Regional gravity field modeling: A comparison of methods

Regional tyngdefeltsmodellering: En sammenligning av metoder

Vegard Ophaug

Norwegian University of Life Sciences
Department of Mathematical Sciences and Technology
Master's thesis 30 credits 2013

Norwegian University of Life Sciences

Department of Mathematical Sciences and Technology

Regional gravity field modeling: A comparison of methods

Regional tyngdefeltsmodellering: En sammenligning av metoder

by

Vegard Ophaug

Master's Thesis in

Geomatics

Ås, May 2013

Supervisor: Dr.-Ing. Christian Gerlach
Norwegian University of Life Sciences and
Bavarian Academy of Sciences and Humanities

Abstract

The geoid is an equipotential surface of Earth's gravity field, coinciding with mean ocean surface level, serving as a reference surface for orthometric heights. In addition to this geodetic application, the geoid finds geophysical uses in that it gives insight into Earth's inner density distribution.

This thesis comprises a literature study on three methods for regional geoid computation. The methods are Stokes integration with various kernel modifications, least-squares collocation and spherical wavelets, or more general spherical radial basis functions. The primary goal of this thesis is to compare these methods, theoretically and numerically with synthetic data.

Theoretically, in the global case, the three methods are equivalent. Regional numerical comparison of Stokes integration and least-squares collocation was done by closed-loop testing using a synthetic gravity field computed from the global geopotential model *EGM2008*. Both methods are practically equal.

The theoretical equivalence of global geoid computation with Shannon radial basis function and spherical harmonics was confirmed numerically.

For the signal representation with radial basis functions, parameter estimation methods were applied. When estimating parameters from observations, inverse problems are encountered. Spherical harmonic analysis and radial basis function analysis by least-squares adjustment are examples of linear inverse problems. Such problems are often ill-conditioned. Hence, the secondary goal of this thesis is to investigate ill-conditioned linear inverse problems as well as possible remedies.

Tikhonov regularization with prior information was applied to several cases of spherical harmonic analysis of a global gravity field computed from *EGM2008*. In turn, synthesized gravity fields were compared to "true" gravity fields computed from the *EGM2008* spherical harmonic coefficients. The L-curve approach for choosing an optimal regularization parameter was applied. The necessity of choosing the regularization parameter correctly, as well as the benefits of correctly applied regularization, was demonstrated.

Global radial basis function analysis by least-squares adjustment using Shannon RBF also formed an ill-conditioned problem that called for regularization. Tikhonov regularization with prior information was applied here as well, generally presenting similar traits as the spherical harmonic analysis cases.

If no regularization was applied, still, mathematically correct RBF coefficients were obtained, in that the solution allowed for the reconstruction of the input data. With the introduction of regularization, physically meaningful RBF coefficients were obtained.

Sammenfatning

Geoiden er en ekvipotensialflate i Jordas tyngdefelt, sammenfallende med midlere havnivå. En geodetisk anvendelse av geoiden er at den er en referanseflate for ortometriske høyder. I tillegg finner geoiden anvendelser i geofysikken, i det at den gir innsyn i Jordas indre massefordeling.

Denne oppgaven omfatter et litteraturstudium av tre metoder for regional geoideberegning. Metodene er Stokes-integrasjon med forskjellige kjernemodifikasjoner, minste kvadraters kollokasjon og sfæriske wavelets, eller mer generelle sfæriske radielle basisfunksjoner. Hovedmålet med oppgaven er å sammenligne disse metodene teoretisk og numerisk med syntetiske data.

Teoretisk, i det globale tilfellet, er de tre metodene ekvivalente. Regionale numeriske sammenligninger av Stokes-integrasjon og minste kvadraters kollokasjon ble gjort i en lukket sløyfe, ved at både syntetiske observasjoner benyttet i metodene og den til sammenligning "sanne" geoiden beregnes fra én og samme globale tyngdemodell, *EGM2008*. For alle praktiske formål viste begge metoder seg å være like.

Den teoretiske ekvivalens mellom en global geoide beregnet med den radielle basisfunksjonen Shannon og sfærisk-harmoniske funksjoner ble bekreftet numerisk.

For signalrepresentasjonen med radielle basisfunksjoner, ble parameterestimeringsmetoder anvendt. Når parametre skal estimeres fra observasjoner, utgjør dette inverse problemer. Sfærisk-harmonisk analyse og analyse med radielle basisfunksjoner ved minste kvadraters metode er eksempler på lineære inverse problemer. Slike problemer er gjerne dårlig stilte. Derfor er oppgavens sekundære mål å undersøke dårlig stilte lineære inverse problemer samt mulige løsninger.

Tikhonov-regularisering med tilleggsinformasjon om de ukjente parametre ble anvendt i flere tilfeller av sfærisk-harmonisk analyse av globale tyngdefelt beregnet fra *EGM2008*. Videre ble tyngdefeltene syntetisert og sammenlignet med "sanne" tyngdefelt beregnet fra de sfærisk-harmoniske koeffisienter i *EGM2008*. L-kurve-metoden for bestemmelse av regulariseringsparameteren ble anvendt. Nødvendigheten av riktig bestemmelse av regulariseringsparameter, samt fordelene ved riktig anvendelse av regularisering, ble demonstrert.

Global analyse med den radielle basisfunksjonen Shannon ved minste kvadraters metode utgjør også et dårlig stilt problem som behøver regularisering. Tikhonov-regularisering med tilleggsinformasjon om de ukjente parametre ble anvendt også her, med lignende trekk som i tilfellene sfærisk-harmonisk analyse.

Dersom ingen regularisering ble anvendt, var det likevel mulig å estimere matematisk korrekte RBF-koeffisienter, i det at inngangssignalet lot seg rekonstruere fra disse. Innføring av regularisering gjorde det mulig å estimere fysisk meningsfulle koeffisienter.

Acknowledgments

First of all I would like to thank my supervisor Christian Gerlach for his guidance, proofreading and endeavoring patience with all my questions.

I am very grateful for the opportunity to visit the Commission for Geodesy and Glaciology at the Bavarian Academy of Sciences and Humanities for two months during this spring and for the practical arrangements made by the commission prior to my visit.

I would also like to thank Katrin Bentel for guidance and ideas with regard to the application of wavelets (or more general radial basis functions).

Thanks to Jack Zeigler for triggering my interest in wavelets in the first place.

A significant part of the practical geoid determination and kernel modification theory in this thesis has its origin in lectures held by Michal Šprlák during fall 2012, and I am very grateful for the time he set aside for these lectures.

I would further like to thank my fellow geomatics "inmate" at the office space for master students, Eirik Oulie Rosbach, for general mental uplift during our long work days, securing a minimum level of outdoor running and some great laughs.

For the rock-solid support and encouragement from my family I am infinitely grateful.

Vegard Ophaug
$\varphi = 59^\circ\,39'\,52''\,\mathrm{N},\ \lambda = 10^\circ\,47'\,40''\,\mathrm{E},\ N = 39.1\,\mathrm{m}$
May 2013

Contents

1 Introduction **1**
1.1 Geophysical applications of the geoid . . . 2
1.2 Thesis objective . . . 4
1.3 Thesis summary . . . 5

2 Fundamentals of Earth's gravity field **7**
2.1 Gravitational potential and gravity . . . 7
2.2 The geodetic Earth model and normal gravity . . . 11
2.3 The linear gravity field . . . 15
2.4 Boundary value problems . . . 17
2.5 Statistics of the gravity field . . . 20

3 Gravity field modeling approaches **27**
3.1 Spherical harmonics . . . 28
3.1.1 Global geopotential models . . . 32
3.1.1.1 A remark on tides . . . 33
3.1.2 Spherical harmonic analysis by least-squares adjustment . . . 35
3.2 Integral formulae . . . 38
3.2.1 Derivation of Stokes's formula . . . 38
3.2.2 Solving Molodensky's problem by analytical continuation . . . 41
3.3 Least-squares collocation . . . 42
3.3.1 Least-squares prediction . . . 42
3.3.2 Generalization to least-squares collocation . . . 43
3.3.3 A remark on collocation with parameters . . . 46
3.4 Radial basis functions . . . 48
3.4.1 Gravity field functionals in terms of radial basis functions . . . 54
3.4.2 Examples of radial basis functions . . . 55
3.4.3 Radial basis function analysis by least-squares adjustment . . . 60
3.4.4 From radial basis functions to wavelets . . . 63
3.5 Model comparison . . . 63
3.5.1 Least-squares collocation versus Stokes integration . . . 64
3.5.2 Stokes integration versus spherical harmonics . . . 66
3.5.3 Spherical harmonics versus radial basis functions . . . 69

4 Practical aspects of gravity field modeling **73**
4.1 Reductions and corrections . . . 73
4.1.1 Topographic reduction schemes . . . 73

4.1.2 Atmospheric correction 83
4.1.3 Ellipsoidal correction 84
4.2 Practical evaluation methods for geoid determination 85
4.2.1 Combination in the frequency domain 86
4.2.2 Combination in the spatial domain 87
4.2.3 Geoid estimators 89
4.2.4 Statistics of geoid estimators 90
4.2.5 Kernel modifications 91
4.2.5.1 Deterministic modification schemes 92
4.2.5.2 Stochastic modification schemes 94
4.2.6 Numerical integration 97
4.2.7 Fast Fourier transform techniques 98

5 Numerical investigations **101**
5.1 Inverse and ill-posed problems 101
5.2 Tikhonov regularization 105
5.3 Examples of ill-conditioned systems and regularization using synthetic data 107
5.4 Agreement check of the different geoid modeling approaches with synthetic data 120
5.4.1 Creation of synthetic data sets 121
5.4.2 Agreement of Stokes numerical integration with spherical harmonic synthesis 124
5.4.3 Agreement of least-squares collocation with Stokes integration and spherical harmonic synthesis 130
5.4.4 Agreement of RBFs with spherical harmonic synthesis 138
5.5 Global RBF analysis and synthesis with Shannon RBF, applying regularization 138

6 Final remarks and outlook **159**

Bibliography **xix**

List of Figures

1.1 The effect of mass anomalies on the geoid 3

2.1 Potential, equipotential surfaces and gravity 9
2.2 Astronomical latitude Φ and longitude Λ, orthometric height H and equipotential surfaces $W = \text{const}$. 10
2.3 Relationship between a sphere with radius $R = a$ and an ellipsoid with semi-major axis a and semi-minor axis b. The flattening of the ellipsoid, $f = \frac{a-b}{a}$ 12
2.4 Geodetic coordinates of an ellipsoid of revolution. 13
2.5 Connection between geodetic and natural coordinates and the deflection of the vertical ε. 13
2.6 Geoid, gravity $\mathbf{g}$, reference ellipsoid and normal gravity $\boldsymbol{\gamma}$ 16
2.7 Geometry of Stokes's and Molodensky's BVP 19
2.8 Signal degree variances c_n computed from the potential coefficients of the satellite-only GGM *GOCO03s*, and the models *Kaula* and *Tscherning/Rapp* 24
2.9 Local covariance function parameters variance C_0, correlation length $C(\psi^0)$ and curvature J 26

3.1 The surface spherical harmonics from left to right: $Y_{5,0}$ (zonal harmonics, $m = 0$), $Y_{5,5}$ (sectorial harmonics, $m = n$) and $Y_{10,5}$ (tesseral harmonics, $n > m,\ m > 0$) 30
3.2 The first few unnormalized Legendre polynomials ($m = 0$), with $t = \cos\theta$ and $\theta = [-\pi, \pi]$ 31
3.3 Geometric quantities concerned in Poisson's integral 39
3.4 Original Stokes's function $S(\psi)$ and modified Stokes's function $F(\psi)$, breaking the singularity at $\psi = 0$ 40
3.5 Graphic display of the least-squares collocation model with a systematic part $\mathbf{Ax}$ and a random part $\mathbf{Bt}$, where $\mathbf{t}$ is a vector containing signal and noise 47
3.6 A regular latitude-longitude grid 50
3.7 A regular latitude-longitude grid with 7° spacing on the unit sphere 51
3.8 A Reuter grid with $\gamma = 30$ 52
3.9 A reuter grid with $\gamma = 30$ on the unit sphere. Note the equidistribution of points. 53
3.10 The Shannon low-pass radial basis function with $N = 250$ in the space and frequency domains. 56
3.11 The Shannon low-pass radial basis function with $N = 250$ in 3D. 57

3.12 The Blackman low-pass radial basis function with $n_1 = 250$ and $n_2 = 450$ in the space and frequency domains. 58
3.13 The Blackman low-pass radial basis function with $n_1 = 250$ and $n_2 = 450$ in 3D. 59
3.14 The Cubic polynomial radial basis function with $n_{\max} = 450$ in the space and frequency domains. 61
3.15 The Cubic polynomial radial basis function with $n_{\max} = 450$ in 3D. 62

4.1 Terrain correction in refined Bouguer reduction 76
4.2 Airy-Heiskanen isostasy model 78
4.3 Helmert's second condensation reduction 79
4.4 Residual terrain model as a difference between the effects of topography and reference topography 80
4.5 Atmospheric layers above and below computation point P 83
4.6 Near (inner) zone σ_0 and distant (far) zone $\sigma - \sigma_0$ in geoid computation . . 87
4.7 Concept of combination in the frequency and spatial domains 89
4.8 Numerical integration on a regular geodetic coordinate grid 97

5.1 Difference between geoid heights at $h = 0\,\text{km}$ and $h = 300\,\text{km}$ 103
5.2 Geoid height signal degree variances at $h = 0\,\text{km}$ and $h = 300\,\text{km}$ 104
5.3 General form of the L-curve. 107
5.4 Synthetic observation data, disturbing potential T_{SHS} at height $h = 0$. . . . 108
5.5 Synthetic observation data, disturbing potential T_{SHS} at height $h = 300\,\text{km}$. 109
5.6 Synthetic observation data, disturbing potential T_{SHS} at height $h = 20000\,\text{km}$. 110
5.7 *Spherical harmonic analysis.* Results from case 1, noise-free observations at $h = 0$. No regularization, $\alpha = 0$. 111
5.8 *Spherical harmonic analysis.* Results from case 2, $\sigma = 20\,\text{m}^2\text{s}^{-2}$ and at $h = 300\,\text{km}$. No regularization, $\alpha = 0$. 112
5.9 *Spherical harmonic analysis.* Results from case 2, $\sigma = 20\,\text{m}^2\text{s}^{-2}$ and at $h = 300\,\text{km}$. Regularization with $\alpha_{\text{large}} = 100$. 114
5.10 *Spherical harmonic analysis.* Results from case 3, $\sigma = 0.1\,\text{m}^2\text{s}^{-2}$ and at $h = 20000\,\text{km}$. No regularization, $\alpha = 0$. 115
5.11 *Spherical harmonic analysis.* Results from case 3, $\sigma = 0.1\,\text{m}^2\text{s}^{-2}$ and at $h = 20000\,\text{km}$. Regularization with $\alpha_{\text{large}} = 1$. 117
5.12 *Spherical harmonic analysis.* L-curve for case 2, corner found by the curvature approach. 118
5.13 *Spherical harmonic analysis.* Results from case 2, $\sigma = 20\,\text{m}^2\text{s}^{-2}$ and at $h = 300\,\text{km}$. Regularization with $\alpha_{\text{L-curve}} = 15.6$. 119
5.14 *Spherical harmonic analysis.* Results from case 2, $\sigma = 20\,\text{m}^2\text{s}^{-2}$ and at $h = 300\,\text{km}$. Regularization with $\alpha = 10$ and $\alpha = 20$. 120
5.15 Signal degree variances c_n computed from the GGM *EGM2008* and the models of *Flury*, *Tscherning/Rapp* and *Kaula* 123
5.16 *East Frisia.* Synthetic gravity anomalies and geoid computed from spherical harmonic synthesis. 126

5.17 *Alpine region.* Synthetic gravity anomalies and geoid computed from spherical harmonic synthesis. . . . 127
5.18 *Stokes integration by 1D-FFT.* Results from geoid computation and comparison, East Frisia. . . . 128
5.19 *Stokes integration by 1D-FFT.* Results from geoid computation and comparison, Alpine region. . . . 129
5.20 *East Frisia* and *Alpine region.* Extended data area with 1° in all directions. 131
5.21 Unmodified and modified cross-covariance functions $C^{N\Delta g}$ and $C^{N^L\Delta g}$ for regional geoid computation. . . . 133
5.22 *East Frisia*, data area. Comparison of unmodified and modified covariance function with Stokes integration in a spherical cap with radius $\psi_0 = 1°$. Note that $N_{LSC}(\psi_0 = 180°)$ corresponds to the original unmodified covariance function. . . . 134
5.23 *Alpine region*, data area. Comparison of unmodified and modified covariance function with Stokes integration in a spherical cap with radius $\psi_0 = 1°$. Note that $N_{LSC}(\psi_0 = 180°)$ corresponds to the original unmodified covariance function. . . . 135
5.24 *Least-squares collocation.* Results from geoid computation and comparison, East Frisia. . . . 136
5.25 *Least-squares collocation.* Results from geoid computation and comparison, Alpine region. . . . 137
5.26 Error of geoid computed from radial basis functions. . . . 139
5.27 Synthetic observation data, disturbing potential T_{SHS} at height $h = 0$, on a 5° latitude-longitude grid (top) and Reuter grid with $\gamma = N_{\max} = 31$ (bottom). The black crosses mark the Reuter grid points. . . . 141
5.28 *RBF analysis.* Results from case 1a, $\sigma = 0$ and at $h = 0$. No regularization. 142
5.29 *RBF analysis.* Results from case 1a, $\sigma = 0$ and at $h = 0$. Regularization with $\alpha_{\mathrm{small}} = 1 \cdot 10^{-8}$. . . . 144
5.30 *RBF analysis.* Results from case 1b, $\sigma = 0$ and at $h = 0$. Regularization with $\alpha_{\mathrm{small}} = 1 \cdot 10^{-8}$. . . . 145
5.31 *RBF analysis.* Results from case 1a, $\sigma = 0$ and at $h = 0$. Regularization with $\alpha_{\mathrm{large}} = 1 \cdot 10^{8}$. . . . 146
5.32 *RBF analysis.* Results from case 2a, $\sigma = 10\,\mathrm{m^2 s^{-2}}$ and at $h = 0$. Regularization with $\alpha_{\mathrm{small}} = 1 \cdot 10^{-8}$. . . . 147
5.33 *RBF analysis.* Results from case 2b, $\sigma = 10\,\mathrm{m^2 s^{-2}}$ and at $h = 0$. Regularization with $\alpha_{\mathrm{small}} = 1 \cdot 10^{-8}$. . . . 148
5.34 *RBF analysis.* Results from case 2b, $\sigma = 10\,\mathrm{m^2 s^{-2}}$ and at $h = 0$. Regularization with $\alpha_{\mathrm{large}} = 100$. . . . 149
5.35 *RBF analysis.* L-curve for case 2b, corner found by the curvature approach. 150
5.36 *RBF analysis.* Results from case 2b, $\sigma = 10\,\mathrm{m^2 s^{-2}}$ and at $h = 0$. Regularization with $\alpha_{\text{L-curve}} = 3.2$. . . . 151
5.37 *RBF analysis.* Results from case 3a, $\sigma = 10\,\mathrm{m^2 s^{-2}}$ and at $h = 300\,\mathrm{km}$. Regularization with $\alpha_{\mathrm{small}} = 1 \cdot 10^{-8}$. . . . 152
5.38 *RBF analysis.* Results from case 3b, $\sigma = 10\,\mathrm{m^2 s^{-2}}$ and at $h = 300\,\mathrm{km}$. Regularization with $\alpha_{\mathrm{small}} = 1 \cdot 10^{-8}$. . . . 153

5.39 *RBF analysis.* Results from case 3b, $\sigma = 10\,\mathrm{m^2 s^{-2}}$ and at $h = 300\,\mathrm{km}$. Regularization with $\alpha_{\mathrm{large}} = 100$. 154
5.40 *RBF analysis.* L-curve for case 3b, corner found by the curvature approach. 155
5.41 *RBF analysis.* Results from case 3b, $\sigma = 10\,\mathrm{m^2 s^{-2}}$ and at $h = 300\,\mathrm{km}$. Regularization with $\alpha_{\text{L-curve}} = 1.48$. 156

List of Tables

2.1 Geodetic reference systems 14
2.2 Boundary conditions for different boundary value problems 17
2.3 Stokes's and simple Molodensky's boundary value problems 18

3.1 Spherical harmonic synthesis of disturbing gravity field functionals 33
3.2 Gravity field functional eigenvalues 34

4.1 Bouguer gravity on the surface of the geoid 75
4.2 Prey reduction scheme by the *remove-restore* principle. 76
4.3 Summary of gravity reduction characteristics 82

5.1 *Spherical harmonic analysis.* Condition numbers for the regularized normal matrix $\mathbf{N} + \alpha\mathbf{K}^{-1}$ for cases 1, 2 and 3. 116
5.2 Omission error for various spherical caps σ_0. E is *EGM2008*, F *Flury*'s model and T/R *Tscherning/Rapp*'s model. 122
5.3 East Frisia, target and data area, $3'$ resolution. Statistics of the synthetic gravity anomalies ($mGal$) and geoid (m). 124
5.4 Alpine region, target and data area, $3'$ resolution. Statistics of the synthetic gravity anomalies ($mGal$) and geoid (m). 125
5.5 *East Frisia.* Statistics of the geoid computed from Stokes's formula and the difference between the geoids (m). 125
5.6 *Alpine region.* Statistics of the geoid computed from Stokes's formula and the difference between the geoids (m). 125
5.7 *East Frisia* and *Alpine region.* Accuracy improvement by extending the data area with 1° in all directions. 130
5.8 *East Frisia* and *Alpine region.* Comparison of unmodified and modified covariance function with Stokes integration in a spherical cap with radius $\psi_0 = 1^\circ$ 132
5.9 *East Frisia.* Statistics of the geoid computed by least-squares collocation and the difference between the geoids (m). 132
5.10 *Alpine region.* Statistics of the geoid computed by least-squares collocation and the difference between the geoids (m). 132
5.11 *Global comparison.* Statistics of the geoid computed by spherical harmonic synthesis, radial basis functions and the difference between the geoids (m). 138
5.12 *RBF analysis.* Condition numbers for the regularized normal matrix $\mathbf{N} + \alpha\mathbf{K}^{-1}$ for cases 1a-b, 2a-b and 3a-b. 140

Chapter 1

Introduction

Geodesy, geomatics and navigation are parts of the *geosciences* and *engineering sciences* (Torge and Müller, 2012). Geodesy may further be divided into *global* geodesy, *geodetic surveys* and *plane surveying*.

Global geodesy concerns the determination of the size and shape of the Earth. This includes the determination of Earth's external gravity field, since it to a large extent shapes the surface of the Earth. Accurate and stable terrestrial and celestial reference systems to relate geodetic measurements to are critical (especially when monitoring temporal changes in the Earth system). Therefore, Earth's orientation in space is also a part of global geodesy.

The relation between global geodesy, geodetic surveys and plane surveying is close, since plane surveys relate to control points established by geodetic surveys, which in turn are linked to reference frames established by global geodesy (Torge and Müller, 2012). Also, applied measurement techniques in global geodesy, geodetic surveying and plane surveying have become more similar. For example, satellite geodesy (e.g., GNSS) is a classical technique in global geodesy, but is now also commonly used in plane surveying.

Rummel et al. (2005) divide into three "pillars of geodesy":

1. *Geometry.* Changes in Earth's position with respect to the celestial reference system as well as Earth's surface geometry and its displacement.
2. *Earth orientation.* Variability of the orientation of Earth's spin axis relative to the stars and changes in Earth's rotation speed.
3. *Gravity field.* Variations in space and time.

Narrowing down, Hofmann-Wellenhof and Moritz (2005) define *physical geodesy* as "the science of determining the figure of the Earth and its gravity field".

The initial awareness that dropped objects fall to the ground and subsequent realization that Earth has a force of attraction, dates way back (Blakely, 1996). Galileo Galilei quantified observations around 1590, and in 1687, Isaac Newton, in his *Principia Mathematica*, proposed that gravitation is a universal property of masses, Earth included.

The gravitational acceleration is a key observable in physical geodesy, and is one of two constituents of observed gravity. The other constituent is the centrifugal force, owing to Earth's rotation. The direction of the gravity vector is the only way to physically define "up" and "down", rendering gravity indispensible for the definition of *heights*.

A common functional of the gravity field is the *geoid*. Mathematician, physicist, astronomer and geodesist Carl Friedrich Gauss (1777-1855) defined the geoid as follows (Torge and Müller, 2012):

"What we call surface of the Earth in the geometrical sense is nothing more than that surface which intersects everywhere the direction of gravity at right angles, and part of which coincides with the surface of the oceans."

Horizontal in the physical sense, the geoid acts as reference surface for the determination of physical heights, e.g., orthometric ones (heights above sea level), and the definition of *vertical reference frames* (height systems).

1.1 Geophysical applications of the geoid

Being a functional of Earth's gravity field, the geoid finds numerous possible applications within geophysics, a few of which will be mentioned in the following.

The static *mean* geoid is commonly of geophysical interest, since it includes both direct and indirect permanent effects of tidal forces (attraction of Sun and Moon), but neglects the time-variable effects, cf. section 3.1.1.1. Such a geoid gives insight into Earth's inner density distribution, since the shape of the geoid is a result of mass anomalies in the Earth interior, cf. Figure 1.1 (*Geodesy for the layman*, 1983).

However, from a mean geoid alone it is impossible to know if a feature is caused by a mass anomaly in Earth's crust, or a considerably larger mass anomaly deeper in the mantle (Wahr, 2009). Geoid signals rapidly varying with time, on the other hand, can usually be attributed to mass changes on Earth surface level. The dedicated satellite gravity mission *Gravity Recovery and Climate Experiment* (GRACE) provides monthly gravity field solutions. Thus, shorter-term cycles in the geoid signal can be analyzed. Such signals most certainly come from a combination of atmosphere and hydrosphere (where the hydrosphere includes oceans, liquid land water masses and frozen land water masses). This is due to the fact that few solid Earth processes vary this rapidly. Consequently, if the tidal forces are considered removed from the signal, solid Earth processes are unlikely to show short-time cycles.

Part of the time-variable gravity field signal are also glacial isostatic adjustment (GIA) effects. For example, when doing mass balance analysis of glaciers, GIA effects must be identified and separated , since they interfere with the signal (Breili, 2013). GIA is the rise of land masses that were depressed by the weight of the continental glaciers during the last glacial period, and comprises geometrical deformations as well as local variations in Earth's gravity field.

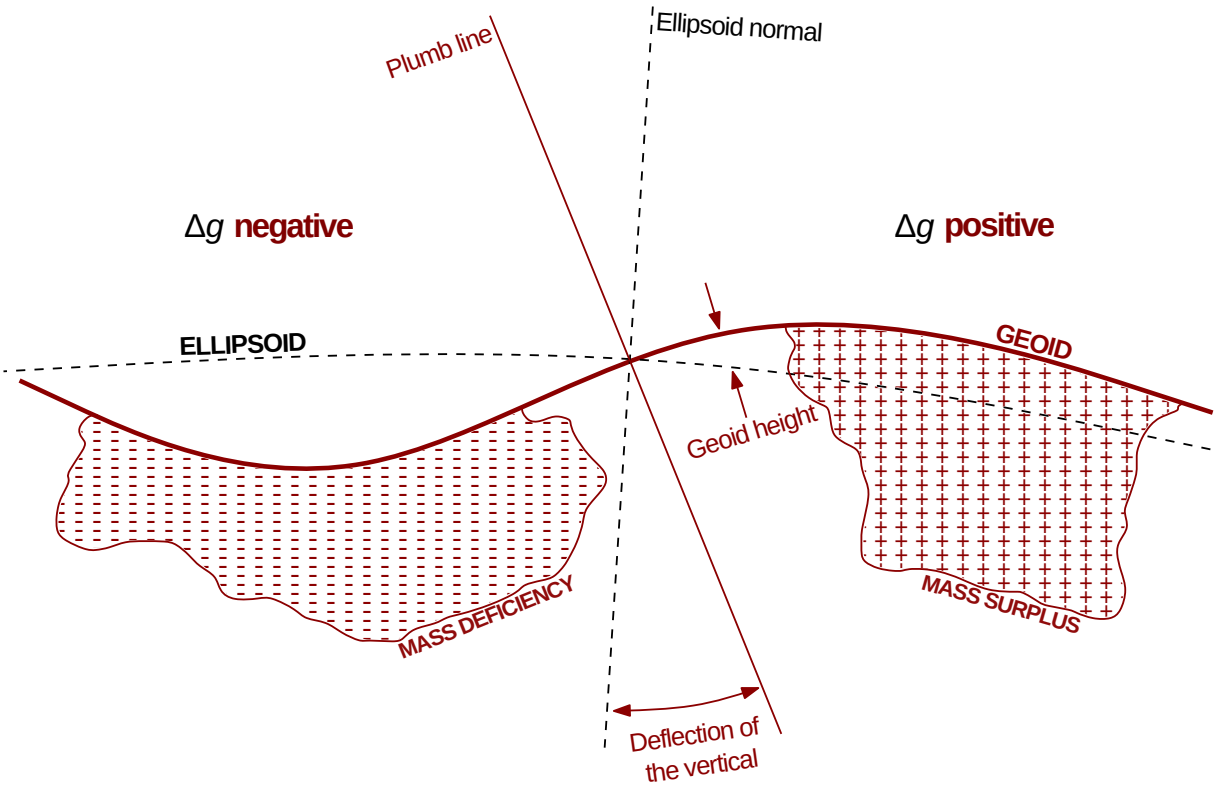

Figure 1.1: *The effect of mass anomalies on the geoid*

A further challenge of analyzing Earth's time-variable gravity field is that such signals are very weak compared with signals due to the relatively immobile major portion of Earth's mass (Wahr, 2009). However, looking at GRACE, geophysical processes causing mass redistribution over 100 km scales are suitable targets for analysis. In particular, the time-variable signals related to water storage on land can contribute to monitoring the accelerated mass loss of the large ice sheets (continental glaciers) of the Antarctic and Greenland, which in turn contributes to sea-level changes.

The mean geoid coincides with "undisturbed" sea level, thus being of particular interest for oceanography (Torge and Müller, 2012). The ocean surface does not coincide with the mean geoid, and the deviations are called sea surface topography (SST). Thus, in order to study SST, an accurate geoid is needed.

As it happens, each of the three pillars of geodesy are important when monitoring sea level (Church et al., 2010). Geometry plays an important role when observing sea-level change. Further, melting polar ice sheets affect Earth's rotation and orientation. In static equilibrium, the sea surface coincides with the geoid, and consequently, mass redistribution due to sea-level change also changes the shape of the geoid.

1.2 Thesis objective

Several methods for modeling the gravity field, or the geoid in particular, exist, based on different (homogeneous or heterogeneous) gravity field quantities. In addition, gravity observations are done on land (terrestrial), on sea and in the air as well as by satellite techniques. Some gravity field modeling approaches are well-suited for global data, other fit regional or local data better. Every method has benefits and downsides, so to compute the best possible gravity field, ideal method combinations are sought.

For example, terrestrial gravity measurements contain the complete signal spectrum, but are in the form of point values, and not globally available. Dedicated satellite gravity missions are well-suited for determining the global geoid, and model the long wavelength part of the signal spectrum particularly well. A combination of these observations is therefore of interest when computing a local or regional geoid.

Among the classical methods for geoid computation are least-squares collocation and Stokes integration with various kernel modifications. Currently, there is an increased interest in radial basis functions for more efficient gravity field modeling.

This thesis investigates and compares classical and new methods for geoid computation, theoretically and numerically with synthetic data, with focus on local to regional applications.

Inverse problems, such as gravity field analysis, are often ill-conditioned. This is notably the case with radial basis function analysis by least-squares adjustment. Therefore, a further objective of this thesis is to look into regularization techniques for solving

ill-conditioned problems, and apply regularization in cases of both spherical harmonic analysis and radial basis function analysis.

1.3 Thesis summary

- This first chapter is an introduction to geodesy and physical geodesy in particular, mentioning only a few of the manifold geophysical applications of the geoid.
- Chapter 2 covers the mathematical fundamentals of Earth's gravity field and goes through the two classical approaches to solve the geodetic boundary problem and their important conceptual differences. Also, an introduction to applied statistics in physical geodesy is given.
- Chapter 3 formulates the different gravity field functional models and compares them theoretically. Treated models are the classical integral formulae, least-squares collocation and radial basis functions.
- Chapter 4 covers important reductions and corrections, which must be considered when a geoid is to be determined in practice. The chapter also comprises different practical approaches to the actual numerical evaluation of the gravity field models.
- Chapter 5 illuminates the theory of inverse and ill-posed problems, as geodetic problems often have this character — as well as practical numerical solutions to such problems, accompanied by examples. This chapter also includes numerical comparisons of the different gravity field modeling approaches with synthetic data, as a supplement to the theoretical comparisons in chapter 3. Results are commented and discussed.
- Chapter 6 closes the thesis with some final summarizing remarks with regards to the results in chapter 5 as well as some comments on recommendations for future work.

Chapter 2

Fundamentals of Earth's gravity field

This chapter is an elaboration of mathematical terms and concepts of Earth's gravity field leading to the geodetic boundary value problem for geoid determination.

2.1 Gravitational potential and gravity

Newton's law of universal gravitation can be written as follows:

$$F = G\frac{mm_0}{r^2} \tag{2.1}$$

where F is a mutual force between a particle of mass m and a particle of mass m_0, separated by a distance r (Blakely, 1996). G is Newton's gravitational constant, $G = 6.6742 \cdot 10^{-11}\,\mathrm{m^3 kg^{-1} s^{-2}}$.

Considering a unit mass m_0 at point P, then dividing equation (2.1) with m_0 yields the gravitational attraction exerted by m at P, a distance r from m:

$$F = G\frac{m}{r^2} \tag{2.2}$$

The gravitational attraction of the Earth is a vector force field, which may be found from the scalar field:

$$V = G\frac{m}{r} \tag{2.3}$$

where V is a *potential* of F, known as gravitational potential, *the work done by the field on the unit mass* (Blakely, 1996).

The components of the gravitational field can be written as (in cartesian coordinates):

$$\mathbf{F} = \left[\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}\right] = grad(V) = \nabla V \tag{2.4}$$

The gravitational field is *conservative*, since the energy at hand in the field is conserved, i.e., no energy is lost or gained when a body is moving from a point P in the field and back to P again.

If we consider point masses continuously distributed over a volume v with density $\rho = \frac{\mathrm{d}m}{\mathrm{d}v}$, where $\mathrm{d}m$ is a mass element and $\mathrm{d}v$ a volume element, then the potential may be written as an integral:

$$V = G \iiint_v \frac{\mathrm{d}m}{r} = G \iiint_v \frac{\rho}{r} \mathrm{d}v \tag{2.5}$$

where r is the distance between the mass element $\mathrm{d}m = \rho \cdot \mathrm{d}v$ and the attracted point P.

The gravitational potential satisfies *Poisson's equation*:

$$\Delta V = \left[\frac{\partial^2 V}{\partial x^2}, \frac{\partial^2 V}{\partial y^2}, \frac{\partial^2 V}{\partial z^2} \right] = -4\pi G\rho \tag{2.6}$$

where the Laplacian Δ of a scalar field is the divergence of the field's gradient, i.e., $\Delta V = \nabla^2 V = div(\nabla V) = div(\mathbf{F})$.

In empty space, $\rho = 0$ and the gravitational potential obeys a simpler differential equation:

$$\Delta V = 0 \tag{2.7}$$

This was shown by French mathematician and astronomer Pierre-Simon Laplace in 1782, and therefore equation (2.7) is called *Laplace's equation* (Blakely, 1996).

Solutions of Laplace's equation are *harmonic* functions, where the simplest example of such a function is the reciprocal distance $1/r$. If the gravitational potential as given in equation (2.5) is inserted into Laplace's equation we get:

$$\Delta V = G\Delta \left[\iiint_v \frac{\rho}{r} \mathrm{d}v \right] = G \iiint_v \rho \Delta \left(\frac{1}{r} \right) \mathrm{d}v = 0 \tag{2.8}$$

The *centrifugal force* is a result of Earth's rotation, and acts on all bodies of mass which rotate with the Earth. The centrifugal force $\mathbf{f}$ may also be found from a potential:

$$\mathbf{f} = \nabla\Phi = \left[\frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y}, \frac{\partial \Phi}{\partial z} \right] \tag{2.9}$$

where, expressed in terms of a constant angular velocity ω about Earth's spin axis:

$$\Phi = \frac{1}{2}\omega^2(x^2 + y^2) \tag{2.10}$$

Gravity is the force composed of both the gravitational force and the centrifugal force, and consequently, the *gravity potential* may be written as:

$$W = V + \Phi = G \iiint_v \frac{\rho}{r} \mathrm{d}v + \frac{1}{2}\omega^2(x^2 + y^2) \tag{2.11}$$

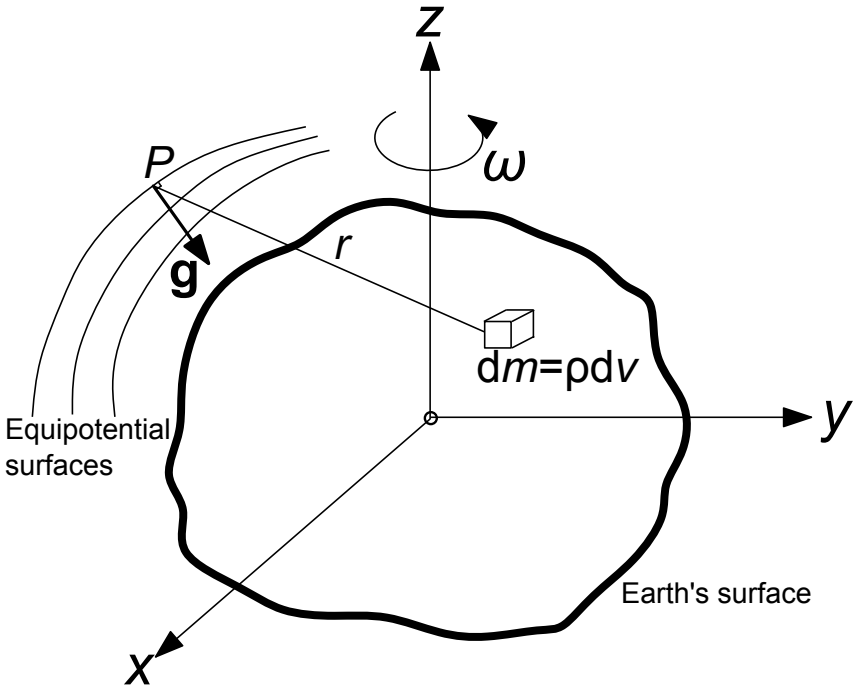

Figure 2.1: *Potential, equipotential surfaces and gravity*

The gradient of W is the *gravity acceleration* (see Figure 2.1):

$$\mathbf{g} = \nabla W = \left[\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}, \frac{\partial W}{\partial z}\right] \tag{2.12}$$

The magnitude g, is generally known as gravity, and measured in Gal, where $1\,\mathrm{Gal} = 0.01\,\mathrm{ms}^{-2}$.

A surface $W(x, y, z) = \text{const.}$ on which the gravity potential remains constant, is termed an *equipotential*, *level* or *geopotential* surface. This property makes an equipotential surface a good vertical reference surface, i.e., a reference for *heights*. The equipotential surface which corresponds to the *mean ocean surface level*, proposed by Gauss as the "mathematical figure of Earth", is what has become known as the *geoid* (Hofmann-Wellenhof and Moritz, 2005). It is defined by:

$$W = W_0 = \text{const.} \tag{2.13}$$

The gravity vector as expressed in equation (2.12) at point P is normal to the equipotential surface through the same point, cf. Figure 2.1. The *lines of force*, *plumb lines* or *verticals* that intersect every equipotential surface at right angles, are slightly curved. The gravity vector is tangent to the plumb line at the same point. The *orthometric height*, or height above sea level, H, is measured along the plumb line, starting on the geoid (Hofmann-Wellenhof and Moritz, 2005):

$$\mathrm{d}W = -g\,\mathrm{d}H \tag{2.14}$$

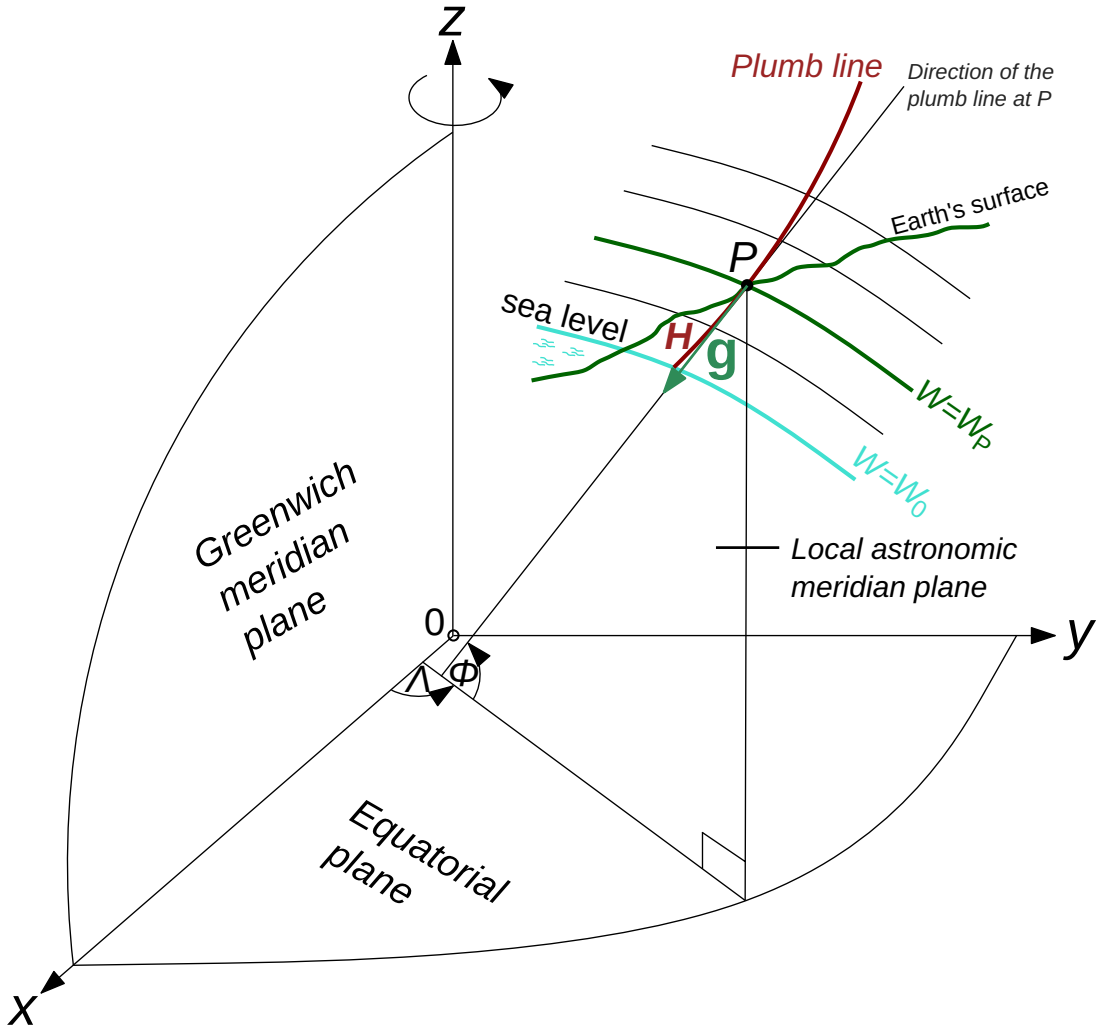

Figure 2.2: *Astronomical latitude* Φ *and longitude* Λ*, orthometric height* H *and equipotential surfaces* $W = \text{const}$.

In equation (2.14), a clear relation between height H and potential W is observed.

"Natural coordinates" Φ, Λ, W may be defined in the gravity field. Astronomical latitude Φ and longitude Λ describe the direction of the vertical in a point P and is observed through astronomical methods, cf. Figure 2.2. The gravity potential W defines P in a system of equipotential surfaces $W = \text{const}$. Thus, the gravity acceleration may also be written as follows:

$$\mathbf{g} = \nabla W = -g \begin{pmatrix} \cos\Phi\cos\Lambda \\ \cos\Phi\sin\Lambda \\ \sin\Phi \end{pmatrix} \tag{2.15}$$

Earth's density distribution, i.e., ρ, is not well known, and therefore equation (2.5) cannot be evaluated directly. If we consider the Earth's exterior (empty space), however,

Laplace's equation, cf. equation (2.8), may be solved and V expanded into a convergent *spherical harmonic* series (in spherical coordinates) (Torge and Müller, 2012):

$$V(r,\theta,\lambda) = \frac{GM}{R}\sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+1}\sum_{m=0}^{n}\bar{P}_{nm}(\cos\theta)\left[\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda\right] \tag{2.16}$$

where $\bar{P}_{nm}(\cos\theta)$ are the *fully normalized associated Legendre functions* of degree n and order m and $\bar{C}_{nm}$ and $\bar{S}_{nm}$ the *fully normalized and dimensionless potential coefficients*, containing the amplitudes of different spectral parts. GM is the product of Newton's gravitational constant and Earth's mass, R a mean Earth radius and r the distance from Earth's center of mass to the computation point P. θ is the colatitude, $\theta = 90° - \varphi$ (Hofmann-Wellenhof and Moritz, 2005).

Equation (2.16) is known as a *spectral decomposition* of the gravity field, corresponding to a *Fourier* expansion of $V(r,\theta,\lambda)$ with basis functions $\bar{P}_{nm}(\cos\theta)\cos m\lambda$ and $\bar{P}_{nm}(\cos\theta)\sin m\lambda$. Determination of $V(r,\theta,\lambda)$ from the amplitudes of the gravity field signal, i.e., the coefficients $\bar{C}_{nm}$ and $\bar{S}_{nm}$ is known as *spherical harmonic synthesis*. Theoretically, $V(r,\theta,\lambda)$ is a continuous function, but the number of potential coefficients is finite, and so the series in equation (2.16) is cropped at some maximal degree $N_{\max}$ (Gerlach, 2003).

The potential coefficient $\bar{C}_{20} = -J_2$ is termed the *dynamical form factor* of the Earth and describes its polar flattening. It is the largest deviation from a spherical Earth, three orders of magnitude larger than the values of the successive coefficients. Furthermore, it follows that a similar spherical harmonic expansion of W is obtained by adding the centrifugal potential Φ to equation (2.16) (Torge and Müller, 2012).

As a consequence of the above, the determination of Earth's gravity field is the main problem of physical geodesy. It will ultimately become apparent that in order to determine the gravity potential function, a *boundary value problem* has to be formulated and solved, i.e., to determine a harmonic function *outside* a surface by the function's value *on* the surface (Gerlach, 2003).

However, we need to simplify Earth's gravity field so that it becomes mathematically feasible.

2.2 The geodetic Earth model and normal gravity

The Earth is slightly flattened at the poles, yielding a discrepancy between polar and equatorial radius of approximately 21 km. An even better approximation of the Earth than a sphere would therefore be an *ellipsoid*, cf. Figure 2.3.

A point on the surface of the ellipsoid may be described by the *geodetic* coordinates φ and λ, cf. Figure 2.4. These coordinates may be extended to a spatial coordinate system by introducing a height h above the ellipsoid, and so we have ellipsoidal geodetic coordinates

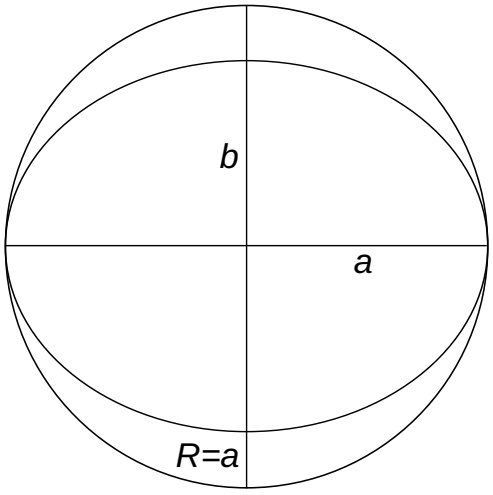

Figure 2.3: *Relationship between a sphere with radius $R = a$ and an ellipsoid with semi-major axis a and semi-minor axis b. The flattening of the ellipsoid, $f = \frac{a-b}{a}$*

(φ, λ, h) in contrast to the already introduced "natural coordinates" (Φ, Λ, W) (Torge and Müller, 2012).

The connection between geodetic and natural coordinates, the first independent of Earth's gravity field and the latter dependent on it, is shown in Figure 2.5. The difference between the direction of the vertical (plumb line) and the direction of the ellipsoid normal is the *deflection of the vertical*, ε.

The geoid has been introduced as an equipotential surface of the actual gravity field. An approximation of the actual gravity field would be to let the ellipsoid have the same mass and rotational velocity as the actual Earth, its surface being an equipotential surface of a *normal* (or *reference*) gravity field. Such an Earth model is also known as a geodetic *reference system*, cf. Table 2.1.

Considering both constituents of the gravity potential W (gravitational and centrifugal potential), the normal potential U may be defined by four parameters: the product of Newton's gravitational constant and the mass of the ellipsoid GM, angular velocity ω and the geometrical parameters a and f. In this way, the external gravity field of the normal ellipsoid can be determined unambiguously, and closed formulae may be derived (Torge and Müller, 2012).

The spherical harmonic expansion of the normal gravity field may be expressed as:

$$U(r,\theta) = \frac{GM}{R}\sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+1}\bar{C}_{n0}\bar{P}_{n0}(\cos\theta) + \frac{1}{2}\omega^2 r^2(\sin\theta)^2 \tag{2.17}$$

From equation (2.17), several important observations are made, due to our choice of mathematical surface. The normal gravity field is rotationally symmetric, and is therefore independent of λ (all longitudes are equivalent). Consequently, all terms with $m > 0$

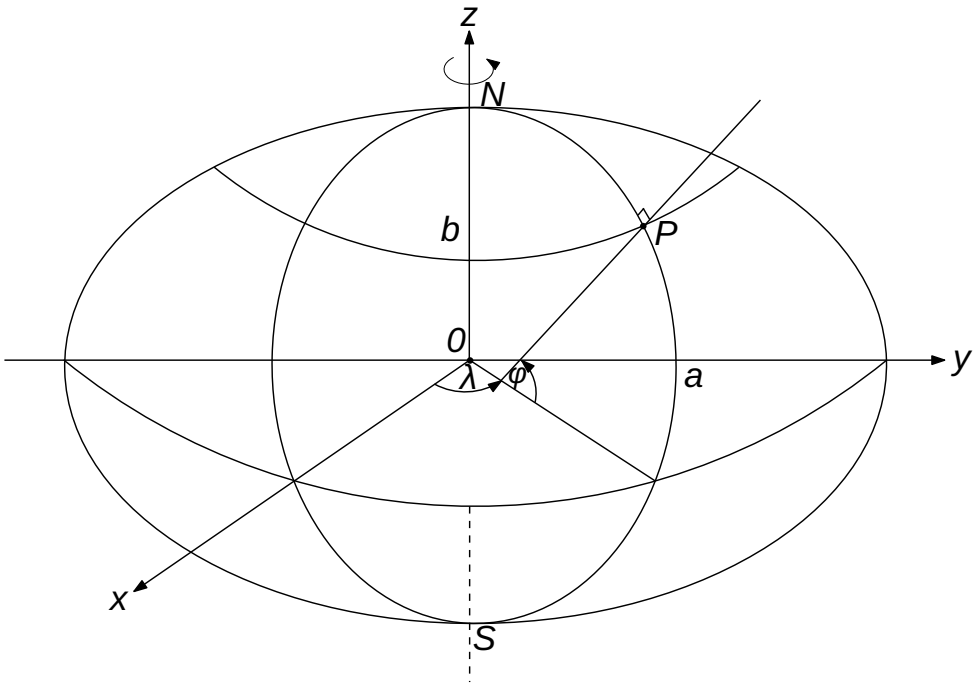

Figure 2.4: *Geodetic coordinates of an ellipsoid of revolution.*

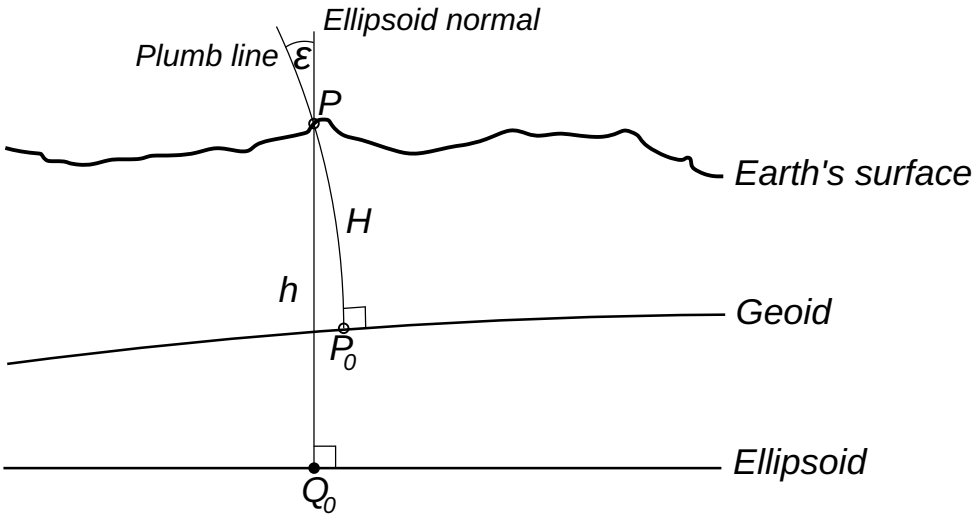

Figure 2.5: *Connection between geodetic and natural coordinates and the deflection of the vertical ε.*

vanish and only *zonal* terms ($m = 0$) remain. Furthermore, the normal gravity field has equatorial symmetry, so only *even degree* zonal terms, i.e., $n = 0, 2, 4, 6, \ldots$, exist (Gerlach, 2009).

As mentioned earlier, the main contribution to the slightly flattened, ellipsoidal form of the Earth comes from the zonal dynamic form factor coefficient J_2. In light of this, the flattening parameter (or "geometric form factor") f of a geodetic reference system is often interchanged with the dynamic form factor. Some contribution comes also from the higher degree coefficients, but because of the smoothness of the normal gravity field the series expansion in equation (2.17) may be truncated at a low degree, i.e., $n > 10$ degree coefficients may be neglected (Gerlach, 2003).

	GM $(10^9\,\mathrm{m^3\,s^{-2}})$	a (m)	f	J_2 (10^{-6})	ω $(10^{-5}\,\mathrm{rad^3\,s^{-1}})$
GRS67	398603	6378160	1/298.247		7.2921151467
WGS84	398600.4418	6378137	1/298.257223563	1081.874	7.292115167
GRS80	398600.5	6378137	1/298.257222101	1082.63	7.292115

Table 2.1: *Geodetic reference systems*

In accordance with equation (2.12), the normal gravity acceleration is formulated as follows:

$$\boldsymbol{\gamma} = \nabla U = \left[\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z}\right] \tag{2.18}$$

Since the pole is closer to the Earth's center of mass, γ_b at the poles will be larger than γ_a at the equator. A *gravity flattening* may be expressed as:

$$f^\star = \frac{\gamma_b - \gamma_a}{\gamma_a} \tag{2.19}$$

Formulae for normal gravity at the poles and at the equator may be found in, e.g., (Hofmann-Wellenhof and Moritz, 2005).

A symmetrical expression for the normal gravity *on* the ellipsoid may be derived, known as *Somigliana's* formula:

$$\gamma_0 = \frac{a\gamma_a(\cos\varphi)^2 + b\gamma_b(\sin\varphi)^2}{\sqrt{a^2(\cos\varphi)^2 + b^2(\sin\varphi)^2}} \tag{2.20}$$

Furthermore, normal gravity at a relatively small ellipsoidal height h may be expressed as:

$$\gamma_h = \gamma_0\left[1 - \frac{2}{a}\left(1 + f + \frac{\omega^2 a}{\gamma_a} - 2f(\sin\varphi)^2\right)h + \frac{3}{a^2}h^2\right] \tag{2.21}$$

An approximation to *Clairaut's* theorem reads:

$$f + f^\star = \frac{5}{2}\frac{\omega^2 a}{\gamma_a} \tag{2.22}$$

Equation (2.22) is interesting, since it shows that the *geometrical* flattening f may be found from the *gravity* flattening $f^\star$. In other words, the Earth's flattening may be found through gravimetry (Hofmann-Wellenhof and Moritz, 2005).

Although now having described a mathematical model of Earth's gravity field which is easier to compute, we still have the general challenge of non-linearity when wanting to determine the gravity field W from observations, in addition to ΔW not being zero because of the centrifugal force. If, however, the remaining deviation between the actual gravity field and the normal gravity field of a level ellipsoid is considered, it will become evident that it is so small that it may be considered *linear* (Hofmann-Wellenhof and Moritz, 2005).

2.3 The linear gravity field

The *anomalous*, *perturbing* or *disturbing* potential is the small difference between the actual geopotential (geoid $W = W_0$) and the normal gravity potential (reference ellipsoid $U = U_0 = W_0$ of same potential), i.e.:

$$T(x, y, z) = W(x, y, z) - U(x, y, z) \tag{2.23}$$

Since the centrifugal part of W is well known we may assume that it is equal to the centrifugal part of U, and the disturbing potential remains a difference in gravitation between Earth and ellipsoid. It consequently is a harmonic function outside the masses, and there it fulfills Laplace's equation, $\Delta T = 0$ (Torge and Müller, 2012).

Subtracting equation (2.17) from W (= (2.16) + Φ), the spherical harmonic expansion of T is obtained:

$$T(r, \theta, \lambda) = \frac{GM}{R} \sum_{n=2}^{\infty} \left(\frac{R}{r}\right)^{n+1} \sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta) \left[\Delta\bar{C}_{nm} \cos m\lambda + \Delta\bar{S}_{nm} \sin m\lambda\right] \tag{2.24}$$

where

$$\Delta\bar{C}_{nm} = \begin{cases} \bar{C}_{nm} - \bar{C}_{nm}^{\text{normal potential}} & m = 0 \ \wedge \ n \in \{2, 4, 6, 8, \ldots, N_{\max}\} \\ \bar{C}_{nm} & \text{else.} \end{cases} \tag{2.25}$$

$$\Delta\bar{S}_{nm} = \bar{S}_{nm} \tag{2.26}$$

The summation in equation (2.24) begins at $n = 2$. The zero-degree term $\left(T_0 = \frac{\delta GM}{r}\right)$ vanishes if the mass M and equatorial radius R of the ellipsoid is equal to the mass and equatorial radius of Earth. This is often assumed, and when defining a geodetic reference system one tries to keep this difference small. If the center of Earth's masses coincides with the origin of the ellipsoid, the first-degree terms vanish (Gerlach, 2003).

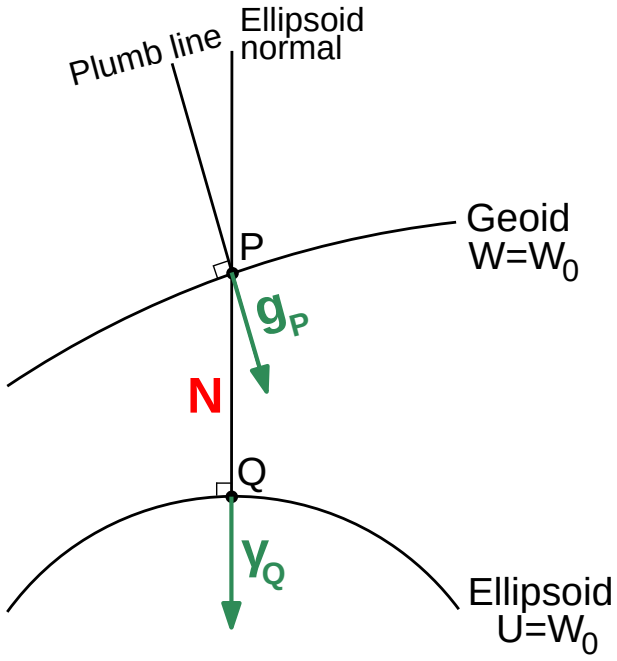

Figure 2.6: *Geoid, gravity* $\mathbf{g}$*, reference ellipsoid and normal gravity* $\boldsymbol{\gamma}$

If the gravity vector $\mathbf{g}$ at point P is considered, as is the normal gravity vector $\boldsymbol{\gamma}$ at point Q, their difference in magnitude is the *gravity anomaly*, cf. Figure 2.6:

$$\Delta g = g_P - \gamma_Q \tag{2.27}$$

Their difference in direction is the already mentioned *deflection of the vertical* with a north-south component ξ and an east-west component η, i.e.:

$$\xi = \Phi - \varphi \tag{2.28}$$

$$\eta = (\Lambda - \lambda)\cos\varphi \tag{2.29}$$

If the vectors $\mathbf{g}$ and $\boldsymbol{\gamma}$ are considered at the same point P, their difference in magnitude is the *gravity disturbance*, $\boldsymbol{\gamma}_P$ at some ellipsoidal height h may be determined by global navigation satellite systems (GNSS):

$$\delta g = g_P - \gamma_P \tag{2.30}$$

The deflection of the vertical is virtually the same in the gravity disturbance case (Hofmann-Wellenhof and Moritz, 2005)

Also apparent in Figure 2.6 is the distance from the reference ellipsoid to the geoid, N, which is the *geoid height* or *geoid undulation*. *Bruns's formula* relates the geoid height to the disturbing potential:

$$N = \frac{T}{\gamma} \tag{2.31}$$

In other words, equation (2.31) gives the *distance* with which the disturbing potential T displaces the surface of the geoid (Boener, 1994). In accordance with Figure 2.5, the following connection between geoid height N, orthometric height H and ellipsoidal height h may be formulated, neglecting the curvature of the plumb line (Torge and Müller, 2012):

$$H = h - N \tag{2.32}$$

GNSS observations yield ellipsoidal height, and in combination with a geoid (i.e., N) heights above sea level may be derived.

It is further assumed that the deflection of the vertical is zero. In addition, a spherical approximation of the linear quantities T, N, Δg and δg is made, justified by the fact that they originate from the small deviation between an ellipsoid and a sphere in the first place and that the flattening is relatively small (0.3 %). The chosen radius for this sphere is defined by equating volume of the reference ellipsoid with the volume of the sphere, e.g., $R_{\text{WGS84}} = \sqrt[3]{a^2 b} = 6371\,\text{km}$. A spherically approximated equation relating measurable gravity to the gravity potential is obtained:

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{R} T \tag{2.33}$$

Equation (2.33) is known as *the fundamental equation of physical geodesy*, and has the form of a differential equation. Δg, however, is not measurable in space, but on the surface of the Earth. Δg may be computed on the geoid after removing or shifting topographic masses inside the Earth interior, cf. section 4.1. If we assume that there are no masses *outside* the geoid, the disturbing potential T fulfills Laplace's partial differential equation ($\Delta T = 0$), and equation (2.33) may be applied as a *boundary condition* for solving it.

2.4 Boundary value problems

In potential theory, when solving an *exterior* boundary value problem (BVP), one tries to determine a function in outer space from the function's value *on a bounding surface* (the boundary condition) and its *behavior in space* (a partial differential equation).

In the beginning of this chapter, the problem of determining Earth's gravity field without knowing the density ρ was observed. If we want to solve the BVP for Earth's *interior*, *Poisson's* equation (2.6) is applied, and density knowledge is assumed. The *exterior* BVP is solved with *Laplace's* equation (2.7), and three different boundary conditions have classically been used, cf. Table 2.2.

Dirichlet BVP	Neumann BVP	Third BVP
V	$\frac{\partial V}{\partial n}$	$V + \frac{\partial V}{\partial n}$

Table 2.2: *Boundary conditions for different boundary value problems*

When the boundary functions are geopotential values, it is a Dirichlet BVP, if they are gravity disturbances it is a Neumann BVP. The fundamental equation of physical geodesy (2.33) is immediately recognized as a boundary condition for the third BVP, i.e., it is a linear combination of T and the partial derivative of T with respect to a surface normal. It has historically been of substantial importance, since it relates observable gravity to the potential, and is therefore termed the *geodetic boundary value problem.*

Stokes's and Molodensky's geodetic boundary value problems

As mentioned earlier, to determine Earth's surface and external gravity field from observations on (or close to) the surface of the Earth, the geodetic boundary value problem must be solved. The classical *Stokes* approach (published by G. G. Stokes in 1849) uses the geoid as surface, while the *Molodensky* approach (published by M. S. Molodensky in 1945) uses the physical surface of the Earth, cf. Table 2.3.

	Field equation	Boundary condition
Stokes's theory	$\Delta T = 0, \quad R^3 - \Omega - \text{geoid}$	$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{r}T\Big\|_{\Omega}$
Molodensky's theory	$\Delta T = 0, \quad R^3 - \Sigma - \text{telluroid}$	$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{r}T\Big\|_{\Sigma}$

Table 2.3: *Stokes's and simple Molodensky's boundary value problems*

The main conceptual difference between the Stokes approach and the Molodensky approach is the choice of boundary surface. If the geoid serves as boundary surface, we will have masses *outside* the geoid (i.e., topography, mass of atmosphere and sea surface topography), and not empty space at all ($\rho \neq 0$!), which is the crucial criterion for fulfilling Laplace's equation, cf. section 3.2.1.

It is then quite tempting to use the physical Earth surface as boundary instead of the geoid, so that we have no topography on the outside of the surface (albeit the atmospheric mass). Such a choice of surface also makes sense since all gravity observations are made on the surface of the Earth. A critical consequence of such a choice is that Earth's surface is no equipotential surface of the gravity field, and the gravity vector does not correspond to the surface normal. Molodensky's problem therefore involves *oblique* derivatives of T (Gerlach, 2009).

Molodensky's problem as it is shown in Table 2.3 is thus a simplified one.

Earth's surface is approximated by the *telluroid* Σ in Molodensky's world. The gravity potential W at point P on Earth's surface S is approximated by the normal potential U at point Q on Σ, cf. Figure 2.7. Consequently, the gravity anomaly Δg becomes the difference between gravity at Earth's surface and normal gravity at a point Q on the

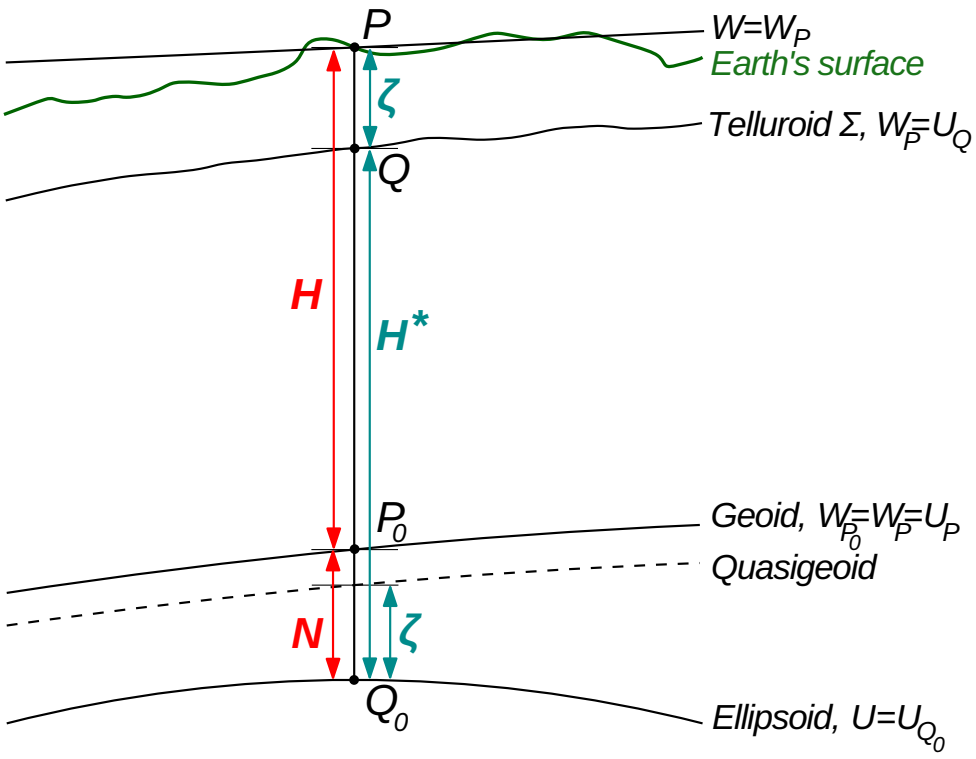

Figure 2.7: *Geometry of Stokes's and Molodensky's BVP*

telluroid. The separation between points P and Q, i.e., between Σ and S, is known as the *height anomaly* ζ. When ζ is plotted above the reference ellipsoid, the *quasigeoid* is obtained. The quasigeoid differs from the geoid by maximum 1 m in mountainous areas and practically coincides with the geoid over the oceans, but is not an equipotential surface of the gravity field (Torge and Müller, 2012).

Analogous to the definition of *orthometric* heights (or height above sea level) H in equation (2.14), *normal heights* $H^\star$ are the distance from Earth's surface along the ellipsoidal normal, starting on the quasigeoid, i.e.:

$$\mathrm{d}W = -\gamma \, \mathrm{d}H^\star \tag{2.34}$$

Moreover, in accordance with equation (2.32), the following holds:

$$H^\star = h - \zeta \tag{2.35}$$

A further clarification of the gravity anomaly in correspondence with Figure 2.7 is considered appropriate. In the case of Stokes it is assumed that gravity is measured *on the geoid*, and we have to reduce the surface value down to the geoid by an approximated so-called *free-air gradient* of normal gravity $\partial\gamma/\partial h \approx 0.3086\,\mathrm{mGal}$ (treated in section 4.1):

$$\Delta g = \left(g_P - \frac{\partial g}{\partial h} H\right) - \gamma_{Q_0} \approx \overbrace{\left(g_P - \frac{\partial \gamma}{\partial h} H\right)}^{g_{P_0}} - \gamma_{Q_0} \tag{2.36}$$

In the case of Molodensky, it is assumed that gravity is measured *on Earth's surface*, and γ_Q is computed from γ_{Q_0} on the ellipsoid:

$$\Delta g = g_P - \overbrace{\left(\gamma_{Q_0} + \frac{\partial \gamma}{\partial h} H^\star\right)}^{\gamma_Q} \tag{2.37}$$

In other words, in case of Stokes, gravity g_P is lowered to the geoid g_{P_0}, in case of Molodensky, normal gravity is upward continued to the telluroid at normal height $H^\star$ above the ellipsoid. As gravity g_P refers to Earth's surface and normal gravity γ_Q refers to the telluroid, the (free-air) gravity anomaly in equation (2.37) refers to the ground.

Bruns's formula expressed in Stokes's theory in equation (2.31) also holds in the Molodensky case, the only difference being the *location* of the computation point (Gerlach, 2009)!:

$$\zeta = \frac{T}{\gamma_Q} \tag{2.38}$$

The quasigeoid is quite similar to the geoid, but the quasigeoid is rougher and not an equipotential surface. To interchange between geoid and quasigeoid, the *separation* between them is of interest. It can be shown that the following holds:

$$N - \zeta = \frac{\bar{g} - \bar{\gamma}}{\bar{\gamma}} H \tag{2.39}$$

where $\bar{\gamma}$ is the mean value of normal gravity along the ellipsoid normal, and $\bar{g}$ a mean value of gravity along the ellipsoid normal (obtained through a *Prey* reduction scheme, cf. section 4.1). H is the orthometric height. $\bar{g} - \bar{\gamma} = \Delta g_B$, or the *Bouguer anomaly* (approximate, because of $\bar{g}$'s density hypothesis). The Bouguer anomaly is a relatively small number, so we may substitute $\bar{\gamma}$ with normal gravity at the surface of the reference ellipsoid, γ_0:

$$N - \zeta \approx \frac{\Delta g_B}{\gamma_0} H \tag{2.40}$$

2.5 Statistics of the gravity field

A statistical description of Earth's gravity field is very useful, not only since the introduction of statistics gives an idea of the *accuracy* or *precision* (or general *quality*) of our measurements and applied methods, but also since it allows the use of *statistical approximation methods* such as least-squares prediction, where assumptions regarding the statistical behavior of our measurements are made in order to predict values at points where no measured value exists. This section relies partly on a section on statistical properties of the gravity field found in (Ophaug, 2013).

Considered is a *stochastic* signal, whose values are randomly spread around the average (or mean) value, which is zero. This means that we are not in position to make any assumptions regarding the signal value at a specific point. Moreover, the statistical properties of a stochastic process should be derived from an infinite number of process realizations. Since only *one* realization of the gravity field is available, the hypothesis of *ergodicity* is required. It states that statistical quantities may also be calculated from average values over just one gravity field realization (Torge and Müller, 2012).

As fundamental gravity field parameter, the gravity anomaly has been chosen for its high-resolution availability over both continents and oceans. The average of the gravity anomaly signal, formed over the whole Earth, is zero, i.e.:

$$\mathcal{M}\{\Delta g\} = \frac{1}{4\pi} \iint_{\sigma} \Delta g \mathrm{d}\sigma = 0 \tag{2.41}$$

where $\mathcal{M}\{\}$ is the average over the whole Earth, corresponding to the integral over the unit sphere σ, divided by its area. However, the integral is zero only if there is no zero-degree term of the spherical harmonic expansion of Δg, cf. section 2.3 (Torge and Müller, 2012).

The global gravity anomaly *signal variance*, is *not* zero, but formulated similar to equation (2.41):

$$\mathcal{M}\{\Delta g^2\} = \frac{1}{4\pi} \iint_{\sigma} \Delta g^2 \mathrm{d}\sigma \tag{2.42}$$

The square root of equation (2.42) yields the *root mean square anomaly*, usually in the range $\mathrm{RMS}\{\Delta g\} = \pm 35\,\mathrm{mGal}$ (Hofmann-Wellenhof and Moritz, 2005).

The statistical quantity termed *covariance* C describes the correlation between two point signal values at different positions in the domain, i.e., it is the average product $\Delta g \Delta g'$ at points P and P' separated by a constant distance s. It thus reflects the average signal behavior, or the tendency of Δg and $\Delta g'$ to have the same size and sign. The covariance function may depend only on the distance s between the two points if the conditions of *homogeneity* (the covariance is independent of the *absolute position* of the point pair in the domain) and *isotropy* (the covariance is independent of the *connection direction* of the point pair) are met (Gerlach, 2003):

$$C(s) = \mathcal{M}\{\Delta g \Delta g'\} \tag{2.43}$$

If the points are situated far apart from each other, their gravity anomaly values are considered uncorrelated, since the gravity anomaly has a local character in general. If the gravity anomalies are uncorrelated (i.e., independent of each other), their covariance is zero. If the distance $s = 0$ we obtain:

$$C(s=0) = C_0 = \mathcal{M}\{\Delta g^2\} \tag{2.44}$$

which is the variance, cf. equation (2.42). In equation (2.44) we observe that as the distance s approaches zero, the gravity anomalies become increasingly equal, and the correlation is strong. As the distance increases, correlation decreases and the gravity anomalies become increasingly independent of each other.

The determination of an exact global covariance function is not an easy task since, in theory, gravity anomalies at all points on Earth's surface must be known. From section 3.1 is clear that the gravity anomaly may be expressed as a spherical harmonic series. The distance s is then converted into the *spherical* distance ψ i.e., $\psi = s/R$ where R is the mean Earth radius. If the spherical harmonic expansion of the gravity anomaly is inserted into equation (2.42) and the orthogonality relations are considered, the following expression holds:

$$C_0^{\Delta g} = \sum_{n=2}^{\infty} \left(\lambda_n^{\mathrm{upw}}(n-1)\right)^2 \sum_{m=0}^{n} \left(\Delta \bar{C}_{nm}^2 + \bar{S}_{nm}^2\right) = \sum_{n=2}^{\infty} c_n^{\Delta g} \tag{2.45}$$

Here, $c_n^{\Delta g}$ are gravity anomaly *signal degree variances*, cf. section 3.1. The individual terms $c_n^{\Delta g}$ of the series represent the contribution of each degree to the total variance. The signal degree variance thus gives information on the signal content of a gravity field functional. Degree variances may be calculated for arbitrary functionals of the gravity field, e.g., the *dimensionless* signal degree variance c_n is linked to gravity anomaly signal degree variance in the following way:

$$c_n^{\Delta g} = \gamma_0^2 (n-1)^2 c_n \tag{2.46}$$

where γ_0 is normal gravity on the surface of the reference ellipsoid.

The covariance function between two signals separated by a spherical distance ψ has the following spherical harmonic expansion:

$$C(\psi) = \sum_{n=2}^{\infty} c_n P_n(\cos\psi) \tag{2.47}$$

Equation (2.47) is as we know point-based. The covariance function of gravity anomaly block mean values $\Delta\bar{g}$ needs to be smoothed with an averaging operator $B(\psi)$, which is used to relate spherical harmonics to mean values of circular areas, i.e., a spherical cap with radius ψ_0. The averaging operator itself has the following spherical harmonic expansion:

$$B(\psi) = \sum_{n=0}^{\infty} \beta_n \frac{2n+1}{4\pi} P_n(\cos\psi) \tag{2.48}$$

The averaging coefficients, $\beta_n(\psi_0)$ are the *Pellinen*-coefficients. The radius ψ_0 (not to be confused with the correlation length ψ^0) is chosen such that the area of the spherical cap equals the area of the usual rectangular blocks for which we have mean gravity anomalies. Recursion formulae exist for the computation of $\beta_n(\psi_0)$ (Gerlach, 2003).

Thus, the covariance function between gravity anomaly mean block values in spherical harmonics yields:

$$\bar{C}(\psi) = \sum_{n=2}^{\infty} \beta_n^2(\psi_0) c_n^{\Delta g} P_n(\cos\psi) \tag{2.49}$$

Several *models* for degree variances exist, since the typical structure of degree variances is that it has high values for the lowest degrees and drops with increasing degree, cf. Figure 2.8. *Kaula*'s model fits well to the dimensionless degree variances (Kaula, 1966):

$$c_n = 0.5 \frac{1.6 \cdot 10^{-10}}{n^3} \tag{2.50}$$

Another well-established model is *Tscherning/Rapp*'s model for gravity anomaly degree variances found in (Tscherning and Rapp, 1974):

$$c_n = s^{n+2} \frac{A(n-1)}{(n-2)(n+B)}, \quad n \geq 3 \tag{2.51}$$

where $A = 425.28\,\mathrm{mGal}^2$, $B = 24$, $s = 0.999617$ and $c_{n=2} = 7.5\,\mathrm{mGal}^2$.

Also worth mentioning is *Flury*'s model for gravity anomaly degree variances (Flury, 2006). This model fits degrees $n \geq 400$ particularly well, and reads (in mGal^2):

$$c_n = \frac{6.8 \cdot 10^7}{(n+0.5)^{3.1}}, \quad n \geq 400 \tag{2.52}$$

From the above it is clear that degree variances may be computed from the potential coefficients of a GGM, and these coefficients also have uncertainties. The degree variances computed from the *errors of the coefficients* are termed *error degree variances*. Since the error degree variances are part of the chosen GGM, they represent a *commission error*. Further, the signal content *above* the maximum degree of the GGM (a "neglected" error) represents an *ommission error*, calculable from a signal degree variance model such as *Tscherning/Rapp* (Gerlach, 2009).

If, for a chosen GGM, the signal degree variances are plotted together with the error degree variances, their crossing point is where the signal-to-noise ratio is one, and it thus defines the resolution of the GGM. Above the corresponding degree, the noise supersedes the signal.

In (Knudsen, 1987), a definition of the local covariance function is reviewed. It reads: "A local covariance function is a special case of a global covariance function where the

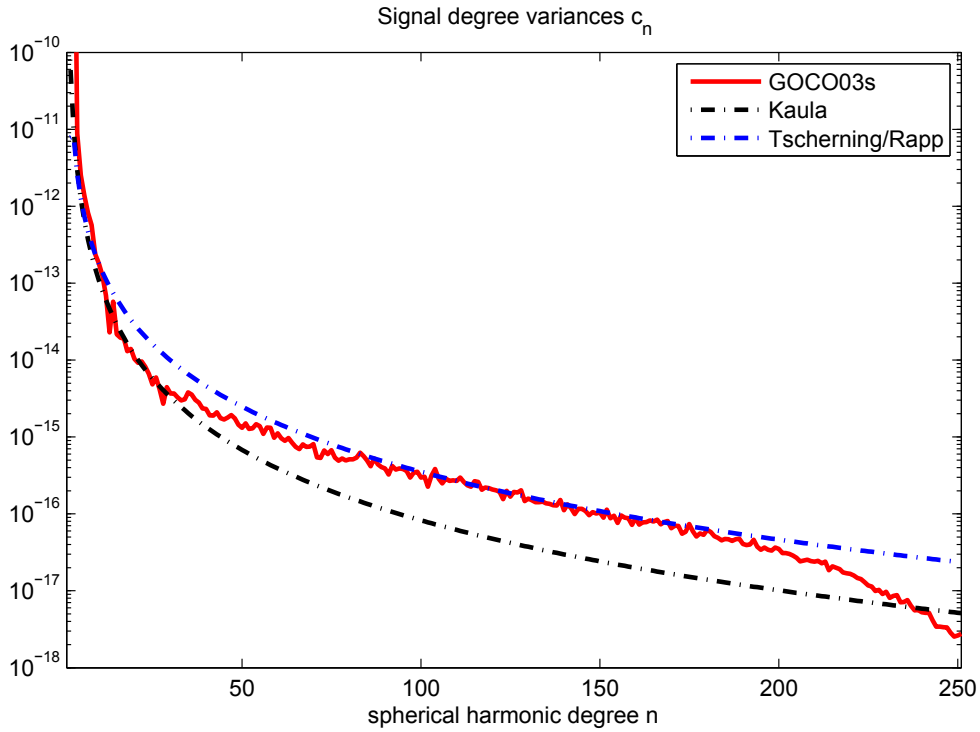

Figure 2.8: *Signal degree variances c_n computed from the potential coefficients of the satellite-only GGM* GOCO03s*, and the models* Kaula *and* Tscherning/Rapp

information content of wavelengths longer than the extent of the local area has been removed, and the information outside, but nearby, the area is assumed to vary in a manner similar to the information within the area."

Topographic reductions (cf. section 4.1) are often applied to smooth the gravity field. A smooth field simplifies prediction and makes the covariance function less dependent on the spatial distribution of the observations. The topography may also be seen as a deterministic part of the observation signal which should be removed before estimating the covariance function.

Still considering spherical harmonics, the low-frequency part of the global covariance function may be removed in the following way ($N_{\max}$ is the maximum degree of the chosen GGM):

$$C_L(\psi) = \bar{C}(\psi) - \sum_{n=2}^{N_{\max}} \beta_n^2(\psi_0) c_n^{\Delta g} P_n(\cos\psi) \tag{2.53}$$

If we consider more local applications, a *planar* local covariance function $C_L(s)$ may suffice. The local covariance function needs to be fitted to empirical values. To estimate a local empirical covariance function, all signal observation pairs x_i and y_i separated by a distance s are considered:

$$C_L(s) = \frac{1}{n} \sum_i x_i \cdot y_i \tag{2.54}$$

Observations are usually *free-air* or rather *Bouguer* anomalies, cf. section 4.1. Since one often considers a limited area, the *mean value* will not necessarily be zero (the observations do not present a stochastic behavior), long-wavelength *trends* may occur and the free-air anomaly correlation with height is accentuated. Assumptions of homogeneity and isotropy are consequently erroneous, and equation (2.54) not applicable for covariance function determination.

One way to attack this problem is to expand the deterministic model. By doing this, the analytical function describes increasingly fine signal structures and the residuals behave increasingly stochastic. If the mean value of the observations considered is not zero, they have to be *centered*. When centering the observations one has to take special care that the mean value is representative for the domain. An example underlining the importance of a representative mean value may be the calculation of an arithmetic mean for an area. A part of the area considered may present dense points that bias the mean value to the values of those points. A covariance function estimated on such a basis will consequently be a biased one.

A trend could present itself as a tilt of the observations in some geographical direction (the observations represent a tilted plane) due to long-wavelength effects. A low-order trend surface may be determined from this, and the observations reduced for it. Trend reduction is also crucial for ensuring the stochasticity of the signal and in principle, any trend function is applicable as long as the observations are reduced by the same trend before being used for prediction. The *remove-restore* technique (also described in section 4.2.1) is applied. Firstly, the observations are reduced by a trend function, the same trend function that was used for deriving the covariance function. Secondly, new point values are predicted and, thirdly, the trend reduction is reverted with respect to the trend function for the new point values.

Subsequently, the discrete values of the local empirical covariance function is fitted to a functional model, in order to derive covariances for arbitrary distances s. Possible models are, e.g., the *exponential* model:

$$C_L(s) = C_0 \cdot e^{-\ln 2 \left(\frac{s}{\psi^0}\right)^2} \tag{2.55}$$

or the *Hirvonen* model:

$$C_L(s) = C_0 \cdot \left(1 + \left(2^{\frac{1}{J}} - 1\right) \left(\frac{s}{\psi^0}\right)^2\right)^{-J} \tag{2.56}$$

where local covariance function parameters C_0, ψ^0 and J are variance, correlation length and the curvature parameters respectively.

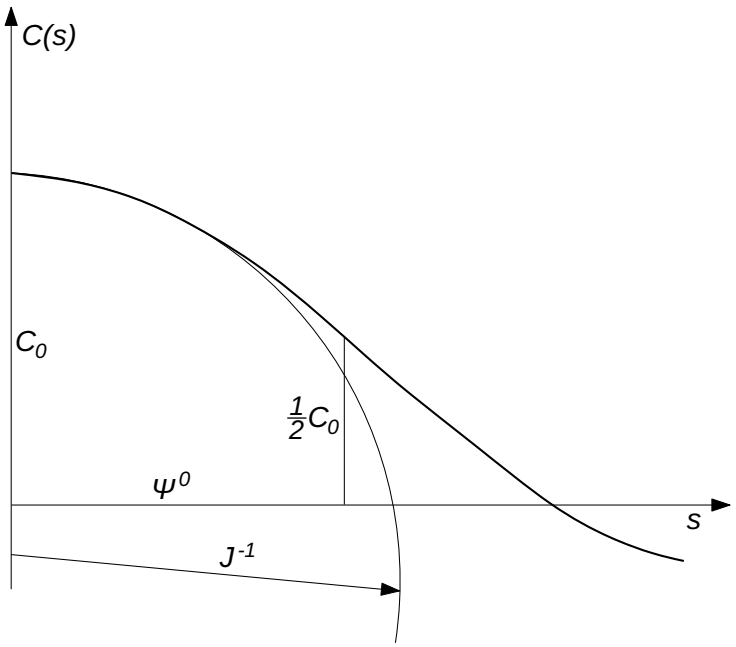

Figure 2.9: *Local covariance function parameters variance C_0, correlation length $C(\psi^0)$ and curvature J*

The local covariance function at "full width at half maximum", i.e., at *correlation length* ψ^0, is formulated as follows:

$$C(\psi^0) = \frac{1}{2}C_0 \tag{2.57}$$

where ψ^0 is chosen as the spherical distance for which $C(\psi^0) = 0.5C_0$. The variance is a measure of the *size* of the signal values, the correlation length a measure of the *roughness* of the signal (Gerlach, 2003).

The *curvature parameter* J is related to the curvature κ of the covariance curve at $s = 0$, $J = \kappa(\psi^0)^2/C_0$. The three parameters C_0, $C(\psi^0)$ and J characterize the local covariance function, cf. Figure 2.9 (Moritz, 1980).

The empirical covariance function can be fitted graphically by testing the different covariance function parameters and seeing which choice of parameters best fits the empirical values.

Chapter 3

Gravity field modeling approaches

The different mathematical models for representing the gravity field, hereunder geoid determination, have different application depending on whether the *global* or a *regional* gravity field is sought.

When the *global* gravity field, or functionals of it, like the geoid N, is sought, the most used model is spherical harmonics, introduced in section 2.1 and typified for the gravitational potential in equation (2.16). A general introduction is given in section 3.1. Spherical harmonic coefficients are not dependent on position at all, and have perfect frequency localization. Consequently, spherical harmonics show no spatial localization, i.e., they have *global support* and show values significantly different from zero at almost every position on the sphere (Eicker, 2008).

Whether the regional or the local gravity field is considered, and a regional or local geoid is to be computed, the same computation scheme applies for both and is explained in detail in section 4.2. Quite basically, the concept is to remove the low-frequency gravity field information from a global geopotential model as well as the high-frequency information from a digital terrain model from measured gravity in a regional or local area and compute the residual part of the gravity field by e.g.:

- The integral formula solutions to the geodetic boundary value problem, cf. section 3.2
- Least-squares collocation, cf. section 3.3.2
- Radial basis functions, cf. section 3.4

It then becomes clear that the difference between a regional or a local geoid is the extent, or resolution. A compromise based on the application is made, since smaller areas (increasingly local) model higher frequencies better (Boener, 1994).

A regional geoid may represent geographical region, and its application may be to derive orthometric heights from ellipsoidal heights determined from GNSS by equation (2.32), thus substituting relatively time-consuming spirit leveling. Global, regional as well as local geoids may aid geophysical and oceanographic research, cf. section 1.1.

3.1 Spherical harmonics

A more general mathematical view of spherical harmonics follows, to make the transition to *other kinds of base functions* more feasible, cf. section 3.4. Any function $f(\theta, \lambda) \in \mathcal{L}_2(\sigma)$ (where $\mathcal{L}_2$ is a space of square integrable functions) on a unit sphere σ can be developed into a series of spherical harmonics, i.e., a sum of all solutions to Laplace's equation:

$$f(\theta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} F_{nm} Y_{nm}(\theta, \lambda) \tag{3.1}$$

Equation (3.1) is known as spherical harmonic synthesis. F_{nm} are the spherical harmonic coefficients (also known as *Stokes coefficients*), a more compact way of expressing the spherical Fourier coefficients. E.g., if the considered gravity signal is the potential V, Stokes coefficients are:

$$F_{nm} = \sqrt{4\pi} \begin{cases} \bar{C}_{nm} & \text{for } m = 0, 1, 2, \ldots, n \\ \bar{S}_{n|m|} & \text{for } m = -n, \ldots, -2, -1 \end{cases} \tag{3.2}$$

Laplace's surface spherical harmonics are defined as follows:

$$Y_{nm}(\theta, \lambda) = \frac{1}{\sqrt{4\pi}} \begin{cases} \bar{P}_{nm}(\cos\theta) \cos m\lambda & \text{for } m = 0, 1, 2, \ldots, n \\ \bar{P}_{n|m|}(\cos\theta) \sin |m|\lambda & \text{for } m = -n, \ldots, -2, -1 \end{cases} \tag{3.3}$$

An important property of spherical harmonics is *orthogonality*, i.e., when two base functions Y_{nm} and Y_{pq} are considered, their integrated product over the sphere is zero if their variables differ:

$$\frac{1}{4\pi} \iint_{\sigma} Y_{nm}(\theta, \lambda) Y_{pq}(\theta, \lambda) \mathrm{d}\sigma = \delta_{np} \delta_{mq} \tag{3.4}$$

where the *Kronecker delta* $\delta_{np} = 1$ for $p = n$ and $\delta_{np} = 0$ for $p \neq n$.

In light of the above, a function $f(\theta, \lambda)$ on the sphere σ as represented in equation (3.1) has a unique representation and the base functions Y_{nm} form an orthonormal basis.

Since the spherical harmonic basis is *orthonormal* the least-squares projection of a scalar function f defined over σ is done by integrating the product of f and the *base functions*. Thus, the spherical harmonic coefficients may be determined by the *spherical Fourier transform*, which is known as *spherical harmonic analysis*:

$$F_{nm} = \iint_{\sigma} f(\theta, \lambda) Y_{nm}(\theta, \lambda) \mathrm{d}\sigma \tag{3.5}$$

The *addition theorem* of spherical harmonics establishes a relationship between the Legendre polynomials $P_n(\cos\psi)$ on the left-hand side and the base functions Y_{nm} on the right-hand side:

$$\frac{2n+1}{4\pi}P_n(\cos\psi) = \sum_{m=-n}^{n} Y_{nm}(\theta',\lambda')Y_{nm}(\theta,\lambda) \tag{3.6}$$

Equation (3.6) may be considered as a generalization of the following trigonometric identity:

$$\cos(\theta'-\theta) = \cos\theta'\cos\theta + \sin\theta\sin\theta' \tag{3.7}$$

where the role of the trigonometric functions on the right-hand side is played by spherical harmonics and that of the left-hand side is played by the Legendre polynomials.

Also, the *norm* of the function f may be expressed in spherical harmonics:

$$\|f(\theta,\lambda)\|^2 = \frac{1}{4\pi}\iint_\sigma f(\theta,\lambda)^2\,\mathrm{d}\sigma = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} F_{nm}^2 = \sum_{n=0}^{\infty} c_n^2 \tag{3.8}$$

where c_n^2 are the *degree variances* of degree n, quite relevant for the statistical description of Earth's gravity field, cf. section 2.5.

As mentioned in section 2.1, the spherical harmonics are adaptable to the geodetic problem of describing functionals of Earth's gravity field. We go from the unit sphere σ, to the exterior of a sphere σ_R with mean Earth radius R and express the gravitational potential as follows:

$$V(r,\theta,\lambda) = \frac{GM}{R}\sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+1}\sum_{m=0}^{n}(\bar{C}_{nm}\bar{\mathcal{R}}_{nm} + \bar{S}_{nm}\bar{\mathcal{S}}_{nm}) \tag{3.9}$$

The factor GM/R renders the potential coefficients dimensionless. The *upward continuation factor* $\left(\frac{R}{r}\right)^{n+1}$ describes the potential at a point (r,θ,λ) in the exterior of the sphere. Further, the spherical harmonic coefficients, or potential coefficients in this special case, are explicitly separated into sine and cosine functions, and the base functions are chosen as follows, cf. Figure 3.1:

$$Y_{nm}(\theta,\lambda)\begin{cases}\bar{\mathcal{R}}_{nm} = \bar{P}_{nm}(\cos\theta)\cos m\lambda \\ \bar{\mathcal{S}}_{nm} = \bar{P}_{nm}(\cos\theta)\sin m\lambda\end{cases} \tag{3.10}$$

where $\bar{P}_{nm}(\cos\theta)$ are the fully normalized Legendre functions of degree n and order m, calculable from, e.g., (Holmes and Featherstone, 2002), $t = \cos\theta$:

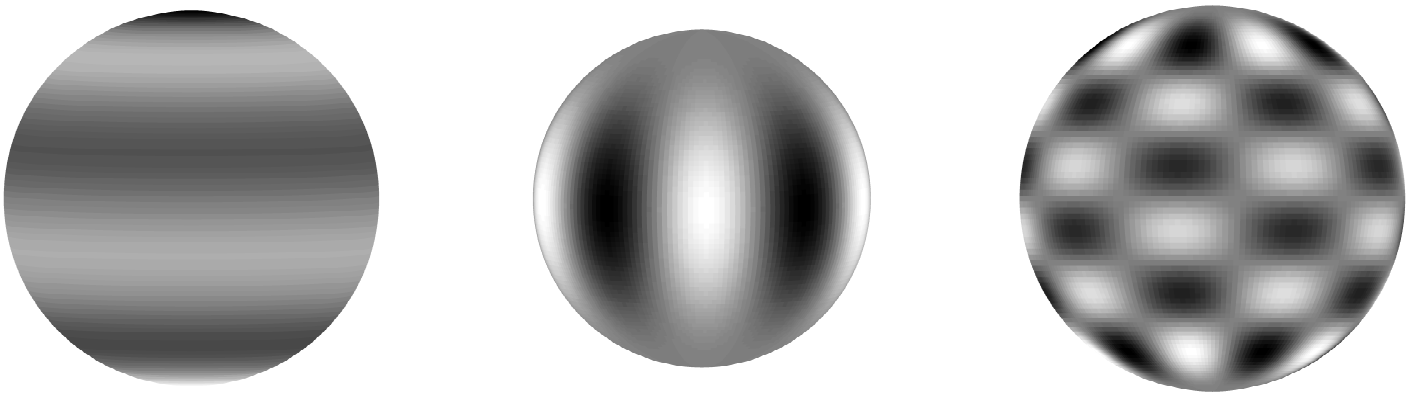

Figure 3.1: *The surface spherical harmonics from left to right: $Y_{5,0}$ (zonal harmonics, $m = 0$), $Y_{5,5}$ (sectorial harmonics, $m = n$) and $Y_{10,5}$ (tesseral harmonics, $n > m,\ m > 0$)*

$$\bar{P}_{n,m}(t) = a_{n,m} t \bar{P}_{n-1,m}(t) - b_{n,m} \bar{P}_{n-2,m}(t) \quad \forall \quad n > m \tag{3.11}$$

for the fully normalized *non-sectorial* (i.e., $n > m$) $\bar{P}_{nm}(t)$, with:

$$a_{n,m} = \sqrt{\frac{(2n-1)(2n+1)}{(n-m)(n+m)}} \tag{3.12}$$

$$b_{n,m} = \sqrt{\frac{(2n-1)(n+m-1)(n-m-1)}{(n-m)(n+m)(2n-3)}} \tag{3.13}$$

The *sectorial* (i.e., $n = m$) $\bar{P}_{nm}(t)$ serve as seed values for equation (3.11), with start values $\bar{P}_{0,0}(t) = 1$ and $\bar{P}_{1,1}(t) = \sqrt{3}u$ with $u = \sin\theta$:

$$\bar{P}_{m,m}(t) = u^m \sqrt{3} \prod_{i=2}^{m} \sqrt{\frac{2i+1}{2i}} \quad \forall \quad m \geq 1 \tag{3.14}$$

For the fully normalized Legendre polynomials (i.e., $m = 0$), the following holds, cf. Figure 3.2:

$$\bar{P}_n(t) = \sqrt{2n+1} P_n(t) = \sqrt{2n+1} \left(\frac{1}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{d}t^n} (t^2 - 1)^n \right) \tag{3.15}$$

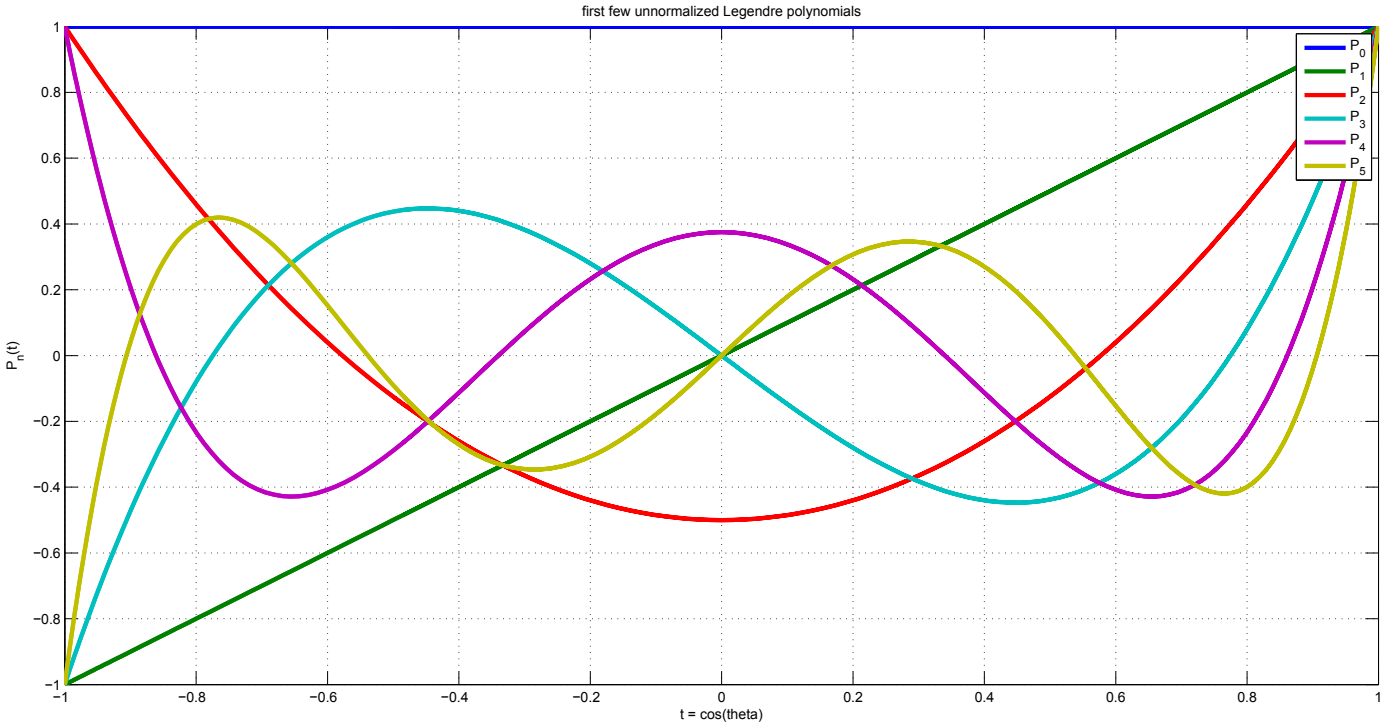

Figure 3.2: *The first few unnormalized Legendre polynomials* ($m = 0$)*, with* $t = \cos\theta$ *and* $\theta = [-\pi, \pi]$

3.1.1 Global geopotential models

A *global geopotential model*, GGM, contains a set of potential coefficients $\bar{C}_{nm}, \bar{S}_{nm}$ up to a maximum degree $N_{\max}$ specifically related to the defining constants GM and R. *Degree-dependent weighting* of the potential coefficients when functionals of the linear gravity field are sought, is done through the different spherical harmonic syntheses, cf. Table 3.1.

GGMs are obtained from one or more of several data sources, e.g., surface gravity anomalies and dedicated satellite gravity missions such as *CHAMP*, *GRACE* and *GOCE*. This leads to the distinction between *satellite-only* (containing data only from satellite observations) and *combined* (containing both low-frequent satellite data and high-frequent surface data such as gravity anomalies) GGMs (Gerlach, 2003).

The potential coefficients are obtained from observations through spherical harmonic analysis, cf. equation (3.5). Its solution based on least-squares adjustment is presented in section 3.1.2.

An example of a satellite-only GGM is the recently (2012) released *GOCO03s* model up to degree $N_{\max} = 250$ based on combination of satellite data sets from *GOCE*, *GRACE*, *CHAMP* and *satellite laser ranging* (SLR). A state-of-the-art combined GGM is *EGM2008*, based on a satellite-only *GRACE* model up to degree $n = 180$ and a global gravity anomaly database with a resolution of 5 arc minutes (corresponding to a maximum degree of $N_{\max} = 2190$).

The degree of a global geopotential model can be related to the spatial resolution on the sphere. If the number of cycles on the sphere generated by the Legendre polynomials is considered, we may observe that the degree number is equal to the number of cycles. An approximation of the wavelength of harmonics n is obtained, and the spatial resolution is half of the wavelength of the smallest spatial feature (Gerlach, 2003), i.e.:

$$\Delta = \frac{\lambda}{2} = \frac{\pi R}{n} \approx \frac{20000\,\mathrm{km}}{n} \tag{3.16}$$

All linear functionals of the gravity field may be expressed in terms of spherical harmonics and the global geopotential models account for the *low-frequency* part of the gravity field. The global *geoid* is an important part of local and regional geoid determination, cf. section 3.2.

The spherical harmonic expansion of the disturbing potential T was presented in equation (2.24). The relations between T and other gravity field functionals allow spherical harmonic expansions of, e.g., the geoid height, the height anomaly, the gravity disturbance and the gravity anomaly. The spherical harmonic expansions are summarized in Table 3.1, where a common *upward continuation operator* is introduced, by placing the geocentric radius r of the computation point in front of the summation instead of the scaling factor R:

$$\lambda_n^{\text{upw}} = \left(\frac{R}{r}\right)^n \tag{3.17}$$

T	$=$	$\frac{GM}{r}$	$\sum_{n=2}^{\infty}$	λ_n^{upw}	$\sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta)\left[\Delta\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda\right]$
ζ	$=$	$\frac{GM}{\gamma r}$	$\sum_{n=2}^{\infty}$	λ_n^{upw}	$\sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta)\left[\Delta\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda\right]$
Δg	$=$	$\frac{GM}{r^2}$	$\sum_{n=2}^{\infty}$	$\lambda_n^{\text{upw}}\ (n-1)$	$\sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta)\left[\Delta\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda\right]$
δg	$=$	$\frac{GM}{r^2}$	$\sum_{n=2}^{\infty}$	$\lambda_n^{\text{upw}}\ (n+1)$	$\sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta)\left[\Delta\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda\right]$
T_{rr}	$=$	$\frac{GM}{r^3}$	$\sum_{n=2}^{\infty}$	$\lambda_n^{\text{upw}}\ (n+1)(n+2)$	$\sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta)\left[\Delta\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda\right]$

Table 3.1: *Spherical harmonic synthesis of disturbing gravity field functionals*

A GGM only provides the potential on the Earth surface, while the geoid, in large parts located *inside* the topographic masses, is not directly computable from a GGM, cf. section 3.2.2 on analytical continuation.

When computing the geoid height N, γ_0 on the ellipsoid may be used as an approximation. In addition, we have topographic masses in the exterior of this surface and density knowledge is assumed. However, by taking the separation between geoid and quasigeoid into account (cf. section 2.4), the geoid height may be computed in the following way (Torge and Müller, 2012):

$$N(r,\theta,\lambda) = \zeta(r,\theta,\lambda) + \frac{\Delta g_B}{\gamma_0} H \tag{3.18}$$

However, the separation between quasigeoid and geoid is also an approximation, dependent on topographic height differences, and not precise in mountainous areas. It therefore holds only in areas where the separation is quite small.

Also in Table 3.1, different *spectral eigenvalues* of the functionals are observed. They allow easy transfer between the different functionals in the spectral domain (Gerlach, 2003), see the spectral eigenvalue summary in Table 3.2. The spectral eigenvalues are of great aid, e.g., when computing cross-covariance functions, cf. section 3.3.2.

3.1.1.1 A remark on tides

Tides also influence Earth's gravity field. They comprise the gravitational attraction of the Sun and Moon on a point in space or on Earth's surface, relative to Earth's center of mass. Tidal forces are always present and therefore included in every gravity observation.

Disturbing gravity field functional	Spectral eigenvalue λ_n
ζ	1
Δg	$(n-1)$
δg	$(n+1)$
T_{rr}	$(n+1)(n+2)$

Table 3.2: *Gravity field functional eigenvalues*

The problem arises in that the masses of Sun and Moon are *outside* the Earth, and Laplace's equation consequently does not hold. We split the tidal effect into a direct effect in that the attraction from Sun and Moon changes the potential, and an indirect effect due to Earth deformation caused by tidal forces also changing the potential (Torge and Müller, 2012). The indirect effect, however, causes Earth mass distribution changes and therefore does not violate Laplace's equation as such.

Tidal forces may be split into time-dependent and time-independent parts, where the time-dependent periodic part is caused by Earth's rotation and the orbits of Sun and Moon, and evens out over time. What must be considered when determining the gravity field is the time-independent part, which is ever present, and therefore termed *permanent tidal deformation* (Torge and Müller, 2012).

If the time-dependent part of the tidal forces is removed from the observed gravity field quantity, we are in the *mean-tide* (MT) system (Petit and Luzum, 2010). The permanent part is also included in this system, i.e., both direct and indirect effects. A MT geoid, for instance, would correspond to a mean ocean surface free of non-gravitational effects (like wind or ocean currents).

If the direct effect of the permanent part is also removed, we are in the *zero-tide* (ZT) system. For the potential, the removal of the permanent direct effect gives ZT potential. The permanent indirect effect related to deformation is still present. For geometry-related quantities such as station positions on the crust, ZT is synonymous with MT (Petit and Luzum, 2010), i.e.:

$$\text{Geopotential:} \qquad \text{MT} \neq \text{ZT} \neq \text{TF} \tag{3.19}$$

$$\text{Geometry:} \qquad \text{MT} = \text{ZT} \neq \text{TF} \tag{3.20}$$

If the indirect effect is also removed, we are in the *tide-free* (TF) system, supposing no Sun and Moon exists and consequently no tidal forces. For the removal of the permanent

indirect effect, however, a model must be applied. Tidal models recommended by the *Internatinal Earth Rotation and Reference System Service* (IERS) use conventional *Love* numbers, which are dimensionless parameters describing the sensitivity of the solid Earth to tidal forces. Results using the IERS models are then in the so-called *conventional tide-free* (CTF) system, and *not* in a TF system (supposing no perturbing bodies at all). If the indirect effect of the permanent tide is to be removed, so-called *fluid limit values* for the Love numbers could be applied in the models (Petit and Luzum, 2010).

For example, in case of GGM *EGM2008*, the difference between ZT and TF value of C_{20} is $-4.1736 \cdot 10^{-9}$ (Petit and Luzum, 2010).

A 1984 IERS recommendation stated that, "the indirect effect due to the permanent yielding of the Earth be not removed", i.e., a ZT system was recommended for quantities associated with Earth's geopotential, and a MT system for quantities related to oceanography, station displacements and alike (Petit and Luzum, 2010). This means that, in practice, geodetic products are given in different tidal systems depending on application (Torge and Müller, 2012).

3.1.2 Spherical harmonic analysis by least-squares adjustment

If the potential $f(r, \theta, \lambda)$ is known, the spherical harmonic coefficients F_{nm} are obtainable through spherical harmonic analysis, cf. equation (3.5). Determining the "output values" $f(r, \theta, \lambda)$ from the "input values" F_{nm}, e.g., spherical harmonic synthesis, is a *forward problem* (Eicker, 2008). Determining the input values F_{nm} from the output values $f(r, \theta, \lambda)$ is known as an *inverse problem.*

Equation (3.5) may be solved through *numerical integration* or *parameter estimation.* Parameter estimation or least-squares adjustment solves for the coefficients F_{nm} (e.g., $\bar{C}_{nm}$,$\bar{S}_{nm}$). The coefficients are in turn used to reconstruct an *approximation* of the field f, increasingly accurate as the number of bands (degrees n) increases.

Advantages of parameter estimation are also underlined in sections 3.3.1 and 3.3.2. We may combine different types of gravity observations in one step, no preparation of the data in terms of interpolation (gridding) is necessary, and a concise frame taking the accuracies of the different observations into account is embedded in the procedure (Bentel et al., 2013). Even upward and downward continuation is directly done, although downward continuation presents itself as an ill-posed problem, cf. section 5.1.

A disadvantage with parameter estimation techniques is that, when the number of observations becomes large, a huge computational effort is needed to invert the matrix product $\left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1} = \mathbf{N}^{-1}$, cf. equation (3.27), and special care has to be taken regarding the structure and organization of the design matrix $\mathbf{A}$.

Generalizing equation (3.1) by going from the unit sphere to a point in the exterior of a sphere σ_R with arbitrary radius R at distance r from the sphere center (cf. equation (3.9)) through the upward continuation operator yields (Bentel et al., 2013):

$$f(r,\theta,\lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{R}{r}\right)^{n+1} F_{nm} Y_{nm}^{R}(r,\theta,\lambda) \tag{3.21}$$

with $Y_{nm}^{R}(r,\theta,\lambda) = \frac{1}{R} Y_{nm}(r,\theta,\lambda)$. We observe that equation (3.21) also holds for points *on* the sphere.

The observation equation of observed signal f and associated error v in terms of spherical harmonics may be formulated as follows (note that $\mathbf{x}$ is used to describe both the position vector and the vector of unknowns):

$$\mathbf{l} + \mathbf{v} = \mathbf{A}\mathbf{x} \tag{3.22}$$

where the expectation $\mathrm{E}\{\mathbf{l}\} = \mathbf{A}\mathbf{x}$ and dispersion (scatter, spread) $\mathrm{D}\{\mathbf{l}\} = \mathbf{Q}_{ll}$. Written out (expressed with position vector $\mathbf{x} = r[\sin\theta\cos\lambda, \sin\theta\sin\lambda, \cos\theta]^T$ for brevity), equation (3.22) reads:

$$\begin{bmatrix} f(\mathbf{x}_1) \\ f(\mathbf{x}_2) \\ \vdots \\ f(\mathbf{x}_L) \end{bmatrix} + \begin{bmatrix} v(\mathbf{x}_1) \\ v(\mathbf{x}_2) \\ \vdots \\ v(\mathbf{x}_L) \end{bmatrix} = \begin{bmatrix} Y_{0,0}^{R}(\mathbf{x}_1) & \cdots & Y_{N_{\max},N_{\max}}^{R}(\mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ Y_{0,0}^{R}(\mathbf{x}_L) & \cdots & Y_{N_{\max},N_{\max}}^{R}(\mathbf{x}_L) \end{bmatrix} \cdot \begin{bmatrix} F_{0,0} \\ F_{1,-1} \\ \vdots \\ F_{N_{\max},N_{\max}} \end{bmatrix} \tag{3.23}$$

where L observations $\mathbf{l}$ and $L \times (N_{\max}+1)^2$ base functions Y_{nm} are considered. We observe that the evaluation of f at point $\mathbf{x}$ is a dot product between the coefficient vector $\mathbf{x}$ and the design matrix $\mathbf{A}$ of evaluated base functions. The above functional model is associated with the following stochastic model:

$$\mathbf{v} \sim N(\mathrm{E}\{\mathbf{v}\} = 0, \mathbf{\Sigma}_{ll}) \tag{3.24}$$

$\mathrm{E}\{\mathbf{v}\} = 0$ expresses that the residual (discrepancy) vector $\mathbf{v}$ has zero expectancy, i.e., the model is *unbiased*. In terms of the weight matrix of the observations:

$$\mathbf{P} = \frac{1}{\sigma_0^2} \mathbf{\Sigma}_{ll}^{-1} = \mathbf{Q}_{ll}^{-1} \tag{3.25}$$

σ_0^2 is the a priori variance of unit weight, relating the weight matrix and the inverted covariance matrix $\mathbf{\Sigma}$. $\mathbf{Q}_{ll}$ is the *cofactor matrix* of the observations. σ_0^2 may also be estimated from observations as an a posteriori variance of unit weight $\hat{\sigma}_0^2$. If both are *statistically* equal, the least-squares adjustment is dependable (Leick, 2004).

Least-squares adjustment is an optimal estimation procedure where the parameters $\mathbf{x}$ are estimated with a minimum-variance optimization criterion:

$$\hat{\mathbf{x}} = \min\{\mathbf{v}^T\mathbf{P}\mathbf{v}\} \tag{3.26}$$

Equation (3.22) may be solved for the unknown parameters (in our case coefficients F_{nm}) in the following way:

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{P}\mathbf{l} \tag{3.27}$$

where $\mathbf{A}^T\mathbf{P}\mathbf{A}$ is known as the *normal matrix* $\mathbf{N}$. For the expectation and dispersion of $\hat{\mathbf{x}}$ we may write (Sneeuw, 2000):

$$\mathrm{E}\{\hat{\mathbf{x}}\} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{P}\mathrm{E}\{\mathbf{l}\} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{P}\mathbf{A}\mathbf{x} = \mathbf{x} \tag{3.28}$$

$$\mathrm{D}\{\hat{\mathbf{x}}\} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1} = \mathbf{N}^{-1} = \mathbf{Q}_{xx} \tag{3.29}$$

where $\mathbf{Q}_{xx}$ is the cofactor matrix of the unknowns. The estimated covariance matrix of the unknowns:

$$\hat{\mathbf{\Sigma}}_{xx} = \hat{\sigma}_0^2\left(\mathbf{A}^T\mathbf{P}\mathbf{A}\right)^{-1} = \hat{\sigma}_0^2\mathbf{Q}_{xx} \tag{3.30}$$

The a posteriori variance of unit weight is computed by:

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^T\mathbf{P}\mathbf{v}}{L - (N_{\max}+1)^2} \tag{3.31}$$

with estimated residuals (or observation errors) $\hat{\mathbf{v}}$:

$$\hat{\mathbf{v}} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{l} \tag{3.32}$$

If the normal equation matrix $\mathbf{N}^{-1}$ exists, the point grid system for the observation is *admissable* and $L \geq (N_{\max}+1)^2$ (the system is overdetermined) (Bentel et al., 2013). Regularization, cf. section 5.2, may be used to solve singular systems.

3.2 Integral formulae

Spherical harmonics, as presented in section 3.1, delivers high-frequent gravity field information at fairly high cost (since the degree of the spherical harmonic expansion must be very high). Representing an alternative approach are the integral formula solutions to the geodetic boundary value problems of Stokes and Molodensky, allowing a point-wise calculation (allowing in principle an arbitrarily high gravity field resolution) of gravity field quantities. Only the data availability (density) and quality act as limits (Torge and Müller, 2012).

In the following, the well-established Stokes integral formula for geoid determination is derived, and subsequently, Molodensky's integral formula for determining height anomalies.

3.2.1 Derivation of Stokes's formula

The transferring from a reference surface to a point in space is known as *upward continuation*. A harmonic function H's value in space, known on the spherical Earth's surface, may be evaluated by *Poisson's integral*, which is a more direct solution to *Dirichlet's* BVP. Moreover, H may be expanded into a spherical harmonic series, with degree one and zero omitted:

$$H_P = \frac{1}{4\pi} \iint_\sigma \left(\frac{-r^3 - R^2 r}{l^3} + 1 + \frac{3R}{r} \cos\psi \right) H \, \mathrm{d}\sigma \tag{3.33}$$

where ψ is the spherical distance, $l = \sqrt{r^2 + R^2 - 2Rr\cos\psi}$. $\iint_\sigma$ is the surface integral over a reference sphere. H is the value of the harmonic function at the variable surface element $R^2 \mathrm{d}\sigma$, H_P refers to the fixed point P, cf. Figure 3.3.

The fundamental equation of physical geodesy, (2.33), may only be regarded as a boundary condition as long as the gravity anomalies Δg are known on the surface. With equation (3.33), gravity anomalies *in space* may be computed and equation (2.33) becomes a differential equation integrable with respect to r. This is possible, since T, in addition to the boundary condition fulfills Laplace's equation $\Delta T = 0$. Equation (2.33) is multiplied by $-r^2$:

$$-r^2 \Delta g(r) = \frac{\partial r^2 T}{\partial r} \tag{3.34}$$

Equation (3.34) is integrated over $[\infty, r]$:

$$r^2 T \Big|_\infty^r = -\int_\infty^r r^2 \Delta g(r) \, \mathrm{d}r \tag{3.35}$$

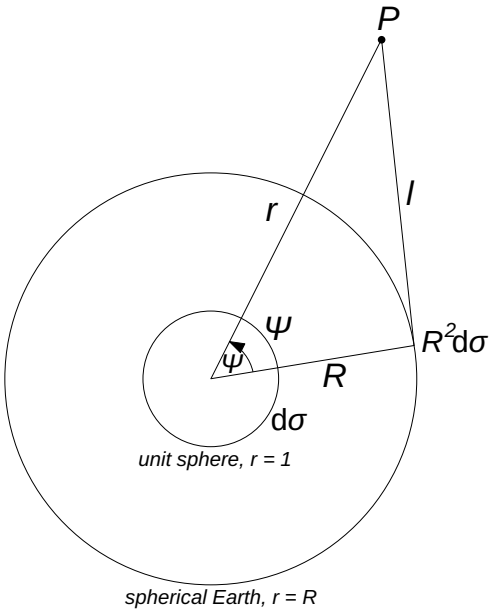

Figure 3.3: *Geometric quantities concerned in Poisson's integral*

The left-hand side of equation (3.35) becomes only r^2T since $\lim_{r\to\infty} r^2T = 0$. Substituting the surface function H with Δg allows to compute T at a point P in space from surface gravity anomalies:

$$r^2T = \frac{R^2}{4\pi}\int_{\infty}^{r}\left[\iint_{\sigma}\left(\frac{-r^3 - R^2r}{l^3} + 1 + \frac{3R}{r}\cos\psi\right)\Delta g\,\mathrm{d}\sigma\right]\mathrm{d}r \tag{3.36}$$

The order of integration is changed, the improper integral evaluated, the spherical distance is expanded into a series and *Pizzetti's formula* is obtained:

$$T(r,\theta,\lambda) = \frac{R}{4\pi}\iint_{\sigma} S(r,\psi)\Delta g\,\mathrm{d}\sigma \tag{3.37}$$

On the geoid, it holds $r = R$, and equation (3.37) becomes radially independent. Applying equation (2.31) *Stokes's formula* is obtained:

$$N = \frac{R}{4\pi\gamma}\iint_{\sigma}\Delta g S(\psi)\,\mathrm{d}\sigma \tag{3.38}$$

where $S(\psi)$ is known as the *Stokes function*, interpreted as a weight function, which expanded into a spherical harmonic series may be expressed as:

$$S(\psi) = \sum_{n=2}^{\infty}\frac{2n+1}{n-1}P_n(\cos\psi) \tag{3.39}$$

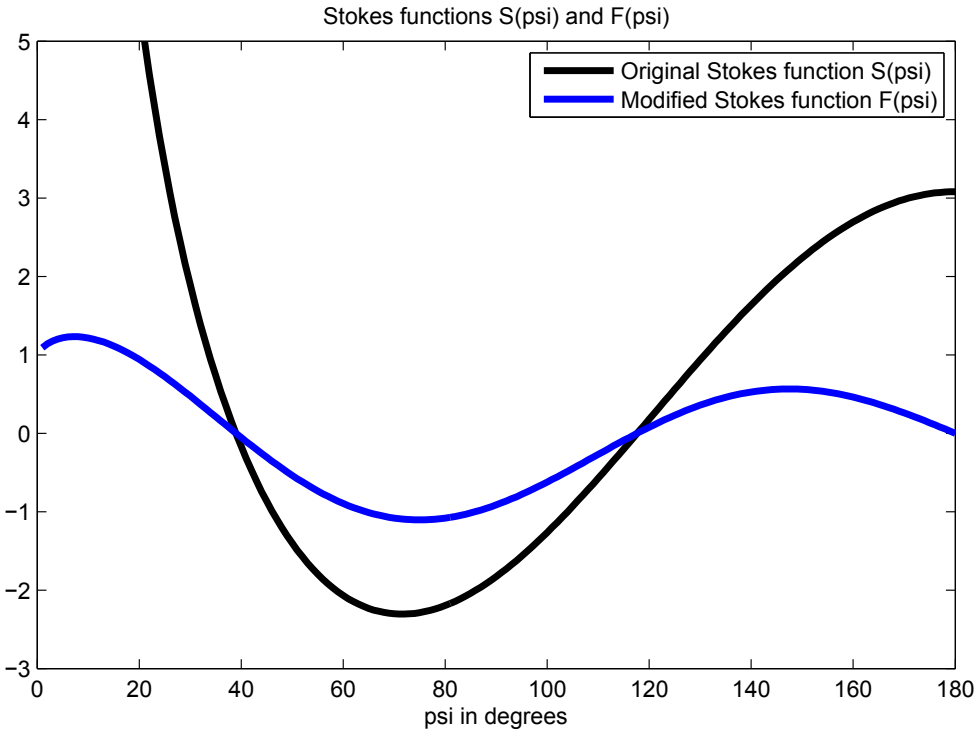

Figure 3.4: *Original Stokes's function $S(\psi)$ and modified Stokes's function $F(\psi)$, breaking the singularity at $\psi = 0$*

The Stokes function also has a closed expression:

$$S(\psi) = \frac{1}{\sin\frac{\psi}{2}} + 1 - 5\cos\psi - 6\sin\frac{\psi}{2} - 3\cos\psi\ln\left(\sin\frac{\psi}{2} + \left(\sin\frac{\psi}{2}\right)^2\right) \tag{3.40}$$

ψ is the spherical distance between the variable coordinates (φ, λ) of the surface element $\mathrm{d}\sigma$ and the computation point (φ', λ'):

$$\cos\psi = \sin\varphi\sin\varphi' + \cos\varphi\cos\varphi'\cos(\lambda' - \lambda) \tag{3.41}$$

Stokes's function is shown in Figure 3.4. Its singularity at $\psi = 0$, giving the gravity anomaly at the computation point infinite weight, may be avoided by modifying Stokes's function in the following way (Hofmann-Wellenhof and Moritz, 2005):

$$F(\psi) = \frac{1}{2}S(\psi)\sin\psi \tag{3.42}$$

Several assumptions have been made throughout the description of the anomalous potential culminating in Stokes's formula for the computation of geoid heights. Equation (3.38) performs an integral over a reference *sphere*, so an error of approximately 0.3 % is introduced (equivalent to Earth's flattening). Moreover, the deflection of the vertical is

assumed to be zero, which may not be completely true in mountainous areas. The mass and radius of the reference ellipsoid is assumed equal to the mass and radius of Earth, e.g., $\frac{\delta GM}{r} = 0$, and also the origin of the reference ellipsoid is assumed to coincide with Earth's center of mass.

However, perhaps the greatest concerns with Stokes's formula is firstly that it assumes that the gravity anomalies Δg are given *on the surface of the geoid*, and secondly that the geoid is a boundary surface where the anomalous potential in the exterior of the surface fulfills Laplace's equation, i.e., is harmonic and $\Delta T = 0$.

The gravity anomalies are measured on the surface of the Earth, not on the geoid, but *topographic reductions* may be performed and the gravity anomalies *downward continued* from the Earth's surface to the geoid, cf. section 4.1.

Since we have topographic masses outside the geoid, it cannot serve as a boundary surface, and Laplace's equation is not fulfilled (but rather Poisson's equation). Topographic reductions can either remove the masses outside the geoid completely, or *shift* the masses inside the geoid, thus ensuring that the geoid is a boundary surface and Laplace's equation is fulfilled. However, by shifting or removing masses, the potential T also changes, and consequently the geoid itself. This indirect effect also needs to be accounted for (Gerlach, 2009).

Several approaches to topographic reductions exist, and they all rely on different hypotheses on the density distribution of the topographic masses. This is exactly what Molodensky sought to avoid.

3.2.2 Solving Molodensky's problem by analytical continuation

As mentioned in section 2.4, the gravity vector does not correspond to Earth's surface normal in the Molodensky approach, and with gravity anomalies measured on Earth's surface, equation 3.38 cannot be applied directly. However, it may be extended so that it holds also in the Molodensky case, by *analytical continuation*.

The method of analytical continuation is based on the *Runge-Krarup* theorem, which states that a harmonic function $T^\star$ *inside* the topographic masses but approximating the potential of the exterior (empty space) T may always be found. $T^\star$ is then called the *harmonic* or *analytical* continuation of T in the topographic interior (Hofmann-Wellenhof and Moritz, 2005).

Analytical continuation thus extends the domain in which a function is defined. However, it does not provide the *true values of that function in the extended domain* (Gerlach, 2009)!

Gravity anomalies measured on the surface of the Earth are analytically *downward* continued to a common level. Further, the analytical continuation of the height anomaly

at that level are computed with Stokes's formula and subsequently analytically *upward* continued to the surface again, expressed as:

$$\zeta = \frac{R}{4\pi\gamma_0} \iint_\sigma S(\psi) \overbrace{\left(\Delta g_P - \frac{\partial g}{\partial h} h\right)}^{\Delta g^\star} \mathrm{d}\sigma + \frac{\partial \zeta}{\partial h} h \tag{3.43}$$

where h is the height above the chosen continuation level, Δg_P the gravity anomalies measured on the surface and $\Delta g^\star$ the analytically downward continued gravity anomalies.

As mentioned, the analytically continued values refer to the exterior potential, and are not true values for the interior. The main goal, however, is achieved by analytical continuation: The gravity anomalies refer to a common height, and that spherical surface can serve as a boundary surface. The direction of the surface normal corresponds with the radial direction. The third BVP is applicable again. The chosen continuation level may be at the height of the computation point P, in which case the height anomaly is already at the surface and the last term in equation (3.43) vanishes (Hofmann-Wellenhof and Moritz, 2005).

Equation (3.43) is actually a *first-order* solution, and the series solution theoretically goes to infinity:

$$\zeta = \frac{R}{4\pi\gamma_0} \iint_\sigma S(\psi) \left(\Delta g + g_1 + g_2 + g_3 + \cdots\right) \mathrm{d}\sigma \tag{3.44}$$

The higher-order *Molodensky terms* are all dependent on the terrain inclination, so that they can be neglected in relatively flat areas (Gerlach, 2009). The first term g_1 may be approximated by a terrain correction, cf. section 4.1.

3.3 Least-squares collocation

3.3.1 Least-squares prediction

The geodetic boundary value problems may be solved in terms of integral formulae over the whole Earth, cf. section 3.2. For instance, Stokes's integral formula presupposes the continuous knowledge of gravity anomalies on the surface of the geoid. It is quite clear, however, that our measurements are restricted to discrete points on land and oceans. To "fill in" values where no measurements exist, *prediction* (interpolation within the range of observed data points and extrapolation beyond that) is essential.

Predicted values, however, are not the true ones, and therefore a prediction method which minimizes the prediction error in a least-squares sense is suitable. Global and local covariance functions provide knowledge of the spatial signal characteristics in order to get optimized estimates on global or local scales.

Assuming the signal covariance function C to be known, an estimate $\Delta\hat{g}_P$ of the gravity anomaly Δg_P at point P from discrete observations Δg_i may be determined as follows:

$$\Delta\hat{g}_P = \mathbf{C}_{Pi}\mathbf{C}_{ij}^{-1}\Delta g_i \tag{3.45}$$

or, written out:

$$\Delta\hat{g}_P = \begin{bmatrix} C_{P1} & C_{P2} & \cdots & C_{Pi} \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1j} \\ C_{21} & C_{22} & \cdots & C_{2j} \\ \vdots & \vdots & & \vdots \\ C_{i1} & C_{i2} & \cdots & C_{ij} \end{bmatrix}^{-1} \begin{bmatrix} \Delta g_1 \\ \Delta g_2 \\ \vdots \\ \Delta g_i \end{bmatrix} \tag{3.46}$$

where $\mathbf{C}_{Pi}$ is the covariance matrix between predicted point P and observations i and $\mathbf{C}_{ij}$ is the covariance matrix between observations. All covariances stem from the same covariance function, only dependent on the distance between the considered points, i.e., $C_{Pi} = C(s_{Pi})$ and $C_{ij} = C(s_{ij})$ (Moritz, 1980).

Analogous, the gravity anomaly block mean values are calculable from point anomalies (or even other block mean values):

$$\Delta\hat{\bar{g}}_P = \bar{\mathbf{C}}_{Pi}\mathbf{C}_{ij}^{-1}\Delta g_i \tag{3.47}$$

In equation (3.47), $\bar{\mathbf{C}}_{Pi}$ is the covariance between point and block mean values.

The accuracy of the predicted values follow from the variance-covariance propagation law (e.g., (Moritz, 1980)):

$$\hat{\mathbf{C}}_{\Delta g_P} = \mathbf{C}_0 - \mathbf{C}_{Pi}\mathbf{C}_{ij}^{-1}\mathbf{C}_{Pi}^T \tag{3.48}$$

3.3.2 Generalization to least-squares collocation

Least-squares prediction as described in section 3.3.1 is actually a specific case of collocation, as collocation in its most general setup is combined estimation, prediction and filtering. Collocation is mainly applied to estimate heterogeneous data, e.g., different quantities of the linear gravity field such as the potential T or geoid heights N *from* heterogeneous data such as gravity anomalies or deflections of the vertical (Moritz, 1980). Decisive for collocation with regard to quality is the determination of a good covariance function as introduced in section 2.5 (Tscherning, 1983).

In section 3.3.1 the covariance function was used to compute gravity anomalies at *any* point on Earth from a known gravity anomaly data set, cf. equation (3.45).

The covariance function is determinable from a convolution integral (Moritz, 1980):

$$C(\psi) = \frac{1}{8\pi^2} \iint_{\sigma} \int_{0}^{2\pi} xy \, \mathrm{d}\alpha \, \mathrm{d}\sigma \tag{3.49}$$

where α is azimuth between gravity field signal value x and y. Integration over the surface σ ensures homogeneity, integration over α ensures isotropy. If equation (3.49) is calculated for all ψ, the covariance function is obtained.

As presented in section 2.5, the covariance function is also determinable by a spherical harmonic expansion, including the signal degree variances, cf. equation (2.47). Through the different *spectral eigenvalues*, signal degree variances of different gravity field quantities are easily obtainable, cf. equation (2.46). This holds also for the upward continuation factor λ_n^{upw}. Consequently, through (2.47), covariance functions of different gravity field quantities are just as easy to compute.

Although several gravity field functionals (e.g., gravity anomalies, gravity disturbances, geoid height, height anomaly or general upward continuation) are linked to the disturbing potential by isotropic spectral eigenvalues, many functionals (e.g., deflections of the vertical) do not have the isotropic operator link. The relatively simple spectral determination of the covariance function is not applicable in these cases, and more complex differential operations must be performed (Gerlach, 2003).

With the spectral eigenvalues, the signal degree variances between *different* gravity field functionals may also be found. A covariance function computed from such degree variances is termed a cross-covariance function. The cross-covariance function does not only describe the correlation of, e.g., gravity anomalies with other gravity anomalies, but also the correlation of, e.g., gravity anomalies and geoid height:

$$C^{N\Delta g}(\psi) = \sum_{n=2}^{\infty} \frac{r}{\gamma_0 (n-1)} c_n^{\Delta g} P_n(\cos\psi) \tag{3.50}$$

With the cross-covariance function in equation (3.50), geoid heights are determinable from gravity anomalies by a generalization of equation (3.45):

$$\hat{N}_P = \mathbf{C}_{Pi}^{N\Delta g} \mathbf{C}_{ij}^{-1} \Delta g_i \tag{3.51}$$

Equation (3.51) is an example of collocation, with the following accuracy, cf. equation (3.48):

$$\hat{\mathbf{C}}_{N_P} = \mathbf{C}_0 - \mathbf{C}_{Pi}^{N\Delta g} \mathbf{C}_{ij}^{-1} (\mathbf{C}_{Pi}^{N\Delta g})^T \tag{3.52}$$

The covariance matrix of observations, $\mathbf{C}_{ij}$, is the sum of the covariance matrices of signal and noise (Moritz, 1980), and is consequently frequently found in a split up notation. If all

measurements of an observed gravity field are assumed uncorrelated, the noise covariance matrix is diagonal.

The general scheme for computing the covariance functions (also holds for the gravity field functionals isotropically linked in the spectral domain as described above) comprises the general kernel function $K(r, r', \psi)$ (a function of the spherical coordinates of two considered points, their latitude and longitude hidden in the spherical distance ψ), containing disturbing potential covariances between two points, from which any cross-covariance function may be determined through the correct functional operator (Moritz, 1980).

For example, linear functionals of geoid height, gravity anomaly and deflections of the vertical (their relation to the disturbing potential) read:

$$N = \frac{1}{\gamma_0} T \tag{3.53}$$

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{r} T \tag{3.54}$$

$$\xi = \frac{1}{\gamma_0 r} \frac{\partial T}{\partial \theta} \tag{3.55}$$

$$\eta = -\frac{1}{\gamma_0 r \sin\theta} \frac{\partial T}{\partial \lambda} \tag{3.56}$$

The covariance and cross-covariance functions are thus obtained by applying the above functionals to the general kernel, some examples follow:

$$C^{N\Delta g} = \frac{1}{\gamma_0} \left(-\frac{\partial K}{\partial r'} - \frac{2}{r'} K \right) \tag{3.57}$$

$$C^{N\xi} = \frac{1}{\gamma_0^2 r'} \frac{\partial K}{\partial \theta'} \tag{3.58}$$

$$C^{N\eta} = -\frac{1}{\gamma_0^2 r' \sin\theta'} \frac{\partial K}{\partial \lambda'} \tag{3.59}$$

$$C^{\Delta g \Delta g} = \left(-\frac{\partial}{\partial r} - \frac{2}{r} \right) \left(-\frac{\partial K}{\partial r'} - \frac{2}{r'} K \right) \tag{3.60}$$

$$= \frac{\partial^2 K}{\partial r \partial r'} + \frac{2}{r'} \frac{\partial K}{\partial r} + \frac{2}{r} \frac{\partial K}{\partial r'} + \frac{4}{r r'} K \tag{3.61}$$

The kernel is dependent on the spherical distance, given as a function of the considered point coordinates (and in geodetic coordinates) in equation (3.41). The partial derivatives

of the kernel function can thus be formulated, e.g., as follows (spherical coordinates, colatitude and longitude):

$$\frac{\partial K}{\partial \theta'} = \frac{\partial K}{\partial \psi}\frac{\partial \psi}{\partial \theta'} \tag{3.62}$$

where formulae for computing partial derivatives such as $\frac{\partial \psi}{\partial \theta'}$ can be found in (Hofmann-Wellenhof and Moritz, 2005).

The strength of collocation is its ability to incorporate several gravity field data sources, e.g., for the computation of geoid heights from gravity anomalies and gravity disturbances one could set up the following collocation equation:

$$\hat{N}_P = \begin{bmatrix} \mathbf{C}_{P1}^{N\Delta g} \cdots \mathbf{C}_{Pi}^{N\Delta g} & \mathbf{C}_{P,i+1}^{N\delta g} \cdots \mathbf{C}_{Pj}^{N\delta g} \end{bmatrix} \begin{bmatrix} C_{11}^{\Delta g\Delta g} & \cdots & C_{1i}^{\Delta g\Delta g} & C_{1,i+1}^{\Delta g\delta g} & \cdots & C_{1j}^{\Delta g\delta g} \\ \vdots & & \vdots & \vdots & & \vdots \\ C_{i1}^{\Delta g\Delta g} & \cdots & C_{ii}^{\Delta g\Delta g} & C_{i,i+j}^{\Delta g\delta g} & \cdots & C_{ij}^{\Delta g\delta g} \\ C_{i+1,1}^{\delta g\Delta g} & \cdots & C_{i+1,i}^{\delta g\Delta g} & C_{i+1,i+1}^{\delta g\delta g} & \cdots & C_{i+1,j}^{\delta g\delta g} \\ \vdots & & \vdots & \vdots & & \vdots \\ C_{j1}^{\delta g\Delta g} & \cdots & C_{ji}^{\delta g\Delta g} & C_{j,i+1}^{\delta g\delta g} & \cdots & C_{jj}^{\delta g\delta g} \end{bmatrix}^{-1} \begin{bmatrix} \Delta g_1 \\ \vdots \\ \Delta g_i \\ \delta g_{i+1} \\ \vdots \\ \delta g_j \end{bmatrix} \tag{3.63}$$

The main obstacle in collocation is that the linear system of equations is as large as the number of observations, making the covariance matrix inversion computationally challenging.

As with the other gravity field modeling methods, the remove-restore approach is applied also when using collocation. The effect of a GGM as well as the topography is removed from the data, and in the end added back to the result. This leaves us with residual data, for which a covariance function must be determined in the area of interest. A regional gravity field approximation is then computed using collocation, we obtain both estimates of the sought gravity field quantity *as well as its error*. The obtained model can be verified by using parts of the data *not* used to compute the model in the first place.

3.3.3 A remark on collocation with parameters

A further generalization of collocation is the introduction of systematic (non-random) functional parameters to the model (Moritz, 1980). The "systematic" or "parametric" part is the well-known $\mathbf{Ax}$ found in, e.g., observation equation (3.22). The random and variable part contains the aforementioned signal and noise. This general model reads:

$$\mathbf{l} + \mathbf{v} = \mathbf{Ax} + \mathbf{Bt} \tag{3.64}$$

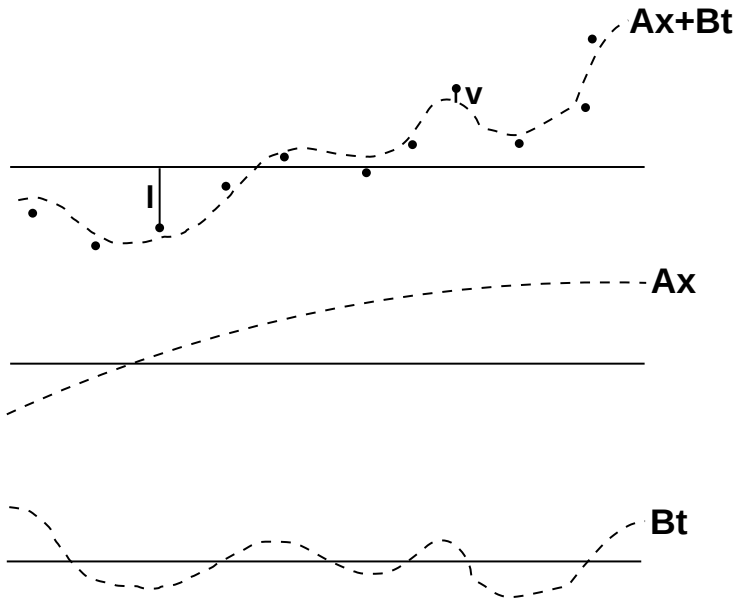

Figure 3.5: *Graphic display of the least-squares collocation model with a systematic part* $\mathbf{Ax}$ *and a random part* $\mathbf{Bt}$*, where* $\mathbf{t}$ *is a vector containing signal and noise*

where $\mathbf{l}$ are the observations with associated errors $\mathbf{v}$, $\mathbf{A}$ is a matrix accounting for the effect of the parameters $\mathbf{x}$ on the observations, the matrix $\mathbf{B}$ contains the functionals of the observations and $\mathbf{t}$ contains both signal and noise. For $\mathbf{A} = 0$, equation (3.64) reduces to $\mathbf{l} + \mathbf{v} = \mathbf{Bt}$, i.e., the observation equation for collocation without parameters (only signal and noise). For $\mathbf{B} = 0$ or $\mathbf{t} = 0$ we obtain $\mathbf{l} + \mathbf{v} = \mathbf{Ax}$, which again is the linear observation equation in usual least-squares adjustment (parameter estimation).

Equation (3.64) is visualized in Figure 3.5. $\mathbf{Ax}$ is a linearized function of parameters $\mathbf{x}$, smooth and slowly varying. The random function $\mathbf{Bt}$ contains signal and noise, oscillating about zero. The curve $\mathbf{Ax} + \mathbf{Bt}$ containing a systematic regular trend and a random part is to be determined by discrete observations $\mathbf{l}$. The noise is an observational error and thus occurs only at observation points (it is discrete), as opposed to the signal, representing continuous irregularities.

The general formulation of collocation could be applied to, e.g., *gravity measurements* (Moritz, 1980). In such a case, $\mathbf{l}$ would be the gravimeter reading, $\mathbf{v}$ a measurement error and the signal would represent the gravity anomaly. Systematic parameters $\mathbf{x}$ could be of different kinds, e.g., the parameters of the normal gravity formula as well as instrument parameters and instrument drift.

3.4 Radial basis functions

As pointed out in sections 3.1 and 3.2, spherical harmonics are global functions, and the series expansion has to be done up to a very high degree to provide the high-frequent information of a considered signal, e.g., the gravity field. We have seen that more locally dependent point-based integral formulae may instead account for that part in addition to their collocation alternative (cf. section 3.3.2).

Another approach is given through a different *choice of base functions* than spherical harmonics. *Radial basis functions* (RBFs) are isotropic (radial-symmetric) functions that store most of their energy in a limited spatial area, i.e., they have the ability to localize in space (Wittwer, 2009). Radial basis functions may have quasi-local support, in that they in spite of their *space*-localizing abilities represent global functions (they are not strictly zero outside their support area, cf. the oscillations in the space domain in, e.g., Figure 3.10). With the RBF approach, we do not only have a vast variety of possible RBFs to choose from, but also have to take their *bandwidth* and *network design* (i.e., a *grid* for positioning the RBFs must be chosen) into account.

RBFs have space-localizing properties that spherical harmonics lack, and consequently, these bases may be assumed more efficient in modeling spatially dependent higher frequencies of a considered signal. The spherical harmonic Stokes coefficients F_{nm} are global parameters, since they do not depend on spatial position, but have optimal *frequency* localization because of their direct relation to the frequency values degree n and order m (Schmidt, 2001). The coefficients of the considered radial basis functions are spatially dependent parameters, since they are a function of the *position vector*! However, the frequency localization is worse than for spherical harmonics, since each coefficient is related to a *frequency band*.

Localization refers to the area in the frequency or space domain in which the function does not vanish, and the function localizes better as this area shrinks (Eicker, 2008). It is actually not possible to have *both ideal frequency and space localization at the same time*, due to *Heisenberg's uncertainty principle*. The better a function localizes in one domain, the worse it localizes in the other.

So-called *Dirac functions* present the direct opposite of spherical harmonics in that they are different from zero at only one single point, i.e., they have optimal space-localizing properties (Bentel et al., 2013). However, frequency localization is non-existent. RBFs can be seen as a compromise between these two extremes, and are consequently very adjustable.

In addition to conventional evaluation of convolution integrals, as found in equation (3.5), by numerical integration, *parameter estimation methods* may be applied, utilizing statistical models.

There is in principle a vast amount of possible RBFs to choose from, as long as they are harmonic kernel functions. Numerical simulations using different RBFs (i.e., varying kernel coefficients B_n in equation (3.74)) under the same conditions were conducted by

Bentel et al. (2013). As pointed out earlier, a good compromise between localization in the frequency domain and the space domain is sought. The signal f has been considered *non band-limited* in the above equations, but in practice it may be *band-limited*, i.e., truncated at some maximum degree N. Bentel et al. (2013) concluded that the frequency localization with respect to the signal bandwidth plays a minor role for a *band-* (or frequency-) limited RBF. For good results in regional gravity field modelling, the spatial behavior of the RBF is of more importance.

An example of a band-limited kernel function is when frequencies higher than a chosen maximum degree N are omitted. The convolution of the signal with such a kernel function will then act as a *low-pass filter*. If a frequency *window* is chosen for the kernel function, the convolution of the signal with the kernel function will act as a *band-pass filter*. Tailoring of the filtering properties of a kernel is done through the choice of kernel coefficients B_n (Bentel et al., 2013).

A general *kernel* (or weight) function may be formulated as:

$$B(\mathbf{x}, \mathbf{x}_P) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi R^2} \left(\frac{R}{r}\right)^{n+1} B_n P_n(\mathbf{r}^T \mathbf{r}_P) \tag{3.65}$$

where the *Legendre* (kernel) coefficients B_n define the kernel and reflect its frequency behavior (Bentel, 2010). The distance between the origin of the RBF on the sphere $\mathbf{x}_P$ and the computation point $\mathbf{x}$, which may be in the exterior of the sphere, is the only variable in the kernel (it is isotropic). For $\mathbf{x}_P = \mathbf{x}$ the kernel is at its maximum.

There is no single conclusive way of selecting the bandwidth of a RBF, but it must be chosen properly so that the RBFs approximate the signal in an optimal way (Wittwer, 2009). Both the spatial and frequency behavior of the RBF must be observed. If the bandwidth is too narrow, the RBFs will approximate the data well, but the quality of the solution between data points will be poor, and a denser grid of RBFs is needed. Conversely, if the bandwidth is chosen too large, the RBFs overlap too much (become too similar) and the normal equation system will be singular and an exaggerated smoothing of the solution may occur. The RBFs will not represent the frequency behavior of the signal very well in such a configuration. *Leakage* of the signal outside the considered area is more of a spatial effect, due to the RBFs not being strictly spatially-localizing but rather band-limited in the frequency domain.

Different methods exist for choosing the RBF bandwidth (Wittwer, 2009). One approach for choosing the bandwidth is to assume the same frequency behavior of the RBFs as the signal (Eicker, 2008). The norm of a signal as presented in equation (3.8) can be interpreted as the energy content or frequency behavior of the signal (Schmidt, 2001). The Legendre coefficients B_n defines the frequency behavior of the kernel, and the norm of $B(\mathbf{x}, \mathbf{x}_P)$ may be written as follows:

$$\|B(\mathbf{x}, \mathbf{x}_P)\|^2 = \sum_{n=0}^{\infty} c_n^2 \tag{3.66}$$

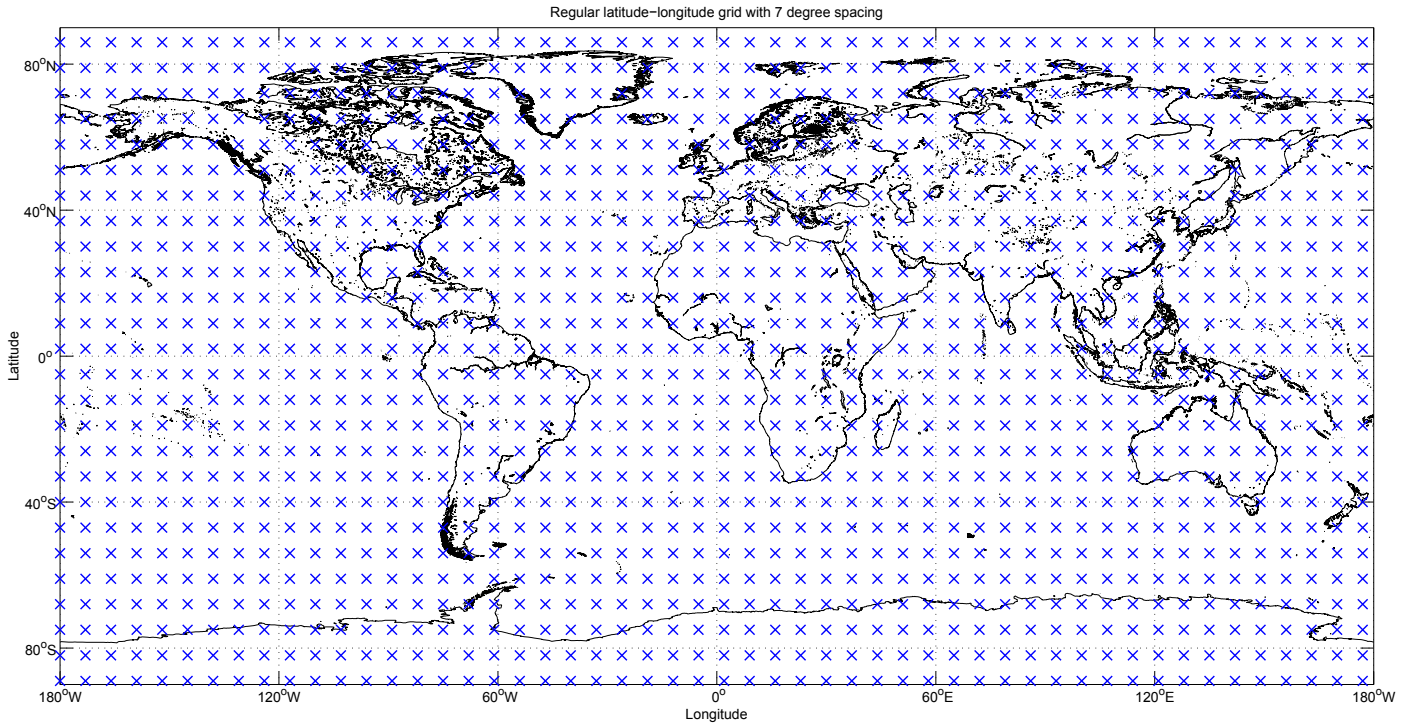

Figure 3.6: *A regular latitude-longitude grid*

where c_n^2 are the signal degree variances of degree n:

$$c_n^2 = \frac{2n+1}{4\pi R^2}|B_n|^2 \tag{3.67}$$

Thus the RBFs may be constructed such that their energy content corresponds to the energy content of the signal.

Another important aspect of regional gravity field modeling are the *boundary effects* that occur in the boundary of the considered area or region (Wittwer, 2009). If observations outside the target area are available, these can be used to reduce boundary effects by extending the data grid to some degrees outside the target or computation area.

The RBFs are dependent on position, and consequently they have to be placed at chosen *locations*. Grid type and associated spacing or level must be chosen such that under- or over-parametrization is avoided (Wittwer, 2009). A number of potentially applicable *point grid systems* exist, the easiest one probably being the equiangular *latitude-longitude* grid with a regular distance (grid spacing) between the latitude and longitude values of each point, cf. figures 3.6 and 3.7 (Bentel, 2010).

Another option is the *Reuter* grid, where the spherical distance between the grid points is almost constant, making it an equidistributed point system, cf. figures 3.8 and 3.9

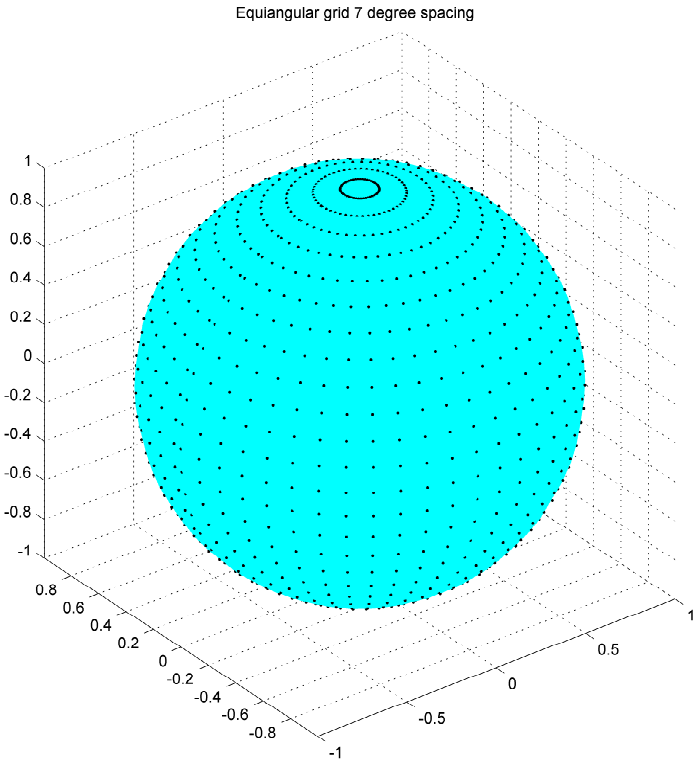

Figure 3.7: *A regular latitude-longitude grid with 7° spacing on the unit sphere*

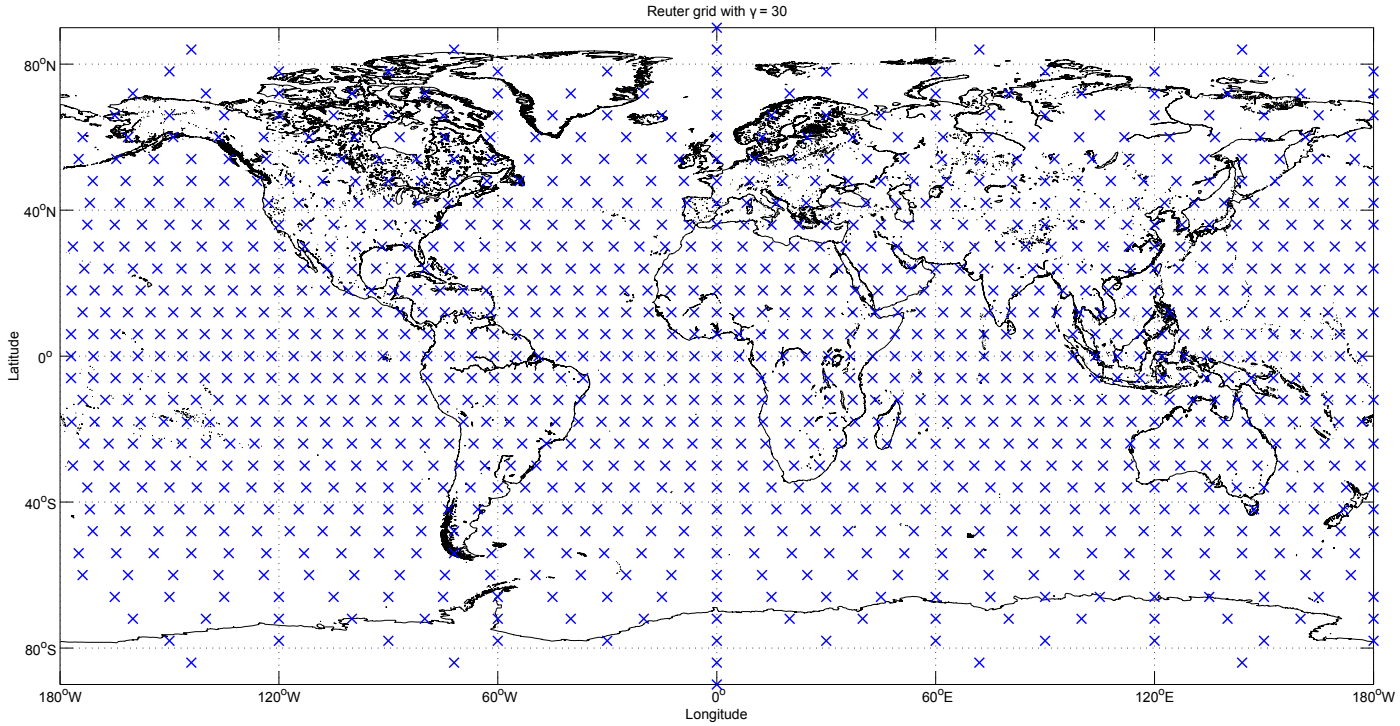

Figure 3.8: *A Reuter grid with* $\gamma = 30$

(Wittwer, 2009). A Reuter grid defined by points expressed in spherical coordinates (θ, λ) for a chosen level γ (to define the number of points) may be constructed as follows:

$$\theta_0 = 0, \lambda_{0,1} = 0 \qquad \text{North Pole} \tag{3.68}$$

$$\Delta\theta = \pi/\gamma \tag{3.69}$$

$$\theta_i = i\Delta\theta, \qquad 1 \leq i \leq \gamma - 1 \tag{3.70}$$

$$\gamma_i = 2\pi / \cos\left((\cos\Delta\theta - (\cos\theta_i)^2)/(\sin\theta_i)^2\right)^{-1} \tag{3.71}$$

$$\lambda_{i,j} = \left(j - \frac{1}{2}\right)(2\pi/\gamma_i), \qquad 1 \leq j \leq \gamma_i \tag{3.72}$$

$$\theta_0 = \pi, \lambda_{\gamma,1} = 0 \qquad \text{South Pole} \tag{3.73}$$

Wittwer (2009) showed that $\gamma \approx N_{\max}$, where $N_{\max}$ is the maximum spherical harmonic degree that can be modeled by RBFs placed in such a grid.

K RBFs are placed at different locations $\mathbf{x}_k$ in a grid system with K points. If the unnormalized Legendre polynomials P_n are used, the intensity values (y-axis in the 2D case, see, e.g., Figure 3.10) are relatively large. However, normalization may be done at a later stage, e.g., in the normalization of the design matrix $\mathbf{A}$.

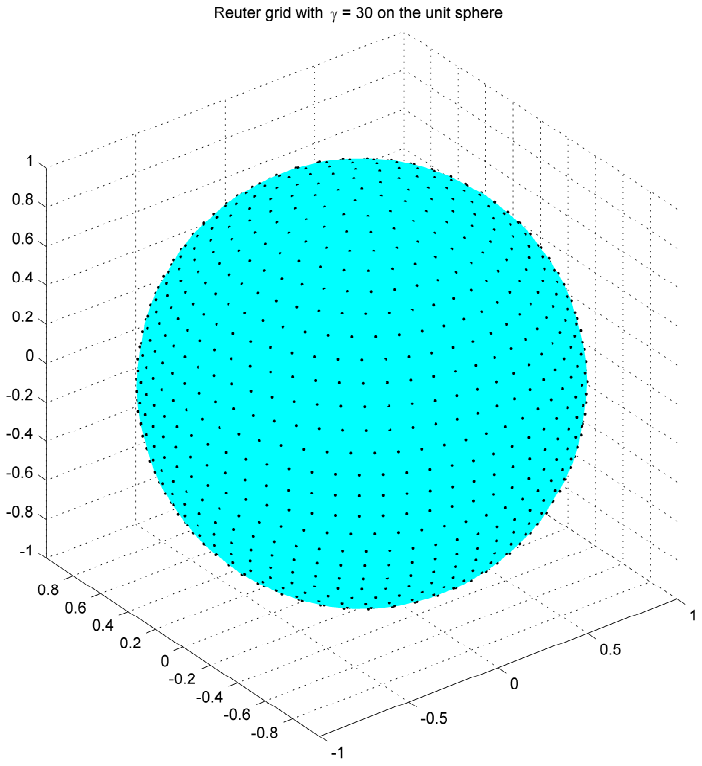

Figure 3.9: *A reuter grid with $\gamma = 30$ on the unit sphere. Note the equidistribution of points.*

The number of RBFs depends only on the resolution achievable with the input data and the size of the target area, and therefore the number of observations does not directly influence the number of RBFs. Wittwer (2009) presents a rule of thumb, stating that "choosing a grid with an amount of basis functions equal to one third the number of observations is a good starting point".

$\mathbf{x}_k$ is the origin of RBF number k, $\mathbf{x}$ is the computation point. If $R = r = 1$, both RBFs and computation points are defined on the unit sphere and equation (3.65) reads:

$$B(\mathbf{x}, \mathbf{x}_k) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} B_n P_n(\mathbf{r}^T \mathbf{r}_k) \tag{3.74}$$

3.4.1 Gravity field functionals in terms of radial basis functions

The above kernel expressions are quite general, in that they hold for *any* signal outside or on the sphere. The direct link to spherical harmonics, considered in detail in section 3.5.3, allows us to model different isotropic gravity field quantities by RBFs in the same manner as it is done by spherical harmonics (Bentel, 2010).

The disturbing potential, for instance, is (in theory) a non band-limited harmonic function that could be represented as an infinite sum of RBF as follows (Wittwer, 2009):

$$T(\mathbf{x}) = \frac{GM}{R} \sum_{k=0}^{\infty} d_k \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \left(\frac{R}{r}\right)^{n+1} B_n P_n(\mathbf{r}^T \mathbf{r}_k) \tag{3.75}$$

where d_k are the RBF coefficients, determinable by, e.g., least-squares adjustment,. This is described in section 3.4.3. In the practical evaluation of equation (3.75), of course the disturbing potential signal is band-limited ($N_{\max} < \infty$), and the number of RBFs k is finite (Wittwer, 2009).

From the RBF expression for the disturbing potential (equation (3.75)) other gravity field functionals may be derived (Wittwer, 2009).

For gravity anomalies, for instance, we set up the well-known link between gravity anomalies and the disturbing potential, i.e., the fundamental equation in spherical approximation:

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{R} T \tag{3.76}$$

For the gravity anomaly representation by a (theoretical) infinite set of RBFs, we get:

$$\Delta g(\mathbf{x}) = -\frac{\partial}{\partial r}T(\mathbf{x}) - \frac{2}{R}T(\mathbf{x}) = \frac{GM}{R^2}\sum_{k=0}^{\infty} d_k \sum_{n=2}^{\infty} \frac{2n+1}{4\pi R^2}(n-1)\left(\frac{R}{r}\right)^{n+1} B_n P_n(\mathbf{r}^T\mathbf{r}_k) \quad (3.77)$$

3.4.2 Examples of radial basis functions

A very simple radial basis function is the *Shannon low-pass function* (Bentel et al., 2013), where all kernel coefficients up to a degree N are set to one and all above are set to zero, cf. figures 3.10 and 3.11. It may be formulated as follows:

$$B_n = 1 \quad \forall \quad n \in [0\,,\, N] \quad (3.78)$$

Further, the *Shannon band-pass function* covers an exact spectral window $[n_1\,,\, n_2]$, leading to larger oscillations in the spatial domain and a negative effect on regional gravity field modelling (Bentel et al., 2013):

$$B_n = \begin{cases} 1 & \forall \quad n \in [n_1\,,\, n_2] \\ 0 & \text{else.} \end{cases} \quad (3.79)$$

Another radial basis function is the *Blackman low-pass function*, cf. figures 3.12 and 3.13. When compared with the Shannon RBF, a smoother drop is observed in the frequency domain. Because of this smoothing, oscillations in the space domain are reduced. The RBFs are global functions, so strong oscillations outside the observation area lead to data leakage (Bentel et al., 2013). The Blackman RBF is adjustable through the frequency parameters n_1 and n_2 defining the "drop window":

$$B_n = \begin{cases} 1 & \text{for} \quad n < n_1 \\ (A\,(n))^2 & \text{for} \quad n = n_1, \ldots, n_2 \\ 0 & \text{for} \quad n > n_2 \end{cases} \quad (3.80)$$

where

$$A(n) = 0.42 - 0.5\cos\left(\frac{2\pi(n-n_2)}{2(n_2-n_1)}\right) + 0.08\cos\left(\frac{4\pi(n-n_2)}{2(n_2-n_1)}\right) \quad (3.81)$$

The kernel coefficients of the *Cubic polynomial* radial basis function are given by a cubic polynomial equal to one at degree zero and equal to zero at maximum degree $n_{\max}$, cf. figures 3.14 and 3.15:

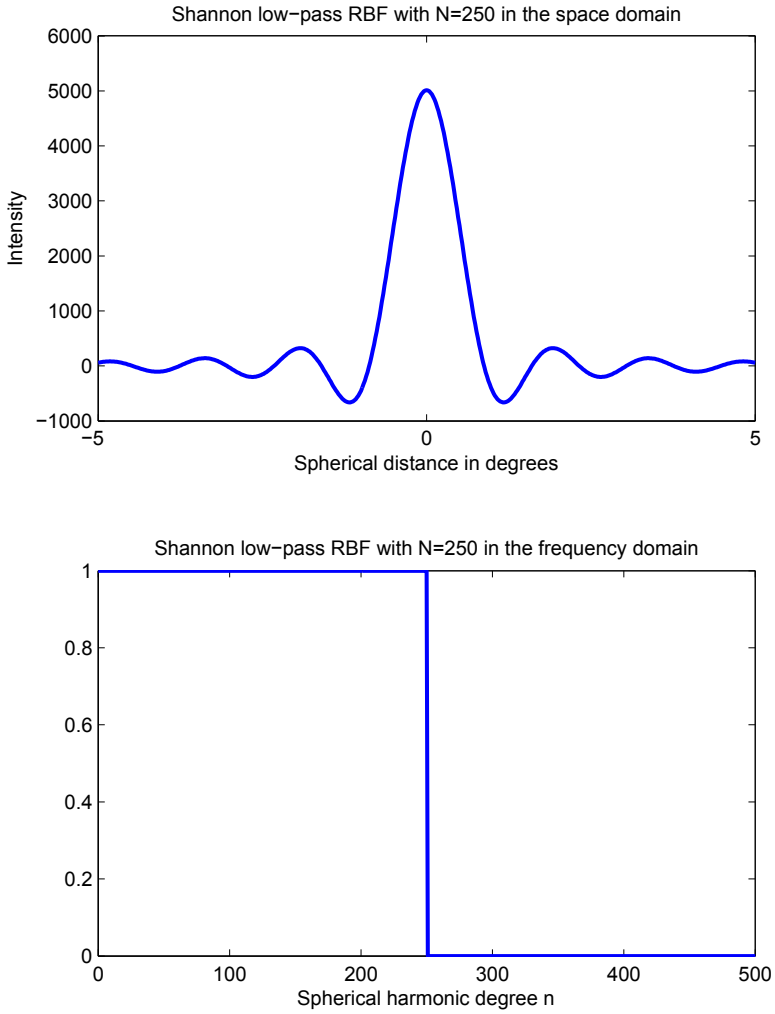

Figure 3.10: *The Shannon low-pass radial basis function with* $N = 250$ *in the space and frequency domains.*

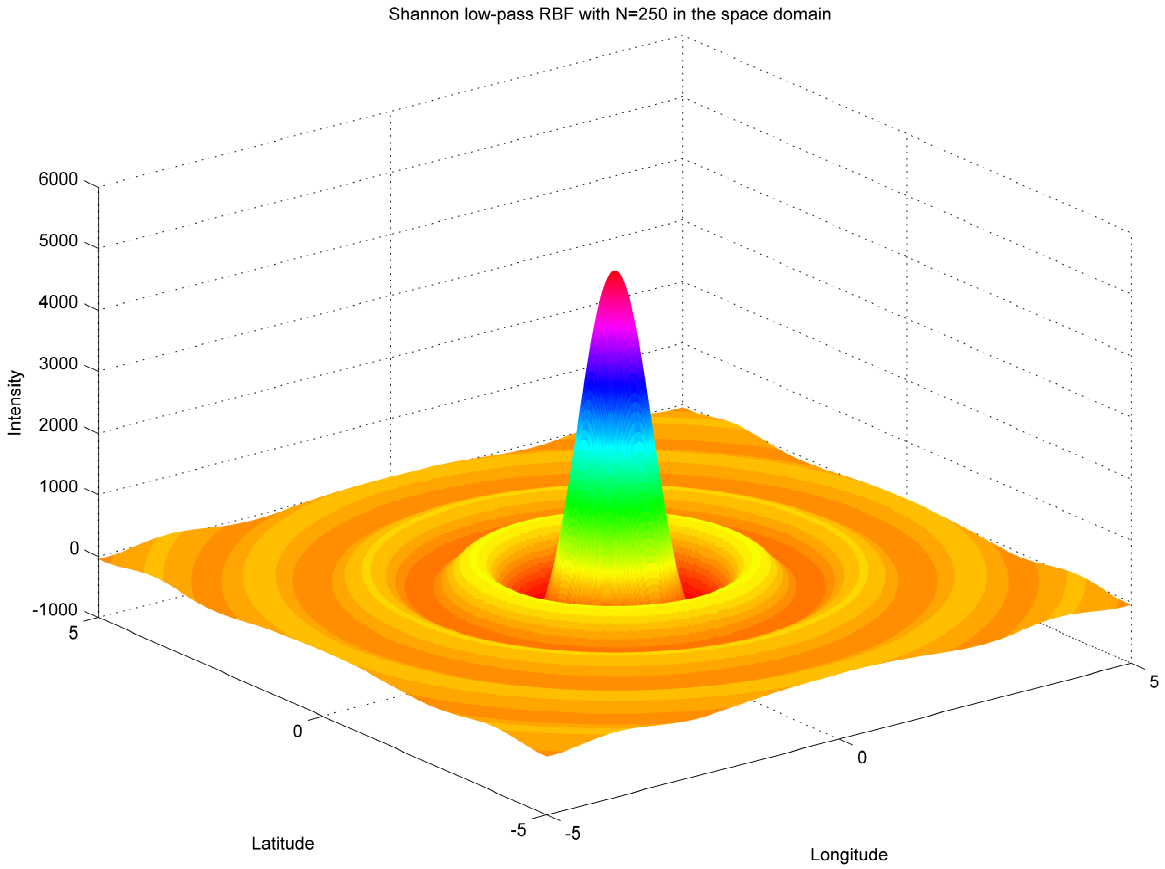

Figure 3.11: *The Shannon low-pass radial basis function with* $N = 250$ *in 3D.*

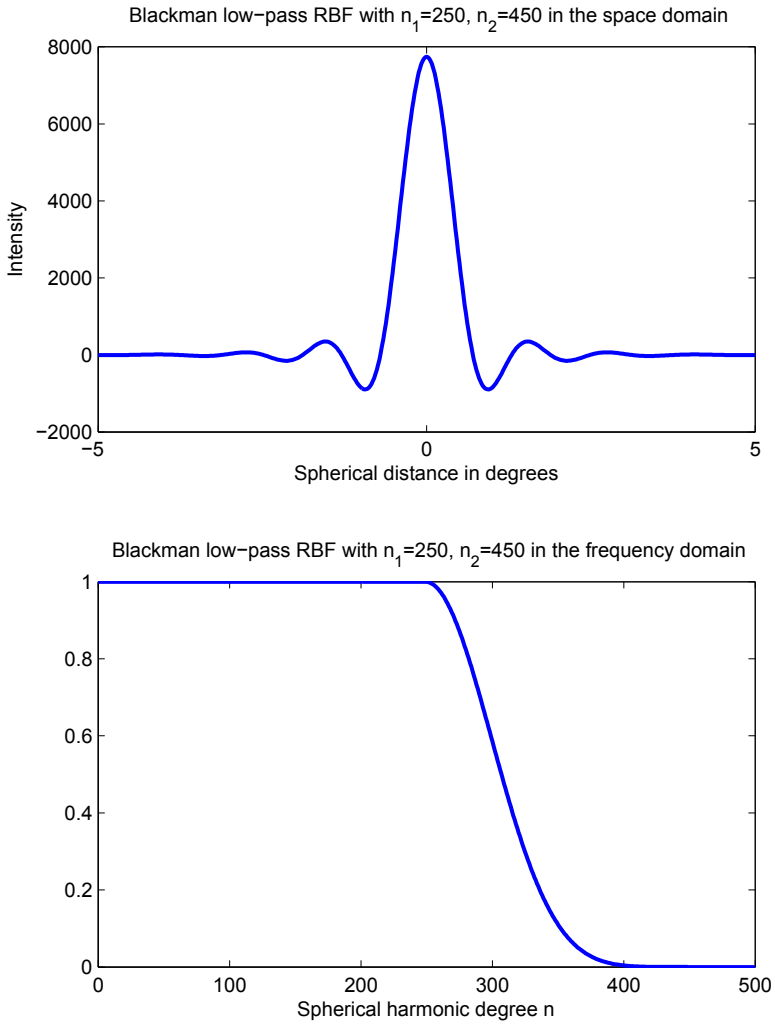

Figure 3.12: *The Blackman low-pass radial basis function with* $n_1 = 250$ *and* $n_2 = 450$ *in the space and frequency domains.*

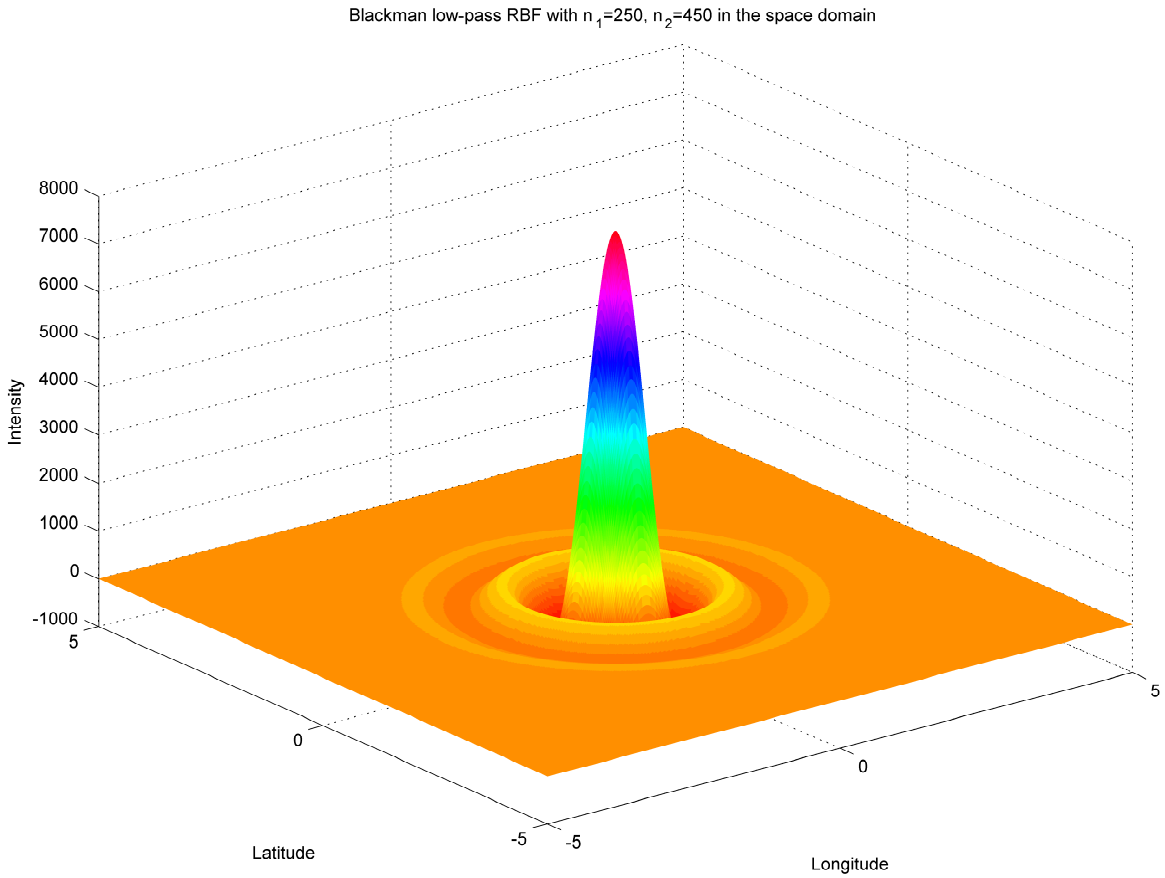

Figure 3.13: *The Blackman low-pass radial basis function with* $n_1 = 250$ *and* $n_2 = 450$ *in 3D.*

$$B_n = \left(1 - \frac{n}{n_{\max}}\right)^2 \left(\frac{2n}{n_{\max}} + 1\right) \tag{3.82}$$

3.4.3 Radial basis function analysis by least-squares adjustment

We may solve for the RBF coefficients d_k by least-squares adjustment (Bentel et al., 2013). Thereby only a finite number of coefficients can be estimated, i.e.:

$$f(\mathbf{x}) = \sum_{k=1}^{\infty} d_k B(\mathbf{x}, \mathbf{x}_k) \approx \sum_{k=1}^{K} d_k B(\mathbf{x}, \mathbf{x}_k) \tag{3.83}$$

The coefficients d_k constitute the RBF part which represents the signal and thus play a similar role as the Stokes coefficients in the spherical harmonic approach.

The low-frequency part of the gravity signal is modeled very well by global spherical harmonic base functions. For the high-frequency part, radial basis functions are better suited. Thus a combination of spherical harmonics and radial basis functions seems reasonable. The gravity signal up to maximum degree $N_{\max}$ is represented by spherical harmonics, while the regional or local features found in the residual gravity signal Δf is modeled with radial basis functions (Bentel et al., 2013):

$$f(\mathbf{x}) = \sum_{n=0}^{N_{\max}} \sum_{m=-n}^{n} F_{nm} Y_{nm}(\mathbf{x}) + \Delta f(\mathbf{x}) \tag{3.84}$$

An observation equation based on equation (3.83) may be formulated:

$$\Delta f(\mathbf{x}) + v(\mathbf{x}) = \sum_{k=1}^{K} d_k B(\mathbf{x}, \mathbf{x}_k) \tag{3.85}$$

In matrix notation, equation (3.85) reads:

$$\Delta \mathbf{f} + \mathbf{v} = \mathbf{A}\mathbf{d} \tag{3.86}$$

or, written out:

$$\begin{bmatrix} \Delta f_1(\mathbf{x}_1^1) \\ \vdots \\ \Delta f_1(\mathbf{x}_i^1) \\ \Delta f_2(\mathbf{x}_j^1) \\ \vdots \\ \Delta f_2(\mathbf{x}_L^1) \end{bmatrix} + \begin{bmatrix} v_1(\mathbf{x}_1^1) \\ \vdots \\ v_1(\mathbf{x}_i^1) \\ v_2(\mathbf{x}_j^1) \\ \vdots \\ v_2(\mathbf{x}_L^1) \end{bmatrix} = \begin{bmatrix} B_1(\mathbf{x}_1^1, \mathbf{x}_1) & \cdots & B_1(\mathbf{x}_1^1, \mathbf{x}_K) \\ \vdots & \ddots & \vdots \\ B_1(\mathbf{x}_i^1, \mathbf{x}_1) & \cdots & B_1(\mathbf{x}_i^1, \mathbf{x}_K) \\ B_2(\mathbf{x}_j^1, \mathbf{x}_1) & \cdots & B_2(\mathbf{x}_j^1, \mathbf{x}_K) \\ \vdots & \ddots & \vdots \\ B_2(\mathbf{x}_L^1, \mathbf{x}_1) & \cdots & B_2(\mathbf{x}_L^1, \mathbf{x}_K) \end{bmatrix} \cdot \begin{bmatrix} d_1 \\ \vdots \\ d_K \end{bmatrix} \tag{3.87}$$

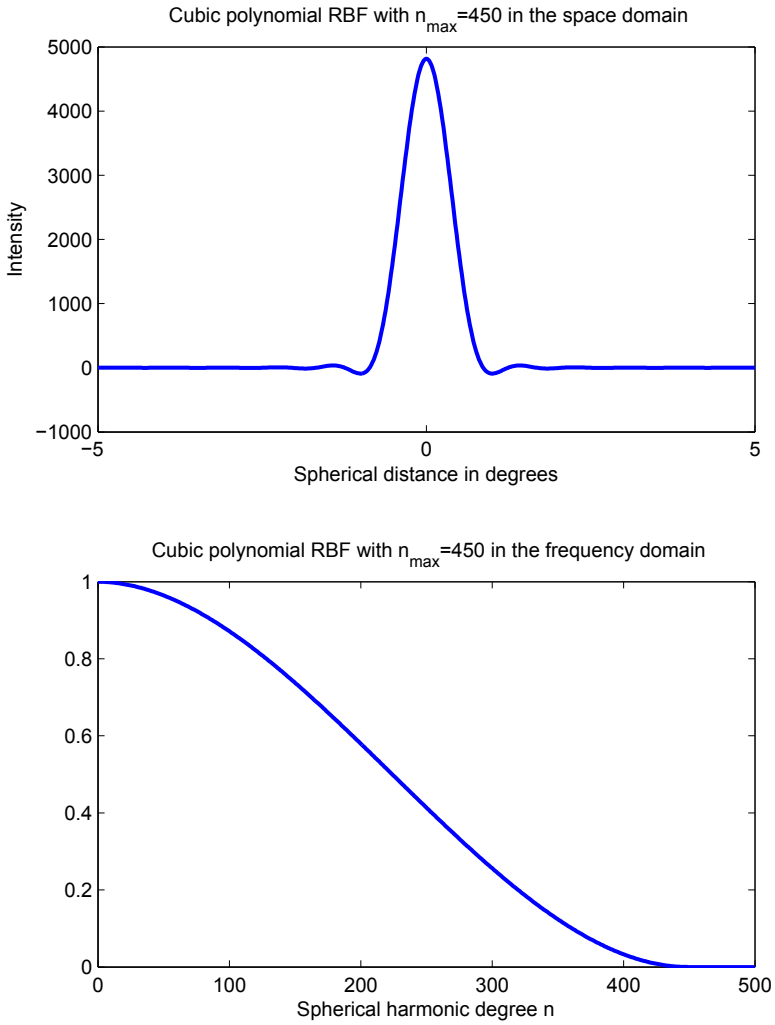

Figure 3.14: *The Cubic polynomial radial basis function with* $n_{max} = 450$ *in the space and frequency domains.*

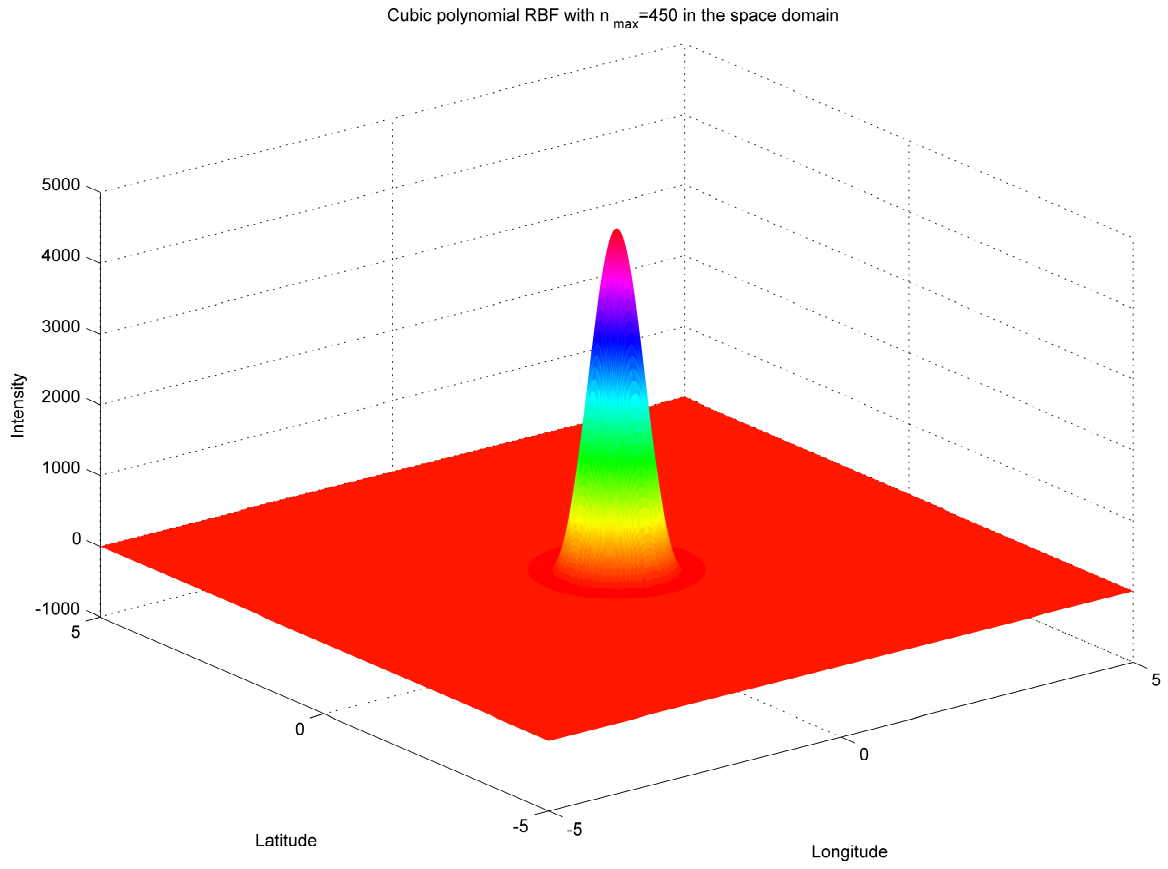

Figure 3.15: *The Cubic polynomial radial basis function with* $n_{max} = 450$ *in 3D.*

Signal frequencies higher than the maximum degree of the kernel are not modeled, but projected into the error vector $\mathbf{v}$ (Bentel et al., 2013). Equation (3.87) also opens up for two (or more) different observation types (e.g., satellite and terrestrial gravity observations), indicated by indices 1 and 2. Superscripted $\mathbf{x}^1$ denote the computation points, while $\mathbf{x}$ without superscript are the RBF origins. In the case of satellite and terrestrial gravity observation types, the kernels B_1 and B_2 would be chosen to cover the respective observation bandwidths.

Further, equation (3.87) is presumably singular, one reason being that *downward continuation* is an ill-posed problem by itself, cf. section 5.1. In addition, normally more RBFs than strictly needed are used and observation sets may contain data gaps. The singularity of the normal equation matrix $\mathbf{N}$ can be overcome by regularization, cf. section 5.2.

3.4.4 From radial basis functions to wavelets

In the above combination approach a lower resolution spherical harmonic model is combined with a regionally refined higher resolution model based on radial basis functions. This may be seen as a *two-scale* approach, and seems applicable to, e.g., a combination of satellite and regionally restricted terrestrial data (Eicker, 2008). However, when for instance both satellite and terrestrial gravity data is considered, the resolution span is so wide that a decomposition of the signal into different *detail* signals might be of interest, i.e., an extension to several scales, which is known as multi-*scale* or multi-*resolution* representation (MRR).

A wavelet representation may be applied in a MRR, since wavelets may be used to decompose a signal into detail signals on different scales, dependent on frequency (Eicker, 2008). In a wavelet representation, the basis functions are made up of both *scaling functions* (radial basis functions) and *wavelet functions*, each wavelet function being the difference between two scaling functions. When wavelets are used for modeling a function on the sphere, they are known as *spherical wavelets*.

At the *German Geodetic Research Institute* (DGFI) in Munich, Germany, regional gravity field modeling incorporating MRR has been investigated through several years, using, e.g., the Shannon and Blackman scaling functions.

3.5 Model comparison

The goal of this section is to show that the different mathematical models for gravity field representation are equal, at least in the global case.

3.5.1 Least-squares collocation versus Stokes integration

In the integral formulae, linear operations are performed on the data directly (Moritz, 1980). This is often solved by numerical integration techniques, involving interpolation for data gridding. In collocation, linear operations are performed on a *kernel function*, and may be solved analytically. Collocation operations are inverse to integral formula operations, e.g., we deal with differentiation instead of integration. Consequently, while inversion is done analytically in the integral formulae, it must be done numerically in collocation, i.e., by inverting the covariance matrix. Also, collocation can handle heterogeneous data as opposed to the integral formulae, handling data of only one type.

However, de Min (1995) showed that Stokes integration and collocation are indeed equivalent methods in the global case. For this comparison purpose, an integration point $P(\varphi_P, \lambda_P)$ is introduced, as well as gravity anomalies at points i for prediction of gravity in any point (or block) Q, cf. section 3.3.1. Stokes's formula and the least-squares prediction formula are presented again:

$$N(P) = \frac{R}{4\pi\gamma} \iint_{\sigma} S(\psi_{PQ}) \Delta g(Q) \, \mathrm{d}\sigma \tag{3.88}$$

$$\Delta g(Q) = \mathbf{C}_{Qi} \mathbf{C}_{ij}^{-1} \Delta g_i \tag{3.89}$$

Given an area of i gravity anomalies Δg_i, gravity anomalies $\Delta g(Q)$ can be predicted in *any point on Earth* Q. It is further assumed that $\Delta g(Q)$ are given for *all points on Earth*, and may be inserted into Stokes's formula as follows:

$$N(P) = \frac{R}{4\pi\gamma} \iint_{\sigma} S(\psi_{PQ}) \Delta g(Q) \, \mathrm{d}\sigma \tag{3.90}$$

$$= \frac{R}{4\pi\gamma} \iint_{\sigma} S(\psi_{PQ}) \mathbf{C}_{Qi} \mathbf{C}_{ij}^{-1} \Delta g_i \, \mathrm{d}\sigma \tag{3.91}$$

$$= \frac{R}{4\pi\gamma} \iint_{\sigma} S(\psi_{PQ}) \mathbf{C}_{Qi} \, \mathrm{d}\sigma \, \mathbf{C}_{ij}^{-1} \Delta g_i \tag{3.92}$$

Since integration is performed globally, a spherical harmonic expansion of the integral in equation (3.92) may be formulated and the decomposition formula applied:

$$\frac{R}{4\pi\gamma}\iint_{\sigma} S(\psi_{PQ})\mathbf{C}_{Qi}\,\mathrm{d}\sigma \tag{3.93}$$

$$= \frac{R}{4\pi\gamma}\iint_{\sigma}\sum_{n=2}^{\infty}\frac{2n+1}{n-1}\bar{P}_n(\cos\psi_{PQ})\sum_{k=2}^{\infty}c_k P_k(\cos\psi_{Qi})\,\mathrm{d}\sigma \tag{3.94}$$

$$= \sum_{n=2}^{\infty}\frac{R}{\gamma(n-1)}c_n P_n(\cos\psi_{Pi}) \tag{3.95}$$

$$= \mathbf{C}^{N\Delta g}(\psi_{Pi}) \tag{3.96}$$

where c_k or rather c_n are the gravity anomaly signal degree variances, as described in section 2.5. Equation (3.96) is immediately recognized as the cross-covariance function presented in equation (3.50). Thus, the expression for $N(P)$ in equation (3.92) becomes

$$N(P) = \mathbf{C}_{Pi}^{N\Delta g}\mathbf{C}_{ij}^{-1}\Delta g_i \tag{3.97}$$

which is the collocation formula, cf. section 3.3.2. To summarize, least-squares collocation performs least-squares prediction over the entire Earth, and applies Stokes's formula to the globally continuous Δg function.

If, as normally is the case, we use Stokes's formula in a regional application, the equivalence is not immediately clear. In practice, Earth is split up into a near zone containing gravity anomaly measurements, and a distant zone accounted for by a GGM. Stokes's formula is applied to a *residual gravity signal* Δg^{RES} in the near zone (where the low-frequent gravity signal from a GGM has been subtracted from the original gravity anomaly, cf. section 4.2). In this case, least-squares prediction is performed over the entire Earth using Δg^{RES}, which means an interpolation in the near zone — but also an extrapolation in the distant zone.

The extrapolation in the distant zone is actually not wanted in the practical approach, since the covariance function used is not necessarily valid outside the original gravity anomaly data area. de Min (1995) suggests to modify the covariance function such that the extrapolation is not taken into account. The topic of kernel modifications, of which the following is a part, is properly introduced and treated in detail in sections 4.2 and 4.2.5.

We go back to equation (3.92), but integrate only in a *near zone* σ_0, defined by a spherical cap with radius ψ_0 around computation point P. By *modifying* Stokes's function (modifying the kernel), we can still write equation (3.92) as a global integral:

$$N(P) = \frac{R}{4\pi\gamma}\iint_{\sigma} S^L(\psi_{PQ})\mathbf{C}_{Qi}\,\mathrm{d}\sigma\,\mathbf{C}_{ij}^{-1}\Delta g_i \tag{3.98}$$

where

$$S^L(\psi_{PQ}) = \begin{cases} S(\psi_{PQ}), & 0^\circ \leq \psi < \psi_0 \\ 0, & \psi_0 \leq \psi \leq \pi \end{cases} \tag{3.99}$$

Equation (3.99) has the following spectral representation:

$$S^L(\psi_{PQ}) = \sum_{n=2}^{\infty} \left(\frac{2}{n-1} - Q_n(\psi_0) \right) \frac{2n+1}{2} P_n(\cos\psi) \tag{3.100}$$

where $Q_n(\psi_0)$ are *Molodensky's truncation coefficients*. With this modification, equation (3.97) may be written as follows:

$$N(P) = \mathbf{C}_{Pi}^{N^L \Delta g} \mathbf{C}_{ij}^{-1} \Delta g_i \tag{3.101}$$

with

$$\mathbf{C}^{N^L \Delta g} = \sum_{n=2}^{\infty} \frac{R}{2\gamma} \left(\frac{2}{n-1} - Q_n(\psi_0) \right) c_n P_n(\cos\psi) \tag{3.102}$$

Thus, the collocation formula, expressed with the modified covariance function, still performs least-squares interpolation and extrapolation over the entire Earth using the gravity values in the near zone, but Stokes's formula is only applied to the continuous gravity function up to integration radius ψ_0 from the computation point (i.e., only the near zone). With this modification, collocation and Stokes integration should be identical (de Min, 1995).

3.5.2 Stokes integration versus spherical harmonics

We may start with the spherical approximation of the fundamental equation of physical geodesy, relating the disturbing potential T to gravity anomalies Δg, cf. equation (2.33). That is, Earth's flattening is neglected and a mean Earth radius R is introduced. Repeated here for convenience, the fundamental equation reads:

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{R} T \tag{3.103}$$

Generally, outside the boundary sphere, we may formulate the spherical harmonic expansion of the disturbing potential T as follows:

$$T(r,\theta,\lambda) = \sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+1}\sum_{m=0}^{n}\left[\Delta\bar{C}_{nm}(T)\cos m\lambda + \bar{S}_{nm}(T)\sin m\lambda\right]\bar{P}_{nm}(\cos\theta) \tag{3.104}$$

Inserting equation (3.104) into equation (3.103) yields for Δg:

$$\Delta g(\theta,\lambda) = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\left[\Delta\bar{C}_{nm}(\Delta g)\cos m\lambda + \bar{S}_{nm}(\Delta g)\sin m\lambda\right]\bar{P}_{nm}(\cos\theta) \tag{3.105}$$

Spectral eigenvalues and dimensioning factors of T and Δg are part of the respective coefficients. Applying the abbreviations $\bar{\mathcal{R}}_{nm}$ and $\bar{\mathcal{S}}_{nm}$ introduced in equation (3.10), the coefficients $\Delta\bar{C}_{nm}(\Delta g)$ and $\bar{S}_{nm}(\Delta g)$ are given by:

$$\begin{Bmatrix}\Delta\bar{C}_{nm}(\Delta g)\\ \bar{S}_{nm}(\Delta g)\end{Bmatrix} = \frac{1}{4\pi}\iint_{\sigma}\Delta g(\theta',\lambda')\begin{Bmatrix}\bar{\mathcal{R}}_{nm}(\theta',\lambda')\\ \bar{\mathcal{S}}_{nm}(\theta',\lambda')\end{Bmatrix} \tag{3.106}$$

where (θ,λ) is our point of interest, and (θ',λ') an arbitrary point.

By inspection of equation (3.104) and equation (3.105) we observe that they are equal, albeit the coefficients. Thus, if a relation between the coefficients is found, the geodetic boundary value problem (equation (3.103)) is solved in terms of Δg. Equation (3.104) and equation (3.105) are inserted into equation (3.103):

$$\begin{Bmatrix}\Delta\bar{C}_{nm}(\Delta g)\\ \bar{S}_{nm}(\Delta g)\end{Bmatrix} = -\frac{2}{R}\begin{Bmatrix}\Delta\bar{C}_{nm}(T)\\ \bar{S}_{nm}(T)\end{Bmatrix} - \frac{\partial}{\partial r}\begin{Bmatrix}\Delta\bar{C}_{nm}(T)\\ \bar{S}_{nm}(T)\end{Bmatrix}, \qquad n = 0,\ldots,\infty \text{ and } m = 0,\ldots,n \tag{3.107}$$

where the derivative of the radial term on the surface of the sphere gives:

$$\frac{\partial}{\partial r}\left(\frac{R}{r}\right)^{n+1} = -\left(\frac{R}{r}\right)^{n+1}\frac{n+1}{r} = -\frac{n+1}{R} \tag{3.108}$$

The relation between T and Δg becomes:

$$\begin{Bmatrix}\Delta\bar{C}_{nm}(T)\\ \bar{S}_{nm}(T)\end{Bmatrix} = \frac{R}{n-1}\begin{Bmatrix}\Delta\bar{C}_{nm}(\Delta g)\\ \bar{S}_{nm}(\Delta g)\end{Bmatrix} \tag{3.109}$$

Thus, for T in terms of Δg:

$$T(\theta,\lambda) = \sum_{n=2}^{\infty}\frac{R}{n-1}\sum_{m=0}^{n}\left[\Delta\bar{C}_{nm}(\Delta g)\bar{\mathcal{R}}_{nm}(\theta,\lambda) + \bar{S}_{nm}(\Delta g)\bar{\mathcal{S}}_{nm}(\theta,\lambda)\right] \tag{3.110}$$

The surface integration expression for the coefficients (equation (3.106)) is interchanged with the summation expression (equation (3.110)):

$$T(\theta,\lambda) = \frac{R}{4\pi}\iint_{\sigma} \Delta g(\theta',\lambda') \sum_{n=2}^{\infty} \frac{1}{n-1} \sum_{m=0}^{n} \quad \times \\ \left[\bar{\mathcal{R}}_{nm}(\theta',\lambda')\bar{\mathcal{R}}_{nm}(\theta,\lambda) + \bar{\mathcal{S}}_{nm}(\theta',\lambda')\bar{\mathcal{S}}_{nm}(\theta,\lambda)\right] \mathrm{d}\sigma \quad (3.111)$$

Legendre's decomposition formula reads (Hofmann-Wellenhof and Moritz, 2005):

$$P_n(\cos\psi) = \frac{1}{2n+1} \sum_{m=0}^{n} \left[\bar{\mathcal{R}}_{nm}(\theta',\lambda')\bar{\mathcal{R}}_{nm}(\theta,\lambda) + \bar{\mathcal{S}}_{nm}(\theta',\lambda')\bar{\mathcal{S}}_{nm}(\theta,\lambda)\right] \quad (3.112)$$

where ψ is the spherical distance between point of interest and arbitrary point. Equation (3.112) inserted into equation (3.111) yields:

$$T(\theta,\lambda) = \frac{R}{4\pi}\iint_{\sigma} \Delta g \quad \overbrace{\sum_{n=2}^{\infty} \frac{2n+1}{n-1} P_n(\cos\psi)}^{\text{Spectral form of Stokes's function}} \quad \mathrm{d}\sigma \quad (3.113)$$

which can also be expressed as:

$$T(\theta,\lambda) = \frac{R}{4\pi}\iint_{\sigma} \Delta g S(\psi)\, \mathrm{d}\sigma \quad (3.114)$$

By applying Bruns's formula, we arrive at Stokes's formula:

$$N(\theta,\lambda) = \frac{R}{4\pi\gamma}\iint_{\sigma} \Delta g S(\psi)\, \mathrm{d}\sigma \quad (3.115)$$

As a result of the above, spherical harmonics and Stokes's formula are equivalent in the global case. This also underlines that, since integration is performed *globally*, globally available Δg are strictly needed.

3.5.3 Spherical harmonics versus radial basis functions

Spherical harmonics may be directly related to radial basis functions (Wittwer, 2009). All radial basis functions are kernel functions that can be represented in an expansion in Legendre polynomials as presented in equation (3.65) (Eicker, 2008). The gravitational potential is a harmonic function, and consequently the radial basis functions are chosen to be harmonic kernel functions.

Again, the spherical harmonic expansion of the disturbing potential T outside the boundary sphere is our starting point, using position vector $\mathbf{x}_P$ at computation point P for brevity:

$$T(\mathbf{x}_P) = \frac{GM}{R} \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} \sum_{m=-n}^{n} F_{nm} Y_{nm}^R(\mathbf{x}_P) \tag{3.116}$$

If the Stokes coefficients F_{nm} are derived from a disturbing potential by spherical harmonic analysis (equation (3.5), in terms of $Y_{nm}^R(\mathbf{x}_P)$), we may write:

$$T(\mathbf{x}_P) = \frac{GM}{R} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{R}{r}\right)^{n+1} \iint_{\sigma_R} T(\mathbf{x}_Q) Y_{nm}^R(\mathbf{x}_Q)\, \mathrm{d}\sigma_R\, Y_{nm}^R(\mathbf{x}_P) \tag{3.117}$$

$$= \iint_{\sigma_R} T(\mathbf{x}_Q) \frac{GM}{R} \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} \frac{1}{R^2} \overbrace{\sum_{m=-n}^{n} Y_{nm}(\mathbf{x}_Q) Y_{nm}(\mathbf{x}_P)}^{\text{Right-hand side of equation (3.6)}} \mathrm{d}\sigma_R \tag{3.118}$$

$$= \iint_{\sigma_R} T(\mathbf{x}_Q) \frac{GM}{R} \sum_{n=0}^{\infty} \frac{2n+1}{4\pi R^2} \left(\frac{R}{r}\right)^{n+1} P_n(\mathbf{r}_P^T \mathbf{r}_Q)\, \mathrm{d}\sigma_R \tag{3.119}$$

where $\mathbf{r}_P^T \mathbf{r}_Q = \cos\psi_{PQ}$ is the scalar product of two unit vectors in the directions of computation point P and arbitrary point Q and ψ_{PQ} is the distance between the two points.

Inserting equation (3.65) into equation (3.119) yields:

$$T(\mathbf{x}_P) = \frac{GM}{R} \iint_{\sigma_R} B(\mathbf{x}_P, \mathbf{x}_Q) T(\mathbf{x}_Q)\, \mathrm{d}\sigma_R \tag{3.120}$$

$$= \frac{GM}{R} (B * T)_{\sigma_R} \qquad \mathbf{x}_Q \in \sigma_R \text{ and } \mathbf{x}_P \in \sigma_R^{\text{exterior}} \tag{3.121}$$

The above equation is known as the *convolution formula*, expressing that the integral of a product of two different functions over a sphere σ_R, where one of the functions must be

dependent on distance only, is a convolution on the sphere. Consequently, the potential outside a sphere can be computed from the potential known on the sphere by a spherical convolution (Bentel, 2010).

In turn, the spherical convolution in equation (3.121) may be expressed as:

$$\frac{GM}{R}\left(B * T\right)_{\sigma_R} = \frac{GM}{R}\sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+1}\sum_{m=-n}^{n} B_n F_{nm} Y_{nm}^{R}(\mathbf{x}_P) \tag{3.122}$$

For proof, see (Bentel, 2010). If the Shannon low-pass RBF is applied, $B_n = 1$ up to a maximum degree N, and equation (3.122) is equivalent to equation (3.116).

The convolution in equation (3.121) is solvable through numerical integration and the field is obtained directly. As presented in section 3.1.2 we can alternatively represent the convolution as a linear combination of the kernel function, treated further in section 3.4.3, now $\mathbf{x}$ is the computation point, and $\mathbf{x}_k$ the origin of RBF number k:

$$\left(B * f\right)(\mathbf{x}) = \sum_{k=1}^{\infty} d_k B(\mathbf{x}, \mathbf{x}_k) \approx \sum_{k=1}^{K} d_k B(\mathbf{x}, \mathbf{x}_k) \tag{3.123}$$

Coefficient transfer between RBFs and spherical harmonics

We may derive spherical harmonic coefficients from RBF coefficients. If we consider a general gravity field signal, its representation by a finite set of RBFs defined on the unit sphere is presented in equation (3.85). When applying the addition theorem to equation (3.85) and compare it with a gravity field signal in spherical harmonics (equation (3.1)), the following connection is observed (Wittwer, 2009):

$$F_{nm} = B_n \frac{4\pi R^2}{2n+1}\sum_{k=1}^{K} d_k \bar{P}_{nm}(\mathbf{r}_k) \tag{3.124}$$

That is:

$$\mathbf{x}_{SH} = \mathbf{A}\mathbf{x}_{RBF} \tag{3.125}$$

or, written out:

$$\begin{bmatrix} F_{00} \\ \vdots \\ F_{N_{\max}N_{\max}} \end{bmatrix} = \begin{bmatrix} B_0\bar{P}_{00}(\mathbf{x}_{k1}) & \cdots & B_0\bar{P}_{00}(\mathbf{x}_{kn}) \\ \vdots & \ddots & \vdots \\ B_{N_{\max}}\frac{4\pi R^2}{2N_{\max}+1}\bar{P}_{N_{\max}N_{\max}}(\mathbf{x}_{k1}) & \cdots & B_{N_{\max}}\frac{4\pi R^2}{2N_{\max}+1}\bar{P}_{N_{\max}N_{\max}}(\mathbf{x}_{kn}) \end{bmatrix} \cdot \begin{bmatrix} d_1 \\ \vdots \\ d_K \end{bmatrix} \tag{3.126}$$

By the variance-covariance propagation law, we get the covariances of the SH coefficients as follows (Leick, 2004):

$$\mathbf{C}_{F_{nm}} = \mathbf{A}\mathbf{C}_{d_k}\mathbf{A}^T \tag{3.127}$$

Thus spectral analysis and direct comparison with spherical harmonics is possible.

Chapter 4

Practical aspects of gravity field modeling

This chapter explores the different practical aspects of gravity field modeling — necessary reductions and corrections as well as practical computation schemes for geoid determination.

4.1 Reductions and corrections

Regardless which approach to geoid determination is chosen, the assumption that there are no masses in the exterior of the surface does not hold. Molodensky's approach does not take the mass of the atmosphere into account. Stokes's approach incorporates a spherical approximation as well as the most important factor needed to be corrected for: no masses outside the geoid is assumed. We must find a way to deal with the remaining topographic masses.

4.1.1 Topographic reduction schemes

Several models for representing the topography and schemes for its reduction exist. Common for most of the methods is that knowledge of Earth crust density is assumed, for which often a constant density (the standard density of the surface rocks of the continental crust, $\rho = 2.67\,\mathrm{g}\cdot\mathrm{cm}^{-3}$) is used.

The magnitude of the so-called *indirect effect* (the fact that a relocation of the masses changes the geoid itself) as well as the smoothness and magnitude of the resulting gravity anomalies are important considerations when choosing a specific gravity reduction scheme (Bajracharya, 2003). The gravity anomalies have different geophysical interpretations dependent on the reductions scheme used.

For a theoretically correct reduction of gravity to the geoid, the vertical gradient of gravity is needed, i.e., $\partial g/\partial H$ (Hofmann-Wellenhof and Moritz, 2005). If g is the observed gravity

value on Earth's surface, then the value g_0 on the surface of the geoid may be obtained through a Taylor expansion:

$$g_0 = g - \frac{\partial g}{\partial H} H \cdots \tag{4.1}$$

where H is the height of the gravity observation station P above the corresponding point P_0 on the geoid. If we further assume that there are no masses outside the geoid and keep only the linear term of the series in equation (4.1) we obtain:

$$g_0 = g + F \tag{4.2}$$

where

$$F = -\frac{\partial g}{\partial H} H \tag{4.3}$$

is the *free-air reduction*. For most applications, the normal gradient of gravity suffices, such that:

$$F = -\frac{\partial \gamma}{\partial h} H = +0.3086 H \qquad [\text{mGal}] \tag{4.4}$$

The *free-air gravity anomaly* is the gravity anomaly corrected for the observation station height (the gravitational effect of topography is neglected):

$$\Delta g_{\text{FA}} = g - \gamma + F \tag{4.5}$$

It refers to the geoid in the Stokes approach and to Earth's surface in the Molodensky approach (i.e., normal gravity on either the ellipsoid or telluroid is applied, respectively), see also Figure 2.7. For the geoid, however, the attraction of the topography must also be accounted for.

The *Bouguer reduction scheme* removes all topographic masses above the geoid, and smooth gravity anomalies are obtained (desired for numerical reasons). The topographic masses are represented by a Bouguer plate. If we further assume that the area around the gravity station P is flat and that the topographic masses between the geoid and Earth's surface have constant density, the attraction of the Bouguer plate (a circular cylinder with thickness H and infinite radius) is:

$$A_B = 2\pi G \rho H = 0.1119 H \qquad [\text{mGal}] \tag{4.6}$$

Equation (4.6) is the *direct effect* of the Bouguer plate on the measured gravity.

By removing the Bouguer plate, the attraction of the observed gravity is subtracted. To complete the reduction, free-air reduction must be applied, and the Bouguer gravity on the geoid is obtained, summarized in Table 4.1.

Measured gravity in P	g
Subtract Bouguer plate A_B	$-0.1119H$
Free-air reduction F	$+0.3086H$
Bouguer gravity in P_0	$g_B = g + 0.1976H$

Table 4.1: *Bouguer gravity on the surface of the geoid*

Ultimately, the Bouguer anomaly is computed as follows:

$$\Delta g_B = g_B - \gamma_0 \tag{4.7}$$

A refined Bouguer reduction takes the deviation of the Bouguer plate from the topography into account (compensates for the plate's lack of shape) through a terrain correction term, illustrated in Figure 4.1. At A we have a mass *surplus* $\Delta m+$ causing an upward attraction which, when removed, leads to an increase in g at P. At B a mass *deficiency* $\Delta m-$ is observed, which when added also leads to an increase in g at P. Thus, the terrain correction term c_P will always be positive. The refined Bouguer gravity anomaly is computed as follows (with the total effect of the topographic masses $A_T = -A_B + c_P$):

$$\Delta g_B = g_P - \gamma_0 + A_T + F \tag{4.8}$$

The terrain correction will be treated later in this section.

The indirect effects of the Bouguer reduction can become large (100 m) and consequently, Bouguer anomalies are not suitable for precise geoid determination (Bajracharya, 2003). Still, they may be used for gravity interpolation, cf. sections 2.5 and 3.3.1.

If g at a point Q *inside* the Earth is sought, a *Prey reduction* scheme may be used (Hofmann-Wellenhof and Moritz, 2005). g_Q can be computed from the observed g_P by the following approximation, cf. Table 4.2:

$$g_Q = g_P + 0.0848(H_P - H_Q) \tag{4.9}$$

In summary, when free-air anomalies are corrected for the topography, Bouguer-anomalies are obtained. They are smoother than the free-air anomalies, but still influenced by the topography (Gerlach, 2003). In mountainous areas, the Bouguer anomalies are strongly negatively correlated with the topography, suggesting that the topographic masses are in some way *compensated* for, i.e., there is a form of mass deficit under the mountains. A

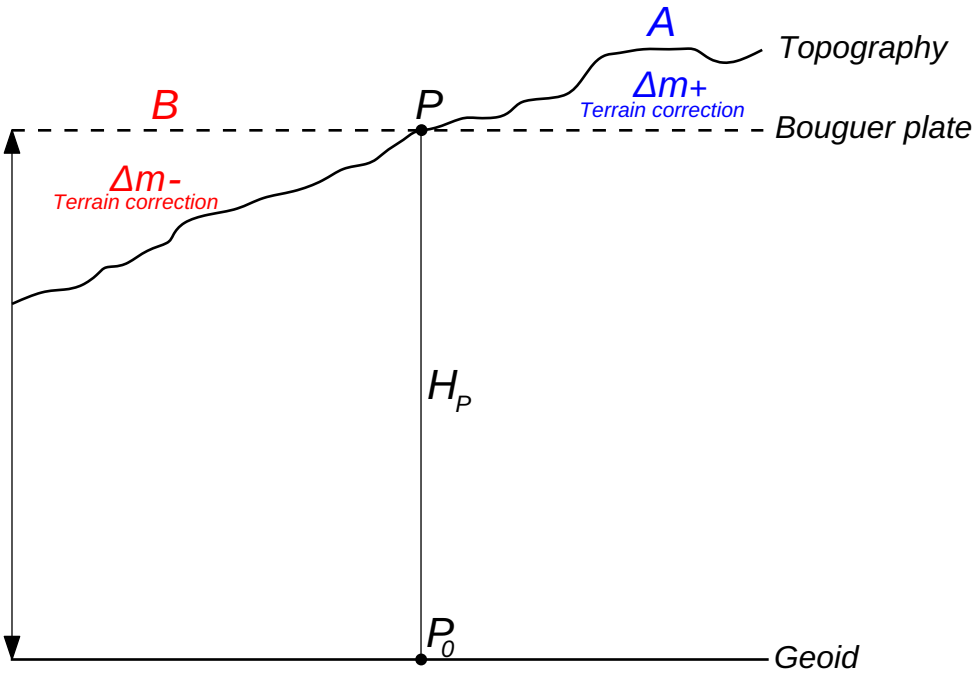

Figure 4.1: *Terrain correction in refined Bouguer reduction*

Measured gravity in P	g
Remove Bouguer plate *above* Q	$-0.1119(H_P - H_Q)$
Free-air reduction	$+0.3086(H_P - H_Q)$
Restore Bouguer plate *above* Q	$-0.1119(H_P - H_Q)$
Gravity in Q	$g_Q = g + 0.0848(H_P - H_Q)$

Table 4.2: *Prey reduction scheme by the* remove-restore *principle.*

similar effect can be seen in the deflections of the vertical, which are smaller than one would expect in mountainous areas (Hofmann-Wellenhof and Moritz, 2005).

As it happens, the Earth is isostatically-compensated, and several theories describing this exist. Generally a static equilibrium is assumed at some depth, where the pressure exerted by the density layers above is constant (Gerlach, 2003). Classical isostasy theories assume that isostatic compensation takes place locally, in vertical columns. The theory of *Airy-Heiskanen* proposes that the thickness of the crust varies, the crust having constant density. In other words, the compensating masses are in the form of "undulations in the crust-mantle interface" (Blakely, 1996). Below mountains, crust with lower density extends into mantle with higher density. Conversely, below deep oceans the mantle extends upward into the crust. The truth, however, is more complicated, and Earth's isostatic-compensation mechanisms vary with location depending on geology. Isostasy is mainly a regional phenomena, and a region may be either uncompensated or compensated at deep levels, e.g., the oceanic trenches and ridges (Forsberg, 1984).

A *topographic-isostatic reduction scheme* formalizes isostasy theories, such as the Airy-Heiskanen model, where the thickness of the *root* (i.e., how deep into the mantle the crust extends) is given by (Forsberg, 1984):

$$t = \frac{\rho}{(\rho_M - \rho_C)} h = \frac{\rho}{\Delta\rho} h \tag{4.10}$$

where ρ is the density of the topography, $\Delta\rho$ the density contrast between crust ρ_C and mantle ρ_M, and the normal density model has crust thickness D, cf. Figure 4.2.

Thus *isostatic gravity anomalies* are formed by removing the effect of the topography *in addition* to restoring the compensated masses in the Earth's crust using, e.g., the Airy-Heiskanen model:

$$\Delta g_{\text{Airy-Heiskanen}} = g_P - \gamma_0 + F - \delta A \tag{4.11}$$

where $\delta A = A_{\text{Topography}} - A_{\text{Airy-Heiskanen}}$ is the direct topographic effect, i.e., the difference between the attraction of topography and the attraction of the compensated masses. Both $A_{\text{Topography}}$ and $A_{\text{Airy-Heiskanen}}$ may be computed by the following integrals, based on Newton's integral, and for a planar Earth σ:

$$A_{\text{Topography}} = G\rho \iint_{\sigma} \left[\int_{0}^{h} \frac{h - H}{r^3} \, \mathrm{d}H \right] \mathrm{d}\sigma \tag{4.12}$$

$$A_{\text{Airy-Heiskanen}} = G\rho \iint_{\sigma} \left[\int_{-D-t-h}^{-D-h} \frac{h - H}{r^3} \, \mathrm{d}H \right] \mathrm{d}\sigma \tag{4.13}$$

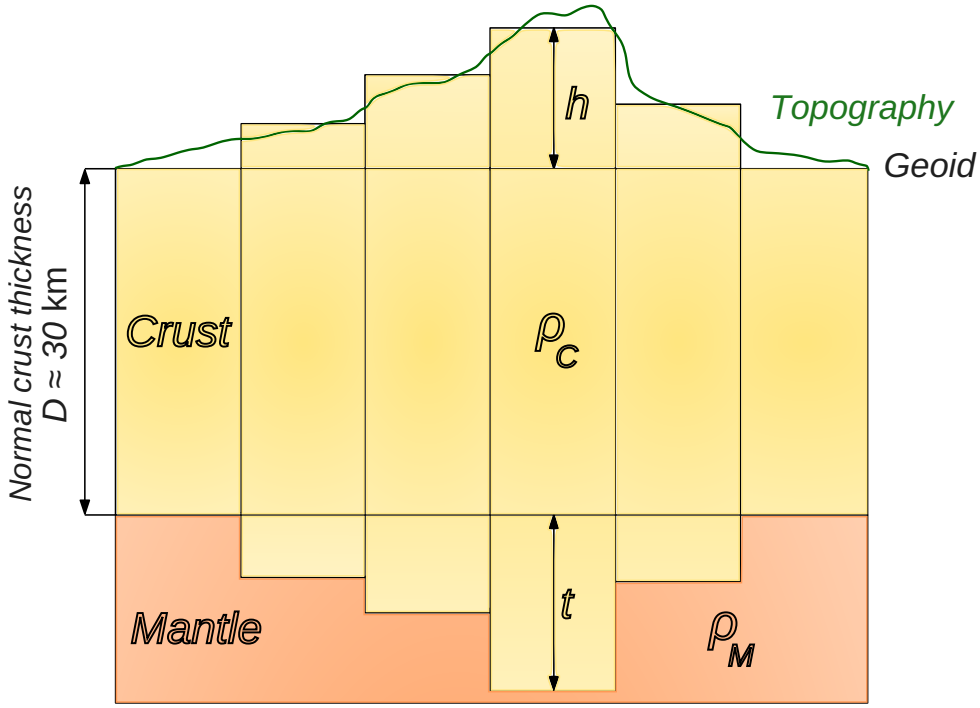

Figure 4.2: *Airy-Heiskanen isostasy model*

To summarize the computation scheme for Airy-Heiskanen anomalies, we start with gravity g_P measured on Earth's surface. Next, the topography above the geoid is removed, and the following effect on gravity $A_{\text{Topography}}$ determined. The masses are shifted inside the geoid, and this mass compensation effect $A_{\text{Airy-Heiskanen}}$ is computed, allowing the determination of δA.

As mentioned above, most of Earth's topography is isostatically compensated, and consequently the isostatic anomalies are usually quite small (Torge and Müller, 2012). Typical uncompensated areas are the boundaries of continental plates and regions experiencing glacial isostatic adjustment (e.g., Scandinavia). A consequence of the nature of isostatic anomalies is that they are of interest to the field of geodynamics (deviations of the isostatic anomalies from zero indicate isostatic imbalance and may reveal geological features). Like Bouguer anomalies, isostatic anomalies may also be successfully applied in gravity interpolation. However, the indirect effects of isostatic anomalies may amount to 10 m, and consequently, they are not recommended for precise geoid determination.

Another reduction scheme which yields small indirect effects, and therefore works well for geoid determination, is *Helmert's condensation method*. It is a classical reduction scheme which is suitable to determine the geoid. In the *first* Helmert condensation method, the topographic masses are shifted along the vertical and condensed at a parallel surface below the geoid. In the *second* condensation method, the masses are condensed onto the geoid directly. The latter method will be introduced here. It is assumed that the condensed layer has a surface density equal to the product of the density and the topographic height,

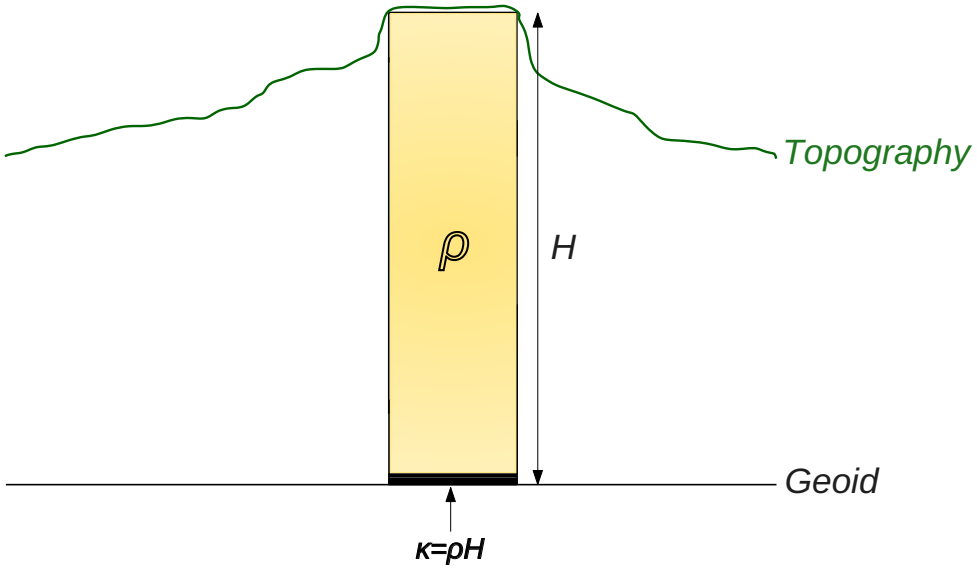

Figure 4.3: *Helmert's second condensation reduction*

cf. Figure 4.3. The attraction of the condensed topography, represented with a Bouguer plate, becomes:

$$A_{\text{Helmert}} = 2\pi G\rho H \tag{4.14}$$

Recalling the refined Bouguer direct topographical effect $A_T = -A_B + c_P$, we get $\delta A = +c_P$. The quantity represents the high frequency of the gravity signal.

Helmert anomalies (also known as Faye anomalies) are thus formed in the following way:

$$\Delta g_{\text{Helmert}} = g_P - \gamma_0 + F + c_P \tag{4.15}$$

Free-air gravity anomalies, cf. equation (4.5), should not be confused with Helmert anomalies, but free-air anomalies can be seen as an approximation to Helmert anomalies at gravity stations on the sea surface or in relatively flat terrain, since the terrain correction term is negligible in those cases (Bajracharya, 2003).

A main purpose of introducing the quasigeoid instead of the geoid, was to avoid assumptions on topographic density, i.e., avoid (a full) topographic reduction. *Residual Terrain Modeling* (RTM) is a common tool for terrain reduction in quasigeoid computation (Forsberg, 1984). The reduction is applied mainly to *smooth* the gravity values, making them easier to interpolate and form block mean values from. It is not a full topographic reduction (containing all frequencies of the topographic spectra), and is applied only on the higher frequencies, i.e., on the *residual* (remaining) terrain model. The RTM anomalies are similar to the isostatic anomalies.

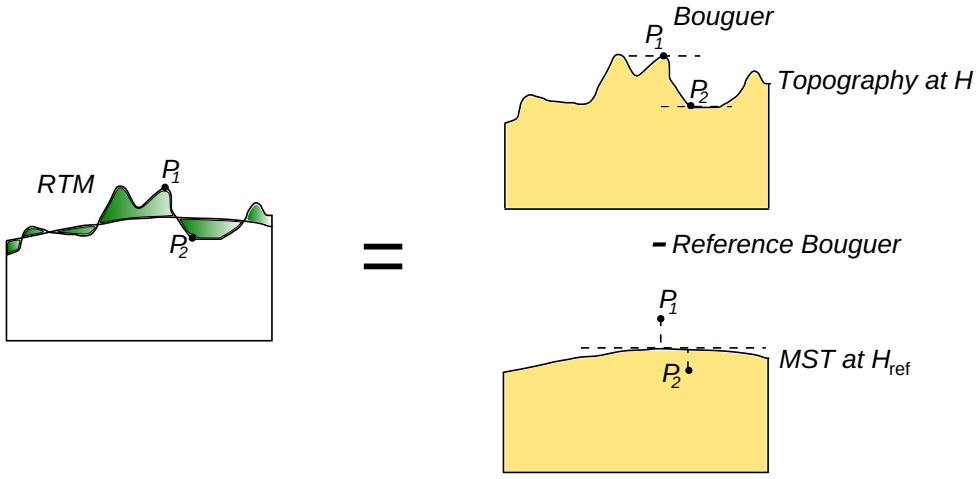

Figure 4.4: *Residual terrain model as a difference between the effects of topography and reference topography*

In RTM we introduce a *mean surface topography* (MST), which generates a gravity field. The residual terrain model refers to the MST, and is the deviation of the MST from Earth's surface. As the RTM values are positive and negative depending on whether the represented areas are above or below the MST respectively, they will even out when the distance from the computation point is large. The MST may be modeled by a spherical harmonic expansion of the topography as follows:

$$H^{\mathrm{MST}} = \sum_{n=0}^{N_{\max}} \sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta) \left[C_{nm}^{\mathrm{MST}} \cos m\lambda + S_{nm}^{\mathrm{MST}} \sin m\lambda \right] \tag{4.16}$$

where C_{nm}^{MST} and S_{nm}^{MST} are the spherical harmonic coefficients of topography from a digital terrain model (DTM), and are thus purely geometrical. $N_{\max}$ is normally chosen equal to the GGM used in the geoid or quasigeoid computation.

The direct effect of the RTM (topography and MST) may be calculated by Newton's integral as introduced for the Airy-Heiskanen approach. It may also be approximated by (Forsberg, 1984):

$$\delta A_{\mathrm{RTM}} = -2\pi G\rho(H - H_{\mathrm{ref}}) + c_P \tag{4.17}$$

In equation (4.17), the first term is the difference between a Bouguer plate computed with the thickness of the computation point height and a Bouguer plate computed with the height of the MST, cf. Figure 4.4.

Also, when applying RTM, we may find ourselves in a situation where the gravity measurements are *inside* the topographic masses after the MST is added. It follows that the requirement of harmonicity is not fulfilled, and a *harmonic correction* needs to be added (Omang et al., 2012).

RTM gravity anomalies are computed as follows:

$$\Delta g_{\mathrm{RTM}} = g_P - \gamma_0 + \delta A_{\mathrm{RTM}} \tag{4.18}$$

The *terrain correction* is generally important in the different gravity reduction schemes, representing the irregular part of the topography (Bajracharya, 2003). The terrain correction integral also stems from Newton's integral for the potential, but does not contain the full topographic signal, only the highest frequencies (Omang and Forsberg, 2000):

$$c_P = G\rho \iint_{\sigma} \left[\int_{H_P}^{h} \frac{h - H_P}{r^3} \, \mathrm{d}H \right] \mathrm{d}\sigma \tag{4.19}$$

There are various ways to compute the terrain correction as given in equation (4.19) (or, indeed the other triple integrals presented in this section). Direct numerical integration in three dimensions (triple summation) may be performed, or an analytical expression for the integral in the brackets (with respect to height H) may be found, and in turn, a two-dimensional numerical integration may be performed (double summation). Alternatively, the computation can be done efficiently in the frequency domain with Fast Fourier Transform techniques, cf. section 4.2.7. Even expressions based on a model of rectangular prisms (bounded by planes parallel to the coordinate planes and defined by coordinates $(x_1, x_2, y_1, y_2, z_1, z_2)$) may be derived (Bajracharya, 2003). The contribution of one prism only (subsequent numerical integration over all prisms necessary) may be formulated as:

$$c_P = G\rho \left| \left| \left| x \ln (y + r) + y \ln (x + r) - z \left(\tan \frac{xy}{zr} \right)^{-1} \right|_{x_1}^{x_2} \right|_{y_1}^{y_2} \Bigg|_{z_1}^{z_2} \tag{4.20}$$

In Molodensky's approach, the terrain correction can approximately replace the g_1 term (Moritz, 1980).

To summarize, terrain reductions are used, even when not theoretically necessary, to smooth the gravity field making it easier to predict gravity values and form block mean values. The classical terrain correction, which is always positive, must be computed for an area at a relatively large distance from the computation point, compared to RTM. The small RTM values with varying positive and negative sign can be neglected at larger distances, making it computationally preferable.

A summary of the treated reduction methods is found in Table 4.3, based on (Bajracharya, 2003).

As mentioned in the introduction of this section, the relocation of the topographic masses changes the gravity field itself, and this *indirect* effect has to be considered. Thus after applying a reduction scheme (the direct effect), we are on a different level surface than the geoid, which is termed the *co-geoid* (Torge and Müller, 2012). The *indirect effect* varies with the applied terrain reduction scheme, but a general setup for the geoid or quasigeoid computation can be formulated:

Reduction scheme	Indirect effects	Geophysical meaning
Bouguer	$\sim 100\,\mathrm{m}$	Yes, correlated with tectonic structures
Airy-Heiskanen	$\sim 10\,\mathrm{m}$	Yes, deeper crustal structures may be revealed
Helmert condensation	$\sim 0.1\,\mathrm{m}$	No
RTM	$\sim 0.5\,\mathrm{m}$ restored terrain effect	No

Table 4.3: *Summary of gravity reduction characteristics*

1. Measured gravity is reduced by the direct effect of the topographic masses (and a direct effect of the dislocated masses may be *added*, cf. Airy-Heiskanen or RTM). Newton's law of gravitation is the starting point for the different computations.

2. The *primary indirect effect* on the potential, due to the aforementioned displacement of masses, must be computed, depending on the reduction method used:

$$\Delta T = T - T_{(\text{Helmert,Airy-Heiskanen,RTM})} \tag{4.21}$$

   where T is the potential of the actual topography and $T_{(\text{Helmert,Airy-Heiskanen,RTM})}$ is the potential of the masses for a chosen reduction scheme.

3. The vertical distance between geoid and co-geoid is computed by Bruns's formula:

$$\Delta N = \frac{\Delta T}{\gamma_0} \tag{4.22}$$

4. The gravity values must be reduced from the co-geoid to the geoid, which is termed a *secondary indirect effect*. A free-air reduction may be used here (Torge and Müller, 2012).

5. The co-geoid N_C is computed using, e.g., Stokes's formula

6. The geoid N is computed by $N = N_C + \Delta N$.

For the RTM reduction scheme, we will have an indirect effect on the quasigeoid, equal to the distance between the actual telluroid and the changed telluroid.

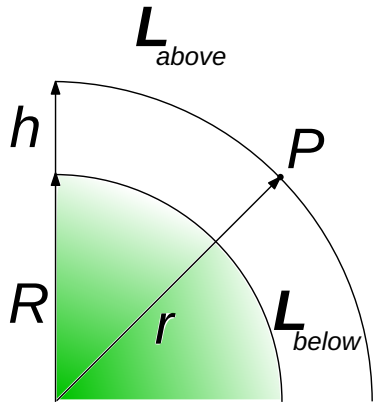

Figure 4.5: *Atmospheric layers above and below computation point P*

4.1.2 Atmospheric correction

Further, the *mass of the atmosphere* has to be considered and reduced for (Gerlach, 2003). In a remove-restore manner, the gravity values are reduced by the atmospheric effect δg_a and ultimately the effect is restored to the geoid. The Earth is spherically approximated and the topography neglected. Density variations are only experienced in the vertical direction. The atmospheric layers *above* a considered point P ($L_{\text{above}}(h > h_P)$) have no gravity effect on P, only the layers *below* P ($L_{\text{below}}(h < h_P)$) are considered, cf. Figure 4.5. The effect comprises:

$$g_a = \frac{G m_{\text{below}}(r)}{r^2} \tag{4.23}$$

where r is the geocentric radius vector of P, $m_{\text{below}}(r)$ the mass of L_{below}.

Important to note when reducing gravity anomalies for the atmospheric effect is that Earth's mass as defined in different geodetic reference systems already *includes* the mass of the atmosphere, i.e., γ_0 includes the atmospheric mass. We get for the atmospheric reduction of the gravity anomaly:

$$\Delta g = g_P - \gamma_0 - \left(-\frac{G m_{\text{above}}(r)}{r^2}\right) = g_P - \gamma_0 + \delta g_a \tag{4.24}$$

A simplified formula for computing δg_a is:

$$\delta g_a = 0.87 \mathrm{e}^{-0.116 H^{1.047}} \tag{4.25}$$

where the result is given in mGal and the orthometric height H of the computation point is in km.

4.1.3 Ellipsoidal correction

Stokes's formula as presented here is a solution to a *spherical* boundary value problem, cf. Table 2.3. The gravity anomaly input data must strictly be given on the sphere, while in reality, gravity anomalies are observed on Earth's surface. As mentioned earlier, the systematic error of neglecting the small flattening of the Earth is about 0.3 % of N.

Solving the ellipsoidal boundary value problem rather than the spherical one can be laborious, with regard to the representation of the disturbing potential T in terms of ellipsoidal harmonics (Fei and Sideris, 2000). Another solution sees Stokes's formula as a *first-order* solution to the ellipsoidal boundary value problem, and introduces a power series approximation up to the order of the square of the eccentricity of the ellipsoid, given by:

$$e^2 = \frac{a^2 - b^2}{a^2} \tag{4.26}$$

Thus for the ellipsoidal reduction of the gravity anomaly, we obtain:

$$\Delta g = \Delta g^0 + e^2 \Delta g^1 \tag{4.27}$$

where $e^2\Delta g^1$ is the ellipsoidal correction term. The gravity input data is given on the ellipsoid, but since the direct transfer of the gravity observable to the disturbing potential T by solving the ellipsoidal boundary value problem is not practical, spherical approximation is introduced. Therefore, as can be seen in (4.27), Δg^1 must be computed in order to transfer the gravity observable Δg to Δg^0. Formulae for the computation of Δg^1 can be found in, e.g., (Fei and Sideris, 2000). In light of this, Stokes's formula takes the following form:

$$N = \frac{R}{4\pi\gamma_0} \iint_\sigma \left(\Delta g - e^2 \Delta g^1\right) S(\psi)\, \mathrm{d}\sigma \tag{4.28}$$

The ellipsoidal correction can be applied to the collocation equation as follows (Moritz, 1980):

$$\hat{N}_P = \mathbf{C}^{N\Delta g^0} \mathbf{C}_{ij}^{-1} \left(\Delta g_i - e^2 \Delta g_i^1\right) \tag{4.29}$$

The ellipsoidal correction can alternatively be computed using spherical harmonic expansions of the disturbing potential and gravity anomalies at the ellipsoid, which is also applicable in a *combined* geoid solution and easily supports different kernel modifications (Claessens and Featherstone, 2007).

4.2 Practical evaluation methods for geoid determination

The previous sections in this thesis suggests that the gravity signal can be split up in parts containing the low frequencies (represented by a GGM), the medium frequencies (represented by regional gravity information) and the high frequencies (represented by the topography, e.g., a digital terrain model) as follows:

$$N = N_{\mathrm{GGM}} + N_{\Delta g} + N_{\text{terrain reduction}} \tag{4.30}$$

Spherical harmonic synthesis, as presented in section 3.1 is suitable for computing N_{GGM}. $N_{\Delta g}$ may be computed by Stokes integration, cf. section 3.2, by the RBF approach using parameter estimation (numerical integration is also possible here), cf. section 3.4, or by collocation, cf. section 3.3.2.

For the computation of a precise geoid, the input gravity data must not only present a high resolution, but also have global coverage. Observed gravity values may account for the high resolution (or the finer structures of the geoid), but are not globally available. GGMs give global geoid solutions, but do not incorporate high resolution data.

The geoid (or quasigeoid) is mostly estimated using the *remove-restore* principle, a *combination in the frequency domain* of the different gravity field sources mentioned above.

Generally, the remove-restore technique can be described in the following steps (Gerlach, 2003):

1. A relatively large part of the gravity signal is *removed* (long-wavelength GGM effect and the effect of topography), its value becomes a relative small residual and easy to grid.
2. Stokes's or Molodensky's formula, collocation or radial basis functions is applied to transform the gravity signal to geoid height or height anomaly.
3. The removed part of the signal is *restored*. The effect of topography depends on the reduction method used, cf. section 4.1.

The remove-restore principle has many advantages (Gerlach, 2003). The computationally intensive integral formulae may be restricted to a limited area where high resolution information is available. The integration areas farther away are covered by the long-wavelength information found in a GGM. The error due to spherical approximation of Stokes's formula is reduced, since the residual gravity signal is only a small part of the whole gravity signal, especially when the high-frequency structures of the topography is additionally reduced. In addition, topographic reductions make the gravity values smoother, and subsequently it is easier to form block mean values needed when evaluating the surface integrals.

4.2.1 Combination in the frequency domain

There are, in principle, two different methods for combining gravity field data: combination in the *frequency* domain, which is the classical remove-restore approach, or combination in the *spatial* domain (Gerlach, 2003). In the following, we assume the application of Stokes's integral formula, but other approaches for modeling the N_2 term, such as collocation, may of course be used as well.

By combination in the frequency (or spectral) domain, the gravity signal is as readily mentioned divided into different frequency parts. The gravity anomaly Δg is reduced for the long- and medium-wavelength parts of the gravity anomaly $\Delta g_{\mathrm{GGM}} = \Sigma_{n=2}^{N_{\max}} \Delta g_n$, cf. Table 3.1, such that only the short-wavelength information remains. As seen above, N_2 may be further reduced for topographic effects. The subsequent integration yields the short-wavelength part of the geoid. The long- and medium-wavelength parts of the geoid are computed by a GGM, and are thus restored as follows:

$$N = N_1 + N_2 + N_3 \tag{4.31}$$

where

$$N_1 = \frac{GM}{\gamma_0 r} \sum_{n=2}^{N_{\max}} \lambda_n^{\mathrm{upw}} \sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta) \left[\Delta\bar{C}_{nm} \cos m\lambda + \bar{S}_{nm} \sin m\lambda\right] \tag{4.32}$$

$$N_2 = \frac{R}{4\pi\gamma_0} \iint_{\sigma_0} (\Delta g - \Delta g_{GGM})\, S(\psi)\, \mathrm{d}\sigma \tag{4.33}$$

$$N_3 = \frac{R}{4\pi\gamma_0} \iint_{\sigma - \sigma_0} (\Delta g - \Delta g_{GGM})\, S(\psi)\, \mathrm{d}\sigma \tag{4.34}$$

The integration of the residual anomalies is done in two separate zones, where the *near zone* is a spherical cap σ_0 with radius (spherical distance) ψ_0 from computation point P, and the *distant zone* comprises $\sigma - \sigma_0$, cf. Figure 4.6. The integration of gravity data in σ_0 with integration radius ψ_0 is a *truncated* integration (Šprlák, 2010). Since the residual signal is only a small portion of the total signal and its influence decreases with increasing distance from the computation point, the residual signal of the far zone (N_3, also termed the *truncation error* or *omission error*) can be neglected if ψ_0 is large enough (Gerlach, 2003). In practice, the size of σ_0 may be chosen such that N_3 is negligible. Consequently integration has to be performed in the near zone only.

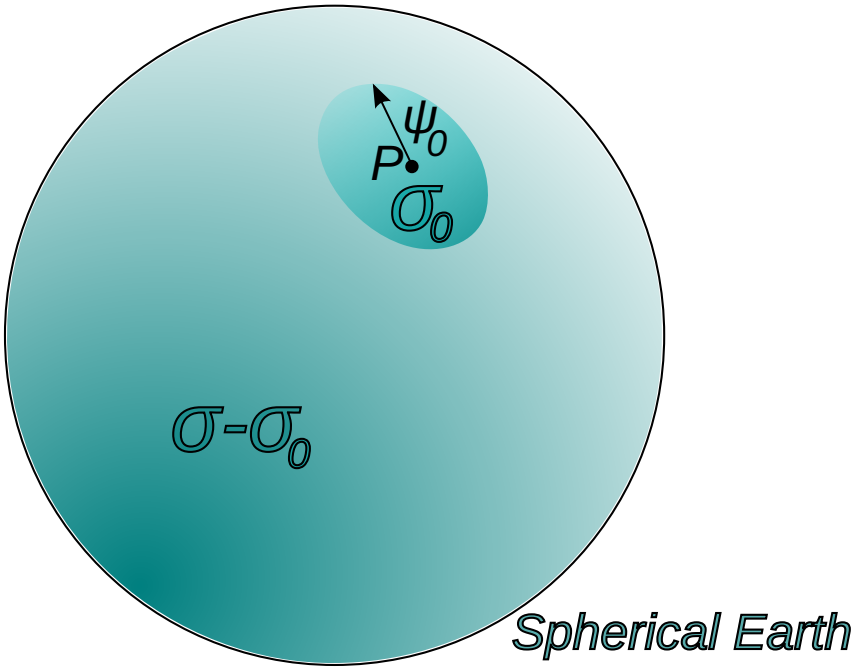

Figure 4.6: *Near (inner) zone σ_0 and distant (far) zone $\sigma - \sigma_0$ in geoid computation*

4.2.2 Combination in the spatial domain

Combination in the spatial domain was originally proposed by Molodensky (Šprlák, 2010). It similarly secures that integration only has to be performed in the near zone, and splits the gravity signal in different spatial parts, such that $N = N_1 + N_2 + N_3$. However, in part N_2, instead of integrating the residual signal, the unreduced gravity anomaly Δg is integrated (the complete spectral bandwidth) (Gerlach, 2003). Only the contribution of the distant zone is missing. The distant zone is split into parts N_1 covering the spectral domain up to degree $N_{\max}$ of a chosen GGM, and N_3 the frequencies above that. Both N_1 and N_3 are filtered in such a way that the near zone in no way contributes to the spherical harmonic syntheses. This is done by looking at the effect of the distant zone:

$$\frac{R}{4\pi\gamma_0} \iint\limits_{\sigma-\sigma_0} S(\psi)\Delta g \, \mathrm{d}\sigma \tag{4.35}$$

Equation (4.35) accounts for the truncation (omission) error, i.e., N_3. A *truncation error kernel* is introduced (Šprlák, 2010):

$$\Delta K(\psi) = \begin{cases} 0, & 0^\circ \leq \psi \leq \psi_0 \\ S(\psi), & \psi_0 < \psi \leq \pi \end{cases} \tag{4.36}$$

The truncation error kernel may also be expressed as a series with Legendre polynomials as base (a Fourier expansion of the error kernel) :

$$\Delta K(\psi) = \sum_{n=2}^{\infty} \frac{2n+1}{2} Q_n(\psi_0) P_n(\cos\psi) \tag{4.37}$$

where $Q_n(\psi_0)$ are Molodensky's *truncation coefficients* (i.e., the Fourier coefficients of the truncation error kernel):

$$Q_n(\psi_0) = \int_0^{\pi} \Delta K(\psi) P_n(\cos\psi) \sin\psi \, \mathrm{d}\psi \tag{4.38}$$

$$= \int_{\psi_0}^{\pi} S(\psi) P_n(\cos\psi) \sin\psi \, \mathrm{d}\psi \tag{4.39}$$

The truncation coefficients are not computed analytically or by numerical methods, but rather by recurrence relations (Šprlák, 2010).

With the truncation error kernel, equation (4.35) may be written as follows:

$$\frac{R}{4\pi\gamma_0} \iint_{\sigma} \Delta K(\psi) \Delta g \, \mathrm{d}\sigma \tag{4.40}$$

Now, equation (4.40) performs an an integration over the *whole* sphere (a global integral), allowing a *spherical harmonic expansion*! Equation (4.40) thus equals:

$$\frac{R}{4\pi\gamma_0} \iint_{\sigma} \sum_{n=0}^{\infty} \frac{2n+1}{2} Q_n(\psi_0) P_n(\cos\psi) \Delta g \, \mathrm{d}\sigma = \frac{R}{2\gamma_0} \sum_{n=2}^{\infty} Q_n(\psi_0) \Delta g_n \tag{4.41}$$

Equation (4.41) is a series formula for the distant zone, and for the geoid height we may write:

$$N_1 = \frac{GM}{\gamma_0 r} \sum_{n=2}^{N_{\max}} \lambda_n^{\mathrm{upw}} \frac{(n-1)}{2} Q_n(\psi_0) \sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta) \left[\Delta\bar{C}_{nm} \cos m\lambda + \bar{S}_{nm} \sin m\lambda \right] \tag{4.42}$$

$$N_2 = \frac{R}{4\pi\gamma_0} \iint_{\sigma_0} \Delta g S(\psi) \, \mathrm{d}\sigma \tag{4.43}$$

$$N_3 = \frac{GM}{\gamma_0 r} \sum_{n=N_{\max}+1}^{\infty} \lambda_n^{\mathrm{upw}} \frac{(n-1)}{2} Q_n(\psi_0) \sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta) \left[\Delta\bar{C}_{nm} \cos m\lambda + \bar{S}_{nm} \sin m\lambda \right] \tag{4.44}$$

Note that the truncation error N_3 in equation (4.44) is completely equivalent to N_3 in equation (4.34).

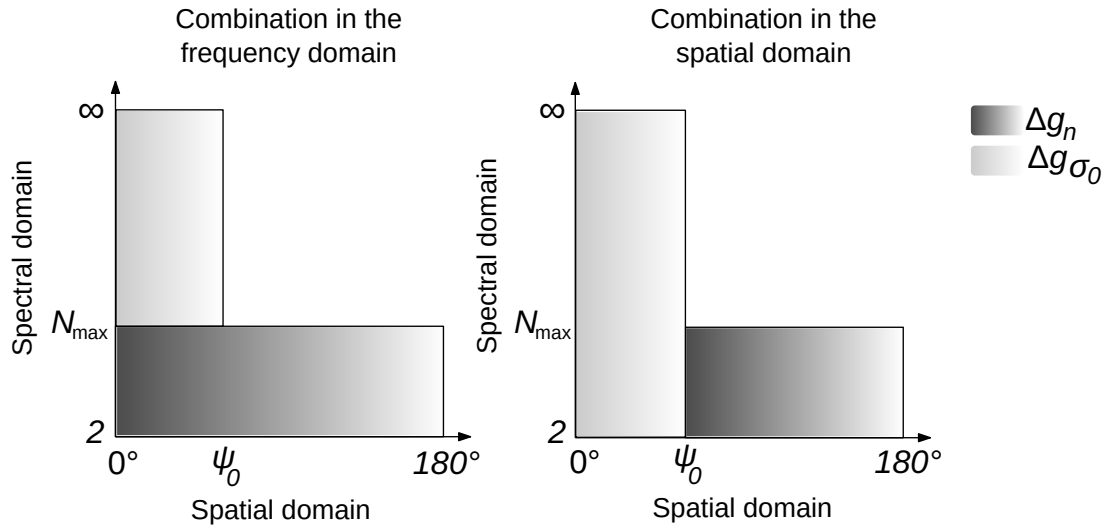

Figure 4.7: *Concept of combination in the frequency and spatial domains*

4.2.3 Geoid estimators

Geoid estimators corresponding to the different combination approaches may be formulated. A geoid estimator is a practical computation estimator for the geoid, based on the combination techniques described above. The following estimators have neglected the N_3 term.

Geoid estimators for the combinations in the frequency (FD) and spatial (SD) domains take the following form (Šprlák, 2010):

$$\tilde{N}^{\mathrm{FD}} = \frac{R}{4\pi\gamma_0} \iint_{\sigma_0} \left[\Delta g - \sum_{n=2}^{N_{\max}} \Delta g_n \right] S(\psi)\, \mathrm{d}\sigma + \frac{R}{2\gamma_0} \sum_{n=2}^{N_{\max}} \frac{2}{n-1} \Delta g_n \tag{4.45}$$

$$\tilde{N}^{\mathrm{SD}} = \frac{R}{4\pi\gamma_0} \iint_{\sigma_0} S(\psi) \Delta g\, \mathrm{d}\sigma + \frac{R}{2\gamma_0} \sum_{n=2}^{N_{\max}} Q_n(\psi_0) \Delta g_n \tag{4.46}$$

Both combinations in the frequency domain (section 4.2.1) and spatial domain (section 4.2.2) are *completely equivalent*, and there is no particular reason to choose one over the other, cf Figure 4.7 (de Min, 1996).

A downside to combination in the spatial domain is that gravity anomalies often are given as block mean values in a regular grid (Gerlach, 2003). The truncation error kernel cuts out a spherical cap around the computation point, thus we might get inconsistencies at the cap edge. However, combination in the spatial domain is well suited for studying error behavior, e.g., estimate the truncation (omission) error, see section 4.2.4.

Ågren and Sjöberg (2004) concluded that the classic remove-restore combination in the frequency domain is quite sensitive to gravity anomaly data correlations (which are often present in a data set) and also sensitive to the lower frequencies of the terrestrial gravity data.

Ellmann (2010) states that although "one would intuitively expect that any numerical error in the integration becomes smaller due to the use of reduced gravity anomalies" studies referred to show that the remove-restore "result is as sensitive to various biases as is the case when Stokes's formula is used with complete anomaly as the integral argument".

4.2.4 Statistics of geoid estimators

The statistical behavior of a geoid estimator is commonly described by the *global mean square error* (GMSE) of the estimator. It measures the global average over the sphere of the square of the error, where the error is the difference between values implied by the estimator and the *true* geoid height as formulated in the original Stokes formula. Ophaug (2013) reviewed the GMSE of the geoid estimator $\tilde{N}^{\mathrm{SD}}$ and found it to be:

$$m^2_{\tilde{N}^{\mathrm{SD}}} = \frac{1}{4\pi}\iint_{\sigma}\left(\tilde{N}^{\mathrm{SD}} - N\right)^2 \mathrm{d}\sigma$$
$$= \overbrace{\left[\frac{R}{2\gamma_0}\right]^2 \sum_{n=2}^{N_{\max}} [Q_n(\psi_0)]^2 \delta c_n}^{\text{Errors of the GGM}} + \overbrace{\left[\frac{R}{2\gamma_0}\right]^2 \sum_{n=N_{\max}+1}^{\infty} [Q_n(\psi_0)]^2 c_n}^{\text{Truncation errors}}$$
$$+ \overbrace{\left[\frac{R}{2\gamma_0}\right]^2 \sum_{n=2}^{\infty} \left[\frac{2}{n-1} - Q_n(\psi_0)\right]^2 \sigma_n^2}^{\text{Terrestrial data errors}} \quad (4.47)$$

The signal degree variances c_n are computed from the potential coefficients of a GGM as follows:

$$c_n = \left(\frac{GM}{r^2}\right)^2 (n-1)^2 \sum_{m=0}^{n} \left(\Delta\bar{C}^2_{nm} + \bar{S}^2_{nm}\right) \quad (4.48)$$

For the degrees above $N_{\max}$ a degree variance model such as *Tscherning/Rapp* (equation (2.51)) may be used. The error degree variances δc_n are computed from the *errors* of the potential coefficients of the same GGM:

$$\delta c_n = \left(\frac{GM}{r^2}\right)^2 (n-1)^2 \sum_{m=0}^{n} \left(\delta(\Delta\bar{C}_{nm})^2 + \delta(\bar{S}_{nm})^2\right) \quad (4.49)$$

As for the terrestrial gravity anomaly variance, it can be estimated as follows (Moritz, 1980):

$$\sigma_n^2 = c_T(1-\mu)\mu^n, \qquad 0 < \mu < 1 \tag{4.50}$$

where c_T and μ are constants estimable from an isotropic covariance function $C(\psi)$, presented here in a closed form:

$$C(\psi) = C_T \left\{ \frac{1-\mu}{\sqrt{1-2\mu\cos\psi+\mu^2}} - (1-\mu) - (1-\mu)\mu\cos\psi \right\} \tag{4.51}$$

With a spherical distance of zero, we get the variance $C_0 = C(0) = c_T\mu^2$, and $C(\psi^0) = \frac{1}{2}c_T\mu^2$, where ψ^0 is the correlation length, cf. section 2.5. c_T and μ may then be found iteratively.

Ellmann (2005) suggests to adapt the value of C_0 that corresponds best to the local terrestrial data to be used for geoid determination. For example, if the gravity anomalies are gridded, their accuracy can be determined by comparing the gridded anomalies with known gravity anomalies. The RMS of the differences between gridded and "true" gravity anomalies may be used for determining C_0. With C_0 determined, σ_n^2 can be computed. One should bear in mind that the quality of the terrestrial data involved (with varying coverage and accuracy) may differ substantially from the *global* average! Therefore, any σ_n^2 model may be questioned and should be treated carefully.

The errors of the GGM represent a commission error, while the truncation error represents an omission error, cf. section 2.5. The GMSE describes how errors in terrestrial and satellite data propagate in the chosen geoid estimator. Individual terms of equation (4.47) may be studied to see how each contribute to the global error. Convergence and filtering properties of a chosen kernel modification can be investigated and information with regard to a suitable integration radius ψ_0 and error propagation of the terrestrial gravity anomalies σ_n can be derived. The GMSE may be used in kernel modification comparison studies, such as (Ågren and Sjöberg, 2004), (Ågren, 2004) or (Šprlák, 2010).

4.2.5 Kernel modifications

The truncation error term, N_3, is often neglected. This does not, of course, mirror reality, since the residual gravity anomalies are not necessarily zero outside the integration domain. Instead one could try to *reduce the truncation error*, which is the main goal with kernel modifications.

Two kernel modification approaches exist:

1. *Deterministic* modifications, focusing on minimizing the truncation error and faster convergence (Šprlák, 2010)

2. *Stochastic* modifications, aiming to reduce the constituents of the GMSE (truncation error, GGM error, terrestrial data error) in a least-squares sense (Ellmann, 2001)

4.2.5.1 Deterministic modification schemes

The original idea of kernel modifications stems from *Molodensky*, which will be presented in the following. The following geoid estimator was proposed:

$$\tilde{N}^{\mathrm{Mol}} = \frac{R}{4\pi\gamma_0}\iint_{\sigma_0}\left[S(\psi) - \sum_{k=2}^{L}\frac{2k+1}{2}s_k P_k(\cos\psi)\right]\Delta g\,\mathrm{d}\sigma \tag{4.52}$$

$$+\frac{R}{4\pi\gamma_0}\iint_{\sigma-\sigma_0}\overbrace{\left[S(\psi) - \sum_{k=2}^{L}\frac{2k+1}{2}s_k P_k(\cos\psi)\right]}^{S^L(\psi)}\Delta g\,\mathrm{d}\sigma \tag{4.53}$$

$$+\overbrace{\frac{R}{4\pi\gamma_0}\iint_{\sigma}\sum_{k=2}^{L}\frac{2k+1}{2}s_k P_k(\cos\psi)\Delta g\,\mathrm{d}\sigma}^{N_1} \tag{4.54}$$

The effect of the distant zone was considered most important, so this term was sought minimized by the use of the Schwarz inequality, which gives the upper bound of the square of an integral. The $\iint_{\sigma-\sigma_0}\Delta g^2\,\mathrm{d}\sigma$ term cannot be minimized, since Δg is our given data. The *kernel*, however, $\iint_{\sigma-\sigma_0}[S^L(\psi)]^2\,\mathrm{d}\sigma$ can be minimized, i.e., a *lowest upper bound for the integral* may be found. The kernel is thus partially differentiated with respect to the coefficients s_k and set equal to zero:

$$\int_{\psi=\psi_0}^{\pi}\left[S(\psi)P_n(\cos\psi) - \sum_{k=2}^{L}\frac{2k+1}{2}s_k P_k(\cos\psi)P_n(\cos\psi)\right]\sin\psi\,\mathrm{d}\psi = 0 \tag{4.55}$$

$$\sum_{k=2}^{L}\frac{2k+1}{2}s_k\overbrace{\int_{\psi=\psi_0}^{\pi}P_k(\cos\psi)P_n(\cos\psi)\sin\psi\,\mathrm{d}\psi}^{e_{nk}(\psi_0)} = \overbrace{\int_{\psi=\psi_0}^{\pi}S(\psi)P_n(\cos\psi)\sin\psi\,\mathrm{d}\psi}^{Q_n(\psi_0)} \tag{4.56}$$

where e_{nk} are termed *Paul's coefficients*, numerically computable by recurrence formulae (Šprlák, 2010). L is often termed the *degree of modification*, and here it is chosen equal to the maximum degree of the chosen GGM. Equation (4.56) represents the system of

equations, solvable for the unknown s_k coefficients. If it is assumed that harmonic coefficients of Δg are available up to a certain degree $N_{\max}$, we may formulate the following geoid estimator (Jekeli, 1981):

$$\tilde{N}^{\mathrm{Mol}} = \frac{R}{4\pi\gamma_0} \iint_{\sigma_0} S^L(\psi)\Delta g \,\mathrm{d}\sigma + \frac{R}{2\gamma_0} \sum_{n=2}^{N_{\max}} s_n \Delta g_n \tag{4.57}$$

Meissl's kernel modification makes the kernel continuous by subtracting the value of Stokes's kernel at the edge of the near zone, $S^{\mathrm{Meissl}}(\psi) = S(\psi) - S(\psi_0), \quad 0^\circ \le \psi \le \psi_0$ (Jekeli, 1981). The following geoid estimator can be formulated:

$$\tilde{N}^{\mathrm{Meissl}} = \frac{R}{4\pi\gamma_0} \iint_{\sigma_0} S^{\mathrm{Meissl}}(\psi)\Delta g \,\mathrm{d}\sigma + \frac{R}{2\gamma_0} \sum_{n=2}^{N_{\max}} Q_n^{\mathrm{Meissl}}(\psi_0)\Delta g_n \tag{4.58}$$

For the truncation coefficients, we may write:

$$Q_n^{\mathrm{Meissl}}(\psi_0) = \int_0^{\pi} \Delta K^{\mathrm{Meissl}}(\psi) P_n(\cos\psi)\sin\psi \,\mathrm{d}\psi \tag{4.59}$$

$$= S(\psi_0) \int_0^{\psi_0} P_n(\cos\psi)\sin\psi \,\mathrm{d}\psi + \overbrace{\int_{\psi_0}^{\pi} S(\psi) P_n(\cos\psi)\sin\psi \,\mathrm{d}\psi}^{Q_n} \tag{4.60}$$

Wong and Gore suggest a kernel modification that removes the low-degree Legendre polynomials from Stokes's kernel, not requiring any minimization. The idea is that when only high frequencies are present in the kernel, the lower frequencies will never be multiplied by our terrestrial Δg data. The lower frequencies are readily available in the harmonic coefficients of Δg_n. Wong and Gore's kernel modification thus acts as a high-pass filter. Their geoid estimator can be written as follows:

$$\tilde{N}^{\mathrm{WG}} = \frac{R}{4\pi\gamma_0} \iint_{\sigma_0} \left[S(\psi) - \sum_{k=2}^{L} \frac{2k+1}{k-1} P_k(\cos\psi) \right] \Delta g \,\mathrm{d}\sigma \tag{4.61}$$

$$+ \frac{R}{4\pi\gamma_0} \iint_{\sigma-\sigma_0} \left[S(\psi) - \sum_{k=2}^{L} \frac{2k+1}{k-1} P_k(\cos\psi) \right] \Delta g \,\mathrm{d}\sigma \tag{4.62}$$

$$+ \frac{R}{4\pi\gamma_0} \iint_{\sigma} \sum_{k=2}^{L} \frac{2k+1}{k-1} P_k(\cos\psi) \Delta g \,\mathrm{d}\sigma \tag{4.63}$$

From the above it is clear that Wong and Gore have made a particular choice of coefficients s_k, i.e., $s_k = 2/(k-1)$. If convolution of Stokes's integral using Wong and Gore's kernel is performed over the whole sphere, the result is identical to the remove-restore approach, where the low frequencies are removed from the gravity anomalies to yield a residual gravity signal. However, in a regional application, if the degree of modification L differs from $N_{\max}$ different results may be expected.

4.2.5.2 Stochastic modification schemes

The following stochastic kernel modification approach has been fronted by Professor L. E. Sjöberg at the Royal Institute of Technology (Kungliga Tekniska Högskolan, KTH) in Sweden since the mid-1980s (Sjöberg, 2003). In the literature, this stochastic modification of Stokes's kernel is therefore often termed the "KTH approach", "method of KTH" or "Sjöberg's approach", e.g., (Ellmann, 2005) or (Ågren, 2004). However, others have also pursued kernel modification utilizing stochastic models, e.g., (Wenzel, 1981).

Sjöberg (2003) proposes the following *optimum* geoid estimator, with modification parameters s_n and b_n:

$$\tilde{N}^{\mathrm{S}} = \frac{R}{4\pi\gamma_0} \iint_{\sigma_0} S^L(\psi)\Delta g \,\mathrm{d}\sigma + \frac{R}{2\gamma_0} \sum_{n=2}^{N_{\max}} b_n \Delta g_n \tag{4.64}$$

where

$$b_n = (Q_n^L + s_n^\star)\frac{c_n}{c_n + \delta c_n}, \qquad 2 \le n \le N_{\max} \tag{4.65}$$

Note that, e.g., for Meissl's deterministic modification, $b_n = 0$ and $s_n = 2S(\psi_0)$, for Molodensky's modification $b_n = s_n$ and for Wong and Gore's modification $b_n = Q_n^{N_{\max}} + s_n$ and $s_n = 2/(n-1)$ (Sjöberg, 2003).

The modified Stokes function is as usual:

$$S^L(\psi) = \overbrace{\sum_{n=2}^{\infty} \frac{2n+1}{n-1} P_n(\cos\psi)}^{\text{Spectral representation of } S(\psi)} - \sum_{n=2}^{L} \frac{2n+1}{2} s_n P_n(\cos\psi) \tag{4.66}$$

The truncation coefficients are:

$$Q_n^L = Q_n - \sum_{k=2}^{L} \frac{2k+1}{2} s_k e_{nk}(\psi_0) \tag{4.67}$$

where Q_n denotes Molodensky's truncation coefficients and $e_{nk}(\psi_0)$ are Paul's coefficients.

With the introduction of spectral errors ε_n^T and ε_n^S of the terrestrial and GGM derived gravity respectively, the geoid estimator can be expressed as follows, in spectral form:

$$\tilde{N}^{\mathrm{S}} = \frac{R}{2\gamma_0} \sum_{n=2}^{\infty} \left(\frac{2}{n-1} - Q_n^L - s_n^\star \right) (\Delta g_n + \varepsilon_n^T) + \frac{R}{2\gamma_0} \sum_{n=2}^{N_{\max}} (Q_n^L + s_n^\star)(\Delta g_n + \varepsilon_n^S) \quad (4.68)$$

The modification parameters $s_n^\star$ are:

$$s_n^\star = \begin{cases} s_n & \text{if } 2 \le n \le L \\ 0 & \text{if } n > L \end{cases} \quad (4.69)$$

An expression for the GMSE of equation (4.68) can be expressed as follows, cf. section 4.2.4:

$$m_{\tilde{N}^{\mathrm{S}}}^2 = \left[\frac{R}{2\gamma_0} \right]^2 \sum_{n=2}^{N_{\max}} (b_n^2 \delta c_n) + \left[\frac{R}{2\gamma_0} \right]^2 \sum_{n=2}^{\infty} [b_n^\star - Q_n^L(\psi_0) - s_n^\star]^2 c_n + \left[\frac{R}{2\gamma_0} \right]^2 \sum_{n=2}^{\infty} \left[\frac{2}{n-1} - Q_n^L(\psi_0) - s_n^\star \right]^2 \sigma_n^2 \quad (4.70)$$

where

$$b_n^\star = \begin{cases} b_n & \text{if } 2 \le n \le L \\ 0 & \text{otherwise} \end{cases} \quad (4.71)$$

Comparing the stochastic approach to Molodensky's kernel modification approach, where Molodensky's approach minimized the *kernel* in a least-squares sense, the stochastic approach minimizes the GMSE. Consequently, the least-squares parameters are obtained by differentiating the GMSE with respect to s_n, $\frac{\partial m_{\tilde{N}^{\mathrm{S}}}^2}{\partial s_n}$. The resulting expression is then equated to zero, and the modification parameters s_n are solved by the method of least-squares from the following linear system of equations:

$$\sum_{r=2}^{L} a_{kr} s_r = h_k, \; k = 2, 3, \ldots, L \quad (4.72)$$

where

$$a_{kr} = \sum_{n=2}^{\infty} E_{nk}E_{nr}C_n + \delta_{kr}C_r - E_{kr}C_k - E_{kr}C_r \tag{4.73}$$

and

$$h_k = \Omega_k - Q_kC_k + \sum_{n=2}^{\infty}(Q_nC_n - \Omega_k)E_nk \tag{4.74}$$

where

$$\Omega_k = \frac{2\sigma_k^2}{k-1} \tag{4.75}$$

$$\delta_{kr} = \begin{cases} 1 & \text{if } k = r \\ 0 & \text{otherwise} \end{cases} \tag{4.76}$$

$$C_k = \sigma_k^2 + \begin{cases} c_k\delta c_k/(c_k + \delta c_k) & \text{if } 2 \leq k \leq N_{\max} \\ c_k & \text{if } k > N_{\max} \end{cases} \tag{4.77}$$

$$E_{nk} = \frac{2k+1}{2}e_{nk}(\psi_0) \tag{4.78}$$

Ellmann (2005) has developed software that computes the stochastic modification parameters s_n and b_n for, e.g., the herein described *optimum* stochastic modification. Ellmann (2012) released an update to the software, allowing computation of gravity anomaly degree variances from high-degree GGMs.

Problems arise when trying to solve the above linear system of equations (Ellmann, 2005). Particularly, the design matrix $\mathbf{A}$ is ill-conditioned, cf. section 5.1. Consequently, a regularization strategy must be chosen for solving the system, and Ellmann (2005) uses the truncated singular value decomposition (T-SVD) approach, see, e.g., (Aster et al., 2005). Ågren (2004) discusses, e.g., why the system of equations is ill-conditioned and needs to be handled by regularization techniques.

Another remark regarding the above described stochastic modification approach, is that *surface* gravity anomalies are used for determining the geoid. *All necessary corrections are subsequently added to* $\tilde{N}^{\mathrm{S}}$, so-called additive corrections. They comprise the combined topographic correction (including the sum of indirect and direct effects of topography), downward continuation effect and a combined atmospheric correction and the ellipsoidal correction. Details can be found in, e.g., (Ellmann, 2001).

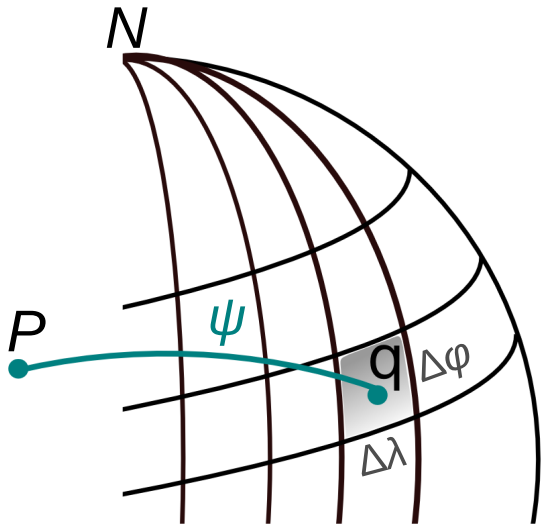

Figure 4.8: *Numerical integration on a regular geodetic coordinate grid*

4.2.6 Numerical integration

Stokes's formula, as an integral formula example, may in geodetic coordinates be written:

$$N(\varphi_P, \lambda_P) = \frac{R}{4\pi\gamma_0} \int_{\lambda_Q=0}^{2\pi} \int_{\varphi_Q=-\frac{\pi}{2}}^{\frac{\pi}{2}} S(\psi) \Delta g(\varphi_Q, \lambda_Q) \cos\varphi_Q \mathrm{d}\varphi_Q \mathrm{d}\lambda_Q \tag{4.79}$$

where (φ_P, λ_P) are the geodetic coordinates of the computation point P, (φ_Q, λ_Q) are the geodetic coordinates of each gravity anomaly $\Delta g(\varphi_Q, \lambda_Q)$. ψ_{PQ} is the distance between these two points. The integral formulae, such as equation (4.79), may be solved numerically (approximation by summation). The Earth may then be subdivided into a regular grid with blocks (or cells) q of size $(\Delta\varphi \times \Delta\lambda)$. Gravity anomalies inside each block are used to form block mean values $\Delta\bar{g}_q$, cf. Figure 4.8. Equation (4.79) becomes:

$$N(\varphi_P, \lambda_P) = \frac{R}{4\pi\gamma_0} \sum_{\varphi_q=\varphi_1}^{\varphi_2} \sum_{\lambda_q=\lambda_1}^{\lambda_2} \Delta\bar{g}_q S(\psi_{Pq}) A_q \tag{4.80}$$

where $A_q = \mathrm{d}\sigma = \cos\varphi_q \Delta\varphi_q \Delta\lambda_q$ is the area of each block. ψ_{Pq} is the spherical distance from the computation point P to the block center, computable from (Hofmann-Wellenhof and Moritz, 2005):

$$\cos\psi_{Pq} = \sin\varphi_P \sin\varphi_q + \cos\varphi_P \cos\varphi_q \cos(\lambda_q - \lambda_P) \tag{4.81}$$

The singularity occurring at $\psi_{Pq} = 0$ can be treated as follows: The innermost zone can be extracted from the solution and treated separately. A good approximation of the contribution of the innermost zone ($0 \leq \psi < \psi_i$), integrated over *polar coordinates* (ψ, α) instead of geodetic coordinates to the singularity, reads (Hofmann-Wellenhof and Moritz, 2005):

$$N_i = \frac{R\psi_i}{\gamma_0}\Delta\bar{g}_i \tag{4.82}$$

where $\Delta\bar{g}_i$ is the gravity anomaly mean block value for the innermost zone. If one deals with block mean values in a geodetic coordinate grid, ψ_i in equation (4.82) is chosen such that the area of the block equals the area of the spherical cap, i.e., $\psi_i = \sqrt{\cos\varphi\Delta\varphi\Delta\lambda/\pi}$ (Gerlach, 2003).

4.2.7 Fast Fourier transform techniques

Alternatively, we can acknowledge that integrals in the form of equation (4.80) form a *convolution* of a spatial function $\Delta\bar{g}_q$ with a weight (or kernel) function $S(\psi_{Pq})$, quite in the same manner as for the general kernel, summarized in equation (3.121). However, the kernel must be a function of distance only, acting as kind of a window, moved over the computation points and element-wise multiplied with the spatial function before all products are summed. Convolution is a computationally intensive procedure in the spatial domain, but in the spectral domain, it is simply a multiplication. Fast Fourier transform techniques (FFT) allow efficient convolution in the spectral domain (Schwarz et al., 1990).

As de Min (1995) pointed out, since FFT-methods are here merely applied for fast computation of convolution integrals, it is not a geoid computation method in itself.

Application of FFT requires the spatial function to be given in a regular grid. Although the width of the block ($\Delta\varphi \times \Delta\lambda$) remains constant, the distance ψ between blocks will vary because of Earth's curvature (Gerlach, 2003). For the distance ψ_{Pq} we may write:

$$\left(\sin\frac{\psi_{Pq}}{2}\right)^2 = \left(\sin\frac{\varphi_P - \varphi_q}{2}\right)^2 + \left(\sin\frac{\lambda_P - \lambda_q}{2}\right)^2 \cos\varphi_P \cos\varphi_q \tag{4.83}$$

From equation (4.83) it is clear that ψ_{Pq} is a function of φ_P and φ_q, and not of $\Delta\varphi$ and $\Delta\lambda$ only, in the considered grid, and a *planar approximation*, with cartesian coordinates $(\Delta x_{Pq}, \Delta y_{Pq})$, may introduced to make the two-dimensional Fast Fourier transform (2D-FFT) applicable. However, the errors due to the approximation of ψ_{Pq} may amount to decimeters in regional applications, and is therefore not suitable when a geoid of centimeter accuracy is sought.

An exact (i.e., identical to numerical integration) one-dimensional FFT solution (1D-FFT) was proposed by Haagmans et al. (1993), where the FFT is just performed along the

latitudes (i.e., one dimension), and the convolution in the meridian (longitude) direction is performed in the spatial domain, so that the solution is limited to a single latitude circle in a single step, i.e., $\Delta\varphi_{Pq} = \text{const}$.

Stokes's integral may be evaluated by the following 1D-FFT formula (Gerlach, 2003):

$$N(\varphi_P) = \frac{R\Delta\varphi\Delta\lambda}{4\pi\gamma_0}\mathcal{F}_1^{-1}\left[\sum_{\varphi_q=\varphi_1}^{\varphi_2}\mathcal{F}_1\{S(\Delta\lambda_{Pq})\}\cdot\mathcal{F}_1\{\Delta\bar{g}_q\cos\varphi_q\}\right] \tag{4.84}$$

where $\mathcal{F}_1$ and $\mathcal{F}_1^{-1}$ denote the 1D-FFT transform operator and its inverse, respectively, and $\Delta\varphi$ and $\Delta\lambda$ are the latitudinal and longitudinal spacing of the computation grid notes, respectively. Equation (4.84) is finally summed over all parallels φ_P.

The Fast Fourier transform implicitly assumes the signal (i.e., Δg) to be periodic, which it actually is if the 1D-FFT is performed along a *whole* parallel ($0 \leq \lambda < 2\pi$) (Gerlach, 2003). In a regional application, however, this requirement is not fulfilled. When FFT is applied in such a case, the signal is implicitly continued to both sides, and on the edges of the considered region, data from the opposite side of the region is mirrored. Therefore *zero-padding* is applied to avoid this. The signal (or spatial function) is continued with zeros in western and eastern directions in an area corresponding to half of the original area. For 2D-FFT zero padding must be done in all directions. Although zero padding also does not reflect reality, since gravity anomalies outside the area of calculation commonly differ from zero, the mirroring effect is eliminated.

With regard to the singularity occurring at $\psi_{Pq} = 0$, again the innermost zone can be extracted from the solution and treated separately, as with numerical integration.

The terrain (topographic) corrections and primary indirect effects (co-geoid to geoid correction) may also be computed by 1D-FFT (Gerlach, 2003).

Any kernel fulfilling the FFT criteria may be introduced into the 1D-FFT. In the remove-restore case, for instance, the residual gravity signal would be introduced instead of the full gravity anomaly, and equation (4.84) combined with a GGM. The modified integration kernels presented in section 4.2.5 may also be introduced into the 1D-FFT, each kernel normally computed before the transformation to the frequency domain.

Chapter 5

Numerical investigations

In this chapter, numerical challenges that occur as a consequence of trying to solve inverse problems as well as possible remedies are explored. This chapter also contains numerical comparisons of the different gravity field modeling approaches (as reviewed theoretically in section 3.5).

5.1 Inverse and ill-posed problems

A problem is called *properly posed* if the solution satisfies the requirements of existence (a solution must exist for arbitrary data), uniqueness (only one solution) and stability (the solution depends continuously on the data) (Moritz, 1980). We have an *ill-posed* problem when one or more of the above requirements are violated. When we go from continuous systems to discrete linear systems, *ill-posed* problems are termed *ill-conditioned* (Aster et al., 2005).

In case of the linear problems described according to $\mathbf{Ax} = \mathbf{l}$, calculating the output $\mathbf{l}$ from a given parameter set $\mathbf{x}$ is the *forward problem*, cf. section 3.1.2. The estimation of parameters $\mathbf{x}$ from observations $\mathbf{l}$, however, is an *inverse problem*. The stability criterion of a properly posed problem implies that the solution $\mathbf{x}$ depends continuously on the data, i.e., that small changes in $\mathbf{x}$ result in small changes in $\mathbf{l}$. For example, in the case of downward continuation, the stability criterion is violated.

In geodesy, we often encounter linear inverse problems, as the parameters to be determined are not directly observable (Eicker, 2008). In turn, inverse problems are often ill-conditioned, for various reasons (Sneeuw, 2000):

- Irregular data distribution or data gaps (e.g., polar gaps in non-polar orbiting satellite observations).
- The observations may not contain sufficient information about the gravity field (e.g., the measuring instrument is not sensitive to all frequencies of the gravity field).
- Downward continuation (as described below) is an ill-posed problem by itself.
- Incorrect choice of radial basis function bandwidth (Bentel et al., 2013).

- The grid for placing the radial basis functions must be chosen properly, to avoid under- or over-parametrization (Wittwer, 2009).

The gravity field signal is damped with increasing distance from Earth's center, since gravitation is proportional to the square of the reciprocal distance, cf. equation (2.1). Upward continuation, as readily described in this thesis, transforms gravity field measurements from one surface to gravity field measurements on a surface at greater distance from the source (Blakely, 1996). In other words, the gravity field in space is determined from the known gravity field on the surface. Consequently, upward continuation is a smoothing operation. The upward continuation operator is repeated here for convenience:

$$\left(\frac{R}{r}\right)^{n+1} \tag{5.1}$$

R is the spherical boundary surface radius (spherical Earth's mean radius) and r is the distance from Earth's center of mass to the computation point. The upward continuation factor could be considered as a weight, and we observe that the weight approaches zero with increasing degree n.

To illustrate the signal dampening at typical gravity satellite height ($\sim 300\,\mathrm{km}$), the global geoid was computed on the Earth surface ($h = 0\,\mathrm{km}$) and at satellite height ($h = 300\,\mathrm{km}$). The GGM *EGM2008* was used. The difference of both geoids can be seen in Figure 5.1. In areas with large variations in topography, differences in geoid heights can amount to about 25 m, which illustrates the smoothing of the signal. In Figure 5.2 the geoid signal degree variances for both cases are plotted. The signal degree variance curve at satellite altitude drops much quicker than the curve at Earth's surface, which means that the signal content in the high degrees is much smaller. In turn, this means that less small-scale features are visible in the signal, i.e., it is smoother.

We experience the dampening of the gravity field signal especially when working with gravity field observations at satellite height. Bringing satellite observations down to the boundary surface is done through downward continuation. It is the opposite, or inverse, of upward continuation. Consequently it is an unstable "unsmoothing" (or roughening) operation, since small observation variations at, e.g., satellite height can cause large and unrealistic variations in the computed surface values (Blakely, 1996). The downward continuation operator reads:

$$\left(\frac{r}{R}\right)^{n+1} \tag{5.2}$$

The singular values of the downward continuation operator approaches infinity with increasing n, i.e., the requirement of stability is violated (Eicker, 2008). Consequently, trying to solve an ill-posed problem such as downward continuation by, e.g., a discrete linear system of equations $\mathbf{Ax} = \mathbf{l}$, will render the system ill-conditioned or singular, the normal equations (3.27) not solvable by the conventional normal matrix inverse $\mathbf{N}^{-1}$, cf. section 3.1.2.

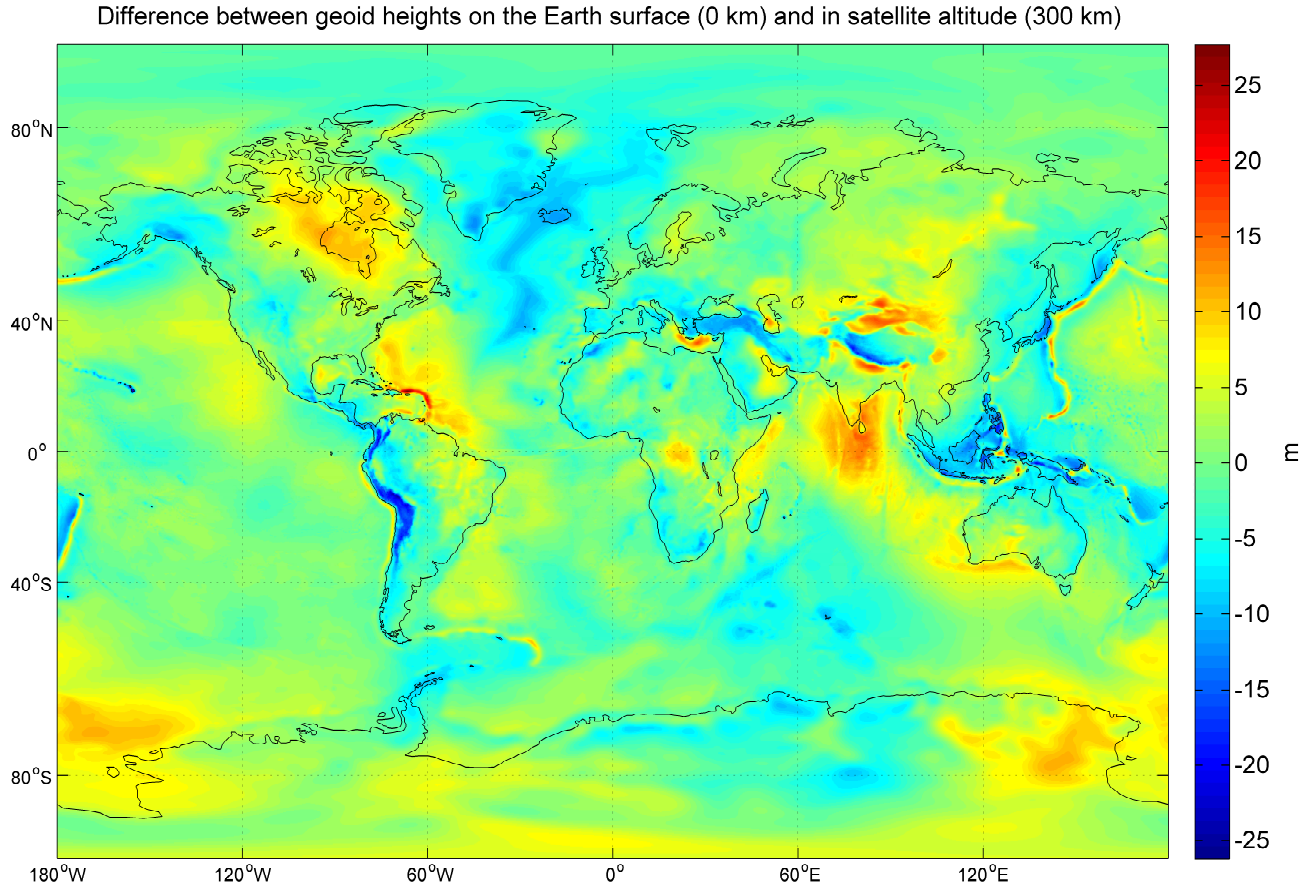

Figure 5.1: *Difference between geoid heights at* $h = 0\,\mathrm{km}$ *and* $h = 300\,\mathrm{km}$

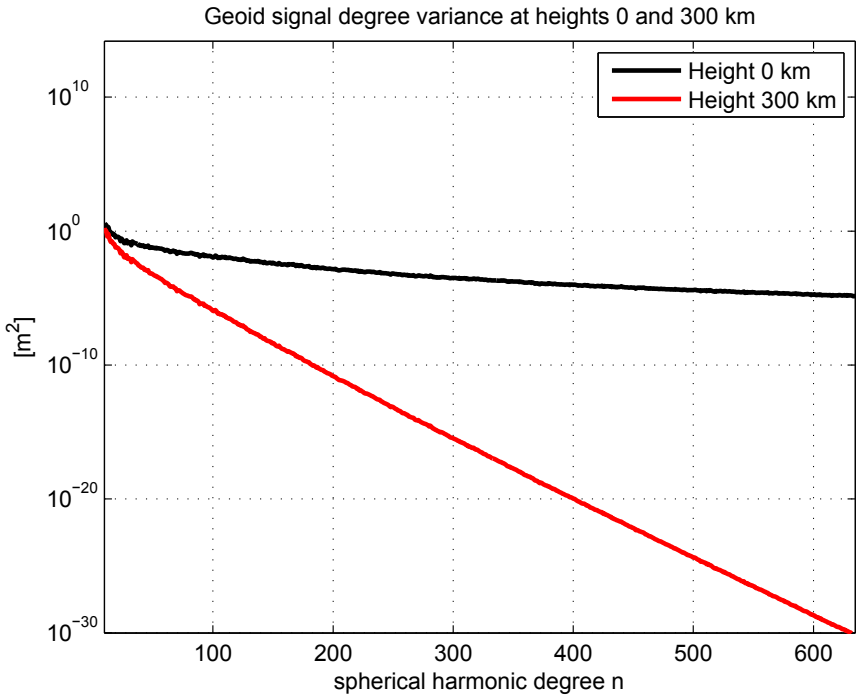

Figure 5.2: *Geoid height signal degree variances at* $h = 0\,\mathrm{km}$ *and* $h = 300\,\mathrm{km}$

A singular system has no unique solution, or the solution is not stable. One way to check if a system is singular is to determine the *condition number*, measuring the instability of the solution. The condition number of a matrix $\mathbf{A}$ is given by (Aster et al., 2005):

$$\mathrm{cond}(\mathbf{A}) = \| \mathbf{A} \| \, \| \mathbf{A}^{-1} \| \tag{5.3}$$

Values of equation (5.3) near one indicates a well-conditioned matrix $\mathbf{A}$, and large values are a characteristic of an ill-conditioned matrix $\mathbf{A}$.

The gradual decay of the eigenvalues of the normal matrix $\mathbf{N}$ towards zero and a large condition number of $\mathbf{N}$ are characteristics of the discrete ill-conditioned system (Sneeuw, 2000). If both characteristics are met, the solution space is not accessible. If the eigenvalues of the normal matrix $\mathbf{N}$ abruptly drop to zero only, the matrix $\mathbf{N}$ has a rank deficiency, which means that the null-space is a part of the solution space.

To be able to solve the system, the normal matrix needs to be regularized (Sneeuw, 2000). Observations $\mathbf{l}$ alone are not sufficient to get a unique solution. For example, one can introduce further solution conditions, such as a priori knowledge of the unknowns $\mathbf{x}$. In the case of gravity field determination, we often have a priori knowledge to a certain extent, so regularization by prior information is possible. If no prior knowledge is at hand, other regularization solutions may be formulated, see, e.g., (Aster et al., 2005).

The common goal of regularization for solving ill-conditioned systems is smoothing the solution of the system. It is however, not wanted too smooth the solution to an extent where important information is lost.

5.2 Tikhonov regularization

A common regularization method for solving inverse and ill-posed problems is *Tikhonov regularization* (Moritz, 1980).

The minimum-variance criterion of least-squares adjustment has readily been mentioned, cf. equation (3.26).

In Tikhonov regularization, the best regularized solution $\hat{\mathbf{x}}_\alpha$ is assumed to be the one that minimizes the *damped* or *regularized* least-squares problem. The norm in equation (3.26) is extended by adding further constraints on the unknown vector $\mathbf{x}$. Assuming prior information is available, the constraints in the extended norm come from prior information on the unknown parameters (Sneeuw, 2000):

$$\hat{\mathbf{x}}_\alpha = \min\{\mathbf{v}^T\mathbf{P}\mathbf{v} + \alpha(\mathbf{x}-\mathbf{x}_0)^T\mathbf{P}_{x_0}(\mathbf{x}-\mathbf{x}_0)\} \tag{5.4}$$

where α is the *regularization parameter*.

The matrix $\mathbf{P}_{x_0}$ is the *regularization* matrix (Sneeuw, 2000). If no a priori information on the unknowns $\mathbf{x}$ is at hand, we can expect them to be zero, i.e., $\mathbf{x}_0 = 0$ with equal variance ($\mathbf{P}_{x_0} = \alpha\mathbf{I}$). In the case of gravity field estimation, signal degree variance models could be applied. Then the coefficients have expectation zero and are assumed to vary according to the signal degree variance, i.e., $\mathbf{x}_0 = 0$ and $\mathbf{P}_{x_0} = \alpha\mathbf{K}^{-1}$, where $\mathbf{K}$ is a diagonal matrix containing the signal degree variances. Another option could be to make use of information from an already existing gravity field solution $\mathbf{x}_0$, and set $\mathbf{P}_{x_0} = \mathbf{Q}_{x_0}^{-1}$, where $\mathbf{Q}_{x_0}$ is the corresponding error cofactor matrix.

Assuming $\alpha = 1$ or incorporated into $\mathbf{P}_{x_0}$ and $\mathbf{P}_{x_0}$ containing both a priori signal and error information, expectation and dispersion of the model may be formulated as follows (Sneeuw, 2000):

$$\mathrm{E}\left\{\begin{pmatrix}\mathbf{l}\\ \mathbf{x}_0\end{pmatrix}\right\} = \begin{pmatrix}\mathbf{A}\\ \mathbf{I}\end{pmatrix}\mathbf{x} \tag{5.5}$$

$$\mathrm{D}\left\{\begin{pmatrix}\mathbf{l}\\ \mathbf{x}_0\end{pmatrix}\right\} = \begin{pmatrix}\mathbf{Q}_{ll} & 0\\ 0 & \mathbf{Q}_{x_0}\end{pmatrix}\mathbf{x} \tag{5.6}$$

We may solve for the unknown parameters as follows:

$$\hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{P}_{x_0}\right)^{-1}\left(\mathbf{A}^T\mathbf{P}\mathbf{l} + \mathbf{P}_{x_0}\mathbf{x}_0\right) \tag{5.7}$$

For the regularized solution in equation (5.7) with $\mathbf{x}_0 = 0$ and $\mathbf{Q}_{x_0}$ containing the signal covariance only, the following equality holds (Moritz, 1980):

$$\left(\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{P}_{x_0}\right)^{-1}\mathbf{A}^T\mathbf{P}\mathbf{l} \equiv \mathbf{Q}_{x_0}\mathbf{A}^T\left(\mathbf{A}\mathbf{Q}_{x_0}\mathbf{A}^T + \mathbf{Q}_{ll}\right)^{-1} \tag{5.8}$$

With signal covariance $x:\ \mathbf{Q}_{x_0} \rightarrow \mathbf{C}_{xx}$, signal covariance $y:\ \mathbf{A}\mathbf{Q}_{x_0}\mathbf{A}^T \rightarrow \mathbf{C}_{yy}$, cross-covariance $xy:\ \mathbf{Q}_{x_0}\mathbf{A}^T \rightarrow \mathbf{C}_{xy}$ and noise covariance $y:\ \mathbf{Q}_{ll} \rightarrow \mathbf{C}_{yy}$, the right-hand side of equation (5.8) is formally the same as the collocation formula (Sneeuw, 2000):

$$\hat{\mathbf{x}} = \mathbf{C}_{xy}(\mathbf{C}_{yy} + \mathbf{D}_{yy})^{-1}\mathbf{l} \tag{5.9}$$

The expectation of $\hat{\mathbf{x}}$ in equation (5.7) can be written as (Sneeuw, 2000):

$$\mathrm{E}\{\hat{\mathbf{x}}\} = \mathbf{R}_y\mathbf{x} + \mathbf{R}_x\mathbf{x} = \mathbf{x} \tag{5.10}$$

with the following definitions:

$$\mathbf{R}_x = \left(\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{P}_{x_0}\right)^{-1}\mathbf{P}_{x_0} \tag{5.11}$$

$$\mathbf{R}_y = \left(\mathbf{A}^T\mathbf{P}\mathbf{A} + \mathbf{P}_{x_0}\right)^{-1}\mathbf{A}^T\mathbf{P}\mathbf{A} \tag{5.12}$$

and $\mathbf{R}_x + \mathbf{R}_y = \mathbf{I}$. $\mathbf{R}_x$ and $\mathbf{R}_y$ are termed *resolution* matrices, measuring the relative contributions of data ($\mathbf{R}_y$) and prior knowledge ($\mathbf{R}_x$) to the solution (Sneeuw, 2000). Note that the resolution matrices are not computed from actual data or actual a priori values, which means that they can be used for evaluating the necessity of regularization with a priori information.

The curve of optimal values of $\| \mathbf{x} \|_2$ versus $\| \mathbf{A}\mathbf{x} - \mathbf{l} \|_2$ plotted on a double logarithmic scale, has a characteristic L shape with a distinct corner, and is therefore known as the *L-curve* (Hansen and O'Leary, 1993), cf. Figure 5.3. The form of the curve is due to $\| \mathbf{x} \|_2$ being a strictly decreasing function of α and $\| \mathbf{A}\mathbf{x} - \mathbf{l} \|_2$ a strictly increasing function of α (Aster et al., 2005). The L-curve may be used for determining the regularization parameter α, since α is assumed to be at the L-curve corner. This implies that a series of solutions for a range of parameters α must be computed. Selecting α by this approach is termed the *L-curve criterion* for choosing a regularization parameter.

Hansen and O'Leary (1993) suggest two different numerical approaches for determining the L-curve corner, stressing, however, that visual inspection of the L-curve is a crucial part of its analysis. One method is to find the point on the curve closest to the origin, another is to choose the point on the L-curve where the *curvature* is at its maximum. If the L-curve function is composed of discrete point-values, a continuous curve has to be fitted to these values and the curvature of the fitted curve computed.

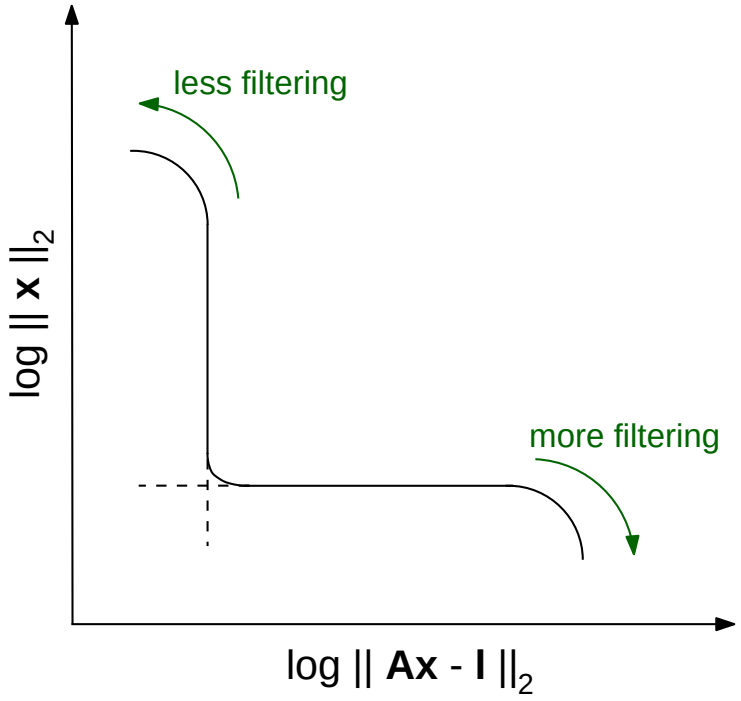

Figure 5.3: *General form of the L-curve.*

5.3 Examples of ill-conditioned systems and regularization using synthetic data

To illustrate the theory presented in sections 5.1 and 5.2, different setups of the problem of estimating spherical harmonic coefficients from an observed global potential field were solved by means of spherical harmonic analysis by least-squares adjustment, cf. section 3.1.2.

Synthetic data was used, allowing closed-loop testing, cf. section 5.4, applying the GGM *EGM2008*. The global disturbing potential was computed by spherical harmonic synthesis, on a global latitude-longitude grid with 5° spacing, up to degree $N_{\max} = 31$, for computational reasons:

$$T_{\mathrm{SHS}} = \frac{GM}{r} \sum_{n=2}^{31} \lambda_n^{\mathrm{upw}} \sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta) \left[\Delta\bar{C}_{nm}^{\mathrm{EGM2008}} \cos m\lambda + \bar{S}_{nm}^{\mathrm{EGM2008}} \sin m\lambda \right] \tag{5.13}$$

In turn, (disturbing) potential coefficients were estimated from these observations by spherical harmonic analysis (least-squares adjustment). Further, the global potential was computed again, by equation (5.13), but with the *estimated coefficients* as argument. Thus, "true" residuals were computable by $T_{\mathrm{est.}} - T_{\mathrm{EGM08}}$. Relative contributions from prior information (regularization) $\mathbf{R}_x$ and from the given data set (observations) $\mathbf{R}_y$ were

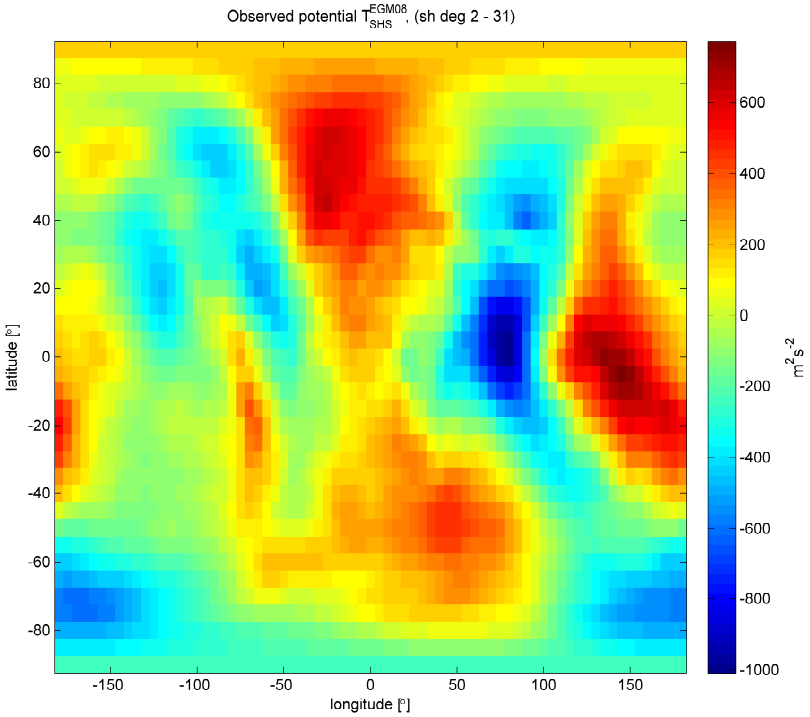

Figure 5.4: *Synthetic observation data, disturbing potential T_{SHS} at height $h = 0$.*

computed. Signal degree variances of "true" *EGM2008* dimensionless potential coefficients were plotted together with the estimated dimensionless potential coefficients, as well as the error degree variances computed from their difference.

In case 1, the observations were assumed to be on Earth's surface, i.e., $h = 0$, and noise-free. In a more realistic case 2, the observations were assumed to be in typical gravity satellite altitude, $h = 300\,\text{km}$. Since this problem in itself does not yield a very ill-conditioned system, it was additionally assumed that the observations are ridden by noise, i.e., $\sigma = 20\,\text{m}^2\text{s}^{-2}$. As it turned out, this problem benefited from regularization, but was numerically solvable without any regularization. Case 3 marked an extreme setup, where it was assumed that the observations were in GPS satellite altitude $h = 20000\,\text{km}$, with $\sigma = 0.1\,\text{m}^2\text{s}^{-2}$. In this setup, no meaningful solutions were obtained without regularization.

The synthetic observation data at different altitudes $h = 0$, $h = 300\,\text{km}$ and $h = 20000\,\text{km}$ are found in figures 5.4, 5.5 and 5.6, respectively.

Results obtained in case 1 are summarized in Figure 5.7. The coefficients are shown in so-called *SC-format*, where the right part of the triangle are the $\Delta\bar{C}_{nm}$ coefficients, starting at $\Delta\bar{C}_{00}$, and the left part of the triangle are the $\bar{S}_{nm}$ coefficients, starting at $\bar{S}_{11}$. With noise-free observations, this is an ideal case, where the exact true field is re-obtained when computing it with the estimated potential coefficients. The residuals are very small, as is

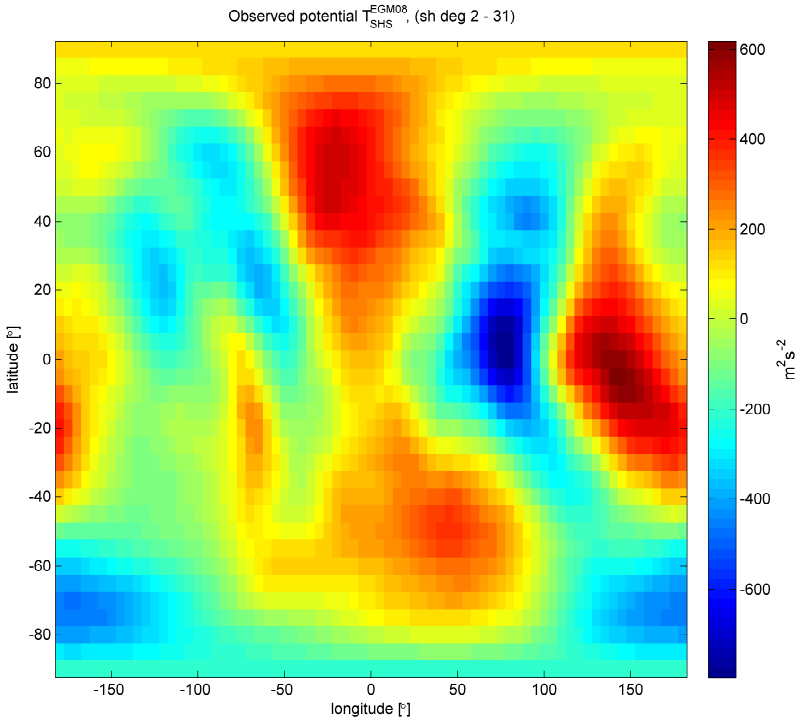

Figure 5.5: *Synthetic observation data, disturbing potential T_{SHS} at height $h = 300\,\mathrm{km}$.*

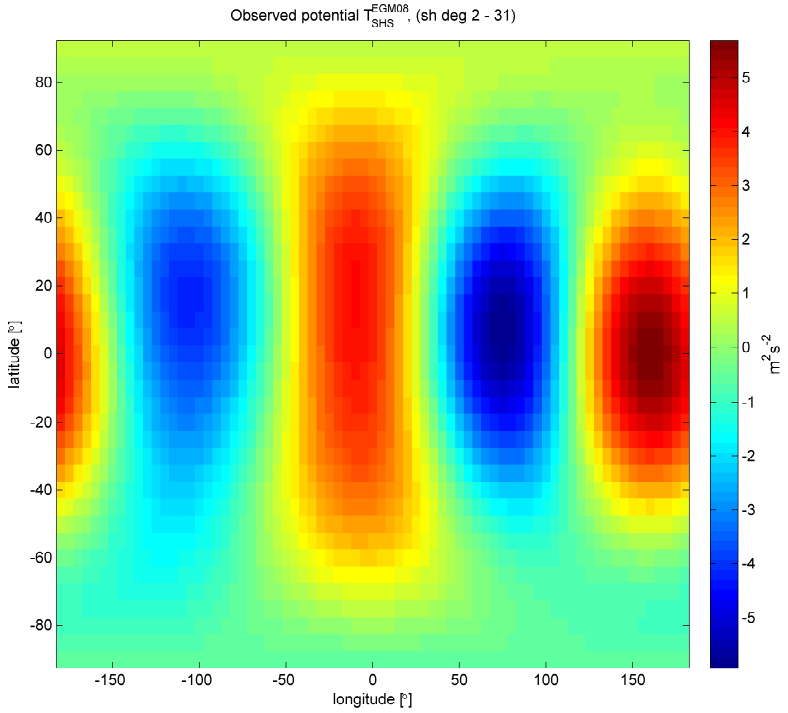

Figure 5.6: *Synthetic observation data, disturbing potential T_{SHS} at height $h = 20000\,\mathrm{km}$.*

the condition number of the normal matrix, summarized for the different cases in Table 5.1. No regularization was applied in this case, i.e., $\alpha = 0$. Consequently, solely the observed data contribute to the solution, which is observed in the resolution matrix $\mathbf{R}_y$. The weights of the regularization matrix $\mathbf{K}^{-1}$ corresponding to the low-degree coefficients $\Delta\bar{C}_{00}$, $\Delta\bar{C}_{10}$, $\Delta\bar{C}_{11}$ and $\bar{S}_{11}$ were set to zero, thus *never* allowing prior information to have an influence on the estimation of these coefficients. The degree variance curve of the estimated coefficients completely overlaps the curve of the true *EGM2008* coefficients, and their error degree variances are small.

Regarding case 2, a solution free of regularization was determined first, summarized in Figure 5.8. Interestingly, the condition number of the normal matrix is still relatively low, cf. Table 5.1. The residuals (observation errors), however, are high, since the data is relatively noisy. This means that the higher degrees cannot be determined. In the degree variances plot it is evident that estimating coefficients above roughly degree 15 is troublesome. The error degree variances curve supersedes the signal curve. This means that we pass the point where the signal-to-noise ratio is one, and it is not possible to separate signal from noise. Since we are in the unique position to see the "true" signal degree variances curve, we see that the estimated coefficients do not mirror reality well above degree 15. There, the solution is unstable.

By introducing a priori information on the unknown coefficients — here they are additional zero observations assumed to vary according to the signal degree variances of *EGM2008* —

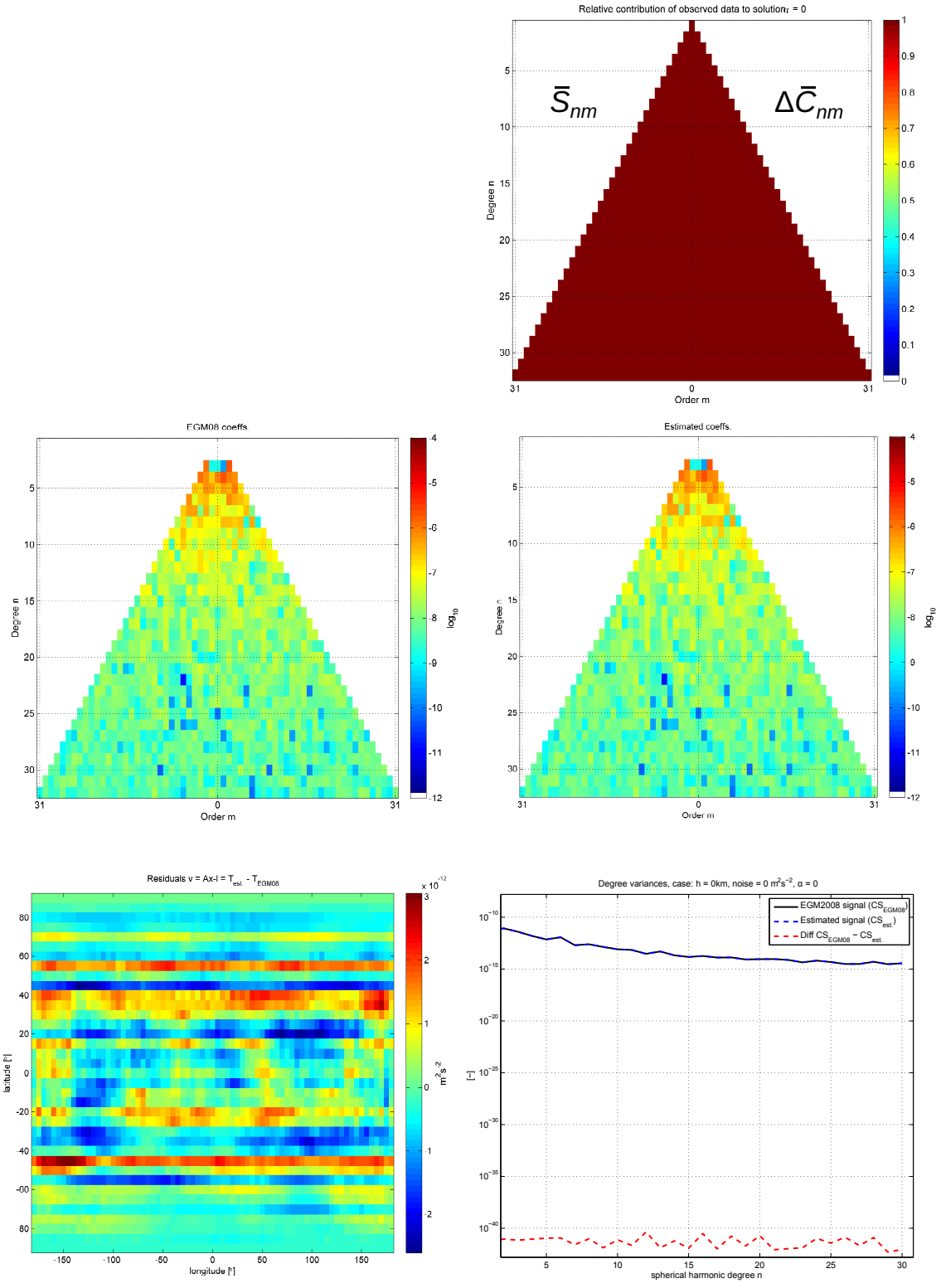

Figure 5.7: Spherical harmonic analysis. *Results from case 1, noise-free observations at* $h = 0$. *No regularization,* $\alpha = 0$.

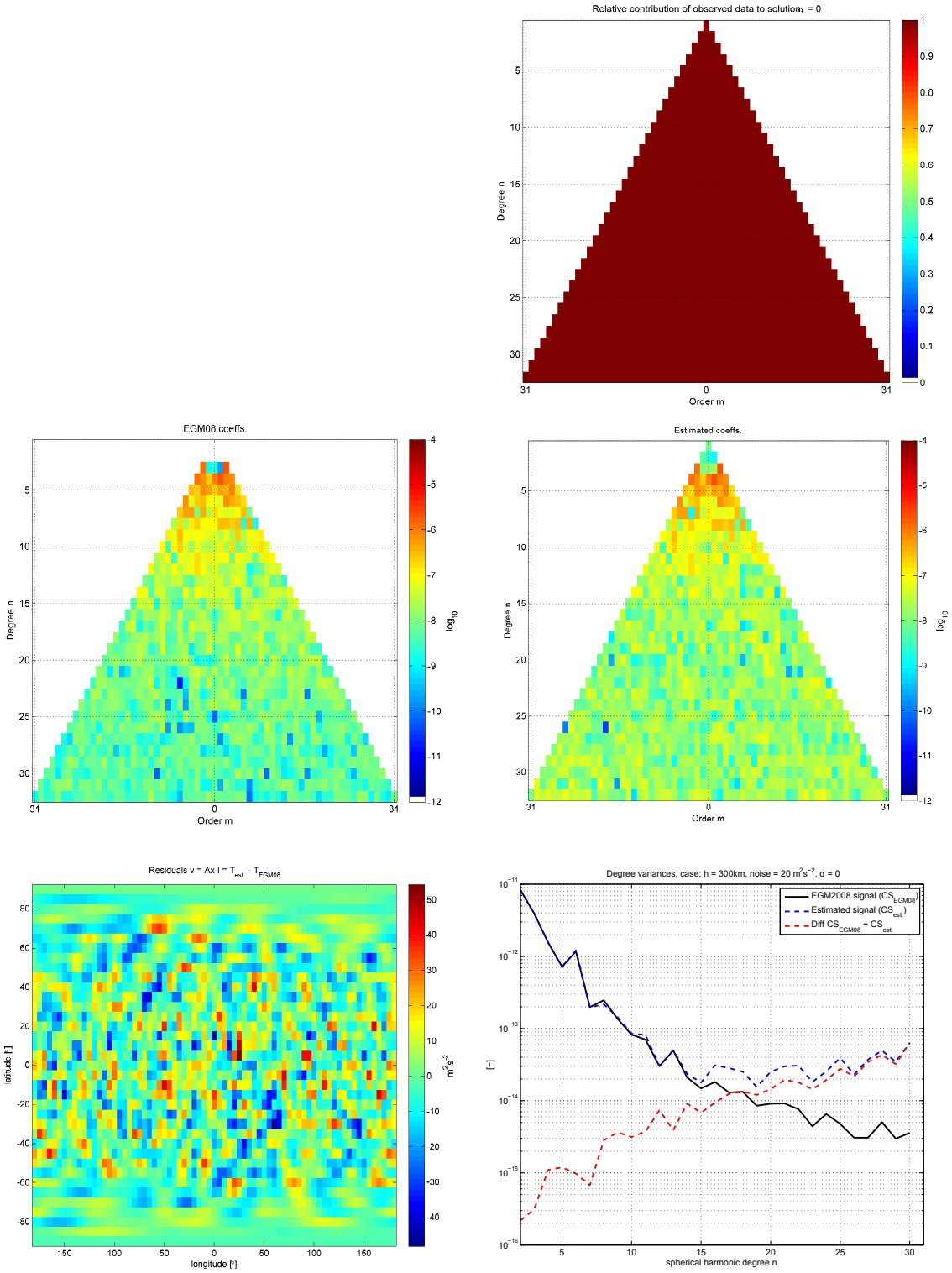

Figure 5.8: Spherical harmonic analysis. *Results from case 2,* $\sigma = 20\,\mathrm{m^2 s^{-2}}$ *and at* $h = 300\,\mathrm{km}$. *No regularization,* $\alpha = 0$.

the solution can be stabilized. The L-curve is computed for a range of α-values. To determine this range, some trial-and-erroring with regularization parameter values was done. A final α_{large} value was chosen such that it was not deemed meaningful with any values above α_{large}. In case 2, a regularization parameter $\alpha_{\text{large}} = 100$ was chosen, based on the results in Figure 5.9.

As expected, the regularized system gives a lower normal matrix condition number than in the previous case, cf. Table 5.1. This is a first indicator that the system is more stable. As mentioned earlier, the solution is stabilized through additional observations (that are zero and with accuracies according to the signal degree variances) for the corresponding degrees.

An important observation in Figure 5.9 is that this solution is too regularized. Looking at resolution matrices $\mathbf{R}_x$ and $\mathbf{R}_y$, regularization influences the solution already at degree 7, with a contribution of roughly 10 %. Already at degree 12, prior information contributes to the solution by 50 %. In the previous case, with no regularization, it was observed that the solution for the estimated coefficients up to roughly degree 15 was stable, and therefore needs no regularization. Therefore, in this case, regularization corrupts the well-conditioned part of the system.

The residuals are also of interest. Whilst trial-and-erroring with regularization parameters even larger than 100, it became evident that the residuals gradually adapts to the signal, and in extreme cases corresponds to the signal. This makes sense looking at the degree variances plot, since that which adapts to the signal degree variances curve are the *errors* of the solution. With a more correctly chosen value of α, the additional zero observations have a *weight* representing an appreciable influence only on the *unstable* higher degrees, as they adapt to the signal curve. In the case of Figure 5.9, also the stable degrees are influenced.

However, it is important to note that, although an error the size of the signal gives a better solution than a solution with unrealistically large errors, the obtained solution (for the influenced degrees) *is not necessarily correct*! It merely means that the errors do not surpass the signal. Correct coefficients are not obtained, but the solution is more stable.

The solution, as evident in the solution plot and degree variances plot, tends towards zero. It increasingly does so with increasing value of α. This is explained by the error degree variances moving upwards with increasing regularization weight — that is, the overall errors become larger. What has to approach zero are therefore the estimated coefficients, and with extreme regularization, the *error for every degree is as large as the signal*. This means that the solution itself approaches zero.

The extremities were underlined in case 3. The condition number of the normal matrix is so high that no sensible solution can be computed at all, cf. Table 5.1 and Figure 5.10. However, the observations are not particularly noisy at $\sigma = 0.1\,\mathrm{m^2 s^{-2}}$, and a relatively small regularization weight α was needed in order to see considerably different results.

The results of the regularized solution of case 3, with $\alpha_{\text{large}} = 1$, is presented in Figure 5.11. The situation is that at degree 7, a 70 % contribution to the solution comes from

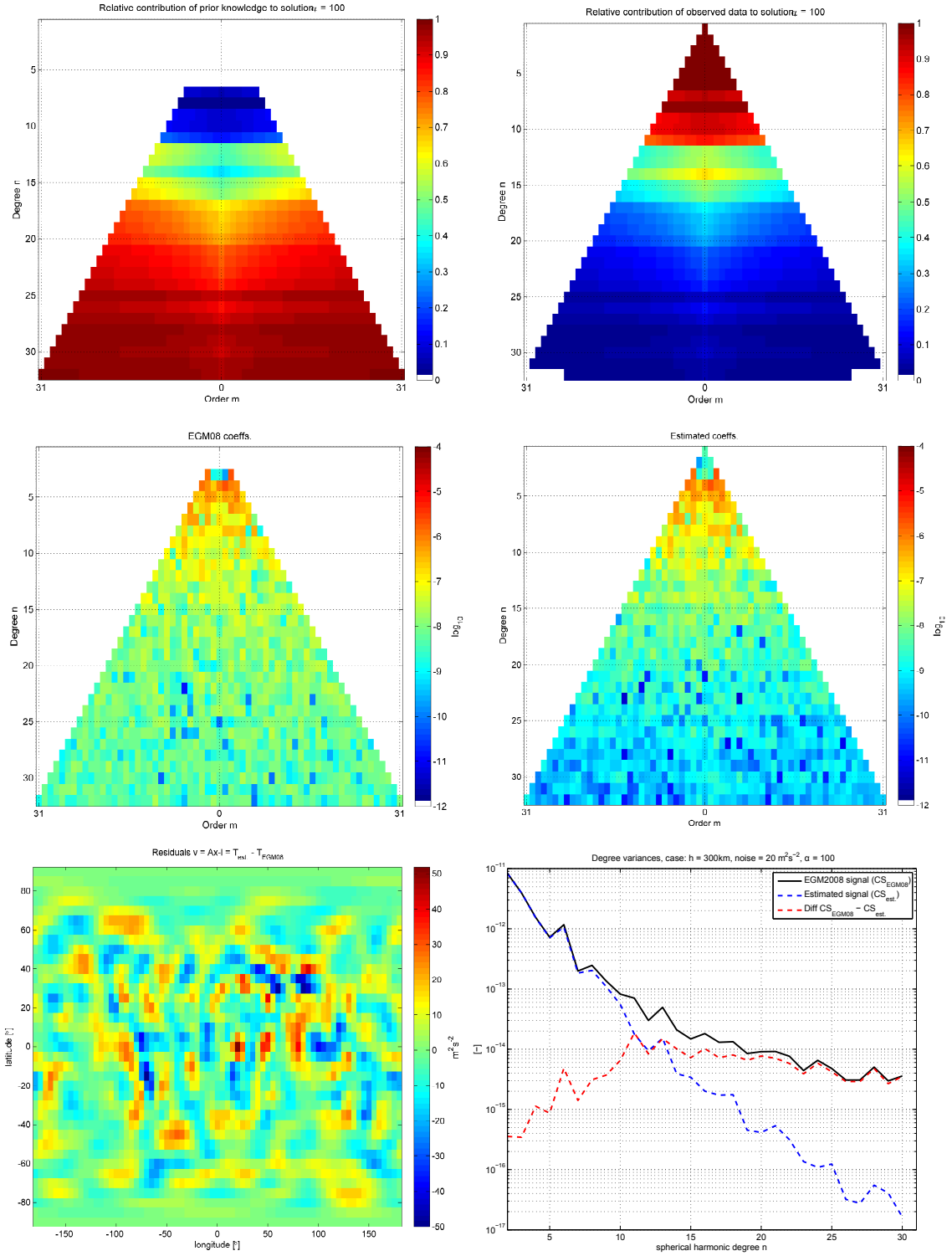

Figure 5.9: Spherical harmonic analysis. *Results from case 2,* $\sigma = 20\,\mathrm{m^2 s^{-2}}$ *and at* $h = 300\,\mathrm{km}$. *Regularization with* $\alpha_{large} = 100$.

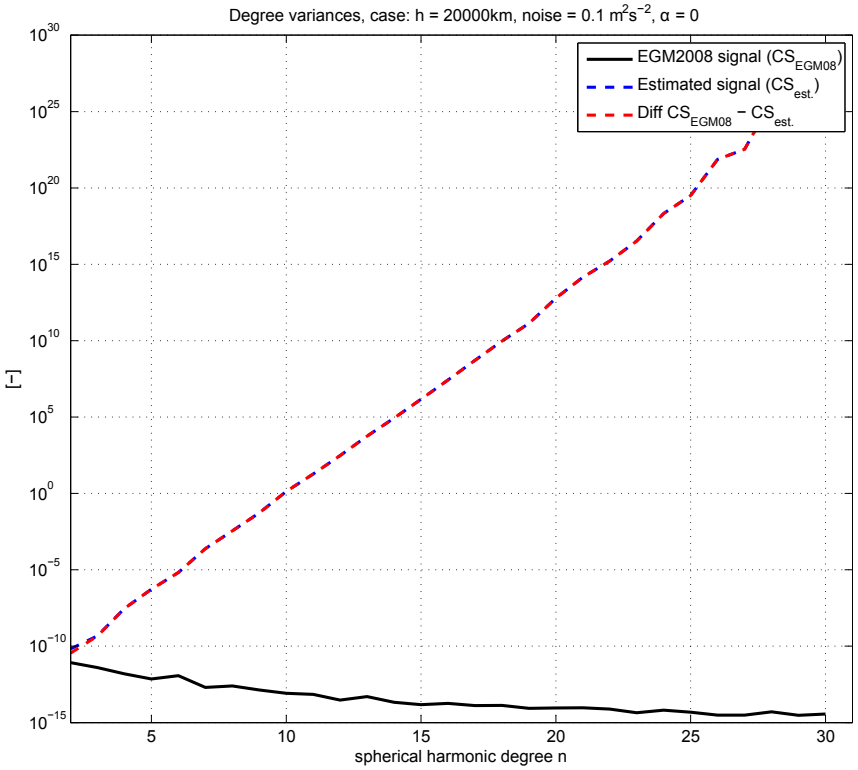

Figure 5.10: Spherical harmonic analysis. *Results from case 3,* $\sigma = 0.1\,\mathrm{m^2 s^{-2}}$ *and at* $h = 20000\,\mathrm{km}$. *No regularization,* $\alpha = 0$.

prior knowledge. A small amount of coefficients is estimable at all, evident in the solution plot and degree variances plot — the solution tends towards zero fast. The observation error is taking the form of the signal, cf. Figure 5.6, and the error degree variance curve corresponds to the signal curve above degree 7. The condition number is considerably reduced, cf. Table 5.1, and the solution stabilized and computable, but evidently not correct for the higher degrees. The observed signal is very smooth, cf. Figure 5.6. However, the residuals in Figure 5.11 presents more short-wavelength features, and the error degree variances curve is relatively small for degrees < 5, so the solution seems meaningful there.

	$\mathrm{cond}(\mathbf{N} + \alpha\mathbf{K}^{-1})$		
	Case 1	Case 2	Case 3
$\alpha = 0$	52	206	$3 \cdot 10^{38}$
α_{large}	–	25	$3 \cdot 10^{7}$
$\alpha_{\text{L-curve}}$	–	37	–

Table 5.1: Spherical harmonic analysis. *Condition numbers for the regularized normal matrix* $\mathbf{N} + \alpha\mathbf{K}^{-1}$ *for cases 1, 2 and 3.*

An attempt to find the best regularization parameter in case 2 using the L-curve approach was made. As mentioned in the above, the range of α was determined from trial regularizations with different α values. The curvature approach for choosing alpha was chosen, where a polygon was fitted to the point values, and the subsequent maximum analytical curvature computed from the polygon. The L-curve is shown in Figure 5.12. A regularization parameter $\alpha_{\text{L-curve}} = 15.6$ was determined using this approach, and the results of regularization with this parameter are found in Figure 5.13.

The results using the regularization parameter $\alpha_{\text{L-curve}}$ are quite good. Up to degree 7, the contribution from actual observations is at 100 %, and stays at 80-90 % up to roughly degree 16. The stable degrees are left practically "untouched". Above degree 20, regularization contributes more than the observations, and the solution there consequently is less correct.

One could argue that, since case 2 is not extremely ill-conditioned and is quite solvable without regularization, it is important not to regularize too much. For comparison with the regularization parameter chosen from the L-curve, case 2 was solved also with $\alpha = 10$ and $\alpha = 20$. Their degree variances plots are shown in Figure 5.14.

With $\alpha = 10$, contribution from regularization is notable from degree 11 and upwards, and the solution does not tend to zero as quickly as the cases of $\alpha_{\text{L-curve}} = 15.6$ and $\alpha = 20$. The error degree variances curve, however, surpasses the signal degree variances curve for the higher degrees.

As for $\alpha = 20$, contribution from regularization is notable from degree 7. The error degree variances curve practically never reaches the signal degree variances curve. The case of $\alpha_{\text{L-curve}} = 15.6$ is a compromise between these two. Most importantly, the two latter

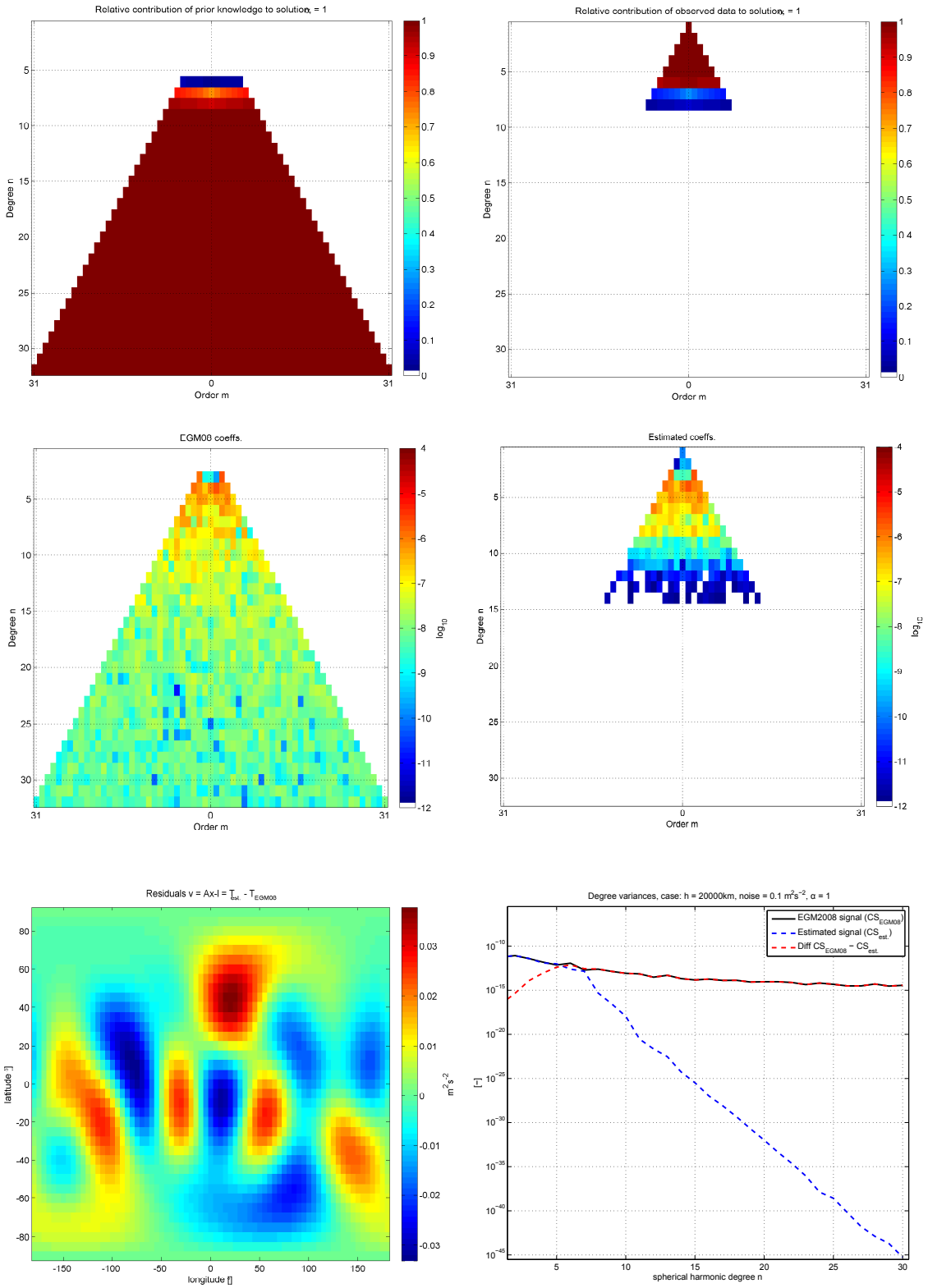

Figure 5.11: Spherical harmonic analysis. *Results from case 3,* $\sigma = 0.1\,\mathrm{m^2 s^{-2}}$ *and at* $h = 20000\,\mathrm{km}$. *Regularization with* $\alpha_{large} = 1$.

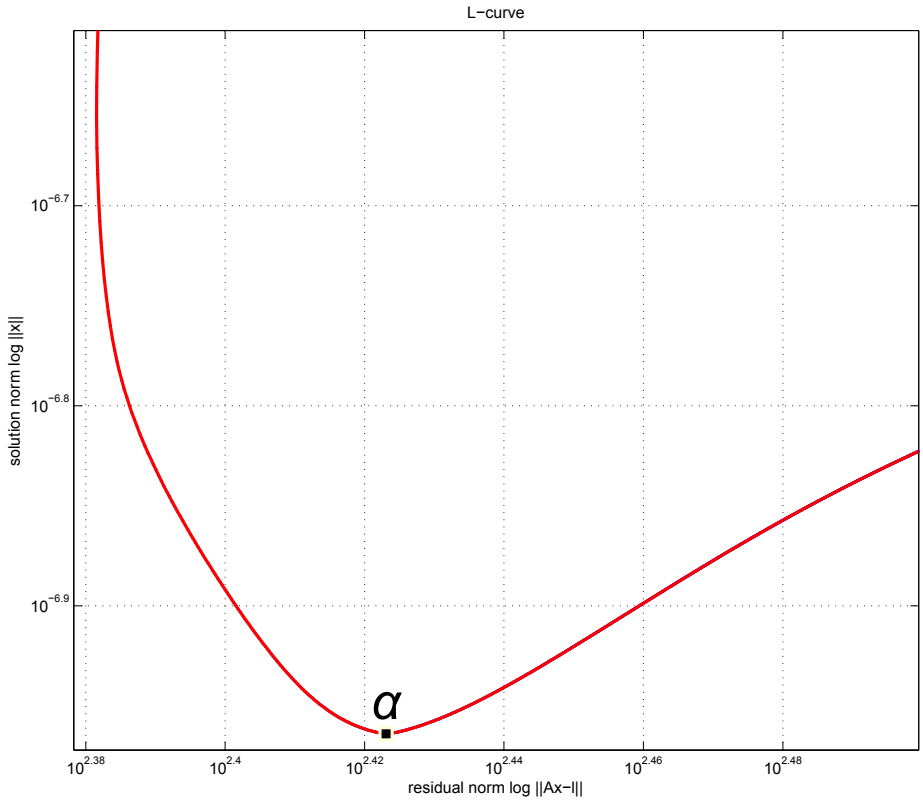

Figure 5.12: Spherical harmonic analysis. *L-curve for case 2, corner found by the curvature approach.*

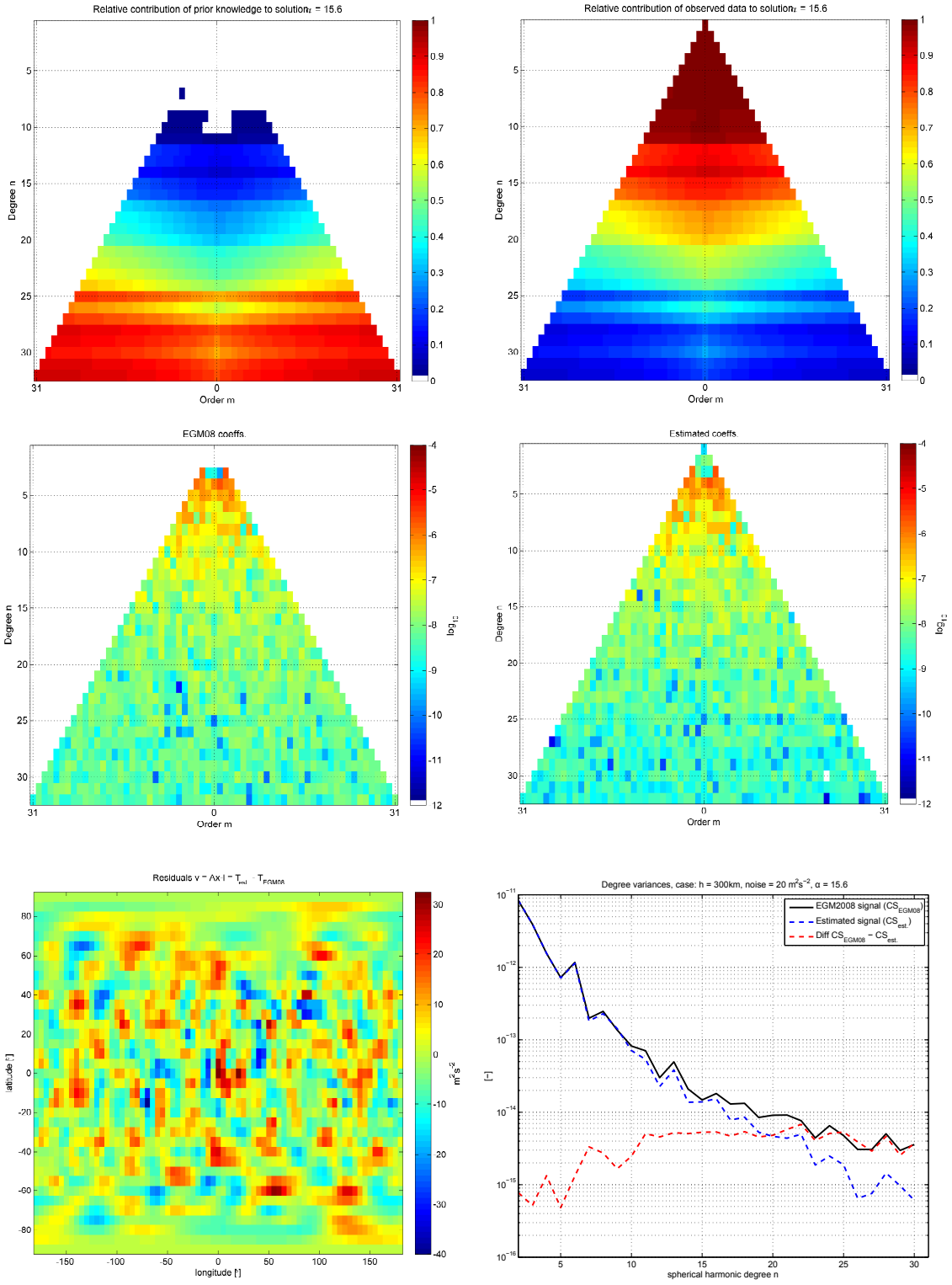

Figure 5.13: Spherical harmonic analysis. *Results from case 2,* $\sigma = 20\,\mathrm{m^2 s^{-2}}$ *and at* $h = 300\,\mathrm{km}$. *Regularization with* $\alpha_{L\text{-}curve} = 15.6$.

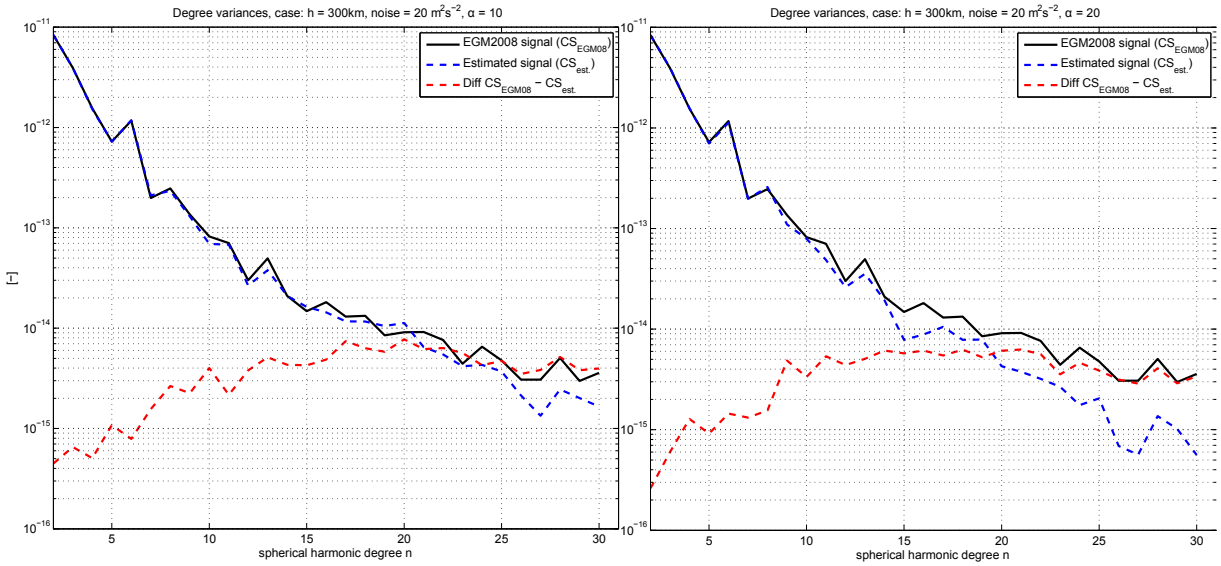

Figure 5.14: Spherical harmonic analysis. *Results from case 2,* $\sigma = 20\,\mathrm{m^2 s^{-2}}$ *and at* $h = 300\,\mathrm{km}$. *Regularization with* $\alpha = 10$ *and* $\alpha = 20$.

choices of alpha also give relatively good results (and are in the vicinity of the L-curve corner).

5.4 Agreement check of the different geoid modeling approaches with synthetic data

Theoretical comparisons between spherical harmonics and other gravity field approaches were conducted in section 3.5, and in this section the goal is to make numerical comparisons using a synthetic gravity field.

If both synthetic observations (e.g., in the form of gravity anomalies) are created and a geoid computed from the same GGM, the gravity anomalies and the geoid will be self-consistent, allowing closed-loop testing of the other geoid determination approaches presented in this thesis. To a great extent, the geoids obtained by the other approaches (using the gravity anomalies as observations) should agree with the synthetic geoid computed from the GGM.

One disadvantage of using a GGM for synthetic data construction is that the synthetic data thus is band-limited, which does not mirror reality. However, with a high-resolution model such as *EGM2008*, the regional gravity field is covered relatively well. Advantages are for instance that the comparison data is in no way affected by noise and we do not have to consider any effects of topography and alike.

5.4.1 Creation of synthetic data sets

With aid of degree variance models, the RMS (Root-Mean-Square, square root of the variance, cf. section 2.5) value of the the omission (or truncation) error (N_3) can be quantified. Assuming the general practical evaluation approach combining long-wavelength data from a GGM with short-wavelength terrestrial observations, we can adjust the size of the *near* zone σ_0 (i.e., alter ψ_0 such that the omission error becomes negligible (or allow further minimization by kernel modifications)). In other words, we can get an impression of how large our observation data area has to be in order to obtain a negligible N_3 term.

A covariance function could be created from analyzing the regional input data. Alternatively, the RMS value of the omission error can be formulated with aid of spherical harmonic degree variances as follows (Gerlach, 2009):

$$\mathrm{RMS}\{N_3\} = \sqrt{\sum_{n_\Delta}^{5000} c_n \lambda_n^2} \tag{5.14}$$

where c_n are the dimensionless degree variances, computed from a GGM or a degree variance model, and λ_n are the *spectral eigenvalues*. 5000 is an arbitrary high figure, to make sure all significant parts of the signal are included.

Attention may further be turned to equation (3.16), relating spherical harmonic degree to the spatial resolution. It may be rewritten as follows:

$$n_\Delta = \frac{20000\,\mathrm{km}}{\Delta} \tag{5.15}$$

where Δ is the extent of the area for which the geoid is to be determined. Then the value n_Δ states the largest resolvable wavelength (or smallest frequency) in that area. For instance, the different regions considered in this thesis range between approximately 150 km to 450 km, yielding the approximate spectral range $40 \leq n_\Delta \leq 130$. Thus, theoretically, a synthetic observation data set created from a GGM could be based on the spherical harmonic spectral window $[n_\Delta\ ,\ N_{\max}]$.

In the synthetic investigations in this thesis, the observation data are gravity anomalies created from the GGM *EGM2008*, with $N_{\max} = 2190$. Further, we say that the long wavelengths are best determinable not from terrestrial gravity measurements, but from a GGM such as *GOCO03s*, covering spherical harmonic degrees up to $N_{\max} = 250$. The synthetic observation data could therefore be created with $n_\Delta = 251$. This would further comply with a real-life situation, since, within near future, the combined *GOCO* models will be superior with regard to representing the long wavelength part of the gravity field accurately.

When computing equation (5.14), the exact spectral range $251 \leq n \leq 2190$ of the input data could be used (since it is known). However, included in *EGM2008* is the effect

of topography. If we assume that topographic reductions have been performed, *Flury*'s model is valid for degrees $n > 400$, cf. equation (2.52) (Gerlach and Pettersen, 2010).

Also, combinations of degree variance models could be used. In Figure 5.15, the signal degree variances of *EGM2008* as well as the models *Flury*, *Tscherning/Rapp* (cf. equation (2.51)), and *Kaula* (cf. equation (2.50)) is plotted. A combination of, e.g., *Tscherning/Rapp* ($251 \leq n \leq 399$) and *Flury* ($400 \leq n \leq 5000$) could be made, cf. Figure 5.15, with the argument that *Flury*'s model fits the higher frequencies of the gravity field better and implicitly assumes topographic reductions have been carried out. Also observable in Figure 5.15 is that the signal degree variances of *EGM2008* are lower than those of *Flury* for approximately $n < 700$. Thus, a combination of *EGM2008* ($251 \leq n \leq 699$) and *Flury* ($700 \leq n \leq 5000$) could also be appropriate.

The RMS values could be useful when creating *synthetic* observations. In tables 5.3 and 5.4 different computational (target) areas are defined, and the synthetic observations are given for somewhat larger areas, theoretically defined by considering the above equation (i.e., how large must the observation data set be to get a small omission error). In Table 5.2 RMS values of N_3 for different near zone caps defined by ψ_0 are computed.

	RMS$\{N_3\}$ (cm)		
ψ_0 (°)	$\mathrm{E}_{251-2190}$	$\mathrm{E}_{251-699} + \mathrm{F}_{700-5000}$	$\mathrm{T/R}_{251-399} + \mathrm{F}_{400-5000}$
1	6.0	6.0	7.8
2	4.3	4.3	5.6
3	3.4	3.4	4.5
4	3.4	3.4	4.4
5	2.8	2.8	3.7
10	2.0	2.0	2.6
15	1.5	1.5	2.0
20	1.1	1.1	1.5
30	0.5	0.5	0.6

Table 5.2: *Omission error for various spherical caps σ_0.* E *is* EGM2008*,* F Flury*'s model and* T/R Tscherning/Rapp*'s model.*

From Table 5.2 it is clear that, theoretically, to preserve a small omission error, our observation area must be relatively large. For practical computation reasons an inner zone with $\psi_0 = 1°$ is chosen, and an awareness of the theoretical omission error of approximately 6 cm is kept. The absolute accuracy of the geoid computed in this configuration would consequently not be on sub-centimeter level, but this is irrelevant for the goal of the numerical investigations in this thesis, which is to compare different computation methods. The omission error is equal for all methods, so we rather look at the accuracy *between* the different methods.

Also, for computational reasons, a general regional observation data resolution of $3'$ is used.

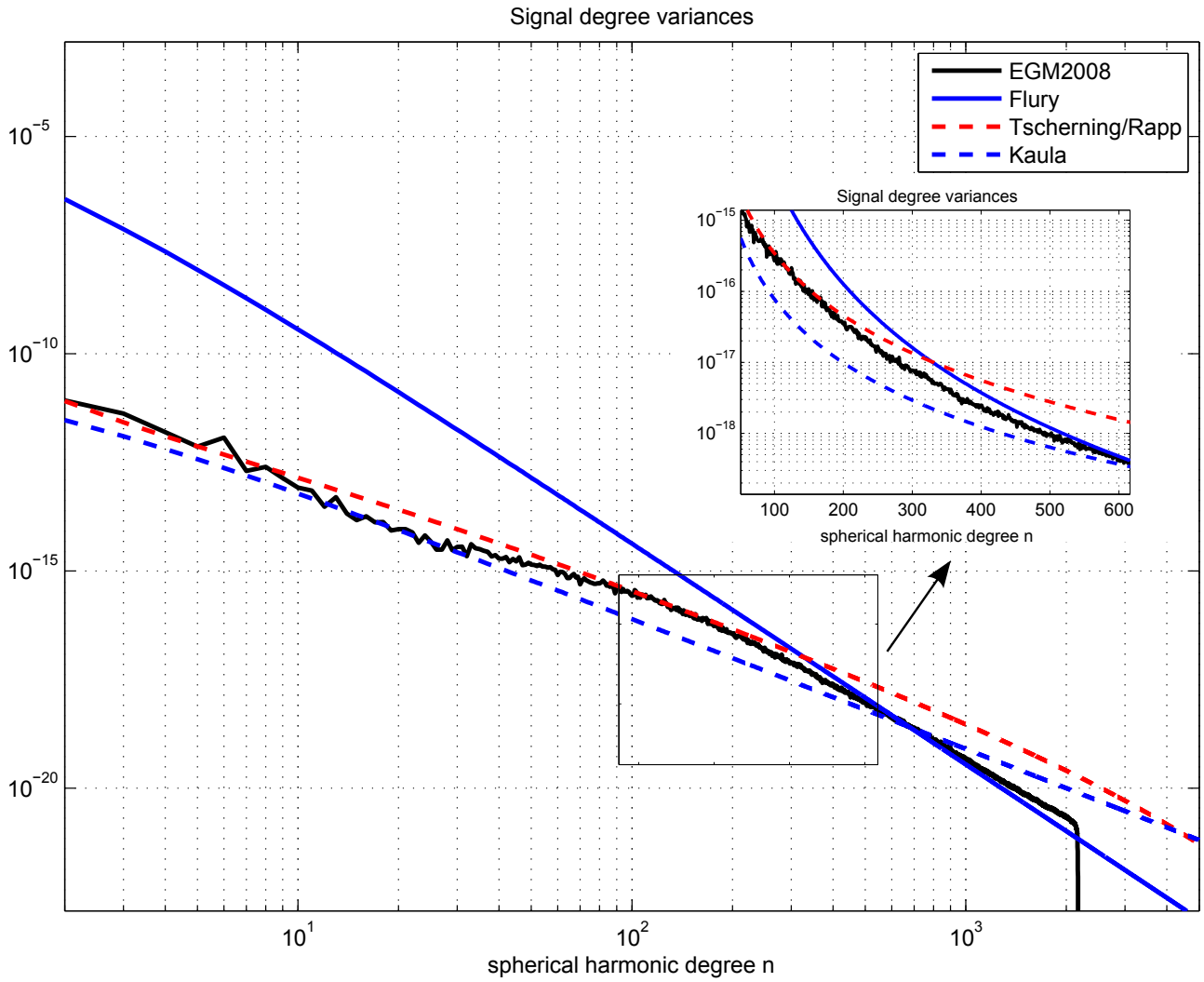

Figure 5.15: *Signal degree variances c_n computed from the GGM* EGM2008 *and the models of* Flury, Tscherning/Rapp *and* Kaula

Since the observation data are gravity anomalies created from the GGM *EGM2008*, the geoid and gravity anomalies are created from spherical harmonic synthesis, cf. Table 3.1, in the spectral range $251 \leq n \leq 2190$:

$$\Delta g_{\mathrm{SHS}} = \frac{GM}{r^2} \sum_{n=251}^{2190} \lambda_n^{\mathrm{upw}} (n-1) \sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta) \left[\Delta \bar{C}_{nm}^{\mathrm{EGM2008}} \cos m\lambda + \bar{S}_{nm}^{\mathrm{EGM2008}} \sin m\lambda \right] \tag{5.16}$$

$$N_{\mathrm{SHS}} = \frac{GM}{r\gamma_0} \sum_{n=251}^{2190} \lambda_n^{\mathrm{upw}} \sum_{m=0}^{n} \bar{P}_{nm}(\cos\theta) \left[\Delta \bar{C}_{nm}^{\mathrm{EGM2008}} \cos m\lambda + \bar{S}_{nm}^{\mathrm{EGM2008}} \sin m\lambda \right] \tag{5.17}$$

Further, we can confirm the computed RMS$\{N_3\}$-values in Table 5.2. The geoid was computed from equation (5.17), once with $\psi_0 = 180°$ and once with $\psi_0 = 1°$. The difference between these two geoids yields the omission (or truncation) error, i.e., the part of the geoid *outside* the spherical cap with radius ψ_0. The RMS value of this difference was computed as follows:

$$\mathrm{RMS}\{N_{SHS}(\psi_0 = 180°) - N_{SHS}(\psi_0 = 1°)\} \approx 6.1\,\mathrm{cm} \tag{5.18}$$

Results of the synthetic gravity anomaly and geoid computations are summarized in tables 5.3 and 5.4 as well as figures 5.16 and 5.17.

	Computation (target) area Δg observation (data) area	$53° \leq \varphi \leq 54°$ and $6.5° \leq \lambda \leq 8.5°$ $52° \leq \varphi \leq 55°$ and $5.5° \leq \lambda \leq 9.5°$		
	max.	min.	mean	RMS
N_{SHS}	0.073	-0.084	0.021	0.043
Δg_{SHS}	11.733	-6.904	0.433	3.308

Table 5.3: *East Frisia, target and data area, 3′ resolution. Statistics of the synthetic gravity anomalies (*mGal*) and geoid (*m*).*

5.4.2 Agreement of Stokes numerical integration with spherical harmonic synthesis

The geoid was computed from the synthetic gravity fields of the different regions by means of 1D-FFT, cf. section 4.2.7. The results are summarized in tables 5.5 and 5.6 as well as figures 5.18 and 5.19.

	Computation (target) area Δg observation (data) area	$47° \leq \varphi \leq 48°$ and $9° \leq \lambda \leq 12°$ $46° \leq \varphi \leq 49°$ and $8° \leq \lambda \leq 13°$		
	max.	min.	mean	RMS
N_{SHS}	0.410	-0.355	-0.016	0.141
Δg_{SHS}	47.755	-63.217	0.218	12.110

Table 5.4: *Alpine region, target and data area, 3′ resolution. Statistics of the synthetic gravity anomalies (*mGal*) and geoid (*m*).*

		max.	min.	mean	RMS
	N_{Stokes}	0.071	-0.107	0.018	0.044
Data area	$N_{\mathrm{SHS}} - N_{\mathrm{Stokes}}$	0.144	-0.072	0.014	0.037
Target area	$N_{\mathrm{SHS}} - N_{\mathrm{Stokes}}$	0.027	-0.015	0.002	0.007

Table 5.5: East Frisia. *Statistics of the geoid computed from Stokes's formula and the difference between the geoids (*m*).*

		max.	min.	mean	RMS
	N_{Stokes}	0.405	-0.353	-0.013	0.140
Data area	$N_{\mathrm{SHS}} - N_{\mathrm{Stokes}}$	0.446	-0.371	0.023	0.085
Target area	$N_{\mathrm{SHS}} - N_{\mathrm{Stokes}}$	0.010	-0.090	-0.003	0.010

Table 5.6: Alpine region. *Statistics of the geoid computed from Stokes's formula and the difference between the geoids (*m*).*

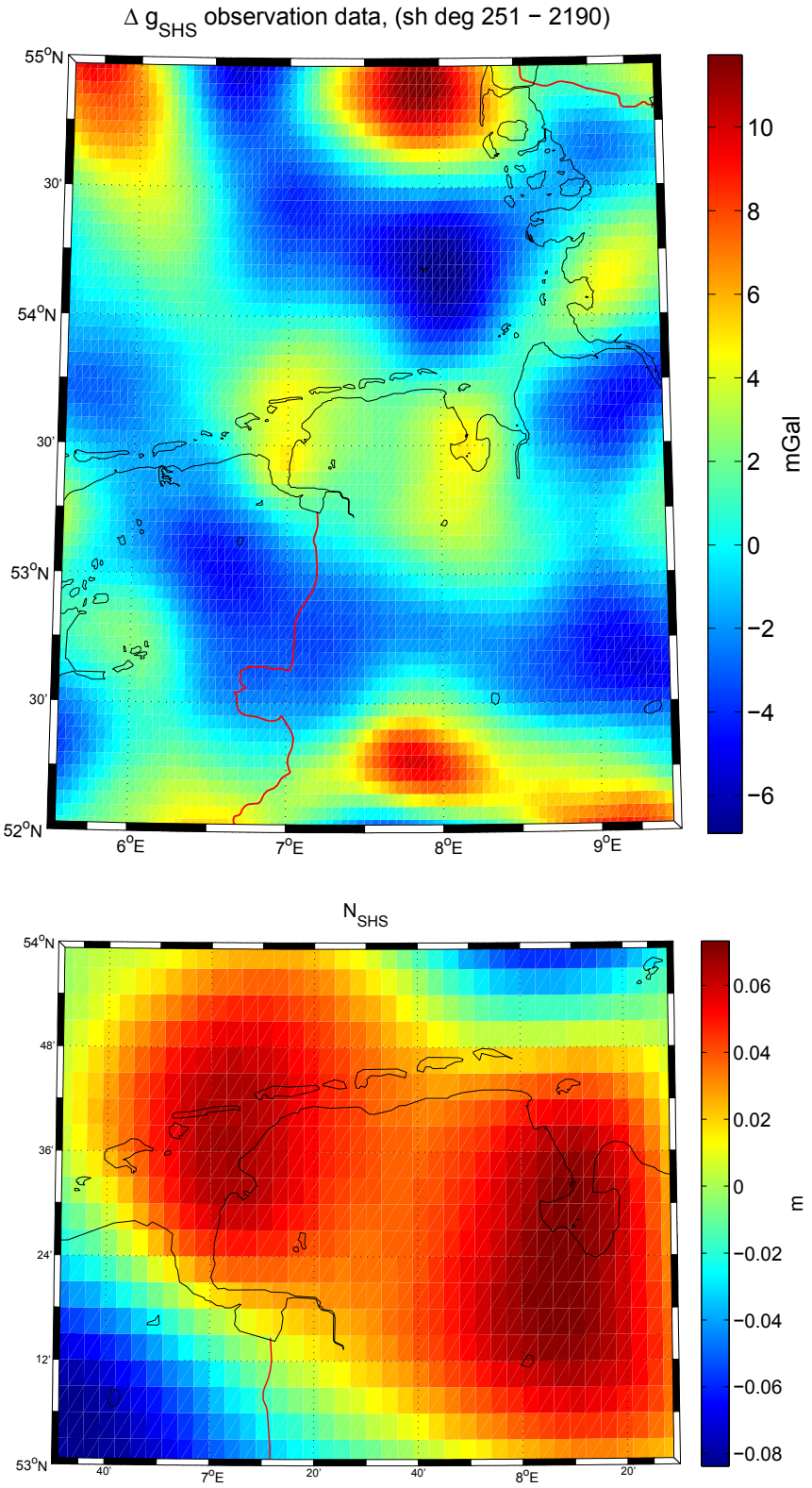

Figure 5.16: East Frisia. *Synthetic gravity anomalies and geoid computed from spherical harmonic synthesis.*

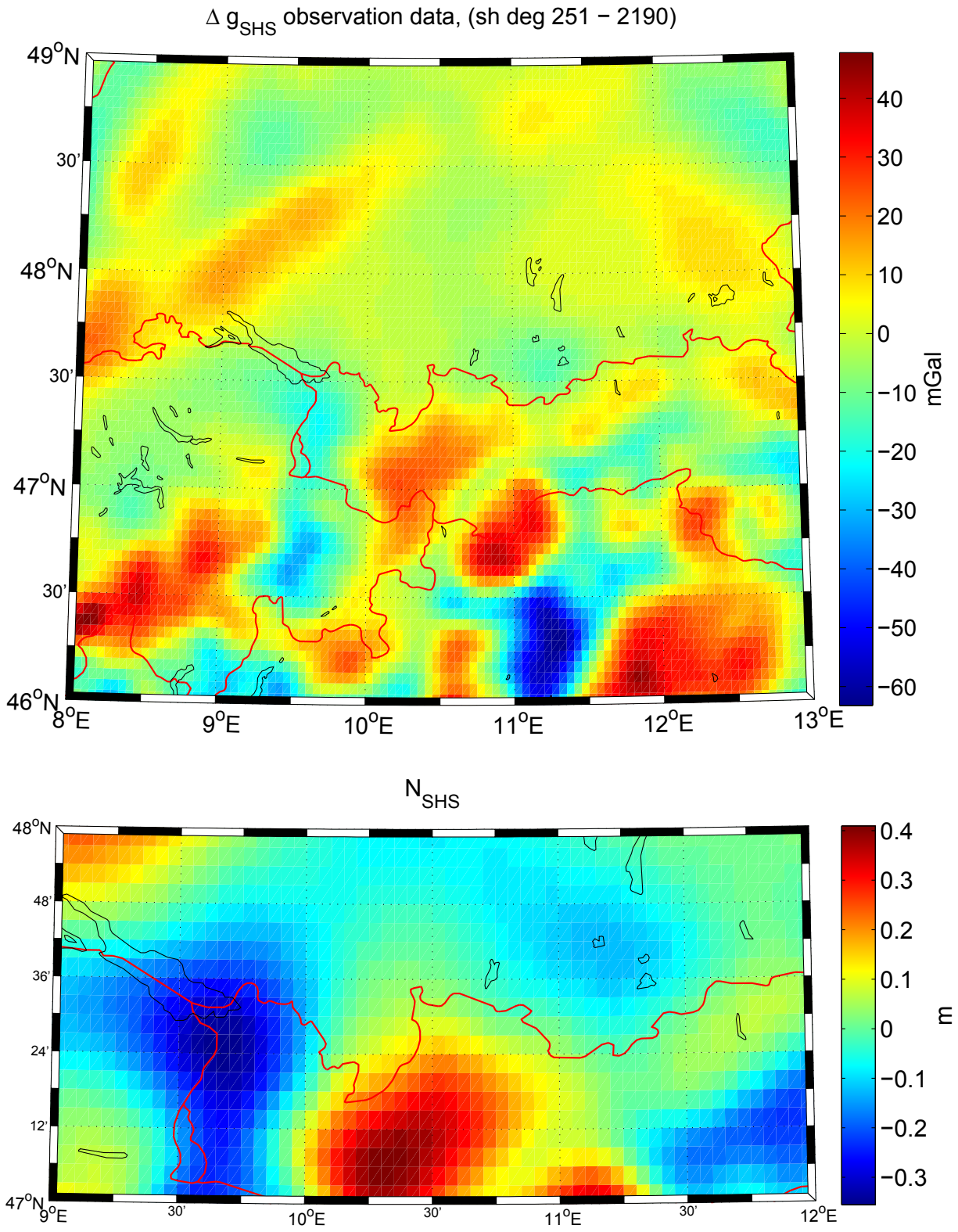

Figure 5.17: Alpine region. *Synthetic gravity anomalies and geoid computed from spherical harmonic synthesis.*

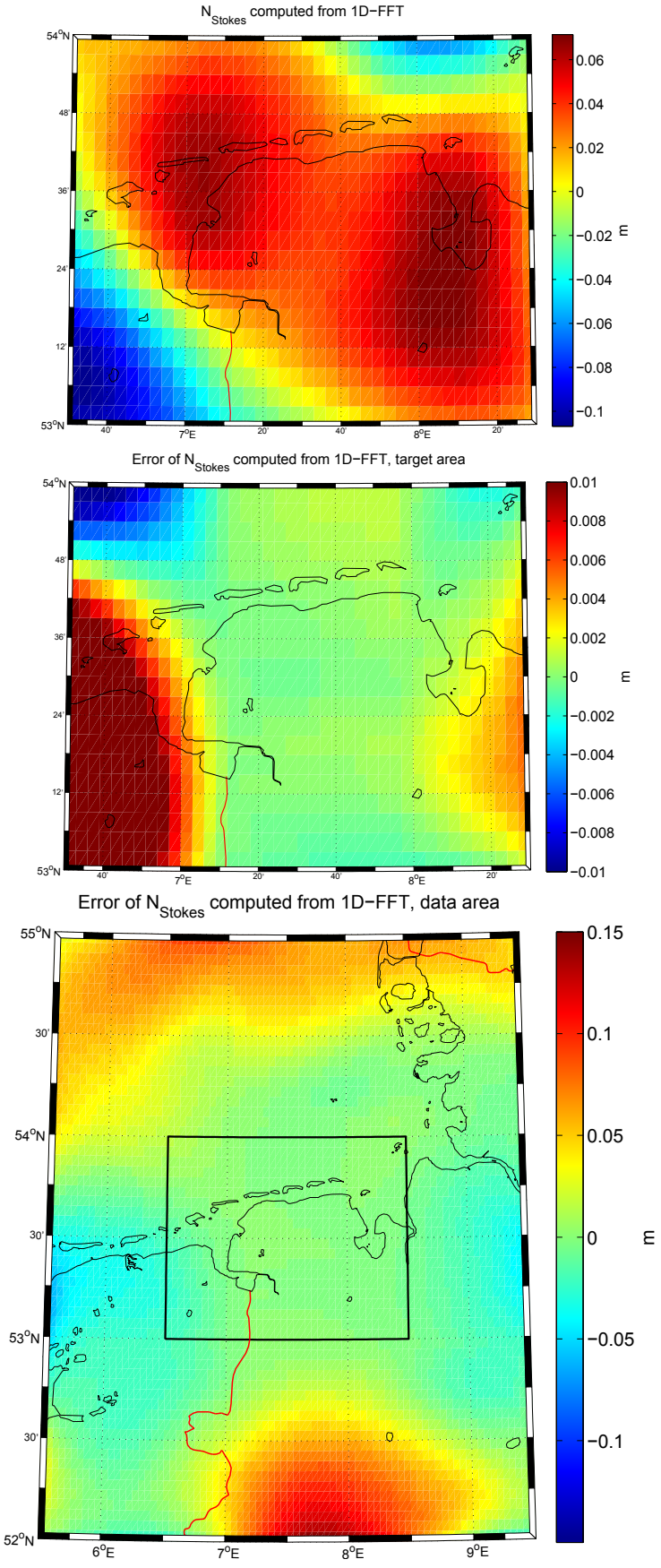

Figure 5.18: Stokes integration by 1D-FFT. *Results from geoid computation and comparison, East Frisia.*

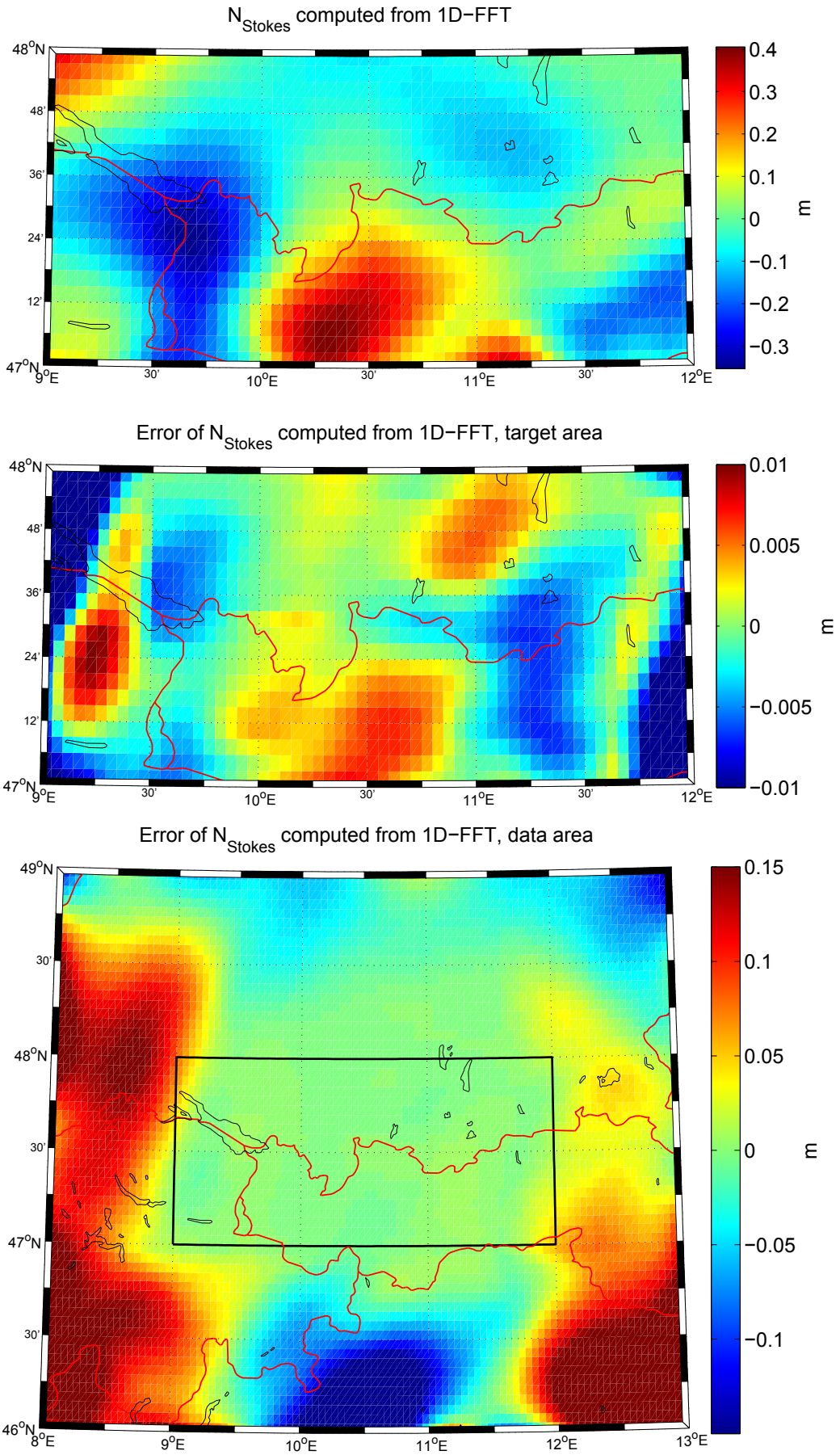

Figure 5.19: Stokes integration by 1D-FFT. *Results from geoid computation and comparison, Alpine region.*

Extending the data area

Observing figures 5.18 and 5.19 it is clear that, with the above choice of data area, there is, to some extent, a negative influence on the accuracies due to edge effects. Extending the data area to remove these edge effects should yield an accuracy improvement. Since it is computationally more demanding, the numerical investigation of accuracy improvement by extending the data area with 1° in all directions is restricted to the case of Stokes integration versus spherical harmonic synthesis. The improved accuracies are presented in table 5.7 and the extended data areas may be seen in Figure 5.20.

	East Frisia		Alpine region	
	Data area	Target area	Data area	Target area
	$51° \leq \varphi \leq 56°$	$53° \leq \varphi \leq 54°$	$45° \leq \varphi \leq 50°$	$47° \leq \varphi \leq 48°$
	$4.5° \leq \lambda \leq 10.5°$	$6.5° \leq \lambda \leq 8.5°$	$7° \leq \lambda \leq 14°$	$9° \leq \lambda \leq 12°$
RMS$\{N_{\text{SHS}} - N_{\text{Stokes}}\}$ (m)	0.0162	0.0005	0.0664	0.0031

Table 5.7: East Frisia *and* Alpine region. *Accuracy improvement by extending the data area with* 1° *in all directions.*

5.4.3 Agreement of least-squares collocation with Stokes integration and spherical harmonic synthesis

The geoid was computed from the synthetic gravity fields of the different regions by means of least-squares collocation, cf. sections 3.3.2 and 3.5.1.

To underline the importance of modifying the covariance function as suggested by de Min (1995), the geoids for East Frisia and the Alpine region are computed with the *original* covariance function and then computed with the *modified* covariance function. The geoids computed with collocation are then compared with the East Frisian and Alpine region geoids computed from spherical harmonic synthesis, i.e., $N_{SHS}(\psi_0 = 1°)$. This corresponds to Stokes integration in a spherical cap with radius $\psi_0 = 1°$. The unmodified covariance function corresponds to global Stokes integration (with extrapolation). Note that $\mathbf{C}^{N\Delta g} = \mathbf{C}^{N^L\Delta g}(\psi_0 = 180°)$. The modified covariance function should also correspond to Stokes integration in a spherical cap with radius $\psi_0 = 1°$.

The cross-covariance functions can be seen in Figure 5.21, and the differences $N_{SHS}(\psi_0 = 1°) - N_{LSC}(\psi_0 = 180°)$ and $N_{SHS}(\psi_0 = 1°) - N_{LSC}(\psi_0 = 1°)$ for East Frisia in Figure 5.22 and the Alpine region in Figure 5.23. The RMS values of the differences are listed in Table 5.8. The results of the geoid computations are summarized in tables 5.9 and 5.10 as well as figures 5.24 and 5.25.

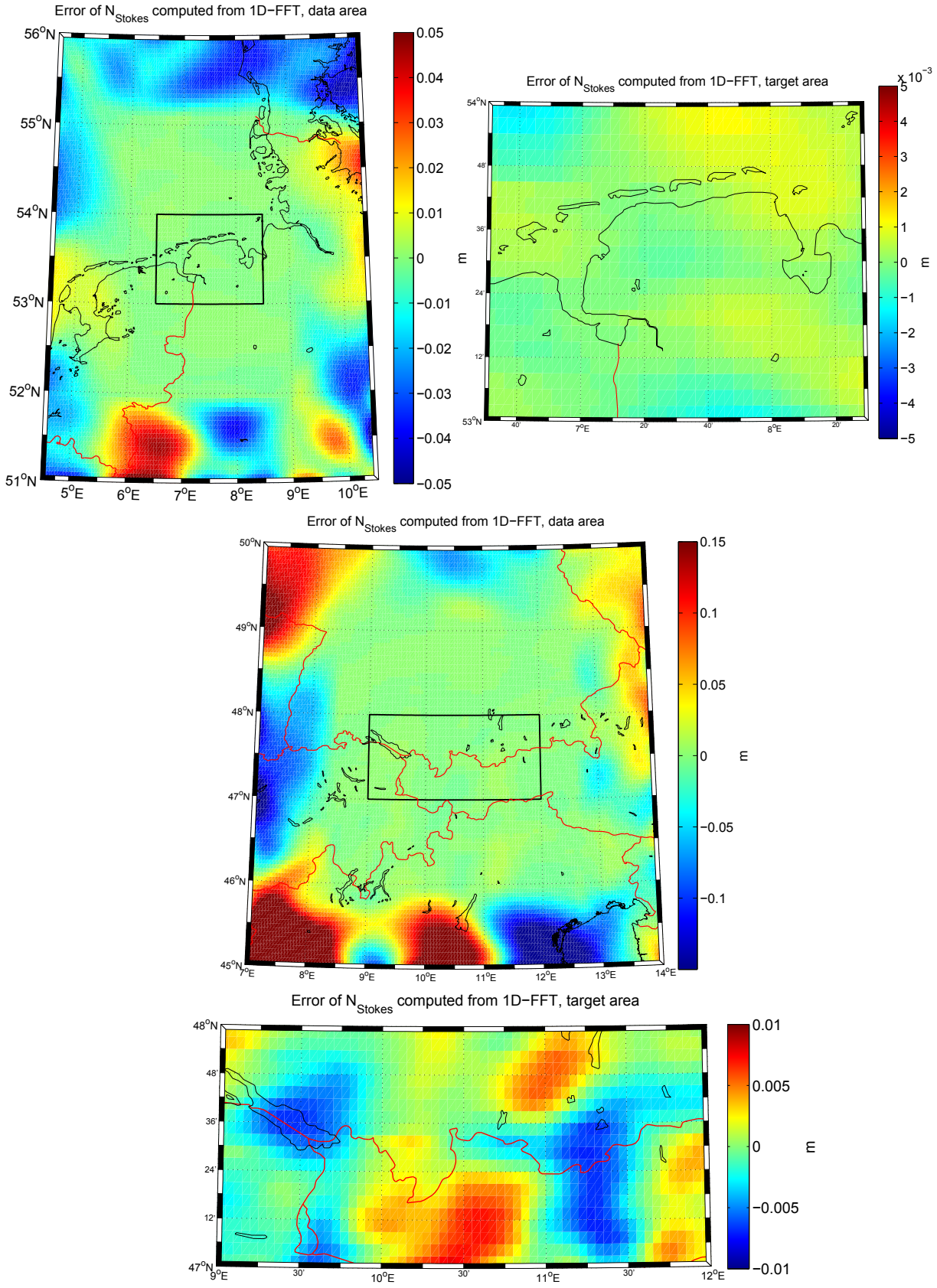

Figure 5.20: East Frisia *and* Alpine region. *Extended data area with* 1° *in all directions.*

	RMS (m)	
	East Frisia	Alpine region
Unmodified covariance function $N_{SHS}(\psi_0 = 1°) - N_{LSC}(\psi_0 = 180°)$	0.028	0.061
Modified covariance function $N_{SHS}(\psi_0 = 1°) - N_{LSC}(\psi_0 = 1°)$	0.015	0.014

Table 5.8: East Frisia *and* Alpine region. *Comparison of unmodified and modified covariance function with Stokes integration in a spherical cap with radius* $\psi_0 = 1°$

		max.	min.	mean	RMS
	N_{Stokes}	0.073	-0.100	0.018	0.043
Data area	$N_{\text{SHS}} - N_{\text{LSC}}$	0.034	-0.069	-0.001	0.015
Target area	$N_{\text{SHS}} - N_{\text{LSC}}$	0.018	-0.000	0.003	0.005

Table 5.9: East Frisia. *Statistics of the geoid computed by least-squares collocation and the difference between the geoids (*m*).*

		max.	min.	mean	RMS
	N_{Stokes}	0.410	-0.355	-0.016	0.141
Data area	$N_{\text{SHS}} - N_{\text{LSC}}$	0.054	-0.059	0.000	0.014
Target area	$N_{\text{SHS}} - N_{\text{LSC}}$	0.032	-0.035	0.000	0.006

Table 5.10: Alpine region. *Statistics of the geoid computed by least-squares collocation and the difference between the geoids (*m*).*

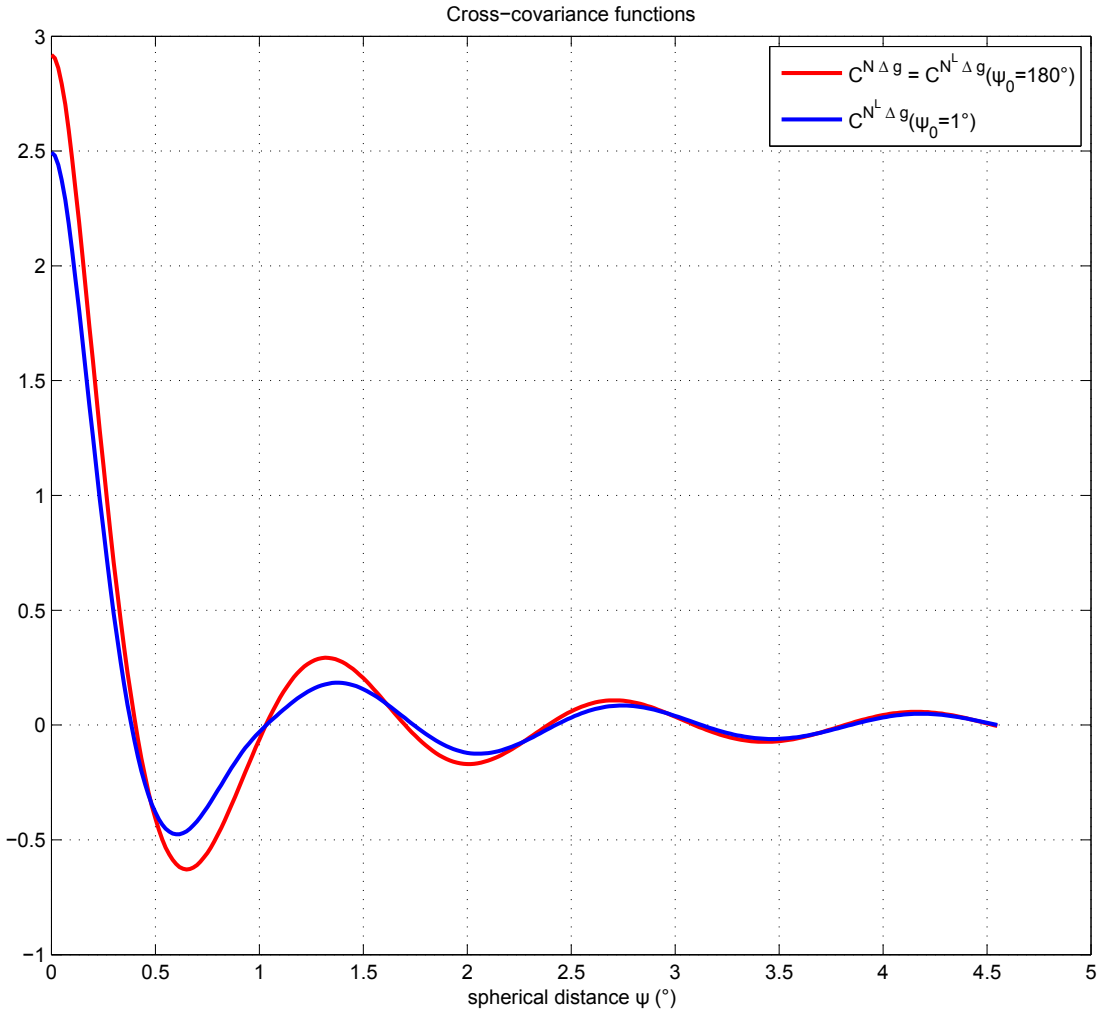

Figure 5.21: *Unmodified and modified cross-covariance functions $C^{N\Delta g}$ and $C^{N^L\Delta g}$ for regional geoid computation.*

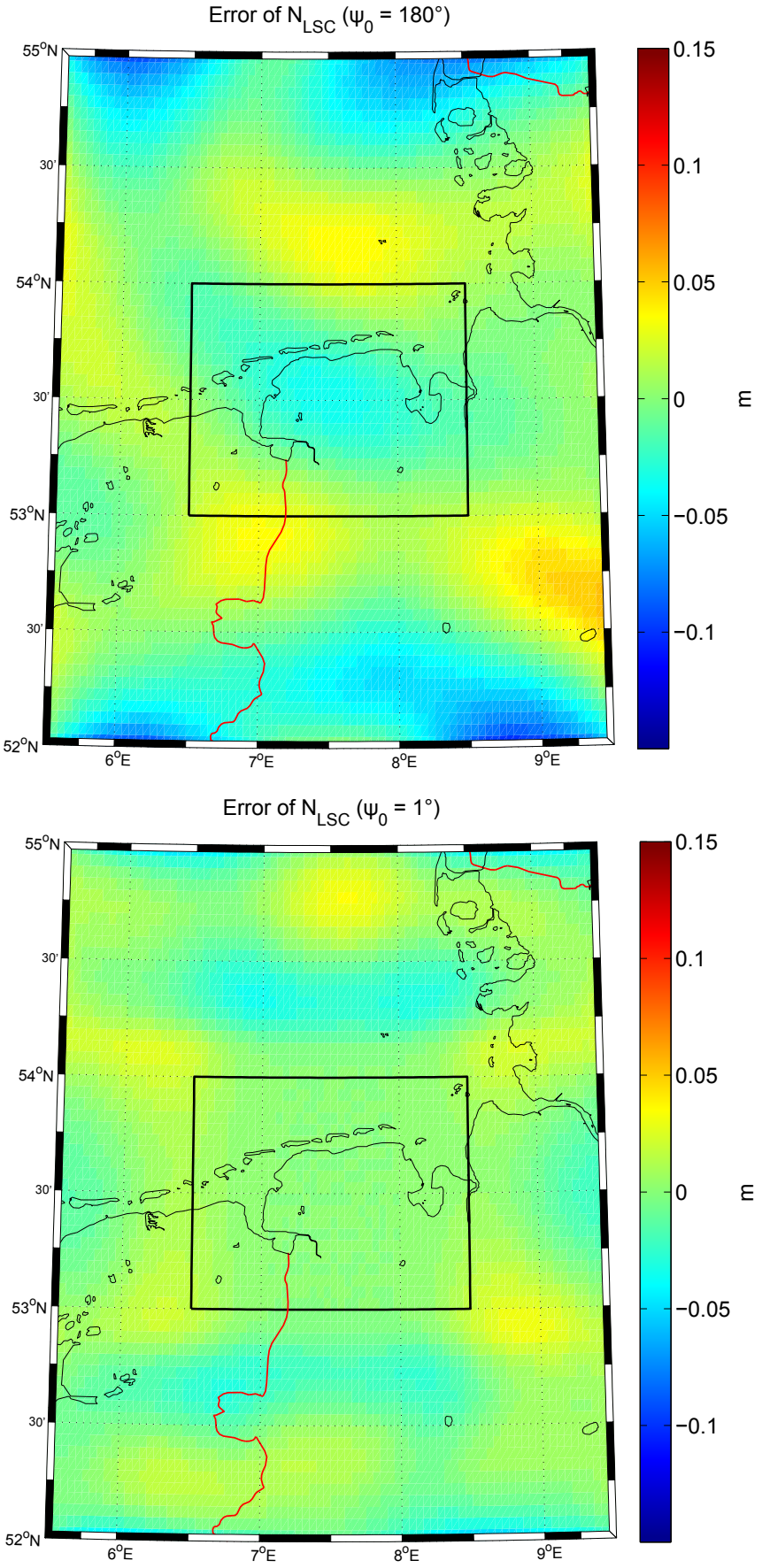

Figure 5.22: East Frisia*, data area. Comparison of unmodified and modified covariance function with Stokes integration in a spherical cap with radius $\psi_0 = 1°$. Note that $N_{LSC}(\psi_0 = 180°)$ corresponds to the original unmodified covariance function.*

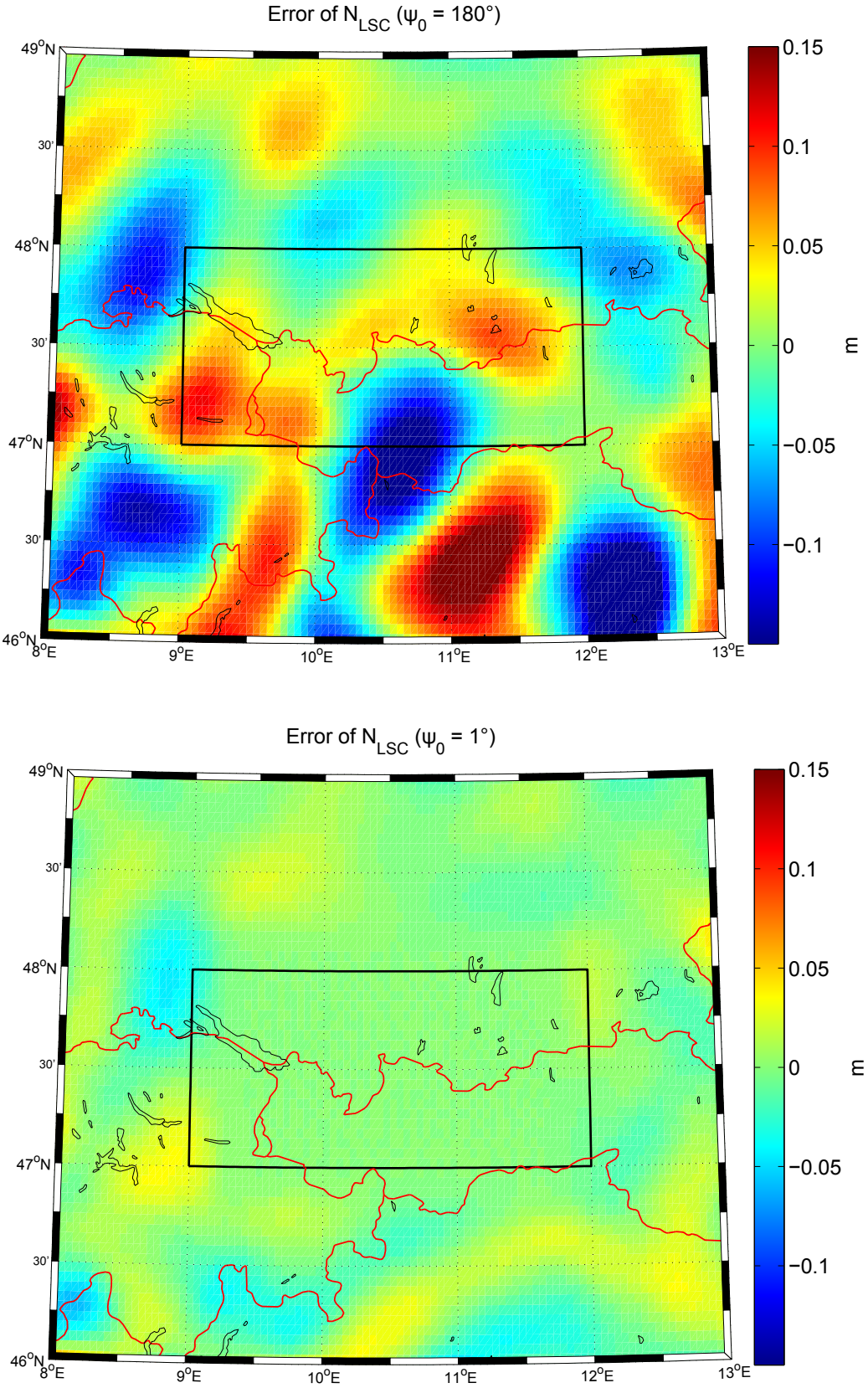

Figure 5.23: Alpine region, *data area. Comparison of unmodified and modified covariance function with Stokes integration in a spherical cap with radius* $\psi_0 = 1°$. *Note that* $N_{LSC}(\psi_0 = 180°)$ *corresponds to the original unmodified covariance function.*

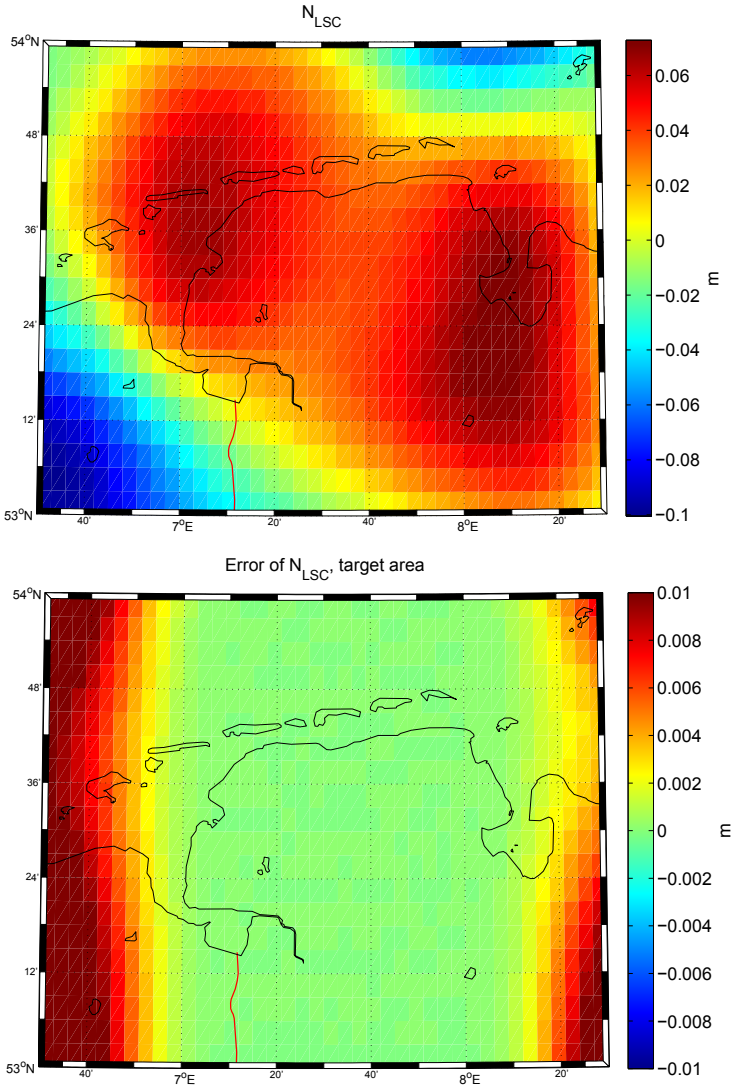

Figure 5.24: Least-squares collocation. *Results from geoid computation and comparison, East Frisia.*

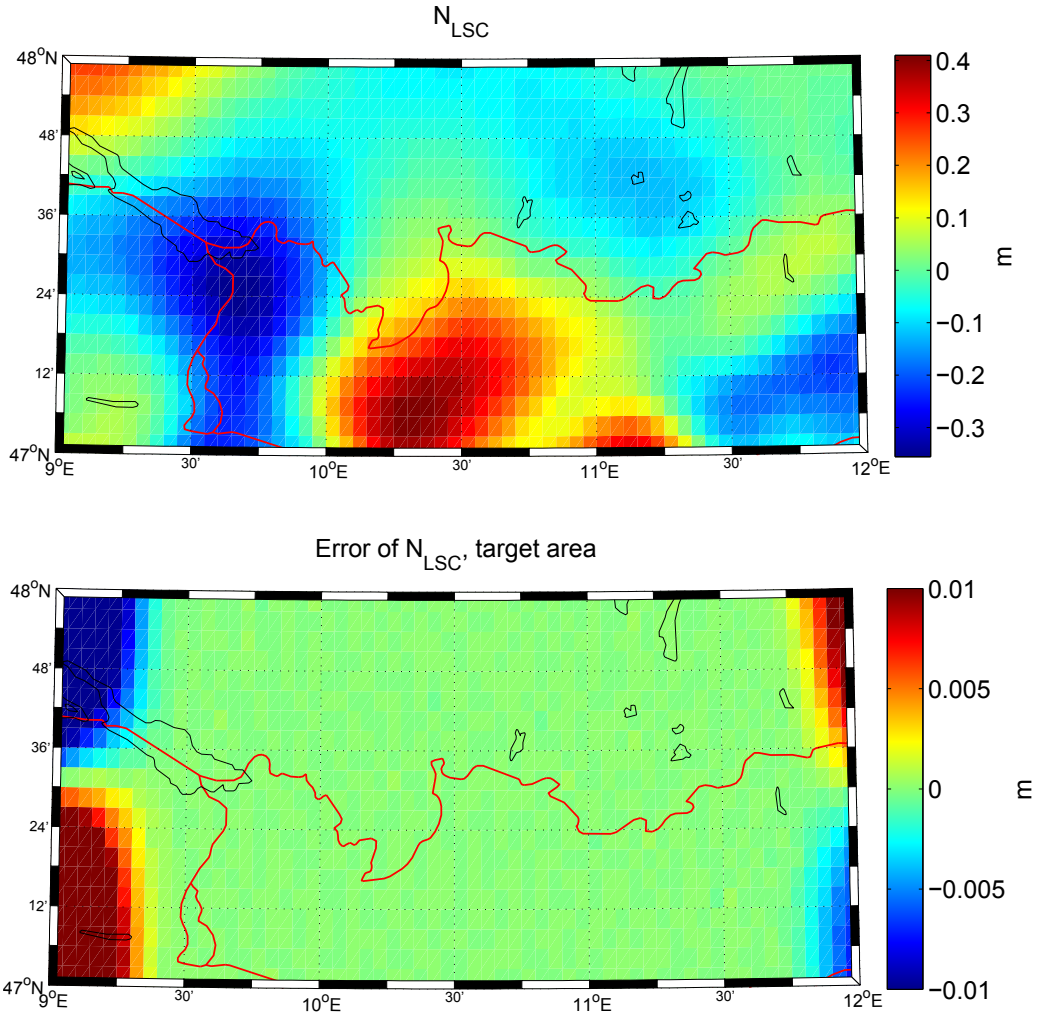

Figure 5.25: Least-squares collocation. *Results from geoid computation and comparison, Alpine region.*

5.4.4 Agreement of RBFs with spherical harmonic synthesis

Global comparison of spherical harmonics with Shannon RBF

In section 3.5.3 it was concluded that, in the global case using *Shannon* RBF, gravity field modeling by spherical harmonic synthesis and radial basis functions are equivalent methods. This was also verified in practical computations. The global geoid was computed on a 5° latitude-longitude grid with spherical harmonic synthesis up to degree 31 ($\sim$ 650 km resolution). Further, the spherical harmonic coefficients were transferred to RBF coefficients according to equation (3.125), and the global geoid synthesized with Shannon low-pass RBFs (equation (3.78) with $N = 31$), placed on the same latitude-longitude grid. Results are summarized in table 5.11 and Figure 5.26.

	max.	min.	mean	RMS
N_{SHS}	79.002	-102.978	-0.849	28.767
N_{RBF}	79.002	-102.978	-0.849	28.767
$N_{\mathrm{RBF}} - N_{\mathrm{SHS}}$	$1.14 \cdot 10^{-12}$	$-1.75 \cdot 10^{-12}$	$-2.49 \cdot 10^{-14}$	$2.56 \cdot 10^{-13}$

Table 5.11: Global comparison. *Statistics of the geoid computed by spherical harmonic synthesis, radial basis functions and the difference between the geoids (*m*).*

5.5 Global RBF analysis and synthesis with Shannon RBF, applying regularization

As opposed to the above case, where the spherical harmonic coefficients were directly transferred to RBF coefficients and subsequently synthesized with Shannon RBF, a synthetic disturbing potential field is both analyzed and synthesized with Shannon low-pass RBF ($N = 31$) in this case, cf. section 3.4.3 on RBF analysis by least-squares adjustment.

Case 1a assumes both observations and RBFs on the Earth surface (i.e., $h = 0$), with noise-free observations. The observations are given on a 5° latitude-longitude grid, and the RBFs are placed on a Reuter grid with $\gamma = N_{\max} = 31$. In case 1b, both observations and RBFs are placed on the Reuter grid.

Cases 2a-b exhibit similar properties as cases 1a-b, only the observation accuracy is set to $\sigma = 10\,\mathrm{m^2 s^{-2}}$.

Cases 3a-b explore the effect of upward continuation, where the observations are given at typical satellite altitude $h = 300$ km, and otherwise similar to cases 2a-b.

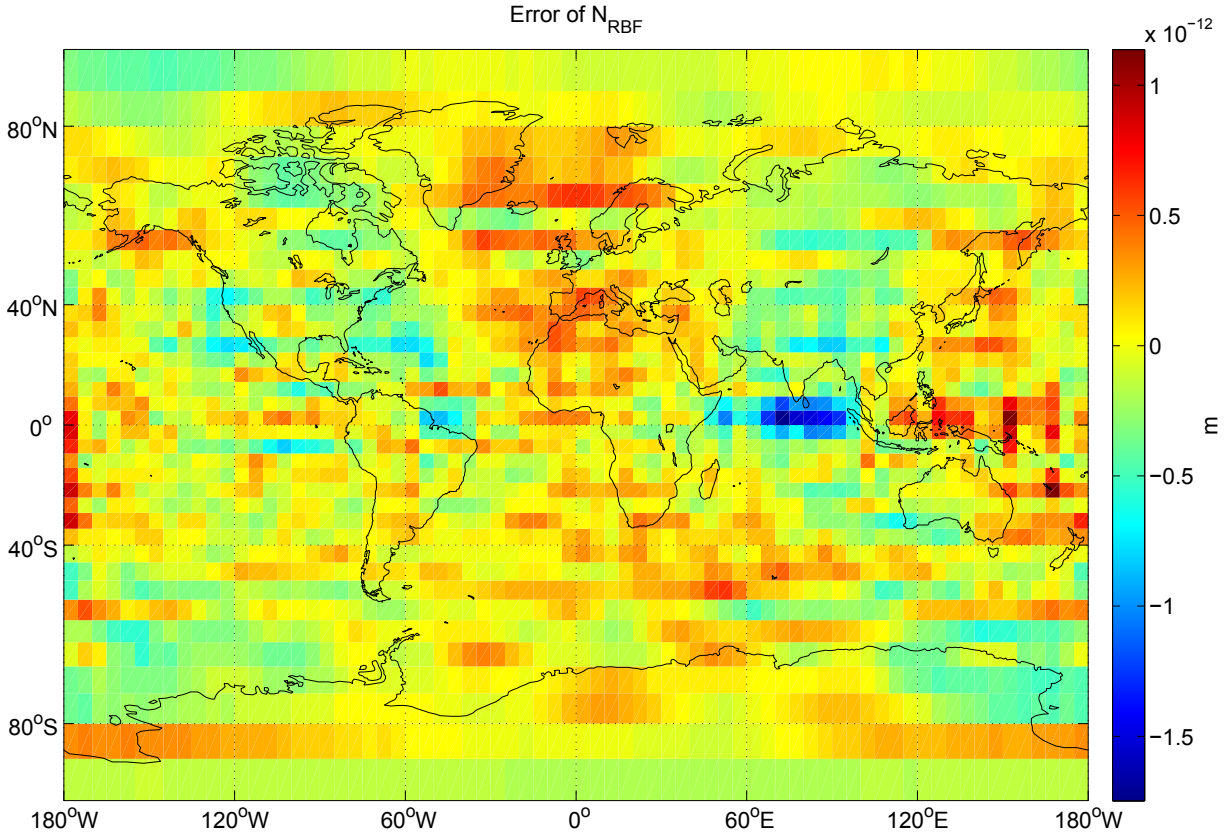

Figure 5.26: *Error of geoid computed from radial basis functions.*

Again, the observations are created with spherical harmonic synthesis up to degree 31, using *EGM2008*. The observed potential at altitude $h = 0$, on a 5° latitude-longitude grid and Reuter grid with $\gamma = N_{\max} = 31$, is shown in Figure 5.27.

In every case, the spherical harmonic coefficients of *EGM2008* were transferred to RBF d-coefficients, for comparison with the estimated d-coefficients obtained in the RBF analysis. Further, the difference between the "true" potential field and the synthesized potential field with the estimated coefficient as argument was computed, i.e., $T_{\text{est.}} - T_{\text{EGM08}}$. Relative contributions from prior information and the observed data were computed. Also, the estimated d-coefficients were transferred to spherical harmonic coefficients, and the signal degree variances of the "true" *EGM2008* disturbing potential coefficients plotted together with the estimated disturbing potential coefficients, as well as the error degree variances computed from their difference. As prior information for the d-coefficients, zero-observations were added, with accuracies corresponding to a RMS-value computed from the "true" d-coefficients.

The condition numbers of the different cases are found in Table 5.12.

	$\text{cond}(\mathbf{N} + \alpha\mathbf{K}^{-1})$					
	Case 1a	Case 1b	Case 2a	Case 2b	Case 3a	Case 3b
$\alpha = 0$	$6 \cdot 10^{18}$	–	–	–	–	–
α_{small}	$7 \cdot 10^{14}$	$4 \cdot 10^{13}$	$6 \cdot 10^{12}$	$4 \cdot 10^{11}$	$2 \cdot 10^{12}$	$9 \cdot 10^{10}$
α_{large}	1.1	–	–	36	–	10
$\alpha_{\text{L-curve}}$	–	–	–	$1 \cdot 10^{3}$	–	599

Table 5.12: RBF analysis. *Condition numbers for the regularized normal matrix* $\mathbf{N} + \alpha\mathbf{K}^{-1}$ *for cases 1a-b, 2a-b and 3a-b.*

Figure 5.28 summarizes the results of case 1a *without* regularization. From Table 5.12 is clear, that the problem is ill-conditioned. The contribution matrices $\mathbf{R}_x$ and $\mathbf{R}_y$ are both computed by the regularized normal matrix inverse, and are therefore not meaningful here. The reason why the relative contribution from prior information, $\mathbf{R}_x$, still seems sensible (no regularization should indeed give no contribution from prior information), this is only since the ill-conditioned regularized normal matrix is multiplied with the regularization parameter (which is zero). The sum of the contribution matrices does not yield the identity matrix, which they should.

The true RBF d-coefficients, transferred from *EGM2008* spherical harmonic coefficients, do resemble the observed signal, which is expected. They can be said to be *physically meaningful*. The estimated d-coefficients, however, do not present these traits. Therefore, it is quite interesting to see the residuals, where the discrepancy between the true disturbing potential and the disturbing potential computed with the seemingly wrong d-coefficients is relatively small. Supporting the residuals is the degree variances plot, revealing small differences between the true signal degree variances and the estimated signal degree variances. It seems that a mathematically correct solution is computable, although the d-coefficients are not physically meaningful. Therefore, it does not seem like

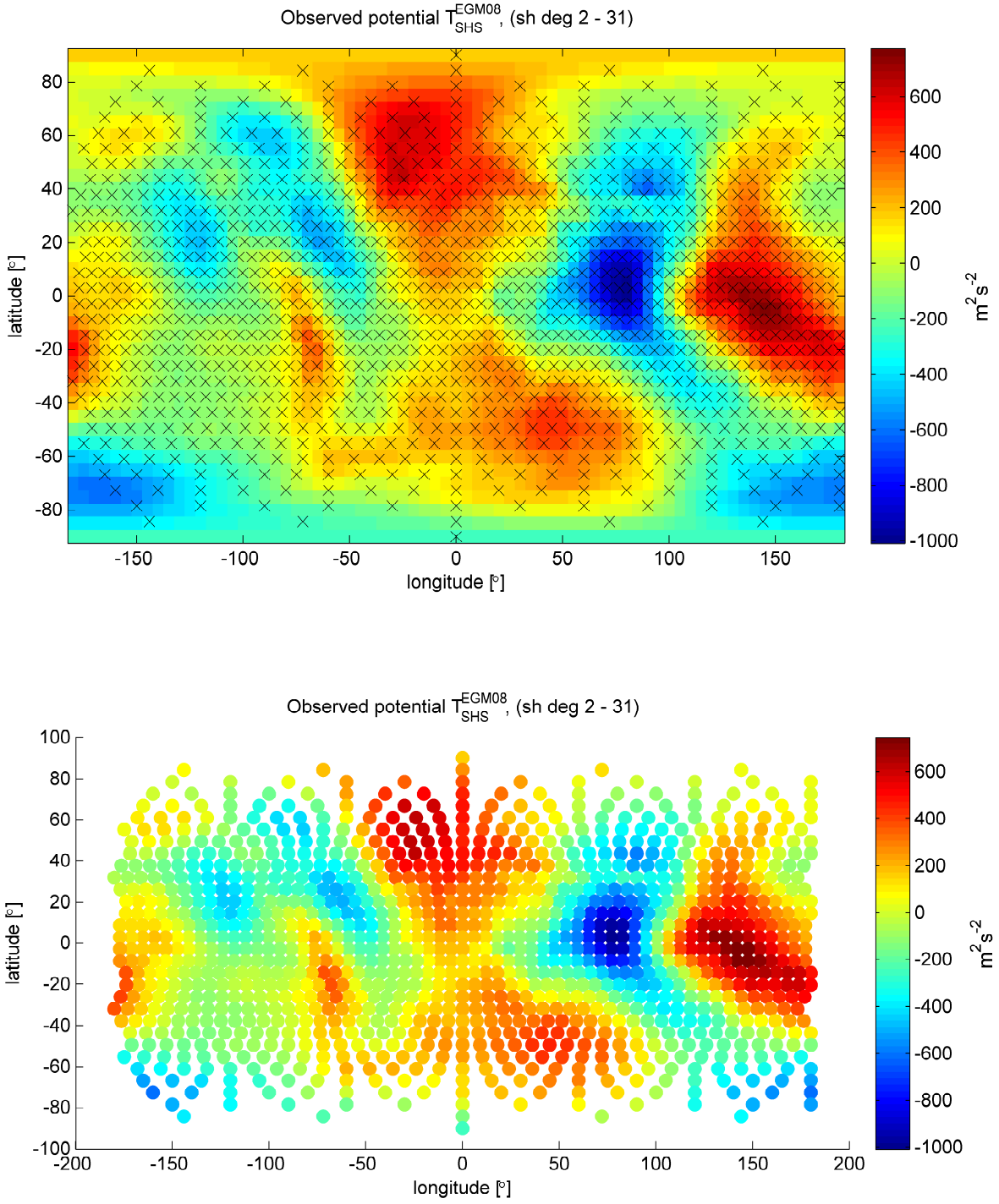

Figure 5.27: *Synthetic observation data, disturbing potential T_{SHS} at height $h = 0$, on a 5° latitude-longitude grid (top) and Reuter grid with $\gamma = N_{max} = 31$ (bottom). The black crosses mark the Reuter grid points.*

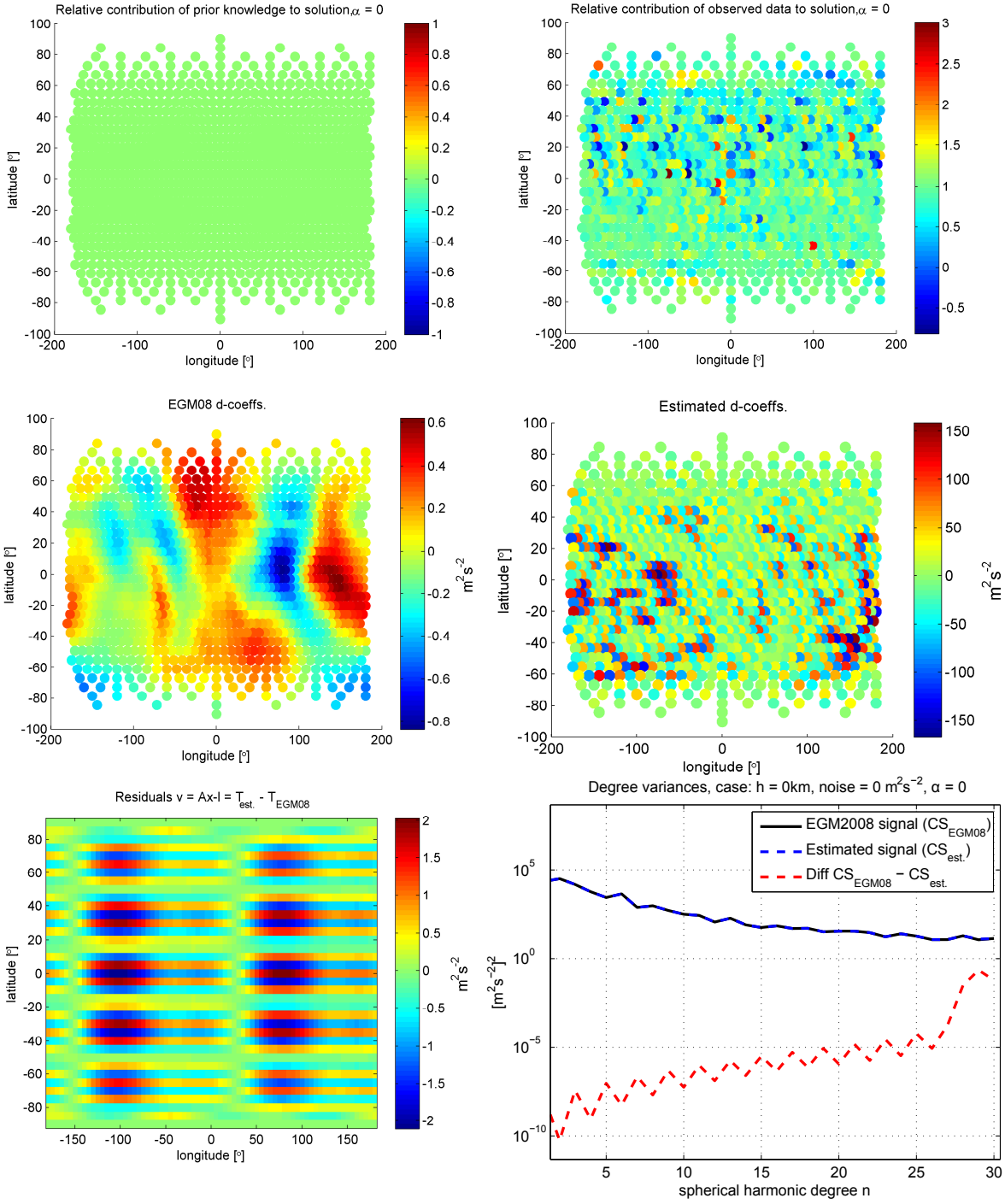

Figure 5.28: RBF analysis. *Results from case 1a,* $\sigma = 0$ *and at* $h = 0$. *No regularization.*

the residuals and the degree variances plot alone are reasonable criteria for estimation quality description.

With a small degree of regularization, the d-coefficients change drastically, cf. Figure 5.29. Still, the problem is ill-conditioned, cf. Table 5.12, but stabilized enough to give a physically meaningful estimate of the d-coefficients. Therefore, this small regularization parameter was used also in the other cases.

In case 1b, summarized in Figure 5.30, the observations are also given on the Reuter grid, where the only notable difference from case 1a is that the residuals are considerably smaller, due to the fact that both noise-free observations and RBFs are placed on the exact same positions.

Generally, similar traits as in the spherical harmonic analysis cases, cf. section 5.3, were observed also in the radial basis function analysis cases. Results of an arbitrarily chosen high degree of regularization in case 1a are summarized in 5.31. A dramatical stabilization of the system is observed in the condition number, cf. Table 5.12. As in the spherical harmonic analysis case with a large degree of regularization, the solution tends towards zero, and the observation error is taking the form of the signal.

Case 2a is summarized in Figure 5.32. Noise has been added to the observations, and consequently, the errors are larger. The estimated signal degree variances plot diverges from the true signal degree variances plot above roughly degree 15. The results from case 2b with same regularization parameter as case 2a are summarized in Figure 5.33. These cases are relatively similar, although the errors seem to be somewhat larger and the solution pattern somewhat more wrong in case 2b.

By trial-and-erroring, an upper limit for the regularization parameter was found in case 2b, for the subsequent L-curve computation. Case 2b results with this large regularization parameter are found in Figure 5.34. The tendency of the solution towards zero can be seen in the estimated d-coefficients plot. Prior information contributes to the solution by roughly 20 %. The residuals begin to take the shape of the signal, and the error degree variances curve moves upwards. The L-curve for case 2b is plotted in Figure 5.35, and the results from case 2b with $\alpha_{\text{L-curve}}$ are found in Figure 5.36. The L-curve solution seems good, with the estimated d-coefficients resembling the signal, the residuals seemingly random and the estimated signal degree variances curve fits better to the true signal degree variances curve.

Results from case 3a, with α_{small} as in the other cases, are found in Figure 5.37. At typical satellite altitude, the error degree variances curve supersedes the true signal degree variances curve at roughly degree 24. Results from case 3b, with α_{small} and both observations and RBFs on the Reuter grid, are found in Figure 5.38. The overall error is larger in case 3b, which can be seen in the residuals plot. Crossing of the error degree variances curve with the true signal degree variances happens at roughly degree 22. The estimated d-coefficients in case 3b present a pattern that does not resemble the true d-coefficient pattern as well as in case 3a.

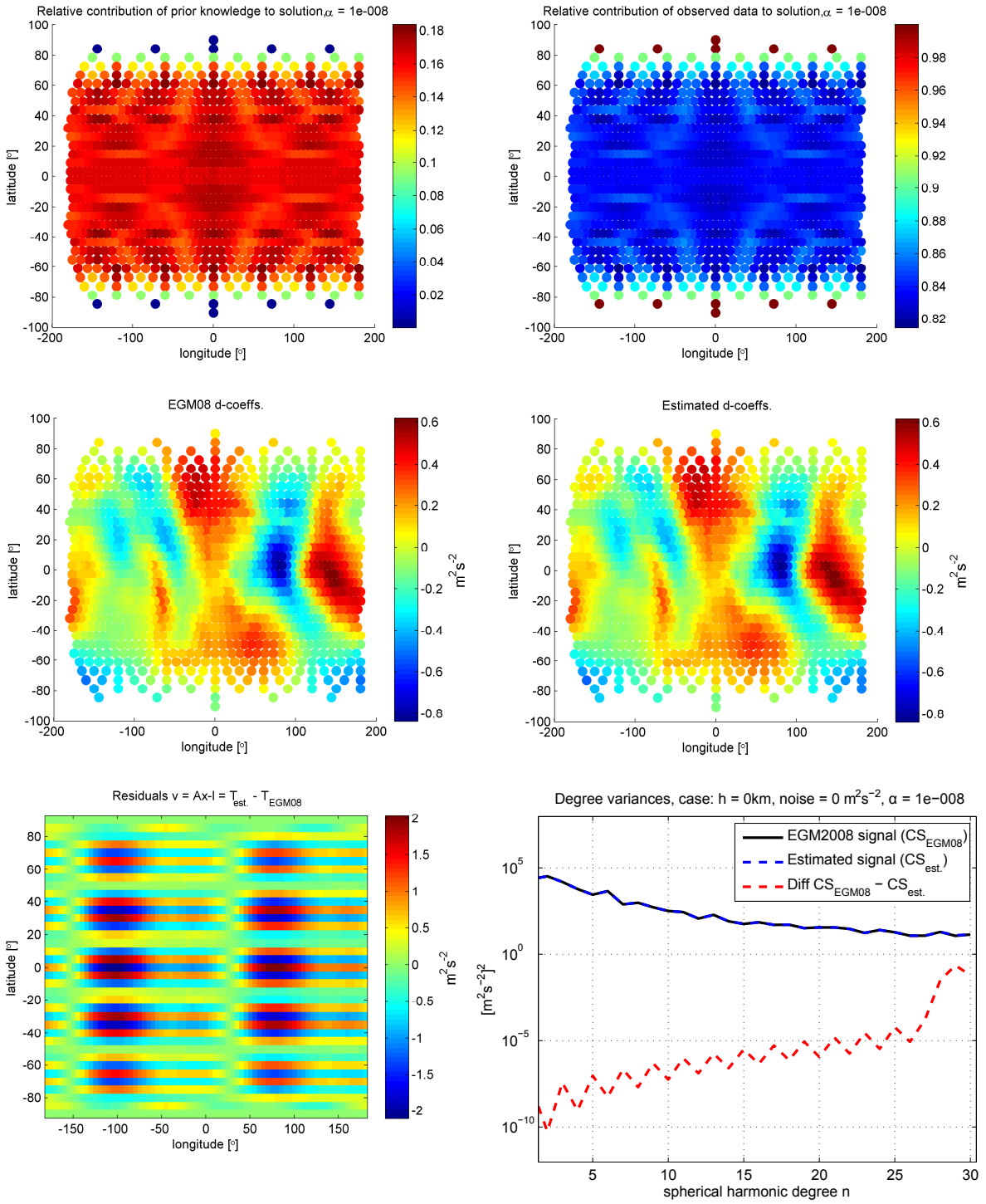

Figure 5.29: RBF analysis. *Results from case 1a,* $\sigma = 0$ *and at* $h = 0$. *Regularization with* $\alpha_{small} = 1 \cdot 10^{-8}$.

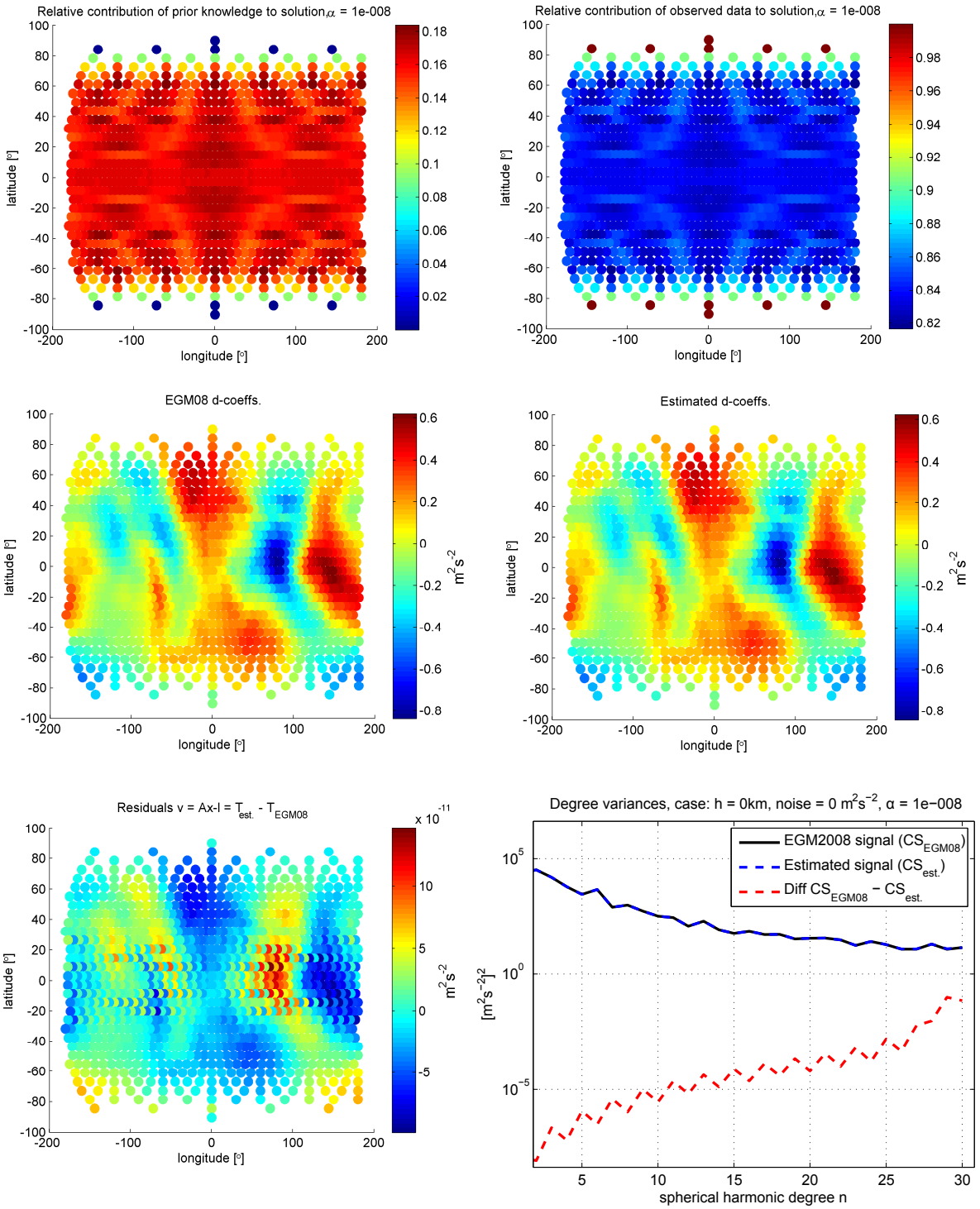

Figure 5.30: RBF analysis. *Results from case 1b, $\sigma = 0$ and at $h = 0$. Regularization with* $\alpha_{small} = 1 \cdot 10^{-8}$.

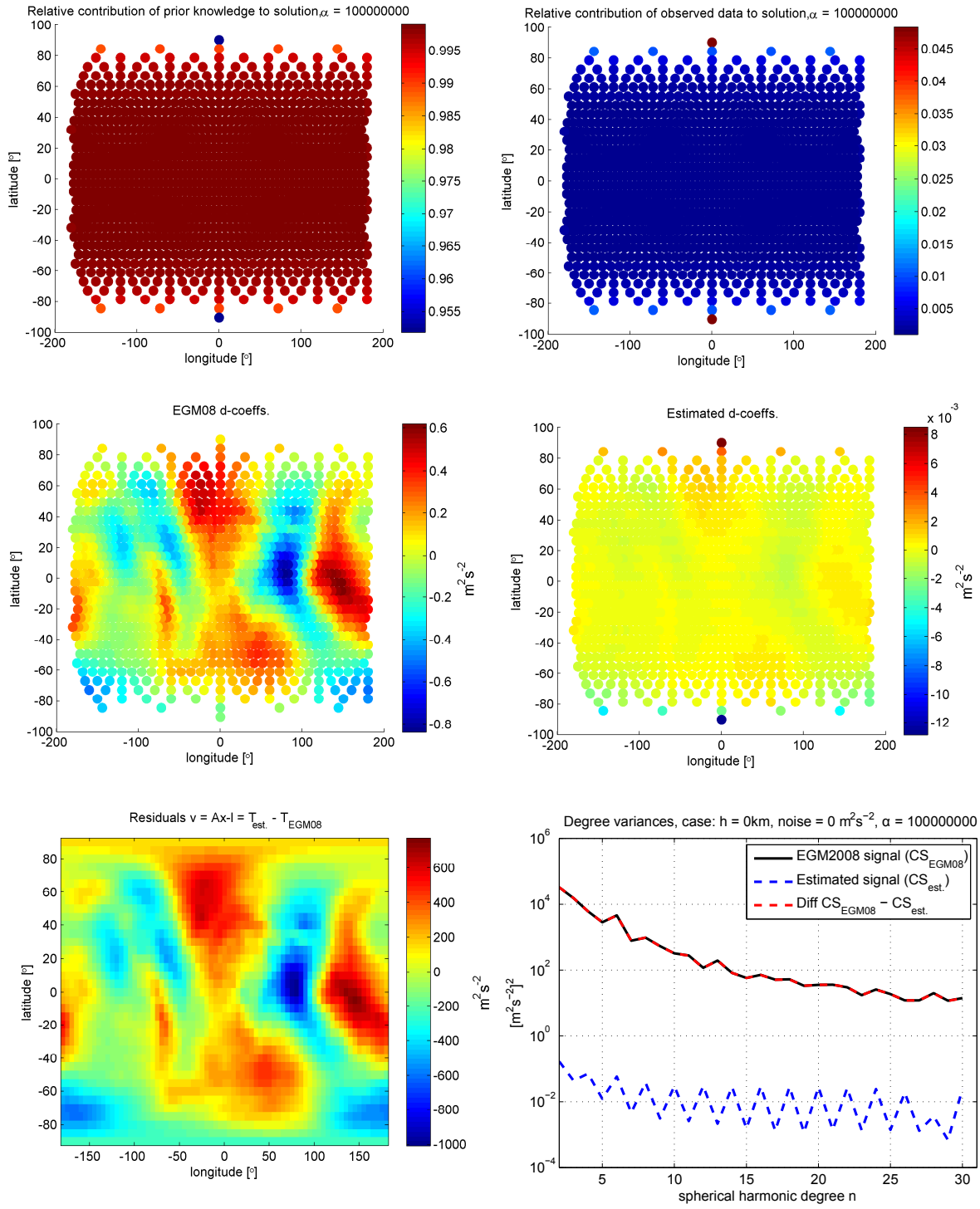

Figure 5.31: RBF analysis. *Results from case 1a,* $\sigma = 0$ *and at* $h = 0$. *Regularization with* $\alpha_{large} = 1 \cdot 10^8$.

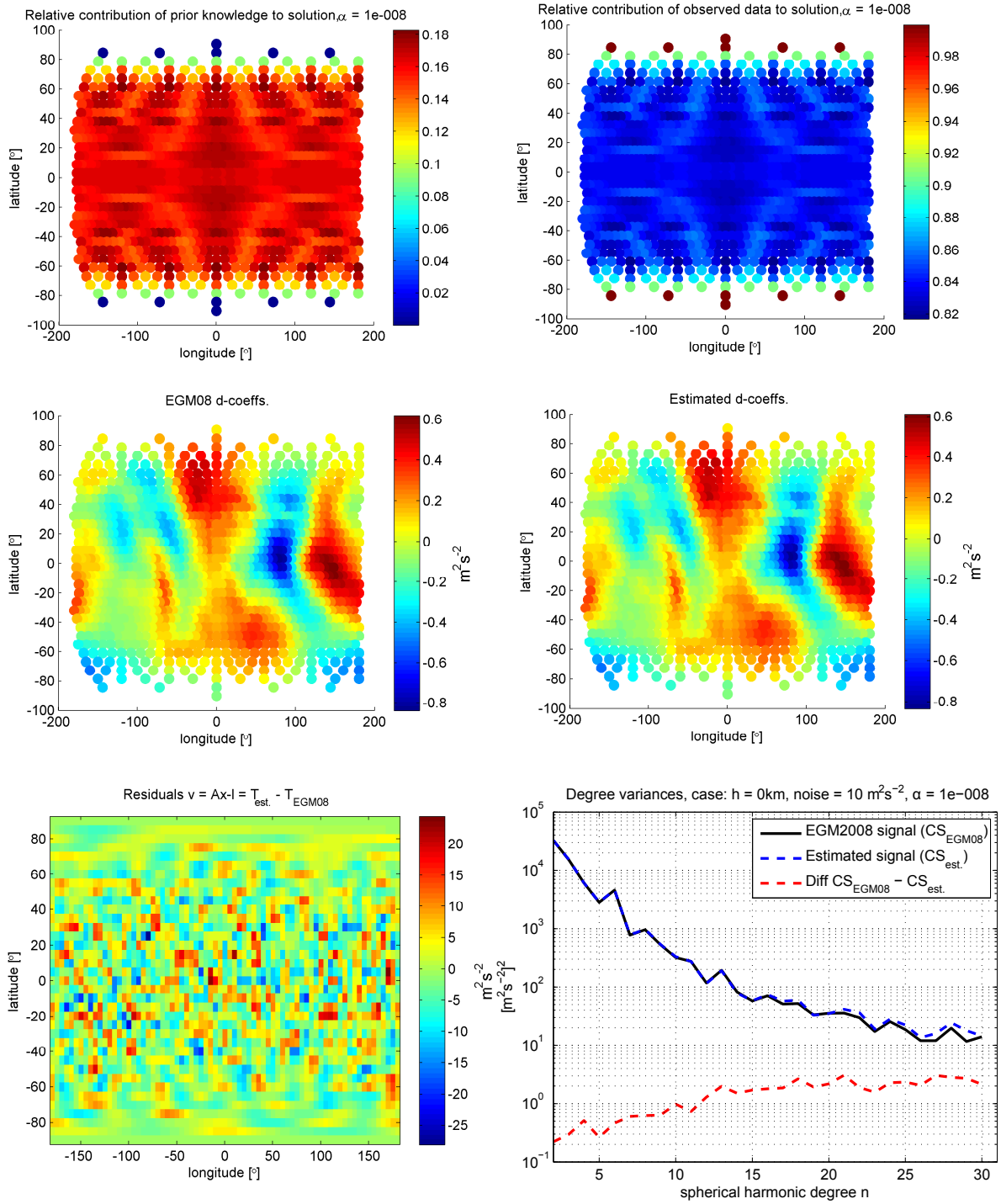

Figure 5.32: RBF analysis. *Results from case 2a,* $\sigma = 10\,\mathrm{m^2 s^{-2}}$ *and at* $h = 0$. *Regularization with* $\alpha_{small} = 1 \cdot 10^{-8}$.

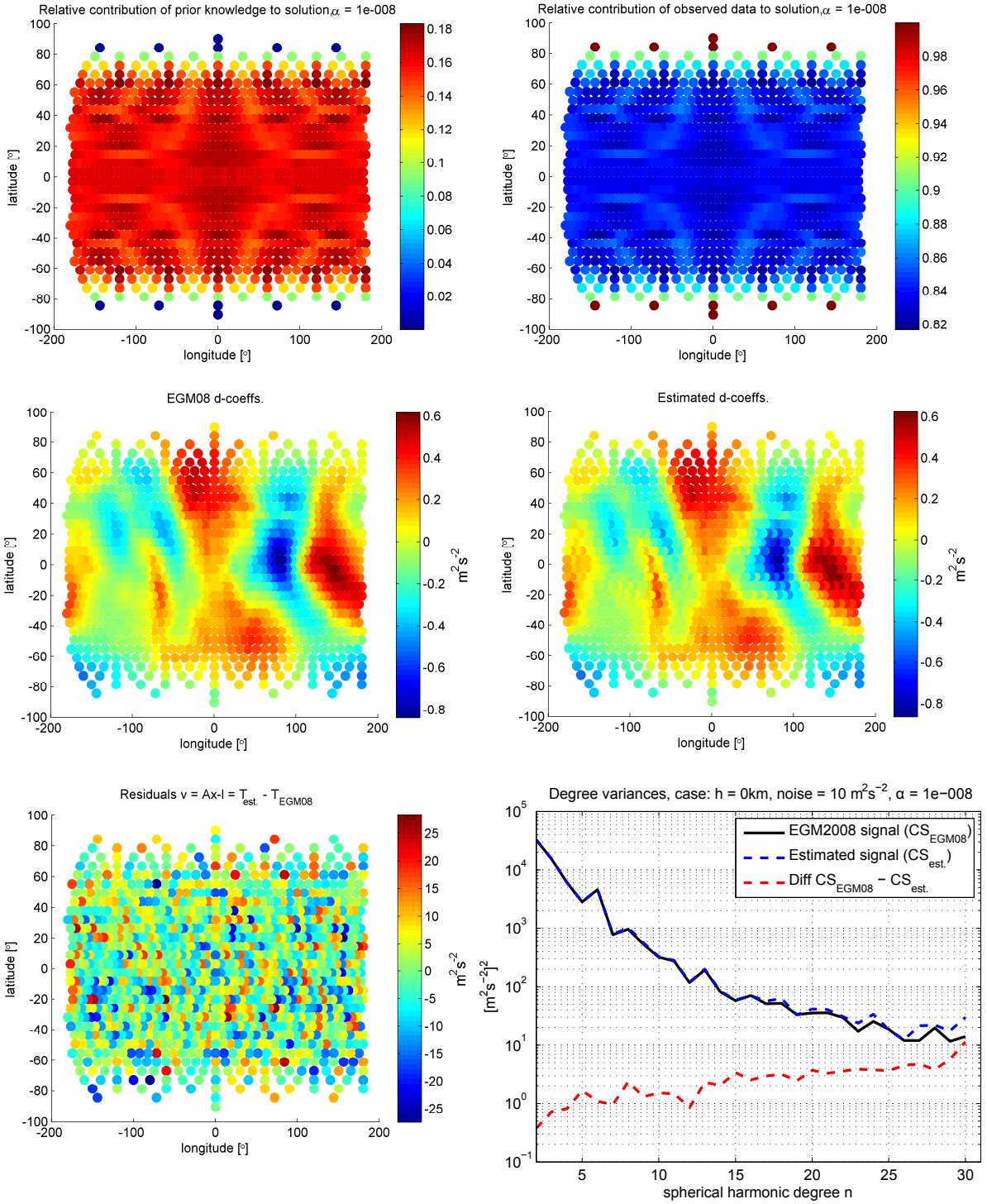

Figure 5.33: RBF analysis. *Results from case 2b,* $\sigma = 10\,\mathrm{m^2 s^{-2}}$ *and at* $h = 0$. *Regularization with* $\alpha_{small} = 1 \cdot 10^{-8}$.

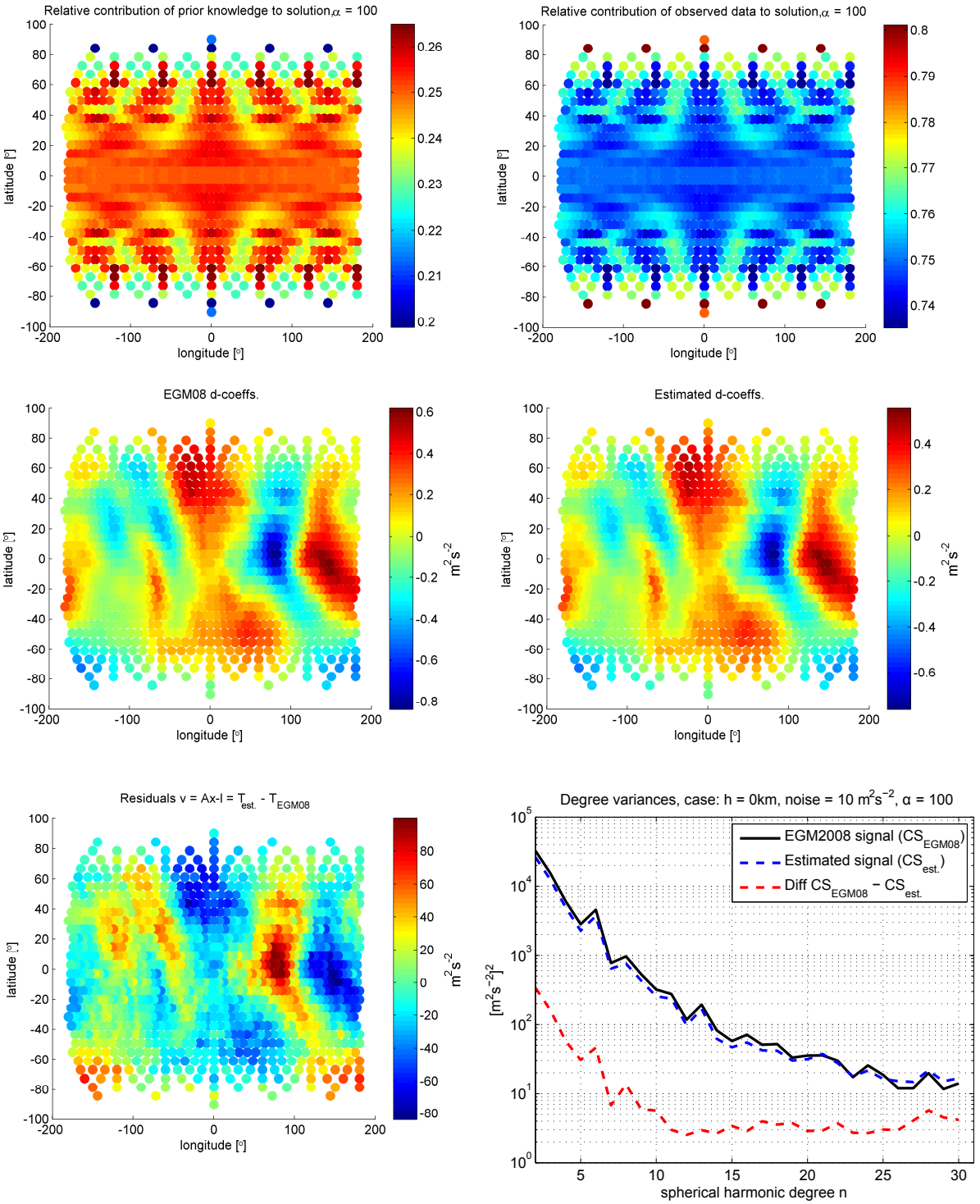

Figure 5.34: RBF analysis. *Results from case 2b,* $\sigma = 10\,\mathrm{m^2 s^{-2}}$ *and at* $h = 0$. *Regularization with* $\alpha_{large} = 100$.

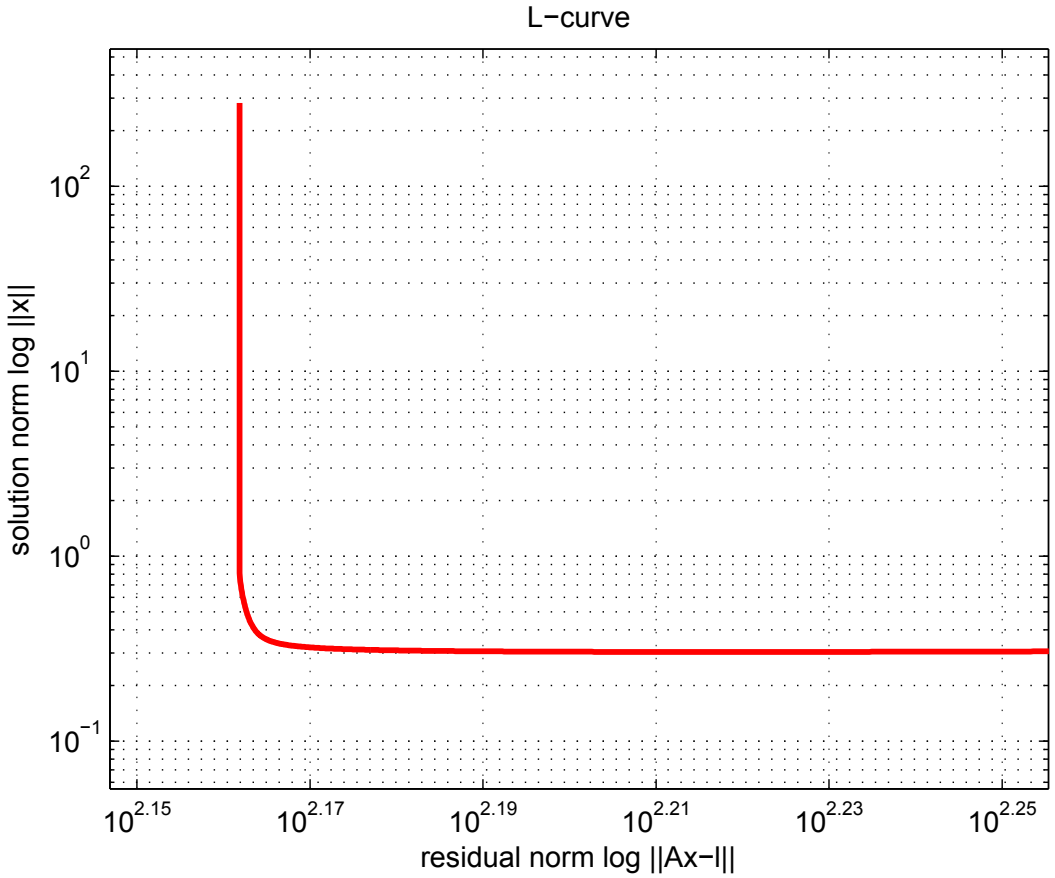

Figure 5.35: RBF analysis. *L-curve for case 2b, corner found by the curvature approach.*

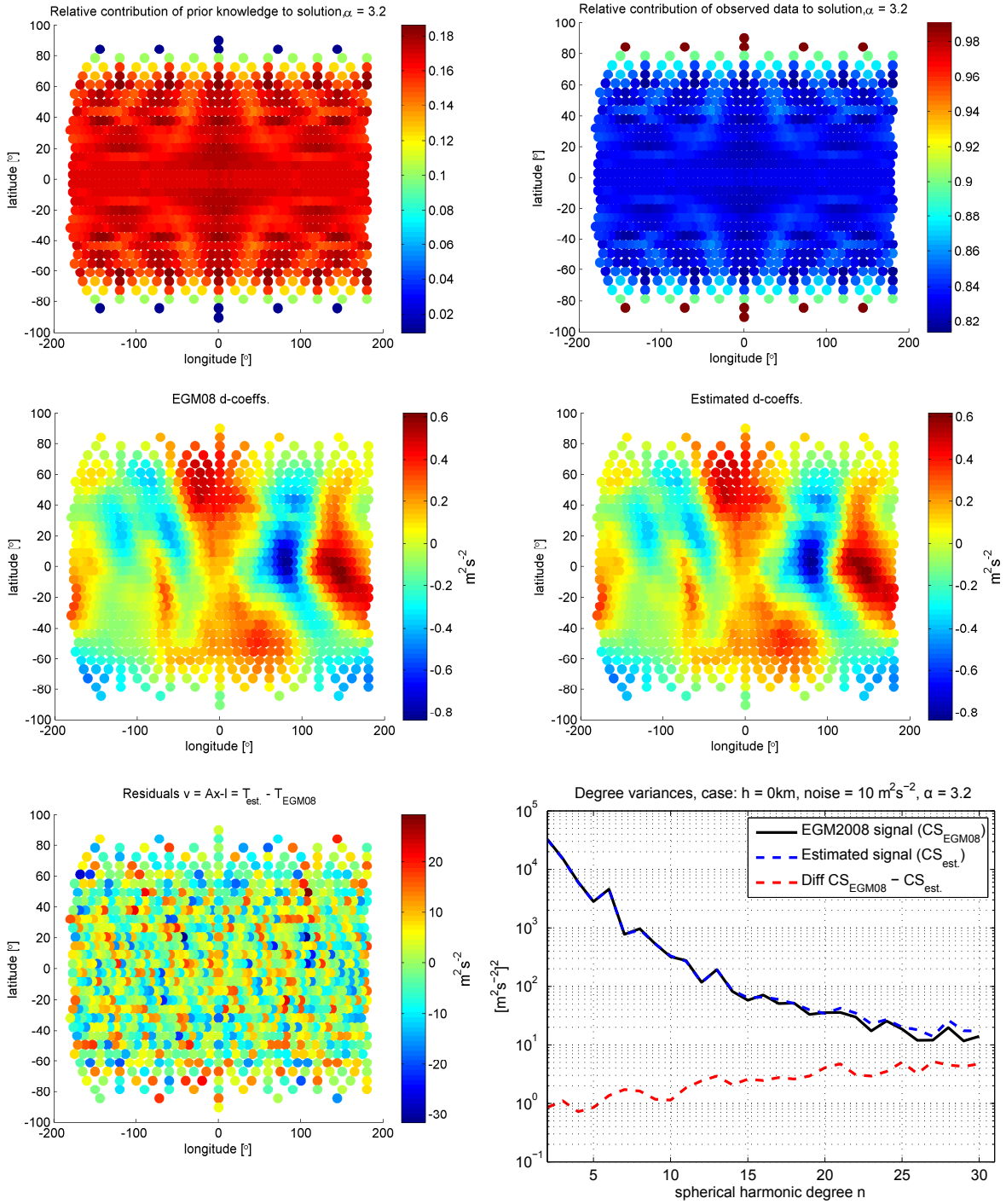

Figure 5.36: RBF analysis. *Results from case 2b,* $\sigma = 10\,\mathrm{m^2 s^{-2}}$ *and at* $h = 0$. *Regularization with* $\alpha_{L\text{-}curve} = 3.2$.

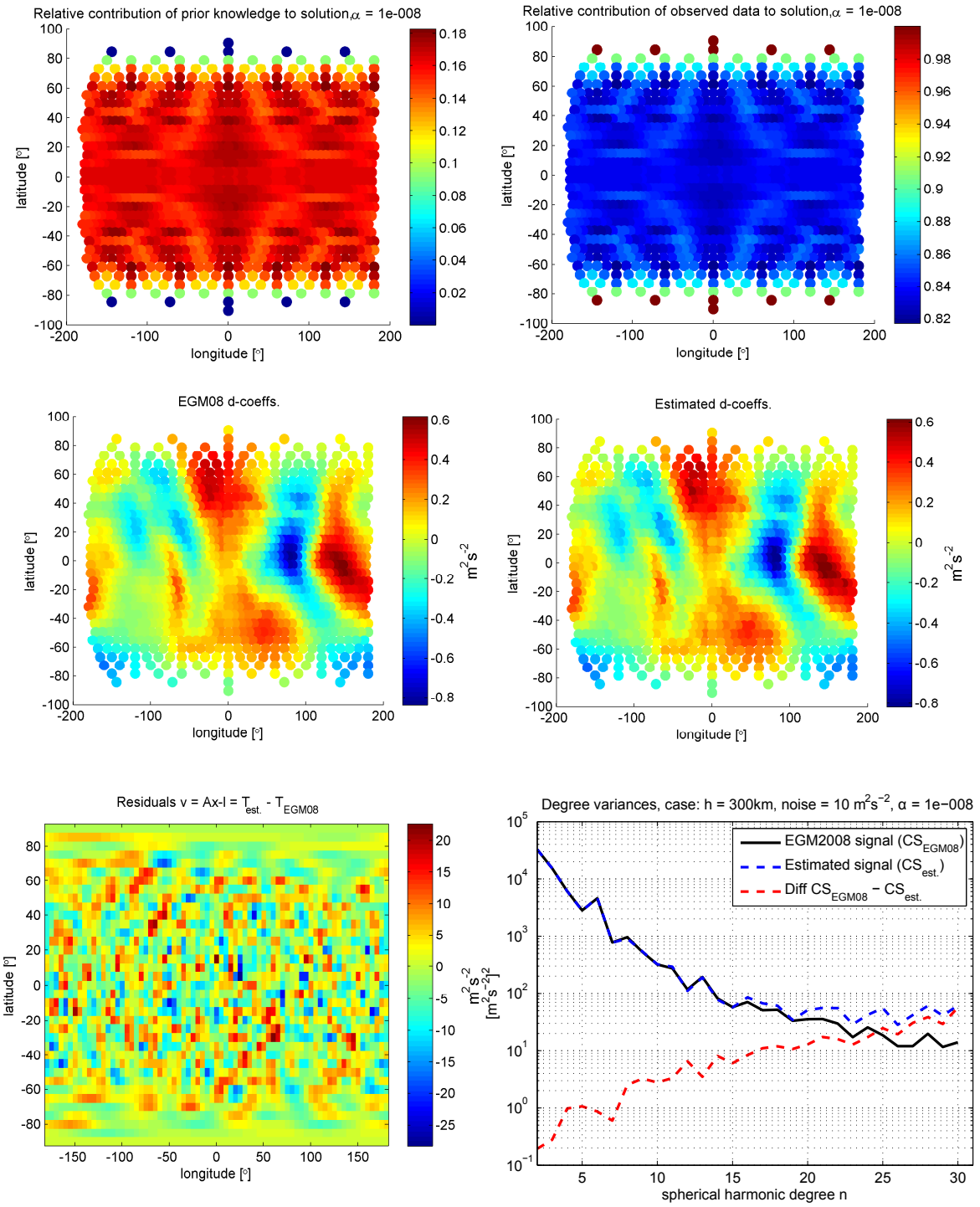

Figure 5.37: RBF analysis. *Results from case 3a,* $\sigma = 10\,\mathrm{m^2 s^{-2}}$ *and at* $h = 300\,\mathrm{km}$. *Regularization with* $\alpha_{small} = 1 \cdot 10^{-8}$.

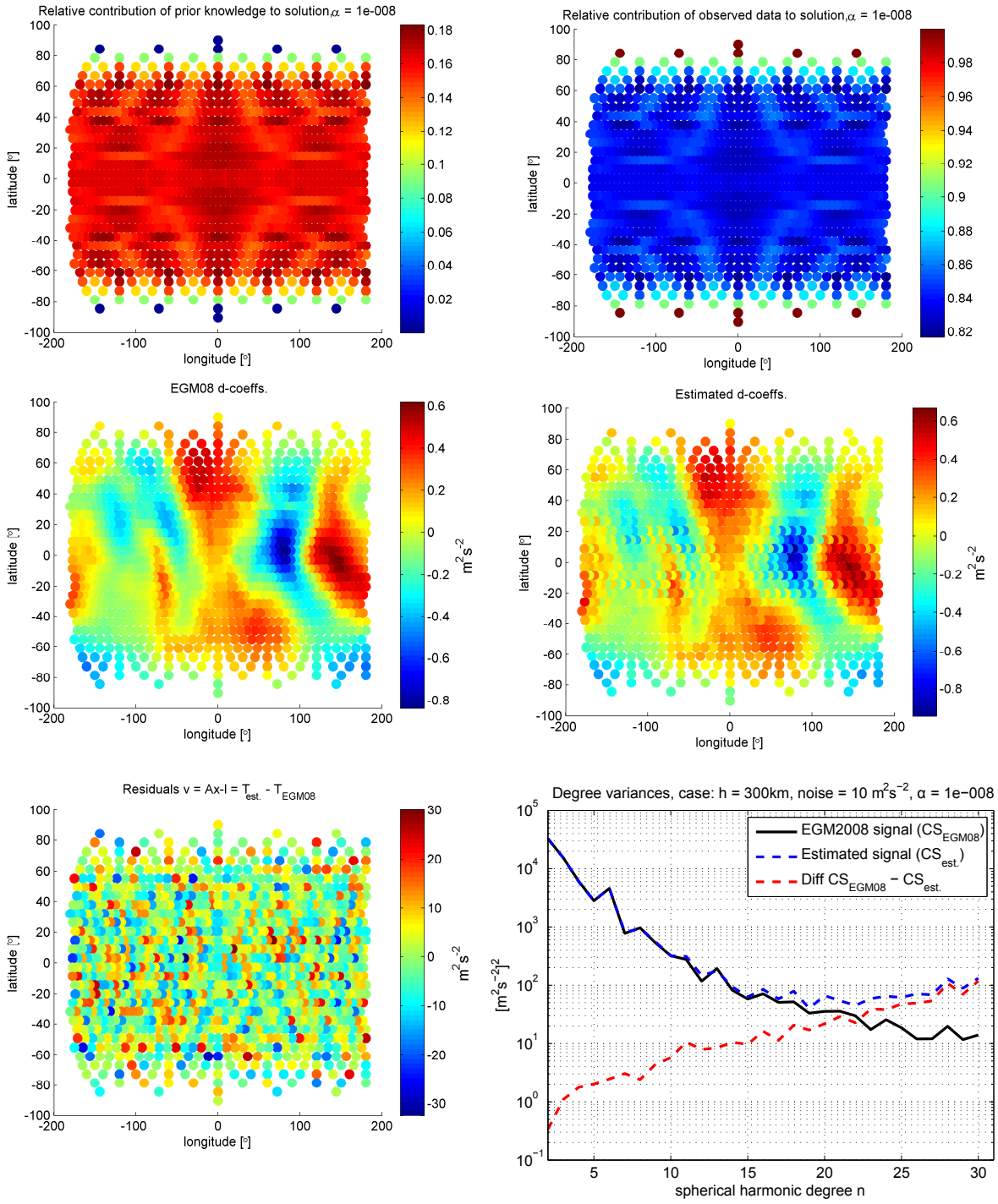

Figure 5.38: RBF analysis. *Results from case 3b,* $\sigma = 10\,\mathrm{m^2 s^{-2}}$ *and at* $h = 300\,\mathrm{km}$. *Regularization with* $\alpha_{small} = 1 \cdot 10^{-8}$.

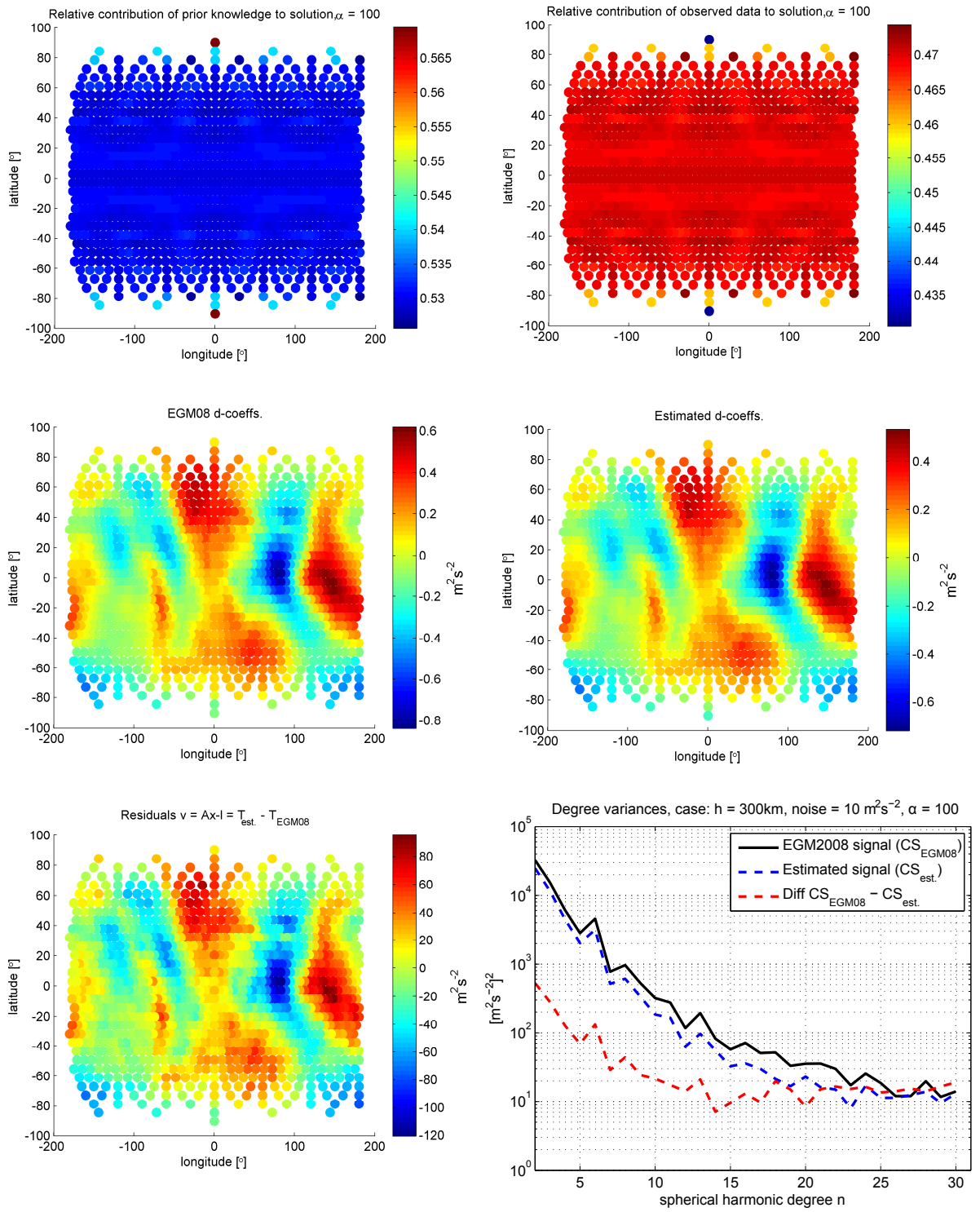

Figure 5.39: RBF analysis. *Results from case 3b,* $\sigma = 10\,\mathrm{m^2 s^{-2}}$ *and at* $h = 300\,\mathrm{km}$. *Regularization with* $\alpha_{large} = 100$.

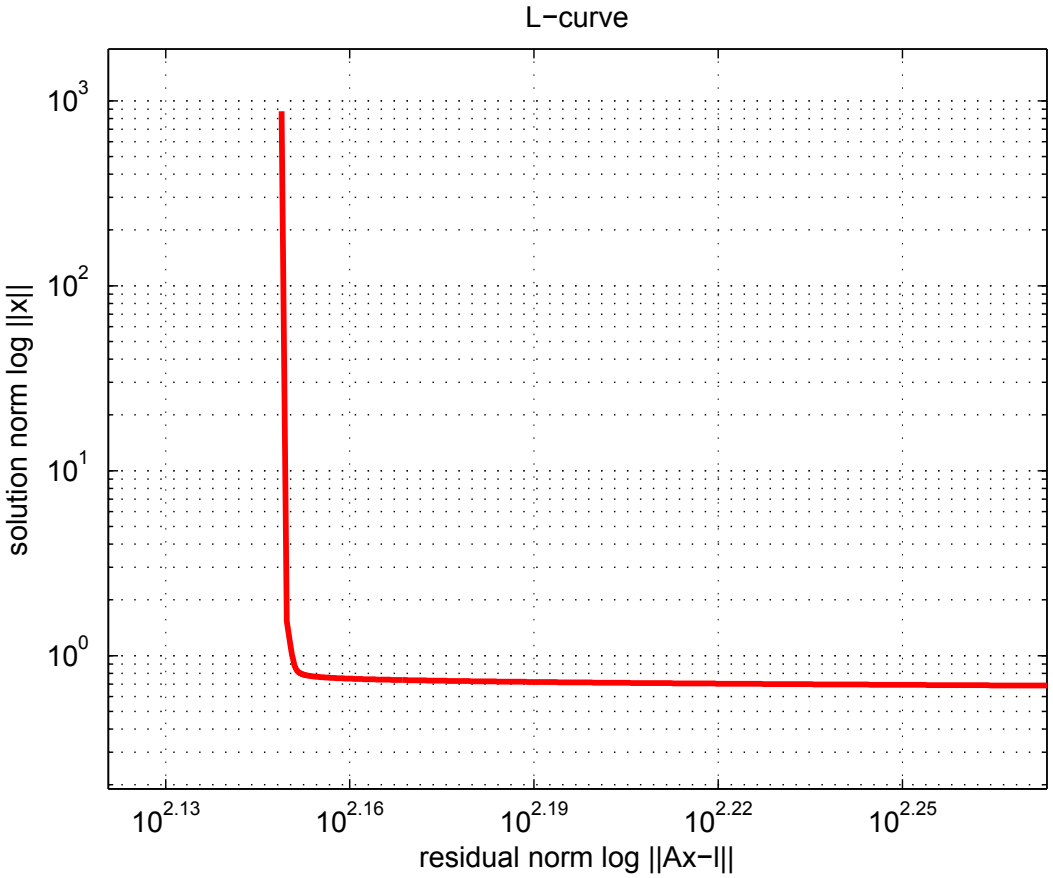

Figure 5.40: RBF analysis. *L-curve for case 3b, corner found by the curvature approach.*

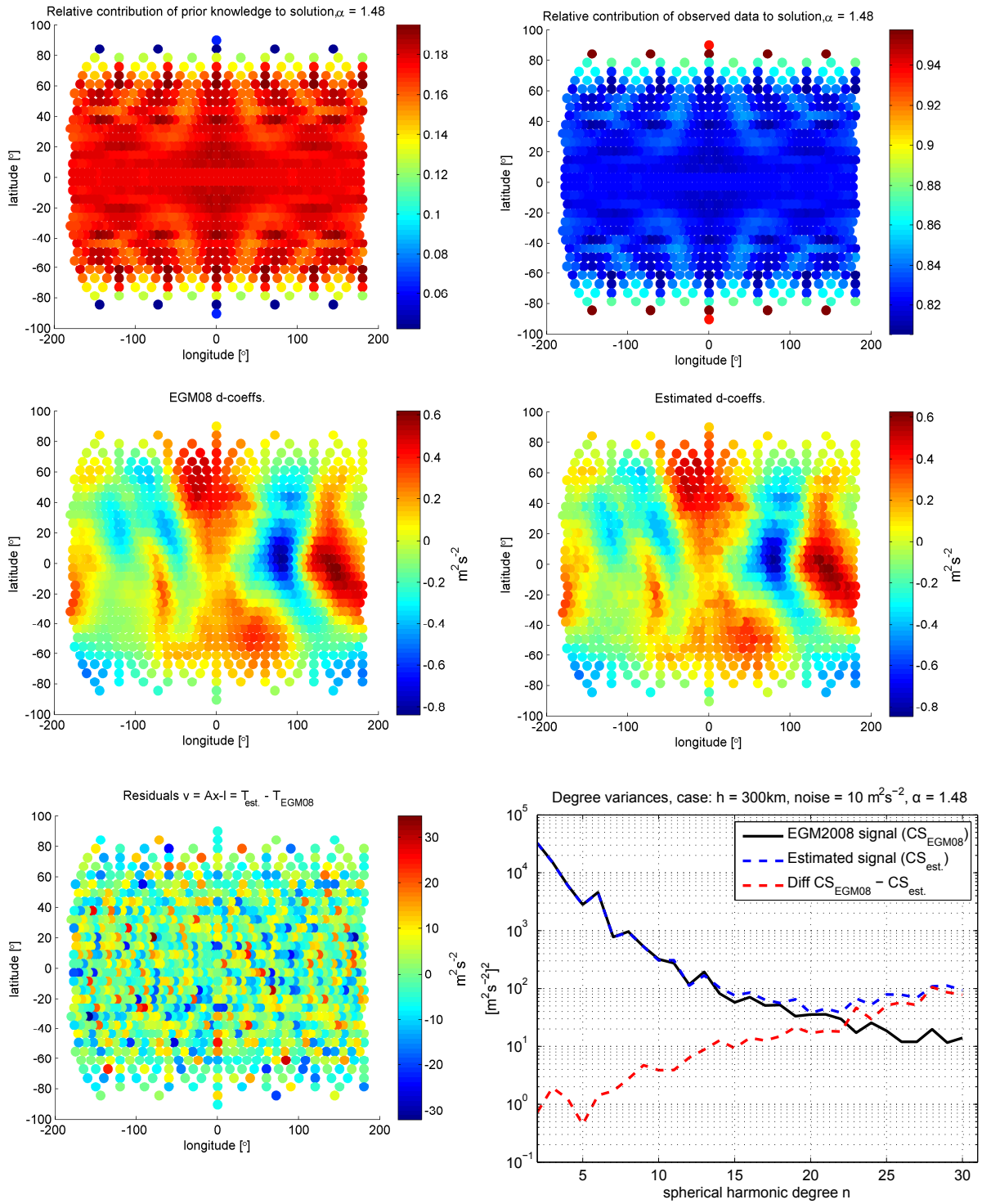

Figure 5.41: RBF analysis. *Results from case 3b,* $\sigma = 10\,\mathrm{m^2 s^{-2}}$ *and at* $h = 300\,\mathrm{km}$. *Regularization with* $\alpha_{L\text{-}curve} = 1.48$.

As in case 2b, case 3b was computed with an L-curve upper limit regularization parameter α_{large}, summarized in Figure 5.39, displaying recognizable traits of over-regularization. The L-curve in case 3b is plotted in Figure 5.40, and the results of case 3b with $\alpha_{\text{L-curve}}$ are found in Figure 5.41. The L-curve solution is certainly an improvement to the solution with α_{small}. Mainly, the estimated d-coefficients now resemble the true signal and the residuals still seem random. The normal matrix condition number suggests that the solution is stable, but above roughly degree 17, the solution is increasingly incorrect, and the errors large.

Solutions for all regularization parameters were not included in this thesis in all cases, mainly for practical space-saving reasons, since the different cases show similar changes from one regularization parameter to the next. The included cases show important differences, one being the difference between mathematically correct and physically meaningful d-coefficients in case 1a. The small regularization parameter yielding physically meaningful d-coefficients was kept for the computation of the other cases.

A rather arbitrarily chosen large regularization parameter was used in case 1a, to illustrate the effect of over-regularization. Cases 2b and 3b, noisy observations on Earth's surface and at satellite altitude, respectively, were chosen as L-curve candidates. The L-curve was not computed for all cases since its plotting is computationally intensive. The goal was to show that the L-curve can be applied also in the radial basis function case.

When computing the L-curve, the largest regularization parameter should not be chosen quite as arbitrarily as it was in case 1a. Therefore, for computational reasons, the large regularization parameter is not as large in the L-curve cases 2b and 3b as in case 1a. In cases 2b and 3b, the regularization parameter range was found by trial-and-erroring.

Chapter 6

Final remarks and outlook

The goal of this thesis was to look at different gravity modeling approaches with focus on regional geoid determination, and compare them. Two well-established models, Stokes integration and least-squares collocation, as well as the more recent application of radial basis functions were considered.

Theoretically, it was shown that, in the global case, the considered models are equal. The theoretical comparisons were supported by numerical investigations with synthetic data.

Numerical comparison of Stokes integration and least-squares collocation with spherical harmonic synthesis was done regionally. Two regions were considered here; East Frisia (where there is little variation in the geoid) and the Alpine region (where there is more variation in the geoid).

East Frisia, with little signal variation, gave the lowest RMS values of the difference between geoids computed by Stokes integration and collocation and the "true" geoid, computed with spherical harmonic synthesis. In case of Stokes integration, the RMS of the difference in geoid heights amounted to 7 mm in the target area. In case of collocation, it amounted to 5 mm.

The Alpine region presents a stronger signal. In case of Stokes integration, the RMS of the difference in geoid heights amounted to 1 cm in the target area. In case of collocation, it amounted to 6 mm.

Although it seems that least-squares collocation yields somewhat better results than Stokes integration, both methods are practically equal.

Remaining discrepancies could be owing to the fact that the synthetic observations are point-values, which spherical harmonic synthesis and least-squares collocation expect. However, Stokes integration, while applied point-wise, expects block mean values. Spherical harmonic synthesis uses the exact observation grid values, while least-squares collocation interpolates and does not give exact values (at least not everywhere). Also, to some extent, differences in the numerical model computations could have an effect.

The performance of Stokes integration improved drastically when the area of observations was extended, and edge effects avoided. The RMS of the difference in geoid heights

amounted to 0.5 mm in the East Frisian target area, and 3.1 mm in the Alpine region target area.

The theoretical global comparison of the Shannon radial basis function with spherical harmonics was confirmed numerically, as the differences between a globally determined geoid using spherical harmonic synthesis and Shannon radial basis function synthesis was in the order of 10^{-12}.

Further, the topic of inverse and ill-posed problems was investigated. Numerical investigations included comparison of applied regularization on different cases of spherical harmonic analysis by least-squares adjustment. Also, the L-curve was plotted, and an optimal regularization parameter α sought.

An ideal case, with noise-free observations, allowed an exact reconstruction of the true gravity field. No regularization was needed, as the normal matrix was well-conditioned.

In the case where observations were at typical satellite altitude and noise-ridden, the normal matrix condition number increased, but stayed well below the level of singularity. The higher degrees could not be determined in this case due to the noisy observations. There, the solution was unstable. With a regularization parameter chosen too large, the system was indeed stabilized, but too regularized, and the well-conditioned part of the system corrupted. Thus, correct coefficients were not obtained, but the solution was more stable. The results with the regularization parameter obtained by the L-curve approach were quite good. The well-conditioned part was left practically "untouched", while the ill-conditioned part was stabilized. Other regularization parameters in the vicinity of the L-curve corner also gave good results.

An extreme case, where observations were assumed to be at GPS satellite altitude, gave a very ill-conditioned system. Without regularization, no sensible solution could be computed at all. Introducing regularization in this case, gave a seemingly correct solution for the low degrees, in accordance with the smoothness of the observations at $h = 20000$ km.

The above regularization problem was transferred directly to the case of Shannon radial basis functions, to investigate regularization effects on radial basis function analysis.

Radial basis function analysis by least-squares adjustment proved to be an ill-conditioned problem even in the ideal case, with noise-free observations and on the Earth surface. In particular, it was observed that, if no regularization was applied, the estimated d-coefficients did not resemble the true signal, like the "true" d-coefficients computed from *EGM2008*. However, when these estimated d-coefficients were introduced into the synthesis, the input signal was re-obtained.

With only a small degree of regularization, the estimated d-coefficients resembled the true signal. Therefore, a distinction between a mathematically correct solution with apparently strange d-coefficients, and a correct solution with *physically meaningful* coefficients was made.

When the solution reflects the signal to a great extent, Bentel et al. (2013) consider it meaningful. While a solution can represent the signal in a mathematical sense, well

capable of representing the signal, only physically meaningful coefficients can be used for, e.g., filtering or transformation between different gravity field functionals.

Otherwise, similar traits as in the spherical harmonic analysis cases were observed also in the radial basis function analysis cases. Regularization with prior information is applicable and gives good results also in the radial basis function analysis case. However, when working with synthetic data, it is evident that the different results (condition numbers, relative contributions, "true" and estimated gravity fields and coefficients, residuals and degree variances plots) must be seen collectively to get a cohesive impression of the "state of the system".

Stokes integration and least-squares collocation may be considered as well-known and extensively treated approaches to gravity field modeling.

A recommendation for further investigations within the topic of this thesis would therefore be to look closer at gravity field modeling with radial basis functions, where an obvious advancement would be to apply regularization in *regional* gravity field modeling. The time frame of this thesis allowed only global gravity field modeling cases with Shannon RBF, where disturbing potential observations were analyzed, and the disturbing potential subsequently synthesized. A natural next step would be to look at radial basis functions in terms of other gravity field functionals, e.g., to analyze gravity anomalies and subsequently synthesize geoid heights with radial basis functions.

Other RBFs are deemed to be better suited for regional gravity field modeling than Shannon RBF. Therefore, further investigations with varying kernels and network design are also of interest.

Bibliography

Aster, R. C., Borchers, B. and Thurber, C. H. (2005), *Parameter Estimation and Inverse Problems*, Elsevier Academic Press.

Bajracharya, S. (2003), Terrain effects on geoid determination, Master's thesis, University of Calgary.

Bentel, K. (2010), 'Regional gravity modeling'. Lectures by Michael Schmidt.

Bentel, K., Schmidt, M. and Gerlach, C. (2013), 'Different Radial Basis Functions and their Applicability for Regional Gravity Field Representation on the Sphere', *International Journal on Geomathematics* .

Blakely, R. J. (1996), *Potential Theory in Gravity & Magnetic Applications*, Cambridge University Press.

Boener, J. H. (1994), Monterey bay geoid, Master's thesis, Naval Postgraduate School.

Breili, K. (2013), 'Geodesy — the foundation for precise geosciences', *Kart og Plan* **73**, 18–31.

Church, J. A., Woodworth, P. L., Aarup, T. and Stanley Wilson, W. e. (2010), *Understanding sea-level rise and variability*, Wiley-Blackwell.

Claessens, S. and Featherstone, W. (2007), 'Ellipsoidal Corrections for Geoid Computations with Modified Kernels', Paper presented to the IUGG 2007 General Assembly, Perugia, Italy.

de Min, E. (1995), 'A comparison of Stokes numerical integration and collocation, and a new combination technique', *Bulletin Géodésique* **69**, 223–232.

de Min, E. (1996), De geoïde voor Nederland, Technical report, Netherlands Geodetic Commission. Publication 34.

Eicker, A. (2008), Gravity Field Refinement by Radial Basis Functions from In-situ Satellite Data, PhD thesis, Universität Bonn.

Ellmann, A. (2001), Least squares modification of Stokes's formula with application to the Estonian geoid, Licentiate thesis, Royal Institute of Technology (KTH), Stockholm, Sweden.

Ellmann, A. (2005), 'Computation of three stochastic modifications of Stokes's formula for regional geoid determination', *Computers & Geosciences* **31**, 742–755.

Ellmann, A. (2010), 'Considerations on the further improvements of regional geoid modeling over the Baltic countries', *Geodezija ir Kartografija / Geodesy and Cartography* **36**, 5–15.

Ellmann, A. (2012), 'Using high-resolution spectral models of gravity anomaly for computing stochastic modifications of Stokes's formula', *Computers & Geosciences* **39**, 188–190.

Fei, Z. L. and Sideris, M. G. (2000), 'A new method for computing the ellipsoidal correction for Stokes's formula', *Journal of Geodesy* **74**, 223–231.

Flury, J. (2006), 'Short-wavelength spectral properties of the gravity field from a range of regional data sets', *Journal of Geodesy* **79**, 624–640.

Forsberg, R. (1984), A study of terrain reductions, density anomalies and geophysical inversion methods in gravity field modelling, Technical report, Department of Geodetic Science and Surveying, The Ohio State University. Tech. Rep. 355.

Geodesy for the layman (1983), Defense Mapping Authority. Tech. Rep. 80-003.

Gerlach, C. (2003), Zur Höhensystemumstellung und Geoidberechnung in Bayern, PhD thesis, Deutsche Geodätische Kommission.

Gerlach, C. (2009), 'Additional notes on gravity field modelling'. Lecture notes to the graduate course in geodesy (GMGD300).

Gerlach, C. and Pettersen, B. R. (2010), 'Validation of GOCE with Terrestrial Gravity Data in Norway', Poster presented to the ESA 2010 Living Planet Symposium, Bergen, Norway.

Haagmans, R., de Min, E. and von Gelderen, M. (1993), 'Fast evaluation of convolution integrals on the sphere using 1D FFT, and a comparison with existing methods for Stokes's integral', *Manuscripta Geodaetica* **18**, 227–241.

Hansen, P. C. and O'Leary, D. P. (1993), 'The use of the L-curve in the regularization of discrete ill-posed problems', *SIAM J. Sci. Comput.* **14**, 1487–1503.

Hofmann-Wellenhof, B. and Moritz, H. (2005), *Physical geodesy*, Springer-Verlag.

Holmes, S. A. and Featherstone, W. E. (2002), 'A unified approach to the Clenshaw summation and the recursive computation of very high degree and order normalized associated Legendre functions', *Journal of Geodesy* **76**, 279–299.

Jekeli, C. (1981), 'Modifying Stokes's Function to Reduce the Error of Geoid Undulation Computations', *Journal of Geophysical Research* **86**, 6985–6990.

Kaula, W. M. (1966), *Theory of Satellite Geodesy*, Blaisdell Publishing Company.

Knudsen, P. (1987), 'Estimation and modelling of the local empirical covariance function using gravity and satellite altimeter data', *Bulletin Géodésique* **61**, 145–160.

Leick, A. (2004), *GPS Satellite Surveying,* 3^{rd} *edition*, Wiley.

Moritz, H. (1980), *Advanced physical geodesy*, Wichmann, Karlsruhe (reprint 2001 by Civil and Environmental Engineering and Geodetic Science, Ohio State University, Columbus, Ohio).

Omang, O. C. D. and Forsberg, R. (2000), 'How to handle topography in practical geoid determination: three examples', *Journal of Geodesy* **74**, 458–466.

Omang, O. C. D., Tscherning, C. C. and Forsberg, R. (2012), 'Generalizing the Harmonic Reduction Procedure in Residual Topographic Modeling', in VII Hotine-Marussi Symposium on Mathematical Geodesy, IAG Symposia 137.

Ophaug, V. (2013), Topics within physical geodesy. Unpublished report, individual course work.

Petit, G. and Luzum, B. e. (2010), IERS Conventions (2010), Technical report, International Earth Rotation and Reference Systems Service (IERS). Tech. Note 36.

Rummel, R., Rothacher, M. and Beutler, G. (2005), 'Integrated Global Geodetic Observing System (IGGOS) — science rationale', *Journal of Geodynamics* **40**, 357–362.

Schmidt, M. (2001), *Grundprinzipien der Wavelet-Analyse und Anwendungen in der Geodäsie*, Shaker Verlag, Aachen. Postdoctoral thesis.

Schwarz, K. P., Sideris, M. G. and Forsberg, R. (1990), 'The use of FFT techniques in physical geodesy', *Geophysical Journal International* **100**, 485–514.

Sjöberg, L. E. (2003), 'A general model for modifying Stokes's formula and its least-squares solution', *Journal of Geodesy* .

Sneeuw, N. (2000), A semi-analytical approach to gravity field analysis from satellite observations, PhD thesis, Deutsche Geodätische Kommission.

Torge, W. and Müller, J. (2012), *Geodesy*, Walter de Gruyter.

Tscherning, C. C. (1983), Determination of a (quasi)geoid for the Nordic Countries from heterogeneous data using collocation, *in* 'Proc. Int. Symp. Figure of the Earth, Moon and other Planets', VÚGTK.

Tscherning, C. C. and Rapp, R. H. (1974), Closed covariance expressions for gravity anomalies, geoid undulations and deflections of the vertical implied by anomaly degree variance models, Technical report, Department of Geodetic Science and Surveying, The Ohio State University. Tech. Rep. 208.

Šprlák, M. (2010), 'Generalized geoidal estimators for deterministic modifications of spherical Stokes function', *Contributions to Geophysics and Geodesy* .

Wahr, J. M. (2009), 'Time-Variable Gravity From Satellites', in Geodesy: Treatise on Geophysics Vol. 3. Elsevier.

Wenzel, H. G. (1981), 'Zur Geoidbestimmung durch Kombination von Schwereanomalien und einem Kugelfunktionsmodell mit Hilfe von Integralformeln', *Zeitschrift für Vermessungswesen* .

Wittwer, T. (2009), Regional gravity field modelling with radial basis functions, PhD thesis, Netherlands Geodetic Commission.

Ågren, J. (2004), Regional Geoid Determination Methods for the Era of Satellite Gravimetry, PhD thesis, Royal Institute of Technology (KTH), Stockholm, Sweden.

Ågren, J. and Sjöberg, L. E. (2004), Comparison of some methods for modifying Stokes's formula in the GOCE era, *in* 'Proceedings of 2nd International GOCE user workshop "GOCE, the geoid and oceanography"', ESA-ESRIN.