. 2010) reportedly achieved and stated an improvement by a factor of $10\,000$ compared to the ACES mission in terms of redshift. Applied to space, the sensitivity of matter wave interferometers will be improved significantly, by $1/T^2$ compared to the $1/T$ we typically have for clocks. This could bring an additional $3-4$ orders of magnitude in sensitivity compared to ground-based results, and is typically claimed for cold-atom clocks in space.

In a similar way, light-pulse atom interferometry, but with atomic point sources, has recently been demonstrated by the Kasevich Group at Stanford University (Dickerson et al. 2013). They report a measured acceleration sensitivity of $6.7 \times 10^{-12}\,g$ (single shot precision on the ground) providing information about rotation, acceleration, and even interferometer imperfections. They also reported the measurement of the Earth's rotation rate with a precision of $200\,\text{nrad/s}$. Since, again, space brings a longer interrogation time T, the results reported by (Dickerson et al. 2013) have the potential to reach $10^{-15}\,g$ and beyond in space. This pushes out significantly the boundaries of space-based mapping of the gravity field of the Earth and can be applied to gravity field mapping of other planets. Atom interferometry provides stability and accuracy that increase with measurement time and this is what we need for low-degree gravity and temporal gravity field variations.

30.5 Relativistic Gravity Field Determination – Towards the mm-Geoid and Unification of Terrestrial Reference Frames for Positioning, Time and Temporal Gravity

Can relativistic measurements of the gravitational potential contribute to global, regional and local gravity field determination? Is there any application of this new technique in reference frame realization for positioning, time and temporal gravity? Can this new geometrical technique unify all three reference frames, for gravity, time and positioning (Svehla et al. 2013b)?

Apart from a number of experiments related to fundamental physics, ACES (Cacciapuoti and Salomon 2009) and STE-QUEST (Schiller et al.) will be the first missions to demonstrate global relativistic geodesy, or determination of the *in situ* gravitational potential differences between ground stations by comparing the frequencies of optical clocks at distant locations. By making use of a dedicated metrology link, it is expected to be able to determine gravitational potential differences between ground stations to an accuracy of several centimetres in terms of geoid height (ACES) and with sub-centimetre accuracy with STE-QUEST. The orbit of the STE-QUEST satellite is designed in such a way that it allows optimal frequency comparison between timing labs on the ground, and at the same time enables measurements of the gravitational potential between perigee and apogee in the proposed elliptic orbit (HEO). In addition, it is expected that optical/microwave metrology links developed for these missions will be installed on board future GNSS and GEO satellites, serving the timing community in the realization of a global reference frame for time, i.e., UTC and TAI.

By comparing the frequencies of optical clocks in the global TAI network, it is expected that STE-QUEST and the forthcoming metrology ground-to-space links in GEO orbit and on board future GNSS satellites, will establish a global reference frame for time and the gravitational potential of the Earth. This reference frame could be used for the realization of TAI (International Atomic Time) as well as to support the realization of a global height system. A global height system is the basis of TAI, since the reference surface to define TAI is again the geoid. The GOCE mission has significantly contributed to the unification of global height systems by increasing the spatial resolution of the satellite-based gravity field of the Earth. However, only optical clocks could provide *in situ* measurements of the gravitational potential. Temporal gravity field maps are provided on a routine basis by the GRACE mission, however, with significantly lower resolution than that of the GOCE static gravity field. For any location on Earth, the omission error due to the high-frequency part of the static gravity model could be considered as constant and the gravitational potential at the location of the atomic clock will change with low frequency. Therefore, it will be very interesting to use the STE-QUEST mission to establish in future a unified terrestrial reference frame for positioning, time and the temporal gravity field of the Earth.

Some 20 years ago, GPS opened up the possibility of having an easy-to-use, three-dimensional positioning system on Earth. Although the coordinates of a point on the Earth can be determined using GPS with an accuracy of 1 mm in a very well-defined international terrestrial system, they are purely geometrical and do not contain any gravity information. This means that height as estimated by GPS is not referring to an equipotential gravity surface, but an arbitrarily chosen mean Earth ellipsoid in Euclidian space. Physical heights are measured along the plumbline, orthogonal to an equipotential surface of the Earth's gravity field. Figure 30.3 depicts three types of leveling: terrestrial, ocean and relativistic leveling. In all three cases the gravitational potential is leveled in a different way: measuring local gravity combined with spirit leveling, sea topography or relativistic frequency offset. By defining the reference gravity potential by W_A and the local gravity potential by W_B, the corresponding orthometric height H_B is related as follows

$$W_B - W_A = -\int_{H_{B'}}^{H_B} g \cdot dH = H_B \cdot \overline{g}_B \tag{30.55}$$
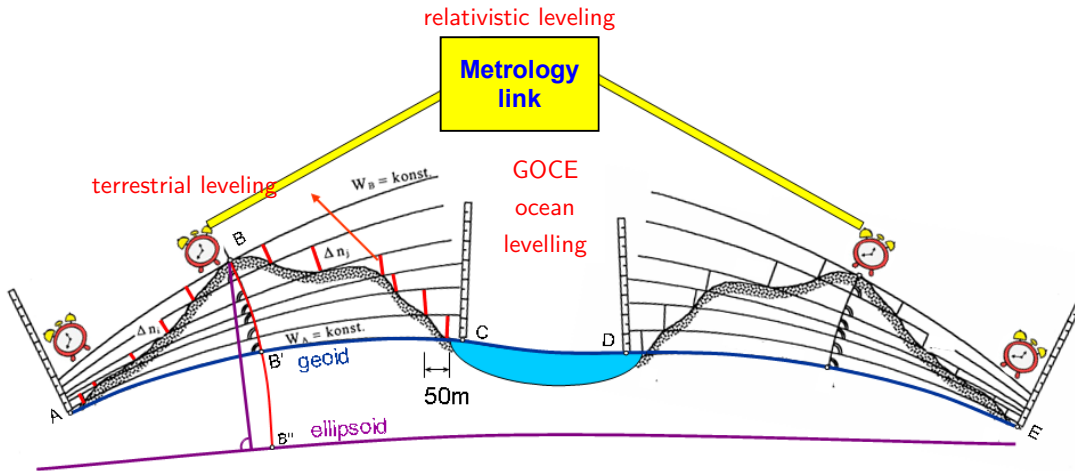
Figure 30.3 Terrestrial, ocean and relativistic leveling - three approaches for the realization of a height system. A new alternative (ocean leveling) is offered by the GOCE geoid in combination with ocean topography measured by satellite altimetry, (Rummel 2012), (Woodworth et al. 2012). However, only optical clocks can provide *in situ* measurements of the gravitational potential differences on a global scale.

The concept was realised by terrestrial leveling, by sequentially measuring gravity g and height differences dn on the surface of the Earth

$$W_B - W_A = -\int_{H_A}^{H_B} g \cdot dn \qquad (30.56)$$

In the classical definition, given by Listing in 1872, a geoid is defined as the particular equipotential surface nearest to mean sea level, see e.g., (Helmert 1884). In these terms, the geoid serves as a reference surface for measuring height and also to define a datum for the gravitational potential. (Bjerhammar 1985) defined a geoid in a relativistic way, i.e., as the surface where precise clocks run at the same rate and this surface is nearest to mean sea level. In a first approximation, the relation between the differences in the clock frequencies f_A and f_B and the gravitational potential is given as

$$\frac{f_A - f_B}{f_A} = \frac{W_B - W_A}{c^2} \qquad (30.57)$$

However, this concept has, as yet, never been realized. Instead, the vertical datum has been historically defined for a country or several countries by determing the mean sea level from observations at tide gauge station(s) taken over a long period of time. In this sense, the geoid was nearest to the mean sea level measured at some particular point on the ocean coastline. However, modern satellite altimetry missions such as Topex/Poseidon, Envisat, Jason-1/2 combined with new geoid models from e.g., GRACE and GOCE show that differences of mean sea level from an equipotential surface may reach up to several metres on a global scale, see e.g., (Fu and Cazenave 2001). Therefore, height systems based on different tide gauge stations may differ in their realisation of the geoid by several metres. On the world height system we refer to (Heck and Rummel 1990), (Rummel et al. 2012) and (Bašić and Rapp 1992).

The difference between the sea level and the reference equipotential surface of the Earth's gravity field that best fits the mean sea surface (geoid) is called Dynamic Ocean Topography (DOT) and it has a magnitude of $\pm 1 - 2$ m on a global scale, (Fu and Cazenave 2001). These deviations are caused by hydrodynamic processes such as variations in water density, currents, wind and atmospheric pressure. It is very important to mention that DOT is a direct measure of the heat and mass transport in the oceans. Following the report of the Intergovernmental Panel on Climate Change on the physical basis of climate change (Solomon et al. 2007),

the stability of large-scale circulation in the oceans in terms of DOT is of the highest relevance in global warming. The establishment of a global height reference system to monitor and observe DOT is a key issue for climate science and geodesy. Traditionally, height reference systems were defined for different countries and continents by the observed sea level, measured at isolated tide gauges and averaged over different time periods. As a result, current local physical height systems and their global separation do not support the observation of climate change on a global scale. Since 2001, the GPS Tide Gauge Benchmark Monitoring - Pilot Project (TIGA-PP) of the IGS has been monitoring vertical motion and geometrical coordinates of the tide gauges throughout the world in the ITRF. With the TIGA project, additional distinction can be made between absolute and relative sea level changes by taking into account the vertical uplift at the tide gauge stations. This provides an important contribution to absolute sea level monitoring by satellite altimetry, in particular to sea level rise and climate change studies. With the GOCE mission, we are, for the first time, in a position to establish a vertical reference system in the ITRF using ocean leveling, see Figure 30.3, (Rummel 2012; Woodworth et al. 2012). For this, the GOCE static gravity model can be combined with mean sea surface topography and dynamic ocean topography derived by satellite altimetry, GNSS-derived geometrical information on tide gauges and in terms of terrestrial physical height systems, gravity anomalies measured all over the world. Such a global vertical frame would serve as a reference frame for the global and uniform monitoring of climate change in terms of large scale ocean circulation and sea level rise, as well as uniform worldwide realization of physical heights for geodesy, surveying, cartography and positioning with GNSS. However, on the other hand, only optical clocks can provide *in situ* measurements of the gravitational potential differences on a global scale and are not limited by the errors of omission as are gravity field mapping missions. Therefore, it is expected that this new technique will find its place in geodesy and will lead to a unification of the terrestrial frame with the reference frames for time and gravity.

New satellite geodesy missions, such as CHAMP, GRACE and GOCE, have opened up a new era in mapping the global gravity field. While the static gravity field models are primarily provided by the CHAMP and GOCE missions, the GRACE mission has given, for the first time, an insight into the temporal variations of the gravity field, see e.g., (Tapley et al. 2004). However, the altitude of a LEO orbit limits the maximum resolution of such gravity field models. The best GRACE gravity field models have achieved a precision and resolution of about $1\,\mathrm{cm}$ over $600\,\mathrm{km}$, see, e.g., (Reigber et al. 2005), whereas in the case of GOCE, gravity models are provided up to a degree and order of $240 - 250$ (Rummel et al. 2011) with a resolution of about $80\,\mathrm{km}$ (half wavelength). However, Figure 30.4 (left) shows that for typical Earth topography with height variations of, e.g., $1000\,\mathrm{m}$ over $30\,\mathrm{km}$ horizontal distance, one may expect variations in the geoid of about $80\,\mathrm{cm}$ (see Figure 30.4 (right)). In mountainous regions, these gradients are significantly higher. Such high-frequency variations in the geoid are difficult to detect for a space gravity mission and require a combination of satellite, airborne and terrestrial gravity measurements including gravity anomalies, deflections of the vertical and GNSS/leveling points. The main disadvantage in measuring height differences by terrestrial leveling is the random walk effect of accumulated errors, since all continental leveling lines (e.g., in Europe, USA, Australia) are built up from short fragments of only e.g., $50 - 100\,\mathrm{m}$.

The so-called cm-geoid is still a challenge for many countries in Europe, especially in mountainous regions, where variations in density do not allow an accurate correction of the measurements, even if alternative definitions of height systems are employed or the Molodensky theory is used. Therefore, there is a need for a new technique in gravity field mapping that will provide a step forward in terms of accuracy and will follow the mm-positioning offered by GNSS.
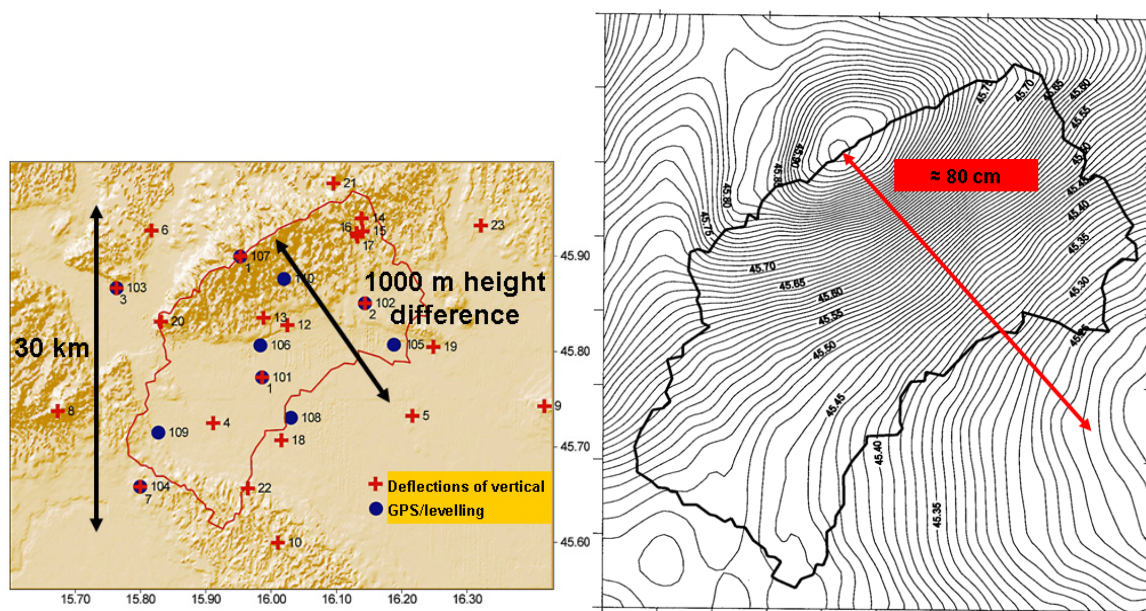
Figure 30.4 Typical topography of the Earth's surface (left) and corresponding geoid variations (right) determined by the combination of astronomically determined deflections of the vertical and a comparison between GPS and leveling heights (Švehla 1997).

The current precision level of regional height systems, in terms of gravity potential differences, is in the order of $1\,\mathrm{m}^2/\mathrm{s}^2$ (10 cm) with inconsistencies between these various systems of up to several $10\,\mathrm{m}^2/\mathrm{s}^2$ (several metres). The actual requirement in the context of GGOS (Global Geodetic Observing System) is $0.01\,\mathrm{m}^2/\mathrm{s}^2$ to $0.1\,\mathrm{m}^2/\mathrm{s}^2$ with permanent, i.e., dynamic monitoring. This requirement for high-precision height monitoring comes from the need to understand, on a global scale, processes such as sea level change, the global and coastal dynamics of ocean circulation, ice melting, glacial isostatic adjustment and land subsidence as well as the interaction of these processes (Rummel 2012). Only by means of monitoring in terms of gravity potential changes at the above level of precision can the change in ocean level be understood as a global phenomenon and data on purely geometric height changes be complemented by information about the associated density or mass changes.

Back in 2003, the International Association of Geodesy (IAG) started a special Inter-Commission Project (ICP1.2) between Commission 1 and Commission 2 under the title ″Vertical Reference Frames″. Making use of the terrestrial and modern satellite observations, ICP1.2 studied the consistent modeling of both geometric and gravimetric parameters to pave the way for the adoption of a Global Vertical Reference System and its realization, a unified Global Vertical Reference Frame (GVRF), see, e.g., (Ihde 2007). However, the realization of the Global Vertical Reference System started with the establishment of a regional vertical reference system in Europe. Since 1994, the IAG Sub-commission for Europe (EUREF) has enhanced the Unified European Leveling Network (UELN) and defined a European Vertical Reference System (EVRS). On the adoption of national geoid models and the work on the unified height system in Europe, see e.g., (Čolić, Pribičević, Švehla 1998). The latest version of EVRS was published under the name EVRF2007, see, e.g., (Sacher et al. 2008). The EVRF2007 datum was defined by 13 datum points distributed over the stable part of Europe and the results of the adjustment are given in geopotential numbers and normal heights, and they are reduced to the zero tidal system (Sacher et al. 2008).

30.6 The State of the Art in the Development of Optical Clocks and Metrology Links

Although the first optical clocks achieved a level of stability and accuracy in the order of 10^{-18}, at the moment there is no operational way to compare the frequencies of optical clocks on a global scale with the same level of stability and accuracy. Clockmakers face a dilemma: The more accurate clocks are, the more difficult it is to compare them. The clocks used by (Rosenband et al. 2008) were located in the same building, at the National Institute of Standards and Technology (NIST) in Boulder, Colorado, and compared using fiber links of a few hundred meters. Comparing clocks that are very far apart presents a different challenge. Precision timing signals between distant laboratories are currently transmitted over microwave networks or by satellites, but these fail at the levels of precision now being achieved with optical clocks. The first ground 5 km free-space coherent optical link over Paris (Djerroud et al. 2010) and the first fiber link over 1000 km in Germany (Predehl et al. 2012) have demonstrated a stability in frequency transfer at the 10^{-19} level and below, whereas a first prototype of the microwave link developed for the ACES mission reached a stability of 10^{-17} in the common view mode. On the other hand, an optical link at the 10^{-18} level and below is under development in Europe for the STE-QUEST mission and an independent development is being carried out in the USA.

The accuracy of measured gravitational redshift is driven by the accuracy of the optical clocks. Over the last decade, we have witnessed an improvement in clock accuracy by a factor of approximately 2 every five years. The latest record in the accuracy of an optical clock has been demonstrated at NIST with a fractional frequency inaccuracy of 8.6×10^{-18} after only 3 hours of averaging time, (Chou et al. 2009). But the really good news is that stability is an order of magnitude better. This clock is based on the quantum logic spectroscopy of an Al^+ ion. The frequency of the clock transition is compared to that of a previously constructed Al^+ optical clock with a statistical measurement uncertainty of 7×10^{-18}. The two clocks exhibit a fractional frequency difference of -1.8×10^{-18}, consistent with the accuracy limit of the older clock. (For more see (Chou et al. 2009).)

Ultra-stable clocks, matter-wave interferometers and atomic lasers based on Bose-Einstein condensation are developing rapidly and it is now conceivable to fly such a clock aboard the International Space Station (ISS) (ACES mission), see (Cacciapuoti and Salomon 2009). Space offers weightlessness and atoms can be cooled to such low temperatures that the Earth's gravity field represents a major perturbing effect on their motion. Microgravity conditions aboard the ISS allow these atoms to be kept in the observation volume (cloud) for several seconds (Cacciapuoti and Salomon 2009), much longer than is possible on the ground. This leads to increased stability and accuracy. Although a frequency stability of $10^{-16} - 10^{-17}$ over one day still does not meet the above requirements for gravity field determination, the latest developments in high-precision
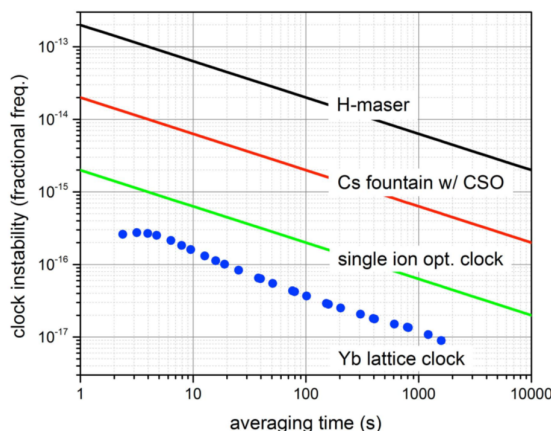

Figure 30.5 State of the art in clock frequency stability with $1.8 \cdot 10^{-18}$ in $20\,000$ s (Ludlow et al. 2013).

optical spectroscopy outperform today's state-of-the-art cesium clocks by one order of magnitude (Udem et al. 2002).

The cesium atomic clock was introduced in 1955, and led to the adoption of the international definition of the SI Second in 1967. The basis of the standard is the absorption of microwave radiation at 9.2 GHz by cesium atoms. Its measurement precision is determined primarily by the narrowness or spectral spread of the absorption in relation to its frequency. Physicists have since speculated that optical absorptions (in the near infrared, visible or ultraviolet) would make better frequency standards and clocks, because of their much higher frequency (approaching 1 million GHz). Essentially, if your clock ticks more quickly in a given time, your precision for time measurements improves with the tick frequency, (Gill 2000).

Atomic clocks became a reality in the mid-1950s with the development of the cesium clock. The time-keeping element in such a clock is a microwave transition in the cesium atom. The first clocks achieved an uncertainty of about 10^{-10}. By steady research and refinement, the clocks were improved until their uncertainty reached the level of about 5×10^{-16}. However, it is generally agreed that major improvements in cesium clocks are no longer likely. Fortunately, a new type of clock is now being realized. The timing element in an atomic clock is the frequency of a transition between energy levels in an atom or ion. The measured precision of the clock is proportional to the transition frequency, assuming that the ability to measure the frequency is
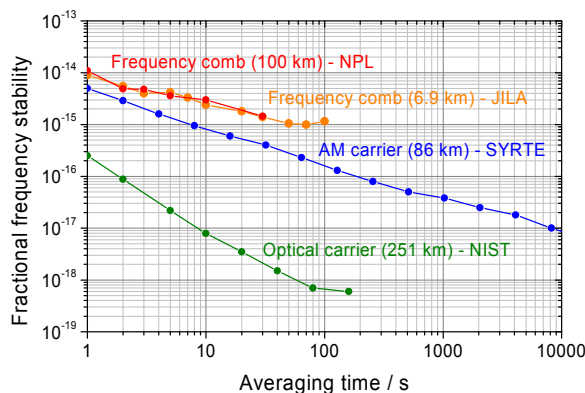

Figure 30.6 State of the art in optical clock comparison using frequency combs, (Gill et al. 2008).

kept constant. Because optical frequencies are higher than microwave frequencies by a factor of $\approx 10^5$, optical clocks hold the potential of being far more precise than the cesium clock (Kleppner 2008).

(Rosenband et al. 2008) reports the comparison of two atomic clocks based on the frequencies of optical transitions in single ions. One clock uses the Al^+ ion, whereas the other uses the Hg^+ ion. They measured the ratio of the frequencies of the two clocks to an uncertainty of 5.2×10^{-17}. This result is among the most precise measurements ever made in physics (in terms of relative precision). The clocks used by (Rosenband et al. 2008) employ a single ion that is confined in a trap by electric fields. The experimental challenge is to approach as closely as possible the ideal of a single particle at rest in space, free from all perturbations and measured as well as quantum mechanics permits. As with every high-precision measurement, the principal challenge was to evaluate the effects of perturbations and sources of uncertainty. Although the sources are quite different for the two clocks, their final uncertainties are approximately the same, yielding an overall uncertainty of 5.2×10^{-17} (Rosenband et al. 2008).

When precision is pushed to new levels, even more subtle effects must be taken into account. For instance, the error budget includes a small contribution at 1×10^{-18}, due to an uncertainty in the gravitational potential of the two clocks. This corresponds to a difference in their altitudes of 1 cm. This heralds one of the most interesting aspects of time keeping with optical clocks: The effects of general relativity that mix time with

gravity are starting to approach a point that will require rethinking the basic concept of "keeping time". (Kleppner 2008)

Ion-based atomic clocks currently achieve the highest accuracy because of their relative freedom from perturbations. However, neutral atom-based atomic clocks offer the advantage of much stronger signals, because the ion clocks use only a single particle, whereas neutral atom clocks typically use tens of thousands of atoms. There are numerous candidates for the new generation of optical atomic clocks, and eventually the second will be redefined based on one of them. However, that is unlikely to happen soon, because currently there is no obvious best choice for an ion or atom optical clock. The advances in optical clocks described by (Rosenband et al. 2008) and (Ludlow et al. 2008) represent a milestone in time keeping because both groups achieved uncertainties that are significantly below those of primary cesium time standards. These state-of-the-art optical atomic clocks rest on developments that stretch back more than 20 years. Enabling technologies include methods for trapping and cooling single ions developed by Wineland and his team in the 1980s; the laser cooling of atoms for which Chu, Cohen-Tannoudji, and Phillips received the Nobel Prize in 1997; the development of methods for ultrahigh optical and ultraviolet spectroscopy of ions by Bergquist and his team in the 1990s; and the invention of the femtosecond frequency comb and optical frequency metrology for which Hänsch and Hall received the Nobel Prize in 2005, (Kleppner 2008).

It will take some time to engineer an optical clock so that it can operate with the reliability and simplicity needed for practical applications, but once the goal is clearly in sight, this sort of engineering can move speedily. The question inevitably arises as to what the next generation of clocks will be useful for. One can point to basic tests such as the constancy of the fundamental constants, and possible applications such as geodesy, (Kleppner 2008). However, the best response to that question is simply to note that, when atomic clocks were invented 50 years ago, nobody was dreaming of GPS. The development of GPS illustrates the truth of the adage that revolutionary technologies are likely to generate revolutionary applications, (Kleppner 2008).

Over the last few years, optical clocks have reported several new records, e.g., (Nicholson et al. 2015) reports a total uncertainty of the JILA Sr clock to 2.1×10^{-18} in fractional frequency units and this gives us the confidence to believe that optical clocks will find an application in geodesy for gravity field determination.

31. References

Abramowitz M, Stegun IA (1965) Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables. J Geophys Res 92:1287–1294

Altamimi Z, Collilieux X & Métivier L, J Geod (2011) 85: 457. doi: 10.1007/s00190-011-0444-4

Anderson JD, Campbell JK, Ekelund JE, et al (2008) Anomalous Orbital-Energy Changes Observed during Spacecraft Flybys of Earth. Phys Rev Lett 100:091102. doi: 10.1103/PhysRevLett.100.091102

Arfken GB, Weber HJ, Weber H-J (1995) Mathematical Methods for Physicists, 4th edn. Academic Press

Arnold D, Dach R, Beutler G, et al (2014) Impact of GNSS Orbit Modelling on Reference Frame Parameters. Luxembourg

Ashby N (2003) Relativity in the Global Positioning System. Living Rev. Relativ. 6:1.doi:10.12942/lrr-2003-1

Balmino G, Borderies N (1978) Gravitational potential of solid bodies in the solar system. Celestial Mechanics and Dynamical Astronomy 17:113–119. doi: 10.1007/BF01371322

Bar-Sever Y, Kuang D (2004) New empirically derived solar radiation pressure model for GPS Satellites. JPL Technaical Report, NASA

Bašić T, Rapp RH (1992) Oceanwide Prediction of Gravity Anomalies and Sea Surface Heights Using Geos-3, Seasat and Geosat Altimeter Data and ETOPO5U Bathymetric Data. OSU Report 416, pp 1-89

Bates DR (1959) Some Problems concerning the Terrestrial Atmosphere above about the 100 km Level. Proc R Soc Lond A 253:451–462. doi: 10.1098/rspa.1959.0207

Bauch A, Achkar J, Bize S, et al (2006) Comparison between frequency standards in Europe and the USA at the $10-15$ uncertainty level. Metrologia 43:109–120. doi: 10.1088/0026-1394/43/1/016

Bauersima I (1983) NAVSTAR Global Positioning System (GPS) I. Berne, Switzerland

Baur O, Grafarend EW (2006) High-Performance GOCE Gravity Field Recovery from Gravity Gradient Tensor Invariants and Kinematic Orbit Information. In: Flury DJ, Rummel PDR, Reigber PDC, et al. (eds) Observation of the Earth System from Space. Springer Berlin Heidelberg, pp 239–253

Baur O (2013) Greenland mass variation from time-variable gravity in the absence of GRACE. Geophys Res Lett 40:4289–4293. doi: 10.1002/grl.50881

Baur O, Bock H, Ditmar P, et al (2013) Comparison of GOCE-GPS gravity fields derived by different approaches. p 9378

Bengtsson L, Robinson G, Anthes R, et al (2003) The Use of GPS Measurements for Water Vapor Determination. Bulletin of the American Meteorological Society 84:1249–1258

Bertiger W, Wu SC (1996) Single frequency GPS orbit determination for low earth orbiters, ION National Technical Meeting, Jan., 1996

Byun SH (2003). Satellite orbit determination using triple-differene GPS carrier phase in pure kinematic mode, J. of Geodesy 76: 569-585

Beutler G (1977) Integrale Auswertung von Satellitenbeobachtungen. SGK, Band 33, Zurich, Schwitzerland

Beutler G, Brockmann E, Gurtner W, et al (1994) Extended Orbit Modeling Techniques at the CODE Processing Center of the International GPS Service for Geodynamics (IGS): Theory and Initial Results. Manuscripta Geodaetica 19:367–386

Bjerhammar A (1985) On a relativistic geodesy. Journal of Geodesy 59:207–220. doi: 10.1007/BF02520327

Bock H (2003) Efficient Methods for Determining Precise Orbits of Low Earth Orbiters Using the Global Positioning System. PhD Thesis, Universität Bern, Schwitzerland

Bock H, Jäggi A, Svehla D, Beutler G, Hugentobler U, Visser P (2007) Precise orbit determination for the GOCE satellite using GPS; Advances in Space Research, Vol. 39, Nr. 10, pp 1638-1647, doi: 10.1016/j.asr.2007.02.053, 2007

Bock H, Jäggi A, Meyer U, et al (2011) GPS-derived orbits for the GOCE satellite. J Geod. doi: 10.1007/s00190-011-0484-9

Bock H, Jäggi A, Beutler G, Meyer U (2014) GOCE: precise orbit determination for the entire mission. J Geod 1–14. doi: 10.1007/s00190-014-0742-8

Boehm J, Niell A, Tregoning P, Schuh H (2006a) Global Mapping Function (GMF): A new empirical mapping function based on numerical weather model data. Geophys Res Lett. doi: 10.1029/2005GL025546

Boehm J, Werl B, Schuh H (2006b) Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data. J Geophys Res 111:406.

Border JS, Kursinski ER (1991) Deep space tracking and frequency standards. Proceedings of the IEEE 45th Annual Symposium on Frequency Control, IEEE 91CH2965-2:594–607. doi: 10.1109/FREQ.1991.145957

Boving D, Droz F, Mosset P, et al (2009) Space Passive Hydrogen Maser, Performances, lifetime data and GIOVE-B related telemetries. In: IGNSS2009 Conference Proceedings

Bowman BR, Kent Tobiska W, Marcos FA, et al (2008a) A New Empirical Thermospheric Density Model JB2008 Using New Solar and Geomagnetic Indices

Bowman BR, Kent Tobiska W, Marcos FA, Valladares C (2008b) The JB2006 empirical thermospheric density model. Journal of Atmospheric and Solar-Terrestrial Physics 70:774–793. doi: 10.1016/j.jastp.2007.10.002

Braxmaier C, Dittus H, Foulon B, et al (2012) Astrodynamical Space Test of Relativity using Optical Devices I (ASTROD I)—a class-M fundamental physics mission proposal for cosmic vision 2015–2025: 2010 Update. Experimental Astronomy 34:181–201. doi: 10.1007/s10686-011-9281-y

Bronstein IN, Semendjajew KA (1996) Teubner-Taschenbuch der Mathematik. Teubner, Stuttgart

Bruinsma S, Loyer S, Lemoine JM, et al (2003) The impact of accelerometry on CHAMP orbit determination. Journal of Geodesy 77:86–93. doi: 10.1007/s00190-002-0304-3

Burša M, Šíma Z, Pick M (1984) Equatorial flattening and principal moments of inertia of the earth. Studia Geophysica et Geodaetica 28:9–10. doi: 10.1007/BF01587106

Cacciapuoti L, Salomon C (2009) Space clocks and fundamental tests: The ACES experiment. The European Physical Journal - Special Topics 172:57–68. doi: 10.1140/epjst/e2009-01041-7

Casotto S (1993) The mapping of Kaula's solution into the orbital reference frame. Celestial Mechanics and Dynamical Astronomy 55:223–241. doi: 10.1007/BF00692511

Cazenave A, Dominh K, Guinehut S, et al (2009) Sea level budget over 2003–2008: A reevaluation from GRACE space gravimetry, satellite altimetry and Argo. Global and Planetary Change 65:83–88. doi: 10.1016/j.gloplacha.2008.10.004

Cazenave A, Llovel W (2010) Contemporary Sea Level Rise. Annual Review of Marine Science 2:145–173. doi: 10.1146/annurev-marine-120308-081105

Chesley SR, Ostro SJ, Vokrouhlický D, et al (2003) Direct Detection of the Yarkovsky Effect by Radar Ranging to Asteroid 6489 Golevka. Science 302:1739–1742. doi: 10.1126/science.1091452

Chou C-W, Hume DB, Koelemeij JCJ, et al (2009) Frequency Comparison of Two High-Accuracy Al+ Optical Clocks

Chou CW, Hume DB, Rosenband T, Wineland DJ (2010) Optical Clocks and Relativity. Science 329:1630–1633. doi: 10.1126/science.1192720

Ciufolini I, Pavlis EC (2004) A confirmation of the general relativistic prediction of the Lense-Thirring effect. Nature 431:958–960. doi: 10.1038/nature03007

Ciufolini I, Paolozzi A, Pavlis E, et al (2012) Testing General Relativity and gravitational physics using the LARES satellite. Eur Phys J Plus 127:1–7. doi: 10.1140/epjp/i2012-12133-8

Clohessy WH, Wiltshire RS (1960) Terminal Guidance System for Satellite Rendezvous. Journal of the Aerospace Sciences 27:653–658. doi: 10.2514/8.8704

Colombo OL (1986) Ephemeris errors of GPS satellites. Bull Géodésique 60:64–84. doi: 10.1007/BF02519355

Cox CM, Chao BF (2002) Detection of a Large-Scale Mass Redistribution in the Terrestrial System Since 1998. Science 297:831–833. doi: 10.1126/science.1072188

Čolić K, Pribičević B, Švehla D (1998) First cm–Geoid in The Republic of Croatia – The Capital City Zagreb Pilot Project. IAG–Symposium: Second Continental Workshop on the Geoid in Europe, Budapest, March 10–14 1998, Reports of the Finnish Geodetic Institute 98:4, Proceedings (Eds. M. Vermeer and J. Adam), pp. 245–249, Masala, Finland

Dach R, Hugentobler U, Schildknecht T, et al (2005) Precise continuous time and frequency transfer using GPS carrier phase. Proceedings of the 2005 IEEE International. pp 329–336

Dach R, Schildknecht T, Hugentobler U, et al (2006) Continuous geodetic time-transfer analysis methods. Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 53:1250–1259. doi: 10.1109/TUFFC.2006.1665073

Davies K (1990) Ionospheric Radio (IEE Electromagnetic Waves Series, Vol. 31). Institution of Electrical Engineers

de Broglie L (1924) Recherches sur la théorie des quanta (Researches on the quantum theory), Thesis (Paris). Ann. Phys. (Paris) 3, 22 (1925)

Debs JE, Robins NP, Close JD (2013) Measuring Mass in Seconds. Science 339:532–533. doi: 10.1126/science.1232923

Dedes GC, Mueller II (1989) Baseline estimation with semidynamic and geometric satellite methods. Bull Geodesique 63:99–114. doi: 10.1007/BF02519145

Delporte J, Mercier F, Laurichesse D, Galy O (2007) Fixing integer ambiguities for GPS carrier phase time transfer. Frequency Control Symposium 2007/21st EFTF. IEEE International. pp 927–932

Delporte J, Mercier F, Laurichesse D, Galy O (2008) GPS Carrier-Phase Time Transfer Using Single-Difference Integer Ambiguity Resolution. International Journal of Navigation and Observation. doi: 10.1155/2008/273785

Desai S (2007) [IGSREPORT-15607] Wk 1449 JPL Analysis Report. http://igscb.jpl.nasa.gov/mail/igsreport/2007/msg00988.html

Dickerson SM, Hogan JM, Sugarbaker A, et al (2013) Multiaxis Inertial Sensing with Long-Time Point Source Atom Interferometry. Phys Rev Lett 111:083001. doi: 10.1103/PhysRevLett.111.083001

Dickey JO, Marcus SL, de Viron O, Fukumori I (2002) Recent Earth Oblateness Variations: Unraveling Climate and Postglacial Rebound Effects. Science 298:1975–1977. doi: 10.1126/science.1077777

Ditmar P, Kuznetsov V, van der Sluijs A, et al (2006) "DEOS_CHAMP-01C_70": a model of the Earth's gravity field computed from accelerations of the CHAMP satellite. Journal of Geodesy 79:586–601. doi: 10.1007/s00190-005-0008-6

Djerroud K, Acef O, Clairon A, et al (2010) Coherent optical link through the turbulent atmosphere. Opt Lett 35:1479–1481. doi: 10.1364/OL.35.001479

Domingos RC, de Moraes RV, De Almeida Prado AFB (2008) Third-Body Perturbation in the Case of Elliptic Orbits for the Disturbing Body. Mathematical Problems in Engineering. doi: 10.1155/2008/763654

Doornbos E, Bruinsma S, Koppenwallner G, et al (2012) Thermospheric density and wind from GOCE thruster activation and accelerometer data. In: EGU General Assembly Conference Abstracts. p 5634

Dow JM, Neilan RE, Gendt G (2005) The International GPS Service: Celebrating the 10th anniversary and looking to the next decade. Advances in Space Research 36:320–326. doi: 10.1016/j.asr.2005.05.125

Droste S, Ozimek F, Udem T, et al (2013) Optical-Frequency Transfer over a Single-Span 1840 km Fiber Link. Phys Rev Lett 111:110801. doi: 10.1103/PhysRevLett.111.110801

Edmonds AR (1960) Angular Momentum in Quantum Mechanics. Princeton University Press, Princeton

Einstein A (1905) Zur Elektrodynamik bewegter Körper. Annalen der Physik 322:891–921. doi: 10.1002/andp.19053221004

Einstein A (1916) Die Grundlage der allgemeinen Relativitätstheorie. Annalen der Physik 354:769–822. doi: 10.1002/andp.19163540702

Elósegui P, Davis JL, Jaldehag RTK, et al (1995) Geodesy using the Global Positioning System: The effects of signal scattering on estimates of site position. J Geophys Res 100:9921–9934

Emeljanov NV, Kanter AA (1989) A method to compute inclination functions and their derivatives. Manuscripta Geodaetica 14:77–83

ESA Portal (2014) Galileo Facts and Figures, http://www.esa.int/Our_Activities/Navigation

Estey LH, Meertens CM (1999) TEQC: The Multi-Purpose Toolkit for GPS/GLONASS Data. GPS Solutions 3:42–49. doi: 10.1007/PL00012778

Feldman U, Landi E, Schwadron NA (2005) On the sources of fast and slow solar wind. Journal of Geophysical Research: Space Physics 110:n/a–n/a. doi: 10.1029/2004JA010918

Fengler MJ, Freeden W, Michel V (2004) The Kaiserslautern multiscale geopotential model SWITCH-03 from orbit perturbations of the satellite CHAMP and its comparison to the models EGM96, UCPH2002_02_0.5, EIGEN-1s and EIGEN-2. Geophysical Journal International 157:499–514. doi: 10.1111/j.1365-246X.2004.02209.x

Fixler JB, Foster GT, McGuirk JM, Kasevich MA (2007) Atom Interferometer Measurement of the Newtonian Constant of Gravity. Science 315:74–77. doi: 10.1126/science.1135459

Fliegel HF, Gallini TE (1996) Solar force modeling of block IIR Global Positioning System satellites. Journal of Spacecraft and Rockets 33:863–866. doi: 10.2514/3.26851

Flohrer C, Otten M, Springer T, Dow J (2011) Generating precise and homogeneous orbits for Jason-1 and Jason-2. Advances in Space Research 48:152–172. doi: 10.1016/j.asr.2011.02.017

Földváry L, Svehla D, Gerlach C, Wermuth M, Gruber T, Rummel R, Rothacher M, Frommknecht B, Peters T, Steigenberger P (2005) Gravity model TUM-2Sp based on the energy balance approach and kinematic CHAMP orbits; Earth Observation with CHAMP, Results from Three Years in Orbit, pp 13-18, Springer, doi : 10.1007/3-540-26800-6_2, 2005

Fortescue P, Swinerd G, Stark J (2011) Spacecraft Systems Engineering. John Wiley & Sons

Friedman AR, Hwang Y-T, Chiang JCH, Frierson DMW (2013) Interhemispheric Temperature Asymmetry over the Twentieth Century and in Future Projections. Journal of Climate 26:5419–5433. doi: 10.1175/JCLI-D-12-00525.1

Fu VL-L, Cazenave A (2001) Satellite Altimetry and Earth Sciences. Academic Press

Ge M, Gendt G, Rothacher M, et al (2007) Resolution of GPS carrier-phase ambiguities in Precise Point Positioning (PPP) with daily observations. J Geod 82:389–399. doi: 10.1007/s00190-007-0187-4

Gendt G (2006) [IGSMAIL-5438]: IGS switch to absolute antenna model and ITRF2005

GEO (2005) The Global Earth Observing System of Systems (GEOSS) - 10-Year Implementation Plan. http://earthobservations.org/

Gerlach C, Földvary L, Švehla D, et al (2003) A CHAMP-only gravity field model from kinematic orbits using the energy integral. Geophys Res Lett 30:2037. doi: 10.1029/2003GL018025

Gerlach C, Sneeuw N, Visser P, Svehla D (2003) CHAMP gravity field recovery using the energy balance approach; Advances in Geosciences, Vol. 1, pp 73-80, ISSN 1680-7340, doi: 10.5194/adgeo-1-73-2003, 2003, Externer Link, Download.

Gerlach C, Sneeuw N, Visser P, Svehla D (2003) CHAMP gravity field recovery with the energy balance approach: first results; in: Reigber, C.; Lühr, H.; Schwintzer, P. (eds.) First CHAMP Mission Results for Gravity, Magnetic and Atmospheric Studies, pp 134-139, doi: 10.1007/978-3-540-38366-6_20, 2003

Giacaglia GEO, Burša M (1980) Transformations of spherical harmonics and applications to geodesy and satellite theory. Studia Geophysica et Geodaetica 24:1–11. doi: 10.1007/BF01628375

Gill E (2006) Together in Space. Potentials and Challenges of Distributed Space Systems. Inaugural speech. TU Delft. TU Delft, The Netherlands

Gill P (2000) Precision measurements: Optical clocks coming of age. Nature 407:579–580. doi: 10.1038/35036680

Gill P, Margolis H, Curtis A, et al (2008) Optical Atomic Clocks for Space - Final Report

Gold K, Bertiger W, Wu S, et al (1994) GPS Orbit Determination for the Extreme Ultraviolet Explorer. Journal of The Institute of Navigation 41:337–352

Goldstein JD (1984) The Effect of Coordinate System Rotations on Spherical Harmonic Expansions: A Numerical Method. J Geophys Res 89:4413–4418

Gonzalez F (2010) Poster on Galileo Attitude from the Galileo Project at ESA. Newcastle, UK

Gooding R, Wagner C (2008) On the inclination functions and a rapid stable procedure for their evaluation together with derivatives. Celestial Mechanics and Dynamical Astronomy 101:247–272. doi: 10.1007/s10569-008-9145-6

Goossens S, Matsumoto K, Liu Q, et al (2010) Lunar gravity field determination using SELENE same-beam differential VLBI tracking data. Journal of Geodesy 1–24

Gordon D, Ma C, MacMillan D, et al (2010) Effects of ICRF2 on the TRF, CRF, and EOP's. In: Reference Frames for Applications in Geosciences (REFAG2010). Paris, France

Gross R, Beutler G, Plag H-P (2009) Integrated scientific and societal user requirements and functional specifications for the GGOS. In: Global Geodetic Observing System. pp 209–224

Guerova G, Bettems J-M, Brockmann E, Matzler C (2006) Assimilation of COST 716 Near-Real Time GPS data in the nonhydrostatic limited area model used at MeteoSwiss. Meteorology and Atmospheric Physics 91:149–164. doi: 10.1007/s00703-005-0110-6

Haines B, Bar-Sever Y, Bertiger W, et al (2004) One-Centimeter Orbit Determination for Jason-1: New GPS-Based Strategies. Marine Geodesy 27:299–318

Hamilton W (1853) Lectures on Quaternions. Hodges and Smith, Royal Irish Academy, Dublin

Hansen JR (1987) Engineer in charge: a history of the Langley Aeronautical Laboratory, 1917-1958. Scientific and Technical Information Office, National Aeronautics and Space Administration

Heck B., Rummel R. (1990) Strategies for Solving the Vertical Datum Problem Using Terrestrial and Satellite Geodetic Data. Sea Surface Topography and the Geoid. International Association of Geodesy Symposia, vol 104. Springer, New York, NY, doi: 10.1007/978-1-4684-7098-7_14

Hedin AE (1991) Extension of the MSIS Thermosphere Model into the Middle and Lower Atmosphere. J Geophys Res 96:1159–1172

Hedin AE (1987) MSIS-86 Thermospheric Model. J Geophys Res 92:4649–4662

Hedin AE, Biondi MA, Burnside RG, et al (1991) Revised Global Model of Thermosphere Winds Using Satellite and Ground-Based Observations. J Geophys Res 96:7657–7688

Hedin AE, Fleming EL, Manson AH, et al (1996) Empirical wind model for the upper, middle and lower atmosphere. Journal of Atmospheric and Terrestrial Physics 58:1421–1447. doi: 10.1016/0021-9169(95)00122-0

Hein G, Rodriguez JA, Wallner S, et al (2007) Envisioning a Future GNSS System of Systems. Inside GNSS March/April 2007:64–72

Heiskanen WA, Moritz H (1967) Physical Geodesy. W.H.Freeman & Co Ltd

Helmert FR (1884) Die mathematischen und physikalischen Theorien der höheren Geodäsie. B. G. Teubner

Henkel P (2008) Zuverlässige Schätzung der Trägerphasenmehrdeutigkeiten bei Galileo. Kolloquium Satellitennavigation, TU München

Hobbs G, Coles W, Manchester RN, et al (2012) Development of a pulsar-based time-scale. MNRAS 427:2780–2787. doi: 10.1111/j.1365-2966.2012.21946.x

Hobson EW (1931) The theory of spherical and ellipsoidal harmonics. Univ. Press, Cambridge

Holme R, de Viron O (2013) Characterization and implications of intradecadal variations in length of day. Nature 499:202–204. doi: 10.1038/nature12282

Holmes SA, Featherstone WE (2002) A unified approach to the Clenshaw summation and the recursive computation of very high degree and order normalised associated Legendre functions. Journal of Geodesy 76:279–299. doi: 10.1007/s00190-002-0216-2

Holzwarth R, Hänsch TW, Sizmann A, et al (2008) High Accuracy Absolute Long Distance Metrology using Ultrafast Lasers: Frequency Combs and Spectral Interference. In: Proceeding from 3rd International Symposium on Formation Flying, Missions and Technologies. ESTEC/ESA, The Netherlands

Hugentobler U, Ploner M, Schildnecht T, Beutler G (1999) Determination of resonant geopotential terms using optical observations of geostationary satellites. Advances in Space Research 23:767–770. doi: 10.1016/S0273-1177(99)00153-2

Hugentobler U, Gruber T, Steigenberger P, Angermann D, Bouman J, Gerstl M, Richter B (2012) GGOS Bureau for Standards and Conventions: Integrated standards and conventions for geodesy; in: Kenyon, S. C.; Pacino, M. C.; Marti, U. J. (eds.) Geodesy for Planet Earth, IAG Symposia, Vol. 136, pp 995-998, Springer, doi: 10.1007/978-3-642-20338-1_124, 2012

Hugentobler U (2012) The Development of the IGS in 2011 The Governing Board's Perspective. IGS Technical Report 2011, pp 3-10, IGS Central Bureau, Jet Propulsion Laboratory, 2012

Hulley GC, Pavlis EC (2007) A ray-tracing technique for improving Satellite Laser Ranging atmospheric delay corrections, including the effects of horizontal refractivity gradients. J Geophys Res. doi: 10.1029/2006JB004834

Hwang C, Tseng TP, Lin T, Svehla D, Schreiner B (2009) Precise orbit determination for the FORMOSAT-3/COSMIC satellite mission using GPS, Journal of Geodesy, Vol. 83, Nr. 5, pp 477-489, Springer, ISSN 0949-7714, doi: 10.1007/s00190-008-0256-3, 2009

Hwang C, Tseng TP, Lin TJ, Svehla D, Hugentobler U, Chao B (2010) Quality assessment of FORMOSAT-3/COSMIC and GRACE GPS observables: analysis of multipath, ionospheric delay and phase residual in orbit determination; GPS Solutions, Vol. 14, Nr. 1, pp 121-131, doi: 10.1007/s10291-009-0145-0, 2010

Iess L, Asmar S, Tortora P (2009) MORE: An advanced tracking experiment for the exploration of Mercury with the mission BepiColombo. Acta Astronautica 65:666–675. doi: 10.1016/j.actaastro.2009.01.049

Ihde J (2007) Consideration of a Global Vertical Reference System (GVRS) in the IERS Conventions. In: Paper presented at the IERS Workshop on Conventions. Sevres

Ijssel J van den, Visser P, Patino Rodriguez E (2003) Champ precise orbit determination using GPS data. Advances in Space Research 31:1889–1895. doi: 10.1016/S0273-1177(03)00161-3

Ilk KH (1983) Ein Beitrag zur Dynamik ausgedehnter Körper : Gravitationswechselwirkung. Beck, München

International Association of Geomagnetism and Aeronomy WGV-MP members, Finlay CC, Maus S, et al (2010) International Geomagnetic Reference Field: the eleventh generation. Geophysical Journal International 183:1216–1230. doi: 10.1111/j.1365-246X.2010.04804.x

Irsigler M (2008) Multipath Propagation, Mitigation and Monitoring in the Light of Galileo and the Modernised GPS. Universität der Bundeswehr München

Jäggi A, Bock H, Prange L, et al (2011) GPS-only gravity field recovery with GOCE, CHAMP, and GRACE. Advances in Space Research 47:1020–1028. doi: 10.1016/j.asr.2010.11.008

Jäggi A, Hugentobler U, Beutler G (2006) Pseudo-Stochastic Orbit Modeling Techniques for Low-Earth Orbiters. Journal of Geodesy 80:47–60. doi: 10.1007/s00190-006-0029-9

James TS, Ivins ER (1997) Global geodetic signatures of the Antarctic ice sheet. J Geophys Res 102:605–633.

Jekeli C, Lee J, Kwon J (2007) On the computation and approximation of ultra-high-degree spherical harmonic series. Journal of Geodesy 81:603–615. doi: 10.1007/s00190-006-0123-z

Ji S, Chen W, Zhao C, et al (2007) Single epoch ambiguity resolution for Galileo with the CAR and LAMBDA methods. GPS Solut 11:259–268. doi: 10.1007/s10291-007-0057-9

Kasevich M (2013) Fundamental Physics Tests by Atom Interferometry. In: STE-QUEST Science Workshop. ESTEC 22-23 May 2013. ESTEC/ESA, The Netherlands

Kaula WM (1961) Analysis of Gravitational and Geometric Aspects of Geodetic Utilization of Satellites. Geophysical Journal International 5:104–133. doi: 10.1111/j.1365-246X.1961.tb00417.x

Kaula WM (1966) Theory of Satellite Geodesy: Applications of Satellites to Geodesy. Blaisdell Publishing Company, Waltham, Massachusettss

Kedar S, Hajj GA, Wilson BD, Heflin MB (2003) The effect of the second order GPS ionospheric correction on receiver positions. Geophys Res Lett 30:1829. doi: 10.1029/2003GL017639

Kleppner D (2008) PHYSICS: A Milestone in Time Keeping. Science 319:1768–1769

Kleusberg A (1980) The similarity transformation of the gravitational potential close to the indentity. Manuscripta Geodaetica 5:241–256

Kodet J, Plötz J, Schreiber U (2013) Co-location of space geodetics techniques in Space and on the ground

Koepf W (1998) Hypergeometric Summation. Vieweg Verlag

Kopeikin SM (2003) The post-Newtonian treatment of the VLBI experiment on September 8, 2002. Physics Letters A 312:147–157(11). doi: 10.1016/S0375-9601(03)00613-3

Kotsakis C A study on the reference frame consistency in recent Earth gravitational models. Journal of Geodesy. doi: 10.1007/s00190-008-0227-8

Kouba J (2004) Improved relativistic transformations in GPS. GPS Solutions 8:170–180. doi: 10.1007/s10291-004-0102-x

Lan S-Y, Kuan P-C, Estey B, et al (2013) A Clock Directly Linking Time to a Particle's Mass. Science 339:554–557. doi: 10.1126/science.1230767

Langley RB (1998) Propagation of the GPS signals. In: Teunissen PJG, Kleusberg A (Eds) GPS for Geodesy, 2nd edn. Springer

Lau L, Cross P (2007) Development and testing of a new ray-tracing approach to GNSS carrier-phase multipath modelling. Journal of Geodesy 81:713–732. doi: 10.1007/s00190-007-0139-z

Laurichesse D, Mercier F (2007) Integer ambiguity resolution on undifferenced GPS phase measurements and its application to PPP. Proceedings of ION GNSS 20th International Technical Meeting of the Satellite Division 839–848

Libbrecht KG (1985) Practical considerations for the generation of large-order spherical harmonics. Solar Physics 99:371–373. doi: 10.1007/BF00157320

LIGO Scientific Collaboration and Virgo Collaboration, Abbott BP, Abbott R, et al (2016) Observation of Gravitational Waves from a Binary Black Hole Merger. Phys Rev Lett 116:061102. doi: 10.1103/PhysRevLett.116.061102

Lucchesi DM, Ciufolini I, Andrés JI, et al (2004) LAGEOS II perigee rate and eccentricity vector excitations residuals and the Yarkovsky–Schach effect. Planetary and Space Science 52:699–710. doi: 10.1016/j.pss.2004.01.007

Luderer G, Pietzcker RC, Bertram C, et al (2013) Economic mitigation challenges: how further delay closes the door for achieving climate targets. Environ Res Lett 8:034033. doi: 10.1088/1748-9326/8/3/034033

Ludlow AD, Zelevinsky T, Campbell GK, et al (2008) Sr Lattice Clock at 1 x 10-16 Fractional Uncertainty by Remote Optical Evaluation with a Ca Clock. Science 319:1805–1808

Ludlow A, Sherman J, Hinkley N, et al (2013) The Yb lattice clock (and others!) at NIST for space-based applications. In: STE-QUEST Science Workshop. ESTEC 22-23 May 2013. ESTEC/ESA, The Netherlands

Luthcke SB, Zelensky NP, Rowlands DD, et al (2003) The 1-Centimeter Orbit: Jason-1 Precision Orbit Determination Using GPS, SLR, DORIS, and Altimeter Data. Marine Geodesy 26:399–421

MacMillan WD (1930) The theory of the potential. McGraw-Hill Book Company, inc

Marchenko AN, Schwintzer P (2003) Estimation of the Earth's tensor of inertia from recent global gravity field solutions. Journal of Geodesy 76:495–509. doi: 10.1007/s00190-002-0280-7

Marini JW, Murray CW (1973) Correction of laser range tracking data for atmospheric refraction at elevations above 10 degrees. Goddard Space Flight Center, Greenbelt, MD

Masters G, Richards-Dinger K (1998) On the efficient calculation of ordinary and generalized spherical harmonics. Geophysical Journal International 135:307–309. doi: 10.1046/j.1365-246X.1998.00622.x

Matsakis D, Lee M, Dach R, et al (2006) GPS Carrier Phase Analysis Noise on the USNO-PTB Baselines. In: International Frequency Control Symposium and Exposition, 2006 IEEE. pp 631–636

Mattioni L, Belloni M, Berthoud P, et al (2002) The Development of a Passive Hydrogen Maser Cclock for the Galileo Navigation System. In proceeding of: 34th Annual Precise Time and Time Interval (PTTI) Meeting:161–170

Mayer-Gürr T, Ilk KH, Eicker A, Feuchtinger M (2005) ITG-CHAMP01: a CHAMP gravity field model from short kinematic arcs over a one-year observation period. Journal of Geodesy 78:462–480. doi: 10.1007/s00190-004-0413-2

Mayer-Gürr T, Eicker A, Ilk KH (2006) Gravity Field Recovery from GRACE-SST Data of Short Arcs. In: Flury DJ, Rummel PDR, Reigber PDC, et al. (eds) Observation of the Earth System from Space. Springer Berlin Heidelberg, pp 131–148

Mayer-Gürr T, Eicker A, Kurtenbach E, Ilk K-H (2010) ITG-GRACE: Global Static and Temporal Gravity Field Models from GRACE Data. In: Flechtner FM, Gruber T, Güntner A, et al. (eds) System Earth via Geodetic-Geophysical Space Techniques. Springer Berlin Heidelberg, pp 159–168

McCarthy DD, Petit G (2004) IERS Conventions (2003) (IERS Technical Note 32). Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main

Mendes VB, Prates G, Pavlis EC, et al (2002) Improved mapping functions for atmospheric refraction correction in SLR. Geophys Res Lett 29:1414. doi: 10.1029/2001GL014394

Mendes VB, Pavlis EC (2004) High-accuracy zenith delay prediction at optical wavelengths. Geophys Res Lett. doi: 10.1029/2004GL020308

Mercier F, Laurichesse D (2007) Receiver/Payload hardware biases stability requirements for undifferenced widelane ambiguity blocking. In: 1st Colloquium on Scientifics and Fundamental Aspects of the Galileo Programme. Toulouse, France

Milani A, Nobili A, Farinella P (1987) Non-Gravitational Perturbations and Satellite Geodesy. Adam Hilger, Bristol

Montenbruck O, Gill E (2000) Satellite Orbits: Models, Methods, and Applications. Springer

Montenbruck O, Gill E (2002) Ionospheric Correction for GPS Tracking of LEO Satellites. The Journal of Navigation 55:293–304. doi: 10.1017/S0373463302001789

Montenbruck O, Kroes R (2003) In-flight performance analysis of the CHAMP BlackJack GPS Receiver. GPS Solutions 7:74–86. doi: 10.1007/s10291-003-0055-5

Montenbruck O, Garcia-Fernandez M, Yoon Y, et al (2009) Antenna phase center calibration for precise positioning of LEO satellites. GPS Solutions 13:23–34. doi: 10.1007/s10291-008-0094-z

Montenbruck O, Hugentobler U, Dach R, et al (2012) Apparent clock variations of the Block IIF-1 (SVN62) GPS satellite. GPS Solut 16:303–313. doi: 10.1007/s10291-011-0232-x

Montenbruck O (2013) MultiGNSS Working Group. In: International GNSS Service Technical Report 2012. IGS Central Bureau, JPL, Pasadena

Montenbruck O, Steigenberger P, Hugentobler U (2014) Enhanced solar radiation pressure modeling for Galileo satellites. J Geod 1–15. doi: 10.1007/s00190-014-0774-0

Moritz H (1980) Advanced physical geodesy. http://adsabs.harvard.edu/abs/1980QB281.M77...... Accessed 13 Apr 2010

Moritz H (1993) Geometry, relativity, geodesy. Wichmann, Karlsruhe

Muellerschoen R, Iijima B, Meyer R, et al (2004) Real-Time Point-Positioning Performance Evaluation of Single-Frequency Receivers Using NASA's Global Differential GPS System,. In: Proceedings of the 17th International Technical Meeting (ION GNSS 2004). Long Beach, CA, USA, pp 1872–1880

Muller H, Peters A, Chu S (2010) A precision measurement of the gravitational redshift by the interference of matter waves. Nature 463:926–929. doi: 10.1038/nature08776

Newton I (1687) Philosophiæ naturalis principia mathematica. Royal Society, London

Nicholson TL, Campbell SL, Hutson RB, et al (2015) Systematic evaluation of an atomic clock at $2 \times 10{-}18$ total uncertainty. Nature Communications 6:6896. doi: 10.1038/ncomms7896

Niell AE (1996) Global mapping functions for the atmosphere delay at radio wavelengths. J Geophys Res 101:3227–3246

Niell AE (2001) Preliminary evaluation of atmospheric mapping functions based on numerical weather models. Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy 26:475–480. doi: 10.1016/S1464-1895(01)00087-4

NOAA (2009) NOAA's National Geophysical Data Center (NGDC). ftp://ftp.ngdc.noaa.gov /STP/SOLAR_DATA/SOLAR_RADIO/FLUX.

Operation CSCAS, Wertz JR (1978) Spacecraft Attitude Determination and Control.

Otsubo T, Appleby G, Gibbs P (2001) GLONASS Laser Ranging Accuracy with Satellite Signature Effect. Surveys in Geophysics 22:509–516

Otsubo T, Appleby GM (2003) System-dependent center-of-mass correction for spherical geodetic satellites. Journal of Geophysical Research: Solid Earth 108:n/a–n/a. doi: 10.1029/2002JB002209

Pail R, Goiginger H, Schuh W-D, et al (2010) Combined satellite gravity field model GOCO01S derived from GOCE and GRACE. Geophysical Research Letters 37:n/a–n/a. doi: 10.1029/2010GL044906

Pail R, Bruinsma S, Migliaccio F, et al (2011) First GOCE gravity field models derived by three different approaches. J Geod 85:819–843. doi: 10.1007/s00190-011-0467-x

Pavlis EC (1985) On the geodetic applications of simultaneous range differences to LAGEOS. J Geophys Res 90:9431–9438. doi: 10.1029/JB090iB11p09431

Pavlis E (2012) Laser Ranging: Scientific Accomplishments of the Past and Requirements for the Future. In: International Technical Laser Workshop 2012. "Satellite, Lunar and Planetary Laser Ranging: characterizing the space segment." Frascati, Italy

Pavlis NK, Weiss MA (2000) The relativistic red shift with 2×10-17 uncertainty at NIST, Boulder, Colorado, USA. In: Frequency Control Symposium and Exhibition, 2000. Proceedings of the 2000 IEEE/EIA International. pp 642–650

Pavlis NK, Weiss MA (2003) The relativistic redshift with 3×10−17 uncertainty at NIST, Boulder, Colorado, USA. Metrologia 40:66–73

Pearlman MR, Degnan JJ, Bosworth JM (2002) The International Laser Ranging Service. Advances in Space Research 30:135–143. doi: 10.1016/S0273-1177(02)00277-6

Peltier WR, Jiang X (1996) Glacial isostatic adjustment and Earth rotation: Refined constraints on the viscosity of the deepest mantle. J Geophys Res 101:3269–3290

Peterseim N, Schlicht A, Stummer C, Yi W (2011) Impact of Cross Winds in Polar Regions on GOCE Accelerometer and Gradiometer Data. In: Proc. of "4th International GOCE User Workshop", Munich, Germany 31 March – 1 April 2011 (ESA SP-696, July 2011). Munich, Germany

Petit G, Tavella P (1996) Pulsars and time scales. Astronomy and Astrophysics 308:290–298

Petit G (1998) Importance of common framework for realization of space-time reference systems. In: Towards an Integrated Global Geodetic Observing system (IGGOS). Springer, pp 1–7

Petit G, Luzum B (2010) IERS Conventions (2010). Verlag des Bundesamts für Kartographie und Geodäsie

Picone JM, Hedin AE, Drob DP, Aikin AC (2002) NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. J Geophys Res 107:1468. doi: 10.1029/2002JA009430

Planck Collaboration, Ade PAR, Aghanim N, et al (2013) Planck 2013 results. XVI. Cosmological parameters.

Pound RV, Rebka GA (1959) Gravitational Red-Shift in Nuclear Resonance. Phys Rev Lett 3:439–441. doi: 10.1103/PhysRevLett.3.439

Predehl K, Grosche G, Raupach SMF, et al (2012) A 920-Kilometer Optical Fiber Link for Frequency Metrology at the 19th Decimal Place. Science 336:441–444. doi: 10.1126/science.1218442

Press HW, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical Recipes 3rd Edition: The Art of Scientific Computing. Cambridge University Press

Rasel E (2013) STE-QUEST Differential Atom Interferometer. In: STE-QUEST Science Workshop. ESTEC 22-23 May 2013. ESTEC/ESA, The Netherlands

Ray J (2002) C1/P1 biases for Leica and Trimble 5700 receivers

Reigber C, Schmidt R, Flechtner F, et al (2005) An Earth gravity field model complete to degree and order 150 from GRACE: EIGEN-GRACE02S. Journal of Geodynamics 39:1–10. doi: 10.1016/j.jog.2004.07.001

Reubelt T, Götzelmann M, Grafarend E (2006) Harmonic Analysis of the Earth's Gravitational Field from Kinematic CHAMP Orbits based on Numerically Derived Satellite Accelerations. In: Observation of the Earth System from Space. pp 27–42

Rievers B, Lämmerzahl C, List M, et al (2009) New powerful thermal modelling for high-precision gravity missions with application to Pioneer 10/11. New J Phys 11:113032. doi: 10.1088/1367-2630/11/11/113032

Riley W (2014) The Stable32 Program for Frequency Stability Analysis. Hamilton Technical Services, Beaufort, SC 29907

Rochat P, Droz F, Wang Q, Froidevaux S (2012) Atomic Clocks and Timing Systems in Global Navigation Satellite Systems

Rodriguez-Solano CJ, Hugentobler U, Steigenberger P (2012) Impact of Albedo Radiation on GPS Satellites. In: Kenyon S, Pacino MC, Marti U (eds) Geodesy for Planet Earth. Springer Berlin Heidelberg, pp 113–119

Rosborough GW, Tapley BD (1987) Radial, transverse and normal satellite position perturbations due to the geopotential. Celestial Mechanics and Dynamical Astronomy 40:409–421. doi: 10.1007/BF01235855

Rosenband T, Hume DB, Schmidt PO, et al (2008) Frequency Ratio of Al+ and Hg+ Single-Ion Optical Clocks; Metrology at the 17th Decimal Place. Science 319:1808–1812

Rothacher M, Mervart L (1996) The Bernese GPS Software ver. 4.0. Astronomical Institute, University of Berne, Switzerland

Rothacher M, Beutler G, Herring TA, Weber R (1999) Estimation of nutation using the Global Positioning System. J Geophys Res 104:4835–4859

Rothacher M, Svehla D (2002) Report on the activities of the subcommission on precise orbit determination for low Earth orbiting satellites; in: Drewes, H.; Bosch, W.; Hornik, H. (eds.) IAG Commission VIII - International Coordination of Space Techniques for Geodesy and Geodynamics (CSTG): Progress Report 2001, CSTG Bulletin, Nr. 17, pp 64-69, Deutsches Geodätisches Forschungsinstitut, 2002

Rothacher M, Schmid R, Steigenberger P, Svehla D, Thaller D (2004) Combination of the space-geodetic techniques for monitoring the Earth's system; EOS Transactions AGU, Fall Meeting Supplement, Abstract G22A-01, Vol. 85, Nr. 47, AGU, 2004

Rothacher M, Beutler G, Behrend D, et al (2009) The future Global Geodetic Observing System. In: Global Geodetic Observing System. ISBN:978-3-642-02686-7, Springer Berlin Heidelberg, pp 237–272, doi: 10.1007/978-3-642-02687-4_9

Rubincam DP (1987) LAGEOS Orbit Decay Due to Infrared Radiation From Earth. J Geophys Res 92:1287–1294

Rummel R, Bosch W, Drewes H (2000) Towards an Integrated Global Geodetic Observing System (IGGOS). International Association of Geodesy Symposia, Vol 120, Springer-Verlag Berlin and Heidelberg, doi: 10.1007/978-3-642-59745-9

Rummel R (2006) Physical Geodesy - Lecture Notes. TU München, Germany

Rummel R, Yi W, Stummer C (2011) GOCE gravitational gradiometry. J Geod 85:777–790. doi: 10.1007/s00190-011-0500-0

Rummel R (2012) Height unification using GOCE. Journal of Geodetic Science 2:355–362

Sacher M, Ihde J, Liebsch G, Mäkinen J (2008) EVRF07 as Realization of the European Vertical Reference System. In: Paper presented at the Symposium of the IAG Sub-commission for Europe (EUREF) in Brussels, June 18-21, 2008. Brussels

Santamaría-Gómez A, Wöppelmann G, Marcos M, et al (2014) Observed differential geocentric sea-level rise between hemispheres over the past hundred years. Luxembourg

Schaer S (1999) Mapping and predicting the earth's ionosphere using Global Positioning System. Schweizerische Geodätische Kommission

Schiller S, Tino G, Gill P, Svehla D, et al Einstein Gravity Explorer–a medium-class fundamental physics mission. Experimental Astronomy. doi: 10.1007/s10686-008-9126-5

Schmid R, Rothacher M (2003) Estimation of elevation-dependent satellite antenna phase center variations of GPS satellites. Journal of Geodesy 77:440–446. doi: 10.1007/s00190-003-0339-0

Schmidt A (1899) Formeln zur Transformation der Kugelfunktionen bei linearer Anderung des Koordinatensystems. Zeitschrift für Mathematik und Physik 44:327–335

Schmidt M, Kusche J, Loon JP, et al (2005) Multiresolution representation of a regional geoid from satellite and terrestrial gravity data. In: Gravity, Geoid and Space Missions. pp 167–172

Shum CK, Abusali, PAM, Lee H, Ogle J, Raney RK, Ries JC, Smith WHF, Svehla D, Zhao C (2008) Orbit determination requirements for ABYSS: a proposed space station science payload; Advances in the Astronautical Sciences, Vol. 130, pp 1207-1218, 2008

Shum CK, Abusali PAM, Kuo CY, Lee H, Ogle J, Raney, RK, Ries JC, Smith WHF, Svehla D, Zhao C (2009) Orbit accuracy requirement for ABYSS: the space station radar altimeter to map global bathymetry; IEEE Geoscience and Remote Sensing Letters, Vol. 6, Nr. 4, pp 653-657, ISSN 1545-598X, doi: 10.1109/LGRS.2009.2012877, 2009.

Sigl R (1985) Introduction to potential theory. Abacus Press

Smith JM, Olver FWJ, Lozier DW (1981) Extended-Range Arithmetic and Normalized Legendre Polynomials. ACM Trans Math Softw 7:93–105. doi: 10.1145/355934.355940

Sneeuw N (2000) A semi-analytical approach to gravity field analysis from satellite observations, PhD Thesis, DGK Reihe C. TU München

Solarić N, Solarić M, Švehla D (2012) Nove revolucionarne mogućnosti u geodeziji koje pružaju otkrića za koja su dobivene Nobelove nagrade za fiziku 2005. i 1997. godine. Geodetski list 66 (89):1–19

Solomon S, Qin D, Manning M, et al (2007) Climate Change 2007 - The Physical Science Basis. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Steigenberger P, Rothacher M, Dietrich R, et al (2006) Reprocessing of a global GPS network. J Geophys Res 111:B05402. doi: 10.1029/2005JB003747

Steigenberger P, Hugentobler U, Loyer S, et al (2014) Galileo orbit and clock quality of the IGS Multi-GNSS Experiment. Advances in Space Research. doi: 10.1016/j.asr.2014.06.030

Švehla D (1995) Orthodrome. Rector's Award for the Best Student Work. University of Zagreb, Faculty of Geodesy, Zagreb

Švehla D (1996) Loxodrome on the Sphere. Rector's Award for the Best Student Work. University of Zagreb, Faculty of Geodesy, Zagreb

Švehla D (1997) Preliminary Determination of Astrogeodetic Geoid of the City of Zagreb. Diploma Thesis, Univerity of Zagreb, Faculty of Geodesy

Švehla D, Rothacher M (2002) Kinematic Orbit Determination of LEOs Based on Zero– or Double–Difference Algorithms Using Simulated and Real SST Data. In: Vistas for Geodesy in the New Millenium. International Association of Geodesy Symposia, vol 125. Springer, Berlin, Heidelberg, pp 322–328, doi: 10.1007/978-3-662-04709-5_53

Švehla D, Rothacher M (2003a) CHAMP double–difference kinematic POD with ambiguity resolution. In: First CHAMP Mission Results for Gravity, Magnetic and Atmospheric Studies. Springer-Verlag, Potsdam, pp 70–77, doi: 10.1007/978-3-540-38366-6_11

Švehla D, Rothacher M (2003b) Kinematic and reduced-dynamic precise orbit determination of low Earth orbiters. Adv. Geosci., 1, 47-56, 2003, doi: 10.5194/adgeo-1-47-2003

Švehla D, Rothacher M (2003c) Kinematic and Reduced–Dynamic Precise Orbit Determination of CHAMP satellite over one year using zero–differences. In: Geophysical Research Abstracts. European Geophysical Society

Švehla D, Rothacher M (2004a) Two Years of CHAMP Kinematic Orbits for Geosciences. European Geosciences Union, 1st General Assembly, 25–30 April 2004, Nice, France. Geophysical Research Abstracts, European Geophysical Society Vol. 6. ISSN:1029–7006

Švehla D, Rothacher M (2004b) Kinematic Orbits for LEO Satellites - a New Product. In: Proceedings from the IGS Workshop 2004: Celebrating a Decade of the International GPS Service - Workshop and Symposium, M. Meindl (Ed.). Astronomical Institute, University of Bern, Bern, Switzerland

Švehla D, Rothacher M (2004c) Formation Flying of LEO Satellites Using GPS. Eos Trans. AGU, 85(47), Fall Meet. Suppl. Abstract SF53A-0735. American Geosciences Union 2004, 13-17 December 2004, San Francisco, US.

Švehla D, Rothacher M (2004d) Notes on the GOCE SST. TU München, Germany

Švehla D, Rothacher M (2005a) Kinematic positioning of LEO and GPS satellites and IGS stations on the ground. Advances in Space Research 36:376–381, doi: 10.1016/j.asr.2005.04.066

Švehla D, Rothacher M (2005b) Kinematic Precise Orbit Determination for Gravity Field Determination. In: A Window on the Future of Geodesy. International Association of Geodesy Symposia, vol 128. Springer, Berlin, Heidelberg, pp 181–188, doi: 10.1007/3-540-27432-4_32

Švehla D (2006) Earth orientation quaternion and modeling the satellite orbit using quaternions. COSPAR Scientic Assembly 2006, Beijing, China

Švehla D, Földváry L (2006) From Kinematic Orbit Determination to Derivation of Satellite Velocity and Gravity Field. In: Observation of the Earth System from Space. Springer, pp 177–192, doi: 10.1007/3-540-29522-4_13

Švehla D, Rothacher M (2006a) Impact of a LEO Formation and a LEO/GPS Dual Constellation on the IGS Products. IGS Workshop 2006 "Perspectives and Visions for 2010 and Beyond". Darmstadt, Germany

Švehla D, Rothacher M, Salomon C, Ziebart M (2006b) GALILEO on board the International Space Station and highly accurate GNSS/ACES time and frequency transfer based on phase clocks. IGS Workshop "Perspectives and Visions for 2010 and Beyond." Darmstadt, Germany

Švehla D, Heinze M, Gruber A, Rothacher M (2006a) Improvements in the Reference Frame Determination using Galileo. DGFI, Munich, Germany

Švehla D, Rothacher M (2006b) Clock Comparison Based on the Phase GPS Satellite Clock and Synergy with the Future ACES Mission. European Frequency and Time Forum EFTF2006. Braunschweig, Germany

Švehla D, Heinze M (2007) Positioning with the four GNSS systems: GPS, GLONASS, GALILEO and BEIDOU based on phase clocks. EGU General Assebly 2007, Vienna, Austria, Geophysical Research Abstracts, Vol. 9, 10412, 2007 SRef-ID: 1607-7962/gra/EGU2007-A-10412

Švehla D (2007a) Navigation System Based on Master Clocks and Two-Way Links in Space. 1st Colloquium Scientific and Fundamental Aspects of the GALILEO Programme. Toulouse, France

Švehla D (2007b) Technical Note: Support to ACES Precise Orbit Determination, TU München

Švehla D (2007c) POD Center Munich. ESTEC/ESA, The Netherlands

Svehla D (2008) Geodesy Part of the ACES Mission: GALILEO on Board the International Space Station. In: ACES and Future GNSS-Based Earth Observation and Navigation, 26 – 27 May 2008. TU München, Germany, in IAPG/FESG-Schriftenreihe, Nr. 28, IAPG, FESG, ISBN (Print) 978-3-934205-27-7, ISSN 1437-8280, 2009

Švehla D, Heinze M, Rothacher M, et al (2008) Combined processing of GIOVE-A and GPS measurements using zero- and double-differences. EGU General Assembly 2008, Vienna, Austria, Geophysical Research Abstracts, Vol. 10, EGU2008-A-11383, 2008 SRef-ID: 1607-7962/gra/EGU2008-A-11383

Švehla D (2008a) From GIOVE and ACES to the GALILEO-2 Concept for Geosciences. In: Proceedings from the 14th General Assembly of WEGENER - WEGENER 2008. Darmstadt, Germany

Švehla D (2008b) Combination of kinematic orbit and analytical perturbation theory for the determination of precise orbit and gravity field. EGU General Assembly 2008, Vienna, Austria, Geophysical Research Abstracts, Vol. 10, EGU2008-A-11428, 2008 SRef-ID: 1607-7962/gra/EGU2008-A-11428

Švehla D (2008c) A novel design for the navigation system and proposal to unify the timing and the positioning system using GIOVE Follow-on. EGU General Assembly 2008, Vienna, Austria, Geophysical Research Abstracts, Vol. 10, EGU2008-A-11022, 2008 SRef-ID: 1607-7962/gra/EGU2008-A-11022

Svehla, D (2009) ACES and future GNSS-based Earth observation and navigation; IAPG/FESG-Schriftenreihe, Nr. 28, Institut für Astronomische und Physikalische Geodäsie, Forschungseinrichtung Satellitengeodäsie, ISBN (Print) 978-3-934205-27-7, ISSN 1437-8280, 2009

Svehla D (2010a) Complete relativistic modelling of the GIOVE-B clock parameters and its impact on POD, track-to-track ambiguity resolution and precise timing applications. IGS Workshop 2010. Newcastle, UK

Svehla D (2010b) Geometry Aapproach to Model Satellite Dynamics Based on Numerical Integration or Analytical Theory. 38th COSPAR Scientic Assembly 2010. Bremen, Germany

Svehla D, Dow JM, Escobar D (2010a) GPS Single Frequency Precise Orbit Determination of LEO Satellites with Cm-Level Accuracy. 38th Assembly of the Committee on Space Research, COSPAR-2010 18-25 July 2010. Bremen, Germany

Svehla D, Flohrer C, Otten M, et al (2010b) Instantaneous Reference Frame Realization by Means of Combination of Space Geodesy Techniques Onboard Jason-2 Satellite. ESA Living Planet Symposium 2010, 28 June 2010 - 2 July 2010, Bergen, Norway

Svehla D, Navarro-Reyes D (2011a) First laser measurements of Europe's Galileo satellites made from Chile. ESA Press Release (Navigation): http://www.esa.int/Our_Activities/Operations/First_laser_measurements_of_Europe_s_Galileo_satellites_made_from_Chile

Svehla D, Peytavi G, Flohrer C, et al (2011b) Geometrical Approach to Monitor Orbit Perturbations of GNSS Satellites for High Precision Positioning. In: 3rd International Colloquium - Scientific and Fundamental Aspects of the Galileo Programme. Copenhagen, Denmark

Svehla D, Floberghagen R, Haagmans R, et al (2012) SLR Measurements of the Forthcoming ESA Earth Observation and Fundamental Physics Missions and Their Applications in the Reference Frames Realization. International Technical Laser Workshop 2012. 5-9.,11.2012. Frascati, Italy, pp. 1-2

Svehla D, Haagmans R, Floberghagen R, et al (2013a) Geometrical SLR Approach for Reference Frame Determination - The First SLR Double-Difference Baseline. IAG Scientific Assembly 2013, Potsdam, Germany, (reviewed for IAG Series 143, Springer)

Svehla D, Rothacher M, Hugentobler U, et al (2013b) STE-QUEST - Space Geodesy Mission for Celestial and Terrestrial Reference Frame Realization. In: STE-QUEST Science Workshop. ESTEC 22-23 May 2013. ESTEC/ESA, The Netherlands

Svehla D, Rothacher M, Hugentobler U, et al (2013c) 10 Years of Precise Geometrical Positioning in Space - Geometrical Model of Solar Radiation Pressure Based on High-Performing Clocks Onboard Galileo. IAG General Assembly 2013, Potsdam, Germany, (reviewed for IAG Series 143, Springer)

Svehla D, Rothacher M, Appleby G, et al (2014) High-Accuracy Double-Difference SLR Approach with GNSS Satellites for Terrestrial Reference Frame Realization. REFAG 2014, Luxembourg

Svehla D, Rothacher M, Hugentobler U, et al. (2014a) Terrestrial and Celestial Reference Frame Realization with Highly Elliptical Orbit - The ESA STE-QUEST Mission. EGU General Assembly 2014, Geophysical Research Abstracts, Vol. 16, EGU2014-7934-2, 2014

Svehla D, Cacciapuoti L, Rothacher M (2015a) Noise model of the Galileo "mm-Clock" - geometrical mapping of the orbit perturbations using a clock on board Galileo satellites. IUGG General Assembly 2015

Svehla D, Rothacher M, Appleby G, et al (2015b) Double-Difference SLR Approach with GNSS, GEO, LAGEOS and LLR. IUGG General Assembly 2015, Prague

Svehla D, Rothacher M, Cacciapuoti L (2016) Thermal Re-Radiation Acceleration in the GNSS Orbit Modelling Based on Galileo Clock Parameters. IGS Workshop 2016, Sydney, Australia

Svehla D, Cacciapuoti L, Petit G, Rothacher M (2017) How Stable is the Primary Clock Onboard Galileo Satellites? First Geometrical Mapping of the Orbit Perturbations Using a Clock in Space. Submitted to Journal of Geodesy (reviewed)

Sneeuw N, Gerlach C, Svehla D, Gruber C (2003) A first attempt at time variable gravity recovery from CHAMP using the energy balance approach; Gravity and Geoid: Proceedings of 3rd Meeting of the International Gravity and Geoid Commission, Thessaloniki, 2002, pp 237-242, ZITI-Publishing, 2003

Sneeuw N, Gerlach C, Földváry L, Gruber T, Peters T, Rummel R, Svehla D (2005) One year of time-variable CHAMP-only gravity field models using kinematic orbits; in: Sansò, F. (eds.) A Window on the Future of Geodesy, IAG Symposia, Vol. 128, pp 288-293, doi: 10.1007/3-540-27432-4_49, 2005

Tapley BD, Watkins MM, Ries JC, et al (1996) The Joint Gravity Model 3. J Geophys Res 101:28,029–28,049

Tapley BD, Bettadpur S, Ries JC, et al (2004) GRACE Measurements of Mass Variability in the Earth System. Science 305:503–505. doi: 10.1126/science.1099192

Tapley B, Ries J, Bettadpur S, et al (2005) GGM02 – An improved Earth gravity field model from GRACE. Journal of Geodesy 79:467–478. doi: 10.1007/s00190-005-0480-z

Teunissen PJG (2001) The probability distribution of the ambiguity bootstrapped GNSS baseline. Journal of Geodesy 75:267–275. doi: 10.1007/s001900100172

Thaller D (2009) Inter-technique combination based on homogeneous normal equation systems including station coordinates, Earth orientation and troposphere parameters. PhD Thesis, TU München. doi: 10.2312/GFZ.b103-08153

Thaller D, Dach R, Seitz M, et al (2010) Combination of GNSS and SLR using satellite co-locations. In: Reference Frames for Applications in Geosciences (REFAG2010). Paris, France

Thaller D, Dach R, Seitz M, et al (2011) Combination of GNSS and SLR observations using satellite co-locations. J Geod 85:257–272. doi: 10.1007/s00190-010-0433-z

Tseng, Tzu-Pang, Hwang Ch, Yang SK (2012) Assessing attitude error of FORMOSAT-3/COSMIC satellites and its impact on orbit determination. Advances in Space Research 49(9):1301–1312, doi: 10.1016/j.asr.2012.02.007

Turyshev SG, Toth VT, Kinsella G, et al (2012) Support for the Thermal Origin of the Pioneer Anomaly. Phys Rev Lett 108:241101. doi: 10.1103/PhysRevLett.108.241101

Udem T, Holzwarth R, Hansch TW (2002) Optical frequency metrology. Nature 416:233–237. doi: 10.1038/416233a

Urschl C, Beutler G, Gurtner W, et al (2007) Contribution of SLR tracking data to GNSS orbit determination. Advances in Space Research 39:1515–1523. doi: 10.1016/j.asr.2007.01.038

Visser P, van den Ijssel J, van Helleputte T, Bock H, Jäggi A, Beutler G, Svehla D, Hugentobler U, Heinze, M (2009) Orbit determination for the GOCE satellite; Advances in Space Research, Vol. 43, Nr. 5, pp 760-768, Elsevier, ISSN 0273-1177, doi: 10.1016/j.asr.2008.09.016 , 2009

Visser P, van den Ijssel J, van Helleputte T, Bock H, Jäggi A, Beutler G, Hugentobler U, Svehla D (2007) Rapid and precise orbit determination for the GOCE satellite; Proceedings of the 3rd International GOCE User Workshop, ESA SP-627 , pp 235-239, ISBN (Print) 92-9092-938-3, ISSN 1609-042X, 2007

Vermeer M (1983) Chronometric levelling. Rep. Finnish Geodetic Inst. 83, 1–7

Vessot RFC, Levine MW, Mattison EM, et al (1980) Test of Relativistic Gravitation with a Space-Borne Hydrogen Maser. Phys Rev Lett 45:2081. doi: 10.1103/PhysRevLett.45.2081

Waller P, Gonzalez F, Binda S, et al (2009) Update on the In-Orbit performances of GIOVE clocks. In: Frequency Control Symposium, 2009 Joint with the 22nd EFTF. IEEE International. pp 388–392

Wang Q, Mosset P, Droz F, Rochat P (2013) Lifetime of space Passive Hydrogen Maser. In: (EFTF/IFC), 2013 pp 973–976

Wenzel HG (1998) Ultra high degree geopotential models GPM98A and GPM98B to degree 1800. In: Proc Joint Meeting Int Gravity Commission and Int Geoid Commission, Budapest. pp 71–80

Wermuth M, Gerlach C, Svehla D, Földváry L (2004) Comparison of different gravity field solution methods applied to CHAMP gravity field modelling; in: Proceedings of the 1st International Workshop on Gravity Field Research, Österreichische Beiträge zu Meteorologie und Geophysik, Heft 31, pp 45-50, Zentralanstalt für Meteorologie und Geodynamik (ZAMG), ISSN 1016-6254, 2004

Wigner EP (1959) Group Theory and Its Application to Quantum Mechanics of Atomic Spectra. Academic Press Inc, New York

Williams GJ, Turyshev S, et al (2009) Lunar Science and Lunar Laser Ranging. A white paper submitted to the Panel on Inner Planets – Mercury, Venus, and the Moon of The National Academies Planetary Science Decadal Survey, US

Williams J, Lightsey EG, Yoon SP, Schutz BE (2002) Testing of the ICESat BlackJack GPS Receiver Engineering Model

Wittwer T, Klees R, Seitz K, Heck B (2008) Ultra-high degree spherical harmonic analysis and synthesis using extended-range arithmetic. Journal of Geodesy 82:223–229. doi: 10.1007/s00190-007-0172-y

Wolf P, Petit G (1995) Relativistic theory for clock syntonization and the realization of geocentric coordinate times. Astronomy and Astrophysics 304:653–661

Wolf P, Bordé CJ, Clairon A, et al (2009) Quantum physics exploring gravity in the outer solar system: the SAGAS project. Exp Astron 23:651–687. doi: 10.1007/s10686-008-9118-5

Woodworth PL, Hughes CW, Bingham RJ, Gruber T (2012) Towards worldwide height system unification using ocean information. Journal of Geodetic Science 2:302–318.

Wöppelmann G, Marcos M, Santamaría-Gómez A, et al (2014) Evidence for a differential sea level rise between hemispheres over the twentieth century. Geophys Res Lett 41:1639–1643. doi: 10.1002/2013GL059039

Wübbena G, Schmitz M, Boettcher G (2006) Near-field Effects on GNSS Sites: Analysis using Absolute Robot Calibrations and Procedures to Determine Correction. IGS Workshop 2010. ESOC/ESA, Darmstadt

Yu N, Kohel JM, Kellogg JR, Maleki L (2006) Development of an atom-interferometer gravity gradiometer for gravity measurement from space. Appl Phys B 84:647–652. doi: 10.1007/s00340-006-2376-x

Yunck T (1993) Coping with the Atmosphere and Ionosphere in Precise Satellite and Ground Positioning. In: Environmental Effects on Spacecraft Positioning and Trajectories (ed A.V. Jones). AGU, Washington, DC, USA

Yunck T, Bertiger W, Wu S, et al (1994) First assessment of GPS-based reduced dynamic orbit determination on TOPEX/POSEIDON. Geophysical Ressearch Letters 21:541–544

Zehentner N, Mayer-Gürr T (2015) Mitigation of ionospheric scintillation effects in kinematic LEO precise orbit determination. In: Geophysical Research Abstracts. Vienna, pp EGU2015-10477–1

Ziebart M, Sibthorpe A, Cross P, et al (2007) Cracking the GPS-SLR Orbit Anomaly. In: Proc. ION-GNSS-2007. pp 2033–2038