

Institut für Astronomische und Physikalische Geodäsie

**A Semi-Analytical Approach to Gravity Field Analysis  
from Satellite Observations**

Nicolaas Sneeuw

Vollständiger Abdruck der von der Fakultät für Bauingenieur- und Vermessungswesen  
der Technischen Universität München zur Erlangung des akademischen Grades eines  
Doktor - Ingenieurs (Dr.-Ing.)  
genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. M. Schilcher

Prüfer der Dissertation:

1. Univ.-Prof. Dr.-Ing. R. Rummel
2. o.Univ.-Prof. Dr. techn. H. Sünkel,  
Technische Universität Graz / Österreich

Die Dissertation wurde am 20. April 2000 bei der Technischen Universität München  
eingereicht und durch die Fakultät für Bauingenieur- und Vermessungswesen  
am 26. Juni 2000 angenommen.



---

## Abstract

THIS work presents a semi-analytical approach to gravity field determination from space-borne observations. Key element is the lumped coefficient formulation, linking gravity field functionals to the unknown gravity field. The observables are lumped coefficients, which are basically Fourier coefficients of the time-series of observations. The unknowns are the spherical harmonic spectral coefficients. The relationship between these different types of spectra is a linear mapping, represented by spectral transfer coefficients. In the semi-analytical approach, these transfer coefficients are derived analytically.

A set of transfer coefficients of several functionals, relevant in dynamical satellite geodesy is denoted as pocket guide. Basic observation techniques of the gravity field functionals, incorporated in this scheme, are:

- high-low satellite-to-satellite tracking (high-low SST). This observable is realized by space-borne GPS tracking on board a low-flying orbiter. It results in three-dimensional accurate and continuous orbit determination.
- low-low satellite-to-satellite tracking (low-low SST). This observable is realized through measurement of the range or range-rate between two low-flying co-orbiting satellites.
- satellite gravity gradiometry (SGG). The spatial derivatives of the components of the gravity vector, i.e. the tensor of second spatial derivatives of the gravitational potential, can be measured by differential accelerometry over short baselines.

The lumped coefficient formulation, together with a set of transfer coefficients, constitutes the observation model. In connection with a stochastic model, it allows for gravity field recovery from satellite observations. In this work, the observation and stochastic model are mainly employed for pre-mission error assessment of any type of gravity field mission. The type of observable, error power spectral density, orbital parameters, mission duration, and so on, are parameters that can be tuned at will in this procedure.

One of the advantages of the semi-analytical approach is the fact that the normal equations, required to infer the unknowns, become a block-diagonal system. In view of the enormous amount of data and of unknowns (e.g. 100 000 coefficients), this model leads to a viable way to arrange data and unknowns, such that computational requirements remain limited.

In particular, this work describes representations of the gravitational potential, on the sphere and along a nominal orbit, leading to the lumped coefficient model. Then a comprehensive set of transfer coefficients for the above functionals is derived. Next, a spectral analysis follows, in which also the stochastic model in the spectral domain is explained. After developing the required least squares theory, including regularization aspects, a number of pre-mission error analysis tools and error representations are presented. Finally, several case studies display the single and combined effects of the above functionals, and of several other parameters.

---

## Zusammenfassung

IN DIESER Arbeit wird ein semi-analytisches Verfahren zur Gravitationsfeldbestimmung aus Satellitenbeobachtungen vorgestellt. Kern des Verfahrens ist die *lumped coefficient* Formulierung, welche eine Verbindung zwischen Gravitationsfeldfunktionalen und den unbekanntem Parametern des Gravitationsfeldes herstellt. Beobachtbare Größen sind hierbei die *lumped coefficients*, die im Wesentlichen die Fourier-Koeffizienten der beobachteten Zeitreihe entlang der Satellitenbahn sind. Die Unbekannten sind die Kugelfunktionskoeffizienten, d.h. das sphärisch-harmonische Spektrum des Gravitationsfeldes. Der lineare Übergang zwischen beiden Arten von Spektren ist durch sogenannte Transferkoeffizienten gegeben, die in dem semi-analytischen Verfahren analytisch bestimmt werden.

Ein Satz solcher Transferkoeffizienten, die für Satellitengeodäsie und Schwerefeldbestimmung relevant sind, wird hier als *pocket guide* bezeichnet. Folgende satellitengestützte Beobachtungsverfahren sind in diesem *pocket guide* aufgenommen:

- Hoch-niedrig-Variante des sogenannten *satellite-to-satellite tracking (high-low SST)*. Dieser Beobachtungstyp wird realisiert durch GPS-Beobachtung an Bord eines niedrigfliegenden Satelliten. Es entsteht eine dreidimensionale hochgenaue und kontinuierliche Bahnbeschreibung.
- Niedrig-niedrig-Variante des *satellite-to-satellite tracking (low-low SST)*. Dieser Beobachtungstyp wird realisiert durch Abstands- oder Geschwindigkeitsbeobachtungen zwischen zwei niedrigfliegenden Satelliten.
- Satellitengradiometrie. Bei diesem Beobachtungstyp werden die 1. räumlichen Ableitungen der Komponenten des Schwerevektors, also die 2. räumlichen Ableitungen des Gravitationspotentials bestimmt. Die Methode kann zum Beispiel realisiert werden durch differenzielle Beschleunigungsmessung über kurze Basislinien.

Die *lumped coefficient* Formulierung stellt, zusammen mit einem Satz von Transferkoeffizienten, das Beobachtungsmodell dar. Unter Einbindung eines stochastischen Modells ist damit die Grundlage für Schwerefeldbestimmung aus Satellitendaten gegeben. In dieser Arbeit werden Beobachtungsmodell und stochastisches Modell hauptsächlich für *a priori* Fehleranalyse eingesetzt. Dabei können Beobachtungstyp, Fehlercharakteristik, Bahnparameter, Missionsdauer usw. beliebig variiert werden. Auf diese Art kann zum Zwecke der Missionsplanung eine Fehlerabschätzung durchgeführt werden.

Einer der Hauptvorteile des semi-analytischen Verfahrens ist die Tatsache, daß Normalgleichungssysteme block-diagonal werden. Angesichts der großen Zahl von Daten und Unbekannten (z.B. 100 000 Kugelfunktionskoeffizienten) führt dieses Verfahren zur Gruppierung von Daten und Unbekannten, so daß selbst der Rechenaufwand mit beschränkten Rechnerkapazitäten bewältigbar ist.

Im Einzelnen beschreibt diese Arbeit verschiedene Darstellungen des Gravitationspotentials auf der Kugel und entlang der Bahn. Letztere führt dann zur *lumped coefficient* Formulierung. Ein umfassender Satz von Transferkoeffizienten wird für die obenerwähnten Beobachtungsgrößen abgeleitet. Eine Spektralanalyse folgt, inklusive einer Darstellung des stochastischen Modells im Spektralbereich. Nach einer Diskussion der benötigten Theorie der Methode der kleinsten

---

Quadrate und der Regularisierung, werden etliche Werkzeuge zur *a priori* Fehlerabschätzung und Möglichkeiten der Fehlerdarstellung präsentiert. Schließlich werden in mehreren Fallstudien Einzeleffekte untersucht, die durch die Wahl des Beobachtungstyps und durch die obengenannten Definitionsparameter auftreten.

# Contents

<b>Abstract</b>	<b>i</b>
<b>Zusammenfassung</b>	<b>ii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 From Analytical to Numerical Techniques . . . . .	1
1.2 Developments in Satellite Technology . . . . .	2
1.3 A Semi-Analytical Approach . . . . .	3
1.4 Outline of This Work . . . . .	4
<b>2 Parametrization of the Geopotential</b>	<b>5</b>
2.1 Representation on the Sphere . . . . .	5
2.2 Representation along the Nominal Orbit . . . . .	7
2.3 Lumped Coefficient Representation . . . . .	11
<b>3 Pocket Guide of Dynamical Satellite Geodesy</b>	<b>13</b>
3.1 Tools for Derivation of Transfer Coefficients . . . . .	14
3.2 Validity of the Linear Model . . . . .	16
3.3 Summary . . . . .	17
<b>4 Functionals of the Geopotential</b>	<b>18</b>
4.1 First Derivatives: Gravitational Attraction . . . . .	18
4.2 Second Derivatives: the Gravity Gradient Tensor . . . . .	20
4.3 Orbit Perturbations . . . . .	22
4.4 Low-Low Intersatellite Range Perturbation . . . . .	24
4.5 Time Derivatives . . . . .	27
4.6 Gradiometry by Accelerometric Low-Low SST . . . . .	28
4.7 Summary . . . . .	30
<b>5 A Spectral Analysis</b>	<b>33</b>

5.1 Spectral Domains . . . . .	34
5.2 The Lumped Coefficient Approach . . . . .	35
5.3 Fourier Domain Mapping . . . . .	38
5.4 Power Spectral Density Modelling . . . . .	41
5.5 Non-Circular Nominal Orbit . . . . .	44
5.6 Summary . . . . .	45
<b>6 Least Squares Error Analysis</b>	<b>47</b>
6.1 Least Squares Estimation and Regularization . . . . .	47
6.2 Contribution and Redundancy . . . . .	51
6.3 Biased Estimation . . . . .	53
6.4 Least Squares Error Simulation . . . . .	56
6.5 Pre-Mission Analysis Types . . . . .	57
6.6 Spectral Error Representation . . . . .	60
6.7 Spatial Error Representation . . . . .	65
6.8 Summary . . . . .	66
<b>7 Applications to Synthetical Satellite Gravity Missions</b>	<b>69</b>
7.1 Non-Isotropic SH Error Spectra . . . . .	69
7.2 Gradiometry . . . . .	73
7.3 Orbit Perturbations . . . . .	81
7.4 Intersatellite Range Perturbations . . . . .	86
7.5 Miscellaneous . . . . .	89
7.6 Summary . . . . .	92
<b>8 Concluding Remarks</b>	<b>94</b>
8.1 Discussion . . . . .	94
8.2 Outlook . . . . .	95
<b>List of Symbols</b>	<b>97</b>
<b>List of Abbreviations</b>	<b>98</b>
<b>A Properties of Inclination Functions</b>	<b>99</b>
A.1 Symmetries of Representation Coefficients . . . . .	99
A.2 Symmetries of Inclination Functions . . . . .	100

<b>B</b>	<b>Block-Diagonal Error Propagation</b>	<b>102</b>
B.1	Full Covariance Propagation . . . . .	102
B.2	Block Covariance Propagation . . . . .	104
B.3	Diagonal Covariance Propagation . . . . .	105
B.4	Isotropic Covariance Propagation . . . . .	105
B.5	Along-Orbit Covariance Propagation . . . . .	105
B.6	Omission Errors on the Sphere . . . . .	106
B.7	Along-Orbit Omission Errors . . . . .	107
	<b>References</b>	<b>108</b>

# 1 Introduction

## 1.1 From Analytical to Numerical Techniques

THE wide-spread use of satellite geodetic techniques and their achievements at present make it hard to believe that this is a development of only a few decades. Before the 1957 Sputnik launch, geodesy was—in a global sense—an underdeveloped science, although less on the theory side than on the data and application side. Geodesy itself was not an integrated discipline but rather divided in sub-disciplines (horizontal control, vertical control, physical geodesy). Both a global datum and a global geoid were practically non-existent, cf. the introduction by Henriksen in (NASA, 1977). Moreover, the terrestrial gravity data-base was extremely sparse.

In his enjoyable book *A tapestry of orbits* King-Hele (1992) gives a personal and colourful account of the early years of space science. During the first decade of satellite geodesy the knowledge of the Earth's gravity field improved dramatically. Soon after the Sputnik launch, the value for the Earth's dynamic flattening  $J_2$  had to be adapted already. Additionally, higher degree even zonal coefficients  $J_{2k}$  were determined. The next achievement was the determination of the Earth's *pear-shape*  $J_3$  and of further odd zonal coefficients  $J_{2k+1}$ , e.g. (O'Keefe, 1959). With more and more satellites in orbit the recovery of low-degree tesseral coefficients became possible. A further milestone was the determination of higher order tesseral harmonics by the analysis of orbital resonance, cf. (Gooding, 1971).

These developments led to several series of gravity field models. In the United States early developments took place at the Smithsonian Astrophysical Observatory, leading to the Smithsonian Standard Earth models (SSE). At NASA's Goddard Space Flight Center the Goddard Earth Models (GEM-series) were initiated. In Europe, a German-French cooperation resulted in a series of GRIM models. In contrast to such so-called comprehensive models, in which all coefficients up to a certain maximum degree  $L$  are determined, the British school computed only selected coefficients from orbital resonance. See (Seeber, 1993) or (Bouman, 1997a) for a historical survey on global gravity field models.

The early coefficient determinations were mainly based on analytical theories of satellite motion. Usually these theories were rooted in celestial mechanics, e.g. (Kaula, 1966), although exceptions occurred, cf. (King-Hele, 1992). Gaposchkin (1978) presents an overview of developments in analytical satellite theory during the early satellite years, including non-gravitational force modelling.

With more and more data coming in and with increasing number of coefficients to be determined, the analytical approach was mostly abandoned in favour of numerical methods. In particular the progress in computer technology contributed

to this development. In modelling satellite-only gravity fields the state-of-art is *brute force* numerical computation, in which the design matrix (i.e. the matrix of partials) is derived by numerical integration of variational equations. The tide of time was definitely against analytical approaches, although the English school adhered to analytical solutions from orbit resonance. To date, these coefficient solutions can compete with modern numerical solutions in comprehensive modelling (King-Hele & Winterbottom, 1994).

## 1.2 Developments in Satellite Technology

CLASSICAL orbit tracking—and subsequent gravity field recovery—is based on methods that provide observations over short time intervals, resulting in low spatial coverage. Moreover, with the exception of optical data, these methods are usually one-dimensional, e.g. satellite laser ranging (SLR), range-rate or Doppler tracking. With the advent of satellite altimetry, a first glimpse of the potential of *continuous* tracking arose. The marine gravity field was mapped with unprecedented resolution.

The concept of *space-borne* GPS tracking extends this idea of continuous orbit tracking. It provides precise orbit determination in three dimensions in a nearly geometric (or kinematic) fashion. The requirements for precise dynamic orbit modelling become less stringent. In (Yunck, 1993), orbit solutions from this type of modelling are called *reduced dynamic*, although the name *augmented kinematic* might be better suited. Other tracking systems evolved during the nineties in France (the DORIS system) and in Germany (PRARE). Both systems, based on range and/or range-rate observations, are also capable of continuous orbit tracking, although the observations are basically one-dimensional.

A further technological development of utmost relevance to space-borne gravity field determination is the use of *accelerometers*. Due to the attenuation of the gravity signal with height, the orbit of any gravity field mission must be as low as possible. At these altitudes, however, the residual atmosphere considerably influences the orbital motion. The corresponding air drag and other non-conservative forces will have to be measured or compensated, in order to separate gravitational from non-gravitational signals. Stated the other way around: the development of accelerometers created the possibility to decrease satellite altitudes to a level, interesting for high resolution gravity field research. Accelerometers are also used, by combining two or more, to measure elements of the gravity gradient tensor. Gravity gradiometry is therefore a logical technological further step.

As a result of these technological developments—continuous tracking and accelerometry—three basic space-borne gravity field methods stand out:

- satellite-to-satellite tracking in high-low mode (high-low SST). The orbit of a low Earth orbiter (LEO) is tracked in three dimensions by means of space-borne GPS tracking.
- satellite-to-satellite tracking in low-low mode (low-low SST). The intersatellite range or range-rate between two co-orbiting LEOs is measured.
- satellite gravity gradiometry (SGG). The differential acceleration between two or more nearby proof masses is measured.

All methods require the use of accelerometers. See (Rummel, 1986b) and (Wakker, 1988) for further historical reference and developments of these technologies.

### 1.3 A Semi-Analytical Approach

AS A consequence of the aforementioned technological developments the problem dimensions in dynamical satellite geodesy (DSG) are enormous nowadays. High resolution gravity field mapping requires many parameters to be determined. The amount of spherical harmonic (SH) coefficients grows quadratically with the maximum degree of development. For a static gravity field up to  $L = 300$ , nearly 100 000 unknowns have to be determined. At the observation side the effect of increasing the maximum degree  $L$  is twofold. On the one hand, the time-sampling becomes denser in order to capture the increasing maximum signal frequency. On the other hand, the mission duration becomes longer to guarantee the proper spectral resolution (in Fourier sense). The amount of data grows quadratically with  $L$  as well.

The tide from analytical to numerical techniques in dynamical satellite geodesy seems to be reversing. Despite the rapid developments in computer hardware technology, the aforementioned dimensions of gravity field determination prompt for the development of semi-analytical techniques. This work sketches the road to such semi-analytical techniques. To this end the so-called *lumped coefficient* (LC) approach is used, which is a bi-spectral approach. It links the Fourier spectrum of observations to the SH spectrum of the geopotential. By introducing the concept of a *nominal orbit* and of linear orbit perturbation theory, this link is linear. As a key property of the semi-analytical approach each individual spherical harmonic order  $m$  generates a separate linear system. This gives rise to a block-diagonal system of normal matrices, of which the maximum block size will be less than  $L \times L$ .

In the LC approach the spectral link between observables and unknowns is given by the *transfer coefficients*. They provide the design matrix in an analytical way. Each type of observable corresponds to a certain transfer coefficient. A set of transfer coefficients for several observables, relevant in DSG, will be referred to as *pocket guide* (PG) of *dynamical satellite geodesy* here. Besides the algorithmic advantages of the semi-analytical approach, it also provides a unified representation of all functionals of the geopotential.

Evaluation of transfer coefficients on the nominal orbit leads to an approximated version of the design matrix. The resulting solution of the vector of unknowns will not be exact either. This problem has to be overcome by corrections to the observations and by iterating the process. Due to the block structure, the iteration can easily be done when the original approximated design matrix is retained. The procedure is comparable to the modified Newton-Raphson iteration, in which the root of a non-linear function is found by iteration, with the derivative evaluated only once (Strang, 1986).

Nevertheless, the correctness of the whole procedure cannot be taken for granted. In particular the non-linearity in orbital motion questions the validity of the semi-analytical approach for functionals of perturbation type, especially SST. This question is discussed more extensively in the corresponding chapter.

In summary, the semi-analytical approach in this work is characterized by:

- lumped coefficient formulation: a unified description of gravitational functionals in the spectral domain
- pocket guide of transfer coefficients: analytical partials
- evaluation of design matrix on nominal orbit: block structure
- $m$ -block normal matrix inversion

Because of the latter aspect, that still involves non-trivial numerical computations, the label *semi-* is retained in the phrase *semi-analytical*.

## 1.4 Outline of This Work

THIS work starts with representations of the Earth's gravitational potential. The conventional parametrization in spherical harmonic series is followed, although in complex-valued quantities. This representation is subsequently rotated into an orbital frame. Similar to Kaula (1966), one arrives at an expression of the geopotential in orbital variables. By rearranging the summations a lumped coefficient representation is introduced naturally.

In the subsequent chapter 4 the pocket guide is introduced. All relevant functionals of the geopotential are modelled in terms of their transfer coefficients. Since transfer coefficients are the building blocks for the design matrix, this chapter defines the observation model. The validity of the linear model, in particular for observables of perturbation type, is discussed.

Chapter 5 deals with spectral aspects: 1D and 2D Fourier interpretations of the lumped coefficients, how to prevent overlapping frequencies, sampling considerations, additional spectral lines in case of eccentric nominal orbits, and so on. Particularly, the noise power spectral density (PSD) is introduced, leading to a frequency-domain stochastic model.

In chapter 6 the process of least squares gravity field recovery is discussed. Since problems in dynamical satellite geodesy tend to be ill-posed, regularization and biased estimation play an important role in this discussion. Several quality measures are developed, that are of help in the analysis of the least squares process: a posteriori covariance matrix, contribution matrix, condition number and correlation measure. In this work, the emphasis is on least squares error analysis, i.e. the pre-mission analysis tool, that determines the expected accuracy of the unknowns without actual data. Only the observation and the stochastic model are required. Much attention is paid to graphical representation of error results.

Finally, in chapter 7 the analysis techniques are applied to an extensive set of simulations. All relevant observation types, high-low SST, low-low SST and SGG are treated in numerous orbit and mission scenarios. Without reference to actual or real missions, these scenarios investigate the effects of the type of observable, single vs. multiple observables, orbital height, inclination and many others more.

## 2 Parametrization of the Geopotential

ONE OF the main aims of dynamical satellite geodesy is the determination of the Earth's gravitational field from observations made using spacecraft. A spacecraft can be used in the following, possibly overlapping, senses (Rummel, 1992; Schneider, 1988):

- i) as far away geometric target,
- ii) as measurement platform,
- iii) as a proof mass, falling in the Earth's gravitational field,
- iv) as a gyroscope, the orbit being an inertial reference.

Since the observables will be functionals of the geopotential field, the first question to be settled is the parametrization of the geopotential.

Section 2.1 will be concerned with the representation of the geopotential on the sphere. In 2.2 this representation will be transformed to the orbit system. In 2.3 the basic linear model of dynamical satellite geodesy is then derived.

### 2.1 Representation on the Sphere

A SATELLITE, revolving around the Earth, samples the field *globally* by nature. Moreover we will be concerned with the geopotential on a global scale. For these reasons the gravitational field will be represented in a spherical harmonic series throughout this work. Such a representation turns out to be of advantage, since spherical harmonics possess the following properties: orthogonality, global support and harmonicity. Because the geopotential fulfills the Laplace equation  $\Delta V = 0$  outside the masses, the harmonicity of the spherical harmonics makes them natural basefunctions to  $V$ . Their orthogonality allows the analysis of the coefficients of the base functions.

The gravitational potential  $V$  is developed into spherical harmonics. For reasons of compactness *complex-valued* quantities are employed here:

$$V(r, \theta, \lambda) = \frac{GM}{R} \sum_{l=0}^{\infty} \left(\frac{R}{r}\right)^{l+1} \sum_{m=-l}^l \bar{K}_{lm} \bar{Y}_{lm}(\theta, \lambda), \quad (2.1)$$

in which

$r, \theta, \lambda$  = radius, co-latitude, longitude

$R$  = Earth's equatorial radius

$GM$  = gravitational constant times Earth's mass

$\bar{Y}_{lm}(\theta, \lambda)$  = surface spherical harmonic of degree  $l$  and order  $m$

$\bar{K}_{lm}$  = spherical harmonic coefficient, corresponding to  $\bar{Y}_{lm}(\theta, \lambda)$ .

The coefficients  $\bar{K}_{lm}$  constitute the *spherical harmonic spectrum* of the function  $V$ . They are the parameters of the gravitational field. The surface spherical harmonics  $\bar{Y}_{lm}(\theta, \lambda)$  are defined in the following way:

$$\bar{Y}_{lm}(\theta, \lambda) = \bar{P}_{lm}(\cos \theta) e^{im\lambda}, \quad (2.2)$$

where the normalized *Legendre* function  $\bar{P}_{lm}(\cos \theta)$  reads:

$$\bar{P}_{lm}(\cos \theta) = \begin{cases} N_{lm} P_{lm}(\cos \theta), & m \geq 0 \\ (-1)^m \bar{P}_{l,-m}(\cos \theta), & m < 0 \end{cases}. \quad (2.3)$$

It follows from this definition that for the complex conjugated of  $\bar{Y}_{lm}(\theta, \lambda)$  it holds:  $\bar{Y}_{lm}^* = (-1)^m \bar{Y}_{l,-m}$ . The quantities denoted by an overbar are (fully) normalized by the factor:

$$N_{lm} = (-1)^m \sqrt{(2l+1) \frac{(l-m)!}{(l+m)!}}. \quad (2.4)$$

Normalized functions and coefficients are obtained from their unnormalized counterparts by:

$$\begin{aligned} \bar{Y}_{lm} &= N_{lm} Y_{lm} \\ \bar{K}_{lm} &= N_{lm}^{-1} K_{lm}. \end{aligned}$$

Notice that the unnormalized version of the spherical harmonic series (2.1) would—apart from the overbars—appear exactly the same. The orthogonality of the base functions is expressed by:

$$\frac{1}{4\pi} \iint_{\sigma} \bar{Y}_{l_1 m_1} \bar{Y}_{l_2 m_2}^* d\sigma = \delta_{l_1 l_2} \delta_{m_1 m_2}. \quad (2.5)$$

**Conventions.** In literature, sometimes the factor  $\frac{1}{4\pi}$  is taken care of in the normalization factor by incorporating a term  $\sqrt{4\pi}$ , eg. (Edmonds, 1957; Ilk, 1983). Another difference between normalization factors, found in literature, is the factor  $(-1)^m$ . It is often used implicitly in the definition of the Legendre functions.

In principle the issues of normalization factors and of complex vs. real are irrelevant. In geodesy, one usually employs real-valued base functions and coefficients, cf. (Heiskanen & Moritz, 1967). The series (2.1) would become:

$$V(r, \theta, \lambda) = \frac{GM}{R} \sum_{l=0}^{\infty} \left(\frac{R}{r}\right)^{l+1} \sum_{m=0}^l \left( \bar{C}_{lm} \cos m\lambda + \bar{S}_{lm} \sin m\lambda \right) \bar{P}_{lm}(\cos \theta), \quad (2.6)$$

with normalization factor:

$$N_{lm} = \sqrt{(2 - \delta_{m0})(2l+1) \frac{(l-m)!}{(l+m)!}}. \quad (2.7)$$

The real- and complex-valued spherical harmonic coefficients, each with their own normalization, are linked by:

$$\bar{K}_{lm} = \begin{cases} (-1)^m (\bar{C}_{lm} - i\bar{S}_{lm})/\sqrt{2}, & m > 0 \\ \bar{C}_{lm}, & m = 0 \\ (\bar{C}_{lm} + i\bar{S}_{lm})/\sqrt{2}, & m < 0 \end{cases},$$

such that  $\bar{K}_{lm} = (-1)^m \bar{K}_{l,-m}^*$ .

## 2.2 Representation along the Nominal Orbit

IN ORDER to link observations, being functionals of  $V$ , to the potential itself, this section will deal with representing  $V$  along the orbit, cf. (Sneeuw, 1992). A key step towards the semi-analytical approach will be the introduction of a *nominal* orbit. The following simplified configuration will be assumed:

- the orbit is circular ( $e = 0$ ), i.e. its radius  $r$  is constant,
- the inclination  $I$  is constant,
- the orbit is secularly precessing due to the flattening term  $J_2 = -\sqrt{5}\bar{K}_{2,0}$ .

This assumption does not necessarily mean that measurements are made on (or reduced to) this nominal orbit. The orbit should be considered as a set of (Taylor-) points, where the observational model is evaluated, cf. (Colombo, 1986; Betti & Sansò, 1989). We will come back to this essential interpretation of the nominal orbit in 3.2.

**Orbit Angular Velocities.** The Earth's oblateness causes secular perturbations in the three angular Kepler elements mean anomaly ( $M$ ), right ascension of the ascending node ( $\Omega$ ), and argument of perigee ( $\omega$ ). The three metric Kepler elements inclination ( $I$ ), semi-major axis ( $a$ ) and eccentricity ( $e$ ) do not change due to the Earth's flattening. From (Kaula, 1966) we take:

$$\dot{\Omega} = -\frac{3}{2}nJ_2\left(\frac{R}{r}\right)^2 \cos I, \quad (2.8a)$$

$$\dot{\omega} = \frac{3}{4}nJ_2\left(\frac{R}{r}\right)^2 [5 \cos^2 I - 1], \quad (2.8b)$$

$$\dot{M} = n + \frac{3}{4}nJ_2\left(\frac{R}{r}\right)^2 [3 \cos^2 I - 1], \quad (2.8c)$$

where zero eccentricity has been assumed already. Since the nominal orbit is circular, the radius  $r$  is used for the semi-major axis  $a$ . However, the perigee is not defined for a circular orbit. To this end, the angle within the orbital plane from the ascending node to the satellite is used. It is the sum of the argument of perigee and the true anomaly, and is denoted as the *argument of latitude* ( $u$ ). Along the circular orbit true and mean anomaly are equal. The precession of the argument of latitude thus becomes:

$$\dot{u} = \dot{\omega} + \dot{M} = n + \frac{3}{2}nJ_2\left(\frac{R}{r}\right)^2 [4 \cos^2 I - 1]. \quad (2.8d)$$

The mean motion  $n$  comes from Kepler's second law ( $n^2 r^3 = GM$ ). The orbit configuration is displayed in fig. 2.1. One further angle will be of importance. It is the longitude of the ascending node ( $\Lambda$ ), i.e. in an Earth-fixed system. The right ascension is defined in an inertial reference frame, with respect to the vernal equinox  $\Upsilon$ . The inertial frame differs from the Earth-fixed one by a rotation about the common  $Z$ -axis by the hour angle of the Greenwich meridian. Thus the Greenwich sidereal time, say  $\theta$ , expressed in units of angle, has to be subtracted:  $\Lambda = \Omega - \theta$ . Its rate is:

$$\dot{\Lambda} = \dot{\Omega} - \dot{\theta} = -\frac{3}{2}nJ_2\left(\frac{R}{r}\right)^2 \cos I - \frac{2\pi}{\text{day}}. \quad (2.8e)$$

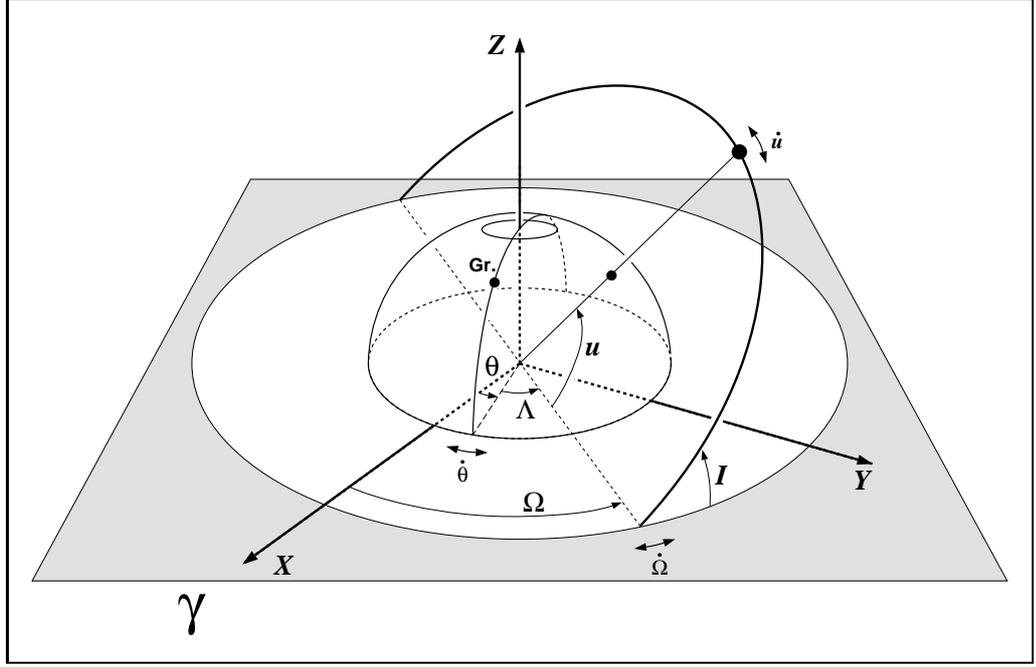


Figure 2.1: Nominal orbit configuration.

The two quantities  $\dot{u}$  and  $\dot{\Lambda}$  are the basic angular velocities—or frequencies—of the nominal orbit.

**Rotation of Spherical Harmonics.** From representation theory, e.g. (Wigner, 1959), and quantum mechanics, e.g. (Edmonds, 1957), the transformation properties of spherical harmonics under rotation are known. Let the Eulerian rotation sequence  $\mathcal{R}(\alpha, \beta, \gamma)$  be defined by:

- i) rotation  $\alpha \in [0; 2\pi)$  about the *initial*  $z$ -axis, followed by
- ii) rotation  $\beta \in [0; \pi]$  about the *new*  $y$ -axis, and finally
- iii) rotation  $\gamma \in [0; 2\pi)$  about the *final*  $z$ -axis.

Spherical harmonics in the original frame  $(\theta, \lambda)$  are a linear combination of harmonics of the same degree  $l$  in the rotated frame  $(\theta', \lambda')$ :

$$\bar{Y}_{lm}(\theta, \lambda) = \sum_{k=-l}^l \bar{D}_{lmk}(\alpha, \beta, \gamma) \bar{Y}_{lk}(\theta', \lambda'), \quad (2.9)$$

with the representation coefficients:

$$\bar{D}_{lmk}(\alpha, \beta, \gamma) = e^{im\alpha} \bar{d}_{lmk}(\beta) e^{ik\gamma}, \quad (2.10)$$

and

$$\bar{d}_{lmk}(\beta) = \left[ \frac{(l+k)!(l-k)!}{(l+m)!(l-m)!} \right]^{\frac{1}{2}} \sum_{t=t_1}^{t_2} \binom{l+m}{t} \binom{l-m}{l-k-t} (-1)^t c^{2l-a} s^a, \quad (2.11)$$

where  $c = \cos \frac{1}{2}\beta$ ,  $s = \sin \frac{1}{2}\beta$ ,  $a = k - m + 2t$ ,  $t_1 = \max(0, m - k)$  and  $t_2 = \min(l - k, l + m)$ . Insertion in the series expression (2.1) yields:

$$V(r, \theta', \lambda') = \frac{GM}{R} \sum_{l=0}^{\infty} \left( \frac{R}{r} \right)^{l+1} \sum_{m=-l}^l \sum_{k=-l}^l \bar{K}_{lm} \bar{D}_{lmk}(\alpha, \beta, \gamma) \bar{Y}_{lk}(\theta', \lambda'), \quad (2.12)$$

which directly shows the transformation rule for the coefficients under rotation:

$$\bar{K}_{lk} = \sum_{m=-l}^l \bar{D}_{lmk}(\alpha, \beta, \gamma) \bar{K}_{lm}.$$

In order to arrive at the required potential expression along the orbit, the Earth-fixed reference frame will be rotated, such that eventually the rotated  $x$ -axis will point towards the satellite, and the orbital plane will be perpendicular to the rotated  $z$ -axis. To this purpose, the following Eulerian rotation sequence is applied, cf. (Giacaglia, 1980; Sneeuw, 1992):

- i) Rotate the Earth-fixed  $x$ -axis (through Greenwich meridian) to the ascending node:  $\Lambda$  (=longitude of the ascending node).
- ii) Tilt the equatorial plane to the orbital plane. Notice that this requires a rotation about the new  $x$ -axis, instead of the  $y$ -axis, using the inclination  $I$ .
- iii) Bring the  $x$ -axis (now through ascending node) to the satellite:  $u$  (=argument of latitude).

In this sequence the second rotation is about the line of nodes, which is an  $x$ -axis. This situation is overcome by pre- and post-rotating with  $\pm\frac{1}{2}\pi$ , cf. (Betti & Sansò, 1989). The representation coefficients to be used now are:

$$\begin{aligned} \bar{D}_{lmk}(\alpha, \beta, \gamma) &:= \bar{D}_{lmk}(\Lambda - \frac{1}{2}\pi, I, u + \frac{1}{2}\pi) \\ &= i^{k-m} \bar{d}_{lmk}(I) e^{i(ku + m\Lambda)}. \end{aligned} \quad (2.13)$$

Using the time-variable elements  $u(t)$  and  $\Lambda(t)$ , the rotation sequence will keep the new  $x$ -axis pointing to the satellite. Its orbital plane will instantaneously coincide with a new equator. The satellite's coordinates reduce to  $\theta' = \frac{1}{2}\pi$  and  $\lambda' = 0$ , so that  $\bar{Y}_{lk}(\theta', \lambda') = \bar{P}_{lk}(0)$ . In principle the third rotation could have been omitted such that the representation coefficient  $\bar{D}_{lmk}(\Lambda - \frac{1}{2}\pi, I, \frac{1}{2}\pi)$  should have been used. In that case the longitude in the new frame would have been  $\lambda' = u$ , leading to the same expression. In both cases the satellite is always on the rotated equator. In the second interpretation the argument of latitude would become the new longitude. In this view the name *argument of latitude* is highly misplaced.

**Coordinates.** The geopotential was expressed in 2.1 as a spherical harmonic series in the spherical coordinates  $\{r, \theta, \lambda\}$ . After the rotations, one ends up with the coordinate system  $\{r, \theta', \lambda'\}$ . Applying the time-dependent rotations to the nominal orbit, these coordinates are not useful, since they become constants. From another viewpoint, the Kepler elements can be considered as coordinates of a 6-dimensional phase space. The circular nominal orbit reduces this space to four phase-space coordinates  $\{r, u, I, \Omega\}$  or  $\{r, u, I, \Lambda\}$ . Eventually, for this specific orbit configuration, the only independent coordinate is the time  $t$ , while all others  $\{r, \dot{u}, u_0, I, \dot{\Lambda}, \Lambda_0\}$  can be considered as orbital constants. This leads to a time-series expression.

A pragmatic approach as to which coordinates are 'best', is probably the most sensible one. Schrama (1989) and Koop (1993), for instance, consider the inclination  $I$  as coordinate, not as orbital constant, leading to three possible coordinate

sets:  $r$  together with two out of  $\{u, \Lambda, I\}$ . In the sequel, the potential series expression will employ the set  $\{r, u, \Lambda\}$  as coordinates. It does not span the same space as  $\{r, \theta, \lambda\}$ , of course, since two cones around the poles are missing for non-polar orbits. However, they do cover the configuration space of the particular preceding orbit. Although  $r$  is a constant, we will keep it as a coordinate for differentiation purposes. For the same purpose one should keep in mind that, although the satellite is always on the rotated equator, the coordinate  $\theta'$  still denotes co-latitude in the rotated system. Furthermore we consider  $I$ , which has no secular rate, as an orbital constant. Inserting the representation coefficients (2.13) and the fact that  $\theta' = \frac{1}{2}\pi, \lambda' = 0$  into (2.12) yields:

$$\begin{aligned} V(r, \theta', \lambda') &:= V(r, u, \Lambda) \\ &= \frac{GM}{R} \sum_{l=0}^{\infty} \left(\frac{R}{r}\right)^{l+1} \sum_{m=-l}^l \sum_{k=-l}^l \bar{K}_{lm} i^{k-m} \bar{d}_{lmk}(I) \bar{P}_{lk}(0) e^{i(ku + m\Lambda)}. \end{aligned}$$

**Inclination Functions.** As a last step a complex-valued *inclination function* is introduced:

$$\bar{F}_{lmk}(I) = i^{k-m} \bar{d}_{lmk}(I) \bar{P}_{lk}(0), \quad (2.14)$$

so that the along-orbit potential is finally reduced to the series:

$$V(r, u, \Lambda) = \frac{GM}{R} \sum_{l=0}^{\infty} \left(\frac{R}{r}\right)^{l+1} \sum_{m=-l}^l \sum_{k=-l}^l \bar{K}_{lm} \bar{F}_{lmk}(I) e^{i(ku + m\Lambda)}. \quad (2.15)$$

The inclination functions (2.14) differ from Kaula's functions  $F_{lmp}(I)$  (Kaula, 1966) in the following aspects:

- they are complex,
- they are normalized by the factor (2.4),
- they make use of the index  $k$ .

The latter aspect is also employed by Emeljanov & Kanter (1989) in their functions  $F_{lm}^k(I)$ . See also the "COMMENTS ON THE NOTATION" in (Gooding & King-Hele, 1989). Legendre functions  $P_{lk}(x)$  are even functions in  $x \in [-1; 1]$  for  $l - k$  even. For  $l - k$  odd the functions are odd, implying a root at the equator:  $P_{lk}(0) = 0$ . To be specific, the unnormalized Legendre functions assume at the equator the value, cf. (Giacaglia, 1980; Sneeuw, 1992):

$$P_{lk}(0) = \begin{cases} 2^{-l} (-1)^{\frac{l-k}{2}} \frac{(l+k)!}{\left(\frac{l-k}{2}\right)! \left(\frac{l+k}{2}\right)!} & \text{for } l - k \text{ even} \\ 0 & \text{for } l - k \text{ odd.} \end{cases}$$

Consequently the inclination functions attain a zero value when  $l - k$  odd. This fact allows the introduction of another index:  $p = \frac{1}{2}(l - k)$ , which is used in Kaula's  $\bar{F}_{lmp}(I)$ . These may be defined analogously, though real, to (Sneeuw, 1992):

$$\bar{F}_{lmp}(I) = N_{lm} d_{lm, l-2p}(I) P_{l, l-2p}(0) (-1)^{p + \text{int}\left(\frac{l-m+1}{2}\right)}.$$

**Remark 2.1** *The  $p$ -index has two advantages: it is positive and it runs in unit steps. The third summation in (2.15) becomes  $\sum_{p=0}^l$ . The major disadvantage*

is that it does not have the meaning of spherical harmonic order (or azimuthal order) in the rotated system anymore. The index  $p$  is not a wavenumber, such as  $k$ . Thus symmetries are lost, and formulae become more complicated, (Gooding & King-Hele, 1989). For instance  $\exp(i(ku + m\Lambda))$  must be written as  $\exp(i((l - 2p)u + m\Lambda))$ . The angular argument seems to depend on 3 indices in that case.

**Remark 2.2** It is not true that inclination functions  $\bar{F}_{lmk}(I)$  do not exist for  $l - k$  odd. Technically speaking, they are just equal to zero. This may seem a trivial remark, but in the case of cross-track inclination functions, to be introduced in 4.1, the situation is reversed. The inclination functions will become zero for  $l - k$  even.

With the basic angles  $u(t) = u_0 + \dot{u}t$  and  $\Lambda(t) = \Lambda_0 + \dot{\Lambda}t$ , we have arrived at an expression for the geopotential along the nominal orbit as a time-series:

$$V(r, u, \Lambda) = V(r, u(t), \Lambda(t)) := V(t).$$

## 2.3 Lumped Coefficient Representation

THE PART  $\exp(i(ku + m\Lambda))$  in (2.15) reminds of a 2D-Fourier series. The coordinates  $u$  and  $\Lambda$  attain values in the range  $[0; 2\pi)$ . Topologically, the product  $[0; 2\pi) \times [0; 2\pi)$  gives a torus, which is the proper domain of a 2D-Fourier series, cf. (Hofmann-Wellenhof & Moritz, 1986). Indeed the potential can be recast in a 2D-Fourier expression, if the following Fourier coefficients are introduced:

$$A_{mk}^V = \sum_{l=\max(|m|, |k|)}^{\infty} H_{lmk}^V \bar{K}_{lm}, \quad (2.16a)$$

with

$$H_{lmk}^V = \frac{GM}{R} \left(\frac{R}{r}\right)^{l+1} \bar{F}_{lmk}(I). \quad (2.16b)$$

With these quantities, the potential expression reduces to the series:

$$V(u, \Lambda) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} A_{mk}^V e^{i\psi_{mk}}, \quad (2.16c)$$

$$\psi_{mk} = ku + m\Lambda. \quad (2.16d)$$

Just like (2.15), the above equations are valid for any orbit, even in case of osculating orbital variables  $u(t)$ ,  $\Lambda(t)$ ,  $I(t)$ ,  $r(t)$ . The 2D-Fourier expression (2.16) makes only sense, though, on the nominal orbit. Only if  $I$  and  $r$  are constant, the  $H_{lmk}^V$  and correspondingly the Fourier coefficients  $A_{mk}^V$  become constant, too.

The Fourier coefficients  $A_{mk}^V$  are usually referred to in literature as *lumped* coefficients, since they are a sum (over degree  $l$ ). All potential coefficients  $\bar{K}_{lm}$  of a specific order  $m$  are lumped in a linear way into  $A_{mk}^V$ . The coefficients  $H_{lmk}^V$  are denoted *transfer* coefficients here. They are also known as *sensitivity* and *influence* coefficients. Both  $A_{mk}^V$  and  $H_{lmk}^V$  are labelled by a super index  $V$ , referring to the geopotential  $V$ .

**Lumped Coefficients.** The word *lumped* merely indicates an accumulation of numbers, e.g. here a linear combination of potential coefficients over degree  $l$ , in general. Nevertheless there exists a host of definitions and notations of lumped coefficients, e.g. real-valued coefficients  $(a_{km}, b_{km})$  in (Engelis, 1988).

Besides again the discussion of real vs. complex quantities, the main difference between the above definition and others is the fact that a circular nominal orbit was chosen as reference. In general, the orbital eccentricity  $e$  must be taken care of. The corresponding potential series must be adapted by a further summation over an index  $q$ , eccentricity functions  $G_{lkq}(e)$ , and an angular argument  $e^{iq\omega}$ , see 5.5 and (Kaula, 1966). The argument of perigee  $\omega$  precesses according to (2.8b). The eccentricity functions are of the order  $\mathcal{O}(e^{|q|})$ , so mostly a few terms, e.g.  $|q| \leq 2$ , are sufficient.

Some authors explicitly denote the expression including the sum over  $q$  as lumped coefficients, e.g. the  $(C^*, S^*)$  in (Wagner & Klosko, 1977), which are a sum over terms with  $\cos(q\omega)$  and  $\sin(q\omega)$ . Especially when the precession of the argument of perigee is slow (shallow resonance) this definition is suitable. Other authors define lumped coefficients that explicitly have an index  $q$ , such as  $(\bar{C}_m^{q,k}, \bar{S}_m^{q,k})$  in (Klokočník, 1988), or  $(\bar{C}_m^{kq}, \bar{S}_m^{kq})$  in (Wnuk, 1988; Moore & Rothwell, 1990). In these cases the transfer coefficients will include a  $q$  as well, e.g.  $Q_{lm}^{kq}$ . Such a definition is more general.

Most lumped coefficients are defined through Kepler elements (inclination and eccentricity functions). An example of lumped coefficients in terms of Hill variables is (Cui & Lelgemann, 1995). A last difference between definitions might be caused by the initial state elements of the angular variable. Let  $\psi_{mk} = \psi_{mk}^0 + \dot{\psi}_{mk}t$ . One can then write  $A_{mk} \exp(i\psi_{mk})$  either as  $[A_{mk} \exp(i\psi_{mk}^0)] \exp(i\dot{\psi}_{mk}t)$ , or as  $A_{mk} [\exp(i(\psi_{mk}^0 + \dot{\psi}_{mk}t))]$ .

An early reference where lumped coefficients are determined and discussed, is (Gooding, 1971). See (Klokočník, Kostelecký & Li, 1990) for a list of lumped coefficients from several resonant orbit perturbations. Also in (Heiskanen & Moritz, 1967) lumped coefficients are discussed; zonal lumped coefficients, to be precise, that include non-linearities.

### 3 Pocket Guide of Dynamical Satellite Geodesy

NOT ONLY the potential, but also its functionals can be represented by a 2D-Fourier series, similar to (2.16). For  $f^\#$ , in which the label  $\#$  represents a specific observable, the spectral decomposition is:

$$f^\# = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} A_{mk}^\# e^{i(ku + m\Lambda)}, \text{ with} \quad (3.1a)$$

$$A_{mk}^\# = \sum_{l=\max(|m|,|k|)}^{\infty} H_{lmk}^\# \bar{K}_{lm}. \quad (3.1b)$$

This basic linear model, especially the derivation of several transfer coefficients  $H_{lmk}^\#$ , is elaborated in 4. See also (Schrama, 1991). Spectral aspects are further treated in 5. In particular, Fourier aspects of (3.1a) are dealt with in 5.3. Section 5.2 covers aspects of spectral mapping in the lumped coefficient approach, based on (3.1b).

By means of the above equations, a linear observation model is established, that links functionals of the geopotential to the fundamental parameters, the spherical harmonic coefficients. The link is in the spectral domain. The elementary building blocks in this approach are transfer coefficients, similar to (2.16b). In conjunction with stochastic modelling, to be described later on, the linear model provides a powerful tool for gravity field analyses. E.g. the recovery capability of future satellite missions can be assessed, or the influence of gravity field uncertainties on other functionals.

A collection of transfer coefficients  $H_{lmk}^\#$  for all relevant functionals—observable or not—will be denoted as *pocket guide* (PG) to dynamical satellite geodesy (DSG). Such a PG reminds of the *Meissl scheme*, cf. (Rummel & Van Gelderen, 1995), which presents the spectral characteristics of the first and second order derivatives of the geopotential. This scheme enables to link observable gravity-related quantities to the geopotential field. A major difference between the PG and the Meissl-scheme is, that the former links SH coefficients to Fourier coefficients, whereas the latter stays in one spectral domain, either spherical harmonic or Fourier. Consequently, the transfer coefficients do not solely depend on SH degree  $l$ . In general, the spherical harmonic orders  $m$  and  $k$  are involved as well. The transfer coefficients can not be considered as eigenvalues of a linear operator, representing the observable, as in the case of the Meissl-scheme.

### 3.1 Tools for Derivation of Transfer Coefficients

THE TWO fundamental types of observables are gravity gradients, from satellite gravity gradiometry (SGG), and orbit perturbations, from satellite-to-satellite tracking (SST). Development of the  $H_{lmk}^{\#}$ 's for all gravity gradient tensor components requires first and second derivatives of the geopotential in all space dimensions. This process is relatively straightforward, cf. 4.1 and 4.2. The  $H_{lmk}^{\#}$ 's, pertaining to observables of orbit perturbation type, employ the first derivatives for the definition of the forcing functions. Moreover, a dynamical model has to be incorporated, cf. 4.3.

**Preferred Coordinate Frame.** In this work a *local Cartesian triad*, co-rotating on the nominal circular orbit with uniform rotation rate, is chosen as preferred frame to express derivatives and orbit perturbations. It is related to the instantaneous rotated geocentric spherical system ( $r, \theta' = \frac{1}{2}\pi, \lambda' = u$ ), described by the rotations in 2.2, followed by a permutation of the axes. This local Cartesian frame is especially suitable for gravity gradiometry in Earth-pointing mode. Also for satellite altimetry (radial orbit perturbation) and for low-low SST (mostly along-track perturbations) this frame is favourable.

Since the nominal orbit is circular, the classical (Serret-) Fresnet triad with tangential, principal normal and binormal base vectors coincides with the system of radial, transverse and normal directions. The latter system follows from rotating the perifocal (or apsidal) frame about the orbit normal by the true anomaly. Table 3.1 and fig. 3.1 summarize the nomenclature of local Cartesian triads and the correspondence of their base vectors.

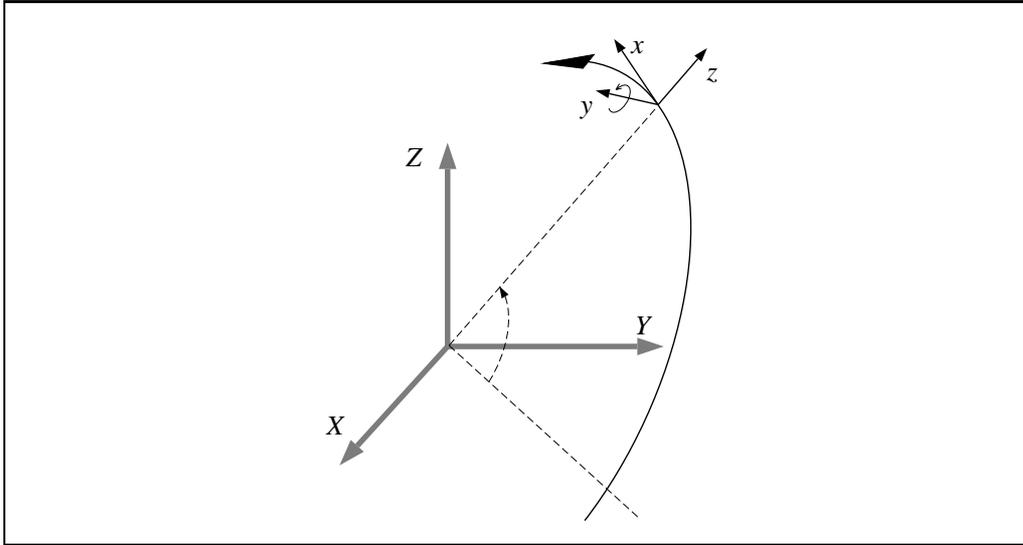
For generalizing to non-circular nominal orbits, one must strictly discern between two triads. Only on circular orbits the tangential and radial vectors are permanently orthogonal. Thus, one local triad would have a real radial axis and a quasi-tangential one. The other would have one axis in the real along-track direction and one quasi-radial. Both have the same cross-track axis. The two triads—real-radial and real-tangential—can be transformed into each other through a rotation about their common cross-track axis, cf. (Schneider, 1988, §11.3.4).

**Table 3.1:** *Base vectors of local Cartesian triads on a circular orbit.*

this work	Fresnet	rotated perifocal
along-track ( $x$ )	tangential ( $t$ )	transverse ( $\tau$ )
cross-track ( $y$ )	binormal ( $b$ )	normal ( $\chi$ or $\eta$ )
radial ( $z$ )	-[principal normal] ( $-n$ )	radial ( $r$ )

Alternatively one may choose an Earth-fixed Cartesian coordinate system as reference. To describe the gravity gradient tensor, especially for space-stable gradiometry, this system is well suited, cf. (Hotine, 1969), (Ilk, 1983) or (Bettadpur, 1991,1995). The (semi-) analytical treatment of orbit-perturbations, however, becomes cumbersome.

**Dynamical Model—Hill Equations.** A pocket guide of DSG must incorporate a dynamical model of satellite motion. Based on this model orbit perturbations are



**Figure 3.1:** The local orbital triad:  $x$  along-track,  $y$  cross-track and  $z$  radial

described. With perturbing forces given in the local Cartesian frame, one choice might be Gauss equations of motion, based either on Kepler elements (Balmino, 1993) or on Hill variables (Cui & Mareyen, 1992). Instead, the linearized *Hill equations* will be employed here, cf. (Hill, 1878; Colombo, 1984; Rummel, 1986a). The abbreviation HE will be used hereinafter.

Hill derived the equations, bearing his name, in order to describe lunar motion in a rotating Cartesian triad. They have been reinvented several times, as documented by the reference list of (Lange & Parkinson, 1966), for numerous purposes. They are used to study motion of binary asteroids, in which case non-linear Hill equations are used, eg. (Chauvineau & Mignard, 1990). They appear in guidance and inertial navigation. Also from rendez-vous problems e.g. (Kaplan, 1976), or from micro-gravity experiments, e.g. (Bauer, 1982), the HE are known. An extension to eccentric orbits and oblate potential fields is derived in (Bauer, 1982). The resulting differential equations (DE), denoted as Tschauner-Hempel equations, have time-dependent coefficients, though.

In geodesy, the HE were introduced by Colombo (1984, 1986) for describing inter-satellite range characteristics, see also (Mackenzie & Moore, 1997). Schrama (1989) used them for a spectral description of radial orbit errors in satellite altimetry. Furthermore they are applied to computation of ephemerides of GPS orbits (Colombo, 1989) and used for pre-mission analysis of low-flying Earth orbiters (Schrama, 1991; Sneeuw, 1994b; Scheinert, 1996).

The HE are derived under the assumption of a spherical potential field, linearized on a circle (Rummel, 1986a). The resulting set of (approximated) equations can be solved exactly. The Earth's oblateness is only incorporated through the introduction of the nominal precessing orbit, not explicitly in the HE themselves.

## 3.2 Validity of the Linear Model

AS MENTIONED in 2.3, a lumped coefficient representation only makes sense on the nominal orbit. In that case, the transfer coefficients and the lumped coefficients become time-independent. Observations, though, are made along the real orbit. Therefore, data and linear model are inconsistent. One way out of this situation is applying *corrections* to the data, based on a priori knowledge. The main part of this correction will be due to height variation. Particularly for gradiometry type of observations, a correction based on existing knowledge, or even on an ellipsoidal reference field, may be sufficient to reconcile the linear model with the data.

Nevertheless, the corrections can never be exact. Exact corrections would imply the knowledge of the unknown gravity field already. Therefore the corrected data will still be slightly wrong, leading to incorrect SH parameters. A further possibility of achieving a consistent combination of linear model and data, is *iteration*. This would be feasible with uncorrected data, too. In each iteration step, the data is corrected with the best available geopotential knowledge, i.e. from the previous iteration step. Subsequently, the geopotential knowledge is updated by estimating new SH coefficients from the data. Note that in every iteration step the linear model is the same: transfer coefficients, evaluated along the nominal orbit.

The nominal orbit may thus be considered a (time-) series of linearization points. Alternatively, this approach may be viewed as the modified Newton-Raphson iteration (Strang, 1986). A non-linear function is evaluated (read: *observation*) and turns out to be non-zero. Instead of evaluating the derivative, or a Jacobian in higher dimensional cases, in this point (read:  $H_{lmk}^{\#}$  *along the actual orbit*) one takes an approximate one ( $H_{lmk}^{\#}$  *along the nominal orbit*). These steps are then iterated without changing the derivative.

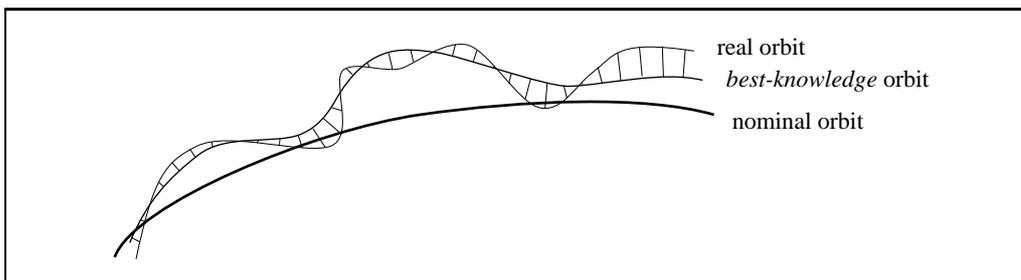
In case of observables of perturbation type, the situation is more complicated. The orbit is the observation, loosely speaking. Moreover, the question seems justified, whether Hill equations—or any linear perturbation theory, for that matter—are sufficiently accurate. In order for a linear perturbation theory to be valid, the perturbations have to remain small. In general, however, deviations of the real from the nominal orbit may become large.

**Reference Orbit.** It must be emphasized that the purpose of a pocket guide is to provide a spectral characterization of the observables. In the case of orbit perturbations e.g., it is *not* the purpose to explain and predict orbital motion itself. This would be a difficult task in view of non-linearity, resonance and the initial state problem. Only due to the separation of individual spectral lines in the frequency domain the spectral transfer of the Hill equations can be used in a sensible way. Instead of correcting data, which may be cumbersome, a reference orbit will be introduced. Also here, iteration will be required.

Suppose, for example, that the Earth's oblateness  $J_2$ , has to be determined. The main periodic effect of  $J_2$  (twice per revolution) e.g. amounts to the order of magnitude of 1 km. However, the purpose will be to improve existing knowledge of the gravity field. So not the size of  $J_2$  itself is relevant, only the level of its

uncertainty. Now consider that one integrates an orbit numerically using a force function that includes an a priori  $J_2$ -value. This orbit is denoted *best-knowledge* orbit. The actual or *real* orbit will deviate from this best-knowledge orbit only marginally, since  $J_2$  is known rather well. Now determination of  $J_2$  becomes determination of  $\Delta J_2$ , for which a linear model will probably do.

This concept is generalized now, see fig. 3.2. The procedure is to construct a numerically integrated orbit with the best a-priori knowledge of SH-coefficients  $\bar{K}_{lm}$ . The differences between real and best-knowledge orbit, due to unmodelled  $\Delta\bar{K}_{lm}$ , are projected now onto the nominal orbit. At this point, the orbit perturbations ( $\Delta x, \Delta y, \Delta z$ ) can safely be modelled by the Hill equations. The procedure is well explained in (Betti & Sansò, 1989). If iteration is necessary, a new orbit has to be integrated after solving the model parameters (gravity field coefficients). This yields new orbit perturbations and requires a new solution of the model parameters (with the old  $H_{lmk}^\#$ ).



**Figure 3.2:** *Nominal orbit (precessing circle), best-knowledge orbit (e.g. numerically integrated) and actual orbit (observed)*

### 3.3 Summary

THE LUMPED coefficient approach leads to a linear observation model. Basic ingredient are transfer coefficients  $H_{lmk}^\#$ , that map the spherical harmonic spectrum onto a Fourier spectrum of the observable along the orbit. The set of transfer coefficients of all relevant geopotential functionals is referred to as pocket guide.

Prerequisite for the lumped coefficient approach is the introduction of a nominal orbit of constant radius and constant inclination. Evaluation of the transfer coefficients with these parameters leads to time-independent lumped coefficients.

The validity of the linear model is enhanced by two counter-measures: data correction and iteration. In case of observables of perturbation type, a reference or best-knowledge orbit has to be introduced for this purpose.

## 4 Functionals of the Geopotential

BASED on the tools of 3.1—spatial differentiation and Hill equations—transfer coefficients are derived in the following sections.

### 4.1 First Derivatives: Gravitational Attraction

SINCE the satellite is in free fall, the gradient of the potential,  $\nabla V$ , is not an observable functional. Only in an indirect way it can be derived, cf. (Bassanino, Migliaccio & Sacerdote, 1991). Nevertheless, the transfer coefficients of the gradient components are of utmost relevance. They are the starting point for the second derivatives and they supply the force function to the dynamic equations. Before applying the gradient operator  $\nabla = [\frac{\partial}{\partial x} \ \frac{\partial}{\partial y} \ \frac{\partial}{\partial z}]^T = [\partial_x \ \partial_y \ \partial_z]^T$  to the geopotential expression (2.15) or to (2.16a)-(2.16d), it is recalled that in the rotated geocentric system  $u$  plays the role of longitude,  $\theta'$  that of co-latitude (although its nominal value is fixed at  $\frac{1}{2}\pi$ ) and  $r$  is the radial coordinate of course. Thus the gradient becomes:

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} = \begin{pmatrix} \frac{1}{r} \frac{\partial}{\partial u} \\ -\frac{1}{r} \frac{\partial}{\partial \theta'} \\ \frac{\partial}{\partial r} \end{pmatrix}.$$

Let the potential be written as  $V = \sum_{lmk} V_{lmk}$ . Then the mechanism for deriving transfer coefficients is explained for the  $x$  and  $z$  components:

$$\begin{aligned} \partial_x V_{lmk} &= \frac{1}{r} \frac{\partial V_{lmk}}{\partial u} = \frac{1}{r} \frac{\partial V_{lmk}}{\partial e^{i\psi_{mk}}} \frac{\partial e^{i\psi_{mk}}}{\partial u} = \frac{ik}{r} V_{lmk}, \\ \partial_z V_{lmk} &= \frac{\partial V_{lmk}}{\partial r} = \frac{\partial V_{lmk}}{\partial (R/r)^{l+1}} \frac{\partial (R/r)^{l+1}}{\partial r} = -\frac{l+1}{r} V_{lmk}. \end{aligned}$$

So the along-track component of the gradient,  $\partial_x V$ , will be characterized by a term  $ik/r$ , and the radial derivative by the usual  $-(l+1)/r$ .

The cross-track component requires special attention. The  $\theta'$ -coordinate is hidden in the inclination function  $\bar{F}_{lmk}(I)$ , (2.14). It is therefore convenient to introduce a cross-track derivative of the inclination function, denoted as  $\bar{F}_{lmk}^*(I)$ , cf. (Sneeuw, 1992):

$$\bar{F}_{lmk}^*(I) = -\frac{\partial \bar{F}_{lmk}(I)}{\partial \theta'} = i^{k-m+2} \bar{d}_{lmk}(I) \frac{d\bar{F}_{lk}(\cos \theta')}{d\theta'} \Big|_{\theta'=\pi/2}.$$

With the parameter  $x = \cos \theta$  the derivatives are:  $\frac{d\bar{P}_{lk}(x)}{dx} = -\frac{d\bar{P}_{lk}(\cos \theta)}{\sin \theta d\theta}$ . At the equator ( $\theta = \pi/2$ , or  $x = 0$ ) no confusion about the  $\sin \theta$  factor can arise. Let the derivative with respect to  $x$  be simply called  $\bar{P}'_{lk}(0)$ , then the *cross-track inclination function* is defined as:

$$\bar{F}_{lmk}^*(I) = i^{k-m} \bar{d}_{lmk}(I) \bar{P}'_{lk}(0). \quad (4.1)$$

When applying e.g. recursion (Z.1.30) from (Ilk, 1983) for unnormalized Legendre functions to the equator, one obtains:

$$(1-x^2) \frac{dP_{lk}(x)}{dx} = \sqrt{1-x^2} P_{l,k+1}(x) - kx P_{lk}(x) \quad \Rightarrow \quad P'_{lk}(0) = P_{l,k+1}(0). \quad (4.2)$$

So the derivative  $\bar{P}'_{lk}$  will be an even function for  $l-k$  odd and an odd one for  $l-k$  even. Thus the cross-track inclination functions will vanish for  $l-k$  even. This would allow the introduction of a Kaula-like cross-track inclination function  $\bar{F}_{lmp}^*(I)$ , with  $p = \frac{1}{2}(l-k-1)$ , cf. (Betti & Sansò, 1989; Koop, 1993).

Other approaches, circumventing the introduction of  $\bar{F}_{lmk}^*(I)$ , exist. Colombo (1986) suggested as cross-track derivative the expression  $(r \sin u)^{-1} \frac{\partial}{\partial I}$ , which shows singularities in  $u$ . See also (Betti & Sansò, 1989; Rummel et al., 1993, A.3.2). Depending on coordinate choice, better worked out in (Koop, 1993) or (Balmino, Schrama & Sneeuw, 1996), other expressions can be derived, e.g. the following singular one:  $(r \cos u \sin I)^{-1} \left( \cos I \frac{\partial}{\partial u} - \frac{\partial}{\partial \Lambda} \right)$ . By multiplying the former by  $\sin^2 u$ , the latter by  $\cos^2 u$  and adding the result, Schrama (1989) derived the *regular* expression:

$$\frac{\partial}{\partial y} = \frac{1}{r} \left[ \sin u \frac{\partial}{\partial I} + \frac{\cos u}{\sin I} \left( \cos I \frac{\partial}{\partial u} - \frac{\partial}{\partial \Lambda} \right) \right],$$

which leads to a corresponding cross-track inclination function:

$$\begin{aligned} \bar{F}_{lmk}^*(I) &= \frac{1}{2} \left[ \frac{(k-1) \cos I - m}{\sin I} \right] \bar{F}_{lm,k-1}(I) - \frac{1}{2} \bar{F}'_{lm,k-1}(I) + \\ &\frac{1}{2} \left[ \frac{(k+1) \cos I - m}{\sin I} \right] \bar{F}_{lm,k+1}(I) + \frac{1}{2} \bar{F}'_{lm,k+1}(I), \end{aligned} \quad (4.3)$$

where the primes denote differentiation with respect to inclination  $I$ .

Although numerical equivalence between the real version of (4.1) and (4.3) could be verified, it was proven analytically in (Balmino et al., 1996) that this last expression consists in fact of a twofold definition:

$$\bar{F}_{lmk}^*(I) = \left[ \frac{(k-1) \cos I - m}{\sin I} \right] \bar{F}_{lm,k-1}(I) - \bar{F}'_{lm,k-1}(I), \quad (4.4a)$$

$$\bar{F}_{lmk}^*(I) = \left[ \frac{(k+1) \cos I - m}{\sin I} \right] \bar{F}_{lm,k+1}(I) + \bar{F}'_{lm,k+1}(I). \quad (4.4b)$$

In summary, the spectral characteristics of the gradient operator in the local triad are given by the following transfer coefficients:

$$\partial_x \quad : \quad H_{lmk}^x = \frac{GM}{R^2} \left( \frac{R}{r} \right)^{l+2} \quad [ik] \quad \bar{F}_{lmk}(I) \quad (4.5a)$$

$$\partial_y \quad : \quad H_{lmk}^y = \frac{GM}{R^2} \left( \frac{R}{r} \right)^{l+2} [1] \quad \bar{F}_{lmk}^*(I) \quad (4.5b)$$

$$\partial_z \quad : \quad H_{lmk}^z = \frac{GM}{R^2} \left( \frac{R}{r} \right)^{l+2} [-(l+1)] \bar{F}_{lmk}(I) \quad (4.5c)$$

**Remark 4.1 (nomenclature)** *The different parts in these transfer coefficients will be denoted in the sequel as dimensioning term containing  $(GM, R)$ , upward continuation term (a power of  $R/r$ ), specific transfer and inclination function part. Especially the specific transfer is characteristic for a given observable.*

According to this nomenclature, the specific transfer of the potential is 1, cf. equation (2.16b). Both  $H_{lmk}^x$  and  $H_{lmk}^z$  show a transfer of  $\mathcal{O}(l, k)$  which is specific to first derivatives in general. Higher frequencies are amplified. The same holds true for  $H_{lmk}^y$ , though hidden in  $\bar{F}_{lmk}^*(I)$ . Equations (4.4) indicate already that  $\bar{F}_{lmk}^*(I) \sim \mathcal{O}(l, k) \times \bar{F}_{lmk}(I)$ . This becomes clearer for the second cross-track derivative, cf. next section. Note also that only the radial derivative is isotropic, i.e. only depends on degree  $l$ . Its specific transfer is invariant under rotations of the coordinate system like (2.9). This is not the case for  $V_x$  and  $V_y$ , when considered as scalar fields.

## 4.2 Second Derivatives: the Gravity Gradient Tensor

**I**N CONTRAST to the first derivatives, the second derivatives of the geopotential field are observable quantities. The observation of these is called *gravity gradiometry*, whose technical realization is described e.g. in (Rummel, 1986a). For a historical overview of measurement principles and proposed satellite gradiometer missions, refer to (Forward, 1973; Rummel, 1986b).

The gravity gradient tensor of second derivatives, or Hesse matrix, reads:

$$\mathbf{V} = \begin{pmatrix} V_{xx} & V_{xy} & V_{xz} \\ V_{yx} & V_{yy} & V_{yz} \\ V_{zx} & V_{zy} & V_{zz} \end{pmatrix}. \quad (4.6)$$

The sub-indices denote differentiation with respect to the specified coordinates. The tensor  $\mathbf{V}$  is symmetric. Due to Laplace's equation  $\Delta V = V_{xx} + V_{yy} + V_{zz} = 0$ , it is also trace-free. In local spherical coordinates  $(r, u, \theta')$  the tensor can be expressed as, e.g. (Koop, 1993, eqn. (3.10)):

$$\mathbf{V} = \begin{pmatrix} \frac{1}{r^2} V_{uu} + \frac{1}{r} V_r & -\frac{1}{r^2} V_{\theta'u} & \frac{1}{r} V_{ur} - \frac{1}{r^2} V_u \\ & \frac{1}{r^2} V_{\theta'\theta'} + \frac{1}{r} V_r & -\frac{1}{r} V_{\theta'r} + \frac{1}{r^2} V_{\theta'} \\ \text{symm.} & & V_{rr} \end{pmatrix}. \quad (4.7)$$

Again, use has been made of the fact that the satellite is always on the rotated equator  $\theta' = \frac{1}{2}\pi$ . With Laplace's equation one can avoid a second differentiation with respect to the  $\theta'$ -coordinate by writing:

$$V_{yy} = -V_{xx} - V_{zz} = -\frac{1}{r^2} V_{uu} - \frac{1}{r} V_r - V_{rr}.$$

As usual, the purely radial derivative is the simplest one. It is spectrally characterized by:  $(l+1)(l+2)/r^2$ . The operator  $\partial_{xx}$  will return the term:  $-[k^2 + (l+1)]/r^2$ . The second cross-track derivative  $\partial_{yy}$  thus gives with Laplace  $[k^2 + (l+1) - (l+1)(l+2)]/r^2 = [k^2 - (l+1)^2]/r^2$ . The *spectral transfer* for  $\partial_{xz}$  becomes:  $[-ik(l+1) - ik]/r^2 = -ik(l+2)/r^2$ . The components  $V_{xy}$  and  $V_{yz}$  make use of  $\partial_{\theta'}$ , which requires the use of  $\bar{F}_{lmk}^*(I)$  again. Starting from the expression for  $V_y$ , one further  $ik/r$ -term is required to obtain  $V_{xy}$ . For  $V_{yz}$  one needs an extra  $[-(l+1) - 1]/r = -(l+2)/r$ . The full set of transfer coefficients, describing the single components of the gravity gradient tensor. is thus given by:

$$\partial_{xx} \quad : \quad H_{lmk}^{xx} = \frac{GM}{R^3} \left(\frac{R}{r}\right)^{l+3} [-(k^2 + l + 1)] \bar{F}_{lmk}(I) \quad (4.8a)$$

$$\partial_{yy} \quad : \quad H_{lmk}^{yy} = \frac{GM}{R^3} \left(\frac{R}{r}\right)^{l+3} [k^2 - (l + 1)^2] \bar{F}_{lmk}(I) \quad (4.8b)$$

$$\partial_{zz} \quad : \quad H_{lmk}^{zz} = \frac{GM}{R^3} \left(\frac{R}{r}\right)^{l+3} [(l + 1)(l + 2)] \bar{F}_{lmk}(I) \quad (4.8c)$$

$$\partial_{xy} \quad : \quad H_{lmk}^{xy} = \frac{GM}{R^3} \left(\frac{R}{r}\right)^{l+3} [ik] \bar{F}_{lmk}^*(I) \quad (4.8d)$$

$$\partial_{xz} \quad : \quad H_{lmk}^{xz} = \frac{GM}{R^3} \left(\frac{R}{r}\right)^{l+3} [-ik(l + 2)] \bar{F}_{lmk}(I) \quad (4.8e)$$

$$\partial_{yz} \quad : \quad H_{lmk}^{yz} = \frac{GM}{R^3} \left(\frac{R}{r}\right)^{l+3} [-(l + 2)] \bar{F}_{lmk}^*(I) \quad (4.8f)$$

The specific transfer is of order  $\mathcal{O}(l^2, lk, k^2)$ , as can be expected for second derivatives. This is also true for  $H_{lmk}^{xy}$  and  $H_{lmk}^{yz}$ , that make use of  $\bar{F}_{lmk}^*(I)$ . Again, the purely radial derivative is the only isotropic component. Adding the specific transfers of the diagonal components yields the Laplace equation in the spectral domain:

$$-(k^2 + l + 1) + k^2 - (l + 1)^2 + (l + 1)(l + 2) = 0.$$

**Cross-Track Gravity Gradient.** An alternative derivation of  $V_{yy}$  could have been obtained directly, i.e. without the Laplace equation, by a second cross-track differentiation. A new inclination function, say  $\bar{F}_{lmk}^{**}(I)$  is required, defined as:

$$\bar{F}_{lmk}^{**}(I) = \frac{\partial^2 \bar{F}_{lmk}(I)}{\partial \theta'^2} = i^{k-m} \bar{d}_{lmk}(I) \bar{P}_{lk}''(0).$$

Now, from recursions (Z.1.38) and (Z.1.44) from (Ilk, 1983, App.), we have for the second latitudinal derivative of the unnormalized Legendre function at the equator:

$$P_{lk}''(0) = P_{l,k+2}(0) - kP_{lk}(0).$$

Together with recursion (Z.1.18):  $P_{l,k+2}(0) = -(l+k+1)(l-k)P_{lk}(0)$ , one obtains:

$$P_{lk}''(0) = [k^2 - l(l+1)]P_{lk}(0).$$

A normalized version of this expression must be inserted in the definition of  $\bar{F}_{lmk}^{**}(I)$  above, yielding the specific transfer  $[k^2 - l(l+1)]$  of the second cross-track derivative  $V_{\theta'\theta'}$ . Since  $V_{yy} = V_{\theta'\theta'}/r^2 + V_r/r$  one ends up with exactly the same

transfer, as derived above with the Laplace equation, namely  $[k^2 - (l + 1)^2]/r^2$ . Moreover, it demonstrates again that  $\bar{F}_{lmk}^*(I)$  is of order  $\mathcal{O}(l, k)$ , since the second cross-track derivative has a transfer of  $\mathcal{O}(l^2, lk, k^2)$ .

**Space-Stable Gradiometry.** The transfer coefficients (4.8) pertain to tensor components in the local triad. Especially for local-level orientations, such as Earth-pointing, these expressions are useful. In principle another orientation can be deduced from them, since a tensor  $\mathbf{V}$  is transformed into another coordinate system by:

$$\mathbf{V}' = R\mathbf{V}R^T,$$

cf. (Koop, 1993), in which  $R$  is the rotation matrix between the two systems. For instance the rotation sequence

$$R = R_z(-\Lambda)R_x(-I)R_z(-u),$$

which is the inverse of the rotations from 2.2, may be used to transform the gravity gradient tensor back into an Earth-fixed reference frame. Note, however, that the angles  $u$  and  $\Lambda$  are time-dependent. The derivation of transfer functions becomes cumbersome. An alternative approach, based on the work of Hotine (1969), is followed by Ilk (1983) and Bettadpur (1991, 1995).

### 4.3 Orbit Perturbations

THE NON-CENTRAL gravity field disturbs the pure Kepler orbit. Thus orbit perturbations convey gravity field information, i.e. they are functionals of the gravitational potential as well. In order to derive their transfer coefficients, a dynamic model of satellite motion is required. In this work orbit perturbations refer to a description of the deviations in the local orbital frame. These deviations are described by the linearized *Hill equations*, cf. 3.1. The Hill equations with harmonic force term read:

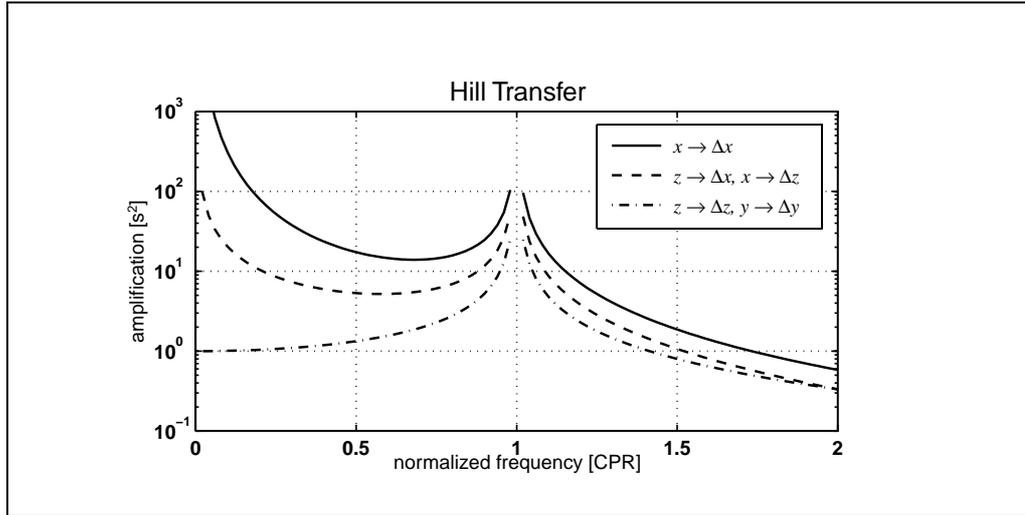
$$\begin{aligned} \ddot{x} + 2n\dot{z} &= f_x = A_x e^{i\omega t} \\ \ddot{y} + n^2 y &= f_y = A_y e^{i\omega t} \\ \ddot{z} - 2n\dot{x} - 3n^2 z &= f_z = A_z e^{i\omega t} \end{aligned} \quad (4.9)$$

with  $n$  the natural orbit frequency (from Kepler's third law  $n^2 r^3 = GM$ ),  $\omega$  the disturbing frequency (not to be confused with the argument of perigee) and  $A_x, A_y, A_z$  amplitudes of the disturbing forces in  $x, y, z$  direction. The out-of-plane equation represents a harmonic oscillator, whereas the in-plane equations for  $x$  and  $z$  are coupled. Strictly speaking, the equations—in particular the orbital rate  $n$ —do not pertain to the situation of the precessing nominal orbit, which includes  $J_2$ -effects:  $n \neq \dot{u}$ . The relative difference, however, is of order  $\mathcal{O}(J_2)$ , cf. equation (2.8d). Thus  $\dot{u}$  and  $n$  will be used interchangeably in the sequel.

The Hill equations can be solved analytically. Since we are mainly interested in the frequency response, a full solution is not required. Full solutions, including resonant and homogeneous parts, may be found in (Scheinert, 1996). The frequency response is given by the particular solution of (4.9) and visualized in

fig. 4.1:

$$\begin{aligned}
 x(t) &= \frac{\omega^2 + 3n^2}{\omega^2(n^2 - \omega^2)} f_x + \frac{2in}{\omega(n^2 - \omega^2)} f_z = \frac{2A_z i \omega n + A_x(\omega^2 + 3n^2)}{\omega^2(n^2 - \omega^2)} e^{i\omega t} \\
 y(t) &= \frac{1}{n^2 - \omega^2} f_y = \frac{A_y}{n^2 - \omega^2} e^{i\omega t} \\
 z(t) &= \frac{-2in}{\omega(n^2 - \omega^2)} f_x + \frac{1}{n^2 - \omega^2} f_z = \frac{A_z \omega - 2in A_x}{\omega(n^2 - \omega^2)} e^{i\omega t}
 \end{aligned} \tag{4.10}$$



**Figure 4.1:** Spectral transfer of Hill equations with harmonic disturbance. The legend explains how the forcing terms  $f_x, f_y, f_z$  are transferred onto the orbit perturbations  $\Delta x, \Delta y, \Delta z$ .

**Remark 4.2 (resonance)** Notice the occurrence of a resonant response to disturbing forces at the zero frequency ( $\omega = 0$ , or DC) and at the natural frequency ( $|\omega| = n$ ). These frequencies have to be discarded from our analyses. Since the linearized Hill equations are employed, this is a critical issue. The magnitude of any resonant disturbance endangers the validity of the HE.

In practice the satellite motion can be disturbed at any frequency. Since the dynamical system (4.9) is linear, the full solution will therefore be a Fourier series of particular solutions (4.10).

Now all ingredients for deriving the transfer coefficients of orbit perturbations are available. The disturbing frequencies  $\omega$  become  $\dot{\psi}_{mk} = k\dot{u} + m\dot{\Lambda}$ . The disturbing force is  $\nabla V$ , so the amplitudes  $A_x, A_y, A_z$  come from (4.5). The specific transfer of the dynamics is directly taken from (4.10).

$$\begin{aligned}
 H_{lmk}^{\Delta x} &= \frac{2(l+1)\dot{\psi}_{mk}n - k(\dot{\psi}_{mk}^2 + 3n^2)}{\dot{\psi}_{mk}^2(\dot{\psi}_{mk}^2 - n^2)} i \frac{GM}{R^2} \left(\frac{R}{r}\right)^{l+2} \bar{F}_{lmk}(I) \\
 H_{lmk}^{\Delta y} &= \frac{1}{n^2 - \dot{\psi}_{mk}^2} \frac{GM}{R^2} \left(\frac{R}{r}\right)^{l+2} \bar{F}_{lmk}^*(I) \\
 H_{lmk}^{\Delta z} &= \frac{(l+1)\dot{\psi}_{mk} - 2kn}{\dot{\psi}_{mk}(\dot{\psi}_{mk}^2 - n^2)} \frac{GM}{R^2} \left(\frac{R}{r}\right)^{l+2} \bar{F}_{lmk}(I)
 \end{aligned}$$

The  $\Delta$ 's are added in the labels to discern orbit perturbations from derivatives. For a further simplification normalized frequencies are introduced, i.e.  $\beta_{mk} = \dot{\psi}_{mk}/n$ . Moreover Kepler's third law can be inserted in order to remove the  $n^2$  terms:

$$\frac{1}{n^2} \frac{GM}{R^2} = \frac{r^3}{GM} \frac{GM}{R^2} = R \left( \frac{R}{r} \right)^{-3}.$$

Consequently the dimensioning term becomes  $R$  and the power of the upward continuation term  $l - 1$ , yielding:

$$\Delta x \quad : \quad H_{lmk}^{\Delta x} = R \left( \frac{R}{r} \right)^{l-1} \left[ i \frac{2(l+1)\beta_{mk} - k(\beta_{mk}^2 + 3)}{\beta_{mk}^2(\beta_{mk}^2 - 1)} \right] \bar{F}_{lmk}(I) \quad (4.11a)$$

$$\Delta y \quad : \quad H_{lmk}^{\Delta y} = R \left( \frac{R}{r} \right)^{l-1} \left[ \frac{1}{1 - \beta_{mk}^2} \right] \bar{F}_{lmk}^*(I) \quad (4.11b)$$

$$\Delta z \quad : \quad H_{lmk}^{\Delta z} = R \left( \frac{R}{r} \right)^{l-1} \left[ \frac{(l+1)\beta_{mk} - 2k}{\beta_{mk}(\beta_{mk}^2 - 1)} \right] \bar{F}_{lmk}(I) \quad (4.11c)$$

Note that in terms of normalized frequency resonance would now occur at  $\beta_{mk} = -1, 0, +1$ . The 'unit' is cycles per revolution (CPR).

**Remark 4.3** *The transfer coefficients, derived from the Hill equations, are consistent with those from other linear perturbation theories up to order zero in eccentricity, e.g. (Rosborough & Tapley, 1987). Analytical equivalence between the  $H_{lmk}^{\Delta z}$  from HE and the one from the linear solution of the Lagrange Planetary Equations was shown in (Schrama, 1989). This was extended to the other components in (Balmino et al., 1996), see also (Balmino, 1993).*

## 4.4 Low-Low Intersatellite Range Perturbation

THE CONCEPT of continuously tracking the distance between orbiting spacecraft for purposes of gravity field determination dates back to the early space era, e.g. (Wolff, 1969; Rummel, Reigber & Ilk, 1978). A historical overview of proofs-of-concept and of proposed missions is given by Wakker (1988). Global geopotential recovery capability has conventionally been studied by means of Kaula's linear perturbation theory, cf. (Kaula, 1983; Wagner, 1983; Schrama, 1986; Sharma, 1995). Also Hill equations have been applied to this end, e.g. (Colombo, 1984; Mackenzie & Moore, 1997). For regional approaches to geopotential recovery, see e.g. (Thalhammer, 1995).

Satellite-to-satellite tracking (SST) is discussed in two modes: high-low SST, in which one satellite flies in a high orbit, the other in a low one, and low-low SST with two LEO's. One way to realize the high-low mode is space-borne GPS-tracking, e.g. (Jekeli & Upadhyay, 1990). However, this concept may as well be considered as 3D orbit tracking of the LEO. In this view the formulae of 4.3 can be applied directly. No effort will be made here to derive transfer coefficients for high-low SST observables.

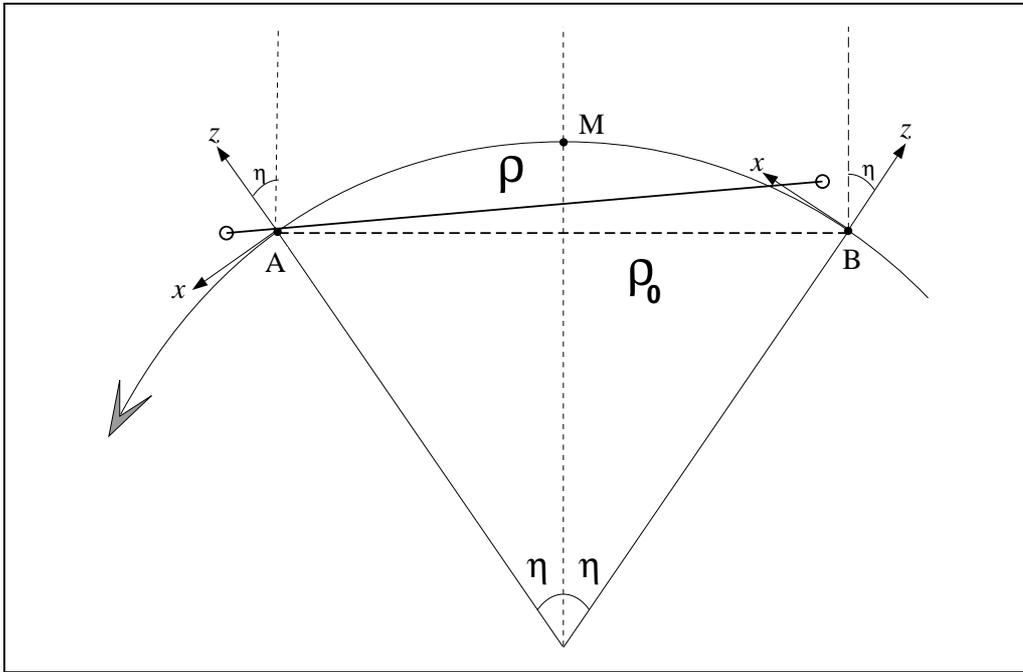
The low-low mode can be realized either by two LEO's on the same nominal orbit, separated in argument of latitude  $u$ , or by two LEO's on two different (but close) nominal orbits with separations in  $\Omega$  and/or  $I$ . The latter type of low-low SST

is described e.g. by Wagner (1983), introducing an average nominal orbit. In (Mackenzie & Moore, 1997) transfer coefficients of various low-low SST options are derived in detail, making use of HE. Since measurements should be made continuously, a height separation between orbits is not considered here. The LEO's would drift apart and lose intervisibility.

In order to demonstrate the principle, the transfer coefficient  $H_{lmk}^{\Delta\rho}$  for the low-low SST observable with both satellites on the same nominal orbit is derived now. Given is a satellite pair, A and B, on the same nominal quasi-circular orbit. The intersatellite distance in geocentric coordinates is:

$$\rho(t) = |\underline{r}_A - \underline{r}_B| = |\underline{r}(t + \tau) - \underline{r}(t - \tau)|,$$

in which the time tag  $t$  refers to the location M, cf. fig. 4.2 and (Colombo, 1984). The nominal separation is  $\rho_0 = 2r_0 \sin \eta$ . The time lag  $\tau$  is connected to the angle  $\eta$  by the angular orbit frequency  $n$ :  $\tau = \eta/n$ . In linear approximation,



**Figure 4.2:** *Low-low SST configuration.*

orbit perturbations  $\Delta \underline{x}_A$  and  $\Delta \underline{x}_B$  couple into the line-of-sight as:

$$\begin{aligned} \Delta \rho(t) &= (\Delta x_A - \Delta x_B) \cos \eta + (\Delta z_A + \Delta z_B) \sin \eta \\ &= (\Delta x(t + \tau) - \Delta x(t - \tau)) \cos \eta + (\Delta z(t + \tau) + \Delta z(t - \tau)) \sin \eta, \end{aligned} \quad (4.12)$$

with  $\Delta \rho = \rho - \rho_0$ . The  $\cos \eta$  and  $\sin \eta$  are due to the fact that the local coordinate system at B is rotated by the angle  $2\eta$  with respect to that of A. The major range perturbations comes from the along-track perturbation difference. The radial perturbations only project onto  $\Delta \rho$  due to the rotation. Cross-track perturbations do not show up in the linear framework with both satellites flying on the same nominal orbit.

With expressions (3.1a) and (3.1b), gravitational periodic orbit perturbations can

be expressed by:

$$\Delta x(t) = \sum_{l,m,k} H_{lmk}^{\Delta x} K_{lm} e^{i\dot{\psi}_{mk}t}, \text{ and } \Delta z(t) = \sum_{l,m,k} H_{lmk}^{\Delta z} K_{lm} e^{i\dot{\psi}_{mk}t}. \quad (4.13)$$

Inserting (4.13) into (4.12) results in:

$$\begin{aligned} \Delta x(t + \tau) - \Delta x(t - \tau) &= \sum_{l,m,k} H_{lmk}^{\Delta x} K_{lm} \left[ e^{i\dot{\psi}_{mk}(t + \tau)} - e^{i\dot{\psi}_{mk}(t - \tau)} \right], \\ \Delta z(t + \tau) + \Delta z(t - \tau) &= \sum_{l,m,k} H_{lmk}^{\Delta z} K_{lm} \left[ e^{i\dot{\psi}_{mk}(t + \tau)} + e^{i\dot{\psi}_{mk}(t - \tau)} \right]. \end{aligned}$$

Reformulating the exponentials yields:

$$e^{i\dot{\psi}_{mk}(t \pm \tau)} = e^{i\dot{\psi}_{mk}t} e^{\pm i\dot{\psi}_{mk}\tau} = e^{i\dot{\psi}_{mk}t} e^{\pm i\eta\beta_{mk}}.$$

Consequently:

$$\begin{aligned} e^{i\dot{\psi}_{mk}(t + \tau)} - e^{i\dot{\psi}_{mk}(t - \tau)} &= e^{i\dot{\psi}_{mk}t} (e^{i\eta\beta_{mk}} - e^{-i\eta\beta_{mk}}) \\ &= e^{i\dot{\psi}_{mk}t} 2i \sin(\eta\beta_{mk}), \end{aligned} \quad (4.14a)$$

$$e^{i\dot{\psi}_{mk}(t + \tau)} + e^{i\dot{\psi}_{mk}(t - \tau)} = e^{i\dot{\psi}_{mk}t} \quad (4.14b)$$

$$= e^{i\dot{\psi}_{mk}t} 2 \cos(\eta\beta_{mk}). \quad (4.14c)$$

Now we can combine all equations in order to obtain an expression for the transfer coefficient  $H_{lmk}^{\Delta\rho}$ . It is:

$$\Delta\rho(t) = \sum_{l,m,k} H_{lmk}^{\Delta\rho} K_{lm} e^{i\dot{\psi}_{mk}t}, \quad (4.15)$$

$$H_{lmk}^{\Delta\rho} = 2i \cos \eta \sin(\eta\beta_{mk}) H_{lmk}^{\Delta x} + 2 \sin \eta \cos(\eta\beta_{mk}) H_{lmk}^{\Delta z}. \quad (4.16)$$

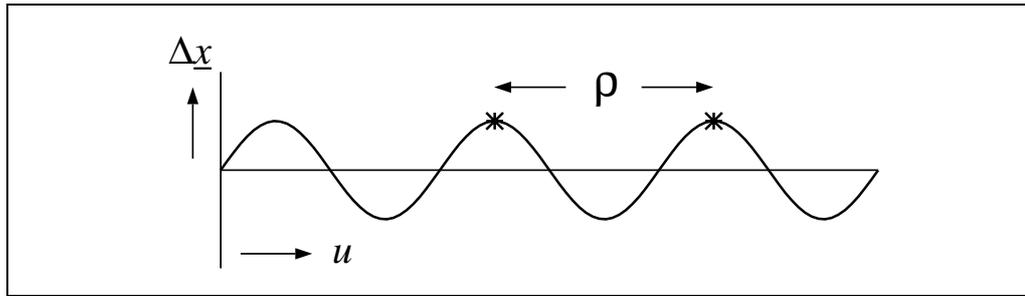
This result is not expanded further, to prevent long formulae. Note that  $H_{lmk}^{\Delta\rho}$  is a linear combination of  $H_{lmk}^{\Delta x}$  and  $H_{lmk}^{\Delta z}$  with dimensionless coefficients. So  $H_{lmk}^{\Delta\rho}$  shares the dimension ( $R$ ) and the upward continuation power ( $l - 1$ ) with its constituents. Moreover, it makes use of the ordinary  $\bar{F}_{lmk}(I)$ .

An approximation can be made by considering that  $\eta$  is usually small. In that case, replacing  $\cos \eta \rightarrow 1$  and  $\sin \eta \rightarrow \eta$ , the transfer coefficient becomes:

$$H_{lmk}^{\Delta\rho} \approx 2i \sin(\eta\beta_{mk}) H_{lmk}^{\Delta x}, \quad (4.17)$$

expressing the fact that the low-low SST observable is mainly a scaled version of the along-track orbit perturbation. For baselines, say, up to  $\rho_0 = 400$  km, the error is only a few percent. The same transfer coefficients are derived in (Wagner, 1983), except for an out-of-plane contribution, and in (Sharma, 1995). A minor difference with the latter is the fact that the  $\sin(\eta\beta_{mk})$ -term shows up as  $\sin(k\eta)$  there. Indeed  $\beta_{mk} = k + m\frac{\dot{A}}{u} \approx k$ , since the frequency ratio for LEO's is about 0.06.

In principle,  $H_{lmk}^{\Delta\rho}$  inherits the resonances from  $H_{lmk}^{\Delta x}$  (and  $H_{lmk}^{\Delta z}$ ). However, the factor  $\sin(\eta\beta_{mk})$ , characteristic for  $\Delta\rho$ , plays an interesting role. For one, it reduces the DC resonance in  $\Delta x$ . In this resonance case the radial perturbation part cannot be neglected anymore. However, in general it can not be assumed that  $\beta_{mk}$  is small. Depending on the separation angle the term  $\sin(\eta\beta_{mk})$  might even become close to zero. In that case it *attenuates* the along-track contribution. As an example, suppose  $\eta = 4^\circ$  and let  $\beta_{mk}$  be approximated by  $k$  again. Then, apart from the case  $k = 0$ , the first attenuation takes place at  $k = 45$ . This implies that the spherical harmonics above  $l = 45$  will possess certain spectral components, that cancel out in the signal  $\Delta\rho$ . Seen alternatively, this is a *common-mode* effect, by which certain frequencies in  $\Delta x$  show up in the same way at locations A and B. This situation is exemplified in fig. 4.3.



**Figure 4.3:** Common mode perturbation at specific wavelength. After (Wolff, 1969, fig. 2).

If one wants to avoid this cancellation of signal, the separation angle should be chosen small enough. The attenuation gets close to zero when  $\beta_{mk}$  is near  $i\pi/\eta$ ,  $i = 0, 1, 2, \dots$ . The DC case ( $i=0$ ) cannot be avoided. But avoiding the first occurrence, i.e.  $\beta_{mk} \approx k < \pi/\eta$ , implies that for a given maximum degree  $L$  the separation should be  $\eta < \pi/L$ . For a gravity field recovery up to degree  $L=90$  e.g.  $\eta$  should be smaller than  $2^\circ$ , equivalent to  $\rho_0 < 450$  km. See also (Mackenzie & Moore, 1997).

**Remark 4.4** The attenuation  $\sin(\eta\beta_{mk})$  does not imply that SH coefficients with degree  $l > \pi/\eta$  cannot be determined in general. Attenuation means that some spectral components of  $\Delta x$  are suppressed in  $\Delta\rho$ . Information about this part of the gravity field is lost.

## 4.5 Time Derivatives

THE TRANSFER coefficients of the fundamental observables of SGG and SST have been presented. However, time derivatives of these functionals may be observable quantities as well, in particular range-rate and range-acceleration perturbations. In this section, the transfer coefficient of the time derivative of a generic functional is derived from its transfer coefficient itself. Let the time-series of the functional  $f^\#$  be written as

$$f^\#(t) = \sum_{l,m,k} H_{lmk}^\# K_{lm} e^{i\psi_{mk}t}. \quad (4.18)$$

Its time-derivative, or *rate*, is simply obtained through differentiation of the exponential, the only term containing time  $t$ . The transfer coefficients of  $\dot{f}^\# = \frac{df^\#}{dt}$  and of  $\ddot{f}^\# = \frac{d^2f^\#}{dt^2}$  become:

$$\begin{aligned} H_{lmk}^{\dot{\#}} &= i\dot{\psi}_{mk} H_{lmk}^\# = in\beta_{mk} H_{lmk}^\#, \\ H_{lmk}^{\ddot{\#}} &= -\dot{\psi}_{mk}^2 H_{lmk}^\# = -n^2\beta_{mk}^2 H_{lmk}^\#. \end{aligned} \quad (4.19)$$

The part  $i\beta_{mk}$  is multiplied to the specific transfer. It expresses high frequency amplification, due to differentiation, cf. the *spatial* derivative (4.5a). The factor  $n$ , though, becomes part of the dimensioning and upward continuation term, since:

$$n = \sqrt{\frac{GM}{r^3}} = \left(\frac{R}{r}\right)^{3/2} \sqrt{\frac{GM}{R^3}}.$$

Equations (4.19) are generally valid. However, they will be useful especially when applied to the perturbations, yielding orbit perturbation rates and range perturbation rates (or just range rates for short), and corresponding accelerations.

**Dimensioning.** A logical choice for dimensioning the rates would be the factor:

$$v_R = n_R R = \sqrt{\frac{GM}{R^3}} R = \sqrt{\frac{GM}{R}}, \quad (4.20)$$

i.e. the linear speed of a satellite at zero altitude (provided a perfectly spherical Earth of radius  $R$  without an atmosphere). The orbital rate  $n_R$  corresponds to a period of nearly 84.5 minutes. The dimensioning of the perturbations ( $R$ ) is multiplied by  $n$ , due to the time-derivative. In order to obtain  $n_R R$  we have to multiply further by

$$\frac{n}{n_R} = \left(\frac{R}{r}\right)^{3/2}.$$

This factor will be incorporated in the upward continuation term. Accelerations will be dimensioned by  $\frac{GM}{R^2}$  of course. The factor  $R$  from the perturbations gets multiplied by  $n^2 = \frac{GM}{r^3}$ . In order to obtain  $\frac{GM}{R^2}$ , a further multiplication by  $(R/r)^3$  is required. Again, this factor will be incorporated in the upward continuation term.

**Remark 4.5** *The transfer coefficients of the range rate (and similarly of the range acceleration) are either obtained by applying  $in\beta_{mk}$  to  $H_{lmk}^{\Delta\rho}$ , or by writing the same linear combination as in (4.16) with  $H_{lmk}^{\Delta\dot{x}}$  and  $H_{lmk}^{\Delta\dot{z}}$ .*

## 4.6 Gradiometry by Accelerometric Low-Low SST

THE LOW-LOW range acceleration  $\Delta\ddot{\rho}$  is related to the along-track gradiometric observable  $V_{xx}$ . Apart from the distance between the two proof masses, the major difference between these two observables is the fact that  $\Delta\ddot{\rho}$  represents a kinematic and  $V_{xx}$  a dynamic acceleration difference. *Kinematic* means here that the two satellites are free falling objects, between which the second time-derivative of its separation distance is observed. Although this range

is the result of dynamics (e.g. Hill equations) and forcing functions ( $\nabla V$ ), the observable itself is a kinematic quantity. *Dynamic* refers to the fact that the observable is a (differential) *specific force*. The proof masses are constrained, since their position changes are nullified within the gradiometer.

Nevertheless, the transfer coefficients of both observables can be related to each other under certain approximations. To this purpose we will investigate the quantity

$$\frac{\Delta\ddot{\rho}}{\rho_0}, \text{ and its transfer } H_{lmk}^{\Delta\ddot{\rho}/\rho_0} = \frac{H_{lmk}^{\Delta\ddot{\rho}}}{\rho_0},$$

for small  $\rho_0$ , and see whether this can be simplified to  $H_{lmk}^{xx}$ . Applying (4.19) to (4.16) yields:

$$\begin{aligned} H_{lmk}^{\Delta\ddot{\rho}} &= -n^2 \beta_{mk}^2 H_{lmk}^{\Delta\rho} \\ &= -2n^2 \beta_{mk}^2 \left[ i \cos \eta \sin(\eta \beta_{mk}) H_{lmk}^{\Delta x} + \sin \eta \cos(\eta \beta_{mk}) H_{lmk}^{\Delta z} \right]. \end{aligned}$$

For very small baselines ( $\rho_0 \rightarrow 0$ ) the following transitions are justified:

$$\begin{aligned} \cos \eta &\rightarrow 1, \quad \cos \eta \beta_{mk} \rightarrow 1 \\ \sin \eta &\rightarrow \eta, \quad \sin \eta \beta_{mk} \rightarrow \eta \beta_{mk} \\ \rho_0 &\rightarrow 2r\eta. \end{aligned}$$

Thus, the following approximation holds:

$$\begin{aligned} H_{lmk}^{\Delta\ddot{\rho}} &\approx -2n^2 \beta_{mk}^2 \left[ i\eta \beta_{mk} H_{lmk}^{\Delta x} + \eta H_{lmk}^{\Delta z} \right] \\ &\approx -2n^2 \beta_{mk}^2 \eta \left[ \frac{-(l+1)\beta_{mk} + k + k\beta_{mk}^2}{\beta_{mk}(\beta_{mk}^2 - 1)} \right] R \left( \frac{R}{r} \right)^{l-1} \bar{F}_{lmk}(I). \end{aligned} \quad (4.21)$$

Upon using  $n^2 = GM/r^3$  and dividing by  $\rho_0 = 2r\eta$ , one obtains

$$\frac{H_{lmk}^{\Delta\ddot{\rho}}}{2r\eta} \approx \frac{GM}{r^4} R \left( \frac{R}{r} \right)^{l-1} \left[ \frac{(l+1)\beta_{mk}^2 - k\beta_{mk}(1 + \beta_{mk}^2)}{\beta_{mk}^2 - 1} \right] \bar{F}_{lmk}(I). \quad (4.22)$$

By rearranging the terms with  $r$  and  $R$ , the dimensioning and the upward continuation reduce to  $GM/R^3$  and  $(R/r)^{l+3}$  respectively, which are usual for gradiometry. More interesting, though, is the fact that the RHS of (4.22) does not contain  $\eta$ . A further simplification is achieved by developing the denominator in series:

$$\frac{\beta^2}{\beta^2 - 1} = 1 + \frac{1}{\beta^2} + \mathcal{O}\left(\frac{1}{\beta^4}\right), \text{ and } \frac{\beta^2 + 1}{\beta^2 - 1} = 1 + \frac{2}{\beta^2} + \mathcal{O}\left(\frac{1}{\beta^4}\right).$$

Insertion of these series in (4.22), while neglecting terms in  $\beta_{mk}^{-2}$ , leads to

$$\frac{H_{lmk}^{\Delta\ddot{\rho}}}{\rho_0} \approx \frac{GM}{R^3} \left( \frac{R}{r} \right)^{l+3} \left[ (l+1) - k\beta_{mk} - \frac{2k}{\beta_{mk}} \right] \bar{F}_{lmk}(I).$$

Finally, employing the approximation  $\beta_{mk} \approx k$ , cf. 4.4, one arrives at

$$\frac{H_{lmk}^{\Delta\ddot{\rho}}}{\rho_0} \approx \frac{GM}{R^3} \left( \frac{R}{r} \right)^{l+3} \left[ l - 1 - k^2 \right] \bar{F}_{lmk}(I). \quad (4.23)$$

Apart from the sign of  $l$ , this transfer coefficient equals  $H_{lmk}^{xx}$ , cf. (4.8). Thus, under the approximations made, the intersatellite range acceleration will closely resemble along-track gradiometry. It must be recalled, though, that (4.23) is only valid under the assumption that the baseline is sufficiently small, while  $\beta_{mk}$  may not be too close to zero, due to the above series development. The latter condition also excludes resonance.

## 4.7 Summary

THE SET of transfer coefficients, described in this chapter, constitute a pocket guide to gravity field functionals along a satellite orbit. They express the spectral transfer—or mapping—from SH coefficients onto lumped coefficients of the functional. They are summarized in tbl. 4.1. For each observable in the left column of this table, the corresponding transfer coefficient is the product of the dimensioning, the specific transfer, the upward continuation term and the appropriate inclination function. The dashes indicate the use of ordinary inclination functions  $\bar{F}_{lmk}(I)$ , the asterisks imply the use of  $\bar{F}_{lmk}^*(I)$ .

Some further conclusions:

- i) Cross-track inclination functions  $\bar{F}_{lmk}^*(I)$  have been introduced by considering the orbit as rotated equator. The cross-track direction is the latitude direction in the corresponding rotated system.
- ii) Transfer coefficients of orbit perturbations have been derived employing Hill equations as dynamic model. Using other dynamics, e.g. the Lagrange Planetary Equations, would have resulted in the same transfer.
- iii) Certain spectral components in the low-low SST signal will be attenuated, due to the term  $\sin \eta \beta_{mk}$  in  $H_{lmk}^{\Delta\rho}$ . This attenuation can be avoided by requiring  $\eta < \pi/L$ . Thus the in-orbit separation angle ( $2\eta$ ) between the two satellites must be smaller than  $2\pi$  divided by the maximum degree. That is, the intersatellite distance should be smaller than the smallest spatial feature to be resolved.
- iv) The intersatellite range acceleration is tightly connected to the along-track in-line gradiometry component. This was shown by simplification of  $H_{lmk}^{\Delta\tilde{\rho}}$  into  $H_{lmk}^{xx}$ .

In linear approximation it will be possible to derive  $H_{lmk}$ 's for any functional of the geopotential. Since formulae tend to grow out of hand, it can be useful to express certain transfer coefficients in terms of more basic building blocks, as was done e.g. for  $H_{lmk}^{\Delta\rho}$ . Examples of other quantities for which transfer coefficients may be derived are perturbations in:

- Kepler elements (Kaula, 1966)
- Hill variables (Cui & Lelgemann, 1995)
- polar coordinates (Exertier & Bonnefond, 1997)
- non-singular elements (Wnuk, 1988)

Table 4.1: Transfer coefficients of several functionals

#	dimensioning	specific transfer	upw. cont.	$\bar{F}$	eqn.
$V$	$GM/R$	1	$l + 1$	-	(2.16b)
$x$	$GM/R^2$	$ik$	$l + 2$	-	(4.5)
$y$		1		*	
$z$		$-(l + 1)$		-	
$xx$	$GM/R^3$	$-(k^2 + l + 1)$	$l + 3$	-	(4.8)
$yy$		$k^2 - (l + 1)^2$		-	
$zz$		$(l + 1)(l + 2)$		-	
$xy$		$ik$		*	
$xz$		$-ik(l + 2)$		-	
$yz$		$-(l + 2)$		*	
$\Delta x$	$R$	$i \frac{2(l + 1)\beta_{mk} - k(\beta_{mk}^2 + 3)}{\beta_{mk}^2(\beta_{mk}^2 - 1)}$	$l - 1$	-	(4.11)
$\Delta y$	$(=\frac{GM}{R^2 n^2})$	$\frac{1}{1 - \beta_{mk}^2}$		*	
$\Delta z$		$\frac{(l + 1)\beta_{mk} - 2k}{\beta_{mk}(\beta_{mk}^2 - 1)}$		-	
$\Delta \rho$		...		-	
$\Delta \dot{x}$	$v_R$	$-\frac{2(l + 1)\beta_{mk} - k(\beta_{mk}^2 + 3)}{\beta_{mk}(\beta_{mk}^2 - 1)}$	$l + \frac{1}{2}$	-	
$\Delta \dot{y}$	$(=\sqrt{\frac{GM}{R}})$	$i \frac{\beta_{mk}}{1 - \beta_{mk}^2}$		*	
$\Delta \dot{z}$		$i \frac{(l + 1)\beta_{mk} - 2k}{\beta_{mk}^2 - 1}$		-	
$\Delta \dot{\rho}$		...		-	
$\Delta \ddot{x}$	$GM/R^2$	$-i \frac{2(l + 1)\beta_{mk} - k(\beta_{mk}^2 + 3)}{\beta_{mk}^2 - 1}$	$l + 2$	-	
$\Delta \ddot{y}$		$\frac{\beta_{mk}^2}{\beta_{mk}^2 - 1}$		*	
$\Delta \ddot{z}$		$-\frac{(l + 1)\beta_{mk}^2 - 2k\beta_{mk}}{\beta_{mk}^2 - 1}$		-	

**Table 4.1:** *continued*

#	dimensioning	specific transfer	upw. cont.	$\bar{F}$	eqn.
$\Delta\ddot{\rho}$		...		-	
$\Delta\rho$		$2i \cos \eta \sin(\eta\beta_{mk}) H_{lmk}^{\Delta x} + 2 \sin \eta \cos(\eta\beta_{mk}) H_{lmk}^{\Delta z}$			(4.16)
$\Delta\dot{\rho}$		$2i \cos \eta \sin(\eta\beta_{mk}) H_{lmk}^{\Delta \dot{x}} + 2 \sin \eta \cos(\eta\beta_{mk}) H_{lmk}^{\Delta \dot{z}}$			(4.19)
$\Delta\ddot{\rho}$		$2i \cos \eta \sin(\eta\beta_{mk}) H_{lmk}^{\Delta \ddot{x}} + 2 \sin \eta \cos(\eta\beta_{mk}) H_{lmk}^{\Delta \ddot{z}}$			

## 5 A Spectral Analysis

IN CHAPTER 2 three basic representations of the geopotential were presented: as a function on the sphere, as time-series along the orbit, and as a lumped coefficient series. In 3 and 4 they were extended to all functionals  $f^\#$  of the geopotential. To each of these representations belongs a different type of spectrum: spherical harmonic, 1D-Fourier and 2D-Fourier, respectively. Schematically, gravity field recovery with the semi-analytical approach is represented by the following set of spectral transformations:

$$\left. \begin{array}{l} \text{signal: } f(t) \xrightarrow{i} f_n \xrightarrow{ii} f_{mk} \\ \text{noise: } n(t) \xrightarrow{iii} \text{PSD} \xrightarrow{iv} \sigma_{mk}^2 \end{array} \right\} \xrightarrow{v} \left\{ \begin{array}{l} f_{lm} \\ \sigma_{lm} \end{array} \right. \quad (5.1)$$

This chapter is devoted to the above mappings, in particular to spectral considerations. The following issues will be discussed for the individual mappings:

- i) This is a 1D Fourier transformation from signal time-series to 1D Fourier spectrum. Under the assumption of a *repeat orbit* the signal becomes periodic and the spectrum discrete. *Aliasing* may occur in case of undersampling. The required sampling rate is discussed in 5.4.
- ii) The transformation between 1D and 2D Fourier spectra is treated in 5.3. The 2D spectrum consists of the lumped coefficients. A second type of *aliasing* appears, in which two distinct 2D frequencies  $\psi_{mk}$  project on the same 1D frequency.
- iii) Also noise time-series must be transformed into the Fourier domain, in order to combine them eventually with the lumped coefficients. The result is a 1D *continuous* error power spectral density (PSD), cf. 5.4.
- iv) The transition from continuous to discrete error spectrum, pertaining to LC's, is made in 5.4. The  $\sigma_{mk}$  are the basis for the stochastic model.
- v) This transformation is the main issue of the LC approach, which is a (forward) spectral mapping from SH domain to 2D Fourier domain. The bounds of the spherical harmonic domain involved, in particular the maximum degree  $L$ , will be a main concern, cf. 5.2. Setting an arbitrary  $L$  will result in (spatial) *leakage* and will cause a certain power loss. The backward transformation, i.e. the recovery of SH coefficients from lumped coefficients is non-unique and will need to be performed by least squares estimation, cf. 6. The observation model is represented by the transfer coefficients  $H_{lmk}^\#$ , the stochastic model comes from the error spectrum  $\sigma_{mk}$ .

Finally, a few remarks on the spectral content in case of a non-circular nominal orbit are given.

All frequencies will be assumed to be in units of [Hz]. The frequency definitions from 2.2, that imply units of [rad/s], need consequently be divided by  $2\pi$ .

## 5.1 Spectral Domains

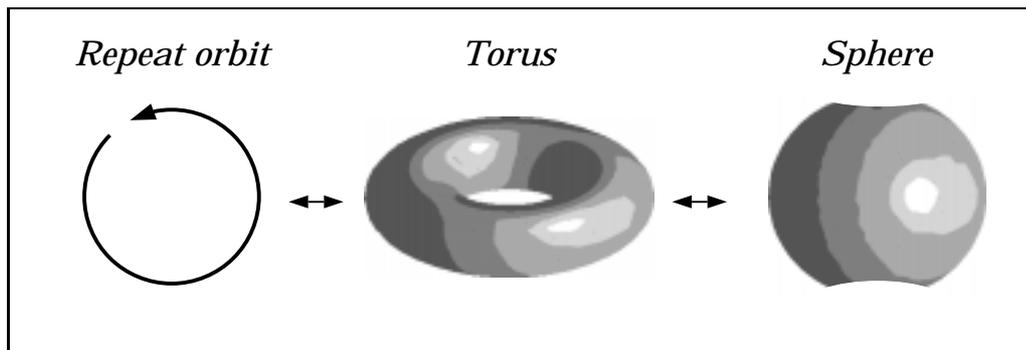
SINCE the sphere is a closed surface, the SH-spectrum is discrete by nature. Rotating the spherical harmonics into the orbital system gives an expression in the angular coordinates  $u$  and  $\Lambda$ , which both make full cycles. The domain, defined by these coordinates,  $[0, 2\pi) \times [0, 2\pi)$ , is topologically a torus. At this point the inclination is arbitrary and the orbit is not degenerate, i.e. does not close on itself. One could also say that for  $t \rightarrow \infty$  the orbit will fill up the whole torus. The corresponding spectrum on this manifold is a 2D Fourier spectrum. It is discrete, since the torus is a closed surface as well. Note that the 1D Fourier spectrum along the non-repeat orbit is discrete, though with irregular spectral line spacing, cf. (5.5) for non-integer  $\alpha$  and  $\beta$ .

**Table 5.1:** *comparing topology*

spatial domain	coordinates	spectral type	amplitude	basefunction
repeat orbit	time $t$	1D Fourier	$A_n$	$e^{i\omega_n}$
torus	$u, \Lambda$	2D Fourier	$A_{mk}$	$e^{i\psi_{mk}}$
sphere	$\theta, \lambda$	spherical harmonic	$\bar{K}_{lm}$	$\bar{Y}_{lm}(\theta, \lambda)$

If the orbit repeats, say after  $\beta$  revolutions and  $\alpha$  nodal days, the orbit will not cover the full torus anymore. As long as the sampling has been dense enough, this may not pose any problem. But in principle the repeat orbit is a degenerate torus. One ends up with a periodic function (time-series), which has a discrete 1D Fourier spectrum with homogeneous spectral line spacing. Still a mapping exists between the 1D and the 2D Fourier spectra, which will be treated in 5.3.

The aliasing problem along the orbit thus covers two issues: (i) aliasing, due to discrete time-sampling, and (ii) aliasing problems, due to the repeat orbit. The former is the classical sampling problem. The latter describes the problems of the closed orbit, in terms of overlapping (and missing) frequencies.



**Figure 5.1:** *The three spatial domains*

The spatial mapping from the 1D to the 2D Fourier domain is relatively easy. It

is governed by the relations:

$$t \mapsto u, \Lambda : u(t) = u_0 + \dot{u}t, \quad \Lambda(t) = \Lambda_0 + \dot{\Lambda}t, \quad (5.2)$$

with the frequencies  $\dot{u}$  and  $\dot{\Lambda}$  from (2.8d) and (2.8e). The step from sphere to torus (and back), though, seems to be more troublesome. Sphere and torus are not *homeomorphic* manifolds, i.e. they are topologically distinct. The way the orbit samples the sphere must be taken into consideration. First of all, one must note that in general not the full surface is considered. An inclination  $I$  will bound the sphere between  $\theta = [\frac{1}{2}\pi - I, \frac{1}{2}\pi + I]$  for prograde and  $\theta = [I - \frac{1}{2}\pi, \frac{3}{2}\pi - I]$  for retrograde orbits. This leaves only a certain *latitude band*. Secondly, this domain is in principle sampled twice, namely by ascending and by descending tracks. A mapping from sphere to torus might be described as follows:

- separate the sphere with upward from the one with downward tracks,
- remove the 4 polar caps, and
- put together the two bands at their northern and southern edges.

The result will be a ‘donut’, around which the orbit will spiral. The ascending tracks eg. at the outer side, and the descending tracks at the inner side of the torus. Though this recipe is non-mathematical, of course, it demonstrates the feasibility of mapping the sphere, as sampled by the satellite orbit, onto a torus. A limiting case occurs for  $I = \frac{1}{2}\pi$ . Indeed, it was shown in (Sneeuw & Bun, 1996) that global spherical harmonic computation may be formulated in terms of 2D Fourier transformations. The Fourier spectrum of  $\bar{P}_{lm}(\cos \theta)$  consists basically of  $\bar{F}_{lmk}(\frac{1}{2}\pi)$ .

## 5.2 The Lumped Coefficient Approach

**I**F GRAVITY field recovery is performed with the time-wise approach in the spectral domain, cf. scheme (5.1), it is referred to as *lumped coefficient approach*. Key element is the mapping  $v$  from SH spectrum to 2D Fourier spectrum, containing the lumped coefficients. Spatially, this is a mapping from sphere to torus. Spectrally, it is represented by the transfer coefficients:

$$A_{mk}^{\#} = \sum_{l=\max(|m|,|k|)}^L H_{lmk}^{\#} \bar{K}_{lm}. \quad (5.3)$$

To obtain a finite-dimensional linear system, the infinite summation over degree  $l$ , i.e. the lumping, is truncated at a certain maximum degree  $L$  here, as compared to (3.1b). Equation (5.3) in matrix notation reads e.g.  $a = H\kappa$ . The inverse mapping  $\kappa = H^{-1}a$  cannot be given explicitly and must be determined by a least squares inversion, cf. 6.

**Block-Diagonal System.** One of the most important properties of (5.3) is that for each order  $m$  a different system of equations  $a = H\kappa$  is obtained. Suppose vector  $\kappa$  would contain the full set of SH-coefficients up to degree  $L$ , then  $H$  would show *block-diagonal* structure. If  $H$  is block-diagonal, so will be the normal matrix  $H^T H$ , cf. (Koop, 1993, fig. 4.1) for a graphical representation. Each block, i.e. each order  $m$  can be treated individually now:  $a^{(m)} = H^{(m)}\kappa^{(m)}$ . Because of

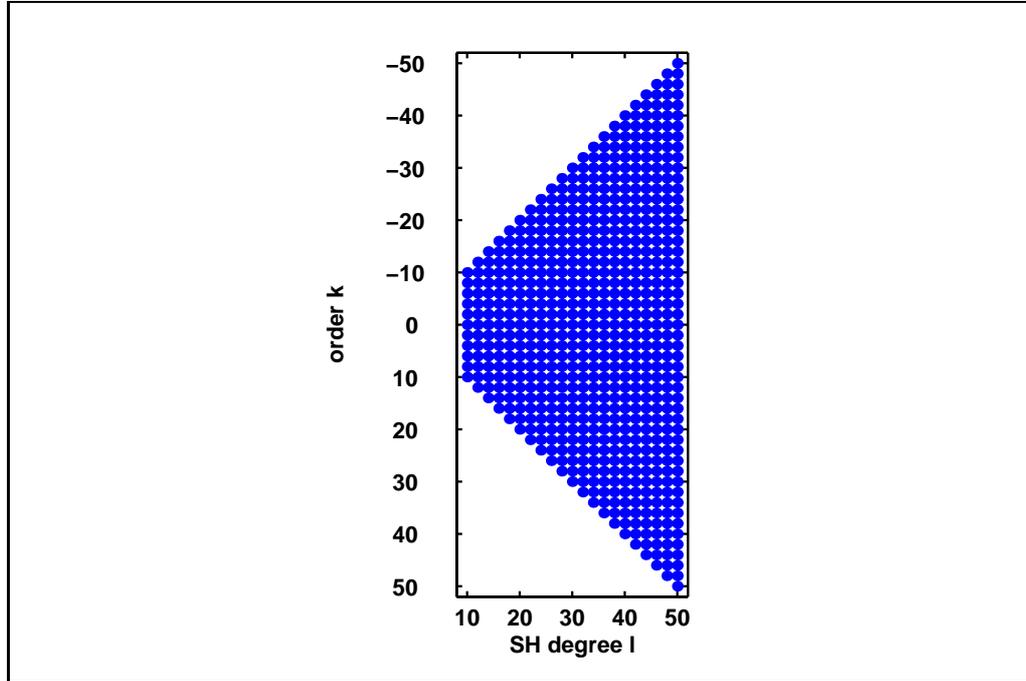
$-L \leq k \leq L$  and  $|m| \leq l \leq L$ , the matrix  $H$  has dimensions  $(2L+1) \times (L-|m|+1)$ , which is easy to manage by relatively simple computer equipment.

The dimensions of these  $m$ -blocks can even be halved by considering the alternating-zero behaviour of the  $H_{lmk}^\#$ . Those using  $\bar{F}_{lmk}(I)$  are zero for  $l-k$  odd, whereas those with  $\bar{F}_{lmk}^*(I)$  vanish for  $l-k$  even. Proper even/odd permutation of columns and rows yields 2 sub-blocks for each order  $m$  of maximum size  $(L+1) \times (\frac{1}{2}(L-|m|+1))$ . In case of transfer coefficients with  $l-k$  even, the following two sets of equations arise for a given order  $m > 0$  (provided that  $L-m$  is even, too):

$$\begin{pmatrix} A_{m,-L} \\ A_{m,-L+2} \\ \vdots \\ A_{m,L} \end{pmatrix} = \begin{pmatrix} H_{m,m,-L} & H_{m+2,m,-L} & \cdots & H_{L,m,-L} \\ H_{m,m,-L+2} & H_{m+2,m,-L+2} & \cdots & H_{L,m,-L+2} \\ \vdots & \vdots & & \vdots \\ H_{m,m,L} & H_{m+2,m,L} & \cdots & H_{L,m,L} \end{pmatrix} \begin{pmatrix} \bar{K}_{m,m} \\ \bar{K}_{m+2,m} \\ \vdots \\ \bar{K}_{L,m} \end{pmatrix},$$

$$\begin{pmatrix} A_{m,-L+1} \\ A_{m,-L+3} \\ \vdots \\ A_{m,L-1} \end{pmatrix} = \begin{pmatrix} H_{m+1,m,-L+1} & H_{m+3,m,-L+1} & \cdots & H_{L-1,m,-L+1} \\ H_{m+1,m,-L+3} & H_{m+3,m,-L+3} & \cdots & H_{L-1,m,-L+3} \\ \vdots & \vdots & & \vdots \\ H_{m+1,m,L-1} & H_{m+3,m,L-1} & \cdots & H_{L-1,m,L-1} \end{pmatrix} \begin{pmatrix} \bar{K}_{m+1,m} \\ \bar{K}_{m+3,m} \\ \vdots \\ \bar{K}_{L-1,m} \end{pmatrix}.$$

Since  $H_{lmk}$  vanishes for  $|k| > l$ , the upper and lower left corners are filled with zeros, such that the non-zero part is trapezium shaped, as exemplified by fig. 5.2.



**Figure 5.2:** Non-zero elements in matrix  $H$  with  $L = 50$ ,  $m = 10$

An immediate consequence of the block structure is, that no correlation between  $m$ -blocks can arise in gravity field solutions, determined by inversion of the block normal matrices. Two coefficients of different order cannot be correlated. Coefficients of the same order are only correlated if they have the same parity. This

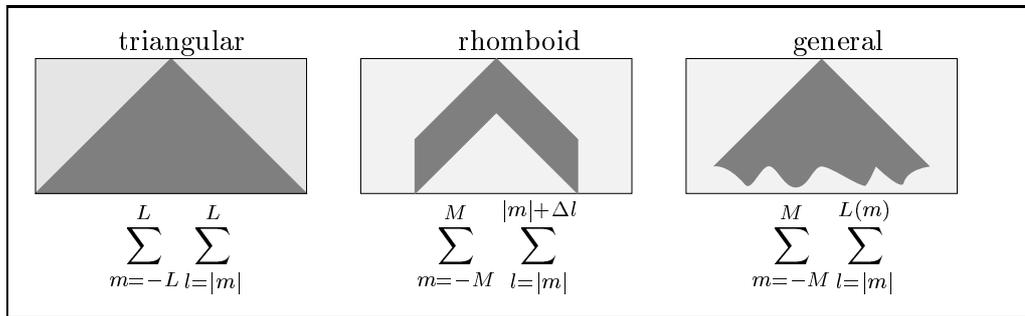
behaviour is even present, though not exact, in satellite-only gravity field models, that were derived under full-normal matrix inversions, cf. (Haagmans & Van Gelderen, 1991).

**Remark 5.1** Viewed in the spatial domain, each  $m$ -block treats the features of the geopotential with zonal wavenumber  $m$ . Each of the two sub-blocks within an  $m$ -block treats the features that are either symmetric or anti-symmetric with respect to the equator.

**Spherical Harmonic Domain.** Another important property, that affects the size of  $H$ , is the fact that the maximum degree  $L$  can be chosen freely. In principle, the summation over degree  $l$  runs till infinity, cf. (3.1b). Setting an (arbitrary)  $L$ , as in (5.3), therefore implies an approximation, characterized by two phenomena:

- Cutting off sharply in the SH domain causes *spatial leakage* on the sphere.
- Neglecting higher degree coefficients induces an *omission error*, i.e. a power loss. Propagated onto the lumped coefficient, the omission error is also referred to as truncation error effect (TEE), cf. (Klokočník et al., 1990).

The conventional choice for cutting off the spherical harmonic domain is a constant  $L$ , resulting in the usual triangular SH domain. But other choices, in which  $L$  may depend on order  $m$ , are valid as well. E.g. the rhomboid domain, with  $L = L(m) = |m| + \Delta l$ , where  $\Delta l$  is a fixed number, yields matrices  $H$  of equal size for all  $m$ .  $\Delta l$  can even be tuned such that  $H$  becomes square. In general,  $L$  is a function of  $m$ , cf. fig. 5.3. Cut-off criterion can be the conditioning of  $H$ , redundancy of  $H$ , size of omission error, amount of leakage to be expected/avoided, and so on.



**Figure 5.3:** choice of spherical harmonic domain (order  $m$  on horizontal, degree  $l$  on vertical axis)

Two effects play a role here. Firstly, the power spectrum of the geopotential decreases for increasing degree, cf. Kaula's rule of thumb (Kaula, 1966). Thus the neglected coefficients will be smaller on average than the ones retained in the model. Secondly, the upward continuation term in  $H_{lmk}^\#$  acts like a low-pass filter. Higher degrees are diminished in size stronger than lower degrees. However, both effects operate in a smooth way. Leakage and omission error can consequently only be reduced in a smooth way by increasing  $L$ . This makes the choice of a proper cut-off criterion a difficult task. The discussion in 6.3 on unresolved parameters elaborates on this problem from the viewpoint of parameter estimation.

**Remark 5.2 (general index ranges)** A band-limited function on the sphere up to degree  $L$ , has a lumped coefficient spectrum along the orbit with  $-L \leq m, k \leq L$  indeed. In the inverse formulation, however, starting from a functional along the orbit, the indices are not necessarily restricted to a fixed range  $[-L, L]$ . The choice of index ranges may depend on similar criteria as in the  $L(m)$  discussion above. The general expression of a functional of the geopotential becomes, see also (3.1):

$$f^\# = \sum_{m=-M}^M \sum_{k=-K}^K \sum_{l=l_0}^L H_{lmk}^\# \bar{K}_{lm} e^{i\psi_{mk}} \quad , \quad l_0 = \max(|m|, |k|) .$$

Thus, a great advantage of the lumped coefficient approach is the separation *and* the flexibility in the spherical harmonic and Fourier spectral domains. Separation refers to the  $m$ -block structure, flexibility relates to the choice of spectral domain  $(L, M, K)$ . A further advantage of the LC approach is the fact that power spectral densities (PSDs) can be used for the stochastic modelling of observations, cf. 5.4. This is especially an advantage over space-wise methods, since instrument accuracies are usually expressed in the spectral domain. They can be transformed into equivalent covariance functions in the time domain. An error description on the sphere, though, will be extremely difficult.

**Initial State Elements.** In the lumped coefficient formulation, so far, it was tacitly assumed that the initial state values of  $\psi_{mk}$  were zero. In general, it is:

$$\psi_{mk}(t) = \psi_{mk}^0 + \dot{\psi}_{mk} t ,$$

with arbitrary  $\psi_{mk}^0$ . This causes deformation of the lumped coefficient spectrum.

$$f^\#(t) = \sum_{m,k} A_{mk}^\# e^{i\psi_{mk}} = \sum_{m,k} \left[ A_{mk}^\# e^{i\psi_{mk}^0} \right] e^{i\dot{\psi}_{mk} t} .$$

The coefficients, pertaining to the spectral lines  $\dot{\psi}_{mk}$  and determined from  $f^\#(t)$ , have to be corrected for the phase information  $\psi_{mk}^0$ . Then the LC formulation (5.3) may be applied.

### 5.3 Fourier Domain Mapping

**M**EASUREMENTS along the orbit produce a time-series. Fourier analysis yields the corresponding 1D Fourier spectrum, which has to be mapped onto the 2D lumped coefficient spectrum. In order to investigate this spectral projection, we will consider the inverse mapping first.

In case the geopotential has maximum degree  $L$ , the lumped coefficient spectrum consists of the spectral lines

$$\dot{\psi}_{mk} = k\dot{u} + m\dot{\Lambda} \quad , \quad \text{with } -L \leq m, k \leq L ,$$

which is naturally discrete. In order to obtain a discrete 1D Fourier spectrum, the orbit must be periodic. A *periodic orbit* is obtained if the basic frequencies are commensurable, i.e. if there exists an integer ratio:

$$-\frac{\dot{u}}{\dot{\Lambda}} = \frac{\beta}{\alpha} , \tag{5.4}$$

also called the repeat ratio. A minus sign has been used in definition (5.4), since the Earth rotation rate is always larger than  $\dot{\Omega}$  and thus  $\dot{\Lambda}$  is always negative. The orbit performs  $\beta$  revolutions in  $\alpha$  nodal days. The relative primes  $\alpha$  and  $\beta$  are sometimes referred to as  $N_d$  (number of nodal days) and  $N_r$  (number of revolutions), respectively (Schrama, 1989).

In reality, such an orbit biting its own tail, is a fiction. Due to drag-forces, unknown gravity field and injection errors, even a drag-free mission can never be designed to be in perfect repeat mode. Periodic orbits can only be maintained by orbit manoeuvres. As *working hypothesis* or as *approximation*, however, the circular repeat orbit is most valuable. In the first place, a repeating orbit guarantees a homogeneous pattern of ground tracks. Apart from a potential aliasing problem it gives the optimal sampling of the Earth for the given inclination. In the second place, a periodic signal will give a discrete spectrum, which is more easily related to the (also discrete) spherical harmonic spectrum.

The 2D spectrum  $\dot{\psi}_{mk}$  can be transformed into a 1D spectrum, making explicit use of the repeat ratio (5.4):

$$\dot{\psi}_{mk} = \dot{u}\left(k + m\frac{\dot{\Lambda}}{\dot{u}}\right) = \dot{u}\left(k - m\frac{\alpha}{\beta}\right) = \frac{\dot{u}}{\beta}(k\beta - m\alpha). \quad (5.5)$$

The lowest frequency larger than DC, i.e. the *spectral resolution*, is  $\Delta\dot{\psi} = \dot{u}/\beta$ . The maximum frequency is:  $\dot{\psi}_{-L,L} = -\dot{\psi}_{L,-L} = L\frac{\alpha+\beta}{\beta}\dot{u}$ . Equation (5.5) defines the spectral mapping from the 2D Fourier onto the 1D Fourier domain, valid for  $\beta/\alpha$ -repeat orbits:

$$(m, k) \mapsto n : \quad \dot{\psi}_n = n\Delta\dot{\psi}, \text{ with } n = k\beta - m\alpha. \quad (5.6)$$

Note that  $(k\beta - m\alpha)$  is a combination of integers only.

The spectral resolution was seen to be  $\Delta\dot{\psi} = \dot{\psi}_1 = \dot{u}/\beta$ . The corresponding period is the repeat period, which is also the *minimum mission duration*:  $T = \beta/\dot{u}$ . Thus  $T$  equals  $\beta$  revolutions. The maximum frequency number is  $N = L(\alpha + \beta)$ . According to the Nyquist criterion, at least  $2N = 2L(\alpha + \beta)$  samples have to be taken within  $T$ . So the Nyquist sampling interval, i.e. the *maximum* interval for which all frequencies can still be resolved without aliasing (of first type), reads:

$$\Delta t_N = \frac{T}{2N} = \frac{T}{2L(\alpha + \beta)} = \frac{\beta}{2L\dot{u}(\alpha + \beta)}. \quad (5.7)$$

Although a repeat orbit gives a homogeneous coverage, it still consists of crossing upward and downward arcs. The *spatial resolution* of the ground-track pattern is hard to define. As a measure, the equatorial distance between two neighbouring tracks can be chosen. Because of the existence of upward *and* downward tracks, this will be a conservative measure. After  $\beta$  revolutions, it holds  $\Delta\lambda = 2\pi/\beta$ . Around the equator, the highest wavenumber is  $L$ . So the Nyquist criterion is met in the spatial domain, if  $\beta > 2L$ .

**Non-Overlap Condition.** Aliasing of second type occurs when two distinct 2D frequencies  $\dot{\psi}_{mk}$  project on the same  $\dot{\psi}_n$ . In that case  $(m, k) \mapsto n$  is not *injective* and the inverse Fourier domain mapping becomes non-unique. On the other

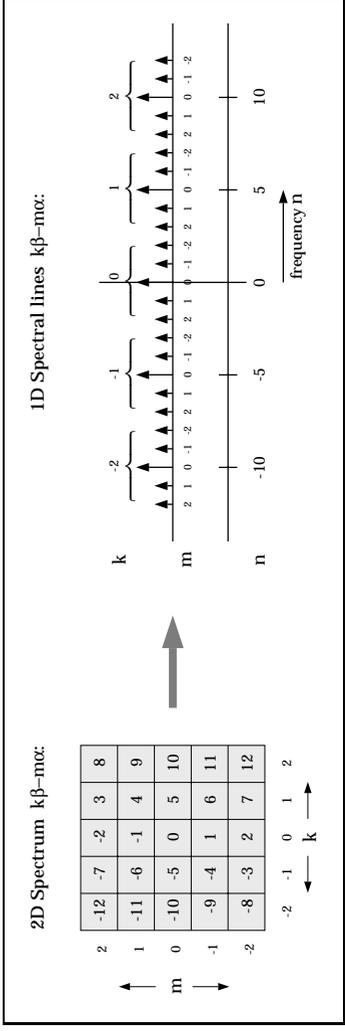
hand, it is not sure in advance whether it is a *surjection* either. Not all spectral numbers  $n$  might be reached by the full set of  $(m, k)$ -combinations. Spectral gaps in the 1D spectrum will be the result, although this will not be a serious problem. So in general the projection will not be *bijective* (one-to-one).

A crude way to derive a non-overlap condition is to inspect the difference in length of the ‘vector’  $\psi_n$  and the size of the ‘matrix’  $\psi_{mk}$ . Knowing that  $-L(\alpha+\beta) \leq n \leq L(\alpha+\beta)$ ,  $-L \leq m, k \leq L$ , the number of spectral lines  $n$  equals  $2L(\alpha+\beta)+1$ . The number of  $(m, k)$ -combinations is  $(2L+1)^2$ . The following minimum condition must hold:

$$(2L+1)^2 \leq 2L(\alpha+\beta) + 1$$

$$\Leftrightarrow 2(L+1) \leq \alpha + \beta.$$

For low Earth orbiters the repeat ratio  $\beta/\alpha$  is around 16, yielding  $\alpha + \beta \approx 1.1\beta$ . Thus a rough minimum condition on  $\beta$  is:  $\beta > 1.9L$ .



**Figure 5.4:** Fourier domain mapping with  $L = 2$ ,  $\alpha = 1$  and  $\beta = 5$ .

It is more sophisticated, though, to see how the projection  $(m, k) \mapsto n$  really works. In fig. 5.4 the left picture represents the 2D spectral lines. The parameters are  $L = 2$ ,  $\alpha = 1$  and  $\beta = 5$ . The right hand part represents the 1D spectrum. The main peaks at  $n = k\beta$  denote integer multiples of the orbital frequency, i.e. they correspond to the effect of zonal coefficients ( $m = 0$ ). It can be seen that for each value of  $k$  the spectral lines are grouped, which is only caused by the fact that  $\beta > \alpha$ . From this picture it is clear already when frequencies could possibly overlap. When  $m$  increases, the spectral lines belonging to  $k$  ‘walk’ towards the group of  $k - 1$ . There are  $\max(m) - \min(m) = 2L$  minor peaks in between the major peaks at  $k\beta$ . For this simple example overlaps would therefore occur if  $2L > \beta$ . Then the minor peaks, belonging to  $k$ , would interfere with those of  $k \pm 1$ . In fig. 5.4 this is easily verified since  $\alpha = 1$ , but in general, one must demand that:

$$\psi_{m_1 k_1} \neq \psi_{m_2 k_2}$$

$$\Leftrightarrow k_1 \beta - m_1 \alpha \neq k_2 \beta - m_2 \alpha$$

$$\Leftrightarrow (k_1 - k_2) \beta \neq (m_1 - m_2) \alpha$$

$$\Leftrightarrow \frac{m_1 - m_2}{k_1 - k_2} \neq \frac{\beta}{\alpha}.$$

Note that  $\alpha$  and  $\beta$  are relative primes, i.e. their greatest common divisor is 1, so that the above case can only occur when  $(m_1 - m_2) = i\beta$  and  $(k_1 - k_2) = i\alpha$

simultaneously, with  $i \in \mathbb{N}/\{0\}$ . This is avoided already when:

$$\beta > 2M. \quad (5.8)$$

In case of the conventional triangular SH domain, this requirement is equivalent to  $\beta > 2L$ , which is in agreement with the above considerations. See also (Schrama, 1989), where the same overlap problem is treated, based on real spectra.

If, however, two frequencies do overlap, the consequence is twofold. One would only see the resulting frequency with the corresponding lumped coefficient:

$$\begin{aligned} \dot{\psi}_n = \dot{\psi}_{m_1 k_1} = \dot{\psi}_{m_2 k_2} &> A_n = A_{m_1 k_1} + A_{m_2 k_2} \\ &= \sum_l (H_{lm_1 k_1} \bar{K}_{lm_1} + H_{lm_2 k_2} \bar{K}_{lm_2}). \end{aligned}$$

The first drawback of this type of aliasing is loss of ‘observations’, possibly causing the system to become underdetermined. The two LCs  $A_{m_1 k_1}$  and  $A_{m_2 k_2}$ , are lumped further into  $A_n$ . But a more serious drawback will be the fact that unknowns of orders  $m_1$  and  $m_2$  will become correlated, since off-diagonal blocks are introduced.

Opposed to aliasing, which has to do with injectivity of the mapping  $(m, k) \mapsto n$ , it might also occur that the mapping is non-surjective. In that case certain 1D frequencies  $\dot{\psi}_n$  are not reached by any  $\dot{\psi}_{mk}$ . For the inverse formulation, this does not pose any problem. Those frequencies are just not considered. Full bijectivity occurs under the following condition, see the ‘crude’ non-overlap condition:

$$(2L + 1)^2 = 2L(\alpha + \beta) + 1 > L = \frac{1}{2}(\alpha + \beta) - 1,$$

which can be verified in the example of fig. 5.4:  $L = \frac{1}{2}(1 + 5) - 1 = 2$  indeed. Note that this is not the same as  $\beta = 2L + 1$ , which would only be true for  $\alpha = 1$ .

## 5.4 Power Spectral Density Modelling

THE LUMPED coefficient approach is a spectral method. Since it is common to characterize instrumental errors spectrally by so-called power spectral densities (PSDs), this is an advantage over spatial or time-wise models. It will be shown, though, that at least for white noise the time-wise and the LC-approach reduce to the same normal equations.

As the name *power spectral density* indicates, a PSD describes the amount of power in an *infinitesimal* spectral band  $df$ . It must be expressed, therefore, in units<sup>2</sup>/Hz. It is defined through the Fourier transform pair between PSD  $S(f)$  and autocorrelation function  $R(\tau)$  of a (weakly) stationary stochastic process (Papoulis, 1965, 10-3):

$$S(f) = \int_{-\infty}^{\infty} R(\tau) e^{-i2\pi f\tau} d\tau \Leftrightarrow R(\tau) = \int_{-\infty}^{\infty} S(f) e^{i2\pi f\tau} df. \quad (5.9)$$

**Remark 5.3** Note that non-stationary processes are excluded from definition (5.9). The noise of phenomena like along-orbit data gaps, thruster firing, and so on, cannot be represented by a PSD.

The total power or variance of the stochastic process is just the spectral integral of  $S(f)$ :

$$R(0) = \sigma^2 = \int_{-f_{\max}}^{f_{\max}} S(f) \, df, \quad (5.10)$$

where  $f_{\max}$  is in principle infinite, though for practical reasons usually finite. When the process is real, both  $S(f)$  and  $R(\tau)$  are real and symmetric. In engineering applications the PSD is therefore usually defined on the positive frequency axis only. The total variance is an integral from DC to  $f_{\max}$ . If PSDs, defined in this way, have to be transformed to the above type, a factor  $\frac{1}{2}$  has to be applied, in order to preserve variance. See also the paragraph on band limitation below.

**Continuous to Discrete.** The error spectrum is continuous, whereas the along-orbit signal consists of discrete spectral lines. For stochastic modelling purposes an error measure pertaining to a single frequency  $\psi_{mk}$  (or  $\beta_{mk}$ ) is required. This error measure is the amount of noise power in the frequency band directly around  $\psi_{mk}$ . Thus the error power spectral *density* must be integrated over a tiny frequency band, whose width is the distance between two neighbouring spectral lines, that is the *spectral resolution*  $\Delta f$ . Let the resulting error variance for a certain spectral line  $f_{mk} = \psi_{mk}$  be called  $\sigma_{mk}^2$ :

$$\sigma_{mk}^2 = \int_{f_{mk} - \Delta f/2}^{f_{mk} + \Delta f/2} S(f) \, df \cong S(f_{mk}) \Delta f. \quad (5.11)$$

Since  $\Delta f$  is a small quantity, the latter quadrature approximates the integral well enough. It is clear that the total variance can now be written:  $R(0) = \sigma^2 = \sum_{m,k} \sigma_{mk}^2$ .

The question arises what the spectral resolution  $\Delta f$  is. For a time-series of length  $T$ , the spectral resolution is generally  $\Delta f = 1/T$ . In the case of a nominal repeat orbit, though, this does not apply directly if the mission length  $T$  and the repeat period  $T_r$  are not equal. The following cases are considered.

- Non-repeat orbit. From a measurement point of view, the spectral resolution is  $1/T$  indeed.  $T$  is the longest possible period. However, if  $\dot{u}$  and  $\dot{\Lambda}$  are not commensurate, the spectral lines  $\beta_{mk}$  are not equally spaced. One would have to consider a frequency dependent spectral resolution  $\Delta f(f_{mk})$ , that equals  $1/T$  on the average.
- Repeat orbit with exactly  $T = T_r = \beta T_u = \alpha T_\Lambda$ . The spectral resolution was seen to be  $\dot{u}/\beta$  in chapter 5. This equals  $1/T_r$ , which is  $1/T$  indeed.
- Repeat orbit with an integer amount of repeats:  $T = nT_r$ ,  $n \in \mathbb{N}/\{0\}$ .  $\Delta f$  remains  $1/T_r$ . But the globe is sampled  $n$  times, loosely speaking. The error variance will diminish by a factor  $n$ . Thus the spectral resolution is effectively  $1/nT_r = 1/T$  again.
- Repeat orbit with  $n \in \mathbb{R}$ . Even in this case, one could maintain that the globe is sampled  $n$  times. It still holds  $\Delta f = 1/T$ .

The case  $n < 1$  is not treated further. The spectral resolution would become too large, colliding with the non-overlap condition (5.8). The diagonal error covariance matrix  $Q$  will thus consist of the discrete, frequency dependent error spectrum:

$$\sigma_{mk}^2 = S(f_{mk})/T. \quad (5.12)$$

**Bandwidth.** In reality the error spectrum is band-limited, since the power  $\sigma^2$  must be finite. From a sampling point of view the maximum frequency  $f_{\max}$  is the *Nyquist* frequency  $f_{\text{Nyq}}$ , which is half the required sampling rate. From the instrumental side band-limitation might be due e.g. to internal low-pass filtering. The total noise variance  $\sigma^2$  is equal to the above integral (5.10) over the full bandwidth.

This is not what happens in gravity field analysis. In principle the degree of the gravitational spectrum goes to infinity. In practice, though, a certain maximum degree  $L$  is chosen, producing a maximum signal frequency of  $\dot{\psi}_{L,-L} \approx L\dot{u}$ , cf. 5.3. Even if the sampling rate is much higher than twice this frequency, the gravity field analysis only considers the error spectrum in the band  $[\dot{\psi}_{-L,L}, \dot{\psi}_{L,-L}]$ :

$$\sigma^2 = \int_{-cL\dot{u}}^{cL\dot{u}} S(f) \, df. \quad (5.13)$$

The factor  $c$  represents the weakness of the spectral cut-off definition and the approximation of the maximum frequency by  $L\dot{u}$ . A choice near 1.5 would be appropriate. After choosing  $L$ , the noise is implicitly band-filtered by the LC approach. If results from this approach have to be compared to the time-wise approach, one should either low-pass filter the along-orbit time-series, or select a proper sampling frequency. Otherwise the time-wise approach is too pessimistic. Modelling of band-limitation in the time-domain is discussed further in (Schuh, 1996).

**A White Noise Example.** Finally, a simple white noise example might elucidate the above PSD-formulae. Suppose we have a constant PSD level, at least in the interesting frequency band:  $S(f) = w$ . Given a mission life-time  $T$  and a sampling rate  $\Delta t$ , one obtains:

$$\begin{aligned} \text{spectral resolution: } \Delta f &= \frac{1}{T}, \\ \text{Nyquist frequency: } f_{\text{Nyq}} &= \frac{1}{2\Delta t}, \\ \text{discrete error spectrum: } \sigma_{mk}^2 &= w\Delta f = \frac{w}{T}, \\ \text{total error power: } \sigma^2 &= 2f_{\text{Nyq}}w = \frac{w}{\Delta t}. \end{aligned}$$

The stochastic model will distribute the total error power  $\sigma^2$  over the observables. Depending on the approach one gets:

$$\begin{aligned} \text{frequency-wise: } \sigma^2 &= \sum_{m,k} \sigma_{mk}^2 = N_f \sigma_{mk}^2, \\ \text{time-wise: } \sigma^2 &= 2f_{\text{Nyq}}c = \frac{2f_{\text{Nyq}}}{\Delta f} \sigma_{mk}^2 = \frac{T}{\Delta t} \sigma_{mk}^2 = N_p \sigma_{mk}^2, \end{aligned}$$

where  $N_f$  refers to the number of discrete spectral lines and  $N_p$  the number of time-samples or points along the orbit. The central connection between both is the relation:

$$\frac{2f_{\text{Nyq}}}{\Delta f} \stackrel{?}{=} \frac{T}{\Delta t} \Leftrightarrow N_f \stackrel{?}{=} N_p. \quad (5.14)$$

If this relation holds, i.e. when the maximum signal frequency equals half the sampling rate indeed, the total noise is distributed over an equal amount of spectral lines and time samples. The time-wise approach and the lumped coefficient approach yield the same results. Relation (5.14) might then be considered a Parseval relation of the stochastic model.

**Remark 5.4** *As far as the lumped coefficient approach is concerned, the sampling rate does not play a role. Only when a comparison to a time-wise approach has to be made, it enters through the condition that Nyquist frequency and maximum signal frequency match.*

## 5.5 Non-Circular Nominal Orbit

IF THE nominal orbit is chosen to be eccentric, the consequence will be a richer spectrum. From (Kaula, 1966, §3.3) we have

$$\frac{1}{r^{l+1}} e^{i[k(\omega + f) + m\Lambda]} = \frac{1}{a^{l+1}} \sum_{q=-\infty}^{\infty} G_{lkq}(e) e^{i[k\omega + (k+q)M + m\Lambda]}, \quad (5.15)$$

with  $G_{lkq}(e)$  the so-called *eccentricity functions*. Since  $G_{lkq}(e) \sim \mathcal{O}(e^{|q|})$ , the summation over  $q$  can be limited to some low integers. A certain functional of the geopotential will now read:

$$f^\# = \sum_l \sum_m \sum_k \sum_q H_{lmk}^\# G_{lkq} \bar{K}_{lm} e^{i\psi_{mkq}}. \quad (5.16)$$

Now the precession of the perigee is slow:  $\dot{\omega}$  corresponds to about 100 days for high inclination LEO's. Hence it should be useful to separate  $\omega$  and write:

$$\begin{aligned} \psi_{mkq} &= k\omega + (k+q)M + m\Lambda \\ &= (k+q)u + m\Lambda - q\omega \\ &= \psi_{m,k+q} - q\omega. \end{aligned}$$

Thus it becomes clear that the old spectrum is augmented by spectral lines  $q\dot{\omega}$  around each former  $\dot{\psi}_{mk}$ . The series (5.16) suggests the use of lumped coefficients with an additional  $q$ -index, e.g.

$$A_{mkq}^\# = \sum_l H_{lmk}^\# G_{lkq} \bar{K}_{lm}. \quad (5.17)$$

See also (Klokočník, 1988; Wnuk, 1988) for similar definitions of real-valued coefficients. Alternatively, by writing  $k+q := p$ , the functional  $f^\#$  can be recast into:

$$\begin{aligned} f^\# &= \sum_m \sum_p \sum_q \sum_l H_{lm,p-q}^\# G_{l,p-q,q} \bar{K}_{lm} e^{-iq\omega} e^{i\psi_{mp}} \\ &= \sum_m \sum_p A_{mp}^\# e^{i\psi_{mp}}, \end{aligned}$$

which looks the same as the usual LC formulation. However,  $A_{mp}^\#$  will be a time-dependent lumped coefficient now:

$$\begin{aligned} A_{mp}^\# &= \sum_q \sum_l H_{lm,p-q}^\# G_{l,p-q,q} \bar{K}_{lm} e^{-iq\omega} \\ &= \sum_q A_{m,p-q,q}^\# e^{-iq\omega}. \end{aligned} \quad (5.18)$$

This definition comes down to the formulation in (Wagner & Klosko, 1977). It is a summation of a few harmonics in  $\omega$ , with amplitudes from (5.17). The main term ( $q=0$ ) equals  $A_{mk}^\#$ .

The introduction of new frequencies, due to the eccentric orbit, mixes up the depicted scheme and ideas of the repeat orbit, with commensurate frequencies  $\dot{u}$  and  $\dot{\Lambda}$ . The requirement that  $\dot{\omega}$  be commensurate as well is too severe to be realistic. Moreover, in order to resolve all frequencies  $\dot{\psi}_{mkq}$ , at least one full revolution of the perigee ( $\sim 100$  days) would be required. This period, and the corresponding spectral resolution, interferes with typical mission lengths and the corresponding spectral resolution  $\dot{u}/\beta$ . However, by virtue of (5.18) and due to the size of  $G_{lkq}(e)$ , one can still employ the ordinary lumped coefficient definition, and consider the time varying part as small correction terms, either a priori with existing gravity field knowledge, or a posteriori by iteration.

## 5.6 Summary

It was seen that the lumped coefficient approach leads to a model that:

- i) is inherently linear.
- ii) yields block-diagonal structure, i.e. treats the information per order  $m$  separately. The  $m$ -blocks will basically be uncorrelated.

From the spectral considerations in this chapter, the following minimum requirements for orbit design are defined. Suppose the geopotential field up to degree  $L$  is to be resolved from a gravity field mission. Suppose furthermore that a  $\beta/\alpha$  repeat orbit is chosen because of homogeneous coverage and discrete spectrum considerations. Then:

- iii) the minimum mission length  $T$  must be  $2L$  revolutions. This requirement comes from the non-overlap condition (5.8), avoiding torus-to-circle aliasing.
- iv) the sampling interval is  $\frac{T}{2L(\alpha+\beta)}$  at most, in order to avoid classical aliasing.
- v) the spatial resolution may be characterized by the equatorial distance between two neighbouring tracks:  $\Delta\lambda = \frac{2\pi}{\beta}$ . This complies with the non-overlap condition above.

Stochastic modelling in the frequency domain requires power spectral densities. It was seen that:

- vi) for practical reasons the spectral resolution  $\Delta f$  equals  $1/T$ , with  $T$  mission duration, even in case of non-repeat orbits and repeated repeat orbits.

- vii) time-wise modelling and frequency-wise modelling lead to the same results in case of white noise.

Further lessons are:

- viii) the non-overlap condition is rather a constraint on the maximum order  $M$  than on maximum degree  $L$ , cf. (5.8).
- ix) if spectral overlap does occur, correlations between orders  $m$  arise.
- x) although the spectral content of the signal along a non-circular nominal orbit is richer, it is preferred to handle eccentricity induced terms by corrections to the conventional lumped coefficients.

## 6 Least Squares Error Analysis

THE POCKET guide, consisting of transfer coefficients of several geopotential functionals, defines the observational model. This would be sufficient to run sensitivity and stability (or eigen-) analyses. In order to perform error simulations, or even to do real data inversion, a stochastic model is required additionally. This chapter is devoted to the stochastic description of observables and unknowns. In terms of the scheme (5.1), this chapter is concerned with the (backward) mapping  $v$ , in which the signal and error SH spectra are inferred from the lumped coefficients and their corresponding error variances.

The first section starts off with a discussion of least squares estimation. In satellite geodesy the problem of gravity field recovery is inherently ill-posed. An inverse can only be attained through inclusion of a priori information. Thus various regularization options and their consequences are discussed.

Least squares error simulation is the main analysis type in this work. The corresponding section clarifies which stochastic information goes in, and what can be expected to be output. Output of the simulations will basically be covariance matrices of the unknowns, although further quality measures are presented. The last sections cover representation, including graphical, of this information.

### 6.1 Least Squares Estimation and Regularization

**Least Squares Estimation.** From 4 the observational model was defined as the linear system  $a = H\kappa$ . This is recast now in the more general and conventional notation  $y = Ax$ , with  $y$  lumped coefficients  $A_{mk}$ ,  $A$  the transfer coefficients  $H_{lmk}^\#$  and  $x$  the vector of unknown SH coefficients  $K_{lm}$ . Since measurements contain noise by nature, the vector of observations is a stochastic variable. In this section stochastic variables are denoted by underscores. The observational model is cast into the following form:

$$E\{\underline{y}\} = Ax, D\{\underline{y}\} = Q_y. \quad (6.1)$$

The operators  $E\{\dots\}$  and  $D\{\dots\}$  denote the mathematical *expectation* and *dispersion* respectively. The matrix  $Q_y$  is the noise covariance matrix of the vector of observations. Alternatively one may write  $\underline{y} = Ax + \underline{e}$ ,  $E\{\underline{e}\} = 0$ , which expresses that the (stochastic) vector of discrepancies  $e$  has zero expectancy, i.e. the model is unbiased. If (6.1) is overdetermined, one seeks those unknowns  $x$  that minimize the weighted discrepancies in quadratic sense:

$$\hat{x} : \min_x \underline{e}^\top P_y \underline{e}, \quad (6.2)$$

which is attained by the *least squares estimator*:

$$\hat{\underline{x}} = (A^T P_y A)^{-1} A^T P_y \underline{y}, \quad (6.3a)$$

with weight matrix of the observations  $P_y = Q_y^{-1}$ . The matrix  $A^T P_y A$  is known as the *normal matrix*  $N$ . Also the name *information matrix* is used, (Bierman, 1977; Strang, 1986). Expectation and dispersion become:

$$E\{\hat{\underline{x}}\} = (A^T P_y A)^{-1} A^T P_y E\{\underline{y}\} = (A^T P_y A)^{-1} A^T P_y A x = x, \quad (6.3b)$$

$$D\{\hat{\underline{x}}\} = (A^T P_y A)^{-1} = N^{-1} = Q_{\hat{x}}. \quad (6.3c)$$

The estimator is unbiased and the a posteriori covariance matrix equals the inverted normal matrix.

**Ill-Posed Problems.** In general the linear problems in satellite geodesy are ill-posed for various reasons, e.g. (Bouman & Koop, 1997). The data-distribution is usually irregular, due to polar gaps (non-polar orbits) or non-continuous data tracking. The observable itself, e.g. a certain single component from the gravity gradient tensor, may contain insufficient information about the gravity field. The downward continuation always will cause instabilities. Also the stochastic model may cause a badly conditioned normal matrix, e.g. when the instrument is unable to measure in the whole spectral domain. The so-called  $1/f$  noise in gradiometry is notorious in this respect.

A vast literature on ill-posed problems exists, rooted in integral equation theory. Opposed to such continuous ill-posed problems, we will be concerned here with discrete ill-posed problems only. Referring to (6.1), the discrete ill-posed problem is characterized by, cf. (Lanczos, 1964; Hansen, 1992):

- a gradual decay of the eigenvalues of  $N$  towards zero,
- a large condition number of  $N$  (ratio largest to smallest singular value).

Both criteria must be met simultaneously. If the eigenvalue spectrum drops off to zero abruptly, the matrix  $N$  is 'just' rank deficient. The solution incorporates a null-space. If both criteria are met, however, one cannot speak of the existence of a null-space anymore. Instead, parts of the solution space are hardly accessible.

If the normal matrix is ill-conditioned, the above norm (6.2) is insufficient to invert  $N$  in a stable way. Therefore the normal matrix has to be regularized. Loosely expressed, the data alone is not sufficient to get a unique solution. One has to impose further conditions on the desired solution, for instance by a priori knowledge of the unknowns. The category that does not employ prior knowledge contains methods like truncated and damped singular value decomposition, maximum entropy regularization, conjugate gradient methods (Hansen, 1992) and generalized ridge regression (Xu & Rummel, 1994). These regularization methods are not pursued here further, since in gravity field determination a priori knowledge actually exists up to a certain level.

**Regularization by Prior Information.** Thus the norm is extended now by further constraints on the unknown vector  $x$ , leading to a so-called *hybrid* norm. These *signal* constraints come from prior information on the unknowns. In general the

side constraint may be nonlinear:  $L(x - x_0)$ , or even a weighted sum of side constraints:

$$\sum_i \alpha_i L_i(x - x_0),$$

in which the  $L_i$  are certain differential operators and  $x_0$  contains a priori values for the unknowns. The parameters  $\alpha_i$  regulate the trade-off between signal and noise constraints. The operators  $L_i$  may be chosen such as to minimize the size of the unknowns (identity operator), or derivatives for instance. The LS problem above is a special case with  $\alpha_i = 0$ . Note that each  $m$ -block may contain its own set of constraints.

Narrowing down to quadratic constraints only, one seeks those unknowns  $x$  that minimize the hybrid norm:

$$\hat{x} : \min_x \left\{ \underline{e}^\top P_y \underline{e} + \alpha (x - x_0)^\top P_{x_0} (x - x_0) \right\}, \quad (6.4)$$

i.e. a simultaneous minimization of discrepancies, due to measurement and model noise, and of signal size. In other words, within the margins set by the observation model and observation noise, the signal closest to  $x_0$ —in RMS sense and weighted by  $P_{x_0}$ —must be found. The symmetric and positive definite matrix  $P_{x_0}$  denotes the *regularization* matrix, that stabilizes the ill-conditioned normal matrix. The following examples may clarify the regularization:

- i) If no further information on the unknowns  $x$  is known a priori, the most simple choice is to expect them to be zero ( $x_0 = 0$ ) and of equal variance ( $P_{x_0} = \alpha I$ ). In combination with the weighted residuals the size of the unknowns ( $x^\top x$ ) is minimized as well.
- ii) In case of gravity field estimation signal degree variances, e.g. Kaula's rule, may be employed as a priori information. This conventional choice expects the coefficients to be zero and to vary according to the signal variance. It leads to  $x_0 = 0$  and  $P_{x_0} = \alpha K^{-1}$ , with  $K$  a diagonal matrix with signal degree-order variances on its diagonal.
- iii) The previous hybrid norm does not make use of existing gravity field solutions, say  $\underline{x}_0$ , with corresponding error-covariance matrix  $Q_x$ . Using this information leads to  $P_{x_0} = Q_{x_0}^{-1}$ . Note that in this case  $\underline{x}_0$  is a random variable as well, which was not strictly the case in the previous examples.
- iv) The next step is to combine both a priori signal and a priori error information into the norm:

$$\hat{x} : \min_x \underline{e}^\top P_y \underline{e} + \alpha_1 (x - \underline{x}_0)^\top P_{x_0} (x - \underline{x}_0) + \alpha_2 x^\top K^{-1} x.$$

For convenience this will still be written as (6.4) above, with the understanding that  $P_{x_0}$  contains both signal and error (co-)variance and that  $x_0$  equals zero for coefficients that are not known a priori. The  $\alpha$  coefficients are either 1 or are assumed to be incorporated in  $P_{x_0}$ .

**Sequential Estimation.** The latter regularization option is easily understood under the framework of *sequential* or *recursive* estimation. The prior knowledge  $\underline{x}_0$  with covariance  $Q_{x_0}$  will be inconsistent with data  $\underline{y}$  with covariance  $Q_y$ . This inconsistency, expressed by the *innovation* or *measurement update*  $\underline{y} - A\underline{x}_0$ ,

necessitates a new solution, that is the unknowns are to be updated. Thus the above regularization can be viewed as a single step in a sequential estimation procedure. It can be extended, as soon as further information comes in.

The sequential estimator is simply derived by extending the original linear model (6.1) by pseudo-observations of the unknowns:  $E\{\underline{x}_0\} = x$ ,  $D\{\underline{x}_0\} = Q_{x_0}$ . Conventionally, the pseudo-observations are zero—hence also known as *zero observations*—with a covariance matrix from a signal degree variance, (Reigber & Ilk, 1976; Schwintzer, 1990; Bouman, 1997b). However, as mentioned before, using full a priori gravity field information yields pseudo-observations  $\underline{x}_0$ , that only happen to become zero for non-existent a priori coefficients, with  $Q_{x_0}$  containing both noise and signal contributions. The augmented linear model reads:

$$E \left\{ \begin{pmatrix} \underline{y} \\ \underline{x}_0 \end{pmatrix} \right\} = \begin{pmatrix} A \\ I \end{pmatrix} x \quad , \quad D \left\{ \begin{pmatrix} \underline{y} \\ \underline{x}_0 \end{pmatrix} \right\} = \begin{pmatrix} Q_y & 0 \\ 0 & Q_{x_0} \end{pmatrix} , \quad (6.5)$$

which is solved by the estimator:

$$\hat{x} = (A^\top P_y A + P_{x_0})^{-1} (A^\top P_y \underline{y} + P_{x_0} \underline{x}_0). \quad (6.6a)$$

Expectation and dispersion of the estimator are:

$$\begin{aligned} E\{\hat{x}\} &= (A^\top P_y A + P_{x_0})^{-1} (A^\top P_y E\{\underline{y}\} + P_{x_0} E\{\underline{x}_0\}) \\ &= (A^\top P_y A + P_{x_0})^{-1} A^\top P_y A x + (A^\top P_y A + P_{x_0})^{-1} P_{x_0} x = x, \end{aligned} \quad (6.6b)$$

$$D\{\hat{x}\} = (A^\top P_y A + P_{x_0})^{-1} = Q_{\hat{x}}. \quad (6.6c)$$

The expectation of (6.6a) equals  $x$  itself, i.e.  $\hat{x}$  is an unbiased estimator.

It is avoided in sequential estimation to build the augmented linear system. Re-processing of old data is unnecessary, as long as the most recent information on  $x$  and  $Q_x$  is available—which also explains the name *information* matrix for the normal matrix. The above solution is rewritten as an update process now, using eqns. (6.6a) and (6.6c), yielding:

$$\text{covariance update: } P_{\hat{x}} = P_{x_0} + A^\top P_y A \quad (6.7a)$$

$$Q_{\hat{x}} = P_{\hat{x}}^{-1} \quad (6.7b)$$

$$\begin{aligned} \text{parameter update: } \hat{x} &= Q_{\hat{x}} (A^\top P_y \underline{y} + P_{x_0} \underline{x}_0) \\ &= Q_{\hat{x}} (P_{\hat{x}} \underline{x}_0 - A^\top P_y A \underline{x}_0 + A^\top P_y \underline{y}) \\ &= \underline{x}_0 + Q_{\hat{x}} A^\top P_y (\underline{y} - A \underline{x}_0) \\ &= \underline{x}_0 + K (\underline{y} - A \underline{x}_0) \end{aligned} \quad (6.7c)$$

$$\text{covariance using } K: Q_{\hat{x}} = (I - KA) Q_{x_0} \quad (6.7d)$$

The latter expression reveals that the estimator uses the innovation  $\underline{y} - A \underline{x}_0$  to update the a priori  $\underline{x}_0$ . It is multiplied by the so called *gain* matrix  $K = Q_{\hat{x}} A^\top P_y$ , cf. (Bierman, 1977; Strang, 1986). The gain matrix determines how new data, i.e. at the level of observations  $\underline{y}$ , flows into the new estimate.

**Collocation.** The regularized solution is formally the same as a *collocation solution* (Reigber & Ilk, 1976; Xu, 1992; Bouman, 1997b). This becomes clear

through the equality, e.g. (Moritz, 1980):

$$(A^T P_y A + P_{x_0})^{-1} A^T P_y \equiv Q_{x_0} A^T (A Q_{x_0} A^T + Q_y)^{-1}. \quad (6.8)$$

This equality is easily proven by first pre-multiplying left and right by  $(A^T P_y A + P_{x_0})$  and then post-multiplying by  $(A Q_{x_0} A^T + Q_y)$ . Note that (6.8) actually describes the gain matrix  $K$ . Strictly speaking, for the collocation interpretation  $x_0$  should be zero and  $Q_{x_0}$  should be a signal covariance only in (6.6a). Moreover, if the following notational substitutions are used:

$$\begin{aligned} \text{signal covariance } x : & \quad Q_{x_0} \rightarrow C_{xx} \\ \text{signal covariance } y : & \quad A Q_{x_0} A^T \rightarrow C_{yy} \\ \text{cross covariance } xy : & \quad Q_{x_0} A^T \rightarrow C_{xy} \\ \text{noise covariance } y : & \quad Q_y \rightarrow D_{yy} \end{aligned}$$

the collocation estimator (6.6a) is more readily expressed by:

$$\hat{\underline{x}}_c = C_{xy} (C_{yy} + D_{yy})^{-1} \underline{y}. \quad (6.9)$$

## 6.2 Contribution and Redundancy

**Relative Contributions.** Let us return to the expectation (6.6b) and rewrite it using the following definitions:

$$R_x = (A^T P_y A + P_{x_0})^{-1} P_{x_0}, \quad (6.10a)$$

$$R_y = (A^T P_y A + P_{x_0})^{-1} A^T P_y A, \quad (6.10b)$$

yielding:

$$E\{\hat{\underline{x}}\} = R_y x + R_x x = x.$$

The above matrices, that obviously add up to the unit matrix:  $R_x + R_y = I$ , are known as *resolution* matrices, (Jackson, 1972,1979; Dong, Herring & King, 1998). They measure the relative contributions of data and prior knowledge to the solution (6.6a).  $R_y$  indicates the share of the data,  $R_x$  that of the prior knowledge. The resolution matrices might as well be regarded as filters through which the vector  $x$  passes to yield the estimator  $\hat{\underline{x}}$ , see also (Bouman, 1997b). Note that the resolution matrix is constructed neither with actual data nor with actual a priori values for the parameters. Thus it is an important tool for evaluating the necessity of regularization by adding a priori knowledge. With this tool those parameters can be identified, that are not or poorly estimated—or updated in sequential adjustment terms—by the data, i.e. those, whose eventual estimates depend fully or heavily on the quality of the a priori knowledge.

**Remark 6.1** *As opposed to the gain matrix, the resolution matrix operates at the level of parameters  $x$ . The relation between both:  $R_y = K A = I - R_x$ . This also leads to:  $Q_{\hat{x}} = R_x Q_{x_0}$ .*

Concentrating in (6.10) on the main diagonal only, yields a contribution measure for individual parameters  $x_i$ . Due to  $R_x = Q_{\hat{x}} P_{x_0}$  and assuming a diagonal  $P_{x_0}$ ,

it is:

$$r_{x_i} = [R_x]_{ii} = \frac{\sigma_{\hat{x}_i}^2}{\sigma_{x_{0,i}}^2}, \quad (6.10c)$$

$$r_{y_i} = [R_y]_{ii} = 1 - r_{x_i}. \quad (6.10d)$$

So the share of a priori knowledge in the  $i$ -th estimated parameter is the ratio of its a posteriori variance to its a priori variance. The share of data is the complement to one. If, for example, the data does not contribute to parameter  $x_i$ , its a posteriori variance will be as good or as bad as before. The variance ratio, i.e. the share of a priori info, becomes one. The data contribution  $r_{y_i}$  becomes consequently zero.

**Redundancy Analysis.** A simple dimensional analysis of the design matrices reveals a certain paradox. Assume the number of observations, i.e. the length of  $y$ , is  $n$  and the number of unknowns equals  $m$ . The design matrix  $A$  has dimensions  $n \times m$ , yielding  $n - m$  degrees of freedom, since  $A$  has full rank. However, for a poorly conditioned  $A$ , this number should be decreased. The augmented linear system (6.5) has dimensions  $(n + m) \times m$ . The augmented system is stable and has full rank. Thus it contains  $n$  degrees of freedom.

The augmented linear system is considered now as a system with  $m$  side constraints, namely from a priori knowledge. Dong & al. (1998) argue that two extreme situations may arise. If the side constraints are weak,  $Q_{x_0} \rightarrow \infty$ , one arrives at a solution  $\hat{x}$  that would be the same as a solution from the original linear model (6.1), having  $n - m$  degrees of freedom. If, however, the a priori knowledge imposes strong side constraints,  $Q_{x_0} \rightarrow 0$ , the solution will be the one from the augmented linear system, having  $n$  degrees of freedom.

Thus the degrees of freedom depend on  $Q_{x_0}$ . It is convenient at this point to speak of the *effective* number of degrees of freedom, say  $n - m + k$ , with  $0 \leq k \leq m$ . Under weak constraints the data determines the solution:  $k \rightarrow 0$ . Strong constraints cause  $k \rightarrow m$ . Thus  $k$  is interpreted as the effective amount of constraints added by the a priori knowledge, or the effective added redundancy. This amount is not necessarily an integer. Now the terms *weak* or *strong* constraints must be measured relative to data content. Therefore it is logical to use the above resolution matrices to define the added redundancy. Noticing that  $R_x + R_y$  equals the unit matrix of dimension  $m$ , and defining  $\text{tr}R_x = r_x$  and  $\text{tr}R_y = r_y$ , one obtains:  $r_x + r_y = m$ . Based on (Dong & al., 1998) the following definition is proposed:

$$k = \text{tr}R_x = r_x = m - r_y. \quad (6.11)$$

This definition assures obviously  $0 \leq k \leq m$ . Moreover, the contributions for individual parameters (6.10c) can be interpreted now as partial redundancies:  $r_x = \sum_{i=1}^m r_{x_i}$ . Note that  $r_x$  has been defined in parameter space. This allows a further interpretation of  $r_x$  as the dimension of the estimated parameter space due to constraints. Its complement  $r_y = m - r_x$  is the effective dimension of estimated parameter space, due to data.

Schwintzer (1990) defines partial redundancies of the pseudo-observations in observation space. His terminology stems from reliability theory. Assume the simple

linear model (6.1) with  $A$  full rank. The estimator  $\hat{x}$  induces estimated observations  $\hat{y} = A\hat{x}$  and residuals  $\hat{e} = y - \hat{y}$ . The sum of squares of weighted residuals have expectation (with variance of unit weight equal to one):

$$\begin{aligned} E\{\hat{e}^\top P_y \hat{e}\} &= n - m = r \\ &= \text{tr}(Q_{\hat{e}} P_y) = \text{tr}(I - Q_{\hat{y}} P_y). \end{aligned} \quad (6.12)$$

The local redundancy number  $r_i$  of a single observation  $y_i$  thus becomes:

$$r_i = 1 - \frac{\sigma_{\hat{y}_i}^2}{\sigma_{y_i}^2},$$

whose sum  $\sum_i r_i$  adds up to  $r$ . The local redundancy number is employed to define the minimal detectable bias of an observation  $y_i$ , which is not pursued here further. Now Schwintzer (1990) argues that for the partial redundancy of pseudo-observations, which are observations too after all, it holds:

$$r_i = 1 - \frac{\sigma_{\hat{x}_i}^2}{\sigma_{x_{0,i}}^2}.$$

Comparison with (6.10) shows the remarkable result  $r_i = 1 - r_{x_i}$ . Schwintzer's partial redundancies are identified as the *complement* of the above contribution definition (6.10c). It is the data contribution, instead of the contribution from a priori information. Schwintzer's misinterpretation becomes clear in view of the above effective redundancy discussion, see also (Bouman, 1997b). The effective redundancy is estimated by inserting  $Q_{\hat{x}}$  from (6.6a) into (6.12). However only the estimated residuals of the  $n$  observations  $\underline{y}$  are considered, i.e. the ones from the  $m$  pseudo-observations  $\underline{x}_0$  are not.

$$\begin{aligned} E\{\hat{e}^\top P_y \hat{e}\} &= \text{tr}(I - Q_{\hat{y}} P_y) \\ &= \text{tr}(I - A Q_{\hat{x}} A^\top P_y) \\ &= n - \text{tr}(Q_{\hat{x}} A^\top P_y A) \\ &= n - r_y = n - m + r_x. \end{aligned} \quad (6.13)$$

Use is made of the identity  $\text{tr}(AB) = \text{tr}(BA)$ . This derivation justifies the choice  $k = r_x$  in the above redundancy discussion.

## 6.3 Biased Estimation

FOR SEVERAL reasons the estimator (6.6a) cannot be considered unbiased in general. These include regularization without prior knowledge, the use of wrong or low-quality prior information, and the existence of unmodelled parameters.

**Biased Estimation.** In case of unwarranted zero pseudo-observations or with wrong a priori information, Xu (1992) argues that the regularization solution (6.6a) must be seen in the framework of biased estimation. Indeed, in case the data does not contribute to the estimator (6.6a), as measured by  $R_y$ , the solution

would be determined by the prior information. It would fully depend on the quality of the a priori values, especially on the validity of the expectation  $E\{\underline{x}_0\} = x$ . In case of zero pseudo-observations, setting  $x_0 = 0$  gives the biased estimator directly. But also in case of wrong prior information, setting  $x_0 = 0$  is allowed in view of the sequential estimation interpretation. Thus one arrives at the biased estimator:

$$\hat{\underline{x}}_b = (A^\top P_y A + P_{x_0})^{-1} A^\top P_y \underline{y}, \quad (6.14a)$$

with expectation and dispersion:

$$\begin{aligned} E\{\hat{\underline{x}}_b\} &= (A^\top P_y A + P_{x_0})^{-1} A^\top P_y E\{\underline{y}\} \\ &= (A^\top P_y A + P_{x_0})^{-1} A^\top P_y A x = R_y x, \end{aligned} \quad (6.14b)$$

$$\begin{aligned} D\{\hat{\underline{x}}_b\} &= (A^\top P_y A + P_{x_0})^{-1} A^\top P_y A (A^\top P_y A + P_{x_0})^{-1} \\ &= R_y Q_{\hat{x}} = Q_{\hat{x}_b}. \end{aligned} \quad (6.14c)$$

Inserting  $I - R_x$  for  $R_y$  renders more explicitly the character of the biased solution:

$$\begin{aligned} E\{\hat{\underline{x}}_b\} &= x - R_x x \neq x, \\ D\{\hat{\underline{x}}_b\} &= Q_{\hat{x}} - R_x Q_{\hat{x}} \neq Q_{\hat{x}}. \end{aligned}$$

From  $R_x \leq I$  and  $R_y \leq I$  it follows  $\|E\{\hat{\underline{x}}_b\}\| \leq \|x\|$ . The estimate (6.14a) is biased towards zero, i.e. smoothed, cf. (Xu, 1992). Consequently the power, or signal variance, of the biased estimate is reduced. Moreover the a posteriori covariance  $Q_{\hat{x}_b}$  will be too optimistic.

The bias in the estimate reads:

$$b = E\{\hat{\underline{x}}_b - x\} = -R_x x. \quad (6.15)$$

In order to account for bias in an a posteriori error measure, Xu (1992) proposes the so-called Mean Square Error matrix, denoted by  $M$  here:

$$M_{\hat{x}_b} = Q_{\hat{x}_b} + b b^\top = Q_{\hat{x}} - R_x Q_{\hat{x}} + R_x x x^\top R_x^\top. \quad (6.16)$$

Naturally the correct unknowns  $x$  are unknown indeed. Therefore it is not feasible to compute (6.15) or (6.16) exactly. At best it may be approximated by substituting the (under-)estimate  $\hat{\underline{x}}_b$  for  $x$ , or taking coefficients from an existing (high degree) gravity field model. In case of approximating the mean square error, one might alternatively insert here a signal variance model, i.e.  $M\{x x^\top\}$ , with  $M\{\dots\}$  the averaging operator. This yields a proper indication of the bias induced error *on the average*. In the special case that the a priori information comes from a signal variance only, i.e. does not include error covariance from previously derived SH models, the mean square error  $M$  (6.16) reduces to  $Q_{\hat{x}}$  again. With  $M\{x x^\top\} = Q_{x_0}$  the bias part becomes:

$$M\{R_x x x^\top R_x^\top\} = (A^\top P_y A + P_{x_0})^{-1} P_{x_0} Q_{x_0} P_{x_0} (A^\top P_y A + P_{x_0})^{-1} = R_x Q_{\hat{x}}.$$

Even in case the a priori information does contain an error variance component, the covariance matrix  $Q_{\hat{x}}$  (6.6c) will represent a more realistic estimate of the mean square error than  $Q_{\hat{x}_b}$ . In spite of being unable to calculate the bias itself, the resolution matrix  $R_x$  is a clear indicator of the level of bias to be expected. Indeed, since  $R_x$  expresses the dependence of the solution on a priori information, it reflects the biasedness towards this a priori information, be it warranted information or not.

**Unresolved Parameters.** A lumped coefficient  $A_{mk}^\#$  is an infinite sum over degree  $l$ , cf. (3.1b). So the observable  $\underline{y}$  contains signal from the full SH spectrum, including the part above maximum degree  $L$ . Note that  $L$  might depend on  $m$ , as explained in 5.1. The data amount is finite, though. Consequently the spectral content for  $l > L$  will in general be mapped onto the estimated coefficients.

The mapping effect is investigated by extending the linear model, cf. (Sneeuw & Van Gelderen, 1997):

$$\mathbb{E}\{\underline{y}\} = (A_1 \ A_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad (6.17)$$

in which  $x_2$ —the unresolved coefficients—are in principle up to infinity although a suitable high degree will be sufficient. The matrix  $A_2$  contains transfer coefficients  $H_{lmk}^\#$  with  $l > L$ . Let  $P_{x_0}$  be zero for simplicity. Taking (6.3a) as an estimator results in:

$$\hat{x}_1 = (A_1^\top P_y A_1)^{-1} A_1^\top P_y \underline{y}, \quad (6.18a)$$

where  $x_1$  is written now, in order to indicate that only coefficients up to  $L$  are estimated. Inserting (6.17) results in:

$$\mathbb{E}\{\hat{x}_1\} = x_1 + (A_1^\top P_y A_1)^{-1} A_1^\top P_y A_2 x_2 \neq x_1, \quad (6.18b)$$

$$\mathbb{D}\{\hat{x}_1\} = (A_1^\top P_y A_1)^{-1} = Q_{\hat{x}_1}, \quad (6.18c)$$

showing that  $x_2$  will project onto the estimate of  $x_1$ . How much, depends on the level of orthogonality in  $A_1^\top P_y A_2$ . Moreover, it is shown that the a posteriori covariance is not influenced by a contribution from  $x_2$ .

A proper terminology for the effect of unestimated  $x_2$  contributions is not self-evident. Firstly, due to  $\mathbb{E}\{\hat{x}_1\} \neq x_1$  it has the character of *biased* estimation. Secondly, since certain spectral components are projected onto others, *aliasing* seems to be a proper name. However, aliasing is usually associated with a clear mapping from single spectral lines onto others. See also the discussion on aliasing (classical and torus-to-orbit) in 5. Here the projection takes place through  $A_1^\top P_y A_2$ , i.e. the  $x_2$  part will be smeared over  $x_1$  in case  $A_1^\top P_y A_2$  is not orthogonal. In this view spectral *leakage* would be a better name. Still, also this name might be misleading, since (6.18b) is due to a spectral cut-off, which leads to spatial leakage in classical Fourier theory.

**Remark 6.2** *In 5.2 two effects of cutting-off the SH domain at maximum degree  $L$  were mentioned: spatial leakage and omission error. The above phenomenon is an additional effect.*

Denoting the  $x_2$ -effect  $a$  for aliasing (with the above terminology reservation in mind), it is:

$$a = \mathbb{E}\{\hat{x}_1 - x_1\} = (A_1^\top P_y A_1)^{-1} A_1^\top P_y A_2 x_2. \quad (6.19)$$

The matrix  $(A_1^\top P_y A_1)^{-1} A_1^\top P_y A_2$  bears resemblance to the resolution or contribution matrices (6.10), although it is neither square nor symmetric. It determines the contribution of  $x_2$ , i.e. the seriousness of the aliasing. Defining

$$R_a = (A_1^\top P_y A_1)^{-1} A_1^\top P_y A_2, \quad (6.20)$$

one obtains  $a = R_a x_2$  with second moment:

$$M\{aa^\top\} = R_a M\{x_2 x_2^\top\} R_a^\top.$$

In analogy to the biased estimation case, the aliasing effect cannot be estimated, since  $x_2$  is unknown. Also here, one can approximate the effect either by inserting an existing high degree gravity field or by employing a signal variance model. An indicator for the seriousness of the aliasing is the matrix  $A_1^\top P_y A_2$ , or, in order to comply with the resolution matrices, the matrix  $R_a$ .

## 6.4 Least Squares Error Simulation

THE PREVIOUS section described the estimation procedure, including several quality measures, of which the a posteriori covariance matrix of estimated unknowns was pre-eminent. Calculation of the covariance matrix, but also of the resolution matrices and of the second moments of bias and alias, does *not* require actual data, though. Input is mainly the linear model ( $A$ ), the stochastic model of the observation ( $Q_y$ ) and the a priori information ( $Q_{x_0}$ ). Based on these input quantities the above a posteriori quality measures can be derived *prior* to observation. This calculation is denoted as *least squares error simulation*. It is an excellent tool for pre-mission analysis, like orbit design or measurement hardware specification.

**Multi-Observables Model.** Suppose that several observation types  $y_n$  are available, either from simultaneous measurements, e.g. a full tensor gradiometry mission, or sequentially, e.g. from different satellite missions. The linear model (6.1) can be extended to

$$E\{\underline{y}_n\} = A_n x, \quad D\{\underline{y}_n\} = Q_n, \quad n = 1, 2, \dots$$

or

$$E\left\{\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \end{pmatrix}\right\} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \end{pmatrix} x, \quad D\left\{\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \end{pmatrix}\right\} = \begin{pmatrix} Q_1 & & \\ & Q_2 & \\ & & \ddots \end{pmatrix}. \quad (6.21)$$

The a priori information—in the form of pseudo-observations—can be just one of these data sets, cf. (6.5). Thus the extended model (6.21) already incorporates regularization.

For each set of observables a normal matrix is formed:

$$N_n = A_n^\top P_n A_n.$$

The least squares solution now reads:

$$\hat{x} = N^{-1} \left( \sum_n A_n^\top P_n \underline{y}_n \right), \quad \text{with } N = \sum_n N_n. \quad (6.22a)$$

Expectation and dispersion are:

$$E\{\hat{x}\} = x, \quad (6.22b)$$

$$D\{\hat{x}\} = N^{-1} = Q_{\hat{x}}. \quad (6.22c)$$

**Dimensioning.** Equations (6.22a) and (6.22c)—and later on (6.23)—show that least squares error simulation mainly consists of operations on normal matrices: stacking, multiplication and inversion. The question arises whether heterogeneous data sources can be assembled in a numerically stable way, considering order of magnitude of signal and of noise. For example, does one have to take special efforts in order to combine gradiometry with orbitography and a priori information?

In 4 dimensioning terms were introduced for each type of functional. These factors allow a direct comparison of heterogeneous observables, expressing them as dimensionless quantities of the same order of magnitude. However, if the observables are made dimensionless, the factors have to be applied to the observation noise as well.

To be concrete, suppose that the observable  $\underline{y}_n$  is made dimensionless by a factor  $d_n$ , e.g. taken from tbl. 4.1, through  $\underline{y}'_n = d_n^{-1}\underline{y}_n$ . The linear model (6.21) transforms into

$$\begin{aligned} E\{\underline{y}'_n\} &= d_n^{-1}A_n x = A'_n x, \\ D\{\underline{y}'_n\} &= d_n^{-2}Q_n = Q'_n. \end{aligned}$$

Now, since the weight matrix becomes  $P'_n = d_n^2 P_n$ , the normal matrices remain:

$$N_n = A_n^T P_n A_n = A_n'^T P_n' A_n'.$$

Thus it is seen that heterogeneous data can be combined and analyzed *without* further dimensioning efforts. The normal matrix is—stated otherwise—a dimensionless object already (in case the unknowns are dimensionless quantities). It can be viewed as a signal-to-noise ratio.

## 6.5 Pre-Mission Analysis Types

**I**N THE preceding sections all the necessary tools for gravity field recovery by least squares inversion have been derived. In 6.4 the goal was narrowed down to pre-mission error assessment. Based on the normal matrix, the following pre-mission analysis tools are relevant.

**I. A Posteriori Covariance.** Inverting the total normal matrix  $N$  yields the covariance matrix  $Q_{\hat{x}}$  of estimated parameters, cf. (6.22c). This is the basic output of LS error simulation. In particular the square root of the diagonal represents the standard deviations  $\sigma_{lm}$  of single coefficients. The full set of  $\sigma_{lm}$  represents the *spherical harmonic error spectrum*. Section 5.4 described the mapping of stochastic information between 1D and 2D Fourier domains. By the process of LS error simulation, the mapping onto the SH domain is achieved.

**II. Redundancy and Contribution.** For each subset  $\underline{y}_n$  a contribution matrix is defined:

$$R_n = N^{-1}N_n, \quad (6.23)$$

for which obviously  $\sum_n R_n = I$  holds. Note again, that the a priori information might be contained in the model already as one of the subsets. Thus the resolution matrix  $R_x$  (6.10a) is just one of the  $R_n$  now.

The diagonal of  $R_n$  returns the partial contribution of the  $n$ 'th subset to single SH coefficients  $\bar{K}_{lm}$ . For each coefficient one obtains a contribution number  $r_{lm}^n$ , the full set of which could be called a SH contribution spectrum from observation set  $\underline{y}_n$ . The trace of  $R_n$  is interpreted, cf. 6.2, as the effective dimension of parameter space. It is a very compact measure of the information content of data  $\underline{y}_n$  to the estimated coefficients of a specific order  $m$ .

Moreover, since the resolution matrices represent information content, they also indicate potential lack of information. As a matter of consequence they signalize those coefficients affected by bias, cf. 6.3. See also the corresponding paragraph on bias and alias assessment.

**III. Eigenanalysis.** The eigenvalue spectrum reveals the level of ill-posedness of the normal matrix. As explained in 6.1, the eigenvalues are expected to decay gradually. If the eigenspectrum drops to zero suddenly rank deficiencies exist. Dependencies between parameters can then be uncovered by the eigenvectors. Note that eigenvalue decomposition is the only analysis tool that can be performed in case  $N$  is not invertible.

A key parameter in assessing the stability of the normal matrix is the *condition number*, which is the ratio between largest and smallest eigenvalues. Its logarithm determines the loss of valid digits due to noise amplification (Lanczos, 1964). Each normal matrix yields one such condition number. In case of even-odd separation of coefficients, two condition numbers pertain to each order  $m$ .

Eigenanalysis comes in through a different path as well, namely in deriving a measure for correlation between unknowns. The SH error spectrum  $\sigma_{lm}$  reflects the diagonal of  $Q_{\hat{x}}$ , i.e. the variances only. The covariance and correlation between estimated parameters are neglected. Ideally, the a posteriori unknowns are uncorrelated, corresponding to a purely diagonal  $Q_{\hat{x}}$ . The question therefore is, to what level the real  $Q_{\hat{x}}$  and its diagonal-only counterpart  $Q_{\hat{x}}^d$  are similar. This question is expressed mathematically by the *generalized eigenvalue problem*:

$$Qx = \lambda Q^d x .$$

The maximum generalized eigenvalue describes the maximum amplification, due to non-similarity, that is, it indicates how serious the neglecting of off-diagonals can be. Ideally it should be  $\lambda_{\max} = 1$ .

Parameters may vary strongly in size as may their error variances. More worthwhile, therefore, would be to compare the *correlation* matrix with its diagonal-only counterpart, the unit matrix. Thus the generalized eigenvalue problem reduces to the following ordinary one:

$$Cx = \lambda C^d x = \lambda x , \text{ with } c_{ij} = \frac{q_{ij}}{\sqrt{q_{ii}q_{jj}}} .$$

The maximum eigenvalue, again, indicates to what level the correlation matrix  $C$  resembles the unit matrix. Ideally  $\lambda_{\max} = 1$  again, but if  $C$  is far from diagonal, it will be large. The maximum eigenvalue is also known as the *norm* of the matrix.

A proper *correlation measure* should become zero if  $C = I$  and should approach one for increasing  $\lambda_{\max}$ . Thus the following quantity is proposed:

$$c = 1 - \frac{1}{\lambda_{\max}} \text{ or } 1 - \frac{1}{\text{norm}(C)}. \quad (6.24)$$

Writing

$$c = \frac{\text{norm}(C) - 1}{\text{norm}(C)} = \frac{\text{norm}(C - I)}{\text{norm}(C)}$$

shows that this measure can also be derived as the ratio of norms of the correlation matrix with its diagonal stripped off versus the norm of  $C$  itself.

**IV. Bias and Alias Assessment.** As remarked in 6.3 the Mean Square Error matrix  $M$  is more realistically represented by  $Q_{\hat{x}} = N^{-1}$  than by  $Q_{\hat{x}_b} = R_y Q_{\hat{x}}$ . Moreover, although the bias contribution remains unknown, the resolution matrix of prior knowledge  $R_x$  indicates bias-involved parameters already. The bias  $b = -R_x x$  and its contribution  $bb^T$  to  $M$  may be derived using:

- i) the estimate  $\hat{x}$  for  $x$ , although it is known to be an under-estimate,
- ii) existing high-degree gravity field models,
- iii) signal degree variances (only applicable to  $bb^T$ ).

A similar assessment is proposed for the aliasing effect, due to unresolved parameters. The matrix  $R_a$  (6.20) indicates the level of cross-spectrum mapping. Note that rather the off-diagonals are to be considered here. The aliasing  $a = R_a x_2$  and its contribution  $aa^T$  to the a posteriori error measure may be derived by the procedures ii) and iii) above.

**V. Sensitivity Analysis.** One analysis type that fits into this list, only due to the similarity of the employed formulae, is *sensitivity analysis*. Based on the formulation

$$f = \sum_{l,m,k} H_{lmk}^f \bar{K}_{lm} e^{i\psi_{mk}},$$

the question arises how sensitive the functional  $f$  is to a single SH coefficient  $\bar{K}_{lm}$ . Since

$$\|e^{i\psi_{mk}}\| = \sqrt{e^{i\psi_{mk}} e^{i\psi_{mk}^*}} = 1,$$

the question reduces to the following: to what extent does  $\bar{K}_{lm}$  contribute to the norm of  $f$  through the transfer  $H_{lmk}^f$ ?

$$\|f_{lm}\| = \left\| \sum_k H_{lmk}^f \right\| \|\bar{K}_{lm}\| \leq \sqrt{\sum_k |H_{lmk}^f|^2} \|\bar{K}_{lm}\| = s_{lm} \|\bar{K}_{lm}\|, \quad (6.25a)$$

$$\text{with } s_{lm} = \sqrt{\sum_k |H_{lmk}^f|^2}. \quad (6.25b)$$

The number  $s_{lm}$  represents the required sensitivity. Equation (6.25a) shows that  $s_{lm}$  is only an upper bound, due to the quadratic summation over  $k$ . Applied to orbit perturbations, Reigber (1989) refers to  $s_{lm}$  correctly as *maximum orbit perturbation*, whereas Balmino (1993) uses the phrase RMS *orbit perturbation*. The latter would better be renamed in RSS (root sum square) orbit perturbation

to underline the implied summation. In matrix notation (6.25b) is nothing else than

$$s_{lm} = \sqrt{\text{diag}(A^T A)}. \quad (6.25c)$$

No stochastic model is involved here. Sensitivity analysis is a forward modelling tool, i.e. a projection of unknowns on the observable, as opposed to an inverse modelling tool, such as the former analysis types. The quantity  $s_{lm}$  does not supply information on recoverability of the corresponding potential coefficients. The SH perturbation in (Reigber, 1989), for example, may not be interpreted in the sense that the coefficients involved can be recovered from SLR measurements to the mentioned satellite. Neither can the measure of separability in (Rummel, 1993) be used to predict which coefficients can be obtained from altimetry. This measure is introduced as a ratio of radial perturbation sensitivity and geoid—or better: sea surface topography—sensitivity. Recoverability and separability are inverse problems with intricate spectral and numerical characteristics.

## 6.6 Spectral Error Representation

THE LEAST squares error simulation, described in 6.4, yields a block-diagonal covariance matrix  $Q_{\hat{x}}$ , or  $\text{COV}\{\bar{K}_{lm}, \bar{K}_{nm}\}$ , of the estimated unknowns. It may either represent the observation noise propagation only, or may also include bias and aliasing effects, cf. 6.4. The block structure remains, though.  $Q_{\hat{x}}$  expresses the basic spectral variance-covariance structure of the SH coefficients. However, it is not apt to further graphical display. Thus, various 2D and 1D error measures are derived now, that are better suited for visualizing and comparison purposes.

**Two-Dimensional SH Error Spectrum.** Taking the diagonal from  $Q_{\hat{x}}$  yields the variances per coefficient:

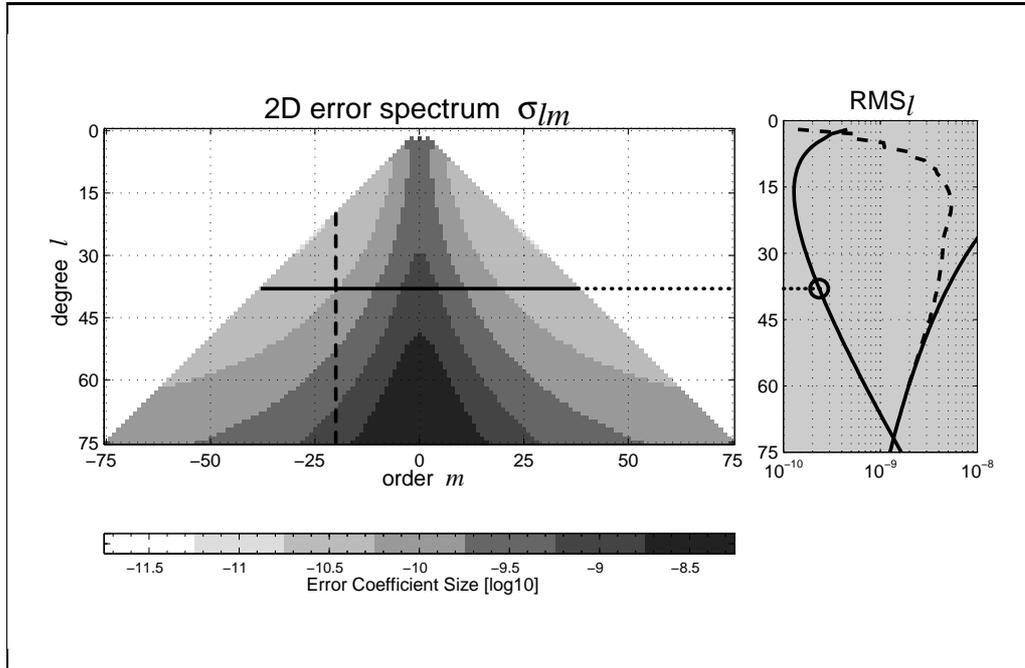
$$\text{diag}(Q_{\hat{x}}) \mapsto \text{VAR}\{\bar{K}_{lm}\} = \sigma_{lm}^2, \quad (6.26)$$

resulting in a 2D spherical harmonic error spectrum  $\sigma_{lm}$ . One could say that from each  $m$ -block of  $Q_{\hat{x}}$  the (square root of the) diagonal is transformed into a column vector, like the dashed line in fig. 6.1. This line signifies at the same time those coefficients that will be correlated: a consequence of the lumped coefficient approach ( $\sum_l$ ). The 2D error spectrum is not necessarily triangular as in fig. 6.1. See the discussion on SH domain options in 5.2 and fig. 5.3.

Although correlations within  $m$ -blocks have been neglected, the 2D error spectrum  $\sigma_{lm}$  constitutes an error measure, that displays a wealth of information. Nevertheless, for comparison purposes the information content is still too much. Especially the quantitative comparison of several SH error spectra—from simulations, real satellite missions, existing gravity field, etc.—requires a more condensed 1D error measure.

**One-Dimensional SH Error Spectrum.** The usual 1D error measure is the *error degree variance*:

$$\sigma_l^2 = \sum_{m=-l}^l \sigma_{lm}^2. \quad (6.27)$$



**Figure 6.1:** SH error spectrum  $\sigma_{lm}$ . The vertical dashed line signifies the coefficients of a certain order  $m$ , that will be correlated. A corresponding 1D error spectrum (RMS or variance) is attained by horizontal averaging. See text for further explanation.

Degree variances  $\sigma_l^2$ , also often denoted as  $\epsilon_l$  (without the square), denote the total error power of a certain degree  $l$ . Due to the unitary transformation of spherical harmonics, cf. (2.9),  $\sigma_l^2$  is invariant under rotations, i.e. it is independent from the definition of the coordinate definition on the sphere. Though theoretically meaningful, this argument is of less use in cases of gravity fields, determined by satellite methods. The nominal satellite orbit configuration does, through the Earth's rotation axis, have a preferred direction. Only in longitude no specific preference exists, leading to a phase uncertainty, say  $\lambda_0$ , in the least squares error analysis. But apart from that, the 2D error spectrum must be considered as basic SH error measure, and the corresponding degree variances (1D error measures in general) a derived quantity.

From the error power per degree a further 1D error measure is derived. Since the number of coefficients per degree is  $(2l + 1)$ , the *mean* variance per degree, also known as *degree-order variance*, is  $\sigma_l^2 / (2l + 1)$ . Its square root will be denoted root-mean-square per degree ( $\text{RMS}_l$ ). It is the average standard deviation to be expected for a single coefficient. Average must be understood as the mean over a specific degree, as graphically explained by the horizontal line in fig. 6.1. We have:

$$\text{RMS}_l = \sqrt{\frac{1}{2l+1} \sigma_l^2} = \sqrt{\frac{1}{2l+1} \sum_{m=-l}^l \sigma_{lm}^2}. \quad (6.28)$$

From fig. 6.1 it is immediately clear that a degree RMS is a rough measure. Only in case of an isotropic error spectrum, i.e. one not depending on order  $m$ , the  $\text{RMS}_l$  would be representative for a single  $\sigma_{lm}$ . The same holds for degree variances.

A less usual, though useful, error representation is the *median per degree*  $\text{MED}_l$ .

It plays a role in cases where the error spectrum  $\sigma_{lm}$  is in tendency isotropic, but with deviations for certain orders. This situation arises e.g. for non-polar orbits, leaving a certain polar gap, cf. 7.1. Only a few coefficients of low order are affected. The degree median then represents an error measure that would be valid if the orbit would have been polar. The median is the middle element of a list of  $\sigma_{lm}$  of the specified degree  $l$ , sorted on size.

**Commission Errors.** Besides considering a certain variance or RMS per degree it is also of interest to consider the cumulative error in a certain bandwidth. This is denoted as the *commission* error, defined as:

$$\text{CUM}_l = \sqrt{\sum_{n=2}^l \sigma_n^2} = \sqrt{\sum_{n=2}^l \sum_{m=-n}^n \sigma_{nm}^2}.$$

It displays the total error power *up to* a certain order  $l$ . Of even more interest would be to consider this commission error for a given gravity field functional. In case the functional  $f^\#$  has eigenvalue  $\lambda_l$ , it is:

$$\text{CUM}_l(f^\#) = \sqrt{\sum_{n=2}^l \lambda_n^2 \sigma_n^2} = \sqrt{\sum_{n=2}^l \lambda_n^2 \sum_{m=-n}^n \sigma_{nm}^2}. \quad (6.29)$$

The commission error thus expresses the noise in a given functional, due to the noise in the coefficients up to a certain order  $l$ . The eigenvalues, or isotropic transfer, may be taken from the Meissl scheme (Rummel & Van Gelderen, 1995). Naturally, the eigenvalues may be applied already to the original 2D error spectrum to yield the corresponding error spectrum of the specific functional:  $\sigma_{lm}^\#$ . For instance the Earth radius  $R$  scales the dimensionless  $\sigma_{lm}$  into an error spectrum of the geoid. The same then holds for the 1D error representations.

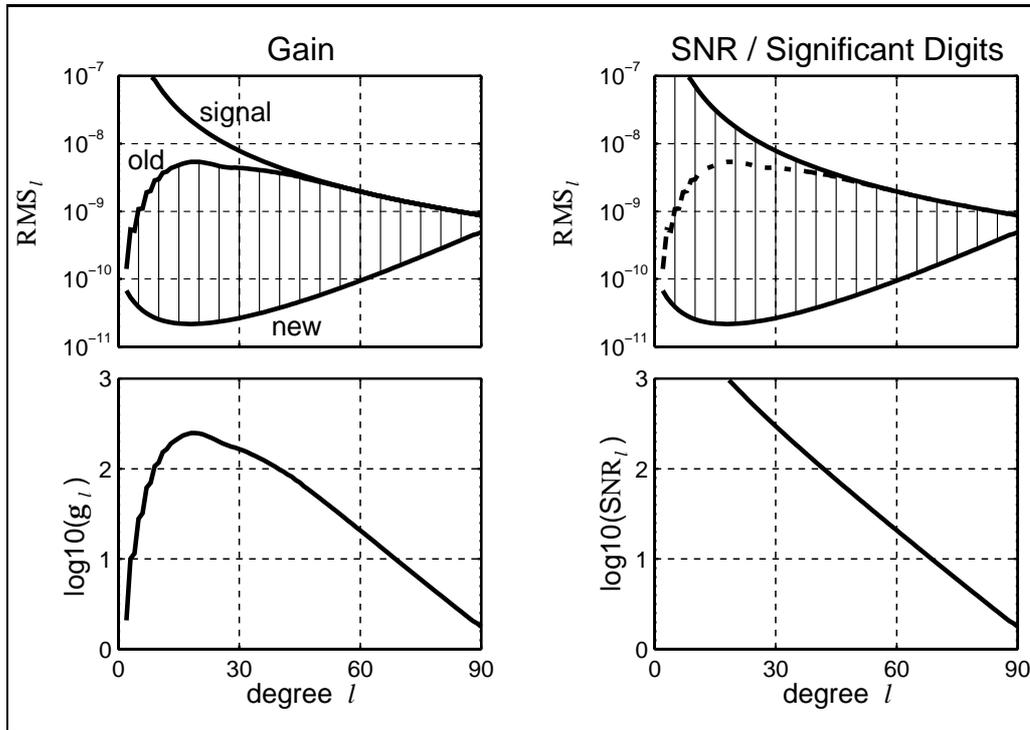
**Relative Error Measures.** The above 1D and 2D error measures are absolute, in that they represent the actual *size* of the noise in the coefficients. For many purposes, though, a relative measure is more worthwhile. The error spectrum might be viewed relative to:

- a signal spectrum, yielding a *signal-to-noise ratio* (SNR),
- another error spectrum, especially an a priori error spectrum, yielding *gain*.

The base 10 logarithm of these relative measures then represents the number of *significant digits* and the gain in significant digits, respectively.

$$\begin{aligned} \text{signal-to-noise: } \text{SNR}_{lm} &= \frac{\text{2D}}{\sigma_{lm}} \frac{|\bar{K}_{lm}|}{\sigma_l}, & \text{SNR}_l &= \frac{\text{1D}}{\sigma_l} \frac{\sqrt{c_l}}{\sigma_l}, \\ \text{gain: } g_{lm} &= \frac{\sigma_{lm}^{\text{old}}}{\sigma_{lm}^{\text{new}}}, & g_l &= \frac{\sigma_l^{\text{old}}}{\sigma_l^{\text{new}}}. \end{aligned} \quad (6.30)$$

The quantity  $c_l$  is the signal degree variance, a quadratic quantity. It can either be a model signal variance, e.g. Kaula or Tscherning-Rapp models, or be computed from a real field:  $c_l = \sum_m \bar{K}_{lm} \bar{K}_{lm}^*$ . The 1D gain and SNR are explained in fig. 6.2. They are, loosely speaking, the *vertical distance* between signal and error (old and new) curves, displayed by the thin vertical lines.



**Figure 6.2:** Graphical explanations of the relative error measures gain (old vs. new errors) and SNR (signal vs. error). See text for further explanation.

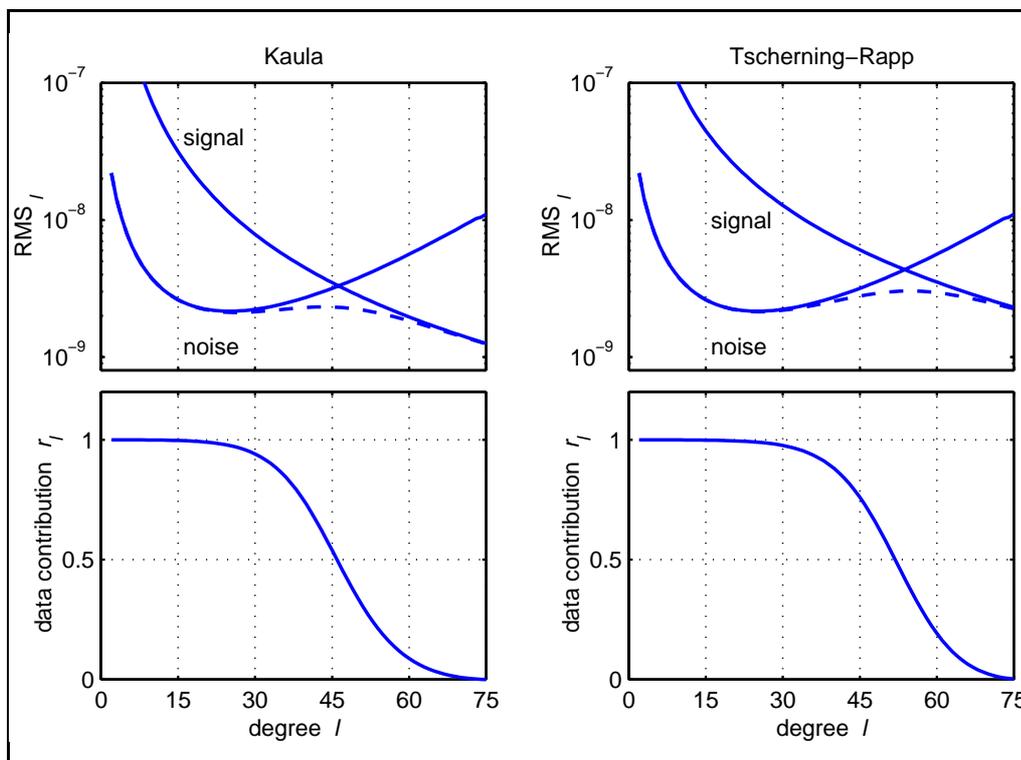
It might occur that the new error spectrum has a larger spectral extent than the older model. In that case the model signal curve takes over the role of the old error curve for the missing spectral part, see RHS of the gain plot. This is consequent, since the old gravity model did not have information in this band, other than that represented by a signal degree variance model. The expectation of the coefficient is zero and their variance equals the signal variance model.

**Resolution.** The signal-to-noise ratio  $\text{SNR}_l$  leads to a working definition of the concept of *resolution* of an error simulation. The resolution, or maximum resolvable degree  $L$ , is defined here by  $\text{SNR}_L = 1$ . It occurs at the cross-over point between signal and noise curves. This definition, though, is very weak for two reasons:

- what is meant by signal?
- what is meant by noise?

In fig. 6.3 the results of some error simulation are displayed, whose parameters do not matter here, except for the fact that  $I = 90^\circ$  was used. The curves labelled ‘noise’ in both top panels represent the  $\text{RMS}_l$  of the simulation. For comparison, curves labelled ‘signal’ are drawn as well. The left panel employs Kaula’s rule of thumb, whereas the right panel uses the Tscherning-Rapp model. The noise curves cross the signal around  $l = 46$  (left) and  $l = 53$  (right). The definition  $\text{SNR}_L = 1$  would lead to different resolutions for different a priori signal models.

The  $\text{RMS}_l$  curve of fig. 6.3 came from an unregularized solution, which could be determined due to the polar orbit. Making use of the signal models as a priori information results in the dashed noise curves. Beyond the previously defined



**Figure 6.3:** Graphical explanation of the concept of resolution. See text for further explanation.

resolutions, the data still seem to contain information on the unknowns, if aided by prior knowledge. Instead of crossing, the signal and noise curves converge. The resolution would move up towards  $l \approx 75$ .

**Remark 6.3** *Alternatively one may use estimated SH coefficients as signal for the purpose of resolution determination. In view of the discussion on biased estimation, 6.3, this approach may lead to a pessimistic result. The biased estimate was shown to be a smoothed one, which yields too low signal  $RMS_l$  values.*

Thus the resolution is seen to depend on the question what exactly is considered noise and signal. The situation becomes even more cumbersome as soon as the error spectrum is anisotropic. In that case resolution cannot be defined anymore as a single number for the whole error spectrum. One number per order  $m$  would be required instead.

Since resolution has to do with information content of the data, the concept of contribution spectrum  $r_{lm}^n$ , cf. 6.5, comes into mind. The bottom panels of fig. 6.3 express the partial contributions of the data (as opposed to the a priori information), though averaged over the orders. That is  $r_l = (2l+1)^{-1} \sum_{m=-l}^l r_{lm}$ . This quantity may be used as an alternative resolution definition, e.g. by setting  $r_L = \frac{1}{2}$ , meaning that the unknowns are equally determined by data and by prior knowledge. With this (arbitrary) choice one arrives at  $L = 46$  (Kaula) and  $L = 53$  (Tscherning-Rapp) again. Any other (arbitrary) choice would result in a different resolution. The question as to what is considered noise, is circumvented. The signal question remains.

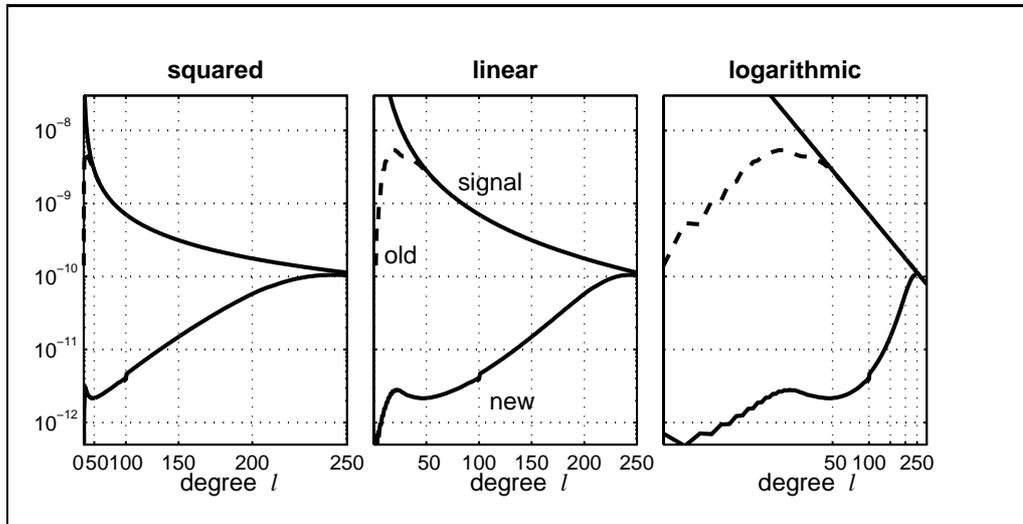


Figure 6.4: Three options for scaling the  $l$  axis.

**Degree Axis Scaling.** The above graphs employed a linear degree axis. If the frequency range becomes large, it may be useful to resort to logarithmic scaling on the degree axis. This is justified from the viewpoint of *length scales*. Degree  $l$  contains—by rule of thumb—spatial scales of  $\pi/l$ , or  $20\,000/l$  km. In this respect a logarithmic scale makes sense. Using a linear scale with  $L=400$  e.g., the first half of the axis contains spatial scales from 20 000 km down to 100 km, the second half only 50–100 km. The logarithmic scaling emphasizes the long wavelengths.

On the other hand, the linear scaling neglects already the quadratic growth of involved coefficients with degree. Only the 2D triangular error spectrum reflects the full spectral content. In 1D this abundance might be underlined by a quadratic degree axis. It very much emphasizes the short spatial scales. Considering these facts and inspecting fig. 6.4, the linear degree axis seems a good compromise.

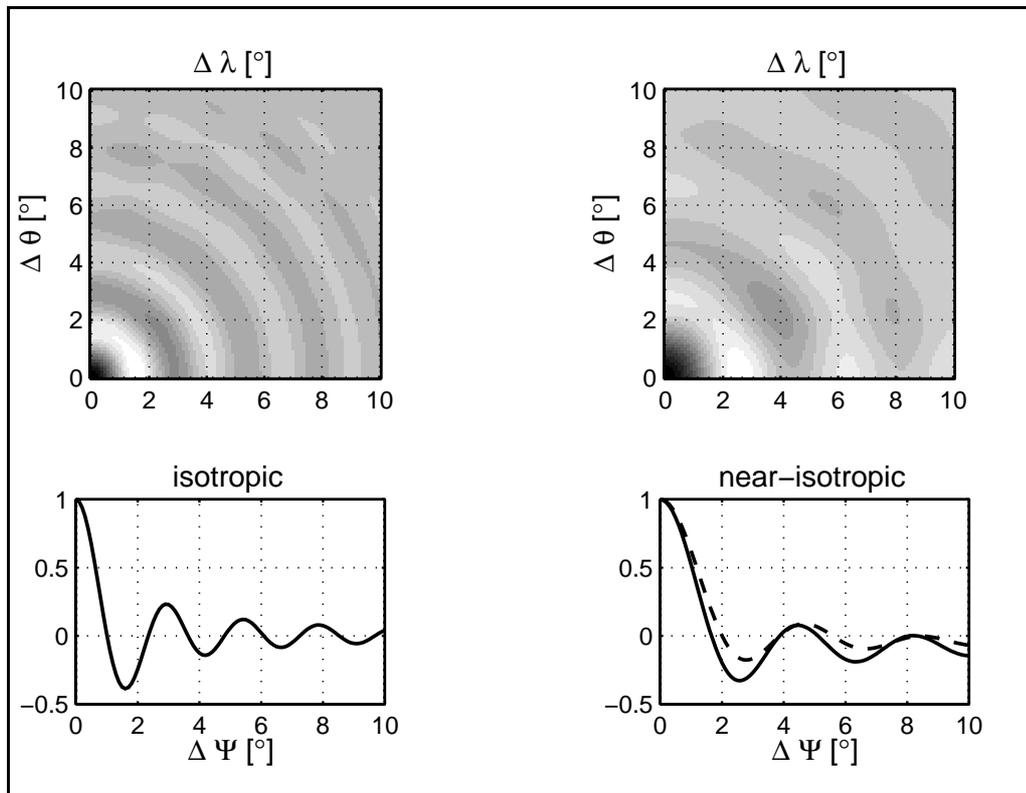
## 6.7 Spatial Error Representation

COMPLEMENTARY to spectral error measures spatial error characteristics are of interest, e.g. gravity field standard deviations over the sphere or geoid covariance functions. As described in B, the full block structure of the normal matrix can be exploited in a formal covariance propagation. Thus the resulting spatial error characteristics do contain the effect of correlation between coefficients of equal order. At the same time the spatial error measures are cumulative by nature, since they employ all  $m$ -blocks in principle. This allows a cross-check with commission errors  $CUM_l$ .

Two basic error representations are relevant:

- i) Covariance functions at a certain latitude circle:  $C(\theta, \Delta\theta, \Delta\lambda)$ ,
- ii) Standard deviation as a function of co-latitude:  $\sigma(\theta)$ .

The formulae in B show that block-structure implies a strictly stationary covariance function in longitude. Hence the  $\Delta\lambda$  argument. In latitude direction this is not the case in general. Formally the covariance function must have the above form, i.e. it pertains to a specific latitude only. Numerical results, however, show



**Figure 6.5:** Example of an isotropic (left) and a nearly isotropic (right) covariance function. The 1D representations below are cross-sections; North–South and East–West in the right hand side case.

more or less a stationary behaviour in latitude direction as well. A further consequence is the fact that a global standard deviation will be a function of co-latitude alone:  $\sigma(\theta) = \sqrt{C(\theta, 0, 0)}$ .

**Isotropy and Stationarity.** In certain circumstances the 2D covariance function becomes isotropic and fully stationary, allowing a 1D representation  $C(\psi)$ , with  $\psi$  spherical distance. In B it is shown that to this end no correlations may exist (variance-only case) and that the SH error spectrum be isotropic. So  $\sigma_{lm}$  is not allowed to vary with  $m$ . These strong demands are not met in practical cases. For many observables the error spectrum is isotropic to a certain extent and the covariance matrix  $Q_{\hat{x}}$  will often be diagonal dominant. The corresponding covariance functions will *in tendency* be isotropic and stationary, see also fig. 6.5.

## 6.8 Summary

THIS chapter described the noise propagation from data level onto the level of estimated parameters. The linear model of dynamic satellite geodesy is ill-posed in general, necessitating regularization of the normal matrices. The regularization consists here of the inclusion of a priori information on the first and second moments of the parameters.

- i) The a priori covariance matrix  $Q_{x_0}$  may consist both of noise covariance

(from existing gravity field models) and of signal variance (from degree variance models).

- ii) Regularization is easily understood in terms of sequential estimation.
- iii) Resolution matrices express the information content from data relative to that from prior knowledge.
- iv) Additionally, the trace of a resolution matrix can be interpreted as partial redundancy.
- v) In case of unwarranted a priori information, the concept of biased information is invoked. The inverted normal matrix remains a proper error measure, though.
- vi) In addition to spatial leakage and omission error, the effect of delimiting the SH domain at a maximum degree  $L$  is a mixture of aliasing and spectral leakage, due to the LS inversion process.

Calculating a posteriori covariance matrices and resolution matrices requires basic operations on normal matrices: addition, multiplication and inversion. These operations (and also eigenanalysis) can be performed without actual data. The process is denoted as *least squares error simulation*.

- vi) LS error simulation is an excellent pre-mission design tool. The influences of orbit configuration, mission profile, observation noise and so on are investigated before satellite launch.
- vii) The main diagonal of the a posteriori covariance matrix (inverted normal matrix) results in a spherical harmonic error spectrum  $\sigma_{lm}$ . It neglects correlations, however.
- viii) Since normal matrices express a signal-to-noise ratio at data level, it is not necessary to make the data of different observable type dimensionless before stacking their corresponding normal matrices.
- ix) Eigenanalysis is employed for determining the level of ill-posedness (through the condition number) and for deriving a correlation measure.
- x) Contrary to least squares error simulation, sensitivity analysis is a forward modelling tool. It conveys information on how single SH coefficients project onto the signal. It does not, however, predict the recoverability of coefficients from the signal. The sensitivity measure  $s_{lm}$  is an upper bound for the amplification.

Although error correlations are neglected already in  $\sigma_{lm}$ , even more condensed error measures are derived for purposes of (graphical) representation.

- xi) From ‘averaging’ over the order  $m$ , the 2D error spectrum  $\sigma_{lm}$  is turned into 1D degree variances and corresponding RMS per degree.
- xii) The 1D error spectrum is only meaningful if its 2D counterpart is close to isotropic.
- xiii) In case the 2D error spectrum is isotropic *in tendency* with only a few deviations, a median per degree is worthwhile.
- xiv) Relative measures (gain and signal-to-noise ratio) are introduced.
- xv) The resolution of an error simulation is weakly determined, depending on how one defines signal and noise. This conclusion may be generalized to gravity field recovery itself.

The a posteriori covariance matrices can be propagated further in order to obtain error representations in the spatial domain.

- xv) The error propagation is able to take care of the full correlation information between unknowns.
- xvi) Block structure of the covariance matrix leads to strict stationarity of the spatial covariance function in longitude direction.

# 7 Applications to Synthetical Satellite Gravity Missions

THE PREVIOUS chapters supplied the necessary ingredients for gravity field error assessment in sufficient detail: the linear observation model (in lumped coefficient formulation), the stochastic model in the frequency domain (using error PSDs) and the method and tools of LS error analysis. By means of numerous case studies this chapter will shed light on the dependency of gravity field errors on choice of observable, measurement noise and orbit characteristic.

As basic error measure the SH error spectrum  $\sigma_{lm}$  is employed, although for purposes of comparison information-by-degree (RMS, cumulative, ...) will be used as well. It will be clear throughout this chapter, though, that 1D error measures cannot be representative in general. To underline this warning, the first section starts off with a number of examples of non-isotropic error spectra.

## 7.1 Non-Isotropic SH Error Spectra

A PROCESS on the sphere is called *homogeneous* if its second moment, i.e. its covariance function, only depends on spherical distance. It depends neither on location (*stationarity*) nor on azimuth (*isotropy*). Spectrally, this is guaranteed if the covariance function is a series of Legendre polynomials with positive coefficients (Obuchov, 1947).

In case of LS error simulation, isotropy requires a spherical harmonic spectrum that only depends on degree  $l$ . As soon as  $\sigma_{lm}$  changes with the order  $m$ , the error spectrum and the corresponding covariance function are non-isotropic. Recall that the order  $m$  is also known as *azimuthal wavenumber*. Non-isotropies may derive from:

- i) spatial and temporal data inhomogeneity, e.g. in satellite-only gravity field models,
- ii) type of observable,
- iii) noise model,
- iv) orbit characteristic, particularly  $I \neq 90^\circ$ .

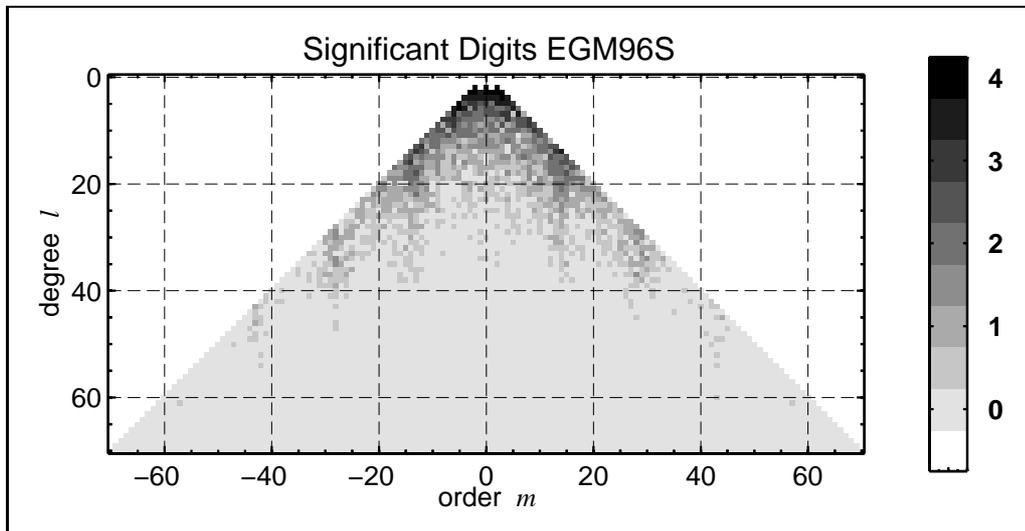
**Satellite-Only Gravity Field Models.** Existing satellite-only gravity models depend on classical tracking data, pre-eminently on satellite laser ranging (SLR). Due to the sparseness of observation stations, the SLR observable does not yield a continuous time-series. It rather produces a geometric configuration, in which the signal content necessarily remains limited to the lower frequencies near 1 CPR.

Exactly on these frequencies the signal is sensitive to (near-) resonance. On the one hand this means amplification, leading to a strong signal. On the other hand it complicates the modelling.

The resolvable coefficients are those that project on the lower Fourier frequencies. In general this comes down to low degree coefficients. Moreover, bands of certain orders  $m$  contribute to the lower frequencies. These are the orders that cause  $\beta_{mk} = \dot{\psi}_{mk}/n$  to be close to 0 or  $\pm 1$ . With (5.5),

$$\dot{\psi}_{mk} = k\dot{u} + m\dot{\Lambda} = \dot{u}(k - m\frac{\alpha}{\beta}),$$

it becomes clear that only orders that are near integer multiples of the repeat ratio  $\beta/\alpha$  are able to cause (near-) resonance. For instance, the satellite Starlette (mean altitude 900 km,  $e = 0.020$ ,  $I = 50^\circ$ ) has  $-\dot{u}/\dot{\Lambda} = 13.8$ . Thus the coefficients of order  $m_r=14$  map onto the low frequencies. Moreover, multiples  $\gamma m_r$  map onto the low frequencies. The integer multiplier  $\gamma$  is denoted as resonance level (Klokočník et al., 1990). The dominant resonance usually occurs for  $\gamma=1$  (Gooding & King-Hele, 1989). Note that  $\gamma=0$  leads to zonal resonances.



**Figure 7.1:** Significant digits of the coefficients of the EGM96S gravity model.

Lambeck (1988; Fig. 6.3) visualizes the SH domain that can be derived from classical orbit perturbation analysis. ‘Classical’ is to be understood in the sense, that the orbit perturbations do not provide a continuous time-series. Apart from the low degree harmonics it is clear that only coefficients from the  $\gamma=0,1,2$  resonance bands are recoverable. Figure 7.1 displays the significant digits, cf. (6.30), from the recent satellite-only gravity field model EGM96S (Lemoine et al., 1998). Due to data from newer and better tracking techniques coefficients from resonance band  $\gamma=3$  and even 4 have been determined. The general picture from Lambeck, though, remains valid in that it shows the limitations of gravity field recovery from classical tracking data.

To test the assertion that the resonance-induced band pattern in fig. 7.1 is not caused by the resonance itself, but rather by the observation geometry, data distribution and corresponding low-frequency limitation, the following simulation has been performed. For all satellites, contributing to the satellite-only model

EGM96S, cf. (Lemoine et al., 1998), it was assumed that the radial orbit perturbation  $\Delta z$  could be measured continuously. However, the corresponding PSDs are far from white, cf. fig. 7.2. They are more or less the inverse of the spectral transfer from the Hill equations, cf. fig. 4.1. Thus it is simulated that the frequencies around 0 and 1 CPR are well observed, whereas the other frequencies degrade rapidly.

The simulation results in the SH error spectrum displayed in fig. 7.3, at the right. From comparison to the EGM96S error spectrum, at the left, one can conclude that most features are present: relatively well determined low degree coefficients and in particular the resonance bands. The  $RMS_l$  of both triangles, not shown here, are comparable.

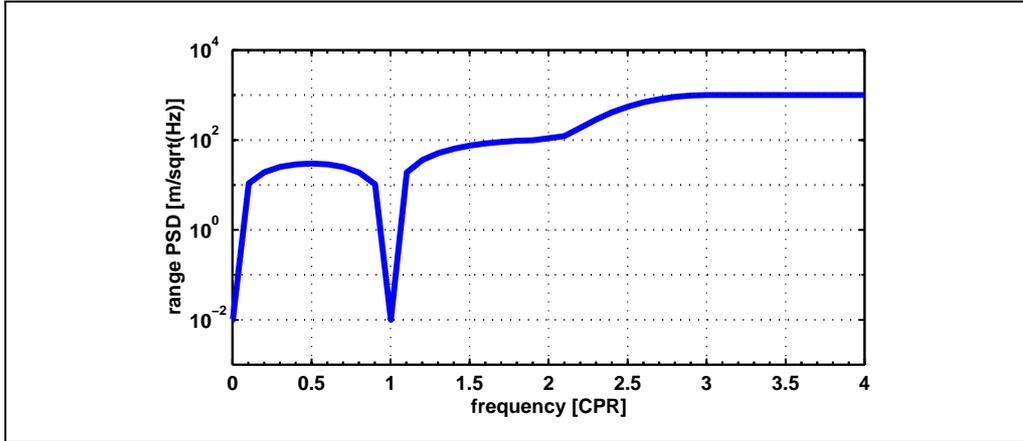


Figure 7.2: Power Spectral Density simulating SLR error spectrum.

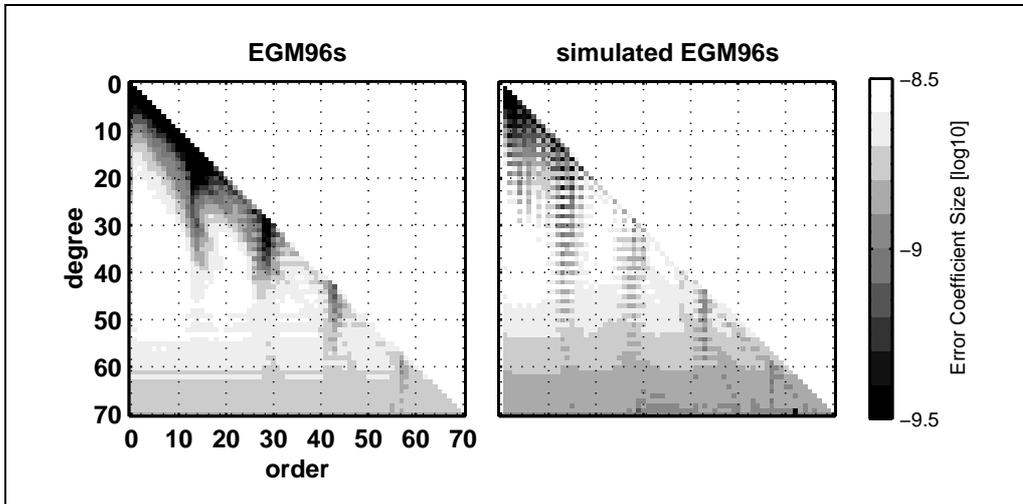
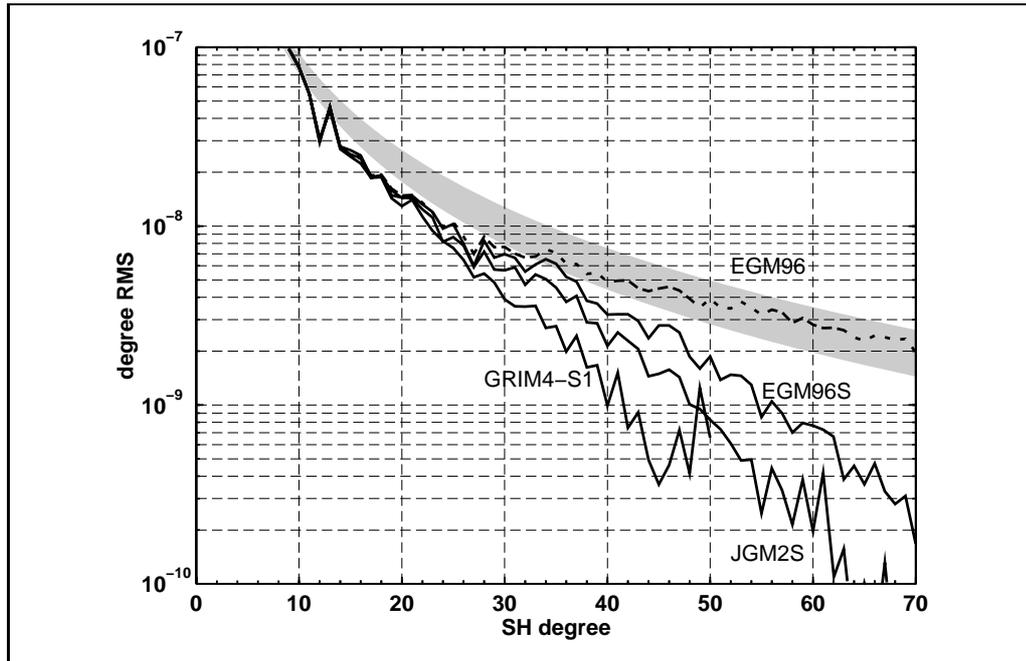


Figure 7.3: EGM96S actual (left) and simulated (right) SH error spectrum.

Being non-isotropic, the SH error spectra of the satellite-only gravity field models are not completely represented by  $RMS_l$  curves. For one reason, though, it is still worthwhile to show them here. In fig. 7.4 the signal degree RMS curves of the satellite-only models GRIM4-S1 (Schwintzer et al., 1992), JGM2S (Nerem et al., 1994) and EGM96S (Lemoine et al., 1998) are displayed. They are compared to the combined model EGM96 and to a signal degree RMS model. The grey area lies between Kaula's rule-of-thumb and the Tscherning-Rapp model. Clearly, the

satellite-only models possess less signal power. Although the models gain more power and resolution over the years, the coefficients are on the average underestimated. This is an expression of the bias, due to regularization, as discussed in 6.3. For a review of global gravity field models, both satellite-only and combination models, refer to (Bouman, 1997a).

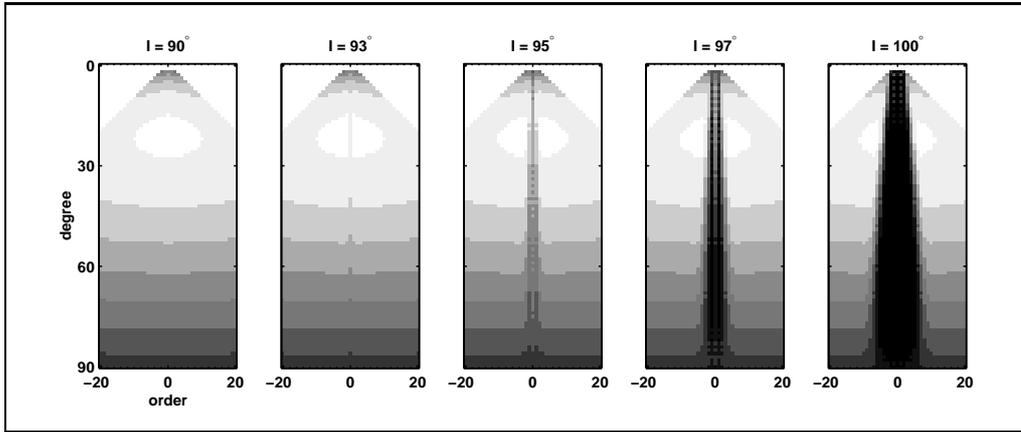


**Figure 7.4:** Signal degree RMS curves of the satellite-only models GRIM4-S1, JGM2S and EGM96S, compared to EGM96 (combination model). The gray band denotes signal from the Kaula (upper limit) and the Tscherning-Rapp (lower limit) models.

**Type of Observable and Noise Model.** The remainder of this chapter will bring about enough case studies, displaying the dependence of gravity recovery on the type of observable. Figure 6.1 may also serve as an example. The error spectrum was based on the  $V_{yy}$  component of the gravity gradient tensor (other parameters are irrelevant at this point). The role of the noise model becomes clear by recalling the spectral mapping procedures in 5.3. Any PSD, other than white, will distribute its noise in an intricate and certainly non-isotropic way over the SH error spectrum. The simulated EGM96S error spectrum clearly demonstrates how non-white PSDs cause non-isotropy.

**Orbit Characteristic.** The orbit design influences the SH error spectrum  $\sigma_{lm}$ . The groundtrack pattern determines the way the Earth is covered, i.e. the way the geopotential is sampled. Orbital height comes into play through the upward continuation operator, cf. remark 4.1. Since this is an isotropic operator, height does not cause anisotropic effects in the SH error spectrum, though. Height effects will be discussed further in 7.2 in the context of the  $V_{zz}$  observable.

Much more important in this respect is the orbital inclination. As seen in fig. 5.1, a non-polar orbit leads to a situation, in which the sphere is sampled homogeneously, except for two polar caps. Such spatial (latitudinal) windowing results



**Figure 7.5:** Polar gap wedges and inclination dependence.

in spectral leakage, in this case in a distortion of the SH spectrum. As shown by Sneeuw and Van Gelderen (1997) especially the low order harmonics are affected. Thus the SH error spectrum becomes non-isotropic due to the polar gaps.

Figure 7.5 shows the evolution of a polar gap wedge in the SH error spectrum for growing polar gap. It was simulated using the  $V_{zz}$  gravity gradient tensor component. Although they are irrelevant at this point, the other parameters are listed in tbl. 7.1. In tendency this component yields a white error spectrum per degree, which is disrupted by the polar gap wedge. In (Sneeuw & Van Gelderen, 1997) the width of the polar gap wedge is given by a rule-of-thumb. The maximum order, affected by it is:

$$m_{\max} = \left| \frac{\pi}{2} - I \right| l,$$

i.e. linearly depending on both  $I$  and  $l$ . Error spectra like the ones in fig. 7.5 demonstrate the usefulness of the 1D error measure  $\text{MED}_l$ , the median per degree. Opposed to a degree variance or  $\text{RMS}_l$  it perfectly represents the error level of the coefficients outside the wedge. Figure 7.6 displays both 1D measures (left-hand plot) and the commission error (right-hand plot), either based on  $\text{RMS}_l$  or on  $\text{MED}_l$ . Whereas the  $\text{MED}_l$  curves follow a typical gradiometry behaviour, see also 7.2, the  $\text{RMS}_l$  curves are too pessimistic, due to the higher noise for  $\sigma_{lm}$ , with  $m < m_{\max}$ .

A formal error propagation from SH error spectrum onto the sphere, e.g. in terms of geoid error, reveals that  $\text{MED}_l$  is representative indeed. In B it is explained that block-diagonal normal matrices lead to a spatial error, depending only on co-latitude  $\theta$ . Figure 7.7 shows the error propagation, corresponding to the error spectrum, on which fig. 7.6 was based. Outside the polar gap the error level assumes the value, implied by  $\text{MED}_l$  and corresponding to its commission error. Inside the polar gap the geoid accuracy deteriorates quickly. However, the effect of the low order coefficients remains confined to the polar gap rather sharply.

## 7.2 Gradiometry

**G**RAVITY gradiometry, i.e. the measurement of second spatial derivatives of the potential, can be realized through differential accelerometry over short

**Table 7.1:** Simulation parameters.  $T = 1$  mo.

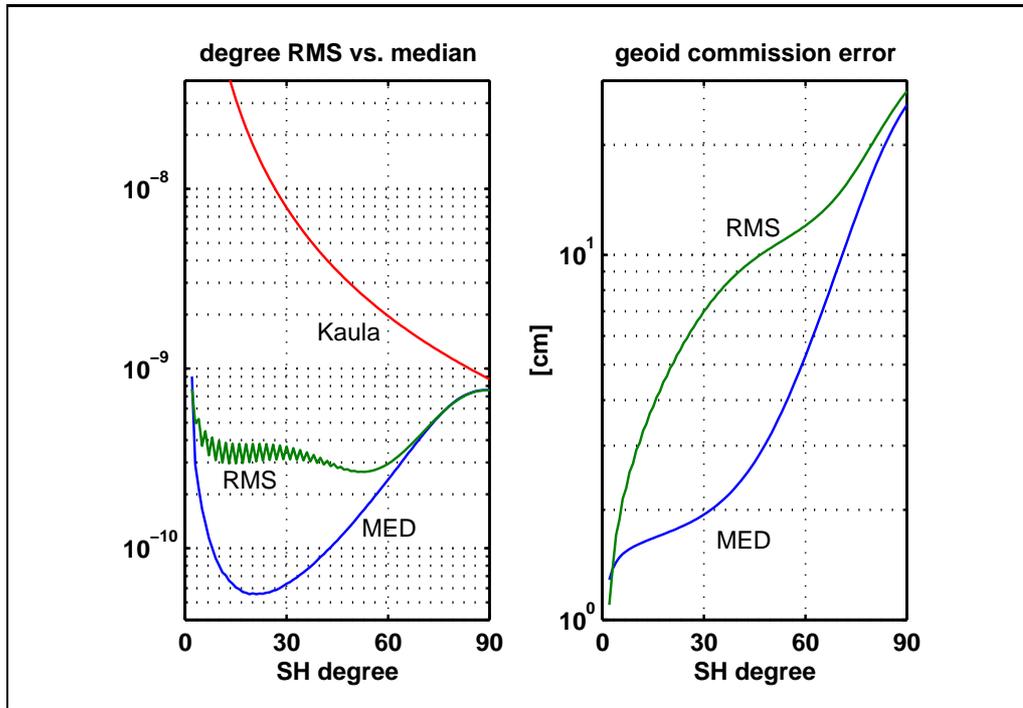
set	obs.	PSD [unit/ $\sqrt{\text{Hz}}$ ]	$h$ [km]	$I$ [ $^\circ$ ]	$L$	$\rho_0$ [km]
EGM96S	$\Delta z$	cf. fig. 7.2	EGM96S satellites, cf. (Lemoine et al., 1998)			
polar gap	$V_{zz}$	$10^{-2}$ E	600	90...100	90	–
gradio	all $V_{ij}$	$10^{-2}$ E	500	90	90 [90]	–
$h$ & reg	$V_{zz}$	$10^{-2}$ E	500...800	90	90	–
hi-lo SST	$\Delta \underline{x}$	$10^{-1}$ m	500	90	50 [50]	–
lo-lo SST1	$\Delta \rho$	$10^{-3}$ m	500	90	90	100
	$\Delta \dot{\rho}$	$5 \cdot 10^{-5}$ m/s				
	$\Delta \ddot{\rho}$	$10^{-6}$ m/s <sup>2</sup>				
lo-lo SST2	$\Delta \rho$	$10^{-3}$ m	500	90	90	400,800,1600
lo-lo SST3	$\Delta \ddot{\rho}$	$10^{-8}$ m/s <sup>2</sup>	500	90	90	1

baselines, (Rummel, 1986a). Thus it makes sense to investigate gravity field recovery capability based on observation of single components of the gravity gradient tensor (4.7). For this purpose a polar orbit is chosen, in order to avoid the polar gap effect. Other simulation parameters are given in tbl. 7.1. They are only chosen for showing the effect of varying single parameters. It is not intended to simulate any real mission here. Conclusions in terms of resolution or geoid accuracy have no meaning beyond the context of this section and the underlying set of parameters.

**Single Tensor Componentents.** Figure 7.8 shows SH error spectra—though only for positive  $m$ —from a least squares error analysis of each tensor component individually. The array of triangles is composed accordingly. Since the orbit is polar and the PSD is chosen white, the SH error spectra purely reflect the result of inverting the transfer coefficients (4.8). Neither polar gap effect nor peculiar phenomena, due to the stochastic model, can show up.

The  $V_{zz}$  error analysis provides a near isotropic SH error spectrum. This is not surprising, since the specific transfer of this term, that is  $(l+1)(l+2)$ , solely depends on degree  $l$ . Only a slight  $m$ -dependency is present, the reason for which is not clear. It might be due to data distribution (more samples toward higher latitudes). It might also be a purely numerical effect. The other components behave non-isotropically. One can conclude from fig. 7.8 that the along-track in-line component  $V_{xx}$  and the off-diagonal  $V_{xz}$  favour zonal (and near-zonal) coefficients. On a polar orbit, along-track sensitivity basically comes down to North-South sensitivity, which explains the aforementioned behaviour qualitatively. Since  $V_{xz}$  contains a radial derivative as well, the appearance of the  $V_{xz}$  error spectrum is midway between  $V_{xx}$  and  $V_{zz}$ .

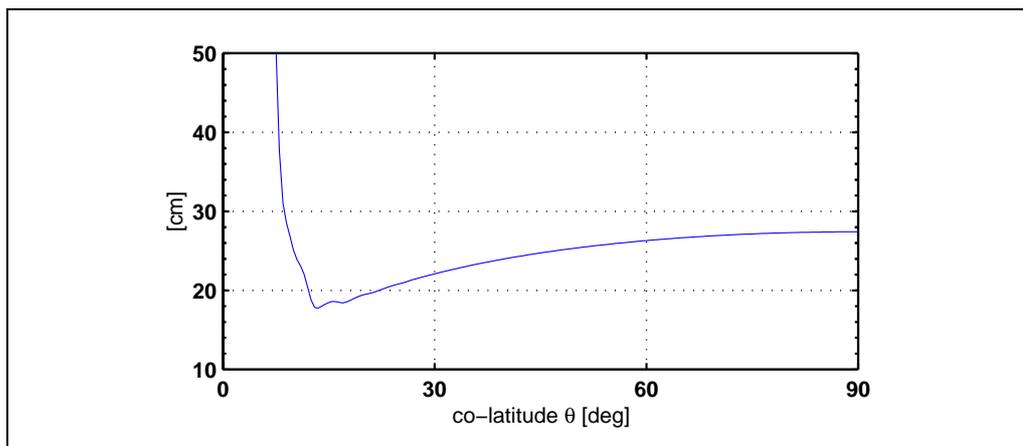
Similarly, the cross-track in-line component  $V_{yy}$  and the off-diagonal  $V_{yz}$  strongly favour sectorial (and near-sectorial) coefficients. The sensitive direction of these components is mainly East-West. Also here the  $V_{yz}$  behaves similar to  $V_{yy}$ , though with a tendency towards  $V_{zz}$ , loosely speaking. The remaining  $V_{xy}$ -term shows bad recovery capability both of near-zonals and near-sectorials.



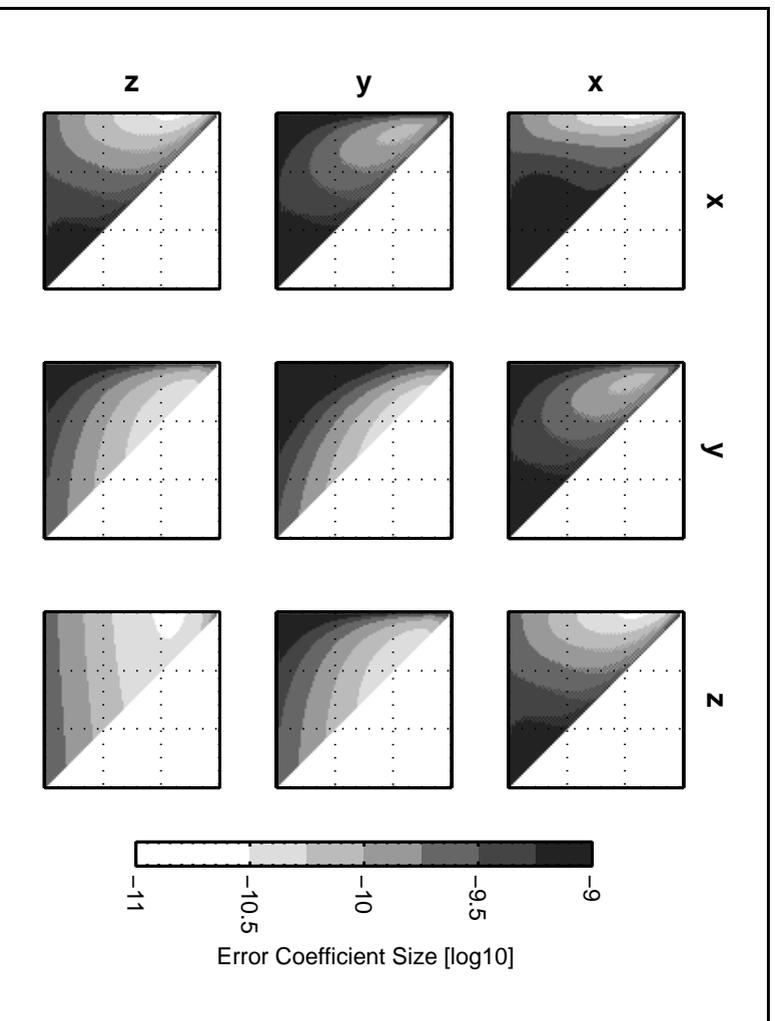
**Figure 7.6:** Demonstration of the difference between  $\text{RMS}_l$  and  $\text{MED}_l$  error measures in case of a  $V_{zz}$  simulation with  $I = 100^\circ$ .

**Remark 7.1** The SH error spectra fig. 7.8 again show the necessity of 2D spectral representation in general. Only for  $V_{zz}$  degree RMS curves are justified.

Not clearly visible from fig. 7.8 is the unsolvability of zonal coefficients from  $V_{xy}$  and  $V_{yz}$ . Coefficients  $\bar{K}_{l,0}$  cannot be determined from these components for the simple reason that  $\bar{F}_{lmk}^*(\frac{1}{2}\pi) = 0$ , cf. (A.7). The signals  $V_{xy}$  and  $V_{yz}$  along a polar orbit do not contain any zonal contribution. It may seem obvious that a gradiometer with cross-track sensitivity will not measure zonal structures—if flown on a polar orbit—since these structures have no variation in that direction. Nevertheless, this reasoning does not hold for  $V_{yy}$ . The reason is the contribution  $\frac{1}{r}V_r$  to  $V_{yy}$ , cf. (4.7).



**Figure 7.7:** Geoid error, propagated from a  $V_{zz}$ -simulation with  $I = 100^\circ$ .

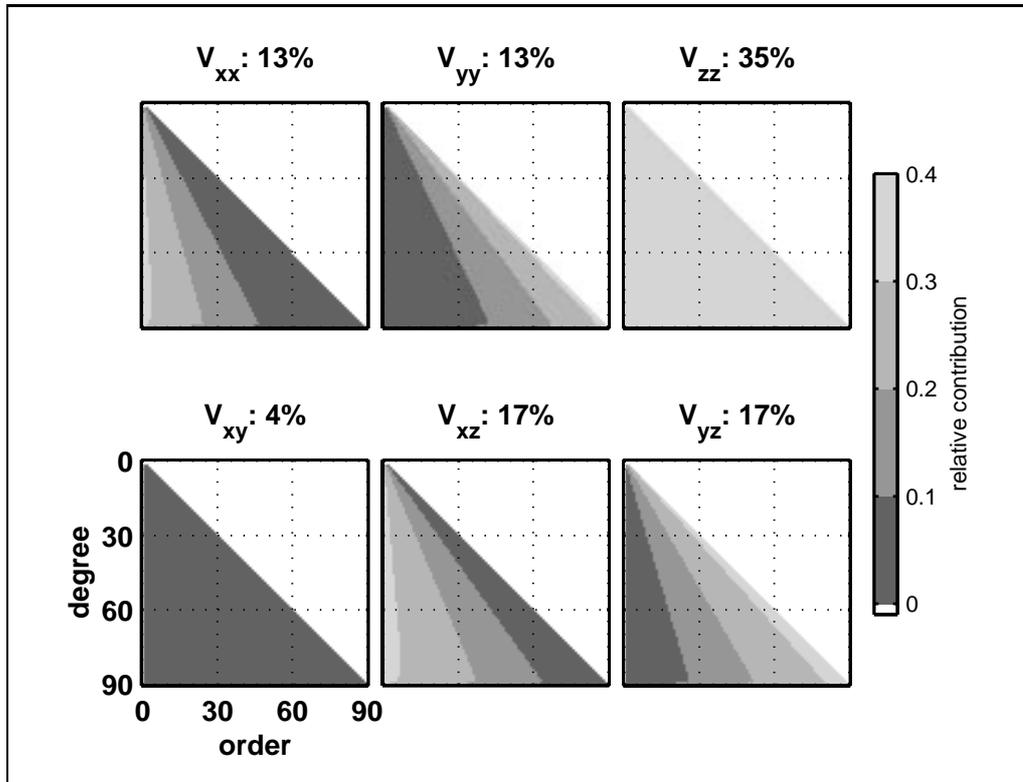


**Figure 7.8:** Matrix of SH error spectra from single component gradiometry. Each triangle has order  $0 \leq m \leq l$  on the horizontal and degree  $0 \leq l \leq L$  on the vertical axis. Symmetry allows to display only half of the full SH error spectra.

In addition to the SH error results in fig. 7.8 of single tensor components, the question arises what these components contribute to a combined solution. Figure 7.9 displays the contribution spectra, cf. 6.5 analysis type II, of the 6 distinct tensor components. The observable  $V_{zz}$  is the strongest component with a contribution of approximately one-third. This contribution is relatively homogeneous over the whole spectral range. The complementary observables  $V_{xz}$  and  $V_{yz}$  together contribute one-third as well. The combination  $\{V_{xx}; V_{yy}; V_{xy}\}$  contributes the rest. Note that this simulation was performed with  $I = 90^\circ$ . The components  $V_{xx}$  and  $V_{xz}$  would suffer most from a non-polar orbit.

Explanation—let alone prediction—of the single component results fig. 7.8 and fig. 7.9 on the basis of transfer coefficients (4.8) is not trivial. As mentioned above, it is not surprising that the  $V_{zz}$  error spectrum is near-isotropic, due to the specific transfer  $(l+1)(l+2)$ . The terms  $V_{xx}$  and  $V_{yy}$  contain  $k^2$  in their specific transfer. Moreover, both of them contain an additional  $(l+1)$  contribution,  $V_{xx}$  linearly and  $V_{yy}$  quadratically. Remarkably, their effect is nearly complementary. For the off-diagonals an analysis is even more cumbersome as soon as  $\bar{F}_{lmk}^*(I)$  is involved. Thus it is concluded that the semi-analytic approach is not too helpful in explaining the particular appearance of SH error spectra from gradiometry.

**Spatial Error Structure.** As a dual representation to the error spectra above, fig. 7.10 displays the error simulation results in the spatial domain, in terms of error covariance functions. They have been computed for a point located at the equator on a grid running from  $-20^\circ$  to  $20^\circ$  in both directions. As shown in B



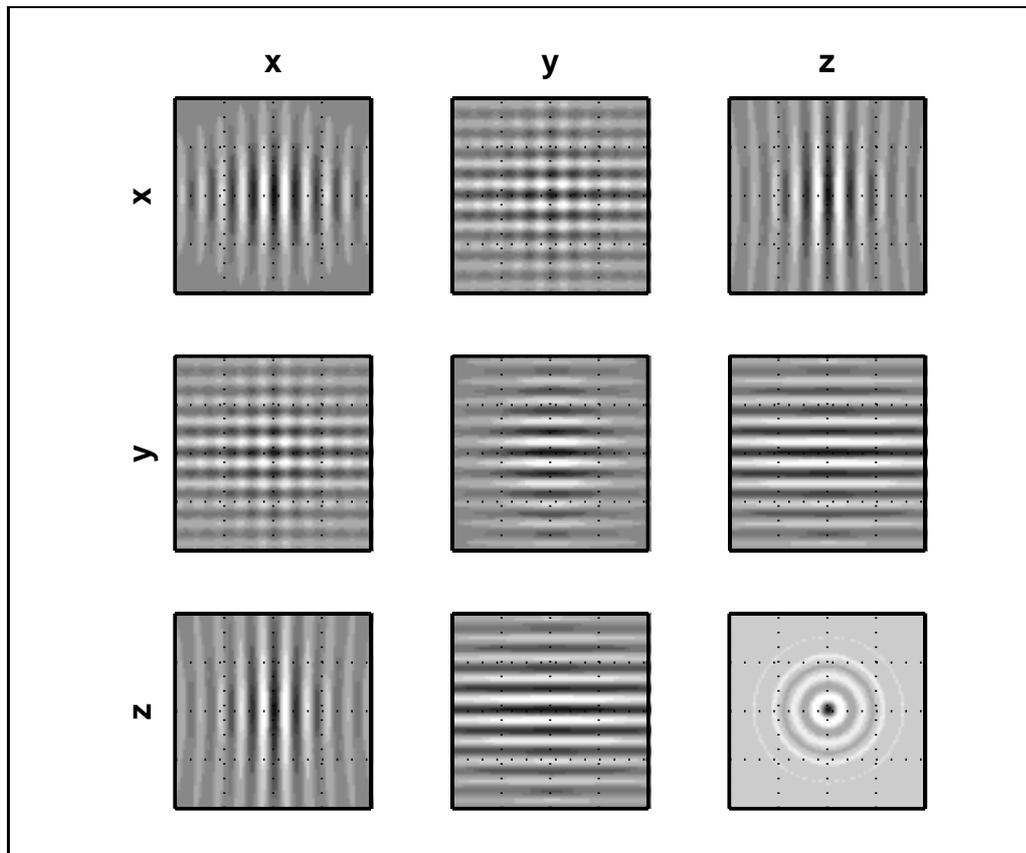
**Figure 7.9:** Contribution or redundancy spectra of single tensor components in a combination solution

these functions are stationary in longitude direction. It is tacitly assumed that stationarity will hold in latitude direction as well, at least to a large extent. The covariance structure in fig. 7.10 would be valid for the whole globe.

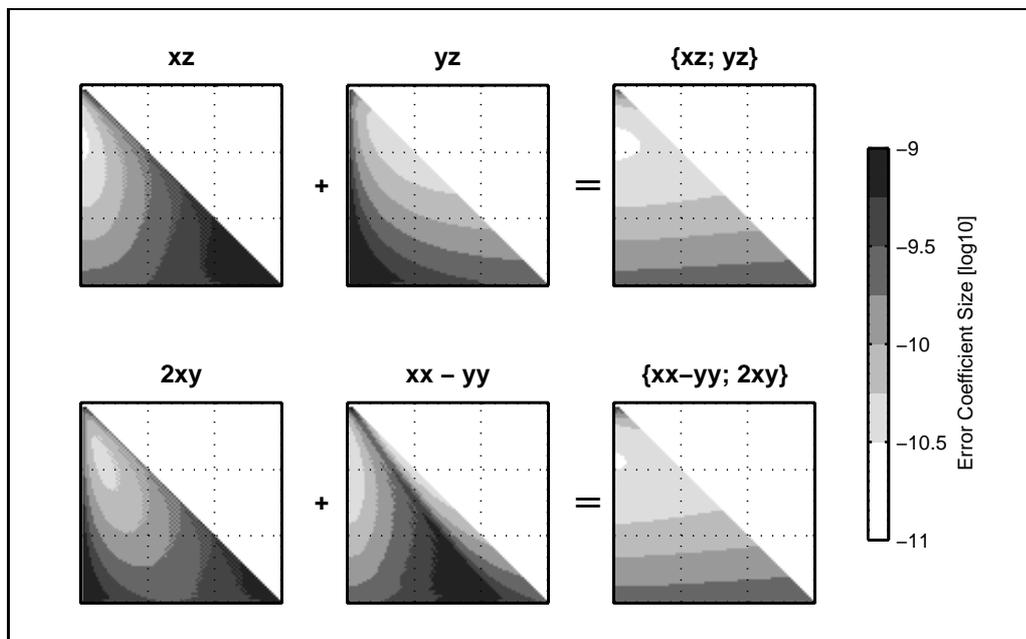
Again,  $V_{zz}$  yields an isotropic result, expressed by the circular appearance of the covariance function. Being able to determine (near-) zonals well,  $V_{xx}$  and  $V_{xz}$  show long correlation distance in North-South direction. Conversely,  $V_{yy}$  and  $V_{yz}$  show long correlations oriented East-West. The  $V_{xy}$  component shows no preferred direction. At first sight it might seem paradox that instrument sensitivity in a certain direction leads to long correlation in the same direction. On the other hand, if the observable does not provide information in the direction perpendicular to the sensitive axis, errors cannot be correlated over long distances in that perpendicular direction.

**Isotropic Combinations.** In their spectral analysis of vectorial and gradiometric potential functionals Rummel and Van Gelderen (1995) emphasize the fact that the extended scheme of eigenvalues pertains to certain combinations of observables. In gradiometric context these combinations are  $\{V_{xz}; V_{yz}\}$  and  $\{V_{xx} - V_{yy}; 2V_{xy}\}$ . The former can be regarded as the radial derivative of the deflection of the vertical. Its eigenvalue is  $l(l+1)$ . The latter is related to the classical torsion balance observable and to the rotating gradiometer concept (Forward, 1973). Its eigenvalue is  $\sqrt{(l-1)l(l+1)(l+2)}$ . Compared to the eigenvalue of  $V_{zz}$ , which is  $(l+1)(l+2)$ , it is seen that for large  $l$  all eigenvalues are approximately  $l^2$ .

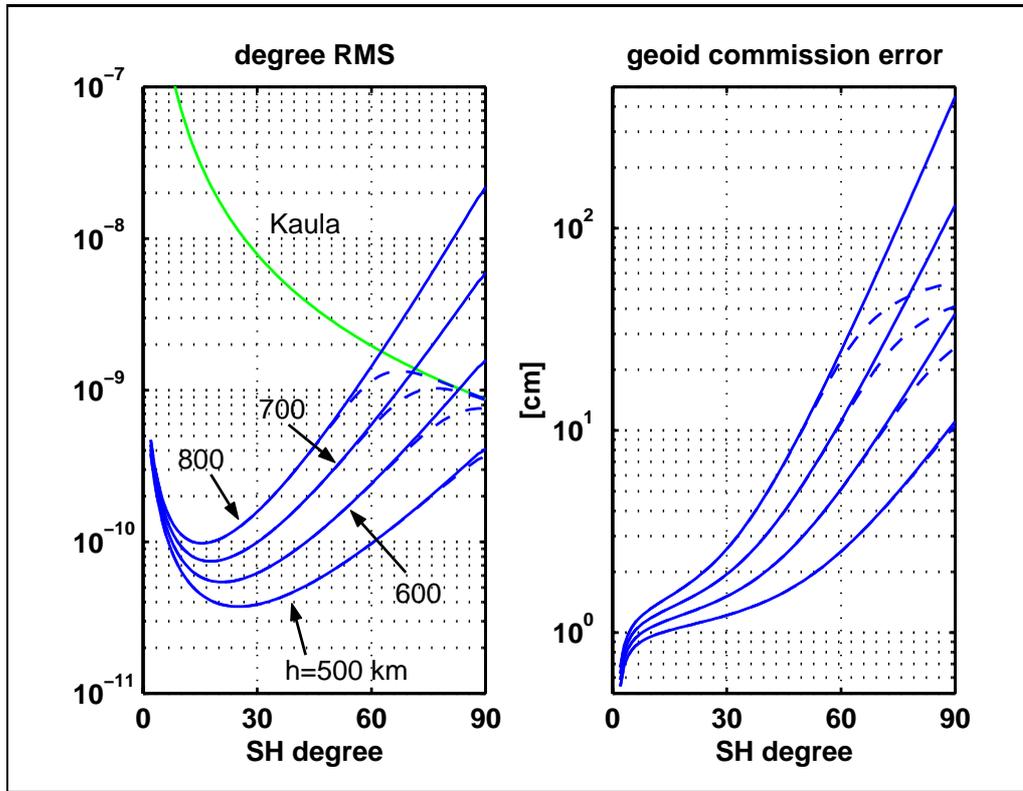
SH error analysis confirms the theoretically predicted behaviour of the aforemen-



**Figure 7.10:** Matrix of spatial covariance functions from single component gradiometry. The covariance functions are computed for the equator on a spatial domain of  $\Delta\lambda, \Delta\theta \in [-20^\circ, 20^\circ]$ . The covariance functions do not share a common colormap, i.e. their grey levels cannot be compared.



**Figure 7.11:** Isotropic combinations  $\{V_{xx} - V_{yy}; 2V_{xy}\}$  and  $\{V_{xz}; V_{yz}\}$ .



**Figure 7.12:** Effect of height variation and regularization (dashed lines) in terms of  $RMS_l$  (left) and  $CUM_l$  (right).

tioned combinations. Figure 7.11 shows the result of combining single observables  $V_{xy}$  and  $V_{yz}$  into  $\{V_{xz}, V_{yz}\}$  (top row) and of  $2V_{xy}$  and  $(V_{xx} - V_{yy})$  into  $\{V_{xx} - V_{yy}; 2V_{xy}\}$  (bottom row). Both combinations yield an isotropic SH error spectrum, comparable to the  $V_{zz}$  error spectrum in fig. 7.8. In either case the individual components are fully complementary to each other. Especially the observable  $(V_{xx} - V_{yy})$  shows an interesting pattern: good recoverability of (near-) zonals and (near-) sectorials and bad recoverability of a band of tesserals in between.

If symmetry of the gravity gradient tensor is employed for determining rotational effects (Rummel, 1986a), a full tensor gradiometer would observe 6 independent gradients:  $V_{xx}$ ,  $V_{yy}$ ,  $V_{zz}$ ,  $V_{xy}$ ,  $V_{xz}$  and  $V_{yz}$ . After grouping into the combinations  $\{V_{zz}\}$ ,  $\{V_{xx} + V_{yy}\}$ ,  $\{V_{xz}; V_{yz}\}$  and  $\{V_{xx} - V_{yy}; 2V_{xy}\}$ , it becomes clear that full tensor gradiometry produces a fourfold complete observation of the Earth. The word *complete* is used here in the sense that the resulting SH error spectra are isotropic.

**Remark 7.2** This conclusion is in line with the result of fig. 7.9, that showed contributions of approximately one-third of  $\{V_{zz}\}$ ,  $\{V_{xx}; V_{yy}; V_{xy}\}$  and  $\{V_{xz}; V_{yz}\}$  respectively. Additional use of the Laplace equation provides a fourth observable.

**Orbit Height.** The effect of orbital height on gravity field recovery is driven by the *upward continuation* operator, whose spectral characteristic is given by  $(R/r)^{l+i}$ , cf. tbl. 4.1. The generic index  $i$  depends on the observable under con-

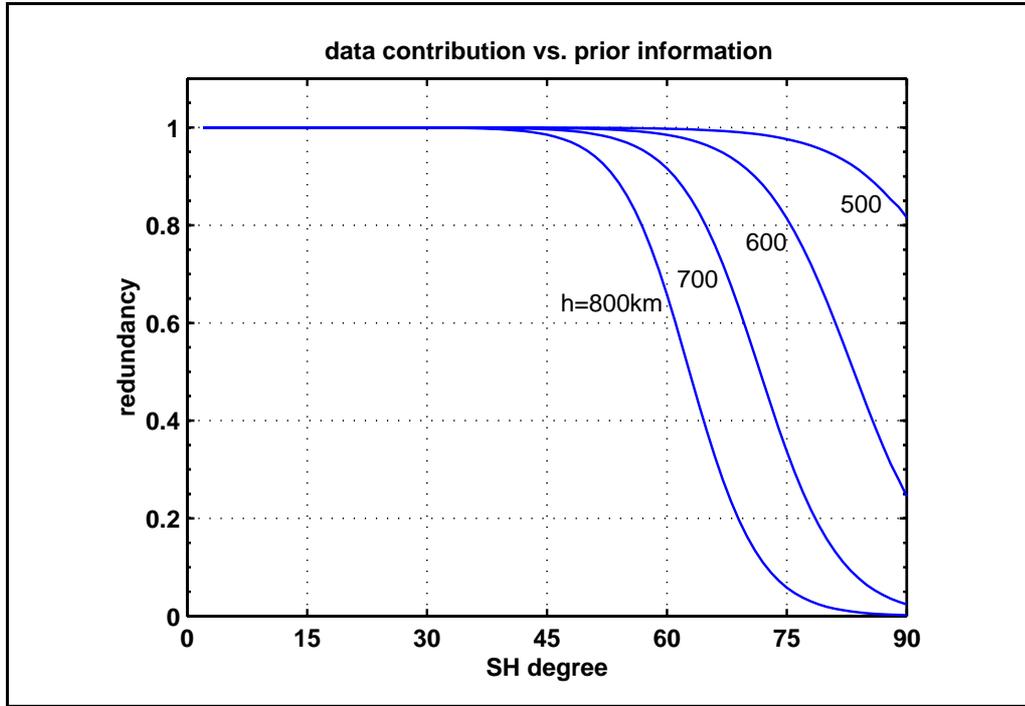


Figure 7.13: Data content under regularization.

sideration. In case of gradiometry we have  $i = 3$ . Although not restricted to gradiometry, upward continuation—being an isotropic operator—is best demonstrated with the isotropic  $V_{zz}$  observable. Its effect is investigated now by means of a simulation on a polar orbit, cf. tbl. 7.1.

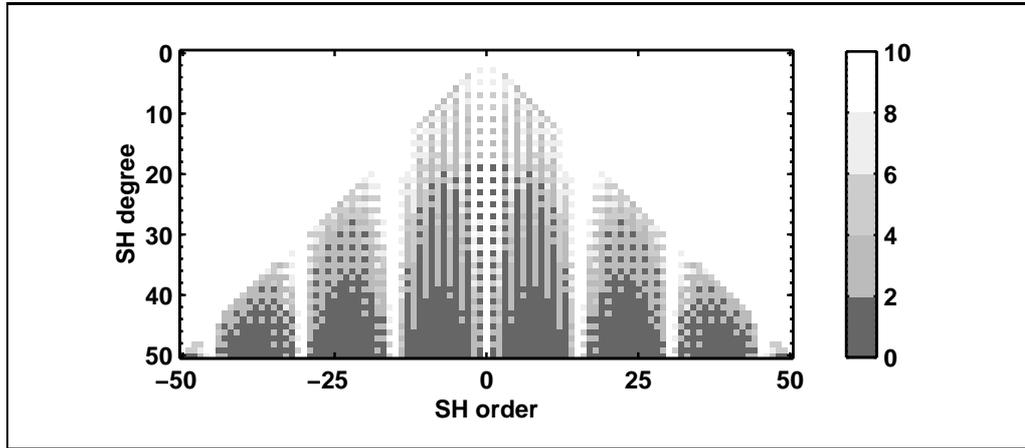
Upward continuation acts as a low-pass filter. Higher degree phenomena are damped progressively with increasing degree. Conversely, the lower the orbit, the better the gravity field recovery for higher degrees. Figure 7.12 clearly displays this behaviour for orbits at 500, 600, 700 and 800 km height. The four solid lines start off at the same point at the lowest degrees, after which they diverge for increasing  $l$ .

This shows that height is crucial to spatial resolution. In this example, an orbit at  $h = 800$  km would produce a gravity field up to  $L \approx 65$ . The 600 km orbit would yield  $L \approx 80$ . The accuracy attenuation with increasing orbital height is proportional to  $(R/r)^{l+3}$ . Thus the ratio in  $\text{RMS}_l$  between two curves should be

$$\left(\frac{R}{r_1}\right)^{l+3} / \left(\frac{R}{r_2}\right)^{l+3} = \left(\frac{r_2}{r_1}\right)^{l+3} = \left(\frac{R+h_2}{R+h_1}\right)^{l+3}. \quad (7.1)$$

At  $l = 90$ , for example, the ratio according to (7.1), with  $h_1 = 500$  km and  $h_2 = 800$  km, becomes 53. From the graph one can deduce a ratio at  $l = 90$  of about  $2 \cdot 10^{-8} / 4 \cdot 10^{-10} = 50$ . Such an estimate does not hold for the geoid commission errors at the right of fig. 7.12, since these are cumulative errors.

**Regularization.** The  $\text{RMS}_l$  curves in fig. 7.12 cross the Kaula curve (signal variance) at a certain point, that is conventionally denoted as *resolution*  $L$ . Beyond this point, where the signal-to-noise ratio (SNR) equals 1, the a posteriori error



**Figure 7.14:** Dimensionless maximum orbit perturbation transfer (cut off at 10). To obtain maximum perturbation signal size, multiply by  $R\bar{K}_{lm}$ .

variance becomes larger than the a priori signal variance. Using the signal variance as a priori information, cf. 6.1, is one way of regularizing the error variance down to the signal level. If a signal model like Kaula's rule is used, which is a power law in  $l$ , the effect is isotropic.

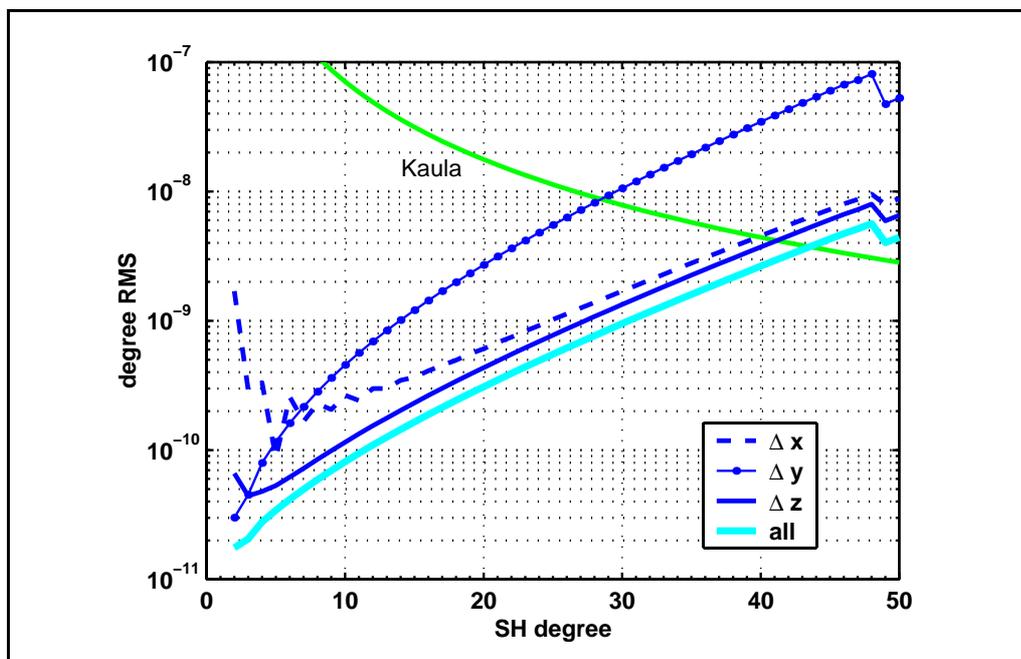
The dashed lines in fig. 7.12 illustrate the effect of regularization. The  $\text{RMS}_l$  curves are reduced and converge to the Kaula curve. A transition zone exists where the regularized  $\text{RMS}_l$  deflects from the regularized one, until it reaches Kaula's rule at a degree higher than the previously defined  $L$ . Apparently the observable still contains information about these high degrees. Only by adding (a priori) knowledge, this information is revealed. After reaching the Kaula curve the error appears to decrease further. Note, however, that  $\text{RMS}_l$  reflects the a posteriori error matrix  $Q_{\hat{x}}$ , (6.6c). It does not contain any bias contribution as discussed in 6.3.

An alternative view on regularization and resolution is presented by fig. 7.13. It shows the data contribution to the solution versus prior information, cf. the redundancy measure (6.10). These curves show the whole range in which regularization is active. The above definition of resolution ( $\text{SNR} = 1$ ) shows up in this graph as a redundancy level of 50 %. Any other redundancy level provides a new definition of *resolution*.

## 7.3 Orbit Perturbations

**T**HIS section assesses the gravity field recovery capability from orbit perturbations. It is assumed that orbit perturbations are known as discrete time-series (at a sufficient sampling rate), in terms of three-dimensional coordinate perturbations. They are expressed in a local co-rotating frame with  $x$  along-track,  $y$  cross-track and  $z$  radial, as explained in fig. 3.1.

This type of observable is realized through space-borne satellite tracking, e.g. GPS. Thus a set of simulations has been performed, denoted as *hi-lo* SST. Table 7.1 lists the parameters for this set. As before, the arbitrariness of these parameters is underlined here. They are chosen for demonstration purposes; not to simulate



**Figure 7.15:** Error degree RMS of individual orbit perturbation components and of combination solution.

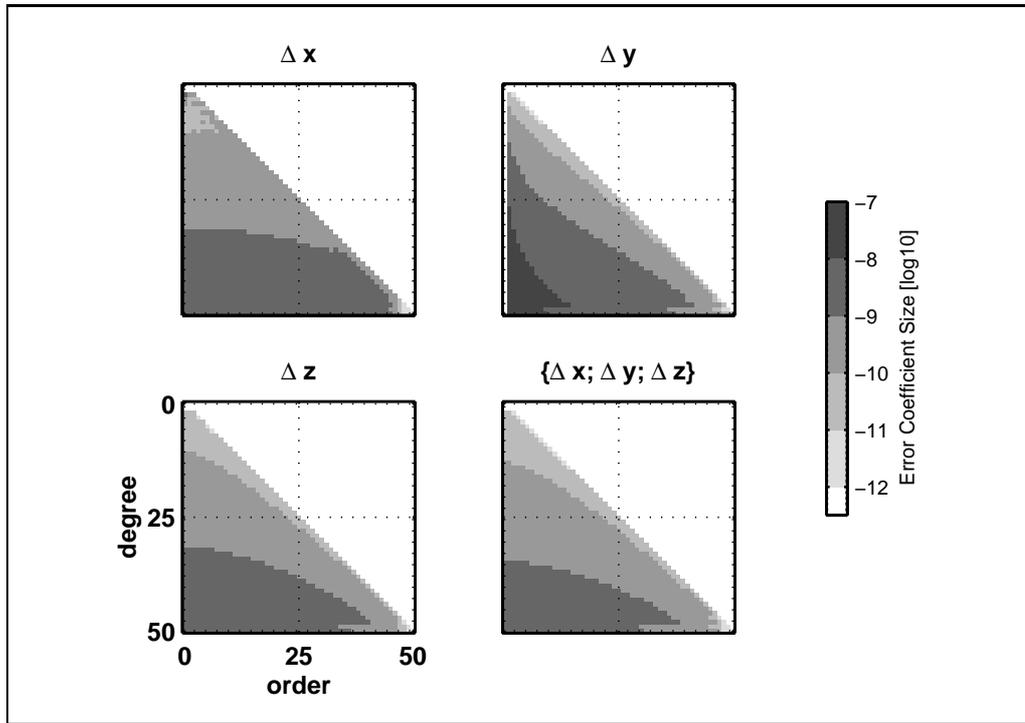
any real or planned mission. Nevertheless they possess some degree of realism concerning observation accuracy and orbit design.

Orbit perturbations are the result of unknown, or unmodelled, gravity field parameters, fed through the dynamic model. The significant digits spectrum in fig. 7.1, for example, reflects which coefficients have been visible in the original data. Especially those in the resonant orders have played a role apparently. For a quick first answer as to which coefficients play a role in the above circumstances, *sensitivity analysis* is used. In fig. 7.14, dimensionless orbit perturbation transfer is shown. In order to derive maximum orbit perturbations from them, the transfer has to be multiplied with potential coefficient magnitude—either real or from Kaula’s rule, e.g.— and dimensioned with the Earth’s radius  $R$ , cf. (Reigber, 1989).

Nevertheless, as noted in 6.5, sensitivity analysis is a *forward* modelling tool. It does not answer the question of recoverability of potential coefficients, which is an *inverse* problem. The results of a least squares SH error analysis are shown in fig. 7.15 (1D) and fig. 7.16 (2D). It shows that gravity field recovery from the radial component yields the lowest RMS. The along-track component performs nearly as good, although especially the lower degrees are worse. The cross-track result is nearly an order of magnitude worse than the radial, apart from the very lowest degrees. The combination solution lies somewhat below the  $\Delta z$ -only curve.

At the highest degrees, the errors seem to drop slightly. This effect is visible both in the 1D and the 2D error spectra. This must be an artifact of the LS inversion, that is not well understood. In combination with a larger maximum degree  $L$  this apparently better accuracy vanishes at the original location. It shifts towards the new maximum degree, see also 7.5.

*Redundancy analysis*, cf. 6.5, reveals the contribution of the individual observ-



**Figure 7.16:** SH error spectra of individual orbit perturbation components and of combination solution.

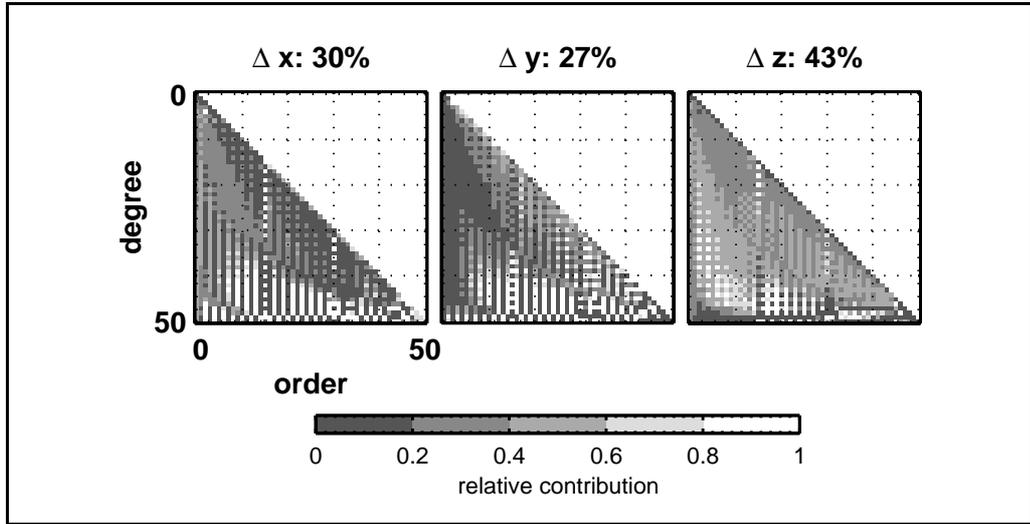
ables to a combination solution. The result, fig. 7.17, is less smooth and less isotropic than the SH error spectra themselves. Especially at the resonant orders, see below, and at the higher degrees this behaviour is strongly present. Apparently a smooth solution—smooth in terms of  $\sigma_{lm}$ —can be obtained from any single component. Nevertheless, for certain coefficients the contribution from other observables is preferred in a combination solution. For the particular simulation *hi-lo* SST the relative strength of  $\Delta x$ ,  $\Delta y$  and  $\Delta z$  is 30%, 27% and 43% respectively. This shows again that the radial orbit perturbation is the strongest component for gravity field recovery.

The error spectra from  $\Delta x$  and  $\Delta z$  have similar characteristics: in general isotropic, although (near-)sectorial coefficients are preferred. This similarity may not come as a surprise. These two are the in-plane components, that are coupled through their differential equations (4.9). The near-isotropy of  $\Delta z$  is explained in (Migliaccio, Sacerdote & Sansò, 1992), where the first time-integral of the  $x$ -equation is substituted in the  $z$ -equation, leading to:

$$\ddot{z} + n^2 z \approx -\Delta g. \quad (7.2)$$

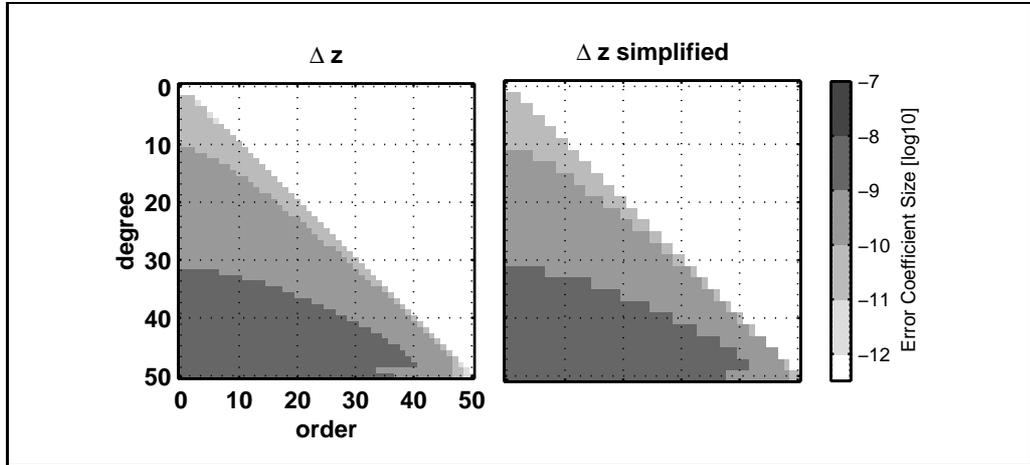
This equation is only approximative, due to a certain simplification in the time-integration of  $x$ . Equation (7.2) says that the radial motion is a harmonic oscillator, driven by gravity anomalies at satellite altitude. At the level of transfer coefficients this can be confirmed if  $H_{lmk}^{\Delta z}$ , cf. (4.11), is simplified by considering that the ratio of frequencies  $\dot{\Lambda}/\dot{u}$  is small. For LEO's it is  $\dot{\Lambda}/\dot{u} \approx 0.06$  (16 revolutions per day). Therefore  $\beta_{mk} = k + m\dot{\Lambda}/\dot{u} \approx k$ . In effect this is the same approximation, as made in (Migliaccio et al., 1992). It results in:

$$H_{lmk}^{\Delta z} \approx R \left( \frac{R}{r} \right)^{l-1} \frac{l-1}{k^2-1} \bar{F}_{lmk}(I), \quad (7.3)$$



**Figure 7.17:** Contribution of individual orbit perturbation components to combination solution.

in which  $(l - 1)$  is the specific transfer of gravity anomalies indeed. Of course, the smaller terms neglected here, are those with order  $m$ . For growing order  $m$ , the approximation becomes worse. This is not the cause, though, for the better accuracy of the (near-)sectorial coefficients, as shown in fig. 7.18. Even if (7.3) is used, the preference for (near-)sectorials remains. This is most probably due to the denominator  $(k^2 - 1)$ .



**Figure 7.18:** SH error spectra of original  $\Delta z$  and of  $\Delta z$  with isotropic approximation.

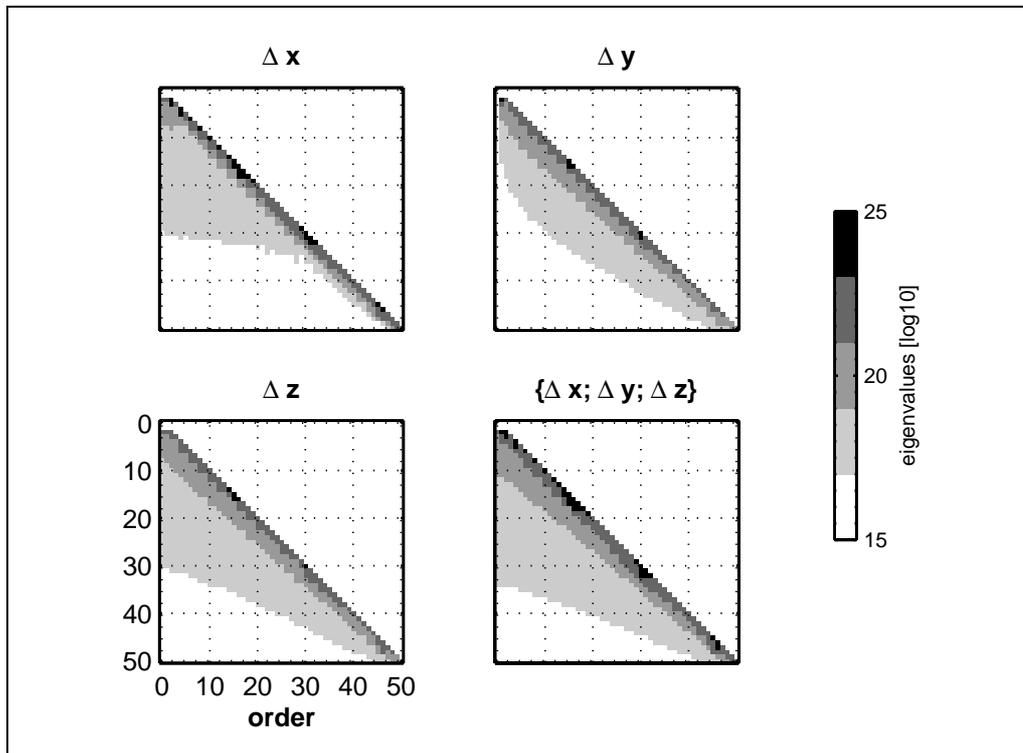
**Resonance.** A striking feature—or better, absence of feature—is the fact that no resonance bands shows up in the SH error spectra. Although slightly visible in the contribution spectra fig. 7.17 this feature is not as marked as could be expected from the sensitivity spectrum fig. 7.14. To understand this situation, the LC formulation (5.3) is recalled here:

$$A_{mk}^{\#} = \sum_{l=l_0}^L H_{lmk}^{\#} \bar{K}_{lm} \quad , \quad l_0 = \max(|m|, |k|) .$$

For each order  $m$  this is a matrix system with a left-hand vector  $y$  of lumped coefficients (running index  $k$ ), a right-hand vector  $x$  of potential coefficients (running index  $l$ ) and a matrix  $A$  with transfer coefficients, shaped like the rotated trapezium shape of fig. 5.2.

When (near-) resonance occurs, a certain lumped coefficient will be large. This is one particular element in  $y$ . The corresponding row in  $A$  will consist of large entries as well. Despite this one row, most information comes from all remaining non-resonant LC's, with corresponding well-behaved matrix entries. Viewed otherwise, all  $\bar{K}_{lm}$  are obtained from all  $A_{mk}$  (of given  $m$ ), of which only one or a few are resonant.

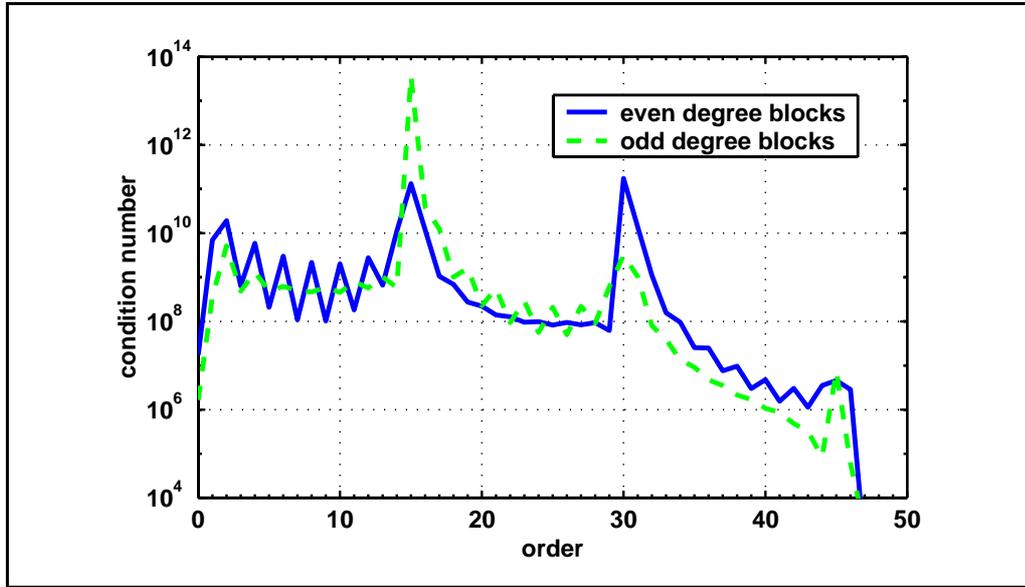
This reasoning is not in contradiction with the discussion on satellite-only models, 7.1. There the gravity field information is obtained to a very large extent from low-frequency, i.e. near-resonant information. This is in contrast to our situation of continuous along-orbit time-series with corresponding wide-band spectra.



**Figure 7.19:** Eigenspectra of the normal matrices ( $m$ -blocks) of simulations hi-lo SST.

Eigenanalysis suggests that the presence of (near-)resonance is numerically unfavourable for gravity field recovery from continuous time-series. Figure 7.19 shows the eigenspectra of the normal matrix blocks. To be precise the  $\log_{10}(\lambda_n)$  are shown. The horizontal axis represents order  $m$  again, but the vertical axis displays an eigenvalue index  $n$ . Eigenvalues are sorted downwards. Since there are as many eigenvalues as unknowns, the eigenspectra can be visualized as SH triangles as well.

The eigenvalues decay rapidly over 10 orders of magnitude. This tells that most information is contained in a few eigenvalues. Moreover, this phenomenon might explain the relative strength of sectorial coefficients. In case of (near-)resonance,



**Figure 7.20:** Condition numbers of the normal matrices ( $m$ -blocks) of simulations hi-lo SST.

which happens around the orders  $\gamma m_r$  with  $\gamma=0,1,2,3$  and  $m_r=15$  in this simulation ( $h = 500$  km:  $\beta/\alpha=15.2$ ), the maximum eigenvalue is considerably larger. The other eigenvalues of that specific order, including  $\lambda_{\min}$  remain similar to their neighbouring orders. Consequently the condition number becomes larger as well. Figure 7.20 shows condition numbers for each order  $m$ , one for the even degree, one for the odd degree  $m$ -blocks. The condition numbers in the resonance bands around  $m = \gamma m_r$  increase up to 4 orders of magnitude. From this viewpoint resonance leads to a numerically unwanted situation.

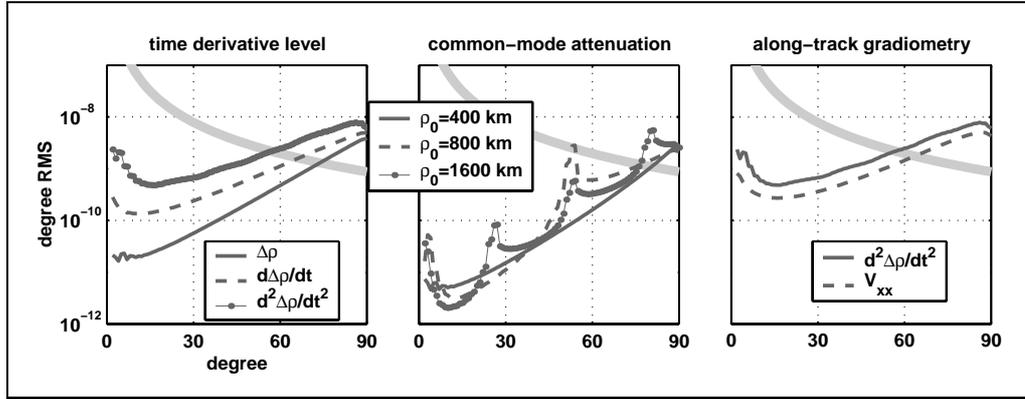
No conditioning peaks seems to appear at  $m = 0$ . However, the resonance at  $m = k = 0$  is an exact one, which has been removed from the equation system. Similarly this can be done for near-resonance, which would reduce the condition number levels. This question remains a trade-off between information disposal and numerical stability of the inverse problem.

## 7.4 Intersatellite Range Perturbations

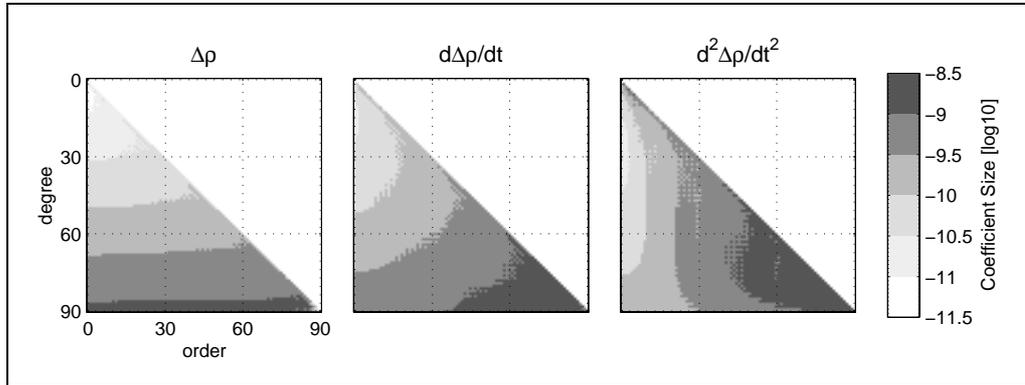
**Lo-Lo SST 1:  $\Delta\rho, \Delta\dot{\rho}$  and  $\Delta\ddot{\rho}$ .** The left panel of fig. 7.21 shows the error results of white noise simulations of the range observables  $\Delta\rho$ , range rate  $\Delta\dot{\rho}$  and range acceleration  $\Delta\ddot{\rho}$ . The noise levels have been chosen such, that around  $l = 90$  the three curves converge. The absolute level, though, is unimportant, since it is scaled by the PSD level itself. More interesting are the slopes of the  $\text{RMS}_l$  curves. The higher the level of time-differentiation, the flatter the curves. Due to the factor  $i\psi_{mk} = in\beta_{mk}$  that comes with each time-derivative, the higher frequencies are amplified. Consequently, the higher degree coefficients will be determined better, which explains the flattening of the error curves shown.

**Remark 7.3** *The same phenomenon would occur for spatial derivatives, in case  $V$  and  $\nabla V$  would be observable functionals. The transfer coefficients of  $V, \nabla V$*

and  $V_{ij}$  are of the order  $\mathcal{O}(l^0, |k|^0)$ ,  $\mathcal{O}(l^1, |k|^1)$  and  $\mathcal{O}(l^2, lk, |k|^2)$ , respectively.



**Figure 7.21:**  $\text{RMS}_l$  of the lo-lo SST simulations 1 (left), 2 (middle) and 3 (right), cf. tbl. 7.1 for simulation parameters. The thick grey line denotes Kaula's rule-of-thumb.



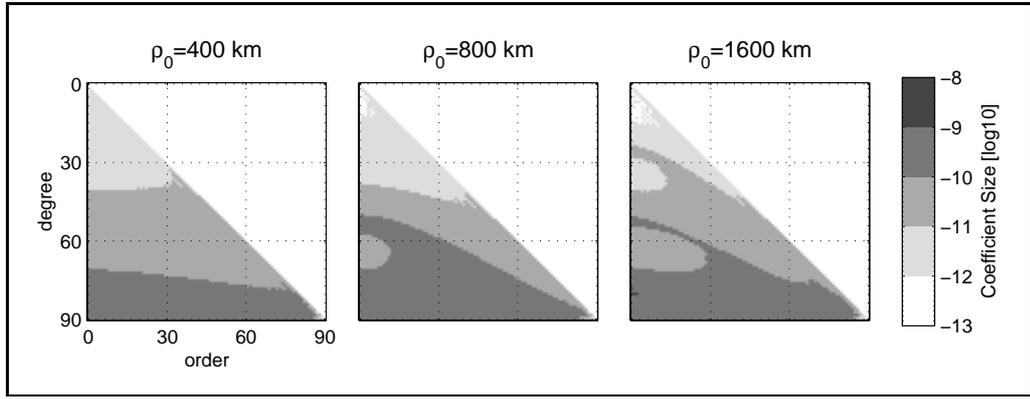
**Figure 7.22:** SH error spectra of the lo-lo SST simulations 1: range, range rate and range acceleration (left-to-right).

The corresponding SH error spectra, cf. Figure 7.22, show that the  $\text{RMS}_l$  curves do not represent the full error spectra. Only the range observable yields an isotropic error spectrum. It is comparable to the error spectra of  $\Delta x$  and  $\Delta z$ , as can be expected from the transfer coefficient (4.16). The sectorials are degraded somewhat, though, such that the error spectrum of  $\Delta\rho$  is even more isotropic. With each time differentiation, the coefficient accuracies concentrate towards the zonals. Although higher order coefficients get worse, the higher degrees benefit from the differentiation, such that on average the  $\text{RMS}_l$  curves flatten out.

A further consequence of the multiplication by  $in\beta_{mk}$  is the appearance of resonance band-like structures at  $m = \gamma m_r$  with  $m_r \approx 15$ .

**Lo-Lo SST 2: Common Mode Attenuation.** In 4.4 the factor  $\sin(\eta\beta_{mk})$  was interpreted as a common-mode attenuation effect. At certain frequencies, both satellites undergo the same motion. The range observable therefore cancels out this common-mode motion. Upon approximating again  $\beta_{mk} \approx k$ , it is:

$$\sin(\eta\beta_{mk}) \approx \sin(k\eta) \approx \sin\left(k\frac{\rho_0}{2R}\right).$$



**Figure 7.23:** SH error spectra of lo-lo SST simulations with different baselines

Attenuation occurs if the sine term becomes zero:

$$k \frac{\rho_0}{2R} \approx n\pi > k \approx n2\pi \frac{R}{\rho_0}, \quad (7.4)$$

with integer  $n$ . Since  $|k| \leq l$  the attenuation effect is to be expected around  $l \approx n2\pi \frac{R}{\rho_0}$ .

The simulation set lo-lo SST2 assesses the gravity field accuracy from range measurements over 400 km, 800 km and 1600 km baselines. The middle panel of fig. 7.21 shows the  $\text{RMS}_l$  of these three simulations. According to the above rule-of-thumb attenuation must be expected around degrees  $100n$ ,  $50n$  and  $25n$ , respectively. With  $L = 90$  the 400 km baseline simulation does not reach its first attenuation at  $n = 1$ . The simulation with  $\rho_0 = 800$  km indeed has a peak of deteriorated  $\text{RMS}_l$  around  $l = 50$ . The longer baseline of 1600 km shows the expected attenuations as well. As fig. 7.23 reveals, the rule-of-thumb of (7.4) does not give the full picture. The SH error spectra show variations with the order  $m$  as well.

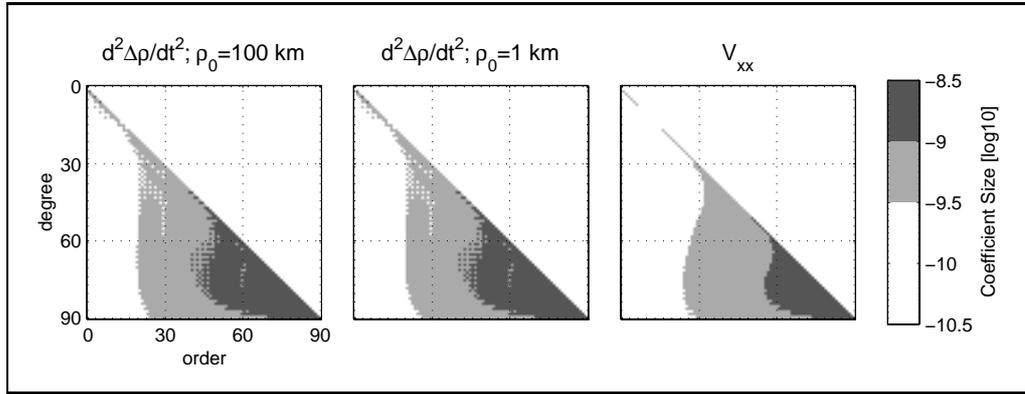
In order to avoid attenuation, the maximum degree must be smaller than the first attenuation band. Thus the following constraint must hold:

$$L < \frac{2\pi R}{\rho_0} = \frac{\pi}{\eta}, \text{ or } \rho_0 < \frac{2\pi R}{L}. \quad (7.5)$$

That leads to the constraint that the baseline  $\rho_0$  must be smaller than the smallest wavelength, to be determined at degree  $L$ . If the baseline is longer than  $\frac{2\pi R}{L}$ , the implication is not that the higher degrees cannot be determined anymore. Only certain bands of SH coefficients will suffer from reduced accuracy, cf. fig. 7.21 and fig. 7.23.

According to the  $\text{RMS}_l$  curves in fig. 7.21, the accuracy of the 800 km baseline simulation at the degrees before the first attenuation is better than that of the 400 km baseline. This is due to the scaling effect of the baseline-length through  $\eta$  in the factor  $\sin(\eta\beta_{mk})$ . See also the next paragraph.

**Lo-Lo SST 3: Range Acceleration and Along-Track Gradiometry.** In the simulation set lo-lo SST3, the assertion of 4.6 is tested. To this end three cases are



**Figure 7.24:** Comparing SH error spectra of two  $\Delta\ddot{\rho}$  simulations to  $V_{xx}$ .

compared:

$$\begin{aligned}\Delta\ddot{\rho} &\sim 10^{-6} \text{ m/s}^2/\sqrt{\text{Hz}}, \text{ with } \rho_0 = 100 \text{ km}, \\ \Delta\ddot{\rho} &\sim 10^{-8} \text{ m/s}^2/\sqrt{\text{Hz}}, \text{ with } \rho_0 = 1 \text{ km}, \\ V_{xx} &\sim 10^{-2} \text{ E}/\sqrt{\text{Hz}},\end{aligned}$$

that all correspond to a PSD of  $1 \text{ E}/\sqrt{\text{Hz}}$ . The  $\text{RMS}_l$  curves of both  $\Delta\ddot{\rho}$  simulations in fig. 7.21 (right panel) are equal. This is due to the small-baseline approximation of  $\Delta\ddot{\rho}/\rho_0$  (4.22), which is invariant to  $\eta$ . Therefore, if the ratio  $\text{PSD}/\rho_0$  of two different small-baseline simulations is equal, the SH error results must be equal as well.

Apart from a constant offset, the range acceleration result compares very well with that of the along-track gradient observable. Also in the SH error triangle, cf. fig. 7.24, the similarity is striking. Only the resonance-like structures of the range observable are absent in gradiometry. It can be concluded that the observation of range accelerations over short baselines approximates along-track gradiometry indeed.

## 7.5 Miscellaneous

**PSD and Mission Length.** Both PSD ( $S$ ) and mission length ( $T$ ) go into the error modelling through the stochastic model, cf. 5.4. From (5.12) the error variance of a given lumped coefficient is  $\sigma_{mk}^2 = S(f_{mk})/T$ . Together with the formalism of least squares error analysis, 6, it can be derived easily how scaling of  $S$  and/or  $T$  translates into scaling of the resulting SH error spectrum  $\sigma_{lm}$ . The variances  $\sigma_{mk}^2$  constitute the diagonal covariance matrix  $Q_y$ . Its inverse, the weight matrix  $P_y = Q_y^{-1}$ , goes into the normal matrix  $N$ . Finally, the inverse of the normal matrix yields the covariance matrix of the unknowns, whose diagonal are the variances  $\sigma_{lm}^2$ .

$$S, T \rightarrow \sigma_{mk}^2 \xrightarrow{\text{diag.}} Q_y \xrightarrow{\text{inv.}} P_y \xrightarrow{A^T P_y A} N \xrightarrow{\text{inv.}} Q_{\hat{x}} \xrightarrow{\text{diag.}} \sigma_{lm}^2.$$

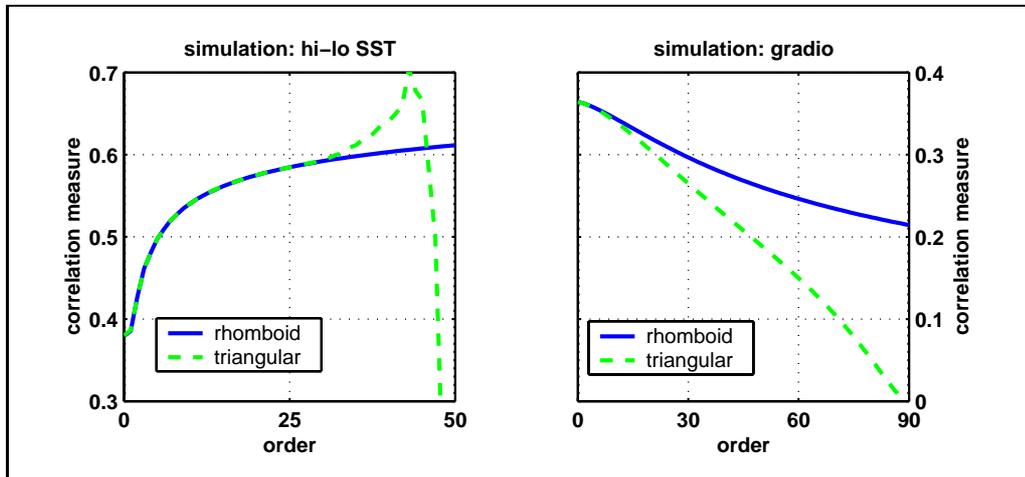
From this reasoning it can be decided that the a posteriori error variance of the unknowns is proportional to PSD, expressed in  $\text{units}^2/\text{Hz}$ , and inversely proportional to mission length. Thus, if a given combination  $S, T$  produces an error

spectrum  $\sigma_{lm}$ , the following scaling table holds:

$$S, T \rightarrow \sigma_{lm} \Rightarrow \begin{cases} S, nT \rightarrow \sigma_{lm}/\sqrt{n} \\ nS, T \rightarrow \sqrt{n}\sigma_{lm} \end{cases}.$$

If the mission duration is multiplied by  $n$ , the gravity field improves by  $\sqrt{n}$ . If the PSD level—in terms of  $\text{units}^2/\text{Hz}$ —becomes higher by a factor  $n$ , the gravity field gets worse by  $\sqrt{n}$ . These simple scaling rules are reproduced by LS error simulation. Results will not be shown graphically here, though.

**Correlations and SH Domain.** An inherent problem in the LC approach is the resulting correlation between coefficients. So far, only error spectra and derived products of LS error simulation have been shown. They represent the main diagonal of  $Q_{\hat{x}}$  only. Equation (6.24) presents a measure for correlation within a covariance matrix. Since it is a scalar, it can only be considered a rough measure. A purely diagonal matrix would yield  $c = 0$ , whereas  $c = 1$  represents high correlation.

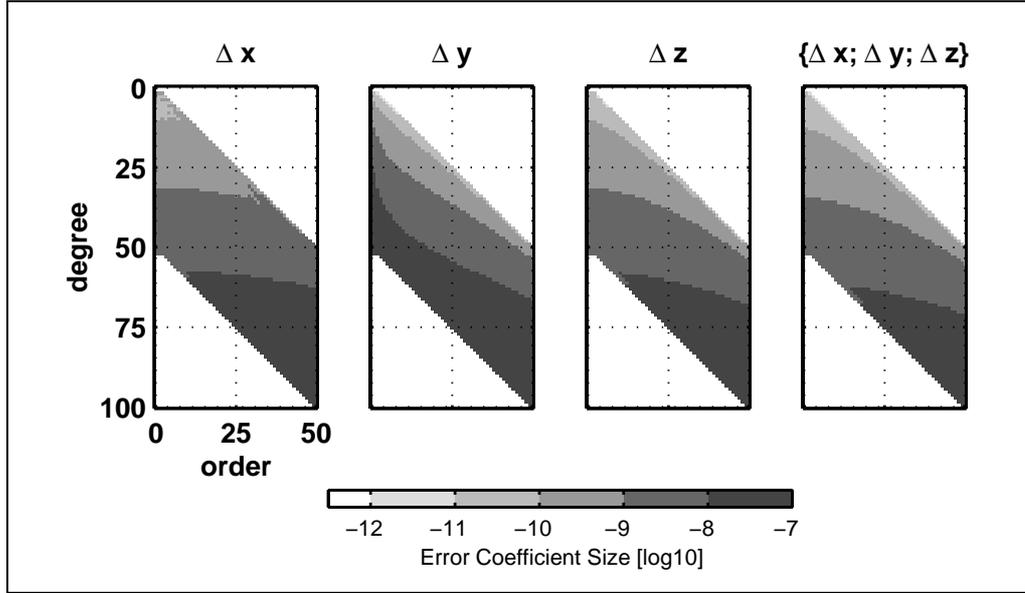


**Figure 7.25:** Correlation measures for each  $m$ -block in the combined hi-lo SST simulation and the combined gradiometer simulation. Both the conventional triangular and the rhomboid SH domain are considered. Note the different scales.

Figure 7.25 visualizes correlation numbers for the combined solutions of the hi-lo SST simulation and for the full-tensor solution of the gradiometer simulation, see the dashed lines. Correlations range from 0 to 0.7. Such high correlations are nothing uncommon in satellite-only gravity field solutions (Haagmans & Van Gelderen, 1991). Both curves start—incidentally—around the same level, somewhat below 0.4. The hi-lo SST correlations increase for increasing order  $m$ , although the  $m$ -blocks reduce in size. The correlations for the gradiometry simulation, on the other hand, decrease towards zero. Thus gradiometry brings about better error decorrelation than orbit perturbation observations.

A tentative explanation of this behaviour can be found in the transfer coefficients themselves. Apart from resonance, the transfer coefficients for orbit perturbation observables are of the order  $\mathcal{O}(l^{-1}, k^{-1})$ , those for gradiometry are  $\mathcal{O}(l^2, k^2)$ . Multiplied with the appropriate upward continuation term, this leaves more variation over  $l$  and  $k$  in case of the gradiometric observables than in case of orbit

perturbations. This will lead to larger variations in the normal matrix and apparently to more diagonal dominance of the inverted normal matrix in case of gradiometry.



**Figure 7.26:** SH error spectra of individual orbit perturbation components and of combination solution on a rhomboid SH domain, see also fig. 7.16.

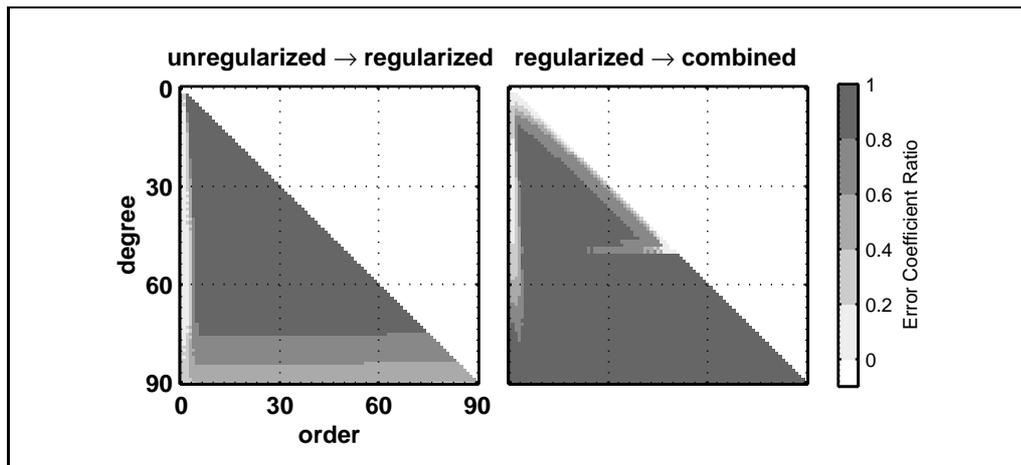
If correlations between SH coefficients of given  $m$  are high, it can be expected that variation of the maximum degree influences the solution, both in terms of error ( $\sigma_{lm}$ ) and of signal ( $\bar{K}_{lm}$ ). In order to test this assertion the correlation numbers have been calculated for the same simulations above on a rhomboid SH domain. The maximum degree depends on the order such that the  $m$ -blocks are of constant size, cf. fig. 5.3. Results are shown in fig. 7.25 too. The hi-lo SST simulation result on the rhomb is equivalent to the triangular domain result up to degree 30, after which both curves deviate. The rhomboid domain, though, shows more consistent behaviour. The gradiometry results start to diverge from the beginning already. The difference increases with order  $m$ .

Extending the SH domain does not affect the a posteriori errors. The error spectra of the hi-lo SST simulation on a rhomboid domain, fig. 7.26, show the same error level up to degree  $L = 50$  as the ones in fig. 7.16. The same behaviour, not shown here, can be found for other observables. The question, whether estimation of coefficients themselves will be affected, cannot be answered by LS error analysis. However, the discussion on unresolved parameters and aliasing, 6.3, already showed that unresolved parameters will project onto resolved coefficients particularly when  $A_1^T P_y A_2$  is non-diagonal. Thus, the correlation measure, found here, also represents the potential level of aliasing.

**Correlations and Multi-Observable Solutions.** In case of many observables, to which also regularization is reckoned, the key to LS error simulation is basically a stacking of normal equation matrices ( $m$ -blocks), cf. the multi-observables model in 6.4. An interesting phenomenon, due to correlations, arises if  $m$ -blocks of distinct maximum degree are added. As an example of this phenomenon the following three error simulation results are presented:

- *unregularized*: the  $V_{zz}$  component from the polar gap simulation with  $I = 97^\circ$
- *regularized*: the same with regularization using Kaula's rule
- *combined*: the *regularized* solution plus the hi-lo SST combination solution

Note that the former two simulations are up to  $L = 90$ , whereas the normal matrices from the hi-lo SST simulation have  $L = 50$ . Figure 7.27 shows the error spectra ratios  $\sigma_{lm}^{\text{reg.}}/\sigma_{lm}^{\text{unr.}}$  (left) and  $\sigma_{lm}^{\text{comb.}}/\sigma_{lm}^{\text{reg.}}$  (right). A ratio below one denotes an improvement of the numerator with respect to the denominator.



**Figure 7.27:** Ratios of regularized vs. unregularized (left) and of regularized vs. combined solution (right) error spectra. A ratio below 1 denotes an improvement in the direction of the arrow. Note the improvement in the right triangle for low orders and degrees above 50.

The left triangle shows that regularization takes effect in the polar gap wedge, due to bad conditioning, and near the resolution. This is in line with plots like fig. 7.6. For the combination solution, the hi-lo SST normal equations are added. But only up to  $L = 50$ . Nevertheless, the right triangle in fig. 7.27 reveals an improvement in the polar gap wedge beyond  $L = 50$  as well. Due to correlation, implied by the lumped coefficient formulation, the higher degree coefficients benefit from new information on the lower degrees.

## 7.6 Summary

- i) The SH error spectrum can be anisotropic for various reasons, among which spatial and temporal data heterogeneity, type of observable, noise model and non-polar orbit. Care must be taken when interpreting 1D spectral error measures like error degree variances or  $\text{RMS}_l$ . In the case of polar gap, the error measure  $\text{MED}_l$  is a useful alternative.
- ii) Spectrally, the effect of a non-polar orbit is noticeable on low orders only. In the spatial domain the propagated errors remain confined to the polar gap itself.
- iii) Through the isotropic combinations  $\{V_{zz}\}$ ,  $\{V_{xx} + V_{yy}\}$ ,  $\{V_{xz}; V_{yz}\}$  and  $\{V_{xx} - V_{yy}; 2V_{xy}\}$  full tensor gradiometry actually yields four independent and spectrally complete potential field observables.

- 
- iv) The best observable for single component gradiometry is the in-line radial term  $V_{zz}$ .
  - v) Provided global coverage by means of a polar orbit, the full gravity field can be recovered from single component gradiometry, except for zonal structures in case of  $V_{xy}$  or  $V_{yz}$ .
  - vi) The semi-analytic approach can hardly predict LS error simulation results, based on  $H_{lmk}^{\#}$  alone.
  - vii) The best component for gravity field recovery from high-low SST is the radial perturbation  $\Delta z$ . The along-track observable  $\Delta x$  results in similar, though somewhat degraded, error behaviour, and the cross-track direction performs worst.
  - viii) The high-low SST simulation showed that conclusions, based on sensitivity analysis, can oppose conclusions from SH error analysis. The former is a forward modelling tool, whereas the latter deals with inverse modelling.
  - ix) Resonance in observables, related to orbit perturbations, leads to the numerically unfavourable situation of increased condition number.
  - x) Gradiometry results in better error decorrelation than orbit perturbation observations.
  - xi) Extending the SH domain does not affect the a posteriori errors.
  - xii) In multi-observable solutions, due to correlation, the higher degree coefficients may benefit from new information on the lower degrees.
  - xiii) Higher derivatives, both temporal and spatial, favour the higher degree harmonics. Therefore, satellite gravity gradiometry and low-low range rate observation—or even better: range accelerometry—are the optimal space methods for high resolution gravity field recovery.
  - xiv) The low-low SST signal is attenuated around degrees that are integer multiples of  $2\pi R/\rho_0$ . In order to avoid this common-mode effect, the baseline length should be chosen smaller than  $2\pi R/L$ , with  $L$  the required resolution. Nevertheless a longer baseline does allow for the full recovery of the higher degrees, though with reduced accuracy for certain coefficients.
  - xv) It was shown that small-baseline range acceleration approximates along-track in-line gradiometry very well.

# 8 Concluding Remarks

## 8.1 Discussion

THE SEMI-ANALYTICAL approach to gravity field recovery—in connection with least squares error simulation techniques—has been shown to be a fast, flexible and powerful pre-mission error assessment tool. Moreover, the same methods, presented in this work, can be applied to actual gravity field recovery from real satellite data. However, this latter application has not been worked out in detail, yet.

The strength of the semi-analytical approach derives from the following characteristics:

- i) The lumped coefficient formulation provides a *linear* relationship between unknowns and observables.
- ii) The corresponding matrix of partials, or design matrix, is formulated *analytically* by means of the transfer coefficients  $H_{lmk}^{\#}$ . There is no need for numerical integration of variational equations if the observables are of perturbation type. The analytical formulation, or any analytical formulation for that matter, provides *insight* into the problem.
- iii) Employing a nominal orbit of constant radius and inclination, the LC formulation leads to *block-diagonal* normal matrices. Maximum block size is  $\frac{1}{2}L \times \frac{1}{2}L$ , with  $L$  the maximum degree of spherical harmonic development.
- iv) The semi-analytical approach presents a *unified* scheme for dynamical satellite geodesy. The pocket guide of transfer coefficients may incorporate any gravity field related observable. Also mechanisms for deriving further  $H_{lmk}^{\#}$  are provided.
- v) The LC formulation is a spectral domain formulation, which allows a straightforward use of error PSDs in the stochastic model.

In order to achieve the LC formulation with block-diagonality, a nominal orbit has to be introduced. Although a prerequisite, this is the main weakness of the semi-analytical approach. Transfer coefficients, i.e. the partials, are to be evaluated on this nominal orbit, leading to an approximation of the design matrix. Although most satellites for gravity field determination purposes have near-circular orbits, height variations of several kilometers will always exist due to residual eccentricity and periodic  $J_2$  perturbations. Two counter-measures to make the observations consistent with the linear model were proposed in 3:

- reduction of the observations, in particular height reduction, and
- iteration of the gravity field recovery, in case of coefficient determination

(not for LS error simulation).

Moreover, the concept of *best-knowledge* reference orbit was introduced in the context of iteration.

## 8.2 Outlook

FURTHER development and use of the semi-analytical approach are required or may be pursued in a number of areas. Obviously, the pocket guide may be extended to other functionals of the geopotential. In particular transfer coefficients  $H_{lmk}^{\#}$  can be derived that are related to orbit perturbations in coordinates or orbit variables, different from the local Cartesian perturbations in this work.

Moreover, the lumped coefficient formulation should be extended as to include eccentric nominal orbits. The constant radius approximation would become an approximation of constant semi-major axis and eccentricity. Extension of the LC formulation would imply an additional summation, i.e. lumped coefficients with a third index. Although this extension has been discussed briefly in 5.5, the full implications concerning sampling rate, mission duration, spectral mapping, and so on, have to be investigated.

A necessary step beyond the error analysis scope of this work is full gravity field recovery. It will be interesting to investigate how well the semi-analytical approach performs in estimating SH coefficients. A central issue will be the validity of the linear model and, connected to it, the convergence behaviour of the iteration process. Especially for observables of perturbation type, this may prove a difficult task. For gradiometry a first successful attempt in this direction is made in (SID, 2000).

Further efforts should be undertaken to validate the results of the semi-analytical approach with other techniques, notably brute-force numerical ones. This applies both to LS error simulation, and to gravity field recovery itself. Apart from the validity of the linear model, a weak point in the LS error simulation is the reliance on (the quality of) a priori stochastic information. In (SID, 2000) an error simulation by means of the semi-analytical approach has been corroborated by a full coefficient determination. Error simulation results, e.g.  $\sigma_{lm}$ , are compared to absolute coefficient differences  $\Delta\bar{K}_{lm}$ . The coefficient recovery, though, was also based on a semi-analytical approach, making use of iteration. Nevertheless, the results compare very well.

The interpretation of lumped coefficients as the 2D Fourier spectrum of a function on a torus contains interesting aspects:

- The torus may be used as a surface to interpolate the data, given as a function of  $u$  and  $\Lambda$ , in case of data gaps and interruptions. This is in contrast to the time-wise approach, where interpolation is difficult.
- The torus is a closed surface, which inherently leads to a discrete spectrum with regular frequency spacing. The concept of a repeat orbit is not needed. The constant radius approximation remains.
- The orientation of the local satellite frame poses no problems, as it may occur in the space-wise approach.
- As opposed to the space-wise approach, ascending and descending tracks are

separated naturally on the torus.

Thus the torus represents an intermediate data processing approach between time-wise and space-wise methods. It combines the positive characteristics of both. These aspects have to be worked out in more detail, though. It can be expected that a torus-wise data processing approach is a viable alternative to time-wise and space-wise methods.

# List of Symbols

## Coordinates

$x, y, z$	Cartesian coordinates in the local orbital frame: along-track, cross-track, radial
$\theta, \lambda, r$	spherical coordinates: co-latitude, longitude, radius
$u, \Lambda$	orbital angular variables or <i>torus</i> coordinates: argument of latitude, longitude of ascending node
$\rho$	inter-satellite range

## Functions

$\bar{Y}_{lm}(\theta, \lambda)$	surface spherical harmonic (complex, normalized)
$\bar{P}_{lm}(\cos \theta)$	fully normalized associated Legendre function
$\bar{d}_{lmk}, \bar{D}_{lmk}$	representation coefficients of the SO(3) group
$\bar{F}_{lmk}(I)$	inclination function (complex, normalized)
$\bar{F}_{lmk}^*(I)$	cross-track inclination function (complex, normalized)

## Coefficients

$\bar{K}_{lm}$	complex SH coefficient (normalized)
$\bar{C}_{lm}, \bar{S}_{lm}$	real SH coefficient (normalized)
$A_{mk}$	lumped coefficient, 2D Fourier coefficient
$H_{lmk}$	transfer coefficient, sensitivity coefficient
$\psi_{mk}, \dot{\psi}_{mk}$	orbital angular variable and signal spectrum

## List of labels to $H_{lmk}$ and $A_{mk}$

#	generic label
$V$	potential
$x, y, z$	1st derivatives in the local orbital frame (see above)
$xx, yy, zz$	main diagonal components of the tensor of 2nd derivatives
$xy, xz, yz$	off-diagonal components of the tensor of 2nd derivatives
$\Delta x, \Delta y, \Delta z$	orbit perturbations in the local orbital frame
$\Delta \dot{x}, \Delta \dot{y}, \Delta \dot{z}$	velocity perturbations in the local orbital frame
$\Delta \ddot{x}, \Delta \ddot{y}, \Delta \ddot{z}$	acceleration perturbations in the local orbital frame
$\Delta \rho, \Delta \dot{\rho}, \Delta \ddot{\rho}$	low-low SST range, range rate and range acceleration perturbations

# List of Abbreviations

CPR.....	Cycles Per Revolution
DSG.....	Dynamical Satellite Geodesy
DC.....	Direct Current (zero frequency)
DE.....	Differential Equation(s)
FFT.....	Fast Fourier Transform
GPS.....	Global Positioning System
HE.....	Hill Equations
LC.....	Lumped Coefficient
LS.....	Least Squares
LEO.....	Low Earth Orbiter
PG.....	Pocket Guide
RMS.....	Root Mean Square
RSS.....	Root Sum Square
SH.....	Spherical Harmonic
SLR.....	Satellite Laser Ranging
SNR.....	Signal-to-Noise Ratio
SST.....	Satellite-to-Satellite Tracking
SGG.....	Satellite Gravity Gradiometry
TEE.....	Truncation Error Effect

# A Properties of Inclination Functions

**I**NCLINATION functions  $\bar{F}_{lmk}(I)$  and cross-track inclination functions  $\bar{F}_{lmk}^*(I)$  are defined by (2.14) and (4.1) respectively. Basic constituents of these functions are the representation coefficients  $\bar{d}_{lmk}$ , given by (2.11). For convenience the definitions are repeated here:

$$\bar{F}_{lmk}(I) = i^{k-m} \bar{d}_{lmk}(I) \bar{P}_{lk}(0), \quad (\text{A.1a})$$

$$\bar{F}_{lmk}^*(I) = i^{k-m} \bar{d}_{lmk}(I) \bar{P}'_{lk}(0), \quad (\text{A.1b})$$

and

$$\bar{d}_{lmk}(I) = \left[ \frac{(l+k)!(l-k)!}{(l+m)!(l-m)!} \right]^{\frac{1}{2}} \sum_{t=t_1}^{t_2} \binom{l+m}{t} \binom{l-m}{l-k-t} (-1)^t c^{2l-a} s^a, \quad (\text{A.2})$$

with  $c = \cos \frac{1}{2}I$ ,  $s = \sin \frac{1}{2}I$ ,  $a = k - m + 2t$ ,  $t_1 = \max(0, m - k)$  and  $t_2 = \min(l - k, l + m)$ . Based on symmetries of the representation coefficients, some elementary properties of inclination functions are derived here.

## A.1 Symmetries of Representation Coefficients

**T**HE COEFFICIENTS  $\bar{d}_{lmk}$  are a matrix representation of rotation around the  $y$ -axis, e.g. (Edmonds, 1957). For each degree  $l$  they constitute a  $(2l+1)$  square matrix, that transforms the spherical harmonics according to (2.9). The resulting matrix is orthonormal, immediately giving rise to the following property:

$$\bar{d}_{lmk}(-I) = \bar{d}_{lk m}(I).$$

On the other hand, using the angle  $-I$  in (A.2) directly leads to:

$$\bar{d}_{lmk}(-I) = (-1)^{k-m} \bar{d}_{lmk}(I).$$

Substituting respectively the combinations  $\{k, m\}$  and  $\{-m, -k\}$  for  $\{m, k\}$  into (A.2) gives a further symmetry:

$$\bar{d}_{lk m}(I) = \bar{d}_{l, -m, -k}(I).$$

A further property can be derived from the observation that the only contribution to a rotation about  $\pi$  in (A.2) comes from the term with  $t = l + m$  in case of  $k = -m$ , cf. (Edmonds, 1957):

$$\bar{d}_{lmk}(\pi) = (-1)^{l+m} \delta_{k, -m}.$$

This is a matrix with alternating plus or minus one on the anti-diagonal. It reflects the fact that spherical harmonics with  $l - m$  odd, i.e. making use of odd Legendre

functions, will change sign when North- and Southpole are interchanged. This relation gives rise to the following symmetry in the argument around  $\frac{1}{2}\pi$ :

$$\bar{d}_{lmk}(\pi - I) = \sum_n \bar{d}_{lmn}(\pi) \bar{d}_{lnk}(-I) = (-1)^{l+m} \bar{d}_{l,-m,k}(-I) = (-1)^{l+k} \bar{d}_{l,-m,k}(I).$$

Combination of the above properties provides the following basic set:

$$\bar{d}_{lmk}(I) = (-1)^{k-m} \bar{d}_{lk m}(I) \tag{A.3a}$$

$$\bar{d}_{lmk}(I) = (-1)^{k-m} \bar{d}_{l,-m,-k}(I) \tag{A.3b}$$

$$\bar{d}_{lmk}(I) = (-1)^{k-m} \bar{d}_{lmk}(-I) \tag{A.3c}$$

$$\bar{d}_{lmk}(I) = (-1)^{l-m} \bar{d}_{l,m,-k}(\pi - I) \tag{A.3d}$$

$$\bar{d}_{lmk}(I) = (-1)^{l-k} \bar{d}_{l,-m,k}(\pi - I) \tag{A.3e}$$

These are symmetries in the indices, (A.3a) and (A.3b), or in the argument, (A.3c)-(A.3e). Only the first 4 properties are independent, the last one is derived from them. Nevertheless, it will be useful for the forthcoming properties of inclination functions.

## A.2 Symmetries of Inclination Functions

FROM (2.3) we know  $\bar{P}_{lk} = (-1)^k \bar{P}_{l,-k}$ . Inserting this, together with (A.3) into definition (A.1a), yields the following basic properties of inclination functions:

$$\bar{F}_{lmk}(I) = (-1)^k \bar{F}_{l,-m,-k}(I) \tag{A.4b}$$

$$\bar{F}_{lmk}(I) = (-1)^{k-m} \bar{F}_{lmk}(-I) \tag{A.4c}$$

$$\bar{F}_{lmk}(I) = (-1)^{l-m} \bar{F}_{l,m,-k}(\pi - I) \tag{A.4d}$$

$$\bar{F}_{lmk}(I) = (-1)^{l-k-m} \bar{F}_{l,-m,k}(\pi - I) \tag{A.4e}$$

Property (A.3a) could not be employed here. Since the inclination is conventionally defined in the range  $[0, \pi)$  property (A.4c) has no further meaning, apart from the conclusion that  $\bar{F}_{lmk}(I)$  must vanish in case  $k - m$  odd. As special cases the zonal inclination functions are produced by setting  $m = 0$ :

$$\bar{F}_{l,0,k}(I) = (-1)^k \bar{F}_{l,0,-k}(I) \tag{A.5b}$$

$$\bar{F}_{l,0,k}(I) = (-1)^l \bar{F}_{l,0,-k}(\pi - I) \tag{A.5d}$$

$$\bar{F}_{l,0,k}(I) = (-1)^{l-k} \bar{F}_{l,0,k}(\pi - I) \tag{A.5e}$$

As a further specification, the case  $I = \frac{1}{2}\pi$  is considered:

$$\bar{F}_{l,0,k}(\frac{1}{2}\pi) = (-1)^k \bar{F}_{l,0,-k}(\frac{1}{2}\pi) \tag{A.6b}$$

$$\bar{F}_{l,0,k}(\frac{1}{2}\pi) = (-1)^l \bar{F}_{l,0,-k}(\frac{1}{2}\pi) \tag{A.6d}$$

$$\bar{F}_{l,0,k}(\frac{1}{2}\pi) = (-1)^{l-k} \bar{F}_{l,0,k}(\frac{1}{2}\pi) \tag{A.6e}$$

Since  $\bar{P}_{lk}(0)$  and therefore  $\bar{F}_{lmk}(I)$  vanishes for  $l - k$  odd, (A.6e) is superfluous. Moreover it allows for the transition from (A.6b) to (A.6d).

If  $\bar{P}_{lk} = (-1)^k \bar{P}_{l,-k}$ , then consequently  $\bar{P}'_{lk} = (-1)^k \bar{P}'_{l,-k}$  as well. This leads to the conclusion that all above properties are valid for cross-track inclination functions too. Note however, that the equivalent of (A.6e) becomes:

$$\bar{F}_{l,0,k}^* \left( \frac{1}{2}\pi \right) = (-1)^{l-k} \bar{F}_{l,0,k}^* \left( \frac{1}{2}\pi \right) = 0. \quad (\text{A.7})$$

Since  $P'_{lk}(0)$  vanishes for  $l - k$  even,  $\bar{F}_{l,0,k}^* \left( \frac{1}{2}\pi \right)$  must vanish as well.

In case of real-valued inclination functions, i.e. with  $m \geq 0$ , properties (A.4b) and (A.4e) are only of interest for the special case  $m = 0$ .

## B Block-Diagonal Error Propagation

**E**RROR analysis of the gravity recovery capability of satellite missions, based on analytical theories and fulfilling some further simplifying assumptions, returns a covariance matrix which is strictly block-diagonal. On the other hand, real gravity field determination from observables to several satellites yields a full covariance matrix, which reveals a block diagonal *predominance*. In both cases it is therefore interesting to investigate the covariance propagation to the sphere of such (exactly or nearly) block-diagonal structures.

Covariance propagation of real or simulated error covariances  $\text{COV}\{\bar{K}_{lm}, \bar{K}_{nm}\}$  onto other geopotential functionals yields the so-called *commission errors*. The same formalism can be applied, though, to investigate the truncation effect of maximum degree  $L$ , yielding the *omission errors*. To that end the error covariances  $\text{COV}\{\bar{K}_{lm}, \bar{K}_{nm}\}$  are replaced by a signal covariance model, describing signal size.

Though we will mainly be concerned with propagation of a block-diagonal covariance matrix, the derivation starts at the level of full covariance propagation. From the resulting formulae, simplifications are made to accommodate block-diagonal structure ( $m=k$ ), purely diagonal structure ( $n=l$  as well) and isotropic behaviour (independence of order  $m$ ), respectively. For each of the 4 cases, both spatial covariances and variances are considered.

### B.1 Full Covariance Propagation

**T**HE COVARIANCE propagation equations are based on a two-step formulation of spherical harmonic synthesis, e.g. (Sneeuw, 1994a), applied to a complex formulation as in (2.1). Similar, though less compact, equations have been derived in real-valued terms in (Haagmans & Van Gelderen, 1991). If convenient, equations will be written in matrix-vector notation as well. However, many indices and variables play a role in the covariance between two points:  $\theta_i, \theta_j, \lambda_i, \lambda_j$ , degrees  $l$  and  $n$ , and orders  $m$  and  $k$ . Thus matrix notation is not a proper tool for such a multitude of ‘dimensions’. Since complex quantities are used, the matrix transpose  $A^T$  is replaced by the Hermitian form  $A^\dagger$ , i.e. transposed *and* complex conjugated.

**Two-Step Spherical Harmonic Synthesis.** The gravitational potential, e.g. equation (2.1), and its (isotropic) functionals are defined in a two-step approach by:

$$A_m(\theta) = \sum_{l=|m|}^L \lambda_l \bar{P}_{lm}(\cos \theta) \bar{K}_{lm} > a = P\Lambda\kappa, \quad (\text{B.1a})$$

$$f(\theta, \lambda) = \sum_{m=-L}^L A_m(\theta) e^{im\lambda} > f = Fa. \quad (\text{B.1b})$$

where  $\lambda_l$  may contain dimensioning, upward continuation, cap-smoothing (e.g. Pellinen's  $\beta_l$  factors) and of course the specific transfer (eigenvalue) of the functional  $f$  under consideration (Rummel & Van Gelderen, 1995). In matrix-vector notation  $P$  is a real matrix, representing the associated Legendre functions.  $\Lambda$  has the  $\lambda_l$  on its diagonal. The matrix  $F$  denotes the discrete Fourier transform, cf. (Strang, 1986).

The matrix notation is misleading in the sense that vector  $a$  in (B.1a) has  $\theta$  (or better: a discretized  $\theta_i$ ) as index, whereas  $a$  in (B.1b) uses order  $m$ . In numerical practice one would make use of a matrix, say  $A$ , having one type of vector  $a$  in its columns, the other type in its rows. For instance: put the resulting  $a$  from (B.1a) for each order  $m$  in consecutive rows of  $A$ , and perform FFT over the columns of  $A$ . FFT over the columns of  $A$  is equivalent to—though faster than—the matrix multiplication  $FA$ . The same problem will show up with the covariances of  $a$ .

**Full Covariance Propagation** from the SH spectral domain  $\text{COV}\{\bar{K}_{lm}, \bar{K}_{nk}\}$  to the spatial domain  $\text{COV}\{f(\theta_1, \lambda_1), f(\theta_2, \lambda_2)\}$  is written now in a two-step formulation. The intermediate product is  $\text{COV}\{A_m(\theta_1), A_k(\theta_2)\}$ . With the abbreviations

$$\begin{aligned} \text{COV}\{A_m(\theta_1), A_k(\theta_2)\} &:= \text{COV}\{A_m^1, A_k^2\} \rightarrow Q_a \\ \text{COV}\{f(\theta_1, \lambda_1), f(\theta_2, \lambda_2)\} &:= \text{COV}\{f_1, f_2\} \rightarrow Q_f \end{aligned}$$

one gets:

$$\begin{aligned} \text{COV}\{A_m^1, A_k^2\} &= \sum_l \sum_n \lambda_l \lambda_n \text{COV}\{\bar{K}_{lm}, \bar{K}_{nk}\} \bar{P}_{lm}(\cos \theta_1) \bar{P}_{nk}(\cos \theta_2) \\ &> Q_a = P \Lambda Q_\kappa \Lambda P^\top, \end{aligned} \quad (\text{B.2a})$$

$$\begin{aligned} \text{COV}\{f_1, f_2\} &= \sum_m \sum_k \text{COV}\{A_m^1, A_k^2\} e^{i(m\lambda_1 - k\lambda_2)} \\ &> Q_f = F Q_a F^\dagger. \end{aligned} \quad (\text{B.2b})$$

Summations are assumed to take place over valid index ranges. They will not be mentioned explicitly in the sequel. The matrix notation is oversimplified again, since  $Q_a$  in (B.2b) has  $m$  and  $k$  as row/column indices with  $\theta_1$  and  $\theta_2$  fixed, and vice versa for  $Q_a$  in (B.2a). Actual computation of spatial covariances  $Q_f$  would make use of a mixed formulation. One possibility would be to fix e.g. point 2. In that case  $\text{COV}\{f_1, f_2\}$  is the spatial covariance function in point  $(\theta_2, \lambda_2)$ . Analogous to the global spherical harmonic synthesis above, the first step would be to compute a vector for each order  $m$ :

$$q = P \Lambda Q_\kappa \Lambda p^\top,$$

with  $p$  one row from the full matrix  $P$  only, corresponding to point 2. The resulting covariance  $q$  is a vector with  $\theta_1$  as index. For each order  $m$ , these vectors are stored in the columns of a matrix, say  $A$  again. The second step would read:

$$Q_f = \sum_k (FA) e^{-ik\lambda_2}.$$

Note that for each order  $k$  a new matrix  $A$  has to be computed. Alternatively, as Haagmans & Van Gelderen (1991) point out, (B.2b) is a 2D Fourier transform. Application of 2D FFT renders a powerful algorithm for covariance computation. In this approach two parallels  $\theta_1$  and  $\theta_2$  would have to be fixed.

The variance  $\text{VAR}\{f(\theta, \lambda)\}$  is attained by setting  $\theta_1=\theta_2=\theta$  and  $\lambda_1=\lambda_2=\lambda$  in (B.2). Its evaluation remains cumbersome. Note that the exponent in (B.2a) reduces to  $e^{i(m-k)\lambda}$ . If  $\text{COV}\{A_m, A_k\}$  would be a Toeplitz matrix, i.e. depending on  $(m-k)$  only, the second step may be reduced further.

## B.2 Block Covariance Propagation

A NUMERICALLY more advantageous situation arises, if the covariance matrix of SH coefficients  $\text{COV}\{\bar{K}_{lm}, \bar{K}_{nk}\}$  is block-diagonal. When only the situation  $k=m$  occurs, one summation disappears. The above formulae reduce to:

$$\text{COV}\{A_m^1, A_m^2\} = \sum_l \sum_n \lambda_l \lambda_n \text{COV}\{\bar{K}_{lm}, \bar{K}_{nm}\} \bar{P}_{lm}(\cos \theta_1) \bar{P}_{nm}(\cos \theta_2) \quad (\text{B.3a})$$

$$\text{COV}\{f_1, f_2\} = \sum_m \text{COV}\{A_m^1, A_m^2\} e^{im(\lambda_1 - \lambda_2)}. \quad (\text{B.3b})$$

The first step of the evaluation algorithm remains the same:

- fix the coordinates of point 2,
- evaluate for each order  $m$  the vector  $q = P \Lambda Q_\kappa \Lambda p^\top$ , and
- store them in the rows of some matrix  $A$ .

The second step however does not require the summation over  $k$  anymore. Moreover, the covariance function only depends on longitude difference  $\Delta\lambda = \lambda_1 - \lambda_2$ . The second step reduces to 1D FFT over the columns of  $A$ , which is equivalent to  $Q_f = FA$ . The block-diagonal structure causes the spatial covariance function to be stationary in longitude. Thus the above algorithm requires to fix one parallel. It then calculates a spatial covariance function, valid for the whole parallel.

Computing the variance  $\text{VAR}\{f\}$  requires, again, to set  $\theta_1=\theta_2=\theta$  and  $\lambda_1=\lambda_2=\lambda$  in (B.3). The terms with longitude in (B.3b) disappear, leaving:

$$\text{VAR}\{f\} = \sum_m \text{VAR}\{A_m(\theta)\}. \quad (\text{B.4})$$

Therefore the spatial variance of  $f(\theta, \lambda)$  is a function of  $\theta$  alone. Now  $\text{VAR}\{A_m(\theta)\}$  is the diagonal of  $\text{COV}\{A_m, A_m\}$ . The variance can thus be evaluated quickly by:

$$\text{VAR}\{f\} = \sum_m \text{diag}(P \Lambda Q_\kappa \Lambda P^\top). \quad (\text{B.5})$$

Since the product of a Legendre function with itself is always an even function with respect to the equator,  $\text{VAR}\{f\}$  is necessarily an even function as well. To summarize, a block-diagonal structure of  $Q_\kappa$  results in:

- i) a covariance function that is stationary in longitude,
- ii) a variance that only varies with latitude,
- iii) a variance that is symmetric in the equator.

### B.3 Diagonal Covariance Propagation

A FURTHER simplification arises if the covariance matrix is (assumed to be) strictly diagonal. The formulae of the previous section can be reduced by setting  $n=l$ , cancelling one summation again:

$$\text{COV}\{A_m^1, A_m^2\} = \sum_l \lambda_l^2 \text{VAR}\{\bar{K}_{lm}\} \bar{P}_{lm}(\cos \theta_1) \bar{P}_{lm}(\cos \theta_2) \quad (\text{B.6a})$$

$$\text{COV}\{f_1, f_2\} = \sum_m \text{COV}\{A_m^1, A_m^2\} e^{im(\lambda_1 - \lambda_2)}. \quad (\text{B.6b})$$

The variance is written in short as:

$$\text{VAR}\{f\} = \sum_l \sum_m \lambda_l^2 \text{VAR}\{\bar{K}_{lm}\} \bar{P}_{lm}(\cos \theta)^2.$$

The evaluation algorithms for variance and covariance use the same matrix implementations as above, with the exception that  $Q_\kappa$  is diagonal now.

### B.4 Isotropic Covariance Propagation

IF THE SH error spectrum depends furthermore on the degree  $l$  alone, it is said to be *isotropic*. An isotropic error spectrum propagates into a spatial error covariance function, that is azimuth independent. By writing (B.6a) and (B.6b) together, setting  $\text{VAR}\{\bar{K}_{lm}\} = \varepsilon_l/(2l+1)$ , and making use of the addition theorem of spherical harmonics, one arrives at:

$$\begin{aligned} \text{COV}\{f_1, f_2\} &= \sum_l \lambda_l^2 \frac{\varepsilon_l}{2l+1} \sum_m \bar{Y}_{lm}(\theta_1, \lambda_1) \bar{Y}_{lm}(\theta_2, \lambda_2) \\ &= \sum_l \lambda_l^2 \varepsilon_l P_l(\cos \psi_{12}). \end{aligned} \quad (\text{B.7})$$

Indeed the covariance function only depends on the spherical distance  $\psi_{12}$ . Thus it is fully homogeneous. The *error degree variance*  $\varepsilon_l$  can be considered as sum of the variances  $\text{VAR}\{\bar{K}_{lm}\}$  of a specific degree  $l$  over all orders  $m$ , see also (Haagmans & Van Gelderen, 1991). Due to  $P_l(1) = 1$ , the variance has the simple expression:

$$\text{VAR}\{f\} = \sum_l \lambda_l^2 \varepsilon_l, \quad (\text{B.8})$$

which is a constant over the sphere. Note that the addition theorem can only be applied in case of the conventional triangular SH domain. All orders  $m$  have to be available for each degree  $l$ .

### B.5 Along-Orbit Covariance Propagation

ERROR covariances of the SH coefficients can be propagated onto covariances of the along-orbit functional in a two-step approach as well. In a first step

they are propagated onto the lumped coefficients. Making use of the definitions (3.1b) and dropping the labels, one simply obtains:

$$A_{mk} = \sum_l H_{lmk} \bar{K}_{lm} > a = H\kappa,$$

$$\text{COV}\{A_{mk}, A_{mp}\} = \sum_l \sum_n H_{lmk} \text{COV}\{\bar{K}_{lm}, \bar{K}_{nm}\} H_{nmp}^* > Q_a = HQ_\kappa H^\dagger.$$

A warning on notation is in place here. The index  $k$  is not used here as the alternative to order  $m$  as in the previous sections. It is just the second index of the lumped coefficient again. As alternative to  $k$  the index  $p$  is used here. Note that  $p$  is not meant here as the  $p = \frac{1}{2}(l - k)$  from Kaula's inclination functions. Due to the LC approach, this propagation starts at the 'block level' already. Further simplifications will not be given here, apart from the variance to variance formulation:

$$\text{VAR}\{A_{mk}\} = \sum_l |H_{lmk}|^2 \text{VAR}\{\bar{K}_{lm}\}. \quad (\text{B.9})$$

In a second step the matrices  $Q_a$  are further propagated onto the along-orbit signal  $f(t)$ . Due to (3.1a) we have:

$$\text{COV}\{f(t_1), f(t_2)\} = \sum_m \sum_k \sum_p e^{i\psi_{mk}(t_1)} \text{COV}\{A_{mk}, A_{mp}\} e^{-i\psi_{mp}(t_2)}. \quad (\text{B.10a})$$

On the nominal orbit  $\psi_{mk}(t)$  may be replaced by  $\dot{\psi}_{mk}t$ . When only variance is propagated, i.e. when  $k=p$ , (B.10a) reduces to:

$$\text{COV}\{f(t_1), f(t_2)\} = \sum_m \sum_k \text{VAR}\{A_{mk}\} e^{i\dot{\psi}_{mk}\tau_{12}}, \quad (\text{B.10b})$$

which only depends on the time difference  $\tau_{12}$ . The covariance function along the orbit is stationary. Including (B.9) one ends up with the variance to variance propagation:

$$\text{VAR}\{f\} = \sum_l \sum_m \sum_k |H_{lmk}|^2 \text{VAR}\{\bar{K}_{lm}\}. \quad (\text{B.10c})$$

The resulting variance  $\text{VAR}\{f\}$  is a constant along the orbit. It is only an approximation to the real error behaviour from the block-diagonal covariance propagation. Nonetheless it is a useful tool for error assessment. For instance, the  $\bar{K}_{lm}$ -errors can be projected onto nominal orbits with varying inclination. Since  $H_{lmk}$  is a function of inclination e.g., the variance  $\text{VAR}\{f\}$  becomes automatically a function of inclination as well.

## B.6 Omission Errors on the Sphere

**E**RROR covariance propagation deals with so-called commission errors. They are error characteristics of the parameters that are projected onto other quantities, dependent on the parameters. Omission errors, on the other hand, are due to limiting the SH domain, for instance by the cut-off degree  $L$  in the conventional triangular domain. They represent loss of higher frequency signal power. So actually the treatment of omission errors does not belong in this appendix. However, the formulae to be used are the same. The main differences are:

- Signal covariances  $E\{\bar{K}_{lm}\bar{K}_{lm}^*\}$  instead of error covariances have to be used.
- The SH domain under consideration is the complement of the domain, used before. In case of the conventional triangular domain, this complement is either defined by  $\sum_{l=L}^{\infty} \sum_{m=-l}^l$  or by  $\sum_{m=-\infty}^{\infty} \sum_{L}^{\infty}$ .

Since the coefficients, let alone their signal covariance, in the complementary domain are unknown, one has to resort to analytical models, e.g. Kaula's rule of thumb (Kaula, 1966) or the Tscherning-Rapp model (Tscherning & Rapp, 1974). These models describe the signal in terms of signal degree variance  $c_l$ , which is the total expected power per degree  $l$ , cf. also the error degree variances  $\varepsilon_l$  above:

$$c_l = \sum_{m=-l}^l E\{\bar{K}_{lm}\bar{K}_{lm}^*\} = (2l+1)E\{\bar{K}_{lm}\bar{K}_{lm}^*\}. \quad (\text{B.11})$$

Degree variance models are isotropic by definition, expressed by the right side of (B.11). For computing the omission errors, formulae (B.7) and (B.8) may thus be applied. They yield:

$$\text{COV}_{\text{om.}}\{f_1, f_2\} = \sum_{L+1}^{\infty} \lambda_l^2 c_l P_l(\cos \psi_{12}), \quad (\text{B.12})$$

$$\text{VAR}_{\text{om.}}\{f\} = \sum_{L+1}^{\infty} \lambda_l^2 c_l. \quad (\text{B.13})$$

Although it is unusual to present omission errors as covariance functions as well, it is a straightforward result of the error propagation formulae. For practical computations,  $\infty$  is replaced by some high degree, e.g. 1000 or 10000.

## B.7 Along-Orbit Omission Errors

**I**N B.5 ERROR covariance was propagated onto lumped coefficients and onto the along-orbit signal. Similarly, the omission errors can be calculated in these domains. To this end the error covariance  $\text{VAR}\{\bar{K}_{lm}\}$  is replaced by the average signal variance  $E\{\bar{K}_{lm}\bar{K}_{lm}^*\} = c_l/(2l+1)$ . The omission errors of the lumped coefficients become:

$$\text{COV}_{\text{om.}}\{A_{mk}, A_{mp}\} = \sum_{l=L+1}^{\infty} H_{lmk} \frac{c_l}{2l+1} H_{lmp}^*, \quad (\text{B.14a})$$

$$\text{VAR}_{\text{om.}}\{A_{mk}\} = \sum_{l=L+1}^{\infty} |H_{lmk}|^2 \frac{c_l}{2l+1}. \quad (\text{B.14b})$$

They are also known as the truncation error effect (TEE), cf. (Klokočník et al., 1990). The omission error along the orbit becomes:

$$\text{COV}_{\text{om.}}\{f(t_1), f(t_2)\} = \sum_{l=L+1}^{\infty} \sum_{m=-l}^l \sum_{k=-l}^l |H_{lmk}|^2 \frac{c_l}{2l+1} e^{i\dot{\psi}_{mk}\tau_{12}}, \quad (\text{B.14c})$$

$$\text{VAR}_{\text{om.}}\{f\} = \sum_{l=L+1}^{\infty} \sum_{m=-l}^l \sum_{k=-l}^l |H_{lmk}|^2 \frac{c_l}{2l+1}. \quad (\text{B.14d})$$

## References

- Balmino G (1993). Orbit Choice and the Theory of Radial Orbit Error for Altimetry, in: R Rummel & F Sansò (eds.), *Satellite Altimetry in Geodesy and Oceanography*, Lecture Notes in Earth Sciences, Vol. 50, Springer-Verlag, Berlin
- Balmino G, E Schrama, N Sneeuw (1996). Compatibility of first-order circular orbit perturbations theories; consequences for cross-track inclination functions, *J. Geodesy* **70**:554–561
- Bassanino M, F Migliaccio, F Sacerdote (1991). A BVP Approach to the Reduction of Spaceborne GPS And Accelerometric Observations, in: O Colombo (ed.), *From Mars to Greenland: Charting Gravity with Space and Airborne Instruments*, IAG Symposium 110, pp 323–337, Springer-Verlag, Berlin
- Bauer HF (1982). Environmental Effects on Micro-Gravity Experiments, *Z. Flugwiss. Weltraumforsch.* **6**:184–194
- Bettadpur SV (1991). *Hotine's nonsingular geopotential formulation and its expression in the perifocal frame*, CSR TM 91-04, Center for Space Research, Austin, Texas
- Bettadpur SV (1995). Hotine's geopotential formulation: revisited, *Bull. Géod.* **69**:135–142
- Betti B, F Sansò (1989). The Integrated Approach to Satellite Geodesy, in: F Sansò and R Rummel (eds.), *Theory of Satellite Geodesy and Gravity Field Determination*, Lecture Notes in Earth Sciences, Vol. 25, Springer-Verlag, Berlin
- Bierman GJ (1977). *Factorization Methods for Discrete Sequential Estimation*, Academic Press, New York
- Bouman J (1997a). *A Survey of Global Gravity Models*, DEOS Report 97.1, Delft Institute for Earth-Oriented Space Research
- Bouman J (1997b). Quality Assessment of Geopotential Models by Means of Redundancy Decomposition?, *DEOS Progress Letters* **97.1**:49–54
- Bouman J, R Koop (1997). Quality Differences between Tikhonov Regularization and Generalized Biased Estimation in Gradiometric Analysis, *DEOS Progress Letters* **97.1**:42–48
- Chauvineau B, F Mignard (1990). Dynamics of Binary Asteroids, *Icarus* **83**:360–381
- Colombo OL (1984). *The Global Mapping of Gravity With Two Satellites*, Publications on Geodesy, New Series **3**, Netherlands Geodetic Commission,
- Colombo OL (1986). Notes on the Mapping of the Gravity Field using Satellite Data, in: H Sünkel (ed.), *Mathematical and Numerical Techniques in Physical Geodesy*, Lecture Notes in Earth Sciences, Vol. 7, Springer-Verlag, Berlin
- Colombo OL (1989). The Dynamics Of Global Positioning System Orbits And The Determination Of Precise Ephemerides, *J. Geophys. Res.* **94**(B7):9167–9182
- Cui Ch, M Mareyen (1992). Gauss's equations of motion in terms of Hill variables and first application to numerical integration of satellite orbits, *manuscr. geod.* **17**:155–163
- Cui Ch, D Lelgemann (1995). Analytical dynamic orbit improvement for the evaluation of geodetic-geodynamic satellite data, *J. Geodesy* **70**:83–97
- Dong D, TA Herring, RW King (1998). Estimating regional deformation from a combination of space and terrestrial geodetic data, *J. Geodesy* **72**:200–214
- Edmonds AR (1957). *Angular Momentum in Quantum Mechanics*, Princeton University

- Press
- Emeljanov NV, AA Kanter (1989). A method to compute inclination functions and their derivatives, *manuscr. geod.* **14**:77–83
- Engelis T (1988). On the simultaneous improvement of a satellite orbit and determination of sea surface topography using altimeter data, *manuscr. geod.* **13**:180–190
- Exertier P, P Bonnefond (1997). Analytical solution of perturbed circular motion: application to satellite geodesy, *J. Geodesy* **71**:149–159
- Forward RL (1973). Review of artificial satellite gravity gradiometer techniques for geodesy, in: G Veis (ed.), *The use of artificial satellites for geodesy and geodynamics*, Athens, Greece, 1973
- Gaposchkin EM (1978). Recent advances in analytical satellite theory, Harvard-Smithsonian Center for Astrophysics, preprint series No. **1074**
- Giacaglia GEO (1980). Transformations of spherical harmonics and applications to geodesy and satellite theory, *Studia geoph. et geod.* **24**:1–11
- Gooding RH (1971). Lumped Fifteenth-order Harmonics in the Geopotential, *Nature, phys. Sci.* **231**(25):168–169
- Gooding RH, DG King-Hele (1989). Explicit forms of some functions arising in the analysis of resonant satellite orbits, *Proc. R. Soc. Lond. A* **422**:241–259
- Haagmans RHN, M van Gelderen (1991). Error Variances-Covariances of GEM-T1: Their Characteristics and Implications in Geoid Computation, *J. Geophys. Res.* **96**(B12):20011–20022
- Hansen PC (1992). Regularization Tools, A Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems, Danish Computing Center for Research and Education, Technical University of Denmark, Lyngby
- Heiskanen WA, H Moritz (1967). *Physical Geodesy*, Freeman, San Francisco
- Hill GW (1878). Researches in the Lunar Theory, *American Journal of Mathematics I* :5–26, 129–147, 245–260
- Hofmann-Wellenhof B, H Moritz (1986). Introduction to Spectral Analysis, in: H Sünkel (ed.), *Mathematical and Numerical Techniques in Physical Geodesy*, Lecture Notes in Earth Sciences, Vol. 7, Springer-Verlag, Berlin
- Hotine M (1969). *Mathematical Geodesy*, ESSA Monograph 2, U.S. Dept. of Commerce, Washington DC
- Ilk K-H (1983). *Ein Beitrag zur Dynamik ausgedehnter Körper — Gravitationswechselwirkung*, Deutsche Geodätische Kommission, Reihe C, Heft 288
- Jackson DD (1972). Interpretation of Inaccurate, Insufficient and Inconsistent Data, *Geophys. J. R. astr. Soc.* **28**:97–109
- Jackson DD (1979). The use of *a priori* data to resolve non-uniqueness in linear inversion, *Geophys. J. R. astr. Soc.* **57**:137–157
- Jekeli C, TN Upadhyay (1990). Gravity Estimation From STAGE, a Satellite-to-Satellite Tracking Mission, *J. Geophys. Res.* **95**(B7):10 973–10 985
- Kaplan H (1976). *Modern Spacecraft Dynamics & Control*, John Wiley & Sons, New York
- Kaula WM (1966). *Theory of Satellite Geodesy*, Blaisdell Publishing Company, Waltham Massachusetts
- Kaula WM (1983). Inference of Variations in the Gravity Field From Satellite-to-Satellite Range Rate, *J. Geophys. Res.* **88**(B10):8345–8349
- King-Hele D (1992). *A tapestry of orbits*, Cambridge University Press
- King-Hele DG, AN Winterbottom (1994). Comparison of geopotential harmonics in comprehensive models with those derived from satellite resonance, 1972–1993, *Planetary and Space Science* **42**(5):359–366
- Klokočník J (1988). Geopotential research mission (GRM): a contribution to the assessment of orbit accuracy, orbit determination and gravity field modelling, *Bull. Astron.*

*Inst. Czechosl.* **39**:45–67

- Klokočník J, J Kostelecký, H Li (1990). Best lumped coefficients from resonances for direct use in earth gravity field modelling, *Bull. Astron. Inst. Czechosl.* **41**:17–26
- Koop R (1993). *Global Gravity Field Modelling Using Satellite Gravity Gradiometry*, Publications on Geodesy, New Series **38**, Netherlands Geodetic Commission,
- Lambeck K (1988). *Geophysical Geodesy: The Slow Deformations of the Earth*, Clarendon Press, Oxford
- Lanczos C (1964). *Linear Differential Operators*, Van Nostrand
- Lange BO, BW Parkinson (1966). Error Equations of Inertial Navigation with Special Application to Orbital Determination and Guidance, in: RL Duncanbe, VG Szebehely (eds.), *Methods in Astrodynamics and Celestial Mechanics, Progr. in Astrodynamics and Aeronautics*, Vol. 17, Academic Press, New York
- Lemoine FG, SC Kenyon, JK Factor, RG Trimmer, NK Pavlis, DS Chinn, CM Cox, SM Klosko, SB Luthcke, MH Torrence, YM Wang, RG Williamson, EC Pavlis, RH Rapp, TR Olson (1998). The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96, NASA/TP—1998–206861
- Mackenzie R, P Moore (1997). A geopotential error analysis for a non planar satellite to satellite tracking mission, *J. Geodesy* **71**:262–272
- Migliaccio F, F Sacerdote, F Sansò (1992). An error analysis technique based on Hill's equations for GPS and accelerometric measurements performed on board a low orbiting satellite, *manuscr. geod.* **17**:257–269
- Moore P, DA Rothwell (1990). A study of gravitational and non-gravitational modelling errors in cross-over differences, *manuscr. geod.* **15**:187–206
- Moritz H (1980). *Advanced Physical Geodesy*, Wichmann, Karlsruhe, Germany
- NASA (1977). *National Geodetic Satellite Program*, NASA SP-365
- Nerem R, F Lerch, J Marshall, E Pavlis, B Putney, B Tapley, R Eanes, J Ries, B Schutz, C Shum, M Watkins, S Klosko, J Chan, S Luthcke, G Patel, N Pavlis, R Williamson, R Rapp, R Biancale, F Nouel (1994). Gravity model development for Topex/Poseidon: Joint Gravity Models 1 and 2, *J. Geophys. Res.* **99**(C12):24 421–24 447
- Obuchov AM (1947). Statistically Homogeneous Random Fields on a Sphere, *Uspechi Mat. Nauk* pp. 196–198 (in Russian)
- O'Keefe JA (1959). Zonal Harmonics of the Earth's Gravitational Field and the Basic Hypothesis of Geodesy, *J. Geophys. Res.* **64**(12):2389–2392
- Papoulis A (1965). *Probability, Random Variables and Stochastic Processes*, McGraw-Hill
- Reigber C, KH Ilk (1976). Vergleich von Resonanzparameterbestimmungen mittels Ausgleichung und Kollokation, *Z. f. Vermessungswesen* **101**(2):59–67
- Reigber C (1989). Gravity Field Recovery from Satellite Tracking Data, in: F Sansò, R Rummel (eds.), *Theory of Satellite Geodesy and Gravity Field Determination, Lecture Notes in Earth Sciences*, Vol. 25, Springer-Verlag, Berlin
- Rosborough GW, BD Tapley (1987). Radial, Transverse and Normal Satellite Position Perturbations due to the Geopotential, *Celest. Mech.* **40**:409–421
- Rummel R, C Reigber, KH Ilk (1978). The use of satellite-to-satellite tracking for gravity parameter recovery, in: *Proc. European Workshop on Space Oceanography, Navigation and Geodynamics*, ESA SP-137, pp. 153-161
- Rummel R (1986a). Satellite Gradiometry, in: H Sünkel (ed.), *Mathematical and Numerical Techniques in Physical Geodesy*, Lecture Notes in Earth Sciences, Vol. 7, Springer-Verlag, Berlin
- Rummel R (1986b). Satellite Gradiometry – A Promising Concept For 25 Years, in: M Schneider (ed.), *Die Arbeiten des Sonderforschungsbereiches 78 Satellitengeodäsie der Technischen Universität München*, Deutsche Geodätische Kommission, Astrono-

- misch-Geodätische Arbeiten, Heft 48
- Rummel R (1992). Geodesy, in: *Encyclopedia of Earth System Science, Vol. 2*, Academic Press
- Rummel R (1993). Principle of Satellite Altimetry and Elimination of Radial Orbit Errors, in: R Rummel, F Sansò (eds.), *Satellite Altimetry in Geodesy and Oceanography*, Lecture Notes in Earth Sciences, Vol. 50, Springer-Verlag, Berlin
- Rummel R, M van Gelderen, R Koop, E Schrama, F Sansò, M Brovelli, F Migliaccio, F Sacerdote (1993). *Spherical harmonic analysis of satellite gradiometry*, Publications on Geodesy, New Series **39**, Netherlands Geodetic Commission
- Rummel R, M van Gelderen (1995). Meissl scheme — spectral characteristics of physical geodesy, *manuscr. geod.* **20**:379–385
- Scheinert M (1996). *Zur Bahndynamik niedrigfliegender Satelliten*, Deutsche Geodätische Kommission, Reihe C, Heft 435
- Schneider M (1988). *Satellitengeodäsie*, BI-Wissenschaftsverlag, 1988
- Schrama EJO (1986). *A study of satellite-to-satellite tracking configuration by application of linear perturbation theory*, Delft University of Technology, Faculty of Geodetic Engineering, Mathematical and Physical Geodesy **86.3**
- Schrama EJO (1989). *The Role Of Orbit Errors In Processing Of Satellite Altimeter Data*, Publications on Geodesy, New Series **33**, Netherlands Geodetic Commission,
- Schrama EJO (1991). Gravity Field Error Analysis: Applications of Global Positioning System Receivers and Gradiometers on Low Orbiting Platforms, *J. Geophys. Res.* **96** (B12):20041–20051
- Schuh WD (1996). *Tailored Numerical Solution Strategies for the Global Determination of the Earth's Gravity Field*, Mitteilungen der geodätischen Institute der TU Graz, **81**
- Schwintzer P (1990). *Sensitivity analysis in least squares gravity field modelling by means of redundancy decomposition of stochastic a priori information*, Deutsches Geodätisches ForschungsInstitut, internal report
- Schwintzer P, C Reigber, W Barth, F-H Massmann, J Raimondo, M Gerstl, A Bode, H Li, R Biancale, G Balmino, B Moynot, J Lemoine (1992). GRIM4 – Globale Erdschwerefeldmodelle, *Z. f. Vermessungswesen* **117**(4):227-247
- Seeber G (1993). *Satellite Geodesy*, Walter de Gruyter
- Sharma J (1995). *Precise Determination of the Geopotential With a Low-Low Satellite-To-Satellite Tracking Mission*, The University of Texas at Austin
- SID consortium (2000). *GOCE End to End Performance Analysis*, Final Report, ESTEC Contract No. 12735/98/NL/GD
- Sneeuw N (1992). Representation coefficients and their use in satellite geodesy, *manuscr. geod.* **17**:117–123
- Sneeuw N (1994a). Global spherical harmonic analysis by least-squares and numerical quadrature methods in historical perspective, *Geophys. J. Int.* **118**:707–716
- Sneeuw N (1994b). Global Gravity Field Error Simulations for STEP-Geodesy, in: *A Major STEP for Geodesy*, Report 1994 of the STEP Geodesy Working Group, pp 45–54, Group, eds. R Rummel and P Schwintzer, GFZ Potsdam, Germany, Nov. 1994
- Sneeuw N, R Bun (1996). Global spherical harmonic computation by two-dimensional Fourier methods, *J. Geodesy* **70**:224–232
- Sneeuw N, M van Gelderen (1997). The Polar Gap, in: F Sansò & R Rummel (eds.), *Geodetic Boundary Value Problems in View of the One Centimeter Geoid*, Lecture Notes in Earth Sciences, Vol. 65, Springer-Verlag, Berlin
- Strang G (1986). *Introduction to Applied Mathematics*, Wellesley-Cambridge Press
- Thalhammer M (1995). *Regionale Gravitationsfeldbestimmung mit zukünftigen Satellitenmissionen (SST und Gradiometrie)*, Deutsche Geodätische Kommission, Reihe C, Heft 437
- Tscherning C, RH Rapp (1974). *Closed Covariance Expressions for Gravity Anoma-*

- lies, *Geoid Undulations, and Deflections of the Vertical Implied by Anomaly Degree Variance Models*, Dept. Geodetic Science, **208**, The Ohio State University, Columbus
- Wagner CA (1983). Direct Determination of Gravitational Harmonics From Low-Low GRAVSAT Data, *J. Geophys. Res.* **88**(B12):10309–10321
- Wagner CA, SM Klosko (1977). Gravitational harmonics from shallow resonant orbits, *Celest. Mech.* **16**:143–163
- Wakker KF (1988). SST principles and state of the art, in: *Study on Precise Gravity Field Determination Methods and Mission Requirements*, Final Report, ESA Contract 7521/87/F/FL
- Wigner EP (1959). *Group Theory and its Applications to the Quantum Mechanics of Atomic Spectra*, Academic Press, New York
- Wolff M (1969). Direct measurements of the Earth's gravitational potential using a satellite pair, *J. Geophys. Res.* **74**:5295–5300
- Wnuk E (1988). Tesseral harmonic perturbations for high order and degree harmonics, *Celest. Mech.* **44**:179–191
- Xu PL (1992). The value of minimum norm estimation of geopotential fields, *Geophys. J. Int.* **111**:170–178
- Xu PL, R Rummel (1994). Generalized ridge regression with applications in determination of potential fields, *manuscr. geod.* **20**:8–20
- Yunck TP (1993). A New Chapter in Precise Orbit Determination, *GPS World* 56–61