Satellite to Satellite Tracking in the Space-wise Approach

Von der Fakultät Luft- und Raumfahrttechnik und Geodäsie
der Universität Stuttgart zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

Vorgelegt von
Mohammad A. Sharifi
aus Abhar, Iran

Hauptberichter: Prof. Dr. sc. techn. Wolfgang Keller

Mitberichter: Prof. Dr.-Ing. Bernhard Heck
Prof. Dr.-Ing. Nico Sneeuw

Tag der mündlichen Prüfung: 13.07.2006

Geodätisches Institut der Universität Stuttgart

2006

Abstract

The launch of the CHAMP mission in 2000 has renewed interest in the recovery of the geopotential field from satellite observations which has been a challenging research issue for decades. It was the first dedicated gravity field mission which was followed by the GRACE spacecrafts.
In the GRACE mission, the High-Low (HL-) and the Low-Low Satellite-to-SatelliteTracking (LL-SST) observations are combined and the resultant observables are expressed in terms of the gravity gradient at the barycenter of two satellites. Each observation at its respective evaluation point can be written in terms of the spherical harmonic coefficients. Consequently, the observations are a sequence of discrete time series which are mathematically related to the unknown coefficients via the corresponding position of the satellites at the evaluation epoch. In this approach, which is called *time-wise approach*, the determination of unknown coefficients becomes possible after plugging the observations into the mathematical model. Fulfilling the sampling theorem, however, leads to a huge linear system of equations with a large number of unknowns.
As an alternative, one can employ the *semi-analytical approach* which is derived from the time-wise approach by imposing some approximations. Observations are still considered as discrete time series on an ideal geometry with a constant radius and/or constant inclination. The coefficients are reordered and then computed via the lumped coefficients or using 2D FFT.
Another alternative is the *space-wise approach* in which the observations are mapped on a specific grid on the mean orbital sphere. In this approach, the observation values are predicted on the grid points and the coefficients are derived by implementation of the *global spherical harmonic analysis* on the gridded observations. Compared to the time-wise approach, the linear system of equations are split into smaller systems which can be solved very easily in ordinary PCs.
In this thesis, the LL-SST problem is formulated both in the semi-analytical as well as the space-wise approach. The space-wise approach is then numerically implemented.
Despite the spirit of modern geodesy to avoid *reduction*, the reduction of observations is required both in the semi-analytical and the space-wise approaches. Different formulations are used for down- or up-ward continuation of observations on the reference geometry. Optimality of the basis functions and their respective parameters is carefully treated by means of the *Genetic Algorithms* (GA).
Optimizing the approximation methods is carefully investigated using the *genetic algorithms*. The idea of *one-leave-out* method or the so-called *residual bootstrap* approach is successfully used in the definition of the object functions. Compared to the classical error criterion, the modified object function results in a better solution.
In order to reduce the linearization and the reduction error, the residual gravity field is recovered. In this study, an adaptive reference orbit is used. Furthermore, the determination of the best fitting reference orbit is expressed as a least squares and an optimization problem.
Indeed, mathematical formulation of the gradiometry approach of the cubic order in terms of Taylor series is derived. The contribution of each individual term to the formulation is analyzed and the formulation is simplified accordingly.
The relative velocity vector is combined with the high-accuracy ranging observations both in the acceleration difference and the gradiometry approaches. Since it is not directly observed it should numerically be derived from GPS observations by means of numerical differentiation. In this regard, a few differentiation algorithms are studied for deriving the relative intersatellite velocity vector.
Furthermore, the recovery of the residual field and computing the relative velocity using the reference field are alternatively utilized to bypass the numerical differentiation. Compared to the numerical differentiation, the alternative methods yield more accurate solution.
The ranging system observations are more accurate than the GPS measurements. Some condition equations are derived for adjusting the low-accuracy observations using the LL-SST measurements. It only improves the cross-track and radial components of the relative position vector whereas the along-track component of the relative velocity benefits from imposing the constraints.
Finally, the previously derived formulation is used for recovery of the residual field. Two different iterative approaches are employed for determination of the residual gravity field using the GRACE *non-invariant observable*. To sum up, the gradiometry approach using a satellite pair is successfully implemented for the recovery of the residual gravity field in the space-wise approach.

Zusammenfassung

Der Start der Kleinsatellitenmission CHAMP im Juli 2000 hat das Interesse an der Bestimmung des Gravitationsfeldes der Erde aus Satellitenbeobachtungen , was bereits seit Jahrzehnten eine Herausforderung für die Wissenschaft darstellt, neu belebt. Diese Pionier-Satellitenmission zur Bestimmung des Schwerefeldes fand ihre Nachfolge in der GRACE Mission. Die GRACE Mission kombiniert die HL- and LL-SST Beobachtungen, um daraus den Schweregradient im Schwerpunkt der Doppelsatellitenkonfiguration abzuleiten. Jede dieser Schweregeradient-Beobachtungen kann als Kugelfunktionsreihe im zugehörigen Auswertepunkt dargestellt werden. Folglich kann die beobachtete Zeitreihe der Schweregradienten mithilfe der zum Beobachtungszeitpunkt bekannten Position des Auswertepunktes mit den unbekannten Koeffizienten der Kugelfunktionsreihe in Verbindung gebracht werden. Diese Vorgehensweise wird gewöhnlich time-wise approach genannt und ermöglicht die Bestimmung der Kugelfunktionskoeffizienten durch Zusammenführung aller Beobachtungen in ein gemeinsames mathematisches Modell. Dies führt bei Erfülltsein des Abttasttheorems zu einem riesigen System linearer Gleichungen mit einer sehr großen Anzahl von Unbekannten.
Alternativ kann auch der sogenannte semi-analytic approach angewandt werden, der unter Hinzunahme von Nebenbedingungen vom time-wise approach abgeleitet ist. In dieser Vorgehensweise werden die Beobachtungen weiterhin als Zeitreihe aufgefaßt, den Orten der Beobachtungen wird aber eine ideale Geometrie mit konstantem Bahnradius und/oder konstanter Bahnneigung unterstellt. Die unbekannten Kugelfunktionskoeffizienten können dann durch 2D FFT aus den sogenannten lumped-coefficients berechnet werden.
Eine weitere Lösunsmöglichkeit ist der sogenannte space-wise approach, bei dem die Beobachtungen auf ein reguläres Gitter einer den mittleren Orbit repräsentierenden Kugel projiziert werden. Mittels globaler Kugelfunktionsanalyse können die unbekannten Kugelfunktionskoeffizienten aus den Gitterwerten abgeleitet werden. Im Vergleich zum time-wise approach zerfällt dabei das riesige Gleichungssystem in mehrere kleinere Systeme, die leicht auf einem Standard-PC gelöst werden können.
In der vorgelegten Dissertation wird das LL-SST Problem sowohl im semi-analytic als auch im space-wise approach formuliert. Der space-wise approach wird dann numerisch implementiert. Entgegen der Tendenz der modernen Geodäsie, Reduktionen zu vermeiden, müssen sowohl im semi-analytic als auch im space-wise approach Reduktionen durchgeführt werden. Für diese notwendigen harmonischen Fortsetzungen der Beobachtungen auf eine Referenzfläche werden verschieden Formulierungen untersucht. Die sorgfältige Anwendung sogenannter Genetischer Algorithmen (GA) ermöglichte eine optimale Wahl der Basisfunktionen und der Parameter für diese harmonischen Fortsetzungen.
Unter Zuhilfenahme Genetischer Algorithmen werden die Approximationsverfahren sorgfältig optimiert. Für die Definition der Zielfunktion werden sowohl die one-leave-out Technik als auch die sogenannte residual-bootstrap Methode erfolgreich eingesetzt. Im Vergleich zum klassischen Fehlerkriterium führen die modifizierte Zielfunktion zu einer besseren Lösung.
Zur Reduktion des Linearisierungs- und Forsetzungsfehlers wird in der Dissertation statt des tatsächlichen Gravitationsfeldes nur das Residualfeld gegenüber eines Referenzfeldes betrachtet. Das Referenzfeld und der darauf basierende Referenzorbit werden adaptiv bestimmt. Die Bestimmung des bestangepassten Referenzorbits wird als Quadratmittel- und als Optimierungsproblem ausgedrückt. Eine mathematische Formulierung des Gravitationsgradienten bis einschließlich Glieder dritter Ordnung wird abgeleitet. Der Beitrag jedes einzelnen Terms in diese Formulierung wird analysiert und als Resultat dieser Analyse wird die Schweregeradientenformulierung dementsprechend vereinfacht.
Sowohl in der Beschleunigungsdifferenzen- als auch in der Gradiometriemethode wird der Geschwindigkeitsvektor mit Abstandsmessungen kombiniert. Da der Geschwindigkeitsvektor nicht direkt gemessen werden kann, muss er durch numerische Differentation aus den GPS Positionen gewonnen werden. Zu diesem Zweck werden verschiedene Differentationalgorithmen studiert und der geeignetste ausgewählt. Um die numerische Differentation zu umgehen, wird alternativ ein Referenzfeld eingeführt und bezüglich dieses Feldes die Relativgeschwindigkeiten bestimmt, um daraus schliesslich das Residualfeld zu berechnen. Im Vergleich zur numerischen Differentation erzielt diese indirekte Methode genauere Ergebnisse.
Die gemessenen Abstandsänderungen sind genauer als die GPS Messungen. Zum Angleich der weniger genauen GPS Beobachtungen an die genaueren SST Beobachtungen werden Bedingungsgleichungen abgeleitet. Es zeigt sich allerdings, dass damit nur die cross-track und die radiale Komponente des Geschwindigkeitsvektors verbessert werden.

Schließlich werden alle vorher abgeleiteten Zwischenergebnisse zur Bestimmung des Residualfeldes zusammengeführt. Es werden zwei verschiedene iterative Verfahren zur Bestimmung des Residualfeldes aus nicht-invarianten GRACE Beobachtungen benutzt.
Zusammenfassend kann gesagt werden, daß in der Dissertation Schweregradieometrie mithilfe eines Satellitenpaares für die Bestimmung des Residualfeldes im space-wise approach erfolgreich eingesetzt wurde.

Acknowledgements

First and foremost, I would like to thank God for life itself. All that I have is due to his grace and I would glorify him for all his blessings.

I would like to express my sincere gratitude to the Ministry of Science, Research and Technology of Iran for providing me the financial support to further my education to this level, without whose assistance this opportunity would not have been realized.

My sincere appreciation is conveyed to my distinguished supervisor *Prof. Dr. sc. techn. Wolfgang Keller* for his continuous guidance, discussions and helpful advice in the PhD. program. Without his encouragement and constant guidance, I could not have finished this dissertation. He was always there to meet and talk about my ideas,to proofread and mark up my papers and chapters, and to ask me good questions to help me think through my problems both analytical and computational.

I would also like to thank *Prof. Dr.-Ing.habil. Dr.tech.h.c.mult Dr.-Ing.E.h.mult Erik W. Grafarend* for his thoughtful suggestions and advice throughout my studies.

A special thanks goes to *Prof. Dr.-Ing. Nico Sneeuw*. Nico has been a friend and mentor. He taught me how to write academic papers, made me a better programmer, and brought out the good ideas in me.

I thank my family: my parents, for giving me life in the first place, for educating me, for unconditional support and encouragement to pursue my interests. My wife, *Parvin Bolouri*, M.D., and my daughter, *Yeganeh*, who got along with a student's life for many years, for their endless support and encouragement. I could not have done it without them.

I wish to thank *M.j. Bolouri* for his amazing patience and time dedicated in correcting my grammar and stylistic errors. I owe a great deal to him.

Last, but not least, I would like to emphasize that many other people helped me with this research directly and indirectly. Unfortunately, I cannot thank everyone explicitly here. Nevertheless, I hereby would like to extend my sincere gratitude to all of them.

Contents

Chapter 1

Introduction

The determination of the gravity field of the Earth is one of the main goals of geodesy. The determination of the Earth flattening by means of artificial satellites, in this concern, was the first considerable achievement. It was then followed by determination of higher degree spherical harmonics coefficients. Until the introduction of satellite altimetry, the Earth's gravity field models were derived only from ground based satellite tracking and surface gravimetry data (Lerch and Wagner, 1981). The finer details were provided by terrestrial gravimetry data whereas the longer wavelength features were supplied by the satellite tracking measurements. On the other hand, from the satellite tracking measurements only, the low-degree coefficients of the gravity field up to degree and order 20 plus some additional zonal harmonics could be recovered (ESA, 1987). Indeed, the ground based stations were not distributed uniformly around the globe. Therefore, the error spectrum of the resultant gravity models were inhomogeneous. Nonuniform distribution of the tracking stations and the lack of global coverage were the shortcomings of the ground based tracking method for the determination of a high accuracy and resolution gravity model. These limitations showed that additional measurement techniques for the recovery of the global geopotential field had to be employed.
On May 14, 1973, the first altimeter was flown on Skylab. Altimetry observations were useful for improving the gravity field knowledge of the ocean areas. Nevertheless, combining the observations with satellite tracking measurements and surface gravimetry data were inadequate for recovery of the gravity field with the required resolution and accuracy. Collecting more data on the continental areas was necessary for recovery of the field to the accuracy levels that were desired. Gathering more information on the continental regions were in particular interest since the knowledge of the gravity field on the continental part was rather poor.
Intensive studies on suitable space techniques for determination of the gravity field have been started since 1970. Two adequate methods have been recognized for being implemented on the gravity field dedicated missions:

- satellite to satellite tracking (SST)
- satellite gravity gradiometry (SGG)

SST was initially used for mapping the near side of the moon by considering the moon as a satellite of the Earth (Muller and Sjogren, 1968). It was then employed for mapping the Earth's gravity field. It was started by different satellites in the HL-mode. However, the results were not promising due to the high altitude of the lower satellite.
The LL-mode, a configuration with two orbiters flying nearly in an identical orbit at the lowest possible altitude, was in particular interest to groups engaged in SST. GRAVSAT and GRM were two American missions where designed to map the Earth's gravity field (White, 1987). Neither GRAVSAT nor GRM missions was realized because of technical difficulties and budget constraints. GAMES mission, an American-French mission in the LL-mode, was the next experiment which never materialized.
The launch of the CHAMP mission in 2000 has renewed interest in the recovery of the geopotential field from satellite observations. It was the first spacecraft in a series of gravity field dedicated missions, to be followed by the GRACE mission. Providing global and high-resolution estimates of the Earth's gravity field and its temporal variations with unprecedented accuracy is the primary science objective

of the GRACE mission. Observing inter-satellite range and range rate by the K-*band Ranging System (*KBR*)* with the highest possible accuracy ($1\,\mu$m/s) is the superiority of the GRACE over the CHAMP mission.
In the GRACE mission, the HL- and LL-SST observations are combined. The resultant observables can be mathematically expressed as the gravitational acceleration differences between the satellite pairs or alternatively as the gravity gradient at the barycenter of two satellites. In the first approach, the mathematical model is a two-point first order spatial derivative of the gravity field whereas the latter case is a one-point second-order derivative of the field.
Despite the conceptual differences between the two formulations, each observation at its respective evaluation point can be written in terms of the spherical harmonic coefficients. Consequently, the observations are a time series which is mathematically related to the unknown coefficients via the corresponding position of the satellites at the evaluation epoch. In this approach or the so-called *time-wise approach*, the determination of unknown coefficients becomes possible after plugging the observations into the mathematical model. Furthermore, the stochastic characteristics of the observations and its propagation into the recovered field can be easily studied. Fulfilling the sampling theorem, however, leads to a huge linear system of equations with an enormous number of unknowns and observations.
As an alternative, one can employ the *semi-analytical approach* which is derived from the time-wise approach by imposing some constraints. Observations are still considered as discrete time series on an ideal geometry with constant radius and/or constant inclination. The coefficients are then computed using the lumped coefficients or 2D FFT.
Another alternative is the *space-wise approach* in which the observations are mapped on a specific grid on the mean orbital sphere. In this approach, the observation values are predicted on the grid points and the coefficients are derived by implementation of the *global spherical harmonic analysis* on the gridded values. Compared to the time-wise approach, the linear system of equations is split into smaller systems which can be solved very easily in ordinary PCs.
Having access to high-performance computers with parallel processors makes the implementation of the time-wise approach possible even up to very high-degree and order. Nevertheless, the semi-analytical and the space-wise approaches are two viable alternatives which results in the same results even with very cheap hardware and less computational skills.
Compared to the time-wise approach, data reduction is the only additional computational step in the semi-analytical and the space-wise approach. We expect an identical solution if this additional step is carefully treated.
The goal of this study is to demonstrate the formulation of the LL-SST problem in the space-wise approach. For completeness, the problem is also formulated in the semi-analytical approach. Indeed, the main concern is to reduce the observations via an optimal procedure.
The dissertation starts by formulation of the LL-SST problem in the brute-force approach in which each observation of the acceleration difference or the gradiometry type is expressed as a linear observation equation. Solving a huge linear system of equations with full normal matrix of size $(N+1)^2 \times (N+1)^2$ is a formidable obstacle on the recovery process. From memory point of view, solution of such a huge linear system is infeasible for approximately $N > 100$ in normal PCs.
Employing supercomputers with a large number of processors is the hardware solution of the problem. On the other hand, one can tackle the problem by imposing some constraints and exploiting the unseen properties of the surface spherical harmonic basis functions. This alternative approach, called *semi-analytical approach*, is comprehensively studied in Chapter 2.
Chapter 2 will then continue with the *space-wise approach*, in which is essential to reduce observations on a specific grid defined on a sphere. The *mean orbital sphere* (MOS) is used as the surface where the observations are reduced.
Compared to the time-wise approach, mapping onto a specified geometry is an additional computational step which should carefully be treated. Therefore, we will closely investigate the observation reduction in Chapter 3.
In order to reduce the truncation and the reduction error, the true observations are replaced by residual gravity observations. The corresponding reference observations are derived by introduction of the respective reference orbit. For minimizing the deviation of the true and the reference orbit, implementation of an adaptive reference orbit as well as the best fitting reference orbit in a least squares sense are investigated in Chapter 4. Furthermore, the problem is formulated as an optimization problem

using *genetic algorithms*.
Different methods for the global spherical harmonic analysis are studied in Chapter 5. They are numerically compared and the most accurate scheme is introduced. It is shown that one can expect perfect recovery if the Gauss grid is employed.
Chapter 6 is devoted to the mathematical formulation of the gradiometry using a satellite pair. The gradiometry observation equation of the first- and cubic-order are formulated. Contribution of each individual component is studied and the formulations are simplified according the insignificance of the minor terms.
Eventually, the simplified gradiometry observation equation of cubic order is employed in Chapter 7 for recovery of the residual field. Due to its non-invariance the GRACE observable has to be analyzed iteratively. Two iterative approaches are implemented and the achieved results are compared. Finally, this study is finished with conclusions and recommendations which are summarized in Chapter 8. The achievements of this study are summarized and the topics for the future studies are listed.

Chapter 2

Space-wise versus Time-wise Approach

With the introduction of new satellite techniques, such as Satellite-to-Satellite tracking (SST) and Satellite Gravity Gradiometry (SGG), a whole new world of investigation has opened up. Besides the existing questions in theory, a number of questions have arisen in numerical computation due to high resolution space missions. Therefore, treatment of the emerging linear system of equations with millions of observations and 10^4– 10^5 of unknown gravity field parameters is a highly demanding task.

Recovery of the gravity field can be formulated using the more-or-less developed *space-wise* and *time-wise* approaches. The space-wise method circumvents the problem, whereas the clever scheme called *semi-analytical approach* should be employed to alleviate the problem in time-wise approach.

In this chapter, the principles of both the time-wise and the space-wise as well as the brute-force and semi-analytical approaches are discussed and pros and cons of each of these strategies will be given.

2.1 Time-wise approach

In time-wise approach, observations are considered as a time-series along the orbit of the Low-Earth Orbiter (LEO). Each individual observation results in a linear equation. Consequently, the geopotential recovery based on direct solution called *brute-force technique* could be a real demanding task particularly for the high resolution geopotential models.

Orthogonality potential of the utilized basis functions, spherical harmonics can be easily exploited to circumvent the numerical problem. To this end, besides the representation of the geopotential in terms of the orbital elements, a few approximations should be imposed on the sequence of the observations. This method which is called *semi-analytical approach*, leads to block-diagonal systems whose solutions are comparatively simple compared to that of a brute-force method. However, observables being observed or projected on an ideal geometric configuration (circle, sphere or torus), constant height or inclination, repeat orbit, evenly-spaced data span are examples of the required conditions (Sneeuw, 2000).

Moreover, as will be seen in the following section, representation in terms of the Kepler elements is the best for the LL-SST observable.

2.1.1 Brute-force approach

The time-wise approach for spherical harmonic analysis of LL-SST data was introduced by Kaula (1983) and independently by Wagner (1983). It has been used by many others either for HL-SST or SGG afterwards (e.g. Rummel et al., 1993).

From the solution of the Laplace equation in the outer-space (outside the attracting masses), the gravitational potential of the Earth can be expressed as an infinite series of spherical harmonics

(Hofmann-Wellenhof and Moritz, 2005):

$$V(\mathbf{r}) \;=\; \frac{GM}{r}\sum_{n=0}^{\infty}\left(\frac{R_{\mathrm{E}}}{r}\right)^{n}\sum_{m=0}^{n}\left[\bar{C}_{nm}\bar{Y}^{\mathrm{c}}_{nm} \;+\; \bar{S}_{nm}\bar{Y}^{\mathrm{s}}_{nm}\right] \tag{2.1}$$

with $\mathbf{r} = (r,\, \theta,\, \lambda)$ the spherical coordinates of an evaluation point in the outer-space ($r > R_{\mathrm{E}}$), GM the gravitational constant times the Earth's mass, R_{E} the mean radius of the Earth, $(\bar{C}_{nm}, \bar{S}_{nm})$ the infinite set of coefficients of the fully normalized surface spherical harmonics $\bar{Y}^{\mathrm{c}}_{nm}$, $\bar{Y}^{\mathrm{s}}_{nm}$. In practice, the series expansion has to be truncated at a maximum degree N, i.e., the upper limit ∞ of the index n is replaced by a finite number N. Moreover, the introduction of a known set of the reference field respective coefficients $(\bar{C}^{(0)}_{nm}, \bar{S}^{(0)}_{nm})$ results in a corresponding expression of the *incremental potential*:

$$\delta V(\mathbf{r}) \;=\; \frac{GM}{r}\sum_{n=0}^{N}\left(\frac{R_{\mathrm{E}}}{r}\right)^{n}\sum_{m=0}^{n}\left[\delta\bar{C}_{nm}\bar{Y}^{\mathrm{c}}_{nm} \;+\; \delta\bar{S}_{nm}\bar{Y}^{\mathrm{s}}_{nm}\right] \tag{2.2}$$

where, $\delta\bar{C}_{nm} = \bar{C}_{nm} - \bar{C}^{(0)}_{nm}$ and $\delta\bar{S}_{nm} = \bar{S}_{nm} - \bar{S}^{(0)}_{nm}$ are the unknowns that should be estimated as the solution of the Earth's gravity recovery problem.
In general, every observation is assumed as a functional of the incremental potential

$$l_p \;=\; \mathcal{L}\delta V(P). \tag{2.3}$$

Therefore, observations for the time-wise approach form a time series of the functionals given along the orbit. Depending on the space sensor, the functional could be *scalar*, *vectorial* or *tensorial*. For instance, the HL-SST observations are formulated as the vectorial functional whereas the LL-SST appears as a scalar one. In SGG case, the problem formulation demands tensorial tools.
Assuming a known disturbing potential for the high-satellite leads to the following functional in HL-SST configuration:

$$\mathcal{L}^{\mathrm{HL}} \;=\; -\,\frac{1}{r}\,\mathbf{e}_{x'}\frac{\partial}{\partial\theta} \;-\; \frac{1}{r\sin\theta}\,\mathbf{e}_{y'}\frac{\partial}{\partial\lambda} + \,\mathbf{e}_{z'}\frac{\partial}{\partial r}. \tag{2.4}$$

where, the set of $(\mathbf{e}_{x'}, \mathbf{e}_{y'}, \mathbf{e}_{z'})$ unit vectors define the *local north-oriented reference frame* (LNRF) at the evaluation point (Keller and Sharifi, 2005). This local system whose origin is a terrestrial or space point is oriented with the z'-axis radially outwards, the x'-axis directed northward and the y'-axis directed westward. Freeden et al. (1999) showed that the problem of developing the geopotential in the outer-space from the given HL-SST information is overdetermined. Therefore, for ease of computation the vectorial and tensorial representation is recast into scalar representation without loss of generality. For instance, the vectorial problem can be decomposed into two scalar components, namely *radial* and *tangential*. The radial component, which has already been widely used (e.g. Schreiner, 1994; Freeden, 1999; Hesse and Gutting, 2003), can be simply derived by mapping the observation vector between the Low-Earth-Orbiting and the high-altitude satellites along the radial direction:

$$\mathcal{L}^{\mathrm{HL}}_{\mathrm{radial}} \;=\; \frac{\partial}{\partial r}. \tag{2.5}$$

Consequently, the HL-SST observation is expressed in $\delta\bar{C}_{nm}$ and $\delta\bar{S}_{nm}$ as:

$$l^{\mathrm{HL}} \;=\; \frac{GM}{r}\sum_{n=0}^{N}\left(\frac{R_{\mathrm{E}}}{r}\right)^{n}\sum_{m=0}^{n} h^{\mathrm{radial}}_{n}(r)\left[\delta\bar{C}_{nm}\bar{Y}^{\mathrm{c}}_{nm} \;+\; \delta\bar{S}_{nm}\bar{Y}^{\mathrm{s}}_{nm}\right]. \tag{2.6}$$

in which, for a circular orbit $h^{\mathrm{radial}}_{n} \;=\; -\frac{n+1}{r}$ can be considered as the *spectral transfer coefficients* which give the spectral link between observables and unknowns (Sneeuw, 2003). As we will see later, they play the key role in the semi-analytical as well as the space-wise approaches. Using the geopotential representation in terms of geocentric coordinates is convenient for brute-force modelling of some space-born data but not for most of them. For instance, the representation is well-suited for the radial component of the HL-SST while it would be rather complicated for the tangential component. For the LL-SST as well as the SGG observations, direct modelling gets even more complicated.
In order to compare formulation complexity of different observables in terms of (r, θ, λ) with those

in the semi-analytical approach, we formulate the tangential and the LL-SST observations in both approaches. For compactness, we express partial derivatives in the following symbolic manner, see Rummel et al. (1993):

$$\delta V^{(a)} = \frac{GM}{r}\sum_{n=0}^{N} h_n^{(a)}\left(\frac{R_{\mathrm{E}}}{r}\right)^n \sum_{m=0}^{n}\left[\delta\bar{C}_{nm}^{(a)}\bar{Y}_{nm}^{c^{(a)}} + \delta\bar{S}_{nm}^{(a)}\bar{Y}_{nm}^{s^{(a)}}\right]. \tag{2.7}$$

where, the a is a symbol notation for expressing one of the partial derivatives which runs over r, θ and λ. Analogously, the symbol ab is used for the second derivatives:

$$\delta V^{(ab)} = \frac{GM}{r}\sum_{n=0}^{N} h_n^{(ab)}\left(\frac{R_{\mathrm{E}}}{r}\right)^n \sum_{m=0}^{n}\left[\delta\bar{C}_{nm}^{(ab)}\bar{Y}_{nm}^{c^{(ab)}} + \delta\bar{S}_{nm}^{(ab)}\bar{Y}_{nm}^{s^{(ab)}}\right]. \tag{2.8}$$

where, ab again a symbol running over rr, $r\theta$, $r\lambda$, $\theta\theta$, $\theta\lambda$ and $\lambda\lambda$. Table (2.1) shows a complete list of the transfer coefficients and the respective representation of the basis functions for the first and the second partial derivatives. Noting that $\bar{Y}_{nm}^{c'} = \frac{d\bar{Y}_{nm}^{c}}{d\theta}$ and $\bar{Y}_{nm}^{c''} = \frac{d^2\bar{Y}_{nm}^{c}}{d\theta^2}$.

Table 2.1: Potential derivatives

ab	$h_n^{(ab)}$	$\delta\bar{C}_{nm}^{(ab)}\bar{Y}_{nm}^{c^{(ab)}}$	$\delta\bar{S}_{nm}^{(ab)}\bar{Y}_{nm}^{s^{(ab)}}$
r	$-\frac{n+1}{r}$	$\delta\bar{C}_{nm}\bar{Y}_{nm}^{c}$	$\delta\bar{S}_{nm}\bar{Y}_{nm}^{s}$
θ	1	$\delta\bar{C}_{nm}\bar{Y}_{nm}^{c'}$	$\delta\bar{S}_{nm}\bar{Y}_{nm}^{s'}$
λ	1	$m\delta\bar{S}_{nm}\bar{Y}_{nm}^{c}$	$-m\delta\bar{C}_{nm}\bar{Y}_{nm}^{s}$
rr	$\frac{(n+1)(n+2)}{r^2}$	$\delta\bar{C}_{nm}\bar{Y}_{nm}^{c}$	$\delta\bar{S}_{nm}\bar{Y}_{nm}^{s}$
$r\theta$	$-\frac{n+1}{r}$	$\delta\bar{C}_{nm}\bar{Y}_{nm}^{c'}$	$\delta\bar{S}_{nm}\bar{Y}_{nm}^{s'}$
$r\lambda$	$-\frac{n+1}{r}$	$m\delta\bar{S}_{nm}\bar{Y}_{nm}^{c}$	$-m\delta\bar{C}_{nm}\bar{Y}_{nm}^{s}$
$\theta\theta$	1	$\delta\bar{C}_{nm}\bar{Y}_{nm}^{c''}$	$\delta\bar{S}_{nm}\bar{Y}_{nm}^{s''}$
$\theta\lambda$	1	$-m\delta\bar{C}_{nm}\bar{Y}_{nm}^{s'}$	$m\delta\bar{S}_{nm}\bar{Y}_{nm}^{c'}$
$\lambda\lambda$	1	$-m^2\delta\bar{C}_{nm}\bar{Y}_{nm}^{c}$	$-m^2\delta\bar{S}_{nm}\bar{Y}_{nm}^{s}$

As the first example, consider the tangential component of the HL-mode:

$$\boldsymbol{\mathcal{L}}_{\text{tangential}}^{\mathrm{HL}} = -\frac{1}{r}\,\mathbf{e}_{x'}\frac{\partial}{\partial\theta} - \frac{1}{r\sin\theta}\,\mathbf{e}_{y'}\frac{\partial}{\partial\lambda} \triangleq \mathcal{L}_{\theta}^{\mathrm{HL}}\,\mathbf{e}_{x'} + \mathcal{L}_{\lambda}^{\mathrm{HL}}\,\mathbf{e}_{y'}. \tag{2.9}$$

with $\boldsymbol{\mathcal{L}}_{\text{tangential}}^{\mathrm{HL}}$ a vectorial and $\mathcal{L}_{\theta}^{\mathrm{HL}}$, $\mathcal{L}_{\lambda}^{\mathrm{HL}}$ scalar functionals. Substituting the corresponding expressions into Eq. (2.9) results in:

$$\boldsymbol{\mathcal{L}}_{\text{tangential}}^{\mathrm{HL}} = -\frac{1}{r}\delta V^{(\theta)}\,\mathbf{e}_{x'} - \frac{1}{r\sin\theta}\delta V^{(\lambda)}\,\mathbf{e}_{y'}. \tag{2.10}$$

It can easily be realized that mapping the tangential component onto the incremental potential cannot be performed using transfer function.

As mentioned earlier, the HL-SST problem is an overdetermined one and can be split into two scalar determined formulations. Consequently, the radial component as the invariant operator can be utilized for the incremental potential formulation. The case of LL-SST is a scalar determined problem whose functional is typically non-invariant. For instance, in the gravitational acceleration difference approach, it reads (Keller and Sharifi, 2005):

$$\mathcal{L}^{\mathrm{LL}}\delta V = \underbrace{\alpha_2\delta V^{(r)} + \frac{\beta_2}{r}\delta V^{(\theta)} + \frac{\gamma_2}{r\sin\theta}\delta V^{(\lambda)}}_{2} - \underbrace{\alpha_1\delta V^{(r)} + \frac{\beta_1}{r}\delta V^{(\theta)} + \frac{\gamma_1}{r\sin\theta}\delta V^{(\lambda)}}_{1} \tag{2.11}$$

with $\mathbf{e}_i = (\alpha_i, \beta_i, \gamma_i)$ the line of sight (LOS) unit vector at the respective LNRF defined at the location

of the front satellite 2, and the rear satellite 1. Substituting δv^r, δv^θ and δv^λ from Table (2.1) results in a complicated formulation. Like the tangential component of the HL-SST, the observable cannot be mapped onto δV just by use of the transfer coefficient.
A viable alternative to the acceleration difference formulation is the gradiometry approach whose linear representation is (Keller and Sharifi, 2005):

$$\mathcal{L}^{\mathrm{SG}}V \;=\; [\mathbf{e}_{\mathrm{B}} \otimes \mathbf{e}_{\mathrm{B}}]^T \left[\; \delta V^{(r)}\mathbf{e}_{z'} + \delta V^{(\theta)}\mathbf{e}_{x'} + \delta V^{(\lambda)}\mathbf{e}_{y'} \;\otimes\; \delta V^{(r)}\mathbf{e}_{z'} + \delta V^{(\theta)}\mathbf{e}_{x'} + \delta V^{(\lambda)}\mathbf{e}_{y'} \;\right]. \tag{2.12}$$

where, $\mathbf{e}_{\mathrm{B}}$ is the LOS unit vector at the barycenter of two LEO satellites. Similar to the acceleration difference approach, the observable is non-invariant and cannot be mapped just by use of the transfer coefficients. Nevertheless, the gradiometry formulation converts a two-point first-order problem to a one-point second-order one. Moreover, the new observable reflects the gravity field curvature along the LOS at the evaluation point. However, as it will be seen in chapter (6), the linear approximation is inaccurate and should be modified in order to achieve a reasonable accuracy.
Therefore, representation of the non-invariant functions like the tangential component in LNRF is cumbersome. However, as will be seen later in this section, the geopotential representation of non-invariant functional in the orbital elements is more convenient. This representation which leads to the semi-analytical approach will be discussed in the following section.
Nevertheless, representation of the gravitational potential in the spherical coordinates and the assumption of the observations as being a time series both in the HL- and LL-SST leads to a linear system of equations (e.g. Reubelt et al., 2003; Keller and Sharifi, 2005). The system can then be solved either by direct or iterative methods.

2.1.2 Semi-analytical approach

Expressing the geopotential as a function of the orbital elements $a, e, i, \Omega, \omega, M$ (*Kepler elements*) alleviates the gradiometry formulation. Therefore, Eq. (2.1) may be transformed to a series expansion in the orbital parameters (Kaula, 1966):

$$\delta V \;=\; \sum_{n=0}^{\infty}\sum_{m=0}^{n} \delta V_{nm} \tag{2.13}$$

$$\delta V_{nm} \;=\; \frac{GM}{r^{n+1}}\, R^n \sum_{p=0}^{n} \bar{F}_{nmp}(i) \sum_{q=-\infty}^{\infty} G_{npq}(e)\, S_{nmpq}(\omega, M, \Omega, \Theta) \tag{2.14}$$

with $\bar{F}_{nmp}(i)$ normalized inclination function, $G_{npq}(e)$ eccentricity function and

$$S_{nmpq} \;=\; \alpha_{nm}\, \cos\psi_{nmpq} \;+\; \beta_{nm}\, \sin\psi_{nmpq}. \tag{2.15}$$

Where,

$$\alpha_{nm} \;=\; \begin{cases} \delta\bar{C}_{nm}, & n-m \quad \text{even} \\ -\delta\bar{S}_{nm}, & n-m \quad \text{odd} \end{cases} \qquad \beta_{nm} \;=\; \begin{cases} \delta\bar{S}_{nm}, & n-m \quad \text{even} \\ \delta\bar{C}_{nm}, & n-m \quad \text{odd} \end{cases} \tag{2.16}$$

and

$$\psi_{nmpq} \;=\; (n-2p)\omega \;+\; (n-2p+q)M \;+\; m(\Omega-\Theta), \qquad \Theta \;=\; \mathrm{GAST} \tag{2.17}$$

More details on the $\bar{F}_{nmp}(i)$ and $G_{npq}(e)$ can be found in Kaula (1966). By introducing a nominal circular orbit ($e = 0$) with a constant inclination ($i = I$) and constant radius ($r = a$) the eccentricity function will drop out[1]:

$$\delta V \;=\; \frac{GM}{R}\sum_{n=0}^{N}\left(\frac{R}{r}\right)^{n+1} \sum_{m=0}^{n}\sum_{p=0}^{n} \bar{F}_{nmp}(I)\left[\bar{\alpha}_{nm}\cos\psi_{nmp} + \bar{\beta}_{nm}\sin\psi_{nmp}\right], \tag{2.18}$$

[1]Since the radial partial derivative will be involved in the problem formulation, we still retain the radius as a parameter in formulation.

where, ψ_{nmp} (Eq. 2.17), can be simplified as:

$$\psi_{nmp} = (n-2p)\omega^{\mathrm{o}} + m\omega_{\mathrm{e}} \tag{2.19}$$

with $\omega^{\mathrm{o}} = \omega + M$ and $\omega_{\mathrm{e}} = \Omega - \Theta$ where superscript o and subscript e are referring to *orbit* and the *Earth*, respectively (Schrama, 1989). The relationship between the *geocentric polar coordinates* (r, θ, λ) and the orbital elements appearing in Eq. (2.18), is depicted in Fig. (2.1).

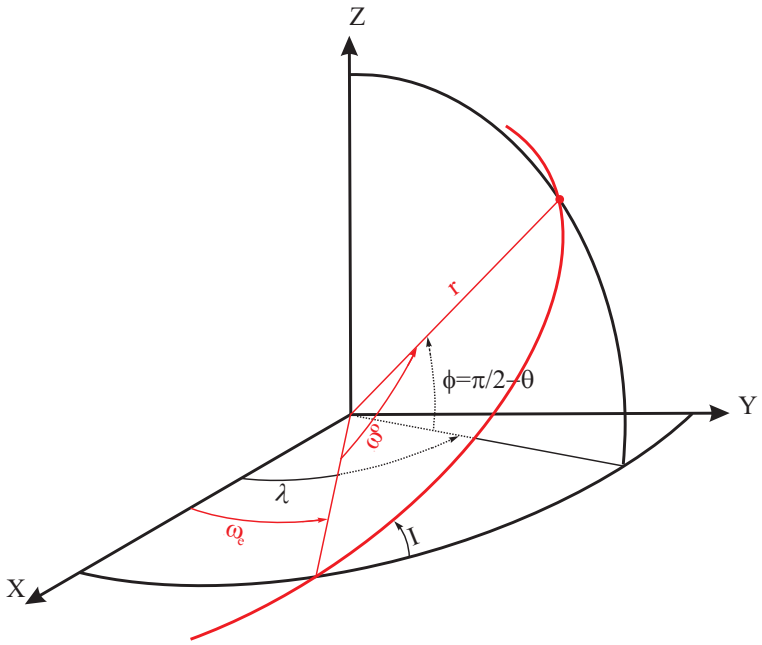


Figure 2.1: Geocentric coordinate system and the orbital elements

Representation of the gravitational potential in terms of the orbital elements by means of group theory, which has already been introduced in (Sneeuw, 1992), is more fundamental and instructive for satellite purposes. One can simply derive the same expression from Eq. (2.18) by introducing a new index k instead of p, with $k = n - 2p$. It is clear that n and k always have the same parity. Furthermore, the index k has a step-size 2 within the bounds of $\pm n$ (Koop, 1993). Therefore:

$$\delta V = \sum_{n=0}^{N} \sum_{m=0}^{n} \sum_{k=-n[2]}^{n} [A_{nmk}(r,I)\cos\psi_{km} + B_{nmk}(r,I)\sin\psi_{km}] \tag{2.20}$$

where,

$$\left.\begin{matrix} A_{nmk} \\ B_{nmk} \end{matrix}\right\} = H_{nmk}(r,I) \left\{\begin{matrix} \alpha_{nm} \\ \beta_{nm} \end{matrix}\right. ,$$

$$H_{nmk}(r,I) = u_n(r)\,\bar{F}^k_{nm}(I),$$

$$u_n = \frac{GM}{r}\left(\frac{R_{\mathrm{E}}}{r}\right)^n$$

and

$$\psi_{km} = k\omega^{\mathrm{o}} + m\omega_{\mathrm{e}}$$

To avoid confusion with the $\bar{F}_{nmp}$-notation, we have introduced a term $\bar{F}^k_{nm}$ where,

$$\bar{F}^k_{nm} = \bar{F}_{nm,\frac{n-k}{2}} = \bar{F}_{nmp}. \tag{2.21}$$

Normalized inclination functions $\bar{F}^k_{nm}$, can be computed in several ways. We will use the recurrence techniques given in Sneeuw (1991a) for computing the function and its partial derivatives. The recursive method, which is based on Emeljanov and Kanter (1989), has been explained in Appendix

A in more detail.
The new index k is more convenient from the computational point of view. To make it clear, let us reorder the summation sequences from n, m, k into m, k, n. Consequently, Eq. (2.20) is recast into:

$$\delta V(r, \omega_{\mathrm{e}}, \omega^{\mathrm{o}}, I) \;=\; \sum_{m=0}^{N} \sum_{k=-N}^{N} \sum_{n=n_1[2]}^{n_2} \left[A_{nmk}(r, I) \cos \psi_{km} \;+\; B_{nmk}(r, I) \sin \psi_{km}\right] \tag{2.22}$$

where,

$$n_1 = \begin{cases} \mid k \mid & \mid k \mid \geq m \\ m & \mid k \mid < m \;\; k - m \text{ even} \\ m+1 & \mid k \mid < m \;\; k - m \text{ odd} \end{cases} \qquad n_2 \;=\; \begin{cases} N & N - k \text{ even} \\ N-1 & N - k \text{ odd} \end{cases}$$

Carrying out the inner loop summation over n results in a double summation formula with the new coefficients $\widetilde{A}_{mk}$ and $\widetilde{B}_{mk}$ which are called *lumped coefficients*:

$$\delta V(r, \omega_{\mathrm{e}}, \omega^{\mathrm{o}}, I) \;=\; \sum_{m=0}^{N} \sum_{k=-N}^{N} \left[\widetilde{A}_{mk}(r, I) \cos \psi_{km} \;+\; \widetilde{B}_{mk}(r, I) \sin \psi_{km}\right]. \tag{2.23}$$

with $\widetilde{A}_{mk}(r, I) \;=\; \sum_n A_{nmk}$ and $\widetilde{B}_{mk}(r, I) \;=\; \sum_n B_{nmk}$. Eq. (2.23) is a Fourier series. One can derive the lumped coefficients by straightforward Fourier transform of Eq. (2.23) and then consider the achieved coefficients as the pseudo-observations. Of course it should be noted here that for full recovery of the geopotential coefficients, certain conditions should be satisfied (Rummel et al., 1993). Since the observables are functionals of the geopotential rather than on the potential itself, let us derive the first and the second partial derivatives of the geopotential as the constituents of satellites' observations. For convenience, we rewrite Eqs. (2.7) and (2.8) with respect to the orbital elements using Eq. (2.20):

$$\delta V^{(a)} \;=\; \sum_{n=0}^{N} \sum_{m=0}^{n} \sum_{k=-n_1[2]}^{n_2} \left[A^{(a)}_{nmk}(r, I) \cos \psi_{km} \;+\; B^{(a)}_{nmk}(r, I) \sin \psi_{km}\right], \tag{2.24}$$

and analogously,

$$\delta V^{(ab)} \;=\; \sum_{n=0}^{N} \sum_{m=0}^{n} \sum_{k=-n_1[2]}^{n_2} \left[A^{(ab)}_{nmk}(r, I) \cos \psi_{km} \;+\; B^{(ab)}_{nmk}(r, I) \sin \psi_{km}\right]. \tag{2.25}$$

with the symbolic notations a and $b \in \{r, \omega_{\mathrm{e}}, \omega^{\mathrm{o}}, I\}$. All the first and second partial derivatives are listed in Table (2.2).
In order to express the SST or SGG observables in terms of the new set of partial derivatives, we define *cartesian orbital coordinate system* (X', X', Z') as follows:

- X'-axis directed towards the ascending node,
- Y'-axis orthogonal to X' in the orbital plane,
- Z'-axis completes the set to a right-handed coordinate system.

The coordinate transformation between the orbital and geocentric coordinate systems is expressed as:

$$\mathbf{r} \;=\; \mathbf{R}_3(\omega_{\mathrm{e}})\, \mathbf{R}_1(I)\, \mathbf{r}'. \tag{2.26}$$

Similar to geocentric polar coordinates r, θ and λ, the respective curvilinear coordinates $r, \omega^{\mathrm{o}}, \theta'$, which is called *orbital polar coordinate system*, are defined. See Fig. (2.2). The relation between cartesian and curvilinear representation is:

$$\begin{array}{l} X' \;=\; r \cos \theta' \cos \omega^{\mathrm{o}} \\ Y' \;=\; r \cos \theta' \sin \omega^{\mathrm{o}} \\ Z' \;=\; r \sin \theta' \end{array} \tag{2.27}$$

Table 2.2: Potential derivatives with respect to the orbital elements

a or ab	$H^{(a)}_{nmk}$ or $H^{(ab)}_{nmk}$	$A^{(a)}_{nmk}$ or $A^{(ab)}_{nmk}$	$B^{(a)}_{nmk}$ or $B^{(ab)}_{nmk}$
r	$-\frac{n+1}{r}H_{nmk}$	$H^{(r)}_{nmk}\alpha_{nm}$	$H^{(r)}_{nmk}\beta_{nm}$
ω_{e}	mH_{nmk}	$H^{(\omega_{\mathrm{e}})}_{nmk}\beta_{nm}$	$-H^{(\omega_{\mathrm{e}})}_{nmk}\alpha_{nm}$
ω^{o}	kH_{nmk}	$H^{(\omega^{\mathrm{o}})}_{nmk}\beta_{nm}$	$-H^{(\omega^{\mathrm{o}})}_{nmk}\alpha_{nm}$
I	H'_{nmk}	$H^{(I)}_{nmk}\alpha_{nm}$	$H^{(I)}_{nmk}\beta_{nm}$
rr	$\frac{(n+1)(n+2)}{r^2}H_{nmk}$	$H^{(rr)}_{nmk}\alpha_{nm}$	$H^{(rr)}_{nmk}\beta_{nm}$
$r\omega_{\mathrm{e}}$	$-m\frac{n+1}{r}H_{nmk}$	$H^{(r\omega_{\mathrm{e}})}_{nmk}\beta_{nm}$	$-H^{(r\omega_{\mathrm{e}})}_{nmk}\alpha_{nm}$
$r\omega^{\mathrm{o}}$	$-k\frac{n+1}{r}H_{nmk}$	$H^{(r\omega^{\mathrm{o}})}_{nmk}\beta_{nm}$	$-H^{(r\omega^{\mathrm{o}})}_{nmk}\alpha_{nm}$
rI	$-\frac{n+1}{r}H'_{nmk}$	$H^{(rI)}_{nmk}\alpha_{nm}$	$H^{(rI)}_{nmk}\beta_{nm}$
$\omega_{\mathrm{e}}\omega_{\mathrm{e}}$	$-m^2H_{nmk}$	$H^{(\omega_{\mathrm{e}}\omega_{\mathrm{e}})}_{nmk}\alpha_{nm}$	$H^{(\omega_{\mathrm{e}}\omega_{\mathrm{e}})}_{nmk}\beta_{nm}$
$\omega_{\mathrm{e}}\omega^{\mathrm{o}}$	$-mkH_{nmk}$	$H^{(\omega_{\mathrm{e}}\omega^{\mathrm{o}})}_{nmk}\alpha_{nm}$	$H^{(\omega_{\mathrm{e}}\omega^{\mathrm{o}})}_{nmk}\beta_{nm}$
$\omega_o\omega_o$	$-k^2H_{nmk}$	$H^{(\omega^{\mathrm{o}}\omega^{\mathrm{o}})}_{nmk}\alpha_{nm}$	$H^{(\omega^{\mathrm{o}}\omega^{\mathrm{o}})}_{nmk}\beta_{nm}$
$\omega_o I$	kH'_{nmk}	$H^{(\omega_o I)}_{nmk}\beta_{nm}$	$-H^{(\omega_o I)}_{nmk}\alpha_{nm}$
II	H''_{nmk}	$H^{(II)}_{nmk}\alpha_{nm}$	$H^{(II)}_{nmk}\beta_{nm}$

Furthermore, we define the *local orbital coordinate Frame* (LORF) x, y, z, which is a well-suited frame for the LL-SST and SGG observation representation. This local system whose origin is a point on a satellite's orbit is oriented with the z-axis radially outwards, the x-axis directed along track and the y-axis cross track such that it is a right-handed system. One can simply prove that:

$$\nabla\delta V = \begin{pmatrix} \delta V^{(x)} \\ \delta V^{(y)} \\ \delta V^{(z)} \end{pmatrix} = \begin{pmatrix} \frac{1}{r\cos\theta'}\delta V^{(\omega^{\mathrm{o}})} \\ \frac{1}{r}\delta V^{(\theta')} \\ \delta V^{(r)} \end{pmatrix}. \tag{2.28}$$

Analogously, the definition of the gravity gradient tensor $\delta\mathbf{G} = \nabla\otimes\nabla^T\delta V$ can be used to derive the second partial derivatives:

$$\mathbf{G} = \begin{pmatrix} \frac{1}{r}\delta V^{(r)} - \frac{\tan\theta'}{r^2}\delta V^{(\theta')} + \frac{1}{r^2\cos^2\theta'}\delta V^{(\omega^{\mathrm{o}}\omega^{\mathrm{o}})} & & \text{symmetric} \\ \frac{1}{r^2\cos\theta'}\delta V^{(\omega^{\mathrm{o}}\theta')} + \frac{\sin\theta'}{r^2\cos^2\theta'}\delta V^{(\omega^{\mathrm{o}})} & \frac{1}{r}\delta V^{(r)} + \frac{1}{r^2}\delta V^{(\theta'\theta')} & \\ -\frac{1}{r^2\cos\theta'}\delta V^{(\omega^{\mathrm{o}})} + \frac{1}{r\cos\theta'}\delta V^{(r\omega^{\mathrm{o}})} & -\frac{1}{r^2}\delta V^{(\theta')} + \frac{1}{r}\delta V^{(r\theta')} & V^{(rr)} \end{pmatrix}. \tag{2.29}$$

For the point on the nominal orbit $\theta' = 0$, the partial derivatives, the second order in particular, take much simpler forms. For instance:

$$\mathbf{G} = \begin{pmatrix} \frac{1}{r}\delta V^{(r)} + \frac{1}{r^2}\delta V^{(\omega^{\mathrm{o}}\omega^{\mathrm{o}})} & & \text{symmetric} \\ \frac{1}{r^2}\delta V^{(\omega^{\mathrm{o}}\theta')} & \frac{1}{r}\delta V^{(r)} + \frac{1}{r^2}\delta V^{(\theta'\theta')} & \\ -\frac{1}{r^2}\delta V^{(\omega^{\mathrm{o}})} + \frac{1}{r}\delta V^{(r\omega^{\mathrm{o}})} & -\frac{1}{r^2}\delta V^{(\theta')} + \frac{1}{r}\delta V^{(r\theta')} & \delta V^{(rr)} \end{pmatrix}. \tag{2.30}$$

Now, representation of the LL-SST observation equation in terms of the new partial derivatives will be more convenient. Since the LOS unit vector in the local orbital frame is simply a unit vector along

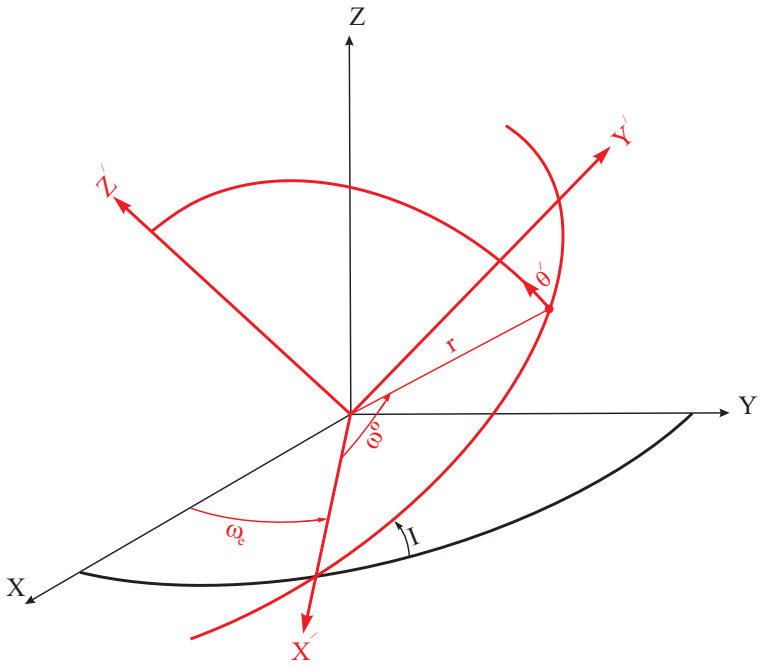


Figure 2.2: Orbital coordinate system

x-axis $\mathbf{e} = [1\,0\,0]^T$, the gravitational acceleration difference functional Eq. (2.11) is recast into:

$$\mathcal{L}^{\mathrm{LL}}\delta V \;=\; \delta V_2^{(x)} - \delta V_1^{(x)} \;=\; [\frac{1}{r}\delta V^{(\omega^{\mathrm{o}})}]_2 - [\frac{1}{r}\delta V^{(\omega^{\mathrm{o}})}]_1. \tag{2.31}$$

Similarly, the linear gradiometry equation Eq. (2.12) reads:

$$\mathcal{L}^{SG}\delta V = \delta V^{(xx)} \;=\; \frac{1}{r}\,\delta V^{(r)} \;+\; \frac{1}{r^2}\,\delta V^{(\omega^{\mathrm{o}}\omega^{\mathrm{o}})}. \tag{2.32}$$

Substituting the corresponding quantities from Table (2.2) into Eqs. (2.31) and (2.32) respectively results in:

$$\mathcal{L}^{\mathrm{LL}}\delta V \;=\; \sum_{m=0}^{N}\sum_{k=-N}^{N} k \sum_{n=n_1[2]}^{n_2} \bar{F}_{nm}^{k}(I)$$

$$\alpha_{nm}\;\; \frac{u_n(r_2)}{r_2}\sin\psi_{km}^2 - \frac{u_n(r_1)}{r_1}\sin\psi_{km}^1 \;\; - \beta_{nm}\;\; \frac{u_n(r_2)}{r_2}\cos\psi_{km}^2 - \frac{u_n(r_1)}{r_1}\cos\psi_{km}^1 \tag{2.33}$$

and

$$\mathcal{L}^{\mathrm{SG}}\delta V \;=\; -\sum_{m=0}^{N}\sum_{k=-N}^{N}\sum_{n=n_1[2]}^{n_2} \frac{k^2+n+1}{r^2}\,[A_{nmk}\cos\psi_{km} + B_{nmk}\sin\psi_{km}] \tag{2.34}$$

In the new representation, the LL-SST functional either in the gravitational difference or in the gradiometric approach is invariant. Consequently, LL-SST observations can be mapped into the geopotential by use of the transfer coefficients. Depending on the boundary utilized as the ideal geometric configuration, there are different possibilities at this step (Sneeuw, 2003).

2.1.3 Representation on a circle: mean circular orbital

To achieve this key element of the semi-analytical approach, the observations should be mapped onto the mean circular orbit (Sneeuw, 2003). Substituting the radial coordinate with the mean radius of the nominal orbit $\bar{a}$, Eq. (2.33) is recast into:

$$\mathcal{L}^{\mathrm{LL}}\delta V \;=\; \sum_{m=0}^{N}\sum_{k=-N}^{N} k \sum_{n=n_1[2]}^{n_2} \bar{F}_{nm}^{k}(I)$$

$$\frac{GMR_{\mathrm{E}}^{n}}{\bar{a}^{n+2}}\left\{\alpha_{nm}\left[\sin\psi_{km}^{2}-\sin\psi_{km}^{1}\right]\;-\;\beta_{nm}\left[\cos\psi_{km}^{2}-\cos\psi_{km}^{1}\right]\right\} \tag{2.35}$$

If we reference the LEO pair to the orbit longitude of their barycenter, we have:

$$\psi_{km}^{(1)}\;=\;\psi_{km}-\frac{\Delta\psi}{2};\qquad\qquad\psi_{km}^{(2)}\;=\;\psi_{km}+\frac{\Delta\psi}{2}, \tag{2.36}$$

where, $\Delta\psi\;=\;k\frac{\Delta s}{\bar{a}}$ with Δs as arc length of the nominal orbit between the pair. Consequently, rewriting Eq. (2.33) leads to an analytical model for LL-SST observations similar to the one described in Wagner (1983):

$$\mathcal{L}^{\mathrm{LL}}\delta V\;=\;\sum_{m=0}^{N}\sum_{k=-N}^{N}\frac{2}{\bar{a}}k\sin\frac{\Delta\psi}{2}\sum_{n=n_1[2]}^{n_2}\bar{F}_{nm}^{k}(I)\frac{GM}{\bar{a}}\quad\frac{R_{\mathrm{E}}}{\bar{a}}\quad^{n}\left[\alpha_{nm}\cos\psi_{km}\;+\;\beta_{nm}\sin\psi_{km}\right]. \tag{2.37}$$

Similar to Eq. (2.23), by performing the inner summation loop over n, one can derive the corresponding expression in terms of lumped coefficients and the respective transfer function as:

$$\mathcal{L}^{\mathrm{LL}}\delta V\;=\;\sum_{m=0}^{N}\sum_{k=-N}^{N}\frac{2}{\bar{a}}\,k\sin\frac{\Delta\psi}{2}\left[\widetilde{A}_{km}\cos\psi_{km}\;+\;\widetilde{B}_{km}\sin\psi_{km}\right]. \tag{2.38}$$

Analogously, the gradiometric equation Eq. (2.34) results in:

$$\mathcal{L}^{\mathrm{SG}}\delta V\;=\;\sum_{m=0}^{N}\sum_{k=-N}^{N}\widetilde{A}_{km}^{SG}\cos\psi_{km}+\widetilde{B}_{km}^{SG}\sin\psi_{km} \tag{2.39}$$

where,

$$\widetilde{A}_{nm}^{\mathrm{SG}}\;=\;-\sum_{n=n_1[2]}^{n_2}\frac{k^2+n+1}{\bar{a}^2}\,A_{nmk} \tag{2.40}$$

$$\widetilde{B}_{nm}^{\mathrm{SG}}\;=\;-\sum_{n=n_1[2]}^{n_2}\frac{k^2+n+1}{\bar{a}^2}\,B_{knm} \tag{2.41}$$

Estimation of the geopotential coefficients

For recovery of the geopotential coefficients $\bar{C}_{nm}$ and $\bar{S}_{nm}$, a linear system of equations based on the observation equations should be set up. The representation of the geopotential in terms of the orbital elements has one more index compared to the geopotential coefficients' indices. Consequently, one can use two different combinations to obtain $\bar{C}_{nm}, \bar{S}_{nm}$

- n– m **combination**
- k– m **combination**.

In the n– m combination, summation over k would be the inner summation loop. For instance, Eq. (2.37) leads to:

$$\mathcal{L}^{\mathrm{LL}}\delta V\;=\;\sum_{m=0}^{N}\sum_{n=n_1[2]}^{n_2}\frac{GM}{\bar{a}}\left(\frac{R_{\mathrm{E}}}{\bar{a}}\right)^{n}$$
$$\alpha_{nm}\left[\sum_{k=-N}^{N}\frac{2}{\bar{a}}k\sin\frac{\Delta\psi}{2}\bar{F}_{nm}^{k}(I)\cos\psi_{km}\right]+\beta_{nm}\left[\sum_{k=-N}^{N}\frac{2}{\bar{a}}k\sin\frac{\Delta\psi}{2}\bar{F}_{nm}^{k}(I)\sin\psi_{km}\right]. \tag{2.42}$$

Similar to the brute-force approach, one can set up the linear system of equations based on Eq. (2.42). Nevertheless, the corresponding normal matrix would be a full matrix whose inversion is a really high demanding computation task (Schrama, 1986). The same holds true for the gradiometry approach. In contrast, the k– m combination overcomes the problem simply by using the orthogonality of the base functions. To achieve this key point of the k– m combination, the following conditions should be fulfilled.

- Data are assumed to be observed regularly along the circular orbit. The objectives can be achieved during the measurement period. However, the data gaps during the maneuvering and sensor failures are inevitable. In reality, the deviation of evaluation points from the circular orbit reaches a few kilometers. The evenly-spaced time series should then be obtained by mapping the irregularly sampled observations. It is assumed that the orbit repetition period (T_r) is long enough that the overall data set yields to very homogeneous and regular reduced data set. The phase ψ_{km} is recast into Eq. (2.43) if the orbital elements are subject to the linear perturbation:

$$\psi_{km}(t) = \psi_{km}(t_0) + \dot{\psi}_{km}(t - t_0) \tag{2.43}$$

  For simplicity, it is supposed that both $\psi_{km}(t_0)$ and t_0 are zero. Therefore,

$$\psi_{km} = \dot{\psi}_{km} t = j\dot{\psi}_{km}\Delta t = j\Delta t(k\dot{\omega}^o + m\dot{\omega}_e). \tag{2.44}$$

  with $\Delta t = \frac{T}{N_p}$ the sampling rate and $j = 1, \cdots, N_p$ where N_p is the total number of evaluation points during the observation period T.

- Again for simplicity, the observations along the orbit are assumed uncorrelated. Hence,

$$\mathbf{P} = \frac{1}{\sigma^2}\mathbf{I}_{N_p \times N_p} \tag{2.45}$$

  with σ the measurement precision.

- The observation period T, the period of revolution $T_r = \frac{2\pi}{\dot{\omega}^o}$ and the nodal period $T_d = \frac{2\pi}{\dot{\omega}_e}$ satisfy the condition:

$$T = \alpha T_d = \beta T_r \tag{2.46}$$

  The orbit must be periodic in order to obtain a discrete 1D Fourier spectrum (Sneeuw, 2000). The orbit will be periodic if the basic frequencies $\dot{\omega}^o$ and $\dot{\omega}_e$ are commensurable. In other words, α and β should be the prime integer (i.e., they have no common divisor).

$$\frac{T_d}{T_r} = -\frac{\dot{\omega}^o}{\dot{\omega}_e} = \frac{\beta}{\alpha} \tag{2.47}$$

  Since the Earth's rotation rate $\dot{\theta}$ is larger than the nodal line precession $\dot{\Omega}_e$ and consequently $\dot{\omega}$ is always negative, the minus sign has been introduced in Eq. (2.47).

- Observations along the nominal orbit are modelled in one dimension while two basic frequencies are involved in each of the observations. The relationship between the 1D spectrum and 2D one reads:

$$\dot{\psi}_{km} = \dot{\omega}^o\left(k + m\frac{\dot{\omega}_e}{\dot{\omega}^o}\right) = \frac{\dot{\omega}^o}{\beta}(k\beta - m\alpha) = l\Delta\dot{\psi} \tag{2.48}$$

  where, $l = k\beta - m\alpha$ is an integer called *frequency number* which in absolute sense reaches at maximum $(\alpha + \beta)N$. $\Delta\dot{\psi} = \dot{\psi}_1$ is the spectral resolution. Therefore, one can define a mapping from 2D to 1D simply by:

$$\dot{\psi}_l = l\dot{\psi}_1.$$

- Shannon's sampling theorem[2] should be fulfilled in order to avoid aliasing namely called *aliasing of the first kind* (Sneeuw, 2000). According to the theorem, the measurement (sampling) frequency must be at least twice the maximum frequency to be measured. Hence,

$$N_p \leq 2(\alpha + \beta)N. \tag{2.49}$$

[2]The sampling theorem is due to Claude Shannon who first discovered it in 1949. Whenever Shannon's sampling theorem is not fulfilled, aliasing occurs.

Therefore, the sampling interval Δt is subject to

$$\Delta t = \frac{T}{N_p} = \frac{T}{2(\alpha+\beta)N} = \frac{T_r}{2\left(\frac{\alpha}{\beta}+1\right)N} =: \Delta t_N \tag{2.50}$$

with Δt_N, the Nyquist sampling interval. For instance, the minimum sampling frequency for the GRACE mission will be about 0.05 Hz if the maximum resolvable degree is equal to 120.

- Perfect signal restoration is guaranteed by satisfying the Shannon's sampling theorem. Nevertheless, it is just valid for the recovery of a digital signal (1D) or an image (2D) when the observables have the same dimension. However, two dimensional geopotential coefficients $\bar{C}_{nm}$ and $\bar{S}_{nm}$ are recovered just based on 1D observations along the nominal orbit. Therefore, the problem of frequencies overlapping via mapping procedure is the second problem that should be treated carefully. The problem which is called *aliasing of second type* has been carefully handled in Sneeuw (2000). The second type of aliasing does not occur if the mapping from $(k,m) \longmapsto n$ is *injective*. It is fulfilled if:

$$2(N+1) \leq \alpha + \beta. \tag{2.51}$$

  One can compute a rough estimate of $\beta > 1.9N$ as the necessary condition to prevent aliasing for Low Earth Orbiters (LEO).

  An alternative interpretation of the second type of aliasing can be given by taking the corresponding normal matrix of the geopotential coefficients in the *time-wise approach in the time domain*. Fulfilling the condition given in Eq. (2.51) leads to a diagonal normal matrix which is the main advantage of the semi-analytical approach (Koop, 1993).

- For obtaining a diagonal normal matrix with minimum bandwidth, the unknown geopotential coefficients ordering should follow:

  1. grouping $\bar{C}_{nm}$ and $\bar{S}_{nm}$ coefficients as two distinct sub-vectors $\mathbf{x}_{\mathrm{C}}$ and $\mathbf{x}_{\mathrm{S}}$ within the vector of unknowns $\mathbf{x}$,
  2. sorting the coefficients within the sub-vectors according to the order (m) into $N+1$ and N diminishing groups $\mathbf{x}_{\mathrm{C^m}}$ and $\mathbf{x}_{\mathrm{S^m}}$corresponding to $\bar{C}_{nm}$ and $\bar{S}_{nm}$ coefficients respectively,
  3. decomposing the respective group for each order into two distinct subgroups corresponding to the even and odd degrees $\mathbf{x}_{\mathrm{C}^{\mathrm{m}}_{\mathrm{e}}}$ and $\mathbf{x}_{\mathrm{C}^{\mathrm{m}}_{\mathrm{o}}}$.

Entries of the normal matrix of the unknown geopotential coefficients can be calculated by multiplying two columns of the design matrix to the corresponding observations. Consider for instance, the observation equations of the gradiometry type. Partial derivatives of the j-th observation with respect to $\bar{C}_{nm}$ and $\bar{S}_{nm}$ for two different cases ($n-$ m even and $n-$ m odd) respectively read:

$$\frac{\partial \mathcal{L}^{SG}\delta V_j}{\partial \bar{C}_{nm}} = u_n(\bar{a}) \sum_k \frac{k^2+n+1}{\bar{a}^2} \bar{F}^k_{nm} \begin{cases} \cos j\dot{\psi}_{km}\Delta t_N & n-m \text{ even} \\ \sin j\dot{\psi}_{km}\Delta t_N & n-m \text{ odd} \end{cases} \tag{2.52}$$

$$\frac{\partial \mathcal{L}^{SG}\delta V_j}{\partial \bar{S}_{nm}} = u_n(\bar{a}) \sum_k \frac{k^2+n+1}{\bar{a}^2} \bar{F}^k_{nm} \begin{cases} \sin j\dot{\psi}_{km}\Delta t_N & n-m \text{ even} \\ -\cos j\dot{\psi}_{km}\Delta t_N & n-m \text{ odd} \end{cases} \tag{2.53}$$

Running the index j over $0, \cdots, N_p-1$ for the N_p observations given on the nominal orbit and multiplying the two different columns corresponding to (n_1,m_1) and (n_2,m_2) pairs results in:

$$\mathbf{N}(x_1,x_2) = \frac{1}{\sigma^2} \sum_{j=0}^{N_p-1} \begin{cases} \frac{\partial \mathcal{L}^{\mathrm{SG}}\delta V_j}{\partial x_1} \frac{\partial \mathcal{L}^{\mathrm{SG}}\delta V_j}{\partial x2} & x_1 = \bar{C}_{n_1m_1} \quad x_2 = \bar{C}_{n_2m_2} \\ \frac{\partial \mathcal{L}^{\mathrm{SG}}\delta V_j}{\partial x_1} \frac{\partial \mathcal{L}^{\mathrm{SG}}\delta V_j}{\partial x_2} & x_1 = \bar{C}_{n_1m_1} \quad x_2 = \bar{S}_{n_2m_2} \\ \frac{\partial \mathcal{L}^{\mathrm{SG}}\delta V_j}{\partial x_1} \frac{\partial \mathcal{L}^{\mathrm{SG}}\delta V_j}{\partial x_2} & x_1 = \bar{S}_{n_1m_1} \quad x_2 = \bar{C}_{n_2m_2} \\ \frac{\partial \mathcal{L}^{\mathrm{SG}}\delta V_j}{\partial x_1} \frac{\partial \mathcal{L}^{\mathrm{SG}}\delta V_j}{\partial x_2} & x_1 = \bar{S}_{n_1m_1} \quad x_2 = \bar{S}_{n_2m_2} \end{cases} \tag{2.54}$$

where, $\mathbf{N}(x_1, x_2)$ is the entry of the normal matrix $\mathbf{N}$ which corresponds to the unknown pair (x_1, x_2). Simply by reordering the summation sequence from j, k_1, k_2 into k_1, k_2, j and considering the orthogonality properties of trigonometric series, the summations given in Eq. (2.54) for the cases $(x_1, x_2) = (\bar{C}_{n_1 m_1}, \bar{S}_{n_2 m_2})$ and $(x_1, x_2) = (\bar{S}_{n_1 m_1}, \bar{C}_{n_2 m_2})$ for every value of n_1, m_1, n_2 and m_2 will be equal to zero. Therefore, the first benefit of considering the aforementioned ordering of the unknown coefficients emerges:

$$\mathbf{N} = \begin{pmatrix} \mathbf{N}_{\bar{C}\bar{C}} & \mathbf{0} \\ \mathbf{0} & \mathbf{N}_{\bar{S}\bar{S}} \end{pmatrix}. \tag{2.55}$$

with $\mathbf{N}_{\bar{C}\bar{C}}$ and $\mathbf{N}_{\bar{S}\bar{S}}$ corresponding to the inner product of $\bar{C}_{nm}$s and $\bar{S}_{nm}$s respectively. According to the rule of trigonometric function product, among the elements of $\mathbf{N}_{\bar{C}\bar{C}}$ and $\mathbf{N}_{\bar{S}\bar{S}}$, the elements with different frequencies will be equal to zero. Hence,

$$\mathbf{N}_{\bar{C}\bar{C}} = \mathbf{N}_{\bar{S}\bar{S}} = 0 \qquad \textbf{if } | \dot{\psi}_{k_1 m_1} | \neq | \dot{\psi}_{k_2 m_2} | . \tag{2.56}$$

Eq. (2.56) implies that the sub-matrices $\mathbf{N}_{\bar{C}\bar{C}}$ and $\mathbf{N}_{\bar{S}\bar{S}}$ themselves will be two block-diagonal matrices if the coefficients are sorted according to order m. Therefore,

$$\mathbf{N}_{\bar{C}\bar{C}} = \begin{pmatrix} \mathbf{N}^0_{\bar{C}\bar{C}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{N}^N_{\bar{C}\bar{C}} \end{pmatrix} \tag{2.57}$$

$$\mathbf{N}_{\bar{S}\bar{S}} = \begin{pmatrix} \mathbf{N}^1_{\bar{S}\bar{S}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{N}^N_{\bar{S}\bar{S}} \end{pmatrix} \tag{2.58}$$

As already mentioned, fulfilling the injectivity condition Eq. (2.51), prevents the aliasing of second kind. Therefore, $| \dot{\psi}_{k_1 m_1} | = | \dot{\psi}_{k_2 m_2} |$ only occurs if, $(k_1, m_1) = (k_2, m_2) = (k, m)$. Consequently, for $m \neq 0$, we have:

$$\mathbf{N}^m_{\bar{C}\bar{C}}(n_1, n_2) = \mathbf{N}^m_{\bar{S}\bar{S}}(n_1, n_2) =$$

$$\frac{N_p}{2\sigma^2}\, u_{n_1}(\bar{a}) u_{n_2}(\bar{a}) \sum_{k=-\min(n_1,n_2)[2]}^{k=\min(n_1,n_2)} \frac{k^2 + n_1 + 1}{\bar{a}^2}\, \frac{k^2 + n_2 + 1}{\bar{a}^2} \bar{F}^k_{n_1 m} \bar{F}^k_{n_2 m}. \tag{2.59}$$

Which is valid both for $n - m$ even and odd. In case of $m = 0$, replacing $k_1 = -k_2 = k$ and $k_1 = k_2 = k$ leads to the identical values for $| \dot{\psi}_{k0} |$. Hence:

$$\mathbf{N}^0_{\bar{C}\bar{C}}(n_1, n_2) = \frac{N_p}{\sigma^2}\, u_{n_1}(\bar{a}) u_{n_2}(\bar{a})$$

$$\sum_{k=-\min(n_1,n_2)[2]}^{k=\min(n_1,n_2)} \frac{k^2 + n_1 + 1}{\bar{a}^2}\, \frac{k^2 + n_2 + 1}{\bar{a}^2} \left[\bar{F}^k_{n_1 0}\ \bar{F}^k_{n_2 0} + \bar{F}^{-k}_{n_1 0}\ \bar{F}^k_{n_2 0} \right]. \tag{2.60}$$

Moreover, $\bar{F}^k_{nm}$ are nonzero if n and k have the same parity. Due to condition $k_1 = k_2 = k$, only those entries of $\mathbf{N}^m_{\bar{C}\bar{C}}$ and $\mathbf{N}^m_{\bar{S}\bar{S}}$ are nonzero which correspond to n_1 and n_2 with the same parity. Therefore, partitioning the $\bar{C}_{nm}$ and $\bar{S}_{nm}$ coefficients into two distinct sets corresponding with n even and odd, forms both $\mathbf{N}^m_{\bar{C}\bar{C}}$ and $\mathbf{N}^m_{\bar{S}\bar{S}}$ as two diagonal matrices:

$$\mathbf{N}^m_{\bar{C}\bar{C}} = \begin{pmatrix} \mathbf{N}^m_{\bar{C}^e\bar{C}^e} & \mathbf{0} \\ \mathbf{0} & \mathbf{N}^m_{\bar{C}^o\bar{C}^o} \end{pmatrix} \tag{2.61}$$

and

$$\mathbf{N}^m_{\bar{S}\bar{S}} = \begin{pmatrix} \mathbf{N}^m_{\bar{S}^e\bar{S}^e} & \mathbf{0} \\ \mathbf{0} & \mathbf{N}^m_{\bar{S}^o\bar{S}^o} \end{pmatrix} \tag{2.62}$$

$\mathbf{N}^{m}_{\bar{C}^e\bar{C}^e}$ and $\mathbf{N}^{m}_{\bar{S}^e\bar{S}^e}$ as well as $\mathbf{N}^{m}_{\bar{C}^o\bar{C}^o}$ and $\mathbf{N}^{m}_{\bar{S}^o\bar{S}^o}$ are identical. Among them either $\mathbf{N}^{0}_{\bar{C}^e\bar{C}^e}$ or $\mathbf{N}^{0}_{\bar{C}^o\bar{C}^o}$ at most has a size of $(\frac{N}{2}+1)\times(\frac{N}{2}+1)$. The size of the others are diminishing by increasing degree n.
Eventually, inversion of a huge normal matrix of size $(N+1)^2\times(N+1)^2$ is replaced by $2N$ times inversion of matrices with comparatively smaller sizes (Koop, 1993). Consequently, the replacement alleviates the solution considerably and reduces computation time enormously.
This approach is usually called *time-wise approach in time domain* whereas as an alternative, the problem can also be solved in *frequency domain* using the lumped coefficients concept (Sneeuw, 2003).

Timewise approach in the frequency domain

For the problem formulation in frequency domain, the observations given along the nominal orbit can be assumed as values of a one-dimensional discretized function. For instance, Eq. (2.39) is recast into:

$$\mathcal{L}^{SG}\delta V \;=\; \sum_{l=-N(\alpha+\beta)}^{N(\alpha+\beta)} A_l^{SG}\cos l\Delta\dot{\psi} \;+\; B_l^{SG}\sin l\Delta\dot{\psi} \tag{2.63}$$

with $A_l^{SG}=\widetilde{A}_{km}^{SG}$ and $B_l^{SG}=\widetilde{B}_{km}^{SG}$. Eq. (2.63) is the finite Fourier series of the $2N(\alpha+\beta)$ sample points measured at:

$$0,\frac{\pi}{N(\alpha+\beta)},\frac{2\pi}{N(\alpha+\beta)},\cdots,\frac{[2N(\alpha+\beta)-1]\pi}{N(\alpha+\beta)},$$

or more succinctly:

$$\omega_p^{\circ}\;=\;\frac{p\pi}{N(\alpha+\beta)}\qquad\qquad p=0,1,2,\cdots,2N(\alpha+\beta)$$

along the mean circle ($r=\bar{a}$). Performing a 1D FFT of the observations over the repeat period leads to the Fourier coefficients A_l^{SG} and B_l^{SG} as the quasi-observations of the unknown coefficients. Using $l=k\beta-m\alpha$, one can recover two dimensional lumped coefficients $\widetilde{A}_{km}^{SG}$ and $\widetilde{B}_{km}^{SG}$ simply by picking up the Fourier coefficient corresponding to each pair of (k,m). Performing block-wise least squares adjustment on the two dimensional lumped coefficients results in the geopotential coefficients $\delta\bar{C}_{nm}$ and $\delta\bar{S}_{nm}$ (Sneeuw, 2003).

2.2 Space-wise approach: representation on a Sphere

In this section, the recovery of the geopotential spherical harmonic coefficients from a set of nearly evenly-spaced observations on the mean orbital sphere is described. In reality, the measurements are observed in a layer of radial distance varying few kilometers around the mean sphere. Therefore, they should be reduced on the mean sphere and interpolated onto the equiangular grid, since the reduced quantities are non-uniformly distributed horizontally.
Mapping the irregular measurements onto a nearly equiangular grid is the crucial point to the space-wise approach. The reduced observations are always contaminated with gridding error. However, the high CPU time and memory demanding brute-force approach can be replaced with the semi-analytical or space-wise approach if the data are mapped on an ideal geometry. Consequently, implementation of an intelligent method for data reduction would be the key point to the space-wise approach which will be addressed in chapter 3.
In contrast to the time-wise approach, the data utilized for the space-wise approach should be point-wise evaluation functional on the incremental potential δV. The acceleration difference between the low orbiters is a non-localized functional on the geopotential while the gradiometric measurements are local by their very nature. Hence, in this section we merely concentrate on the gradiometric representation of the LL-SST observations.
Furthermore, the space-wise approach is merely applicable to *invariant* functionals. It is the single biggest obstacle to the space-wise approach and would be the main disadvantage of the space-wise approach in comparison with the time-wise approach. Even in case we avail ourselves to invariant

functionals, we can employ *Banach's fixed-point* theorem and recover the geopotential using even the *non-invariant functionals* using an *iterative algorithm.*

2.2.1 Mathematical formulation

Consider a function $f(r,\theta,\lambda)$ which is an invariant functional $\mathcal{L}$ on the incremental potential:

$$f(r,\theta,\lambda) \;=\; \mathcal{L}\;\delta v \;=\; \sum_{n=0}^{N}\sum_{m=0}^{n} h_n(r)u_n(r)\left[\delta\bar{C}_{nm}\bar{Y}^{\mathrm{c}}_{nm} \;+\; \delta\bar{S}_{nm}\bar{Y}^{\mathrm{s}}_{nm}\right]. \tag{2.64}$$

Observations on the mean orbital sphere ($r=\bar{r}$) or reduced onto the sphere read:

$$f(\theta,\lambda) \;=\; \sum_{n=0}^{N}\sum_{m=0}^{n} h_n(\bar{r})u_n(\bar{r})\left[\delta\bar{C}_{nm}\bar{Y}^{\mathrm{c}}_{nm} \;+\; \delta\bar{S}_{nm}\bar{Y}^{\mathrm{s}}_{nm}\right]. \tag{2.65}$$

Swapping the n- and m-summations and introducing a new set of coefficients $A^f_m(\theta)$ and $B^f_m(\theta)$ called *latitude lumped coefficients* leads to (Sneeuw, 1994):

$$f(\theta,\lambda) \;=\; A^f_0(\theta) \;+\; \sum_{m=1}^{N} A^f_m(\theta)\cos m\lambda \;+\; \sum_{m=1}^{N} B^f_m(\theta)\sin m\lambda \tag{2.66}$$

where,

$$\begin{pmatrix} A^f_m(\theta) \\ B^f_m(\theta)\end{pmatrix} \;=\; \sum_{n=m}^{N} h_n(\bar{r})u_n(\bar{r})\begin{pmatrix}\delta\bar{C}_{nm}\\ \delta\bar{S}_{nm}\end{pmatrix}\bar{P}_{nm}(\cos\theta), \tag{2.67}$$

Eq. (2.66) can be interpreted as the *finite Fourier series* of the $2N$ sample points measured at:

$$\lambda_p \;=\; \frac{p\pi}{N}, \qquad\qquad p=0,1,2,\cdots,2N-1$$

along the parallels ($\theta=\text{const}$). Noting that the number of longitudinal data points is assumed equal to the total number of latitude lumped coefficients ($=N+1+N-1=2N$). The finite Fourier series for function $f(\theta,\lambda_p)$ sampled at λ_p is:

$$f(\theta,\lambda) \;=\; \frac{a^f_0(\theta)}{2} \;+\; \sum_{m=1}^{N-1}\left[a^f_m(\theta)\cos m\lambda + b^f_m(\theta)\sin m\lambda\right] \;+\; \frac{a^f_N(\theta)}{2}\cos N\lambda, \tag{2.68}$$

where,

$$a^f_m(\theta) \;=\; \frac{1}{N}\sum_{p=0}^{2N-1} f(\theta,\lambda_p)\cos m\lambda_p \qquad\qquad m=0,1,\cdots,N \tag{2.69}$$

$$b^f_m(\theta) \;=\; \frac{1}{N}\sum_{p=0}^{2N-1} f(\theta,\lambda_p)\sin m\lambda_p \qquad\qquad m=1,2,\cdots,N-1 \tag{2.70}$$

The Fourier coefficients $a^f_m(\theta)$ and $b^f_m(\theta)$ are the quasi-observations of the linear system of equations represented by Eq. (2.67).

$$A^f_0(\theta) \;=\; \frac{a^f_0(\theta)}{2}, \qquad\text{and}\qquad A^f_N(\theta) \;=\; \frac{a^f_N(\theta)}{2}$$

$$A^f_m(\theta) \;=\; a^f_m, \quad\text{and}\quad B^f_m(\theta) \;=\; b^f_m, \quad m=1,2,\cdots,N-1$$

Noting that the coefficients $B^f_0(\theta)=B^f_N(\theta)=0$.
To recover the coefficients $\bar{C}_{nm}$ and $\bar{S}_{nm}$, the corresponding linear system of equations is inverted by a block-wise least squares adjustment (Sneeuw, 2003). This mapping can be written in matrix

notation as:

$$\left.\begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\right\} = \mathbf{P}\boldsymbol{\Gamma}\left\{\begin{matrix}\mathbf{c}\\ \mathbf{s}\end{matrix}\right. \tag{2.71}$$

where,

$$\mathbf{a} = \begin{pmatrix} A_m^f(\theta_0) \\ \vdots \\ A_m^f(\theta_{N_\theta-1}) \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} B_m^f(\theta_0) \\ \vdots \\ B_m^f(\theta_{N_\theta-1}) \end{pmatrix} \tag{2.72}$$

$$\mathbf{P} = \begin{pmatrix} \bar{P}_{mm}(\cos\theta_0) & \bar{P}_{m+1m}(\cos\theta_0) & \cdots & \bar{P}_{Nm}(\cos\theta_0) \\ \bar{P}_{mm}(\cos\theta_1) & \bar{P}_{m+1m}(\cos\theta_1) & \cdots & \bar{P}_{Nm}(\cos\theta_1) \\ \vdots & \vdots & \cdots & \vdots \\ \bar{P}_{mm}(\cos\theta_{N_\theta-1}) & \bar{P}_{m+1m}(\cos\theta_{N_\theta-1}) & \cdots & \bar{P}_{Nm}(\cos\theta_{N_\theta-1}) \end{pmatrix} \tag{2.73}$$

and matrix of the transfer coefficients, $\boldsymbol{\Gamma}$:

$$\boldsymbol{\Gamma} = \begin{pmatrix} h_m(\bar{r})u_m(\bar{r}) & 0 & \cdots & 0 \\ 0 & h_{m+1}(\bar{r})u_{m+1}(\bar{r}) & \cdots & 0 \\ 0 & 0 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_N(\bar{r})u_N(\bar{r}) \end{pmatrix}, \tag{2.74}$$

finally,

$$\mathbf{c} = \begin{pmatrix} \bar{C}_{mm} \\ \vdots \\ \bar{C}_{Nm} \end{pmatrix} \qquad \mathbf{s} = \begin{pmatrix} \bar{S}_{mm} \\ \vdots \\ \bar{S}_{Nm} \end{pmatrix} \tag{2.75}$$

Moreover, N_θ is the number of latitudinal grid points which is defined based on the selected grid on the mean sphere. However, the number of observations $A_m^f(\theta)$ or $B_m^f(\theta)$ at least should be equal to the number of the unknowns $\bar{C}_{nm}$ $\bar{S}_{nm}$ at most ($\min(N_\theta) = N + 1$). Different types of gridding will be discussed in chapter 5 in more detail.
In general, the number of linear equations of type (2.71) is larger than or equal to the number of coefficients. Therefore, the solution is obtained by the method of least squares,

$$\widehat{\mathbf{c}} = \left(\boldsymbol{\Gamma}^T\mathbf{P}^T\mathbf{P}\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\Gamma}^T\mathbf{P}^T\mathbf{a} \tag{2.76}$$

$$\widehat{\mathbf{s}} = \left(\boldsymbol{\Gamma}^T\mathbf{P}^T\mathbf{P}\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\Gamma}^T\mathbf{P}^T\mathbf{b} \tag{2.77}$$

As it will be seen in chapter 5, the estimated coefficients are erroneous due to non-orthogonality of discretized Legendre function (Sneeuw, 1994). The solution can be improved either by modification of the estimator or by considering the Gauss grid rather than the equiangular one. It will be explained in more detail in chapter 5.

Case Study: LL-SST problem

A succinct expression of Eq. (2.12) is:

$$\delta V^{(xx)} = \mathbf{e}_{\mathrm{B}}^T\mathbf{G}\mathbf{e}_{\mathrm{B}} \tag{2.78}$$

where, $\mathbf{G}$ is the Hotine-Marussi tensor and contains the second partial derivatives of the geopotential. The LOS unit vector $\mathbf{e}_{\mathrm{B}}$, can be expressed in terms of the unit vector's azimuth α, as shown in Fig. (2.3), in the local north-oriented reference frame

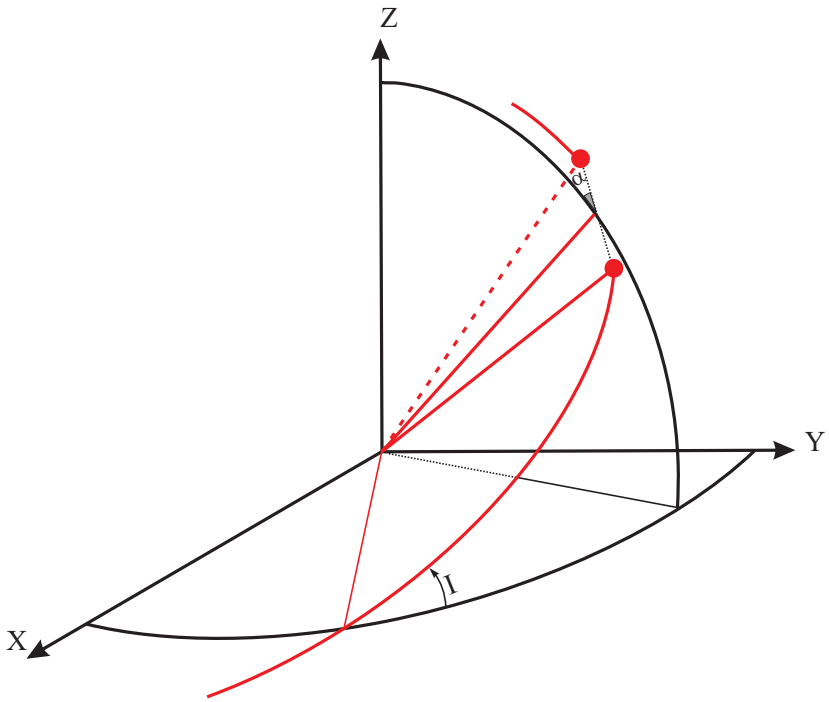


Figure 2.3: Definition of intersatellite azimuth

$$\mathbf{e}_{\mathrm{B}} \approx \begin{pmatrix} \sin\alpha \\ \pm\cos\alpha \\ \frac{\delta r}{\rho} \end{pmatrix} \tag{2.79}$$

with δr as the difference between the radii of the satellites at the evaluation point. The second component of the vector $\mathbf{e}_{\mathrm{B}}$ with plus sign is corresponding to ascending arcs whereas the minus sign is used for the descending arcs.
The azimuth of the intersatellite vector, i.e., the angle, is counted clockwise from North, between the direction of intersatellite vector and the local meridian passing through the barycenter (Vermeer, 1990). In an inertial frame it reads,

$$\sin\alpha \;=\; \frac{\cos I}{\sin\theta}. \tag{2.80}$$

The alternative is to choose the nominal azimuth of the satellite equal to the Earth-fixed azimuth α_E, which is obtained from α by adding a *sidewind correction* (Vermeer, 1990).

$$\alpha_E \;=\; \arcsin\left(\frac{\cos I}{\sin\theta}\right) \;-\; \arctan\left(\frac{\sqrt{\sin^2\theta - \cos^2 I}}{\frac{n}{\dot{\Theta}} - \cos I}\right) \tag{2.81}$$

For a mission with a polar inclination Eq. (2.81) is recast into:

$$\tan\alpha_E \;=\; -\frac{\sin\theta}{\frac{n}{\dot{\Theta}}} \tag{2.82}$$

which is the formalism utilized in Rummel et al. (1993). In the literature (Albertella et al. (1995)), an apparently different expression can be found for the computation of azimuth α_E,

$$\sin\alpha_E \;=\; \frac{\frac{n}{\dot{\Theta}}\frac{\cos I}{\sin\theta} - \sin\theta}{\sqrt{\left(\frac{n}{\dot{\Theta}}\right)^2 + \sin^2\theta - 2\frac{n}{\dot{\Theta}}\cos I}} \tag{2.83}$$

Despite of their appearance, identical results are obtained by both formulations. Again for polar orbit for instance, the same relationship can be derived by simplifying Eq. (2.83).
The radial component of the unit vector is significantly smaller than the others. Since one satellite is chasing the other nearly in the same orbit and the differences in the radii of the two satellites are relatively small compared to the intersatellite distance. Therefore, Eq. (2.78) is recast into Eq. (2.84) by neglecting the radial component of the unit vector,

$$\delta V^{(xx)} = \cos^2\alpha \mathbf{G}_{22} \pm \sin 2\alpha \mathbf{G}_{23} + \sin^2\alpha \mathbf{G}_{33}. \tag{2.84}$$

All evaluation points are located in the outer space where the Laplace equation ($\nabla^2 \delta V = 0$) is valid. Thus

$$\delta V^{(xx)} = \cos^2\alpha (-\mathbf{G}_{11} - \mathbf{G}_{33}) \pm \sin 2\alpha \mathbf{G}_{23} + \sin^2\alpha \mathbf{G}_{33}. \tag{2.85}$$

Consequently, Eq. (2.85) is recast into

$$\delta V^{(xx)} = -\cos^2\alpha \mathbf{G}_{11} \pm \sin 2\alpha \mathbf{G}_{23} - \cos 2\alpha \mathbf{G}_{33}. \tag{2.86}$$

Based on the reduced observations onto the mean orbital sphere, a linear system of equations can be set up using Eq. (2.86). The system can be solved by either using direct inversion or iterative methods. A brute force approach can also be used for the achieved linear system corresponding to the reduced observations. Compared to the time-wise approach, the reduced observations' respective linear system consists of less equations but with a normal matrix of the same size. It can be an advantage for vectorized programming algorithm.
In order to recover the geopotential coefficients using the space-wise approach, the functional should be isotropic. As seen in Eq. (2.84), the observation equation of the GRACE mission is non-isotropic and even different for the ascending and descending tracks. Therefore, the mathematical model cannot be implemented directly and should be modified. As will be discussed in chapter 6, the contribution of $\mathbf{G}_{23}$ into Eq. (2.86) is negligible, except in the polar regions. Neglecting the second term on the right-hand-side of Eq. (2.86) and dividing both sides of the equation by $-\cos^2\alpha$, leads to

$$-\frac{\delta V^{(xx)}}{\cos^2\alpha} = \mathbf{G}_{11} + \frac{\cos 2\alpha}{\cos^2\alpha}\mathbf{G}_{33}. \tag{2.87}$$

The left-hand-side of the equation is considered as quasi-observation. The right-hand-side of Eq. (2.87) contains respectively *invariant (homogenous and isotropic)* and *non-invariant (inhomogeneous and non-isotropic)* components, where the invariant component is the dominant, see chapter 7. Keller (1994) proposed an iterative solution based on the Banach fixed point theorem for recovery of the geopotential using the non-invariant functional.

Iterative recovery scheme based on Banach's fixed point theorem

The basic data processing scheme, corresponding to this iterative formulation, consists of the following steps:

1. Splitting the observable into invariant and non-invariant components Eq. (2.87),

2. Computing the non-invariant constituent based on a priori information, i.e. available Earth's gravity field models

$$-\frac{\delta V^{(xx)}}{\cos^2\alpha} - \frac{\cos 2\alpha}{\cos^2\alpha}\mathbf{G}^0_{33} = \mathbf{G}_{11}, \tag{2.88}$$

3. Setting up the linear system of equations (Eq. 2.71) of the invariant component

$$u_n = \frac{GM}{r}\left(\frac{R_E}{r}\right)^n \tag{2.89}$$

$$h_n = \frac{(n+1)(n+2)}{r^2}, \tag{2.90}$$

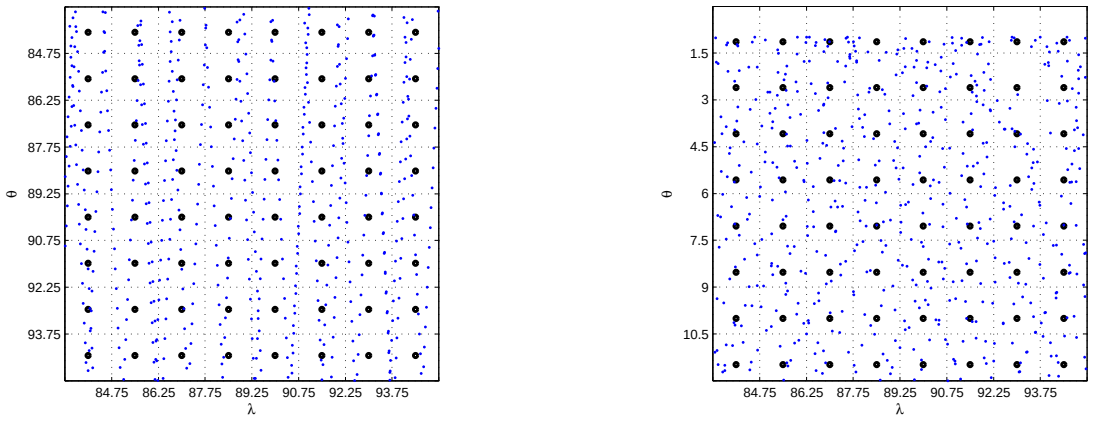
4. Estimating the geopotential coefficients based on the global harmonic analysis of the invariant component on the mean sphere Eqs. (2.76) and (2.77),

5. And iterating steps 2-5 based on the estimated solution to mitigate the non-invariant part initialization error.

Convergence of the scheme for cross-track in-line observable of STEP mission investigated in Keller (1994). As it will be seen in Chapter 7, this leads to a convergent iterative scheme for the GRACE observable.

2.2.2 Observation distribution

From a mathematical point of view, the observation density and the ground-track coverage are the key issues for all prediction methods. Moreover, data reduction quality, to a great extent depends on the roughness of the topography beneath the evaluation point. In satellite geodesy, the distribution pattern of observations is a complicated function of the satellite inclination and periodicity of the orbit, whereas the data density is a function of sampling rate and mission duration. Extending the duration of observations beyond the orbit repetition period of repeating orbit improves merely the observation density along the orbit arcs while for the non-repeat orbit, it improves both the pattern and the density. For instance, the equatorial and polar ground track of the GRACE mission for a one- and a four-month spans of the observations are depicted in Figs. (2.4) and (2.5). Compared



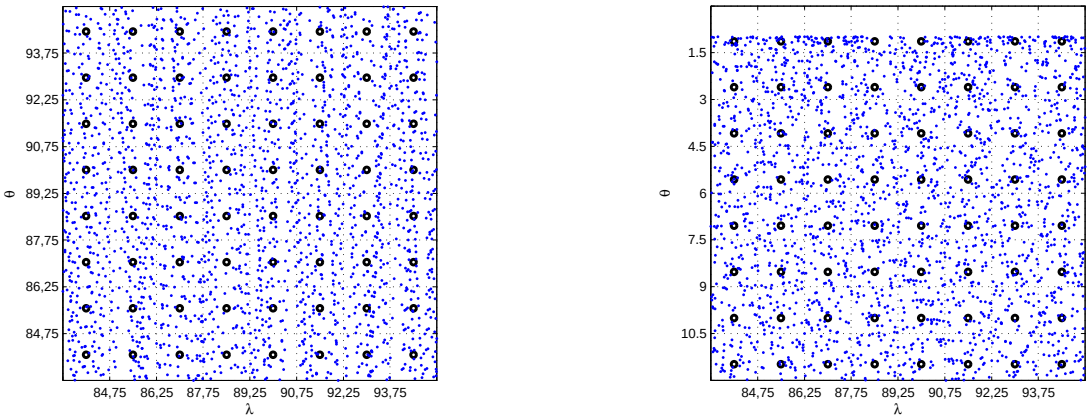
(a) Equatorial (b) Polar

Figure 2.4: Typical ground-track patterns of a one-month orbit with 10 Hz sampling (dots: observation points and circles: grid points)

to Fig. (2.4), four months of the observations generate a relatively high density data distribution. However, the along-track patterns are still clearly visible in later case, see Fig. (2.5). It should also be noted that data distribution for the polar and near-polar orbits around the poles is relatively dense and randomly scattered, whereas that for the equatorial regions is more directional along the ascending and descending arcs.
Fig. (2.6) shows data point distribution within $1.5^\circ \times 1.5^\circ$ block size for the four-month observations sampled at 10 Hz. As seen in Fig. (2.6), the cell with the minimum number contains 15 points and it reaches 55 points at most.

2.3 Space-wise versus time-wise — A comparison

The time-wise and the space-wise approaches were discussed in the first two sections of this chapter. The first and second partial derivatives of the geopotential in general and some combinations in



(a) Equatorial (b) Polar

Figure 2.5: Typical ground-track patterns of a four-month orbit with 10 Hz sampling (dots: observation points and circles: grid points)

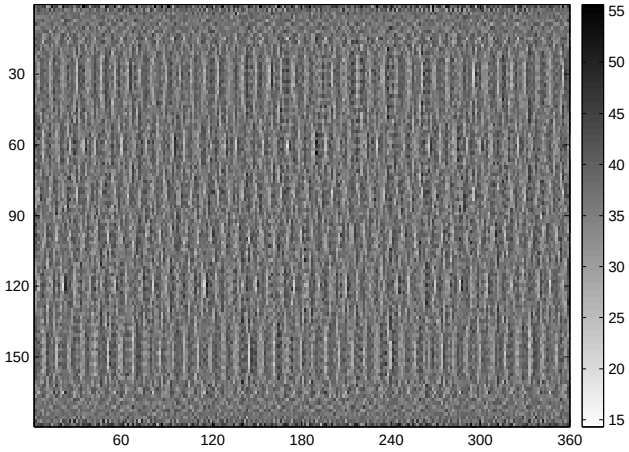


Figure 2.6: A four-month span of the GRACE real data distribution

particular were described. Moreover, the time-wise approach was split into brute-force and semi-analytical approaches. In this section, different methods are compared and a trade-off between pros and cons of each of these techniques is addressed.

The time-wise in time domain solution as well as the two-stage modelling leading to the lumped coefficients are the intermediate alternatives. Smaller linear systems of equations (at most $(N+1)/2$) and faster computation because of FFT implementation are the main advantages of the alternative techniques. Having a repeat orbit with a completed period and the continuous stream of data at the specified sampling rate (Δt_N) given along the nominal orbit, are the prerequisites for the time-wise approach both in time and in frequency domains. An iterative solution is required to mitigate the approximation error.

Alternatively, the space-wise approach, mapping the orbital observations on the mean orbital sphere, and recovering the coefficients using the global spherical harmonic analysis, can be considered. The method can be implemented directly for the isotropic function while the non-isotropic observations must be treated using an iterative solution. Moreover, the data should be mapped onto a specific grid (the Gauss grid, see chapter 5) in order to minimize the consequence of non-orthogonality of the Legendre functions. However, the orbit could be either a repeat or a non-repeat one and data gaps have no effect on the quality of the estimated coefficients as long as the sphere is covered sufficiently

dense. Due to the presence of interpolation error, an iterative solution is inevitable both for isotropic and non-isotropic functions.
The time-wise and space-wise approaches from theoretical and practical point of view are complementary. The former is closely related to satellite geodesy whereas the latter is rooted in geodetic boundary value problems.

Chapter 3

Observation Reduction

In general, it is essentially impossible to collect data in an evenly-spaced spatial grid. Except for the brute-force method, a representation of the originally irregularly-spaced data points on a regular grid is required. Gridding is thus one of the first and most crucial steps in the semi-analytical and the space-wise approaches which were introduced to alleviate numerical difficulties of the brute-force technique.
The problem of data reduction is a special case of approximation, in which a set of scattered observations are used to generate an evenly-spaced data grid in the observational space or in a specific subspace of the original space. Consequently, the gridding error is a function of the approximation method, basis functions, data distribution and density, and the grid spacing.
In this chapter, the observation reduction using *polynomials, harmonic polynomials, radial basis functions and rational* functions, as the potential gridding techniques, are described and their performance compared on two different sets of the simulated data: one for an invariant and one for non-invariant function, respectively. Moreover, the approximation methods are optimized using numerical optimization techniques.

3.1 Approximation methods

In general, suppose a set of n_o distinct irregularly-distributed sampling points [1] $\mathbf{x}_o$, $o = 1, \cdots, n_o$, and the sampled values $s_o = f(\mathbf{x}_o) + \epsilon_o$, where f is the underlying function (isotropic or non-isotropic) and ϵ_o the sampling error. The task is to find the function's value on a finite and discrete set of nearly regularly-spaced grid nodes $\mathbf{x}_g$ by approximating the function with a reasonable and computable function $\widetilde{f}$.

3.1.1 Polynomial approximation

To build up the approximating function $\widetilde{f}$, different basis functions can be employed. Among them, the polynomial basis as the simplest ones are expressed as:

$$\Pi^3 = \left\{1, x, y, z, x^2, xy, xz, y^2, yz, z^2, x^3, \cdots\right\}. \tag{3.1}$$

According to *Weierstrass approximation theorem*,

> *If $f(\mathbf{x})$ is a continuous real-valued function on $[\mathbf{x}_1, \mathbf{x}_2]$, and if any $\varepsilon > 0$ is given, then there exists a polynomial $\widetilde{f}$ on $[\mathbf{x}_1, \mathbf{x}_2]$ such that:*
>
> $$| f(\mathbf{x}) - \widetilde{f}(\mathbf{x}) | < \varepsilon$$
>
> *for all $\mathbf{x} \in [\mathbf{x}_1, \mathbf{x}_2]$. In other words, any continuous function on a closed and bounded interval, can be uniformly approximated on that interval by polynomials to any degree of*

[1] In general, we assume 3D sampling points whose coordinates are given either in the Cartesian (x, y, z) or in the spherical coordinates (r, θ, λ).

accuracy (Jeffreys and Jeffreys, 1988).

The corresponding approximating polynomial function is,

$$\widetilde{f}(x,y,z) \;=\; \sum_{i=0}^{d}\sum_{j=0}^{d-i}\sum_{k=0}^{d-i-j} a_{ijk}x^i y^j z^k. \tag{3.2}$$

with d, the maximum degree of the approximating polynomial function. a_{ijk}s are the unknown coefficients which are estimated by setting up a linear system of equations based on the discrete observed values of the function on the evaluation points. The system will be under-determined if the total number of unknown coefficients is less than the number of data points:

$$df(d) \;=\; n_o - u(d) = n_o - \left(\frac{d^3+6d^2+11d}{6}+1\right) \geq 0. \tag{3.3}$$

The higher the polynomial degree, the lower the degree of freedom becomes. Fig. (3.1) shows how the degree of freedom rapidly decreases by increasing the maximum degree of the basis functions.

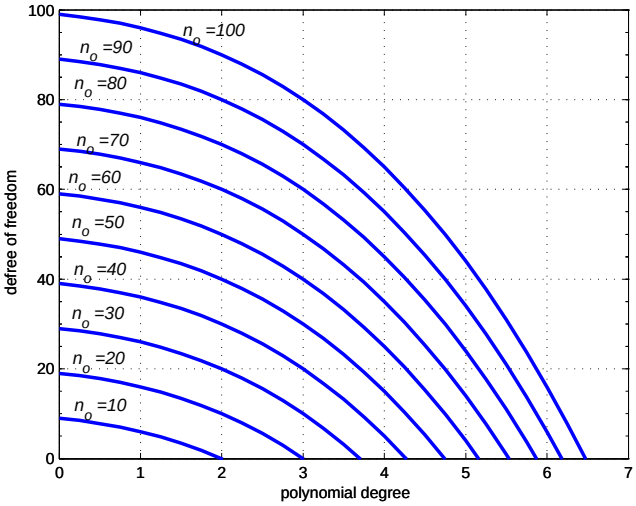


Figure 3.1: Maximum polynomial degree versus degree of freedom

From the classical point of view, the best approximating polynomial will be the best approximation if

$$\|f(\mathbf{x}) - \widetilde{f}(\mathbf{x})\|$$

reaches the minimum for the choice. In practice, the underlying function is known just at the data points. The criteria will then be modified to select the optimal approximating function as:

$$\|\mathbf{v}\| \;=\; \|f(\mathbf{x}_o) - \widetilde{f}(\mathbf{x}_o)\| \;=\; \sum_{o=1}^{n_o}[s_o - \widetilde{f}(\mathbf{x}_o)]^2,$$

with $\mathbf{v}$ as the vector of residuals[2]. Increasing the maximum degree of the approximating function in one hand, decreases the norm of the residuals at the data points and in the other hand, an generates artificial oscillation within the data points. Therefore, the cubic polynomial function is usually the high-degree polynomial function which is used for approximation.
Furthermore, the approximating function of the low-degree polynomial represents the low frequency signals. Moreover, the instability of normal matrix is the numerical problem which appears when the data points are widely spread on the domain. Implementation of the patch-wise polynomials

[2]A more realistic criteria can be set up using some check-points which can be selected from the data points. They are not involved in the coefficients' estimation process and are used just for checking the quality of fitting.

prevents the instability problem and represents the high-frequency feature of the underlying function to a great extent.

To provide coordinates of the evaluation points for the patch-wise approximation, a local coordinate system is used. For instance, for the space-wise approach, the LNRF defined at the patch mid-point may be the more convenient local frame since the coefficients indicate the partial derivatives of the observable along the coordinate lines, the parallels and the meridians. It is more practical to transform the coordinate differences, i.e. the data points minus the respective patch mid-point coordinates, rather than the coordinates themselves.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} -\frac{X_g Z_g}{r^2 p} & -\frac{Y_g Z_g}{r^2 p} & \frac{p}{r^2} \\ -\frac{Y_g}{p^2} & \frac{X_g}{p^2} & 0 \\ \frac{X_g}{r} & \frac{Y_g}{r} & \frac{Z_g}{r} \end{pmatrix} \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix}, \tag{3.4}$$

where,

$$\begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} = \begin{pmatrix} X - X_g \\ Y - Y_g \\ Z - Z_g \end{pmatrix},$$

and $p = \sqrt{X_g^2 + Y_g^2}$, and $r = \sqrt{X_g^2 + Y_g^2 + Z_g^2}$.

Analogously, for the time-wise in the frequency domain and torus approach, LORF will be the more convenient local frame due to the aforementioned reason. The transformation equations $(x, y, z) = \mathbf{J}(\omega^{\mathrm{o}}, \omega_{\mathrm{e}}, I)(\Delta X, \Delta Y, \Delta Z)$ are given by:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\cos\omega_{\mathrm{e}} \sin\omega^{\mathrm{o}} - \sin\omega_{\mathrm{e}} \cos\omega^{\mathrm{o}} \cos I & -\sin\omega_{\mathrm{e}} \sin\omega^{\mathrm{o}} + \cos\omega_{\mathrm{e}} \cos\omega^{\mathrm{o}} \cos I & \cos\omega^{\mathrm{o}} \sin I \\ \sin\omega_{\mathrm{e}} \sin I & -\cos\omega_{\mathrm{e}} \sin I & \cos I \\ \cos\omega_{\mathrm{e}} \cos\omega^{\mathrm{o}} - \sin\omega_{\mathrm{e}} \sin\omega^{\mathrm{o}} \cos I & \sin\omega_{\mathrm{e}} \cos\omega^{\mathrm{o}} + \cos\omega_{\mathrm{e}} \sin\omega^{\mathrm{o}} \cos I & \sin\omega^{\mathrm{o}} \sin I \end{pmatrix} \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix}, \tag{3.5}$$

where, ω_{e} and ω^{o} refer to the mid-point of the corresponding patch. By assuming a known inclination I and using Eqs. (3.6) and (3.7), one can derive the transformation matrix entries (e.g. Karrer, 2000):

$$\sin\omega^{\mathrm{o}} = \frac{Z}{r \sin I} \tag{3.6}$$

and

$$\tan\omega_{\mathrm{e}} = \frac{\cos\omega^{\mathrm{o}} - Y \sin\omega^{\mathrm{o}} \cos I}{Y \cos\omega^{\mathrm{o}} - X \sin\omega^{\mathrm{o}} \cos I} \tag{3.7}$$

Although the patch-wise polynomial approximation represents more features and leads to a stable normal matrix, the achieved results are not convincing.

In the polynomial modelling, just the geometrical properties of the function were considered. Incorporating the underlying function's physical process may improve the estimation accuracy.

3.1.2 Harmonic and biharmonic polynomials

As mentioned in Chapter 2, every observation can be written as a functional on the gravity field. Some of these functionals are harmonics while the others are not. For instance, $r^2 T_{rr}$ is harmonic in the outer space whereas $T_{\theta\theta}$ is a typical non-invariant one. It is strongly recommended to incorporate any additional available information on the underlying function in the interpolation and approximation process. For example, implementation of the harmonicity condition results in the *Harmonic approximating polynomials* $\widetilde{f}^H(\mathbf{x})$,

$$\nabla^2 \widetilde{f}^H(\mathbf{x}) = 0 \tag{3.8}$$

similarly, for the *biharmonic* polynomials,

$$\nabla^2 \left(\nabla^2 \widetilde{f}^B(\mathbf{x}) \right) = \nabla^4 \widetilde{f}^B(\mathbf{x}) = 0. \tag{3.9}$$

For example, a biharmonic function of three variables for instance, reads (Hardy, 1983):

$$\frac{\partial^4 \widetilde{f}^B}{\partial x^4} + \frac{\partial^4 \widetilde{f}^B}{\partial y^4} + \frac{\partial^4 \widetilde{f}}{\partial z^4} + 2\frac{\partial^4 \widetilde{f}^B}{\partial x^2 \partial y^2} + 2\frac{\partial^4 \widetilde{f}^B}{\partial y^2 \partial z^2} + 2\frac{\partial^4 \widetilde{f}^B}{\partial y^2 \partial z^2} \;=\; 0 \tag{3.10}$$

Implementation of the constraints Eqs. (3.8) and (3.9) in general case Eq. (3.2), is rather complicated. Hence, we apply the constraints for special cases. Nevertheless, implementation of the Laplace equation Eq. (3.2) results in a polynomial of degree $(d-2)$ whose coefficients should be equal to zeros. Therefore, the number of equations of the type functional constraint is equal to the number of a $(d-2)$ order polynomial function,

$$s^H = \frac{(d-2)^3 + 6(d-2)^2 + 11(d-2)}{6} + 1, \tag{3.11}$$

and the corresponding degree of freedom is:

$$df^H \;=\; n_o - u + s^H = n_o - (d+1)^2. \tag{3.12}$$

For instance, the cubic polynomial approximating function will be harmonic if the following linear conditions are fulfilled:

$$\begin{aligned} a_{200} + a_{020} + a_{002} &= 0 \\ 3a_{300} + a_{120} + a_{102} &= 0 \\ a_{210} + 3a_{030} + a_{012} &= 0 \\ a_{201} + a_{021} + 3a_{003} &= 0 \end{aligned} \tag{3.13}$$

The additional equations increase the linear system's degree of freedom. Moreover, the underlying function physical characteristic is also considered. We can then expect to obtain a higher accuracy for the grid values.

The polynomial functions, up to the third degree, are biharmonic. Use of the higher degree biharmonic polynomials leads to some additional constraints which may increase the gridding accuracy. Similar to the harmonic polynomial, it can be easily proved that the number of functional constraints is

$$s^B = \frac{(d-4)^3 + 6(d-4)^2 + 11(d-4)}{6} + 1, \tag{3.14}$$

and the corresponding degree of freedom is

$$df^B \;=\; n_o - u + s^B = n_o - 2(d^2+1). \tag{3.15}$$

For instance, the fourth order polynomial yields to the following constraint

$$3\left(a_{400} + a_{040} + a_{004}\right) \;+\; \left(a_{220} + a_{202} + a_{002}\right) \;=\; 0. \tag{3.16}$$

The biharmonic polynomial approximation analogous with the harmonic approximation increases the number of equations. However, the number of equations is comparatively less than that of the harmonic approximation. Nevertheless, in both cases, the number of unknowns is a quadratic function of d, while it is a cubic function in the case of ordinary polynomial approximation. Fig. (3.2) shows a maximum polynomial degree versus degree of freedom, both for the harmonic and the biharmonic polynomials. Compared to the polynomial approximation, fitting the high-degree harmonic and biharmonic polynomials can be performed using much less data points.

The additional linear functional constraints can be combined with the observation equations and solved in an augmented linear system (e.g. Wells, 1994),

$$\widehat{\mathbf{x}} \;=\; \mathbf{N}^{-1}\left[\mathbf{I} + \mathbf{A}_c^T\left(\mathbf{A}_c\mathbf{N}^{-1}\mathbf{A}_c\right)^{-1}\mathbf{A}_c\mathbf{N}^{-1}\right]\mathbf{u}, \tag{3.17}$$

where, $\mathbf{N}$ and $\mathbf{u}$ are the normal equation matrix and the known vector of the polynomial approximation respectively. $\mathbf{A}_c$ is the first order design matrix of the constraint equations.

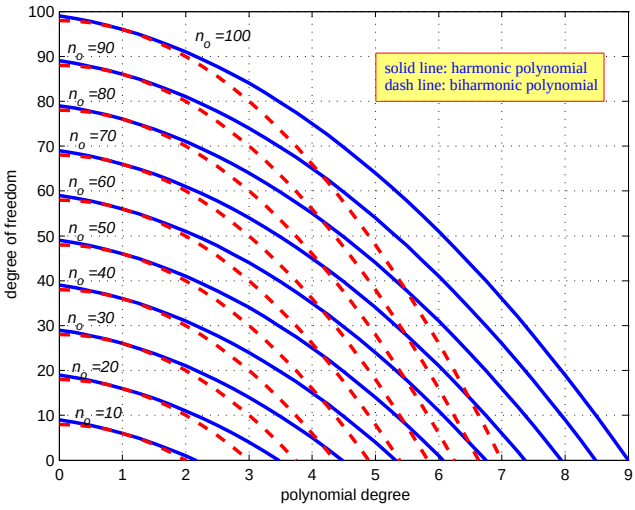


Figure 3.2: Maximum harmonic and biharmonic polynomial degree versus *df*

In order to compare the proposed gridding methods with the other techniques, numerical analysis based on the same data sets, have been performed. The achieved results are presented in section 3.4.4.

3.1.3 Rational approximation

Mathematically, tendency of the polynomials to oscillate is the main disadvantage of polynomial approximation. For the high-degree polynomials, the oscillation exceeds the average bounds of approximation error (Tao and Hu, 2001). The *rational approximation (*RA*)* scheme is the alternative which is significantly smoother and can spread the approximation error more evenly between data points. Moreover, the polynomial approximation results in an unacceptable solution in the presence of finite discontinuities (outside of the approximation domain), while the RA solution is not far out (Burden and Faires, 1997; Tao and Hu, 2001).

Definition: Let f be the previously defined function. Let d_p and d_q be non-negative integers. Then, the class $R(d_p, d_q)$ of (ordinary) rational functions is the class of functions $r(\mathbf{x})$ which can be represented in the following form:

$$r(\mathbf{x}) = \frac{p(\mathbf{x})}{q(\mathbf{x})} \tag{3.18}$$

where,

$$p(\mathbf{x}) = \sum_{i=0}^{d_P} \sum_{j=0}^{d_P-i} \sum_{k=0}^{d_P-i-j} a_{ijk} x^i y^j z^k$$

$$q(\mathbf{x}) = \sum_{i=0}^{d_q} \sum_{j=0}^{d_q-i} \sum_{k=0}^{d_q-i-j} b_{ijk} x^i y^j z^k$$

where, a_{ijk} and b_{ijk} are real numbers and $q(\mathbf{x}) \neq 0$. The rational approximation problem consists of finding an element $r \in R(d_p, d_q)$ satisfying:

$$r(\mathbf{x}_o) = s_o \qquad o = 1, 2, \cdots, n_o \geq u(d_p) + u(d_q), \tag{3.19}$$

where, $u(d)$ is the number of unknown coefficients of a d-order polynomials. In general, existence of a solution is not guaranteed. Nevertheless, it is not hard to prove that at least one nontrivial solution always exists for the following combined equations which is equivalent to the rational

approximation problem (Wuytack, 1974): [3]

$$s_o\, q(\mathbf{x}_o) \;-\; p(\mathbf{x}_o) \;=\; 0. \tag{3.20}$$

The linear system of Eq. (3.20) is a system of n_o linear homogeneous equations in the $u(d_p) + u(d_q)$ unknowns a_{ijk} and b_{ijk}.

Although Eq. (3.20) is linear, the initial values of the coefficients are required to set up the linear system of equations. Therefore, the model should be solved iteratively (Tao and Hu, 2001).
Alternatively, the model can be reduced into a parametric model without loss of generality. Assuming $b_{000} \neq 0$ and dividing both the numerator and the denominator by this coefficient, leads to:

$$s_o \;=\; p'(\mathbf{x}_o) - s_o\, q'(\mathbf{x}_o), \tag{3.21}$$

where,

$$p'(\mathbf{x}) \;=\; \sum_{i=0}^{d_p} \sum_{j=0}^{d_p-i} \sum_{k=0}^{d_p-i-j} \frac{a_{ijk}}{b_{000}}\, x^i y^j z^k$$

and

$$q'(\mathbf{x}) \;=\; \sum_{i=0}^{d_q} \sum_{\substack{j=0 \\ i\wedge j\wedge k\neq 0}}^{d_q-i} \sum_{k=0}^{d_q-i-j} \frac{b_{ijk}}{b_{000}}\, x^i y^j z^k$$

Equations (3.21) can be assumed as a linear system of observation equations which can be solved directly. Herein, we implement the modified equation of the rational approximation.
Despite of the noteworthy advantages of the rational model, the instability of the first design matrix of the normal equations and a relatively large number of the unknown coefficients are the main disadvantages of the formulation. In order to stabilize the normal matrix inversion, regularization of the respective normal matrix is required. However, it is interesting to note that the formulation is fortunately insensitive to the regularization parameter and can be easily treated.
As it will be seen in section (3.4.4), the rational gridding method yields promising results. However, because of the large number of unknowns, more data points are required even for a low-degree rational model.

3.1.4 Radial basis functions

For one-dimensional data, the polynomial and rational approximations result in non-singular linear systems if the data points are distinct. However, existence of the solution is not guaranteed in general for higher dimensions. For instance in 3D, it yields a singular normal matrix if the data points are along an arbitrary line. This problem is common to all approximation methods with basis functions independent of the data points.
The problem can be bypassed by following a theoretically different approach. Instead of using linear combinations of the polynomial basis functions, which are independent of the data points, a linear combination of translates of a radially symmetric (about its center) single base function is employed. The technique which is called *radial basis functions* (RBF) method, was originated by one of the greatest geodesists *Richard L. Hardy*. Hardy (1971), as the pioneer of the RBF technique, applied a special class of RBF called *multiquadrics* (MQ) on geodetic data. Ability to handle arbitrarily distributed data and be easily generalized from one to multi-dimensional spaces, as well as the adjustable roughness capability, have made the method be of particular interest to many research and engineering fields of application. Numerical solution of partial differential equations; artificial neural networks (e.g. Dudnik and Bidyuk, 2003), medical imaging (e.g. Carr et al., 1997), surface and volume visualization(e.g. Levoy, 1988; Nielson et al., 1991), quantum hydrodynamics (e.g. Trahan and Wyatt, 2003) are some examples of where the RBF method has recently been employed.

[3] Wuytack (1974) showed the existence of a solution for the case of interpolation. However, if the solution exists for interpolation, its existence is fully guaranteed for approximation.

Multiquadric method's definition: For the already given set of n_o distinct data (source) points $\mathbf{x}_o$, the MQ interpolant is given by:

$$\widetilde{f}^{MQ}(\mathbf{x}) \;=\; \sum_{k=1}^{n_o} \alpha_k \, \varphi_k(\mathbf{x}, \mathbf{x}_k, c^2) \;=\; \sum_{k=1}^{n_o} \alpha_k \, \sqrt{c^2 + \|\mathbf{x} - \mathbf{x}_k\|^2}, \tag{3.22}$$

with $\|\cdot\|$ as the Euclidian norm. The expansion coefficients α_k are the unknown coefficients associated with n_o nodal points, and c^2 is the *shape parameter* of the radial basis φ_k which defines the basis functions sharpness (Hardy, 1990). Assume c^2 is known, the unknown coefficients are then the solution of the following symmetric linear system:

$$\mathbf{A}_{n_o \times n_o} \boldsymbol{\alpha}_{n_o \times 1} \;=\; \mathbf{l}_{n_o \times 1}, \tag{3.23}$$

where, the entries of $\mathbf{A}$ are given by $a_{ij} \;=\; \varphi_j(\mathbf{x}_i, c^2)$.

Originally, Hardy (1971) proposed the method for topographic map generation using irregularly-sampled height data. Later, he utilized the method for approximating the Earth's gravity field both for the exterior and interior of the Earth (e.g. Hardy and Göpfert, 1975; Hardy and Nelson, 1986a,b). Moreover, he applied the method for solving purely geometrical problems (Hardy, 1972, 1977), see Hardy (1990, and references therein) for more details. Furthermore, the MQ method was used by other researchers and scientists in other scientific disciplines. Hardy also introduced *Inverse Multiquadric* (IMQ) method with radial function of the type $\varphi_k \;=\; \left(c^2 + \|\mathbf{x}_k - \mathbf{x}\|^2\right)^{\frac{-1}{2}}$.
In parallel, new methods similar to MQ method, developed independently in most cases, started emerging. It was followed by the *thin-plate-spline* (TPS) method with the basis functions of the type $\varphi \;=\; \|\mathbf{x} - \mathbf{x}_k\|^2 \log\|\mathbf{x} - \mathbf{x}_k\|$ and $\varphi \;=\; \|\mathbf{x} - \mathbf{x}_k\|^3$ (Duchon, 1976). Schagen (1979, 1980) introduced yet another radial function namely called the *Sum of Gaussians* (SOG) $\varphi \;=\; e^{-\|\mathbf{x}-\mathbf{x}_k\|/2\sigma_k^2}$.
Franke (1979) carried out one of the most important studies on a vast number of interpolation methods including MQ. Therein, he compared about 30 different methods' performances on 6 different test functions sampled on 3 different sampling rates. The MQ method provided the best approximation in 13 of the 18 and was followed by TPS as the best approximating method in the others. Although the numerical study indicated the superiority of the MQ technique over TPS and SOG, it was suffering from lack of theoretical foundation.
Eventually, mathematical foundations of the MQ as well as the IMQ were provided by Micchelli (1986). In addition, he also presented sufficient conditions to guarantee non-singularity of other types of radial interpolating functions. Furthermore, his results were completely compatible with the already given conditions for TPS by (Duchon, 1976).
Micchelli's results had a great impact on the generalization of the MQ technique into a new class of interpolating functions which has been known as the *radial basis functions* (RBF). Using translates of a single homogenous and isotropic base function $\varphi(\rho)$ that depends only on the Euclidian distance from its center ($\rho_k \;=\; \|\mathbf{x}-\mathbf{x}_k\|$), is the rationale behind this general method. According to Micchelli's results, the method is applicable to multidimensional interpolation with a fully guaranteed nonsingular normal matrix.

RBF method's definition: Assume a set of n_o distinct data points $\mathbf{x}_k \in \mathbb{R}^n$ and their respective observed scalar values $f(\mathbf{x}_k)$. The RBF interpolant is defined as:

$$\widetilde{f}^{\mathrm{RBF}}(\mathbf{x}) \;=\; \sum_{k=1}^{n_o} \alpha_k \varphi_k(\rho) \tag{3.24}$$

where, $\varphi_k(\rho)$ is the interpolant kernel function. The expansion coefficients α_k are determined using the liner system of equations of Eq. (3.23) type, with the normal matrix entries of the type $a_{i,j} = \varphi_i(\rho_j) = \|\mathbf{x}_j - \mathbf{x}_i\|$.

In contrast to Eq. (3.24) which is a *pure radial sum* namely called basic RBF, the *augmented* RBF can be defined as:

$$\widetilde{f}^{\mathrm{aug}}(\mathbf{x}) \;=\; \sum_{k=1}^{n_o} \alpha_k \varphi_k(\rho) \;+\; \sum_{i=1}^{u_m} \beta_i p_i(\mathbf{x}), \tag{3.25}$$

with $p_i(\mathbf{x}) \in \Pi_m^n$ (n-dimensional polynomial basis of degree m). The basic RBF is purely deformable and thus limited and there are no global influences. The additional polynomial function absorbs the contribution of the long wavelengths if there are any. Besides the radial component coefficients u_m polynomial coefficients should also be estimated. Consequently, the linear system of interpolation conditions is an underdetermined one with n_o equations and $n_o + u_m$ unknowns. Therefore, u_m additional equations are required to set up a determined linear system. The u_m additional conditions required to obtain a unique solution are typically chosen as (Hickernell and Hon, 1998):

$$\sum_{k=1}^{n_o} \alpha_k \, p_i(\mathbf{x}_k) \, = \, 0, \tag{3.26}$$

or in matrix form,

$$\mathbf{P}^T \boldsymbol{\alpha} \, = \, \mathbf{0}. \tag{3.27}$$

Combination of the interpolation conditions and the additional equations is recast into the following linear system of equations (Ling, 2003),

$$\begin{pmatrix} \mathbf{A} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{l} \\ \mathbf{0} \end{pmatrix}, \tag{3.28}$$

where, $\mathbf{A}$ is the first design matrix of the interpolation condition $\widetilde{f}(\mathbf{x}_i) = s_i$. It is interesting to note that the maximum degree of the polynomial component m depends on the polynomial precision ν, required by the user. Hickernell and Hon (1998) for instance, showed that a polynomial function of degree $m = \frac{(\nu+1)(\nu+2)}{2}$ is required to obtain the polynomial precision ν for modelling 2D surface wind data.

Some common examples of the RBF kernels that lead to a uniquely solvable method, are given in Table (3.1). As seen in Fig. (3.3), various RBFs were compared with identical shape parameter ($c = 0.1$). The inverse multiquadric RBF is more localized as compared to the others. As the figure implies, both the inverse multiquadric and the Gaussian RBFs are monotonically decreasing with distance increase, whereas the multiquadric RBF is increasing monotonically with increasing distance from the source center (source point).

Table 3.1: Some of the more commonly employed RBF kernels

Type of Kernel	$\varphi(\rho)$	
Hardy's Multiquadric (MQ)	$\sqrt{\rho^2 + c^2}$	Infinitely smooth RBF
Inverse Multiquadric (IMQ)	$\frac{1}{\sqrt{\rho^2+c^2}}$	Infinitely smooth RBF
Inverse Quadratic (IQ)	$\frac{1}{\rho^2+c^2}$	Infinitely smooth RBF
Thin-Plate Spline (TPS)	$\rho^2 \log(\rho)$	piecewise smooth RBF
Gaussian	$e^{-c^2\rho^2}$	Infinitely smooth RBF
Linear	ρ	piecewise smooth RBF
Cubic	ρ^3	piecewise smooth RBF

Furthermore, since the underlying function is usually a functional of the incremental potential, the infinitely smoothed RBFs either in basic or augmented forms yield nearly identical results. Therefore, we retain the basic inverse multiquadric RBF.
As it will be seen in 3.4.4, for the test functions, the inverse multiquadric RBF obtains a comparatively higher gridding accuracy. Consequently, we retain the inverse multiquadric RBF and no comparison will be performed on the observation gridding.

MQ and IMQ in the spherical coordinates: Representation of the MQ or IMQ interpolant in terms of the spherical coordinates (r, θ, λ) is a specific representation which is more convenient

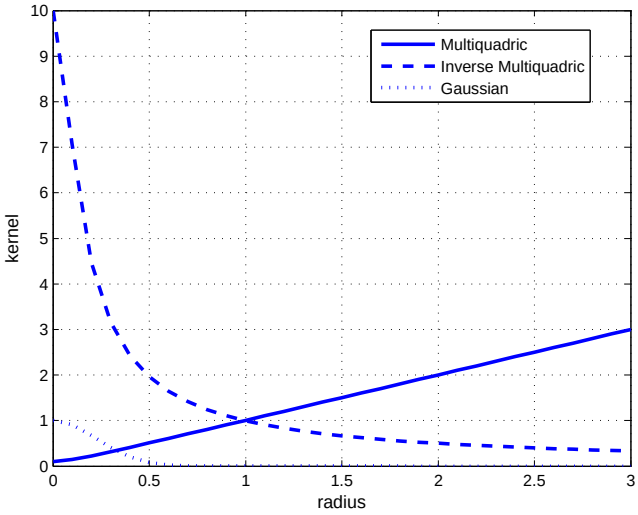


Figure 3.3: Comparison of infinitely smoothed RBFs

for the gravity field applications. The MQ kernel for instance, in three variables (x, y, z) is:

$$\varphi_k \;=\; \sqrt{\rho^2(\mathbf{x}, \mathbf{x}_k) + c^2} \;=\; \sqrt{(x - x_k)^2 + (y - y_k)^2 + (z - z_k)^2 + c^2} \tag{3.29}$$

ρ^2 in terms of spherical coordinates reads,

$$\rho^2(r, r_k, \psi_k) \;=\; \sqrt{r^2 + r_k^2 - 2rr_k \cos\psi_k} \tag{3.30}$$

where,

$$\cos\psi_k \;=\; \cos\theta\cos\theta_k + \sin\theta\sin\theta_k\cos(\lambda - \lambda_k). \tag{3.31}$$

Using Eq. (3.30), transformation of the coordinates from the spherical into Cartesian is bypassed. Moreover, if the grid points as well as the source points are assumed on the sphere with radius $\bar{r}$, then,

$$\varphi(\bar{r}, \psi_k) \;=\; 2\bar{r}\sqrt{\sin^2\frac{\psi_k}{2} + c'^2}. \tag{3.32}$$

One can simply scale the observations and represent the MQ interpolation conditions as:

$$s'_i \;=\; \sum_{k=1}^{n_o} \alpha_k \sqrt{\sin^2\frac{\psi_{i,k}}{2} + c'^2}, \tag{3.33}$$

with $s'_i = \frac{s_i}{2\bar{r}}$, $\psi_{i,k}$ the spherical distance between i and k, and $c'^2 = \frac{c}{2\bar{r}}$. Similarly, the MQ or IMQ interpolant can be expressed just as a function of the spherical distance ψ.

RBFs kernel with adaptive shape parameter: Usually, a single parameter c^2 is used as the shape parameter of the kernels. For instance, Shul'min and Mitel'man (1974) proposed the following formula for the single shape parameter computation,

$$c^2 \;=\; \frac{1}{n_o(n_o - 1)} \sum_{i=1}^{n_o}\sum_{j=1}^{n_o} \rho_{ij}^2, \tag{3.34}$$

which is just the mean quadratic distance between the data points and independent from the underlying function. Therefore, we cannot expect to obtain a reasonable accuracy for all types of functions.
Ruprecht and Müller (1995) proposed using individual values c_k^2 for each data point p_k, com-

puted from the distances to the nearest neighbor. For instance, for MQ kernel,

$$\varphi(\rho_k, c_k) \;=\; \sqrt{\rho_k^2 + c_k^2}; \qquad c_k \;=\; \min\{\rho_{i,k}\} \qquad i = 1, 2, \cdots n_o \tag{3.35}$$

The modification generates more flexible interpolation kernels by taking the local distribution of the data point. However, even with this modified flexible kernel, the underlying function has no contribution to the interpolation process.

Having discussed the virtues of radial basis functions, they are by no means infallible. In the case of infinitely smoothed RBFs as IMQ, selection of an appropriate shape parameter is paramount to good results. In order to optimize the shape parameter by considering the observed values besides their relative distributions, we will consider the problem as a constrained optimization problem.

3.2 Optimal approximation

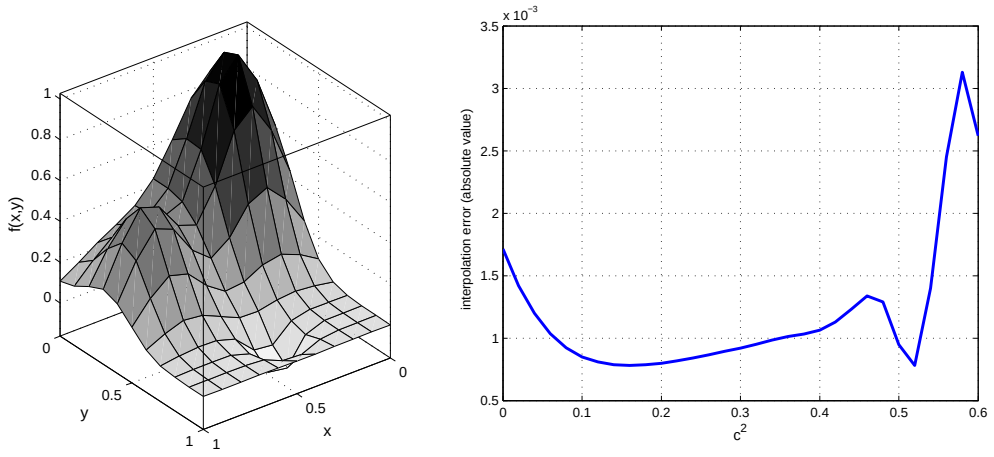
We discussed different algorithms for gridding the scattered data on a regular grid. As already mentioned, the gridding accuracy of the localized RBFs depends on the selection of shape parameter. In other words, the method should be optimized in order to obtain a reasonable solution. A similar problem is also hidden in the polynomial-based approximations. For instance, there is no clear clue for selection of the polynomial appropriate order. Selection of an optimal set of polynomial basis is yet another question. Therefore, the polynomial-based approximations should also be optimized by selecting the appropriate basis and the optimal order. Unlike the interpolation, a posteriori estimate of the error statistics is the byproduct of the approximation process. However, they are useless since they indicate the model misfit just at the observation points.
To bring it to the point, consider one of the most frequently used test function (Franke, 1979):

$$\begin{aligned} f(x,y) \;=\; & 0.75\exp\left[-\tfrac{(9x-2)^2+(9y-2)^2}{4}\right] + 0.75\exp\left[-\tfrac{(9x+1)^2}{49} + \tfrac{(9y+1}{10}\right] \\ & +0.50\exp\left[-\tfrac{(9x-7)^2+(9y-3)^2}{4}\right] - 0.20\exp\left[(9x-4)^2 - (9y-7)^2\right]. \end{aligned} \tag{3.36}$$

Assume the function is sampled with spacing $(\Delta x, \Delta y) = (0.1, 0.1)$. The problem is to find the function at an arbitrary point, for instance $P(0.65, 0.65)$. Depending on the shape parameter, a different value is obtained. The function and interpolation errors at point P are shown in Fig. (3.4). As it can be seen in the figure, even with a really high density data, the interpolation error changes significantly with the shape parameter variation. This simple example also shows the existence of local minima for the shape parameter. Therefore, as mentioned earlier, an optimal choice of the parameter c^2 is required for obtaining reasonable interpolation accuracy by means of RBFs. Moreover, we should bear in mind that the error could be a non-monotonic function. In general, the interpolation or approximation problem can be formulated as a constrained optimization problem with an object function g and the linear observation equations as the subject functions:

$$\begin{cases} \mathbf{A}\mathbf{x}_1 \;=\; \mathbf{l} \\ g(\mathbf{x}_1, \mathbf{x}_2) \rightsquigarrow \min \end{cases} \tag{3.37}$$

where, $\mathbf{x}_1$ and $\mathbf{x}_2$ are the unknown basis functions coefficients and the additional model parameters like c^2, respectively.
In the approximation problem, norm of the error function $e(\mathbf{x}) = \widetilde{f}(\mathbf{x}) - f(\mathbf{x})$ is usually considered as the object function g. Since $f(\mathbf{x})$ is unknown, the error function can only be discretely estimated at the observation points.
The estimation will be realistic if the function is evaluated at some *check points*, i.e. the observation points which are not involved in the unknown estimation process. However, in most cases of the Earth's gravity field determination it is almost impossible to split the data points into two parts because of the low number of observation points. For such cases, the method called *one-leave-out method* can be implemented.



(a) Function (b) interpolation error at $P(0.65, 0.65)$

Figure 3.4: Franke test function and interpolation error versus the shape parameter

3.2.1 One-leave–out method

Consider the given data set with n_o observations. Instead of excluding a group of observations as the check points, a single observation is removed from the observation vector. In other words, the interpolant is constructed just based on $n_o - 1$ source points. Let us assume $\widetilde{f}^{[k]}$ as the approximating function which has been set up based on the data points, except k^{th} point. The error function can be evaluated at point $\mathbf{x}_k$ based on $\widetilde{f}^{[k]}$. Performing the evaluation for each of n_o data points, yields a $n_o \times 1$ vector whose entries are a rough estimate of the error function at the observation points,

$$\widetilde{\mathbf{e}} = \begin{pmatrix} \widetilde{e}^{[1]} \\ \vdots \\ \widetilde{e}^{[k]} \\ \vdots \\ \widetilde{e}^{[n_o]} \end{pmatrix} = \begin{pmatrix} \widetilde{f}^{[1]} - s_1 \\ \vdots \\ \widetilde{f}^{[k]} - s_k \\ \vdots \\ \widetilde{f}^{[n_o]} - s_{n_o} \end{pmatrix} \tag{3.38}$$

Consequently, the first formulation of the optimization problem can be introduced as:

$$\begin{cases} \mathbf{A}\mathbf{x}_1 = \mathbf{l} \\ \|\widetilde{\mathbf{e}}\| \rightsquigarrow \min \end{cases} \tag{3.39}$$

Rippa (1999) tested the object functions defined by the L_1 and the L_2 norms of $\widetilde{\mathbf{e}}$ and found that there is a very slight advantage in using the object function defined by the L_1 norm. Then,

$$\begin{cases} \mathbf{A}\mathbf{x}_1 = \mathbf{l} \\ \sum_{k=1}^{n_o} \mid \widetilde{e}^{[k]} \mid \rightsquigarrow \min \end{cases} \tag{3.40}$$

Clearly, there is no dependency between the observation and the respective estimated error at the point. Hence, we virtually generate an independent criterion which is the main advantage of the proposed method.

In gridding problems, we are interested in evaluating the underlying function accurately at the grid points rather than at the data points. Accordingly, the object function can be modified to involve the interpolation error at the grid points as well:

$$\begin{cases} \mathbf{A}\mathbf{x}_1 = \mathbf{l} \\ \sum_{k=1}^{n_o} \mid \widetilde{e}^{[k]} \mid + \sum_{k=1}^{n_o} \sum_{j=1}^{n_I} \mid \widetilde{e}_j^{[k]} \mid \rightsquigarrow \min \end{cases} \tag{3.41}$$

The term $\widetilde{e}_j^{[k]}$ is the approximation error at the grid point $j \in \{1, 2, \cdots, n_I\}$, which is estimated in the absence of the k^{th} observation. However, $\widetilde{e}_j^{[k]}$ is inestimable since the underlying function is unknown at the grid point. From a statistical point of view, the mean value is the best estimation of a repeatedly measured quantity:

$$\widetilde{e}_j^{[k]} = \widetilde{f}_j^{[k]} - \frac{1}{n_o} \sum_{i=1}^{n_o} \widetilde{f}_j^{[i]} \tag{3.42}$$

Hence, Eq. (3.42) gives a reasonable estimate of the error value at the grid point j.
Although the aforementioned method of object function, to a certain extent, presents the reality, it is a cumbersome procedure particularly for RBF interpolation. For instance, n_o normal matrices of size $(n_o - 1) \times (n_o - 1)$ should be inverted for an individual value of the shape parameter. To alleviate this heavy computation, we utilize two different schemes for RBF interpolation and the approximation techniques, respectively.

3.2.2 RBFs interpolation

Rippa (1999) proposed an innovative single-step computation scheme for the RBFs interpolation methods. It is called single-step scheme since the unknown coefficients is computed only once. The computational scheme corresponding to this one-stage estimation consists of the following steps:

1. Compute the radial basis coefficients α_k based on the full set of observation,

$$\boldsymbol{\alpha} = \mathbf{A}^{-1}\mathbf{l} \tag{3.43}$$

2. Set up a new system of equations defined as:

$$\boldsymbol{\gamma}^{[k]} = \mathbf{A}^{-1}\mathbf{i}^{[k]} \tag{3.44}$$

   where, $\mathbf{i}^{[k]}$ is the k^{th} column of the $n_o \times n_o$ identity matrix and $\boldsymbol{\gamma}^{[k]}$ is the k^{th} column of $\mathbf{A}^{-1}$ which is determined simply by Eq. (3.44).

3. Calculate the error vector entries $e^{[k]}$ using

$$e^{[k]} = \frac{\boldsymbol{\alpha}(k)}{\boldsymbol{\gamma}^{[k]}(k)} \tag{3.45}$$

   which is the ratio of the k^{th} elements of the two vectors.

4. Evaluate the vector $\boldsymbol{\alpha}^{[k]}$ defined as:

$$\boldsymbol{\alpha}^{[k]} = \boldsymbol{\alpha} - e^{[k]}\boldsymbol{\gamma}^{[k]} \tag{3.46}$$

   The $n_o \times 1$ vector $\boldsymbol{\alpha}^{[k]}$ is the expansion coefficient in which the k^{th} observation contribution is removed.

5. Calculate the interpolation point values using

$$\widetilde{f}_j^{[k]} = \mathbf{A}_I \boldsymbol{\alpha}^{[k]}; \qquad \text{for } j = 1, 2, \cdots n_I \tag{3.47}$$

   where, $\mathbf{A}_I$ is the design matrix corresponding to the interpolation points.

6. Iterate steps $2 - 5$ for $k = 1, 2, \cdots, n_o$

7. Compute the interpolation error at the interpolation point using Eq. (3.42).

8. Evaluate the object function Eq. (3.41)

From a computational point of view, this algorithm is very efficient since the normal matrix is inverted just once and the residual quantities are somehow the by-products of the original linear system of

equations. Obviously, steps 4 and 5 will then be ignored if the basic object function rather than the modified one is used.
We have developed a realistic object function which can be utilized for optimization of the shape parameter. As it will be will seen in section (3.3.1), optimization of single variable function is the easiest case of optimization. This, also can be another hidden advantage of RBFs.

3.2.3 Optimal polynomial approximation

In general, assume $\widetilde{f}^{PB}$ as the polynomial-based approximating function either harmonic, biharmonic or rational. For simplicity, we also assume that the equations of the type functional constraints have been inserted into the observation equations. Hence, the problem is how to perform the one-leave-out technique without repeating the estimation process. In other words, the problem is how to eliminate the contribution of a single observation from the estimated unknown vector.
Based on the detailed formulation given in Appendix B, the following computational scheme is proposed for a single-stage estimation procedure:

1. Estimate the unknown coefficients based on all n_o observations,

$$\widehat{\boldsymbol{\xi}} = \mathbf{N}^{-1}\mathbf{u}. \tag{3.48}$$

2. Compute the k^{th} element's contribution to $\mathbf{N}$ and $\mathbf{u}$:

$$\mathbf{N}_k = \mathbf{A}_k^T\mathbf{A}_k \tag{3.49}$$

$$\mathbf{u}_k = \mathbf{A}_k^T s_k, \tag{3.50}$$

   where, $\mathbf{A}_k$ is the k^{th} row of the design matrix $\mathbf{A}$.

3. Calculate the k^{th} observation contribution to the estimated unknown vector,

$$\widehat{\Delta\boldsymbol{\xi}} = \left(\mathbf{I}_{n_o\times n_o} + \mathbf{N}^{-1}\mathbf{N}_k\right)\mathbf{N}^{-1}\left(\mathbf{u}_k - \mathbf{N}\,\widehat{\boldsymbol{\xi}}\right). \tag{3.51}$$

4. Evaluate the unknown vector in the absence of the k^{th} observation,

$$\hat{\boldsymbol{\xi}}^{[k]} = \widehat{\boldsymbol{\xi}} - \widehat{\Delta\boldsymbol{\xi}}. \tag{3.52}$$

5. Calculate the approximation error at the observation point, P_k

$$\hat{e}_k = \mathbf{A}_k\widehat{\boldsymbol{\xi}}^{[k]}. \tag{3.53}$$

6. Evaluate the approximating function $\widetilde{f}$ at the grid points,

$$\widetilde{f}_j^{[k]} = \widetilde{f}(\hat{\boldsymbol{\xi}}^{[k]}, \mathbf{x}_j); \qquad \text{for } j = 1, 2, \cdots, n_I \tag{3.54}$$

7. Iterate steps $2-6$ for $k = 1, 2, \cdots, n_o$.

8. Calculate the approximation error at the grid point using Eq. (3.42).

9. Evaluate the object function Eq. (3.41).

As seen, the design matrix and the known vector of the normal equations are computed once and the rest of computation is performed based on the already derived quantities ($\mathbf{A}$, $\widehat{\boldsymbol{\xi}}$ and $\cdots$).
If the basic form of the object function Eq. (3.40) is used, the single-step estimation algorithm can be modified accordingly (ignoring steps 6 and 8).
Although optimization of the RBFs interpolation and the approximation problems are based on the minimization of an identical object function, they are conceptually different. In RBF interpolation a definite number of translates of a single-base function is used and the object function should be

minimized by an appropriate choice of the shape parameter. Therefore, the optimization problem is a single variable optimization which can easily be treated by either *local* or *global* numerical optimization methods, while the optimal approximation should be obtained by selecting the most suitable polynomial basis. Because of this basic difference, a particular optimization scheme is required. This will be discussed in section (3.3.2).
Nevertheless, considering the minimization of the object function (Eq. 3.41) rather than the norm of residuals for a definit set of basis functions leads to an optimization problem which cannot be solved by the classical least squares approach. Thus, problem should be solved by either local or global numerical multi-variable optimization techniques.

3.3 Numerical optimization

As discussed in the previous section, the error function $e(\mathbf{x})$, in its basic or modified form should be minimized by an appropriate selection of one or more independent variables. In other words, we want to find the value of those variables where the error function takes up a minimum value. We can then evaluate the underlying function at any interesting point based on the corresponding value of the variable(s).
A minimum (being the lowest point) can be either the lowest in a finite neighborhood *(local minimum)* or truly the minimum value of the function *(global minimum)*. The search methods are divided into the local and global minimization techniques correspondingly. We have to use the global method for minimization of the function since the truly minimum of the error function is of particular interest. However, finding a global minimum compared to the local one is, in general, a very difficult task. Nevertheless, the local minimum can also yield the global minimum either if the minimum is unique or if the neighborhood is correctly defined to include the global minimum. In the latter, enough knowledge of the physical process about the object function should be available which is not the case for the error function. Consequently, by implementing the local techniques for minimization, the error function is assumed unimodal.
Simplicity of the local methods as well as the published literature on the topic (Hickernell and Hon, 1998; Rippa, 1999; Trahan and Wyatt, 2003) are reasons to begin with the local method although the correctness of the underlying assumption is not fully guaranteed.

3.3.1 Golden section search technique (local search method)

Nearly all classical optimization methods can be utilized for local minimization. In general, the methods using the derivatives are somewhat more powerful than those using only the function (Press et al., 1988). Since there is no additional information on the function derivatives, we retain the algorithms using the function only; and since the problem is just a one-dimensional minimization problem, the *golden section search* method as the simplest technique is implemented.
The golden section search is a technique for finding the extremum (minimum or maximum) of a mathematical function, by successively narrowing the brackets by upper bounds and lower bounds. The idea is the same as the bisection method for solving nonlinear equations except it needs three points rather than two. Suppose we have three points a, b and c with $a < b < c$. If $f(a) > f(b)$ and $f(b) < f(c)$, the interval (a, c) is sure to contain a minimum of $f(x)$, and we say that the points (a, b, c) *bracket the minimum*. Figure 3.5 illustrates the technique for finding a minimum. The value of $f(x)$ has been evaluated at the three points: a, b, and c. Since $f(b)$ is smaller than either $f(a)$ or $f(c)$, it is clear that a minimum lies inside the interval from a to c.
The next step in the minimization process is to evaluate the function at a new value of d, namely d either between a and b or b and c. It is most efficient to choose d somewhere inside the largest interval, i.e. between a and b, then we evaluate $f(d)$. If $f(b) > f(d)$, then the new bracketing triple of points is (a, d, b); otherwise the new bracketing triple is (d, b, c). In all cases, the middle point of the new triple is the abscissa whose ordinate is the best minimum we have achieved so far; see Fig. (3.6). The process of narrowing the bracket is continued until the two outer points of the bracket triple get close enough to each other. In other words, the distance between those two points becomes negligible. By introducing new points and comparing the function values, we can systematically reduce the size of the interval bracketing the minimum which is supposed to be within the neighborhood $]a, b[$. The question is, how to choose the value of d?

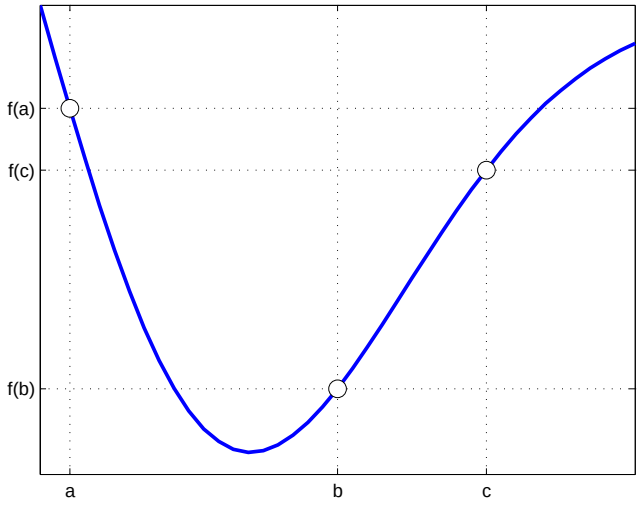
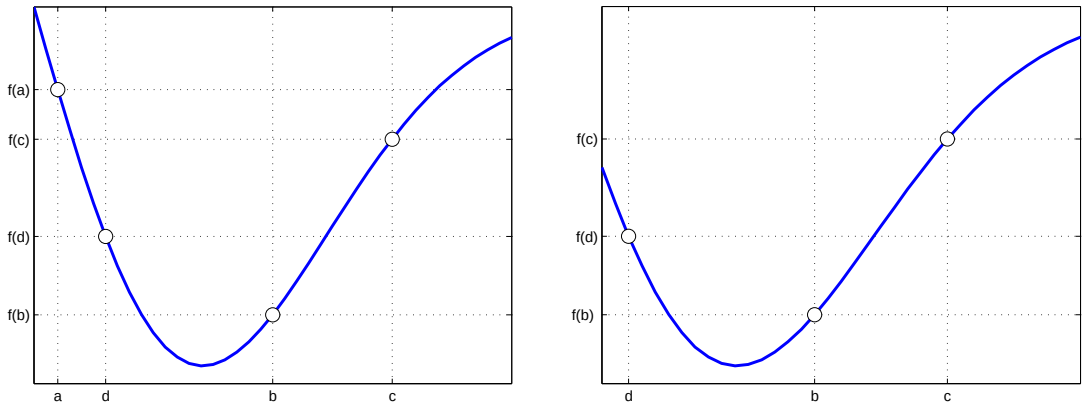


Figure 3.5: Bracketing a minimum



(a) A new point (b) A new bracket

Figure 3.6: Successive bracketing

The selection scheme is based on the idea of *golden ratio* which was first studied by ancient mathematicians due to its frequent appearance in geometry. Two quantities are said to be in the golden ratio, if *the whole is to the larger as the larger is to the smaller*, i.e. if:

$$\frac{c-a}{b-a} = \frac{b-a}{c-b}. \tag{3.55}$$

Equivalently, they are in the golden ratio if the ratio of the larger one to the smaller one equals the ratio of the smaller one to their difference:

$$\frac{b-a}{c-b} = \frac{c-b}{(b-a)-(c-b)}. \tag{3.56}$$

After multiplying the first equation by $\frac{b-a}{c-b}$ or the second equation by $\frac{(b-a)-(c-b)}{c-b}$, both of these equations are seen to be equivalent to:

$$\left(\frac{b-a}{c-b}\right)^2 - \frac{b-a}{c-b} = 1. \tag{3.57}$$

and hence $\frac{b-a}{c-b}$, which is typically symbolized by τ, is called the golden ratio:

$$\tau \doteq \frac{b-a}{c-b} = \frac{\sqrt{5}-1}{2} = 0.61803.$$

or equivalently,

$$\frac{b-a}{c-a} = \frac{\tau}{1+\tau} = 0.38197.$$

The relationship can be used similarly, the new point d should be selected such that, d should be the same fraction of the way from b to c as was b from a to c, i.e.,

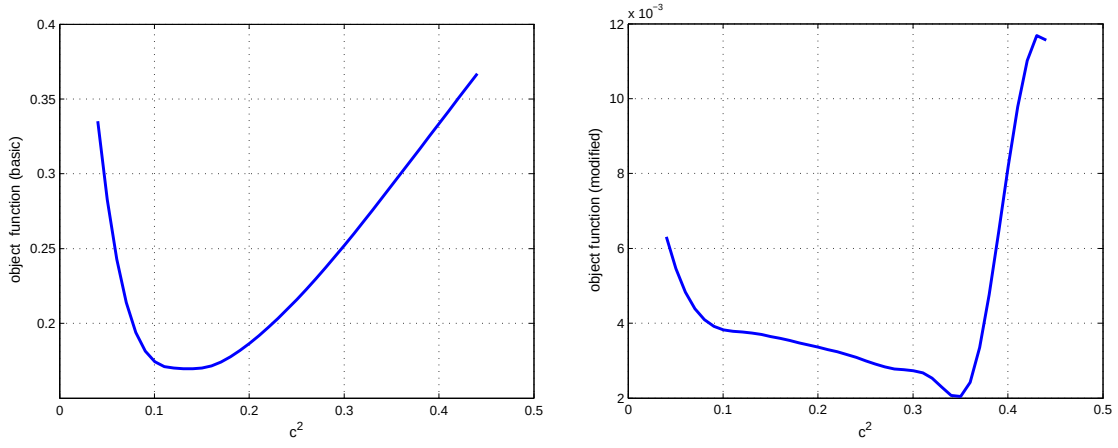
$$\frac{d-b}{c-b} = \frac{b-a}{c-a} = 0.38197. \tag{3.58}$$

Conclusively, at each iteration, the distance of a new point measuring from the central point of the triplet is a fraction 0.38197 to the larger of the two intervals.

If we start out with a bracketing triplet whose segments are not in the golden ratios, the procedure of choosing successive points at the golden mean point of the larger segment will quickly converge us to the proper, self-replicating ratios. The golden section search guarantees that each new function evaluation will (after self-replicating ratios have been achieved) bracket the minimum to an interval just 0.61803 times the size of the preceding interval (Press et al., 1988).

Besides the unimodal assumption, the method will work if we are able to bracket the minimum in the first place. Therefore, the initial bracketing is the essential part of the minimization method. From the theoretical point of view, there is no general rule for performing the initial bracketing perfectly. Nevertheless, we can start bracketing the minimum with an initial interval and proceed to find a reasonable triplet points. For instance, if two initial distinct points a and b define the downhill direction of the object function ($f(b) < f(a)$ for $a < b$), the interval is expanded by a dilatation factor (i.e., golden ratio). In other words, the first two initial points are the left and the middle points of the bracketing triplet and the process should be continued to find a high third point (Press et al., 1988).

To show the method's performance, Franke's test function i.e., Fig. (3.4) should be recalled. To show superiority of the modified object function (Eq. 3.41) over the basic one (Eq. 3.40), both functions were implemented and the achieved results depicted in Fig. (3.7). It can be seen in Fig. (3.7), the



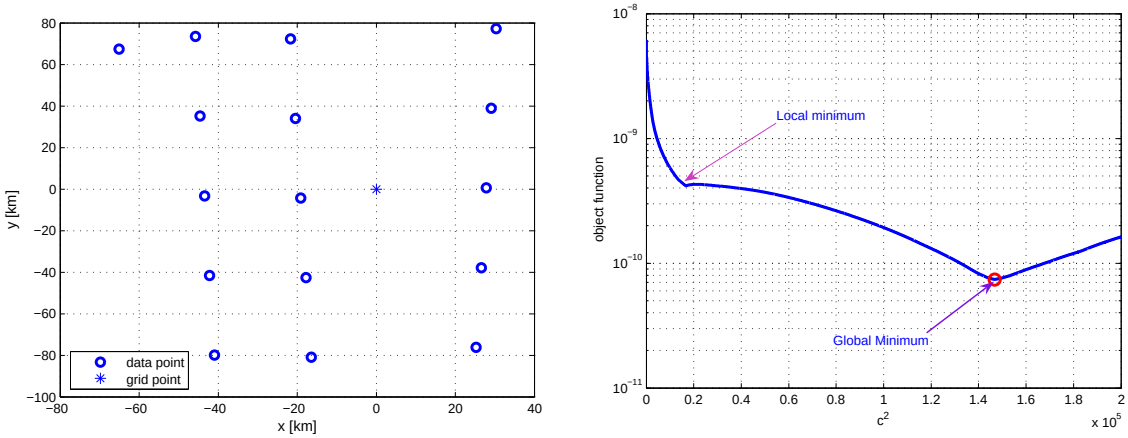
(a) Basic object function (Eq. 3.40)

(b) Modified object function (Eq. 3.41)

Figure 3.7: The basic and modified object function of Franke's test function

modified object function generally leads to a better solution. However, comparison of Figs. (3.4) and (3.7) indicates some similarities between the basic object function and the original error function. Nevertheless, we utilize the modified object function since it yields a comparatively better solution.

In order to calculate an initial value, we use Eq. (3.34). For the given data set, the shape parameter $c^2 = 0.4033$ is obtained which is obviously not the optimal solution. By feeding for instance c^2 and 0.1 c^2 and correspondingly 0.0063 and 0.0087 as the object function's values into the bracketing routine, we obtain $a = -0.54701$, $b = 0.040333$ and $c = 0.40333$ as the initial bracketing triplet of points. Initialization of the golden section search method with the first minimum bracketing points leads to an optimal solution. Implementation of the modified object function leads to $c^2 = 0.3549$ as the optimal shape parameter, whereas the parameter is about 0.1369 for the basic object function. One can see the achieved results validating simply by considering Fig. (3.7).
Luckily, the object function of the preceding example was a unimodal function and consequently could be resolved by the golden section search method. The method or any other similar techniques work as long as the unimodality condition is fulfilled. However, in reality, there is no way to guarantee unimodality of the object function. Therefore, we cannot expect to obtain a truly minimum solution by implementation of the local algorithms.
To clarify the matter, consider a set of the along-track observations T_{xx}, in a polar orbit which are horizontally distributed around the grid point $(0, 0, 0)$ as shown in Fig. (3.8). As it is seen, the data points are nearly distributed uniformly. Similar to the previous example, the object function is computed for a wide range of the shape parameter. As seen, the shape function has two minima 16700 and 146814.562. The shape parameter obtained by Eq. (3.34) is 8626.744 where it is closer



(a) Data points 2D distribution

(b) Object function

Figure 3.8: Exitance of the local and global minimum

to the local minimum rather than the global one. Therefore, the object function is not a unimodal one. Consequently, the local minimum search method might be trapped in the local minimum which unfortunately has a higher chance due to closeness of the rough estimate to the local one.
This example aimed to clarify how the optimization process could be violated due to lack of information on the initial guess of the shape parameter. Therefore, we have to employ a global search method which can easily bridge the local minima if there is any. Herein, we implement the *Genetic algorithms* which will be discussed in detail in the following section.

3.3.2 Genetic algorithms (global optimization method)

As already stated, the local optimization methods are based on the concept of stepping downhill from an arbitrary initial point. They just differ in deciding how to proceed but do not improve the algorithm's ability to find the global minimum instead of the local one.
Besides the *genetic algorithm* pioneered by Holland (1975), *simulated annealing* (Kirkpatrick et al., 1983), *particle swarm optimization* (Parsopoulos and Vrahatis, 2002), *ant colony optimization* (Dorigo and Gambardella, 1997), and *evolutionary algorithms* (Schwefel, 1993) are some outstanding methods which have been used for seeking the global minimum. These methods create new points in the feasible solution space by applying operators to the current solution and statistically moving toward more

optimal solution in the feasible search space. The methods are intelligent surveying the large search space for just a finite search. Since the methods do not require the underlying function's derivatives, they can thus be implemented on the non-continuous object function and discrete variables.
To find the global minimum rather than the local one, two techniques must be used to make the optimization method efficient: *exploration* to investigate new and unknown areas in the search space; and *exploitation* to make use of knowledge found at already visited points to help find better points. The first specification helps the algorithm jump out of a local minimum, whereas the second leads to an efficient search. These two requirements are contradictory which can be perfectly fulfilled simultaneously. For instance, the simulated annealing method only deals with a single candidate at each time and so does not build up an overall picture of the search space. Consequently, no information is saved from the already visited points to guide the selection of new points (Whitely, 1993).
Genetic algorithms (GAs) are adaptive search algorithms which are based on the genetic process of biological organism and the evolutionary ideas of natural selection. As such, they represent an intelligent exploitation of random search used to optimize an underlying function. Although randomized, GAs by no means are random, instead they exploit the previously visited points information to direct the search into regions with possibly better solution within the search space. In other words, GAs seek the optimum solution as natural populations evolve over many generations according to the principle of natural selection and *survival of the fittest*, first clearly stated by by *Charles Darwin* in *The Origin of species*.
The method was developed by Holland (1975) in the 1960s and 1970s and finally popularized by one of his students, David E. Goldberg. Holland's original idea was summarized in his book (Goldberg, 1989). Jong (1975) was the first to demonstrate the usefulness of the GAs for function optimization and did the first concentrated research on the selection of GAs parameters. Since then, many more researchers have carried out researches and hence lots of books and articles have been published. The followings are the main reasons why GAs are of particular interest to engineers and scientists (Haupt and Haupt, 2004):

- Optimize with continuous or discrete variables,
- Do not require derivatives information,
- Enable simultaneous search from a wide sampling of the feasible solutions,
- Deal with a large number of variables,
- Are well suited for parallel computers,
- Optimize variables with extremely complex object functions and can easily jump out of a local minimum,
- Provide a list of optimum solutions instead of a single solution,
- Provide possibility to carry out the optimization with the encoded variables, and
- Provide ability to work with simulated data, real observations, or analytical functions.

Despite of these inevitable advantages of GAs over the traditional ones it may not be the best method for every problem. For instance, for the unimodal functions of a few variables, the traditional methods outperform the GAs. Moreover, in real time applications, ant colony optimization algorithms have an advantage over the GAs since the ant colony algorithm can be run continuously and adapted to changes in real time.
What follows is a summary of issues involved in the GAs field. Herein, we will just present a general overview on the basic principles of the method. Good introduction to the topic, of various sizes, include books (e.g. Jong, 1975; Goldberg, 1989; Michalewicz, 1996; Haupt and Haupt, 2004), reports (e.g. Whitely, 1993), and introductory papers such as (e.g. Beasley et al., 1993; Busetti).

Genetic algorithms overview

GAs are based on the genetic processes of biological organisms. Over many generations, natural populations evolve according to the principles of natural selection and *survival of fitness*. By mimicking this process, GAs can evolve solutions to real world problem if they are suitably encoded.
In nature, since the resources are limited, the individuals in a population are in inevitable competition for food, water and shelter and even to attract a mate. The fittest individuals which are most successful in surviving and attracting a mate will have more offsprings compared the poorly performed individuals. Therefore, the genes from the individuals with higher fitness will spread to an increasing number of individuals in each successive generation. This evolutionary process may combine good characteristics from different ancestors and generate *superfit* offspring with high compatibility even higher than that of the parents.
As mentioned earlier, a suitable representation of the problem is the primary step of the GAs implementation. It begins with the definition of the object function or namely the *fitness function* as a function (s), of the optimization variables. The variables (known as *genes*) are joined together to form a string of values, often called *chromosome*. For instance, the shape parameter is the only gene of the fitness function of the MQ interpolation, while the number of genes in case of the polynomial fitting equals the number of the polynomial basis u_m. Consequently, MQ chromosomes consist of a single gene, whereas those in the polynomial approximation have been made of u_m genes. Usually a *genotype* is referred to the set of variables represented by a particular chromosome. The genotype contains the information required to construct an organism which is referred to as *phenotype*. For instance, in MQ interpolation, the optimal shape parameter is the genotype, while the optimum solution is the phenotype.
Natural behavior is directly followed in GAs. They work with a population of individuals where each individual represents a possible solution to the underlying problem. A fitness score is assigned to each individual according to its fitness. The highly fit individuals are given opportunity to *reproduce*, by *cross breeding* with other individuals in the population. Consequently, a set of new individuals as *offspring* are produced. Also, the lesser fit individuals of the population have smaller chance to reproduce, and more likely *die out*.
The selected highly fit individuals are reproduced as the new generation of possible solutions which contain a higher proportion of the characteristics by the good members of the previous generation. In this way, good characteristics of the individuals are mixed, exchanged and gradually spread throughout the individuals from generation to generation. The consequence of the more fit individuals' mating is exploring the most promising area of the search space of the underlying problem. The *evolutionary process* leads to an optimal solution to the problem if the GA has been coded appropriately.
An optimization problem using GAs can be coded as either a *binary* or a *continuous* problem. The same rules are used for genetic recombination and natural selection in both representations. The only difference is the genes' representation. In the *binary genetic algorithms* (BGAs), the genes are represented as encoded binary strings and computation will be performed using the binary strings. The *continuous genetic algorithms* (CGAs) use the genes as they are. Depending on the application, one may suit better than the other. For instance, the optimum shape parameter calculation problem can be solved either by the BGAs or CGAs while for the optimum basis optimum basis functions selection, BGAs are the only applicable algorithms. Therefore, we are mainly focused on the BGAs and give a short description on the differences to clarify the CGAs. Much of the information given here is from Haupt and Haupt (2004).

The binary genetic algorithms

As in every optimization method, the GAs start with definition of the genes or chromosomes, fitness function and the fitness [4]. Assume we wish to optimize a function of k variables $g(x_1, x_2, \cdot, x_k)$. Considering each of these variables as a gene, results in the chromosome as a k variable row vector:

$$\text{Chromosome} = [\ x_1 \quad x_2 \quad \cdots \quad x_k\ ]. \tag{3.59}$$

[4]The term fitness, which is used in the GAs literature, is referred to the output of the object function. The fitness as well as the fitness function are equivalent to the term cost and cost or object function. Although the term cost is specifically used for the maximization problem, one can easily adopt it for minimization problem. Herein, we will use fitness instead of cost.

Table 3.2: Sample examples of the 2D chromosomes in the binary and continuous form

Binary	Decimal
11111000011100001011101101 01	(0.9705, 0.1829)
01011000110101111111011100001	(0.3470, 0.9825)
10100100001100110101100000110	(0.6414, 0.8364)
00101011010011110011100110 10	(0.1691, 0.8063)

and correspondingly

$$\text{Fitness function} \;=\; g\,(\text{Chromosome})\,. \tag{3.60}$$

Suppose further that each variable x_i can take a value from a domain $\mathbb{D}_i = [x_i^{\text{lb}} \;\; x_i^{\text{hb}}] \in \mathbb{R}$. We wish to optimize the function g with some required precision; d decimal digits.

In order to solve the problem using binary representation, the continuous variables should be converted into binary and vice versa. Quantization samples a continuous range of values and categorizes the sample into non-overlapping subranges. It is clear that to achieve the requested precision, in each domain $\mathbb{D}_i$ should be cut into $(x_i^{\text{hb}} - x_i^{\text{lb}}) \cdot 10^d$ equal size ranges. Thus, the smallest integer m_i which reads (Michalewicz, 1996):

$$(x_i^{\text{hb}} \;-\; x_i^{\text{lb}}) \cdot 10^d \leq 2^{m_i} - 1$$

is the length of the binary string which clearly satisfies the precision requirement. Equivalently,

$$m_i \;=\; \text{ceil}\left[\frac{\log\left(10^d \Delta_i + 1\right)}{\log 2}\right], \tag{3.61}$$

with $\Delta_i \;=\; x_i^{hb} \;-\; x_i^{lb}$ and *ceil* as the function which rounds the value to the next highest integer. Therefore, each chromosome has a binary string of length $n_{\text{chr}} \;=\; \sum_{i=1}^{k} m_i$.

For decoding a binary string into the corresponding decimal number, the following formula can be used

$$x_i \;=\; x_i^{lb} + \text{decimal}\{(\text{gene})_i\} \cdot \frac{\Delta_i}{2^{m_i} - 1} \tag{3.62}$$

To clarify the process, consider the Frnake test function given in Eq. (3.36). It is a two variable x, y where $x, y \in [0,\, 1]$. Let us assume that the required precision is four decimal places for each variable. The domain of both variables has a length 1. Therefore, the domain should be divided into at least 10000 equal size ranges. This means that each of these variables can be represented with the required accuracy with a gene of length 14,

$$2^{13} < 10000 \leq 2^{14}.$$

Consequently, the total length of a chromosome (solution vector) is then $2 \times 14 = 28$ bits; the first 14 bits for x and the remaining 14 for y. Table (3.2) shows some randomly generated examples of chromosomes in the binary as well as in the continuous forms. It is also interesting to note here that the GAs work with the binary encoding while the fitness function often requires continuous form of the individuals.

Besides the fitness and fitness function definitions, the GAs' parameters also should be defined. Depending on the structure of the algorithm, different parameters are considered, *size of the initial population, selection and mutation rate* and *convergence criteria* to name a few. The performance of the GAs is closely dependent on the setting of these parameters. Jong (1975) was the first who thoroughly studied the influence of the parameter selection on the GAs' performance. We will describe each of these parameters as we proceed.

The initial population is the initial step of the GAs. The population consists of n_{pop} randomly generated individuals represented by n_{pop} chromosomes of size n_{chr} each. In other words, the initial population is a matrix of size $N_{\text{pop}} \times N_{\text{chr}}$ whose entries are either 0 or 1. For instance, a randomly generated initial population for Franke's test function given in Eq. (3.36) and their respective fitness are tabulated in Table 3.3. The initial population was considered of size 16 with the individuals of size 28, 14 for each.

Table 3.3: The initial individuals and their respective fitness

chromosome	x	y	Fitness
11110110100110111011110110011	0.96332	0.93286	0.00019
10001000100101111001010110	0.53354	0.89593	−0.03358
10001110110110100101101011	0.55802	0.5886	0.05821
00001001011000101100111111111	0.03662	0.70311	0.00682
00100011001101011100111011111	0.13752	0.45114	0.27686
11000101110100101001100000101	0.77275	0.64878	0.06857
11101000000011000101010000110	0.90649	0.08240	0.19991
01110110011001010001111110110	0.46249	0.28066	0.40617
00001000101001111000010111110	0.03375	0.88079	0.00031
11011111011110100101001001111	0.87298	0.58054	0.12388
11011101000101011101111111010	0.86364	0.46841	0.30795
00011100110011100010011010010	0.11249	0.53775	0.10111
10010110011110001100011111010	0.58780	0.19496	0.36932
00111000010101001010111001	*0.22005*	*0.16755*	*1.04000*
00010010010011011110010110000	0.07147	0.47415	0.17522
10000011001110101011111101110	0.51260	0.68644	−0.05619

As seen in Table (3.3), the most fitted chromosome with the fitness of about 1.04000 is on the top. Accidentally, the chromosome is very close to the global maximum of the test function, see Fig. (3.9).

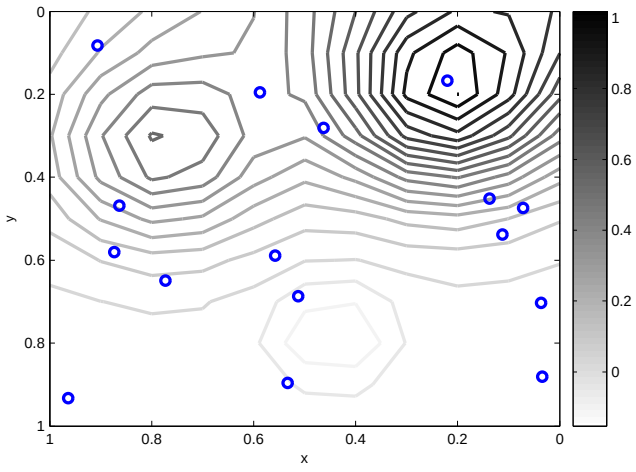


Figure 3.9: The initial individuals overlayed on Franke's test function

It should be noted here that the initial population should be large enough to prevent *premature convergence* of the GAs. On the other hand, the bigger the number the slower the process converges. However, the more the optimization parameters the larger the required population. Jong (1975) empirically showed that for most optimization problems, the initial population of size $50-100$ could lead to the optimum solution. Furthermore, Schaffer et al. (1989) showed a good performance with the initial population sizes in the range $20-30$ by implementing of the GAs on ten different functions. Survival of the fittest or selection of the best individuals is the next step in the GAs to form the *mating pool*. It can also be interpreted as selecting the best for continuation of the process and discarding the individuals with lower fitness to make room for the new offspring.
The manner in which individuals are selected for the mating pool is determined by the *reproduction operator* (Telfar, 1994). There are different methods for implementing this operator,but for correct convergence it must be performed carefully. Otherwise, the process will converge slowly or lead to a

Table 3.4: Roulette wheel selection

chromosome	q_i	r_i	selected chromosome
1	0.05252	0.95013	15
2	0.10326	0.23114	5
3	0.15883	0.60684	11
4	0.21169	0.48598	9
5	0.27874	0.89130	15
6	0.33485	0.76210	13
7	0.39785	0.45647	8
8	0.47169	0.01850	1
9	0.52421	0.82141	14
10	0.58322	0.44470	8
11	0.65190	0.61543	11
12	0.70972	0.79194	14
13	0.78162	0.92181	15
14	0.88873	0.73821	13
15	0.95044	0.17627	4
16	1	0.40571	8

premature solution. *Roulette wheel* and *tournament selection* are the most frequently used techniques for most GAs. We have numerically implemented these standard methods. The basic principles of the methods are:

Roulette wheel The selection of the individuals for the mating pool is based on the cumulative probability q_i which is defined based on the relative fitness. Since in general the fitness function values could be either negative or positive, an arbitrary constant value C can be added to the original fitness values to obtain positive fitness values. Consider f_i as the modified fitness of the i^{th} individual. Then,

$$q_i = \frac{1}{\sum_{j=1}^{n_{\text{pop}}} f_j} \sum_{k=1}^{i} f_k. \tag{3.63}$$

The cumulative probabilities q_i are used in selecting the chromosome. A vector of random numbers r_i of size n_{pop} between 0 and 1 is generated. The first chromosomes with cumulative probabilities greater than the random numbers are selected for the mating pool. This type of selection is equivalent to the spinning of the roulette wheel n_{pop} times and each time select a single chromosome for the mating pool (Michalewicz, 1996).

Let us consider the initial population and their corresponding fitness values. We can simply derive positive fitness by adding on a unit to all values. The modified fitness values as well as the cumulative probabilities for the population are given in Table (3.4).

As seen in the Table, chromosomes with higher fitness have been selected even twice throughout the selection procedure. It means that they have a higher chance to mate and survive.

In this method that the individuals are weighed based on their fitness, is usually called *fitness-based roulette wheel*. Weighting is also possible using the rank of the individual within the population. This alternative method is namely called *rank-based roulette wheel* (Haupt and Haupt, 2004).

Tournament selection From the initial generation, groups of individuals of size n_T are selected randomly and the fittest individual in the tournament is placed in the mating pool. All selected individuals are returned to the population and another tournament is carried out. The process is repeated until the mating tool is full. It can be shown that a binary tournament which leads to nearly the same results is achieved by ranking procedure (Haupt and Haupt, 2004).

Let us consider the previous example. Performing a tournament selection with $n_T = 3$ for instance results in a mating pool with individuals listed in Table (3.5). Similar to the roulette

Table 3.5: Tournament selection

Tournament	Group	Fittest
1	(15, 2, 5)	5
2	(8, 4, 6)	8
3	(16, 12, 13)	13
4	(13, 15, 5)	13
5	(4, 1, 10)	10
6	(16, 2, 14)	14
7	(10, 3, 1)	10
8	(6, 7, 16)	7
9	(14, 3, 8)	14
10	(16, 15, 9)	15
11	(6, 14, 15)	14
12	(2, 10, 11)	11
13	(13, 15, 1)	13
14	(16, 3, 10)	10
15	(2, 12, 10)	10
16	(12, 11, 1)	11

wheel method, the individuals with higher fitness have been selected more than once by tournament selection method. For instance, chromosome 14 has emerged three times whereas the weak chromosomes like 1 or 4 or 9 have died out.

All selection methods including the aforementioned methods have advantages and disadvantages and any individual method can be made to produce similar results to other ones. However, as long as the selection method is implemented carefully, any of the above methods will give reasonable results.

Mating

Now, the mating pool of size n_{pop} is obtained by implementation of the selection techniques. *Mating* is the creation of one or more offsprings from the parents appeared in the mating pool. Throughout this step namely called *crossover*, genetic information from two parents is combined to produce an offspring with characteristic from both. Parents are usually two individuals which are chosen at random from the mating pool. Like in the selection process, there are different methods for performing the crossover. Among them, *single-point* or *simple* crossover is the standard example of a crossover technique. In this simple reproduction method, each parent has its chromosome cut randomly at a fixed position and the tail chromosome segments are swapped to generate two new offsprings, see Fig. (3.10). First, the binary code to the left of the crossover point of the parents goes to the left of that crossover point to the offsprings. Second, parent 1 and 2 pass their binary code to the right of the crossover point to offsprings 2 and 1 respectively. To remain true to the evolutionary analogy, there should exist a possibility for any given generation to compete with the future generations. Therefore, the crossover probability p_c, usually is a number between 0.6 and 1.0 (Beasley et al., 1993).
Consider the earlier example. Assuming $p_c = 0.6$ leads to recombination of genetic information of approximately 10 chromosomes out of 16. We can select the candidate for crossover randomly. For instance, the individuals 12, 7, 11, 9, 8, 6, 13, 4, 5, 16 were selected randomly as parents which were equivalent to the individuals 11, 10, 14, 14, 7, 14, 13, 13, 10, 11 in the original population. A random pairing of the potential parents yields $(14, 7)$, $(13, 11)$, $(11, 13)$, $(14, 14)$, and $(10, 10)$. Since only the pairs with distinct individuals can generate new offspring, we perform crossover operator only on the first two pairs. To find the crossover points, we generate 2 random numbers between 1 and the $N_{\text{chr}} = 28$ each of those corresponding to crossover point in each of 2 parents. Table (3.6) shows randomly selected pairs and their respective randomly chosen crossover points. Therefore, from the mating pool the 8^{th}, 11^{th}, 12^{th}, and 13^{th} individuals will be replaced with the generated offspring.

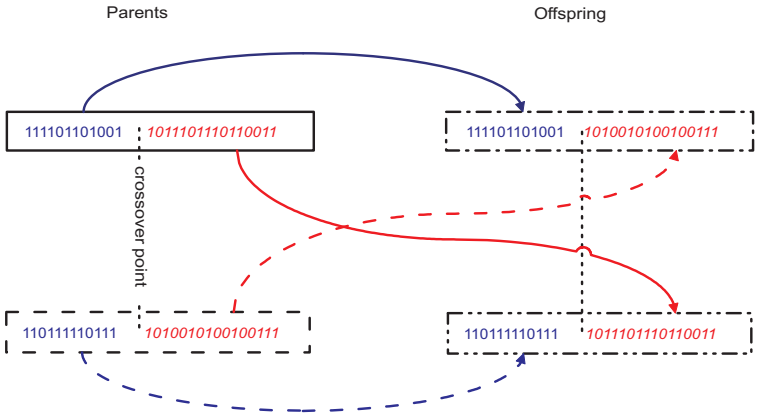


Figure 3.10: Single-point crossover

Table 3.6: Pairing and mating process of simple crossover

Pairs	Crossover point	offspring
(14, 7)	24	00111000010101001010110110
		11101000000011000101010010 01
(13, 11)	8	10010110000101011101111110 10
		11011101011110001100011110 10

Simple crossover is usually quite effective, however in some cases more sophisticated algorithms have outperformed the single-point crossover. A discussion of other alternatives of the crossover operator is given by Reeves (1993).
The simplest alternative is obtained by a direct expansion of the single-point crossover. The method which is called *multi-point* crossover can be developed by partitioning the parents' chromosomes into k segments rather than 2. Eshelman et al. (1989) investigation showed that multi-point crossover tends to help for a better exchange of the genetic information compared to the single-point crossover. The *generalized* or *uniform* crossover proposed by Syswerdar (1989) as the generalized form of the multi-point method, *string-of-change* operator proposed independently by Booker (1987) and Fairley (1991) and *threshold* crossover suggested by Sirag and Weisser (1987) are the alternatives which could also be utilized for mating. Some are generally applicable while the others are not appropriate for certain problems.

Mutation

In order to maintain diversification in the population, *mutation* is the secondary reproductive operator which is applied after crossover. It provides a small amount of random search in the GAs and keeps them from being converging prematurely. *Simple mutation* as the simplest form of mutation alters a gene in the list of offspring with some small probability p_m. In the BGAs, single point mutation amounts to flipping a randomly selected gene's value from 0 to 1 and vice versa. Figure (3.11) schematically shows a possible outcome of applying single point mutation to a binary individual.
Mühlenbein (1992) and Bäck (1993) concluded, independently of each other, that the optimal mutation rate in the case of uni-modal problem is inversely proportional to the length of chromosomes for a fixed mutation rate throughout the run. Furthermore, Bäck (1993) showed that mutation probability other than $p_m = \frac{1}{n_{\text{chr}}}$ for the multi-modal functions is more adequate.
Permutation mutation includes a set of alternative mutation operators. It includes:

- *Position-based* operator: removes gene at location i from present position and reinserts at location j,
- *swap mutation*: swaps two randomly selected genes at locations i and j, and

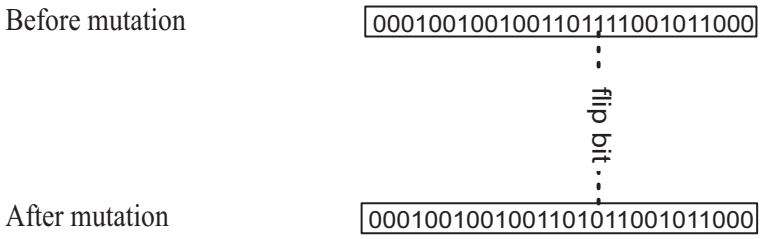


Figure 3.11: Single point mutation

- *Scramble mutation*: randomly rearranges the genes between two arbitrary loci on the chromosome.

The extended mutation operators are depicted in Fig. (3.12).

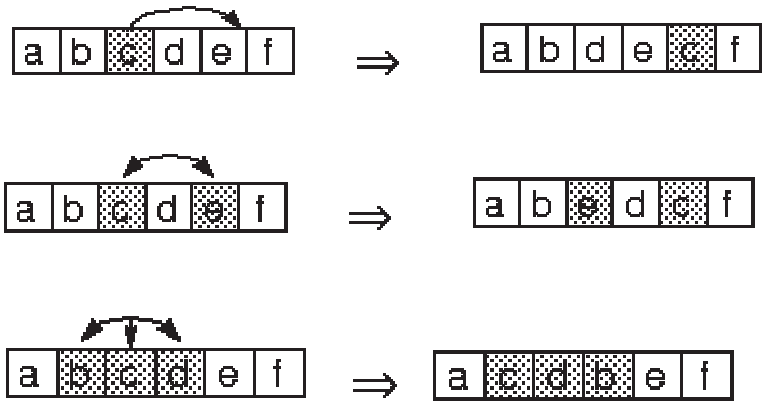


Figure 3.12: Permutation mutation; from top to bottom: the position-base, swap and scramble mutation (from Hondroudakis et al. (1996))

The mutation operator plays the secondary role in the new generation of reproduction. However, it would be very useful for exploring new and unknown areas. It is also interesting to note that the fittest individual is considered as *elite* solution will not be mutated due to *elitism* (Haupt and Haupt, 2004).
Let us consider the previous example. As seen in Table (3.6), four individuals $(8, 11, 12$ and $13)$ were replaced by the generated offspring. As shown in Table (3.3), the individual number 14 emerged twice in the mating pool can be considered. Therefore, either the 6^{th} or 9^{th} individual can be kept constant and the others amount to the mutation operator. Considering $n_{\text{chr}} = 28$ and population size $n_{\text{pop}} = 16$ results in just one mutated individual. The 7^{th} bit of the 8^{th} individual is the randomly selected candidate for mutation. The population after mutation or the so called *second generation* with their respective locations and fitness are tabulated in Table (3.7).
As it is inferred from Table (3.7), the process moves toward the maximum. For instance, there was no negative fitness whereas before we had two individuals with negative fitness. Although the maximum was the same, the average fitness increased from 0.1903 to 0.3926. The process tendency to climb up the hill is shown in Fig. (3.13).
Most of the initial individuals with lower fitness have died out and new individuals with better fitness have appeared.
A path through the components of the GAs is depicted in Fig. (3.14) which is similar to the flow chart given by Haupt and Haupt (2004).
In order to improve the solution, the process described is iterated. For instance, after 100 iterations, the global maximum of the function ($f_{\max} = 1.0414$) has been found at location $(0.21242, 0.17091)$. The process could also be stopped earlier since the fittest solution of the first iteration has not been too far from the final solution. For instance, the same setting of the GA leads to $f_{\max} = 1.0394$ after merely 50 iterations. The difference is about 0.2% which seems negligible.

Table 3.7: The second generation and the corresponding fitness

chromosome	x	y	fitness
00100011001101011100011011111	0.1375	0.4511	0.2769
01110110011001010001111110110	0.4625	0.2807	0.4062
10010110011110001100011111010	0.5878	0.1950	0.3693
10010110011110001100011111010	0.5878	0.1950	0.3693
11011111011110100101001001111	0.8730	0.5805	0.1239
00111000010101001010101110001	0.2200	0.1676	1.0400
11011111011110100101001001111	0.8730	0.5805	0.1239
00111010010101001010101110110	0.2279	0.1674	1.0368
00111000010101001010101110001	0.2200	0.1676	1.0400
00010010010011011111001011000	0.0715	0.4742	0.1752
11101000000011000101010010001	0.9065	0.0826	0.2000
10010110000101011101111111010	0.5863	0.4684	0.2015
11011101011110001100011111010	0.8652	0.1950	0.3635
11011111011110100101001001111	0.8730	0.5805	0.1239
11011111011110100101001001111	0.8730	0.5805	0.1239
11011101000101011101111111010	0.8636	0.4684	0.3079

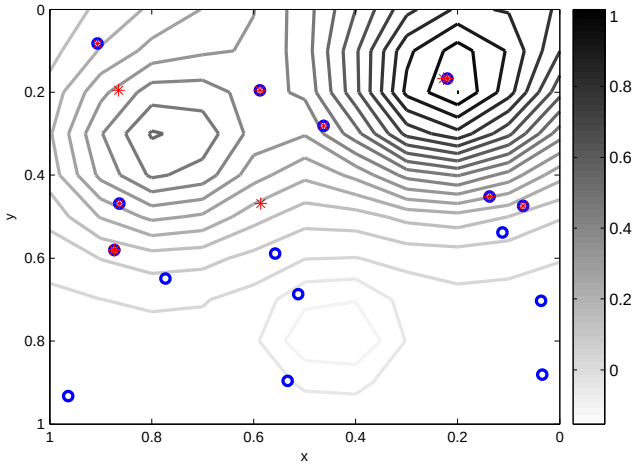


Figure 3.13: The initial and second generations overlayed on the test function (∘: initial population ∗: second generation)

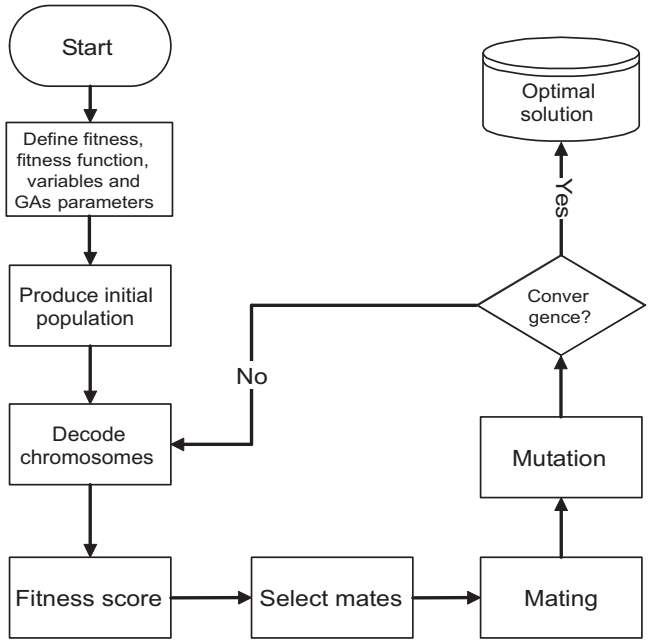


Figure 3.14: The Binary Genetic Algorithms (BGAs) flow chart

Figure (3.15) shows the successive improvements of the fitness function. As seen, there is nearly no improvement as the iterations higher than 50. Consequently, instead of iterating the process up to a maximum iteration, it can be run until it has improved.

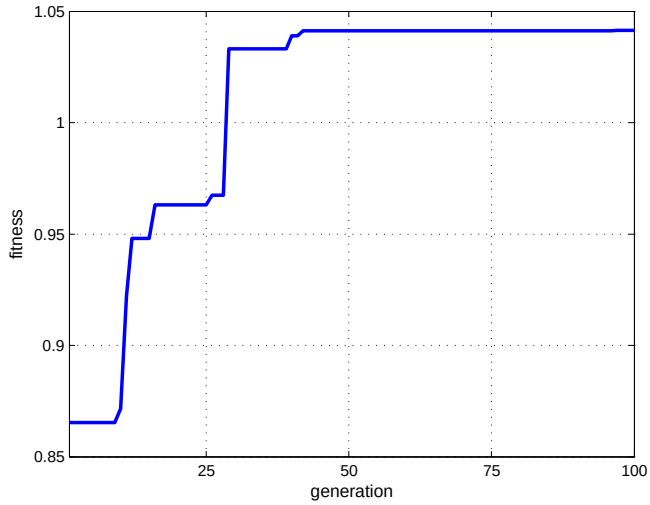


Figure 3.15: The test function fitness throughout 100 generations

It is interesting to note that there have been debates for decades as whether both crossover and mutation are necessary for reproduction. In general, it is good to have both crossover and mutation since they both have separate roles. In other words, there is cooperation *and* competition between them (Eiben and Smith, 2003). Crossover is *explorative* which makes a big jump to an area somewhere in between the location of two parents in the search space. In contrast, mutation is *exploitive* which creates random small diversions around the location of the parents. Furthermore, only crossover can combine information from two parents while mutation can introduce new *alleles*[5].
We have managed to describe the basic rules of the BGAs and clarify them with implementation of the

[5]An alternative form of a gene that is located at a specific position on a specific chromosome.

rules on the Franke's test function. We intentionally selected the function rather simple to illustrate the concept more easily. Implementation of the algorithm on a rather difficult problem, the RBF interpolation and polynomial-based approximation, will be presented in section (3.4.4).

The continuous genetic algorithms

The binary representation seems very simple, but it is not the only way to represent every problem. Without doubt the binary GAs are the well-suited codings for the naturally quantized variables. For instance, the polynomial basis functions selection by very nature is binary and we employ the representation in section (3.4.3). However, problems with continuous variables can be treated directly without quantization of the variable which leads to quantization error. Although the error can be minimized by utilizing longer chromosomes, it considerably slows down the algorithm. So, it is more logical to represent the variables with floating-point numbers. Although the binary representation is more relevant, the following advantages of the *continuous genetic algorithms* (CGAs) or also called *real-valued* GAs are inevitable (Haupt and Haupt, 2004)

- machine precision instead of the limited precision of the binary representation,
- less memory storage compared to the binary coding,
- faster iteration due to excluding the decoding prior to the scoring the fitness.

The basic structure of the CGAs is similar to that of the BGAs, except the decoding step which has been left out in the continuous representation, see Fig. (3.16).

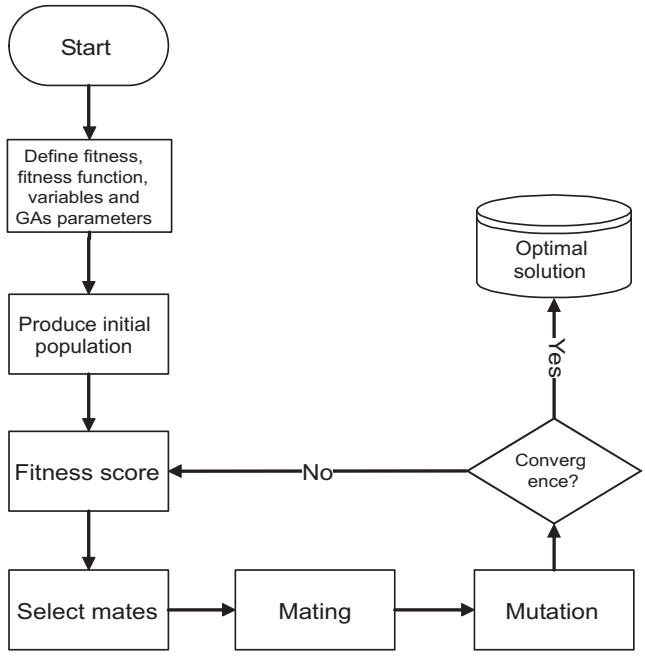


Figure 3.16: The Continuous Genetic Algorithms (CGAs) flow chart

Despite the similarities, there are some differences in the implementation of the predefined operators. In this section, we briefly describe how the continuous version of the operators can be derived from the already mentioned operators.

Genes and chromosomes: We no longer need to consider how many bits are necessary to accurately present a value. Instead, each variable is represented as a single real-valued gene. Correspondingly, each chromosome is an array of floating-point numbers of size equal to the number of genes (variables).

Initial population: The initial population consists of n_{pop} random individuals with uniform distribution of size k rather than n_{chr}. If the generated individuals are normalized, the unnormalized values are given by:

$$x_i = x_i^{\text{lb}} + \Delta_i \; \tilde{x}_i, \tag{3.64}$$

where, $\tilde{x}_i$ is the i^{th} normalized chromosome.

Selection and pairing: Since the algorithms implemented in the binary mode are performed based on the individuals' numbers and their respective fitness, they can directly be utilized in the continuous representation.

Mating: Besides the n-point and uniform crossover operators implemented in the binary representation, the *arithmetic recombination* is the alternative reproduction scheme in the continuous coding. It is defined as:

$$O \; = \; \kappa \; P_1 \; + \; (1 - \kappa) P_2, \tag{3.65}$$

where, κ is a uniform random number in $[0, 1]$ and $i, j \in \{1, 2, \cdots, n_{\text{pop}}\}$. The parameter κ can be constant, variable or selected randomly for each generation. Furthermore, the arithmetic crossover will be namely called

- *single arithmetic* if the operator is applied just on a randomly chosen gene. For instance, running the operator on two individuals P_1 and P_2 at a randomly selected locus ($m \leq k$) yields the following offspring

$$O_1 = [P_{1,1}, P_{1,2} \cdots P_{1,m-1}, \kappa P_{1,m} + (1 - \kappa) P_{2,m}, P_{1,m+1}, \cdots, P_{1,k}], \tag{3.66}$$

  and reverse for the other child,

$$O_2 = [P_{2,1}, P_{2,2} \cdots P_{2,m-1}, \kappa P_{2,m} + (1 - \kappa) P_{1,m}, P_{2,m+1}, \cdots, P_{2,k}]. \tag{3.67}$$

- *simple arithmetic* if all genes after the locus m are combined

$$O_1 = [P_{1,1}, P_{1,2} \cdots P_{1,m-1}, \kappa P_{1,m} + (1 - \kappa) P_{2,m}, \cdots, \kappa P_{1,k} + (1 - \kappa) P_{2,k}], \tag{3.68}$$

  and accordingly,

$$O_2 = [P_{j,1}, P_{2,2} \cdots P_{2,m-1}, \kappa P_{2,m} + (1 - \kappa) P_{1,m}, \cdots, \kappa P_{2,k} + (1 - \kappa) P_{1,k},]. \tag{3.69}$$

- *whole arithmetic* or *flat crossover* if the command is performed on the whole genes. This operator, which is the most commonly used arithmetic operator, is defined as

$$O_1 = \kappa P_1 \; + \; (1 - \kappa) P_2 \tag{3.70}$$

  and the reverse for the other offspring,

$$O_2 = (1 - \kappa) P_1 \; + \; \kappa P_2 \tag{3.71}$$

The flat crossover, pioneered by Radcliffe (1990), is a special case of a general form of crossover operators, namely called *blend crossover*. The method was proposed by Eshelman and Schaffer (1993) as an extended version of the flat crossover operator. In other words, the flat crossover just generates offspring on the search space within the parents, whereas the blend operator extends the segment on both sides determined by a user specified parameter α. Hence, the arithmetic recombination operator is redefined as:

$$O_1 = \gamma P_1 + (1 - \gamma) P_2, \tag{3.72}$$

and the reverse for the other

$$O_2 = (1 - \gamma) P_1 + \gamma P_2. \tag{3.73}$$

where, $\gamma \; = \; (1 + 2\alpha)\kappa - \alpha$. The new formulation which is denoted by BLX-α allows the offspring

Table 3.8: The initial population in the continuous representation

# chromosome	x	y	fitness
1	0.70150	0.01955	0.28289
2	0.05836	0.40792	0.29729
3	0.30008	0.61401	0.03444
4	0.04759	0.26234	0.62397
5	0.90832	0.84139	0.00199
6	0.28705	0.29142	*0.78417*
7	0.86285	0.22317	0.40117
8	0.38970	0.22593	0.64104
9	0.44265	0.28323	0.43710
10	0.41871	0.82921	−0.15218
11	0.97891	0.55251	0.08628
12	0.60167	0.50736	0.16688
13	0.19724	0.60303	0.04959
14	0.57567	0.66155	0.01201
15	0.96686	0.58662	0.06825
16	0.67585	0.92239	0.00000

to sweep a wider segment than the segment defined by the parents. One can easily show that

$$P_1 - \alpha(P_2 - P_1) \leq O_{1,2} \leq P_2 + \alpha(P_2 - P_1). \tag{3.74}$$

As Eq. (3.74) indicates, α plays an important role in the exploration of new areas in the search space. For instance, setting $\alpha = 0.0$ leads to exploitation of the area within the location of parents, while $\alpha = 0.5$ results in the exploration of the exterior space about half of the inner-space around each parent besides the exploitation on the interior space. Exploration of the outer-space may only be possible via mutation if the arithmetic crossover is implemented.

Mutation As already mentioned, the function of mutation is to keep diversity of a population and promote searching in the search space. *Uniform mutation* is the most commonly used mutation operator in the continuous representation. Implementation of this operator on the mating pool is equivalent to adding a vector of uniformly generated random numbers to a randomly selected individual with a mutation probability p_m. Adding normally distributed random numbers is the alternative mutation operator which is called *Gaussian mutation*.

Let us consider the previous example with the same setting of the GA parameters. The initial population consists of 16 real-valued chromosomes of size 2. The population with the respective fitness is given in Table (3.8). Compared to the binary representation, the initial population is relatively far from the fittest. Nevertheless, the population is distributed almost uniformly on the search area. To build up the mating pool, one of the implemented selection algorithms for the binary coding can be used. The Tournament selection, for instance, was applied and the selected individuals were listed in Table (3.9). We assume crossover probability $p_c = 0.6$ and carry out the flat crossover. Randomly selected parents are $16, 7, 4, 1, 8, 13, 3, 5, 11$ and 15 which are equivalent to the individuals $2, 4, 6, 7, 8$ and 12 in the original population. The selected parents randomly paring yields the following three pairs $(2, 6)$, $(4, 8)$, $(7, 12)$. The generated offspring are

$$\begin{cases} O_1 = \kappa_1 x_2 + (1-\kappa_1)x_6 \\ O_2 = (1-\kappa_1)x_2 + \kappa_1 x_6 \\ O_3 = \kappa_2 x_4 + (1-\kappa_2)x_8 \\ O_4 = (1-\kappa_2)x_4 + \kappa_2 x_8 \\ O_5 = \kappa_3 x_7 + (1-\kappa_3)x_{12} \\ O_6 = (1-\kappa_3)x_7 + \kappa_3 x_{12} \end{cases} \tag{3.75}$$

Table 3.9: Tournament selection in the continuous coding

Tournament	Group	Fittest
1	(9, 3, 6)	6
2	(11, 16, 9)	9
3	(12, 4, 13)	4
4	(10, 4, 8)	8
5	(14, 8, 1)	8
6	(1, 9, 4)	4
7	(14, 16, 12)	12
8	(16, 15, 2)	2
9	(5, 7, 6)	6
10	(14, 12, 2)	2
11	(7, 2, 4)	4
12	(12, 1, 13)	1
13	(13, 7, 11)	7
14	(16, 2, 10)	2
15	(15, 3, 8)	8
16	(12, 10, 15)	12

$\kappa_{1,2,3} = 0.23, 0.58, 0.85$ are three random numbers which have drawn randomly in $[0, 1]$. The achieved offspring are tabulated in Table (3.10). As inferred from the Table, performing the crossover generated

Table 3.10: The generated offspring using the flat crossover

Parent 1	Parent 2	Offspring	fitness
$(0.058365, 0.40792)$	$(0.28705, 0.29142)$	$(0.23445, 0.31821)$	0.7600
		$(0.11096, 0.38113)$	0.4472
$(0.047597, 0.26234)$	$(0.3897, 0.22593)$	$(0.19128, 0.24705)$	0.9550
		$(0.24602, 0.24123)$	0.9596
$(0.86285, 0.22317)$	$(0.60167, 0.50736)$	$(0.82367, 0.2658)$	0.4933
		$(0.64085, 0.46473)$	0.2677

super fit individuals whose fitness are considerably higher than their parents.
The first iteration will end up with performing the mutation process. Let us replace one of the individuals with a randomly generated chromosome. For instance, individual number 2 is mutated and replaced by $(0.3706, 0.5079)$. Consequently, the second generation will consist of 10 old individuals, 6 offspring and 1 mutated chromosome. The new population with respective fitness values are tabulated in Table (3.11). The individuals' fitness improved significantly. For instance, there are no negative values and also the mean fitness has been increased from 0.2334 to 0.5228 which indicates a meaningful improvement. Iterating the process, 100 times for instance, will hopefully yield the sought-after fittest chromosome. The continuous algorithm found $(0.2103, 0.1712)$ with the respective fitness 1.0414 as the fittest individual. Compared to the binary representation, the location is slightly different, however the fitness is the same. The emerging differences might be due to the presence of the quantization error.
Of course, there are many more sophisticated selection methods, crossover and mutation operators. Detailed discussion of all methods are far beyond the scope of this dissertation. The interested reader can refer to (e.g. Holland, 1975; Schwefel, 1993; Michalewicz, 1996; Eiben and Smith, 2003; Haupt

Table 3.11: The second generation

# chromosome	x	y	fitness
1	0.11096	0.38113	0.44723
2	0.37066	0.50787	0.12398
3	0.04759	0.26234	0.62397
4	0.24602	0.24123	*0.95965*
5	0.38970	0.22593	0.64104
6	0.19128	0.24705	0.95500
7	0.60167	0.50736	0.16688
8	0.23445	0.31821	0.76005
9	0.28705	0.29142	0.78417
10	0.05836	0.40792	0.29729
11	0.04759	0.26234	0.62397
12	0.70150	0.01955	0.28289
13	0.82367	0.26580	0.49330
14	0.05836	0.40792	0.29729
15	0.38970	0.22593	0.64104
16	0.64085	0.46473	0.26775

and Haupt, 2004).

Why do they work?

It is claimed that via the operations of reproduction, crossover and mutation over successive generations, the GAs will converge towards the global optimum. One question may occur to anyone's mind. *Why do these simple operators set up a fast, useful and robust technique for global optimization?*
As stated earlier, implementation of the operators on the best partial solutions combines the genetic information of previous generations in a randomized manner and generates new solutions. Explicitly, the mating pool is formed by selecting the individuals based on their fitness. Implicitly, the selected individuals in common, contain good information which can be recognized by comparing the structure of the selected chromosomes. These similarities could then be exploited to construct new solutions that will hopefully also share these good characteristics.
A *schema* describes a subset of string with similarities at certain loci. A schema is built by introducing a wild card symbol ($\star$) into the alphabet of genes. Schemata are basically pattern matching devices, matching particular strings from binary alphabet in every position, i.e., 1s match 1s, 0s match 0s and $\star$s match either. For instance, let us consider the strings and schemata of the length 9. The schema

$$00\star111\star01 \tag{3.76}$$

matches four strings:

$$\begin{aligned} &00\mathbf{0}111\mathbf{0}01\\ &00\mathbf{0}1111\mathbf{1}01\\ &00\mathbf{1}1110\mathbf{0}1\\ &00\mathbf{1}1111\mathbf{1}01 \end{aligned} \tag{3.77}$$

The *schema theorem* proposed by Holland (1975) will be formulated on the basis of two properties which are used to distinguish between different schemata. The properties are

Schemata order ($o(S)$): The number of 0 and 1 positions, i.e., *fixed* positions, presented in the schema,

Schemata length ($\delta(S)$): The distance between the most left and the most right fixed string position.

In the above schema, the order and defining length are 7 and 8 respectively.
A schema's *fitness* (f_S) is another useful property which is used to illustrate the GAs performance. Fitness of a schema in each generation is defined as the average fitness of all strings in the population matched by the schema. In the mating pool, the number of strings matching the schema will be increased if the ratio of the schema's fitness to the average fitness of the population is larger than one. Otherwise, it will die out. In other words, an *above average* schema receives an increasing number of strings in the mating pool generation, a *below average* schema receives decreasing number of strings, and an *average* schema stays on the same level (Michalewicz, 1996).
Furthermore, implementation of the crossover operator may destroy a schema. The shorter the length the higher the survivability is. In other words, a schema with above-average fitness and short defining length is the most probable one to be seen still in the offspring. Conclusively, the defining length of a schema plays the main role in the probability of its destruction and survival (Michalewicz, 1996).
Mutation is the next operator which could destruct a schema. As discussed earlier, the mutation operator randomly changes a single position within a chromosomes with probability p_m. For a schema to survive, all the fixed positions of a schema must remain unchanged. It is clear that the lower the order the higher the survivability probability becomes. In other words, number of the lower order above-average schemata with short defining lengths will be increased in next generations. Summing up the achieved results leads to a theorem which has been introduced by Holland (1975).

Schema Theorem: *Short, low-order, above-average schemata receive exponentially increasing trials in subsequent generations of a genetic algorithm.*

Building Blocks Hypothesis proposed by Goldberg (1989) is an intermediate result of the theorem which tries to explain information exchange during crossover:

Building blocks hypothesis: *The GAs seek near-optimal performance through the juxtaposition of short, low-order, high-performance schemata, called the building blocks.*

Using either the theorem or the hypothesis, one can interpret the GAs work by discovering, emphasizing and recombination of good building blocks of solution in highly parallel fashion. In other words, the GAs differ from random search methods since the algorithms are based on a systematic procedure rather than on a simple stochastic sampling.

3.4 Application of the genetic algorithms in approximation

Although the basic principles of the GAs are very simple, various setups of the method are used in a large number of scientific and engineering problems. Optimization, automatic programming, machine learning, economic models, immune system models, ecological phenomena behavior modelling, population genetics models, interaction between evolution and learning, and social system modelling are some typical examples of the GAs applications (Mitchell, 1998).
The GAs have been widely used in numerical optimization (e.g. Moerder and Pamadi, 1994; Charbonneau, 2002; Vincent, 2003). Herein, we have implemented the GAs for optimization of the approximation function either in the selection of the basis functions or an appropriate shape parameter. Although all the problems can be coded by binary representation, it is the well-suited coding for basis functions selection whereas for the optimal shape parameter estimation, the continuous representation is superior.

3.4.1 Optimum shape parameter estimation

The accuracy of the RBF interpolation methods depends extremely upon the value chosen for the shape parameter c^2. In general, the nature of the scene distortions and the number and distribution of control points are the key factors for the shape parameter selection (Fogel and Tinney, 1996). In other words, it is problem-dependent. The value of c^2 determines whether the interpolating surfaces will be degenerated conics ($c^2 = 0$) or hyperboloids (Hardy, 1990).
There are numerous published methods for generating an appropriate value for c^2. Besides Eq. (3.34) which has also been cited in Hardy (1990), Hardy (1977) recommended the following equation for c^2:

$$c^2 = 0.665\ d^2 \tag{3.78}$$

where, d is the grid spacing of the data points. In reality, data points usually have irregular distribution. Obviously, the formula should be reinterpreted. For the scattered data set, d has been interpreted as the mean distance to its nearest neighbor. Consequently, different schemes have been introduced for computation of the mean distance. Considering the bounding circle of diameter D of the data points has been introduced by Franke (1979). If the circle is divided into n equal circular areas, they have the identical radius $\frac{D}{2\sqrt{n}}$. It can be used as an approximation for the mean distance between two neighboring points:

$$c^2 = \left(0.815\ \frac{D}{2\sqrt{n}}\right)^2 \tag{3.79}$$

However, Franke (1979) has also proposed a modified version of Eq. (3.79) in which the constant parameter has been replaced by a weight parameter W

$$c^2 = \left(W\ \frac{D}{2\sqrt{n}}\right)^2, \tag{3.80}$$

where, $W = \frac{NPPR}{10}$ with NPPR defined as *Number of Points Per Region*. This general form has been reduced to its special form by replacing NPPR= 25 in the scattered data interpolation literature (e.g. Wolfberg, 1990; Fogel and Tinney, 1996)

$$c^2 = \left(1.25\frac{D}{\sqrt{n}}\right)^2. \tag{3.81}$$

Foley (1987), as an alternative, utilized a similar value for the shape parameter c^2 based on the area of the bounding rectangle to the data

$$c^2 = \left(4.0\frac{A}{\sqrt{n}}\right)^2, \tag{3.82}$$

where, A is the area of the bounding rectangle $((x_{\max} - x_{\min}) \cdot (y_{\max} - y_{\min}))$.

As seen in the previous formulae, the shape parameter c^2 is computed merely based on the horizontal distribution of the data points. Carlson and Foley (1991) presented an algorithm for determining the shape parameter in the three-dimensional case. For a 3D data set, they showed that c^2 was a strong function of the third components z_i. Moreover, they observed that the shape parameter was essentially independent of both the number and the horizontal locations (x_i, y_i) which was contradictory to most previous researches. The basic shape parameter computation scheme, corresponding to Carlson and Foley's method, consists of the following steps:

- Scale the data points to the unit cube by setting

$$\bar{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \qquad \bar{y}_i = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}}, \qquad \bar{z}_i = \frac{z_i - z_{\min}}{z_{\max} - z_{\min}} \tag{3.83}$$

- Compute the least squares bivariate quadratic polynomial fit to the data $(\bar{x}_i, \bar{y}_i, \bar{z}_i)$

$$q(\bar{x}_i, \bar{y}_i) =: a_0 + a_1\bar{x} + a_2\bar{y} + a_3\bar{x}^2 + a_4\bar{x}\bar{y} + a_5\bar{y}^2 = \bar{z}_i \tag{3.84}$$

- Compute the average residual

$$s^2 = \frac{1}{n}\sum_{i=1}^{n_o}[\bar{z}_i - q(\bar{x}_i, \bar{y}_i)] \tag{3.85}$$

- Define the shape parameter as:

$$c^2 = \frac{1}{(1 + 120s^2)^2} \tag{3.86}$$

- Form the interpolating function

$$G(\bar{x},\bar{y}) = \sum_{i=1}^{n_o} \alpha_i \sqrt{(\bar{x}-\bar{x}_i)^2 + (\bar{y}-\bar{y}_i)^2 + c^2} \tag{3.87}$$

where,

$$\bar{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \qquad \bar{y} = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \tag{3.88}$$

Carlson and Foley numerical analysis showed a dramatic improvement. Nevertheless, the shape parameter is estimated empirically.
Kansa (1990a,b) proposed a variable shape parameter c_i^2 for each observation point and correspondingly the MQ interpolating function of the form:

$$\varphi(\rho_i) \;=\; \sqrt{\rho_i^2 + c_i^2}. \tag{3.89}$$

where, c_i^2 has the empirical form

$$c_i^2 \;=\; c_{\min}^2 \left(\frac{c_{\max}^2}{c_{\min}^2}\right)^{\frac{i-1}{n_o-1}}, \tag{3.90}$$

where, $c_{\min}^2$ and $c_{\max}^2$ are two input parameters. Although very accurate results were obtained, the question on the determination of $c_{\min}^2$ and $c_{\max}^2$ for different problem was not answered. Moreover, the underlying function was not involved in the computation of the shape parameters.
As already discussed, a similar approach with the shape parameters given in Eq. (3.35) has been implemented by Ruprecht and Müller (1995) for image warping.
Obviously, it is impossible to develop a general formula for the shape parameter computation. As discussed, both the underlying function and the data points distribution should be considered to obtain an optimal or a near-optimal shape parameter. Kansa and Carlson (1992) formulated the problem differently. Besides the observation points (n_o points), they assumed the function to be known at n_e additional points. These additional points were considered as the *evaluation points*. Consequently, a *misfit function* is defined as:

$$Z = \sum_{j=1}^{n_e} \|f_j^e - \hat{f}(\mathbf{x}_j)\|^2, \tag{3.91}$$

where, f_j^e is the value of the underlying function at the j^{th} evaluation point. $\hat{f}(\mathbf{x}_j)$ is the computed value using MQ or IMQ interpolation function at the evaluation point $\mathbf{x}_j$. The goal is to find the set of n_o shape parameters, c_j^2, which minimizes the misfit function. In other words, the following non-linear system of normal equations yields an optimal solution:

$$\frac{\partial Z}{\partial \mathbf{c}} \;=\; \mathbf{0}. \tag{3.92}$$

Although Kansa's formulation was a great step ahead, it could not be implemented easily since

- besides the data points, n_e additional evaluation points were required to define the misfit function,
- the shape parameter is the solution of the non-linear system.

As mentioned earlier, Rippa (1999) introduced the idea of using the one-leave-out method to define the misfit function which resolved the first problem of Kansa's method. He also utilized a two-step numerical procedure to overcome the second problem. In section (3.1.4) we introduced an alternative formulation to Rippa's method which yields a lower interpolation error. Both Rippa and we have utilized a constant shape parameter rather than the variable one. Although we have numerically investigated the implementation of variable shape functions, we have retained the constant shape parameter because of its negligible difference.

As discussed in section (3.3.1), implementation of the gradient-based local optimization problem for selection of the RBF interpolation's shape parameter may lead to an inappropriate solution due to presence of local optimum or unsuccessful minimum bracketing. The GAs algorithms are the suitable optimization methods because they are global optimum seeker and are never trapped in local minima because of the high capability in the exploration of new areas.

Mathematical formulation of the problem is the same as that stated in section (3.3.1). Despite Rippa's method, we don't need any apriori information of the shape parameter for successfully bracketing the minimum. Moreover, since by nature the GAs are parallelized, they sweep the search space very effectively.

The problem is a non-linear one-dimensional optimization problem of the type constraint. Since the GAs solve the problem numerically, one never senses whether the object function is linear or non-linear. In other words, for each individual the corresponding coefficients are computed by setting up a linear system of equations at the observation points. The algorithm proceeds to reach an optimal shape parameter value based on the given data set. Eventually, the interpolant function coefficients α_i are estimated using the estimated optimal shape parameter. As will be seen in section (3.4.4), implementation of the algorithm improves the interpolation accuracy dramatically.

3.4.2 Optimum augmented RBFs

As seen in section (3.1.4), a polynomial function can be added to the RBF interpolating function to remove the long wavelength signals. The question is: what would be the optimum degree of the augmented polynomial? Hickernell and Hon (1998) presented an empirical relationship for two-dimensional wind data. They showed that a polynomial function of degree $m = \frac{(\nu+1)(\nu+2)}{2}$ is required to obtain polynomial precision ν for modelling 2D surface wind data. So far, no general formulation has been addressed for the augmented RBFs.

In order to define an optimal polynomial augmented function, the GAs are the most appropriate tool. In other words, a binary string can be used for including or excluding a particular base function. For ease of operation, let us consider a two-dimensional problem with an augmented polynomial of degree 3. The set of basis functions is

$$\mathbf{\Pi}_3^2 = \left\{1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3\right\}. \tag{3.93}$$

Each base function can be considered as a gene of length 1. Correspondingly, a binary chromosome of length 10 can be used to represent the genes together. Among the basis functions, only the members whose genes are equal to 1 are included. Figure (3.17), shows an arbitrary chromosome and its respective approximating function. In this simple example just 6 basis functions out of 10 will be

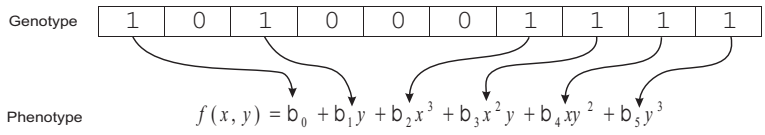


Figure 3.17: An example of genotype and the corresponding phenotype

considered. The aim is to find the optimum combination of basis functions which results in the optimum solution. The solution will be optimal if it yields the maximum fitness. The fitness function can be simply defined as either Eq. (3.40) or Eq. (3.41).

It is also interesting to note that the optimization of augmented polynomial function and the shape parameter can be formulated in a single optimization problem using binary genetic algorithms. In other words, one can combine the shape parameter gene with an appropriate length with the polynomial basis functions' genes and try to find the fittest individual via the BGAs. In general, consider

Eq. (3.94) with the object function given in Eq. (3.41)

$$\begin{cases} \begin{pmatrix} \mathbf{A} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{l} \\ \mathbf{0} \end{pmatrix}, \\ \sum_{k=1}^{n_o} \mid \widetilde{e}^{[k]} \mid + \sum_{k=1}^{n_o} \sum_{j=1}^{n_I} \mid \widetilde{e}_j^{[k]} \mid \rightsquigarrow \min \end{cases} \tag{3.94}$$

As already mentioned, the number of the additional unknowns β_i is u_m. If a binary string of length n_g is required for representation of the shape parameter with enough accuracy, the individuals will be chromosomes with $n_g + u_m$ bits, see Fig. (3.18). Hence, by implementation of the GAs, one can look for an optimum solution in a $n_o + u_m$ dimensional search space. It should be noted here that

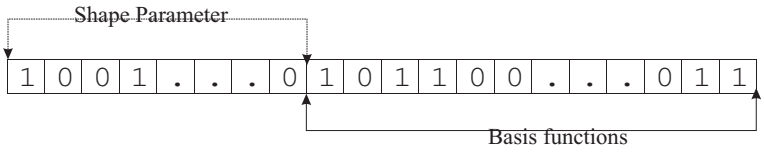


Figure 3.18: An example of combined chromosome

the shape parameter's gene should be decoded for the fitness function evaluation. Furthermore, the basis functions should be selected according to the bit values corresponding to the basis functions. In other words, the numbers β_i and correspondingly the number of functional constraint equations (Eq. 3.27) equal the number 1 within the basis function gene.
Compared to the empirical formula given in Hickernell and Hon (1998), our formulation has the following advantages:

- It can be implemented even in higher dimensions.
- The shape parameter and the polynomial function are optimized simultaneously.
- It is possible to select the basis functions based on the underlying function behavior.

As discussed earlier, every observation of the gravity field can be reduced into a functional on the incremental potential by removing the respective low-frequency signals of the observations. Therefore, we utilize basic RBFs interpolation in which the shape parameter will be the only optimization parameter. As illustrated in the previous section, the problem can be solved either in binary or continuous method.

3.4.3 Optimization of the approximating function

Selecting the suitable interpolating/approximating polynomial function is the main problem which occurs in the function implementation. Zhong (1997) implemented the idea for the interpolation of GPS geoid heights and showed how mismodelling affected the estimated results in the polynomial surface fitting. He utilized the *F-test* to determine polynomial basis function significance. Like every statistical hypothesis testing, the algorithm works if some prerequisites are fulfilled. For instance, removing the outliers is one of those primary requests which has been mentioned in Zhong (1997). However, in most surface fitting problems, the possible blunder cannot be detected using the classical outlier detection methods. Consequently, the correctness of the derived results by this statistical method is not guaranteed.
Herein, we have developed an algorithm for selection of the optimal polynomial basis functions using the GAs. The idea is implemented both for polynomial and rational interpolation/approximation.

Best fitting basis functions selection

In order to obtain an unbiased minimum variance approximation, the mathematical model should contain just *few but significant* basis functions (Zhong, 1997). As demonstrated in section (3.4.2), the BGAs are powerful schemes for telling the significant and insignificant basis apart. The same scheme can be utilized for polynomial approximation. The setup of the genetic coding is the same except the

shape parameter gene which is missing in the polynomial approximation. Therefore, chromosomes are binary string of length $u(d)$ (the unknown coefficients of a polynomial of degree d).
The fitness function can also be defined based on either the observation points (Eq. 3.40) or using both the observation and the evaluation points (Eq. 3.41). The aim is to find the fittes chromosomes or the optimal combination of the polynomial basis functions which leads to minimum variance solution.
To clarify the idea, let us consider the following bivariate function:

$$f(x, y) = 1 + 2x + 3y + 4x^2 + 5xy + 6y^2 + 7x^2y \tag{3.95}$$

The discrete values of the function have been given on an evenly-spaced grid (10×10) points. The problem is to define

- optimum degree of the approximating function, and
- optimal selection of the polynomial basis functions.

For simplicity, the fitness function is assumed equivalent to that of the method of least squares (Euclidian norm of the residuals at the observation points). Having 100 two-dimensional observation points enables us to utilize a polynomial of degree 12 at most. We have started the approximation with a polynomial of degree 12 and gradually decreased the polynomial degree to 3. The degree of the polynomial approximation decreased since the value of the high-degree (higher than 3) basis functions corresponding genes were zero. Consequently, we set up the GA representation for a cubic polynomial approximating function. For instance, the initial value was depicted in Fig. (3.19). Evolution of the

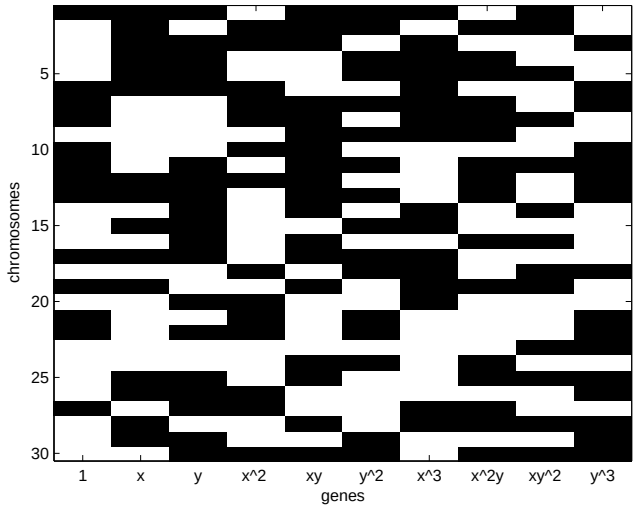


Figure 3.19: The initial population of $f(x, y)$

initial population via reproduction operator reached the fittest individual. As shown in Fig. (3.20), the basis functions of the original function and the recovered one are identical. To recover the

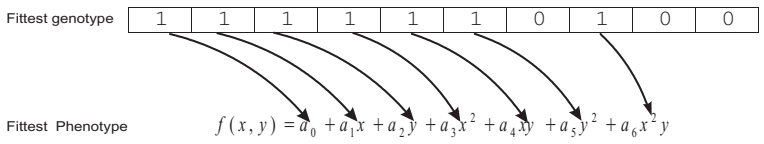


Figure 3.20: The fittest individual

corresponding coefficients, one can use either the continuous GAs or even the ordinary least squares. The latter scheme was selected since the problem was linear and could be solved straightforwardly using the method of least squares.

It is interesting to note that estimating the unknown coefficients based on the fitness function equivalent to Eq. (3.40) or Eq. (3.41) is nearly impossible using the classical least squares. In this case, the problem can be easily treated by using the CGAs. Although the problem will be solved via an evolutionary process instead of a single step solution, it undoubtedly results in a better solution because of setting up the fitness function based on the estimated residuals at the evaluation points rather than the observation ones.

Optimal rational approximation and simultaneous coefficient estimation (Nonlinear Adjustment)

The appropriate rational function can also be selected using the BGAs. However, we can also use the CGAs to alleviate the rational coefficient estimation which is a non-linear estimation problem in its original form (3.18). Simultaneous solution of both the basis function selection and the coefficients estimation is possible if both the BGAs and CGAs are nested. Obviously, running the GA process for each individual throughout the whole generation is rather time consuming. As an alternative, the following multi-stage solution is recommended:

- expressing the interpolating function in the form of the parametric observation equation (3.21) instead of the combined model (3.18),
- running the BGAs to select the optimal set of basis function using either Eq. (3.40) or Eq. (3.41),
- rewriting the mathematical model in its original form Eq. (3.18) and running the CGAs to estimate the coefficients of the non-linear model.

An obvious advantage of the multi-step solution is the comparatively less computation time compared to the simultaneous solution. Furthermore, the coefficients are estimated using the complete mathematical model without any simplification.
As already stated, the rational approximation in its original form results in a non-linear adjustment. Using the GAs either in binary or in continuous representation, one can solve a non-linear problem without linearization. Moreover, there is no need to use any initial value for the unknown coefficients since the problem is solved numerically. The idea can be generalized for also solving other non-linear models. In other words, the GAs is a potentially powerful scheme to numerically estimate the solution of non-linear models.

3.4.4 Comparison of the gridding methods

In order to compare the performance of the already discussed gridding algorithms, the *harmonic* f_{H} and *non-harmonic* f_{N} test functions were used,

$$f_{\mathrm{H}} = r^2 \delta V^{(rr)}, \tag{3.96}$$

$$f_{\mathrm{N}} = \frac{1}{r}\,\delta V^{(r)} + \frac{1}{r^2}\,\delta V^{(\theta\theta)}. \tag{3.97}$$

The functions were computed based on the incremental potential field ($\delta V = V_{\mathrm{EGM96}} - V_{\mathrm{EIGEN2}}$). The coefficients up to 120 from both fields were used to compute the test functions along the GRACE integrated orbit with 0.2 Hz sampling frequency. The simulated observations were split into 120×240 blocks, the size of $1.5^\circ \times 1.5^\circ$. The number of data points within the cells were in the range of $7 - 40$ points, see Fig. (3.21). Moreover, the functions were calculated on the cell mid-points on the mean orbital sphere in order to analyse the gridding error of the various methods. The functions on the sphere are shown Figs. (3.22) and (3.23). The observations were reduced on the mean sphere using the *inverse multiquadric* RBF, *nearest neighbor*, it harmonic polynomials and *rational* algorithms; and the achieved results for the harmonic and the non-harmonic functions are depicted in Figs. (3.24) and (3.25) respectively.
In both cases, the IMQ method with the optimal shape parameter leads to pronouncedly better gridded observations. The nearest neighbor followed the IMQ RBF again for both functions. The error of the nearest neighbor was nearly twice that of the IMQ; and the harmonic and rational approximations were the worst respectively. Of course, one could expect a higher griding errors for the two latter cases

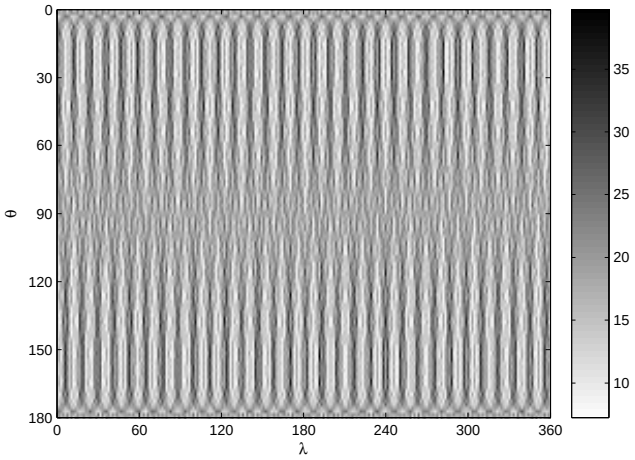


Figure 3.21: Distribution of a one-month span of the GRACE mission simulated observations

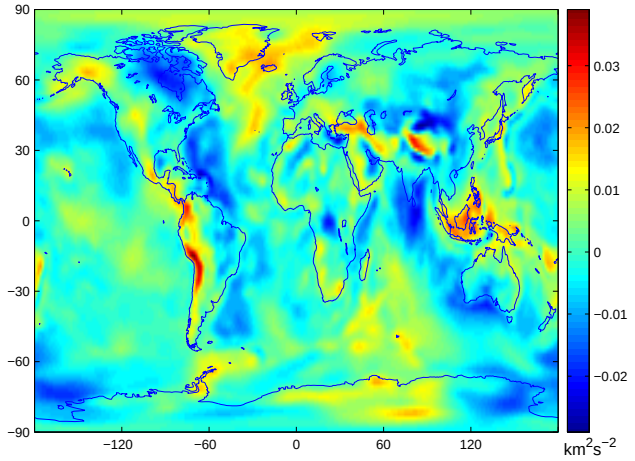


Figure 3.22: Harmonic test function (f_{H})

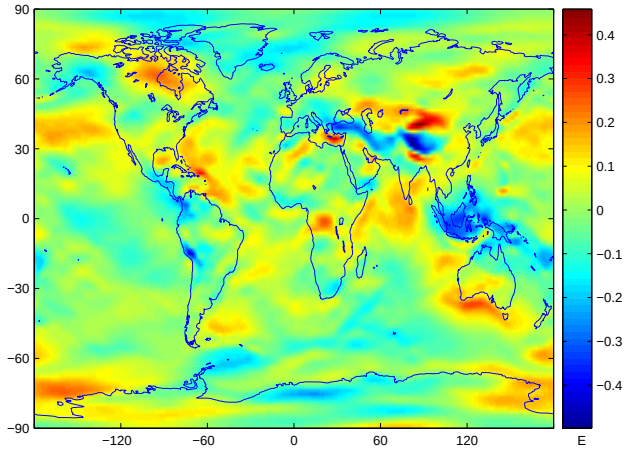


Figure 3.23: Non-harmonic test function (f_{N})

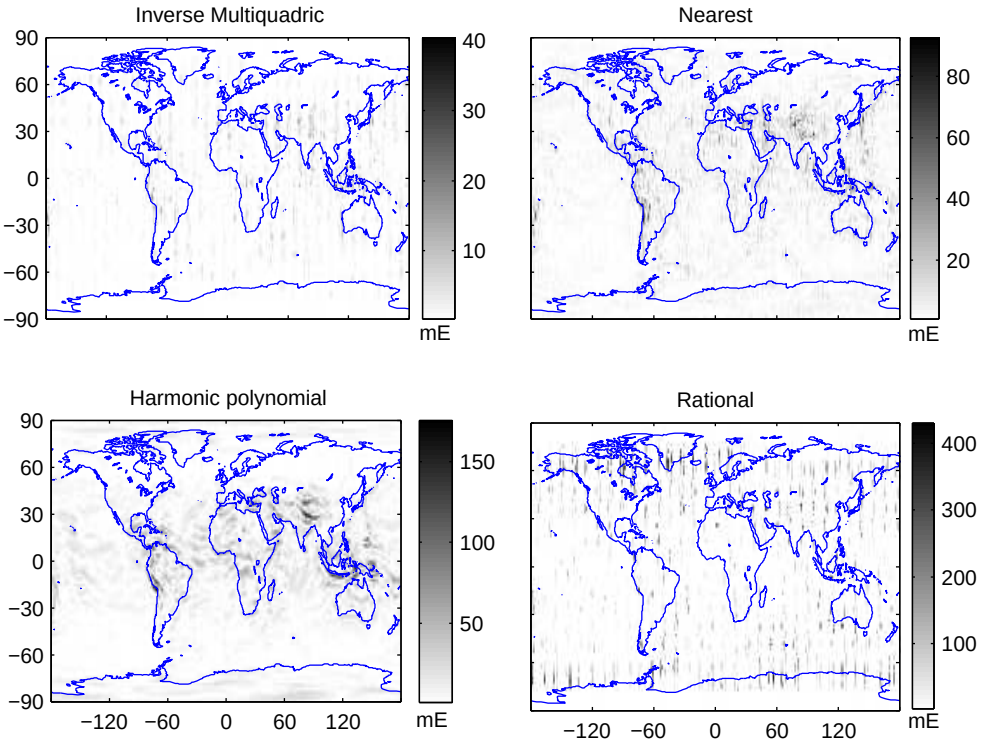


Figure 3.24: Interpolation error of the harmonic test function (f_{H})

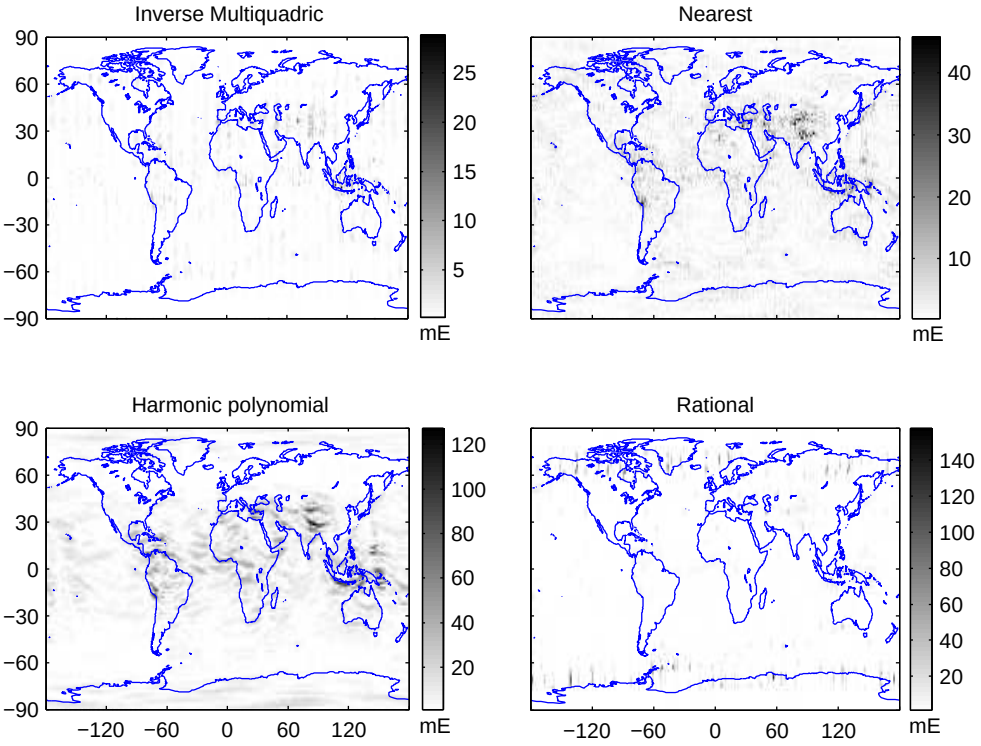


Figure 3.25: Interpolation error of the non-harmonic test function (f_{N})

because of low number of data points and consequently the low-degree approximating polynomials within the cells. The weakness has been highlighted in the mountainous areas since the underlying functions are relatively rougher in the regions. Furthermore, the rational approximation was even worse than the harmonic polynomial since the number of the unknown coefficients is twice those of the harmonic polynomial approximation. The results will improve if the number of the data points within the cells are increased and the polynomial basis either in the harmonic polynomial or in the rational approximation is optimized using the genetic algorithms. Specifically, the results were tabulated in Tables (3.12) and (3.13) in terms of the basic statistics.

Table 3.12: Interpolation error statistics of the harmonic test function

Method	min [mE]	max	mean	std.
Inv. multiquadric	-20.84	40.75	0.01	1.38
Nearest neighbor	-84.58	93.59	0.03	5.89
Harmonic polynomial	-154.27	176.82	0.3399	13.06
Rational	-397.13	433.72	-0.04	32.00

Table 3.13: Interpolation error statistics of the non-harmonic test function

Method	min [mE]	max	mean	std.
Inv. multiquadric	-29.05	12.36	-0.00	0.82
Nearest neighbor	-46.15	44.27	-0.02	3.12
Harmonic polynomial	-114.07	128.39	-0.30	9.78
Rational	-119.79	159.12	-0.02	5.05

As mentioned earlier, IMQ yields the best standard deviations of 1.38 mE and 0.82 mE for harmonic and non-harmonic functions respectively., while the rational approximation statistics were 32 mE and 5 mE.
As seen in Table (3.12), the harmonic polynomial gave a better solution compared to the rational approximation. However, its performance was inferior to that of IMQ and even nearest neighbor. Low-performance was expected since we had just a few observation points within each block. For instance, on average there were 18 observations per cell. Consequently, the total number of the coefficients reached 18 at most which was equivalent to selecting *linear* or *quadratic polynomials*. Taking the local structure of the underlying functions into consideration could convince us that the linear or quadratic rational approximations is inadequate for modelling the local rough structure.

3.5 Summary

In this chapter, different interpolation/approximation methods were used for reduction of observations on a reference surface. Besides the mathematical formulation of the algorithms, an optimization scheme was developed for selecting the optimal interpolation or approximation function.
The optimal RBF interpolator outperformed the other methods. As discussed, *one-leave-out* method as a particular type of *bootstrap* technique was employed based on the observation and interpolation points. Consequently, the shape parameter was determined in a way that the best values in the sense of least squares were estimated at the interpolation points.
Besides the optimal selection of the shape parameter of the RBF interpolators, the BGAs was introduced for deriving an optimal augmented RBF interpolant.

Chapter 4

The Best Fitting Reference Orbit

In general, every observed quantity in the gravity field of the Earth can be written as a functional on the field. For the determination of the field based on the terrestrial data, the respective operator is applied on the normal field and the derived values as the normal field contribution are subtracted from the observations. Although the same scheme can be employed for some extraterrestrial data, it is inconvenient for some others. For instance, satellite gradiometry using satellite pairs is the most ideally suited candidate for doing the job by a formation flying mission. However, the inter-satellite relative velocity vector is a numerically derived constituent of the observable which may spoil the highly accurate observed quantities and break down this brilliant idea. On the other hand, the quantity cannot be written as a functional of the gravity field. Therefore, the corresponding reference quantity should be directly derived using the so called *reference orbit* determination. The reference orbit employment cancels the problematic term and facilitates the whole process tremendously. Furthermore, in order to minimize the linearization error of the gradiometric observation equation, we have to replace the observed quantities correspondingly with the incremental ones.
Since the reference field is just an approximation of the real field, the reference orbit deviates from the real orbit. Approximate gravitational and incomplete non-gravitational forces modelling have also played a secondary role. This deviation monotonously increases as a function of time even with the same initial values for both orbits. Consequently, the respective reference quantities are by a few order of magnitude off, which renders the introduction of the reference orbit completely obsolete. The deviations will decrease if the dynamic reference orbit is replaced by the reduced dynamic orbit.
In this chapter, first we address different methods of the numerical orbit integration. Second, *adaptive solution*, *least squares approach* and *the Genetic approach* are introduced to obtain reduced dynamic orbit.

4.1 Numerical orbit integration

The basic task in the orbit determination process is to derive the state vector (position, velocity and acceleration) of a satellite from observations or a priori information. In other words, the process requires a propagator to obtain the satellite state vector at some time before or after a given state by solving the equations of motion:

$$\ddot{\mathbf{r}} = -\frac{\mu}{\|\mathbf{r}\|^3}\,\mathbf{r} + \mathbf{P}, \tag{4.1}$$

where, $\mathbf{r}$ and $\ddot{\mathbf{r}}$ are the position and acceleration vectors in the inertial frame respectively. μ is the geocentric gravitational constant, and $\mathbf{P}$ is the perturbation forces per unit mass (Kaula, 1966). The perturbed satellite motion integration can be worked out either *analytically* or *numerically*. To integrate the equations analytically which is called *general perturbations* (GP), the perturbation forces are approximated by an expression, which can be analytically integrated. In contrast, to solve the equations numerically, a full force model is considered at a particular position and they are then used as initial conditions for a step-wise integration (Seeber, 2003). This method is usually defined *special perturbations* (SP) (Vallado, 1997).
Compared to SP integrators, GP propagators are faster since they use an analytical solution. On the

other hand, analytical solutions are less accurate than is possible with numerical integrators because of approximating the perturbation forces. Moreover, perturbations caused by non-gravitational forces are discontinuous and thus difficult to model with an analytical expressions (Seeber, 2003). Today, the accuracy requirements for tracking space objects have increased to the point where analytical integrators are no longer adequate. Therefore, the numerical integration is always implemented for orbit determination and spacecraft tracking.
There are numerous methods for performing numerical integration, and many are employed in orbit determination. Depending on the utilized algorithms, they fall into various categories. Integrating to the next point depends on the number of points used; they are called either *single-step* or *multi-step* integrators. Integrators may have either *fixed* or *variable* step size. Moreover, whether the multi-step integration is carried out from epoch or the step-by-step, it is divided into *summed* and *non-summed* groups. Finally, the *double-integration* methods derive the position vector directly from the acceleration vector, whereas the *single-integration* integrators find velocity-given acceleration, and position-given velocity (Berry, 2004).
Furthermore, orbit determination is known as Cowell's formulation if the computation is performed by numerically solving the equation of motion (Eq. 4.1). An alternative formulation is Encke's formulation which involves only integrating the perturbation forces and adding the achieved results to the known solution of the unperturbed motion (two-body problem). Herein, we have considered the Cowell type integration though the algorithms can be implemented on Encke's formulation as well.
In the following sections, we present some of the most commonly used integration methods of type single- and multi-step integration. Refer to Berry (2004) for the discussion and more details on each individual scheme.

4.1.1 Single-step integrators

Integrators that use the state at the current point $\boldsymbol{\xi}_n$ to integrate forward to the next point $\boldsymbol{\xi}_{n+1}$ are called single-step integrators, because they use information from a single point. Examples of single-step integrators include *Euler's method* and *Runge-Kutta*'s. Being used as the low-accuracy integration methods, the single-step methods will also be used to produce the initial backpoints of the multi-step integrators. Moreover, they have a simpler form, so some basic concepts of numerical integration can be easily illustrated using single-step methods. Hence, we begin with fairly simple methods.

Euler's method

Euler's method is the simplest single-step integrator which is seldom used in practice. However, the simplicity of its derivation is an advantage which makes the demonstration of the more advanced techniques possible in the easiest way. The aim of the method is to obtain an approximating solution $\boldsymbol{\xi}(t)$ to a problem of the form (Burden and Faires, 1997)

$$\frac{d\boldsymbol{\xi}}{dt} = \mathbf{f}(t, \boldsymbol{\xi}), \qquad a \leq t \leq b \tag{4.2}$$

subject to the initial condition:

$$\boldsymbol{\xi}(a) = \boldsymbol{\xi}_0.$$

The solution $\boldsymbol{\xi}(t)$ is discrete rather than a continuous one. It is just an approximation to $\boldsymbol{\xi}(t)$ at some discrete points called *mesh points*. The mesh points $\{t_0, t_1, \cdots, t_N\}$ are defined by dividing the interval into N equal subintervals of length $h = \frac{b-a}{N}$ called *step size*,

$$t_i = a + ih, \qquad i = 1, 2, \cdots, N. \tag{4.3}$$

Therefore, the object can be reworded to obtain an approximation to the problem at the mesh points. We use Taylor's expansion to compute $\boldsymbol{\xi}(t_{i+1})$ using the function value and its derivative at the preceding mesh point (*linear approximation*):

$$\boldsymbol{\xi}(t_{i+1}) \approx \boldsymbol{\xi}(t_i) + \frac{d\boldsymbol{\xi}}{dt}(t_{i+1} - t_i), \tag{4.4}$$

since $\boldsymbol{\xi}(t)$ satisfies the differential equation (4.2) and $h = t_{i+1} - t_i$,

$$\boldsymbol{\xi}_{i+1} \approx \boldsymbol{\xi}_i + h\,\mathbf{f}(t_i, \boldsymbol{\xi}_i). \tag{4.5}$$

This algorithm is known as *Euler's method* for solving initial value problems. The formulation is modified to start with the initial condition since the value of the function is given only at t_0

$$\boldsymbol{\xi}_i \approx \boldsymbol{\xi}_0 + h\sum_{j=0}^{i-1} \mathbf{f}(t_j, \boldsymbol{\xi}_j). \tag{4.6}$$

The *truncation error* of the method is of the order of h^2 since just the linear representation of Taylor's expansion of the function is used for approximation. This error is usually called *local error*, because it is the error at one step (Berry, 2004). Therefore, the method is locally correct to first order since the local error is of the second-order. Inaccurate approximation of the function leads to incorrect initial conditions and consequently to a wrong value at the next step. This build-up of the local error is accumulative and usually called *global error*. It is one order lower than the local error, so the method is globally correct to zeroth order. Therefore, the solution is getting worse and worse as the evaluation point moves outwards from the initial point.

Since Euler's method is derived by using the linear approximation of Taylor's expansion, adding higher terms of the expansion is our first attempt to improve the method's accuracy. For instance, including the quadratic term of the expansion result in:

$$\boldsymbol{\xi}_i \approx \boldsymbol{\xi}_0 + h\sum_{j=0}^{i-1} \mathbf{f}(t_j, \boldsymbol{\xi}_j) + \frac{h^2}{2}\sum_{j=0}^{i-1} \dot{\mathbf{f}}(t_j, \boldsymbol{\xi}_j), \tag{4.7}$$

where, $\dot{\mathbf{f}} = \frac{d\mathbf{f}}{dt}$ is derived by differentiation. In general,

$$\boldsymbol{\xi}_i \approx \boldsymbol{\xi}_0 + \sum_{k=1}^{n} \frac{h^k}{k!}\sum_{j=0}^{i-1} \mathbf{f}^{(k)}(t_j, \boldsymbol{\xi}_j), \tag{4.8}$$

where, $\mathbf{f}^k$ is the k^{th} derivative and is derived by successive differentiations.

Implementation of Eq. (4.8) or known as the *Taylor's method of order n*, dramatically reduces the local error. The higher the order the higher order the local error becomes. Nevertheless, the method has the disadvantage of requiring the computation and evaluation of the derivatives of $\mathbf{f}$. For most practical applications, it is a complicated and time-consuming procedure, so the Taylor methods as the generalized form of Euler's algorithm are not used.

Runge-Kutta methods

As discussed previously, the main disadvantage of the Taylor methods was the need to compute and evaluate the derivatives of $\mathbf{f}(t, \boldsymbol{\xi}(t))$. Runge-Kutta methods are a family of single-step integrators which have the high-order local truncation error without requesting the derivatives. They have the general form of:

$$\boldsymbol{\xi}_{i+1} = \boldsymbol{\xi}_i + h\,\boldsymbol{\chi} \tag{4.9}$$

where, $\boldsymbol{\chi}$ is a linear function of slope estimates over the subinterval $t_{i+1} - t_i$

$$\boldsymbol{\chi} = \sum_{j=1}^{m} \alpha_j\,\mathbf{k}_j, \tag{4.10}$$

$\mathbf{k}_j$ are the slope estimates which are found by evaluation of $\mathbf{f}(t, \boldsymbol{\xi})$ at some particular points in the interval, m denotes the order of the integrator and α_i s are weighting constants.

A specific Runge-Kutta method known as the *midpoint method* is simply derived from Euler's method

by evaluation of $\mathbf{f}(t,,\boldsymbol{\xi})$ at the interval midpoint rather than the most left point

$$\boldsymbol{\xi}_{i+1} = \boldsymbol{\xi}_i + h\ \mathbf{f}(t_i + \frac{h}{2},\ \boldsymbol{\xi}_i + \frac{h}{2}\ \mathbf{f}(t_i, \boldsymbol{\xi}_i)). \tag{4.11}$$

Implementation of the midpoint method is equivalent to the Euler method of step size $\frac{h}{2}$, so the local truncation error may decrease to a certain extent. The midpoint method is an example of a first-order Runge-Kutta method ($\alpha_1 = 1$, $\mathbf{k}_1 = \mathbf{f}(t_i + \frac{h}{2}, \boldsymbol{\xi}_i + \frac{h}{2}\mathbf{f}(t_i, \boldsymbol{\xi}_i))$).
A simple form of a second-order Runge-Kutta method is obtained by modification of the Euler's method in the from (Burden and Faires, 1997, pg. 280):

$$\boldsymbol{\xi}_{i+1} = \boldsymbol{\xi}_i + \frac{h}{2}\left[\mathbf{f}(t_i, \boldsymbol{\xi}_i) + \mathbf{f}(t_{i+1}, \boldsymbol{\xi}_i + h\mathbf{f}(t_i, \boldsymbol{\xi}_i))\right] \tag{4.12}$$

equivalent to:

$$\alpha_1 = \alpha_2 = \frac{1}{2}, \qquad \mathbf{k}_1 = \mathbf{f}(t_i, \boldsymbol{\xi}_i), \qquad \mathbf{k}_2 = \mathbf{f}(t_{i+1}, \boldsymbol{\xi}_i + h\mathbf{f}(t_i, \boldsymbol{\xi}_i))$$

The other important second-order Runge-Kutta method is *Heun's method* in which the slope at the beginning of the interval is averaged with the slope resulting from an Euler step to give a better estimate of the solution. It takes the form:

$$\boldsymbol{\xi}_{i+1} = \boldsymbol{\xi}_i + \frac{h}{4}\mathbf{f}(t_i, \boldsymbol{\xi}_i) + \frac{3h}{4}\mathbf{f}(t_i + \frac{2}{3}h, \boldsymbol{\xi}_i + \frac{2}{3}h\mathbf{f}(t_i, \boldsymbol{\xi}_i)) \tag{4.13}$$

or

$$\alpha_1 = \frac{1}{4}, \qquad \alpha_2 = \frac{3}{4}, \qquad \mathbf{k}_1 = \mathbf{f}(t_i, \boldsymbol{\xi}_i), \qquad \mathbf{k}_2 = \mathbf{f}(t_i + \frac{2}{3}h, \boldsymbol{\xi}_i + \frac{2}{3}h\mathbf{f}(t_i, \boldsymbol{\xi}_i))$$

The most commonly used Runge-Kutta method is the *fourth order* which utilizes the weighted average of four slope estimates. It is given by the form

$$\boldsymbol{\xi}_{i+1} = \boldsymbol{\xi}_i + h\left(\frac{1}{6}\mathbf{k}_1 + \frac{1}{3}\mathbf{k}_2 + \frac{1}{3}\mathbf{k}_3 + \frac{1}{6}\mathbf{k}_4\right) \tag{4.14}$$

where,

$$\mathbf{k}_1 = \mathbf{f}(t_i,\ \boldsymbol{\xi_i}), \quad \mathbf{k}_2 = \mathbf{f}(t_i + \frac{h}{2},\ \boldsymbol{\xi_i} + \frac{h}{2}\ \mathbf{k}_1), \quad \mathbf{k}_3 = \mathbf{f}(t_i + \frac{h}{2},\ \boldsymbol{\xi_i} + \frac{h}{2}\ \mathbf{k}_2), \quad \mathbf{k}_4 = \mathbf{f}(t_{i+1},\ \boldsymbol{\xi_i} + h\ \mathbf{k}_3)$$

The local truncation error of the method is of order four, and consequently leads to a better approximation. It requires the function $\mathbf{f}(t, \boldsymbol{\xi})$ to be evaluated four times at each step, whereas the modified Euler's method as well as Heun's method require just two evaluations per step. The higher the order the more the number of evaluations per step is. The relationship between the number of evaluations per step and the order of the local truncation error has been derived by Butcher (1985). He showed that the local truncation error of the methods of order less than five is equal to the number of evaluations per step, whereas for the higher orders the local truncation error is of less orders. Therefore, the low-order (< 5) methods with smaller step size are superior to the higher order ones with a larger step size.
If necessary, the higher-order Runge-Kutta integration methods can also be used. For instance, the fifth through eighth-order Runge-Kutta methods have been formulated in Fehlberg (1974).
The previous Runge-Kutta methods are the prototypes of the general form given by Eq. (4.9), where $\mathbf{k}_j$ s are used to form the increment function $\boldsymbol{\chi}$. In general, the increment function of order m is defined as (Götzelmann, 2003):

$$\boldsymbol{\chi} = \sum_{j=1}^{m} \alpha_j\ \mathbf{f}\left(t_i + \beta_j h,\ \boldsymbol{\xi}_i + \sum_{l=1}^{m} \gamma_{jl}\mathbf{k}_j\right), \tag{4.15}$$

where, the coefficients α_j, β_j and γ_{jl} define each method. These coefficients are selected such that the order of local truncation error is as high as possible (Montenbruck and Gill, 2005). Usually, they

are determined such that the following condition equations are fulfilled

$$\sum_{j=1}^{m} \alpha_j = 1, \qquad \beta_1 = 1, \qquad \sum_{l=1}^{k-1} \gamma_{kl} = \beta_k. \qquad \text{for} \quad 2 \leq k \leq m \tag{4.16}$$

These condition equations generate $m+1$ linear equations, while the total number of unknown coefficients is $\frac{m(m-1)}{2} + 2m - 1$ (Schäfer, 2000). To derive all the coefficients some $\frac{m(m+1)}{2} - 2$ linear equations in terms of the unknown coefficients are required. To derive the requested number of equations, usually the local truncation is expressed in terms of step-size h using the Taylor expansion, where the corresponding coefficients can be written in terms of the unknown coefficients. To guarantee a certain local truncation error order, the respective Taylor's coefficients have to vanish (Schäfer, 2000). It yields $\frac{m(m+1)}{2} - 2$ condition equations which in the combination with Eqs. (4.16) solve the unknown coefficients. A detailed example of the coefficients determination for low-order Runge-Kutta methods has been given in Butcher (1963).
Besides using the Runge-Kutta methods of higher order, one can utilize the *variable-step* integration methods to meet some tolerance for the local error. Mathematically, they are formulated as:

$$\boldsymbol{\xi}_{i+1} = \boldsymbol{\xi}_i + h_i \boldsymbol{\chi}(t_i, \boldsymbol{\xi}_i, h_i), \tag{4.17}$$

where, h_i is the step size of the i^{th} step. The step size is adjusted so that the estimated error of the next step will be roughly equal to the tolerance ϵ.
The local error is estimated by comparing two integrations made of different orders. The Runge-Kutta-Fehlberg method is one popular single-step method of the variable-step type in which the fourth- and fifth-order Runge-Kutta formulae are used for the local error computation (Fehlberg, 1974). The basic computation scheme, corresponding to this technique, consists of the following stages (Burden and Faires, 1997):

1. Problem initialization including introducing an initial step-size h,
2. Computation of $\boldsymbol{\xi}_{i+1}$ using the fourth- and fifth-order Runge-Kutta methods:

$$\boldsymbol{\xi}_{i+1}^{(V)} = \boldsymbol{\xi}_i + h\left[\frac{16}{135}\mathbf{k}_1 + \frac{6656}{12825}\mathbf{k}_3 + \frac{28561}{56430}\mathbf{k}_4 - \frac{9}{50}\mathbf{k}_5 + \frac{2}{55}\mathbf{k}_6\right], \tag{4.18}$$

$$\boldsymbol{\xi}_{i+1}^{(IV)} = \boldsymbol{\xi}_i + h\left[\frac{25}{216}\mathbf{k}_1 + \frac{1408}{2565}\mathbf{k}_3 + \frac{2197}{4104}\mathbf{k}_4 - \frac{1}{5}\mathbf{k}_5\right], \tag{4.19}$$

where,

$$\begin{aligned}
\mathbf{k}_1 &= \mathbf{f}(t_i, \boldsymbol{\xi}_i) \\
\mathbf{k}_2 &= \mathbf{f}(t_i + \frac{h}{4}, \boldsymbol{\xi}_i + \frac{1}{4}\mathbf{k}_1) \\
\mathbf{k}_3 &= \mathbf{f}(t_i + \frac{3h}{8}, \boldsymbol{\xi}_i + \frac{3}{32}\mathbf{k}_1 + \frac{9}{32}\mathbf{k}_2) \\
\mathbf{k}_4 &= \mathbf{f}(t_i + \frac{12h}{13}, \boldsymbol{\xi}_i + \frac{1932}{2197}\mathbf{k}_1 - \frac{7200}{2197}\mathbf{k}_2 + \frac{7296}{2197}\mathbf{k}_3) \\
\mathbf{k}_5 &= \mathbf{f}(t_i + h, \boldsymbol{\xi}_i + \frac{439}{216}\mathbf{k}_1 - 8\mathbf{k}_2 + \frac{3680}{513}\mathbf{k}_3 - \frac{845}{4104}\mathbf{k}_4) \\
\mathbf{k}_6 &= \mathbf{f}(t_i + \frac{1}{2}h, \boldsymbol{\xi}_i - \frac{8}{27}\mathbf{k}_1 + 2\mathbf{k}_2 - \frac{3544}{2565}\mathbf{k}_3 + \frac{1859}{4104}\mathbf{k}_4 - \frac{11}{40}\mathbf{k}_5).
\end{aligned} \tag{4.20}$$

3. Setting of the local error estimate $\delta\boldsymbol{\xi} = \frac{1}{h}\left(\boldsymbol{\xi}_{i+1}^{(V)} - \boldsymbol{\xi}_{i+1}^{(IV)}\right)$.
4. Comparing $\|\delta\boldsymbol{\xi}\|$ and ϵ. The estimate is acceptable and the step size is used for the next step if $\|\delta\boldsymbol{\xi}\| \leq \epsilon$. Otherwise, the step size is updated by:

$$h_i = 0.84\left(\frac{\epsilon}{\|\delta\boldsymbol{\xi}\|}\, h\right)^{\frac{1}{4}}, \tag{4.21}$$

and the calculation is repeated using h_i.

For the problem with complicated functions $\mathbf{f}(t, \boldsymbol{\xi})$, most of the total run-time of an integration is spent on the evaluation of the slope estimates. From the computational aspect, an integrator with a small total evaluation number is preferable, while maintaining a specified accuracy. In the single-step Runge-Kutta methods, it is impossible to minimize this number; and in order to retain a specified accuracy, we should either use the small step size low-order methods or the high-order with a larger step size. In any case, the multiplication of evaluations per step and the number of subintervals (= total number of evaluations) cannot be reduced.
In multi-step methods, the number of evaluations per step is independent of the order. Therefore, the problem is overwhelmed by implementation of the multi-step integrators.

4.1.2 Multi-step integrator

The previously discussed methods were called single-step methods since the approximation for the mesh point t_{i+1} was only based on the previous mesh point and the function evaluations between the points using the mesh point t_i. The evaluations have no direct contribution to the evaluation of the underlying function in future approximations. In other words, a number of evaluations have been carried out in each step. They are used just once and are discarded by moving to next step.
Instead of setting up the approximation at t_{i+1} based on the approximate solution at t_i and the evaluations in between, one can formulate the approximate solution $\boldsymbol{\xi}_{i+1}$ as a function of the previous solutions at the mesh points $t_0, t_1, t_2, \cdots, t_i$. Since the global error is increased with i, it seems reasonable to develop methods that use this more accurate previous information. Moreover, there is no need to evaluate the function within the subinterval over which the solution is being approximated. Of course, for the first few points there is no history available, so the single-step methods inevitably are used.
Because of using more than one previous mesh points called *back-points* to approximate the function at the next step, these methods are called *multi-step* methods. Depending on the number of back-points being used for the forward approximation, the multi-step methods of different step/order are defined. An *m-step* method for finding the approximation $\boldsymbol{\xi}_{i+1}$ is represented as (Burden and Faires, 1997)

$$\boldsymbol{\xi}_{i+1} = \boldsymbol{\xi}_i + h \sum_{k=0}^{m} \alpha_{m-k}\ \mathbf{f}(t_{i+1-k},\, \boldsymbol{\xi}_{i+1-k}) \tag{4.22}$$

where, α_j s are constants. With $\beta_m = 0$, the new approximation is explicitly derived using previous mesh points, whereas $\beta_m \neq 0$ yields an implicit expression for the new approximation. Therefore, the multi-step methods are divided into *explicit* and *implicit* methods. Although the explicit and implicit methods can be used independently, they are usually used simultaneously. Using the approximate solutions at the mesh points $t_0, t_1, \cdots, t_i$, one can predict the approximation at the mesh point t_{i+1}. The predicted value can then be used in an implicit method to derive a refined prediction. Therefore, the multi-step methods with simultaneous implementation of the explicit and implicit integrators are known as the *predictor-corrector* algorithms.
The second term in the right-hand-side of Eq. (4.22) is the increment function of the multi-step method, equivalent to the $\boldsymbol{\chi}$ function in the single-step method. From the theoretical point of view, it is equivalent to:

$$\boldsymbol{\chi} = \frac{1}{h} \int_{t_i}^{t_{i+1}} \mathbf{f}(t, \boldsymbol{\xi}(t)) dt. \tag{4.23}$$

Since $\boldsymbol{\xi}(t)$ is unknown for $t \in (t_i\ t_{i+1}]$, this integral cannot be evaluated. To overcome the problem, one can replace the unknown function $\boldsymbol{\xi}(t)$ by an interpolating function, e.g., the polynomial increment function of degree m, $\mathbf{p}^m(t)$. Hence,

$$\widetilde{\boldsymbol{\chi}}^m = \frac{1}{h} \int_{t_i}^{t_{i+1}} \mathbf{p}^m(t) dt, \tag{4.24}$$

or equivalently:

$$\widetilde{\boldsymbol{\chi}}^m = \int_0^1 \mathbf{p}^m(\sigma) d\sigma, \tag{4.25}$$

where, $\sigma = \frac{t-t_i}{h}$ (Montenbruck and Gill, 2005). The polynomial function's coefficients are computed using the values of $\mathbf{f}$ at the back-points. For instance, to derive the three-step integrator consider a third-order polynomial function $\mathbf{p}^3(\sigma)$

$$\mathbf{p}(\sigma) \;=\; \beta_0 + \beta_1\sigma + \beta_2\sigma^2 + \beta_3\sigma^3. \tag{4.26}$$

Substituting Eq. (4.26) into Eq. (4.25) leads to the following simple expression for the increment function

$$\widetilde{\boldsymbol{\chi}}^3 \;=\; \beta_0 \;+\; \frac{1}{2}\,\beta_1 \;+\; \frac{1}{3}\,\beta_2 \;+\; \frac{1}{4}\,\beta_3. \tag{4.27}$$

The unknown coefficients can then be simply computed using the function values at the current and three previous mesh points $\mathbf{f}_i$, $\mathbf{f}_{i-1}$, $\mathbf{f}_{i-2}$ and $\mathbf{f}_{i-3}$,

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 6 & 0 & 0 & 0 \\ 11 & -18 & 9 & -2 \\ 6 & -15 & 12 & -3 \\ 1 & -3 & 3 & -1 \end{pmatrix}\begin{pmatrix} \mathbf{f}_i \\ \mathbf{f}_{i-1} \\ \mathbf{f}_{i-2} \\ \mathbf{f}_{i-3} \end{pmatrix} \tag{4.28}$$

Replacing the coefficients into Eq. (4.27) yields the cubic increment function,

$$\widetilde{\boldsymbol{\chi}}^3 \;=\; \frac{1}{24}\left(55\mathbf{f}_i \;-\; 59\mathbf{f}_{i-1} \;+\; 37\mathbf{f}_{i-2} \;-\; 9\mathbf{f}_{i-3}\right). \tag{4.29}$$

and consequently:

$$\widetilde{\boldsymbol{\xi}}^3_{i+1} \;=\; \boldsymbol{\xi}_i + \frac{1}{24}\left(55\mathbf{f}_i \;-\; 59\mathbf{f}_{i-1} \;+\; 37\mathbf{f}_{i-2} \;-\; 9\mathbf{f}_{i-3}\right). \tag{4.30}$$

Equation (4.30) is the explicit four-step integrator which is called *four-step Adams-Bashforth predictor*. Similarly, one can derive lower order Adams-Bashforth predictors using low-order polynomials and less backpoints.

The idea can be generalized to the m-step Adams-Bashforth predictor formulation. To achieve this goal, a polynomial approximation $\mathbf{p}^m$ of order $m-1$ that interpolates m points

$$(t_{i-m+1}, \mathbf{f}_{i+m-1}), \cdots (t_i, \mathbf{f}_i)$$

is given by Newton's formula (Montenbruck and Gill, 2005)

$$\mathbf{p}^m(t) \;=\; \mathbf{p}^m(t_i + \sigma h) \;=\; \sum_{j=0}^{m-1} (-1)^j \binom{-\sigma}{j} \nabla^j \mathbf{f}_i \tag{4.31}$$

where, ∇^j stands for the j-order *backward differences*. Inserting $\mathbf{p}^m$ into Eq. (4.24) gives the increment function of the m-step Adams-Bashforth method

$$\widetilde{\boldsymbol{\chi}}^m_{AB} \;=\; \sum_{j=1}^{m} \alpha_{mj}\,\mathbf{f}_{i-m+j}, \tag{4.32}$$

where, α_{mj} s are the unknown coefficients obtained by using a simple relation (Montenbruck and Gill, 2005):

$$\alpha_{mj} \;=\; (-1)^{m-j} \sum_{l=m-j}^{m-1} \beta_l \binom{l}{m-j}. \tag{4.33}$$

Numerical values of the coefficients β_l are obtained from the following simple recurrence formula

$$\beta_l \;=\; 1 - \sum_{k=0}^{l-1} \frac{1}{l+1-k}\,\beta_k \tag{4.34}$$

With $\beta_0 = 1$, one can easily derive the previously derived coefficients for the low-order Adams-Bashforth methods using Eqs. (4.33) and (4.34).

As stated earlier, the explicit form of the increment function is set up based on the given information

at the back-points up to t_i and the function is utilized for the prediction of the function at $t_{i+1} = t_i + h$. In other words, the approximation is obtained throughout an extrapolation process where the solution cannot be expected to be very good.

The implicit multi-step methods include the value of the function at the evaluation point $f(t_{i+1})$ into the approximation process and result in a better approximation. The most commonly used type of the multi-step implicit method is the *Adams-Moulton* method which is the implicit version of the Adams-Bashforth explicit formulation. As with the Adams-Bashforth method, the polynomial $\mathbf{p}^m$ passing through

$$(t_{i-m+2}, \mathbf{f}_{i+m-2}), \cdots (t_{i+1}, \mathbf{f}_{i+1})$$

is considered. The implicit increment function, equivalent to Eq. (4.32), is defined: (Montenbruck and Gill, 2005)

$$\boldsymbol{\chi}_{AM}^{m} = \sum_{j=1}^{m} \alpha_{mj}^{*} \, \mathbf{f}_{i+1-m+j} \tag{4.35}$$

where the unknown coefficients α_{mj}^{*} are computed in a two-stage computation. First, the numerical values of β_l^* are computed using a recurrence relation similar to Eq. (4.34)

$$\beta_l^* = -\sum_{k=0}^{l-1} \frac{1}{l+1-k} \, \beta_k^*, \tag{4.36}$$

where, $\beta_0^* = 0$ is used for the initialization. Finally, the coefficients α_{mj}^{*} are derived using:

$$\alpha_{mj}^{*} = (-1)^{m-j} \sum_{l=m-j}^{m-1} \beta_l^* \binom{l}{m-j} \tag{4.37}$$

for $j = 1, 2, \cdots, m$ (Gerigorieff, 1977).

Apparently, Eq. (4.35) and (4.32) are similar; however, they are different from implementation point of view. The Adams-Moulton formulation is implicitly a function of the unknown quantity $\boldsymbol{\xi}_{i+1}$. Therefore, the implicit formulation can be used in an iterative solution rather than a single stage one. Alternatively, the Adams-Bashforth method can be used to predict the function and the predicted value is improved using the Adams-Moulton algorithm. Consequently, these two algorithms are complementary and usually known as the *predictor-corrector* method. The following five-stage computational algorithm is employed to obtain a reasonable solution:

1. Predict an initial estimate of the solution at t_{i+1} using the Adams-Bashforth formula (*Prediction*)
$$\boldsymbol{\xi}_{i+1}^{P} = \boldsymbol{\xi}_i + h\widetilde{\boldsymbol{\chi}}_{AB}^{m}$$
2. Evaluate the corresponding function value using the initial estimate (*Evaluation*)
$$\mathbf{f}_{i+1}^{P} = \mathbf{f}\left(t_{i+1}, \boldsymbol{\xi}_{i+1}^{P}\right)$$
3. Correct the estimate using the Adams-Moulton corrector (*Correction*)
$$\boldsymbol{\xi}_{i+1}^{C} = \boldsymbol{\xi}_i + h\widetilde{\boldsymbol{\chi}}_{\mathrm{AM}}^{m}$$
4. Evaluate the function corrected estimate (*Evaluation*)
$$\mathbf{f}_{i+1}^{C} = \mathbf{f}\left(t_{i+1}, \boldsymbol{\xi}_{i+1}^{C}\right)$$
5. Iterate the third and fourth steps until the difference of two subsequent corrected values is negligible.

The computational algorithm is usually called $\mathrm{PE(CE)}^n$, where n stands for the number of iterations until the required accuracy is obtained. For most practical applications, $n = 1$ seems adequate since

each iteration costs another function evaluation.

There are many more numerical methods which have already been applied in scientific and engineering applications. Berry (2004) has done a comprehensive study on different numerical methods of integration. Detailed discussion of all the numerical methods are far beyond the scope of this thesis. Interested readers can refer to (e.g. Maury and Segal, 1969; Lundberg, 1981; Berry, 2004; Montenbruck and Gill, 2005).

In the absence of the non-gravitational forces, the equation of motion (Eq. 4.1) is a second-order, nonlinear, ordinary differential equation where the function $\mathbf{P}$ is a smooth, periodic function representing the forces of gravity acting upon the satellite. As mentioned previously, second-order initial value problems can be directly solved by *Class* II *integration method* (Lundberg, 1981). Alternatively, it may be reduced from a set of second-order ordinary differential equations to a set of first-order equations, thus allowing the option of integrating (4.1) with a *Class I method*. An equivalent representation of the equation of motion as a first-order initial value problem is

$$\dot{\boldsymbol{\xi}}(t) = \begin{pmatrix} \dot{\mathbf{r}}(t) \\ \ddot{\mathbf{r}}(t) \end{pmatrix} = \begin{pmatrix} \dot{\mathbf{r}}(t) \\ \nabla V(\mathbf{r}(t)) \end{pmatrix} = \mathbf{f}(t, \boldsymbol{\xi}(t)), \tag{4.38}$$

subject to:

$$\boldsymbol{\xi}_0 = \boldsymbol{\xi}(t_0) = \begin{pmatrix} \mathbf{r}(t_0) \\ \dot{\mathbf{r}}(t_0) \end{pmatrix}. \tag{4.39}$$

Now, the numerical orbit integration task is equivalent to the solution of a first-order vector-valued (6×1) initial value problem which can be solved using one of the previously discussed methods. The selection of a particular numerical method is subject to the accuracy and cost of using the numerical method, where the cost is usually measured in terms of computer time.

Herein, Satellite Orbit Synthesis Program *SOSP2* written by Götzelmann (2003) is used for orbit integration. A twelve-step predictor-corrector algorithm has been implemented for performing the task. For initialization of the multi-step algorithm, the Runge-Kutta method of order 16 has been used.

4.2 The best fitting reference orbit

As discussed in Chapter (2), the original observations are replaced by incremental observations. Ideally, the reference observations derived from the reference orbit integration should be computed at the location of the original observations. However, the real and the reference orbits deviate from each other due to the difference in the Earth's real and reference gravitational fields. Consequently, the reference positions are different from the original observation locations. Differences in the locations of the reference observations introduce an error into the residual observations which is called *location error*. In order to minimize the location error, the reference orbit should be as close to the real orbit as possible.

A fictitious satellite in the reference field imitates the real one if the gravitational fields and the initial conditions are identical. Finding a reference gravitational potential which exactly represents the geopotential is the unknown solution. Consequently, it results in an implicit problem which should be solved iteratively. This idea can independently be used for the determination of the gravity field which is called *variational approach*.

An appropriate selection of the initial conditions can reduce the location error. However, using an identical set of initial conditions may lead to an inappropriate reference orbit when the gravitational fields are different. In the subsequent sub-sections, we propose three different methods for selecting the initial conditions such that the total misfit of two orbits is minimized.

4.2.1 The adaptive reference orbit

As stated in the previous section, the orbit determination problem is reduced into an initial value problem and then solved numerically by considering the initial conditions as well as the governing

forces:

$$\begin{cases} \dot{\boldsymbol{\xi}}(t) = \mathbf{f}(t, \boldsymbol{\xi}(t)) \\ \boldsymbol{\xi}(t_0) = \boldsymbol{\xi}_0 \end{cases} \tag{4.40}$$

Let us assume a given reference gravitational potential. The goal is to generate the corresponding reference orbit in such a way that the difference between the real and reference orbits or equivalently the location error is minimized. Furthermore, assume that at time t_0 both orbits coincide. Since the governing gravitational real and reference potentials are different, the two orbits are diverging as a function of time

$$\Delta\boldsymbol{\xi}(t) = \boldsymbol{\xi}(t) - \boldsymbol{\xi}^{\text{ref}}(t), \tag{4.41}$$

where, $\boldsymbol{\xi}^{\text{ref}}$ refers to the reference state vector. Depending on how good the approximation of the reference field is, the differences behave differently. However, the differences linearly increase, see Fig. (4.1). EMG96 and EIGEN2 both up to degree and order 120 were considered as the pseudo-

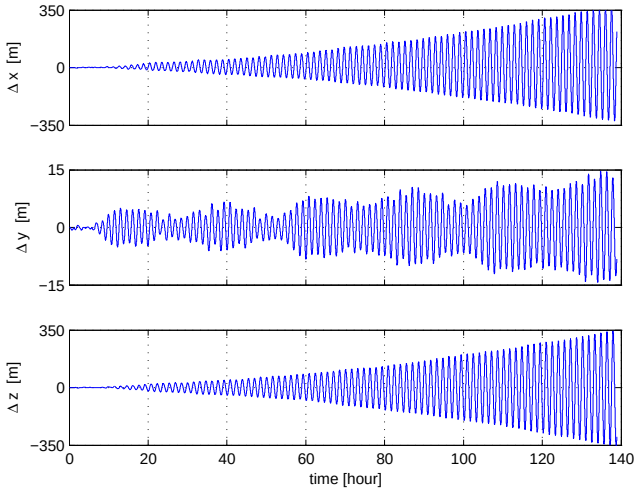


Figure 4.1: Coordinate differences (pseudo-real orbit (EGM96) minus reference orbit (EIGEN2))

real and the reference field. Although the differences at the beginning were negligible, they linearly increase and reach a few hundred meters just in a few days. Therefore, the reference orbit should be modified such that the differences do not exceed some limits.

Instead of integrating the whole reference orbit using just one set of the initial conditions, one can split the orbit into smaller time intervals and update the initial conditions at the beginning of each sub-interval. Depending on the accuracy requirement, the number of sub-intervals is specified. For instance, instead of integrating a one-month span of the orbit at once, we can split the time span into 30 sub-intervals and perform the integration for each day using the updated initial conditions. This method yields the so-called *adaptive reference orbit* whose maximum deviation from the pseudo-real occurs at the end of each sub-interval. Undoubtedly, the maximum diurnal deviation is significantly smaller than that for the monthly integration.

In order to compare the idea of the adaptive reference orbit, the previous example was performed using daily update initial conditions. As seen in Fig. (4.2), the differences decreased dramatically where the maximum deviations were about a few ten meters.

The simplicity of the implementation is the main advantage of the adaptive method where the differences can be bounded without doing any extra computation. The only additional computation is the single-step integration which should be performed for each sub-interval individually. In contrast, the adaptive reference orbit has a discontinuity at the beginning of each sub-interval which is created artificially because of successive updating. However, for the SST of the type LL-SST, the coordinate differences between the satellite pair rather than the absolute position vectors are used. As a consequence, the reference observations will be affected negligibly if the position vector of both satellites jumped identically. Figure (4.3) shows the differences in the intersatellite relative position vector $\delta\mathbf{r}$. As shown in Fig. (4.3), the coordinate differences deviated from the pseudo-real orbit about a

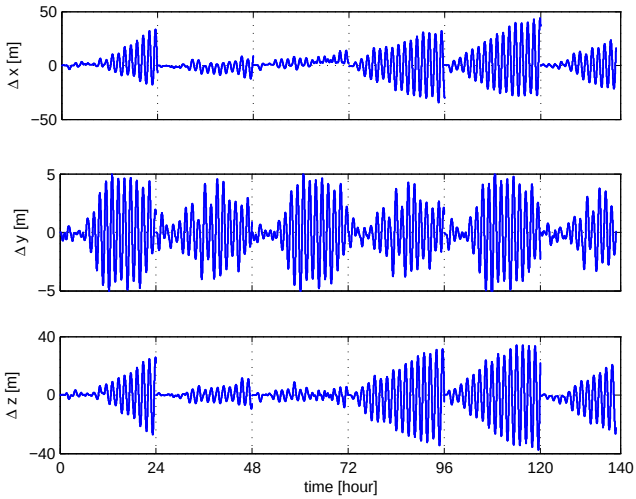


Figure 4.2: Adaptive coordinate differences (pseudo-real orbit (EGM96) minus daily adaptive reference orbit (EIGEN2))

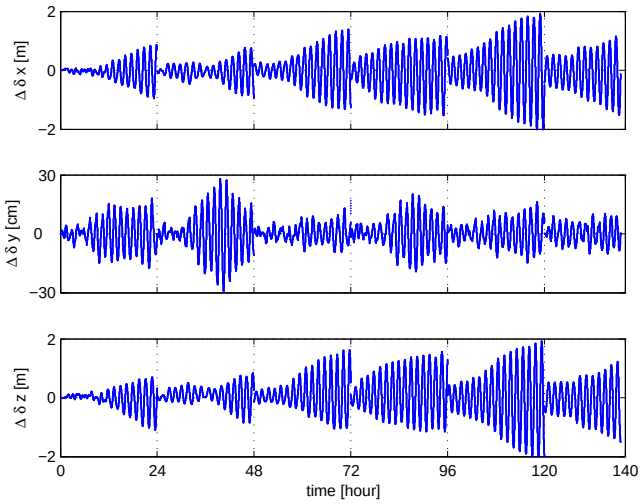


Figure 4.3: Intersatellite relative position vector differences based on daily adaptive reference orbit (EIGEN2)

few meters in the worst case. Hopefully, the effect of the remaining discrepancies on the reference observation vector is negligible. For instance, consider the LL-SST observation of the gradiometric type (Keller and Sharifi, 2005)

$$\mathcal{L}^{\mathrm{SG}}\delta V \approx \frac{\ddot{\varrho}}{\varrho} + \frac{\dot{\varrho}^2}{\varrho^2} - \frac{\|\delta\dot{\mathbf{r}}\|^2}{\varrho^2} \tag{4.42}$$

where the intersatellite range, range-rate, and range-acceleration are derived using the intersatellite relative position, velocity and acceleration vector

$$\begin{aligned} \varrho &= \sqrt{\delta\mathbf{r}^T\,\delta\mathbf{r}} \\ \dot{\varrho} &= \frac{1}{\varrho}\delta\mathbf{r}^T\,\delta\dot{\mathbf{r}} \\ \ddot{\varrho} &= \frac{1}{\varrho}\delta\mathbf{r}^T\,\delta\ddot{\mathbf{r}} - \frac{\dot{\varrho}^2}{\varrho} + \frac{\|\delta\dot{\mathbf{r}}\|^2}{\varrho}. \end{aligned} \tag{4.43}$$

These quantities were computed based on the adaptive reference orbit and the achieved results were depicted in Fig. (4.4). As shown in the figure, there is just a small jump in the intersatellite range between two subsequent subintervals and no discontinuities can be seen in the range-rate and range-acceleration.

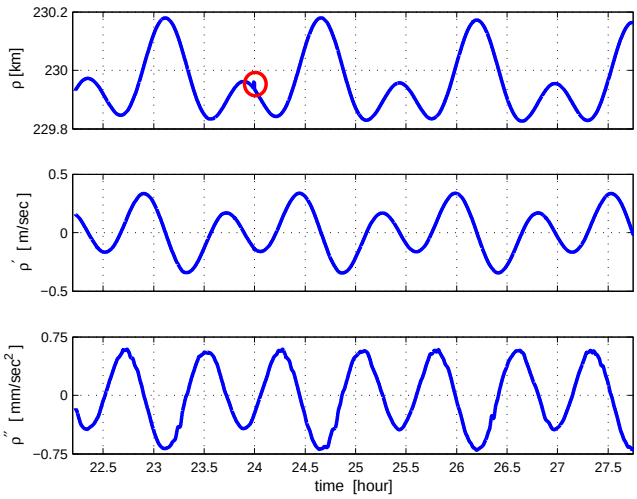


Figure 4.4: The reference intersatellite range, range-rate and range- acceleration derived from the adaptive reference orbit

Finally, the observation of the gradiometric type, the right-hand-side of Eq. (4.42) was calculated and the achieved results are shown in Fig. (4.5). Fortunately, the artificial discontinuities completely disappeared in the gradiometric observation and even in the residual observations.
Alternatively, the discontinuities of the residual observations can also be treated using the cycle slip repairing methods. For instance, the region where discontinuity occurs, may be modelled by a low-degree polynomial $p(t)$ (Keller, 2005). The coefficients of the polynomial are then determined using the neighboring points.
To sum up, the carried out numerical analysis showed that the appeared discontinuities vanished for the relative observations in the LL-SST mode.

4.2.2 The least squares approach

In the previous approach, we integrated the orbit in smaller subintervals in order to circumvent the linearly increasing differences. Although the differences were zero at the beginning, they increased linearly and reached their maxima at the end of each subinterval. It may be better to uniformly distribute the differences over the interval. In other words, the V-*shaped* pattern of the differences is changed in a such way that the deviation of two orbits remains constant as much as possible. To this

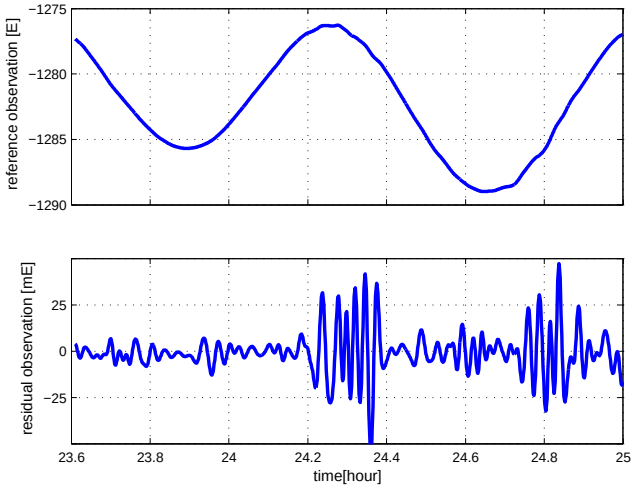


Figure 4.5: The adaptive reference observations

end, the initial conditions corresponding to the reference field rather than the pseudo-real one should be used. Hence,

$$\dot{\boldsymbol{\xi}}^{\mathrm{ref}}(t) \;=\; \mathbf{f}^{\mathrm{ref}}(t, \boldsymbol{\xi}^{\mathrm{ref}}(t)), \qquad \boldsymbol{\xi}^{\mathrm{ref}}(t_0) \;=\; \boldsymbol{\xi}_0^{\mathrm{ref}} \tag{4.44}$$

In the current formulation instead of using $\boldsymbol{\xi}_0$ as the initial values, the initial values corresponding to the reference orbit ($\boldsymbol{\xi}_0^{\mathrm{ref}}$) are used. The question is how we can suitably determine $\boldsymbol{\xi}_0^{\mathrm{ref}}$ which results in a uniform deviation.

The idea of the best fitting orbit in the sense of least squares of the deviations has been employed by Ballani (1988). In this method, we assume the given initial state vector as the approximate value of the reference initial conditions and try to find the corrections

$$\boldsymbol{\xi}_0^{\mathrm{ref}} \;=\; \boldsymbol{\xi}_0 \;+\; \delta\boldsymbol{\xi}. \tag{4.45}$$

Therefore, the problem is to find an appropriate correction to the initial estimate of the initial state vector in a way that the deviation of two orbits is minimized. Consequently, it can be interpreted as an optimization problem as follows

$$\begin{cases} \boldsymbol{\xi}_i \;+\; \Delta\boldsymbol{\xi}_i \;=\; \boldsymbol{\xi}^{\mathrm{ref}}(t_i, \boldsymbol{\xi}_0^{\mathrm{ref}}) & \text{for } i = 0, 1, \cdots, N \\ \sum_{j=1}^{N} \|\Delta\boldsymbol{\xi}_j\| \quad \rightsquigarrow \quad \min \end{cases} \tag{4.46}$$

The subject functions are N (total number of the integration points) non-linear equations in terms of $\boldsymbol{\xi}_0^{\mathrm{ref}}$ or equivalently $\delta\boldsymbol{\xi}$. Assume the sought-after correction is small enough such that the linearization of the equations is accurate enough. Then,

$$\boldsymbol{\xi}^{\mathrm{ref}}(t_i, \boldsymbol{\xi}_0^{\mathrm{ref}}) \;=\; \boldsymbol{\xi}^{\mathrm{ref}}(t_i, \boldsymbol{\xi}_0) \;+\; \frac{\partial \boldsymbol{\xi}_i^{\mathrm{ref}}}{\partial \boldsymbol{\xi}_0^{\mathrm{ref}}}\, \delta\boldsymbol{\xi}. \tag{4.47}$$

Inserting the linearized form of the state vector into Eq. (4.46) is recast into:

$$\begin{cases} \Delta\boldsymbol{\xi}_i \;=\; \frac{\partial \boldsymbol{\xi}_i^{\mathrm{ref}}}{\partial \boldsymbol{\xi}_0^{\mathrm{ref}}}\, \delta\boldsymbol{\xi} - \left(\boldsymbol{\xi}_i - \boldsymbol{\xi}^{\mathrm{ref}}(t_i, \boldsymbol{\xi}_0)\right) \\ \sum_{j=1}^{N} \|\Delta\boldsymbol{\xi}_j\| \quad \rightsquigarrow \quad \min \end{cases} \tag{4.48}$$

or equivalently,

$$\begin{cases} \mathbf{d} \;=\; \mathbf{A}\,\delta\boldsymbol{\xi} \;-\; \delta\mathbf{l} \\ \|\mathbf{d}\| \quad \rightsquigarrow \quad \min \end{cases} \tag{4.49}$$

with, the misfit vector $\mathbf{d} = [d_i] = [\Delta\boldsymbol{\xi}_i]_{6N\times 1}$, the first-order design matrix $\mathbf{A} = \left[\frac{\partial\boldsymbol{\xi}_i^{\text{ref}}}{\partial\boldsymbol{\xi}_0^{\text{ref}}}\right]_{6N\times 6}$ and the observation misclosure vector $\delta\mathbf{l} = [\delta l_i] = [\boldsymbol{\xi}_i - \boldsymbol{\xi}_{i0}^{\text{ref}}]$. Applying the method of least squares yields

$$\widehat{\delta\boldsymbol{\xi}} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{P}\delta\mathbf{l}. \tag{4.50}$$

where, $\mathbf{P}$ is the weight matrix of the observation equations which can be set to the identity matrix $\mathbf{I}_{6N\times 6N}$.
Except the design matrix $\mathbf{A}$, the remaining vectors and matrices can easily be computed. As defined, the design matrix entries are the partial derivatives of the state vectors with respect to the initial state vector. The following are two applicable schemes for deriving the partial derivatives:

Numerical differentiation The derivative of the state vector $\boldsymbol{\xi}_i^{\text{ref}}$ with respect to the initial state elements $\xi_{k_0}^{\text{ref}}$ is:

$$\frac{\partial\boldsymbol{\xi}_i^{\text{ref}}}{\partial\xi_{k_0}^{\text{ref}}} \approx \frac{\boldsymbol{\xi}(t_i,\boldsymbol{\xi}_0 + d\xi_k\mathbf{e}_k) - \boldsymbol{\xi}(t_i,\boldsymbol{\xi}_0)}{d\xi_k}, \tag{4.51}$$

where, $d\xi_k$ is an increment of the k^{th} element of the initial state vector and $\mathbf{e}_k$ is a 6×1 vector which is defined as:

$$\mathbf{e}_k(j,1) = \begin{cases} 1 & \text{if} \quad j = k \\ 0 & \text{if} \quad j \neq k \end{cases}$$

The formulation is valid if the increment is infinitesimal. For instance, 1 m and 1 m/s have been proposed respectively as the position and velocity increment by Götzelmann (2003). By running two times, integration and implementation of the formula, entries of the design matrix are calculated.

Analytical solution The *analytical scheme* as the alternative to the numerical method has been introduced by Ballani (1988). The state vector at the mesh point t_i is computed using the differential equation Eq. (4.44). Therefore, it can be used to express the partial derivative of the state vector at time t_i with respect to the initial state vector. Differentiating Eq. (4.44) with respect to $\boldsymbol{\xi}_0^{\text{ref}}$ yields:

$$\frac{\partial\dot{\boldsymbol{\xi}}_i^{\text{ref}}}{\partial\boldsymbol{\xi}_0^{\text{ref}}} = \frac{\partial}{\partial\boldsymbol{\xi}_0^{\text{ref}}}\mathbf{f}(t_i,\boldsymbol{\xi}_i^{\text{ref}}) = \begin{pmatrix} \frac{\partial\dot{\mathbf{r}}_i}{\partial\boldsymbol{\xi}_0^{\text{ref}}} \\ \frac{\partial\ddot{\mathbf{r}}_i}{\partial\boldsymbol{\xi}_0^{\text{ref}}} \end{pmatrix}. \tag{4.52}$$

The equation is recast into the following first-order differential equations (Götzelmann, 2003)

$$\dot{\boldsymbol{\zeta}}_i = \mathbf{g}(\boldsymbol{\zeta}_i) = \begin{pmatrix} \frac{\partial\dot{\mathbf{r}}_i}{\partial\boldsymbol{\xi}_0^{\text{ref}}} \\ \mathbf{G}_i\,\frac{\partial\mathbf{r}_i}{\partial\boldsymbol{\xi}_0^{\text{ref}}} \end{pmatrix}, \tag{4.53}$$

where,

$$\boldsymbol{\zeta}_i = \begin{pmatrix} \frac{\partial\mathbf{r}_i}{\partial\boldsymbol{\xi}_0^{\text{ref}}} \\ \frac{\partial\dot{\mathbf{r}}_i}{\partial\boldsymbol{\xi}_0^{\text{ref}}} \end{pmatrix}$$

They are usually called *variation equations*. The determination of $\boldsymbol{\zeta}$ can be formulated as an initial value problem by combining Eq. (4.53) with an appropriate set of initial conditions. One can simply derive the initial condition at time $t = t_0$ as:

$$\boldsymbol{\zeta}_0 = \boldsymbol{\zeta}(t_0) = \begin{pmatrix} \frac{\partial\mathbf{r}_0}{\partial\boldsymbol{\xi}_0^{\text{ref}}} \\ \frac{\partial\dot{\mathbf{r}}_0}{\partial\boldsymbol{\xi}_0^{\text{ref}}} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_{3\times 3} & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \mathbf{I}_{3\times 3} \end{pmatrix} \tag{4.54}$$

consequently,

$$\begin{cases} \dot{\boldsymbol{\zeta}}_i = \mathbf{g}(\boldsymbol{\zeta}_i) \\ \boldsymbol{\zeta}(t_0) = \boldsymbol{\zeta}_0 \end{cases} \tag{4.55}$$

The problem can be solved numerically using one the previously discussed numerical integration schemes. The achieved results are the design matrix entries.

As a representative example, consider the previous example (Fig. 4.1). For a full month of observations the deviations in x, y and z components reached nearly 2 km, 0.1 km and 2 km respectively. The original sampling rate was 0.2 Hz. For ease of operation, they were down sampled to 0.03 Hz. By implementation of the method of least squares the deviations of x and z components reduced to nearly 600 m whereas the y component remained unchanged, see Fig. (4.6). It should be noted that the design matrix of the linearized system were derived using the closed-form formulas of the Kepler orbit (Goodyear, 1965; Feng, 2001). For convenience, observations were down sampled to 0.0001 Hz

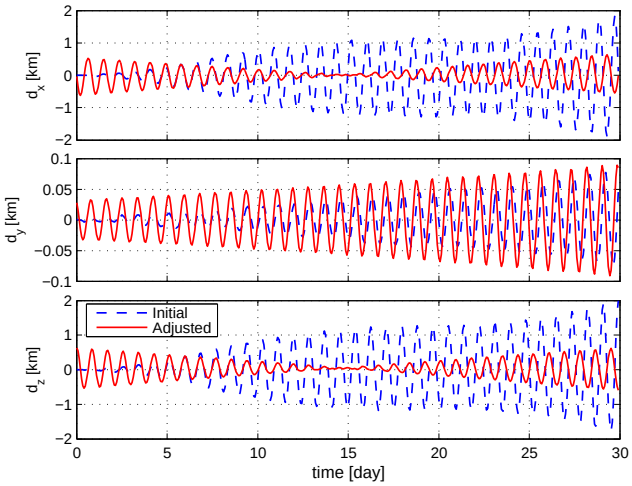


Figure 4.6: The butterfly pattern of the deviations

in Fig. (4.6).

4.2.3 The genetic approach

As discussed previously, better initial values are estimated by minimization of the real and reference orbit discrepancies in the least squares approach. Since the optimization is performed based on the linearized equations, it is required that

- the first guess of the initial values be close enough to the optimal ones, and
- a good approximation of the gravity field be known.

Alternatively, the problem can be solved using the non-linear optimization methods, e.g., the *genetic algorithms*. Using the GAs, initialization can be easily bypassed. Therefore, the GAs can be utilized as an optimizer to solve the problem in its non-linear form Eq. (4.46).
As an intelligent alternative, the *genetic approach* can also be used for modelling the reference orbit deviations. Instead of minimization of the mean of the deviations squared, we will model the discrepancies as a non-linear time-dependent model. In general, the differences are decomposed into three different components defined as:

$$\Delta\boldsymbol{\xi}(t) \;=\; \mathbf{p}^m(t) \;+\; \mathbf{q}(t)\,\mathbf{t}_1(t) + \mathbf{r}(t)\,\mathbf{t}_2(t) \tag{4.56}$$

which, $\mathbf{p}$ is an m-order polynomial function representing the polynomial trend of the deviations. To prevent the unfair fluctuation, its degree is selected as low degree as possible. $\mathbf{q}$ is a non-linear function which demonstrates the behavior of the non-periodic amplitude of the periodic fluctuation of the differences, whereas the remaining periodic differences are formulated as $\mathbf{r}$. $\mathbf{t}_1$ and $\mathbf{t}_2$ are the periodic constituents of the differences. For simplicity, the norm of the differences $\|\Delta\boldsymbol{\xi}\|$ rather than

the individual components of the difference vector may be expressed as a scalar function of type Eq. (4.56)

$$\|\Delta\boldsymbol{\xi}(t)\| \;=\; P^m(t) \;+\; Q(t)\,t_1(t) + R(t)\,t_2(t) \tag{4.57}$$

The dramatic increase of the differences have been expressed by $P^m(t)$ and $Q(t)$. The differences will be bounded if these components disappear. For clarity, consider the differences shown in Fig. (4.1). As seen in Fig. (4.7), a fifth order polynomial is a reasonable estimate of the deviations' trend. Although the remaining part still tends to grow as a function of time, it is significantly smaller than the total deviations. The non-linear behavior of the modulation is expected since the error of initial

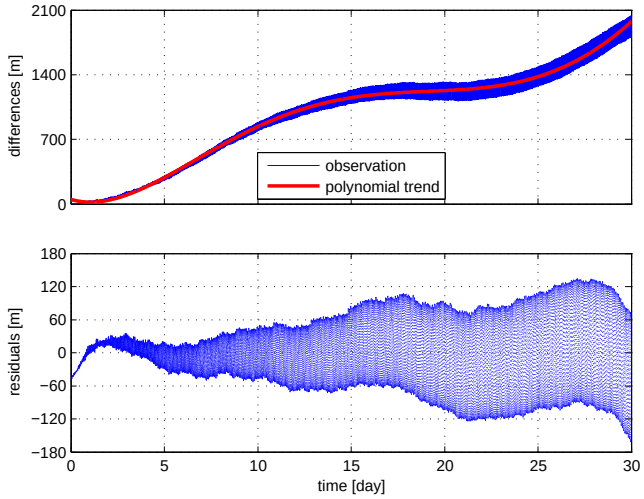


Figure 4.7: Removing the polynomial trend

conditions is non-linearly propagated on the state vector at the subsequent mesh-points.
A very simple approach for limiting the differences is to select the initial state vector such that the polynomial trend disappears. In other words, the differences can be approximately modelled as a polynomial function. The initial solution vector is nearly optimal if:

$$P^m(t) = a_0. \tag{4.58}$$

Practically, it is impossible to fulfill the condition. However, it can be minimized by selecting appropriate initial conditions. Hence, the goal is to choose the initial state vector $\boldsymbol{\xi}_0$ such that the polynomial function will be minimized over the integration period. Hence, it can be formulated as a *multi-objective* optimization problem

$$\begin{cases} \|\Delta\boldsymbol{\xi}(t)\| \;=\; \sum_{j=0}^{m} a_j(\boldsymbol{\xi}_0)t^j \\ a_j \quad \rightsquigarrow \quad \min \end{cases} \tag{4.59}$$

Replacing multiple object functions by the *sum of weighted cost function* is the most straightforward approach utilized for multi-objective optimization problem (Haupt and Haupt, 2004)

$$\begin{cases} \|\Delta\boldsymbol{\xi}(t)\| \;=\; \sum_{j=0}^{m} a_j(\boldsymbol{\xi}_0)t^j \\ \sum_{j=0}^{m} p_j a_j \quad \rightsquigarrow \quad \min \end{cases} \tag{4.60}$$

whose general representation is the *minimum norm solution*

$$\begin{cases} \|\Delta\boldsymbol{\xi}(t_i,\boldsymbol{\xi}_0)\| \;=\; \sum_{j=0}^{m} a_j(\boldsymbol{\xi}_0)t_i^j \\ \|\mathbf{x}\| \;=\; \mathbf{x}^T\mathbf{P}_\mathbf{x}\mathbf{x} \quad \rightsquigarrow \quad \min \end{cases} \tag{4.61}$$

where, the diagonal matrix $\mathbf{P}_\mathbf{x}$ is the weight matrix of the solution vector $\mathbf{x}$ (a_is). This method yields an optimal solution with optimizing the weighted sum of the object functions. The problem of the

type minimum norm could be solved analytically if the object function was explicitly written in terms of the initial solution. Since the relationship between the unknown coefficients and the initial state vector is unknown, one can alternatively use numerical optimization methods.
The GAs, as powerful tools for numerical optimization, can be implemented. The aim is to estimate the initial state vector such that the object function is minimized. The initial population can be generated by randomly altering the initial state vector which is selected equivalently to the state vector of the real orbit at time t_0. The problem is then continued till the minimum deviation is obtained.
For obtaining a better solution, the periodic components should also be considered. As already stated, the remaining differences are split into two periodic components respectively with non-linear and periodic amplitude functions. Experimentally, the non-linear amplitude function can be suitably modelled using a *hyperbolic* function, whereas the periodic part is simply replaced by a trigonometric function. Thus, the general formulation is recast into

$$Q(t)r_1(t) \;=\; \alpha_1 \sec h(\omega t)\sin(\varpi_1 t + \vartheta_1), \qquad R(t)r_2(t) \;=\; \alpha_2 \sin(\varpi_2 t + \vartheta_2) \tag{4.62}$$

Consider the previous example and remove a polynomial trend of degree 5. The achieved results of the decomposition is depicted in Fig. (4.8). The periodic differences will be bounded if the range

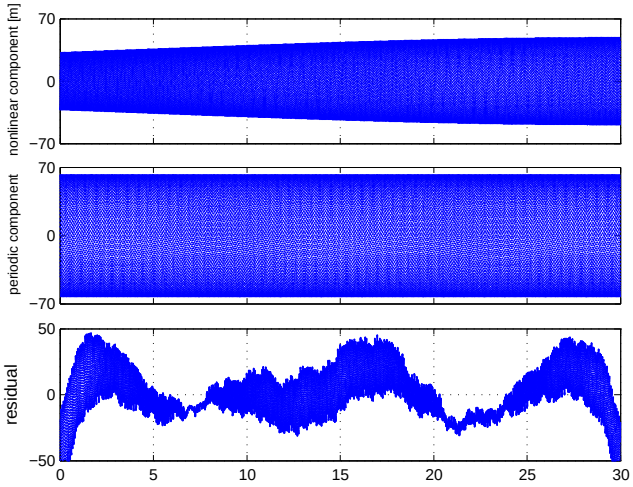


Figure 4.8: Decomposition of the periodic part

of $Q(t)$ remains constant. It can be achieved if $\alpha_1 = 0$ is fulfilled. However, it seems to be a very optimistic assumption. Instead of fulfilling the condition as a functional constraint, we can try to minimize the coefficient. Consequently, the object function is modified for limiting the non-linear periodic component besides the polynomial trend

$$\begin{cases} \|\Delta\boldsymbol{\xi}(t_i,\boldsymbol{\xi}_0)\| \;=\; \sum_{j=0}^{m} a_j(\boldsymbol{\xi}_0)t_i^j \;+\; \alpha_1(\boldsymbol{\xi}_0)\sec h(\omega(\boldsymbol{\xi}_0)t_i)\sin(\varpi_1(\boldsymbol{\xi}_0)t_i + \vartheta_1(\boldsymbol{\xi}_0)) \\ \|\mathbf{x}_a\| \;=\; \mathbf{x}_a^T\mathbf{P}_a\mathbf{x}_a \qquad \rightsquigarrow \qquad \min \end{cases} \tag{4.63}$$

where, $\mathbf{x}_a$ is the augmented unknown vector which includes both the polynomial and non-linear periodic functions coefficients (α_1). Compared to the previous formulation (Eq. 4.61), the latter formulation is rather complicated because of the non-linearity of the model. In other words, the polynomial coefficients can easily be estimated for each individual of the population using the ordinary method of least squares, whereas in the latter model it is impossible. Therefore, the solution of the non-linear model requires nested GA implementation which is very highly demanding from the computational aspect. Therefore, the periodic component with non-linear amplitude can be included if it is really necessary. Otherwise, we should try to remove a polynomial trend by selecting an appropriate set of initial conditions.
It should also be noted here that the main difficulty of using the method of weighted sum is determination of appropriate values for p_{ii}. Different weight settings may lead to completely different optimal

solutions. The Non-Pareto and Pareto-based approaches as well as the Niche induction techniques and the struggle genetic crowding algorithm are alternative schemes which have been widely used in the literature (e.g. Fonseca and Flemming, 1995; Deb, 1999; Andersson, 2001).
Compared to the least squares approach, the genetic scheme

- tries to minimize the differences by minimizing the model parameter $\mathbf{x}$,
- is more flexible due to the minimization of multiple object functions rather than the residual vector,
- has a relatively small linear system of equations as compared to the least squares approach.
- does not require partial derivatives of the differences with respect to the initial state vector.

In contrast, the least squares approach solves the problem analytically. It may then be superior since it needs less CPU time.

4.3 Summary

Replacing the full observation with its respective residual part is essential for reducing different errors in the formulation of gradiometry problem in the space-wise approach. The intersatellite rang, rage-rate and range-acceleration as well as the intersatellite velocity vector are the quantities which are involved in the gradiometry observations. In order to remove the reference field's contribution, the respective quantities should be generated using the reference field.
The observables can be directly calculated if the dynamic orbits of the GRACE satellites are generated for observation duration. Due to the mentioned advantages of the numerical integrations, they were employed in this study. First part of this chapter dedicated to different numerical orbit integration algorithms. the multi-step predictor-corrector method was selected for generating the GRACE dynamic orbit.
In order to minimize the deviation of dynamic and kinematic orbits three different schemes were developed. Among them the adaptive and least squares approaches were implemented. The least squares approach changes the V-shaped pattern of the deviations into butterfly one whereas we keep the deviations in an acceptable limits by updating the initial values in the adaptive method.

Chapter 5

Global Spherical Harmonic Analysis and Synthesis

The terms analysis and synthesis originate from Greek and mean literally *to loosen up* and *to put together* respectively (Ritchey, 1996). These terms are used within most modern scientific disciplines to denote similar investigative procedures.
The decomposition of a function on the sphere into a series of global spherical harmonics is known as *global spherical harmonic analysis* (hereafter GSHA). In other words, throughout the global spherical analysis a given function is expanded into spherical harmonics and the corresponding coefficients are estimated. In contrast, the evaluation of the function in the space domain using the coefficients is called *global spherical harmonic synthesis* (hereafter GSHS).
From theoretical point of view, since the surface spherical harmonics set up a set of 2D orthogonal basis functions, performing the analysis is relatively straightforward. In practice, only discrete sampling is possible and the integrals should be replaced by summations. The λ-summation retains the orthogonality while the θ-integration does not. Therefore, in continuous and discrete representation the solution of the λ- integration and summation are similar, whereas that of the θ is completely different.
If the data are evenly-spaced, it is of practical benefit to utilize a two-stage computation which is called *two-step analysis*. Observations are transformed to the latitude-dependent coefficients using 1D Fourier transformation in first step. Since the orthogonality is preserved, it leads to unique solution. The resultant coefficients are then mapped to the spherical harmonic coefficients in second step. There exists no unique solution for second step since there is no longer orthogonality.
In this chapter, firstly we define the GSHS and GSHA in the continuous case. Subsequently, the first step of the GSHA in the discrete case is described. Thirdly, least squares, quadrature and $2D$ Fourier methods are explained as different schemes for performing the second analysis step. Finally, an overview of the methods are given in this chapter.

5.1 Theoretical background

In the outer space (outside the attracting masses), the gravitational potential fulfills the Laplace equation (Hofmann-Wellenhof and Moritz, 2005)

$$\nabla^2 V(r,\theta,\lambda) \;=\; r^2\frac{\partial^2 V}{\partial r^2} + 2r\frac{\partial V}{\partial r} \;+\; \nabla_S^2 V \;=\; 0. \tag{5.1}$$

where, ∇_S^2 is called *surface Laplace* or *Beltrami operator*. In the latter representation, the Laplace equation is split into radial and surface components. As stated earlier, the surface harmonics are eigenfunctions of the surface Laplace operator. Therefore, the gravitational potential can be written as the product of the radial function and the surface harmonics

$$V(r,\theta,\lambda) = R(r)\bar{Y}_{nm}(\theta,\lambda). \tag{5.2}$$

Consequently, the single three-variable partial differential equation (PDE) can be reduced into the radial and surface PDEs

$$\begin{cases} r^2R'' + 2rR' = n(n+1)R \\ \nabla_S^2 \bar{Y}_{nm} = -n(n+1)\bar{Y}_{nm} \end{cases} \tag{5.3}$$

where, $n(n+1)$ is the associated eigenvalue of the surface Laplace operator [1].
The first equation is the Cauchy-Euler equation. Since $-n(n+1) < 1$, then the solution of the equation in the outer space is given as

$$R(r) = r^{-(n+1)}. \tag{5.4}$$

Therefore, the solution of Eq. (5.2) is

$$V = \frac{1}{r^{n+1}} \bar{Y}_{nm}(\theta, \lambda) \tag{5.5}$$

which is called *solid spherical harmonics of degree n*. Any linear combination of the solution is a solution of the equation since the Laplacian is a linear operator

$$V = \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \bar{Y}_{nm}(\theta, \lambda) \tag{5.6}$$

The same is true for the surface Laplace operator. Hence, the surface solution can also be expressed as a linear combination of the individual solutions. Then,

$$V(r, \theta, \lambda) = \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \sum_{m=0}^{n} \left[\bar{C}_{nm} \cos(m\lambda) + \bar{S}_{nm} \sin(m\lambda)\right] \bar{P}_{nm}(\cos\theta). \tag{5.7}$$

The unknown coefficients $\bar{C}_{nm}$ and $\bar{S}_{nm}$ are usually called *spherical harmonic coefficients* which introduce physical properties of the underlying function. In practice, the function is assumed to be bandlimited, so the upper limit (infinity) is replaced by an upper-bound N

$$V(r, \theta, \lambda) = \sum_{n=0}^{N} \frac{1}{r^{n+1}} \sum_{m=0}^{n} \left[\bar{C}_{nm} \cos(m\lambda) + \bar{S}_{nm} \sin(m\lambda)\right] \bar{P}_{nm}(\cos\theta). \tag{5.8}$$

5.1.1 Continuous representation

In the previous section, the expansion of a harmonic function in terms of *solid spherical harmonics* ($\frac{\bar{Y}_{nm}}{r^{n+1}}$) is fully described. For simplicity, consider an arbitrary bandlimited function $f(\theta, \lambda)$ which is known on the sphere. It can be expanded into a series of surface spherical harmonics (Hofmann-Wellenhof and Moritz, 2005)

$$f(\theta, \lambda) = \sum_{n=0}^{N} \sum_{m=0}^{n} \left[\bar{C}_{nm} \cos(m\lambda) + \bar{S}_{nm} \sin(m\lambda)\right] \bar{P}_{nm}(\cos\theta). \tag{5.9}$$

Knowing the coefficients, the function can be evaluated at any arbitrary point in the space domain. Evaluation of a function by synthesizing the spherical harmonic basis functions with the corresponding spherical coefficients is called *global spherical harmonic synthesis* (GSHS).
To calculate the expansion coefficients, one can use the orthogonality relations of the spherical harmonics. Multiplying both sides of Eq. (5.9) by surface spherical harmonics and integrating over the sphere yields

$$\begin{pmatrix} \bar{C}_{nm} \\ \bar{S}_{nm} \end{pmatrix} = \frac{1}{4\pi} \int_{\Omega} \int f(\theta, \lambda) \bar{P}_{nm}(\cos\theta) \begin{pmatrix} \cos m\lambda \\ \sin m\lambda \end{pmatrix} d\Omega. \tag{5.10}$$

[1] If $\mathcal{L}$ is a linear operator on a function space, then f is an eigenfunction for $\mathcal{L}$ and ν is the associated eigenvalue whenever $\mathcal{L}f = \nu f$ (Weisstein)

The main advantage of the normalized surface spherical harmonics implementation is seen in the simplicity of Eq. (5.10). The computation of the spherical harmonics expansion coefficients of a given function is usually called *Global Spherical Harmonics Analysis (*GSHA*)*.
Although both the GSHS and the GSHA formulae are straightforward, they can be made even more efficient from the computational point of view. As seen in Eqs. (5.9) and (5.10), the computation along the parallel can be carried out using one-dimensional *Fourier Transform*. Therefore, splitting the computation into two one-dimensional summations or integrations makes the whole process more convenient.
Interchanging the summation is the first step of the two-step computation. One can easily prove that the following two summations are equivalent (e.g. Rizos, 1979)

$$\sum_{n=0}^{N}\sum_{m=0}^{n} \Leftrightarrow \sum_{m=0}^{N}\sum_{n=m}^{N} \tag{5.11}$$

Using this swapped summation, the single-step GSHS (Eq. 5.9) is split into *two-step* analysis formulation (Kampes, 1998)

Step I:

$$\begin{pmatrix} A_m(\theta) \\ B_m(\theta) \end{pmatrix} = \sum_{n=m}^{N} \bar{P}_{nm} \begin{pmatrix} \bar{C}_{nm} \\ \bar{S}_{nm} \end{pmatrix} \tag{5.12}$$

Step II:

$$f(\theta,\lambda) = \sum_{m=0}^{N} \left[A_m \cos m\lambda + B_m \sin m\lambda \right]. \tag{5.13}$$

Equivalently, the single-step GSHA (Eq. 5.10) is replaced by the following two-step computation algorithm (Sneeuw, 1994)

Step I:

$$\begin{pmatrix} A_m(\theta) \\ B_m(\theta) \end{pmatrix} = \frac{1}{(1+\delta_{m0})\pi} \int_{\lambda=0}^{2\pi} f(\theta,\lambda) \begin{pmatrix} \cos m\lambda \\ \sin m\lambda \end{pmatrix} d\lambda \tag{5.14}$$

Step II:

$$\begin{pmatrix} \bar{C}_{nm} \\ \bar{S}_{nm} \end{pmatrix} = \frac{1+\delta_{m0}}{4} \int_{\theta=0}^{\pi} \begin{pmatrix} A_m(\theta) \\ B_m(\theta) \end{pmatrix} \bar{P}_{nm}(\cos\theta) \sin\theta d\theta. \tag{5.15}$$

Equations (5.13) and (5.14) can be interpreted as Fourier transform with latitude-dependent coefficients. Therefore, the computation along each parallel is treated as Fourier transform which can be performed very efficiently. Moreover, the computation along the meridians can be carried out independently for each individual order m. It consequently makes parallelization of the computation possible.
Compared to the single-step formulation, the two-step approach is more efficient. However, it cannot be implemented in reality since only discrete values of the underlying function are known. Therefore, the continuous representation should be modified correspondingly.

5.2 Discretized function with an evenly–spaced observations

Consider the function $f(\theta,\lambda)$ is sampled at some discrete points (θ_i,λ_j). In general, assume the observations are evenly distributed along the parallels

$$\lambda_j = j\Delta\lambda = j\frac{2\pi}{N_\lambda}, \qquad j = 0,1,\cdots,N_\lambda - 1.$$

For the moment, suppose the data spacing along the meridians is also constant. If the poles are also included, the latitudinal interval $\Delta\theta$ reads

$$\Delta\theta = \frac{\pi}{N_\theta - 1}.$$

If the *center points* instead of the grid points are used, the observation spacing along the meridians is

$$\Delta\theta = \frac{\pi}{N_\theta}.$$

In both cases, i runs from 0 to $N_\theta - 1$ and the total number of observations is $N_\theta\ N_\lambda$. To obtain an *equiangular grid* ($\Delta\theta = \Delta\lambda$), the number of longitudinal grid lines should be twice that of the latitudinal ones. Therefore, for the grid including the poles $N_\lambda = 2(N_\theta - 1)$ and the center point observations $N_\lambda = 2N_\theta$. It is assumed that the observations are dense enough to fulfill the sampling theorem for the recovery of signals up to N.
Representation of the discretized function in terms of the spherical harmonic coefficients is the same as the continuous representation. For instance, the function $f(\theta, \lambda)$ at the grid point (θ_i, λ_j) is synthesized as

$$f(\theta_i, \lambda_j) = \sum_{n=0}^{N} \sum_{m=0}^{n} \left[\bar{C}_{nm} \cos m\lambda_j + \bar{S}_{nm} \sin m\lambda_j\right] \bar{P}_{nm}(\cos\theta_i) \tag{5.16}$$

which is the same as Eq. (5.9). Therefore, the two-step discrete synthesis can be performed analogously

Step I:

$$\begin{pmatrix} A_m(\theta_i) \\ B_m(\theta_i) \end{pmatrix} = \sum_{n=m}^{N} \bar{P}_{nm}(\cos\theta_i) \begin{pmatrix} \bar{C}_{nm} \\ \bar{S}_{nm} \end{pmatrix} \tag{5.17}$$

Step II

$$f(\theta_i, \lambda_j) = \sum_{m=0}^{N} \left[A_m(\theta_i) \cos m\lambda_j + B_m(\theta_i) \sin m\lambda_j\right] \tag{5.18}$$

From the computational aspect, matrix-vector representation as an alternative is very useful for discrete GSHS computation. Consider the matrix $\mathbf{P}$ and the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{s}$ and $\mathbf{f}$ which are defined as follows:

$$\mathbf{c} = \begin{pmatrix} \bar{C}_{mm} \\ \bar{C}_{m+1m} \\ \vdots \\ \bar{C}_{Nm} \end{pmatrix}_{N-m+1\times 1} \qquad \mathbf{s} = \begin{pmatrix} \bar{S}_{mm} \\ \bar{S}_{m+1m} \\ \vdots \\ \bar{S}_{Nm} \end{pmatrix}_{N-m+1\times 1}$$

$$\mathbf{P} = \begin{pmatrix} P_{mm}(t_0) & P_{m+1m}(t_0) & \cdots & P_{Nm}(t_0) \\ P_{mm}(t_2) & P_{m+1m}(t_2) & \cdots & P_{Nm}(t_2) \\ \vdots & \cdots & \ddots & \vdots \\ P_{mm}(t_{N_\theta-1}) & P_{m+1m}(t_{N_\theta-1}) & \cdots & P_{Nm}(t_{N_\theta-1}) \end{pmatrix}_{N_\theta \times N-m+1}$$

$$\mathbf{a}_m = \begin{pmatrix} A_m(\theta_0) \\ A_m(\theta_2) \\ \vdots \\ A_m(\theta_{N_\theta-1}) \end{pmatrix}_{N_\theta\times 1} \qquad \mathbf{b}_m = \begin{pmatrix} B_m(\theta_0) \\ B_m(\theta_2) \\ \vdots \\ B_m(\theta_{N_\theta-1}) \end{pmatrix}_{N_\theta\times 1}$$

Consequently, the discrete GSHS can be represented as two successive linear mappings, each of which is equivalent to Eq. (5.17) and (5.18)

Step I:

$$\begin{pmatrix} \mathbf{a}_m \\ \mathbf{b}_m \end{pmatrix} = \mathbf{P} \begin{pmatrix} \mathbf{c} \\ \mathbf{s} \end{pmatrix} \tag{5.19}$$

Step II:

$$\mathbf{F} = \mathbf{AC} + \mathbf{BS} \tag{5.20}$$

where,

$$\mathbf{F} = [f_{ij}]_{N_\theta\times 2N} \qquad f_{ij} = f(\theta_i, \lambda_j)$$

$$\mathbf{A} = [\ \mathbf{a}_0 \quad \mathbf{a}_1 \quad \cdots \quad \mathbf{a}_N\ ] \qquad \mathbf{B} = [\ \mathbf{b}_0 \quad \mathbf{b}_1 \quad \cdots \quad \mathbf{b}_N\ ]$$

and

$$\mathbf{C} = [c_{pq}]_{N+1\times 2N} \qquad c_{pq} = \cos(p-1)\lambda_{q-1}$$
$$\mathbf{S} = [s_{pq}]_{N+1\times 2N} \qquad s_{pq} = \sin(p-1)\lambda_{q-1}$$

where, $p = 1, 2, \cdots, N+1$ and $q = 1, 2, \cdots, 2N$. Equations (5.19) and (5.20) are set just based on the matrix and vector multiplication. The computation can be accelerated by vectorization of the computation. It yields a relatively fast computation algorithm which is comparable to the Fourier based method (Sneeuw, 1994).

As seen in Eqs. (5.16),(5.17) and (5.18), both in the one- and the two-step discrete GSHS are carried out independently from the observation grid. However, the continuous GSHA has been formulated based on the orthogonality of the surface harmonics which are defined continuously on the sphere. For the discrete representation, the underlying functions are sampled at some finite locations. Therefore, the GSHA should be revisited accordingly. In other words, the observation grid should be defined such that discrete orthogonality relationship can be set up in order to achieve maximum computational efficiency.

Along the parallels, every surface harmonic is a periodic function of period 2π. In the discrete representation, N_λ equally spaced points $\lambda_j = j\Delta\lambda$ define a discrete function of period $N_\lambda = 2N$ which can be represented as a vector of length N_λ. Correspondingly, the inner product of the discrete periodic functions is defined as the inner product of two vectors (Kampes, 1998)

$$\sum_{j=0}^{N_\lambda-1} \cos m_1\lambda_j \cos m_2\lambda_j = \begin{cases} 0 & \text{if } m_1 \neq m_2 \\ 2N & \text{if } m_1 = m_2 = 0 \text{ or } \frac{N_\lambda}{2} \\ N & \text{elsewhere} \end{cases} \tag{5.21}$$

$$\sum_{j=0}^{N_\lambda-1} \sin m_1\lambda_j \sin m_2\lambda_j = \begin{cases} 0 & \text{if } m_1 \neq m_2 \\ 0 & \text{if } m_1 = m_2 = 0 \text{ or } \frac{N_\lambda}{2} \\ N & \text{elsewhere} \end{cases} \tag{5.22}$$

$$\langle \sin m_1\lambda, \cos m_2\lambda \rangle = \sum_{j=0}^{N_\lambda-1} \sin m_1\lambda_j \cos m_2\lambda_j = 0 \quad \forall\, m_1 \text{ and } m_2 \tag{5.23}$$

Consequently, similarly to the continuous case the latitude dependent Fourier coefficients $a_m(\theta_i)$ and $b_m(\theta_i)$ can be easily computed as (Sneeuw, 1994)

$$\begin{pmatrix} A_m(\theta_i) \\ B_m(\theta_i) \end{pmatrix} = \frac{1}{N(1+\delta_{m0}+\delta_{mN})} \sum_{j=0}^{2N-1} f(\theta_i, \lambda_j) \begin{pmatrix} \cos m\lambda_j \\ \sin m\lambda_j \end{pmatrix}. \tag{5.24}$$

Equivalently, it can be represented using vector-matrix notation as

$$\begin{pmatrix} \mathbf{a}_m \\ \mathbf{b}_m \end{pmatrix} = \frac{1}{N(1+\delta_{m0}+\delta_{mN})} \mathbf{F} \begin{pmatrix} \mathbf{c}_m^T \\ \mathbf{s}_m^T \end{pmatrix}, \tag{5.25}$$

where, $\mathbf{c}_m$ and $\mathbf{s}_m$ are the m^{th} row of matrices $\mathbf{C}$ and $\mathbf{S}$ respectively.

However, there is no way to define the latitudinal distribution of data which retains the surface harmonics orthogonality along the meridians

$$\langle \bar{P}_{n_1 m}(\cos\theta_i), \bar{P}_{n_2 m}(\cos\theta_i) \rangle = \begin{cases} 0 & n_2 - n_1 \text{ odd} \\ ? & \text{elsewhere} \end{cases} \tag{5.26}$$

Therefore, mapping the latitude dependent coefficients $\mathbf{a}_m$ and $\mathbf{b}_m$ into $\bar{C}_{nm}$ and $\bar{S}_{nm}$ is not uniquely defined.

The followings are the relevant computational schemes for performing the second step of the discrete GSHA.

5.2.1 Least squares method

As shown in Eq. (5.19), mapping $\mathbf{c}$ and $\mathbf{s}$ into $\mathbf{a}_m$ and $\mathbf{b}_m$ is of a linear type which has been expressed by the mapping matrix $\mathbf{P}$. Therefore, the inverse mapping can also be represented as a vector-valued linear function

$$\begin{pmatrix}\widehat{\mathbf{c}}\\ \widehat{\mathbf{s}}\end{pmatrix} = \mathbf{P}^{-}\begin{pmatrix}\mathbf{a}_m\\ \mathbf{b}_m\end{pmatrix} \tag{5.27}$$

where, $\mathbf{P}^{-}$ stands for some generalized inverse (Sneeuw, 1994). Substituting Eq. (5.19) into Eq. (5.27) yields

$$\begin{pmatrix}\widehat{\mathbf{c}}\\ \widehat{\mathbf{s}}\end{pmatrix} = \mathbf{P}^{-}\mathbf{P}\begin{pmatrix}\mathbf{c}\\ \mathbf{s}\end{pmatrix}. \tag{5.28}$$

The matrix $\mathbf{P}^{-}\mathbf{P}$ is usually called *resolution in the model space*. Alternatively, one can set up an equivalent criterion called *resolution in the observation space*

$$\begin{pmatrix}\widehat{\mathbf{a}}_m\\ \widehat{\mathbf{b}}_m\end{pmatrix} = \mathbf{P}\mathbf{P}^{-}\begin{pmatrix}\mathbf{a}_m\\ \mathbf{b}_m\end{pmatrix}. \tag{5.29}$$

where, the matrix $\mathbf{P}\mathbf{P}^{-}$ is the resolution in the observation space. In order to obtain an unbiased estimation, the resolution matrix either in the model or in the observation space should be close to the identity matrix $\mathbf{I}$ (Gubbins, 2004).
The maximum length of a single vector $\mathbf{c}$ or $\mathbf{s}$ equivalent to the maximum number of the harmonic coefficients with constant order is equal to $\max(N-m+1) = N+1$. The number of equations should always be more than or equal to number of the unknowns ($N_\theta \geq N+1$), which is a constraint on the number of observation points along the meridians.
Implementation of the criteria in the observation space is equivalent to the solution of the over-determined linear system of equations in the sense of *the method of least squares*

$$\begin{pmatrix}\widehat{\mathbf{c}}\\ \widehat{\mathbf{s}}\end{pmatrix} = \underbrace{\left(\mathbf{P}^T\mathbf{P}\right)}_{\mathbf{N}}{}^{-1}\mathbf{P}^T\begin{pmatrix}\mathbf{a}_m\\ \mathbf{b}_m\end{pmatrix}. \tag{5.30}$$

Entries of the design matrix of normal equations ($\mathbf{N}$) are the inner products of Legendre functions in the discrete case

$$\mathbf{N}(n_1, n_2) = \sum_{i=0}^{N_\theta - 1} \bar{P}_{n_1 m}(t_i)\bar{P}_{n_2 m}(t_i). \tag{5.31}$$

Therefore, the entries represent the orthogonality of the functions in the discrete case. Theoretically, if the orthogonality had been retained the matrix $\mathbf{N}$ would have been equal to $2(2-\delta_{m0})\mathbf{I}$. Unfortunately, the matrix is far off from the theoretically expected values. For instance, the normal matrix of the least squares scheme for $N_\theta = 120$, $N = 120$ and $m = 0$ is depicted in Fig. (5.1). It is clearly visible in the structure of the matrix that

- the diagonal elements are significantly larger than $2(2-\delta_{m0})$.
- the diagonal elements have widely leaked into the neighboring off-diagonal elements.
- the validity of the orthogonality relationship is clearly visible for $n_2 - n_1 =$ odd.

Nevertheless, the resolution matrix perfectly fulfills the criterion $\mathbf{P}\mathbf{P}^{-} = \mathbf{I}$, see Fig. (5.2).
In oder to clarify the effect of normal matrix deviation from the theoretically expected structure, the incremental potential $\delta V = V^{120}_{\text{EGM96}} - V^{120}_{\text{EIGEN2}}$ is simulated on a sphere on a regular grid $1.5^\circ \times 1.5^\circ$. The recovered spectrum is compared to the original one and the differences are shown in Fig. (5.3). The differences are about the level of the original signal. They are more pronounced for the zonal harmonics. The effect of non-orthogonality is large so they cannot be ignored. Obtaining an inaccurate solution is the main disadvantage of least squares. Nevertheless, it is very flexible for handling various kinds of observations with an arbitrary distribution. Furthermore, the GSHA can be implemented on observations with an unevenly latitudinal distribution using the least squares method.

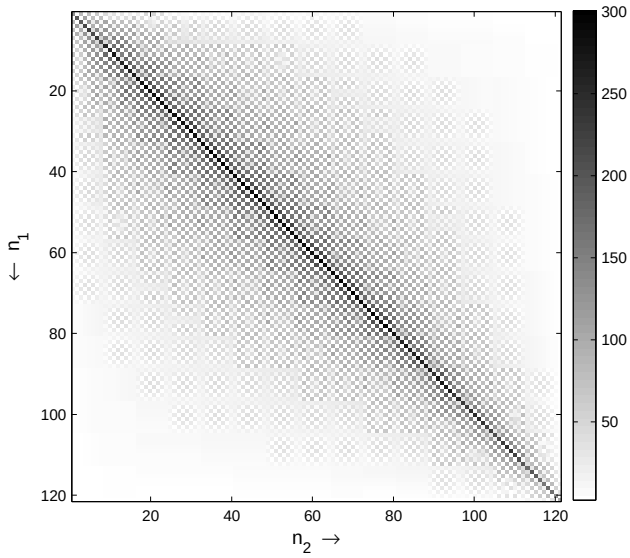


Figure 5.1: Structure of normal matrix in the least squares approach ($N_\theta = 120$, $N = 120$, $m = 0$)

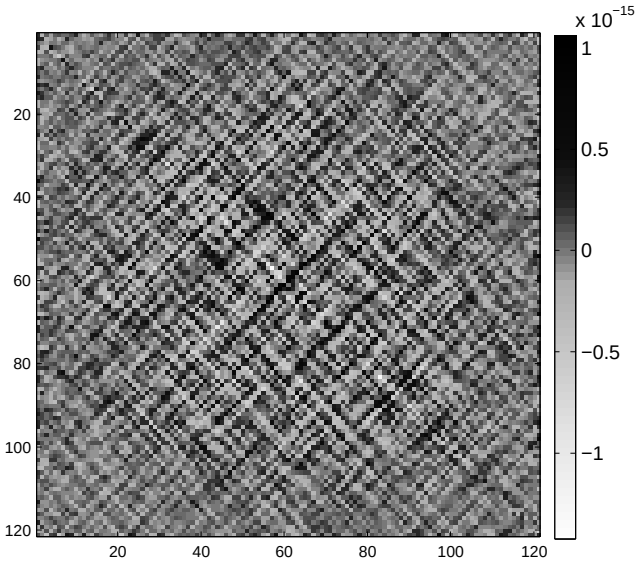


Figure 5.2: The resolution matrix in the observation space minus identity matrix ($N_\theta = 120$, $N = 120$, $m = 0$)

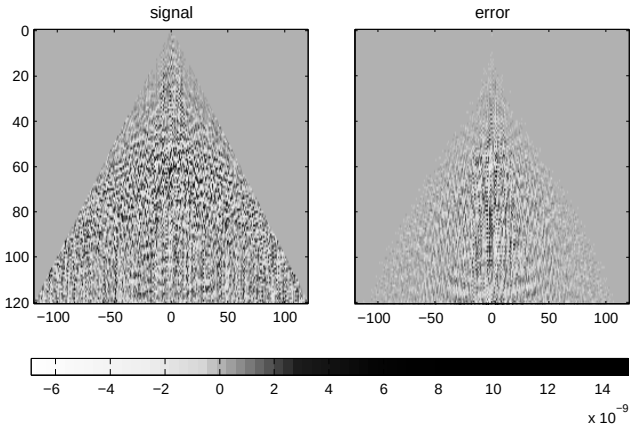


Figure 5.3: Original signal and error of the recovered spectrum using least squares scheme

In order to perform the mapping efficiently, different modifications have been introduced. Compared to the least squares method, the followings are well-known alternative methods which yield comparatively better solutions.

5.2.2 Approximate quadrature/weighted least squares

As seen in Fig. (5.1), both diagonal and off-diagonal elements of the normal matrix $\mathbf{N}$ differ from the identity matrix. By neglecting the non-zero off-diagonal elements, one can improve the normal equations by down-weighting the observations. In other words, instead of setting up the normal matrix with unit weight matrix we can assume the data points are independent but have different weights

$$\mathbf{N}_w = \mathbf{P}^T\mathbf{W}\mathbf{P} \tag{5.32}$$

where, the weight matrix $\mathbf{W}$ is defined as

$$w_{ij} = \kappa \begin{cases} w_i & i = j \\ 0 & i \neq j \end{cases}$$

with κ as variance factor. The coefficients w_i can be considered as the additional parameters which absorb the effect of non-orthogonality of the Legendre functions in the discrete case. Hence, they should be estimated such that the normal matrix is close to the identity matrix as much as possible.

The problem can be deemed as an optimization problem with a solution space of dimension N_θ. The goal is then to find the optimum solution which yields $\mathbf{N}_w = \kappa\mathbf{I}$. The GAs as a powerful multidimensional optimizer can be effectively employed. Nevertheless, it might be rather time consuming due to high-dimensional search space.

Alternatively, an estimate of the unknowns can be derived analytically by discretization of integral formulas corresponding to the continuous GSHA (Eq. 5.15)

$$\begin{pmatrix} \bar{C}_{nm} \\ \bar{S}_{nm} \end{pmatrix} \doteq \frac{1+\delta_{m0}}{4} \sum_{i=0}^{N_\theta-1} \begin{pmatrix} A_m(\theta_i) \\ B_m(\theta_i) \end{pmatrix} \bar{P}_{nm}(\cos\theta_i) \sin\theta_i \Delta\theta \tag{5.33}$$

or in vector-matrix notation,

$$\begin{pmatrix} \widehat{\mathbf{c}} \\ \widehat{\mathbf{s}} \end{pmatrix} = \frac{1+\delta_{m0}}{4} \mathbf{P}^T\mathbf{W} \begin{pmatrix} \mathbf{a}_m \\ \mathbf{b}_m \end{pmatrix}. \tag{5.34}$$

for an evenly distributed data points, it reads (Sneeuw, 1994)

$$\kappa = \Delta\theta = \frac{\pi}{N_\theta}; \qquad w_i = \sin\theta_i \tag{5.35}$$

whereas assuming a uniformly random distribution of observations leads to (Ellsaesser, 1966)

$$\kappa = \frac{2}{\sum_{k=1}^{N_\theta} \sin\theta_k}; \qquad w_i = \sin\theta_i. \tag{5.36}$$

In both cases, $\sin\theta_i$ is defined as the weight of the observations located on the colatitude θ_i. By comparing Eqs. (5.30) and (5.34), one can conclude that the normal matrix has been replaced by $\frac{(1+\delta_{m0})}{4}\mathbf{I}$.
As seen in Fig. (5.4), by introducing the weight matrix the normal matrix was scaled to the identity matrix. To compare the effect of the data distribution as well as the weighting methods, evenly-spaced as well as randomly uniform distributed data were employed. Using both weighing methods gave nearly the same results. However, deviations of the normal matrix corresponding to the randomly distributed data were significant for both weighting methods. Although the implementation of the

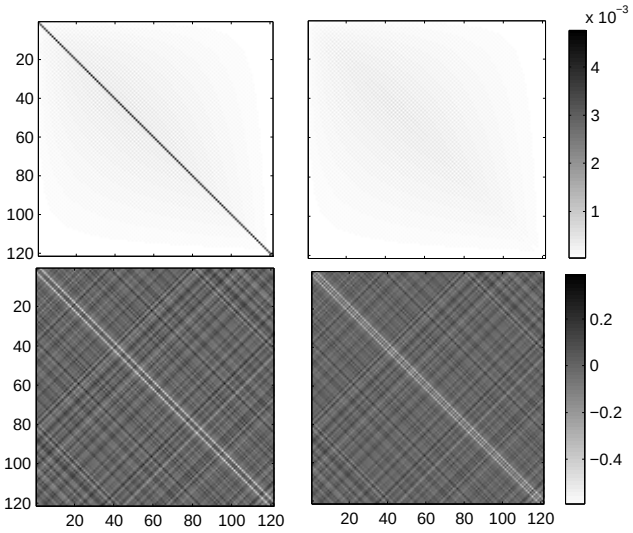


Figure 5.4: Normal matrix deviation from the identity matrix for evenly spaced and randomly uniform distributed data (UL: evenly spaced data ($\Delta\theta = \frac{\pi}{N_\theta}$) UR: evenly spaced data ($\Delta\theta = \frac{2}{\sum_{k=0}^{N_\theta-1}\sin\theta_k}$) LL: randomly uniform distributed data ($\Delta\theta = \frac{\pi}{N_\theta}$) LR: randomly uniform distributed data ($\Delta\theta = \frac{2}{\sum_{k=0}^{N_\theta-1}\sin\theta_k}$))

weight matrix derived from discretization of the continuous formulas has dramatically improved the normal matrix, it still suffers from non-orthogonality. The GSHA of observations of potential type utilized in previous example was repeated using the weighted least squares method. As shown in Fig. (5.5), the down-weighting algorithm effectively reduces error of the recovered spectrum. However,

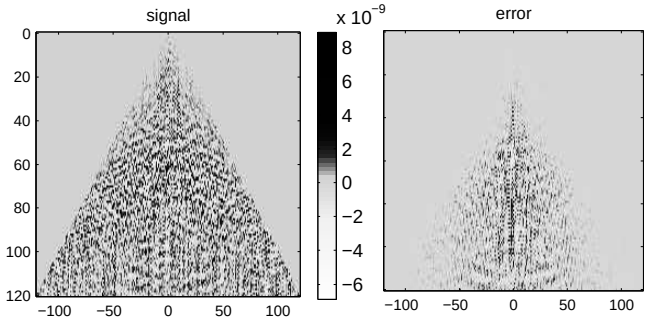


Figure 5.5: Original signal and error of the recovered spectrum using weighted least squares scheme

the recovered signal was erroneous even for the ideal situation (absence of errors). We proceed the

algorithm modification with *Neumann's methods* which are also called *exact methods* due to their perfect performance.

5.2.3 Neumann's methods

In the weighted least squares approach, the integral formulas (Eq. 5.15) was discretized and the second step of GSHA was evaluated in a weighted summation form. Although significant improvement was achieved it was erroneous. Neumann (1838) showed that an appropriate set of weights can yield full orthogonality for the discrete case. In other words, Eq. (5.15) can be replaced by a weighted summation defined as:

$$\begin{pmatrix} \bar{C}_{nm} \\ \bar{S}_{nm} \end{pmatrix} = \frac{1+\delta_{m0}}{4} \sum_{i=0}^{N_\theta - 1} w_i^{\mathrm{Neu}} \bar{P}_{nm}(\cos\theta_i) \begin{pmatrix} a_m(\theta_i) \\ b_m(\theta_i) \end{pmatrix} \tag{5.37}$$

where, w_i^{Neu} are the Neumann weights which should fulfill the following relationships for obtaining perfect orthogonality (Sneeuw, 1994)

$$\sum_{i=1}^{N_\theta} w_i^{\mathrm{Neu}} t_i^n = \int_{-1}^{1} t^n \, dt = \begin{cases} \frac{2}{n+1} & n \text{ even} \\ 0 & n \text{ odd} \end{cases} \tag{5.38}$$

where, $t_i = \cos\theta_i$ and n runs over $\{0, 1, \cdots, N_\theta - 1\}$. Validity of the given orthogonality relationship has been proved in (e.g. Neumann, 1887; Lense, 1954; Sneeuw, 1994).
Equation (5.38) is a determined linear system of equations

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ t_0 & t_1 & t_2 & \cdots & t_{N_\theta-1} \\ t_0^2 & t_1^2 & t_2^2 & \cdots & t_{N_\theta-1}^2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ t_0^{N_\theta-1} & t_1^{N_\theta-1} & t_2^{N_\theta-1} & \cdots & t_{N_\theta-1}^{N_\theta-1} \end{pmatrix} \begin{pmatrix} w_0^{\mathrm{Neu}} \\ w_1^{\mathrm{Neu}} \\ w_2^{\mathrm{Neu}} \\ \vdots \\ w_{N_\theta}^{\mathrm{Neu}} \end{pmatrix} = \begin{pmatrix} \frac{2}{1} \\ 0 \\ \frac{2}{3} \\ \vdots \\ \frac{2}{N_\theta - 1} \end{pmatrix} \tag{5.39}$$

Design matrix of the system is called *Vandermonde matrix* which is nonsingular for N_θ distinct points

$$\Delta(t_0, t_1, \cdots, t_{N_\theta-1}) = \prod_{\substack{i,j \\ j>i}}^{N_\theta - 1} (t_j - t_i). \tag{5.40}$$

Therefore, the system has a unique solution in the N_θ variables.
It should be noted here that the normal matrix corresponding to the Neumann weight matrix ($\mathbf{P}^T\mathbf{W}^{\mathrm{Neu}}\mathbf{P}$) is replaced by $2(2-\delta_{m0})\mathbf{I}$. Entries of the normal matrix are the products of Legendre functions which are in most of degree $2N$. Therefore, it is possible to expect the substitution of the integration by summation to be exact if the degree of the approximating polynomial is equal to $2N$ at least. Therefore, $N_\theta = 2N + 1$ is the minimum number of latitudinal observation points for uniquely determining the Neumann weights. There is no restriction on the selection of parallels except they should be $2N + 1$ distinct parallels.
Linear systems with a Vandermonde coefficient matrix are known to be numerically ill-conditioned even for a relatively small value of N_θ (see Press et al., 1992, sec. 3.5). Consequently, they cannot be calculated using linear inversion of the system. An alternative scheme for computation of the Neumann weights has been addressed in Sneeuw (1994).
Moreover, it should be considered that $2N(2N + 1)$ observations are required for the determination of $(N+1)^2$ unknown spherical harmonic coefficients. It is nearly four times the number of unknowns. This is a disadvantage of this method which is usually known as the *First Neumann's Method*.
Compared to the weighted least squares method, the first Neumann's method has some additional constraint which leads to an exact solution. In both methods, there is no restriction on the latitudinal distribution of observations. However, the number of observations in the first Neumann's method is

nearly two times the number of observations to be used in the weighted least squares scheme.
In the context of numerical integration, the number of data points can be reduced if the points for evaluation are selected in an optimal, rather than equally spaced, manner. The nodes $\theta_0, \theta_1, \cdots, \theta_{N_\theta - 1}$ in the integration interval $[0, \pi]$ and weights w_i^{Neu}, are chosen to minimize the expected error obtained in performing the approximation

$$\int_0^\pi \begin{pmatrix} A_m(\theta) \\ B_m(\theta) \end{pmatrix} \bar{P}_{nm}(\cos\theta) \sin\theta d\theta \approx \sum_{i=0}^{N_\theta - 1} w_i^{\text{Neu}} \bar{P}_{nm}(\cos\theta_i) \begin{pmatrix} A_m(\theta_i) \\ B_m(\theta_i) \end{pmatrix} \tag{5.41}$$

Neumann (1887) showed that if the roots of the Legendre function of degree $N+1$ are selected as θ_is, the number of parallels is reduced from $2N+1$ to $N+1$. This intelligent selection of θ-grid which is called *Gauss grid* yields an exact orthogonality for the Legendre functions using a minimum number of discrete points on the sphere.
Besides the computation of the Neumann's weights, the abscissas t_i should also be calculated with sufficient precision. Different methods have been proposed in the literature for the computation of the abscissas (e.g. Davis and Rabinowitz, 1984; Press et al., 1992; S. Zhang, 1996).
A comparison between the equally-spaced and the Gauss grids is shown in Fig. (5.6). The difference between the two grids decreases as the maximum spherical harmonic degree increases.

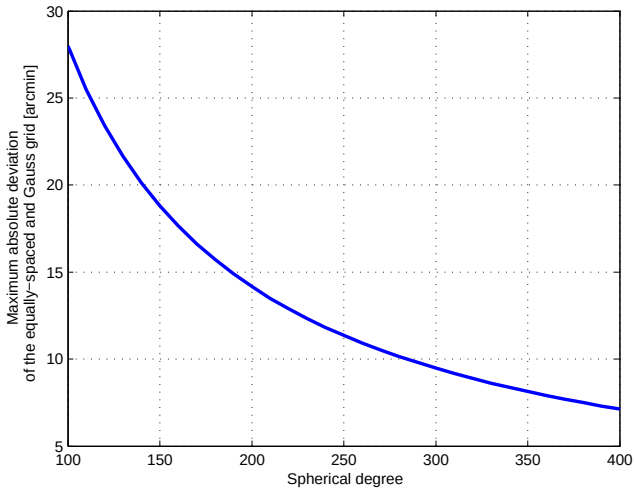


Figure 5.6: Maximum absolute deviation of the equally-spaced and the Gauss grid

If the zeros of the Legendre polynomial of degree $N+1$ are computed, the Neumann weights can be computed using either Eq. (5.39) or one of the following formulas (Atkinson, 1993, Pg. 187)

$$w_i^{\text{Neu}} = \frac{-2}{(n+2)P'_{L+1}(t_i)P_{L+2}(t_i)} \tag{5.42}$$

or (Krylov, 1962)

$$w_i^{\text{Neu}} = \frac{2}{(1-t_i^2)(P'_{N+1}(t_i))^2} \tag{5.43}$$

where, the prime shows the derivative of P_{N+1} with respect to t. In Sneeuw (1994), the following formula has been presented as an alternative:

$$w_i^{\text{Neu}} = \frac{2(1-t_i^2)}{\left[(N+1)P_N(t_i)\right]^2} \tag{5.44}$$

in which just the function itself rather than the derivative is used.
As soon as the θ-grid and computed the corresponding weights are defined, the second step of GSHA can be performed straightforwardly. This method is usually called the *Second Neumann's Method*

in which both the abscissas and the weights are subject to constraints. As previously mentioned, reducing the number of observations by a factor of two is the main advantage of using the Gauss grid. Let us consider the previously utilized potential data set, except that this time the data are given on the Gauss grid. As shown in Fig. (5.7), a perfect result is obtained by implementation of the second Neumann's method. To compare the effect of introducing different weights on discrete orthogonality

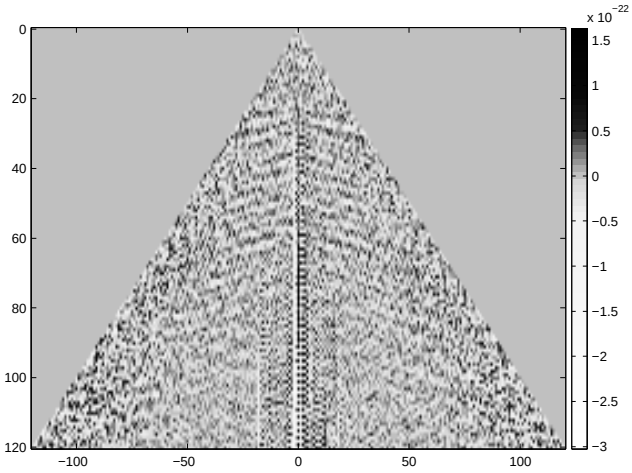


Figure 5.7: Error spectrum of the spherical harmonics recovered using the second Neumann's method

relationship, the normal matrix is depicted in Fig. (5.8). As expected, the second Neumann's method obtained the best results followed by weighted least squares.

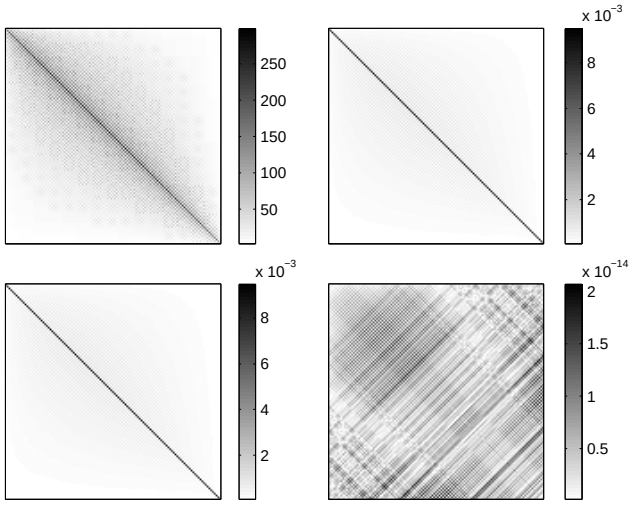


Figure 5.8: Deviation of the normal matrix from the Identity matrix (UL: Least squares UR: Weighted least squares ($w_i = \frac{\Delta\theta}{N}\sin\theta_i$) LL: Weighted least squares ($w_i = \frac{2}{\sum_{k=0}^{N_\theta-1}\sin\theta_k}\sin\theta_i$) LR: Second Neumann's method)

Eventually, in Table (5.1) a comparison of different methods is listed. The number of observations, θ-grid, weights and the main advantage and disadvantage of the methods are compared.

5.3 Summary

Representation of the geopotential and the geomagnetic fields are two interesting applications of the spherical harmonics. In this chapter, we reviewed theoretical background of the spherical harmonics.

Table 5.1: Comparison of the methods applied for discrete GSHA

Method	N_θ	θ-grid	w_i	advantage	disadvantage
LS	$N+1$	arbitrary	1	flexible	no orthogonality
WLS	$N+1$	arbitrary	Eq. (5.35)	arbitrary grid	approximate orthogonality
Neumann I	$2N+1$	arbitrary but distinct	Eq. (5.39)	perfect orthogonality	No. of observations!
Neumann II	$N+1$	Gauss grid	Eq. (5.44)	perfect orthogonality	restricted θ-grid

Furthermore, continuous representation of GSHA and GSHS were formulated.
For practical applications, discrete representation of the GSHA and GSHS is needed since observations are always given in some discrete points. Implementation of the GSHS is independent of data type, however, GSHA is completely different form continuous to discrete type. In discrete representation, lack of orthogonality of the surface spherical harmonics in θ-direction makes the implementation of the GSHA rather complicated.
Least squares, approximate quadrature and the Neumann methods have been introduced for analysis of discrete data on a sphere. As compared in Table (5.1), the least squares method yields the worst results since the lack of orthogonality is completely ignored. Introducing the approximate weights in the weighted least squares approach improves the accuracy of the recovered coefficients to a certain extent. On the other hand, Neumann's methods can successfully handel the problem by imposing constraints on weighting procedure, the number and the location of observations. Eventually, an exact solution can be obtained using minimum number of observations by implementation of the second Neumann method.

Chapter 6

Mathematical Formulation of the LL-SST Problem

Gravity field recovery using space-borne observations has entered a new era by launching the Earth's gravity field dedicated missions. The German spacecraft CHAMP as the first satellite of a new generation has portrayed the gravity field with much more detail and higher precision. It was followed by the joint German-American low-low SST mission GRACE which is capable of measuring the gravity field even with higher resolution and accuracy. Equipping the pair of spacecraft with highly accurate intersatellite ranging system enables the mission to detect the temporal variation of the field as well.
Having the opportunity of being continuously tracked by the positioning satellites GPS enables us to track LEO spacecrafts nearly continuously in the space. From the kinematical point of view, the spacecrafts can then be considered as the gravity sensors since their motion is mainly governed by the gravity field. Due to attenuation of the field with altitude this high-low configuration has a limited resolution. In order to extend the limit, relative motion of the LEO satellites should be measured. This idea has been realized in the GRACE mission using two coplanar LEO satellites whose relative distance is measured with an ultra-precise microwave system.
Combination of the high-low with low-low SST information observed by the GRACE satellites provides us an invaluable source of information. The mission is considered as the differential gravimeter and the respective mathematical model is addressed in the first section. Alternatively, the mission can be viewed as a one-axis long arm gradiometer. The second part of this chapter is dedicated to the formulation of GRACE as a gradiometer. The potential problems are explained and some solutions are introduced to tackle the problems.

6.1 Satellite to satellite tracking

In this section, a link between GRACE observables and the geopotential spherical harmonic coefficients is established. The GRACE mission consists of two identical LEO satellites which fly on an identical orbit; one approximately 220 ± 50 km ahead of the other. Accurate geodetic positions of each satellite in terms of the geocentric position vector ($\mathbf{r}$) are provided by Black-Jack GPS receivers mounted on the satellites. In addition, complete velocity vectors $\dot{\mathbf{r}}$ of the two satellites can also be numerically derived from GPS observations. Thus, complete state vectors consisting of the position and velocity vectors of the two satellites will be available. Indeed, the intersatellite biased range ϱ and the range-rate $\dot{\varrho}$ can be measured accurately by means of the K-band ranger.
The position vectors and their time derivatives yield the time series of the intersatellite position and velocity vector $\Delta\mathbf{r}$, $\Delta\dot{\mathbf{r}}$. Furthermore, the intersatellite acceleration along the line of sight ($\ddot{\varrho}$) will be numerically obtained from the intersatellite range-rate during the data pre-processing phase. The relationship between the range, range-rate and range-acceleration and relative velocity and acceleration vectors can be derived following (Rummel et al., 1978):

$$\langle \frac{\Delta\mathbf{r}}{\varrho}, \Delta\ddot{\mathbf{r}} \rangle \;=\; \ddot{\varrho} \;+\; \frac{1}{\varrho}\left[\|\Delta\dot{\mathbf{r}}\|^2 - \dot{\varrho}^2\right]. \tag{6.1}$$

or equivalently

$$\Delta\ddot{\mathbf{r}}^{\mathrm{LOS}} = \ddot{\varrho} + \frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho} - \frac{\dot{\varrho}^2}{\varrho} \tag{6.2}$$

where $\Delta\ddot{\mathbf{r}}^{\mathrm{LOS}}$ is the projection of the relative acceleration vector onto the *line of sight* vector ($\mathbf{e}$). In general, the observed acceleration vector consists of the gravitational and the nuisance accelerations. Assume the nuisance component which includes the gravitational components (except the Earth's gravitational acceleration) and non-gravitational have been removed. Then, the remaining part, i.e. the Earth's gravitational acceleration can be expressed as a linear functional on the geopotential

$$\mathcal{L}^{\mathrm{LL}}V = \ddot{\varrho} + \frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho} - \frac{\dot{\varrho}^2}{\varrho}. \tag{6.3}$$

where $\mathcal{L}^{\mathrm{LL}}V$ has been defined in Eq. (2.11). Equation (6.3) links the GRACE observable to the geopotential in terms of acceleration differences. With this formulation, one can interpret the GRACE mission as a differential gravimeter. It is the simplest form of the formulation which has been widely studied for three decades (e.g. Wolff, 1969; Rummel et al., 1978; Rummel, 1980; Blaha, 1992; Jekeli, 1999; Garcia, 2002, 2003; Keller and Sharifi, 2005).
Representation of the LL-SST as a linear functional on the gravity field ($\mathcal{L}^{\mathrm{LL}}$) connects straightforwardly the GRACE observable to the spherical harmonics. As mentioned in Chapter 2, it is the well-suited form for the brute-force approach. However, it will result in a huge linear system of equations which cannot be handled computationally.

6.2 Satellite gradiometry

In order to alleviate the mentioned computational difficulties the spacewise approach can be employed. In contrast to the timewise approach, in the spacewise approach the observables are considered as evaluational functionals which are evaluated at some definite locations. Therefore, it is more convenient to express observables as functions with explicit connection to the field at the evaluation point.
Switching from the first partial derivative of the Earth's gravity field to the second is an alternative formulation of the LL-SST which suits better to the spacewise approach. In other words, combining the two SST concepts, as shown in Fig. 6.1, makes the twin satellites to appear as a very accurate one-component gradiometer. One can show that the accuracy of this virtual one-dimensional gradiometer is about $10^{-6}\,\mathrm{E}/\sqrt{\mathrm{Hz}}$, see Rummel (2003) for more details. This unique characteristic of GRACE is a motivation to switch from the first derivatives of the gravitational potential to the second derivatives of the field. In other words, we write the observation equation (Eq. 6.3) as a function of the gravitational acceleration gradient components instead of the gravitational potential gradient.

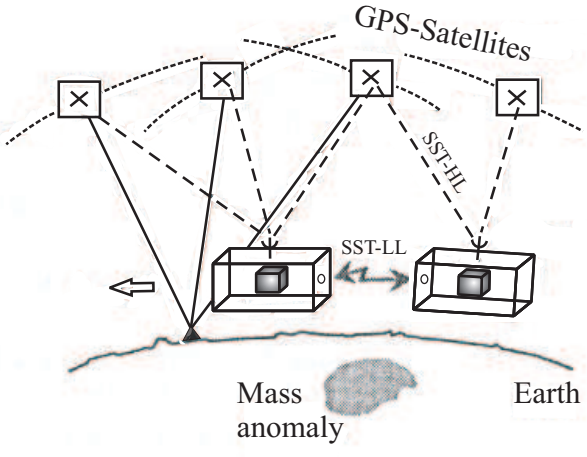


Figure 6.1: Gradiometry with the GRACE twin satellites (from Rummel et al. (2002))

To derive the respective mathematical formulas, we expand the gravitational acceleration at the two satellites' respective positions around the mid-point using Taylor expansion. Subtracting the resultant expression yields (Sharifi and Keller, 2005):

$$\Delta\Gamma = \sum_{j=1}^{\infty} \frac{2^{2-2j}}{(2j-1)!} (\nabla^{2j-1} \otimes \Gamma_{\mathrm{B}}^{T}) \cdot \Delta\mathbf{r}^{2j-1}, \tag{6.4}$$

where, $\otimes$ is Kronecker product symbol and Γ_{B} is the gravitational acceleration at the barycenter of the configuration. The left-hand-side of Eq. (6.4) is a two-point first order quantity, whereas the right-hand-side is a one-point higher order (at least second order) one. Consequently, inserting Eq. (6.4) into Eq. (6.3) results in the sought-after formulation.

Obviously, the expansion (Eq. 6.4) contains partial derivatives of the Earth's gravitational potential higher than the second order. Including all the partial derivatives makes the mathematical model rather complicated. In (Sharifi and Keller, 2005) and (Keller and Sharifi, 2005) it has been shown that even the linear approximation of the expansion can be modified in a way that yields a reasonable accuracy.

Besides the modified linear formulation the cubic approximation of the problem is also presented. In order to simplify the formulation, the observable is expressed as a functional on an incremental potential. More simplifications will be carried out by neglecting subordinate constituents.

6.2.1 Linear approximation

The linear approximation can be derived from Eq. (6.4) by setting $\max(j) = 1$:

$$\Delta\Gamma \;=\; (\nabla \otimes \Gamma^{T}) \cdot \Delta\mathbf{r} = \mathbf{G}\Delta\mathbf{r} \;+\; \Delta_{\mathrm{lin}}, \tag{6.5}$$

with Δ_{lin} the linearization error and $\mathbf{G}$ the Earth's gravitational gradient tensor. Inserting Eq. (6.5) into Eq. (6.3) and dividing both sides of the equation by ϱ results in Eq. (6.6), which is called *linear gradiometry equation*:

$$\mathbf{e}_B^T \mathbf{G} \mathbf{e}_{\mathrm{B}} \;+\; \Delta_{\mathrm{lin}} \;=\; \frac{\ddot{\varrho}}{\varrho} + \frac{\dot{\varrho}^2}{\varrho^2} - \frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho^2}. \tag{6.6}$$

Replacing Eq. (6.6) into Eq. (2.78) yields the observation equation of linear gradiometry

$$V^{(xx)} \;+\; \Delta_{\mathrm{lin}} \;=\; \frac{\ddot{\varrho}}{\varrho} + \frac{\dot{\varrho}^2}{\varrho^2} - \frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho^2}. \tag{6.7}$$

Substituting Eq. (2.79) into Eq. (6.7) gives

$$\left(\frac{\delta r}{\rho}\right)^2 G_{11} + 2\frac{\delta r}{\rho}\cos\alpha\, G_{12} \pm 2\frac{\delta r}{\rho}\sin\alpha\, G_{13} +$$

$$+\cos^2\alpha\, G_{22} \pm \sin 2\alpha\, G_{23} + \sin^2\alpha\, G_{33} \;+\; \Delta_{\mathrm{lin}} \;=\; \frac{\ddot{\varrho}}{\varrho} + \frac{\dot{\varrho}^2}{\varrho^2} - \frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho^2}. \tag{6.8}$$

Due to relatively long baseline of the GRACE virtual gradiometer ($\approx 220\,\mathrm{km}$) the linearization error for the full gravity field is significant (Keller and Sharifi, 2005). However, the error can be noticeably decreased if an incremental potential is used.

In order to show the validity of the idea, we consider EGM96 and EIGEN2 as two representative examples of the Earth's gravity field. In order to simulate an aliasing-free observation vector, both satellites' orbits were integrated based on the EGM96 and EIGEN2 up to degree and order 120 as the real and reference orbit respectively. Consequently, the residual observation δl is calculated as

$$\delta l = l^{120} \;-\; l_{\mathrm{ref}}^{120} \;= \frac{1}{\varrho^2}\left(\ddot{\varrho}\varrho + \dot{\varrho}^2 - \|\Delta\dot{\mathbf{x}}\|\right) - \frac{1}{\varrho_{\mathrm{ref}}^2}\left(\ddot{\varrho}\varrho + \dot{\varrho}^2 - \|\Delta\dot{\mathbf{x}}\|\right)_{\mathrm{ref}}. \tag{6.9}$$

On the other hand, the left-hand-side of Eq. (6.8) excluding Δ_{lin}, contains full terms of the linear

approximation. Let us denote the corresponding quantity for the incremental potential by $\delta V_{\mathrm{f}}^{(xx)}$

$$\delta V_{\mathrm{f}}^{(xx)} = \left(\frac{\delta r}{\varrho}\right)^2 T_{11} + 2\frac{\delta r}{\varrho}\cos\alpha\, T_{12} \pm 2\frac{\delta r}{\varrho}\sin\alpha\, T_{13} + \cos^2\alpha\, T_{22} \pm \sin 2\alpha\, T_{23} + \sin^2\alpha\, T_{33} \quad (6.10)$$

It has been computed directly using the spherical harmonics coefficient differences δc_{nm} and δs_{nm},

$$\begin{pmatrix} \delta c_{nm} \\ \delta s_{nm} \end{pmatrix} = \begin{pmatrix} c_{nm} \\ s_{nm} \end{pmatrix}_{\mathrm{EGM96}} - \begin{pmatrix} c_{nm} \\ s_{nm} \end{pmatrix}_{\mathrm{EIGEN2}}. \quad (6.11)$$

Eventually, $\Delta_{\mathrm{lin}} = \delta l - \delta V_f^{(xx)}$ were computed for a one-month span of the orbit and the achieved results depicted in Fig. (6.2).

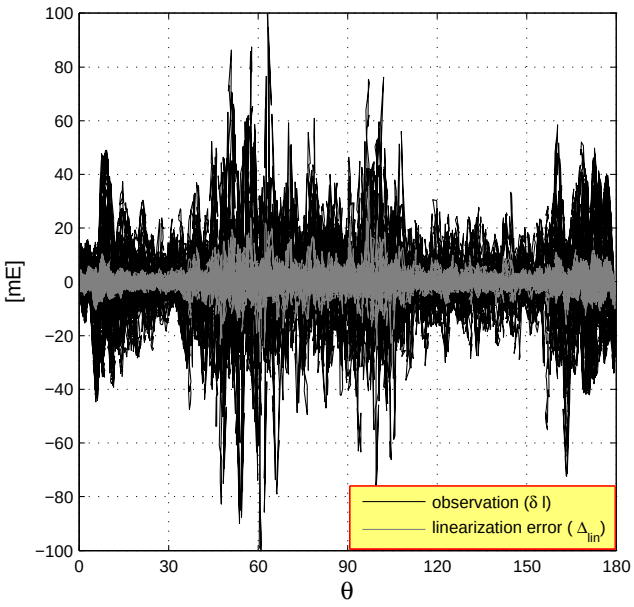


Figure 6.2: One-month span of observations and the respective linearization error with respect to colatitude (θ)

As seen, despite pronounced fluctuations of the error around its mean value (zero), standard deviation seems acceptable (≈ 2.5 mE). Nevertheless, we should try either to amplify the signal or dampen the error to strengthen the S/N ratio.

In Eq. (6.10), some terms are dominant and the others can be neglected without significant loss of accuracy. For instance, one can derive a simplified expression by ignoring all the off-diagonal entries as well as the first diagonal element

$$\delta V^{(xx)} = \cos^2\alpha T_{22} + \sin^2\alpha T_{33}. \quad (6.12)$$

In order to clarify the neglected terms' contribution, differences between $\delta V_{\mathrm{f}}^{(xx)}$ and $\delta V^{(xx)}$ were depicted in Fig. (6.3). As we see in Fig. (6.3), the differences are highly pronounced in the polar regions. Contribution of the off-diagonal elements are significant at the polar region. Consequently, neglecting those terms degrades the formulations in the high-latitude regions.

To analyse the influence of those neglected terms individually, we start with the first diagonal component as the most dominant entry in the Marussi tensor. $T_{11} = \delta V^{(rr)}$ and its respective contribution to $\delta V_{\mathrm{f}}^{(xx)}$ were depicted in Fig. (6.4). Despite the significant values of $\delta V^{(rr)}$, its contribution to the along-track component is negligible due to the special geometrical configuration of the GRACE mission. Moreover, $\delta V^{(rr)}$ is about 300 times greater than the off-diagonal elements, see Rummel (2003). Consequently, the other off-diagonal elements can be ignored by replacing $\frac{\delta r}{\rho} = 0$. Hence,

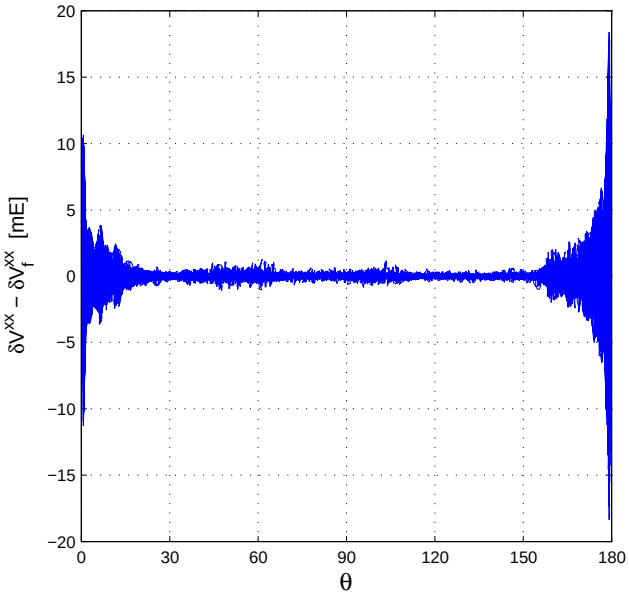


Figure 6.3: Mathematical model simplification error

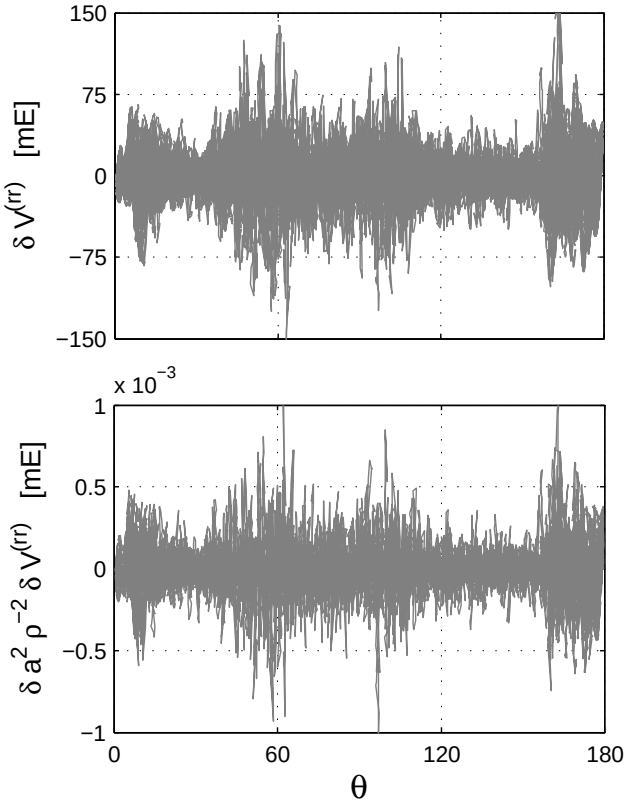


Figure 6.4: $\delta V^{(rr)}$ and its contribution to $\delta V_f^{(xx)}$

Eq. (6.10) can be recast into

$$\delta V_{\pm}^{(xx)} = \cos^2\alpha\, T_{22} \pm \sin 2\alpha\, T_{23} + \sin^2\alpha\, T_{33} \tag{6.13}$$

Subscript $\pm$ have been used due to the presence of T_{23} with two different signs, where the $+$ sign is valid for ascending tracks and the $-$ sign for descending ones. To justify whether the substitution is allowed, $\delta V_{\pm}^{(xx)} - \delta V_f^{(xx)}$ was depicted in Fig. (6.5). Comparing Figs. (6.3) and (6.5) shows that the

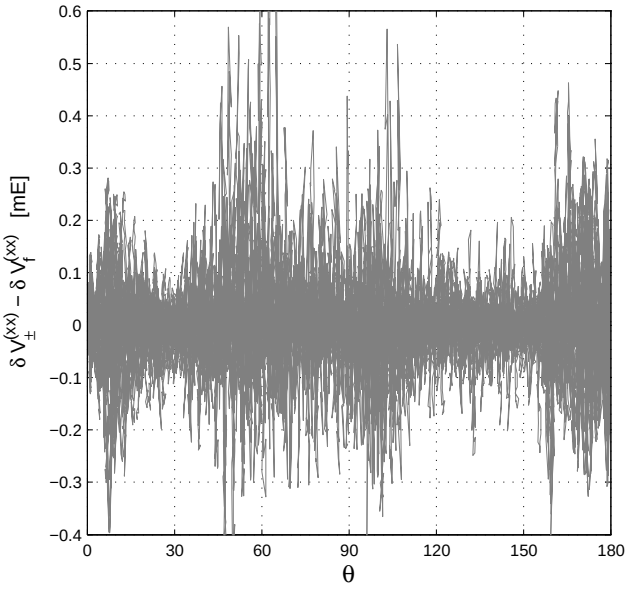


Figure 6.5: Deviation of $\delta V_{\pm}^{(xx)}$ from $\delta V_f^{(xx)}$

significant contribution of the mathematical simplification error (Fig. 6.3) is due neglecting T_{23}. This figure brings us to the conclusion that it is sufficient to include just T_{23} with its respective coefficients if we would like to utilize a fairly complete mathematical model. In other words, the other terms in $\delta V_{\mathrm{f}}^{(xx)}$ (Eq. 6.10) can be ignored.

By adding T_{23} to the mathematical model, we have to distinguish the ascending and descending arcs. For instance, they can be represented on two spheres; one for the data on ascending tracks (*ascending observations*) and one for the data on descending tracks (*descending observations*). In this approach, which is called *Rosborough approach*, we can use either ascending or descending observations to perform GSHA (e.g. Rosborough, 1986; Sneeuw, 2003).

Alternatively, both ascending and descending observations are usually combined into spatially *mean* and *variable* contributions: $\delta V_{\mathrm{m}}^{(xx)}$ and $\delta V_{\mathrm{v}}^{(xx)}$.

$$\delta V_{\mathrm{m}}^{(xx)} = \frac{1}{2}(\delta V_{+}^{(xx)} + \delta V_{-}^{(xx)}) = \cos^2\alpha\, T_{22} + \sin^2\alpha\, T_{33} \tag{6.14}$$

$$\delta V_{\mathrm{v}}^{(xx)} = \frac{1}{2}(\delta V_{+}^{(xx)} - \delta V_{-}^{(xx)}) = \sin 2\alpha\, T_{23} \tag{6.15}$$

We use $\delta V_{\mathrm{m}}^{(xx)}$ for the GSHA since it contains the dominant part of the signal, while $\delta V_{\mathrm{v}}^{(xx)}$ is zero except at the polar regions, see Figs. (6.6) and (6.7).

By definition, to derive both the mean and variable observations the respective components δV_{+}^{xx} and δV_{-}^{xx}, should be measured at the same location. Otherwise, they should be mapped to the common points before inserting into Eqs. (6.14) and (6.15). In reality, one can rarely find the ascending and descending observations measured in the same location. Therefore, implementation of an appropriate prediction scheme is inevitable for realization of either the mean or variable observations.

The same approach is valid for mapping the ascending and descending observation on an equiangular grid on the *mean orbital sphere* (MOS). In other words, each of those data sets should be individually

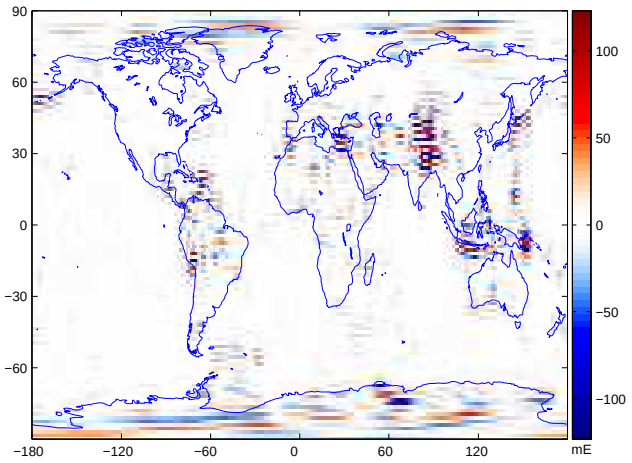


Figure 6.6: $\delta V_{\mathrm{m}}^{(xx)}$ of the linear gradiometry equation

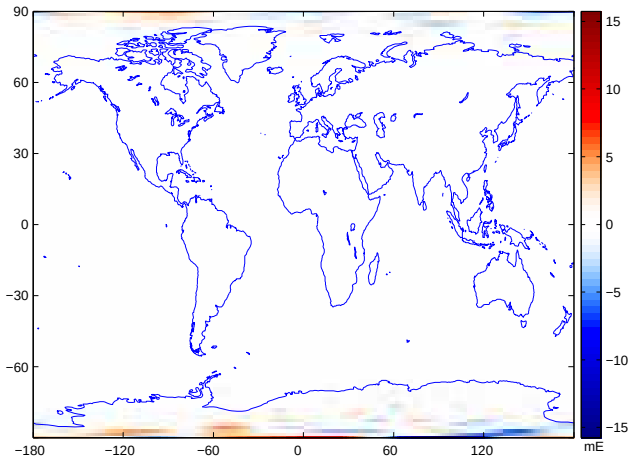


Figure 6.7: $\delta V_{\mathrm{v}}^{(xx)}$ of the linear gradiometry equation

projected on the grid and the averaging is carried out based on the predicted ascending and descending values at the grid points.

From a statistical point of view, δl can be predicted on the grid points with reasonable accuracy if an adequate number of the ascending and descending observations are well-distributed over the area surrounding the respective grid points. Moreover, if the number of the ascending observations are the same as the descending ones, we can expect the predicted values based on a one-step prediction (using both the ascending and descending observations simultaneously) are close enough to the mean signal.

To get an idea about the observation distribution, the previously utilized data set is considered. The number of ascending and descending observations within $1.5° \times 1.5°$ blocks on the mean orbital sphere counted. The achieved results as well as their sum and differences were depicted in Fig. (6.8). Indeed, the respective basic statistics were tabulated in Table (6.1). As seen both in Fig. (6.8) and Table

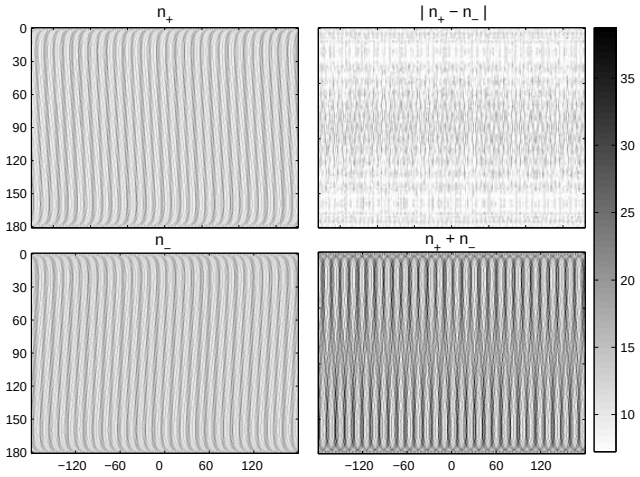


Figure 6.8: Ascending and descending observation distribution of one month of the simulated GRACE data

Table 6.1: Ascending and descending observations statistics

quantity	statistic			
	min	max	mean	std.
n_+	2	20	9	3.8
n_-	1	20	9	3.8
$\lvert n_+ - n_- \rvert$	0	16	3.7	3.1
$n_+ + n_-$	7	39	18	5.9

(6.1) ascending and descending observations were distributed uniformly. Indeed, one can expect that the differences reduce using a longer arc of the orbit.

At this step, let us assume the resulting errors due to the differences in the number of ascending and descending observations are hopefully negligible.

To sum up, the linearization error is the dominant systematic error. Compared to the linearization error, neglecting the radial component disturbs the observations negligibly. However, the off-diagonal element $\mathbf{G}_{23}$ should be considered. Therefore, we utilize the cubic approximation of the gradiometry equation for the diagonal elements as well as the $(2, 3)$ entry.

6.2.2 Cubic approximation

In order to investigate the higher terms' contribution, a more complete form of the gradiometry equation is presented. To obtain a better solution the *Cubic* term of Taylor expansion (Eq. 6.4) is included. Consequently, the observable δl is approximated with Marussi-tensor elements and their

partial derivatives up to the second order

$$V_{\mathrm{fc}}^{(xx)} \approx \mathbf{e}_{\mathrm{B}}^{T} \cdot \mathbf{G} \cdot \mathbf{e}_{\mathrm{B}} + \frac{\varrho^2}{24}\, \mathbf{e}_{\mathrm{B}}^{T} \otimes \mathbf{e}_{\mathrm{B}}^{T} \otimes \mathbf{e}_{\mathrm{B}}^{T} \cdot (\nabla \otimes \nabla \otimes \mathbf{G}) \cdot \mathbf{e}_{\mathrm{B}} \tag{6.16}$$

where the subscript fc stands for the *full cubic* representation of the gradiometry equation. Ignoring the respective radial terms and substituting $\mathbf{e}_{\mathrm{B}}$ from Eq. (2.79) into Eq. (6.16) yields

$$\begin{aligned} \delta V_{\pm c}^{(xx)} &= \cos^2\alpha\, T_{22} \pm \sin 2\alpha\, T_{23} + \sin^2\alpha\, T_{33} \\ &+ \tfrac{\varrho^2}{24\bar{r}^2} \Big(\cos^4\alpha\, T_{22}^{(\theta\theta)} + \tfrac{\sin^2 2\alpha}{4}\, T_{33}^{(\theta\theta)} + \tfrac{\sin^2 2\alpha}{4\sin^2\theta}\, T_{22}^{(\lambda\lambda)} + \tfrac{\sin^4\alpha}{\sin^2\theta}\, T_{33}^{(\lambda\lambda)} \\ &\pm \tfrac{\sin 2\alpha \cos^2\alpha}{\sin\theta}\, T_{22}^{(\theta\lambda)} \pm \tfrac{\sin 2\alpha \sin^2\alpha}{\sin\theta}\, T_{33}^{(\theta\lambda)} \mp \tfrac{\sin 2\alpha \cos^2\alpha}{2\sin\theta\tan\theta}\, T_{22}^{(\lambda)} \\ &\mp \tfrac{\sin 2\alpha \sin^2\alpha}{2\sin\theta\tan\theta}\, T_{33}^{(\lambda)} + \tfrac{\sin^2 2\alpha}{\sin\theta}\, T_{23}^{(\theta\lambda)} - \tfrac{\sin^2 2\alpha}{2\sin\theta\tan\theta}\, T_{23}^{(\lambda)} \\ &\pm \tfrac{\sin 2\alpha \sin^2\alpha}{\sin^2\theta}\, T_{23}^{(\lambda\lambda)} \pm \sin 2\alpha \cos^2\alpha\, T_{23}^{(\theta\theta)} \Big). \end{aligned} \tag{6.17}$$

Also it should be noted that $T_{22}^{(\lambda)}$ and $T_{33}^{(\lambda)}$ have an opposite sign compared to the others. In other words,− and + correspond to ascending and descending tracks respectively, while for the other terms of the expansion the reverse is true.

Decomposition of the observable $\delta V_{\pm c}^{(xx)}$ into the mean and variable components pronouncedly simplifies the equation.

$$\begin{aligned} \delta V_{mc}^{(xx)} &= \cos^2\alpha\, T_{22} + \sin^2\alpha\, T_{33} \\ &+ \tfrac{\varrho^2}{24\bar{r}^2} \Big(\cos^4\alpha\, T_{22}^{\theta\theta} + \tfrac{\sin^2 2\alpha}{4}\, T_{33}^{\theta\theta} + \tfrac{\sin^2 2\alpha}{4\sin^2\theta}\, T_{22}^{\lambda\lambda} + \tfrac{\sin^4\alpha}{\sin^2\theta}\, T_{33}^{\lambda\lambda} \\ &+ \tfrac{\sin^2 2\alpha}{\sin\theta}\, T_{23}^{\theta\lambda} - \tfrac{\sin^2 2\alpha}{2\sin\theta\tan\theta}\, T_{23}^{\lambda} \Big). \end{aligned} \tag{6.18}$$

$$\begin{aligned} \delta V_{vc}^{(xx)} &= \sin 2\alpha\, T_{23} \\ &+ \tfrac{\varrho^2}{24\bar{r}^2} \Big(\tfrac{\sin 2\alpha \cos^2\alpha}{\sin\theta}\, T_{22}^{\theta\lambda} + \tfrac{\sin 2\alpha \sin^2\alpha}{\sin\theta}\, T_{33}^{\theta\lambda} - \tfrac{\sin 2\alpha \cos^2\alpha}{2\sin\theta\tan\theta}\, T_{22}^{\lambda} \\ &- \tfrac{\sin 2\alpha \sin^2\alpha}{2\sin\theta\tan\theta}\, T_{33}^{\lambda} + \tfrac{\sin 2\alpha \sin^2\alpha}{\sin^2\theta}\, T_{23}^{\lambda\lambda} + \sin 2\alpha \cos^2\alpha\, T_{23}^{\theta\theta} \Big). \end{aligned} \tag{6.19}$$

Similar to the linear approximation, the mean and variable components of the cubic approximation are two alternative representations of the original observable. The mean observable will be used for the GSHA because of its dominant contribution.

In order to evaluate the included cubic term, we compute the cubic term contribution to $\delta V^{(xx)}$. It can be easily computed just by subtracting the right-hand-side of Eq. (6.18) from Eq. (6.14) which is equal to the second term on the right-hand-side of Eq. (6.18). The achieved results were plotted in Fig. (6.9).

As seen in Fig. (6.9), the differences are highly pronounced in the mountainous area like the Himalayas. In other words, some correlation can be easily sensed between the differences and topography which seems reasonable. Without doubt, implementation of the cubic approximation results in a better solution. However, direct manipulation of the cubic term seems very complicated. In the subsequent Chapter, we will propose an alternative scheme for treating the cubic terms which is superior from a computational point of view.

6.3 Inter-satellite velocity determination

As stated previously, combination of the HL– with LL-SST observables is the unique opportunity which is realized by launching the GRACE mission. Different sampling rate and level of accuracy of two configuration observations implies that data fusion should be performed carefully.

The range-rate observation $\dot{\varrho}$ is the key observable of the GRACE mission because of its comparatively high accuracy. On the other hand, the relative velocity of the GRACE satellites should be derived

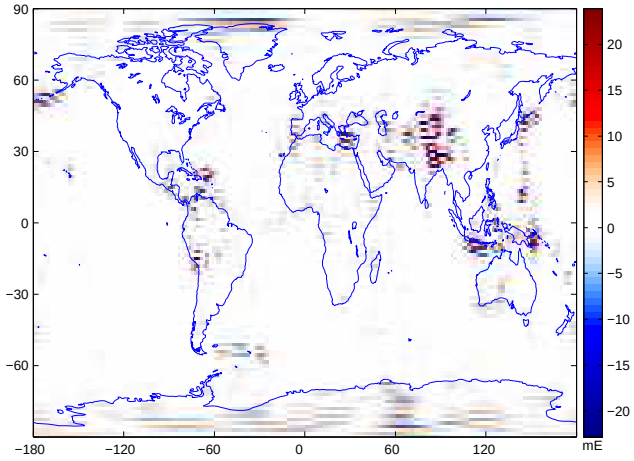


Figure 6.9: Differences between the linear and cubic approximation of the mean signal

by numerical differentiation which by its very nature amplifies the observation noise. Therefore, one may spoil the range data if the relative velocity is not appropriately computed.

Reducing the observables by means of a suitable reference orbit and computing the velocity term using the reference field are two alternatives to numerical differentiation which are addressed in the following subsections. Two different numerical differentiation algorithm are employed and the results are compared to that of the alternatives.

6.3.1 Numerical differentiation

The Newton differentiator

The problem of one-dimensional interpolation and integration usually involves the approximation of a function $f(x)$ by a suitably chosen n^{th} degree polynomial $\widetilde{f}(x)$ such that $f_i = f(x_i) = \widetilde{f}(x_i)$ for $i = 0, 1, \cdots, n$. Although the polynomial is represented in different forms these $n+1$ distinct function values define a unique polynomial.

Similarly one can utilize the polynomial function for computation of the function derivatives. However, there is no need to find the polynomial function $\widetilde{f}(x)$ if the values of the derivatives at the mesh points $x_0, x_1, \cdots, x_n$ is required. In this particular case and for the evenly spaced data with spacing h, the k-th derivative of the function $f(x)$ is

$$f^{(k)}(x_i) \;=\; h^{-k} \sum_{j=0}^{n} a_{ij}^{ks} f(x_j) \;+\; \mathcal{O}(h^s) \tag{6.20}$$

where $f^{(k)}(x_i)$ is the k-th derivative of the function at the mesh point x_i and the truncation error $\mathcal{O}(h^s)$ is assumed to be of degree s. With the current representation, the derivative of the function is expressed as a linear combination of the function values at the mesh points. Consequently, the derivative of the function can be simply computed if the coefficients a_{ij}^{ks} are determined. Herein, we follow Gregory (1957)'s elegant scheme for the determination of the coefficients.

Let assume the function $f(x)$ possesses a continuous $k+s$ derivative for $x \in [x_0, x_n]$ for a specified order of derivative k and truncation s. The Taylor expansion of the function at the point x_j about the mesh point x_i is

$$f(x_j) \;=\; \sum_{t=0}^{k+s-1} (j-i)^t \, \frac{h^t}{t!} \, f^{(t)}(x_i) \;+\; \frac{(j-i)^{k+s} h^{k+s}}{(k+s)!} \, f^{(k+s)}(\xi_{ij}) \tag{6.21}$$

with $\xi_{ij} \in [x_i \,,\, x_j]$. If we expand the function for $j = 0, 1, \cdots, n$ and then form a linear combination

of the expansions we obtain

$$\sum_{j=0}^{n} a_{ij}^{ks} f(x_j) \;=\; \sum_{j=0}^{n} a_{ij}^{ks} \sum_{t=0}^{k+s-1} (j-i)^t \;\frac{h^t}{t!}\, f^{(t)}(x_i) \;+\; \sum_{j=0}^{n} a_{ij}^{ks} \frac{(j-i)^{k+s} h^{k+s}}{(k+s)!}\, f^{(k+s)}(\xi_{ij}). \tag{6.22}$$

and swapping the summations yields,

$$\sum_{j=0}^{n} a_{ij}^{ks} f(x_j) \;=\; \sum_{t=0}^{k+s-1} \frac{h^t}{t!} \sum_{j=0}^{n} a_{ij}^{ks} (j-i)^t\, f^{(t)}(x_i) \;+\; \frac{h^{k+s}}{(k+s)!} \sum_{j=0}^{n} a_{ij}^{ks} (j-i)^{k+s}\, f^{(k+s)}(\xi_{ij}). \tag{6.23}$$

Dividing both sides of Eq. (6.23) by h^k and replacing

$$b_{ij}^{ks} \;=\; \sum_{j=0}^{n} a_{ij}^{ks} (j-i)^t, \qquad\qquad t = k+s \tag{6.24}$$

results in

$$h^{-k} \sum_{j=0}^{n} a_{ij}^{ks} f(x_j) \;=\; \sum_{t=0}^{k+s-1} \frac{h^{t-k}}{t!}\, b_{ij}^{ks}\, f^{(t)}(x_i) \;+\; E_i^{ks} \tag{6.25}$$

where

$$E_i^{ks} \;=\; \frac{h^s}{(k+s)!} \sum_{j=0}^{n} a_{ij}^{ks} (j-i)^{k+s}\, f^{(k+s)}(\xi_{ij}). \tag{6.26}$$

In order to obtain an equivalent formula given in Eq. (6.20), we determine the unknown coefficients a_{ij}^{ks} in a way that

$$b_{ij}^{ks} \;=\; k!\delta_{tk}. \tag{6.27}$$

As a consequence, Eq. (6.25) can be recast into the following simple form

$$f^{(k)}(x_i) \;=\; h^{-k} \sum_{j=0}^{n} a_{ij}^{ks} f(x_j) \;-\; E_i^{ks}. \tag{6.28}$$

where $-E_i^{ks}$ is equivalent to the truncation error $\mathcal{O}(h^s)$. Replacing Eq. (6.27) into Eq. (6.24) yields a linear system of equations

$$\sum_{j=0}^{n} a_{ij}^{ks} (j-i)^t \;=\; k!\delta_{tk} \tag{6.29}$$

where $t = 0, 1, \cdots, k+s-1$. An equivalent representation of the system in matrix notation is

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ (-i) & (1-i) & (2-i) & \cdots & (n-i) \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ (-i)^k & (1-i)^k & (2-i)^k & \cdots & (n-i)^k \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ (-i)^{t_m} & (1-i)^{t_m} & (2-i)^{t_m} & \cdots & (n-i)^{t_m} \end{pmatrix} \begin{pmatrix} a_{i0}^{ks} \\ a_{i1}^{ks} \\ a_{i2}^{ks} \\ a_{i3}^{ks} \\ \vdots \\ a_{in}^{ks} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ k! \\ \vdots \\ 0 \end{pmatrix}, \tag{6.30}$$

where $t_m = k+s-1$. The system consists of $k+s$ equations in $n+1$ unknowns. The system is overdetermined if the number of equations is greater than that of the unknowns,

$$n+1 \geq k+s.$$

Using this relationship one can determine the minimum number of mesh points for a given k and s. In the other hand, for a certain number of mesh points and a specified derivative order k the order of the truncation error is defined.

The system with unique solution ($n+1 = k+s$) is the most commonly used solution. In this

particular case, the uniqueness of solution is guaranteed because of the non-singular Vandermonde normal matrix (Gregory, 1957).
The system can be set up for $i = 0, 1, \cdots, n$ and the respective coefficients can be derived for each individual mesh point. As a representative example, the coefficients of the nine-point differentiator ($n = 9$) for the first- and second-order differentiation were determined and the achieved coefficients were given in $\mathbf{A}^{18}$ and $\mathbf{A}^{27}$ respectively.

$$\mathbf{A}^{18} = \begin{pmatrix} -\frac{106}{39} & 8 & -14 & \frac{56}{3} & -\frac{35}{2} & \frac{56}{5} & -\frac{14}{3} & \frac{8}{7} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{223}{140} & \frac{7}{2} & -\frac{7}{2} & \frac{35}{12} & -\frac{7}{4} & \frac{7}{10} & -\frac{1}{6} & \frac{1}{56} \\ \frac{1}{56} & -\frac{2}{7} & -\frac{19}{20} & 2 & -\frac{5}{4} & \frac{2}{3} & -\frac{1}{4} & \frac{2}{35} & -\frac{1}{168} \\ -\frac{1}{168} & \frac{1}{14} & -\frac{1}{2} & -\frac{9}{20} & \frac{5}{4} & -\frac{1}{2} & \frac{1}{6} & -\frac{1}{28} & \frac{1}{280} \\ \frac{1}{280} & -\frac{4}{105} & \frac{1}{5} & -\frac{4}{5} & 0 & \frac{4}{5} & -\frac{1}{5} & \frac{4}{105} & -\frac{1}{280} \\ -\frac{1}{280} & \frac{1}{28} & -\frac{1}{6} & \frac{1}{2} & -\frac{5}{4} & \frac{9}{20} & \frac{1}{2} & -\frac{1}{14} & \frac{1}{168} \\ \frac{1}{168} & -\frac{2}{35} & \frac{1}{4} & -\frac{2}{3} & \frac{5}{4} & -2 & \frac{19}{20} & \frac{2}{7} & -\frac{1}{56} \\ -\frac{1}{56} & \frac{1}{6} & -\frac{7}{10} & \frac{7}{4} & -\frac{35}{12} & \frac{7}{2} & -\frac{7}{2} & \frac{223}{140} & \frac{1}{8} \\ \frac{1}{8} & -\frac{8}{7} & \frac{14}{3} & -\frac{56}{5} & \frac{35}{2} & -\frac{56}{3} & 14 & -8 & \frac{106}{39} \end{pmatrix}. \tag{6.31}$$

$$\mathbf{A}^{27} = \begin{pmatrix} \frac{208}{71} & \frac{-481}{35} & \frac{621}{20} & \frac{-2003}{45} & \frac{691}{16} & \frac{-141}{5} & \frac{631}{53} & \frac{-103}{35} & \frac{12}{37} \\ \frac{12}{37} & \frac{1}{79} & \frac{-83}{40} & \frac{153}{40} & \frac{-180}{49} & \frac{47}{20} & \frac{-39}{40} & \frac{5}{21} & \frac{-1}{39} \\ \frac{-1}{39} & \frac{39}{70} & \frac{-57}{62} & \frac{1}{10} & \frac{9}{16} & \frac{-37}{90} & \frac{7}{40} & \frac{-3}{70} & \frac{1}{214} \\ \frac{1}{214} & \frac{-4}{59} & \frac{29}{40} & \frac{-59}{45} & \frac{11}{16} & \frac{-1}{40} & \frac{-1}{51} & \frac{1}{140} & \frac{-1}{1120} \\ \frac{-1}{1120} & \frac{1}{79} & \frac{-1}{10} & \frac{4}{5} & \frac{-121}{85} & \frac{4}{5} & \frac{-1}{10} & \frac{1}{79} & \frac{-1}{1120} \\ \frac{-1}{1120} & \frac{1}{140} & \frac{-1}{51} & \frac{-1}{40} & \frac{11}{16} & \frac{-59}{45} & \frac{29}{40} & \frac{-4}{59} & \frac{1}{214} \\ \frac{1}{214} & \frac{-3}{70} & \frac{7}{40} & \frac{-37}{90} & \frac{9}{16} & \frac{1}{10} & \frac{-57}{62} & \frac{39}{70} & \frac{-1}{39} \\ \frac{-1}{39} & \frac{5}{21} & \frac{-39}{40} & \frac{47}{20} & \frac{-180}{49} & \frac{153}{40} & \frac{-83}{40} & \frac{1}{79} & \frac{12}{37} \\ \frac{12}{37} & \frac{-103}{35} & \frac{631}{53} & \frac{-141}{5} & \frac{691}{16} & \frac{-2003}{45} & \frac{621}{20} & \frac{-481}{35} & \frac{208}{71} \end{pmatrix} \tag{6.32}$$

Using the matrices $\mathbf{A}^{18}$ and $\mathbf{A}^{27}$ both the first- and the second-order derivatives of the underlying function can be determined for all mesh points. However, the error increases by moving toward the borders point. In other words, the middle mesh point has the lowest differentiation error.
In practice, for evaluation of the derivatives at each mesh point a symmetric window of size n is considered and the derivatives are only computed for the center points. Equations (6.33) and (6.34) for instance, represent the respective scalar formulas for the 9-point first- and second-order differentiator.

$$h\, f^{(1)}(x_4) = \frac{1}{280} f(x_0) - \frac{4}{105} f(x_1) +$$
$$\frac{1}{5} f(x_2) - \frac{4}{5} f(x_3) + \frac{4}{5} f(x_5) - \frac{1}{5} f(x_6) + \frac{4}{105} f(x_7) - \frac{1}{280} f(x_8) \tag{6.33}$$

$$h^2\, f^{(2)}(x_4) = -\frac{1}{1120} f(x_0) + \frac{1}{79} f(x_1) -$$
$$\frac{1}{10} f(x_2) + \frac{4}{5} f(x_3) - \frac{121}{85} f(x_4) + \frac{4}{5} f(x_5) - \frac{1}{10} f(x_6) + \frac{1}{79} f(x_7) - \frac{1}{1120} f(x_8) \tag{6.34}$$

These formulas can also be derived from the Newton high-resolution interpolation formula by differentiation. So, they can correspondingly be called the *Newton differentiators*.
Numerical analysis for evaluation of the Newton differentiator's performance in comparison with other differentiators will be carried out later in subsection (6.3.3).

Differentiating spline

Let us recall the problem setting used in the Newton differentiation. Discrete values of the unknown function $f(x)$ are given at some finite mesh points $x_0 < x_1 < x_2 < \cdots < x_n$. The is goal is to

compute the underlying function derivatives exactly at the mesh points.
As expressed previously, the function can be approximated by means of a polynomial function of degree n. Instead of approximating the underlying function globally one can use n piecewise polynomials of degree m which satisfy the following conditions:

1. for $x \in [x_i\,,\, x_{i+1}]$, $S(x) = S_i(x)$,
2. $S_i(x_i) = f(x_i)$ and $S_i(x_{i+1}) = f(x_{i+1})$,
3. derivatives up to order $(m-1)$ exist and are continuous at the internal mesh points

$$S_i^{(m-1)}(x_i) \;=\; S_{i+1}^{(m-1)}(x_i),$$

The function $S(x)$ is then called a *spline* of degree m. The *Cubic spline* ($m = 3$) is the most popular form of spline. In this particular case, we have $4n$ unknown coefficients for n subintervals. According to the second property of spline, the total number of observation equations is $2n$ equations. Moreover, the continuity condition of the derivatives up to order 2 ($= 3-1$) at the internal mesh points ($n-1$ points) sets up $2(n-1)$ equations of the type functional constrained. Therefore, we can set up an underdetermined linear system with $2n+2(n-1)$ equations in $4n$ unknowns. There are various cubic splines depending on how to fix the two free conditions. The frequently used choices are:

- the *natural cubic spline* by imposing $S_1^{(2)}(x_0) \;=\; S_n^{(2)}(x_n) = 0$,
- the *clamped cubic spline* defined by $S_1^{(1)}(x_0) \;= f^{(1)}(x_0)$ and $S_n^{(1)}(x_n) \;= f^{(1)}(x_n)$.

Herein, we use the natural cubic spline. The underlying function and its first- and second-order derivatives are known as soon as the unknown coefficients are determined.
To set up the linear system of equations let us begin with the continuity of the second order derivative. Since the function itself is cubic, the second-order derivative is a linear function (Moritz, 1978)

$$S_i^{(2)}(x) = a_i \left(\frac{x_{i+1}-x}{h}\right) \;+\; a_{i+1}\left(\frac{x-x_i}{h}\right), \tag{6.35}$$

for $i = 0, 1, \cdots, n-1$. Then,

$$\left.\begin{aligned} S_i^{(2)}(x_i) \;&=\; a_i \\ S_i^{(2)}(x_{i+1}) \;&=\; a_{i+1} \end{aligned}\right\} \quad \Rightarrow \quad S_{i-1}^{(2)}(x_{i+1}) \;=\; S_i^{(2)}(x_i) \;=\; a_i. \tag{6.36}$$

which implies continuity of the second-order derivative at the junction point of two subsequent subintervals $i-1$ and i. Equation (6.35) is recast into the following equation with twice integration

$$S_i(x) \;=\; a_i\frac{(x_{i+1}-x)^3}{6h} \;+\; a_{i+1}\frac{(x-x_i)^3}{6h} \;+\; b_i(x_{i+1}-x) \;+\; c_i(x-x_i), \tag{6.37}$$

where the additional terms are written in this special form for convenience. Using the function values at the mesh points generates two observation equations

$$\left.\begin{aligned} S_i(x_i) \;&=\; a_i\tfrac{h^2}{6} + b_i h \;=\; f_i \\ S_i(x_{i+1}) \;&=\; a_{i+1}\tfrac{h^2}{6} + c_i h \;=\; f_{i+1} \end{aligned}\right\} \quad \Rightarrow \quad \begin{aligned} hb_i \;&=\; f_i - a_i\tfrac{h^2}{6} \\ hc_i \;&=\; f_{i+1} - a_{i+1}\tfrac{h^2}{6} \end{aligned} \tag{6.38}$$

consequently,

$$\begin{aligned} S_i(x) \;=\; a_i\frac{(x_{i+1}-x)^3}{6h} \;+\; a_{i+1}\frac{(x-x_i)^3}{6h} \;+ \\ \left(f_i - a_i\frac{h^2}{6}\right)\frac{(x_{i+1}-x)}{h} \;+\; \left(f_{i+1} - a_{i+1}\frac{h^2}{6}\right)\frac{(x-x_i)}{h}. \end{aligned} \tag{6.39}$$

Immediately, its first-order derivatives at the respective mesh points are determined as

$$S_i^{(1)}(x_i) \;=\; -\frac{h}{3}a_i \;-\; \frac{h}{6}a_{i+1} \;+\; \frac{f_{i+1}-f_i}{h} \qquad S_i^{(1)}(x_{i+1}) \;=\; \frac{h}{3}a_{i+1} \;+\; \frac{h}{6}a_i \;+\; \frac{f_{i+1}-f_i}{h} \tag{6.40}$$

Imposing the continuity condition of the first-order derivatives of two subsequent subintervals $i-1$ and i leads to

$$a_{i-1} + 4a_i + a_{i+1} = \frac{6}{h^2}\left(f_{i-1} - 2f_i + f_{i+1}\right). \tag{6.41}$$

For $i = 1, 2, \cdots, n-1$ we can set up a linear system with $n-1$ equations in $n+1$ unknowns. Using Eq. (6.36), two free unknowns can be fixed by imposing the natural spline condition

$$\left.\begin{aligned} S_0^{(2)}(x_0) &= a_0 = 0 \\ S_{n-1}^{(2)}(x_n) &= a_n = 0 \end{aligned}\right\} \quad \Rightarrow \quad a_0 = a_n = 0. \tag{6.42}$$

Therefore, $n-1$ unknowns coefficients $(a_1, a_2, \cdots, a_{n-1})$ remain which can be determined using the following uniquely solution linear system of equations.

$$\begin{pmatrix} 4 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 4 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 4 & 1 \\ 0 & \cdots & & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} f_0 - 2f_1 + f_2 \\ f_1 - 2f_2 + f_3 \\ \\ \vdots \\ f_{n-2} - 2f_{n-1} + f_n \end{pmatrix} \tag{6.43}$$

or simply,

$$\mathbf{A}\,\mathbf{a} = \mathbf{b}. \tag{6.44}$$

The design matrix $\mathbf{A}$ is a tridiagonal, symmetric, positive definite and strictly diagonally dominant matrix. Therefore, the existence of a unique solution is guaranteed (Barannyk, 2005).
When the coefficients have been determined, the analytical form of first- and second-order derivatives of the underlying function can be approximated by a quadratic and linear function respectively

$$S_i^{(1)}(x) = \frac{a_{i+1}}{2h}\left(x - x_i\right)^2 - \frac{a_i}{2h}\left(x_{i+1} - x\right)^2 + \frac{1}{h}\left(f_{i+1} - f_i\right) - \frac{h}{6}\left(a_{i+1} - a_i\right) \tag{6.45}$$

and

$$S_i^{(2)}(x) = \frac{a_{i+1}}{h}\left(x - x_i\right) + \frac{a_i}{h}\left(x_{i+1} - x\right). \tag{6.46}$$

These two equations, the so called *differentiator spline*, are the analytical form of the derivatives. However, the derivatives are required at the mesh points which can be represented in much simpler forms

$$S_i^{(1)}(x_i) = -\frac{h}{6}\left(2a_i + a_{i+1}\right) + \frac{1}{h}(f_{i+1} - f_i), \tag{6.47}$$

and the coefficient a_i is the second derivative at the mesh point i. However, approximation of the second order derivatives with a linear function may provide inaccurate estimation. Instead, one can use double first-order differentiation to improve the second derivatives.
The spline differentiators and the Newton differentiation method are numerically compared in the subsection (6.3.3) .

6.3.2 Adjusting GPS observation with the range measurements

Besides an independent differentiation scheme, one can utilize the range observations within the differentiation process to obtain higher accuracy. At each evaluation point, the intersatellite range-rate ($\dot{\varrho}$) observed by the range defines the following condition equation

$$\hat{\varrho}\dot{\varrho} - \widehat{\Delta\mathbf{r}}^T\widehat{\Delta\dot{\mathbf{r}}} = 0. \tag{6.48}$$

Where $\widehat{\Delta\mathbf{r}}$ and $\widehat{\Delta\dot{\mathbf{r}}}$ are the adjusted relative position and velocity vectors. Due to the ranger highly-accurate range-rate measurement $\dot{\varrho}$ can be assumed error-free. The condition equation is a nonlinear equation which takes the following form by writing the intersatellite range in terms of coordinate differences

$$\sqrt{\widehat{\Delta\mathbf{r}}^T\widehat{\Delta\mathbf{r}}}\,\dot{\varrho} - \widehat{\Delta\mathbf{r}}^T\widehat{\Delta\dot{\mathbf{r}}} = 0. \tag{6.49}$$

For linearization of the equation, an initial value for $\Delta\dot{\mathbf{r}}$ is required which can be derived by numerical differentiation. The correction of the position vector in the sense of least squares is

$$\widehat{\mathbf{V}}_p = \frac{\varrho\dot{\varrho} - \Delta\mathbf{r}^T\Delta\dot{\mathbf{r}}}{\|\Delta\dot{\mathbf{r}}\|^2 + \beta\varrho^2 - \dot{\varrho}^2}\left(\Delta\dot{\mathbf{r}} - \frac{\dot{\varrho}}{\varrho}\Delta\mathbf{r}\right) \tag{6.50}$$

and equivalently the velocity vector correction,

$$\widehat{\mathbf{V}}_{\mathrm{v}} = \frac{\varrho\dot{\varrho} - \Delta\mathbf{r}^T\Delta\dot{\mathbf{r}}}{\|\Delta\dot{\mathbf{r}}\|^2 + \beta\varrho^2 - \dot{\varrho}^2}\,\beta\Delta\mathbf{r} \tag{6.51}$$

where β is the ratio of the relative velocity accuracy to that of the position.
A similar condition equation can be developed for the range-acceleration. As we will see, imposing the range-rate condition equation totally corrects the observations along the orbit and no more constraint is required. On the other hand, imposing a constraint on the range-acceleration does not improve the perpendicular components. Therefore, we just set up the condition equation for the range-rate and ignore any additional condition equations.

6.3.3 Numerical comparison of the numerical differentiators

In this subsection, the previously developed differentiators are implemented on some real GRACE observations. Both GPS and the ranger data are taken from the mission observations within the days $230 - 260$ in the year 2003. Some pre-analysis has been carried out on the observation for removing the blunders and adding the corrections.
In order to find an optimal window for the Newton differentiator, the process using windows with different lengths were carried out and the achieved results depicted in Fig. (6.10).

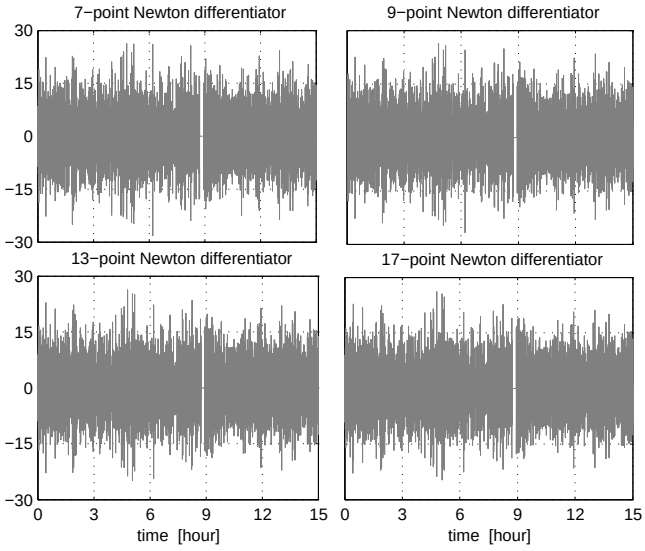


Figure 6.10: Comparison of the Newton differentiators

No superiority was observed between the differentiators. Furthermore, comparing the basic statistics of the output of the differentiations (Table 6.2) showed a slight improvement by increasing window size. As a middle solution, the 9-point Newton differentiator will be employed for the later computations. In order to cross-validate different numerical differentiators, we use the range-rate and range-acceleration observations. Different numerical differentiations are applied on the range-rate observations and the achieved results compared with the range-acceleration ($\ddot{\varrho}$).
As seen in Fig. (6.11), both methods gave nearly the same results. However, the spline method outperforms the Newton interpolation method in the presence of irregularities. Data gaps for instance cause serious problem for the Newton differentiator while these can be easily handled in the spline differentiation.

Table 6.2: The basic statistics of the Newton differentiators comparison

method	min [μGal]	max [μGal]	mean [μGal]	std. [μGal]
7-point	-28	26	0.00	6.82
9-point	-27	26	0.00	6.68
13-point	-25	26	0.00	6.52
17-point	-25	27	0.00	6.44

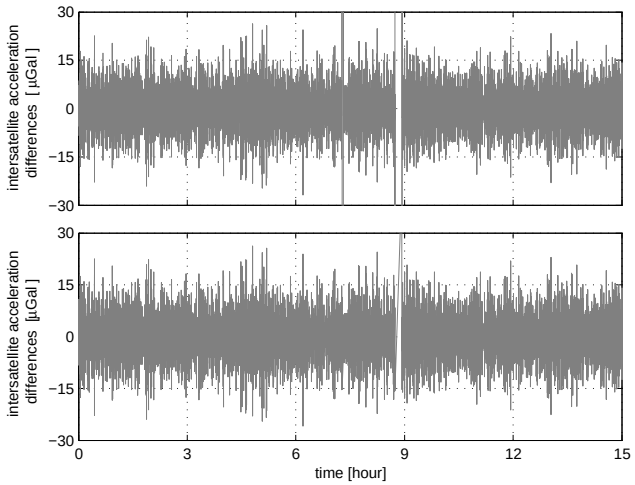


Figure 6.11: Differences between the numerically derived and given intersatellite range-acceleration (top: Newton differentiator; bottom: Spline differentiation)

The expected accuracy of the differentiation methods are about 6 μGal which seems to be a reasonable accuracy compared to other source of errors. Of course, the accuracy of differentiation is better than 6 μGal since the given range-acceleration is assumed error-free.
For evaluation of the GPS observation accuracy, one can compare the quantities from range observations which can be derived numerically from GPS data. The intersatellite range-rate for instance, is one of those observables which is observed by the ranger and can be numerically derived from GPS measurements.
As depicted in Fig. (6.12), it is obvious that the intersatellite range-rate derived from the high-low configuration is comparatively inaccurate compared to the directly observed one. The range-rate is observed by the ranger system with an accuracy of about 1 μm/s whereas that of for the numerically derived one is 200 μm/s.

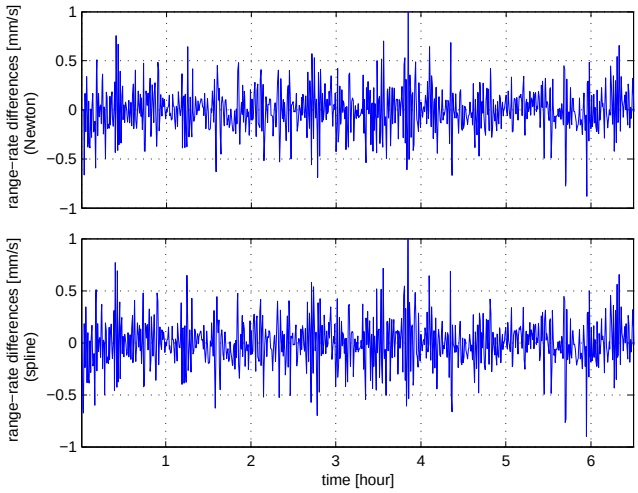


Figure 6.12: Differences of the numerically derived and directly observed intersatellite range-rate (top: Newton differentiator; bottom: Spline differentiation)

However, $\dot{\varrho}^2/\varrho$ and $\dot{\varrho}^2/\varrho^2$ are the quantities which are used as the components of the observation vectors in the acceleration difference and gradiometry approach. As an example, the difference of the numerically derived and directly observed quantity for the gradiometry approach was drawn in Fig. (6.13). The differences, in contrast to the range-rate deviations, are significantly lower because of the long arm-length of the GRACE gradiometer.
Analogously, one can compare the numerically derived and directly observed $\dot{\varrho}^2/\varrho$. As expressed in Eq. (6.3), it is directly involved in the acceleration difference observation. Dividing by the intersatellite range confines the differences of the μGal level, see Fig. (6.14)
There are two possibilities for deriving the second-order derivatives:

- applying second-order differentiator to derive $\ddot{\varrho}(t)$ from $\varrho(t)$, or
- performing double first-order differentiation $\varrho(t) \rightarrow \dot{\varrho}(t) \rightarrow \ddot{\varrho}(t)$

From a theoretical point of view they are identical if the function is known analytically. In practice, however the function is only sampled at some discrete points. Moreover, the observations are contaminated with random errors. Therefore, they may behave differently in practice.
To figure out any possible superiority of the methods, the intersatellite range-accelerations were computed from GPS observations by means of the aforementioned differentiation schemes. Both the Newton differentiator as well as spline differentiation were employed and residuals of the achieved results were depicted in Figs. (6.15) and (6.16). Additionally, the basic statistics of the differences were tabulated in Table (6.3).
As the above figures showed, the double first-order differentiator gave higher accuracy compared to the second-order differentiation methods. The basic statistics also implied that double first-order differentiation obviously outperformed the second-order differentiation. It is due to the fact that in

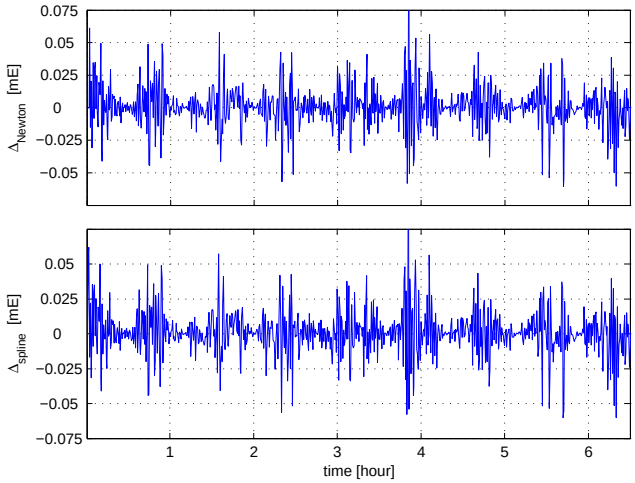


Figure 6.13: Differences of the numerically derived and directly observed $\dot{\varrho}^2/\varrho^2$ (top: $\Delta_{Newton} = \frac{\dot{\varrho^c}^2_{Newton}}{\varrho^2} - \frac{\dot{\varrho}^2}{\varrho^2}$ bottom: $\Delta_{Spline} = \frac{\dot{\varrho^c}^2_{Spline}}{\varrho^2} - \frac{\dot{\varrho}^2}{\varrho^2}$)

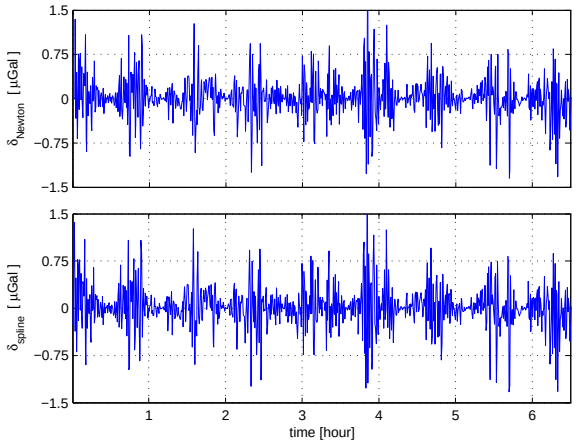


Figure 6.14: Differences of the numerically derived and directly observed $\dot{\varrho}^2/\varrho$ (top: $\delta_{Newton} = \frac{\dot{\varrho^c}^2_{Newton}}{\varrho} - \frac{\dot{\varrho}^2}{\varrho}$ bottom: $\delta_{Spline} = \frac{\dot{\varrho^c}^2_{Spline}}{\varrho} - \frac{\dot{\varrho}^2}{\varrho}$)

Table 6.3: Comparison of the numerical differentiation schemes for the acceleration computation

method	min $[mGal]$	max $[mGal]$	mean $[mGal]$	std. $[mGal]$
double first-order Newton	−3.92	8.10	0.00	1.10
double first-order spline	−4.30	5.53	0.00	1.12
second-order Newton	−10.63	12.17	0.00	2.16
second-order spline	−18.45	19.20	−0.01	3.57

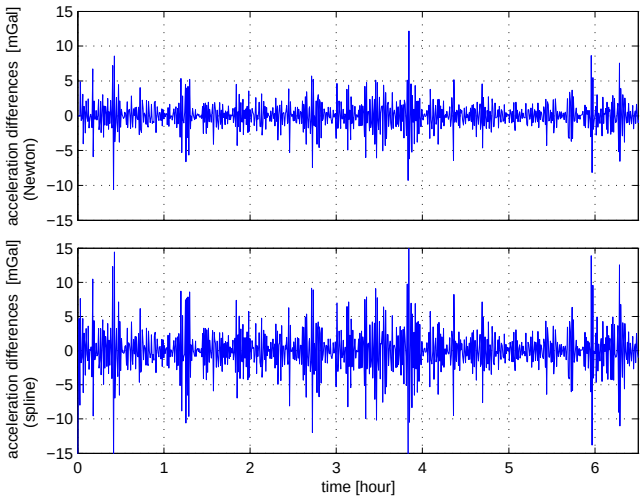


Figure 6.15: Differences between the numerically derived and directly observed acceleration (top: second-order Newton differentiator; bottom: second-oder spline differentiation)

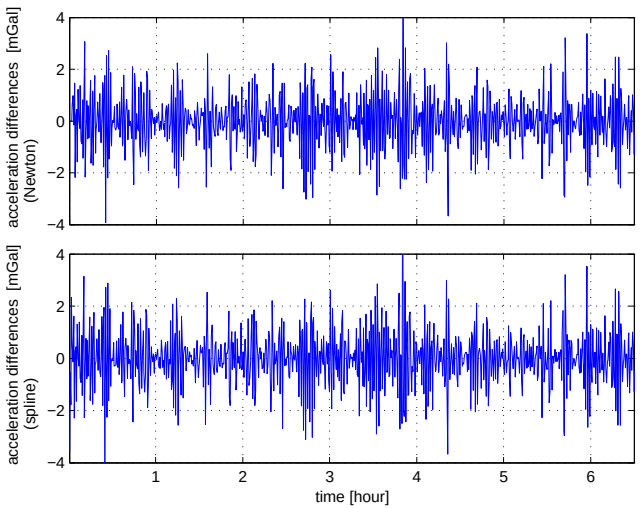


Figure 6.16: Differences between the numerically derived and directly observed acceleration (top: double first-order Newton differentiator; bottom: double first-oder spline differentiation)

the double first-order differentiation the observation noise level is decreased twice because of a double averaging process.
Furthermore, observations with standard deviation of about 80 μGal seem satisfactorily accurate. As Fig. (6.17) shows, both the range-rate and range-acceleration are reasonably smooth functions, so one can expect such promising results by numerical differentiation.

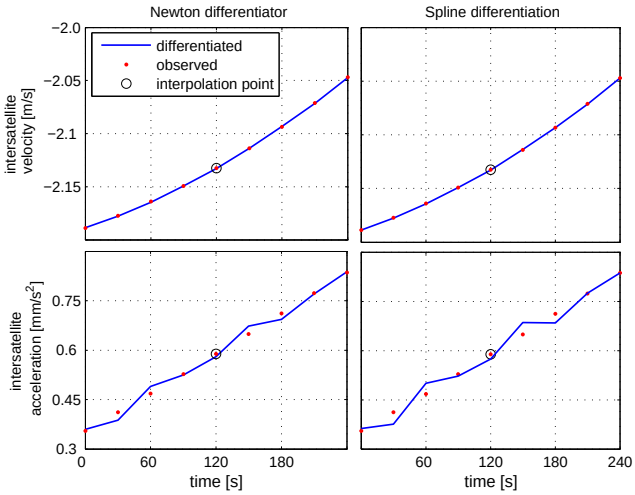


Figure 6.17: The numerically derived versus directly observed intersatellite range-rate and range-acceleration

More interesting results are obtained for the observation of gradiometric type. Despite the acceleration difference approach, $\ddot{\varrho}/\varrho$ is used in the gradiometry observation equation (Eq. 6.6). Thanks to the GRACE for its relatively long intersatellite range which reduces the noise level dramatically by dividing the range-acceleration by the intersatellite range. As Fig. (6.18) shows, obtaining gradiometry observations with standard deviation of about 50 mE is also possible *merely* with GPS observations.

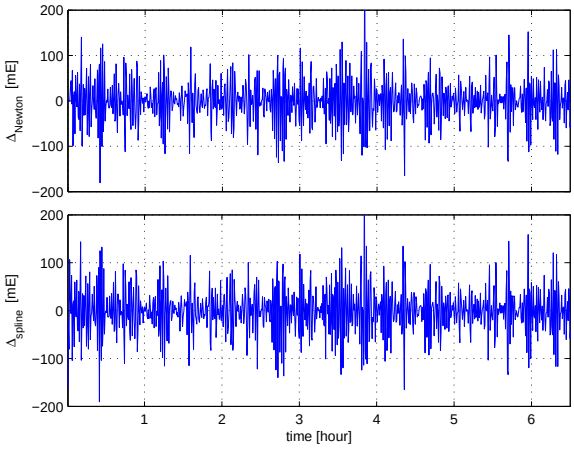


Figure 6.18: Differences of the numerically derived and directly observed $\ddot{\varrho}/\varrho$ (top: $\Delta_{Newton} = \frac{\ddot{\varrho}^c_{Newton}}{\varrho} - \frac{\ddot{\varrho}}{\varrho}$ bottom: $\Delta_{Spline} = \frac{\ddot{\varrho}^c{}_{Spline}}{\varrho} - \frac{\dot{\varrho}}{\varrho}$)

In order to evaluate the effect of imposing the constraint given in Eq. (6.48) on the GPS observations, the numerically derived range-rate was adjusted by the ranger measurements. The achieved results were depicted in Fig. (6.19). As shown in Fig. (6.19), implementation of the condition equation confined the differences to zero.
$\|\Delta\dot{\mathbf{r}}\|^2/\varrho^2$ is the required quantity which should be numerically derived from GPS observations. To see

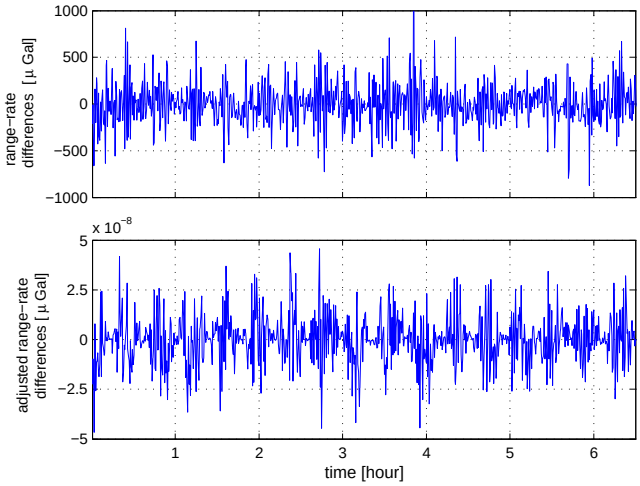


Figure 6.19: Differences of the numerically derived intersatellite range-rate with and without observation adjustment

the effect of the constraint, the quantity was computed with and without imposing the constraint on the observations and the results were plotted in Fig. (6.20). Improvement of $\|\Delta\dot{\mathbf{r}}\|^2/\varrho^2$ compared to

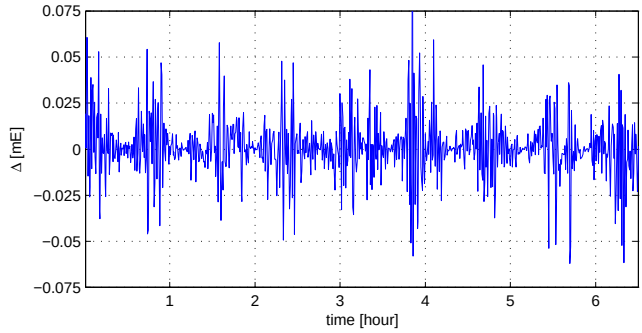


Figure 6.20: Differences of the numerically derived intersatellite range-rate with and without observation adjustment ($\Delta = \|\Delta\dot{\mathbf{r}}\|^2/\varrho^2 - \|\widehat{\Delta\dot{\mathbf{r}}}\|^2/\hat{\varrho}^2$)

that for the range-rate is not significant. It is due to the fact that the condition equation is imposed along the LOS. Consequently, one can expect the correction mainly along the LOS (see Fig. 6.19) and nearly no improvement on the other directions, see Figs. (6.21) and (6.22).
Figure (6.21) shows that adding the condition equation makes some corrections on the cross-track and the out-of-plane components. The along-track correction is nearly negligible. In contrast, the along-track velocity (Fig. 6.22) is the main component which benefits from the observations adjustment.

Although significant improvement has not been achieved by employing the condition equation it may be useful for detection of the outliers.
To set up the observation equations for recovery of the geopotential, $\|\Delta\dot{\mathbf{r}}\|^2$ is required. It should be derived by implementation of the numerical differentiator on $\Delta\mathbf{r}$ which is derived from GPS observations. It cannot be directly compared since there is no equivalent observable in the ranging system. Another investigation is performed based on the GRACE satellites dynamic orbits although considering the previous evaluations ensures obtaining the quantity with reasonable accuracy from GPS measurements.
For dynamic orbit evaluation only the spline differentiator is implemented since both the Newton and spline first-order differentiator gave nearly the same results on kinematic orbit.
Our first comparison was done on range-rate and intersatellite velocity vector derived from the kinematic and dynamic orbits. As Fig. (6.23) shows, deviations of the kinematic and dynamic rage-rate

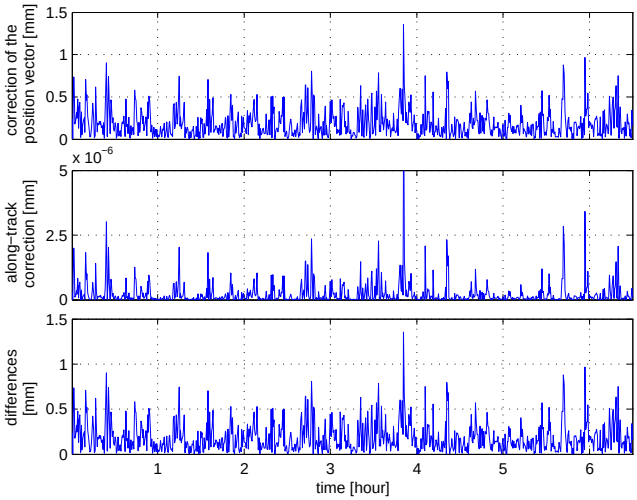


Figure 6.21: Relative position vector correction and its respective decomposition

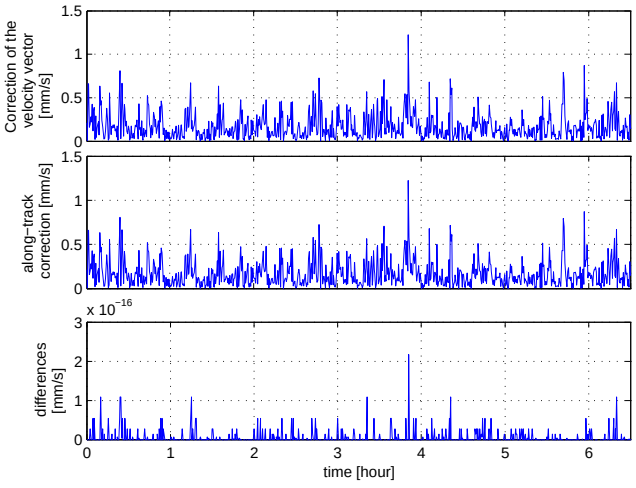


Figure 6.22: Relative velocity vector correction and its respective decomposition

and intersatellites velocity are less than $50\,\mu$m/s. Similarly, $\frac{\dot{\varrho}^2}{\varrho}$ and $\frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho}$ were compared and the achieved results depicted in Fig. (6.24).

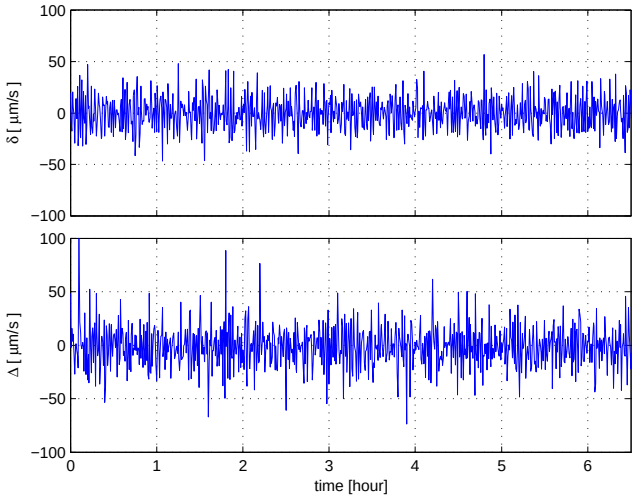


Figure 6.23: Differences of the range-rate and norm of the relative velocity vector $\|\Delta\dot{\mathbf{r}}\|$, based on the dynamic orbit of the GRACE satellites (top: $\delta = \dot{\varrho}_c - \dot{\varrho}_{dyn}$ bottom: $\Delta = \|\Delta\dot{\mathbf{r}}\|_c - \|\Delta\dot{\mathbf{r}}\|_{dyn}$)

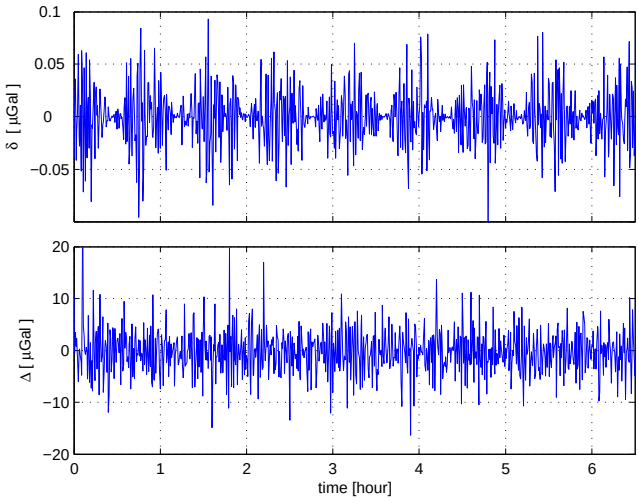


Figure 6.24: Differences of $\dot{\varrho}^2/\varrho$ and $\|\Delta\dot{\mathbf{r}}\|^2/\varrho$, based on the dynamic orbit of the GRACE satellites (top: $\delta = \frac{\dot{\varrho}_c^2}{\varrho} - \frac{\dot{\varrho}_{dyn}^2}{\varrho}$ bottom: $\Delta = \frac{\|\Delta\dot{\mathbf{r}}\|_c^2}{\varrho} - \frac{\|\Delta\dot{\mathbf{r}}\|_{dyn}^2}{\varrho}$)

Compared to Fig. (6.23), differences of $\frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho}$ was comparatively higher than that of $\frac{\dot{\varrho}^2}{\varrho}$. In other words, the amplification of noise level in the later case is dominantly higher than the other.
The last comparison of this part was done on the acceleration differences between the kinematic and dynamic orbits. Similar to Fig. (6.24), the differences of $\dot{\varrho}^2/\varrho^2$ was much lower than that of $\|\Delta\dot{\mathbf{r}}\|^2/\varrho^2$.

Comparing Figs. (6.13) and (6.25) shows that the differences of $\dot{\varrho}^2/\varrho^2$ for the kinematic orbit is nearly 50 times larger than that of for the dynamic orbit. With a pessimistic estimation we can expect 10 mE accuracy if the same ratio is valid for $\|\Delta\dot{\mathbf{r}}\|^2/\varrho^2$.
Furthermore, comparing Figs. (6.24) and (6.25) proves the superiority of the *satellite gradiometry using a satellite pair* over the acceleration difference which has been previously considered by (e.g. Heß and Keller, 1999; Keller and Sharifi, 2005; Sharifi and Keller, 2005).

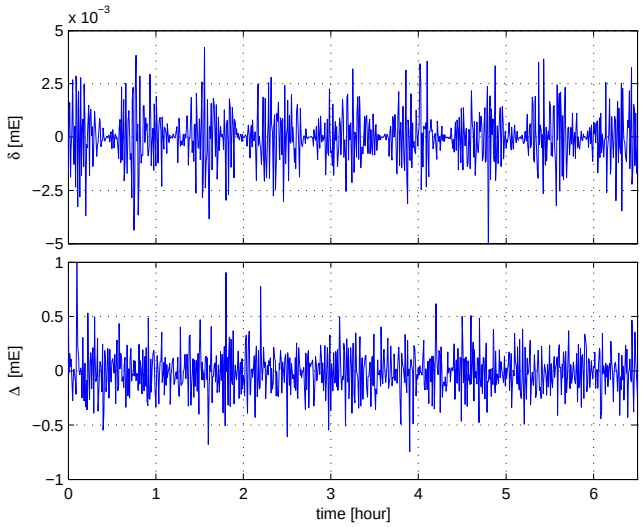


Figure 6.25: Differences of $\dot{\varrho}^2/\varrho^2$ and $\|\Delta\dot{\mathbf{r}}\|^2/\varrho^2$, based on the dynamic orbit of the GRACE satellites (Up: $\delta = \frac{\dot{\varrho}_c^2}{\varrho^2} - \frac{\dot{\varrho}_{dyn}^2}{\varrho^2}$ Down: $\Delta = \frac{\|\Delta\dot{\mathbf{r}}\|_c^2}{\varrho^2} - \frac{\|\Delta\dot{\mathbf{r}}\|_{dyn}^2}{\varrho^2}$)

6.4 Reference orbit employment

The numerical investigation carried out in the previous subsection showed that it was nearly impossible to obtain $\|\Delta\dot{\mathbf{r}}\|$ with reasonable accuracy even using sophisticated numerical differentiators. In order to bypass the problem, we are interested to find an alternative which yields better accuracy. The idea of using an appropriate reference field to generate a reference dynamic orbit and recovery of the residual gravity field is an alternative which has been proposed by many authors, see e.g. Hajela (1978), Rummel (1980) and Blaha (1992). In the following paragraphs we illustrate the theory of the idea and justify the applicability of the method for the GRACE mission.

As will become apparent, the range-acceleration ($\ddot{\varrho}$) is the dominant constituent of the acceleration difference as well as the gradiometry approach. The acceleration between two satellites with relative position vector $\Delta\mathbf{r}$ is expressed as (Rummel, 1980)[1]

$$\ddot{\varrho} = \langle \Delta\ddot{\mathbf{r}}\,,\, \mathbf{e}\rangle + \langle \Delta\dot{\mathbf{r}}\,,\, \dot{\mathbf{e}}\rangle. \tag{6.52}$$

Hajela (1978) showed that the second term on the right-hand-side, the problematic term in our approach, can be ignored for the HL-SST. However, for the LL-SST it is completely different so that their absolute values are nearly equal, see Fig. (6.26).

Instead of working with the *true* field, Rummel (1980) has suggested use of the residual field in LL-SST mode. By introducing a reference field and considering the equation of motion one can generate the respective dynamic orbit and derive the dynamic observations correspondingly

$$\ddot{\varrho}_d = \langle \Delta\ddot{\mathbf{r}}_d\,,\, \mathbf{e}_d\rangle + \langle \Delta\dot{\mathbf{r}}_d\,,\, \dot{\mathbf{e}}_d\rangle. \tag{6.53}$$

the residual range-acceleration can be derived by subtracting Eq. (6.53) from Eq. (6.52),

$$\delta\ddot{\varrho} = (\Delta\ddot{\mathbf{r}} - \Delta\ddot{\mathbf{r}}_d)^T\ \mathbf{e} + \frac{1}{\varrho}\left[\left(\|\Delta\dot{\mathbf{r}}\|^2 - \|\Delta\dot{\mathbf{r}}\|_d^2\right) - (\dot{\varrho} - \dot{\varrho}_d)\right] \tag{6.54}$$

in which we assume $\mathbf{e} = \mathbf{e}_d$ and $\varrho = \varrho_d$ to bring the relationship to convenient form. In Rummel (1980), more approximations have been used to simplify the second term on the right-hand-side of Eq. (6.54). Therein, the average magnitude of the first and second term has been compared and the second term has been ignored due to its negligible value.

[1]Equations (6.52) and (6.2) are identical although they are apparently different. For simplicity, the alternative representation is introduced in this part.

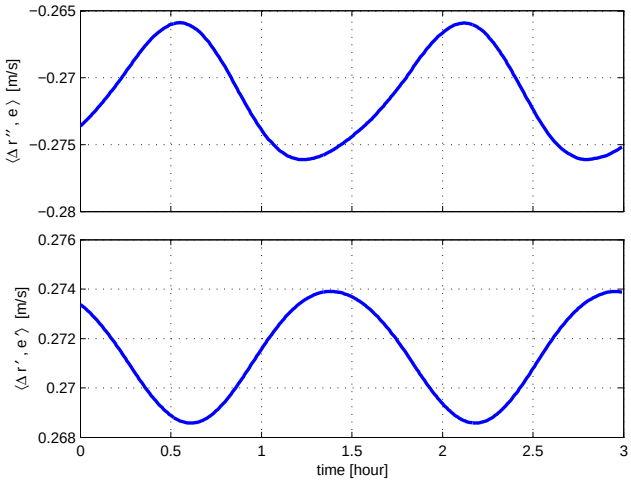


Figure 6.26: A comparative example of the first and second term on the right-hand-side of Eq. (6.52) (top: $\langle \Delta\ddot{\mathbf{r}}\,,\,\mathbf{e}\rangle$ bottom $\langle \Delta\dot{\mathbf{r}}\,,\,\dot{\mathbf{e}}\rangle$)

Herein, we alternatively analyse the observations of gradiometry type defined as

$$\frac{1}{\varrho}\Delta\ddot{\mathbf{r}}^T\,\mathbf{e} \;=\; \frac{\ddot{\varrho}}{\varrho} \;+\; \frac{1}{\varrho^2}\left(\dot{\varrho}^2 \;-\; \|\Delta\dot{\mathbf{r}}\|^2\right). \tag{6.55}$$

For the true field it can be derived using a combination of the ranger and GPS observations whereas that of for the reference field can be computed by means of dynamic orbit integration. Similar to the acceleration difference approach the second term on the right-hand-side of Eq. (6.55) cannot be ignored. As will become apparent, however, the quantity for the residual field is negligible.
For cross validation, the idea was applied on the simulated as well as real observations. The simulated true and reference observations were obtained by means of numerical integration using the EGM96 and EIGEN2 fields respectively. The gravitational force was considered as the only governing force for both fields. Different components were compared and the results were depicted in Fig. (6.27).

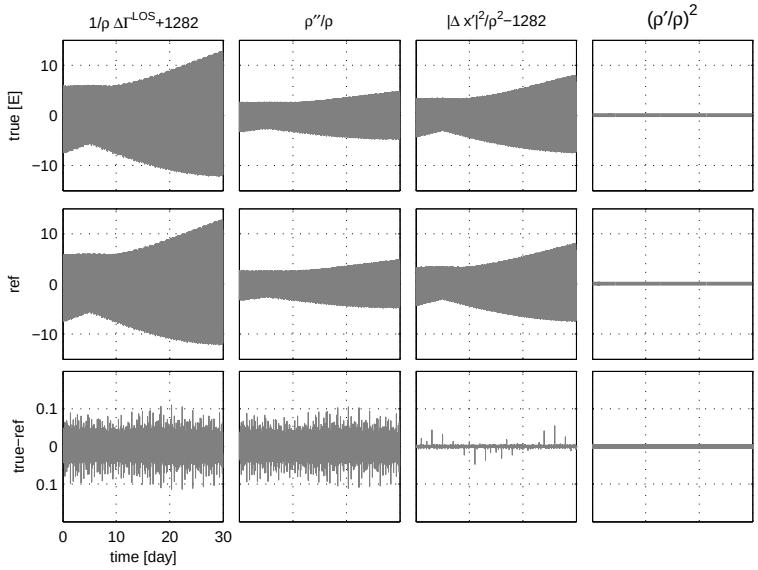


Figure 6.27: A one-month span of the true, reference and residual simulated observations

As the Figure shows, the intersatellite range-rate contribution to the gradiometry observation is comparatively much smaller than that of $\|\Delta\dot{\mathbf{r}}\|$ and the range-acceleration even for the true observations.

Nevertheless, the range-acceleration is the dominant component of the residual observations since it is the clearly pronounced constituent of the residual observations, see Fig. (6.27). In Figs. (6.27)-(6.29), the first-order time derivative is identified by a prime, and the second-order time derivative by a double prime.
For better understanding, the observations as well as the respective components for a two-revolution time span of the satellites are shown in Fig. (6.28). Since the adaptive reference orbit (Subsect. 4.2.1)

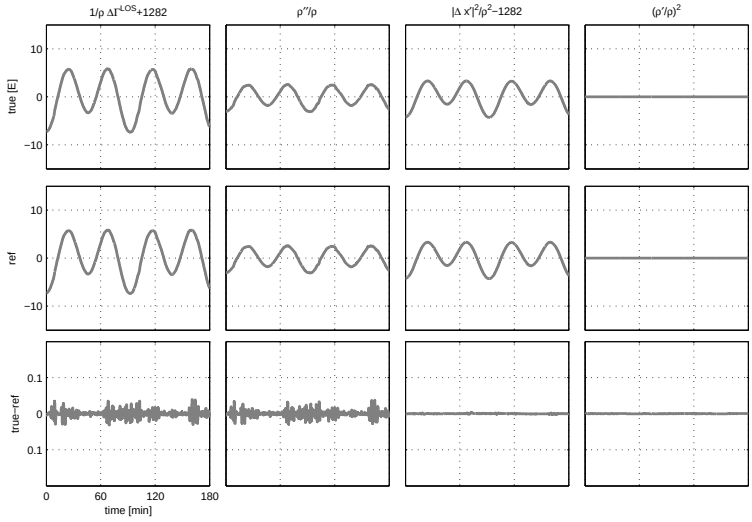


Figure 6.28: A three-hour span of the true, reference and residual simulated observations

has been used for the reference orbit a series of periodic blunders has appeared in the residual component $\delta\|\Delta\dot{\mathbf{r}}\|$ $(= \|\Delta\dot{\mathbf{r}}\| - \|\Delta\dot{\mathbf{r}}\|_d)$. They can be removed by smoothing the orbit at the updating points or by implementation of the least squares technique (Subsect. 4.2.2) for each subinterval.
Similarly, the analysis was carried out using the GRACE real as well as the computed dynamic observations for a period of one month (230-260/2003). To derive $\|\Delta\dot{\mathbf{r}}\|$ the spline differentiator was implemented on the satellites' kinematic orbit. The true, reference and residual observations were compared in Fig. (6.29).

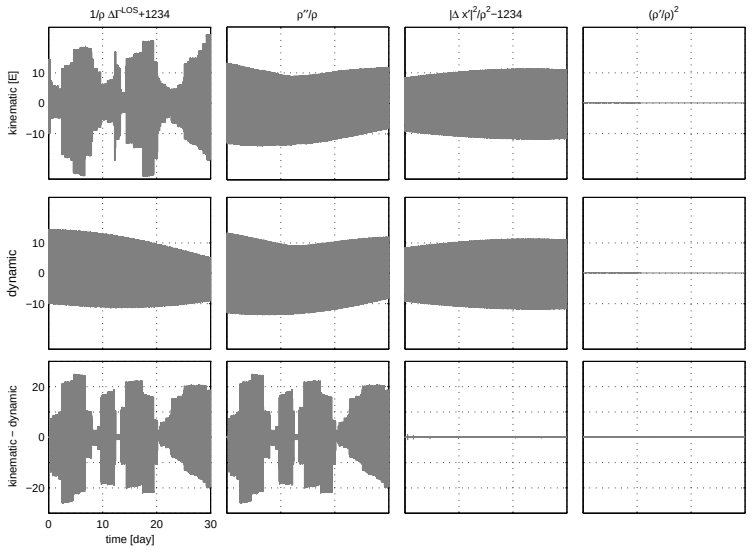


Figure 6.29: A one-month span of the true, reference and residual observations of real type

Compared to the simulated observations, the differences were significantly larger for the real obser-

vations. However, the patterns of the observations were identical, i.e., the range-acceleration was the main component of the residual field. In other words, the contribution of the second term on the right-hand-side of Eq. (6.55) was comparatively smaller than that of the first term.
Besides the numerical investigation performed on the acceleration differences given in Rummel (1980), our numerical analysis showed that the gradiometry mathematical model can be simplified by replacing the true field with the residual one. In other words, one can utilize the residual gravity field to bypass the problem of numerical differentiation required for computation of the relative velocity vector. Consequently, the residual observation equation of the gradiometry type (Eq. 6.9) can be simplified as

$$\delta l^{SG} = \frac{1}{\varrho}\,(\ddot{\varrho} - \ddot{\varrho}_d) = \frac{\delta\ddot{\varrho}}{\varrho}\,. \tag{6.56}$$

The new observation equations are set up only using the range-acceleration. In the following subsection, an alternative scheme is introduced in which both the range-rate and the range-acceleration are involved.
Similar analysis can be performed for the acceleration difference approach. The simplified observation equation is expressed as

$$\delta l^{\mathrm{LL}} = (\ddot{\varrho} - \ddot{\varrho}_d) = \delta\ddot{\varrho}\,. \tag{6.57}$$

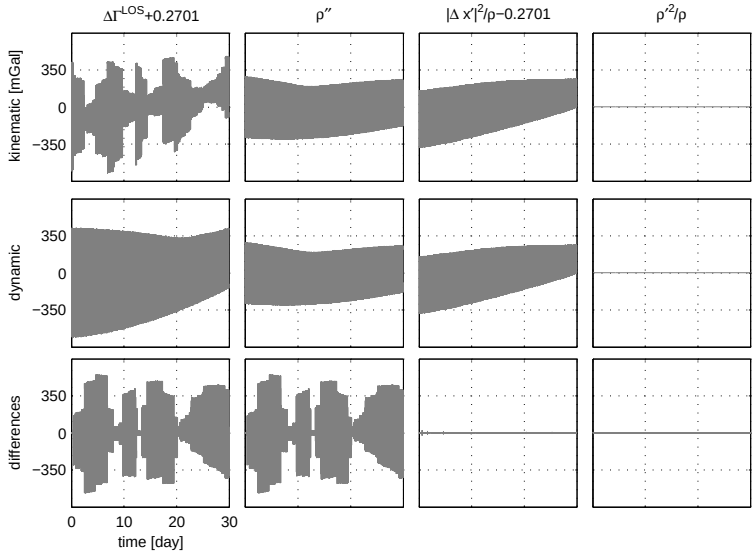


Figure 6.30: A one-month span of real true, reference and residual observations of the acceleration difference type

6.4.1 Computation of the relative velocity vector using the reference field

In the previous subsection we proposed the idea of replacing the true observation with the range-acceleration residuals corresponding to the sought-after residual gravity field. Using the aforementioned algorithm, the numerical differentiation problem can be bridged.
Alternatively, instead of the kinematic orbit we can use the dynamic orbit to derive the relative velocity vector. In other words, we obtain $\Delta\dot{\mathbf{r}}$ by numerical integration based on the reference field instead of implementing numerical differentiation on the kinematic orbit. The resultant velocity vector is then used to set up the true observation equations. The gradiometry observation equation for instance, is recast into the so-called *gradiometry quasi observation*

$$l_Q^{\mathrm{SG}} = \frac{\ddot{\varrho}}{\varrho} + \frac{\dot{\varrho}^2}{\varrho^2} - \frac{\|\Delta\dot{\mathbf{r}}\|_d^2}{\varrho^2} - \frac{1}{\varrho}\Delta\Gamma_{0l}^{\mathrm{LOS}}. \tag{6.58}$$

The only difference between the quasi-observation and the respective true observation equation is the

replacement of the relative velocity vector by its analogue in the dynamic orbit. It should be noted here that the last term on the right-hand-side of Eq. (6.58) is added to reduce the linearization error (Sharifi and Keller, 2005).
Analogously, the quasi observations of the acceleration difference can be written as

$$l_Q^{\mathrm{LL}} = \ddot{\varrho} + \frac{\dot{\varrho}^2}{\varrho} - \frac{\|\Delta\dot{\mathbf{r}}\|_d^2}{\varrho}. \tag{6.59}$$

The better the reference field, the smaller discrepancies are achieved. However, instead of solving the problem once, we can start with an arbitrary Earth's gravity model and improve the estimate iteratively. Clearly, this iterative process benefits the sequential approach.
Fig. (6.31), as a representative example for instance, shows the differences $\frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho^2} - \frac{\|\Delta\dot{\mathbf{r}}\|^2_{\mathrm{ref}}}{\varrho^2_{\mathrm{ref}}}$ for a one-month span of the GRACE satellites. EGM96 up to degree and order 120 and EIGEN2 up to 30, 90 and 120 were assumed as the pseudo-real model and the reference models respectively. As seen in

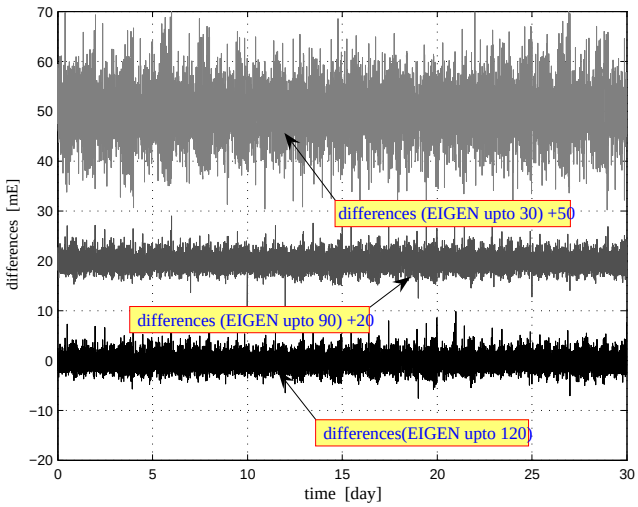


Figure 6.31: Differences between the quasi and the pseudo-real observations

Fig. (6.31) and Table (6.4), the differences are about the observation noise although for $n = 30$ the differences have a wider range. The idea has also been tested using real observations. The intersatellite

Table 6.4: The basic statistics of the simulated quasi gradiometry observation differences for different reference fields

Ref. model	min $[mE]$	max $[mE]$	mean $[mE]$	std $[mE]$
$n = 30$	-19.81	19.95	-0.08	4.18
$n = 90$	-9.95	8.98	-0.00	1.12
$n = 120$	-7.60	9.82	-0.00	1.20

relative velocity vector for a one-month $(230 - 260/2003)$ span of the kinematic and dynamic orbit of the GRACE satellites were computed. The differences were very small and comparatively negligible, see Fig. (6.32).
The differences can be neglected though they are relatively larger than that of the simulated observations. Therefore, instead of directly computing the relative velocity vector it can be modelled by means of dynamic orbit integration.
As a comparative example, residuals of the quasi observations generated by removing (Eq. 6.56) and modelling the relative velocity vector by means of the reference field has been compared in Fig. (6.33). Similarly, one can perform the comparison for the acceleration difference approach. As

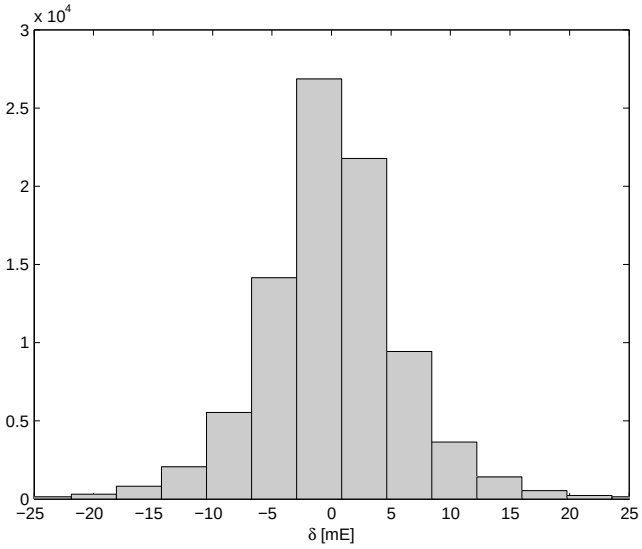


Figure 6.32: Differences of $\frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho^2}$ computed for a one-month span of the kinematic and dynamic orbit $(\delta = \left(\frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho^2}\right)_{\text{kin}} - \left(\frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho^2}\right)_d)$

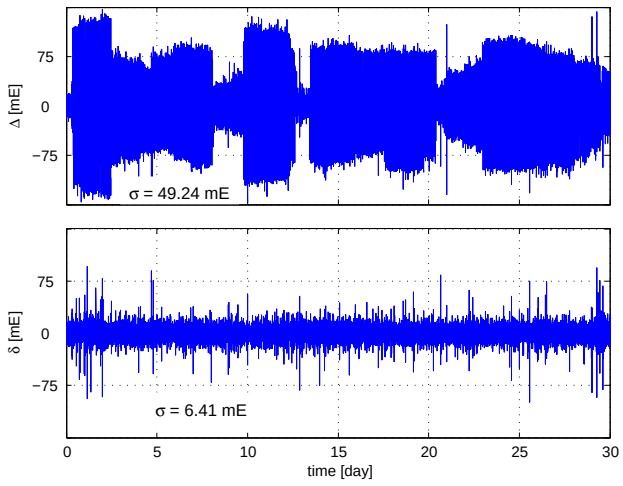


Figure 6.33: Residuals of the gradiometry quasi observations generated by removing and modelling the relative velocity vector (top: $\Delta := \frac{1}{\varrho^2}\left(\delta\dot{\varrho} - \delta\|\Delta\dot{\mathbf{r}}\|^2\right)$ bottom: $\delta := \left(\frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho^2}\right) - \left(\frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho^2}\right)_d$)

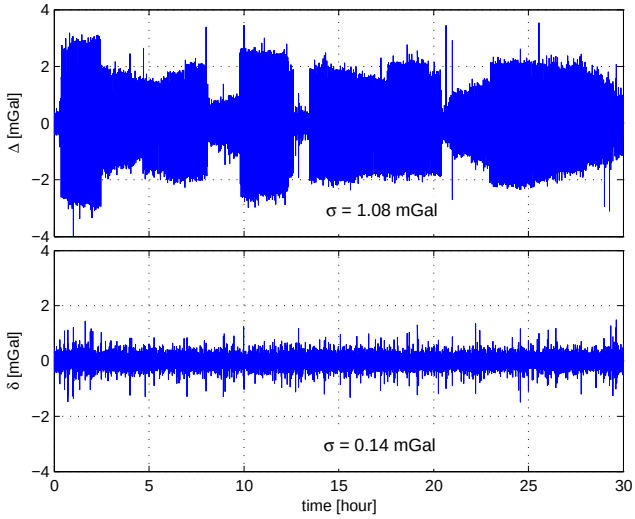


Figure 6.34: Residuals of the quasi observations of the acceleration difference type generated by removing and modelling the relative velocity vector (top: $\Delta := \frac{1}{\varrho}\left(\delta\dot{\varrho} - \delta\|\Delta\dot{\mathbf{r}}\|^2\right)$ bottom: $\delta := \left(\frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho}\right) - \left(\frac{\|\Delta\dot{\mathbf{r}}\|^2}{\varrho}\right)_d$)

seen, the differences are about few *mGal* for the observations of the reduced type (Eq. 6.57). They are significantly smaller for the quasi observations (Eq. 6.59). For a comparative analysis the basic statistics of the differences are tabulated in Table (6.5).

Table 6.5: The basic statistics of the reduced and quasi observations residuals

observation	min	max	mean	std.
Reduced Acc. diff. (Eq. 6.57)	−3.98 mGal	3.52 mGal	−0.00 mGal	1.078 mGal
Quasi Acc. diff. (Eq. 6.59)	−1.49	1.49	0.000	0.137
Reduced Grad. (Eq. 6.56)	−147 mE	147 mE	0.27 mE	49.24 mE
Quasi Grad. (Eq. 6.58)	−99	95	0.00	6.41

To sum up, the modelling scheme is superior to the removing approach because:

- the true observations can be directly utilized for the recovery of the gravity field,
- both the range-rate and range-acceleration are used,
- numerical differentiation is no more needed,
- higher accuracy can be obtained without any extra computation, and
- both removing and modelling the velocity term give a better solution for the gradiometry approach.

6.5 Observation decomposition

As discussed in chapter (5), spherical harmonic analysis in the space-wise approach is only applied to recover the spherical harmonic coefficients of the *invariant* functions. For the representation of the geopotential in spherical coordinates, the evaluation functional of the gravitational potential and its

radial derivatives are the only isotropic functionals which are defined on the Earth's gravitational potential. On the other hand, the GRACE observables both the acceleration differences and gradiometry type are *non-invariant* functionals. Therefore, the GSHA cannot be applied directly on the GRACE observables.

6.5.1 Decomposition of the gradiometry observation with linear approximation

For simplicity, neglect the arc-dependent constituent ($\pm \sin 2\alpha T_{23}$) or consider the mean contribution. Imposing the Laplace equation on Eq. (6.12) results in

$$\delta V_{\mathrm{m}}^{(xx)} = -\cos^2\alpha\,(T_{11}+T_{33}) + \sin^2\alpha T_{33} = -\cos^2\alpha T_{11} - \cos 2\alpha T_{33}. \tag{6.60}$$

Inserting Eq. (6.60) into Eq. (6.7) and dividing both sides of the equation by $-\cos^2\alpha$ yields the corresponding linear invariant and non-invariant components

$$\begin{aligned} f_{\mathrm{in}} &= \delta V^{(rr)} \\ f_{\mathrm{non}} &= \frac{\cos 2\alpha}{\cos^2\alpha}\left(\frac{\delta V^{(\lambda\lambda)}}{r^2\sin^2\theta} + \frac{\delta V^{(r)}}{r} + \frac{\cot\theta}{r^2}\,\delta V^{(\theta)}\right) \end{aligned} \tag{6.61}$$

where the invariant component is the dominant entry of the Marussi-tensor. As a representative example, $\delta V^{(xx)}$ as well as its respective components f_{in} and f_{non} have been calculated based on the already utilized incremental potential field (EGM96-EIGEN2). Figures (6.35),(6.36) and (6.37) show the observation and its decomposition. As seen in Fig. (6.36), the invariant component f_{in} imitates

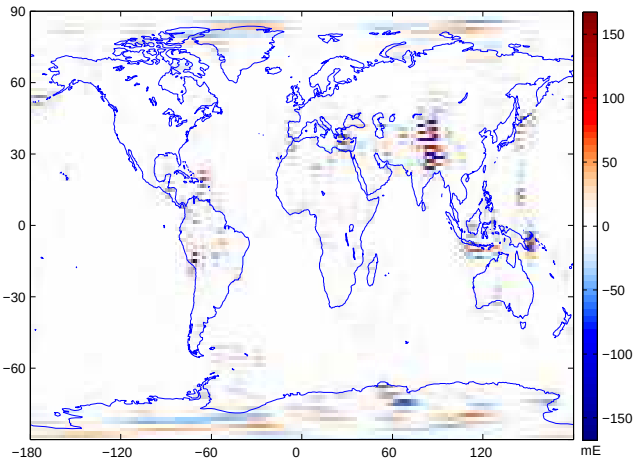


Figure 6.35: Original signal

the total signal pattern fairly well. Besides, the non-invariant component is considerably lower than the invariant one. Therefore, the iterative solution of the GSHA of the observations based on the Banach's fixed point principle will hopefully converge to the original signal.
A precise formulation will be achieved if the arc-dependent component is included. The invariant components are identical but the non-invariant function is respectively defined as

$$f_{\mathrm{non}}^{\pm} = \frac{\cos 2\alpha}{\cos^2\alpha}$$

$$\left(\frac{\delta V^{(\lambda\lambda)}}{r^2\sin^2\theta} + \frac{\delta V^{(r)}}{r} + \frac{\cot\theta}{r^2}\,\delta V^{(\theta)}\right) \pm \frac{2\tan\alpha}{r^2\sin^2\alpha}\left(-\sin\theta\,\delta V^{(\lambda\theta)} + \cos\theta\,\delta V^{(\lambda)}\right). \tag{6.62}$$

The arc-dependent component can be removed from either the in situ observations or the downward observations on the sphere. Therefore, two different approaches emerge for removing the non-invariant part:

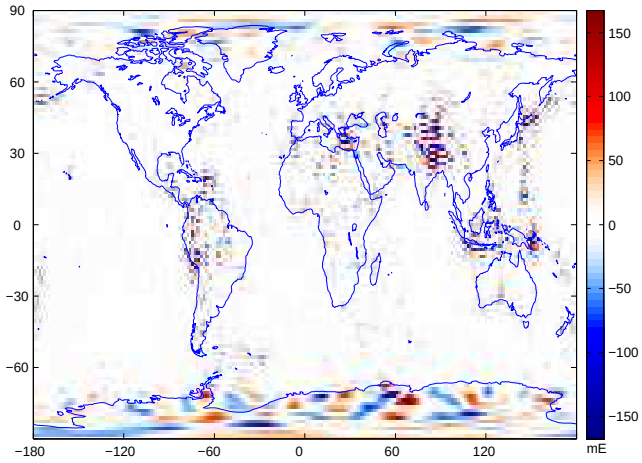


Figure 6.36: Invariant component (f_{in})

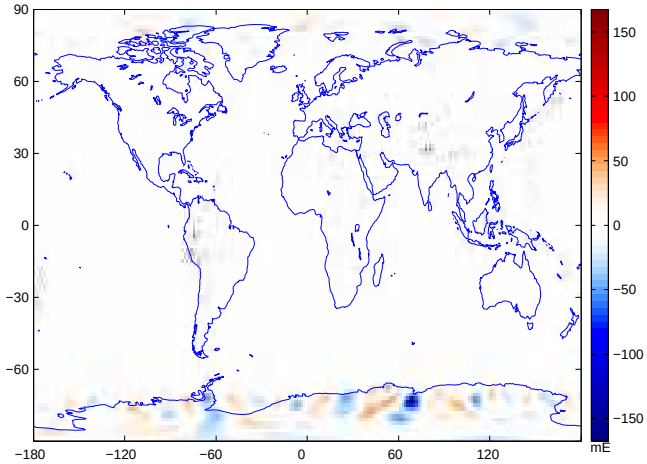


Figure 6.37: Non-invariant component (f_{non})

- Mapping the observations on the mean orbital sphere and removing the non-invariant component from the observation on the mean orbital sphere (*approach I*),

- Removing the non-invariant component at the satellite altitude and mapping the invariant part on the mean orbital sphere(*approach II*), see Fig. (6.38).

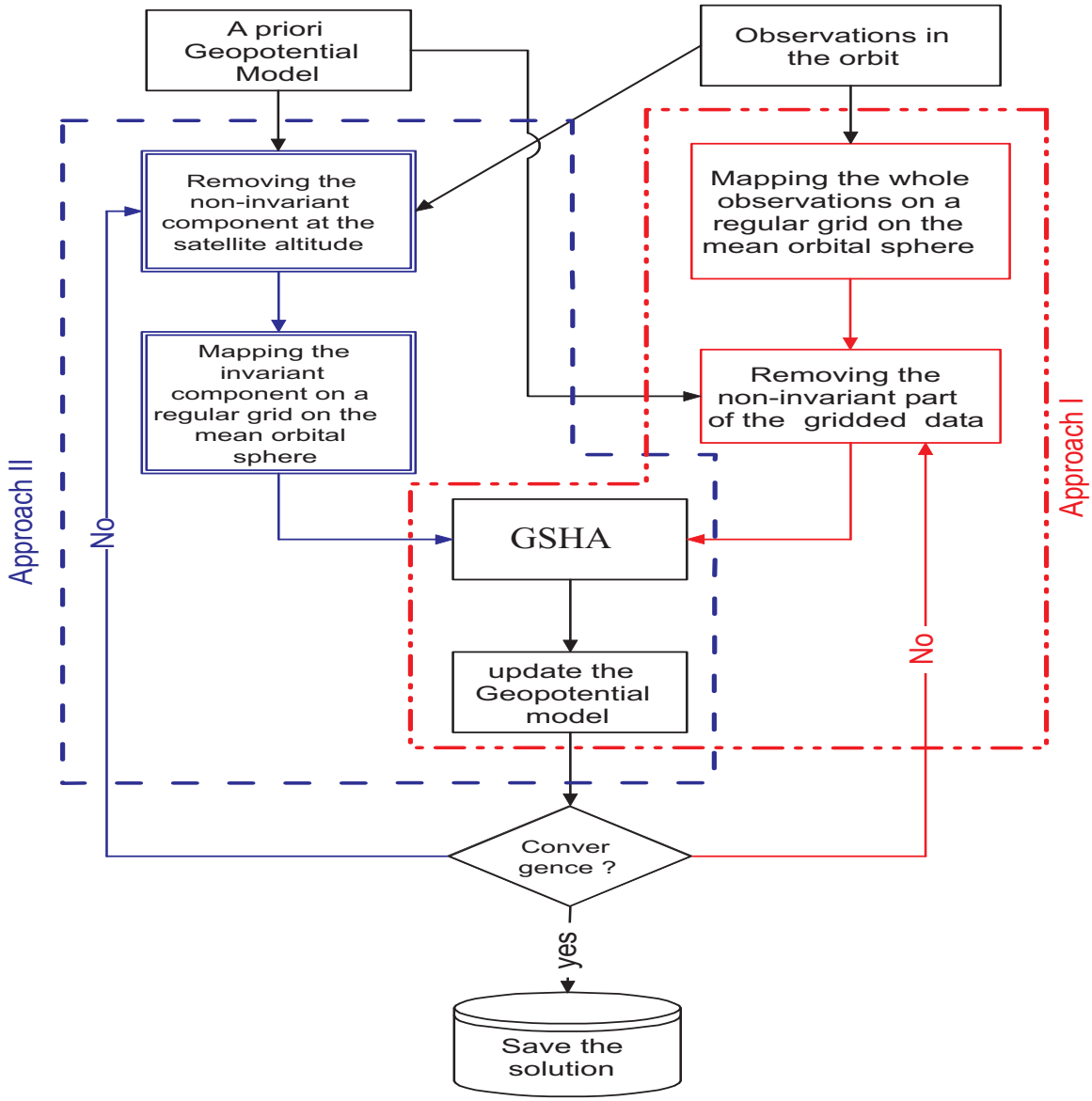


Figure 6.38: Flowchart of the decomposition of the original observation into invariant and non-invariant component

The second approach is superior since the non-invariant component as the nuisance part is removed before downward continuation. Consequently, one can expect lower reduction or gridding error. Moreover, one can correctly remove the non-invariant arc-dependent component's contribution. On the other hand, downward continuation is carried out once for the first approach whereas that of is equal to the number of iterations for the second approach.

6.5.2 Decomposition of the gradiometry observation with cubic approximation

Decomposition of the gradiometry observations with cubic approximation is similar to the linear type except the non-invariant component is rather complicated. Inserting the Laplace equation into Eq. (6.18) leads to the following decomposition

$$f_{\text{in}}^{\text{mc}} \;=\; T_{11} = \delta V^{(rr)} \tag{6.63}$$

$$\begin{aligned} f_{\text{non}}^{\text{mc}} \;=\; & \tfrac{\cos 2\alpha}{\cos^2\alpha}\, T_{33} - \tfrac{\varrho^2}{24\bar{r}^2}\Big(\cos^2\alpha\, T_{22}^{(\theta\theta)} + \sin^2\alpha\, T_{33}^{(\theta\theta)} + \\ & \tfrac{\sin^2\alpha}{\sin^2\theta}\, T_{22}^{(\lambda\lambda)} + \tfrac{\tan^2\alpha\sin^2\alpha}{\sin^2\theta}\, T_{33}^{(\lambda\lambda)} + \tfrac{4\sin^2\alpha}{\sin\theta}\, T_{23}^{(\theta\lambda)} - \tfrac{2\sin^2\alpha}{\sin\theta\tan\theta}\, T_{23}^{(\theta)}\Big) \end{aligned} \tag{6.64}$$

Assuming $\alpha = k\pi \quad k = 0, 1$ results in much simpler representations. For instance, Eq. (6.17) is recast into

$$\delta V_{\text{sc}}^{(xx)} \;=\; T_{22} + \frac{\varrho^2}{24r^2}\, T_{22}^{(\theta\theta)}. \tag{6.65}$$

The corresponding invariant and non-invariant components read

$$f_{\text{in}}^{\text{sc}} \;= -\delta V^{(rr)} - \frac{1}{r}\,\delta V^{(r)}, \tag{6.66}$$

$$f_{\text{non}}^{\text{sc}} \;= \delta V^{(rr)} + \frac{2}{r}\,\delta V^{(r)} + \frac{1}{r^2}\,\delta V^{(\theta\theta)} + \frac{\varrho^2}{24r^2}\left(\frac{1}{r^2}\,\delta V^{(\theta\theta\theta\theta)} + \frac{1}{r}\,\delta V^{(r\theta\theta)}\right) \tag{6.67}$$

Besides the first and second partial derivatives of fully normalized associated Legendre functions $\bar{P}_{nm}$, the fourth partial derivative is required for cubic approximation. Computation of the fourth derivative as well as the first and second ones can be performed by means of recurrence relationships. Using a set of simulated data, the Earth's gravity field is recovered using decomposition of the GRACE observations in the subsequent chapter.

6.6 Summary

This chapter started with the formulation of the SST problem in the low-low mode. Both linear and cubic approximation of the gradiometry approach were formulated and compared numerically.
Being the intersatellite relative velocity involved in the gradiometry observations was the motivation for devoting the second part of the chapter to numerical differentiation. The Newton and the spline differentiator were used for differentiation. They were compared numerically both for the first and second order differentiations.
Besides directly deriving the relative velocity vector from GPS observations, two different alternatives were proposed for bypassing the numerical differentiation. Using the residual observations and modelling the relative velocity vector based on the available geopotential models were the alternatives. Numerical analysis showed the later approach could be superior to the others due to its higher accuracy.
Furthermore, the gradiometry observations were decomposed into the invariant and non-invariant components for the purpose of GSHA. Two different iterative approaches were developed for recovery of the gravity field using the GRACE data. The proposed schemes are employed in the subsequent chapter for recovery of the field using synthetic observation.

Chapter 7

Recovery of the Earth's Gravity Field

Recovery of the Earth's gravity field with an unprecedented accuracy is the primary scientific objective of the GRACE mission. Although the Earth's gravitational acceleration has the main impact on the satellites motion the tidal as well as the non-gravitational accelerations perturb the space-borne observations. In order to portray the Earth's gravity field, the nuisance accelerations' contribution should be removed from the observations using either appropriate models or any auxiliary observations.
Contribution of the tidal accelerations as deterministic components are usually removed from the satellites' observables. In contrast, the effect of the non-gravitational acceleration i.e., the atmospheric drag and solar radiation pressure, are observed by the satellite onboard accelerometers and removed from the measurements at the preprocessing step.
after removing the nuisance components, different approaches can be employed for the recovery of the field. In the previous chapter (chapter 6), the GRACE mission observations were formulated as the gradiometry observations. As mentioned in chapter 5, spherical harmonic coefficients can be obtained from the reduced gradiometry observations on the mean orbital sphere by means of global spherical harmonic analysis.
In this chapter, the numerical results achieved by the implementation of the aforementioned approach are presented. For cross-validation of the developed algorithm, synthetic data is used.

7.1 Data processing

The Earth's gravity field recovery process using the satellite data consists of few steps. A general flowchart of theses steps and their interconnection is shown in Fig. (7.1).
In general, the process can be divided into two independent parts:

Derivation of the true observations: The satellites' kinematic orbits and the ranging information are the main observations of the mission. The ranging observations i.e., range-rate and range-acceleration are two constituents of the observations vector. Besides, they can be used for detection of the hidden outliers in GPS observations by imposing Eq. (6.49) on the observations at each evaluation point. Furthermore, the satellite onboard accelerometers measure the non-gravitational accelerations. The star cameras observations are used for transformation of the accelerometer observations from the *satellite body-fixed frame* to the inertial frame i. e., the celestial reference frame (Ditmar et al., 2005).

Computation of the reference observations: Despite the kinematic part, the reference observations are derived by implementation of the existing models for the gravity field, and the tidal accelerations. Compared to the kinematic observations the reference observations only contain the contribution of the gravitational accelerations. Moreover, the reduced dynamic orbit is a joint product of the dynamic and kinematic observations. Herein, it is considered as an input.

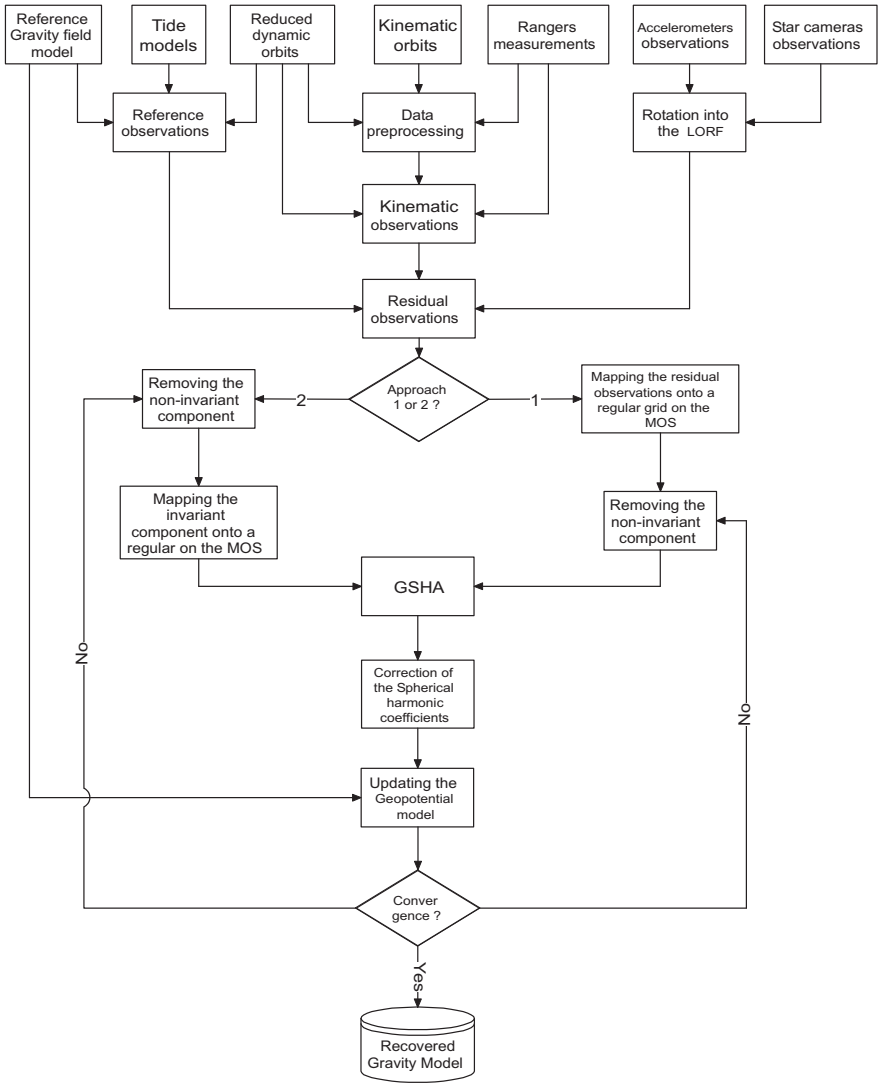


Figure 7.1: Flowchart of data processing

As explained in section (6.3), use of the reduced-dynamic relative velocity vector results in a higher accuracy in comparison with the relative kinematic velocity derived by numerical differentiation. Therefore, the reduced-dynamic relative velocity vector is used for derivation of the kinematic observations. Furthermore, the reduced-dynamic orbits will be used as the blunder detector in the kinematic orbit.

When the kinematic and the reference observations are available, the residual observations are computed

$$\delta l = \frac{\delta \ddot{\varrho}}{\varrho} + \frac{\dot{\varrho}^2 - \dot{\varrho}^2_{ref}}{\varrho^2} - \frac{1}{\varrho}\langle \Delta \mathbf{a}, \mathbf{e} \rangle, \tag{7.1}$$

where $\Delta \mathbf{a}$ is the non-gravitational acceleration difference between the GRACE satellites. Sharifi and Keller (2005) showed that working with the residual field is more convenient for the gradiometry approach since the linearization error is small and can be ignored. The residual observations just contain the residual gravity field contribution. Hence, they form the observation vector which can be used for the recovery of the residual gravity field.
As stated in section (6.5.1), two approaches can be used for obtaining the invariant component on a

regular grid on the mean orbital sphere. In the first approach, the residual observations are mapped on a regular grid on the mean sphere and the non-invariant component is subtracted from the gridded residual observations. In the second approach, the non-invariant component is removed from the in situ observations and the remaining part i.e., the invariant component is reduced on the regular grid on the mean sphere. In both cases, the initial guess of the non-invariant component is assumed zero.

7.2 Numerical results

To validate the proposed iterative recovery algorithm, EGM96 and EIGEN2 were assumed as the pseudo-real and the reference gravity fields respectively. Using the residual field (Fig. 7.2), various observables were synthesized on the mean orbital sphere as well as the satellite altitude. For ease of

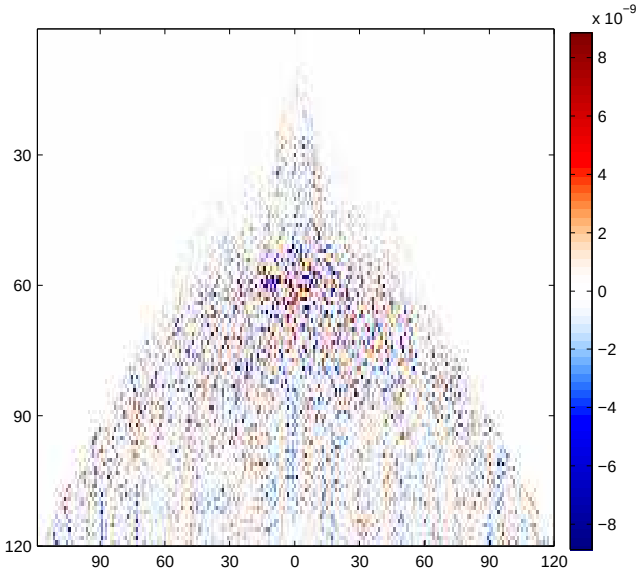


Figure 7.2: Spectrum of the residual field (EGM96-EIGEN2)

operation, the tidal accelerations as well as the non-gravitational ones were neglected.

7.2.1 Simulated along-track gravity gradient on the mean orbital sphere

The along-track gravity gradient component for an inclined orbit (Eq. 6.60) was the first synthesized observable. The signal and its respective decomposition were depicted in Figs. (6.35), (6.36) and (6.37).
For initialization of the iterative approach, the non-invariant component was assumed equal to zero. The process iterated 50 times and the recovered signal spectrum's residuals were plotted in Fig. (7.3). From Fig. (7.3), it can be inferred that the iterative solution successfully extracted the non-invariant part from the total signal. In other words, the non-invariant component was gradually removed from the total signal, see Fig. (7.4). Consequently, the more iterations the better solution is achieved.
Comparing Figs. (7.4) and (6.37) implies that the non-invariant component has been removed from the along-track component up to 95 %.
Furthermore, the absolute and relative errors of the estimated signals were calculated and achieved results were shown Fig. (7.5). They confirmed the previous figure. As seen in Fig. (7.5), the relative error was less than 10% for $n \leq 100$. It reached 15% in its worse case.
In order to analyse convergence of the iterative solution, $\epsilon = \frac{\|\mathbf{x}^{n+1}-\mathbf{x}^{n}\|_{L_2}}{\|\mathbf{x}\|_{L_2}}$ was computed for each iteration and the results were drawn in Fig. (7.6). As expected, ϵ dramatically decreased for the first few iterations and the improvement was not pronounced for the higher iterations.
Representation of the achieved results in terms of error degree variances is another criterion for judgement. In Fig. (7.7), error degree variances were depicted for the first few iterations. Similar to

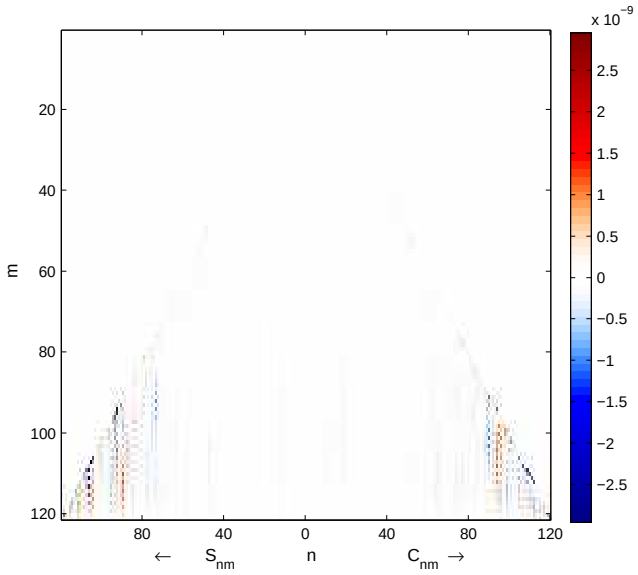


Figure 7.3: The recovered signal spectrum's residuals after 50 iterations

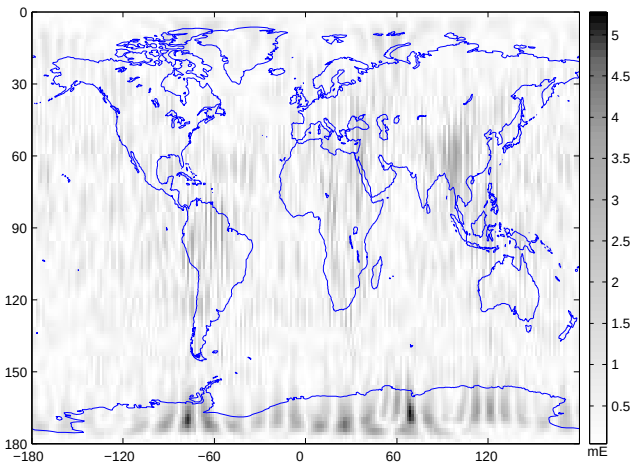


Figure 7.4: The non-invariant component's residuals (absolute values) after 50 iterations

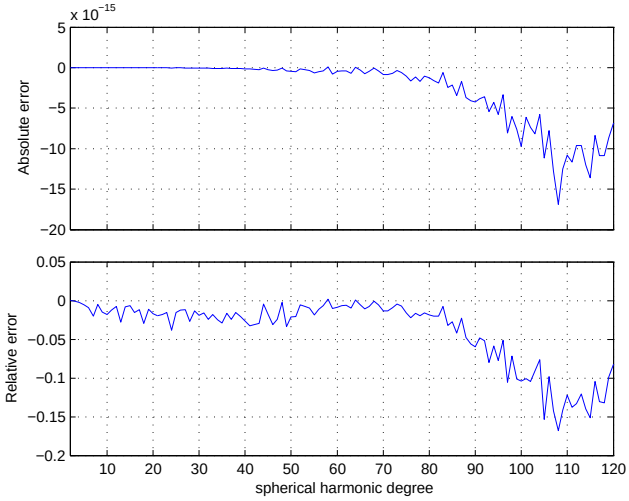


Figure 7.5: Absolute and relative error of the estimated spherical harmonic coefficients after 50 iterations

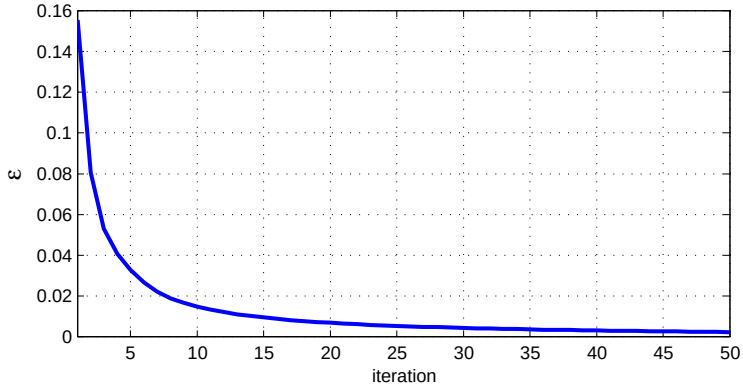


Figure 7.6: convergence rate of the iterative process ($\epsilon = \frac{\|\mathbf{x}^{n+1}-\mathbf{x}^n\|}{\|\mathbf{x}\|}$)

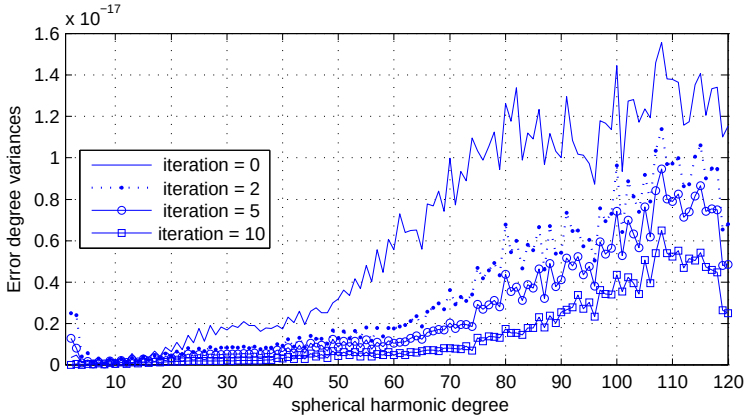


Figure 7.7: Error degree variances for the first few iterations

Fig. (7.6), highly pronounced differences between the error curves, showing significant improvement, are observed at the first few iterations.
This simple example proves the iterative algorithm of the first type (Approach I) is able to remove the non-invariant component part successfully. Of course, it is far from reality due to the absence of the errors which contaminate the observations in practice. The next simulations become more realistic by involving the potential error sources in the space-wise approach.

7.2.2 Analysis of the simulated residual observations (linear approximation)

In order to investigate the performance of the first approach for the GRACE mission observable, the residual observations δl (7.1) were simulated at the satellite's orbits. In the linear gradiometry approach, the observed quantity is written as a linear functional on the gravity field $\delta V^{(xx)}$ (Eq. 6.60). Consequently, the previously implemented spherical harmonic analysis procedure can be employed for the simulated observations as soon as they have been downward continued on the mean orbital sphere.
The simulated observations at the satellite altitude were reduced onto the Gauss grid defined on the mean orbital sphere as

$$\begin{cases} \theta_i &= \arccos(x_i) \\ \lambda_j &= -\frac{2\pi}{N_{\max}} + \frac{\pi}{N_{\max}} j \end{cases} \tag{7.2}$$

where, x_i are the zeros of $P_{N_{\max}+1}$ and $N_{\max} = 120$. Before performing the analysis, the reduced observations were compared with the simulated values at the Gauss grid and the differences were plotted in Fig. (7.8). The linear approximation is inaccurate at the polar regions. Therefore, the de-

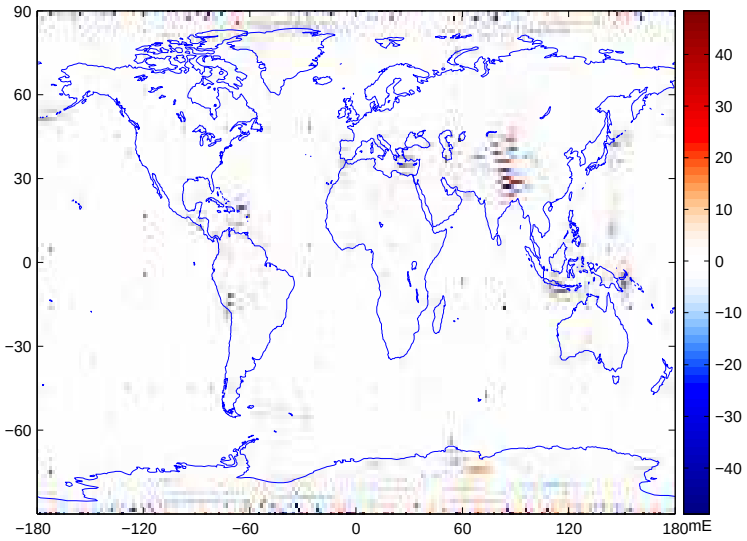


Figure 7.8: Observations deviations from $\delta V^{(xx)}$ (filtered)

viations of the residual observation from the along-track component are really meaningful. Therefore, the residual observations whose deviations were more than 50 mE were replaced by the along-track component at the respective grid points. The basic statistics of the deviations were tabulated in Table (7.1).

Table 7.1: Basic statistics of the reduced δl with computed values on the Gauss grid

Observable	Statistic			
	min	max	mean	std.
δl	−49.61 mE	49.20 mE	0.005 mE	3.70 mE

The GSHA has been carried out on the gridded observations and the achieved results, after 30 iterations, have been shown in Fig. (7.9). With the achieved results after the first few iterations, one

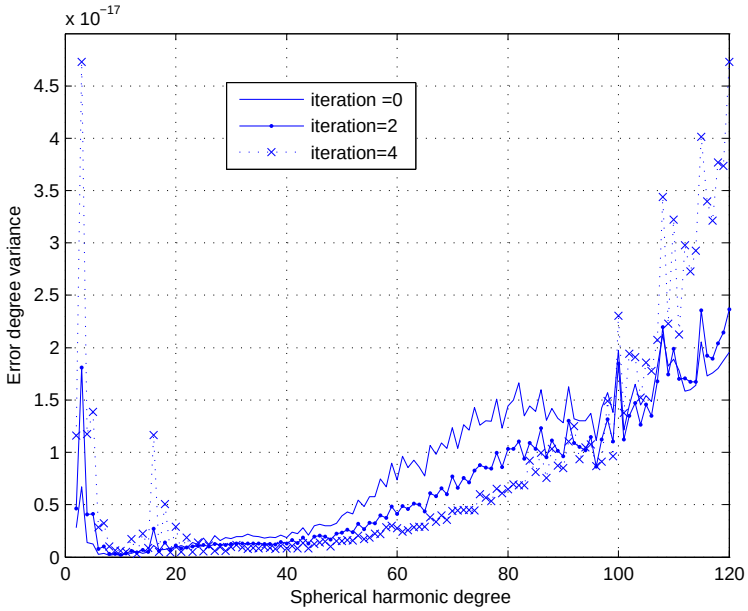


Figure 7.9: Error degree variances for the first few iterations of the simulated observations analysis

could expect divergence of the iterative solution. Consequently, it has led us to halt the process just after a few iterations.
Let us have a deeper look at the signal $\delta V^{(xx)}$, to find out why the process has converged for the first case i.e., simulated along-track on the MOS, whereas it has diverged for the second one i.e., the reduced observations. Fig. (7.10) shows four different classifications of $\delta V^{(xx)}$. As seen, less than 1%

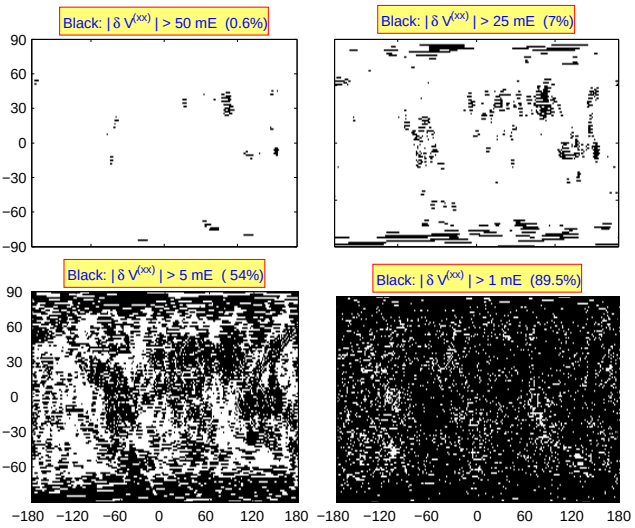


Figure 7.10: $\delta V^{(xx)}$ classification

of the data were larger than 50 mE. Just around 54% of the data were larger than 5 mE. This figure brings us to the conclusion that the GSHA based on such a rather weak signal can be easily violated by noise as well as disturbing signal.
The followings are the potential error sources:

1. Aliasing/omission error
2. Utilizing the nominal instead of real orbit ($\sin\alpha \approx \frac{\cos I}{\sin\theta}$)
3. Interpolation error
4. Linearization error

Among them, the linearization error as well as the interpolation error have been discussed in Chapters 3 and 6 respectively. In the following paragraphs, we briefly explain the two others.

Aliasing error

Real observations contain all frequencies up to $N = \infty$ while we just consider the signals with frequencies up to a certain limit ($N = N_{max}$). The presence of the signal parts higher than N_{max} influences the recovery process especially at the high-degree components. To have a rough estimate of the error, the GRACE observable $V^{(xx)}$ was simulated using $EGM96$ up to 120 and 360. The differences can be interpreted as an approximate estimation of the aliasing error if we consider $EGM96$ up to 360 as the pseudo-real gravity field model. The achieved results were sketched in Fig. (7.11).

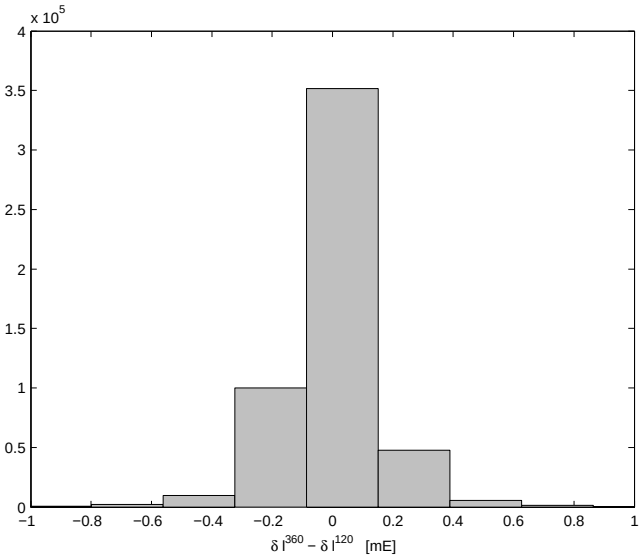


Figure 7.11: Aliasing error (l^N observations of the Eq. (6.6) type generated using EGM96 up to degree and order N)

Compared to the previously investigated error, the aliasing error seems to be out of question at least as long as the observations are influenced by the linearization error.

The nominal instead of the real orbit

As mentioned earlier, $\mathbf{e}$ given in Eq. (2.79) is valid for the nominal orbit with a constant inclination I. In reality, this assumption is violated to some extent due the presence of the disturbing forces. To be sure of achieving satisfying accuracy even by using the nominal inclination (89.5°), the real inclination deviations from the nominal one were plotted in Fig. (7.12). As seen, the maximum deviation is just 20 arcsec which seems to be negligible.
To sum up, the already mentioned errors were compared and their basic statistics tabulated in Table (7.2).
Compared to the others, the linearization error is noticeable. Therefore, we should focus on the linearization error to improve the grided observations.
In order to reduce the linearization error, the higher order terms of Taylor's expansion of the gradiometry equation (Eq. 6.4) should be considered. As explained in Chapter 6, the cubic approximation is the alternative, which yields comparatively higher accuracy.

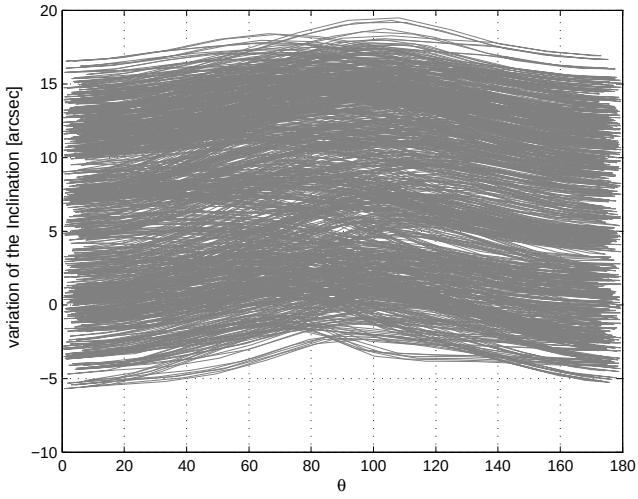


Figure 7.12: Variation of the orbit inclination around the nominal inclination (89.5 °)

Table 7.2: Basic statistics of different type of errors

quantity	statistic			
	min (mE)	max (mE)	mean (mE)	std. (mE)
δl_{120} (signal)	−107.7	105.9	0.00	11.67
$\delta l^{360} - \delta l^{120}$ (aliasing)	−1.75	1.81	0.00	0.15
$\delta l^{120} - \delta l^{120}_{red}$ (reduction)	−4.99	4.88	0.00	1.19
Δ_{lin} (linearization)	−29.89	35.65	0.00	2.39
$\delta V^{xx} - \delta V^{xx}_f$ (model)	−18.37	18.41	0.00	0.99
$\delta V^{xx}_{\pm} - \delta V^{xx}_f$ (model)	−0.51	0.74	0.00	0.06

reduced observation: $\delta l^{120}_{red} := \left(\frac{\ddot{\varrho}}{\varrho}\right)^{120}_{EGM96} - \left(\frac{\ddot{\rho}}{\rho}\right)^{120}_{EIGEN2}$

Table 7.3: The ascending and descending non-invariant functions and their respective mean and differences basic statistics

quantity	min [mE]	max [mE]	mean [mE]	std [mE]
Ascending	−106.55	72.17	−0.09	9.89
Descending	−107.50	71.41	−0.09	10.03
Mean	−107.03	71.79	−0.09	9.92
Differences	−15.59	32.40	0.00	1.80

Moreover, the arc-dependent terms of the non-invariant constituent have been assumed to be dropped out by averaging. To examine the validity of this assumption, the non-invariant function corresponding to the ascending and descending arcs and their respective mean and differences for the Gauss grid were computed. The achieved results and their basic statistics were shown in Fig. (7.13) and Table (7.3). Although the differences between the ascending and descending component are small

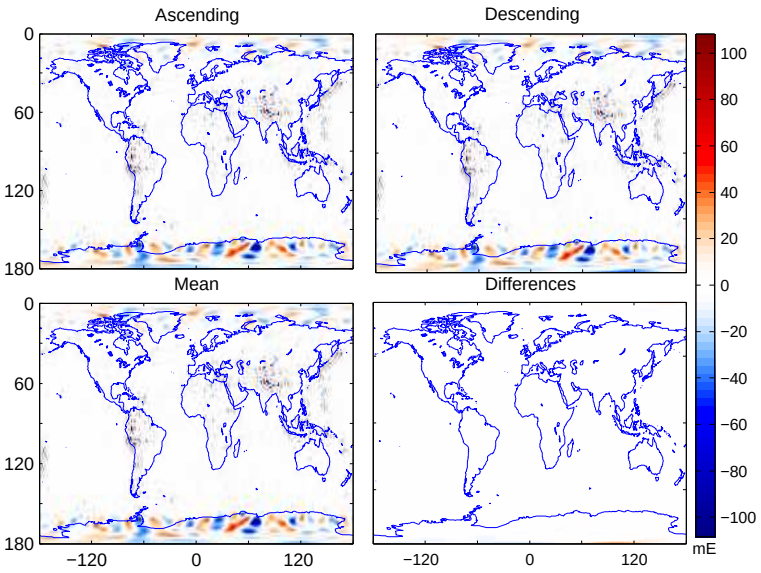


Figure 7.13: The ascending and descending non-invariant functions and their respective mean and differences on the MOS

they may influence the solution. For the sake of completeness, we consider the arc-dependent term as well.

7.2.3 Analysis of the simulated residual observations (cubic approximation)

Compared to the linear approximation, computation of the cubic approximation of the non-invariant function requires more information. The intersatellite range for instance, should also be downward continued on the mean orbital sphere if the first iterative approach is employed. However, the non-invariant component can be removed at the satellite altitude (*approach II*). Hence, the invariant component is downward continued on the MOS.
Furthermore, using the gradiometry equation of the cubic-order reduces the truncation error of the expansion (Eq. 6.4), see Fig. (7.14). Consequently, the residual invariant function can be replaced by the invariant function corresponding to the disturbing potential. In other words, we can just remove an ellipsoidal reference field's contribution from the invariant function. In comparison with the residual invariant function, the signal to noise ratio will be amplified.

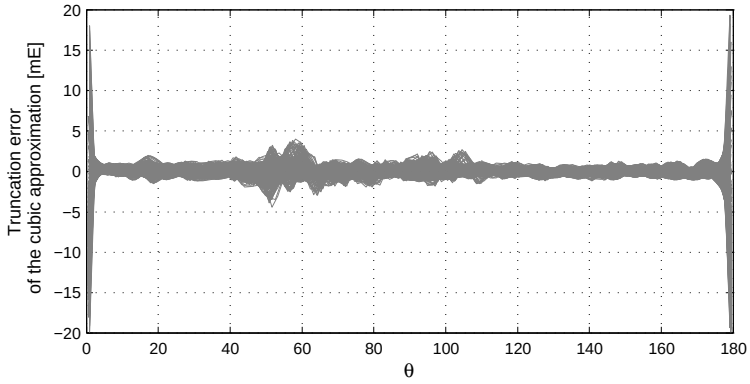


Figure 7.14: Truncation error of the cubic approximation of the gradiometry equation corresponding with the disturbing potential

As shown in Fig. (7.1), the non-invariant component can be removed from the observations at the satellites altitude. Since the computation is carried out in the satellite orbit the ascending and descending arcs are distinguished and there is no need to use the mean observations. After removing the non-invariant component the invariant component (remaining part) is reduced on the MOS. Furthermore, the invariant function can be replaced by the harmonic function

$$f_{\text{in}}^{\text{H}} = r^2 \delta V^{(rr)}. \tag{7.3}$$

Therefore, *analytical continuation* can also be used for downward continuation of the function onto the MOS (Hofmann-Wellenhof and Moritz, 2005).
Dividing both sides on Eq. (6.17) by $-\cos^2\alpha$ is recast into

$$f_{\text{in}}^{\pm\text{c}} = \delta V^{(rr)} \tag{7.4}$$

$$\begin{aligned} f_{\text{non}}^{\pm\text{c}} &= f_{\text{non}}^{m\text{c}} \mp 2\tan\alpha\, T_{23} \\ &\pm \frac{\varrho^2}{24r^2}\left(\frac{\sin 2\alpha}{2\sin\theta\tan\theta}\, T_{22}^{(\lambda)} + \frac{\sin^2\alpha\tan\alpha}{\sin\theta\tan\theta}\, T_{33}^{(\lambda)} - \frac{\sin 2\alpha}{\sin\theta}\, T_{22}^{(\theta\lambda)} - \right. \\ &\left. \frac{2\sin^2\alpha\tan\alpha}{\sin\theta}\, T_{33}^{(\theta\lambda)} - \frac{2\sin^2\alpha\tan\alpha}{\sin^2\theta}\, T_{23}^{(\lambda\lambda)} - \sin 2\alpha\, T_{23}^{(\theta\theta)}\right) \end{aligned} \tag{7.5}$$

where, $f_{\text{non}}^{m\text{c}}$ is the mean non-invariant function of the cubic-order (Eq. 6.64). The additional terms on the right-hand-side of Eq. (7.5) are the contribution of the arc-dependent components.
In order to evaluate the performance of the cubic approximation in the iterative approach, the already simulated observations are considered. The non-invariant component and the ellipsoidal reference field are computed using the EIGEN2 model. Therefore, the invariant function corresponding to the disturbing gravity field is reduced on the MOS.
The reduced observations were plugged into the GSHA processor and the achieved results were plotted in Fig. (7.15). Errors of the recovered coefficients were depicted in terms of the error degree variances for the first few iterations. As shown in Fig. (7.15), the medium wavelengths $(10-110)$ were recovered successfully just by a few iterations. However, the frequencies lower than 10 and higher than 110 were erroneous. In other words, the method could successfully recover the medium wavelengths.
In the time-wise approach, the high-degree coefficients are recovered efficiently by introducing the regularization parameter (Ditmar et al., 2005). For the recovery of the high-degree coefficients, a similar modification may be helpful in the space-wise approach. Keeping the low-degree coefficients constant, the problem of the recovery of the low-degree wavelengths can also be bypassed.
Furthermore, Fig. (7.16) shows the original invariant and non-invariant signals and their respective residuals after three iterations.
The iterative approach could successfully remove the remaining part of the non-invariant function and recover the spherical harmonic coefficients. The remaining part of the non-invariant function is

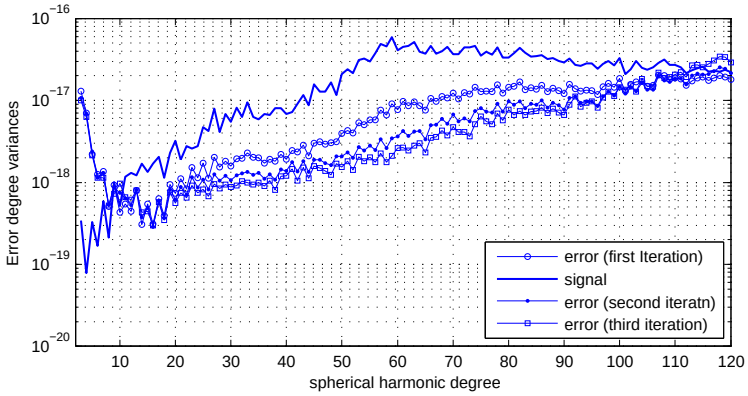


Figure 7.15: The error degree variances for the first few iteration using the *approach II*

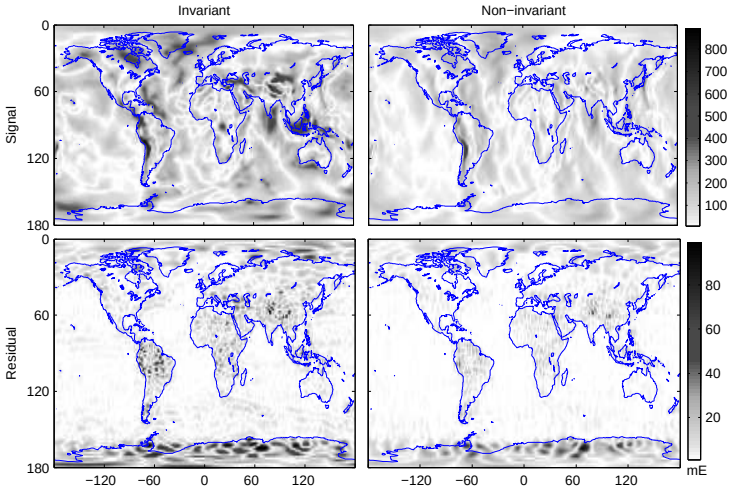


Figure 7.16: The original invariant and non-invariant and their respective residuals after 3 iterations

less than 10% of the original non-invariant function.

7.3 Summary

We have implemented the proposed recovery schemes for determination of the gravity field using simulated observations for a configuration of the GRACE type mission.
For the observation of the gradiometry type between a pair of satellites either the residual observation or the high-order approximation of the equation has to be used. For the time-wise approach, the residual observations with linear gradiometry observation equation can successfully be used (Sharifi and Keller, 2005). However, in the space-wise approach, one has to remove the non-invariant component before applying the GSHA. Incomplete removal of the non-invariant component and the presence of the linearization error make the first approach (approach I) inefficient in the space-wise approach. The second approach (approach II) has alternatively been employed. In contrast to the first approach, the third order approximation of the gradiometry equation has been used. Consequently, one can expect smaller truncation error compared to the linear approximation.
Moreover, the non-invariant component has been removed at the satellite altitude. Therefore, the arc-dependent terms have been correctly subtracted from the observations. Additionally, removing an ellipsoidal reference field instead of a high-degree field has led us to an amplified invariant signal. As the numerical comparison showed, the latter approach outperforms the first one in the space-wise representation.

Chapter 8

Summary, Achievements, Conclusions and Recommendations

This chapter summarizes the work contained in the dissertation, outlines the major contribution, lists the main achievements, and addresses a number of topics that deserves additional investigation.

8.1 Summary

This dissertation was started by formulation of the LL-SST problem in the brute-force approach. It followed by the semi-analytical approach obtained by imposing some constraints on the brute-force approach. The problem was formulated as 1D FFT problem by assuming the orbit of repeat type with uninterrupted stream of observations and mapping the observations onto the nominal orbit (*the time-wise approach in the frequency domain*). Alternatively, recovery of the gravity field using the LL-SST observations reduced into $N+1$ determined linear system of equations with $\frac{N}{2}+1$ unknowns at most by exploiting the orthogonality of the basis functions (*the time-wise in the time domain*).

Representation on a torus can be achieved by assuming a constant inclination ($I = I_0$) besides a constant radius for the nominal orbit. Having a full periodic function in both directions has made the old difficult problem quite simple.

The semi-analytical approach then proceeded with the space-wise approach by mapping the observations on the mean orbital sphere. From a numerical point of view, it was the simplest representation among the others though its formulation became rather complicated due to the special configuration of the GRACE mission.

Mapping the observations onto an ideal geometric configuration with constant radius and /or inclination was common both to the semi-analytical and space-wise approach. In order to handle the greatest single obstacle of the approaches, we did a comprehensive study on scattered data modelling. Optimization of the basis functions in the polynomial as well as the rational approximation and optimal selection of the shape parameter of the RBF interpolation methods were closely studied by means of the GAs.

In order to minimize the linearization as well as the numerical differentiation error the full gravity field was superseded by the residual field. In order to minimize the deviation of the reference dynamic orbit corresponding with the true one, we studied the idea of the best fitting reference orbit.

We then focused on the global spherical harmonic analysis on sphere. Different methods were presented and their performance compared numerically.

The chapter entitled as *mathematical formulation of LL-SST problem* was devoted to the formulation of the LL-SST problem in the space-wise approach. Because of the direct relation of the gradiometry observations with the gravity field geometry, we developed a mathematical formulation for the field recovery for the gradiometry type. Besides the linear approximation, the cubic approximation of the gradiometry equation was formulated and numerically compared to the linear approximation. Moreover, two different approaches were addressed for obtaining the observation vector without implementation of the numerical differentiation.

Subsequently, we reviewed decomposition of the anisotropic observations into the invariant and non-invariant components. After observation decomposition, two different iterative algorithms were developed for the recovery of the residual field Stokes's coefficients.
Finally, the developed iterative schemes were implemented in chapter 7 and their performance compared from the numerical point of view.
To sum up, the main achievements and the recommendation for the future studies are summarized in the following sections.

8.2 Achievements

This dissertation made the following contributions:

- ✓ In this study, the GAs have been applied successfully for the determination of the shape parameter of the radial basis functions. It can be generalized to every kind of RBF although herein it is just applied for the inverse multiquadric interpolation.

- ✓ The one-leave out approach or the *standard bootstrap method* has been utilized for setting up the object function of the interpolation and approximation optimization problem. In Rippa (1999), the method has been used for the multiquadric interpolation in the local optimization sense. The idea is elegantly generalized to the approximation problem.

- ✓ The object function of the RBF shape parameter optimization problem as well as the polynomial approximation scheme has been generalized such that the objective function includes the interpolation points. From the numerical aspect, including the interpolation point prevents the undesirable oscillations of the high-degree polynomials.

- ✓ The optimization problem of the RBF interpolation as well as the polynomial approximation have been modelled in the global rather than the local sense. Consequently, the global optimum rather than the local solution is achieved.

- ✓ The augmented multiquadric interpolation method can be chosen in an optimal manner by means of the GAs. Both the shape parameter and the additional polynomial components can be optimized using the binary representation of the GAs.

- ✓ Similar to the selection of the optimal augmented multiquadric interpolator, the rational interpolation/approximation function is optimized using the binary representation of the GAs.

- ✓ In Keller (1994), the first iterative approach has been introduced based on the Banach fixed-point theorem. The second approach has alternatively been developed for recovery of the gravity field using the anisotropic functionals. Compared to the first approach, the second one is superior since the non-invariant component can be correctly removed by distinguishing the function on ascending and descending arcs. Furthermore, higher gridding accuracy will be expected since the non-invariant component as the nuisance part has been removed.

- ✓ The intersatellite relative velocity vector is numerically derived from the GRACE satellite positions. Compared to the K–band measurements, the GPS observations are inaccurate. Moreover, the numerical differentiation is a unfavorable numerical process since it amplifies the observation noise. Using the residual gravity observations has been suggested by Rummel (1980). He showed that the residual observation vector can be set up using just the range-acceleration differences. Alternatively, replacing the relative velocity vector with the corresponding vector derived from the reduced dynamic orbit is proposed. It has been numerically shown that the latter approach outperforms the former one.

- ✓ The gradiometry observation equation of the third order has been developed. Compared to the linear approximation, it yields higher accuracy since the truncation error is significantly smaller than that of the linearization. Consequently, a reasonable accuracy can be obtained even by subtracting an ellipsoidal reference field's contribution which is inaccurate in the linear approximation.

8.3 Conclusion and recommendations

- Instead of mapping the observations on the sphere they can be reduced on a torus. Compared to the sphere, observation equations can be expressed rigorously in the torus frame. Furthermore, 2D FFT can be employed since both ω^o and ω_E complete a full period on torus.
- The same approach can be implemented for the recovery of the gravity field using GOCE data. The radial component $V^{(rr)}$ is an invariant functional on the gravity field. Therefore, it can directly be transferred into the geopotential coefficients by mapping the observations on the mean orbital sphere and performing the GSHA. The other observables (G_{22}, G_{33} and the off-diagonal elements) of the GOCE mission can be solved using the iterative approaches.
- In Montesinos et al. (2005), the GAs have been proposed for 3 D inversion of the gravity data. It is recommended to employ a similar approach for the local gravity field modelling and determination of the local temporal variations the field.
- The bootstrap method introduced by Efron (1979) can be used for transformation of the stochastic behavior of the observations into the estimated coefficients in the GSHA process.
- The Earth's topography, without exception, can be seen in the estimated residuals of the gridding methods. Moreover, the same pattern can be seen in the observations decomposition. Therefore, removing the contribution of the topography in space-borne observations generates a much smoother residual field. Consequently, we can expect to achieve higher accuracy both in downward continuation as well as in the decomposition. More details on the removing the effect of the topography potential can be found in (e.g. Wild and Heck, 2004; Heck and Wild, 2005).
- Similar to the time-wise approach, the estimation of the high-degree coefficients requires particular care in the space-wise approach. The problem is bypassed by introducing a *regularization parameter* in the time-wise approach. An investigation should be carried out on the applicability of an analogous scheme in the space-wise method.
- The method of least squares prediction has proven its flexibility and high-performance in the Earth's gravity field determination. The LL-SST problem can also be formulated effectively in the least squares collocation sense. Both the acceleration difference and the gradiometry observation are transferred into the gravitational potential on the boundary.
- The analytical continuation can be successfully employed in the second approach for reducing the invariant component on the specified grid. The continuation is well-defined since the reduction is carried out in the outer space.
- Spherical harmonic analysis in terms of unevenly spaced observations has been implemented for geomagnetic data (e.g. Whaler and Gubbins, 1981; Mochizuki et al., 1997). It is a potential alternative to the GSHA using the gridded data. Performing numerical analysis will clarify the pros and cons of the global harmonic analysis based on the gridded and scattered observations.

Bibliography

A. Albertella, F. Migliaccio, and F. Sansò. Application of the concept of biorthogonal series to a simulation of a gradiometric mission. In F. Sansò, editor, *Geodetic Theory Today*, volume 114 of *IAG Symposia*, pages 350–361. IAG, Springer-Verlag, 1995.

J. Andersson. *Multiobjective Optimization in Engineering Design*. PhD thesis, Linköpings University, Division of Fluid and Mechanical Engineering Systems, Linköping, Sweden, 2001.

K. E. Atkinson. *Elementary Numerical Analysis*. John Wiley and Sons, 1993.

L. Ballani. Partielle Ableitungen und Variationsgleichungen zur Modellierung von Satellitenbahnen und Parameterbestimmung. *Vermessungstechnik*, 36(6):192–194, 1988.

L. Barannyk. Numerical methods, 2005. Lecture notes.

T Bäck. Optimal mutation rates in genetic search. In *Proceedings of the 5th International Conference on Genetic Algorithms*, pages 2–8, San Francisco, CA, USA, 1993. Morgan Kaufmann Publishers Inc. ISBN 1-55860-299-2.

D. Beasley, D. R. Bull, and R. R. Martin. An overview of genetic algorithms: Part 1, fundamentals. *University Computing*, 15(4):170–181, 1993.

M. W. Berry. *A variable-step double-integration multi-step integrator*. PhD thesis, Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, Virginia, April 2004.

G. Blaha. Refinement of the atellite-to-satellite line-of-sight acceleration model in a residual gravity field. *Manuscripta Geodaetica*, 17:321–333, 1992.

L. Booker. Improving search in genetic algorithms. In *Genetic Algorithms and Simulated Annealing*, pages 61–73. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987.

R. L. Burden and J. D. Faires. *Numerical Analysis*. Cole Publ. Co., sixth edition, 1997. ISBN 0-534-95532-0.

F. Busetti. Genetic algorithms overview. URL `citeseer.ist.psu.edu/busetti01genetic.html`.

J. C. Butcher. Coefficients for the study of runge-kutta integration processes. *J. Austral. Math. Soc*, 3:185–201, 1963.

J. C. Butcher. The non-existence of ten stage eighth order explicit runge-kutta methods. *BIT Numerical Mathematics*, 25(3):521–540, 1985.

R. E. Carlson and T. A. Foley. The parameter r^2 in multiquadric interpolation. *Compute. Math. Applic*, 21(9):29–42, 1991.

J. C. Carr, W. R. Fright, and R.K. Beatson. Surface interpolation with radial basis functions for medical imaging. *IEEE Transactions on Medical Imaging*, 16(1):96–107, 1997.

P. Charbonneau. An introduction to genetic algorithms for numerical optimization. Technical Report 450+IA, NCAR, Boulder, Colorado, March 2002.

P. F. Davis and P. Rabinowitz. *Methods of Numerical Integration*. Academic Press, San Diego, second edition, 1984.

K. Deb. Multi-objective genetic algorithms: Problem difficulties and construction of test problems. *Evolutionary Computation*, 7(3):205–230, 1999.

P. Ditmar, V. Kuznetsov, A. A. van Eck van der Sluijs, E. Schrama, and R. Klees. Deos-champ-01c-70: a model of the earth's gravity field computed from accelerations of the champ satellite. *Journal of Geodesy*, 79(10):??, 2005.

M. Dorigo and L. M. Gambardella. Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, 1(1):53–66, 1997.

J. Duchon. Interpolation des fonctions de deux variables suivant le principl de la flexion des plaques minces. *R.A.I.R.O. Analysi numer.*, 10:5–12, 1976.

O. V. Dudnik and P. I. Bidyuk. Application of radial basis functions in neural networks for prognosis of economic parameters. *Journal of Automation and Information Sciences*, 35(4):39–45, 2003.

B. Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, Jan 1979.

A. E. Eiben and J. E. Smith. *Introduction to Evolutionary Computing*. Natural Computing Series. Springer, 2003. ISBN 3-540-40184-9.

H. W. Ellsaesser. Expansion of hemispheric meteorological data in antisymmetric surface spherical harmonic (laplace) series. *J. Appl. Metheorology*, 5:263–276, 1966.

N. V. Emeljanov and A. A. Kanter. A method to compute inclination functions and their derivatives. *Manuscripta Geodaetica*, 14:77–83, 1989.

ESA. Study of a satellite-to-satellite tracking gravity mission. Technical report, ESTEC/Contract No. 6557/85/NLP, 1987.

L. J. Eshelman, R. Caruana, and J. D. Schaffer. Biases in the crossover landscape. In *Proceedings of the 3rd International Conference on Genetic Algorithms table of contents*, pages 10–19, San Francisco, CA, USA, 1989. Morgan Kaufmann Publishers Inc.

L. J. Eshelman and J. D. Schaffer. Real-coded genetic algorithms and interval schemata. In L. Darrell Whitley, editor, *Foundation of Genetic Algorithms*, volume 2, pages 187–202, San Mateo, 1993. Morgan Kaufmann.

A. Fairley. Comparison of methods of choosing the crossover point in the genetic crossover operation. Technical report, Department of Computer Science, University of Liverpool, 1991.

E. Fehlberg. Classical seventh-, sixth-, and fifth-order runge-kutta-nystrom formulas with stepsize control for general second-order differential equations. Technical Report NASA-TR-R-432, NASA Center for AeroSpace Information (CASI), 1974.

Y. Feng. An alternative orbit integration algorithm for gps-based precise leo autonomous navigation. *GPS Solution*, 5(2):1–11, 2001.

D. N. Fogel and L. R. Tinney. Image registration using multiquadric functions, the finite element method, bivariate mapping polynomials and thin plate spline. Technical Report 96-1, National Center for Geographic Information and Analysis, 1996.

T. A. Foley. Interpolation and approximation of 3-d and 4-d scattered data. *Compute. Math. Applic*, 13:711–740, 1987.

C. M. Fonseca and P. J. Flemming. An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary computation*, 3(1):1–16, 1995.

R. Franke. A critical comparison of some methods for interpolation of scattered data. Technical Report NPS-53-79-003, Naval Postgraduate School, Monterary, California, 1979.

W. Freeden. *Multiscale Modelling of Spaceborne Geodata*. B. G. Teubner, Leipzig, 1999.

W. Freeden, O. Glockner, and M. Thalhammer. Multiscale gravitational field recovery from gps-satellite-to-satellite tracking. *Studia Geophysica et Geodaetica*, 43:229–264, 1999.

R. V. Garcia. Local geoid determination from grace mission. Technical Report 460, Ohio State University, Columbus, 2002.

R. V. Garcia. Efficient global gravity determination from satellite-to-satellite tracking (sst). Technical Report 467, Ohio State University, Columbus, 2003.

R. D. Gerigorieff. *Numerik gewöhnlicher Differentialgleichungen 2*. Teubner Verlag, Stuttgart, 1977.

D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Pub. Co., 1989.

W. H. Goodyear. Completely general closed-form solution for coordinates and partial derivatives of the two-body problem the. *Astronomical Journal*, 70(3):189–192, 1965.

R. T. Gregory. A method for deriving numerical differentiation formulas. *The American MAthematical Monthly*, 64(2):79–82, 1957.

M. Götzelmann. Short-arc Bahnanpassung bei niedrig fliegenden Satelliten. Master's thesis, Institute of Geodesy, Faculty of Aerospace Engineering and Geodesy, University of Stuttgart, 2003.

D. Gubbins. *Time Series Analysis and Inverse Theory for Geophysicists*. Cambridge University Press, 2004. ISBN 0521819652.

D. P. Hajela. Improved procedures for the recovery of 5° mean gravity anomalies from ats-6/geos-3 staellite-to-satellite range-rate observations. Technical Report 276, Ohio State University, Columbus, 1978.

R. L. Hardy. Multiquadric equations of topography and other irregular surfaces. *J. Geophy. Res.*, 76 (1905-1915), 1971.

R. L. Hardy. Analytical topographic surfaces by spatial intersection. *Photogramm. Engng Rem. Sens.*, 38:452–458, 1972.

R. L. Hardy. Least squares prediction. *Photogrammetric Engineering Remote Sensing*, 43:475–492, 1977.

R. L. Hardy. The biharmonic potential and its applications. *Technical Papers of the American Congress on Surveying and Mapping*, 1983.

R. L. Hardy. Theory and applications of the multiquadric-biharmonic method. *Computers and Mathematical Applications*, 19:163–208, 1990.

R. L. Hardy and W. M. Göpfert. Least squares prediction of gravity anomalies, geoidal undulations, and deflections of the vertical with multiquadric harmonic functions. *Geophys. Res. Lett*, 10:423–426, 1975.

R. L. Hardy and S. A. Nelson. Approximation of Newtonian potentials. In *Approximation Theory V*, Proc. 5th Int. Symp., pages 375–378, College Station, Texas, 1986a.

R. L. Hardy and S. A. Nelson. A multiquadric biharmonic representation and approximation of disturbing potential. *Geophys. Res. Lett.*, 13:18–21, 1986b.

R. L. Haupt and S. E. Haupt. *Practical genetic algorithms*. John Wiley & Sons, Inc., New York, NY, USA, second edition, 2004. ISBN 047-45565-2.

D. Heß and W. Keller. Gradiometrie mit GRACE, Teil I. *ZfV*, 124:137–144, 205–211, 1999.

B. Heck and F. Wild. Topographic reductions in satellite gravity gradiometry based on a generalized condensation model. In F. Sansó, editor, *A Window on the Future of Geodesy*, volume 128 of *IAG Symposia*, pages 294–299, 2005.

K. Hesse and M. Gutting. Smoothing splines in multiscale geopotential determination from satellite data. *Berichte der Arbeitsgruppe Technomathematik (AGTM-Reports)*, Report No. 255:827–846, 2003.

F. J. Hickernell and Y. C. Hon. Radial basis function approximation of the surface wind field from scattered data. *Internat. J. Appl. Sci. Comput.*, 4(3):221–247, 1998.

B. Hofmann-Wellenhof and H. Moritz. *Physical Geodesy*. Springer, 2005.

J. H. Holland. *Adaptation in natural and artificial systems*. Ann Arbor: university of Michigan Press, 1975.

A. Hondroudakis, J. Malard, and G. V. Wilson. An introduction to genetic algorithms using rpl2, May 1996. Student notes.

H. Jeffreys and B.s. Jeffreys. *Methods of Mathematical Physics*. Cambridge University Press, Cambridge, England, third edition, 1988.

C. Jekeli. The determination of gravitational potential differences from satellite-to-satellite tracking. *Celest. Mech. Dynam. Astron.*, 75:85–101, 1999.

K. A. De Jong. *Analysis of the behavior of a class of genetic adaptive systems*. PhD thesis, University of Michigan, 1975.

B. Kampes. Analysis of evenly distributed data on the sphere: a comparison of least squares, quadrature and 2d-fourier methods. Master's thesis, Delft university of technology, 1998.

E. J. Kansa. Multiquadrics- a scattered data aproximation scheme with applications to computational fluid dynamics: II. solutions to parabolic, hyperbolic, and elliptic partial differential equations. *Comput. Math. Appl.*, 19:147–161, 1990a.

E. J. Kansa. Multiquadrics-a scattered data aproximation scheme with applications to computational fluid dynamic: I. surface approximations and partial derivate estimates. *Comput. Math. Appl.*, 19: 127–145, 1990b.

E. J. Kansa and R. E. Carlson. Improved accuracy of multiquadric interpolation using variable shape parameter. *Computer Math. Applic.*, 24(12):99–120, 1992.

M. Karrer. Schwerefeldbestimmung aus Satellitengradiometrie (Torus approach). Master's thesis, Technische Universität München, 2000.

W. M. Kaula. *Theory of Satellite Geodesy*. BLAISDELL Publishing Company, 1966.

W. M. Kaula. Inference of variations in the gravity field from satellite-to-satellite range rate. *Journal of Geophysical Research*, 88:8345–8349, 1983.

W. Keller. Solving the step-observation equation using banach's fixed-point principle. In H. sünkel and I. Marson, editors, *Gravity and Geoid*, volume 113 of *IAG Symposia*, pages 117–130. IAG, Springer-Verlag, 1994.

W. Keller. Observation techniques in satellite geodesy, 2005. Lecture notes.

W. Keller and M. A. Sharifi. Satellite gradiometry using a satellite pair. *Journal of Geodesy*, 78: 544–557, 2005.

S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220, 4598(4598):671–680, 13 May 1983. URL `citeseer.ist.psu.edu/kirkpatrick83optimization.html`.

R. Koop. Global gravity field modelling using satellite gravity gradiometry. New Series 38, Netherlands Geodetic Commission, 1993.

V. I. Krylov. *Approximate calculation of integrals.* Macmillan, 1962.

J. Lense. Kugelfunktionen. pages 230–242, 1954. Geest and Portig.

F. J. Lerch and C. A. Wagner. Goddard earth models for oceanographic applications. *Marine Geodesy*, 5(2):145–187, 1981.

Marc Levoy. Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8 (3):29–37, 1988.

L. Ling. *Radial Basis Functions in Scientific Computating.* PhD thesis, Simon Fraser University, 2003.

J. B. Lundberg. Multistep integration formulas for the numerical integration of the satellite problem. Technical Report NASA-CR-164193, NASA Center for AeroSpace Information, 1981.

J. L. Maury and G. P. Segal. Cowell type numerical integration as applied to satellite orbit computation. Technical Report N6926703, NASA, 1969.

H. Mühlenbein. How genetic algorithms really work: Mutation and hillclimbing. In *second international conference on Parallel Problem Solving From Nature*, 1992.

C. A. Micchelli. Interpolation of scattered data: distance matrices and conditionally positive definite functions. *Constr. Approx.*, 2:11–22, 1986.

Z. Michalewicz. *Genetic algorithms + data structures = evoultion programs.* Springer-Verlag, 3rd rev. and extended ed. edition, 1996.

M. Mitchell. *An Introduction to Genetic Algorithms.* The MIT press, 1998. ISBN 0-262-63185-7.

E. Mochizuki, Y. Yokoyama, I. Shimizu, and Y. Hamano. Spherical harmonic analysis in terms of unevenly distributed observation points and its applications to geomagnetic data. *Journal of Geomagnetism and Geoelectricity*, 49(8):1013–1033, 1997.

D. D. Moerder and B. N. Pamadi. Constrained minimization of smooth functions using a genetic algorithm, November 1994. NASA Technical paper 3329.

O. Montenbruck and E. Gill. *Satellite Orbits Models, Methods and Applications.* Springer, 3rd edition, 2005. ISBN 3-540-67280-X.

F. G. Montesinos, J. Arnoso, and R Vieira. Using a genetic algorithm for 3-d inversion of gravity data in Fuerteventura (canary islands). *International Journal of Earth Sciences*, 94(2):301–316, 2005.

H. Moritz. *Approximation Methods in Geodesy*, chapter Introduction to interpolation and approximation, pages 1–46. Wichmann, 1978.

P. M. Muller and W. L. Sjogren. Mascons: Lunar mass concentrations. *Science*, 161:680–684, 1968.

F. Neumann. Über eine neue Eigenschaft der Laplaceschen $y^{(n)}$ und ihre Anwendung zur analytischen Darstellung derjenigen Phänomene, welche Functionen der geographischen Länge und Breite sind. *Schumachers Ann. Nachr.*, 15:313–325, 1838. (Reprinted in Math. Ann., 14, p. 567).

F. Neumann. Vorlesungen über die Theorie des Potentials und der Kugelfunctionen. pages 135–154, 1887. Teubner.

G. M. Nielson, T. A. Foley, B. Hamann, and D. Lane. Visualizing and modeling scattered multivariate data. *IEEE Computer Graphics and Applications*, 11(3):47–55, 1991.

K. E. Parsopoulos and M. N. Vrahatis. Recent approaches to global optimization problems through particle swarm optimization. *Natural Computing*, 1(2-3):235–306, June 2002.

W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C.* Cambridge University Press, 1988.

W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in FORTRAN.* Cambridge University Press, 1992.

N. J. Radcliffe. *Genetic Neural Networks on MIMD computers.* PhD thesis, University of Edinburg, Edinburg, U. K., 1990.

C. R. Reeves. *Modern Heuristic Techniques for Combinatorial Problems*, chapter Genetic Algorithms, pages 151–196. John Wiley & Sons, Inc, New York, NY, USA, 1993.

T. Reubelt, G. Austen, and EW. Grafarend. Harmonic analysis of the earth's gravitational field by means of semi-continuous ephemeris of a low earth orbiting gps-tracked satellite. case study: Champ. *Journal of Geodesy*, 77:257–278, 2003.

S. Rippa. An algorithm for selecting a good value for the parameter c in radial basis function interpolation. *Advances in Computational Mathematics*, 11:193–210, 1999.

T. Ritchey. Analysis and synthesis on scientific method based on a study by bernhard riemann. *Systems Research*, 8(4):21–41, 1996. revised.

C. Rizos. An efficient computer technique for the evaluation of geopotential from spherical harmonic models. *Aust. J. Geod. PHoto Surv.*, (31):161–169, 1979.

G. W. Rosborough. Satellite orbit perturbations due to the geopotential. Technical Report CSR-86-1, Center for Space Research, The university of Texas at Austin, 1986.

R. Rummel. Geoid height, geoid height differences, and mean gravity anomalies from low-low satellite-to-satellite tracking- an error analysis. Technical Report 306, Ohio State University, Columbus, 1980.

R. Rummel. How to climb the gravity wall. *Space Science Reviews*, 108:1–14, 2003.

R. Rummel, G. Balmino, J. Johannessen, P. Visser, and P. Woodworth. Dedicated gravity field missions - principles and aims. *J. Geodynamics*, 33:3–20, 2002.

R. Rummel, CH. Reigber, and K. H. Ilk. The use of satellite-to-satellite tracking for gravity parameter recovery. In *Proc. of the European workshop on Space Oceanography, Navigation and Geodynamics*, number SP-137, pages 153–161. ESA, 1978.

R. Rummel, M. van Gelderen, R. Koop, E. Schrama, F. Sanso, M. Brovelli, F. Migliaccio, and F. Sacerdote. Spherical harmonic analysis of satellite gradiometry. New Series 39, Netherlands Geodetic Commission, 1993.

D. Ruprecht and H. Müller. Image warping with scattered data interpolation. *IEEE Computer Graphics and Applications*, 15(2):37–43, 1995.

J. Jin S. Zhang. *Computation of Special Functions.* Wiley, 1996. ISBN 0-471-11963-6.

J. D. Schaffer, R. A. Caruana, L. J. Eshelman, and R. Das. A study of control parameters affecting online performance of genetic algorithms for function optimization. In J. D. Schaffer, editor, *Proceedings of the Third International Conference on Genetic Algorithms*, pages 51–60, San Mateo, CA, 1989. Morgan Kaufmann.

I. P. Schagen. Interpolation in two dimensions-a new technique. *J. Inst. Math. Appl.*, 23:53–59, 1979.

I. P. Schagen. The use of stochastic processes in interpolation and approximation. *Intern. J. Comput. Math. B*, 8:63–76, 1980.

C. Schäfer. *Space Gravity Spectroscopy. The sensitivity analysis of GPS-tracked satellite missions (case study CHAMP).* PhD thesis, University of Stuttgart, 2000.

E. Schrama. A study of a satellite-to-satellite tracking configuration by application of linear perturbation theory. Internal note 86.3, TU Delft, 1986.

E. J. O. Schrama. The role of orbit errors in processing of satellite altimeter data. New Series 33, Netherlands Geodetic Commission, 1989.

M. Schreiner. *Tensor Spherical Harmonics and Their application in Satellite Gradiometry.* PhD thesis, University of Kaiserslautern, Geomathematics Group, 1994.

H. P. Schwefel. *Evolution and Optimum Seeking: The Sixth Generation.* John Wiley & Sons, Inc., New York, NY, USA, 1993. ISBN 0471571482.

G. Seeber. *Satellite Geodesy.* Walter de Gruyter GmbH, second edition, 2003. ISBN 3110175495.

M. A. Sharifi and W. Keller. Grace gradiometer. In C. Jekeli, L. Bastos, and J. Fernandes, editors, *Gravity Geoid and Space Mission*, volume 129, pages 42–47. IAG, Springer, 2005.

M. V. Shul'min and Y. Y. Mitel'man. The multiquadric method of approximating a topographic surface. *Geodesy Mapp. Photogramm*, 16:13–17, 1974. translated from Russian for AGU, ACSM and ASP (1977).

D. J. Sirag and P. T. Weisser. Towards a unified thermodynamic genetic operator. In J. J. Grefenstette, editor, *Proceedings of the Second International Conference on Genetic Algorithms*, pages 116–122, 1987.

N. Sneeuw. Global spherical harmonic analysis by least-squares and numerical quadrature methods in historical perspective. *Geophys. J. Int.*, 1994.

N. Sneeuw. Dynamical satellite geodesy on the torus: Block-diagonality from a semi-analytical approach. In M.G. Sideris, editor, *Gravity, Geoid, and Geodynamics*, IAG Symposia. IAG, 2000.

N. Sneeuw. Space-wise, time-wise, torus and rosborough representations in gravity field modelling. volume 108 of *Space Sciences Series of ISSI*, pages 37–46. International Space Science Institute, Kluwer Academic Publisher, 2003.

N. J. Sneeuw. Inclination functions: Group theoritical background and recursive algorithms. Internal 91.2, Mathematical and Physical Geodesy, Faculty of Geodetic Engineering, TU Delft, 1991a.

N. J. Sneeuw. Representation coefficients and their use in satellite geodesy. *Manuscripta Geodaetica*, 17:117–123, 1992.

G. Syswerdar. Uniform crossover in genetic algorithms. In *Proceedings of the 3rd International Conference on Genetic Algorithms table of contents*, pages 2–9, San Francisco, CA, USA, 1989. Morgan Kaufmann Publishers Inc.

C. V. Tao and Y. Hu. A comprehensive study of the rational function model for photogrammetric processing. *Photogrammetric Engineering and Remote Sensing*, 67(12):1347–1357, 2001.

G. Telfar. Generally applicable heuristics for global optimisation: An investigation of algorithm performance for the euclidean traveling salesman problem. Master's thesis, Victoria University of Wellington, October 1994.

C. J. Trahan and R. E. Wyatt. Radial basis function interpolation in the quantum trajectory method: optimization of the multi-quadric shape parameter. *Journal of Computational Physics*, 185(1):27–49, 2003.

D. A. Vallado. *Fundamentals of Astrodynamics and Applications.* McGraw-Hill, New York, 1997.

M. Vermeer. Observable quantities in satellite gradiometry. *Bulletin Geodesique*, 64(4):347–361, 1990.

J. Vincent. Introduction to numerical optimisation using genetic algorithms. Technical report, School of Design, Engineering and Computing, Bournemouth University, 2003.

C. A. Wagner. Direct determination of gravitational harmonics from low-low gravsat data. *Journal of Geophysical Research*, 88:10309–10321, 1983.

E. W. Weisstein. Eigenfunction. URL `http://mathworld.wolfram.com/Eigenfunction.html`. From MathWorld–A Wolfram Web Resource.

D. Wells. Functional analysis applied to least squares in geomatics. lecture notes, 1994.

K. A. Whaler and D. Gubbins. Spherical harmonic analysis of the geomagnetic field: An example of a linear inverse problem. *Geophys. J. Roy. Astron. Soc.*, 65:645–693, 1981.

L. K. White. Simulation and analysis of a geopotential research mission. Technical Report CSR-87-2, Center for Space Research, 1987.

D. Whitely. A genetic algorithm tutorial. Technical Report CS-93-103, Colorado State University, 1993.

F. Wild and B. Heck. Effects of topographic and isostatic masses in satellite gravity gradiometry. In *The Geoid and Oceanography*, Proc. Second International GOCE User Workshop GOCE, Frascati/Italy, March 2004. ESA-ESRIN.

G. Wolfberg. *Digital image wrapping.* IEEE Computer Society Press, Los Alamitos, California, 1990.

M. Wolff. Direct measurements of the earth's gravitational potential using a satellite pair. *J. Geophys Res.*, 74:5295–5300, 1969.

L. Wuytack. On some aspects of the rational interpolation problem. *SIAM Journal of Numerical Analysis*, 11(1):52–60, 1974.

D. Zhong. Robust estimation and optimal selection of polynomial parameters for the interpolation of gps geoid heights. *Journal of Geodesy*, 71:552–561, 1997.

Appendix A

Inclination functions

There are several methods for normalized inclination function $\bar{F}^k_{nm}$ computation (Sneeuw, 1991a).

Direct evaluation The function is computed by use of rigorous formulas. Nevertheless, it leads to instability due to numerical overflow and /or underflow. Furthermore, the number of repeated computations is a waste of computer time.

Fourier method It is quite powerful however, computation would be rather slow because of executing the whole procedure for evaluation of an individual pair of (n, m).

Recursive approach Different recurrence procedures have been proposed for the evaluation of function. Among them, Emeljanov and Kanter's method (Emeljanov and Kanter, 1989)is preferred because of its stability up to high-degree and higher computational efficiency.

The Emeljanov and Kanter's Method
Since computation of the $\bar{F}^k_{nm}$ has been clearly described in (Sneeuw, 1991a), for ease of convenience we explain it very briefly.

Introduction of intermediary functions: The recursion runs over the auxiliary functions E^k_{nm} [1] rather than the functions $\bar{F}^k_{nm}$ themselves. Inclination functions and the intermediary ones fulfill

$$\bar{F}^k_{nm} \;=\; \left[\sin\frac{I}{2}\right]^{|k-m|} E^k_{nm}. \tag{A-1}$$

Initialization: First we compute the value of the auxiliary function for $n = N$, maximum degree and order up to which the inclination functions are required,

$$E^{-N}_{N0} \;=\; (-1)^{C(\frac{n+1}{2})}\,\sqrt{2N+1}\,R_N, \tag{A-2}$$

with C the Ceil operator which round the operand to the nearest integers greater than or equal to the operand. R_N is computed using the recurrence relation

$$R_n \;=\; \frac{2n-1}{n}\,\cos\frac{I}{2}\,R_{n-1}, \tag{A-3}$$

with starting value $R_0 = 1$.

Recursion over m:

$$E^{-N}_{Nm} \;=\; \frac{(-1)^{N-m+1}}{\cos\frac{I}{2}}\sqrt{\frac{(2-\delta_{m,0})(N-m+1)}{(2-\delta_{m-1,0})(N+m)}}\;E^{-N}_{N,m-1} \tag{A-4}$$

[1] In Sneeuw (1991a) A^k_{nm} stands for the intermediary function while we have already reserved it for the lumped coefficients.

Recursion over k: Three different recurrence relationships are used for different value of the index k,

$$E^k_{km} = (-1)^{k-m+1} \cos\frac{I}{2} \sqrt{\frac{4k^2-1}{k^2-m^2}}\, E^{k-1}_{k-1,m}; \qquad k > m \tag{A-5}$$

$$E^k_{mm} = \cos\frac{I}{2}\, E^{k-1}_{mm} \begin{pmatrix} \frac{2}{m+k} \\ \frac{m-k+1}{2} \end{pmatrix}^{m-k \text{ even}}_{m-k \text{ odd}}; \qquad -m < k \leq m \tag{A-6}$$

$$E^k_{-km} = \frac{(-1)^{k+m-1}}{\cos\frac{I}{2}} \sqrt{\frac{(k-1)^2-m^2}{(2k-1)(2k-3)}}\, E^{k-1}_{-k+1,m}; \qquad k \leq -m \tag{A-7}$$

Recursion over n:

$$E^k_{nm} = (-1)^{n-m}\, a_{nk} \sqrt{\frac{4n^2-1}{n^2-m^2}}\, \frac{n(n-1)\cos I - mk}{2(n-1)}\, E^k_{n-1,m} +$$

$$b_{nk} \sqrt{\frac{(2n+1)[(n-1)^2-m^2]}{(2n-3)(n^2-m^2)}}\, \frac{n[(n-1)^2-k^2]}{n-1}\, E^k_{n-2,m} \tag{A-8}$$

where

$$a_{nk} = \begin{bmatrix} \frac{4}{n^2-k^2} \\ 1 \end{bmatrix}^{n-k \text{ even}}_{n-k \text{ odd}} \qquad b_{nk} = \begin{bmatrix} \frac{1}{n^2-k^2} \\ \frac{1}{(n-1)^2-k^2} \end{bmatrix}^{n-k \text{ even}}_{n-k \text{ odd}}$$

For initialization of Eq. (A-8), we use Eqs. (A-5), (A-6) and (A-7) on the one hand and set $E^k_{k-1,m} = E^k_{m-1,m} = E^k_{-k-1,m} = 0$ on the other hand.

Furthermore, Eq. (A-8) is singular for $(n, m, k) = (1, 0, 0)$. To avoid this singularity Sneeuw (1991a) recommended to compute the corresponding function directly by use of rigorous formula (Kaula, 1966)

$$\bar{F}^0_{10}(I) = E^0_{10}(I) = -\frac{\sqrt{3}}{2}\, \cos I. \tag{A-9}$$

For $n \geq 2$ the recurrence relation Eq. (A-8) can be used without any problem.

These recurrence relations can also be used for computation of the first and second partial derivatives of the Inclination functions. However, herein we will just utilize the along-track component which contains just the function itself. Therefore, the formulas are not repeated here. The interested reader can refer to Sneeuw 1991a and the references therein.

Appendix B

A single observation elimination

The linear system of equations based on n and $n-1$ observations are

$$\mathbf{A}_n \boldsymbol{\xi}_n = \mathbf{l}_n \tag{B-1}$$

and

$$\mathbf{A}_{n-1} \boldsymbol{\xi}_{n-1} = \mathbf{l}_{n-1} \tag{B-2}$$

Least squares solution of Eqs. B-1 and B-2 can be written as

$$\mathbf{A}_n^T \mathbf{A}_n \hat{\boldsymbol{\xi}}_n = \mathbf{A}_n^T \mathbf{l}_n \rightarrow \mathbf{N}_n \hat{\boldsymbol{\xi}}_n = \mathbf{u}_n \rightarrow \hat{\boldsymbol{\xi}}_n = \mathbf{N}_n^{-1} \mathbf{u}_n \tag{B-3}$$

and

$$\mathbf{A}_{n-1}^T \mathbf{A}_{n-1} \hat{\boldsymbol{\xi}}_{n-1} = \mathbf{A}_{n-1}^T \mathbf{l}_{n-1} \rightarrow \mathbf{N}_{n-1} \hat{\boldsymbol{\xi}}_{n-1} = \mathbf{u}_{n-1} \tag{B-4}$$

As seen in Eq. (B-3), the solution is a function of $\mathbf{N}_n^{-1}$ and $\mathbf{u}_n$. Similarly, the same equation can be written for the later case. Nevertheless, we would like to express $\hat{\boldsymbol{\xi}}_{n-1}$ as a function of $\hat{\boldsymbol{\xi}}_n$, $\mathbf{N}_n^{-1}$ and the excluded observation's respective row in the design matrix $\mathbf{a}_n$, and its value l_n. In order to derive the sought after expression, let us decompose the design matrix and the observation vector as follows:

$$\mathbf{A}_n = \begin{bmatrix} \mathbf{A}_{n-1} \\ \mathbf{a}_n \end{bmatrix} \rightarrow \mathbf{N}_{n-1} = \mathbf{N}_n - \mathbf{a}_n^T \mathbf{a}_n = \mathbf{N}_n - \mathbf{N}_n^{[n]} \tag{B-5}$$

similarly

$$\mathbf{l}_n = \begin{bmatrix} \mathbf{l}_{n-1} \\ l_n \end{bmatrix} \rightarrow \mathbf{u}_{n-1} = \mathbf{u}_n - \mathbf{a}_n^T l_n = \mathbf{u}_n - \mathbf{u}_n^{[n]} \tag{B-6}$$

$\mathbf{N}_n^{[n]}$ and $\mathbf{u}_n^{[n]}$ are the l_n observation's contribution to the design matrix $\mathbf{N}_n$ and the known vector $\mathbf{u}_n$ of the normal equations, respectively. Substituting Eqs. (B-5) and (B-6) into Eq. (B-4) leads to

$$\left(\mathbf{N}_n - \mathbf{N}_n^{[n]}\right) \hat{\boldsymbol{\xi}}_{n-1} = \mathbf{u}_n - \mathbf{u}_n^{[n]} \tag{B-7}$$

or simply

$$\hat{\boldsymbol{\xi}}_{n-1} = \left(\mathbf{N}_n - \mathbf{N}_n^{[n]}\right)^{-1} \left(\mathbf{u}_n - \mathbf{u}_n^{[n]}\right) \tag{B-8}$$

Although Eq. (B-8) expresses the solution vector $\mathbf{x}_{n-1}$ in term of $\hat{\boldsymbol{\xi}}_n$, $\mathbf{N}_n$, $\mathbf{a}_n$ and l_n, it is not the ideal solution that we are looking for. Because, $(\mathbf{N}_n - \mathbf{N}_n^{[n]})$ should be inverted n times which seems to be rather a difficult task for the large number of n. Alternatively, one may utilize Taylor expansion to avoid n times inversion.

$$\hat{\boldsymbol{\xi}}_{n-1} \doteq \left(\mathbf{N}_n^{-1} + \mathbf{N}_n^{-1} \mathbf{N}_n^{[n]} \mathbf{N}_n^{-1}\right) \left(\mathbf{u}_n - \mathbf{u}_n^{[n]}\right) = \left(\mathbf{I}_{n\times n} + \mathbf{N}_n^{-1} \mathbf{N}_n^{[n]}\right) \left(\hat{\boldsymbol{\xi}}_n - \mathbf{N}_n^{-1} \mathbf{u}_n^{[n]}\right) \tag{B-9}$$

$\mathbf{I}_{n\times n}$ is an $n \times n$ identity matrix.
Eq. (B-9) seems to be the ideal sought after equation. However, numerical analysis shows that the contributions of the high-order terms of Taylor expansion to the unknown vector are significant. In

other words, linear expansion of the normal matrix inversion (Eq. B-8) results in inaccurate solution. Therefore, the higher order terms of the expansion should be included in order to achieve a reasonably accurate solution. One should investigate up to which degree of the expansion is required.
As an alternative, we propose to estimate the solution differences instead of the solution vector itself. Subtraction of Eq. (B-3) from Eq. (B-7) results in

$$\mathbf{N}_n \left(\hat{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_{n-1}\right) + \mathbf{N}_n^{[n]} \hat{\boldsymbol{\xi}}_{n-1} = \mathbf{u}_n^{[n]} \tag{B-10}$$

or

$$\mathbf{N}_n \left(\hat{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_{n-1}\right) - \mathbf{N}_n^{[n]} \left(\hat{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_{n-1}\right) + \mathbf{N}_n^{[n]} \hat{\boldsymbol{\xi}}_{n-1} = \mathbf{u}_n^{[n]} \tag{B-11}$$

By substituting $\hat{\Delta\boldsymbol{\xi}} = \left(\hat{\boldsymbol{\xi}}_n - \hat{\boldsymbol{\xi}}_{n-1}\right)$ in Eq. (B-11), the following equation holds

$$\left(\mathbf{N}_n - \mathbf{N}_n^{[n]}\right) \hat{\Delta\boldsymbol{\xi}} = \mathbf{u}_n^{[n]} - \mathbf{N}_n^{[n]} \hat{\boldsymbol{\xi}}_n. \tag{B-12}$$

Finally,

$$\hat{\Delta\boldsymbol{\xi}} = \left(\mathbf{N}_n - \mathbf{N}_n^{[n]}\right)^{-1} \left(\mathbf{u}_n^{[n]} - \mathbf{N}_n^{[n]} \hat{\boldsymbol{\xi}}_n\right). \tag{B-13}$$

Using Taylor expansion of the normal matrix inversion, Eq. (B-13) is recast into

$$\hat{\Delta\boldsymbol{\xi}} \doteq \left(\mathbf{I}_{n\times n} + \mathbf{N}_n^{-1} \mathbf{N}_n^{[n]}\right) \mathbf{N}_n^{-1} \left(\mathbf{u}_n^{[n]} - \mathbf{N}_n^{[n]} \hat{\boldsymbol{\xi}}_n\right). \tag{B-14}$$

One can simply derive the corresponding expression for $\hat{\mathbf{x}}_{n-1}$ using the following relation.

$$\hat{\boldsymbol{\xi}}_{n-1} = \hat{\boldsymbol{\xi}}_n - \hat{\Delta\boldsymbol{\xi}} = \left(\mathbf{I}_{n\times n} + \mathbf{N}_n^{-1} \mathbf{N}_n^{[n]}\right) \left(\hat{\boldsymbol{\xi}}_n - \mathbf{N}_n^{-1} \mathbf{u}_n^{[n]}\right) + \mathbf{N}_n^{-1} \mathbf{N}_n^{[n]} \mathbf{N}_n^{-1} \mathbf{N}_n^{[n]} \hat{\boldsymbol{\xi}}_n \tag{B-15}$$

Compared to Eq. (B-9), one more term has appeared on the right-hand side of Eq. (B-15). It hence renders $\hat{\boldsymbol{\xi}}_{n-1}$ estimation with higher accuracy possible.