

Department of Spatial Sciences

**Solutions to Ellipsoidal Boundary Value Problems
for Gravity Field Modelling**

Sten Johan Claessens

**This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University of Technology**

September 2006

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature:

Date:

ABSTRACT

The determination of the figure of the Earth and its gravity field has long relied on methodologies that approximate the Earth by a sphere, but this level of accuracy is no longer adequate for many applications, due to the advent of new and advanced measurement techniques. New, practical and highly accurate methodologies for gravity field modelling that describe the Earth as an oblate ellipsoid of revolution are therefore required. The foundation for these methodologies is formed by solutions to ellipsoidal geodetic boundary-value problems.

In this thesis, new solutions to the ellipsoidal Dirichlet, Neumann and second-order boundary-value problems, as well as the fixed- and free-geodetic boundary-value problems, are derived. These solutions do not rely on any spherical approximation, but are nevertheless completely based on a simple spherical harmonic expansion of the function that is to be determined. They rely on new relations among spherical harmonic base functions. In the new solutions, solid spherical harmonic coefficients of the desired function are expressed as a weighted summation over surface spherical harmonic coefficients of the data on the ellipsoidal boundary, or alternatively as a weighted summation over coefficients that are computed under the approximation that the boundary is a sphere.

Specific applications of the new solutions are the computation of geopotential coefficients from terrestrial gravimetric data and local or regional gravimetric geoid determination. Numerical closed-loop simulations have shown that the accuracy of geopotential coefficients obtained with the new methods is significantly higher than the accuracy of existing methods that use the spherical harmonic framework. The ellipsoidal corrections to a Stokesian geoid determination computed from the new solutions show strong agreement with existing solutions. In addition, the importance of the choice of the reference sphere radius in Stokes's formula and its effect on the magnitude and spectral sensitivity of the ellipsoidal corrections are pointed out.

ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to my supervisor Prof Will Featherstone, for giving me the opportunity to pursue the degree of Doctor of Philosophy under his guidance, and for his advice and support during this research. The funding of my scholarship from Australian Research Council grant DP0211827 is also duly acknowledged.

Secondly, I am grateful to Dr Simon Holmes, formerly at the Department of Spatial Sciences of Curtin University of Technology and currently at SGT in Greenbelt, USA, for providing very-high-degree and order data sets for the numerical tests performed in this research.

In addition, the use of the software for spherical harmonic synthesis and analysis written by Dr Kurt Seitz of the University of Karlsruhe, Germany, and my co-supervisor Dr Michael Kuhn, was vital in the numerical verification of the newly derived solutions to ellipsoidal boundary-value problems, and I would therefore like to thank both authors.

Moreover, I am grateful to the examiners of this thesis, Prof Lars Sjöberg of the Royal Institute of Technology, Sweden, and Prof Nico Sneeuw of the University of Stuttgart, Germany, for their thorough reviews and useful comments.

Thanks are also extended to all current and former members of the Department of Spatial Sciences, who all together create such an inspiring and enjoyable working atmosphere.

Finally, I would like to thank my parents and sister for their support in my decision to migrate from The Netherlands to Australia, and especially my wife Deavi, for her unconditional love and understanding.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
LIST OF FIGURES	xv
LIST OF TABLES	xvii
LIST OF ACRONYMS	xviii
LIST OF SYMBOLS	xix
1. INTRODUCTION	1
1.1 Historical overview of physical geodesy	1
1.2 Ellipsoidal geodetic boundary-value problems	3
1.2.1 Conditions in boundary-value problems	3
1.2.2 Geodetic boundary conditions	4
1.3 Solutions to ellipsoidal geodetic boundary-value problems	5
1.3.1 Solutions using boundary integral equations	6
1.3.2 Solutions using spherical harmonic expansions	7
1.3.3 Solutions using ellipsoidal harmonic expansions	8
1.4 Research objectives	9
1.4.1 Analytic solutions	10
1.4.2 Spherical harmonic approach	10
1.4.3 Terrestrial gravimetric data	12
1.5 Thesis structure	13
2. BACKGROUND TO ELLIPSOIDAL PHYSICAL GEODESY	15
2.1 The geometry of the ellipsoid	15
2.1.1 Coordinate systems	15
2.1.2 Parameters with respect to the ellipsoid	18
2.1.3 Coordinate transformations	21
2.2 The Earth's gravity field	22
2.2.1 Newton's universal law of gravitation	22

2.2.2	Gravity	23
2.2.3	The external gravity field: A harmonic vector field	24
2.2.4	The reference gravity field	26
2.3	Geodetic boundary-value problems	29
2.3.1	Types of boundary-value problems	30
2.3.2	Geodetic boundary conditions	32
2.3.3	Approximation effects in geodetic boundary conditions	37
2.4	Harmonic base functions	43
2.4.1	Solid spherical harmonics	43
2.4.2	Surface spherical harmonics.....	46
2.4.3	Ellipsoidal harmonics.....	47
2.5	Summary	48
3.	A SOLUTION TO THE ELLIPSOIDAL DIRICHLET BOUNDARY-VALUE PROBLEM IN SPHERICAL HARMONICS	50
3.1	Solution strategies to ellipsoidal boundary-value problems	50
3.1.1	The ellipsoidal harmonics method.....	51
3.1.2	The upward-continuation method	52
3.1.3	The ellipsoidal integration method	53
3.1.4	The coefficient transformation method.....	53
3.2	New relations among spherical harmonic base functions.....	55
3.2.1	Basic recursive relations.....	55
3.2.2	New relations	56
3.2.3	Computation of Legendre weight functions.....	57
3.2.4	Numerical accuracy of Legendre weight functions.....	61
3.2.5	First- and second-order derivatives of spherical harmonics	64
3.3	The coefficient transformation method.....	66
3.3.1	Transformation from solid to surface spherical harmonic coefficients	66
3.3.2	Transformation from surface to solid spherical harmonics.....	70
3.4	Numerical closed-loop simulation.....	73
3.4.1	Computation of surface spherical harmonic coefficients.....	74
3.4.2	Computation of solid spherical harmonic coefficients	81
3.5	Summary	85

4. SPECTRAL RELATIONS BETWEEN GRAVITY FIELD QUANTITIES: THE ELLIPSOIDAL MEISSL SCHEME	86
4.1 Relation between the disturbing potential and its normal derivative	86
4.1.1 The radial derivative of the disturbing potential	86
4.1.2 The derivative with respect to the ellipsoidal normal	88
4.1.3 The normal derivative of the disturbing potential	89
4.2 Relation between the disturbing potential and its second normal derivative	95
4.2.1 The second radial derivative of the disturbing potential	96
4.2.2 The second derivative with respect to the ellipsoidal normal	97
4.2.3 The second normal derivative of the disturbing potential	98
4.3 Relation between the disturbing potential and the gravity anomaly	101
4.3.1 Spherical and constant radius approximation	102
4.3.2 The rigorous ellipsoidal relation	103
4.4 The Meissl scheme	107
4.4.1 The spherical Meissl scheme	108
4.4.2 The ellipsoidal Meissl scheme	110
4.5 Summary	116
5. ANALYTIC APPROACHES TO THE CONSTRUCTION OF A GLOBAL SPHERICAL HARMONIC GEOPOTENTIAL MODEL	118
5.1 Methodologies	118
5.1.1 Least-squares estimation	119
5.1.2 Numerical quadrature	121
5.1.3 The ellipsoidal harmonics method	122
5.1.4 Alternative numerical quadrature methods	124
5.2 Method 1: Upward continuation	125
5.2.1 Solution strategy	126
5.2.2 Expansion of the correction terms	127
5.2.3 Computation of the geopotential coefficients	129
5.3 Method 2: Ellipsoidal integration	131
5.3.1 Solution strategy	132
5.3.2 Expansion of the integration kernels	133
5.3.3 Computation of geopotential coefficients	136

5.4	Method 3: Coefficient transformation	137
5.5	Comparison of the upward continuation, ellipsoidal integration and coefficient transformation methods with existing methods	138
5.5.1	Comparison of weights	139
5.5.2	Closed-loop simulation	143
5.5.3	Comparisons in the space domain	145
5.5.4	Computation of very-high-degree and -order geopotential coefficients	152
5.6	Summary	154
6.	ELLIPSOIDAL CORRECTIONS TO GRAVIMETRIC GEOID COMPUTATIONS	155
6.1	The spherical approximation to geoid computation	155
6.1.1	Definition of spherical approximation	156
6.1.2	Stokes's formula	158
6.2	Existing methods for the computation of ellipsoidal corrections	159
6.2.1	Types of ellipsoidal corrections	160
6.2.2	Ellipsoidal corrections from a global integration	160
6.2.3	Ellipsoidal corrections from spherical harmonic geopotential coefficients	163
6.3	The surface harmonics method to ellipsoidal corrections	164
6.3.1	Derivation of the correction formulas	164
6.3.2	The zero- and first-degree coefficients	166
6.3.3	Ellipsoidal corrections for various reference spheres	168
6.3.4	Ellipsoidal corrections with varying reference sphere radius	173
6.4	Ellipsoidal corrections in view of the remove-compute-restore (RCR) approach to geoid computation	176
6.4.1	The remove-compute-restore approach	176
6.4.2	High-frequency ellipsoidal corrections	179
6.4.3	Magnitude of ellipsoidal corrections in various bandwidths	181
6.5	Summary	186
7.	SUMMARY, CONCLUSIONS AND OUTLOOK	187
7.1	Summary	187

7.2	Conclusions.....	188
7.3	Outlook.....	193
	REFERENCES	196
	APPENDIX A: RELATION AMONG FULLY NORMALISED LEGENDRE WEIGHT FUNCTIONS	214
	APPENDIX B: REARRANGEMENT OF SUMMATION ORDER	216
	APPENDIX C: BINOMIAL SERIES EXPANSION	219

LIST OF FIGURES

Figure 2.1	Geocentric (θ), geodetic (ϑ) and reduced (β) co-latitude	17
Figure 2.2	The spherical polar (r, θ, λ) and geodetic (h, ϑ, λ) coordinate systems .	17
Figure 2.3	Categorisation of vector fields by investigation of their curl and divergence	25
Figure 2.4	Relations between the Earth surface, mean sea level, the ellipsoid and the geoid: the ellipsoidal height is measured from the ellipsoid to the Earth surface, the geoid height from the ellipsoid to the geoid, the sea surface topography from the geoid to mean sea level (all three along the ellipsoidal normal) and the orthometric height from the geoid to the Earth surface along the plumbline of gravity, where all four measures are positive in the external direction.....	27
Figure 2.5	The deflections of the vertical on the surface of the ellipsoid	41
Figure 3.1	Schematic overview of the computation of solid SHCs \bar{f}_{nm}^R from function values on the ellipsoid $f(r_e, \theta, \lambda)$ using the 1. upward-continuation, 2. ellipsoidal integration, and 3. coefficient transformation methods ..	54
Figure 3.2	Recursive computation of cosinusoidal LWFs F_{nm}^{ij}	58
Figure 3.3	Recursive computation of cosinusoidal LWF F_{nm}^{ij} for even j	60
Figure 3.4	Average absolute relative error ς_n (Equation 3.47) per degree n in the computation of fully normalised cosinusoidal LWFs in dual mode (Equation 3.39) for $\theta = 30^\circ$, $\theta = 45^\circ$ and $\theta = 60^\circ$, where the functions are computed in IEEE double precision	62
Figure 3.5	The size of the first ten terms in the binomial series in Equation (3.65) relative to the total sum in percentage for degrees $0 \leq n \leq 2160$. The dashed line indicates the boundary described by Equation (3.66), increased by 0.5 so that it crosses the border between two terms where those terms are equal.	68

Figure 3.6 The absolute size of the first fifteen terms in the binomial series in Equation (3.65) for degrees $0 \leq n \leq 2160$. The dashed line indicates the boundary described by Equation (3.66), increased by 0.5 so that it crosses the border between two terms where those terms are equal. The solid lines indicate until which term the series should be evaluated to ensure a relative accuracy of at least e^4 , e^2 and 1 respectively (from left to right). 68

Figure 3.7 Full matrix form of Equation (3.75) for SHCs to degree and order n . . . 71

Figure 3.8 Logarithm of the ratio between $|\lambda_{nm0}|$ and $\sum_{i=-50, i \neq 0}^{50} |\lambda_{nmi}|$ for all degrees and orders up to 720, where the line starting at degree 520 indicates for which pairs of degree and order the ratio is closest to unity 72

Figure 3.9 Numerical tests comparing the surface SHCs (left), the function values on the ellipsoid (centre) and the solid SHCs (right) via different ways of computation 75

Figure 3.10 Degree variances of surface SHCs \bar{T}_{nm}^{re} (black line), their differences with the degree variances of the solid SHCs \bar{T}_{nm}^R (red line), and the differences between the surface SHCs from the synthesis-analysis procedure (Equation 3.83) and those from the new transformation formula (Equation 3.81) (green line), where the summation over i was performed from -20 to 20 76

Figure 3.11 Absolute relative difference between surface harmonic coefficients \bar{T}_{nm}^{re} computed using the synthesis-analysis procedure (Equation (3.83) and the new transformation formula (Equation 3.81) for $n + m$ is even (black) and $n + m$ is odd (red), where the summation over i was performed from -20 to 20 78

Figure 3.12 Disturbing potential of the Earth's gravity field from the EGM96 geopotential model and GRS80 reference gravity field (units in m^2s^{-2} ; Robinson projection) 78

Figure 3.13	Differences between the disturbing potential from a spherical harmonic synthesis of the EGM96 geopotential coefficients and surface spherical harmonic coefficients obtained via a coefficient transformation up to degree and order 360 for the region of Indonesia (units in m^2s^{-2} ; Mercator projection)	80
Figure 3.14	Differences between the disturbing potential from a spherical harmonic synthesis of surface spherical harmonic coefficients obtained via synthesis-analysis procedure and via a coefficient transformation up to degree and order 340 for the region of Indonesia (units in m^2s^{-2} ; Mercator projection)	80
Figure 3.15	Degree variances of solid SHCs \bar{T}_{nm}^R (black line), and their absolute differences with the degree variances of the solid SHCs \bar{T}_{nm}^R obtained after a forward and backward transformation, where the backward transformation is performed with 20 iterations (red line) and 50 iterations (green line). The summation over i was performed from -20 to 20 and the degree variance differences of zero were set to 10^{-32}	82
Figure 3.16	Degree variances of solid SHCs \bar{T}_{nm}^R (black line), and their absolute differences with the degree variances of the solid SHCs \bar{T}_{nm}^R obtained after a forward and backward transformation, where the backward transformation is performed with 120 iterations and the summation over i was performed from -50 to 50	84
Figure 4.1	Normal derivative to the ellipsoid	89
Figure 4.2	Values of $ \lambda_{n00} $ (black) and $\sum_{i=-\infty, i \neq 0}^{\infty} \lambda_{n0i} $ (green) from Equation (3.74), as well as $ \lambda_{n00}(d_h T, T) $ (red) and $\sum_{i=-\infty, i \neq 0}^{\infty} \lambda_{n0i}(d_h T, T) $ (blue) from Equation (4.34), indicating the diagonal dominance of the weights below degree $n = 520$ for both cases	94
Figure 4.3	Spherical Meissl scheme (adapted from Rummel and Van Gelderen 1995)	108
Figure 4.4	Ellipsoidal Meissl scheme	110
Figure 4.5	Connections between the spherical and ellipsoidal Meissl schemes (black dashed lines), given by Equations (3.73), (4.33) and (4.57)	111

Figure 4.6	Logarithm of the ratio between $ \lambda_{nm0}(d_h T, T) $ and $\sum_{i=-50, i \neq 0}^{50} \lambda_{nmi}(d_h T, T) $ for a transformation between surface SHCs of a function and its normal derivative on the ellipsoid ($0 \leq n \leq 720$)	115
Figure 5.1	Schematic overview of three methods to compute spherical harmonic geopotential coefficients from gravity anomalies on the ellipsoid	125
Figure 5.2	Convergence rate in the computation of the weights for the case $i = 0$ (summation over k in Equation (5.63)): number of terms with an absolute value higher than e^4 for all pairs of degree n and order m ...	140
Figure 5.3	Convergence rate in the computation of the coefficients (summation over i in Equation (5.62)): number of terms with an absolute value higher than e^4 for all pairs of degree n and order m	140
Figure 5.4	Weight functions $\lambda_{nmi}(T, \tilde{T})$ for zonal harmonics ($m = 0$) with $i = 0$ (solid lines), and $ i = 1$ (dashed lines) for the methods of Cruz (1986) and Sjöberg (2003c) (green), the method of Petrovskaya et al. (2001) (blue), the new ellipsoidal integration (ei) method (black) and the new coefficient transformation (ct) method (red). The weights of $ i = 2$ are also shown for the coefficient transformation method (dotted red line).	142
Figure 5.5	The degree variances of EGM96 (black) and the absolute errors in the degree variances after a closed-loop simulation using the spherical approximation (red), the methods by Cruz (1986) and Sjöberg (2003c) (green), the method by Petrovskaya et al. (2001) (blue), the ellipsoidal harmonics method by Jekeli (1988) (magenta), the new ellipsoidal integration method (yellow) and the new coefficient transformation method (brown)	144
Figure 5.6	Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the spherical approximation for $0 \leq n \leq 360$ (units in m; Robinson projection)	146

Figure 5.7	Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the method of Cruz (1986) for $20 \leq n \leq 340$ (units in m; Robinson projection).....	146
Figure 5.8	Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the method of Petrovskaya et al. (2001) for $20 \leq n \leq 340$ (units in m; Robinson projection).....	147
Figure 5.9	Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the ellipsoidal integration method for $20 \leq n \leq 340$ (units in m; Robinson projection).....	147
Figure 5.10	Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the coefficient transformation method for $20 \leq n \leq 340$ (units in m; Robinson projection).....	148
Figure 5.11	Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the ellipsoidal harmonics method by Jekeli (1988) for $20 \leq n \leq 340$ (solution provided by Holmes (2004, pers. comm.)) (units in m; Robinson projection).....	148
Figure 5.12	Relative accuracy of SHCs \bar{T}_{nm}^R computed from gravity anomalies using Jekeli's (1988) method (black) and the coefficient transformation method (red) compared to the SHCs resulting from a synthesis-analysis procedure for (a) $n = 30$, (b) $n = 90$, (c) $n = 180$ and (d) $n = 300$	151
Figure 5.13	Weight functions $\lambda_{nmi}(T, \tilde{T})$ for zonal harmonics ($m = 0$) with $i = 0$ (black), $ i = 1$ (red), $ i = 2$ (green) and $ i = 3$ (blue) for the ellipsoidal integration method (solid lines) and the coefficient transformation method (dashed lines).....	153

Figure 6.1	Geoid surface computed from Equation (6.1) using the EGM96 global geopotential model and GRS80 reference ellipsoid (units in m; Robinson projection)	157
Figure 6.2	Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 1$ (units in m; Robinson projection)	167
Figure 6.3	Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 340$ and a reference sphere with radius $R = a$ (units in m; Robinson projection)	170
Figure 6.4	Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 340$ and a reference sphere with radius $R = b$ (units in m; Robinson projection)	170
Figure 6.5	Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 340$ and a reference sphere with radius $R = (a^2b)^{1/3}$ (units in m; Robinson projection)	171
Figure 6.6	Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 340$ and a reference sphere with radius $R = a + 20\text{km}$ (units in m; Robinson projection)	171
Figure 6.7	Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 340$ and $R = r_e$ (units in m; Robinson projection)	174
Figure 6.8	From top to bottom: Degree variances of the solid SHCs \overline{T}_{nm}^R (black) and of the surface SHCs $\overline{\delta T}_{nm}^{r_e}$ for $R = a + 20\text{km}$ (yellow), $R = b$ (green), $R = a$ (red), $R = (a^2b)^{1/3}$ (blue) and $R = r_e$ (brown)	175
Figure 6.9	Ellipsoidal correction term δN_2 (Equation 6.38) from the EGM96 geopotential model with degrees $90 \leq n \leq 340$ and a reference sphere with radius $R = a$ (units in m; Mercator projection)	182

Figure 6.10	Ellipsoidal correction term δN_2 (Equation 6.38) from the EGM96 geopotential model with degrees $90 \leq n \leq 340$ and a reference sphere with radius $R = (a^2b)^{1/3}$ (units in m; Mercator projection)	182
Figure 6.11	Ellipsoidal correction term δN_2 (Equation 6.38) from the EGM96 geopotential model with degrees $90 \leq n \leq 340$ and a reference sphere with radius $R = r_e$ (units in m; Mercator projection).....	183
Figure 6.12	Ellipsoidal correction term δN_2 (Equation 6.38) from the EGM96 geopotential model with degrees $20 \leq n \leq 340$ and a reference sphere with radius $R = r_e$ (units in m; Mercator projection).....	184
Figure 6.13	Geoid computed from the EGM96 geopotential model with degrees $20 \leq n \leq 340$ (units in m; Mercator projection)	184
Figure B-1	Schematic overview of a rearrangement of summation order from indices i and j to indices k and l	217

LIST OF TABLES

Table 2.1	Expressions for the ellipsoidal radius (r_e), meridian radius of curvature (ρ) and prime vertical radius of curvature (ν) as a function of geocentric (θ), geodetic (ϑ) and reduced (β) co-latitude	19
Table 2.2	Expressions for the sine, cosine and tangent of the ellipsoidal deflection angle (ϕ) as a function of geocentric (θ), geodetic (ϑ) and reduced (β) co-latitude	20
Table 2.3	Coordinate transformation from spherical polar (r, θ, λ), geodetic (h, ϑ, λ) and ellipsoidal (u, β, λ) coordinates to Cartesian (x, y, z) coordinates ..	21
Table 2.4	Coordinate transformation from Cartesian (x, y, z) coordinates to spherical polar (r, θ, λ), geodetic (h, ϑ, λ) and ellipsoidal (u, β, λ) coordinates on the surface of the ellipsoid	21
Table 2.5	Parameters of the Geodetic Reference System 1980 GRS80 (from Moritz, 1980)	29
Table 3.1	Average absolute relative error ς_j per index j (Equation 3.46) in the computation of fully normalised and non-normalised cosinusoidal LWFs in single mode (Equation 3.33) or dual mode (Equation 3.39) for $\theta = 30^\circ$, $\theta = 45^\circ$ and $\theta = 60^\circ$, where the functions are computed in IEEE double precision	63
Table 3.2	Global statistics of the comparison between the values of the disturbing potential on the ellipsoid from a spherical harmonic synthesis of surface SHCs obtained via a coefficient transformation and 1) a spherical harmonic synthesis of the EGM96 geopotential coefficients up to degree and order 360, and 2) a spherical harmonic synthesis of the surface coefficients obtained via a synthesis-analysis procedure up to degree and order 340 (in m^2s^{-2})	79

Table 4.1	Eigenvalues forming the diagonal elements in the matrices of the spherical Meissl scheme (Figure 4.3).....	109
Table 5.1	Statistics of the errors in the geoid height from solid SHCs of the disturbing potential computed using various methods for $20 \leq n \leq 340$ (units in m), where <i>abs. mean</i> is the average of the absolute values of the errors.....	149
Table 6.1	Statistics of the ellipsoidal corrections for a reference sphere with radius R equal to semi-minor axis b , ‘mean’ radius $(a^2b)^{1/3}$, semi-major axis a and semi-major axis $a + 20\text{km}$, obtained from geopotential model EGM96 and degrees $0 \leq n \leq 340$ synthesised on a 30’ grid (units in m)	172
Table 6.2	Global statistics of the ellipsoidal corrections for a reference sphere with radius R equal to semi-minor axis b , ‘mean’ radius $(a^2b)^{1/3}$, semi-major axis a and ellipsoidal radius r_e , obtained from geopotential model EGM96 and degrees $90 \leq n \leq 340$ synthesised on a 30’ grid (units in mm)	185
Table 6.3	Global statistics of the ellipsoidal corrections for a reference sphere with radius R equal to semi-minor axis b , ‘mean’ radius $(a^2b)^{1/3}$, semi-major axis a and ellipsoidal radius r_e , obtained from geopotential model EGM96 and degrees $20 \leq n \leq 340$ synthesised on a 30’ grid (units in mm)	185

LIST OF ACRONYMS

ALF	Associated Legendre Function	44
BVP	Boundary-Value Problem	2
CHAMP	Challenging Mini-satellite Payload	12
CODATA	[US] Committee on Data for Science and Technology	23
EGM	Earth Geopotential Model	11
EIGEN	European Improved Gravity Model of the Earth by New Techniques	11
GGM	Global Gravity Model	12
GOCE	Gravity Field and Steady-State Ocean Circulation Explorer ...	12
GPM	Global Potential Model	11
GPS	Global Positioning System	4
GRACE	Gravity Recovery and Climate Experiment	12
GRS	Geodetic Reference System	23
IEEE	Institute of Electrical and Electronic Engineers	62
LSC	Least-Squares Collocation	7
LWF	Legendre Weight Function	57
OSU	Ohio State University	125
RCR	Remove-Compute-Restore	9
SHC	Spherical Harmonic Coefficient	7
SHF	Spherical Harmonic Function	43
SOR	Successive Over-Relaxation	73

LIST OF SYMBOLS

A	design matrix	120
a	semi-major axis of an ellipsoid	16
\vec{a}	acceleration vector	23
a_n	arbitrary constant in ellipsoidal integration method	53
b	semi-minor axis of an ellipsoid	16
c	scale factor between semi-major axis and spherical radius	66
$\overline{d_r T_{nm}^R}$	surface SHCs of first radial derivative of disturbing potential ..	87
$\overline{d_h T_{nm}^{r_e}}$	surface SHCs of first normal derivative of disturbing potential ..	90
$\overline{d_r^2 T_{nm}^R}$	surface SHCs of 2 nd radial derivative of disturbing potential ...	96
$\overline{d_h^2 T_{nm}^{r_e}}$	surface SHCs of 2 nd normal derivative of disturbing potential ..	100
E	linear eccentricity of an ellipse	18
$E(\theta)$	auxiliary function in the ellipsoidal correction	178
\mathcal{E}	ellipsoidal surface	53
e	first numerical eccentricity of an ellipse	18
e'	second numerical eccentricity of an ellipse	18
e_γ	function of ellipsoidal flattening and gravity flattening	104
$F(\alpha, \beta, \gamma, x)$	hypergeometric function	48
F_{nm}^{ij}	cosinusoidal LWFs	55
\overline{F}_{nm}^{ij}	fully normalised sinusoidal LWFs	55
\vec{F}	force vector	22
\mathcal{F}	arbitrary function	31
\mathcal{F}^R	vector containing solid SHCs of harmonic function	70
\mathcal{F}^{r_e}	vector containing surface SHCs of harmonic function	70
f	flattening of an ellipse	18
$f(r, \theta, \lambda)$	harmonic function outside an ellipsoid	50
\overline{f}_{nm}^R	fully normalised solid SHCs of harmonic function	50
$\overline{f}_{nm}^{r_e}$	fully normalised surface SHCs of harmonic function	53
\hat{f}_{nm}^u	ellipsoidal harmonic coefficients of harmonic function	51
G	universal gravitational constant	23

g	magnitude of gravity	34
\vec{g}	gravity vector	24
H	orthometric height	33
h	ellipsoidal height	15
h, ϑ, λ	geodetic coordinates	15
i, j, k, l, p, q	summation indices	34
K_{nm}^{ij}	sinusoidal LWFs	57
\overline{K}_{nm}^{ij}	fully normalised sinusoidal LWFs	57
L_{nmk}	auxiliary function in ellipsoidal harmonics method	122
M	mass	22
m	order of spherical harmonic expansion	43
\overline{m}	geodynamic parameter	132
N	geoid height	33
\tilde{N}	spherically approximated geoid height	157
N_1	geoid height contribution from global model in the RCR approach	176
N_2	geoid height contribution from local data in the RCR approach .	176
N_{\max}	maximum available degree of global geopotential model	121
\overline{N}_{nm}^{ij}	fully normalised LWF for first-order derivatives of ALFs	65
n	degree of spherical harmonic expansion	43
\vec{n}	normal unit vector of a point on a surface	31
n_{\max}	maximum degree used in RCR geoid determination	176
O	origin of coordinate system	17
\mathcal{O}	Landau symbol	37
P	point in space	16
P'	point on the geoid	33
\mathbf{P}	covariance matrix	120
$P_{nm}(\theta)$	associated Legendre function	44
$\overline{P}_{nm}(\theta)$	fully normalised associated Legendre function	44
Q	point in space	23
Q'	point on an ellipsoid	28
$Q_{nm}(\theta)$	associated Legendre functions of the second kind	47
q_{in}	Pellinen smoothing factor	121

R	spherical radius	38
R_s	spherical radius at satellite altitude	108
\mathcal{R}	arbitrary radius	113
\overline{R}_{nm}^{ij}	fully normalised LWF for second-order derivatives of ALFs ...	65
r	geocentric radius	15
r, θ, λ	spherical polar coordinates	15
$r_e(\theta)$	ellipsoidal radius for co-latitude θ	18
$r_s(\theta)$	ellipsoidal radius at satellite altitude for co-latitude θ	110
S	auxiliary potential function	132
\mathcal{S}	closed surface	31
\overline{S}_{nm}	real normalised ALFs of the second kind	122
$S(\psi)$	spherical Stokes kernel	159
$S_e(\theta, \lambda, \theta', \lambda')$	ellipsoidal Stokes kernel	161
s	arc length	19
T	disturbing potential	26
\tilde{T}	spherically approximated disturbing potential	130
T_{nm}^R	solid SHCs of disturbing potential	43
\overline{T}_{nm}^R	fully normalised solid SHCs of disturbing potential	45
$\overline{T}_{nm}^{r_e}$	fully normalised surface SHCs of disturbing potential	47
\tilde{T}_{nm}^R	spherically approximated solid SHCs of disturbing potential ..	121
\hat{T}_{nm}^u	ellipsoidal harmonic coefficients of disturbing potential	47
\mathcal{T}	vector containing SHCs of disturbing potential	109
t_n	degree variances of disturbing potential	75
U	normal potential	26
u, β, λ	ellipsoidal coordinates	15
V	gravitational potential	25
\mathcal{V}	volume	30
\vec{v}	arbitrary vector	24
W	gravity potential	26
W_0	gravity potential at the geoid (constant)	26
w, z	two-dimensional Cartesian coordinates	16
x, y, z	three-dimensional Cartesian coordinates	15
\vec{x}	3D position vector	30

$Y_{nm}(\theta)$	spherical harmonic function	43
$\bar{Y}_{nm}(\theta)$	fully normalised spherical harmonic function	45
α_{nk}	auxiliary ellipsoidal weight function	69
β	reduced co-latitude	15
β_{nmij}	auxiliary function in coefficient transformation	91
Γ	gamma-function	219
γ	magnitude of ellipsoidal normal gravity	27
$\tilde{\gamma}$	magnitude of spherical normal gravity	38
γ_a	normal gravity at the equator	28
γ_b	normal gravity at the poles	28
γ_{nmij}	auxiliary function in coefficient transformation	92
Δg	gravity anomaly	35
$\widetilde{\Delta g}$	spherically approximated gravity anomaly	42
$\overline{\Delta g}$	block mean spherically approximated gravity anomaly	119
$\Delta g_{n_{\max}}$	contribution to gravity anomaly from global model	177
$\overline{\Delta g}_{nm}^{r_e}$	fully normalised surface SHCs of gravity anomaly	102
$\widehat{\Delta g}_{nm}^u$	ellipsoidal harmonic coefficients of gravity anomaly	122
$\Delta \mathcal{G}^{r_e}$	vector containing surface SHCs of gravity anomaly	120
$\Delta \sigma$	surface area on a sphere	119
δ_{ij}	Kronecker delta	45
δg	gravity disturbance	34
$\widetilde{\delta g}$	spherically approximated gravity disturbance	42
δN	ellipsoidal correction to geoid height	160
δN_2	ellipsoidal correction to geoid height from local data	180
δS	ellipsoidal correction to Stokes's kernel	160
δT	ellipsoidal correction to disturbing potential	165
$\overline{\delta T}_{nm}^{r_e}$	surface SHCs of ellipsoidal correction to disturbing potential ...	165
$\delta \Delta g$	ellipsoidal correction term to gravity anomaly	126
ϵ	quartic eccentricity of an ellipse	18
ϵ_c	upward-continuation correction term	126
ϵ_e	boundary condition correction term	126
ϵ_{nmij}	auxiliary function in coefficient transformation	105
ζ_{nmij}	auxiliary function in coefficient transformation	100

η	east-west deflection of the vertical	40
η_{nmij}	auxiliary function in coefficient transformation	100
Θ	total deflection of the vertical	39
θ	geocentric co-latitude	15
ϑ	geodetic co-latitude	15
κ	arbitrary constant	31
κ_{nmij}	auxiliary function in coefficient transformation	100
Λ	matrix containing weight functions λ_{nmi}	70
λ	geodetic longitude	15
λ_{nmi}	ellipsoidal weight functions	70
$\lambda_{nmi}^{\text{uc}}$	ellipsoidal weight functions in upward-continuation method	130
$\lambda_{nmi}^{\text{ei}}$	ellipsoidal weight functions in ellipsoidal integration method	136
$\lambda_{nmi}^{\text{ct}}$	ellipsoidal weight functions in coefficient transformation method	137
μ	terrestrial gravitational constant ($= GM$)	27
μ_{in}	auxiliary function in ellipsoidal integration method	135
ν	prime vertical radius of curvature	19
ν_{in}	auxiliary function in ellipsoidal integration method	134
ξ	north-south deflection of the vertical	40
ξ_{in}	auxiliary function in ellipsoidal integration method	134
ρ	meridian radius of curvature	19
ρ_{in}	auxiliary function in ellipsoidal integration method	135
ϱ	mass density	25
σ	unit sphere	45
σ_{in}	auxiliary function in ellipsoidal integration method	135
ς	numerical rounding error or truncation error	61
τ_{nmi}	auxiliary function in ellipsoidal corrections	163
Φ	scalar potential	31
ϕ	ellipsoidal deflection angle	20
χ_{nk}	auxiliary function in upward-continuation method	128
$\vec{\chi}$	centrifugal acceleration	23
ψ	spherical distance	158
Ω	centrifugal potential	26
ω	angular velocity of the Earth's rotation	23

1. INTRODUCTION

This thesis investigates the new field of what is herein termed *ellipsoidal physical geodesy*, which is concerned with the rigorous determination of the figure of the Earth and its external gravity field using an ellipsoidal reference model. In practice, many geodesists still rely on approximate methodologies using spherical models for gravity field modelling, such as Stokes's (1849) integral for geoid computation. However, the introduction of new measurement techniques, and increased accuracy and spatial distribution of traditional observations calls for more accurate theoretical methodologies. An overview of contributions to the field of physical geodesy and the solution strategies chosen in this research are presented in this introductory chapter.

1.1 Historical overview of physical geodesy

The determination of the size and shape of the Earth has been a topic of research since the Greek philosopher Eratosthenes conducted an experiment to calculate the Earth's circumference from solar observations around 250 B.C., but the influence of gravity on the Earth's shape was largely unknown until I. Newton in 1687 derived his universal law of gravitation (Newton, 1729). Newton was also the first to realise that the Earth's shape is closer to an oblate ellipsoid of revolution flattened at the poles, than to a sphere. This was confirmed several decades later by astro-geodetic expeditions undertaken by the French Academy of Sciences.

The origin of physical geodesy as a field of study can be attributed to A. C. Clairaut (e.g., Dragomir et al., 1982), who in 1743 published his theorem that relates the flattening of the Earth to the magnitude of gravity at the poles and the equator. This theorem first showed that the geometry of the Earth can be derived from purely dynamical quantities obtained by gravity measurements (Heiskanen and Moritz, 1967,

p.75).

At the beginning of the nineteenth century, the notion that the figure of the Earth is actually a level surface departing from the ellipsoid, usually credited to C. F. Gauss, arose, and this level surface was named the geoid by J. B. Listing in 1872 (e.g., Moritz, 1990). Developments in potential theory by J. Lagrange, P. S. Laplace and others also revealed that the Earth's gravity field can, due to its harmonic properties, be expressed as a scalar potential field (e.g., MacMillan, 1958). This facilitates the determination of the Earth's external potential from gravity measurements on the surface of the Earth; the geodetic boundary-value problem (BVP), where the boundary condition is formed by the so-called fundamental equation of physical geodesy (e.g., Heiskanen and Moritz, 1967, p. 86). This fundamental equation is a differential equation relating the disturbing potential of the Earth's gravity field to the gravity anomaly.

Probably the most important contribution to physical geodesy was made by Stokes (1849), who found a solution to the geodetic BVP expressing the disturbing potential of the Earth's gravity field as an integral over gravity anomalies weighted by a Green's function, often called the Stokes kernel. However, Stokes's integral only provides a solution to the spherical approximation of the fundamental equation of physical geodesy, where the ellipsoidal reference gravity model is approximated by a spherical reference gravity model to simplify the differential equation.

The error resulting from a spherical approximation is often assumed to be of the order of the flattening of the Earth ($\sim 0.3\%$), but may be even larger due to the combined effect of subsequent approximations (e.g., Sansò and Tscherning, 2003; Tscherning, 2004). For a geoid-ellipsoid separation of 100 m, the error induced by a spherical approximation is therefore in the order of 30 cm and possibly larger. This level of accuracy is not sufficient for many present-day applications.

The accuracy and coverage of gravity measurements are nowadays of high enough quality to provide geoid and gravity field models with relative accuracies well within the order of the flattening of the Earth, which has led to the global aim to compute a geoid model accurate to 1 cm (Rapp, 1997a). Naturally, highly accurate modelling of

the geoid and the Earth's external gravity field is only possible if the required rigorous theoretical methodologies exist. This calls for a solution of a BVP where the boundary is formed by an ellipsoid of revolution, i.e., an ellipsoidal BVP.

1.2 Ellipsoidal geodetic boundary-value problems

BVPs are concerned with the determination of a function in a certain region of space from given conditions on the boundary of the region (Mackie, 1965). Ellipsoidal BVPs thus arise when the boundary is formed by an ellipsoid. As the figure of the Earth can closely (< 100 m) be approximated by an oblate ellipsoid of revolution, ellipsoidal BVPs form the foundation of what is here called ellipsoidal physical geodesy.

1.2.1 Conditions in boundary-value problems

Any BVP consists of a set of conditions that apply to the function under investigation. Common conditions are the harmonicity condition, the regularity or radiation condition, the continuation condition, and most importantly, the boundary condition (e.g., Mackie, 1965; Lehmann, 1999a; Fei, 2000).

The harmonicity condition is given by Laplace's differential equation and is a prerequisite for any BVP in physical geodesy (e.g., Sigl, 1985). The regularity condition prescribes the behaviour of the function at infinity (Lehmann, 1999a) and is therefore of particular importance in BVPs where the domain is external, which is always the case in geodetic applications since the Earth's external gravity field is of interest. In the case of the gravity potential, the regularity condition states that the potential approaches zero for distances going to infinity. The continuation condition is often not explicitly stated in geodetic BVPs, but it is often assumed that the potential is continuous and at least twice differentiable.

The boundary condition is the most important condition, since it connects the function of interest to a known quantity on the boundary, and it typically takes the form of an

ordinary or partial differential equation. The boundary conditions that are of interest in physical geodesy have been studied by many authors (e.g., Grafarend et al., 1985; Heck, 1991; Sansò, 1995). Different geodetic boundary conditions can be constructed, depending mainly on the observational data available, and the accuracy required.

1.2.2 Geodetic boundary conditions

A common categorisation of boundary conditions is a division between fixed and free conditions, depending on whether the size and shape of the boundary are known or not (e.g., Heck, 1991). In the classical approach, the boundary is assumed unknown, giving rise to a free BVP (e.g., Grafarend and Niemeier, 1971; Sansò, 1981). However, with the advent of precise satellite positioning from, for example, the Global Positioning System (GPS), fixed BVPs (e.g., Bjerhammar and Svensson, 1983; Grafarend et al., 1985) are of growing importance.

An additional category is the scalar-free BVP (e.g., Sacerdote and Sansò, 1986), where the horizontal position is assumed known, but the vertical is not. The rationale behind this is that the accuracy of the horizontal position is less critical than the vertical, while it can be measured more accurately using classical triangulation measurements or satellite positioning. The scalar problem not only contains fewer degrees of freedom than the vector free problem, the errors due to linearisation are also approximately one order of magnitude smaller (Heck, 1989). In this thesis, both the fixed- and the scalar-free BVPs are considered.

Due to the introduction of new observation techniques, non-classical boundary conditions can be formed that allow for the combination of different observation types. An example is the altimetry-gravimetry problem which combines satellite altimetry data on the oceans with gravity data on the continents (e.g., Svensson, 1988; Keller, 1996; Lehmann, 1999b; Grebenitcharsky and Sideris, 2005). Multiple measurement techniques can also give rise to overdetermined geodetic BVPs (e.g., Sacerdote and Sansò, 1985; Rummel et al., 1989; Van Gelderen and Rummel, 2001). Naturally, new dedicated satellite gravity field missions also provide new observation types, albeit not

on the same boundary as the terrestrial data.

In practice, however, the gravimetric BVPs are used most frequently in the treatment of terrestrial data. Satellite altimetry data can be transformed into gravity anomalies, and thus incorporated in the fixed- or scalar-free gravimetric BVP (e.g., Rapp and Bašić, 1992; Andersen et al., 1996; Hwang and Parsons, 1996). Multiple data sources that give rise to overdetermined problems are usually combined to provide higher spatial resolution, or to provide a verification of the solution. Therefore, the focus in this thesis is on the gravimetric BVPs. However, an extension of the methodologies developed here to other BVPs can be established as well.

1.3 Solutions to ellipsoidal geodetic boundary-value problems

A possible categorisation of solutions to BVPs is by the mathematical techniques used to rewrite the boundary condition. The BVPs occurring in physical geodesy, either with spherical or ellipsoidal boundaries, are usually solved in one of two ways:

1. Solutions using boundary integral equations based on Green's second identity (e.g., Martensen and Ritter, 1997);
2. Solutions using a representation of the gravity field in harmonic base functions (e.g., MacRobert, 1967),

although it should be mentioned that a combination of both is also possible (e.g., Sjöberg, 2003c). With all approaches, a solution can be sought either in an analytic or in a deterministic way. The second approach can be subdivided in solutions that use spherical harmonic base functions and solutions that use ellipsoidal harmonic base functions. This gives rise to three possible approaches, and the application of these three approaches in ellipsoidal physical geodesy are discussed in the next three Sections.

1.3.1 Solutions using boundary integral equations

Investigations towards the solution of ellipsoidal geodetic BVPs commenced more than 100 years after Stokes published his spherical solution. The works of Sagrebin (1956), Molodenskii et al. (1962), Bjerhammar (1962, 1966), Brovar (1964), Koch (1968, 1969), and Mather (1973) provide the first solutions, with highly varying practicality. All of these early solutions are based on a reformulation of the boundary condition as an integral equation by implementation of Green's second identity.

More recent applications of boundary integral equations are given by Arnold (1986, 1987, 1989, 1990), Yurkina (2002) and Brovar et al. (2001). Wang (1999) uses the boundary integral approach to provide ellipsoidal corrections to gravity anomalies from satellite altimetry, and Sideris et al. (2002), Fei (2000) and Fei and Sideris (2000) apply it to obtain a generalisation of Stokes's and Hotine's integral kernels for an ellipsoidal boundary.

A similarity between all the above-mentioned methods is that none of them provide a theoretically exact solution. Instead, they generally approximate the boundary condition to the level of the first numerical eccentricity of the ellipsoid.

The numerical evaluation of the boundary integral equation can be performed using various deterministic methods such as finite difference methods, finite element methods or boundary element methods (e.g., Ladyzhenskaya, 1985). Finite element methods were introduced to physical geodesy by Meissl (1985), but are rarely used in practice, as are finite difference methods. Boundary element methods, contrary to the other two approaches, only require a discretisation of the boundary instead of the whole external domain (Klees, 1997), and are therefore more useful in physical geodesy. Examples of their application to ellipsoidal physical geodesy are given by Ritter (1998) and Lehmann and Klees (1999).

1.3.2 Solutions using spherical harmonic expansions

More practical solutions to ellipsoidal BVPs can be found using harmonic base functions. They mainly thank their practicality to the wide consensus that harmonic expansions provide an efficient representation of the external gravity field. The most simple case is a spherical harmonic expansion of the Earth's gravity field. The computation of such a spherical harmonic model from terrestrial gravity data requires the solution to an ellipsoidal BVP.

Spherical harmonic coefficients can be computed numerically using a least-squares estimation or least-squares collocation (LSC) (e.g., Sansò and Tscherning, 2003). This numerical procedure is often used to compute coefficients of low degree and order, but is less suitable for higher degrees. The reason for this is, first of all, the rapidly increasing computational load of the method for increasing degrees, but more important is the fact that the numerical methods are susceptible to aliasing effects in degrees close to the Nyquist frequency of the spherical harmonic expansion (Gleason, 1989a; Jekeli, 1996). Analytical solutions can overcome this problem.

Cruz (1985, 1986) and Pavlis (1988) were among the first to investigate the analytic computation of spherical harmonic coefficients from gravity data on the ellipsoid, and this method was refined by Petrovskaya et al. (2001). In their approach, the data on the ellipsoid is upward-continued to a sphere and corrections to the boundary condition are applied to transform the problem into a spherical BVP. An alternative procedure, which is in fact based on a boundary integral equation, is suggested in Sjöberg (2003c).

Once a spherical harmonic expansion of the Earth's gravity field is available, it can be utilised to obtain ellipsoidal corrections to various gravity field functionals, in particular to geoid heights, by implementing the spherical harmonic coefficients (SHCs) into the rigorous ellipsoidal and spherically approximated fundamental equation of physical geodesy. This approach was first outlined by Hotine (1969) and further developed by Lelgemann (1970), Pellinen (1982) and Moritz (1989).

Formulas specifically tailored to geoid computation based on the spherical harmonic approach are derived in Rapp (1981), Sjöberg (2003a,b, 2004b), Heck and Seitz (2003) and Ellmann and Sjöberg (2004). Hipkin (2004) applies the spherical harmonic approach to the computation of gravity anomalies on the ellipsoid from geopotential coefficients, and argues its use in local geoid computation in combination with a satellite gravity field model.

It should, however, be noted that all of the above solutions, like the boundary integral solutions, are restrained to an order of relative accuracy equal to the square of the first numerical eccentricity of the ellipsoid. This degrades the accuracy of ellipsoidal corrections to geoid heights, but more seriously compromises the application of the spherical harmonic approach to the computation of geopotential coefficients, as was pointed out by Gleason (1988, 1989b). This is the major weakness of the solutions using spherical harmonic expansions. However, it will be shown in the remainder of this thesis that this weakness can now be taken away, allowing to profit from the simplicity and practicality of the approach.

1.3.3 Solutions using ellipsoidal harmonic expansions

Basic ellipsoidal BVPs, such as the Dirichlet and Neumann problems (e.g., Kellogg, 1929; Heiskanen and Moritz, 1967), can be solved if the function under consideration is expanded into a series of ellipsoidal harmonic base functions (e.g., Yu et al., 2003). These solutions can also be applied to several geodetic BVPs, such as the fixed- and free-geodetic BVPs (e.g., Heck, 1991).

Ellipsoidal harmonic expansions have long been used as an aid in the construction of a reference gravity field (e.g., Heiskanen and Moritz, 1967, p. 44). Thông and Grafarend (1989), Sona (1995) and Garnier and Barriot (2001) argue their use as a representation of the Earth's gravity field, and this has led to investigations towards solutions of the ellipsoidal BVPs by means of ellipsoidal harmonic expansions. Yu and Cao (1996) and Jinghai and Xiaoping (1997) provide solutions to the free gravimetric and the altimetry-gravimetry BVPs, respectively, as a combination of ellipsoidal harmonic coefficients.

Thông (1996), Martinec and Grafarend (1997a,b) and Martinec (1998) derive explicit expressions of the integration kernel in this problem, as a generalisation of the Stokes kernel. The practical implementation of the ellipsoidal Stokes kernel is discussed by Ardestani and Martinec (2000), who also study the numerical contribution of the ellipsoidal correction in a remove-compute-restore (RCR) geoid computation (Ardestani and Martinec, 2001, 2003a,b), and by Safari et al. (2005) who apply the ellipsoidal harmonic approach to a local geoid computation in Iran.

The ellipsoidal harmonics approach can, paradoxically, also be used to compute spherical harmonic coefficients of the Earth's external gravity field. This follows from the fact that an exact transformation between spherical and ellipsoidal harmonic coefficients is possible (e.g., Hotine, 1969; Buchdahl et al., 1977; Jekeli, 1988; Dechambre and Scheeres, 2002). It provides an alternative way to construct a spherical harmonic geopotential model from terrestrial gravity data, which is elaborated upon in, e.g., Rapp and Pavlis (1990). This approach is also approximative, because the fundamental equation of physical geodesy that forms the boundary condition in this problem is approximated to the level of the square of the first numerical eccentricity of the ellipsoid.

The major drawback of the ellipsoidal harmonics approach is the complicated nature of the underlying associated Legendre functions of the second kind (cf. Hobson, 1931; Heiskanen and Moritz, 1967). The numerical stability of functions of the second kind of high degree and order is a point of concern (e.g., Sona, 1995).

1.4 Research objectives

Even though several theoretical contributions to ellipsoidal physical geodesy have been made (see Section 1.3), the practical application of ellipsoidal theories remains limited. This can presumably, at least partly, be prescribed to the complicated nature of the formulas involved. Therefore, the general aim of this research is to *find practical and highly accurate methodologies to model the figure of the Earth and its external gravity field from geodetic observations by solving ellipsoidal BVPs.*

The solution strategy followed to attain this research objective is based on the literature review in Section 1.2. It is aimed for to find *analytic solutions* using the *spherical harmonic approach* from *terrestrial gravimetric data*, profiting from the practicality of spherical harmonics, yet improving on the poor accuracy of existing solutions that hampers their use in some applications, such as the construction of geopotential models. A more extensive rationale behind this solution strategy is explained next.

1.4.1 Analytic solutions

The emphasis in this thesis will be on the derivation of analytic solutions to ellipsoidal BVPs, even though deterministic solutions based on fairly straightforward general principles also exist (e.g., Gruber, 2001). However, the computational requirements of these solutions are high, although this can, for example, be reduced by the use of semi-analytic methods (e.g., Sneeuw, 2000). Although the optimal approach may be different for various specific applications, full analytic solutions are often more efficient and are therefore widely used. An example is the widespread use of (adaptations of) Stokes's integral formula in local and regional gravimetric geoid determination.

Secondly, in many cases, numerical methodologies to gravity field modelling suffer from inaccuracies that can be circumvented by analytic methods. For example, the computation of spherical harmonic coefficients of the Earth's external gravity field can be computed from terrestrial gravity data using least-squares estimation. This method suffers from inaccuracies due to aliasing effects in the coefficients of high degree and order (close to the Nyquist frequency) (Gleason, 1989a), which can be overcome by the use of analytic methods (Lemoine et al., 1998, p. 8-44).

1.4.2 Spherical harmonic approach

It can be seen from the literature review in Section 1.2 that many contributions to ellipsoidal physical geodesy are based on the framework of an ellipsoidal harmonic expansion. This is not surprising, since ellipsoidal harmonics form a natural set of

base functions for data on an ellipsoid. In practice, however, the use of ellipsoidal harmonics has remained very minor. The few ellipsoidal harmonic representations of the Earth's gravity field that have been constructed, were all directly based on a transformation from existing spherical harmonic models (e.g., Sansò and Sona, 2001; Ardalan and Grafarend, 2001b,c), which clearly underlines the dominant position that spherical harmonics take in the field of physical geodesy.

A representation of the Earth's gravity field in spherical harmonic base functions is more practical and still the standard today. The vast majority of recent Earth gravity field models, such as for example EGM96 (Lemoine et al., 1998), the GPM98 models (Wenzel, 1998), GRIM5-S1 (Biancale et al., 2000), the GGM models (Tapley et al., 2003, 2005), and the EIGEN models (Reigber et al., 2002a, 2005) are represented as spherical harmonic series expansions.

The seemingly lack of interest in the use of ellipsoidal harmonics can at least partly be explained by the difficulties that occur in their computation (e.g., Sona, 1995). The computation of spherical harmonic functions on the other hand is fairly straightforward (e.g., Rizos, 1979; Koop and Stelpstra, 1989; Masters and Richards-Dinger, 1998), and recently developed algorithms (Holmes and Featherstone, 2002) provide a fast and stable way to compute spherical harmonic functions up to very high degree and order (at least until degree 2600). Therefore, in this work the use of spherical harmonics is taken as a point of departure in the development of theories to deal with ellipsoidal effects in the computation of the Earth's gravity field.

Solutions to specific ellipsoidal BVPs completely derived in the framework of spherical harmonics have been published before (e.g., Cruz, 1986; Petrovskaya et al., 2001; Heck and Seitz, 2003). In these methods, ellipsoidal computation of geoid heights and/or spherical harmonic geopotential coefficients are presented as weighted summations over spherically approximated coefficients, and are simpler and more practically applicable than methods based on ellipsoidal harmonics expansions. Their accuracy, however, can generally not match those of ellipsoidal harmonic methods, although this is less critical for the computation of geoid heights than it is for recovery of the full external gravity

field. The ideal is to obtain a practical solution in the spherical harmonic framework that can match or even supersede the accuracy of the methods based on ellipsoidal harmonics.

1.4.3 Terrestrial gravimetric data

A lot of interest in recent years has been focussed on satellite gravity missions such as the Challenging Mini-Satellite Payload (CHAMP) mission (Reigber et al., 2002b), the Gravity Recovery and Climate Experiment (GRACE) (Tapley et al., 2004) and the Gravity Field and Steady-State Ocean Circulation Explorer (GOCE) (Drinkwater et al., 2003). Observations from these gravity missions will greatly improve the accurate modelling of the Earth's gravity field, but there are two main reasons why the focus in this thesis is on terrestrial gravimetric data.

First of all, the orbits of satellites in the aforementioned gravity missions are, without exception, chosen to be nearly spherical. The eccentricities of the orbits are in the order of 0.004 for CHAMP, 0.0006 for GRACE and 0.002 for GOCE, which is in all three cases smaller than the eccentricity of the Earth's ellipsoid (0.007). Moreover, even though the orbits are slightly elliptical, the satellite tracks do not form the surface of an ellipsoid. This is due to the fact that the Earth's geocentre is at one of the focal points of the elliptical orbit, rather than in the centre of the orbit, and also due to the fact that the orbits are not polar. Therefore, the issues faced due to orbital ellipticity are of a very different nature compared to the ones faced in gravity field modelling from terrestrial data.

Secondly, satellite gravity missions can only sense the long-wavelength structure of the gravity field. Terrestrial gravity data are still needed to provide the fine structure of the field. In fact, the advance that satellite gravity missions provide calls for more accurate modelling of terrestrial data in order to be able to fully profit from the highly accurate modelling of the long-wavelength field that satellite missions offer (cf. Kern et al., 2003). Since numerical methods perform best for the long wavelengths and analytical methods are especially useful for the shorter wavelengths, analytic methods

are of the highest importance for the modelling from terrestrial gravity data.

The merger of terrestrial and satellite data is an important issue nonetheless, and should be regarded a task of ellipsoidal physical geodesy as well. It is, however, outside the scope of this thesis.

1.5 Thesis structure

In Chapter 2, the physical and mathematical theories that form the foundation of ellipsoidal physical geodesy are presented, where specific attention is given to the formulas relevant to the methodologies derived in the remainder of the thesis. It contains the basics of the geometry of the ellipsoid, as well as the physical foundations of gravity, the derivation of the gravimetric boundary conditions, and the representation of the gravity field in series of harmonic base functions. Some new formulas with respect to the geometry of the ellipsoid are also provided (in Section 2.1.2), as well as new formulas to assess secondary effects in geodetic boundary conditions in Section 2.3.3. Other well-known formulas are presented with an eye on their use in ellipsoidal physical geodesy in general and their use in this work in particular.

Chapter 3 describes a new solution to the most simple of the ellipsoidal BVPs, the ellipsoidal Dirichlet problem. The most interesting aspect of this method is that it is theoretically exact and completely described in terms of spherical harmonic base functions. It relies on relations among spherical harmonics (Claessens, 2005) that are derived in Section 3.2. These new relations form the key, not only to the solution of the Dirichlet problem, but also to solutions to other problems presented in Chapters 4 to 6. A numerical closed-loop simulation in Section 3.4 confirms the validity of the new solution to the ellipsoidal Dirichlet problem.

In Chapter 4, the solution to the ellipsoidal Dirichlet problem derived in Chapter 3 is extended to also provide solutions to the ellipsoidal Neumann and second-order BVPs, and in particular the fixed- and scalar-free-geodetic BVPs for computing the

Earth's gravity field from terrestrial gravity and satellite positioning measurements. It is subsequently shown that the gravity field can also be resolved if second normal derivatives or gravity anomalies are provided on the surface of the ellipsoid. This culminates in the construction of an ellipsoidal generalisation of the so-called Meissl scheme (Rummel and Van Gelderen, 1995) that connects various gravity field quantities in an ellipsoidal framework.

Specific applications of the newly derived theories are provided in Chapters 5 and 6. Chapter 5 discusses the analytic computation of spherical harmonic geopotential coefficients from terrestrial gravity data, a critical step in the construction of a high-resolution global gravity field model. It is shown that two existing methodologies of limited accuracy can be extended to much higher orders of accuracy by a rigorous derivation based on the new relations among spherical harmonics in Section 3.2. A third and novel methodology, based on the solution to the scalar-free-geodetic BVP derived in Chapter 4, is also introduced. The three newly derived methods are compared numerically, also to various existing methods, in Section 5.5.

A second application of solutions to ellipsoidal geodetic BVPs is local or regional gravimetric geoid computation, which is discussed in Chapter 6. It is shown that the spectral relation between the disturbing potential and the gravity anomaly derived in Chapter 4 can be used to compute ellipsoidal corrections to Stokes's (1849) integral formula, which is based on spherical and constant radius approximations. The magnitude of the ellipsoidal corrections is determined numerically, also in view of the widely used remove-compute-restore (RCR) technique for geoid modelling, and it is shown that the choice of the radius of the reference sphere has a substantial impact on the magnitude and spectral sensitivity of the ellipsoidal corrections.

Chapter 7 provides a summary of the overall results and an itemised overview of the main conclusions. In addition, an outlook on possible future research is provided.

2. BACKGROUND TO ELLIPSOIDAL PHYSICAL GEODESY

The field of ellipsoidal physical geodesy, which this thesis investigates, combines the classical fields of physical and geometrical geodesy, while realising that the figure of the Earth is more closely approximated by an oblate ellipsoid of revolution (in short denoted ‘ellipsoid’ throughout this thesis) than by a sphere. The concepts of physical geodesy form its primary foundation, but the geometrical properties of the ellipsoid also play an important role. The geometry of the ellipsoid is therefore discussed in Section 2.1, and some new formulas that assist in its application to ellipsoidal physical geodesy are derived. Section 2.2 provides background to the mathematical modelling of the gravity field of the Earth, and in Sections 2.3 and 2.4, the theoretical background to its computation and the representation of the Earth’s gravity field in space by harmonic base functions are outlined, respectively.

2.1 The geometry of the ellipsoid

The geometry of the ellipsoid and the coordinate systems with respect to the ellipsoid are highly important in ellipsoidal physical geodesy. The formulas for various ellipsoidal parameters take on different forms in different coordinate systems.

2.1.1 Coordinate systems

A point in space can naturally be denoted by Cartesian coordinates (x,y,z) , but there are several other coordinate systems that are more useful in many applications. In ellipsoidal physical geodesy, three different coordinate systems are used:

1. Spherical polar coordinates (r,θ,λ) ,
2. Geodetic coordinates (h,ϑ,λ) ,
3. Ellipsoidal coordinates (u,β,λ) .

The longitude λ is the same for all systems and is, by convention, measured from the Greenwich meridian. The latitudinal coordinates are shown in Figure 2.1 and the geocentric radius r and height h of an arbitrary point P are shown in Figure 2.2. A point on the surface of the ellipsoid is completely defined by the longitude λ and the latitudinal coordinate θ , ϑ or β . Thus, only two coordinates are needed to describe a point on the ellipsoidal surface, which immediately shows the main advantage of these coordinate systems over 3D Cartesian coordinates.

The relations among the three different latitudinal coordinates can all be found from the defining equation of the ellipse, which is

$$\frac{w^2}{a^2} + \frac{z^2}{b^2} = 1 \quad (2.1)$$

where a and b are the semi-major and semi-minor axes of the ellipse respectively, and w and z are the two-dimensional Cartesian coordinates aligned with these axes, where $w^2 = x^2 + y^2$. Thus, the only difference with a circle is that the z -coordinate is scaled by a factor $\frac{b}{a}$. From Figure 2.1, it can then be seen that the relation between the geocentric co-latitude θ and the reduced co-latitude β is (e.g., Bomford, 1971, p. 564)

$$\tan \theta = \frac{a}{b} \tan \beta \quad (2.2)$$

A relation between the geodetic co-latitude ϑ and the reduced co-latitude β can be found from the parametric form of the definition of the ellipse

$$w = a \sin \beta \quad z = b \cos \beta \quad (2.3)$$

From this definition, an equation for the slope of the tangent of the ellipsoid can be found

$$\frac{dz}{dw} = \frac{dz/d\beta}{dw/d\beta} = -\frac{b}{a} \tan \beta \quad (2.4)$$

while, by definition, the slope of the tangent is also related to the geodetic co-latitude ϑ , as can be seen in Figure 2.1

$$\frac{dz}{dw} = -\tan \vartheta \quad (2.5)$$

Combining Equations (2.2), (2.4) and (2.5), the three different definitions of latitude are related by (e.g., Bomford, 1971, p. 564)

$$\tan \theta = \frac{a}{b} \tan \beta = \frac{a^2}{b^2} \tan \vartheta \quad (2.6)$$

Equation (2.6) plays a crucial role in many applications involving the ellipsoid.

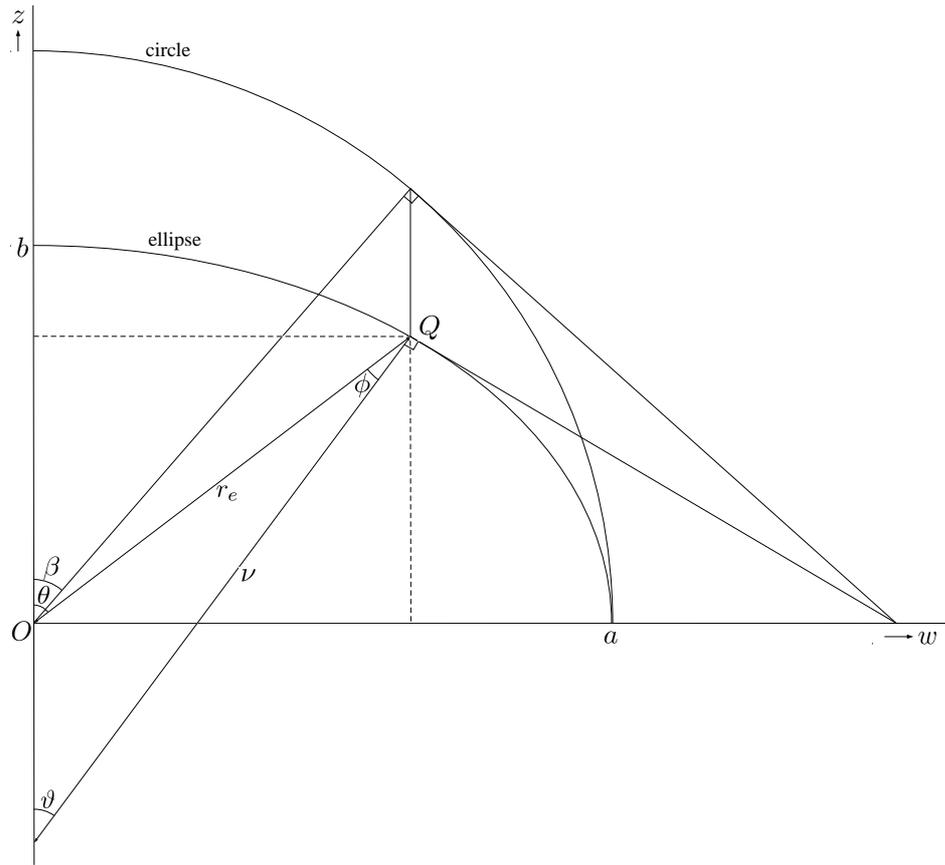


Figure 2.1: Geocentric (θ), geodetic (ϑ) and reduced (β) co-latitude

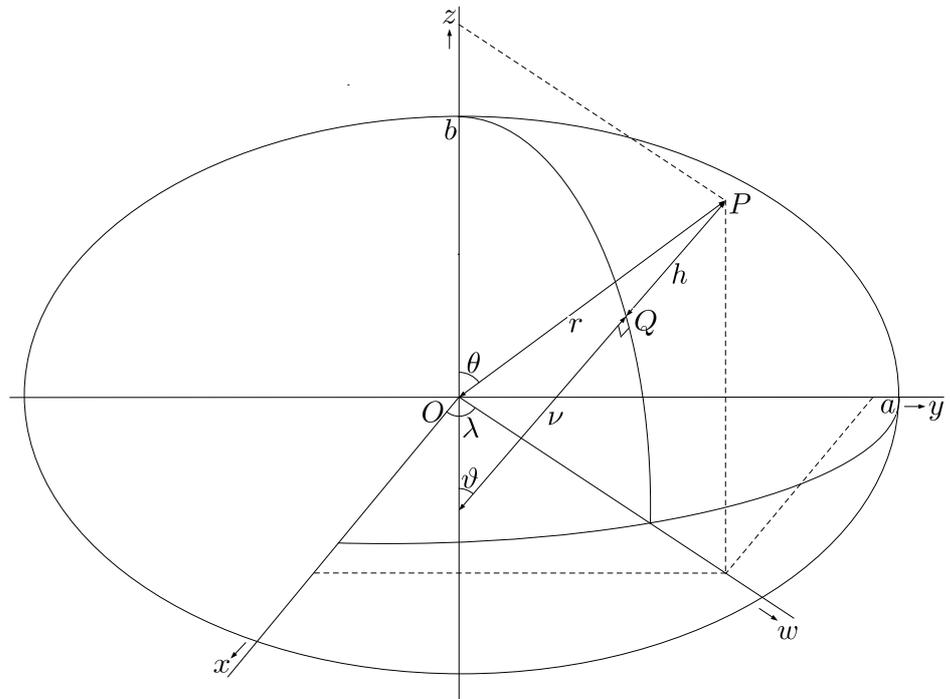


Figure 2.2: The spherical polar (r, θ, λ) and geodetic (h, ϑ, λ) coordinate systems

2.1.2 Parameters with respect to the ellipsoid

The shape of the ellipsoid is usually characterised by one of the following parameters (e.g., Bomford, 1971):

$$\text{flattening} \quad f = \frac{a - b}{a} \quad (2.7)$$

$$\text{first numerical eccentricity} \quad e^2 = \frac{a^2 - b^2}{a^2} \quad (2.8)$$

$$\text{second numerical eccentricity} \quad e'^2 = \frac{a^2 - b^2}{b^2} \quad (2.9)$$

$$\text{linear eccentricity} \quad E^2 = a^2 - b^2 \quad (2.10)$$

In addition, for the purpose of simplifying the equations in the following chapters, a new type of eccentricity is defined here

$$\text{quartic eccentricity} \quad \epsilon^4 = \frac{a^4 - b^4}{a^4} \quad (2.11)$$

The quartic eccentricity can easily directly be related to the first numerical eccentricity by

$$\epsilon^4 = e^2(2 - e^2) \quad (2.12)$$

All the above parameters are solely dependent on the shape of the ellipsoid, and are thus independent of position. Several other parameters with respect to the ellipsoid, which are not only dependent on the shape, but also on the latitude of a point on the surface of the ellipsoid, will be used extensively in the investigation of the geodetic BVPs. The equations for these parameters in terms of spherical polar, geodetic and ellipsoidal coordinates are therefore given below. Recall that all parameters given here are independent of longitude λ , due to the rotational symmetry of the ellipsoid.

The ellipsoidal radius

The ellipsoidal radius r_e is the distance from the geometric centre of the ellipsoid (not the foci) to a point on the surface of the ellipsoid. In Cartesian coordinates, its expression simply follows from Pythagoras's theorem

$$r_e = \sqrt{w^2 + z^2} \quad (2.13)$$

	spherical polar	geodetic	ellipsoidal
r_e	$a\sqrt{\frac{1-e^2}{1-e^2\sin^2\theta}}$	$a\sqrt{\frac{1-\epsilon^4\cos^2\vartheta}{1-e^2\cos^2\vartheta}}$	$a\sqrt{1-e^2\cos^2\beta}$
ρ	$\frac{a(1-\epsilon^4\sin^2\theta)^{\frac{3}{2}}}{\sqrt{1-e^2}(1-e^2\sin^2\theta)^{\frac{3}{2}}}$	$\frac{a(1-e^2)}{(1-e^2\cos^2\vartheta)^{\frac{3}{2}}}$	$\frac{a(1-e^2\sin^2\beta)^{\frac{3}{2}}}{\sqrt{1-e^2}}$
ν	$a\sqrt{\frac{1-\epsilon^4\sin^2\theta}{(1-e^2)(1-e^2\sin^2\theta)}}$	$\frac{a}{\sqrt{1-e^2\cos^2\vartheta}}$	$a\sqrt{\frac{1-e^2\sin^2\beta}{1-e^2}}$

Table 2.1: Expressions for the ellipsoidal radius (r_e), meridian radius of curvature (ρ) and prime vertical radius of curvature (ν) as a function of geocentric (θ), geodetic (ϑ) and reduced (β) co-latitude

Insertion of the parametric definition of the ellipse (Equation 2.3) into Equation (2.13) gives an expression in ellipsoidal coordinates and by subsequently using Equation (2.6), expressions in spherical polar and geodetic coordinates can also be found. All these expressions are given in Table 2.1.

The principal radii of curvature

The principal radii of curvature of a point on the surface of the ellipsoid are defined as the radius of curvature in the north-south direction (ρ), and the radius of curvature in the east-west direction (ν). An expression for ρ , also called the meridian radius of curvature, follows from the definition of the radius of curvature, which is the ratio of the infinitesimal arc length to the corresponding change in the so-called tangential angle. In this case, the tangential angle is the geodetic co-latitude, thus (e.g., Schwarz, 1996; Dragomir et al., 1982)

$$\rho = \frac{ds}{d\vartheta} = \sqrt{\left[\left(\frac{dw}{d\beta}\right)^2 + \left(\frac{dz}{d\beta}\right)^2\right]} \frac{d\beta}{d\vartheta} \quad (2.14)$$

Applying Equations (2.3) and (2.6), expressions for ρ can be found in spherical polar, geodetic and ellipsoidal coordinates, and these can be found in Table 2.1.

Expressions for ν , also called the prime vertical radius of curvature, can be found more

easily. Due to the rotational symmetry of the ellipsoid, the prime vertical radius is equal to the distance from the surface of the ellipsoid to the z -axis, along the ellipsoidal normal. An equation for ν can easily be derived from Figure 2.1

$$\nu = \frac{w}{\sin \vartheta} \quad (2.15)$$

Applying Equations (2.3) and (2.6), expressions for ν can be found in spherical polar, geodetic and ellipsoidal coordinates and these are shown in Table 2.1.

The ellipsoidal deflection angle

Another important parameter in the derivations in the following chapters is the difference between the geocentric co-latitude θ and the geodetic co-latitude ϑ , which is here named the ellipsoidal deflection angle ϕ , since it gives the deflection of the geocentric radial direction with respect to the ellipsoidal normal (cf. Figure 2.1)

$$\phi = \theta - \vartheta \quad (2.16)$$

Using Equations (2.6), (2.16) and various trigonometric identities, expressions for ϕ can be found that only depend on any of the geocentric (θ), geodetic (ϑ) or reduced (β) co-latitude. Expressions for the sine, cosine and tangent of the deflection angle are shown in Table 2.2.

	spherical polar	geodetic	ellipsoidal
$\sin \phi$	$\frac{e^2 \sin \theta \cos \theta}{\sqrt{1 - e^4 \sin^2 \theta}}$	$\frac{e^2 \sin \vartheta \cos \vartheta}{\sqrt{1 - e^4 \cos^2 \vartheta}}$	$\frac{e^2 \sin \beta \cos \beta}{\sqrt{1 - e^2 + e^4 \sin^2 \beta \cos^2 \beta}}$
$\cos \phi$	$\frac{1 - e^2 \sin^2 \theta}{\sqrt{1 - e^4 \sin^2 \theta}}$	$\frac{1 - e^2 \cos^2 \vartheta}{\sqrt{1 - e^4 \cos^2 \vartheta}}$	$\frac{\sqrt{1 - e^2}}{\sqrt{1 - e^2 + e^4 \sin^2 \beta \cos^2 \beta}}$
$\tan \phi$	$\frac{e^2 \sin \theta \cos \theta}{1 - e^2 \sin^2 \theta}$	$\frac{e^2 \sin \vartheta \cos \vartheta}{1 - e^2 \cos^2 \vartheta}$	$\frac{e^2 \sin \beta \cos \beta}{\sqrt{1 - e^2}}$

Table 2.2: Expressions for the sine, cosine and tangent of the ellipsoidal deflection angle (ϕ) as a function of geocentric (θ), geodetic (ϑ) and reduced (β) co-latitude

2.1.3 Coordinate transformations

Spherical polar, geodetic and ellipsoidal coordinates of an arbitrary point P in space can easily, though not always uniquely, be related to Cartesian coordinates. Formulas for the computation of Cartesian coordinates can be found in Table 2.3. The inverse relations are shown in Table 2.4, but it should be noted that the formulas for the geodetic co-latitude (ϑ), the ellipsoidal height (h), and the ellipsoidal parameter (u) only hold for points on the ellipsoid. For any point in space, the expression for u is

$$u = \sqrt{\frac{1}{2}(r^2 - E^2) + \frac{1}{2}\sqrt{(r^2 - E^2)^2 + 4E^2z^2}} \quad (2.17)$$

Iterative and approximate formulas for the computation of the geodetic co-latitude (ϑ), and thereby also the ellipsoidal height (h), from Cartesian coordinates are given by, e.g., Bowring (1976). A closed expression for points anywhere in space was long

	spherical polar	geodetic	ellipsoidal
x	$r \sin \theta \cos \lambda$	$(\nu + h) \sin \vartheta \cos \lambda$	$\sqrt{u^2 + E^2} \sin \beta \cos \lambda$
y	$r \sin \theta \sin \lambda$	$(\nu + h) \sin \vartheta \sin \lambda$	$\sqrt{u^2 + E^2} \sin \beta \sin \lambda$
z	$r \cos \theta$	$(\nu(1 - e^2) + h) \cos \vartheta$	$u \cos \beta$

Table 2.3: Coordinate transformation from spherical polar (r, θ, λ) , geodetic (h, ϑ, λ) and ellipsoidal (u, β, λ) coordinates to Cartesian (x, y, z) coordinates

spherical polar	geodetic	ellipsoidal
$r = \sqrt{x^2 + y^2 + z^2}$	$h = 0$	$u = b$
$\theta = \arccos \frac{z}{r}$	$\vartheta = \arccos \sqrt{\frac{r^2 - a^2}{r^2 e^2 - a^2 \varepsilon^4}}$	$\beta = \arccos \frac{z}{u}$
$\lambda = \arctan \frac{y}{x}$	$\lambda = \arctan \frac{y}{x}$	$\lambda = \arctan \frac{y}{x}$

Table 2.4: Coordinate transformation from Cartesian (x, y, z) coordinates to spherical polar (r, θ, λ) , geodetic (h, ϑ, λ) and ellipsoidal (u, β, λ) coordinates on the surface of the ellipsoid

thought impossible (e.g., Heiskanen and Moritz, 1967, p. 183), but exact formulas based on the solution of quartic equations exist, as was first discovered by Paul (1973). More efficient exact algorithms have been derived since and can for example be found in Borkowski (1987), Jones (2002), Vermeille (2002) or Zhang et al. (2005).

2.2 The Earth's gravity field

Like all branches of physics and applied mathematics, physical geodesy consists of mathematical models of reality that are based on empirical evidence. One of the most famous of these models is Newton's universal law of gravitation, which forms the foundation of physical geodesy. This law of gravitation was revised by Einstein's theory of general relativity, but Newton's simpler formulation is generally considered of sufficient accuracy for modelling the relatively weak gravitational field of the Earth. Based on the universal law of gravitation, mathematical properties of the Earth's gravitational field can be extracted.

2.2.1 Newton's universal law of gravitation

In order to study the Earth's gravity field, it is important to know in which way the gravitational force between objects is related to other quantities, such as the masses of the objects and the distance between the objects. This relationship is given by Newton's universal law of gravitation, which was originally based on observations of celestial bodies (Newton, 1729) and, somewhat reformulated, states that:

Every particle of matter in the universe attracts every other particle, with a force whose direction is that of the line joining the two, and whose magnitude is directly as the product of their masses, and inversely as the square of their distance from each other.

Thus, in mathematical terms, the force \vec{F} exerted on a point mass at P with mass M_P

with respect to another point mass at Q with mass M_Q equals

$$\vec{F}_P = -\frac{GM_P M_Q}{r_{PQ}^3} \vec{r}_{PQ} \quad (2.18)$$

where \vec{r}_{PQ} is the vector between points P and Q , r_{PQ} is its length, and G is the universal gravitational constant. The value of G has been determined empirically. The 2002 CODATA (Committee on Data for Science and Technology) recommended value of the gravitational constant is (Mohr and Taylor, 2005)

$$G = (6.6742 \pm 0.001) \times 10^{-11} \text{m}^3 \text{kg}^{-1} \text{s}^{-2} \quad (2.19)$$

The gravitational acceleration \vec{a} caused by this force field can be easily found by applying Newton's second law of motion ($\vec{F} = M\vec{a}$)

$$\vec{a}_P = -\frac{GM_Q}{r_{PQ}^3} \vec{r}_{PQ} \quad (2.20)$$

These universal formulas form the basis for every study of the Earth's gravitational field, and are therefore of prime importance in ellipsoidal physical geodesy.

2.2.2 Gravity

Gravitational acceleration induced by the masses of the Earth forms the major part of the acceleration observed by gravimetric instruments, but there are other effects that play a role. The most important of these is the centrifugal acceleration $\vec{\chi}$ due to the rotation of the Earth, which in the spherical polar coordinate system reads

$$\vec{\chi} = \omega^2 r \begin{pmatrix} \sin \theta \cos \lambda \\ \sin \theta \sin \lambda \\ 0 \end{pmatrix} \quad (2.21)$$

where ω is the angular velocity of the Earth's rotation. The numerical value of ω is determined by space geodetic techniques and its numerical value in the definition of the Geodetic Reference System 1980 (GRS80) (Moritz, 1980) can be found in Table 2.5. The velocity of the Earth's rotation changes slightly over time (e.g., Lambeck, 1980; Gross, 2001), but since the length-of-day variation is only in the order of several milliseconds, its effect is at least seven orders of magnitude smaller than the centrifugal

acceleration itself. The Earth's rotation can therefore safely be treated as a constant for gravity field modelling.

There are a number of other time-variable gravity effects, such as tidal gravitational acceleration primarily caused by the Sun and the Moon, and inertial acceleration of rotation caused by precession, nutation and polar motion. These effects will not be considered in this work, as they are at least three orders of magnitude smaller than the ellipsoidal effects (e.g., Groten, 1979).

Gravity is defined as the vector sum of the gravitational and the centrifugal acceleration

$$\vec{g} = \vec{a} + \vec{\chi} \quad (2.22)$$

and will in this thesis be considered time-invariant, although seasonal time-variations in the Earth's gravity field have recently been revealed by satellite observations from GRACE (Wahr et al., 2004) and CHAMP (Sneeuw et al., 2005).

2.2.3 The external gravity field: A harmonic vector field

The Earth's gravitational field is a vector field, because every point in space has an absolute magnitude and a direction. Vector fields are often categorised by making use of two important differential operators, the *divergence* and the *curl*. The divergence of vector \vec{v} at a point is the dot product of the gradient operator and that vector ($\nabla \cdot \vec{v}$), whereas the curl of vector \vec{v} is defined as the cross-product of the gradient operator and the vector ($\nabla \times \vec{v}$). Categorisation of vector fields is obtained by asserting whether the divergence and curl of the vectors are zero or non-zero (see Figure 2.3).

If the curl of a vector field is equal to zero at all points, the field is called irrotational or conservative. In this case, the vector field is called a Poisson field. It can be shown using Newton's universal law of gravitation that the Earth's gravitational field, if considered stationary, is a Poisson field

$$\nabla \times \vec{a} = 0 \quad (2.23)$$

For a formal proof of Equation (2.23), see, e.g., Sigl (1985). Poisson fields have the

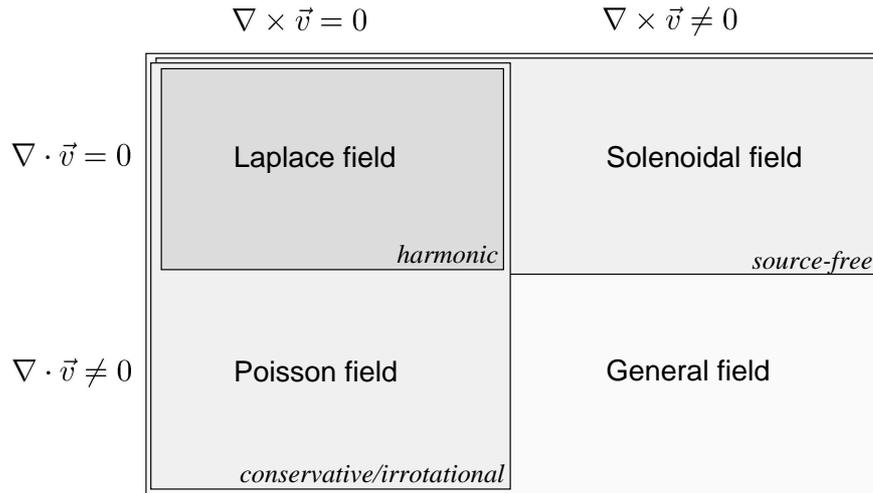


Figure 2.3: Categorisation of vector fields by investigation of their curl and divergence

important property that they can always be expressed as a gradient field. This means that at every point, the vector \vec{a} can be written as the gradient of some scalar function V , which here is called the gravitational potential

$$\vec{a} = \nabla V \quad (2.24)$$

when \vec{a} is considered continuous and V continuously differentiable, which holds for the external gravitational field (e.g., Kellogg, 1929). This is extremely useful, since it reduces the three-dimensional vector field to a one-dimensional scalar field.

If the divergence of a vector field is zero, the field is called solenoidal. A Poisson field is not necessarily solenoidal; the divergence of a gradient field is, in general, a function of position

$$\nabla \cdot \vec{a} = \nabla \cdot \nabla V = \nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} \quad (2.25)$$

where ∇^2 is the Laplacian differential operator. Using Newton's law (Equation 2.18), it can be shown that the divergence at a point in the gravitational field is a function of the mass-density ρ at that same point

$$\nabla \cdot \vec{a} = 4\pi G\rho \quad (2.26)$$

Equation (2.26) is known as Poisson's differential equation. A mathematical proof for this formula is provided in, e.g., Sigl (1985).

Thus, outside the masses of the Earth and its atmosphere, where the density is equal to zero, the gravity field is a solenoidal field (see Figure 2.3). From Equations (2.23) and (2.26), it can be seen that the external gravity field is a Laplace field that can be described by a harmonic function, the scalar gravitational potential V , which obeys Laplace's differential equation

$$\nabla^2 V = 0 \quad (2.27)$$

As discussed in Section 2.2.2, the Earth's gravity field comprises the gravitational field and the acceleration field induced by the Earth's rotation. The potential of the gravity field W can thus be separated in the gravitational potential V , plus a rotational component Ω

$$W = V + \Omega \quad (2.28)$$

where

$$\Omega = \frac{1}{2}\omega^2 r^2 \sin^2 \theta \quad (2.29)$$

It can be seen from Equation (2.29) that the rotational potential Ω does not satisfy Laplace's differential equation. Therefore, the Earth's gravity field is, contrary to the gravitational field, not harmonic outside the surface of the Earth.

However, the gravity field of the Earth is conservative, which allows for the definition of the rotational potential. This means that the Earth's gravity field can be viewed as given by surfaces of equal potential. The geoid, depicted in Figure 2.4, is here defined as the surface with a certain potential W_0 that best fits, in a least-squares sense, global mean sea level. Estimates of W_0 are given by, e.g., Burša (1995) and Grafarend and Ardalan (1997).

2.2.4 The reference gravity field

To simplify the determination of the Earth's gravity field, it is often separated into a reference (or normal) gravity field with potential U and a remaining disturbing (anomalous) field with potential T at any point P

$$W_P = U_P + T_P \quad (2.30)$$

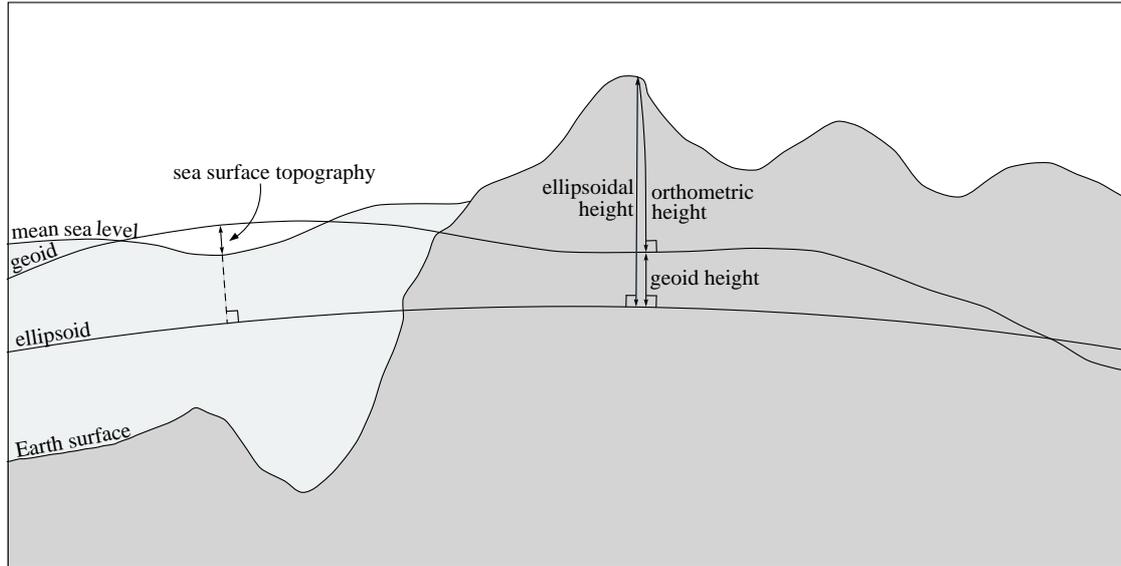


Figure 2.4: Relations between the Earth surface, mean sea level, the ellipsoid and the geoid: the ellipsoidal height is measured from the ellipsoid to the Earth surface, the geoid height from the ellipsoid to the geoid, the sea surface topography from the geoid to mean sea level (all three along the ellipsoidal normal) and the orthometric height from the geoid to the Earth surface along the plumbline of gravity, where all four measures are positive in the external direction

The reference field U should be chosen in such a way that it is close to the true gravity field, but mathematically simply defined. Since the disturbing field will then be small, it can usually be considered linear (or linearised), which simplifies the construction of methodologies to compute the Earth's gravity field.

The most simple form of a reference gravity field is a spherical one, where the Earth is assumed to be a point mass, a concentric shell or a sphere with no lateral density variations. The magnitude of the reference gravity vector $\vec{\gamma}$ induced by such a field at any point P outside the sphere at a distance r_P from its origin can simply be derived from Newton's universal law of gravitation (Equation 2.20)

$$\gamma_P = \frac{\mu}{r_P^2} \quad (2.31)$$

where μ is the terrestrial gravitational constant, which is simply the product of the gravitational constant G and the mass of the reference Earth model. Since this reference gravity field is a harmonic vector field, the reference potential U can also be derived

$$U_P = \frac{\mu}{r_P} \quad (2.32)$$

Due to the radial symmetry of Equation (2.32), the equipotential surfaces of any given potential value U are all spheres. This definition of the reference field excels in simplicity, but because the centrifugal potential Ω is not modelled by the reference potential field, the disturbing potential T is non-harmonic. Secondly, the equipotential surfaces of the Earth's gravity field, in particular the geoid, are much closer described by an ellipsoid of revolution. For these reasons, a rotating ellipsoidal reference gravity field is usually adopted, with potential U resulting from the sum of the gravitational potential due to an assumed total mass of the reference Earth and the centrifugal potential due to a reference rotation velocity.

A reference gravity field that has equal potential at all points on the ellipsoidal surface can be constructed, and the magnitude of the gravity at a point Q' on the surface of the ellipsoid induced by this field can be found by the Somigliana-Pizetti formula (e.g., Heiskanen and Moritz, 1967, Eq. 2-78)

$$\gamma_{Q'} = \frac{a\gamma_a \sin^2 \vartheta + b\gamma_b \cos^2 \vartheta}{\sqrt{a^2 \sin^2 \vartheta + b^2 \cos^2 \vartheta}} \quad (2.33)$$

where γ_a and γ_b are the magnitude of gravity at the equator and at the poles respectively, which can be computed from the reference mass, reference rotation velocity and major and minor axes of the ellipsoid (e.g., Heiskanen and Moritz, 1967; Moritz, 1980). The values of a , γ_a and γ_b for GRS80 (Moritz, 1980) are shown in Table 2.5, and the semi-minor axis b can be derived from the semi-major axis a and the flattening f through Equation (2.7).

Equation (2.33) is given as a function of geodetic co-latitude (ϑ), but can through application of Equation (2.6) also be expressed as a function of geocentric co-latitude (θ) and reduced co-latitude (β)

$$\gamma_{Q'} = \frac{a^3\gamma_b \cos^2 \theta + b^3\gamma_a \sin^2 \theta}{\sqrt{a^2 \cos^2 \theta + b^2 \sin^2 \theta} \sqrt{a^4 \cos^2 \theta + b^4 \sin^2 \theta}} \quad (2.34)$$

$$\gamma_{Q'} = \frac{a\gamma_b \cos^2 \beta + b\gamma_a \sin^2 \beta}{\sqrt{a^2 \cos^2 \beta + b^2 \sin^2 \beta}} \quad (2.35)$$

The reference gravity vector on the surface of the ellipsoid points in the opposite direction to the ellipsoidal external normal. In the remainder of this thesis, the normal gravity and its derivatives will only be required for points on the ellipsoid, and therefore the subscript Q' is, in most instances, dropped from hereon.

description	symbol	value
semi-major axis	a	6378137 m
inverse flattening	f^{-1}	298.257222101
terrestrial gravitational constant	μ	$3.986005 \cdot 10^{14} \text{ m}^3\text{s}^{-2}$
angular rotation velocity	ω	$7.292115 \cdot 10^{-5} \text{ s}^{-1}$
normal potential at ellipsoid	U_0	$6.2636860850 \cdot 10^7 \text{ m}^2\text{s}^{-2}$
gravity at equator	γ_a	$9.7803267715 \text{ ms}^{-2}$
gravity at pole	γ_b	$9.8321863685 \text{ ms}^{-2}$

Table 2.5: Parameters of the Geodetic Reference System 1980 GRS80 (from Moritz, 1980)

The derivative of γ with respect to the external normal, here notated as dh , can be found as a function of the principal radii of curvature and the angular rotation velocity, using a formula generally attributed to Bruns (Heiskanen and Moritz, 1967, p. 70)

$$\frac{\partial\gamma}{\partial h} = -\gamma\left(\frac{1}{\rho} + \frac{1}{\nu}\right) - 2\omega^2 \quad (2.36)$$

The reference gravity and its normal derivative play an important role in the derivation of the fundamental formulas that allow for the determination of the Earth's gravity field from gravity measurements, as will be shown in Section 2.3.2.

2.3 Geodetic boundary-value problems

The determination of the Earth's external gravity field and the geoid from gravity-related measurements requires the solution of a BVP, since measurements are generally provided on a surface, whereas the unknown gravity field is required in the complete three-dimensional external space. The boundary conditions of the most often used geodetic BVPs are derived here and the approximations are discussed.

Since the focus of this thesis is on ellipsoidal effects, the boundary conditions are here derived on the surface of the ellipsoid. Effects due to the downward-continuation of

observed gravity from the surface of the Earth to the geoid or the ellipsoid and the treatment of the effect of the topographic mass are not taken into account. Atmospheric and tidal effects are also assumed to be properly modelled.

The downward-continuation of gravity is a common step in gravity field and geoid modelling and its details are ubiquitous in the literature (e.g., Martinec and Vaníček, 1994; Vaníček et al., 1996, 1999; Nahavandchi and Sjöberg, 1998). It forms an important part of the Stokes (1849) approach and the Hotine (1969) approach to geoid modelling. This introduces an indirect effect to the geoid, and the resulting equipotential surface is called the cogeoid or regularised geoid (e.g., Heiskanen and Moritz, 1967), but here the term geoid will be used throughout. The Stokes approach is distinctly different from the Molodensky (1962) approach, where gravity is not downward-continued. This approach does not yield geoid heights, but telluroid or quasigeoid heights, and will not be considered here.

The topographic effects in ellipsoidal physical geodesy were recently investigated by, e.g., Finn et al. (2002), Ardalan and Safari (2004), Novák and Grafarend (2005) and Sjöberg (2004a), the latter of whom concludes that the error introduced by the commonly applied spherical approximations is practically negligible.

2.3.1 Types of boundary-value problems

Depending on the nature of the available boundary conditions (i.e., depending on the available data), different BVPs can be distinguished, but other conditions also need to be defined. Firstly, since the exterior gravitational field is a Laplace field, the Laplace equation forms the harmonicity conditions in every geodetic BVP

$$\nabla^2 T(\vec{x}) = 0, \quad \vec{x} \in \mathcal{V} \quad (2.37)$$

where \mathcal{V} is the space external to the masses of the Earth. Secondly, note that the disturbing potential T can only be determined up to a constant, because it is the gradient of a vector. This constant is usually set to zero, which yields the regularity

condition occurring in every geodetic BVP

$$T(\vec{x}) \rightarrow 0, \quad |\vec{x}| \rightarrow \infty \quad (2.38)$$

Alongside these two conditions, the boundary condition gives information about the potential and/or a derivative thereof on a boundary \mathcal{S} outside the Earth's masses. Some standard BVPs include (Kellogg, 1929; Sigl, 1985):

$$\begin{aligned} \text{Dirichlet:} \quad T(\vec{x}) &= \mathcal{F}, & \vec{x} \in \mathcal{S} \\ \text{Neumann:} \quad \frac{\partial T(\vec{x})}{\partial n(\vec{x})} &= \mathcal{F}, & \vec{x} \in \mathcal{S} \\ \text{Robin:} \quad \frac{\partial T(\vec{x})}{\partial n(\vec{x})} + \kappa T(\vec{x}) &= \mathcal{F}, & \vec{x} \in \mathcal{S} \end{aligned} \quad (2.39)$$

where \mathcal{F} is a quantity assumed to be known, $n(\vec{x})$ is the normal to the surface \mathcal{S} , and κ is an arbitrary constant.

Green's theorems

Green's integral theorems relate the gravitational potential of a continuous mass-density distribution in a limited area to integrals over the normal derivative of the potential and the potential itself at the boundary. They provide the formulas to compute the Earth's exterior gravitational potential from observational data on the boundary, without the need to know the interior density distribution of the Earth and thereby facilitate the solution of geodetic BVPs. Since the Earth's interior density distribution is not likely to be known in the near future, modelling of the gravity field of the Earth will in practice always require a solution to a BVP.

Green's first integral theorem can be used to determine which boundary conditions a potential function must obey in order to be uniquely determined. It reads (e.g., Kellogg, 1929; Sigl, 1985)

$$\iiint_{\mathcal{V}} (\Phi_a \nabla^2 \Phi_b + \nabla \Phi_a \cdot \nabla \Phi_b) d\mathcal{V} = \iint_{\mathcal{S}} \Phi_a \frac{\partial \Phi_b}{\partial n} d\mathcal{S} \quad (2.40)$$

where Φ_a and Φ_b are non-identical potential functions. With the use of this theorem, it can be proved that the Dirichlet, Neumann and Robin BVPs are all uniquely determined (Sigl, 1985).

Green's second integral theorem follows directly from the first theorem by exchanging the scalar functions Φ_a and Φ_b and taking the difference with Equation (2.40) (e.g., Kellogg, 1929; Sigl, 1985)

$$\iiint_{\mathcal{V}} (\Phi_a \nabla^2 \Phi_b - \Phi_b \nabla^2 \Phi_a) d\mathcal{V} = \iint_S \left(\Phi_a \frac{\partial \Phi_b}{\partial n} - \Phi_b \frac{\partial \Phi_a}{\partial n} \right) d\mathcal{S} \quad (2.41)$$

Equation (2.41) is alternatively known as Green's reciprocity theorem. It is often utilised to solve various BVPs (Kellogg, 1929). It also follows from this theorem that the Earth's gravity field can be computed exterior to the Earth's surface if the potential and its normal derivative are known on the surface, because application of this formula to the unbounded domain external of the Earth's masses provides a function for the disturbing potential in any point P in space from information on the boundary \mathcal{S}_Q

$$T_P = \frac{1}{4\pi} \iint_S \left(T_Q \frac{\partial r_{PQ}^{-1}}{\partial n} - \frac{1}{r_{PQ}} \frac{\partial T_Q}{\partial n} \right) d\mathcal{S}_Q \quad (2.42)$$

where all masses outside the Earth are neglected or mathematically moved inside. This solution requires knowledge of the disturbing potential and its normal derivative on the boundary.

2.3.2 Geodetic boundary conditions

Various geodetic boundary conditions can be defined for different types of observations, or combinations of observations. A boundary condition, in general, relates the observed quantity to the gravity field of the Earth, usually represented by the disturbing potential T . The three main quantities that can be observed (directly or indirectly) are geoid heights, gravity disturbances and gravity anomalies. Their relations to the disturbing potential, which cannot be observed, give rise to the first, second and third geodetic BVPs respectively.

The first geodetic boundary-value problem

Geoid heights can be observed over the oceans from time-averaged satellite radar altimetry, after instrument corrections and corrections for sea surface topography and

tides are applied (e.g., Fu and Cazenave, 2001; Vergos et al., 2005). On land, geoid heights follow from the difference between the geometric ellipsoidal height h and the orthometric height H (see Figure 2.4). However, this commonly used relation is not exact, since the orthometric height is measured along the plumbline of the gravity field rather than along the ellipsoidal normal.

Ellipsoidal heights can be observed with a satellite positioning system and orthometric heights are usually derived from geodetic spirit levelling, although it should be noted that true orthometric heights cannot be known unless the mean value of the Earth's gravity acceleration along the plumbline within the topography is known (e.g., Tenzer et al., 2005). In addition, differences between local vertical datums introduce inconsistencies in orthometric heights.

The relation between geoid heights N and the disturbing potential T is given by Bruns's equation, which forms the boundary condition in the first geodetic BVP and also plays an important role in the derivation of the third geodetic BVP.

The geoid height, or geoid undulation, N is defined as the distance between a point on the geoid P' and a point on the ellipsoid Q' in a straight line along the ellipsoidal normal, as in, e.g., Heiskanen and Moritz (1967, p. 83). Points P' and Q' thus have the same geodetic latitude and longitude. Note, however, that other definitions exist. For example, Martinec (1998) defines the geoid height as the length of the plumbline of the reference gravity field, which is a curved line, and thus situates point Q' at the same plumbline as point P' . Since the curvature of the plumbline of the reference gravity field is smaller than the inverse of the Earth's radius, the difference between both concepts is at the sub-micrometre level.

The simplest form of Bruns's equation occurs when the gravity potential of the geoid W_0 is chosen equal to the known (and constant) reference potential at the ellipsoid $U_{Q'}$, in which case according to Equation (2.30)

$$T_{P'} = U_{Q'} - U_{P'} \quad (2.43)$$

The difference between the reference gravity potential on the ellipsoid and on the geoid

can be written as a Taylor series around point Q'

$$T_{P'} = - \sum_{j=1}^{\infty} \frac{1}{j!} \left. \frac{\partial^j U}{\partial h^j} \right|_{Q'} N^j \quad (2.44)$$

When only the first term in the series on the right-hand side of Equation (2.44) is taken into account, a simple solution emerges, bearing in mind that the derivative of the reference potential with respect to the ellipsoidal normal equals the magnitude of the reference gravity

$$N = \frac{T_{P'}}{\gamma_{Q'}} \quad (2.45)$$

and this is called Bruns's formula (e.g., Heiskanen and Moritz, 1967, p. 85). It forms the boundary condition in the first geodetic BVP, which according to Equation (2.39) is a Dirichlet problem. The error introduced by the truncation of the summation in Equation (2.45) is evaluated in Section 2.3.3.

The second geodetic boundary-value problem

The second geodetic BVP is based on the gravity disturbance δg , which is defined as the difference between the magnitudes of gravity and reference gravity at the same point. For a point Q' on the surface of the ellipsoid, the gravity disturbance thus reads

$$\delta g = g_{Q'} - \gamma_{Q'} \quad (2.46)$$

Computation of gravity at point Q' requires a downward-continuation of gravity observed at the Earth's surface to the surface of the ellipsoid. This is possible when the ellipsoidal height h of the point, at which gravity is observed, is known, which is in practice not the case for most gravity observations. Alternatively, the ellipsoidal height can be approximated if the orthometric height H and an a-priori estimate of the geoid height N are available.

The magnitude of gravity at point Q' is the derivative of the potential in the direction of the gravity vector, which is opposite to the external normal to the equipotential surface of the Earth's gravity field that runs through Q' . The external normal to the equipotential surface is here symbolised by dH , since the orthometric height H is measured along the plumbline. Equation (2.46) can thus also be written as

$$\delta g = - \left. \frac{\partial W}{\partial H} \right|_{Q'} + \left. \frac{\partial U}{\partial h} \right|_{Q'} \quad (2.47)$$

The difference in the direction of the gravity vector and the ellipsoidal normal is called the deflection of the vertical. If this deflection of the vertical is neglected, approximating $\partial W/\partial H$ by $\partial W/\partial h$, the gravity disturbance can be related to the disturbing potential T via a simple relation

$$\delta g = -\left.\frac{\partial T}{\partial h}\right|_{Q'} \quad (2.48)$$

The error introduced by neglecting the deflection of the vertical is assessed in Section 2.3.3. Equation (2.48) forms the boundary condition in the second geodetic BVP, also called the fixed-geodetic BVP, which is a Neumann problem (see Equation 2.39). Since the gravity disturbance is considered at the ellipsoid, the second geodetic BVP with the boundary condition in Equation (2.48) is an ellipsoidal BVP. The computation of geoid heights from gravity disturbances defined by Equation (2.48) is called the Hotine (1969) approach to geoid modelling.

The third geodetic boundary-value problem

When the ellipsoidal height is not available, as is the case for most gravity observations, the gravity at point Q' that appears in Equation (2.46) cannot be computed, but only estimated if an a-priori value of the geoid height is known. Alternatively, gravity on the ellipsoid at point Q' can be replaced by gravity on the geoid at point P' . This leads to the definition of the gravity anomaly Δg

$$\Delta g = g_{P'} - \gamma_{Q'} \quad (2.49)$$

The difference between the gravity anomaly and the gravity disturbance can be evaluated using a Taylor series expansion with respect to the geoid height N

$$\Delta g - \delta g = g_{P'} - g_{Q'} \approx \gamma_{P'} - \gamma_{Q'} = \sum_{j=1}^{\infty} \frac{1}{j!} \left.\frac{\partial^j \gamma}{\partial h^j}\right|_{Q'} N^j \quad (2.50)$$

The error due to the approximation in Equation (2.50) does not exceed the error that occurs when the summation in Equation (2.50) is truncated after the first term (Moritz, 1989), as is common practice. This approximation would not have been required if the gravity disturbance would have been defined at point P' on the geoid, but here the definition on the ellipsoid is retained to acquire a boundary condition on the ellipsoid.

The magnitude of the error introduced by the truncation of Equation (2.50) after the first term is discussed in Section 2.3.3. The gravity anomaly then becomes

$$\Delta g = \delta g + \frac{\partial \gamma}{\partial h} N \quad (2.51)$$

The boundary condition in the case of the gravity anomaly can now be found by inserting Equations (2.45) and (2.48) into Equation (2.51)

$$\Delta g = -\frac{\partial T}{\partial h} \Big|_{Q'} + \frac{1}{\gamma} \frac{\partial \gamma}{\partial h} T_{Q'} \quad (2.52)$$

where $T_{P'}$ in Bruns's equation (Equation 2.45) is replaced by $T_{Q'}$, which is consistent with the linear approximation of the Taylor series (e.g., Moritz, 1989). The gravity anomaly in Equation (2.52) depends solely on functionals at point Q' on the ellipsoid, and is therefore herein called the gravity anomaly on the ellipsoid. Note that Equation (2.52) is not exactly similar to the fundamental equation of physical geodesy as defined by Heiskanen and Moritz (1967, Eq. 2-148), where the disturbing potential T and its normal derivative are defined on the geoid. Due to the fact that the gravity anomaly is here defined on the ellipsoid, the BVP of which Equation (2.52) forms the boundary condition is an ellipsoidal BVP, which can be solved without further approximations (see Chapters 3 and 4).

The BVP of solving the disturbing potential from gravity anomalies is called the third geodetic BVP, or alternatively the free-geodetic BVP. However, the latter name is not very appropriate under the current definition of the gravity anomaly on the ellipsoid, because this definition has essentially fixed the boundary to the ellipsoid. The boundary is thus no longer free, as it is under the more common definition of the fundamental equation of physical geodesy (Heiskanen and Moritz, 1967, Eq. 2-148). Inspection of Equation (2.39) shows that this third geodetic BVP is a Robin problem. The computation of geoid heights from the gravity anomalies in Equation (2.52) is called the Stokes (1849) approach to geoid modelling.

More rigorous relations between the gravity anomaly and the disturbing potential exist, which do not neglect the deflection of the vertical (e.g., Moritz, 1989; Heck, 1991; Vaníček et al., 1999) or that do not truncate the Taylor series (Grafarend et al., 1999). These more general formulas have, however, never been solved rigorously. All existing

solutions of the third ellipsoidal geodetic BVP take Equation (2.52) as a point of departure, and even approximate this equation further to facilitate its solution (e.g., Stokes, 1849). In this thesis, further approximation of Equation (2.52) is avoided. The influence of the approximations already made in the derivation of Equations (2.45), (2.48) and (2.52) are discussed in detail in the following Section.

2.3.3 Approximation effects in geodetic boundary conditions

Approximation effects in geodetic boundary conditions are here defined as the theoretical differences between the classical first-order spherical approximations and their rigorous counterparts. Recall that effects due to the downward-continuation of gravity on the surface of the Earth to the geoid and the removal of the effect of the topographic mass are not taken into account here, and that atmospheric and tidal effects are also assumed to be properly modelled. Still, there are three approximation effects remaining that follow directly from the derivations of the boundary conditions in Equations (2.45), (2.48) and (2.52). Their magnitudes and possible ways to model them are summarised below.

The linearisation effect

The linearisation effect affects both the first and the third geodetic BVPs and it stems from the truncation of the Taylor series in Equations (2.44) and (2.50) after the first term. Grafarend et al. (1999) provide an algorithm to compute the geoid height N from Equation (2.44) using the solution of a multivariate homogeneous polynomial (Grafarend et al., 1996), and numerical results of this formulation are given in Ardalan and Grafarend (2001a). This algorithm gives a more accurate version of Bruns's equation by adding terms with higher powers of the disturbing potential

$$N = \frac{T}{\gamma} - \frac{1}{2\gamma^3} \frac{\partial \gamma}{\partial h} T^2 + \left[\frac{1}{2\gamma^5} \left(\frac{\partial \gamma}{\partial h} \right)^2 - \frac{1}{\gamma^4} \frac{\partial^2 \gamma}{\partial h^2} \right] T^3 + \mathcal{O}(T^4) \quad (2.53)$$

Since the magnitude of the higher-order terms decreases rapidly, the error in Bruns's equation (Equation 2.45) can be estimated from the second term. The Earth's disturbing potential has an absolute maximum of approximately $1000 \text{ m}^2\text{s}^{-2}$, which means that

the maximum error in the geoid height due to linearisation effect in Bruns's equation is approximately 1.5 mm.

It is worth noticing that a refinement to Bruns's equation simpler than Equation (2.53) can also be derived. When the reference potential in the nonlinear terms in the Taylor series of Equation (2.44) is approximated by the spherical reference potential (Equation 2.32), the summation of the terms with j running from 2 to ∞ is given by an alternating power series

$$\begin{aligned} T &= \gamma N - \sum_{j=2}^{\infty} \frac{1}{j!} \frac{\partial^j \tilde{U}}{\partial h^j} N^j \\ &= \gamma N - \frac{\mu}{R} \sum_{j=2}^{\infty} (-1)^j \left(\frac{N}{R}\right)^j \end{aligned} \quad (2.54)$$

where R is the spherical radius and where the spherical approximation of the reference potential has been indicated by a tilde. The infinite sum in Equation (2.54) can be expressed as a closed analytic expression. When the disturbing potential T is multiplied by N/R and subsequently added to T , all cubic and higher-order terms cancel out, resulting in

$$T = \gamma N - \frac{\mu}{R^2} \left(\frac{N^2}{R+N}\right) = \frac{(\gamma - \tilde{\gamma})N^2 + R\gamma N}{R+N} \quad (2.55)$$

The term containing the square of the geoid undulation on the right-hand side of Equation (2.55) will always be small. In the case of the GRS80 ellipsoidal reference potential, the contribution of this quadratic term is only approximately 5×10^{-5} percent of the total disturbing potential, equivalent to around 50 μm in geoid undulation at most. Disregarding this term, the solution for the geoid undulation N can then be found as Bruns's equation multiplied by a factor close to unity

$$N = \frac{T}{\gamma} \left(\frac{R}{R - T/\gamma}\right) \quad (2.56)$$

This generalisation of Bruns's equation is exact if a spherical reference gravity field (Equation 2.32) is applied, and improves on the accuracy of the original Bruns's equation by approximately three orders of magnitude in case of the GRS80 reference field.

Since the simple Bruns's equation (Equation 2.45) is used in the derivation of Equation (2.52), the third geodetic BVP also contains a linearisation effect. Moreover, a second

linearisation error is apparent due to the truncation of the Taylor series in Equation (2.50). Grafarend et al. (1999) give the generalised fundamental equation without linearisation effect as

$$\Delta g = -\frac{\partial T}{\partial h} + \frac{1}{\gamma} \frac{\partial \gamma}{\partial h} T + \frac{1}{2\gamma^3} \left[\frac{\partial^2 \gamma}{\partial h^2} - \left(\frac{\partial \gamma}{\partial h} \right)^2 \right] T^2 + \mathcal{O}(T^3) \quad (2.57)$$

Applying the principle of spherical approximation to the nonlinear terms in the Taylor series in Equation (2.50) and inserting Equation (2.56) gives the same series, but in this case all terms after the second term vanish. Since this formulation is exact for a spherical reference gravity field, the linearisation effect in the fundamental equation can be given as a closed analytical expression when this spherical reference field is applied throughout.

When an ellipsoidal reference field is applied, an exact closed expression cannot be found, but in this case higher-order terms can be neglected because the linearisation effect is in this case much smaller

$$\Delta g = -\frac{\partial T}{\partial h} + \frac{1}{\gamma} \frac{\partial \gamma}{\partial h} T + \frac{T^2}{\mu} + \mathcal{O}(T^3) \quad (2.58)$$

From both the formulations in Equations (2.57) and (2.58) it can be estimated that the error in the original fundamental equation of physical geodesy (Equation 2.52) due to the linearisation effect has an absolute maximum of approximately $0.3\mu\text{Gal}$ ($1 \text{ Gal} \equiv 100 \text{ m/s}^2$), which is negligible considering the accuracy of absolute gravimeters, which is in the order of $1\mu\text{Gal}$ (Faller, 2002), and its effect on the geoid height is well below 1 mm. Further investigations into the linearisation effect of the third geodetic BVP are made in, e.g., Seitz et al. (1994) and Heck and Seitz (1995).

The deflection effect

The deflection effect affects the second and third geodetic BVPs and stems from neglecting the deflection of the vertical Θ . Various definitions of the deflection of the vertical exist, the most common of which is Helmert's definition (Jekeli, 1999), which defines the vertical deflection as the difference in direction between the gravity vector and the ellipsoidal normal. On the surface of the ellipsoid this is equal to the direction difference between the gravity vector and the reference gravity vector.

It should be noted that some authors consider the deflection effect as a component of the ellipsoidal effect (e.g., Vaníček et al., 1999). This approach is not followed here. Instead, the approach of Jekeli (1981) is followed, where the deflection effect in the second and third geodetic BVPs is given by the difference between the derivatives of the gravity potential in the direction of the normal gravity vector and in the direction of the true gravity vector on the ellipsoid. In Vaníček et al. (1999), this effect is included inside the correction that accounts for the approximation of the gravity disturbance by the derivative of the disturbing potential in the inward radial direction.

Jekeli (1981, p.125) and Cruz (1986, p.3) provide distinctly different estimates for the magnitude of the deflection effect, which can presumably be explained by an unwarranted approximation of the latitudinal derivative of the normal potential in Jekeli's (1981) derivation. Here, the approach of Jekeli (1981) is followed, but with a simple correction, confirming the result of Cruz (1986).

The deflection of the vertical is usually decomposed into a north-south component (ξ) and an east-west component (η). It follows from Figure 2.5 that the north-south and east-west deflections are related to the total deflections by

$$\tan^2 \Theta = \tan^2 \xi + \tan^2 \eta \quad (2.59)$$

Because the total deflection does generally not exceed 70 arcseconds (Bomford, 1971), Equation (2.59) can safely be approximated by

$$\Theta^2 = \xi^2 + \eta^2 \quad (2.60)$$

It also follows from Figure 2.5 that the derivative in the direction of the gravity vector can be expressed in terms of geodetic coordinates using the direction cosines of dH with respect to the normal directions that follow from geometric consideration

$$\frac{\partial}{\partial H} = \cos \Theta \left(\frac{\partial}{\partial h} - \tan \xi \frac{\partial}{\rho \partial \vartheta} + \tan \eta \frac{\partial}{\nu \sin \vartheta \partial \lambda} \right) \quad (2.61)$$

Inserting the gravity potential W and approximating the direction cosines in the horizontal directions gives an expression for the error introduced by neglecting the deflection effect (Jekeli, 1981)

$$\frac{\partial W}{\partial h} - \frac{\partial W}{\partial H} = (1 - \cos \Theta) \frac{\partial W}{\partial h} + \xi \frac{\partial W}{\rho \partial \vartheta} - \eta \frac{\partial W}{\nu \sin \vartheta \partial \lambda} \quad (2.62)$$

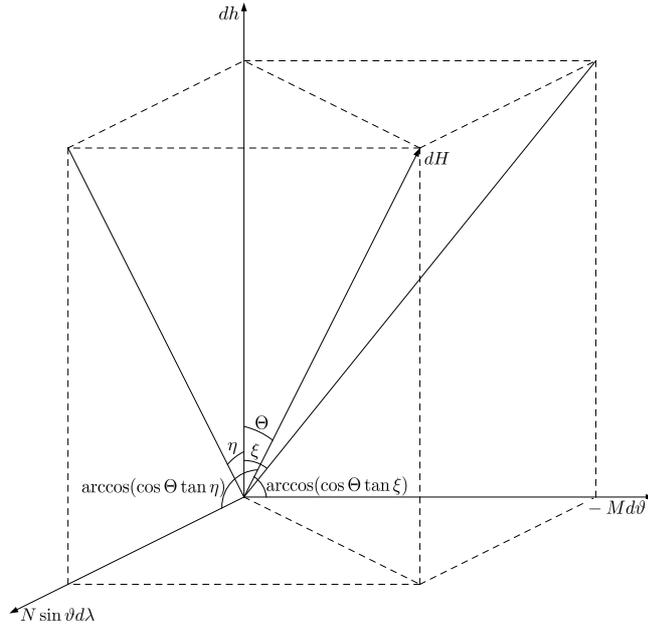


Figure 2.5: The deflections of the vertical on the surface of the ellipsoid

Further approximating the principal radii of curvature (Equations 2.14 and 2.15) by the spherical radius R and using the approximate relations for the deflections of the vertical (Jekeli, 1999)

$$\xi \approx \frac{1}{\gamma R} \frac{\partial T}{\partial \vartheta} \quad \text{and} \quad \eta \approx -\frac{1}{\gamma R \sin \vartheta} \frac{\partial T}{\partial \lambda} \quad (2.63)$$

it follows that Equation (2.62) can be approximated by

$$\frac{\partial W}{\partial h} - \frac{\partial W}{\partial H} = (\cos \Theta - 1)g + \xi \frac{\partial U}{R \partial \vartheta} - \eta \frac{\partial U}{\nu \sin \vartheta \partial \lambda} + \xi^2 \gamma + \eta^2 \gamma \quad (2.64)$$

Since the normal potential U is constant on the surface of the ellipsoid, the derivatives of U with respect to the geodetic latitude and longitude equal zero. An approximation of the deflection error can thus be obtained

$$\frac{\partial W}{\partial h} - \frac{\partial W}{\partial H} = \frac{1}{2} \gamma \Theta^2 \quad (2.65)$$

where the first term on the right-hand side of Equation (2.64) is approximated by $-\frac{1}{2} \gamma \Theta^2$. Total deflections as large as 70 arcseconds have been measured near Mount Everest (Bomford, 1971, p. 528), for which case the deflection error, according to Equation (2.65), is approximately 56 μGal . The deflection error can be properly modelled if an approximation of the vertical deflection is available, and this can, for example, be computed a priori from a global geopotential model (Jekeli, 1999).

The ellipsoidal effect

The ellipsoidal effect is the largest of the three approximation effects in BVPs. It is caused by the approximation of the derivative of the disturbing potential and the reference gravity with respect to the ellipsoidal normal by the radial derivative in the second and third geodetic BVP, and by the approximation of the ellipsoidal reference gravity by a spherical one in the third geodetic BVP. These approximations are usually applied to facilitate their solution. In these spherical approximations, indicated by tildes, the second geodetic boundary condition (Equation 2.48) reads

$$\widetilde{\delta g} = -\frac{\partial T}{\partial r} \quad (2.66)$$

and the third geodetic boundary condition (Equation 2.52) reads

$$\widetilde{\Delta g} = -\frac{\partial T}{\partial r} - \frac{2}{r}T \quad (2.67)$$

The error introduced by these spherical approximations is of the order of the flattening of the Earth, i.e., approximately 0.3%. A geoid computation based on Equation (2.67) leads to an error in the geoid height of several decimetres.

An extra approximation is often made by replacing the geocentric radius r by a spherical radius R , and this is known as the constant radius approximation. Equation (2.67) then becomes

$$\widetilde{\Delta g} = -\frac{\partial T}{\partial r} - \frac{2}{R}T \quad (2.68)$$

It will be shown in the remainder of this thesis that the errors resulting from this constant radius approximation can become much greater than the order of the Earth's flattening in certain applications, such as the analytic computation of a global spherical harmonic geopotential model from terrestrial gravity observations.

The primary aim of ellipsoidal physical geodesy is to correct for or avoid these errors, by solving the rigorous boundary conditions in Equations (2.45), (2.48) and (2.52) for an ellipsoidal boundary. Theoretically exact solutions to these boundary conditions are derived in the following chapters for the first time, in the convenient spherical harmonic framework, in accordance with the objective of this study (Section 1.4).

2.4 Harmonic base functions

The Earth's gravity field can be represented in many ways, the most popular of which by far is a representation as a series of solid spherical harmonic base functions (SHFs), often simply called spherical harmonics (e.g., Hobson, 1931; MacRobert, 1967). An alternative to spherical harmonics are ellipsoidal harmonics (e.g., Hobson, 1931), which play an important role in many methodologies in ellipsoidal physical geodesy, but are seldom used as a representation of the Earth's gravity field (see Section 1.4.2).

2.4.1 Solid spherical harmonics

Solid spherical harmonic expansions provide a solution to Laplace's differential equation (Equation 2.27), which in spherical polar coordinates reads (e.g., Hobson, 1931; MacRobert, 1967)

$$r^2 \frac{\partial^2 T}{\partial r^2} + 2r \frac{\partial T}{\partial r} + \cot \theta \frac{\partial T}{\partial \theta} + \frac{\partial^2 T}{\partial \theta^2} + \csc^2 \theta \frac{\partial^2 T}{\partial \lambda^2} = 0 \quad (2.69)$$

where \csc denotes the cosecant (the multiplicative inverse of the sine). It is shown here for the disturbing potential T , but naturally holds for any harmonic function. The solution to Equation (2.69) can be found by separation of variables (e.g., Dragomir et al., 1982; Sigl, 1985) and leads to two series solutions, one of which diverges for $r \rightarrow \infty$ and one of which diverges for $r \rightarrow 0$. The latter can be used to represent the Earth's external gravity field, since according to Newton's universal law of gravitation (Equation 2.18), the gravity field tends to zero for $r \rightarrow \infty$. This solution reads

$$T(r, \theta, \lambda) = \sum_{n=0}^{\infty} \left(\frac{R}{r} \right)^{n+1} \sum_{m=-n}^n T_{nm}^R Y_{nm}(\theta, \lambda) \quad (2.70)$$

where T_{nm}^R are the spherical harmonic coefficients (SHCs) of degree n and order m , and R is the radius of some reference sphere. The inclusion of R is in fact not necessary, since it is a constant, but it provides a useful scaling to the coefficients T_{nm}^R because the absolute value of the term $r^{-(n+1)}$ rapidly decreases with increasing degrees n . The use of the superscript R in the SHCs is not common, but is applied here to show that the values of the SHCs depend on the choice of the reference sphere.

In Equation (2.70), Y_{nm} are the SHFs, which are defined as

$$Y_{nm}(\theta, \lambda) = P_{n|m|}(\cos \theta) \begin{cases} \cos m\lambda & \text{for } m \leq 0 \\ \sin m\lambda & \text{for } m > 0 \end{cases} \quad (2.71)$$

where P_{nm} are the associated Legendre functions (ALFs). These are the solutions to the associated Legendre differential equation (e.g., Abramowitz and Stegun, 1972)

$$\frac{\partial^2 P_{nm}}{\partial \theta^2} + \cot \theta \frac{\partial P_{nm}}{\partial \theta} + [n(n+1) - m^2 \csc^2 \theta] P_{nm} = 0 \quad (2.72)$$

Several recurrence formulas among ALFs are known, such as (e.g., MacRobert, 1967; Abramowitz and Stegun, 1972)

$$P_{nm} = (2n-1) \sin \theta P_{n-1, m-1} + P_{n-2, m}, \quad m \geq 1 \quad (2.73)$$

$$P_{nm} = \frac{2n-1}{n-m} \cos \theta P_{n-1, m} - \frac{n+m-1}{n-m} P_{n-2, m}, \quad m \leq n-1 \quad (2.74)$$

$$P_{nm} = 2(m-1) \cot \theta P_{n, m-1} - (n-m+2)(n+m-1) P_{n, m-2}, \quad m \geq 2 \quad (2.75)$$

These equations can be used to compute the functions of degree $n \geq 2$, based on the initial values

$$P_{0,0}(\cos \theta) = 1 \quad (2.76)$$

$$P_{1,0}(\cos \theta) = \cos \theta \quad (2.77)$$

$$P_{1,1}(\cos \theta) = \sin \theta \quad (2.78)$$

where the functions with $m > n$ are defined as equal to zero.

ALFs are often fully normalised, making the average square value of the SHFs over the sphere equal to unity, by applying a scale factor based on the degree n and order m of the function

$$\bar{P}_{nm} = P_{nm} \begin{cases} \sqrt{2n+1} & \text{for } m = 0 \\ \sqrt{2(2n+1) \frac{(n-m)!}{(n+m)!}} & \text{for } m \neq 0 \end{cases} \quad (2.79)$$

although it should be noted that alternative normalisation procedures, such as the Gram-Schmidt semi-normalisation, also exist. The fully normalised ALFs in Equation (2.79) obey similar, but slightly different, recurrence relations than the unnormalised ones (Equations 2.73 - 2.75). Algorithms to compute them, avoiding numerical instabilities for large degrees and orders, are for example provided by Koop and Stelpstra (1989) and Holmes and Featherstone (2002).

Introducing the fully normalised ALFs into Equation (2.71) yields fully normalised SHFs \bar{Y}_{nm} . The normalisation ensures that the average square value of these fully normalised harmonics over the sphere is equal to unity

$$\frac{1}{4\pi} \int_{\sigma} \bar{Y}_{nm}(\theta, \lambda) \bar{Y}_{n'm'}(\theta, \lambda) d\sigma = \delta_{nn'} \delta_{mm'} \quad (2.80)$$

where δ_{ij} is the Kronecker delta, which is equal to 1 for $i = j$ and 0 otherwise. Equation (2.80) is an important relation since it shows that the SHFs form a set of orthogonal base functions on the sphere. Application of these fully normalised harmonics to represent a function, as in Equation (2.70), yields different SHCs and these are therefore also notated with an overbar

$$T(r, \theta, \lambda) = \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (2.81)$$

The coefficients \bar{T}_{nm}^R can be computed from the data on the sphere of radius R , which follows from the orthogonality of the SHFs on the sphere (Equation 2.80)

$$\bar{T}_{nm}^R = \frac{1}{4\pi} \int_{\sigma} T(R, \theta, \lambda) \bar{Y}_{nm}(\theta, \lambda) d\sigma \quad (2.82)$$

Thus, Equation (2.81) represents the harmonic function T anywhere in the exterior space and its coefficients can be derived from function values on a sphere. It is therefore a solution to the spherical Dirichlet BVP. The radius R in Equations (2.81) and (2.82) can theoretically be any real positive value. In many cases in physical and satellite geodesy, R is chosen equal to the semi-major axis of the Earth's ellipsoid a .

The convergence of a spherical harmonic expansion of the Earth's disturbing potential has long been a point of discussion (e.g., Krarup, 1969; Sjöberg, 1980, 1984; Jekeli, 1981, 1982; Colombo, 1982; Moritz, 1989). It is fairly straightforward to show that convergence can be guaranteed anywhere outside the smallest sphere that completely encloses the masses of the Earth, the so-called Brillouin sphere (e.g., Moritz, 1989). However, it is not elementary that the spherical harmonic expansion converges everywhere outside the Earth's surface, or everywhere outside a reference ellipsoid, even if the topography and atmosphere are assumed to be mathematically moved inside the ellipsoid.

The disturbing potential can be approximated arbitrarily well anywhere outside the Earth's masses or the ellipsoid, i.e., also inside the Brillouin sphere, according to the

Runge-Krarup theorem (Moritz, 1989). Therefore, convergence is often tacitly assumed in practice, despite the warnings of Sjöberg (1980, 1984) and Jekeli (1981, 1982) that the Runge-Krarup theorem cannot be invoked to justify use of a truncated spherical harmonic series of the disturbing potential, because the arbitrarily close approximation of the disturbing potential does not hold term by term.

Numerical computations by Jekeli (1982) have shown that the effect of divergence of the total spherical harmonic series is < 1 mm for geoid heights and $< 10\mu\text{Gal}$ for gravity anomalies using an expansion up to degree 300, which is practically negligible. The size of the error due to a possible divergence on or close to the Earth's surface for high-degree spherical harmonic expansions ($n \geq 300$) was identified by Sjöberg (1984) as an open question, which to the best of this author's knowledge remains unsolved today. However, the issue of convergence of spherical harmonic expansions falls outside the scope of this research, and the error due to possible divergence is assumed negligible in the remainder of this thesis.

2.4.2 Surface spherical harmonics

A surface spherical harmonic expansion is, in principle, a special case of a solid spherical harmonic expansion, but its properties allow for a broader use. For T residing on the sphere of radius R , Equation (2.81) reduces to

$$T(R, \theta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (2.83)$$

Since only a function on a surface is described, the expansion in Equation (2.83) is in this case called a *surface* spherical harmonic expansion. This expansion does, contrary to the solid harmonic expansion, not only hold for harmonic functions, but can be used to represent any arbitrary continuous function.

Moreover, surface SHCs are not restricted to a function defined on a sphere, as is for example pointed out by Jekeli (1988). This is due to the fact that the formula for the surface spherical harmonic expansion only depends on geocentric latitude and longitude, not on the geocentric distance. Thus, a surface harmonic expansion can be

applied to a function on a surface if any point on that surface is uniquely defined by its geocentric longitude and latitude. Such a surface is called a star-shaped surface, of which a more formal definition is given by Grafarend and Engels (1994).

A function defined on an ellipsoid of revolution, for example, can also be expanded into surface SHFs

$$T(r_e, \theta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \bar{T}_{nm}^{r_e} \bar{Y}_{nm}(\theta, \lambda) \quad (2.84)$$

The coefficients $\bar{T}_{nm}^{r_e}$ can, even though the function is defined on the ellipsoid, still be computed from an integration over the unit sphere

$$\bar{T}_{nm}^{r_e} = \frac{1}{4\pi} \int_{\sigma} T(r_e, \theta, \lambda) \bar{Y}_{nm}(\theta, \lambda) d\sigma \quad (2.85)$$

due to the orthogonality of the SHFs on the sphere (Equation 2.80). It is not a surface integration, but can be interpreted as a transform, where every pair of co-latitude and longitude uniquely identifies one point on the ellipsoid (Jekeli, 1988).

The fact that surface spherical harmonic expansions can be applied to functions on an ellipsoid is very useful in the field of ellipsoidal physical geodesy. It will prove of vital importance in the derivations of new solutions to ellipsoidal BVPs and new accurate gravity field modelling techniques described in the remainder of this thesis.

2.4.3 Ellipsoidal harmonics

Ellipsoidal harmonics can be derived in a very similar way to solid spherical harmonics, the difference being that Laplace's differential equation (Equation 2.27) is now transformed into ellipsoidal coordinates (u, β, λ) (e.g., Hobson, 1931)

$$(u^2 + E^2) \frac{\partial^2 T}{\partial u^2} + 2u \frac{\partial T}{\partial u} + \cot \theta \frac{\partial T}{\partial \beta} + \frac{\partial^2 T}{\partial \beta^2} + \frac{u^2 + E^2 \cos^2 \beta}{(u^2 + E^2) \sin^2 \beta} \frac{\partial^2 T}{\partial \lambda^2} = 0 \quad (2.86)$$

As in the case with spherical polar coordinates, there are two solutions to this equation, and the one that tends to zero for $u \rightarrow \infty$ reads

$$T(u, \beta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \frac{Q_{nm}\left(\frac{i u}{E}\right)}{Q_{nm}\left(\frac{i b}{E}\right)} \hat{T}_{nm}^u Y_{nm}(\beta, \lambda) \quad (2.87)$$

where Q_{nm} are ALFs of the second kind with complex argument and the circumflex in \hat{T}_{nm}^u indicates that these ellipsoidal harmonic coefficients differ from the spherical harmonic coefficients. The division over $Q_{nm}(ib/E)$ is in fact not required, but ensures that the coefficients \hat{T}_{nm}^u are real values.

The ALFs of the second kind Q_{nm} are, like the ALFs P_{nm} , a solution to the associated Legendre differential equation (Equation 2.72), and they too obey recurrence relations (Hobson, 1931; Sona, 1995). However, these recursive relations cannot be used to compute ALFs of the second kind due to numerical instabilities, even below degree $n = 20$ (Sona, 1995).

Several alternative procedures to compute Q_{nm} have been developed. Hobson (1931) derives an expression using the hypergeometric function F

$$Q_{nm}(z) = (-i)^m \frac{n!(n+m)!2^{2n+m+1}}{(2n+1)!z^{n+m+1}} (z^2 + 1)^{\frac{m}{2}} F\left(m + \frac{1}{2}; n + m + 1; n + \frac{3}{2}; \frac{1}{z^2}\right) \quad (2.88)$$

Other methods include the classical perturbative approach (Heck, 1991) and the limit layer approach (Sona, 1995), which are both based on approximations of the Legendre differential equation to order e^2 .

Equation (2.88) can also be used to derive a transformation between the spherical harmonic coefficients T_{nm}^R and the ellipsoidal harmonic coefficients \hat{T}_{nm}^u (Buchdahl et al., 1977; Jekeli, 1988; Dechambre and Scheeres, 2002). The appearance of slowly converging hypergeometric series in the transformation formulas, however, hinders their practicality.

2.5 Summary

In this Chapter, the formulas that form the foundation of ellipsoidal physical geodesy were derived. The geometry of the geodetic reference ellipsoid and the coordinate systems associated with it were discussed, and two new parameters – the quartic eccentricity (Equation 2.12) and the ellipsoidal deflection angle (Equation 2.16) – were introduced. However, the major part of this Chapter deals with the physical aspect of

ellipsoidal physical geodesy. The Earth's external gravitational field was shown to be harmonic, which means that it can be expressed as a potential field and represented by a spherical or ellipsoidal harmonic expansion. The computation of the Earth's external gravity field from gravity observations on its surface requires the solution of a geodetic boundary-value problem (BVP).

Three geodetic BVPs were derived here, with three different boundary conditions that relate the observations to the gravity potential (Equations 2.45, 2.48 and 2.52) based on various observation types. These boundary conditions are constructed on the surface of an ellipsoid, resulting in so-called ellipsoidal BVPs. Here and throughout this thesis, it is assumed that the Earth's topography and atmosphere are mathematically removed, and that their indirect effect on the gravity field is accounted for. The theory presented here is independent of their treatment.

The magnitude of two approximations made in the derivation of the boundary condition – the linearisation effect and the deflection effect – were found not to exceed $0.3\mu\text{Gal}$ and $56\mu\text{Gal}$, respectively, and formulas to model them were derived (Equations 2.58 and 2.65). A third and fourth approximation often applied in practice are the spherical and constant radius approximations, and these together are generally called the ellipsoidal effect, the magnitude of which is larger than that of the first two approximations (see Chapters 5 and 6). The remainder of this thesis is focussed on the rigorous solution to ellipsoidal boundary-value problems avoiding the latter two approximations.

3. A SOLUTION TO THE ELLIPSOIDAL DIRICHLET BOUNDARY-VALUE PROBLEM IN SPHERICAL HARMONICS

In this Chapter, a solution to the ellipsoidal Dirichlet boundary-value problem (BVP) is derived in terms of the simple spherical harmonic functions (SHFs). This results in a novel approach to the computation of solid spherical harmonic coefficients (SHCs) from function values on the surface of the reference ellipsoid. In Section 3.1, several strategies towards the solution of the problem are discussed, after which the novel approach is outlined in detail in the remaining Sections. A vital step in the derivation involves a new relation among associated Legendre functions (ALFs) and SHFs, which is derived in Section 3.2 (cf. Claessens, 2005). Finally, a numerical closed-loop simulation is performed to confirm the validity of the new approach.

3.1 Solution strategies to ellipsoidal boundary-value problems

The external ellipsoidal Dirichlet BVP is to determine a harmonic function $f(r, \theta, \lambda)$ anywhere in the space outside the ellipsoid, from function values \mathcal{F} on the surface of the ellipsoid

$$f(r, \theta, \lambda) = \mathcal{F} \quad \text{for} \quad r = r_e, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \lambda < 2\pi \quad (3.1)$$

where the function is harmonic outside the ellipsoid

$$\nabla^2 f(r, \theta, \lambda) = 0 \quad \text{for} \quad r > r_e, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \lambda < 2\pi \quad (3.2)$$

and approaches zero at infinity, obeying the asymptotic regularity condition

$$f(r, \theta, \lambda) \sim \mathcal{O}\left(\frac{1}{r}\right) \quad \text{for} \quad r \rightarrow \infty, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \lambda < 2\pi \quad (3.3)$$

Since the function $f(r, \theta, \lambda)$ is harmonic outside the ellipsoid, it can in this (unbounded) region be represented by a solid spherical harmonic expansion (Equation 2.81)

$$f(r, \theta, \lambda) = \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} \sum_{m=-n}^n \bar{f}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (3.4)$$

where the solid SHCs \bar{f}_{nm}^R can be computed from function values given continuously on a sphere with radius R

$$\bar{f}_{nm}^R = \frac{1}{4\pi} \int_{\sigma} f(R, \theta, \lambda) \bar{Y}_{nm}(\theta, \lambda) d\sigma \quad (3.5)$$

or alternatively from function values given on any other sphere of radius R'

$$\bar{f}_{nm}^R = \frac{1}{4\pi} \left(\frac{R'}{R} \right)^{n+1} \int_{\sigma} f(R', \theta, \lambda) \bar{Y}_{nm}(\theta, \lambda) d\sigma \quad (3.6)$$

since it follows from Equation (3.4) that

$$\bar{f}_{nm}^R = \left(\frac{R'}{R} \right)^{n+1} \bar{f}_{nm}^{R'} \quad (3.7)$$

However, if the function values are not given on a sphere, the computation of solid SHCs becomes more complicated. This is due to the fact that the the orthogonality of SHFs on the sphere (Equation 2.80) does not hold on the ellipsoid. This has been the main hindrance to the solution of the ellipsoidal BVPs in the spherical harmonic framework. Several strategies can be followed to solve the ellipsoidal BVP. Four methods that present the solution in terms of solid SHCs are discussed below.

3.1.1 The ellipsoidal harmonics method

An obvious approach to the solution of the ellipsoidal BVP is to first write the solution in terms of an ellipsoidal harmonic expansion (Equation 2.87)

$$f(u, \beta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \hat{f}_{nm}^u \frac{Q_{nm}(i\frac{u}{E})}{Q_{nm}(i\frac{b}{E})} Y_{nm}(\beta, \lambda) \quad (3.8)$$

The ellipsoidal harmonic coefficients \hat{f}_{nm}^u can then be transformed into SHCs \bar{f}_{nm}^R using the transformation derived by Jekeli (1988).

This approach contains some disadvantages. First of all, the numerical stability of very high degree ellipsoidal harmonic base functions is not well-known. Sona (1995) pinpoints several difficulties in the computation of ellipsoidal harmonics, and the stable computation of ellipsoidal harmonics over a wide range of degrees, orders and latitudes remains an open area for future research. The numerical computation of SHFs, on the

other hand, has received much more interest (e.g., Colombo, 1981; Koop and Stelpstra, 1989), and stable algorithms are in existence (e.g., Holmes and Featherstone, 2002). Secondly, the transformation from ellipsoidal harmonic coefficients to SHCs is complicated and time-consuming, since the formulas involve slowly converging hypergeometric series (see also Section 1.3.3) .

3.1.2 The upward-continuation method

Another method sometimes used in practice (e.g., Rapp, 1977; Cruz, 1986; Petrovskaya et al., 2001) is the upward-continuation method, where the data on the ellipsoid are first upward-continued to an enveloping sphere, i.e., a sphere that circumscribes the ellipsoid with radius equal to the semi-major axis of the ellipsoid ($R = a$), using a Taylor series expansion. Function values on the sphere can be obtained using

$$\begin{aligned} f(R, \theta, \lambda) &= f(r_e, \theta, \lambda) + \left. \frac{\partial f}{\partial r} \right|_{r_e} (R - r_e) + \frac{1}{2} \left. \frac{\partial^2 f}{\partial r^2} \right|_{r_e} (R - r_e)^2 + \dots \\ &= \sum_{i=0}^{\infty} \frac{1}{i!} \left. \frac{\partial^i f}{\partial r^i} \right|_{r_e} (R - r_e)^i \end{aligned} \quad (3.9)$$

Equation (3.5) can then be applied to yield the solid SHCs.

This is, in fact, not a strict solution to the ellipsoidal Dirichlet BVP, since radial derivatives of the function on the ellipsoid are also required. When the eccentricity of the ellipsoid is small, the Taylor series in Equation (3.9) converges rapidly and a truncation after the term containing the first radial derivative may give a solution of sufficient accuracy, depending on the application. Thus, this method can, under certain conditions, yield a satisfactory solution, but will, as well as the function values on the ellipsoid, at least require their first-order radial derivative. An additional condition for this method is that the derivatives are continuous in the space between the ellipsoid and the sphere, but this is guaranteed by the harmonicity condition if the sphere is completely outside the ellipsoid.

3.1.3 The ellipsoidal integration method

A third method is based on Green's second integral theorem (Equation 2.41)

$$\iiint_{\mathcal{V}} (f^* \Delta f - f \Delta f^*) d\mathcal{V} = \iint_{\mathcal{S}} \left(f^* \frac{\partial f}{\partial h} - f \frac{\partial f^*}{\partial h} \right) d\mathcal{S} \quad (3.10)$$

which relates a three-dimensional integration over a volume \mathcal{V} to a two-dimensional integration over surface \mathcal{S} , using two non-identical potential functions f and f^* . Sjöberg (1988) shows that this theorem can be used to find an expression for the solid SHC \bar{f}_{nm}^R using an integration over the ellipsoid

$$\bar{f}_{nm}^R = \frac{1}{4\pi R(2n+1)a_n} \iint_{\mathcal{E}} \left(f \frac{\partial f^*}{\partial h} - f^* \frac{\partial f}{\partial h} \right) d\mathcal{E} \quad (3.11)$$

where a_n is an arbitrary constant, \mathcal{E} is the ellipsoidal surface and f^* is defined as

$$f^* = \left[\left(\frac{R}{r_e} \right)^{n+1} + a_n \left(\frac{r_e}{R} \right)^n \right] \bar{Y}_{nm}(\theta, \lambda) \quad (3.12)$$

It follows from Equation (3.11) that, for this solution, the derivative of the function f with respect to the ellipsoidal normal needs to be known, and this is the major drawback of this method. Furthermore, the normal derivative of the function f^* needs to be found, which is possible in a theoretically exact way, but makes this approach more complicated than the upward-continuation method. However, it avoids the truncation error that will always be apparent in practical application of the upward-continuation method.

3.1.4 The coefficient transformation method

A second analytic approach that does not require any information, other than the function values on the surface of the ellipsoid (in contrast to the upward-continuation and ellipsoidal integration approaches), is possible. This is a novel approach to solve the ellipsoidal Dirichlet BVP, based on the fact that function values on the ellipsoidal surface can be described as a series of surface SHFs (Equation 2.84)

$$f(r_e, \theta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \bar{f}_{nm}^{r_e} \bar{Y}_{nm}(\theta, \lambda) \quad (3.13)$$

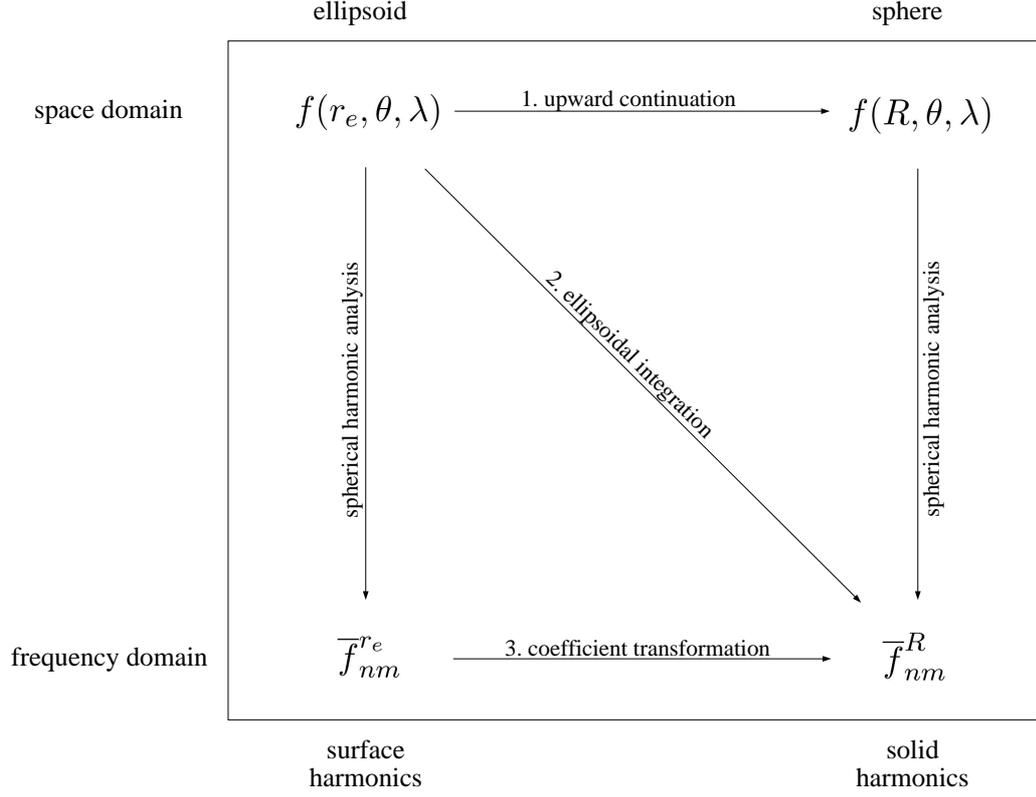


Figure 3.1: Schematic overview of the computation of solid SHCs \bar{f}_{nm}^R from function values on the ellipsoid $f(r_e, \theta, \lambda)$ using the 1. upward-continuation, 2. ellipsoidal integration, and 3. coefficient transformation methods

where the surface SHCs $\bar{f}_{nm}^{r_e}$ can be computed from an integration over the unit sphere

$$\bar{f}_{nm}^{r_e} = \frac{1}{4\pi} \int_{\sigma} f(r_e, \theta, \lambda) \bar{Y}_{nm}(\theta, \lambda) d\sigma \quad (3.14)$$

The surface SHCs can then be transformed into solid SHCs, as follows. Comparing Equations (3.4) and (3.13) gives

$$\sum_{n=0}^{\infty} \sum_{m=-n}^n \left(\frac{R}{r_e}\right)^{n+1} \bar{f}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \bar{f}_{nm}^{r_e} \bar{Y}_{nm}(\theta, \lambda) \quad (3.15)$$

A schematic overview of this approach, as well as the upward-continuation and ellipsoidal integration approaches, is shown in Figure 3.1. The SHFs \bar{Y}_{nm} form a set of orthogonal base functions on the sphere, which would allow for a one-to-one comparison between the coefficients of each pair of degree n and order m , if the presence of the ellipsoidal radius r_e on the left-hand side of Equation (3.15) would not have disturbed this orthogonality. The key to finding a transformation between the solid SHCs \bar{f}_{nm}^R and the surface SHCs $\bar{f}_{nm}^{r_e}$ therefore lies in the ability to move the dependence on

latitude that is present inside the ellipsoidal radius into the SHFs \bar{Y}_{nm} . This can be achieved using a new relation among SHFs presented below.

3.2 New relations among spherical harmonic base functions

Many relations among SHFs exist, such as for example the product-sum formula, which writes the product of two SHFs as a weighted summation over SHFs of equal order (e.g., Giacaglia and Burša, 1980; Hwang, 1995). Such a weighted summation can also be achieved for the product of an SHF with any trigonometric function. Here, these known relations will be extended to hold for the product of SHFs with arbitrary powers of the sine and cosine of the latitude. A more detailed description and additional relations are provided in Claessens (2005).

3.2.1 Basic recursive relations

Equation (2.74) can be transformed in such a way that the product of the cosine and an ALF $P_{nm}(\cos \theta)$ is expressed as a weighted sum over two other ALFs, where the weights F are independent of co-latitude θ

$$\cos \theta P_{nm} = F_{nm}^{-1} P_{n-1,m} + F_{nm}^1 P_{n+1,m} \quad (3.16)$$

where

$$F_{nm}^{-1} = \frac{n+m}{2n+1} \quad \text{and} \quad F_{nm}^1 = \frac{n-m+1}{2n+1} \quad (3.17)$$

Equation (3.16) can be transformed into a fully normalised form using Equation (2.79), yielding

$$\cos \theta \bar{P}_{nm} = \bar{F}_{nm}^{-1} \bar{P}_{n-1,m} + \bar{F}_{nm}^1 \bar{P}_{n+1,m} \quad (3.18)$$

where

$$\bar{F}_{nm}^{-1} = \sqrt{\frac{(n+m)(n-m)}{(2n-1)(2n+1)}} \quad (3.19)$$

$$\bar{F}_{nm}^1 = \sqrt{\frac{(n-m+1)(n+m+1)}{(2n+1)(2n+3)}} \quad (3.20)$$

From Equation (2.71), it can be seen that Equations (3.16) and (3.18) are also valid when the ALFs P_{nm} and \overline{P}_{nm} are replaced by the SHFs Y_{nm} and \overline{Y}_{nm} respectively, since the ALFs within the summation are of equal order m . Equations (2.73) and (2.75) can be rewritten in similar fashion (Claessens, 2005).

3.2.2 New relations

As noted in, for example, Moritz (1989), Equation (3.16) can be applied twice to obtain a relation when the cosine is squared. Replacing the ALFs by SHFs, this results in

$$\cos^2 \theta Y_{nm} = F_{nm}^{-2} Y_{n-2,m} + F_{nm}^0 Y_{nm} + F_{nm}^2 Y_{n+2,m} \quad (3.21)$$

where

$$F_{nm}^{-2} = \frac{(n+m)(n+m-1)}{(2n-1)(2n+1)} \quad (3.22)$$

$$F_{nm}^0 = \frac{2n^2 - 2m^2 + 2n - 1}{(2n-1)(2n+3)} \quad (3.23)$$

$$F_{nm}^2 = \frac{(n-m+1)(n-m+2)}{(2n+1)(2n+3)} \quad (3.24)$$

Relations for fully normalised SHFs can also easily be derived in the same way. In the case of full normalisation, it follows from Equation (2.79) that Equation (3.21) remains the same, though with overbars over the SHFs \overline{Y}_{nm} and the functions \overline{F}_{nm}^i , which in this case become

$$\overline{F}_{nm}^{-2} = \sqrt{\frac{[(n-1)^2 - m^2][n^2 - m^2]}{(2n-3)(2n-1)^2(2n+1)}} \quad (3.25)$$

$$\overline{F}_{nm}^0 = \frac{2n(n+1) - 2m^2 - 1}{(2n-1)(2n+3)} \quad (3.26)$$

$$\overline{F}_{nm}^2 = \sqrt{\frac{[(n+1)^2 - m^2][(n+2)^2 - m^2]}{(2n+1)(2n+3)^2(2n+5)}} \quad (3.27)$$

In general, for any power of the cosine term, its product with an ALF (or similarly with an SHF) can be expressed as a weighted summation over ALFs or SHFs of equal degree

$$\cos^j \theta \overline{Y}_{nm} = \sum_{i=-j,2}^j \overline{F}_{nm}^{ij} \overline{Y}_{n+i,m}, \quad j \in \mathbb{N} \quad (3.28)$$

where the function \overline{F}_{nm}^{ij} depends on degree n and order m solely and \mathbb{N} is the set of natural numbers. The summation runs from $-j$ to j in steps of 2.

Since the sum of the squares of the sine and cosine of the same angle equals unity, it follows from Equation (3.21) that

$$\sin^2 \theta \overline{Y}_{nm} = \overline{K}_{nm}^{-2} \overline{Y}_{n-2,m} + \overline{K}_{nm}^0 \overline{Y}_{nm} + \overline{K}_{nm}^2 \overline{Y}_{n+2,m} \quad (3.29)$$

where

$$\overline{K}_{nm}^{-2} = -\overline{F}_{nm}^{-2}, \quad \overline{K}_{nm}^0 = 1 - \overline{F}_{nm}^0, \quad \overline{K}_{nm}^2 = -\overline{F}_{nm}^2 \quad (3.30)$$

Equation (3.29) can be generalised, leading to an equation similar to Equation (3.28), but which only holds for even values of j

$$\sin^j \theta \overline{Y}_{nm} = \sum_{i=-j,2}^j \overline{K}_{nm}^{ij} \overline{Y}_{n+i,m}, \quad j \in \mathbb{E} \quad (3.31)$$

where \mathbb{E} is the set of even natural numbers.

Since the functions \overline{F}_{nm}^{ij} and \overline{K}_{nm}^{ij} weight the ALFs or SHFs in the summation in Equations (3.28) and (3.31), they are herein called Legendre weight functions (LWFs). The functions \overline{F}_{nm}^{ij} give the relation for the case of the power of a cosine multiplied by the ALFs and are therefore called cosinusoidal LWFs, whereas \overline{K}_{nm}^{ij} are here called sinusoidal LWFs.

3.2.3 Computation of Legendre weight functions

Now the question arises how to compute the LWFs \overline{F}_{nm}^{ij} and \overline{K}_{nm}^{ij} in Equations (3.28) and (3.31). A recursive relation can be found by investigation of the scheme shown in Figure 3.2. The LWFs of index j can be obtained by applying Equation (3.16) to the previous level of index $j - 1$. For example, LWF $\overline{F}_{nm}^{-2,2}$ is multiplied by the SHF $\overline{Y}_{n-2,m}$ and when this is multiplied with another cosine of θ , this term becomes, with Equation (3.16)

$$\overline{F}_{nm}^{-2,2} \overline{F}_{n-2,m}^{-1,1} \overline{Y}_{n-3,m} + \overline{F}_{nm}^{-2,2} \overline{F}_{n-2,m}^{1,1} \overline{Y}_{n-1,m} \quad (3.32)$$

Thus, $\overline{F}_{nm}^{-3,3} = \overline{F}_{nm}^{-2,2} \overline{F}_{n-2,m}^{-1,1}$ and $\overline{F}_{nm}^{-1,3} = \overline{F}_{nm}^{-2,2} \overline{F}_{n-2,m}^{1,1}$ plus an extra term from $\overline{F}_{nm}^{0,2}$ (see Figure 3.2). In general, the arrows pointing left in Figure 3.2 indicate a multiplication

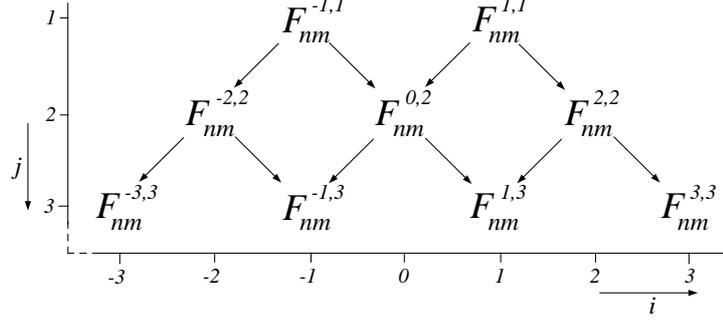


Figure 3.2: Recursive computation of cosinusoidal LWFs F_{nm}^{ij}

of the LWF with $\overline{F}_{n+i-1,m}^{-1,1}$ and the arrows pointing right indicate a multiplication of the LWF with $\overline{F}_{n+i-1,m}^{1,1}$. This leads to the following recursive relation

$$\overline{F}_{nm}^{ij} = \overline{F}_{nm}^{i-1,j-1} \overline{F}_{n+i-1,m}^{1,1} + \overline{F}_{nm}^{i+1,j-1} \overline{F}_{n+i+1,m}^{-1,1} \quad (3.33)$$

where \overline{F}_{nm}^{ij} are set to zero for $|i| > j$ or $n < m$. This recursive relation and all relations below hold for the ranges $j \geq 2$ and $-j \leq i \leq j$, unless otherwise stated, where it should be noted that i and j are either both even or both odd numbers.

Equation (3.33) can also be found in a purely algebraic way, as follows

$$\begin{aligned} \cos^j \theta \overline{Y}_{nm} &= \cos \theta \cos^{j-1} \theta \overline{Y}_{nm} \\ &= \sum_{p=-j+1,2}^{j-1} \overline{F}_{nm}^{p,j-1} \cos \theta \overline{Y}_{n+p,m} \\ &= \sum_{p=-j+1,2}^{j-1} \overline{F}_{nm}^{p,j-1} \sum_{q=-1,2}^1 \overline{F}_{n+p,m}^{q,1} \overline{Y}_{n+p+q,m} \end{aligned} \quad (3.34)$$

Substituting p by $i - q$, this results in

$$\cos^j \theta \overline{Y}_{nm} = \sum_{i=-j,2}^j \sum_{q=-1,2}^1 \overline{F}_{nm}^{i-q,j-1} \overline{F}_{n+i-q,m}^{q,1} \overline{Y}_{n+i,m} \quad (3.35)$$

and upon comparison with Equation (3.28), a relation for the LWF \overline{F}_{nm}^{ij} emerges as

$$\overline{F}_{nm}^{ij} = \overline{F}_{nm}^{i+1,j-1} \overline{F}_{n+i+1,m}^{-1,1} + \overline{F}_{nm}^{i-1,j-1} \overline{F}_{n+i-1,m}^{1,1} \quad (3.36)$$

which is exactly the same as Equation (3.33). However, if the terms $\cos \theta$ and $\cos^{j-1} \theta$

in Equation (3.34) are reversed, another recursive relation appears as

$$\begin{aligned}
\cos^j \theta \bar{Y}_{nm} &= \cos^{j-1} \theta \cos \theta \bar{Y}_{nm} \\
&= \sum_{q=-1,2}^1 \bar{F}_{nm}^{q,1} \sum_{p=-j+1,2}^{j-1} \bar{F}_{n+q,m}^{p,j-1} \bar{Y}_{n+p+q,m} \\
&= \sum_{i=-j,2}^j \sum_{q=-1,2}^1 \bar{F}_{nm}^{q,1} \bar{F}_{n+q,m}^{i-q,j-1} \bar{Y}_{n+i,m}
\end{aligned} \tag{3.37}$$

Thus

$$\bar{F}_{nm}^{ij} = \bar{F}_{nm}^{-1,1} \bar{F}_{n-1,m}^{i+1,j-1} + \bar{F}_{nm}^{1,1} \bar{F}_{n+1,m}^{i-1,j-1} \tag{3.38}$$

Equations (3.36) and (3.38) can both be used to compute any cosinusoidal LWFs from the initial values in Equation (3.17). Equation (3.36) has the advantage that the LWFs of one solitary pair of degree n and order m can be evaluated very quickly since only the initial values ($j = 1$) need to be known for other degrees. These initial values, $\bar{F}_{n+i+1,m}^{-1,1}$ and $\bar{F}_{n+i-1,m}^{1,1}$, can be directly computed from Equations (3.19) and (3.20). Thus, only a recursion over i and j is needed. Equation (3.38), on the other hand, requires knowledge of LWFs of degrees $n - 1$ and $n + 1$ with index $j - 1$ ($\bar{F}_{n-1,m}^{i+1,j-1}$ and $\bar{F}_{n+1,m}^{i-1,j-1}$). These LWFs, in turn, depend upon LWFs of degree $n - 2$, n and $n + 2$, with index $j - 2$. A recursive scheme based Equation (3.38) should therefore compute many LWFs of various degrees with indices up to $j - 1$ before the LWFs of higher index j can be computed. However, if all LWFs from degree and order zero onwards are desired, Equation (3.38) can provide a slightly more efficient recursion than Equation (3.36), because the initial values do not need to be stored or computed as often.

If one is only interested in LWFs of even j , as will prove to be the case in the solution of ellipsoidal BVPs, other recursive relations can be found by taking the case $j = 2$ as initial values. A recursive scheme is shown in Figure 3.3 and the mathematical relation can easily be formulated

$$\bar{F}_{nm}^{ij} = \sum_{k=-1}^1 \bar{F}_{nm}^{i-2k,j-2} \bar{F}_{n+i-2k,m}^{2k,2} \tag{3.39}$$

Also in this case, an alternative expression can be found along similar lines as Equations (3.37) to (3.38), which is

$$\bar{F}_{nm}^{ij} = \sum_{k=-1}^1 \bar{F}_{nm}^{2k,2} \bar{F}_{n+2k,m}^{i-2k,j-2} \tag{3.40}$$

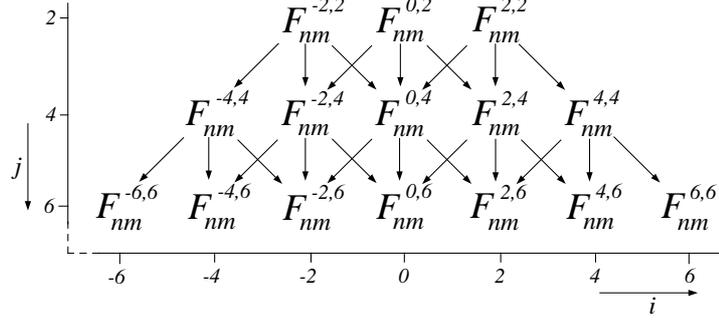


Figure 3.3: Recursive computation of cosinusoidal LWF F_{nm}^{ij} for even j

Recursive computation based on Equation (3.39) or (3.40) is here named dual mode computation of LWFs, whereas computation of LWFs using Equation (3.36) or (3.38) is named single mode computation. The dual mode is more efficient, since it avoids having to compute the coefficients of odd index j in the process.

The recursive relations for the computation of the cosinusoidal LWFs \overline{F}_{nm}^{ij} (Equations 3.33, 3.38, 3.39 and 3.40) hold for both fully normalised and unnormalised LWFs. More importantly, the sinusoidal LWFs \overline{K}_{nm}^{ij} can also be computed using the same recursive relations in both the fully normalised and unnormalised case, simply replacing \overline{F}_{nm}^{ij} by K_{nm}^{ij} or \overline{K}_{nm}^{ij} . However, the initial values in Equations (3.19) and (3.20) are different, and because of this, the fully normalised LWFs \overline{F}_{nm}^{ij} and \overline{K}_{nm}^{ij} possess some interesting properties that can facilitate their computation. Firstly, the fully normalised LWFs are symmetric with respect to order m , contrary to the non-normalised case, as can be seen from Equations (3.19) and (3.20). Secondly, the two initial values are connected by a simple equation

$$\overline{F}_{nm}^{1,1} = \overline{F}_{n+1,m}^{-1,1} \quad (3.41)$$

This relation can be generalised to

$$\overline{F}_{nm}^{ij} = \overline{F}_{n+i,m}^{-i,j} \quad (3.42)$$

A proof of Equation (3.42) based on mathematical induction can be found in Appendix A. Equation (3.42) also holds for the fully normalised sinusoidal LWFs \overline{K}_{nm}^{ij} , as follows from the relationship between the initial values of the cosinusoidal and sinusoidal LWFs in Equation (3.30). This relation can reduce the necessary computation time approximately by a factor of two, since only LWFs with $i \geq 0$ need to be computed.

The LWFs of negative index i then follow directly from Equation (3.42). Moreover, using Equation (3.42), Equations (3.36) and (3.39) can be formulated in a slightly more convenient way

$$\overline{F}_{nm}^{ij} = \overline{F}_{nm}^{i+1,j-1} \overline{F}_{n+i,m}^{1,1} + \overline{F}_{nm}^{i-1,j-1} \overline{F}_{n+i,m}^{-1,1} \quad (3.43)$$

$$\overline{F}_{nm}^{ij} = \sum_{k=-1}^1 \overline{F}_{nm}^{i+2k,j-2} \overline{F}_{n+i,m}^{2k,2} \quad (3.44)$$

where the degrees of the LWFs no longer depend upon the summation index k .

3.2.4 Numerical accuracy of Legendre weight functions

The computation of the sinusoidal and cosinusoidal LWFs can be performed in many ways using any of Equations (3.33), (3.38), (3.39) or (3.40). Numerical tests were performed to assess the errors in the computation of the LWFs for different values of θ and j . In the routine used to compute the LWFs, the functions were all computed on-the-fly, i.e., they were not stored in memory. This is the most practical approach, especially considering that for values of j up to 32, the LWFs outnumber the ALFs by more than a factor of 10^3 . For example, for a certain co-latitude θ , degree and order up to 360 and index j up to 32, there are 69,000,096 LWF values compared to the 65,341 ALF values.

In the numerical tests, the validity of Equations (3.28) and (3.31) is examined, and the measure used is an absolute relative error ς_{nmj} , defined as

$$\varsigma_{nmj}(\theta) = \frac{\left| \cos^j \theta \overline{P}_{nm} - \sum_{i=-j,2}^j \overline{F}_{nm}^{ij} \overline{P}_{n+i,m} \right|}{\cos^j \theta \overline{P}_{nm}} \quad (3.45)$$

This error measure also includes numerical errors in the ALFs \overline{P}_{nm} . For the computation of these ALFs, the forward column algorithm by Holmes and Featherstone (2002) was used, which has a relative precision below 10^{-12} up to $n = 2700$ for all values of θ tested.

Since the error measure ς_{nmj} depends on four variables (degree n , order m , index j and co-latitude θ) and also on the algorithm used to compute the LWFs, it is not

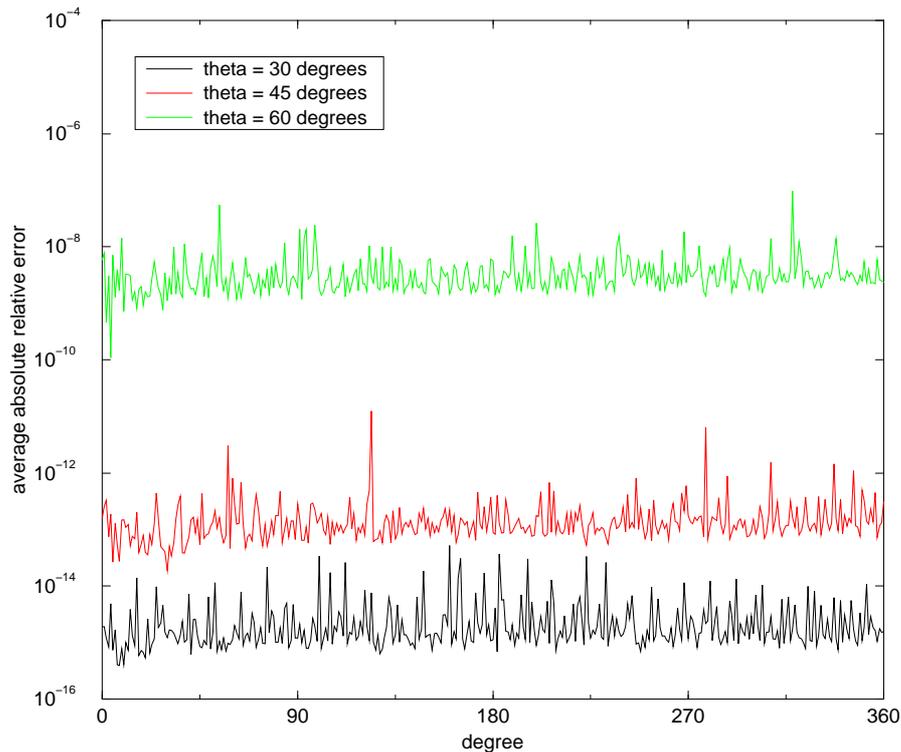


Figure 3.4: Average absolute relative error ς_n (Equation 3.47) per degree n in the computation of fully normalised cosinusoidal LWFs in dual mode (Equation 3.39) for $\theta = 30^\circ$, $\theta = 45^\circ$ and $\theta = 60^\circ$, where the functions are computed in IEEE double precision

straightforward to assess the influence of any of the variables independently. To obtain a useful tool for comparison, the average is taken over degrees n and orders m up to 360

$$\varsigma_j(\theta) = \frac{1}{65,341} \sum_{n=0}^{360} \sum_{m=0}^n \varsigma_{nmj}(\theta) \quad (3.46)$$

and the average over order m and index j up to 32

$$\varsigma_n(\theta) = \frac{1}{32(n+1)} \sum_{j=0}^{32} \sum_{m=0}^n \varsigma_{nmj}(\theta) \quad (3.47)$$

The measure ς_n is shown in Figure 3.4 for degrees up to 360 and $\theta = 30^\circ$, $\theta = 45^\circ$ and $\theta = 60^\circ$, where the LWFs were computed in 64-bit IEEE double precision (~ 16 significant digits). It can be seen that the relative error becomes larger for increasing θ . In addition, it can be seen from Figure 3.4 that the error ς_n is roughly independent of the degree n , which makes the average error over degree and order ς_j (Equation 3.46) a useful tool.

The values of ς_j for both fully normalised and non-normalised LWFs using either Equa-

θ	j	fully normalised		non-normalised	
		single	dual	single	dual
30°	2	$2.6 \cdot 10^{-16}$	$2.4 \cdot 10^{-16}$	$2.8 \cdot 10^{-16}$	$2.5 \cdot 10^{-16}$
	4	$5.4 \cdot 10^{-16}$	$4.9 \cdot 10^{-16}$	$5.1 \cdot 10^{-16}$	$4.6 \cdot 10^{-16}$
	8	$1.2 \cdot 10^{-15}$	$1.1 \cdot 10^{-15}$	$1.1 \cdot 10^{-15}$	$1.1 \cdot 10^{-15}$
	16	$3.6 \cdot 10^{-15}$	$3.2 \cdot 10^{-15}$	$3.2 \cdot 10^{-15}$	$3.1 \cdot 10^{-15}$
	32	$2.4 \cdot 10^{-14}$	$2.2 \cdot 10^{-14}$	$2.3 \cdot 10^{-14}$	$2.2 \cdot 10^{-14}$
45°	2	$1.1 \cdot 10^{-15}$	$1.0 \cdot 10^{-15}$	$1.1 \cdot 10^{-15}$	$6.8 \cdot 10^{-16}$
	4	$2.4 \cdot 10^{-15}$	$1.3 \cdot 10^{-15}$	$1.2 \cdot 10^{-15}$	$2.1 \cdot 10^{-15}$
	8	$3.2 \cdot 10^{-15}$	$4.5 \cdot 10^{-15}$	$3.8 \cdot 10^{-15}$	$5.7 \cdot 10^{-15}$
	16	$3.4 \cdot 10^{-14}$	$3.0 \cdot 10^{-14}$	$3.2 \cdot 10^{-14}$	$3.7 \cdot 10^{-14}$
	32	$3.9 \cdot 10^{-12}$	$3.4 \cdot 10^{-12}$	$3.8 \cdot 10^{-12}$	$3.6 \cdot 10^{-12}$
60°	2	$3.9 \cdot 10^{-16}$	$3.8 \cdot 10^{-16}$	$2.7 \cdot 10^{-17}$	$2.7 \cdot 10^{-17}$
	4	$1.4 \cdot 10^{-15}$	$1.3 \cdot 10^{-15}$	$1.5 \cdot 10^{-15}$	$1.5 \cdot 10^{-15}$
	8	$1.3 \cdot 10^{-14}$	$1.1 \cdot 10^{-14}$	$1.5 \cdot 10^{-14}$	$1.5 \cdot 10^{-14}$
	16	$1.7 \cdot 10^{-12}$	$1.7 \cdot 10^{-12}$	$1.6 \cdot 10^{-12}$	$1.6 \cdot 10^{-12}$
	32	$6.5 \cdot 10^{-08}$	$5.5 \cdot 10^{-08}$	$4.2 \cdot 10^{-08}$	$3.8 \cdot 10^{-08}$

Table 3.1: Average absolute relative error ς_j per index j (Equation 3.46) in the computation of fully normalised and non-normalised cosinusoidal LWFs in single mode (Equation 3.33) or dual mode (Equation 3.39) for $\theta = 30^\circ$, $\theta = 45^\circ$ and $\theta = 60^\circ$, where the functions are computed in IEEE double precision

tion (3.33) (single mode) or Equation (3.39) (dual mode) to compute the LWFs are shown in Table 3.1 for $\theta = 30^\circ$, $\theta = 45^\circ$ and $\theta = 60^\circ$. For the extremes $\theta = 0^\circ$ and $\theta = 90^\circ$, the values of ς_j are below 10^{-16} for all j up to 32.

It can be seen from Table 3.1 that the error ς_j is very similar for the computation algorithms using Equations (3.33) and (3.39), and although it is not shown in Table 3.1, algorithms based on Equations (3.38) and (3.40) give very similar results. The choice of algorithm can therefore mainly be based on computational speed, which is roughly twice as fast for the dual mode compared to the single mode.

It also follows from Table 3.1 that the error generally increases with increasing index j , and this effect is most profound for the case that $\theta = 60^\circ$. An explanation for this is that a high power of the cosine of θ , as in the left-hand side of Equation (3.28), becomes

very small for θ close to 90° . The right-hand side of Equation (3.28) must naturally result in the same small value, but this requires the relatively large components of different index i to almost cancel out. It is a well-known fact that subtraction of almost equal relatively large values leads to a degradation of accuracy, and this is what causes the larger errors in Table 3.1 for high j and $\theta = 60^\circ$. It is important to realise that this does, however, not mean that the LWFs are of lower accuracy in this case.

Not surprisingly, the accuracy of sinusoidal LWFs \overline{K}_{nm}^{ij} shows a very similar pattern to the cosinusoidal LWFs \overline{F}_{nm}^{ij} , although ‘opposite’ with respect to θ , i.e. the behaviour of the sinusoidal LWFs for $\theta = 30^\circ$ matches that of the cosinusoidal LWFs for $\theta = 60^\circ$ and vice versa. This follows naturally from the basic trigonometric identity that the cosine of an angle equals the sine of its complement.

3.2.5 First- and second-order derivatives of spherical harmonics

The product of first- or second-order derivatives of ALFs or SHFs with respect to θ with the sine or cosine of θ can also be expressed as a weighted summation over ALFs or SHFs of equal order m respectively. These expressions follow directly from well-known relations equating the derivatives of ALFs to a summation over ALFs (e.g., Hobson, 1931; MacRobert, 1967; Abramowitz and Stegun, 1972). The first-order derivative can be expressed as a summation over ALFs of equal degree n , or as a summation over ALFs of equal order m (e.g., Abramowitz and Stegun, 1972). The case with ALFs of equal order m is of most interest, since in this case the ALFs can be replaced by SHFs. The first derivative of the normalised SHF with respect to θ reads

$$\frac{\partial \overline{Y}_{nm}}{\partial \theta} = n \cot \theta \overline{Y}_{nm} - (n + m) \csc \theta \overline{Y}_{n-1,m} \quad (3.48)$$

Multiplying both sides of Equation (3.48) by the sine of co-latitude θ , a weighted summation over two SHFs of equal order m can be obtained

$$\begin{aligned} \sin \theta \frac{\partial \overline{Y}_{nm}}{\partial \theta} &= n \cos \theta \overline{Y}_{nm} - (n + m) \overline{Y}_{n-1,m} \\ &= (n \overline{F}_{nm}^{-1,1} - n - m) \overline{Y}_{n-1,m} + n \overline{F}_{n,m}^{1,1} \overline{Y}_{n+1,m} \end{aligned} \quad (3.49)$$

Multiplication of Equation (3.49) by the cosine of the co-latitude θ yields

$$\sin \theta \cos \theta \frac{\partial \bar{Y}_{nm}}{\partial \theta} = \sum_{i=-2,2}^2 \bar{N}_{nm}^{i,2} \bar{Y}_{n+i,m} \quad (3.50)$$

where the functions $\bar{N}_{nm}^{i,2}$ follow from a combination of the cosinusoidal LWFs

$$\bar{N}_{nm}^{-2,2} = (n\bar{F}_{nm}^{-1,1} - n - m)\bar{F}_{n-1,m}^{-1,1} \quad (3.51)$$

$$\bar{N}_{nm}^{0,2} = (n\bar{F}_{nm}^{-1,1} - n - m)\bar{F}_{n-1,m}^{1,1} + n\bar{F}_{nm}^{1,1}\bar{F}_{n+1,m}^{-1,1} \quad (3.52)$$

$$\bar{N}_{nm}^{2,2} = n\bar{F}_{nm}^{1,1}\bar{F}_{n+1,m}^{1,1} \quad (3.53)$$

By using Equations (3.28) and (3.31), Equation (3.50) can be generalised to hold for sine and cosine to the arbitrary power j . This only holds for odd powers j , since Equation (3.31) only holds for even powers

$$\sin^j \theta \cos^j \theta \frac{\partial \bar{Y}_{nm}}{\partial \theta} = \sum_{i=-2j,2}^{2j} \bar{N}_{nm}^{i,2j} \bar{Y}_{n+i,m} \quad (3.54)$$

where the functions $\bar{N}_{nm}^{i,2j}$ can be computed from any of the same recursive relations as the sinusoidal and cosinusoidal LWFs \bar{F}_{nm}^{ij} and \bar{K}_{nm}^{ij} (Equations 3.33, 3.38, 3.39 and 3.40).

The second-order derivatives of ALFs and SHFs follow from the characteristic differential equation (e.g., Hobson, 1931)

$$\frac{\partial^2 \bar{Y}_{nm}}{\partial \theta^2} = -\cot \theta \frac{\partial \bar{Y}_{nm}}{\partial \theta} - [n(n+1) - m^2 \csc^2 \theta] \bar{Y}_{nm} \quad (3.55)$$

Multiplying this with the square of the sine of co-latitude θ gives

$$\sin^2 \theta \frac{\partial^2 \bar{Y}_{nm}}{\partial \theta^2} = -\sin \theta \cos \theta \frac{\partial \bar{Y}_{nm}}{\partial \theta} - [n(n+1) \sin^2 \theta - m^2] \bar{Y}_{nm} \quad (3.56)$$

Upon implementation of Equations (3.31) and (3.50), it can be seen that second-order derivatives of ALFs or SHFs can also be written as a weighted summation over ALFs or SHFs of equal order m

$$\sin^2 \theta \frac{\partial^2 \bar{Y}_{nm}}{\partial \theta^2} = \sum_{i=-2,2}^2 \bar{R}_{nm}^{i,2} \bar{Y}_{n+i,m} \quad (3.57)$$

where the weights are again solely dependent on the degree n and order m

$$\bar{R}_{nm}^{-2,2} = -\bar{N}_{nm}^{-2,2} - n(n+1)\bar{K}_{nm}^{-2,2} \quad (3.58)$$

$$\overline{R}_{nm}^{0,2} = -\overline{N}_{nm}^{0,2} - n(n+1)\overline{K}_{nm}^{0,2} + m^2 \quad (3.59)$$

$$\overline{R}_{nm}^{2,2} = -\overline{N}_{nm}^{2,2} - n(n+1)\overline{K}_{nm}^{2,2} \quad (3.60)$$

Naturally, Equation (3.57) can be multiplied by the square of the cosine of co-latitude θ by implementation of Equation (3.28). After a rearrangement of summation order (Appendix B) it becomes

$$\sin^2 \theta \cos^2 \theta \frac{\partial^2 \overline{Y}_{nm}}{\partial \theta^2} = \sum_{i=-4,2}^4 \overline{R}_{nm}^{j,4} \overline{Y}_{n+i,m} \quad (3.61)$$

where

$$\overline{R}_{nm}^{j,4} = \sum_{j=-2,2}^2 \overline{R}_{nm}^{j,2} \overline{F}_{nm}^{i-j,2} \quad (3.62)$$

Equation (3.61) can be raised to higher powers of sine and cosine using Equations (3.28) and (3.31). However, this only holds for even powers since Equation (3.31) also holds for even powers only. The relations for the case of higher powers of sine and cosine are not needed in the remainder of this thesis, and will therefore not be shown in detail here.

3.3 The coefficient transformation method

The new relations among SHFs derived in Section 3.2 can be used to construct formulas for the transformation between solid and surface SHCs. However, Equation (3.4) must first be manipulated to obtain a form that allows for the insertion of these new relations. This can be achieved by applying a binomial series expansion to the inverse ellipsoidal radius (r_e^{-1}).

3.3.1 Transformation from solid to surface spherical harmonic coefficients

Inserting the equations for the ellipsoidal radius in spherical polar coordinates (r , θ , λ) from Table 2.1 into Equation (3.4) gives

$$f(r_e, \theta, \lambda) = \sum_{n=0}^{\infty} \left(\frac{c}{\sqrt{1-e^2}} \right)^{n+1} (1 - e^2 \sin^2 \theta)^{\frac{n+1}{2}} \sum_{m=-n}^n \overline{f}_{nm}^R \overline{Y}_{nm}(\theta, \lambda) \quad (3.63)$$

where c is a scale factor between the semi-major axis of the ellipsoid a and the reference sphere R

$$c = \frac{a}{R} \quad (3.64)$$

The term $(1 - e^2 \sin^2 \theta)^{\frac{n+1}{2}}$ can be expanded into a binomial series (see Appendix C)

$$(1 - e^2 \sin^2 \theta)^{\frac{n+1}{2}} = \sum_{k=0}^{\infty} (-1)^k \binom{\frac{n+1}{2}}{k} e^{2k} \sin^{2k} \theta \quad (3.65)$$

The power of the expression on the left-hand side of Equation (3.65) is a natural number for odd n , but it is not for even n . Therefore, for odd n , the summation on the right-hand side of Equation (3.65) will have a limited number of non-zero terms, whereas for even n , the number of non-zero terms is infinite.

The series in Equation (3.65) will always converge, since $e^2 \sin^2 \theta < 1$ (see Appendix C). Also, because $e^2 \sin^2 \theta$ is always positive, the series will always be alternating. Then, according to Equation (C-11), the absolute values of the terms in the binomial series will start to decrease if

$$k > \frac{(n+1)e^2 \sin^2 \theta - 2}{2(1 + e^2 \sin^2 \theta)} \quad (3.66)$$

The convergence is thus slowest for points at the equator ($\theta = \frac{\pi}{2}$), but more importantly it is slower for increasing degrees n . Figure 3.5 shows the convergence rates in this case for degrees n up to 2160. The absolute values of the terms in the binomial series (Equation 3.65) start to decrease above the dashed line in Figure 3.5, as follows from Equation (3.66).

The truncation error ς of the binomial series can be approximated using Equation (C-12)

$$\varsigma < \frac{(n+1)^k}{2^k k!} e^{2k} \sin^{2k} \theta \quad (3.67)$$

As stated, the error is highest for points at the equator. Figure 3.6 shows how many terms need to be taken into account to ensure a relative accuracy of 1, e^2 or e^4 for degrees n up to 2160. The number of terms needed to achieve a given accuracy rises almost linearly with the degree n , and the rate of this rise is very similar for any desired accuracy. For the very high degree $n = 2160$, 16 terms are needed to obtain a relative accuracy of at least one, 22 terms are needed to obtain a relative accuracy of e^2 and 27 terms are needed to obtain a relative accuracy of at least e^4 .

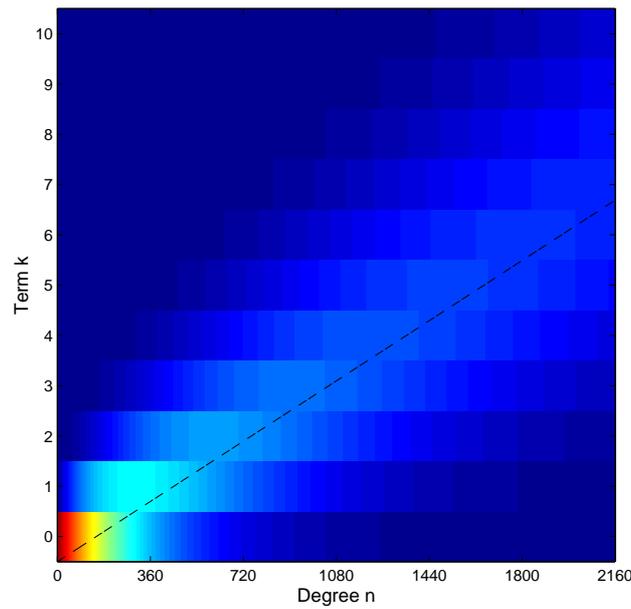


Figure 3.5: The size of the first ten terms in the binomial series in Equation (3.65) relative to the total sum in percentage for degrees $0 \leq n \leq 2160$. The dashed line indicates the boundary described by Equation (3.66), increased by 0.5 so that it crosses the border between two terms where those terms are equal.

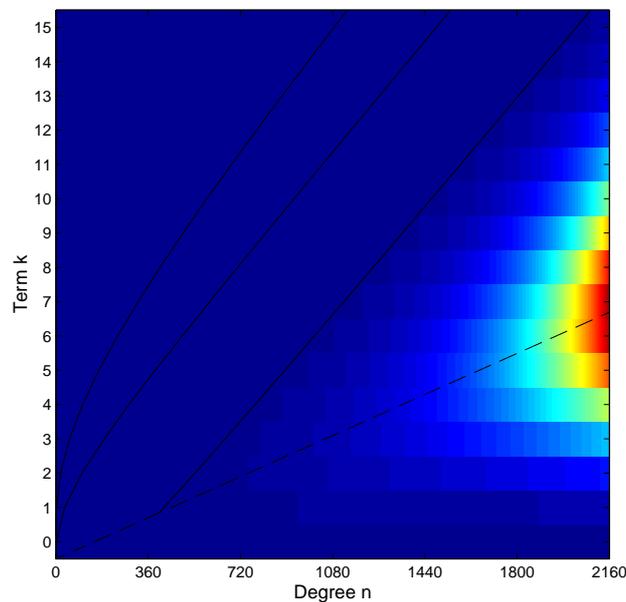


Figure 3.6: The absolute size of the first fifteen terms in the binomial series in Equation (3.65) for degrees $0 \leq n \leq 2160$. The dashed line indicates the boundary described by Equation (3.66), increased by 0.5 so that it crosses the border between two terms where those terms are equal. The solid lines indicate until which term the series should be evaluated to ensure a relative accuracy of at least e^4 , e^2 and 1 respectively (from left to right).

Inserting the binomial series in Equation (3.65) into the spherical harmonic expansion in Equation (3.63) gives

$$f(r_e, \theta, \lambda) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \alpha_{nk} \sin^{2k} \theta \sum_{m=-n}^n \bar{f}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (3.68)$$

where

$$\alpha_{nk} = \left(\frac{c}{\sqrt{1-e^2}} \right)^{n+1} (-1)^k \binom{\frac{n+1}{2}}{k} e^{2k} \quad (3.69)$$

The term α_{nk} depends upon the degree n , the index of the binomial series k , the square of the eccentricity of the ellipsoid e^2 and the scale factor c that relates the reference sphere R to the semi-major axis of the ellipsoid a (Equation 3.64). It is thus independent of the longitude and latitude, i.e., the position of the point where the function f is to be determined. Only the sine of the co-latitude θ and the SHFs \bar{Y}_{nm} in Equation (3.68) depend upon position.

Now, the sinusoidal LWFs \bar{K}_{nm}^{ij} (Equation 3.31) can be inserted into Equation (3.68), because it contains the product of an even power $\sin \theta$ with a SHF

$$f(r_e, \theta, \lambda) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \alpha_{nk} \sum_{m=-n}^n \bar{f}_{nm}^R \sum_{i=-k}^k \bar{K}_{nm}^{2i,2k} \bar{Y}_{n+2i,m}(\theta, \lambda) \quad (3.70)$$

This introduces yet another summation, but now all dependency on position is centered in the SHFs \bar{Y}_{nm} . This is very useful, since it allows for a direct comparison with a series of surface SHFs (Equation 3.13) relating the surface and solid SHCs by an expression that is independent of latitude. First, a rearrangement of the summation order (see Appendix B) is required to ensure that the SHFs do not depend on summation index i

$$f(r_e, \theta, \lambda) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \sum_{i=-k}^k \alpha_{n-2i,k} \sum_{m=-n}^n \bar{f}_{n-2i,m}^R \bar{K}_{n-2i,m}^{2i,2k} \bar{Y}_{nm}(\theta, \lambda) \quad (3.71)$$

Comparing Equations (3.71) and (3.13), a relation between solid and surface SHCs can be found. Since all dependence on latitude is within the SHFs, which form a set of orthogonal base functions, the coefficients of each pair of degree n and order m can be equated individually

$$\bar{f}_{nm}^{r_e} = \sum_{k=0}^{\infty} \sum_{i=-k}^k \alpha_{n-2i,k} \bar{K}_{n-2i,m}^{2i,2k} \bar{f}_{n-2i,m}^R \quad (3.72)$$

Reversing the summations over indices i and k (see Appendix B) gives the final transformation formula as an infinite weighted summation

$$\bar{f}_{nm}^{r_e} = \sum_{i=-\infty}^{\infty} \lambda_{nmi} \bar{f}_{n-2i,m}^R \quad (3.73)$$

where the weights also follow from an infinite summation

$$\lambda_{nmi} = \sum_{k=|i|}^{\infty} \alpha_{n-2i,k} \bar{K}_{n-2i,m}^{2i,2k} \quad (3.74)$$

Although the derivation of these weights is somewhat cumbersome, the presentation of the final formula (Equation 3.73) as a simple weighted summation is very convenient for practical application. However, this equation does not yet provide an outright solution to the the ellipsoidal Dirichlet BVP, because this would require the computation of solid SHCs from surface SHCs, i.e., the inverse to Equation (3.73).

3.3.2 Transformation from surface to solid spherical harmonics

The transformation from surface to solid SHCs can be found through application of the inverse of Equation (3.73). This inverse can be found if the summation for all coefficients up to a certain degree n and order m is written in matrix form.

$$\mathcal{F}^{r_e} = \Lambda^{r_e R} \mathcal{F}^R \quad (3.75)$$

The vectors \mathcal{F}^{r_e} and \mathcal{F}^R contain all solid and surface SHCs up to degree and order n respectively. If the coefficients are ordered in an efficient way, the matrix $\Lambda^{r_e R}$ becomes block-diagonal (see Figure 3.7). The coefficients in the vectors are first sorted by order m , and subsequently sorted by even and odd degrees n to achieve the block-diagonality. Since the matrix $\Lambda^{R r_e}$ is square, Equation (3.75) can easily be inverted

$$\mathcal{F}^R = \Lambda^{R r_e} \mathcal{F}^{r_e} \quad (3.76)$$

where $\Lambda^{R r_e}$ is the matrix inverse of $\Lambda^{r_e R}$

$$\Lambda^{R r_e} = (\Lambda^{r_e R})^{-1} \quad (3.77)$$

$$\begin{pmatrix} \overline{f}_{0,0}^e \\ \overline{f}_{2,0}^e \\ \overline{f}_{4,0}^e \\ \vdots \\ \overline{f}_{n,0}^e \\ \overline{f}_{1,0}^e \\ \overline{f}_{3,0}^e \\ \vdots \\ \overline{f}_{n-1,0}^e \\ \overline{f}_{2,1}^\theta \\ \overline{f}_{4,1}^\theta \\ \vdots \\ \overline{f}_{n,1}^e \\ \vdots \\ \overline{f}_{n-1,n-1}^e \\ \overline{f}_{n,n}^e \end{pmatrix} = \begin{pmatrix} \lambda_{0,0,0} & \lambda_{0,0,-1} & \lambda_{0,0,-2} & \cdots & \lambda_{0,0,-\frac{n}{2}} \\ \lambda_{2,0,1} & \lambda_{2,0,0} & \lambda_{2,0,-1} & \cdots & \lambda_{2,0,-\frac{n-2}{2}} \\ \lambda_{4,0,2} & \lambda_{4,0,1} & \lambda_{4,0,0} & \cdots & \lambda_{4,0,-\frac{n-4}{2}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{n,0,\frac{n}{2}} & \lambda_{n,0,\frac{n-2}{2}} & \lambda_{n,0,\frac{n-4}{2}} & \cdots & \lambda_{n,0,0} \\ \lambda_{1,0,0} & \lambda_{1,0,-1} & \cdots & \lambda_{1,0,-\frac{n-2}{2}} \\ \lambda_{3,0,1} & \lambda_{3,0,0} & \cdots & \lambda_{3,0,-\frac{n-4}{2}} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n-1,0,\frac{n-2}{2}} & \lambda_{n-1,0,\frac{n-4}{2}} & \cdots & \lambda_{n-1,0,0} \end{pmatrix} \begin{pmatrix} \overline{f}_{0,0}^R \\ \overline{f}_{2,0}^R \\ \overline{f}_{4,0}^R \\ \vdots \\ \overline{f}_{n,0}^R \\ \overline{f}_{1,0}^R \\ \overline{f}_{3,0}^R \\ \vdots \\ \overline{f}_{n-1,0}^R \\ \overline{f}_{2,1}^R \\ \overline{f}_{4,1}^R \\ \vdots \\ \overline{f}_{n,1}^R \\ \vdots \\ \overline{f}_{n-1,n-1}^R \\ \overline{f}_{n,n}^R \end{pmatrix}$$

Figure 3.7: Full matrix form of Equation (3.75) for SHCs to degree and order n

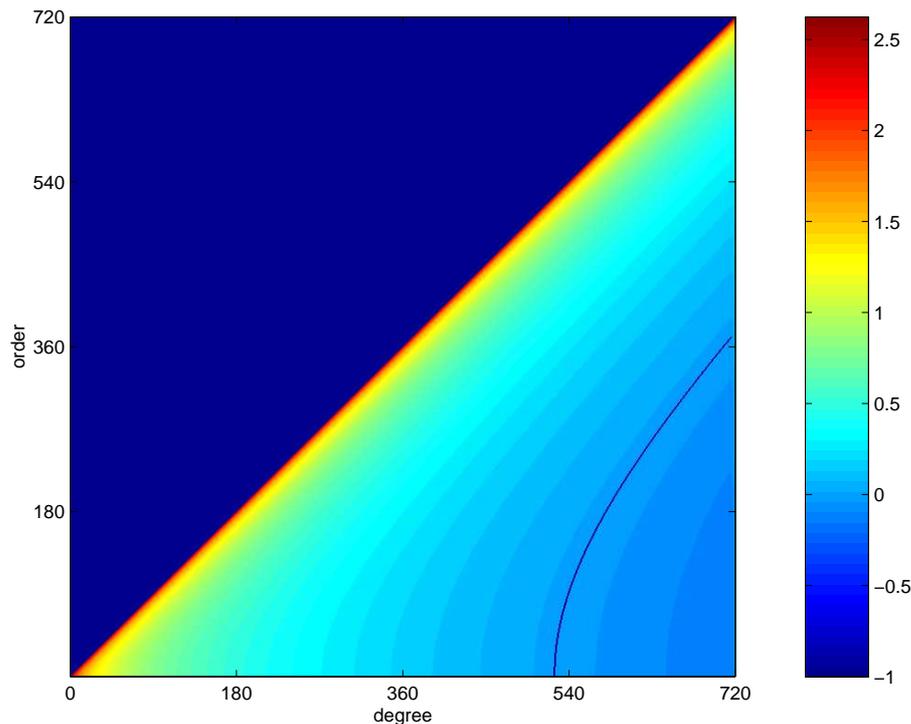


Figure 3.8: Logarithm of the ratio between $|\lambda_{nm0}|$ and $\sum_{i=-50, i \neq 0}^{50} |\lambda_{nmi}|$ for all degrees and orders up to 720, where the line starting at degree 520 indicates for which pairs of degree and order the ratio is closest to unity

The fact that matrix $\Lambda^{r_e R}$ is block-diagonal means that the inversion can be performed in a numerically efficient way, if the matrix is well-conditioned, which is always the case if the matrix is diagonally dominant, i.e., if

$$|\lambda_{nm0}| \geq \sum_{i=-\infty, i \neq 0}^{\infty} |\lambda_{nmi}| \quad (3.78)$$

It can be seen from Figure 3.8 that this will not be the case for all pairs of degree n and order m . If the spherical harmonic expansions exceed degree and order 520, diagonal dominance is lost and it must be examined separately whether the matrix is well-conditioned. It can be expected from the matrix structure that the matrix becomes less well-conditioned with increasing degree n . The increasing importance of the off-diagonal weights is, naturally, a direct effect of the behaviour of the binomial series (Equation 3.65), which is shown in Figures 3.5 and 3.6. The weights λ_{nmi} where $i \neq 0$, increase with increasing degree n due to the presence of n inside the binomial coefficient in Equation (3.65).

An alternative to the rigorous matrix inversion, which despite the block-diagonal structure is numerically time-consuming, is an iterative approach, where the surface SHCs are used as a first estimate for the solid SHCs

$$\bar{f}_{nm}^{R(k)} = \frac{1}{\lambda_{nm0}} \left(\bar{f}_{nm}^{r_e} - \sum_{i=-\infty, i \neq 0}^{\infty} \lambda_{nmi} \bar{f}_{n-2i,m}^{R(k-1)} \right) \quad (3.79)$$

with as an initial approximation

$$\bar{f}_{nm}^{R(0)} = \bar{f}_{nm}^{r_e} \quad (3.80)$$

This is, in principle, the Jacobi iteration method for the solution of a linear system of equations (e.g., Strang, 1986). The convergence of this iterative solution can be accelerated by using the Gauss-Seidel iteration or the successive overrelaxation (SOR) methods (e.g., Strang, 1986). However, convergence can only be guaranteed for any initial estimate of the coefficients \bar{f}_{nm}^R if the matrix $\Lambda^{r_e R}$ is diagonally dominant. Thus, below approximately degree 520, the iterative method will converge, and a rigorous matrix inversion can be avoided. However, from degree 520 onwards, divergence may occur for coefficients of low order m .

It is, however, not certain that the iteration will diverge when the diagonal dominance is lost. In this case, the availability of a close initial approximation may become crucial. The convergence of the Jacobi iteration with the initial estimate given in Equation (3.80), as well as the overall numerical performance of the coefficient transformation, is determined for the case of the disturbing potential of the Earth's gravity field using a closed-loop simulation, as follows.

3.4 Numerical closed-loop simulation

The coefficient transformation solution to the ellipsoidal Dirichlet BVP is tested numerically using a closed-loop simulation (e.g., Ledin, 2001). In this test, the EGM96 global geopotential model (Lemoine et al., 1998) to degree and order 360 is used as input data. Naturally, a more recent global geopotential model, such as GGM02 (Tapley et al., 2005) or EIGEN-02S (Reigber et al., 2005), could also have been used. However, the accuracy of the geopotential coefficients is of little concern here, since

this test merely serves to confirm the validity of the BVP solution. It is thus used as a synthetic test data set (cf. Tziavos, 1996), which does not need to be an exact representation of reality, as long as it is reasonably realistic (e.g., Pail, 1999; Novák et al., 2001; Claessens, 2003; Holmes, 2003). EGM96 is assumed to provide a realistic representation of the Earth's disturbing potential.

EGM96 consists of a set of solid SHCs of the Earth's gravity potential W up to degree and order 360. Solid SHCs of the disturbing potential T can be obtained by subtraction of the solid SHCs of a reference gravity field. The GRS80 reference field, defined in Table 2.5, is here used for this purpose. For the maximum degree and order of 360, the iteration in Equation (3.79) for the inverse transformation should in theory converge, as follows from Figure 3.8. The convergence beyond degree 360 is also tested, using the GPM98A geopotential model to $n = 1800$ (Wenzel, 1998).

3.4.1 Computation of surface spherical harmonic coefficients

The solid SHCs of the disturbing potential comprising the EGM96 coefficients can be transformed into a set of surface SHCs defined on the ellipsoid using Equation (3.73)

$$\bar{T}_{nm}^{r_e} = \sum_{i=-\infty}^{\infty} \lambda_{nmi} \bar{T}_{n-2i,m}^R \quad (3.81)$$

where the weights λ_{nmi} can be computed using Equation (3.74).

This new transformation formula is tested by a synthesis-analysis procedure. First, a grid of values of the disturbing potential is computed on the surface of the ellipsoid from the solid SHCs, using a spherical harmonic synthesis

$$T(r_e, \theta, \lambda) = \sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^{n+1} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (3.82)$$

where the radius of the reference sphere is equal to the semi-major axis of the GRS80 ellipsoid. The grid size chosen here is $30'$, to coincide with the spectral resolution of EGM96 up to degree $n = 360$. This is to ensure that both the spherical harmonic expansion and the discretised spatial grid contain information up to the same frequency.

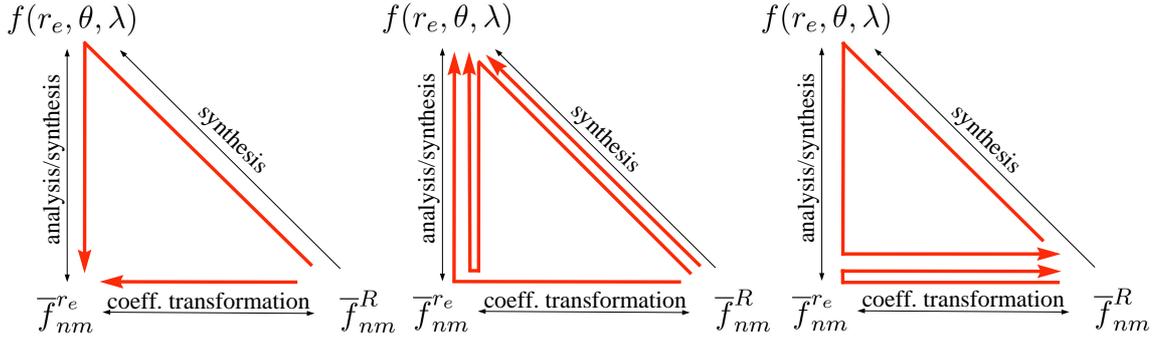


Figure 3.9: Numerical tests comparing the surface SHCs (left), the function values on the ellipsoid (centre) and the solid SHCs (right) via different ways of computation

Subsequently, a spherical harmonic analysis is performed to expand the grid of function values into a series of surface SHFs. This step involves a discretised version of Equation (2.85)

$$\begin{aligned} \bar{T}_{nm}^{re} &= \frac{1}{4\pi} \int_{\sigma} T(r_e, \theta, \lambda) \bar{Y}_{nm}(\theta, \lambda) d\sigma \\ &\approx \frac{1}{4\pi} \sum_{i=0}^{180} \sum_{j=1}^{360} T(r_{e,i}, \theta_i, \lambda_j) \bar{Y}_{nm}(\theta_i, \lambda_j) \end{aligned} \quad (3.83)$$

Naturally, despite the band-limited nature of the test data, the synthesis-analysis procedure is not error-free. Numerical tests by Rapp (1986) have shown that a synthesis-analysis procedure cannot exactly reproduce the same set of coefficients, because the orthogonality of the SHFs no longer holds in discrete form, and the discrepancies can reach several centimetres in terms of geoid height.

The complete test setup is displayed in Figure 3.9. First, a comparison is made in the spectral domain between the surface SHCs resulting from the transformation (Equation 3.81) and the surface SHCs resulting from the synthesis-analysis (Equation 3.83). To simplify the comparison, degree variances t_n (signal power) were computed for both sets of coefficients

$$t_n = \sum_{m=-n}^n (\bar{T}_{nm}^{re})^2 \quad (3.84)$$

The degree variances of the surface SHCs and the differences between the two sets of coefficients are shown in Figure 3.10. It can first be seen that the degree variances of the test data (black line) generally decrease with increasing degree n , due to the

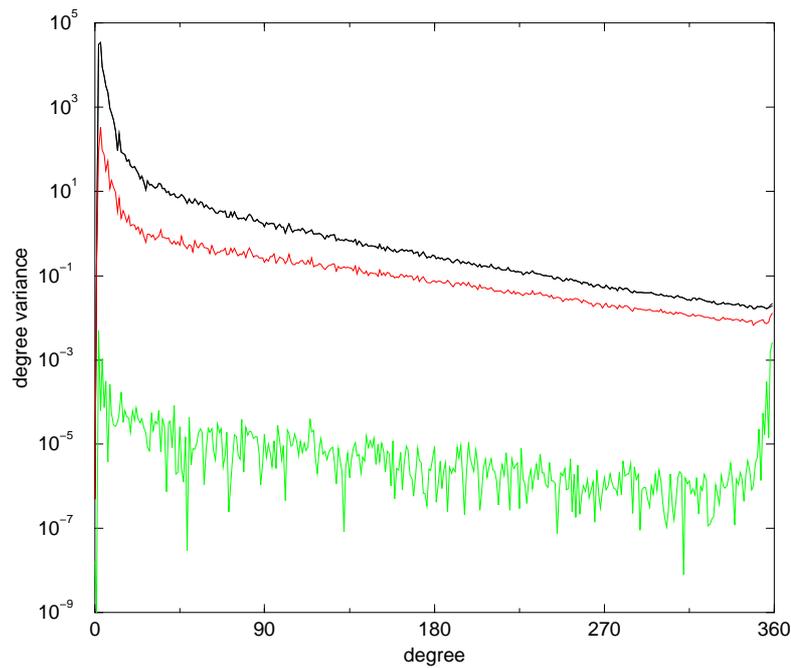


Figure 3.10: Degree variances of surface SHCs $\overline{T}_{nm}^{r_e}$ (black line), their differences with the degree variances of the solid SHCs \overline{T}_{nm}^R (red line), and the differences between the surface SHCs from the synthesis-analysis procedure (Equation 3.83) and those from the new transformation formula (Equation 3.81) (green line), where the summation over i was performed from -20 to 20

smoothness of the external potential surfaces of the Earth's gravity field. Therefore, the magnitude of the absolute differences between the solid spherical harmonic degree variances and the surface spherical harmonic degree variances from data on the ellipsoid (red line in Figure 3.10) decreases as well. However, the gap between both lines also decreases, and this indicates that the relative difference between both sets of coefficients increases. For degree 360, the relative difference between the solid degree variances and the surface degree variances is approximately 50%, which shows the inaccuracy of the constant radius approximation.

When the solid SHCs are transformed using Equation (3.81), the agreement with the surface SHCs becomes much better, which can be seen from the fact that the green line in Figure 3.10 is below the red line. In this transformation, the summation over i was performed from -20 to 20 to ensure that the truncation error is smaller than the numerical round-off error. The absolute differences between the degree variances com-

puted from the coefficients \bar{T}_{nm}^{re} from Equation (3.81) and Equation (3.83) are generally five orders of magnitude smaller than the degree variances themselves, resulting in a relative error of approximately 0.001%. Only for degrees very close to the maximum degree of 360, the errors are larger, generally starting to increase from degree ~ 345 onwards. This is not surprising, since the transformation formula (Equation 3.81) requires solid SHCs \bar{T}_{nm}^R of lower and higher degree n . Since coefficients beyond degree 360 are not available, they are set to zero, which causes large errors for coefficients of degrees close to 360. One should therefore always be aware of the fact that the SHCs obtained from the transformation method will be less accurate in the vicinity of the maximum degree of the input coefficients.

Another feature that can be noticed in Figure 3.10 is the strong fluctuation of the green line, which indicates that the accuracy of the transformation method varies widely per degree n . This is somewhat surprising, since the degree variances comprise a summation over a number of SHCs, which would expectedly smoothen out the occasional numerical spike. The fluctuations can partly be explained by a peculiar pattern in the relative accuracy of the SHCs themselves (see Figure 3.11). It can be seen that SHCs for which $n + m$ is odd are of very high accuracy, with relative errors in the order of 10^{-11} , whereas SHCs for which $n + m$ is even display lower accuracies, with relative errors in the order of 10^{-4} . This behaviour cannot be explained by numerical inaccuracies in the transformation from solid to surface SHCs, but are more likely to be caused by errors in the synthesis-analysis procedure.

To investigate the influence of the differences between the SHCs on the function values on the ellipsoid, a comparison in the spatial domain is performed as well (see centre Figure 3.9), where the grid values from the direct synthesis of the solid SHCs (Equation 3.82) are compared to grid values derived from a synthesis of the surface SHCs that were computed via the coefficient transformation

$$T(r_e, \theta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \bar{T}_{nm}^{re} \bar{Y}_{nm}(\theta, \lambda) \quad (3.85)$$

The values of the disturbing potential on the ellipsoid from Equation (3.82) are shown in Figure 3.12 and the differences between these values and the disturbing potential from Equation (3.85) are shown in Figure 3.13 for the region of Indonesia. This region

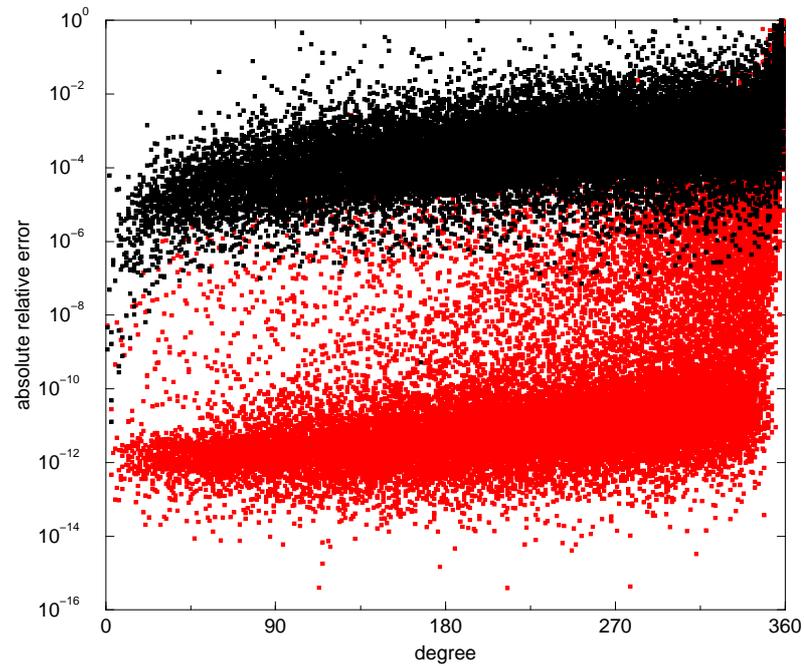


Figure 3.11: Absolute relative difference between surface harmonic coefficients \overline{T}_{nm}^{re} computed using the synthesis-analysis procedure (Equation (3.83)) and the new transformation formula (Equation 3.81) for $n+m$ is even (black) and $n+m$ is odd (red), where the summation over i was performed from -20 to 20

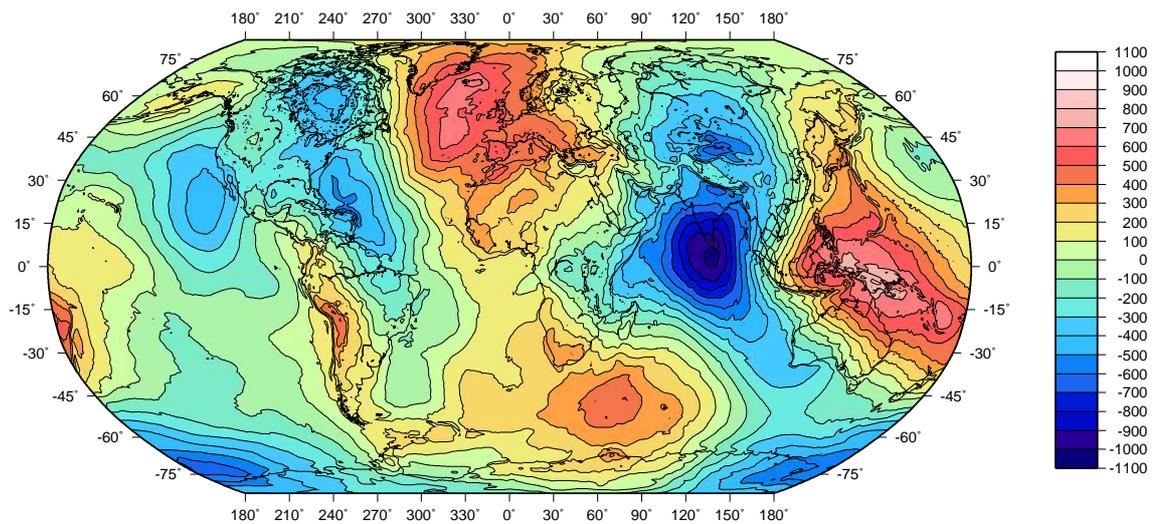


Figure 3.12: Disturbing potential of the Earth's gravity field from the EGM96 geopotential model and GRS80 reference gravity field (units in m^2s^{-2} ; Robinson projection)

	comparison 1	comparison 2
minimum difference	-0.625	-0.060
maximum difference	0.641	0.067
average difference	$-8.302 \cdot 10^{-6}$	$-5.872 \cdot 10^{-5}$
average absolute difference	0.111	$6.119 \cdot 10^{-4}$

Table 3.2: Global statistics of the comparison between the values of the disturbing potential on the ellipsoid from a spherical harmonic synthesis of surface SHCs obtained via a coefficient transformation and 1) a spherical harmonic synthesis of the EGM96 geopotential coefficients up to degree and order 360, and 2) a spherical harmonic synthesis of the surface coefficients obtained via a synthesis-analysis procedure up to degree and order 340 (in m^2s^{-2})

was selected because the largest differences occur around the equator. Some global statistics of this comparison can be found in Table 3.2. Not surprisingly, the differences in Figure 3.13 are completely dominated by short-wavelength errors resulting from the large errors in the SHCs of degree close to the maximum degree of 360. In addition, it should be noted that a comparison of this kind will always contain (small) errors due to the fact that a solid harmonic expansion and a surface harmonic expansion up to the same degree and order do not contain the same information, meaning that the omission errors will not be equal.

To avoid this effect, the grid values resulting from a spherical harmonic synthesis of the surface SHCs that were obtained via a transformation are compared to a synthesis of the surface SHCs that were obtained via the synthesis-analysis procedure. Secondly, in order to avoid the obvious errors in the SHCs close to the maximum degree of 360, the syntheses are performed up to degree and order 340. The results of this comparison are shown in Figure 3.14 for the region of Indonesia, and the global statistics are shown in Table 3.2. It can be seen from Table 3.2 that especially the average absolute difference has decreased considerably with respect to the first comparison, and is approximately equivalent to $60 \mu\text{m}$ in geoid height. The absolute maximum error is considerably larger at 7 mm in geoid height, but this is still relatively insignificant with respect to the accuracy of gravity field modelling in practice, and falls within the global aim to compute geoid heights accurate to 1 cm (e.g., Rapp, 1997a). Curiously, the maximum

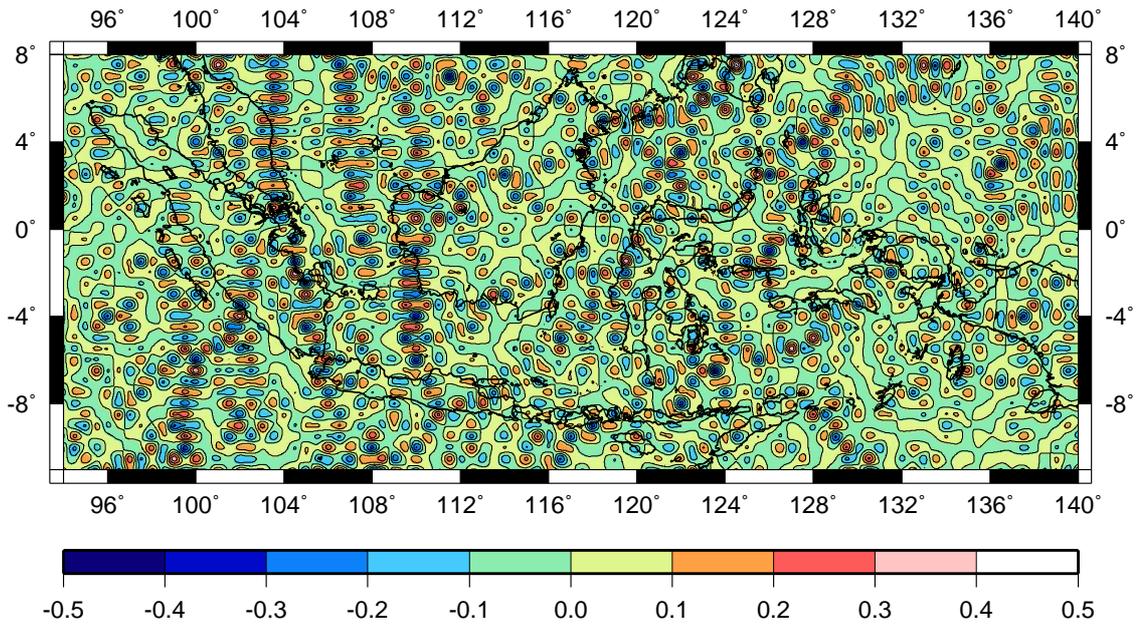


Figure 3.13: Differences between the disturbing potential from a spherical harmonic synthesis of the EGM96 geopotential coefficients and surface spherical harmonic coefficients obtained via a coefficient transformation up to degree and order 360 for the region of Indonesia (units in m^2s^{-2} ; Mercator projection)

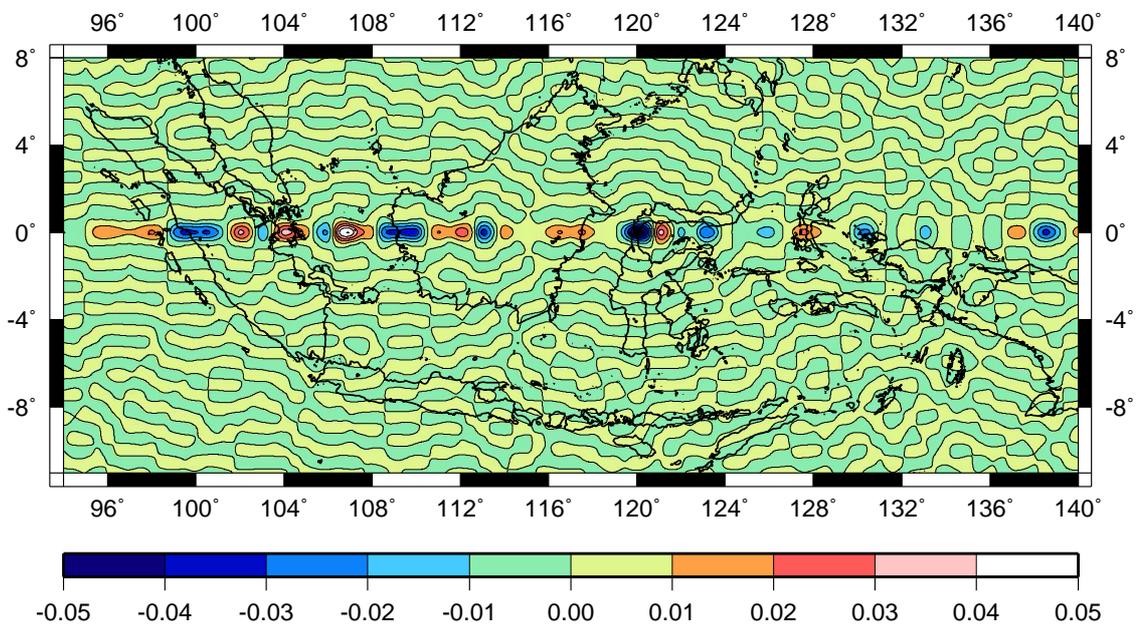


Figure 3.14: Differences between the disturbing potential from a spherical harmonic synthesis of surface spherical harmonic coefficients obtained via synthesis-analysis procedure and via a coefficient transformation up to degree and order 340 for the region of Indonesia (units in m^2s^{-2} ; Mercator projection)

error occurs exactly at the equator, and it can be seen from Figure 3.14 that there is a distinct band of relatively large errors at the equator, whereas the errors in all other areas are much smaller.

The difference in accuracy between the SHCs for which $n + m$ is odd and those for which $n + m$ is even can now be explained. The SHFs for which $n + m$ is odd are all zero at the equator, and therefore the SHCs for which $n + m$ is odd have no influence on function values at the equator. On the other hand, SHCs for which $n + m$ is even do not influence function values on the equator only, but all over the globe. It can therefore be concluded that the larger differences at the equator cannot be caused by numerical errors in the transformation from solid to surface SHCs, because these would result in larger errors at all points on the ellipsoid. They must be caused by errors in the synthesis-analysis procedure instead.

Notwithstanding the larger differences at the equator, the magnitude of the differences is small, which proves beyond doubt that the coefficient transformation procedure is an adequate method for spherical harmonic models of the Earth's gravity field up to degree and order 340. The degradation of accuracy with increasing degree n shown in Figures 3.10 and 3.11 is small, suggesting that the method may also be applied successfully in spherical harmonic series of higher degree and order, although expansions of very high degree and order are likely to suffer larger inaccuracies. Moreover, it can be derived from the short-wavelength nature of the errors in the spatial domain that functions with higher power in the higher frequencies, such as for example gravity disturbances or gravity anomalies, will perform relatively worse in the spatial domain.

3.4.2 Computation of solid spherical harmonic coefficients

In Section 3.4.1, the 'forward' transformation from solid SHCs to surface SHCs is verified numerically with the use of the EGM96 geopotential model. The 'inverse' transformation is more complicated, since it requires an inversion of the weight matrix (Equations 3.76 and 3.77), or alternatively, an iterative procedure (Equations 3.79 and 3.80).

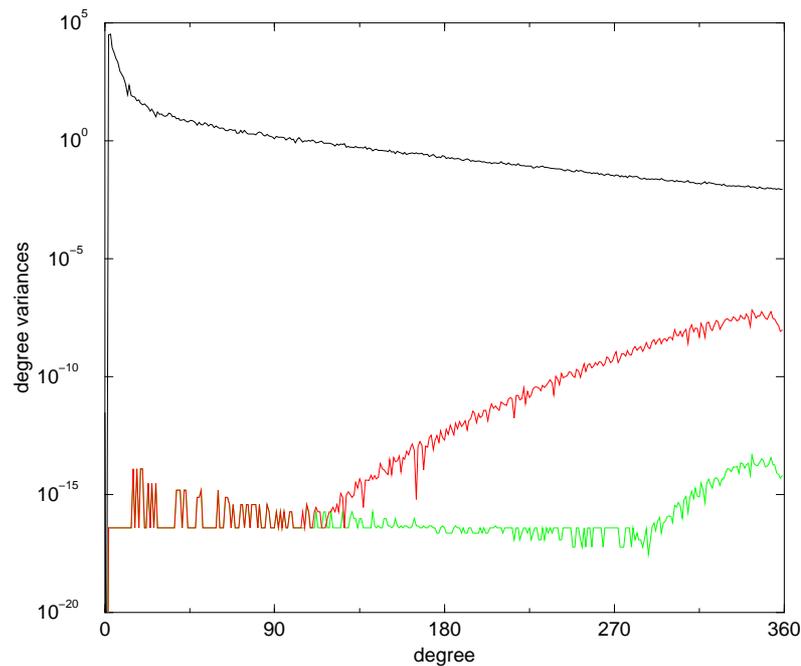


Figure 3.15: Degree variances of solid SHCs \overline{T}_{nm}^R (black line), and their absolute differences with the degree variances of the solid SHCs \overline{T}_{nm}^R obtained after a forward and backward transformation, where the backward transformation is performed with 20 iterations (red line) and 50 iterations (green line). The summation over i was performed from -20 to 20 and the degree variance differences of zero were set to 10^{-32}

The iterative procedure is computationally less demanding than the matrix inversion. The convergence of the iteration can, however, not be guaranteed beyond degree 520 (see Figure 3.8). It will therefore be important that the initial approximation is a close estimate. Unfortunately, the surface SHC that is used as the initial approximation will become a worse estimate for increasing degree n , which follows directly from the increasing importance of the off-diagonal weights.

The inverse transformation is first investigated using the EGM96 spherical harmonic expansion of the Earth's disturbing potential field. Figure 3.15 shows the differences between the degree variances of the original EGM96 geopotential model, and the degree variances after a forward and backward transformation (Equations 3.73 and 3.79). It can be seen that convergence is achieved for all coefficients up to degree and order 360, demonstrating the functionality of the iterative backward transformation. The number of iterations required to obtain an optimal relative accuracy of 10^{-16} increases

for increasing degree n . A total of 50 iterations suffices to obtain this accuracy using the simple Jacobi iteration.

When backward-transforming the surface SHCs that were obtained from a synthesis-analysis procedure, the relative differences between the resulting solid SHCs and the original SHCs shows a very similar pattern to Figure 3.11, i.e., the coefficients for which $n + m$ is odd are of significantly higher accuracy than the coefficients for which $n + m$ is even. A spatial comparison of the solid SHCs \overline{T}_{nm}^R also shows a similar pattern to the spatial comparison of the surface coefficients \overline{T}_{nm}^{re} , which is not surprising since the validity of the backward transformation is already confirmed by Figure 3.15.

It should be noted that the computational efficiency of the coefficient transformation is much higher than that of the synthesis-analysis procedure. Up to degree and order 360, the inverse transformation with 50 iterations was of the order of 10^2 times faster than the synthesis-analysis procedure, although this does of course strongly depend on the algorithms used for the spherical harmonic synthesis and analysis. Not surprisingly, the forward transformation is even faster than the inverse, and a transformation from solid to surface SHCs up to degree and order 360 can be performed in several seconds on a Pentium IV 2.4 GHz PC.

Finally, the convergence beyond degree 360 is investigated by a forward and backward transformation using the GPM98A geopotential model to degree and order 720 with 120 iterations. The difference between the original and the resulting degree variances is shown in Figure 3.16. It can be seen that the accuracy of the transformed SHCs starts to decrease around degree 300, just as it does in Figure 3.15 for the case of 50 iterations. Thus, the extra iterations performed in the creation of Figure 3.16 do not enhance the accuracy of the coefficients.

The order of magnitude of the coefficient differences in Figure 3.16 increases steadily from degree 300 to approximately degree 520, where the relative error of the transformed SHCs is in the order of 1%. Then, a sudden change takes place. The differences start to increase more rapidly, crossing the degree variances of the geopotential model itself, which means that the accuracy of the transformed coefficients becomes

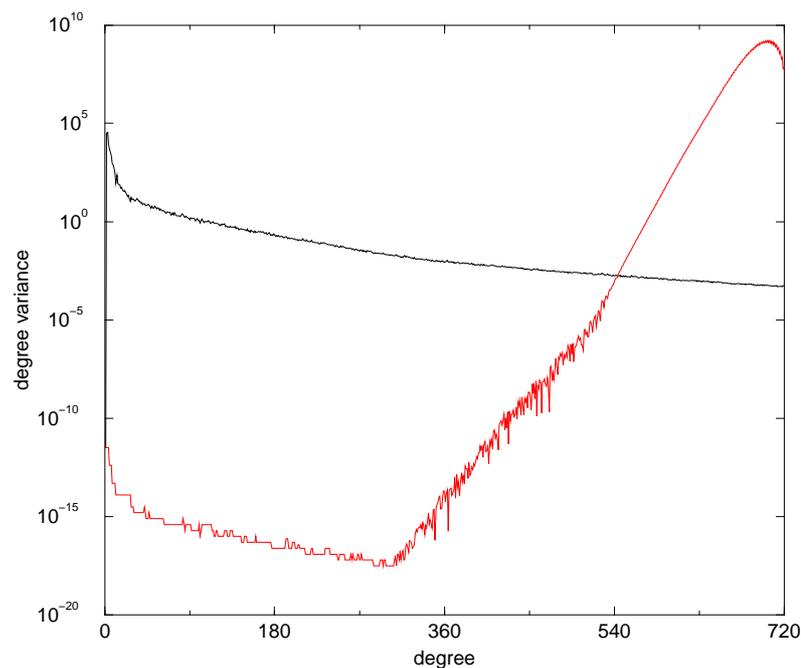


Figure 3.16: Degree variances of solid SHCs \overline{T}_{nm}^R (black line), and their absolute differences with the degree variances of the solid SHCs \overline{T}_{nm}^R obtained after a forward and backward transformation, where the backward transformation is performed with 120 iterations and the summation over i was performed from -50 to 50

very poor.

In addition, the differences fluctuate significantly less beyond degree 520. It is no coincidence that this sudden change takes place at degree 520, which was earlier identified as the degree where diagonal dominance of the weights is lost, and the backward transformation might diverge for SHCs of low order m . Obviously, divergence of the iterative procedure in the backward transformation of low order SHCs causes the decreased accuracy of the degree variances. In conclusion, the iterative backward transformation in Equation (3.79) cannot be used beyond degree 520, and loss of accuracy already occurs from degree 300 onwards, as the diagonal dominance of the weight functions λ_{nmi} weakens. The inverse transformation can nevertheless yield highly accurate results if a spherical harmonic expansion to degree 360 is sufficient.

3.5 Summary

In this Chapter, a new solution to the ellipsoidal Dirichlet BVP was derived in the framework of spherical harmonic expansions. Solid SHCs of a harmonic function can be computed from function values on an oblate ellipsoid of revolution via a transformation from surface SHCs. The transformation between surface and solid SHCs consists of a weighted summation, where the derivation of the weights relies heavily on new relations among ALFs and SHFs. These new relations express the product of an ALF or SHF with an arbitrary power of the sine or cosine of co-latitude θ as a summation over ALFs or SHFs of equal order m respectively, multiplied by the herein-called LWFs.

Several numerical comparisons have confirmed the validity of the transformation from solid to surface SHCs and its inverse for an expansion of the Earth's disturbing potential for degree and order ≤ 360 . The errors in SHCs for which $n+m$ is even were significantly larger than those in SHCs for which $n+m$ is odd. However, this effect was shown to be completely caused by numerical errors at the equator in the synthesis-analysis test procedure, and is thus not caused by the new transformation formulas.

If the influence of the expected errors in the SHCs of degrees close to the maximum degree of the expansion are omitted, SHCs can be computed with an average relative accuracy in the order of 10^{-11} . A degradation of accuracy of approximately 1 order of magnitude for an increase in degree n of ~ 200 can be observed. It should, however, be noted that the numerical errors are expected to be higher for an ellipsoid with higher eccentricity than the Earth, and for coefficients of higher degree. The convergence of the iterative inverse transformation can be guaranteed up to degree and order 520.

4. SPECTRAL RELATIONS BETWEEN GRAVITY FIELD QUANTITIES: THE ELLIPSOIDAL MEISSL SCHEME

The relations among different gravity field quantities are much simpler in the spectral domain than they are in the space domain, as pointed out in Meissl (1971). However, the relations derived there only hold in spherical approximation. When the ellipticity of the Earth is taken into account, a representation of gravity quantities as a surface spherical harmonic expansion is still possible. The spectral relations between them become more complicated, but they exist nevertheless. In this Chapter, the spectral relationships between various gravity quantities on the ellipsoid are derived for the first time. These allow for the solution of the ellipsoidal Neumann and second-order BVPs, as well as the second and third geodetic BVPs. Finally, the derived relations are combined in a new model, the ellipsoidal Meissl scheme, as a generalisation of the spherical Meissl scheme devised by Rummel and Van Gelderen (1992, 1995) and extended by, e.g., Grafarend (2001) and Nutz (2002).

4.1 Relation between the disturbing potential and its normal derivative

This Section describes the derivation of the spectral relations between the SHCs of the disturbing potential and that of its normal derivative at the ellipsoid. This relation is very straight-forward in spherical approximation, where the normal derivative is simply the radial derivative, but is more complicated for the ellipsoidal case.

4.1.1 The radial derivative of the disturbing potential

The radial derivative of the fully normalised solid spherical harmonic expansion (Equation 2.81) can be found by a simple differentiation. The radial derivative of the dis-

turbing potential on an ellipsoid with geocentric radius r_e can thus be written in terms of solid SHCs \bar{T}_{nm}^R

$$\left. \frac{\partial T}{\partial r} \right|_{r_e} = -\frac{1}{R} \sum_{n=0}^{\infty} (n+1) \left(\frac{R}{r_e} \right)^{n+2} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (4.1)$$

The radial derivative of the disturbing potential can alternatively be expanded into a series of surface spherical harmonics

$$\left. \frac{\partial T}{\partial r} \right|_{r_e} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \overline{d_r T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (4.2)$$

where $\overline{d_r T}_{nm}^R$ are fully normalised surface SHCs. If in Equation (4.1) the ellipsoidal radius r_e is approximated by the spherical radius R – the so-called *constant radius approximation* (e.g., Heck, 1991) – the term $(R/r_e)^{n+1}$ vanishes and a simple relation between the coefficients in both series can be found. Since in both Equations (4.1) and (4.2), all dependence on position is then located inside the SHFs, it follows from the orthogonality of these functions on the sphere that the coefficients $\overline{d_r T}_{nm}^R$ read (e.g., Heiskanen and Moritz, 1967, p. 88)

$$\overline{d_r T}_{nm}^R = -\frac{n+1}{R} \bar{T}_{nm}^R \quad (4.3)$$

Thus, there is a one-to-one relationship between both types of coefficients. However, this is not the case if the constant radius approximation is not applied.

An exact solution is possible nevertheless, using a procedure very similar to the transformation from solid to surface SHCs provided in Chapter 3. The explicit formula for the ellipsoidal radius in geodetic coordinates (Table 2.1) can be inserted into Equation (4.1), yielding

$$\frac{\partial T}{\partial r} = -\frac{1}{R} \sum_{n=0}^{\infty} (n+1) \left(\frac{c}{\sqrt{1-e^2}} \right)^{n+2} (1-e^2 \sin^2 \theta)^{\frac{n+2}{2}} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (4.4)$$

where the scale factor c (Equation 3.64) is again introduced. Thus, in analogy with Equation (3.73), the spectral relationship between the disturbing potential and its radial derivative becomes

$$\overline{d_r T}_{nm}^R = \sum_{i=-\infty}^{\infty} \lambda_{nmi} (d_r T, T) \bar{T}_{n-2i,m}^R \quad (4.5)$$

where

$$\lambda_{nmi}(d_r T, T) = -\frac{n+1}{R} \sum_{k=|i|}^{\infty} \alpha_{n-2i+1,k} \overline{K}_{n-2i+1,m}^{2i,2k} \quad (4.6)$$

The notation of $d_r T$ and T between brackets indicates that the function $\lambda_{nmi}(d_r T, T)$ transforms SHCs of the disturbing potential T into SHCs of the radial derivative of the disturbing potential $d_r T$, and this notation will be used from hereon. More interesting than the radial derivative is the derivative with respect to the ellipsoidal normal, since this derivative appears in the ellipsoidal Neumann and Robin BVPs, as well as in the second and third geodetic BVPs. The normal derivative on the ellipsoid can be split into the radial term above, plus a latitudinal term, as follows.

4.1.2 The derivative with respect to the ellipsoidal normal

The normal derivative on the ellipsoid can be found via the well-known chain rule of differentiation

$$\frac{\partial T}{\partial h} = \frac{\partial r}{\partial h} \frac{\partial T}{\partial r} + \frac{\partial \theta}{\partial h} \frac{\partial T}{\partial \theta} \quad (4.7)$$

The derivatives of the geocentric radius r and the geocentric co-latitude θ with respect to the external ellipsoidal normal can be found by evaluating the relations among Cartesian, spherical polar and geodetic coordinates (Table 2.3). The distance from a point in space to the geocentre can be written in geodetic coordinates

$$r = \sqrt{(\nu + h)^2 \sin^2 \vartheta + (\nu(1 - e^2) + h)^2 \cos^2 \vartheta} \quad (4.8)$$

Differentiation with respect to the ellipsoidal normal gives

$$\frac{\partial r}{\partial h} = \frac{\nu(1 - e^2 \cos^2 \vartheta) + h}{r} \quad (4.9)$$

which on the surface of the ellipsoid ($h = 0$) reduces to

$$\left. \frac{\partial r}{\partial h} \right|_{r_e} = \frac{\nu}{r_e} (1 - e^2 \cos^2 \vartheta) = \cos \phi \quad (4.10)$$

where the relations for ν , r_e and $\cos \phi$ from Tables 2.1 and 2.2 were used.

A similar procedure can be applied to find the normal derivative of the geocentric co-latitude θ . It can be derived from the formulas in Table 2.3 that the geocentric

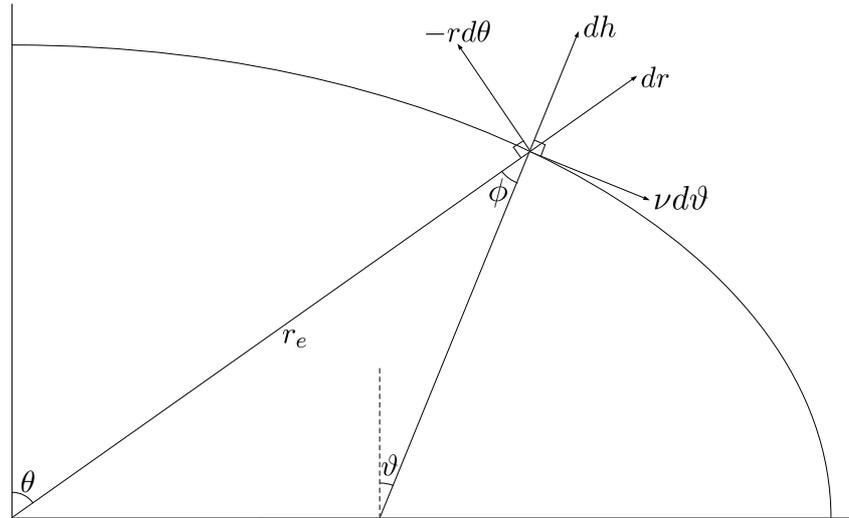


Figure 4.1: Normal derivative to the ellipsoid

co-latitude can be expressed as a function of the geodetic coordinates

$$\theta = \arctan \left(\frac{\nu + h}{\nu(1 - e^2) + h} \tan \vartheta \right) \quad (4.11)$$

Therefore, the derivative of the geocentric co-latitude with respect to the external ellipsoidal normal becomes, after some simplifications (not approximations)

$$\frac{\partial \theta}{\partial h} = \frac{-e^2 \nu \tan \vartheta}{(\nu(1 - e^2) + h)^2 + (\nu + h)^2 \tan^2 \vartheta} \quad (4.12)$$

which on the surface of the ellipsoid reduces to

$$\left. \frac{\partial \theta}{\partial h} \right|_{r_e} = \frac{-e^2 \tan \vartheta}{\nu(1 - e^2 + \tan^2 \vartheta)} = -\frac{\sin \phi}{r_e} \quad (4.13)$$

Insertion of Equations (4.10) and (4.13) into Equation (4.7) gives the disturbing potential (or any other function) on the ellipsoid

$$\frac{\partial T}{\partial h} = \cos \phi \frac{\partial T}{\partial r} - \sin \phi \frac{\partial T}{r_e \partial \theta} \quad (4.14)$$

Equation (4.14) can also be found geometrically, considering Figure 4.1. However, the algebraic derivation was followed because it can more easily be extended to higher order derivatives, as will be shown in Section 4.2.2.

4.1.3 The normal derivative of the disturbing potential

According to Equation (4.14), the normal derivative with respect to the ellipsoid can be found from a combination of the radial and the latitudinal derivatives. The latitudinal

derivative in terms of solid SHCs \bar{T}_{nm}^R reads

$$\left. \frac{\partial T}{\partial \theta} \right|_{r_e} = \sum_{n=0}^{\infty} \left(\frac{R}{r_e} \right)^{n+1} \sum_{m=-n}^n \bar{T}_{nm}^R \frac{\partial \bar{Y}_{nm}}{\partial \theta} \quad (4.15)$$

Now, Equations (4.1) and (4.15) can be inserted into Equation (4.14) to find the normal derivative of the disturbing potential in terms of the solid SHCs \bar{T}_{nm}^R

$$\left. \frac{\partial T}{\partial h} \right|_{r_e} = -\frac{1}{R} \sum_{n=0}^{\infty} \left(\frac{R}{r_e} \right)^{n+2} \sum_{m=-n}^n \bar{T}_{nm}^R \left((n+1) \cos \phi \bar{Y}_{nm} + \sin \phi \frac{\partial \bar{Y}_{nm}}{\partial \theta} \right) \quad (4.16)$$

Alternatively, the normal derivative of the disturbing potential can be expressed in terms of surface SHCs

$$\left. \frac{\partial T}{\partial h} \right|_{r_e} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \overline{d_h T}_{nm}^{r_e} \bar{Y}_{nm} \quad (4.17)$$

The aim is to find a relation between the solid SHCs \bar{T}_{nm}^R and the surface SHCs of the normal derivative on the ellipsoidal surface $\overline{d_h T}_{nm}^{r_e}$. A comparison of the two is only possible in the spectral domain if both can be expressed in terms of the same set of base functions. However, two obstacles are present in Equation (4.16) that prevent the direct comparison of coefficients. Firstly, Equation (4.16) contains terms that depend on latitude, which disturbs the orthogonality of the base functions. Secondly, the presence of the latitudinal derivative of the SHFs impedes the direct comparison of \bar{T}_{nm}^R and $\overline{d_h T}_{nm}^{r_e}$.

Equation (4.16) consists of two main parts, one part containing the SHF \bar{Y}_{nm} and another part containing its first derivative with respect to θ . Part one will be evaluated first. The terms that depend on the latitude can be expanded into binomial series (see Appendix C). This was shown for the term $(R/r_e)^{n+2}$ in Section 4.1.1, and the same strategy can be followed for the other terms.

The cosine of ϕ can be expanded into a product of two binomial series (see Appendix C) using the formulas in Table 2.2

$$\begin{aligned} \cos \phi &= (1 - \epsilon^4 \sin^2 \theta)^{-\frac{1}{2}} (1 - e^2 \sin^2 \theta) \\ &= \left(\sum_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} \epsilon^{4k} \sin^{2k} \theta \right) \left(\sum_{k=0}^1 (-1)^k e^{2k} \sin^{2k} \theta \right) \end{aligned} \quad (4.18)$$

The two binomial series can be combined in one through a Cauchy multiplication (see Appendix B). After applying the relation between the first numerical eccentricity e and

the quartic eccentricity ϵ in Equation (2.12), the following relation appears

$$\cos \phi = \sum_{k=0}^{\infty} (-1)^k e^{2k} (2 - e^2)^k \sum_{l=k-1}^k \binom{-\frac{1}{2}}{l} \sin^{2k} \theta \quad (4.19)$$

Many ways can be followed to proceed from here. Since the term in Equation (4.19) is multiplied by the SHF \bar{Y}_{nm} in Equation (4.16), the sinusoidal LWFs can be introduced by implementating Equation (3.31). This moves the dependence on latitude apparent in Equation (4.19) into the spherical harmonics, which is required for the comparison of \bar{T}_{nm}^R and $\bar{d}_h T_{nm}^{r_e}$.

However, it follows from Equation (4.16) that Equation (4.19) still needs to be multiplied by the term $(R/r_e)^{n+2}$. This term can also be written as a power series over the sine of co-latitude θ , and it is more efficient to combine both power series before inserting the LWFs. This can be seen from the fact that the insertion of the LWFs introduces an extra summation sign, whereas a Cauchy multiplication introduces none. It is therefore chosen to introduce the LWFs after the Cauchy multiplication has taken place.

In Section 4.1.1, it was derived that the term $(R/r_e)^{n+2}$ can be expressed as a power series

$$\left(\frac{R}{r_e}\right)^{n+2} = \sum_{k=0}^{\infty} \alpha_{n+1,k} \sin^{2k} \theta \quad (4.20)$$

Now, the two series in Equations (4.19) and (4.20) can be combined using a Cauchy multiplication (see Appendix B)

$$\left(\frac{R}{r_e}\right)^{n+2} \cos \phi = \sum_{j=0}^{\infty} \sum_{k=0}^j \alpha_{n+1,j-k} (-1)^k e^{2k} (2 - e^2)^k \sum_{l=k-1}^k \binom{-\frac{1}{2}}{l} \sin^{2j} \theta \quad (4.21)$$

After multiplication of Equation (4.21) with the SHFs \bar{Y}_{nm} , the sinusoidal LWFs (Equation 3.31) can be inserted. This gives

$$\left(\frac{R}{r_e}\right)^{n+2} \cos \phi \bar{Y}_{nm} = \sum_{j=0}^{\infty} \sum_{i=-j}^j \beta_{nmij} \bar{Y}_{n+2i,m} \quad (4.22)$$

where

$$\beta_{nmij} = \bar{K}_{nm}^{2i,2j} \sum_{k=0}^j \alpha_{n+1,j-k} (-1)^k e^{2k} (2 - e^2)^k \sum_{l=k-1}^k \binom{-\frac{1}{2}}{l} \quad (4.23)$$

It can be seen from Equation (4.23) that the function β_{nmij} in Equation (4.22) only depends on degree n , order m , the first numerical eccentricity of the ellipsoid e , and the scale factor c (Equation 3.64). It is therefore independent of latitude and longitude.

In a similar fashion, the second part of Equation (4.16) can be rewritten. First, the expression for $\sin \phi$ in spherical polar coordinates (see Table 2.2) can be inserted

$$\sin \phi \frac{\partial \bar{Y}_{nm}}{\partial \theta} = (1 - \epsilon^4 \sin^2 \theta)^{-\frac{1}{2}} e^2 \sin \theta \cos \theta \frac{\partial \bar{Y}_{nm}}{\partial \theta} \quad (4.24)$$

The term $(1 - \epsilon^4 \sin^2 \theta)^{-\frac{1}{2}}$ can be expanded into a binomial series, and furthermore Equation (3.50) can be inserted to obtain

$$\sin \phi \frac{\partial \bar{Y}_{nm}}{\partial \theta} = e^2 \sum_{j=0}^{\infty} (-1)^j \binom{-\frac{1}{2}}{j} \epsilon^{4j} \sin^{2j} \theta \sum_{l=-1}^1 \bar{N}_{nm}^{2,2l} \bar{Y}_{n+2l,m} \quad (4.25)$$

When Equation (4.25) is multiplied by the term $(R/r_e)^{n+2}$ and Equation (4.20) inserted, after a Cauchy multiplication of the two power series (see Appendix B), it becomes

$$\left(\frac{R}{r_e}\right)^{n+2} \sin \phi \frac{\partial \bar{Y}_{nm}}{\partial \theta} = e^2 \sum_{j=0}^{\infty} \sum_{k=0}^j \alpha_{n+1,j-k} (-1)^k \binom{-\frac{1}{2}}{k} \epsilon^{4k} \sin^{2j} \theta \sum_{l=-1}^1 \bar{N}_{nm}^{2l,2} \bar{Y}_{n+2l,m} \quad (4.26)$$

Now, the sinusoidal LWFs (Equation 3.31) can be inserted and a formula where all dependence on latitude is centred inside the SHFs can be derived

$$\begin{aligned} \left(\frac{R}{r_e}\right)^{n+2} \sin \phi \frac{\partial \bar{Y}_{nm}}{\partial \theta} &= e^2 \sum_{j=0}^{\infty} \sum_{k=0}^j \alpha_{n+1,j-k} (-1)^k \binom{-\frac{1}{2}}{k} \epsilon^{4k} \\ &\times \sum_{i=-j}^j \bar{K}_{n+2l,m}^{2i,2j} \sum_{l=-1}^1 \bar{N}_{nm}^{2l,2} \bar{Y}_{n+2i+2l,m} \end{aligned} \quad (4.27)$$

Equation (4.27) is slightly more complicated than Equation (4.22), derived for the first part of Equation (4.16), because the degree of the SHF depends on three summation indices. A solution of similar form as Equation (4.22) can be obtained nevertheless, because the summation order over i and l can be rearranged (see Appendix B). Equation (4.27) can thus be written as

$$\left(\frac{R}{r_e}\right)^{n+2} \sin \phi \frac{\partial \bar{Y}_{nm}}{\partial \theta} = \sum_{j=0}^{\infty} \sum_{i=-j-1}^{j+1} \gamma_{nmij} \bar{Y}_{n+2i,m} \quad (4.28)$$

where

$$\gamma_{nmij} = e^2 \sum_{k=0}^j \alpha_{n+1,j-k} (-1)^k \binom{-\frac{1}{2}}{k} \epsilon^{4k} \sum_{l=\max(-j,i-1)}^{\min(j,i+1)} \bar{K}_{n+2l,m}^{2(i-l),2j} \bar{N}_{nm}^{2l,2} \quad (4.29)$$

The function γ_{nmij} depends on the same parameters as β_{nmij} , i.e., degree n , order m , the first numerical eccentricity of the ellipsoid e , and the scale factor c . Equations (4.22) and (4.28) are of similar form, the main difference being that the summation over i contains two extra terms in the latter, namely $i = -j - 1$ and $i = j + 1$. Note, however, that β_{nmij} is zero for these two cases, due to the fact that the sinusoidal LWFs $\overline{K}_{nm}^{2i,2j}$ are defined as zero for $|i| > j$.

Equations (4.22) and (4.28) can thus easily be added together and implemented in Equation (4.16). Finally, an expression for the derivative of the disturbing potential with respect to the ellipsoidal normal can be found

$$\left. \frac{\partial T}{\partial h} \right|_{r_e} = -\frac{1}{R} \sum_{n=0}^{\infty} \sum_{m=-n}^n \overline{T}_{nm}^R \sum_{j=0}^{\infty} \sum_{i=-j-1}^{j+1} [(n+1)\beta_{nmij} + \gamma_{nmij}] \overline{Y}_{n+2i,m} \quad (4.30)$$

where all dependence on longitude and latitude is inside the SHFs \overline{Y}_{nm} . The summation order over n and i can be rearranged (see Appendix B) to yield an expression where the SHFs solely depend on degree n and order m

$$\left. \frac{\partial T}{\partial h} \right|_{r_e} = -\frac{1}{R} \sum_{n=0}^{\infty} \sum_{m=-n}^n \sum_{j=0}^{\infty} \sum_{i=-j-1}^{j+1} [(n-2i+1)\beta_{n-2i,mij} + \gamma_{n-2i,mij}] \overline{T}_{n-2i,m}^R \overline{Y}_{nm} \quad (4.31)$$

Subsequently, the summation over indices i and j can be rearranged so that the summation over i , upon which the SHCs depend, can be brought outside the summation over j

$$\left. \frac{\partial T}{\partial h} \right|_{r_e} = -\frac{1}{R} \sum_{n=0}^{\infty} \sum_{m=-n}^n \sum_{i=-\infty}^{\infty} \sum_{j=|i|-1}^{\infty} [(n-2i+1)\beta_{n-2i,mij} + \gamma_{n-2i,mij}] \overline{T}_{n-2i,m}^R \overline{Y}_{nm} \quad (4.32)$$

This allows for a direct comparison with a surface harmonic expansion of the normal derivative of the disturbing potential. The surface SHCs $\overline{d}_h \overline{T}_{nm}^{r_e}$ can now directly be computed from a weighted summation over the geopotential coefficients \overline{T}_{nm}^R of equal order m

$$\overline{d}_h \overline{T}_{nm}^{r_e} = \sum_{i=-\infty}^{\infty} \lambda_{nmi}(d_h T, T) \overline{T}_{n-2i,m}^R \quad (4.33)$$

where

$$\lambda_{nmi}(d_h T, T) = -\frac{1}{R} \sum_{j=|i|-1}^{\infty} [(n-2i+1)\beta_{n-2i,mij} + \gamma_{n-2i,mij}] \quad (4.34)$$

Note that Equation (4.33) is almost identical to Equation (4.5), the only difference being that the radial derivative is replaced by the normal derivative to the ellipsoid.

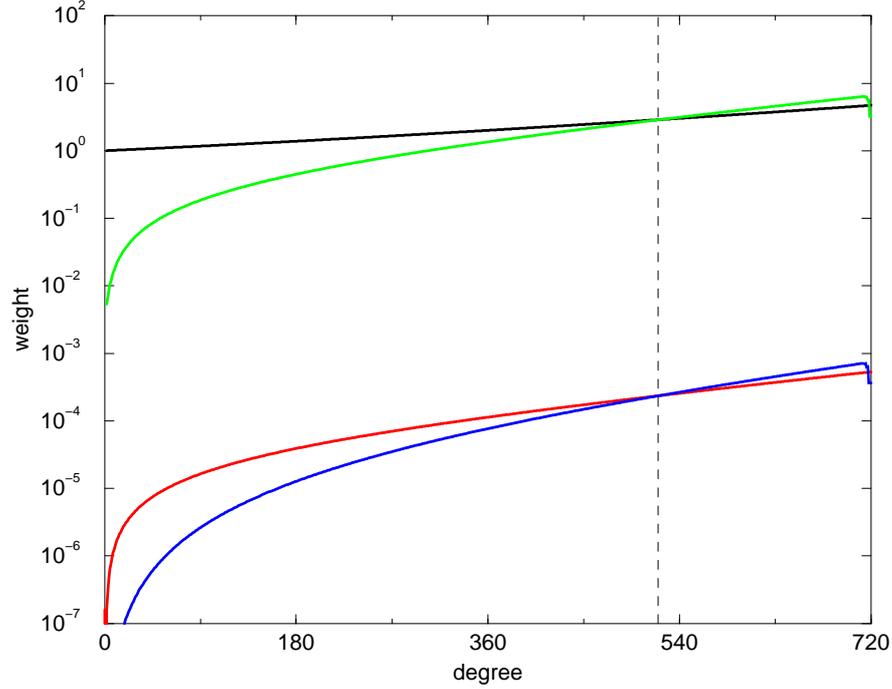


Figure 4.2: Values of $|\lambda_{n00}|$ (black) and $\sum_{i=-\infty, i \neq 0}^{\infty} |\lambda_{n0i}|$ (green) from Equation (3.74), as well as $|\lambda_{n00}(d_h T, T)|$ (red) and $\sum_{i=-\infty, i \neq 0}^{\infty} |\lambda_{n0i}(d_h T, T)|$ (blue) from Equation (4.34), indicating the diagonal dominance of the weights below degree $n = 520$ for both cases

It is a generalisation and theoretical improvement of the simple spherical and constant radius approximation in Equation (4.3). A similar generalisation can be derived for the second normal derivative of the disturbing potential, as will be shown later in Section 4.2.

As for the solution to the ellipsoidal Dirichlet BVP in Chapter 3, the inverse of Equation (4.33) can be found via a rigorous matrix inversion, or via an iterative procedure similar to Equation (3.79). Unfortunately, just as in the case of the solution to the ellipsoidal Dirichlet BVP, the iteration will not converge for some SHCs beyond degree 520. This is because the weights $\lambda_{nmi}(d_h T, T)$ are no longer diagonally dominant beyond degree 520, i.e., the weights do no longer obey Equation (3.78) beyond degree 520, as is shown for the zonal weights ($m = 0$) in Figure 4.2. The underlying reason is the same as in the weights λ_{nmi} in Equation (3.79), and the cause lies in the functions α_{nk} , which contain binomial coefficients that are dependent on the spherical harmonic degree n . All other binomial coefficients apparent in β_{nmij} and γ_{nmij} do not contain the degree

n or order m .

The inverse of the relation in Equation (4.33) computes solid SHCs \overline{T}_{nm}^R from surface SHCs $\overline{d}_h T_{nm}^{re}$. Since the solid SHCs describe the disturbing potential T in the whole space exterior to the ellipsoid, the inverse of Equation (4.33) allows us to compute the disturbing potential anywhere outside the ellipsoid from its derivative on the surface of the ellipsoid (represented by a set of surface SHCs). The transformation does not only hold for the disturbing potential T , but for any function that is harmonic outside the ellipsoid, and the inverse of Equation (4.33) is therefore a solution to the ellipsoidal Neumann BVP. The second geodetic BVP (or fixed geodetic BVP) with the boundary condition given by Equation (2.48) is an example of an ellipsoidal Neumann BVP, and the transformations derived here thus provide its solution.

4.2 Relation between the disturbing potential and its second normal derivative

The second normal derivative of the disturbing potential on the ellipsoid describes the vertical gravity gradient, i.e., the gradient of gravity in the direction of the ellipsoidal normal. The gravity gradient has received a lot of attention in recent years (e.g., Klees et al., 2000; Drinkwater et al., 2003), because the technique of satellite gravity gradiometry directly measures the second derivatives of the Earth's potential. In addition, vertical gravity gradients are important for the reduction of gravity measured on the Earth's surface and the computation of orthometric heights (cf. Tenzer et al., 2005). In the latter application, the gravity gradient along the plumbline is needed, but this is in practice often approximated by the gravity gradient in the direction of the ellipsoidal normal (e.g., Vaníček and Krakiwsky, 1986, p.497).

4.2.1 The second radial derivative of the disturbing potential

The second radial derivative of the disturbing potential is the gravity gradient in the direction towards the Earth's centre (here assumed to coincide with the centre of the ellipsoid). The second radial derivative of a solid spherical harmonic expansion follows from a differentiation of Equation (4.1) with respect to r_e , and the second radial derivative of the disturbing potential can thus be written as follows

$$\left. \frac{\partial^2 T}{\partial r^2} \right|_{r_e} = \frac{1}{R^2} \sum_{n=0}^{\infty} (n+1)(n+2) \left(\frac{R}{r_e} \right)^{n+3} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (4.35)$$

This can also be expanded into a series of surface SHFs

$$\left. \frac{\partial^2 T}{\partial r^2} \right|_{r_e} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \overline{d_r^2 T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (4.36)$$

In constant radius approximation, where $R = r_e$, the term $(R/r_e)^{n+3}$ equals unity and a simple relation between the coefficients in both series can be found (e.g., Rummel and Van Gelderen, 1992)

$$\overline{d_r^2 T}_{nm}^R = \frac{(n+1)(n+2)}{R^2} \bar{T}_{nm}^R \quad (4.37)$$

Thus, as in the case of the first radial derivative in constant radius approximation, there is a one-to-one relationship between both type of coefficients. However, this is not the case if the constant radius approximation is not applied, but an exact solution can be found along similar lines as the derivation in Section 4.1.1.

The explicit formula for the ellipsoidal radius in spherical polar coordinates (Table 2.1) and the scale factor c (Equation 3.64) can be inserted into Equation (4.35)

$$\left. \frac{\partial^2 T}{\partial r^2} \right|_{r_e} = \frac{1}{R^2} \sum_{n=0}^{\infty} (n+1)(n+2) \left(\frac{c}{\sqrt{1-e^2}} \right)^{n+3} (1-e^2 \sin^2 \theta)^{\frac{n+3}{2}} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (4.38)$$

Equation (4.38) is very similar to the case of the first radial derivative in Equation (4.4), so can be treated along the same lines. The spectral relationship between the disturbing potential and its radial derivative thus becomes

$$\overline{d_r^2 T}_{nm}^R = \sum_{i=-\infty}^{\infty} \lambda_{nmi} (d_r^2 T, T) \bar{T}_{n-2i,m}^R \quad (4.39)$$

where

$$\lambda_{nmi}(d_r^2 T, T) = \frac{(n+1)(n+2)}{R^2} \sum_{k=|i|}^{\infty} \alpha_{n-2i+2,k} K_{n-2i+2,m}^{2i,2k} \quad (4.40)$$

Obtaining a similar relation for the second derivative with respect to the ellipsoidal normal is more complicated, but it is possible nonetheless, as is shown in the remainder of this Section. The first step in this procedure is the derivation of a general relation writing the second derivative with respect to the ellipsoidal normal in terms of first- and second-order derivatives with respect to the geocentric radius and latitude.

4.2.2 The second derivative with respect to the ellipsoidal normal

The second derivative along the ellipsoidal normal can be expressed as a combination of radial and latitudinal derivatives by applying the chain rule of differentiation to Equation (4.7)

$$\begin{aligned} \frac{\partial^2 T}{\partial h^2} &= \frac{\partial}{\partial h} \left(\frac{\partial r}{\partial h} \frac{\partial T}{\partial r} + \frac{\partial \theta}{\partial h} \frac{\partial T}{\partial \theta} \right) \\ &= \frac{\partial^2 r}{\partial h^2} \frac{\partial T}{\partial r} + \left(\frac{\partial r}{\partial h} \right)^2 \frac{\partial^2 T}{\partial r^2} + \frac{\partial^2 \theta}{\partial h^2} \frac{\partial T}{\partial \theta} + \left(\frac{\partial \theta}{\partial h} \right)^2 \frac{\partial^2 T}{\partial \theta^2} + 2 \frac{\partial \theta}{\partial h} \frac{\partial r}{\partial h} \frac{\partial^2 T}{\partial r \partial h} \end{aligned} \quad (4.41)$$

The second normal derivative of the ellipsoidal radius can be found by differentiation of Equation (4.9) with respect to h

$$\frac{\partial^2 r}{\partial h^2} = \frac{1}{r} - \frac{1}{r^3} (\nu(1 - e^2 \cos^2 \vartheta) + h)^2 = \frac{1}{r} \left[1 - \left(\frac{\partial r}{\partial h} \right)^2 \right] \quad (4.42)$$

For a point on the surface of the ellipsoid, after implementation of Equation (4.10), this becomes

$$\left. \frac{\partial^2 r}{\partial h^2} \right|_{r_e} = \frac{\sin^2 \phi}{r_e} \quad (4.43)$$

The second normal derivative of the geocentric latitude can be found by differentiation of Equation (4.12) with respect to h

$$\begin{aligned} \frac{\partial^2 \theta}{\partial h^2} &= \frac{2e^2 \tan \vartheta (\nu + h)(1 - e^2 + \tan \vartheta)}{[(1 - e^2)^2 + \tan^2 \vartheta]^2} \\ &= \frac{2(\nu + h)(1 - e^2 + \tan \vartheta)}{e^2 \nu \tan \vartheta} \left(\frac{\partial \theta}{\partial h} \right)^2 \end{aligned} \quad (4.44)$$

For a point on the surface of the ellipsoid, after implementation of Equation (4.13), this becomes

$$\left. \frac{\partial^2 \theta}{\partial h^2} \right|_{r_e} = \frac{2(1 - e^2 + \tan^2 \vartheta) \sin^2 \phi}{e^2 r_e^2 \tan \vartheta} = \frac{2 \sin \phi \cos \phi}{r_e^2} \quad (4.45)$$

Thus, inserting Equations (4.10), (4.13), (4.43) and (4.45) into Equation (4.41) gives the second normal derivative of the disturbing potential on the surface of the ellipsoid

$$\left. \frac{\partial^2 T}{\partial h^2} \right|_{r_e} = \cos^2 \phi \frac{\partial^2 T}{\partial r^2} + \sin^2 \phi \left(\frac{\partial T}{r_e \partial r} + \frac{\partial^2 T}{r_e^2 \partial \theta^2} \right) + \sin 2\phi \left(\frac{\partial T}{r_e^2 \partial \theta} - \frac{\partial^2 T}{r_e \partial r \partial \theta} \right) \quad (4.46)$$

which will allow for an expression of the second normal derivative in terms of SHCs, described next.

4.2.3 The second normal derivative of the disturbing potential

To compute the second normal derivative of the disturbing potential from geopotential coefficients, the second derivatives with respect to the geocentric radius and latitude are required in terms of solid SHCs. The second radial derivative of the disturbing potential follows from differentiation of Equation (4.1) with respect to r

$$\left. \frac{\partial^2 T}{\partial r^2} \right|_{r_e} = \frac{1}{R^2} \sum_{n=0}^{\infty} (n+1)(n+2) \left(\frac{R}{r} \right)^{n+3} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm} \quad (4.47)$$

The second derivative with respect to the geocentric latitude follows from differentiation of Equation (4.15) with respect to θ

$$\left. \frac{\partial^2 T}{\partial \theta^2} \right|_{r_e} = \sum_{n=0}^{\infty} \left(\frac{R}{r_e} \right)^{n+1} \sum_{m=-n}^n \bar{T}_{nm}^R \frac{\partial^2 \bar{Y}_{nm}}{\partial \theta^2} \quad (4.48)$$

The cross-derivative can be derived in two ways, either from a differentiation of Equation (4.1) with respect to the geocentric latitude, or by a differentiation of Equation (4.15) with respect to the geocentric radius. Both give the same result:

$$\left. \frac{\partial^2 T}{\partial r \partial \theta} \right|_{r_e} = -\frac{1}{R} \sum_{n=0}^{\infty} (n+1) \left(\frac{R}{r} \right)^{n+2} \sum_{m=-n}^n \bar{T}_{nm}^R \frac{\partial \bar{Y}_{nm}}{\partial \theta} \quad (4.49)$$

The second derivative of the disturbing potential with respect to the ellipsoidal normal can now be found by inserting Equations (4.1), (4.15), (4.47), (4.48) and (4.49) into Equation (4.46)

$$\begin{aligned} \left. \frac{\partial^2 T}{\partial h^2} \right|_{r_e} = \frac{1}{R^2} \sum_{n=0}^{\infty} \left(\frac{R}{r_e} \right)^{n+3} \sum_{m=-n}^n \bar{T}_{nm}^R \left\{ (n+1)[n+2 - (n+3) \sin^2 \phi] \bar{Y}_{nm} \right. \\ \left. + (n+2) \sin 2\phi \frac{\partial \bar{Y}_{nm}}{\partial \theta} + \sin^2 \phi \frac{\partial^2 \bar{Y}_{nm}}{\partial \theta^2} \right\} \quad (4.50) \end{aligned}$$

This expression for the second normal derivative contains many terms that are dependent on latitude. As in the case of the first derivative, all dependence on latitude can be moved into the spherical harmonic base functions \bar{Y}_{nm} to allow for a direct comparison with a surface spherical harmonic expansion of the second derivatives.

Equation (4.50) can be rewritten to move the dependence on co-latitude θ into the SHFs, and express the derivatives of the SHFs as a function of the SHFs themselves as well, in a very similar way to the treatment of the first derivative of the disturbing potential in Section 4.1.3. The trigonometric functions of the ellipsoidal deflection angle ϕ apparent in Equation (4.50) are first expanded into one or two binomial series (see Appendix C), using the relations provided in Table 2.2 and some trigonometric identities

$$\begin{aligned}\sin^2 \phi &= e^2 \sin^2 \theta \cos^2 \theta (1 - \epsilon^4 \sin^2 \theta)^{-1} \\ &= e^2 \sin^2 \theta \cos^2 \theta \sum_{k=0}^{\infty} (-1)^k \binom{-1}{k} \epsilon^{4k} \sin^{2k} \theta\end{aligned}\quad (4.51)$$

$$\begin{aligned}\sin 2\phi &= 2e^2 \sin \theta \cos \theta (1 - e^2 \sin^2 \theta)(1 - \epsilon^4 \sin^2 \theta)^{-1} \\ &= 2e^2 \sin \theta \cos \theta \left[\sum_{k=0}^1 (-1)^k e^{2k} \sin^{2k} \theta \right] \left[\sum_{k=0}^{\infty} (-1)^k \binom{-1}{k} \epsilon^{4k} \sin^{2k} \theta \right]\end{aligned}\quad (4.52)$$

Subsequently, the products of binomial series can be combined using a Cauchy multiplication (see Appendix B), and when multiplied with the first- or second-order derivatives of the SHFs, the new relations derived in Section 3.2 can be inserted.

In the multiplication of $\sin 2\phi$ (Equation 4.52) with the first derivative of the spherical harmonics, Equation (3.50) can be inserted

$$\sin 2\phi \frac{\partial \bar{Y}_{nm}}{\partial \theta} = 2e^2 \sum_{j=0}^{\infty} (-1)^j e^{2j} \sum_{k=j-1}^j (2 - e^2)^k \binom{-1}{k} \sin^{2j} \theta \sum_{i=-1}^1 \bar{N}_{nm}^{2i,2} \bar{Y}_{n+2i,m} \quad (4.53)$$

In the case of the second derivative of the spherical harmonics, Equation (3.61) can be applied

$$\sin^2 \phi \frac{\partial^2 \bar{Y}_{nm}}{\partial \theta^2} = e^2 \sum_{k=0}^{\infty} (-1)^k \binom{-1}{k} \epsilon^{4k} \sin^{2k} \theta \sum_{i=-2}^2 \bar{R}_{nm}^{2i,4} \bar{Y}_{n+2i,m} \quad (4.54)$$

Equation (4.50) also contains a term where $\sin^2 \phi$ is multiplied by the SHF itself. Equations (3.21) and (3.29) can be inserted to account for the terms $\cos^2 \theta$ and $\sin^2 \theta$

apparent in Equation (4.51). After a Cauchy multiplication of the series in Equations (3.21) and (3.29), it is derived that

$$\sin^2 \phi \bar{Y}_{nm} = e^2 \sum_{k=0}^{\infty} (-1)^k \binom{-1}{k} \epsilon^{4k} \sin^{2k} \theta \sum_{i=-2}^2 \sum_{j=-1}^1 \bar{K}_{nm}^{2i,2} \bar{F}_{nm}^{2(i-j),2} \bar{Y}_{n+2i,m} \quad (4.55)$$

From here, the derivation follows exactly the same principles as the derivation in the case of the first normal derivative of the disturbing potential in Section 4.1.3. Equations (4.53) to (4.55) can be pre-multiplied by the term $(R/r_e)^{n+3}$, which can be expanded into a binomial series similar to Equation (4.20). The two series can be combined using a Cauchy multiplication, after which the sinusoidal LWFs \bar{K}_{nm}^{ij} can be inserted. This moves all dependence on latitude inside the SHFs \bar{Y}_{nm} .

When the summation order is rearranged such that the SHFs are independent of any summation indices other than n and m , the series can be compared to a surface spherical harmonic expansion of the second-order derivatives of the disturbing potential on the ellipsoid

$$\left. \frac{\partial^2 T}{\partial h^2} \right|_{r_e} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \overline{d_h^2 T}_{nm}^{r_e} \bar{Y}_{nm}(\theta, \lambda) \quad (4.56)$$

The surface SHCs $\overline{d_h^2 T}_{nm}^{r_e}$ can eventually be expressed as a weighted summation over the solid SHCs \bar{T}_{nm}^R

$$\overline{d_h^2 T}_{nm}^{r_e} = \sum_{i=-\infty}^{\infty} \lambda_{nmi} (d_h^2 T, T) \bar{T}_{n-2i,m}^R \quad (4.57)$$

where the weights follow from an infinite summation

$$\lambda_{nmi} (d_h^2 T, T) = \frac{1}{R^2} \sum_{j=|i|-2}^{\infty} \left[(n+1)(n+2)\delta_{0i} - (n-2i+1)(n-2i+3)\kappa_{n-2i,mij} \right. \\ \left. + (n-2i+2)\zeta_{n-2i,mij} + \eta_{n-2i,mij} \right] \quad (4.58)$$

The weights only depend on the degree n , order m , the first numerical eccentricity of the ellipsoid e , the reference sphere R , and the scale factor c , as can be seen from the explicit formulas for the terms in Equation (4.58)

$$\zeta_{nmij} = 2e^2 \sum_{k=0}^j \alpha_{n+2,j-k} (-1)^k e^{2k} \sum_{l=k-1}^k (2-e^2)^l \binom{-1}{l} \sum_{p=\max(-j,i-1)}^{\min(j,i+1)} \bar{K}_{n+2p,m}^{2(i-p),2j} \bar{N}_{nm}^{2p,2} \quad (4.59)$$

$$\eta_{nmij} = e^2 \sum_{k=0}^j \alpha_{n+2,j-k} (-1)^k \epsilon^{4k} \binom{-1}{k} \sum_{l=\max(-j,i-2)}^{\min(j,i+2)} \bar{K}_{n+2l,m}^{2(i-l),2j} \bar{R}_{nm}^{2l,4} \quad (4.60)$$

$$\kappa_{nmij} = e^2 \sum_{k=0}^j \alpha_{n+2,j-k} (-1)^k \epsilon^{4k} \binom{-1}{k} \sum_{l=\max(-j,i-2)}^{\min(j,i+2)} \overline{K}_{n+2l,m}^{2(i-l),2j} \overline{K}_{nm}^{2l,2} \sum_{p=-1}^1 \overline{F}_{nm}^{2(i-p),2} \quad (4.61)$$

Equations (4.59) to (4.61) are the contributions to the weight functions of the terms in Equations (4.53) to (4.55) respectively.

The inverse of Equation (4.57) computes solid SHCs of the disturbing potential (or any other function that is harmonic outside the ellipsoid), from surface SHCs of its second derivative on the ellipsoid. It therefore forms a solution of the second-order ellipsoidal BVP in the spectral domain.

The numerical stability and efficiency of this solution for the case of the Earth's disturbing potential are not explicitly tested here, but based on the appearance of the term $\alpha_{n+2,j-k}$ in Equations (4.59) to (4.61), difficulties in the inversion for high degrees ($n > 520$) can be expected. Since the spectral sensitivity of these high degrees of the second normal derivative of the disturbing potential is more than 99% (e.g. Schwarz, 1985), this severely hinders the practicality of the solution in applications involving the Earth's gravity field.

4.3 Relation between the disturbing potential and the gravity anomaly

In Section 2.3.2, the relation between the disturbing potential T and the gravity anomaly Δg was derived in the space domain. This relation is given by the fundamental equation of physical geodesy (Equation 2.52)

$$\Delta g = -\frac{\partial T}{\partial h} + \frac{1}{\gamma} \frac{\partial \gamma}{\partial h} T \quad (4.62)$$

and the level of approximation of Equation (4.62) was assessed in Section 2.3.3. In this Section, the relation in the spectral domain will be derived. This relation is well-known for the spherical and constant radius approximation (e.g., Heiskanen and Moritz, 1967, p. 89), and is derived in Section 4.3.1. Without application of these approximations, the relation becomes more complicated, since a one-to-one relation between SHCs is no longer possible. However, the relation can be found using the new relations among

SHFs derived in Section 3.2, and this more rigorous relation is derived here for the first time.

4.3.1 Spherical and constant radius approximation

The gravity anomaly on the ellipsoid can be expanded into a series of surface SHFs (Equation 2.83)

$$\Delta g = \sum_{n=0}^{\infty} \sum_{m=-n}^n \overline{\Delta g}_{nm}^{r_e} \overline{Y}_{nm} \quad (4.63)$$

The coefficients $\overline{\Delta g}_{nm}^{r_e}$ can simply be related to the geopotential coefficients \overline{T}_{nm}^R in spherical and constant radius approximation, as follows. The spherical approximation includes two steps. First, the normal derivative to the ellipsoid is approximated by the radial derivative, which transforms Equation (4.62) into

$$\Delta g = -\frac{\partial T}{\partial r} + \frac{1}{\gamma} \frac{\partial \gamma}{\partial r} T \quad (4.64)$$

Secondly, the reference gravity γ is approximated by a spherical reference gravity $\tilde{\gamma}$. This spherical reference gravity is simply the gravity induced by a point mass or a homogeneous sphere (of which the radius does not exceed the ellipsoidal radius at the point of interest) with a total mass equal to that of the Earth, which is independent of longitude and latitude (Equation 2.31). The radial derivative of the spherical reference gravity can be found by differentiation, and inserted into Equation (4.64)

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{r_e} T \quad (4.65)$$

The error in the gravity anomaly introduced by the spherical approximation is well-known to be of the order of the flattening f of the ellipsoid (≈ 0.003) (e.g., Heiskanen and Moritz, 1967; Rummel, 1985).

Inserting the spherical harmonic representations of the disturbing potential (Equation 2.81) and its radial derivative (Equation 4.1) into Equation (4.65) gives an expression for the gravity anomaly in terms of geopotential coefficients

$$\Delta g = \frac{1}{R} \sum_{n=0}^{\infty} \sum_{m=-n}^n (n-1) \left(\frac{R}{r_e}\right)^{n+2} \overline{T}_{nm}^R \overline{Y}_{nm} \quad (4.66)$$

The simplest solution occurs when the ellipsoidal radius r_e is set equal to the radius of the reference sphere R , i.e. when the constant radius approximation is applied. This causes the term $(R/r_e)^{n+2}$ to equal unity and Equation (4.66) reduces to

$$\Delta g = \sum_{n=0}^{\infty} \sum_{m=-n}^n \frac{n-1}{R} \overline{T}_{nm}^R \overline{Y}_{nm} \quad (4.67)$$

and all dependence on latitude is now contained inside the SHFs. Due to the orthogonality of the SHFs on the sphere, a one-to-one relation between surface SHCs of the gravity anomaly and geopotential coefficients can be obtained upon comparison of Equations (4.63) and (4.67) (e.g., Heiskanen and Moritz, 1967, p. 89)

$$\overline{\Delta g}_{nm}^{r_e} = \frac{n-1}{R} \overline{T}_{nm}^R \quad (4.68)$$

Alternatively, similar to the case of the first derivative of the disturbing potential, the term $(R/r_e)^{n+2}$ can be expanded into a binomial series after which an expression occurs with all dependence on latitude inside the SHFs

$$\Delta g = -\frac{1}{R} \sum_{n=0}^{\infty} \sum_{m=-n}^n \overline{Y}_{nm} \sum_{k=0}^{\infty} \sum_{i=-k}^k (n-1-2i) \alpha_{n+1-2i,k} \overline{K}_{n-2i,m}^{2i,2k} \overline{T}_{n-2i,m}^R \quad (4.69)$$

The errors introduced by the spherical and constant radius approximation will be assessed in Chapter 5. These errors can be revealed because a rigorous derivation without these approximations is possible, as is presented next.

4.3.2 The rigorous ellipsoidal relation

While the approximations to Equation (4.62) made above do simplify the solution, it will be shown here that they are not strictly necessary. A rigorous derivation, largely based on formulas derived in previous Sections, is possible. In Section 4.1, it was shown that the spectral relation between the disturbing potential and its derivative with respect to the ellipsoidal normal can be derived rigorously. It remains to be proven that the second term in Equation (4.62), containing the reference gravity γ , can be expressed in a similar way, and this is shown next.

According to Equation (2.36), the derivative of the reference gravity can be expressed as a function of the principal radii of curvature ρ and ν and the angular velocity of the

Earth's rotation ω . Division of this formula by the normal gravity γ gives

$$\frac{1}{\gamma} \frac{\partial \gamma}{\partial h} = -\frac{1}{\rho} - \frac{1}{\nu} - 2 \frac{\omega^2}{\gamma} \quad (4.70)$$

The terms on the right-hand side of Equation (4.70) can all be expressed as a function of the geocentric latitude. Moreover, they can all be expressed in such a way that they can be expanded into converging binomial series, which allows for the insertion of the LWFs derived in Section 3.2.

For example, the inverse of the meridian radius of curvature ρ can be written as (see Table 2.1)

$$\frac{1}{\rho} = \frac{b}{a^2} (1 - e^2 \sin^2 \theta)^{\frac{3}{2}} (1 - \epsilon^4 \sin^2 \theta)^{-\frac{3}{2}} \quad (4.71)$$

The two terms between brackets in Equation (4.71) can be written as a binomial series (see Appendix C)

$$\frac{1}{\rho} = \frac{b}{a^2} \left[\sum_{j=0}^{\infty} (-1)^j \binom{\frac{3}{2}}{j} e^{2j} \sin^{2j} \theta \right] \left[\sum_{k=0}^{\infty} (-1)^k \binom{-\frac{3}{2}}{k} \epsilon^{4k} \sin^{2k} \theta \right] \quad (4.72)$$

and these two binomial series can be combined using a Cauchy multiplication (see Appendix B)

$$\frac{1}{\rho} = \frac{b}{a^2} \sum_{j=0}^{\infty} (-1)^j e^{2j} \sum_{k=0}^j \binom{\frac{3}{2}}{k} \binom{-\frac{3}{2}}{j-k} (2 - e^2)^{j-k} \sin^{2j} \theta \quad (4.73)$$

A power series for the inverse of the prime vertical radius of curvature ν can be found in a very similar way

$$\begin{aligned} \frac{1}{\nu} &= \frac{b}{a^2} (1 - e^2 \sin^2 \theta)^{\frac{1}{2}} (1 - \epsilon^4 \sin^2 \theta)^{-\frac{1}{2}} \\ &= \frac{b}{a^2} \left[\sum_{j=0}^{\infty} (-1)^j \binom{\frac{1}{2}}{j} e^{2j} \sin^{2j} \theta \right] \left[\sum_{k=0}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} \epsilon^{4k} \sin^{2k} \theta \right] \\ &= \frac{b}{a^2} \sum_{j=0}^{\infty} (-1)^j e^{2j} \sum_{k=0}^j \binom{\frac{1}{2}}{k} \binom{-\frac{1}{2}}{j-k} (2 - e^2)^{j-k} \sin^{2j} \theta \end{aligned} \quad (4.74)$$

The term $2\omega^2/\gamma$ in Equation (4.70) can also be expressed as a power series. For this purpose, Somigliana's formula for normal gravity in geodetic coordinates (Equation 2.34) first needs to be written in an alternative form

$$\gamma = \gamma_b (1 - e^2 \sin^2 \theta)^{-\frac{1}{2}} (1 - \epsilon^4 \sin^2 \theta)^{-\frac{1}{2}} (1 - e_\gamma^2 \sin^2 \theta) \quad (4.75)$$

where

$$e_\gamma^2 = 1 - \frac{\gamma_a}{\gamma_b}(1 - e^2)^{\frac{3}{2}} \quad (4.76)$$

which is obviously much smaller than 1, since $\gamma_a \approx \gamma_b$ and $e^2 \ll 1$. The inverse of the reference gravity needed in Equation (4.70) can therefore be written as the product of three fast-converging binomial series (see Appendix C)

$$\begin{aligned} \frac{1}{\gamma} &= \frac{1}{\gamma_b}(1 - e^2 \sin^2 \theta)^{\frac{1}{2}}(1 - \epsilon^4 \sin^2 \theta)^{\frac{1}{2}}(1 - e_\gamma^2 \sin^2 \theta)^{-1} \\ &= \frac{1}{\gamma_b} \left[\sum_{i=0}^{\infty} (-1)^i \binom{\frac{1}{2}}{i} e^{2i} \sin^{2i} \theta \right] \left[\sum_{j=0}^{\infty} (-1)^j \binom{\frac{1}{2}}{j} \epsilon^{4j} \sin^{2j} \theta \right] \left[\sum_{k=0}^{\infty} e_\gamma^{2k} \sin^{2k} \theta \right] \end{aligned} \quad (4.77)$$

The three binomial series can be combined using a Cauchy multiplication (see Appendix B). Combining the first two series and subsequently combining it with the third, results in a single power series expression

$$\begin{aligned} \frac{1}{\gamma} &= \frac{1}{\gamma_b} \left[\sum_{j=0}^{\infty} \sum_{i=0}^j (-1)^j \binom{\frac{1}{2}}{i} \binom{\frac{1}{2}}{j-i} (2 - e^2)^{j-i} e^{2j} \sin^{2j} \theta \right] \left[\sum_{k=0}^{\infty} e_\gamma^{2k} \sin^{2k} \theta \right] \\ &= \frac{1}{\gamma_b} \sum_{j=0}^{\infty} \sum_{k=0}^j (-1)^k e^{2k} e_\gamma^{2(j-k)} \sum_{i=0}^k \binom{\frac{1}{2}}{i} \binom{\frac{1}{2}}{k-i} (2 - e^2)^{k-i} \sin^{2j} \theta \end{aligned} \quad (4.78)$$

Pre-multiplying Equation (4.78) by the term $(R/r_e)^{n+1}$ and subsequently applying a Cauchy multiplication of the power series, gives

$$\left(\frac{R}{r_e}\right)^{n+1} \frac{1}{\gamma} = \frac{1}{\gamma_b} \sum_{j=0}^{\infty} \sum_{k=0}^j \alpha_{n,j-k} \sum_{l=0}^k (-1)^l e^{2l} e_\gamma^{2(k-l)} \sum_{p=0}^l \binom{\frac{1}{2}}{p} \binom{\frac{1}{2}}{l-p} (2 - e^2)^{l-p} \sin^{2j} \theta \quad (4.79)$$

Then, Equations (4.73), (4.74) and (4.78) can be inserted into Equation (4.70) and a simple summation of the three power series gives

$$\begin{aligned} \left(\frac{R}{r_e}\right)^{n+1} \frac{1}{\gamma} \frac{\partial \gamma}{\partial h} &= - \sum_{j=0}^{\infty} \sum_{k=0}^j \alpha_{n,j-k} \sum_{l=0}^k \left(\frac{b}{a^2} (-1)^k (2 - e^2)^{k-l} e^{2k} \left[\binom{\frac{3}{2}}{l} \binom{-\frac{3}{2}}{k-l} + \right. \right. \\ &\quad \left. \left. \binom{\frac{1}{2}}{l} \binom{-\frac{1}{2}}{k-l} \right] + 2 \frac{\omega^2}{\gamma_b} (-1)^l e^{2l} e_\gamma^{2(k-l)} \sum_{p=0}^l \binom{\frac{1}{2}}{p} \binom{\frac{1}{2}}{l-p} (2 - e^2)^{l-p} \right) \sin^{2j} \theta \end{aligned} \quad (4.80)$$

This series can be used to obtain an expression for the second term of the fundamental equation of physical geodesy (Equation 4.62) in terms of the solid SHCs of the disturbing potential. All dependence on latitude can be centred inside the SHFs by inserting the sinusoidal LWFs \bar{K}_{nm}^{ij} (Equation 3.31)

$$\frac{1}{\gamma} \frac{\partial \gamma}{\partial h} T = \sum_{n=0}^{\infty} \sum_{m=0}^n \bar{T}_{nm}^R \sum_{j=0}^{\infty} \sum_{i=-j}^j \varepsilon_{nmij} \bar{Y}_{n+2i,m} \quad (4.81)$$

where

$$\begin{aligned} \varepsilon_{nmij} = & -\overline{K}_{nm}^{2i,2j} \sum_{k=0}^j \alpha_{n,j-k} \sum_{l=0}^k \left(\frac{b}{a^2} (-1)^k (2-e^2)^{k-l} e^{2k} \left[\binom{\frac{3}{2}}{l} \binom{-\frac{3}{2}}{k-l} + \right. \right. \\ & \left. \left. \binom{\frac{1}{2}}{l} \binom{-\frac{1}{2}}{k-l} \right] + 2 \frac{\omega^2}{\gamma_b} (-1)^l e^{2l} e_{\gamma}^{2(k-l)} \sum_{p=0}^l \binom{\frac{1}{2}}{p} \binom{\frac{1}{2}}{l-p} (2-e^2)^{l-p} \right) \end{aligned} \quad (4.82)$$

A similar expression for the first term of the fundamental equation of physical geodesy, which is simply the normal derivative of the disturbing potential, was already derived in Section 4.1 and given in Equation (4.32).

Equations (4.32) and (4.81) can be summed to obtain a formula for the gravity anomaly in terms of the SHCs \overline{T}_{nm}^R

$$\Delta g = \sum_{n=0}^{\infty} \sum_{m=0}^n \overline{T}_{nm}^R \sum_{j=0}^{\infty} \sum_{i=-j-1}^{j+1} \left(\varepsilon_{nmij} + \frac{n+1}{R} \beta_{nmij} + \frac{1}{R} \gamma_{nmij} \right) \overline{Y}_{n+2i,m} \quad (4.83)$$

A rearrangement of the summation order (see Appendix B) is applied to ensure that the SHFs do not depend on the summation indices i and j , which facilitates the comparison with the surface SHCs of gravity anomalies

$$\Delta g = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{j=0}^{\infty} \sum_{i=-j-1}^{j+1} \left(\varepsilon_{n-2i,m,i,j} + \frac{n-2i+1}{R} \beta_{n-2i,m,i,j} + \frac{1}{R} \gamma_{n-2i,m,i,j} \right) \overline{T}_{n-2i,m}^R \overline{Y}_{nm} \quad (4.84)$$

This expression can, due to the orthogonality of the SHFs on the sphere, be directly compared to surface SHCs from the gravity anomalies on the ellipsoid (Equation 4.63)

$$\overline{\Delta g}_{nm}^{r_e} = \sum_{j=0}^{\infty} \sum_{i=-j-1}^{j+1} \left(\varepsilon_{n-2i,m,i,j} + \frac{n-2i+1}{R} \beta_{n-2i,m,i,j} + \frac{1}{R} \gamma_{n-2i,m,i,j} \right) \overline{T}_{n-2i,m}^R \quad (4.85)$$

Reshuffling the summation order once more provides the final formula containing a weighted summation over SHCs of the disturbing potential from $-\infty$ to ∞

$$\overline{\Delta g}_{nm}^{r_e} = \sum_{i=-\infty}^{\infty} \lambda_{nmi}(\Delta g, T) \overline{T}_{n-2i,m}^R \quad (4.86)$$

where the weights $\lambda_{nmi}(\Delta g, T)$ can be computed from another infinite summation

$$\lambda_{nmi}(\Delta g, T) = \sum_{j=|i|-1}^{\infty} \left(\varepsilon_{n-2i,m,i,j} + \frac{n-2i+1}{R} \beta_{n-2i,m,i,j} + \frac{1}{R} \gamma_{n-2i,m,i,j} \right) \quad (4.87)$$

The inverse of Equation (4.86) is a solution to the third geodetic BVP (or free-geodetic BVP) with the boundary condition given in Equation (2.52), since the solid SHCs \bar{T}_{nm}^R define the disturbing potential in the complete exterior of the ellipsoid.

The weights $\lambda_{nmi}(\Delta g, T)$ are very similar to the weights $\lambda_{nmi}(d_h T, T)$ that define the spectral relation between the disturbing potential and its first derivative with respect to the ellipsoidal normal, especially for high degrees n . The weights of high degrees are dominated by the term α_{nk} , which defines the continuation between the sphere and the ellipsoid. Therefore, the iterative method to find the inverse of Equation (4.86), similar to Equation (3.79), can only be guaranteed to converge for SHCs up to degree 520, just as in the case of the solutions to the ellipsoidal Dirichlet, Neumann and fixed geodetic BVP.

4.4 The Meissl scheme

The spectral relations between several gravity field quantities can be combined in a diagram, often called the Meissl scheme. The principles of the spectral theory of gravity quantities were formulated by Meissl (1971) and these were combined into a concise scheme by Rummel and Van Gelderen (1992, 1995). The main part of the scheme relates surface SHCs of the disturbing potential with coefficients of its first and second-order radial derivatives on spheres at different altitudes. This scheme has been extended to also include derivatives in latitudinal and longitudinal direction by using vector or tensor spherical harmonic expansions instead of the common scalar spherical harmonic expansion (Rummel and Van Gelderen, 1995). It has been further expanded to include vertical deflections and gravity disturbances (Grafarend, 2001) and forward gravity field modelling using a Newtonian integration (Kuhn and Featherstone, 2003), while Keller (2002) discusses the Meissl scheme when pseudo-differential operators are applied.

All of the above contributions only apply for gravity field functionals given on a sphere. An extension of the Meissl scheme to a not necessarily spherical geometry is provided by

Nutz (2002) using wavelet techniques. A second limitation of the original Meissl scheme is that it only considers radial, longitudinal and latitudinal derivatives. Derivatives with respect to the ellipsoidal normal are not included, even though this derivative features in many BVPs. Bölling and Grafarend (2005) derive relations between the disturbing potential and its derivatives in the ellipsoidal domain. They use the framework of ellipsoidal harmonics to derive the ellipsoidal spectral properties. Here, the Meissl scheme will be generalised to include relations between the disturbing potential and its normal derivatives on the ellipsoid, in the framework of the simple and most commonly used scalar spherical harmonic expansion.

4.4.1 The spherical Meissl scheme

The spectral relations between the disturbing potential and its first- and second-order radial derivative are shown in Equations (4.3) and (4.37). Furthermore, from the spherical harmonic series (Equation 2.70) it follows that the relation between two SHCs referenced to two spheres of different radii, say R and R_s , is given by a simple one-to-one relation

$$\overline{T}_{nm}^{R_s} = \left(\frac{R}{R_s}\right)^{n+1} \overline{T}_{nm}^R \tag{4.88}$$

The spectral relations between the different functionals can be shown in the form of a diagram (see Figure 4.3). This diagram was published in Rummel and Van Gelderen (1995) and Rummel (1997).

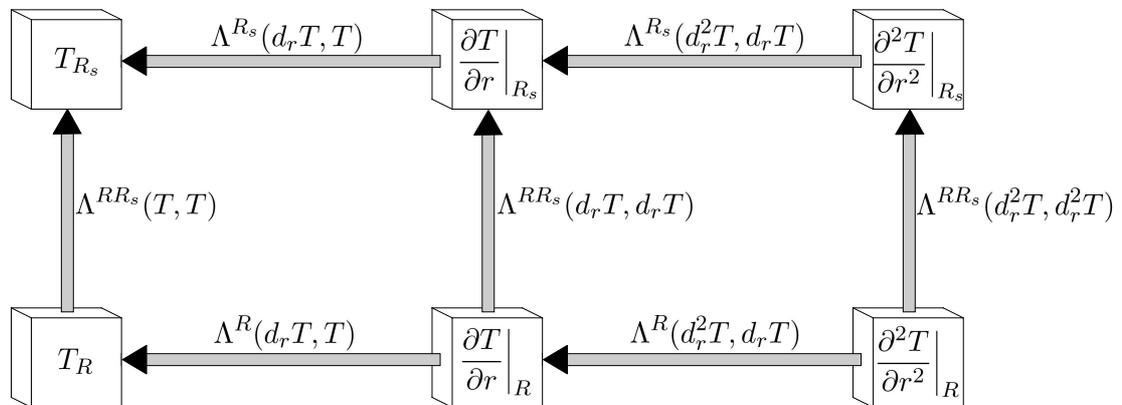


Figure 4.3: Spherical Meissl scheme (adapted from Rummel and Van Gelderen 1995)

matrix	eigenvalue	matrix	eigenvalue
$\Lambda^R(d_r T, T)$	$\frac{-R}{n+1}$	$\Lambda^{RR_s}(T, T)$	$\left(\frac{R}{R_s}\right)^{n+1}$
$\Lambda^{R_s}(d_r T, T)$	$\frac{-R_s}{n+1}$	$\Lambda^{RR_s}(dT, dT)$	$\left(\frac{R}{R_s}\right)^{n+2}$
$\Lambda^R(d_r^2 T, d_r T)$	$\frac{-R}{n+2}$	$\Lambda^{RR_s}(d^2 T, d^2 T)$	$\left(\frac{R}{R_s}\right)^{n+3}$
$\Lambda^{R_s}(d_r^2 T, d_r T)$	$\frac{-R_s}{n+2}$		

Table 4.1: Eigenvalues forming the diagonal elements in the matrices of the spherical Meissl scheme (Figure 4.3)

The arrows in Figure 4.3 point in the direction of decreased power in the high frequencies. It shows that the lower frequencies become more dominant for higher altitudes and for lower-order derivatives. The capital lambdas accompanying the arrows are the matrices that connect the vectors of surface SHCs. For example, the relation between the disturbing potential at a sphere of radius R and at a sphere of radius R_s reads

$$\mathcal{T}^{R_s} = \Lambda^{RR_s}(T, T)\mathcal{T}^R \quad (4.89)$$

where the matrix notation introduced in Section 3.3.2 is again used.

Equation (4.89) is, in fact, nothing else than Equation (4.88), although the surface SHCs are now put together to form a vector. Since Equation (4.88) provides a one-to-one relation between the coefficients, it follows that the matrix $\Lambda^{RR_s}(T, T)$ will be diagonal. Written out more explicitly, Equation (4.89) reads

$$\begin{pmatrix} \bar{T}_{00}^{R_s} \\ \vdots \\ \bar{T}_{nn}^{R_s} \end{pmatrix} = \begin{pmatrix} \left(\frac{R}{R_s}\right) & & \\ & \ddots & \\ & & \left(\frac{R}{R_s}\right)^{n+1} \end{pmatrix} \begin{pmatrix} \bar{T}_{00}^R \\ \vdots \\ \bar{T}_{nn}^R \end{pmatrix} \quad (4.90)$$

The elements on the diagonal of $\Lambda^{RR_s}(T, T)$ can be viewed as eigenvalues, since Equation (4.90) represents an eigenvalue analysis of, in this case, the upward-continuation or Poisson operator (e.g., Rummel and Van Gelderen, 1995). All matrices Λ in Figure 4.3 are diagonal, and the eigenvalues that form the diagonal elements are provided in

Table 4.1. They follow simply from Equations (4.3), (4.37) and (4.88). In the case that the quantities are not given on a sphere, but on an ellipsoid, the matrices will no longer be diagonal, since no one-to-one relation between the surface spherical harmonic coefficients exists. However, also in this case the transformation matrix can be determined, as is shown next.

4.4.2 The ellipsoidal Meissl scheme

The relations derived in Sections 4.1 and 4.2 allow for the construction of an ellipsoidal Meissl scheme, because the spectral relations are now known for the disturbing potential and its derivatives. A scheme very similar to the spherical Meissl scheme (Figure 4.3) can therefore be set up, where the spherical radii R and R_s are replaced by the ellipsoidal radii r_e and r_s , and where the radial derivatives are replaced by derivatives with respect to the ellipsoidal normal. This ellipsoidal Meissl scheme is shown in Figure 4.4. The two ellipsoids with radii r_e and r_s do not necessarily need to have the same flattening or eccentricity, although from a numerical point of view the eccentricity of both should be small to avoid large truncation errors in the binomial series (see Appendix C).

The matrices that connect the various quantities can be found from the final transformation formulas derived in Sections 4.1 and 4.2 (Equations 4.33 and 4.57). However,

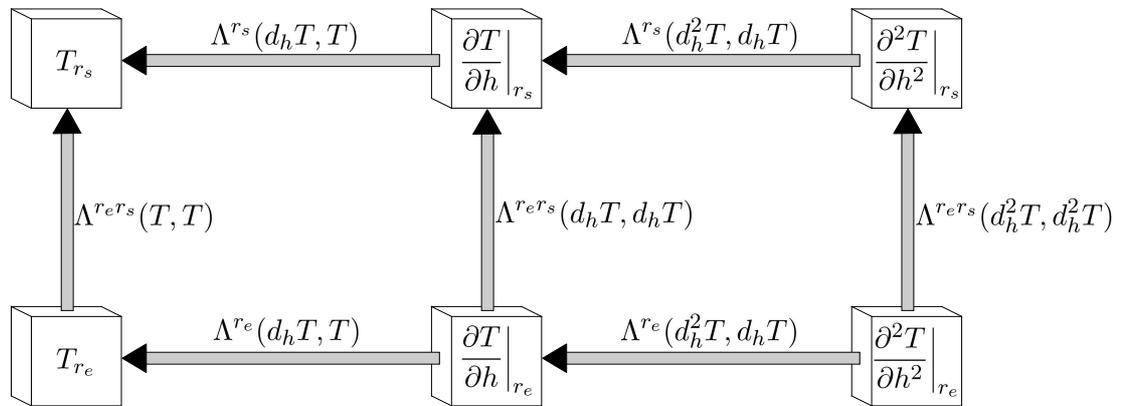


Figure 4.4: Ellipsoidal Meissl scheme

these formulas relate the surface SHCs of normal derivatives on the ellipsoid to solid SHCs \overline{T}_{nm}^R , which feature in the spherical Meissl scheme, but not in the ellipsoidal Meissl scheme. The matrices of the ellipsoidal Meissl scheme can therefore only be found in two steps, always first transforming from one ellipsoidal quantity to the solid SHCs before a transformation to another ellipsoidal quantity can be performed. The relation between the ellipsoidal and the spectral domain is given in Equation (3.73) for the disturbing potential and by Equations (4.33) and (4.57) for the first- and second-order derivatives, respectively. The spherical and ellipsoidal Meissl schemes can thus be connected, as is shown in Figure 4.5.

Equation (3.73) becomes for the disturbing potential T in matrix notation

$$\mathcal{T}^{r_e} = \Lambda^{r_e R}(T, T)\mathcal{T}^R \quad (4.91)$$

Furthermore, Equation (4.33) that gives the spectral relation between the disturbing potential and its first normal derivative on the ellipsoid reads, in matrix notation

$$d\mathcal{T}^{r_e} = \Lambda^{r_e R}(T, dT)\mathcal{T}^R \quad (4.92)$$

The vectors \mathcal{T}^R and $d\mathcal{T}^{r_e}$ contain all surface SHCs up to a certain degree and order of the disturbing potential T and its normal derivative $\partial T/\partial h$, respectively.

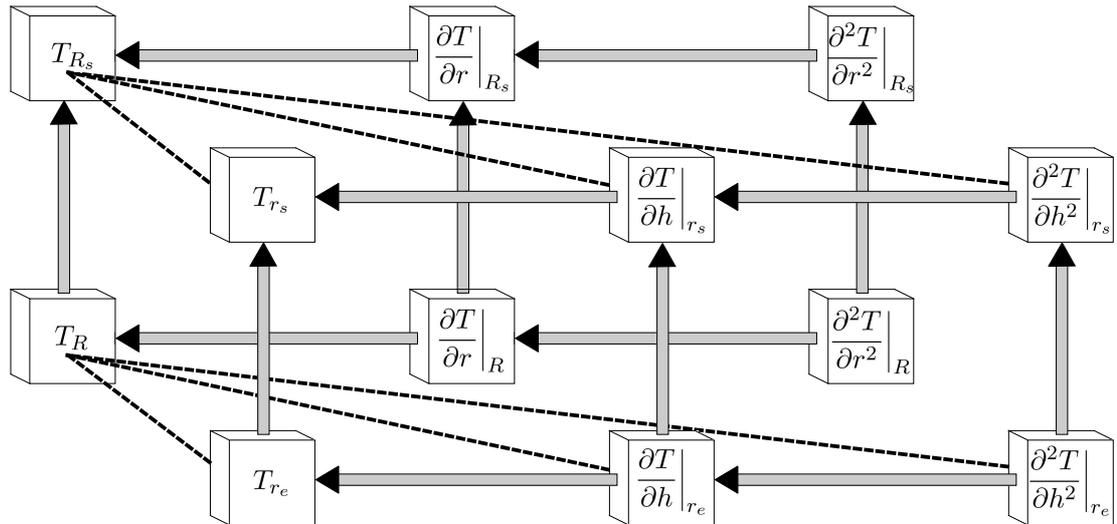


Figure 4.5: Connections between the spherical and ellipsoidal Meissl schemes (black dashed lines), given by Equations (3.73), (4.33) and (4.57)

As long as both the disturbing potential and its derivative are expanded into a series of equal maximum degree and order, the matrix $\Lambda^{r_e R}(T, dT)$ will be square. The inverse of Equation (4.93) can thus simply be found by a matrix inversion

$$\mathcal{T}^R = \Lambda^{Rr_e}(dT, T)d\mathcal{T}^{r_e} \quad (4.93)$$

where

$$\Lambda^{Rr_e}(dT, T) = [\Lambda^{r_e R}(T, dT)]^{-1} \quad (4.94)$$

Now, Equation (4.93) can be inserted into Equation (4.91), and the matrix $\Lambda^{r_e}(dT, T)$ in the ellipsoidal Meissl scheme (Figure 4.4) follows from a matrix product

$$\mathcal{T}^{r_e} = \Lambda^{r_e}(dT, T)d\mathcal{T}^{r_e} \quad (4.95)$$

where

$$\Lambda^{r_e}(dT, T) = \Lambda^{r_e R}(T, T)\Lambda^{Rr_e}(dT, T) \quad (4.96)$$

A similar procedure can be followed to acquire the other matrices in the ellipsoidal Meissl scheme in Figure 4.4. For example, the second derivative of the disturbing potential with respect to the ellipsoidal normal is related to the disturbing potential itself by Equation (4.57), which in matrix notation reads

$$\mathcal{T}^R = \Lambda^{Rr_e}(d^2T, T)d^2\mathcal{T}^{r_e} \quad (4.97)$$

Taking the inverse relation and inserting Equation (4.93) gives the matrix $\Lambda^{r_e}(d^2T, dT)$ in the ellipsoidal Meissl scheme

$$\Lambda^{r_e}(d^2T, dT) = [\Lambda^{Rr_e}(d^2T, T)]^{-1}\Lambda^{Rr_e}(dT, T) \quad (4.98)$$

As in the original Meissl diagram (Figure 4.3), the connection between quantities at different altitudes can also be established in the ellipsoidal case. However, continuation from one surface to another cannot be derived from surface spherical harmonic expansions only. The surface spherical harmonic expansion at one ellipsoid should therefore first be transformed into a solid spherical harmonic expansion, and subsequently be back-transformed at the level of the other ellipsoid. For example, the transformation from $\bar{T}_{nm}^{r_e}$ to $\bar{T}_{nm}^{r_s}$ can be found by first transforming to \bar{T}_{nm}^R using Equation (4.91). Then, the inverse transformation, where r_e is replaced by r_s can be inserted

$$\mathcal{T}^{r_e} = \Lambda^{r_e r_s}(T, T)\mathcal{T}^{r_s} \quad (4.99)$$

where

$$\Lambda^{r_e r_s}(T, T) = \Lambda^{r_e R}(T, T) [\Lambda^{r_s R}(T, T)]^{-1} \quad (4.100)$$

The upward-continuation of the first- and second-order derivatives of the disturbing potential from r_e to r_s can be derived similarly, by a forward and backward transformation.

From the derivations above, it can be deduced that all quantities in the full Meissl scheme are connected in a similar fashion, with the general formula being

$$d^i \mathcal{T}^{\mathcal{R}_1} = \Lambda^{\mathcal{R}_1 \mathcal{R}_2}(d^j T, d^i T) d^j \mathcal{T}^{\mathcal{R}_2} \quad \text{with } 0 \leq i \leq 2, \quad 0 \leq j \leq 2 \quad (4.101)$$

where the radii \mathcal{R}_1 and \mathcal{R}_2 can be either R , R_s , r_e or r_s . To somewhat shorten the notation, the superscript \mathcal{R}_2 is in some of the above formulas dropped from $\Lambda^{\mathcal{R}_1 \mathcal{R}_2}(d^j T, d^i T)$ if it is equal to \mathcal{R}_1 .

The matrix $\Lambda^{\mathcal{R}_1 \mathcal{R}_2}(d^j T, d^i T)$ that connects the two quantities can be found from a multiplication of known matrices by a transformation via the quantity $d^k T$ on a surface with radius \mathcal{R}_3

$$\Lambda^{\mathcal{R}_1 \mathcal{R}_2}(d^j T, d^i T) = \Lambda^{\mathcal{R}_1 \mathcal{R}_3}(d^k T, d^i T) \Lambda^{\mathcal{R}_3 \mathcal{R}_2}(d^j T, d^k T) \quad (4.102)$$

with $0 \leq i \leq 2, \quad 0 \leq j \leq 2, \quad 0 \leq k \leq 2$

where it also holds that

$$\Lambda^{\mathcal{R}_1 \mathcal{R}_2}(d^j T, d^i T) = [\Lambda^{\mathcal{R}_2 \mathcal{R}_1}(d^i T, d^j T)]^{-1} \quad (4.103)$$

This allows the computation of any transformation matrix between two quantities in the combined spherical-ellipsoidal Meissl scheme in Figure 4.5 from known matrices. Equations (4.96), (4.98) and (4.100) are examples thereof.

Using Equation (4.102) and the relations between surface SHCs of various functionals and the solid SHCs of the disturbing potential in Equations (3.73), (4.33), (4.57) and (4.88), all matrices that connect the functionals in the ellipsoidal Meissl scheme (Figure 4.4) can be derived. However, in all cases a transformation from and to the sphere is required. This means that in all cases, a forward and backward transformation is

required, and the divergence of the iterative backward transformation will limit the practical use of the relations. However, there is an alternative and more efficient way to find the matrices that connects two quantities on the same ellipsoidal surface in the ellipsoidal Meissl scheme (Figure 4.4). This will be shown here for the relation between the surface SHCs of the disturbing potential and the SHCs of its normal derivative.

Comparison of the surface and solid spherical harmonic expansions (Equations 2.81 and 2.84) shows that the relation between the solid and surface SHCs for a certain latitude is given by

$$T_{nm}^R = \left(\frac{r_e}{R}\right)^{n+1} \bar{T}_{nm}^{r_e} \quad (4.104)$$

This relation is not very useful for a coefficient transformation, since the desired set of coefficients would be different for every different latitude, and only applicable for synthesis at that specific latitude. For this reason, the transformation between both sets of coefficients derived in Section 3.3 first moves the dependence on latitude into the spherical harmonic base functions \bar{Y}_{nm} . In this case, Equation (4.104) will prove useful nonetheless, because it can be inserted into Equation (4.16), yielding

$$\left.\frac{\partial T}{\partial h}\right|_{r_e} = -\frac{1}{R} \sum_{n=0}^{\infty} \frac{R}{r_e} \sum_{m=-n}^n \bar{T}_{nm}^{r_e} \left((n+1) \cos \phi \bar{Y}_{nm} + \sin \phi \frac{\partial \bar{Y}_{nm}}{\partial \theta} \right) \quad (4.105)$$

The only difference with Equation (4.16), besides the change from solid to surface SHCs $\bar{T}_{nm}^{r_e}$, that the term (R/r_e) is not raised to the power $n+2$.

A relation between the surface SHCs of the disturbing potential and those of its derivative can therefore be derived along exactly the same lines as those laid out in Section 4.1.3, and the only difference between the weights λ_{nmi} is that the functions $\alpha_{n+1,k}$ that appear in the expressions of the functions β_{nmij} and γ_{nmij} (Equations 4.23 and 4.29) are changed to $\alpha_{0,k}$, which according to Equation (3.69) become

$$\alpha_{0,k} = \frac{c}{\sqrt{1-e^2}} (-1)^k \binom{\frac{1}{2}}{k} e^{2k} \quad (4.106)$$

The replacement of $\alpha_{n+1,k}$ by $\alpha_{0,k}$ greatly facilitates the computation, because now the binomial coefficients are no longer dependent on the degree n , which means that the weights of λ_{nmi} of $i \neq 0$ (the off-diagonal weights) do not significantly increase for higher degrees n . The matrix containing the weights is therefore strongly diagonally

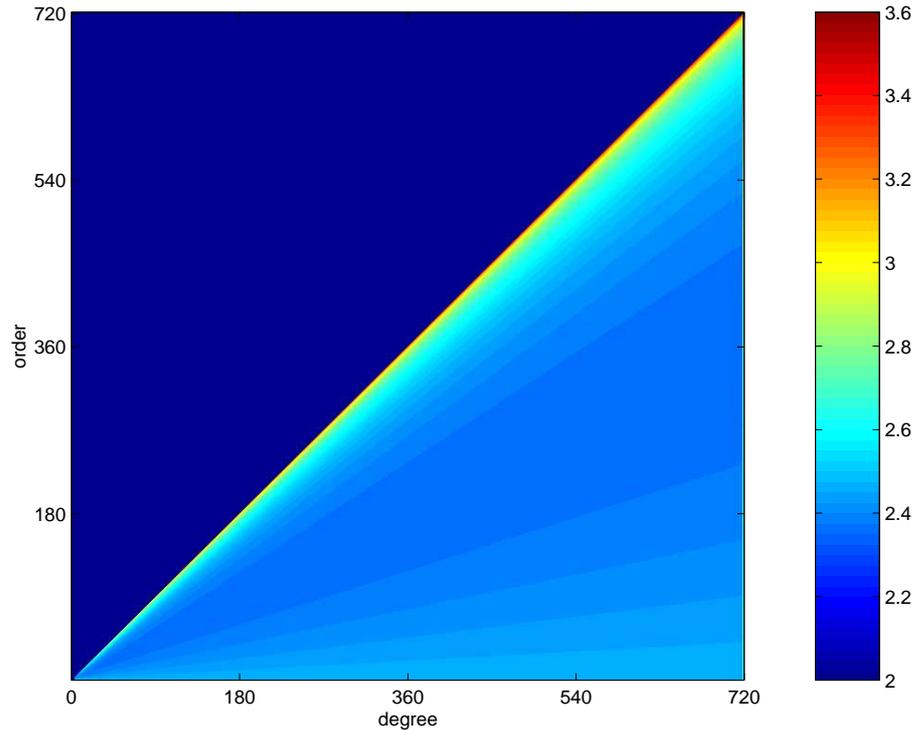


Figure 4.6: Logarithm of the ratio between $|\lambda_{nm0}(d_h T, T)|$ and $\sum_{i=-50, i \neq 0}^{50} |\lambda_{nmi}(d_h T, T)|$ for a transformation between surface SHCs of a function and its normal derivative on the ellipsoid ($0 \leq n \leq 720$)

dominant, i.e., obeying Equation (3.78), even when SHCs of very high degree and order are taken into account, which is shown in Figure 4.6.

This not only makes the forward transformation more efficient, it also ensures that iterative procedures such as Jacobi iteration, or alternatively Gauss-Seidel or SOR methods (e.g., Strang, 1986), in the backward transformation will always converge. The rate of convergence of these iterative methods will be high, due to the strong diagonal dominance displayed in Figure 4.6, with every element on the diagonal being at least two orders of magnitude larger than the sum of all other elements in the same row.

The spectral relation between the surface SHCs of the disturbing potential and its second derivative can be obtained along the same principles, and all horizontal transformations in the ellipsoidal Meissl scheme (Figure 4.4) can thus be performed very efficiently in both directions. Transformations from the sphere to the ellipsoid are less

efficient due to the increase in the off-diagonal terms for higher degrees n , but only contain a forward transformation and do therefore not contain the problems encountered in inverse transformations. Only transformation from an ellipsoidal surface to another surface (either the sphere or another ellipsoid) requires the computation of the inverse of a non-diagonally dominant matrix, which can only be performed efficiently up to degree and order 520 for an ellipsoid with the eccentricity of the Earth ($e^2 \approx 0.0067$).

4.5 Summary

Based on the transformation from solid to surface spherical harmonic coefficients (SHCs) of a function described on an ellipsoid derived in Chapter 3, the transformation to surface SHCs of the first- and second-order derivatives of the function were derived as well, principally providing solutions to the Neumann and second-order ellipsoidal BVPs. In addition, the spectral ellipsoidal relation between solid SHCs of the disturbing potential and surface spherical harmonic coefficients of the gravity anomaly was derived based on the rigorous fundamental equation of physical geodesy (Equation 4.62). The latter provides a solution to the third geodetic boundary-value problem. In all of the transformations derived here, the SHCs of the unknown quantity are expressed as a weighted summation over SHCs of the known quantity with equal order m .

The newly derived relations were combined to form an ellipsoidal Meissl scheme, as an extension of the spherical Meissl scheme introduced by Rummel and Van Gelderen (1995). This scheme provides the spectral relation between the disturbing potential and its first- and second-order derivatives on any ellipsoid, also allowing transformations from one ellipsoid to another. The transformations are expressed as a multiplication of a block-diagonal matrix with a vector containing surface SHCs of one quantity, yielding a vector containing surface SHCs of a second quantity.

Admittedly, the ellipsoidal Meissl scheme is more complicated than the spherical Meissl scheme, where a one-to-one relation exists between the SHCs (i.e., the matrix becomes

diagonal in spherical and constant radius approximation). From a numerical point of view, the main difficulties in the ellipsoidal Meissl scheme occur in the transformation from the ellipsoid to another surface, which requires inversion of a non-diagonally dominant matrix, and can only be solved efficiently up to degree and order 520.

5. ANALYTIC APPROACHES TO THE CONSTRUCTION OF A GLOBAL SPHERICAL HARMONIC GEOPOTENTIAL MODEL

One application where the formulas derived in Chapter 4 can be used is the analytic computation of spherical harmonic geopotential coefficients, i.e., a spherical harmonic analysis, from terrestrial gravity anomalies. Several approaches to the computation of geopotential coefficients from gravity anomalies can be taken, and these are discussed in Section 5.1. Three of these approaches provide very similar formulas, namely as a weighted summation over coefficients that are computed under spherical and constant radius approximation. These are investigated in more detail in Sections 5.2 to 5.4.

One of these so-called weighted summation methods, based on the spectral relations derived in Section 4.3 has never been applied before for this purpose. The other two methods are here derived more rigorously than before, using the new relations among associated Legendre functions (ALFs) derived in Section 3.2. The newly derived formulas are compared numerically to various existing methods, in Section 5.5.

5.1 Methodologies

The computation of solid spherical harmonic coefficients (SHCs) of the disturbing potential can be performed in various ways that can roughly be divided into two categories (Lemoine et al., 1998):

1. Least-squares estimation/collocation,
2. Numerical quadrature.

Both methods are explained in brief below. More detailed derivations can be found in, e.g., Colombo (1981) or Sneeuw (1994), and a discussion on their advantages and disadvantages is given by, e.g., Pavlis et al. (1996). The least-squares techniques are

based on a solution of a system of normal equations, and the numerical quadrature approach is based on analytic solutions of a geodetic boundary-value problem (BVP). The latter forms the primary focus of this Chapter. Several strategies can be followed to derive a numerical quadrature procedure, and the most commonly used is the ellipsoidal harmonic method discussed in Section 5.1.3. Three alternative methods are derived and discussed in Sections 5.2 to 5.4.

5.1.1 Least-squares estimation

The least-squares estimation and collocation (LSC) techniques compute all coefficients up to a certain degree and order simultaneously from a set of gravity-related observations, solving a system of linearised normal equations (e.g., Lemoine et al., 1998; Sansò and Tscherning, 2003). These techniques are formally equivalent in the case of band-limited data (e.g., Moritz, 1989), and the principles of least-squares estimation are briefly explained here.

The least-squares estimation method requires that several ellipsoidal corrections are applied to the gravity anomalies prior to the computation of solid SHCs of the disturbing potential (Pavlis, 1988; Rapp and Pavlis, 1990), so that the spectral relation between gravity anomalies and disturbing potential in spherical approximation (Equation 4.68) can be applied

$$\widetilde{\Delta g}(\theta, \lambda) = \frac{1}{R} \sum_{n=0}^{\infty} (n-1) \left(\frac{R}{r_e}\right)^{n+1} \sum_{m=-n}^n \overline{T}_{nm}^R \overline{Y}_{nm}(\theta, \lambda) \quad (5.1)$$

Given an equiangular grid of mean gravity anomalies $\overline{\Delta g}$, Equation (5.1) forms a set of linear equations (Colombo, 1981)

$$\overline{\Delta g}(r_{e,i}, \theta_i, \lambda_j) = \frac{1}{\Delta\sigma_i R} \sum_{n=2}^{N_{\max}} (n-1) \left(\frac{R}{r_{e,i}}\right)^{n+1} \sum_{m=-n}^n \overline{T}_{nm}^R \int_{\theta_i}^{\theta_{i+1}} \int_{\lambda_j}^{\lambda_{j+1}} \overline{Y}_{nm} d\lambda \sin\theta d\theta \quad (5.2)$$

where $\Delta\sigma_i$ is the surface area of one grid block

$$\Delta\sigma_i = \frac{1}{(\lambda_{j+1} - \lambda_j)(\cos\theta_i - \cos\theta_{i+1})} \quad (5.3)$$

The only unknowns in Equation (5.2) are the SHCs \bar{T}_{nm}^R , which can be computed via a least-squares adjustment (e.g., Lemoine et al., 1998)

$$\mathcal{T}^R = (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} \Delta \mathcal{G}^{r_e} \quad (5.4)$$

where \mathcal{T}^R is a vector containing the SHCs \bar{T}_{nm}^R , $\Delta \mathcal{G}^{r_e}$ is a vector containing the block-averaged gravity anomalies, \mathbf{A} is the design matrix that contains the remaining terms of Equation (5.2) and \mathbf{P} is the variance-covariance matrix of the observations.

The main advantage of the least-squares approach, besides its simplicity, is that it allows for the simultaneous computation of internal error estimates, which after calibration can serve as a measure of the accuracy of the coefficients. It should, however, be noted that not all global gravity field models computed using the least-squares approach are calibrated, and that error estimates are not always released in full.

A disadvantage of the least-squares approach is that large systems of linear equations become ill-conditioned, which necessitates special treatment to compute the least-squares solution. This problem is more serious in the case of satellite data, as opposed to terrestrial data (e.g., Lemoine et al., 1998). The solution can be improved when a regularisation method, such as Kaula's regularisation or Tikhonov regularisation (Tikhonov and Arsenin, 1977), is applied (e.g., Kusche and Klees, 2002).

A second disadvantage is the required computation time for coefficients of high degree and order, since it requires the inversion of a matrix with a dimension equal to the total number of coefficients. For example, for computation of coefficients to degree and order 360, a matrix of dimension 65,341 needs to be inverted. This disadvantage can be overcome by applying some restrictions to the grid sampling intervals and data weights (Colombo, 1981), at the expense of the rigour exercised in the development of the model (Lemoine et al., 1998), or by applying a semi-analytical approach (Sneeuw, 2000). Both these techniques make the design matrix block-diagonal, which facilitates its inversion. Increased computer power due to the advent of technology also slightly reduces this disadvantage.

A third disadvantage of the least-squares estimation procedure is that it is suscepti-

ble to aliasing effects in the high degrees, decreasing the accuracy of the computed geopotential coefficients (Gleason, 1989a; Jekeli, 1996). This results from the fact that the data are assumed band-limited, while in reality they are not, and that the least-squares estimation technique is unable to compute coefficients beyond the Nyquist degree defined by the resolution of the data. Analytic methods discretised into a numerical quadrature procedure can be employed to compute the high-frequency part of the spectrum. This strategy was, for example, followed in the construction of the EGM96 global geopotential model (Lemoine et al., 1998), and will also be followed in the construction of EGM06 (Holmes, 2004, pers. comm.).

5.1.2 Numerical quadrature

The numerical quadrature approach is based on the expression of surface SHCs as an integral over functions values on the ellipsoid (Equation 2.85)

$$\overline{\Delta g}_{nm}^{r_e} = \frac{1}{4\pi} \int_{\sigma} \Delta g \overline{Y}_{nm} d\sigma \quad (5.5)$$

Implementing the spherical spectral relation between the gravity anomalies and the disturbing potential (Equation 4.68), an expression for the solid SHCs \overline{T}_{nm}^R emerges

$$\tilde{T}_{nm}^R = \frac{R}{4\pi(n-1)} \int_{\sigma} \Delta g \overline{Y}_{nm} d\sigma \quad (5.6)$$

where the tilde over T_{nm}^R indicates that a spherical approximation is again applied. Equation (5.6) can be discretised in a manner suggested by Colombo (1981) for an equiangular grid of area means of gravity anomalies

$$\tilde{T}_{nm}^R = \frac{R}{4\pi(n-1)} \sum_{i=0}^{N_{\max}-1} \frac{1}{q_{in}} \sum_{j=0}^{2N_{\max}-1} \overline{\Delta g}_{ij} \int_{\theta_i}^{\theta_{i+1}} \int_{\lambda_j}^{\lambda_{j+1}} \overline{Y}_{nm} \sin \theta d\theta d\lambda \quad (5.7)$$

where q_{in} are smoothing factors described by Pellinen (1966).

Equation (5.7) cannot be used when high accuracy is required, due to the presence of the spherical approximation (e.g., Rapp and Pavlis, 1990). There are several ways to obtain a higher level of accuracy avoiding or correcting for the spherical approximation. Four

methods can be distinguished, based on the four solutions to the ellipsoidal Dirichlet BVP discussed in Section 3.1. The first method is the ellipsoidal harmonics method, which differs from the other three methods, because it utilises an ellipsoidal harmonic expansion to account for the fact that the gravity anomalies in Equation (2.52) are defined on the ellipsoid, whereas the other methods, described in Section 5.1.4, avoid the use of ellipsoidal harmonics.

5.1.3 The ellipsoidal harmonics method

In the ellipsoidal harmonics method, a fully normalised ellipsoidal harmonic expansion is applied to the gravity anomalies on the surface of the ellipsoid in ellipsoidal coordinates, multiplied by the ellipsoidal radius r_e (Gleason, 1988)

$$r_e \Delta g(u, \beta, \lambda) = R \sum_{n=0}^{\infty} \sum_{m=-n}^n \frac{\overline{Q}_{nm}(i\frac{u}{E})}{\overline{Q}_{nm}(i\frac{b}{E})} \widehat{\Delta g}_{nm}^u \overline{Y}_{nm}(\beta, \lambda) \quad (5.8)$$

The multiplication with r_e is required because the gravity anomaly function is not harmonic in the exterior space, but under the spherical definition of the gravity anomaly (Equation 2.67), $r_e \Delta g$ is (e.g., Heiskanen and Moritz, 1967, p. 90). The definition of the gravity anomaly (Equation 2.67) is here assumed to be valid not only on the surface of the ellipsoid, but in the whole exterior space. This extension from 2D to 3D is questioned by Vaníček et al. (2004), who suggest that it may not be possible to derive the 3D gravity anomaly (Equation (2.52) where point Q can be located anywhere in the space exterior to the ellipsoid), from the 2D gravity anomaly (Equation 2.49).

The product of the ellipsoidal radius and the gravity anomaly $r_e \Delta g$ may also be expanded into a series of solid spherical harmonic functions (SHFs)

$$r_e \Delta g(r, \theta, \lambda) = R \sum_{n=0}^{\infty} \sum_{m=-n}^n \left(\frac{R}{r}\right)^{n+1} \overline{\Delta g}_{nm}^R \overline{Y}_{nm}(\theta, \lambda) \quad (5.9)$$

The SHCs $\overline{\Delta g}_{nm}^R$ can be computed from the ellipsoidal harmonic coefficients $\widehat{\Delta g}_{nm}^u$ via the transformation derived by Jekeli (1988)

$$\overline{\Delta g}_{nm}^R = \sum_{k=0}^{\text{int}(\frac{n-|m|}{2})} \frac{L_{nmk}}{\overline{S}_{n-2k,m}(\frac{b}{E})} \widehat{\Delta g}_{n-2k,m}^u \quad (5.10)$$

where

$$L_{nmk} = \left(\frac{E}{R}\right)^{2k} \frac{(-1)^k (n-k)! (2n-4k+1)!}{k! (n-2k)! (2n-2k+1)!} \sqrt{\frac{(2n-4k+1)(n-m)!(n+m)!}{(2n+1)(n-2k-m)!(n-2k+m)!}} \quad (5.11)$$

and \bar{S}_{nm} is a normalisation of the ALFs of the second kind \bar{Q}_{nm} that avoids the use of complex numbers

$$\bar{S}_{nm}\left(\frac{u}{E}\right) = \left(\frac{R}{E}\right)^{n+1} \frac{i^{n+1} (2n+1)!}{2^n n!} \sqrt{\frac{\delta_{m0} + 1}{2(2n+1)(n-m)!(n+m)!}} \bar{Q}_{n,|m|}\left(i\frac{u}{E}\right) \quad (5.12)$$

which can alternatively be written in the form of a hypergeometric expansion (Jekeli, 1988).

This transformation allows for the computation of solid SHCs of the gravity anomaly from ellipsoidal harmonic coefficients that follow from an integration over the sphere

$$\widehat{\Delta g}_{nm}^u = \frac{1}{4\pi R} \int_{\sigma} r_e \Delta g(b, \beta, \lambda) \bar{Y}_{nm}(\beta, \lambda) d\sigma \quad (5.13)$$

The solid SHCs of the disturbing potential can finally be obtained using the spherical spectral relation between gravity anomalies and disturbing potential (Equation 4.68). A spherical approximation of the boundary condition is thus still required, and the ellipsoidal harmonic transformation only accounts for the fact that the gravity anomalies are given on an ellipsoid and not on a sphere. The approximation of the boundary condition, i.e., the use of Equation (2.67) instead of (2.52), is usually rectified by applying ellipsoidal corrections to the gravity anomalies prior to processing (Pavlis, 1988; Rapp and Pavlis, 1990). The spectral relation between gravity anomalies and disturbing potential on the ellipsoid, derived in Section 4.3, can alternatively be used.

The final formulas for the computation of geopotential coefficients from an equiangular grid of block mean gravity anomalies $\widetilde{\Delta g}$ that are corrected for the use of the spherical boundary condition (Pavlis, 1988; Rapp and Pavlis, 1990) follow from Equations (5.7), (5.10) and (5.13)

$$\begin{aligned} \bar{T}_{nm}^R &= \frac{1}{4\pi} \sum_{i=0}^{N_{\max}-1} r_{e,i} \sum_{k=0}^{\text{int}(\frac{n-|m|}{2})} \frac{L_{nmk}}{(n-2k-1)q_{i,n-2k} \bar{S}_{n-2k,m}\left(\frac{b}{E}\right)} \\ &\times \sum_{j=0}^{2N_{\max}-1} \widetilde{\Delta g}(r_{e,i}, \theta, \lambda) \int_{\theta_i}^{\theta_{i+1}} \int_{\lambda_j}^{\lambda_{j+1}} \bar{Y}_{nm} \sin \theta d\theta d\lambda \end{aligned} \quad (5.14)$$

A distinct advantage of this ellipsoidal harmonics method over the least-squares estimation procedure is that it allows for the computation of estimates of the SHCs \bar{T}_{nm}^R above the maximum degree defined by the resolution of the grid of gravity anomalies. Moreover, the ellipsoidal harmonics method avoids the problems associated with the inversion of an ill-conditioned matrix that burdens the least-squares estimation procedure. The same advantages over least-squares estimation also hold for all alternative numerical quadrature methods, described next.

5.1.4 Alternative numerical quadrature methods

Three alternatives to the ellipsoidal harmonics method described in Section 5.1.3 that do not involve the use of ellipsoidal harmonics all represent the geopotential coefficients as a weighted summation over spherically approximated geopotential coefficients. The spherically approximated coefficients are defined as the coefficients computed using the spherical and constant radius approximation, i.e., an approximation of the boundary condition in Equation (2.52) by Equation (2.68) and neglecting the term $(R/r)^{n+1}$ in the spherical harmonic expansion (Equation 2.81), which results in Equation (5.7).

An important difference between these three methods and the ellipsoidal harmonics method is that the gravity anomalies do not need to be corrected before processing, because the three alternative methods are based on the rigorous boundary condition (Equation 2.52) instead of the spherical approximation in Equation (2.68). In addition, for the three alternative methods, the boundary condition relating the gravity anomaly to the disturbing potential only needs to hold on the ellipsoid, and not in the whole exterior space. As mentioned in Section 5.1.3, this extension from 2D to 3D has been questioned by Vaníček et al. (2004).

The three methods for the computation of spherical harmonic geopotential coefficients \bar{T}_{nm}^R from gravity anomalies Δg are schematically shown in Figure 5.1. The classical method upward-continues the gravity data from the ellipsoid to the sphere circumscribing the ellipsoid ($R = a$), and this is here called the upward-continuation method. The second method is the ellipsoidal integration method, where the geopotential coef-

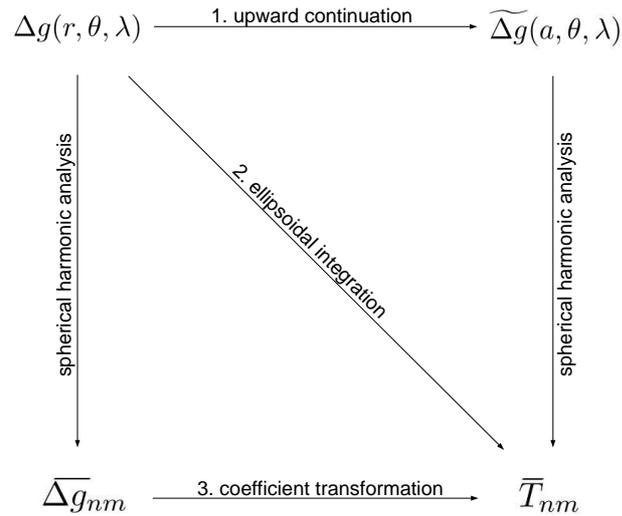


Figure 5.1: Schematic overview of three methods to compute spherical harmonic geopotential coefficients from gravity anomalies on the ellipsoid

ficients are computed from an integration over the ellipsoid, using a formula based on Green's second identity (Equation 2.41). The third method, here called the coefficient transformation method, has never before been proposed for the purpose of computation of geopotential coefficients. It first involves an expansion of the gravity anomalies on the ellipsoid into a series of surface SHFs, and then transforms these into geopotential coefficients, using the transformation derived in Section 4.3.

5.2 Method 1: Upward continuation

The oldest approach to the analytic computation of geopotential coefficients \overline{T}_{nm}^R from gravity anomalies on the ellipsoid was derived by Rapp (1977) and Cruz (1985, 1986) and refined by Petrovskaya (1995) and Petrovskaya et al. (2001). It was, for example, used in the construction of the OSU86E/F geopotential models (Rapp and Cruz, 1986). The method has so far only been derived involving approximations to the boundary condition of order e^2 , but this was proven insufficient by Gleason (1988). Here, the rigorous derivation is presented.

5.2.1 Solution strategy

Geopotential coefficients can be computed from an integration over a sphere, if the disturbing potential is known on the surface of that sphere

$$\bar{T}_{nm}^R = \frac{1}{4\pi} \int_{\sigma} T(R, \theta, \lambda) \bar{Y}_{nm} d\sigma \quad (5.15)$$

From the spherical spectral relation between the disturbing potential and the gravity anomaly (Equation 4.68), it can be derived that the geopotential coefficients can also be computed from an integration over the gravity anomalies

$$\bar{T}_{nm}^R = \frac{R}{n-1} \frac{1}{4\pi} \int_{\sigma} \widetilde{\Delta g}(R, \theta, \lambda) \bar{Y}_{nm} d\sigma \quad (5.16)$$

where the tilde over the gravity anomaly Δg indicates that a spherical approximation was applied. Thus, the approximated gravity anomalies $\widetilde{\Delta g}(R, \theta, \lambda)$ on the sphere are needed, whereas gravity anomalies on the ellipsoid are provided.

A correction to the gravity anomalies needs to be applied to be able to use them in Equation (5.16)

$$\widetilde{\Delta g}(R, \theta, \lambda) = \Delta g(r_e, \theta, \lambda) + \delta\Delta g \quad (5.17)$$

The term $\delta\Delta g$ contains corrections for the spherical boundary condition ε_e , as well as a continuation of the gravity anomaly from the ellipsoid to the sphere ε_c (Cruz, 1986)

$$\delta\Delta g = \varepsilon_e + \varepsilon_c \quad (5.18)$$

where the correction term ε_e equals the difference between the rigorous fundamental equation of physical geodesy (Equation 2.52) and the spherically approximated fundamental equation of physical geodesy (Equation 2.67)

$$\varepsilon_e = \frac{1}{\gamma} \frac{\partial \gamma}{\partial h} T - \frac{\partial T}{\partial h} + \frac{\partial T}{\partial r} + \frac{2}{r} T \quad (5.19)$$

and the correction term ε_c consists of a Taylor series expansion to provide the upward-continuation from the ellipsoid to the sphere

$$\varepsilon_c = \sum_{i=1}^{\infty} \frac{1}{i!} \left. \frac{\partial^i \widetilde{\Delta g}}{\partial r^i} \right|_{r=R} (r_e - R)^i \quad (5.20)$$

Both correction terms can be expanded into a series of solid SHCs of the disturbing potential, as follows.

5.2.2 Expansion of the correction terms

Equation (5.20) contains the spherically approximated gravity anomaly, which can be written in terms of solid SHCs of the disturbing potential using Equation (4.66)

$$\widetilde{\Delta g} = -\frac{2}{r_e}T - \frac{\partial T}{\partial r} = \frac{1}{R} \sum_{n=0}^{\infty} (n-1) \left(\frac{R}{r_e}\right)^{n+2} \sum_{m=-n}^n \overline{T}_{nm}^R \overline{Y}_{nm} \quad (5.21)$$

However, the i -th radial derivative of this gravity anomaly with respect to the geocentric radius r is required for evaluation of Equation (5.20). Differentiating Equation (5.21) with respect to r once gives

$$\frac{\partial \widetilde{\Delta g}}{\partial r} = -\frac{1}{R^2} \sum_{n=0}^{\infty} (n-1)(n+2) \left(\frac{R}{r_e}\right)^{n+3} \sum_{m=-n}^n \overline{T}_{nm}^R \overline{Y}_{nm} \quad (5.22)$$

A general relation for the i -th derivative can then easily be deduced

$$\frac{\partial^i \widetilde{\Delta g}}{\partial r^i} = (-1)^i \frac{1}{R^{i+1}} \sum_{n=0}^{\infty} (n-1) \frac{(n+1+i)!}{(n+1)!} \left(\frac{R}{r_e}\right)^{n+2+i} \sum_{m=-n}^n \overline{T}_{nm}^R \overline{Y}_{nm} \quad (5.23)$$

The complete Taylor series in Equation (5.20) also includes the i -th power of the difference between the ellipsoidal radius r_e and the radius of the sphere R . Inserting the formula for the ellipsoidal radius (Table 2.1) and the scale factor c (Equation 3.64), this can be rewritten as

$$(r_e - R)^i = (-ca)^i \left[1 - \frac{\sqrt{1-e^2}}{c} (1 - e^2 \sin^2 \theta)^{-\frac{1}{2}} \right]^i \quad (5.24)$$

The term between the square brackets will be $\ll 1$, since the scale factor c is always close to 1 and $e^2 \ll 1$. This term is taken to the power of an integer i , which means it can be expressed as a finite binomial series (see Appendix C)

$$(r_e - R)^i = (-ca)^i \sum_{j=0}^i (-1)^j \binom{i}{j} \left(\frac{\sqrt{1-e^2}}{c}\right)^j (1 - e^2 \sin^2 \theta)^{-\frac{j}{2}} \quad (5.25)$$

However, because the term between the square brackets in Equation (5.24) contains the difference between two almost equal values, the resulting binomial series will be an alternating one where the absolute values of the individual terms become very large, especially for high values of i . Since these alternating terms almost cancel out to a value very close to zero, the relative numerical accuracy for terms of high index i is

expected to be poor. This is due to the loss of significant digits when two large similar values are subtracted from one another. The terms of high index i become especially relevant for high degrees ($n > 360$), since the absolute value of the i -th derivative of the gravity anomaly in Equation (5.23) increases with increasing n .

The term $(1 - e^2 \sin^2 \theta)^{-j/2}$ in Equation (5.25) can also be expressed as a binomial series, yielding

$$(r_e - R)^i = (-ca)^i \sum_{j=0}^i (-1)^j \binom{i}{j} \left(\frac{\sqrt{1-e^2}}{c} \right)^j \sum_{k=0}^{\infty} (-1)^k \binom{-j/2}{k} e^{2k} \sin^{2k} \theta \quad (5.26)$$

Then, inserting Equations (5.23) and (5.26) into Equation (5.20) gives an expression for the correction term ε_c

$$\varepsilon_c = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \chi_{nk} \sin^{2k} \theta \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm} \quad (5.27)$$

where

$$\chi_{nk} = (-1)^k e^{2k} \frac{(n-1)}{(n+1)!} \sum_{i=1}^{\infty} \frac{1}{i! R} (n+1+i)! \sum_{j=0}^i (-1)^j \binom{i}{j} \left(\frac{\sqrt{1-e^2}}{c} \right)^j \binom{-j/2}{k} \quad (5.28)$$

The dependence on latitude in Equation (5.27) can now be completely moved into the SHFs by application of the sinusoidal Legendre weight functions (LWFs) (Equation 3.31)

$$\varepsilon_c = \sum_{n=0}^{\infty} \sum_{m=-n}^n \bar{T}_{nm}^R \sum_{k=0}^{\infty} \chi_{nk} \sum_{j=-k}^k \bar{K}_{nm}^{2j,2k} \bar{Y}_{n+2j,m} \quad (5.29)$$

and after a rearrangement of the summation order (see Appendix B) to make the SHFs independent of index i , this becomes

$$\varepsilon_c = \sum_{n=0}^{\infty} \sum_{m=-n}^n \bar{Y}_{nm} \sum_{j=0}^{\infty} \sum_{i=-j}^j \chi_{n-2i,j} \bar{K}_{n-2i,m}^{2i,2j} \bar{T}_{n-2i,m}^R \quad (5.30)$$

A similar type of expression can also be derived for the correction term ε_e by inserting Equations (4.69) and (4.84) into Equation (5.19)

$$\begin{aligned} \varepsilon_e = & \sum_{n=0}^{\infty} \sum_{m=0}^n \bar{Y}_{nm} \sum_{j=0}^{\infty} \sum_{i=-j-1}^{j+1} \left(\varepsilon_{n-2i,m,i,j} + \frac{n-2i+1}{R} \beta_{n-2i,m,i,j} + \right. \\ & \left. \frac{1}{R} \gamma_{n-2i,m,i,j} - \frac{n-1-2i}{R} \alpha_{n+1-2i,j} \bar{K}_{n-2i,m}^{2i,2j} \right) \bar{T}_{n-2i,m}^R \end{aligned} \quad (5.31)$$

Now, the total correction to the gravity anomaly can be found by addition of Equations (5.30) and (5.31)

$$\delta\Delta g = \sum_{n=0}^{\infty} \sum_{m=0}^n \bar{Y}_{nm} \sum_{j=0}^{\infty} \sum_{i=-j-1}^{j+1} \left[\varepsilon_{n-2i,m,i,j} + \frac{n-2i+1}{R} \beta_{n-2i,m,i,j} + \frac{1}{R} \gamma_{n-2i,m,i,j} + \left(\chi_{n-2i,j} - \frac{n-1-2i}{R} \alpha_{n+1-2i,j} \right) \bar{K}_{n-2i,m}^{2i,2j} \right] \bar{T}_{n-2i,m}^R \quad (5.32)$$

After interchanging the summations over i and j (see Appendix B), Equation (5.32) reads

$$\delta\Delta g = \sum_{n=0}^{\infty} \sum_{m=-n}^n \bar{Y}_{nm} \sum_{i=-\infty}^{\infty} \lambda_{nmi}(\delta\Delta g, T) \bar{T}_{n-2i,m}^R \quad (5.33)$$

where

$$\lambda_{nmi}(\delta\Delta g, T) = \sum_{j=|i|-1}^{\infty} \left[\varepsilon_{n-2i,m,i,j} + \frac{n-2i+1}{R} \beta_{n-2i,m,i,j} + \frac{1}{R} \gamma_{n-2i,m,i,j} + \left(\chi_{n-2i,j} - \frac{n-1-2i}{R} \alpha_{n+1-2i,j} \right) \bar{K}_{n-2i,m}^{2i,2j} \right] \quad (5.34)$$

Equation (5.34) provides the weights for a transformation between surface SHCs of the gravity anomaly corrections $\delta\Delta g$ and solid SHCs of the disturbing potential T .

5.2.3 Computation of the geopotential coefficients

The computation of solid SHCs \bar{T}_{nm}^R follows from inserting Equation (5.17) into Equation (5.16)

$$\bar{T}_{nm}^R = \tilde{T}_{nm}^R + \frac{R}{n-1} \frac{1}{4\pi} \int_{\sigma} \delta\Delta g \bar{Y}_{nm} d\sigma \quad (5.35)$$

where the tilde over T_{nm}^R indicates that these coefficients result from a spherical approximation (Equation 5.7). Inserting Equation (5.33) into Equation (5.35) gives

$$\bar{T}_{nm}^R = \tilde{T}_{nm}^R + \frac{R}{n-1} \frac{1}{4\pi} \int_{\sigma} \sum_{n'=0}^{\infty} \sum_{n'=-m'}^{m'} \bar{Y}_{n'm'} \sum_{i=-\infty}^{\infty} \lambda_{n'm'i}(\delta\Delta g, T) \bar{T}_{n'-2i,m'}^R \bar{Y}_{nm} d\sigma \quad (5.36)$$

where primes were added to the degree and order to be able to differentiate between the two instances of degree and order occurring in this formula.

The order of summation and integration can be reversed, and since only the spherical harmonics depend on longitude and latitude, all other terms can be taken outside the

integration. It then follows from the orthogonality of SHFs on the sphere (Equation 2.80) that the integration vanishes

$$\bar{T}_{nm}^R = \tilde{T}_{nm}^R + \frac{R}{n-1} \sum_{i=-\infty}^{\infty} \lambda_{nmi}(\delta\Delta g, T) \bar{T}_{n-2i,m}^R \quad (5.37)$$

There are several ways to solve for the solid SHCs \bar{T}_{nm}^R from Equation (5.37), and the most straightforward method is through an iterative approach outlined by Cruz (1986). As an initial estimate, the SHCs $\bar{T}_{n-2i,m}^R$ can be replaced by the spherically approximated SHCs $\tilde{T}_{n-2i,m}^R$. This results in a weighted summation over the spherically approximated SHCs

$$\bar{T}_{nm}^R = \sum_{i=-\infty}^{\infty} \lambda_{nmi}^{\text{uc}}(T, \tilde{T}) \tilde{T}_{n-2i,m}^R \quad (5.38)$$

where

$$\lambda_{nmi}^{\text{uc}}(T, \tilde{T}) = \begin{cases} 1 + \frac{R}{n-1} \lambda_{nm0}(\delta\Delta g, T) & \text{for } i = 0 \\ \frac{R}{n-1} \lambda_{nmi}(\delta\Delta g, T) & \text{for } i \neq 0 \end{cases} \quad (5.39)$$

and where the superscript uc is added to indicate that the upward-continuation approach is applied. Equation (5.37) can then be solved iteratively.

In the derivation of Cruz (1986), the summation in Equation (5.38) only runs from -3 to 3 . Since the more rigorous infinite summation derived here depends on the new, and thus at that time unknown, sinusoidal LWFs, Cruz (1986) derived lengthy explicit formulas to achieve this solution. The errors caused by the truncation of all terms of $|i| > 3$ are, however, smaller than the errors caused by approximations to the order of e^2 that Cruz introduced in the correction terms ε_c and ε_e (but that were avoided in this derivation), as will be proven in Section 5.5.1.

Cruz's (1986) derivation was improved upon by Petrovskaya et al. (2001), but the improvement did not involve the inclusion of terms higher than order e^2 in the correction terms. The improvement is instead based on a different iteration strategy, which can be obtained by moving the term \bar{T}_{nm}^R inside the summation on the right-hand side of Equation (5.37) to the left-hand side, which in the formulation used in this thesis

becomes

$$\bar{T}_{nm}^R = \left(1 - \frac{R}{n-1} \lambda_{nm0}(\delta\Delta g, T)\right)^{-1} \left(\tilde{T}_{nm}^R + \frac{R}{n-1} \sum_{i=-\infty, i \neq 0}^{\infty} \lambda_{nmi}(\delta\Delta g, T) \bar{T}_{n-2i, m}^R\right) \quad (5.40)$$

When the SHCs \bar{T}_{nm}^R on the right-hand side of Equation (5.40) are approximated by the spherically approximated coefficients, Equation (5.38) can again be obtained. However, the weight functions $\lambda_{nmi}^{\text{uc}}$ are in this case slightly different from Equation (5.39):

$$\lambda_{nmi}^{\text{uc}}(T, \tilde{T}) = \begin{cases} \left(1 - \frac{R}{n-1} \lambda_{nm0}(\delta\Delta g, T)\right)^{-1} & \text{for } i = 0 \\ \frac{R \lambda_{nmi}(\delta\Delta g, T)}{n-1 - R \lambda_{nm0}(\delta\Delta g, T)} & \text{for } i \neq 0 \end{cases} \quad (5.41)$$

In Petrovskaya et al. (2001), the formulas are derived with an accuracy of order e^2 , and the summation in Eq. (5.38) only runs from -1 to 1 . Inclusion of terms with $|i| > 1$ would not have provided a more accurate solution, given the level of accuracy of order e^2 . The numerical performance of Cruz's (1986) and Petrovskaya's (2001) methods are outlined in Section 5.5 and compared to the two other methods that are derived in Sections 5.3 and 5.4.

5.3 Method 2: Ellipsoidal integration

The second method to compute geopotential coefficients from terrestrial gravity anomalies involves an integration over the ellipsoidal surface, based on the integral formula by Sjöberg (1988). This method was proposed by Fan (1989) and Sjöberg (2003c), and yields a similar type of weighted summation as the upward-continuation approach in the previous Section (Equation 5.38). In Fan's (1989) formulas, the summation runs from -3 to 3 and in Sjöberg's (2003c) formulas from -1 to 1 , which introduces an error due to the series truncation, but a general derivation with unlimited summation based on the ellipsoidal integration methods was derived by Claessens and Featherstone (2005).

5.3.1 Solution strategy

The ellipsoidal integration method is founded on a formula for the geopotential coefficients based on Green's second identity (Equation 2.41). Sjöberg (1988) has shown that the solid SHCs of the Earth's disturbing potential can be computed from an integration over the ellipsoid \mathcal{E}

$$\bar{T}_{nm}^R = \frac{1}{4\pi R(2n+1)a_n} \int_{\mathcal{E}} \left(T \frac{\partial S}{\partial h} - S \frac{\partial T}{\partial h} \right) d\mathcal{E} \quad (5.42)$$

where S is the function

$$S = \left[\left(\frac{R}{r_e} \right)^{n+1} + a_n \left(\frac{r_e}{R} \right)^n \right] \bar{Y}_{nm} \quad (5.43)$$

and a_n is an arbitrary constant depending solely on the spherical harmonic degree n . Sjöberg (1988) identifies $a_n = (n-1)/(n+2)$ as a suitable choice for the computation of solid SHCs of the disturbing potential from gravity anomalies, and this choice was utilised in the numerical investigations in Section 5.5.

The integration over the ellipsoidal surface \mathcal{E} can be transformed into an integration over the sphere σ using a relationship between the ellipsoidal infinitesimal elements and its spherical counterpart (Sjöberg, 1988)

$$d\mathcal{E} = \frac{r_e^2 d\sigma}{\cos \phi} \quad (5.44)$$

Fan (1989) and Sjöberg (2003c) approximate this relation to the order of e^2 , and also approximate the function S and its derivatives with respect to the ellipsoidal normal to the same order of accuracy. These approximations are acceptable for the computation of low-degree and -order SHCs, but severely compromise the accuracy of coefficients of higher degrees and order ($n > 180$), as will be shown in Section 5.5.1.

Moreover, the approximation depends on the numerical coincidence that the geodynamic parameter \bar{m} is approximately half the size of e^2 (Moritz, 1989, Equation (39.12)). Hipkin (2004) argues against this approach, because it is not general and does not lend itself to higher-order approximations. In this study, an alternative procedure is followed.

Because the normal derivative of the disturbing potential is present in both the formula for computation of the geopotential coefficients (Equation 5.42) and in the boundary condition (Equation 2.52), the gravity anomaly can be inserted directly into Equation (5.42), and with Equations (4.14) and (5.44) this gives

$$\bar{T}_{nm}^R = \frac{1}{4\pi R(2n+1)a_n} \int_{\sigma} \frac{r_e^2 S}{\cos \phi} \Delta g + \left(r_e^2 \frac{\partial S}{\partial r} - r_e \tan \phi \frac{\partial S}{\partial \theta} - \frac{r_e^2 S}{\cos \phi} \frac{1}{\gamma} \frac{\partial \gamma}{\partial h} \right) T d\sigma \quad (5.45)$$

The integration kernel contains four parts that are either multiplied by the gravity anomaly Δg or the disturbing potential T , namely

$$\frac{r_e^2 S}{\cos \phi}, \quad r_e^2 \frac{\partial S}{\partial r}, \quad r_e \tan \phi \frac{\partial S}{\partial \theta} \quad \text{and} \quad \frac{r_e^2 S}{\cos \phi} \frac{1}{\gamma} \frac{\partial \gamma}{\partial h}$$

These four parts can all be expanded into binomial series, after which the dependence on the co-latitude θ can be shifted into the SHFs solely, using the sinusoidal LWFs (Equation 3.31).

5.3.2 Expansion of the integration kernels

Combining Equations (5.43) with relations for the ellipsoidal radius r_e and the cosine of the ellipsoidal deflection angle ϕ from Tables 2.1 and 2.2, respectively, gives for the first of the four parts in Equation (5.45)

$$\begin{aligned} \frac{r_e^2 S}{\cos \phi} = a^2 (1 - \epsilon^4 \sin^2 \theta)^{\frac{1}{2}} & \left[c^{-n-1} (1 - e^2)^{-\frac{n-1}{2}} (1 - e^2 \sin^2 \theta)^{\frac{n-3}{2}} \right. \\ & \left. + c^n a_n (1 - e^2)^{\frac{n+2}{2}} (1 - e^2 \sin^2 \theta)^{-\frac{n+4}{2}} \right] \bar{Y}_{nm} \end{aligned} \quad (5.46)$$

The three parts in Equation (5.46) containing the sine of co-latitude θ can be expressed as binomial series, which will always be convergent due to the small eccentricity of the Earth (see Appendix C)

$$\begin{aligned} \frac{r_e^2 S}{\cos \phi} = a^2 \sum_{i=0}^{\infty} (-1)^i \binom{-\frac{1}{2}}{i} \epsilon^{4i} \sin^{2i} \theta & \left[c^{-n-1} (1 - e^2)^{-\frac{n-1}{2}} \sum_{j=0}^{\infty} (-1)^j \binom{\frac{n-3}{2}}{j} \right. \\ & \left. \times e^{2j} \sin^{2j} \theta + c^n a_n (1 - e^2)^{\frac{n+2}{2}} \sum_{k=0}^{\infty} (-1)^k \binom{-\frac{n+4}{2}}{k} e^{2k} \sin^{2k} \theta \right] \bar{Y}_{nm} \end{aligned} \quad (5.47)$$

Subsequent summation and Cauchy multiplication of the Taylor series (see Appendix B) yields an expression as one single series over the square of the sine of the co-latitude

θ

$$\frac{r_e^2 S}{\cos \phi} = \sum_{i=0}^{\infty} \nu_{in} \sin^{2i} \theta \bar{Y}_{nm} \quad (5.48)$$

where the term ν_{in} follows from a finite summation over binomial functions

$$\begin{aligned} \nu_{in} = a^2 e^{2i} \sum_{k=0}^i (2 - e^2)^{i-k} \binom{\frac{1}{2}}{i-k} \left[c^{-n-1} (1 - e^2)^{-\frac{n-1}{2}} \binom{\frac{n-3}{2}}{k} \right. \\ \left. + c^n a_n (1 - e^2)^{\frac{n+2}{2}} \binom{-\frac{n+4}{2}}{k} \right] \end{aligned} \quad (5.49)$$

The form of Equation (5.48) allows for the application of the sinusoidal LWFs (Equation 3.31), but first, the other parts of Equation (5.45) can be rewritten using a similar procedure of binomial expansion and Cauchy multiplication.

For the second part in Equation (5.45), the radial derivative of the function S with respect to r can be found by differentiation of Equation (5.43)

$$\left. \frac{\partial S}{\partial r} \right|_{r_e} = \left[-\frac{n+1}{R} \left(\frac{R}{r_e} \right)^{n+2} + n R a_n \left(\frac{r_e}{R} \right)^{n-1} \right] \bar{Y}_{nm}(\theta, \lambda) \quad (5.50)$$

After implementation of the equation for the ellipsoidal radius r_e (Table 2.1), a formula can be found that can be expanded into binomial series

$$\begin{aligned} r_e^2 \left. \frac{\partial S}{\partial r} \right|_{r_e} &= \left[a^2 c^{n-2} n a_n (1 - e^2)^{\frac{n+1}{2}} (1 - e^2 \sin^2 \theta)^{-\frac{n+1}{2}} \right. \\ &\quad \left. - c^{-n-1} (n+1) (1 - e^2)^{-\frac{n}{2}} (1 - e^2 \sin^2 \theta)^{\frac{n}{2}} \right] a \bar{Y}_{nm} \end{aligned} \quad (5.51)$$

As in the first part in Equation (5.45), summation and Cauchy multiplication of the binomial series (see Appendix B) yields an expression as one single series over the square of the sine of the co-latitude θ

$$r_e^2 \left. \frac{\partial S}{\partial r} \right|_{r_e} = \sum_{i=0}^{\infty} \xi_{in} \sin^{2i} \theta \bar{Y}_{nm} \quad (5.52)$$

where the term ξ_{in} , like the term ν_{in} , follows from a finite summation over binomial functions

$$\xi_{in} = a e^{2i} \sum_{k=0}^i \left[c^{n-2} n a_n a^2 (1 - e^2)^{\frac{n+1}{2}} \binom{-\frac{n+1}{2}}{k} - c^{-n-1} (n+1) (1 - e^2)^{-\frac{n}{2}} \binom{\frac{n}{2}}{k} \right] \quad (5.53)$$

For the third part in Equation (5.45), the latitudinal derivative of the function S (Equation 5.43) must be obtained. This can be done by first introducing the formula for

the ellipsoidal radius (Table 2.1) into Equation (5.43) and subsequently differentiating with respect to the geocentric co-latitude, yielding

$$\left. \frac{\partial S}{\partial \theta} \right|_{r_e} = \left[-(n+1) \left(\frac{R}{r_e} \right)^{n+1} + na_n \left(\frac{r_e}{R} \right)^n \right] \tan \phi \bar{Y}_{nm} + \left[\left(\frac{R}{r_e} \right)^{n+1} + a_n \left(\frac{r_e}{R} \right)^n \right] \frac{\partial \bar{Y}_{nm}}{\partial \theta} \quad (5.54)$$

Combining Equation (5.54) and the equations for r_e (Table 2.1) and $\tan \phi$ (Table 2.2) provides an equation that expresses the third part of Equation (5.45) as a function of θ . Then, as in the first two parts, the terms containing $\sin \theta$ can be expanded into binomial series, which after insertion of Equation (3.50) leads to

$$r_e \tan \phi \left. \frac{\partial S}{\partial \theta} \right|_{r_e} = \sum_{i=0}^{\infty} (\mu_{in} \sin^{2i} \theta \bar{Y}_{nm} + \sigma_{in} \sum_{j=-1}^1 \bar{N}_{nm}^{2j,2} \sin^{2i} \theta \bar{Y}_{n+2j,m}) \quad (5.55)$$

where

$$\mu_{in} = ae^{2i+4} \sum_{j=1}^{\min(i,2)} (2i-2j-3) \sum_{k=0}^i \binom{-\frac{3}{2}}{i-k} \left[(n+1)c^{-n-1}(1-e^2)^{-\frac{n}{2}} \binom{\frac{n-1}{2}}{k} - nc^n a_n (1-e^2)^{\frac{n+1}{2}} \binom{-\frac{n+2}{2}}{k} \right] \quad (5.56)$$

and

$$\sigma_{in} = ae^{2i+2} \sum_{k=0}^i \binom{-\frac{3}{2}}{i-k} \left[c^{-n-1}(1-e^2)^{-\frac{n}{2}} \binom{\frac{n+1}{2}}{k} - c^n a_n (1-e^2)^{\frac{n+1}{2}} \binom{-\frac{n}{2}}{k} \right] \quad (5.57)$$

The fourth part in Equation (5.45) is equal to the first (Equation 5.48), multiplied by the term $\gamma^{-1} \partial \gamma / \partial h$. Combining the relations for r_e and $\cos \phi$ from Tables 2.1 and 2.2 with Equation (4.80), the following formula can be found for the fourth and final part

$$\frac{r_e^2 S}{\cos \phi} \frac{1}{\gamma} \frac{\partial \gamma}{\partial h} = \sum_{i=0}^{\infty} \rho_{in} \sin^{2i} \theta \bar{Y}_{nm} \quad (5.58)$$

where

$$\rho_{in} = - \sum_{k=0}^i \nu_{i-k,n} \left\{ \frac{\sqrt{1-e^2}}{a} e^{2i} \sum_{j=0}^k (2-e^2)^j \left[\binom{-\frac{3}{2}}{k} \binom{\frac{3}{2}}{k-j} + \binom{-\frac{1}{2}}{k} \binom{\frac{1}{2}}{k-j} \right] - 2 \frac{\omega^2}{\gamma_b} \sum_{j=0}^k e^{2(k-j)} \binom{\frac{1}{2}}{k-j} \sum_{l=0}^k e^{2(k-l)} (2-e^2)^{k-l} \left(1 - \frac{b^3 \gamma_a}{a^3 \gamma_b} \right)^{2l} \binom{-1}{l} \binom{\frac{1}{2}}{k-l} \right\} \quad (5.59)$$

The functions in Equations (5.49), (5.53), (5.56), (5.57) and (5.59) all depend only on the spherical harmonic degree n and the parameters of the reference gravity field.

5.3.3 Computation of geopotential coefficients

Inserting Equations (5.48), (5.52), (5.55) and (5.58) into Equation (5.45), followed by insertion of Equation (3.31) and interchanging of the order of summation and integration gives

$$\begin{aligned} \bar{T}_{nm}^R = & \frac{1}{4\pi R(2n+1)a_n} \left[\sum_{i=0}^{\infty} \nu_{in} \sum_{j=-i}^i \bar{K}_{nm}^{ij} \int_{\sigma} \Delta g \bar{Y}_{n-2j,m} d\sigma \right. \\ & + \sum_{i=0}^{\infty} (\xi_{in} - \mu_{in} - \rho_{in}) \sum_{j=-i}^i \bar{K}_{nm}^{ij} \int_{\sigma} T \bar{Y}_{n-2j,m} d\sigma \\ & \left. - \sum_{i=0}^{\infty} \sigma_{in} \sum_{k=-1}^1 \bar{N}_{nm}^{2k,2} \sum_{j=-i}^i \bar{K}_{n-2k,m}^{ij} \int_{\sigma} T \bar{Y}_{n-2(j+k),m} d\sigma \right] \end{aligned} \quad (5.60)$$

The integrals in Equation (5.60) are equal to those in the spherical approximation, where it holds that

$$\tilde{T}_{nm}^R = \frac{1}{4\pi} \int_{\sigma} T \bar{Y}_{nm}(\theta, \lambda) = \frac{R}{4\pi(n-1)} \int_{\sigma} \Delta g \bar{Y}_{nm}(\theta, \lambda) d\sigma \quad (5.61)$$

where use was made of the relationship between the disturbing potential T and the gravity anomaly Δg in the spherical spectral domain (Equation 4.68). After using Equations (5.61) and rearrangement of the summation order, the final compact formula for the computation of geopotential coefficients from gravity anomalies on the ellipsoid becomes

$$\bar{T}_{nm}^R = \sum_{i=-\infty}^{\infty} \lambda_{nmi}^{ei}(T, \tilde{T}) \tilde{T}_{n-2i,m}^R \quad (5.62)$$

where

$$\begin{aligned} \lambda_{nmi}^{ei}(T, \tilde{T}) = & \frac{1}{R(2n+1)a_n} \sum_{k=|i|}^{\infty} \left[\left(\frac{n-1}{R} \nu_{kn} + \xi_{kn} - \mu_{kn} - \rho_{kn} \right) \bar{K}_{nm}^{ki} \right. \\ & \left. - \sigma_{kn} \sum_{j=\max(i-k,-1)}^{\min(i+k,1)} \bar{N}_{nm}^{2j,2} \bar{K}_{n-2j,m}^{k,i-j} \right] \end{aligned} \quad (5.63)$$

and where the superscript ei was added to indicate that the weights are computed via the ellipsoidal integration method.

Thus, the geopotential coefficients can be obtained from a weighted summation over the spherically approximated coefficients of the same order m , as in the upward-continuation method. The weights here also depend solely on the degree and order

of the coefficient and the reference gravity field, and not on latitude or longitude. The formula for the weights $\lambda_{nmi}^{\text{ei}}(T, \tilde{T})$ (Equation 5.63) is slightly more complicated than that of the weights $\lambda_{nmi}^{\text{uc}}(T, \tilde{T})$ (Equation 5.41), but iterations are not required here.

5.4 Method 3: Coefficient transformation

The third method for the computation of geopotential coefficients from terrestrial gravity anomalies is arguably the simplest. It follows directly from the spherically approximated and rigorous spectral relations between the disturbing potential and the gravity anomalies (Equations 4.68 and 4.86 respectively)

$$\begin{aligned} \overline{\Delta g}_{nm}^{r_e} &= \sum_{i=-\infty}^{\infty} \lambda_{nmi}(\Delta g, T) \overline{T}_{n-2i,m}^R \\ &= \frac{n-1}{R} \tilde{T}_{nm}^R \end{aligned} \quad (5.64)$$

The relation between the solid SHCs of the disturbing potential and their counterparts in spherical approximation can therefore be simply derived

$$\tilde{T}_{nm}^R = \frac{R}{n-1} \sum_{i=-\infty}^{\infty} \lambda_{nmi}(\Delta g, T) \overline{T}_{n-2i,m}^R \quad (5.65)$$

The inverse relation can then be found via a matrix inversion, or via an iterative procedure such as the Jacobi, Gauss-Seidel or SOR method (see Section 3.3.2).

This result is very similar to that of Method 1 in Section 5.2, and both methods require an iterative procedure to obtain the solution. The main difference is that the continuation between the sphere and the ellipsoid is here not treated separately using a Taylor expansion, but incorporated inside the transformation between surface and solid SHCs. This simplifies the formulas for the weight functions.

The inverse of Equation (5.65) reads

$$\overline{T}_{nm}^R = \sum_{i=-\infty}^{\infty} \lambda_{nmi}^{\text{ct}}(T, \tilde{T}) \tilde{T}_{n-2i,m}^R \quad (5.66)$$

where the superscript *ct* indicates that the weight functions $\lambda_{nmi}^{\text{ct}}(T, \tilde{T})$ result from a coefficient transformation. These functions can be computed via a matrix inversion,

when the weights $\frac{R}{n-1}\lambda_{nmi}(\Delta g, T)$ are inserted into the matrix in Figure 3.7. The inverse of this matrix will then contain the weights $\lambda_{nmi}^{\text{ct}}(T, \tilde{T})$. This rigorous matrix inversion is numerically stable up to at least degree $n = 1800$ (see Section 5.5.4).

However, it is more efficient to compute the solid SHCs \bar{T}_{nm}^R via an iterative procedure as in Method 1, where the functions $\lambda_{nmi}^{\text{ct}}(T, \tilde{T})$ are defined by

$$\lambda_{nmi}^{\text{ct}}(T, \tilde{T}) = \begin{cases} \frac{n-1}{R\lambda_{nm0}(\Delta g, T)} & \text{for } i = 0 \\ \frac{R}{n-1}\lambda_{nmi}(\Delta g, T) & \text{for } i \neq 0 \end{cases} \quad (5.67)$$

Due to the loss of diagonal dominance of the matrix for $n > 520$, convergence of this iterative procedure for the computation of solid SHCs beyond degree 520 cannot be guaranteed (see Section 3.4).

Obviously, as in Equations (5.39) and (5.41), in spherical approximation the functions $\lambda_{nmi}^{\text{ct}}(T, \tilde{T})$ reduce to 1 for $i = 0$ and to 0 for $i \neq 0$.

5.5 Comparison of the upward continuation, ellipsoidal integration and coefficient transformation methods with existing methods

The newly derived ellipsoidal integration (Section 5.3) and coefficient transformation (Section 5.4) methods are compared numerically to the existing formulations of Cruz (1986), Petrovskaya et al. (2001), Sjöberg (2003c) and Jekeli (1988). Cruz's and Petrovskaya's methods are approximations of the upward-continuation approach derived rigorously in Section 5.2, Sjöberg's approach is an approximation of the rigorous formulation of the ellipsoidal integration approach in Section 5.3, and Jekeli's approach is the ellipsoidal harmonics method described in Section 5.1.3.

A comparison between Cruz's (1986), Jekeli's (1988) and Fan's (1989) methods was performed by Nord (1992), using a simple theoretical ellipsoidal Earth model to spherical harmonic degree and order 180. Here, a comparison between the weight functions in the various methods is first made in Section 5.5.1, followed by a closed-loop simulation using EGM96 geopotential coefficients up to degree and order 360, which provides a

more realistic comparison than that of Nord (1992), in Sections 5.5.2 and 5.5.3. A view at the accuracy of various methods beyond degree and order 360 is shown in Section 5.5.4.

5.5.1 Comparison of weights

A first and most simple comparison is that of the weight functions $\lambda_{nmi}^{\text{uc}}$, $\lambda_{nmi}^{\text{ei}}$ and $\lambda_{nmi}^{\text{ct}}$. Theoretically, these functions should not be equal even in the rigorous formulation, since the upward-continuation and coefficient transformation methods require an iterative procedure to obtain the solid SHCs \overline{T}_{nm}^R . However, since both Cruz (1986) and Petrovskaya et al. (2001) accept the result after the first iteration as an acceptable solution, a comparison of the weights is justified. Thus, the differences between $\lambda_{nmi}^{\text{uc}}$, $\lambda_{nmi}^{\text{ei}}$ and $\lambda_{nmi}^{\text{ct}}$ provide an indication about the level of error that can be expected in each method. The ellipsoidal harmonics method of Jekeli (1988) cannot be included in this comparison, since it does not contain a weighted summation over spherically approximated coefficients. Other methods of comparison that include the ellipsoidal harmonics method are described in Sections 5.5.2 and 5.5.3.

The computation of the weight functions $\lambda_{nmi}(T, \tilde{T})$ in all three approaches described in Sections 5.2 to 5.4 contains a converging infinite summation. The convergence rate is important for the practical evaluation, and is shown for the computation of $\lambda_{nmi}^{\text{ei}}$ (Equation 5.63) in Figure 5.2. It can be seen that the convergence is generally slower for increasing degrees n . Evaluation of only the first two terms ($k = 0$ and $k = 1$) is clearly insufficient especially for the high degrees ($n > 360$), as could already have been expected from the numerical results in Figures 3.5 and 3.6. The more rigorous formulation derived here is therefore of vital importance to obtain accurate results. As can be seen from Figure 5.2, more than 25 terms of the summation in Equation (5.63) need to be taken into account to reach a relative accuracy of e^4 for $n > 2000$, but this does not provide a major computational burden. The computation of all weights $\lambda_{nmi}(T, \tilde{T})$ for $0 \leq m \leq n \leq 2160$ and $0 \leq i \leq 25$ takes in the order of one hour on a Pentium IV 2.4GHz PC.

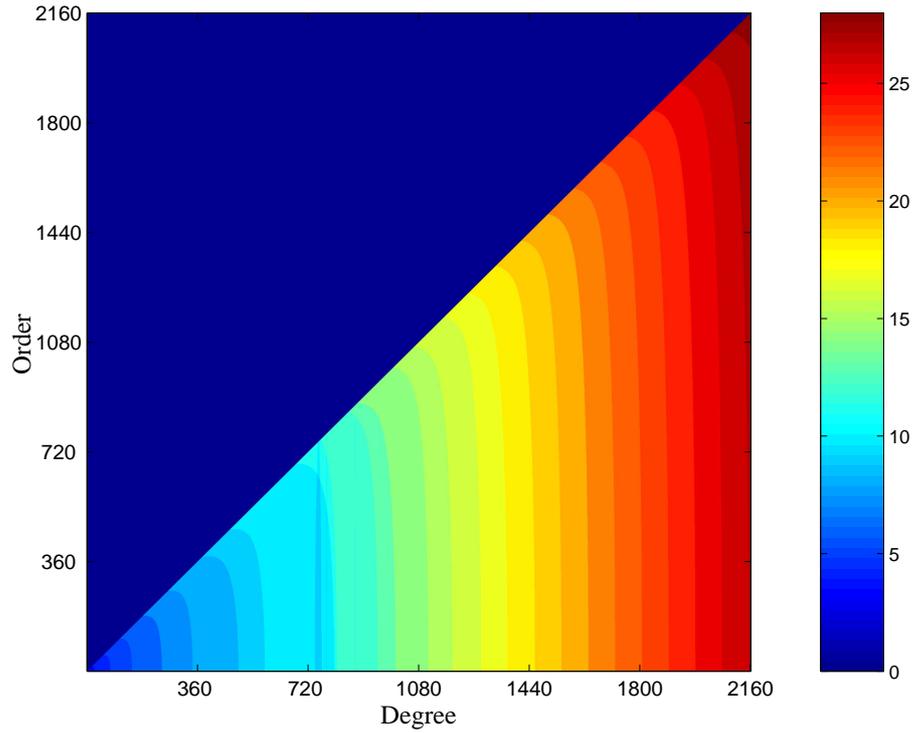


Figure 5.2: Convergence rate in the computation of the weights for the case $i = 0$ (summation over k in Equation (5.63)): number of terms with an absolute value higher than e^4 for all pairs of degree n and order m

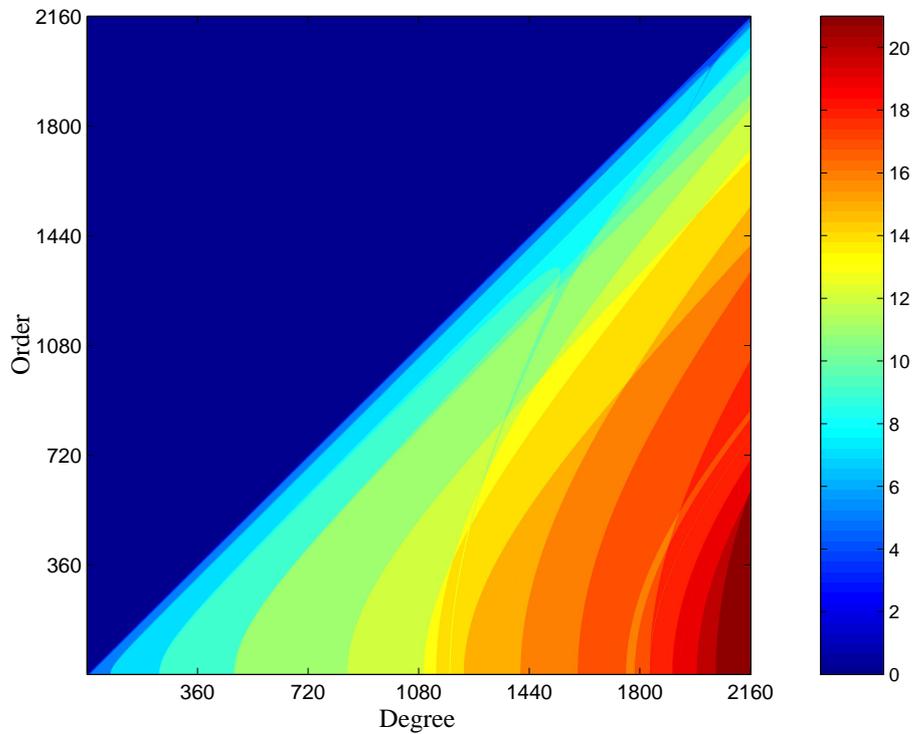


Figure 5.3: Convergence rate in the computation of the coefficients (summation over i in Equation (5.62)): number of terms with an absolute value higher than e^4 for all pairs of degree n and order m

The computation of the solid SHCs of the disturbing potential also contains an infinite summation for all three methods (see Equations 5.38, 5.62 and 5.66). The convergence rate in all three methods is very similar, and is shown for the ellipsoidal integration method in Figure 5.3. The pattern in Figure 5.3 indicates that the number of required terms increases for an increase in the difference between degree n and order m . For the case $n = m$ (the sectorial harmonics), the convergence is very fast, and for the case $m = 0$ (the zonal harmonics), the convergence is slowest. This can be explained intuitively from the geometry of the ellipsoid, as follows.

The sectorial harmonics do not change sign along a meridian. Any variation in latitudinal direction is therefore mainly represented by zonal and tesseral harmonics, whereas the sectorial harmonics mainly account for variations in longitudinal direction. A transformation between an ellipsoid of revolution and a sphere alters the shape of the meridians, but not of the parallels. Therefore, zonal surface SHCs on an ellipsoid and on a sphere will differ most, and sectorial surface SHCs will differ least.

Therefore, the zonal SHCs will require the largest corrections, and can be viewed as a ‘worst case’. The zonal weights $\lambda_{n,0,i}(T, \tilde{T})$ are shown in Figure 5.4 for $i = 0$, $|i| = 1$ and $|i| = 2$, for various methods. The methods of Cruz (1986) and Sjöberg (2003c) give very similar weights, presented by the green lines in Figure 5.4, which decrease almost linearly. This occurs both for the primary weights ($i = 0$) and the secondary weights ($|i| = 1$), where it should be noted that the weights for $i = 1$ and $i = -1$ are not exactly the same, but agree to such an extent that the difference cannot be distinguished in Figure 5.4. The similarity between Cruz’s and Sjöberg’s methods can be explained by the fact that both approximate the fundamental equation of physical geodesy (Equation 2.52) to the order of the square of the first eccentricity of the ellipsoid e^2 . The weights in Petrovskaya’s method do not show the linear dominance present in the other two methods, because the relations between $\lambda_{nmi}^{\text{uc}}(T, \tilde{T})$ and $\lambda_{nmi}(\delta\Delta g, T)$ in Equation (5.41) are nonlinear.

The rigorous forms of the ellipsoidal integration method and the coefficient transformation method derived in Sections 5.3 and 5.4 respectively, are also shown in Figure

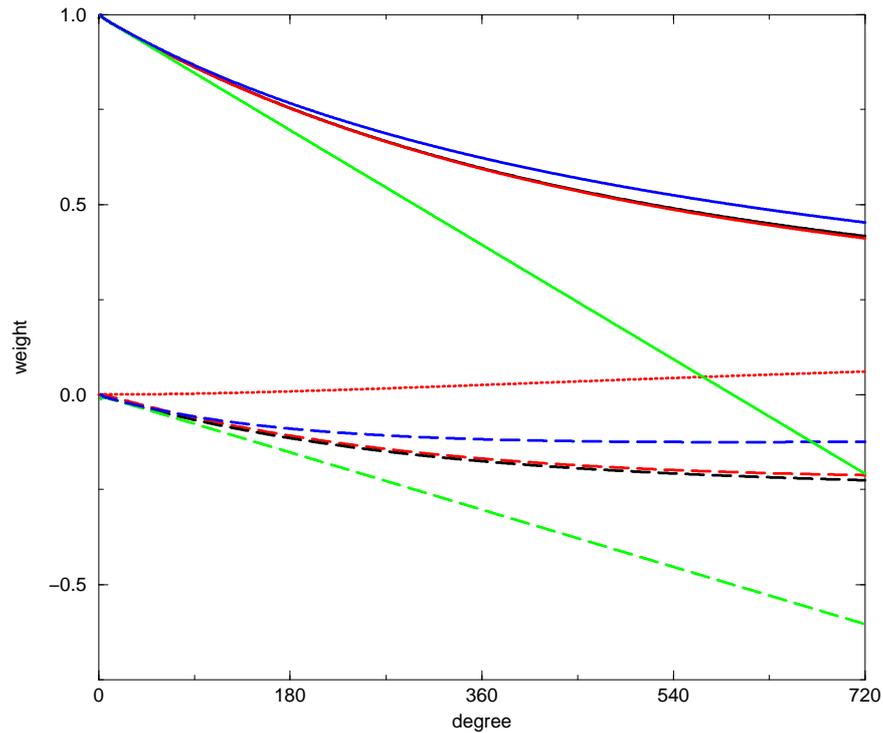


Figure 5.4: Weight functions $\lambda_{nmi}(T, \tilde{T})$ for zonal harmonics ($m = 0$) with $i = 0$ (solid lines), and $|i| = 1$ (dashed lines) for the methods of Cruz (1986) and Sjöberg (2003c) (green), the method of Petrovskaya et al. (2001) (blue), the new ellipsoidal integration (ei) method (black) and the new coefficient transformation (ct) method (red). The weights of $|i| = 2$ are also shown for the coefficient transformation method (dotted red line).

5.4 and give very similar results. For the primary weights ($i = 0$), the numerical differences between both methods are very small, but the differences for the secondary weights ($|i| = 1$) are slightly larger. Still, the agreement between these two methods is far greater than that with any of the other methods. This is a first indication that the accuracy of both the rigorous ellipsoidal integration and coefficient transformation methods surpasses that of the other methods.

Figure 5.4 also shows the weights for $i = 2$ given by the coefficient transformation method. Sjöberg's (2003c) and Petrovskaya's (2001) approximate methods neglect terms of $|i| > 1$, but it can be seen that these weights in fact become significant for the higher degrees ($n > 360$). Cruz's (1986) method does include this term, as well as the weights of index $|i| = 3$. However, it can be deduced from Figure 5.4 that the errors in the terms of $i = 0$ and $|i| = 1$ in Cruz's method are larger than the contribution of the terms of higher index $|i|$.

Furthermore, it can be seen that for all methods, the correction to the spherically approximated SHCs is smallest for the low degrees, where the primary weights are close to one and all other weights close to zero. For all methods, the weights deviate further from these nominal values for increasing degree n . This corresponds to the observation from Figure 5.3 that the corrections to the SHCs increase in size for increasing difference between degree n and m . Obviously, the size of the corrections also depends on the size of the spherically approximated SHCs themselves, and even for very low degrees, the correction to SHCs \tilde{T}_{nm}^R can become relatively large if $\tilde{T}_{nm}^R \ll \tilde{T}_{n-2,m}^R$ and/or $\tilde{T}_{nm}^R \ll \tilde{T}_{n+2,m}^R$.

5.5.2 Closed-loop simulation

The various methods for the computation of geopotential coefficients are tested using a closed-loop simulation (cf. Ledin, 2001). As in Section 3.4, the coefficients of the EGM96 geopotential model (Lemoine et al., 1998) are used to create a grid of gravity anomalies on the ellipsoid, and these gravity anomalies serve as input for the methods under investigation. First, the spherically approximated coefficients are computed from the grid of gravity anomalies through a discretised spherical harmonic analysis (Equation 5.7). The radius of the reference sphere R is in this numerical simulation set equal to the semi-major axis of the ellipsoid a , since this is also the case in the existing methods of Cruz (1986), Jekeli (1988), Petrovskaya et al. (2001) and Sjöberg (2003c).

In Figure 5.5, the degree variances of EGM96 are shown, along with the absolute errors in degree variances of various methods. These absolute errors are defined as the absolute difference between the EGM96 degree variances and the degree variances after spherical harmonic synthesis on the ellipsoid and spherical harmonic analysis using various methods described in Sections 5.1 to 5.4. It can be seen that the absolute errors in the degree variances of the spherically approximated coefficients contain almost as much power in the high degrees as the original signal itself. In other words, the spherical approach is highly inaccurate, especially for the high degrees.

Cruz's (1986) method clearly improves on the spherical approximation and gives an

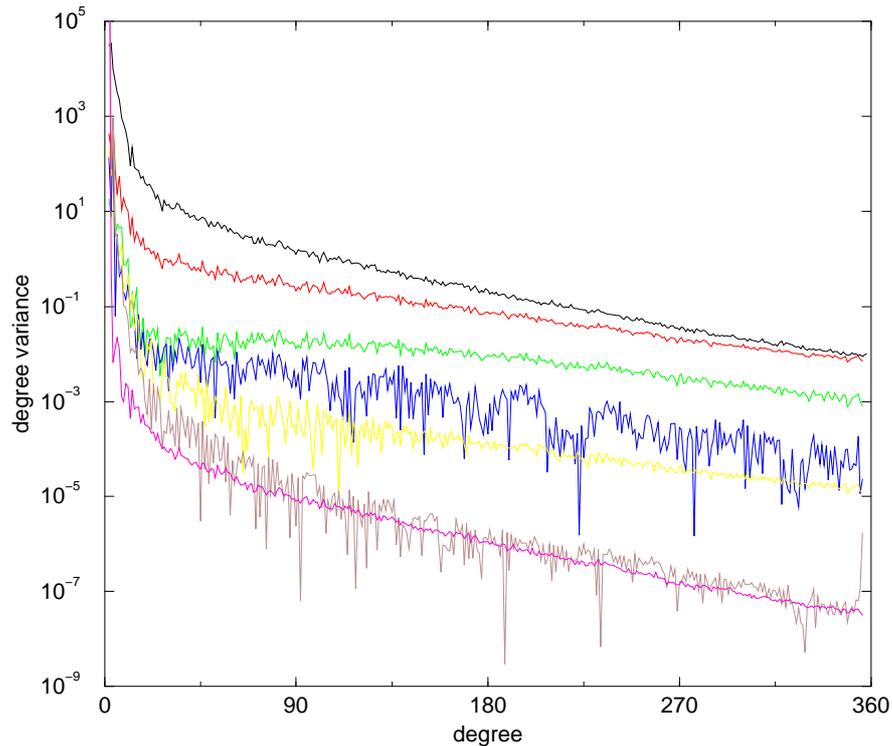


Figure 5.5: The degree variances of EGM96 (black) and the absolute errors in the degree variances after a closed-loop simulation using the spherical approximation (red), the methods by Cruz (1986) and Sjöberg (2003c) (green), the method by Petrovskaya et al. (2001) (blue), the ellipsoidal harmonics method by Jekeli (1988) (magenta), the new ellipsoidal integration method (yellow) and the new coefficient transformation method (brown)

error spectrum very similar to Sjöberg’s (2003c) method. It can be seen in Figure 5.5 that the difference between the logarithmic degree variances of the original EGM96 and the logarithm of the absolute errors of Cruz’s and Sjöberg’s methods decreases with increasing degree n . This indicates that these two methods contain the largest errors in the high degrees, which was already acknowledged by Sjöberg (2003c), and could also be expected from the weights in Figure 5.4. Around degree 360, the difference between the logarithmic degree variances is around one, which translates to a relative error of approximately 10%.

Petrovskaya’s (2001) method improves on Cruz’s method by approximately a factor of 10, but its accuracy varies more strongly for different degrees, as can be seen from Figure 5.5. This behaviour can also be seen in the ellipsoidal integration and coefficient transformation methods derived in Sections 5.3 and 5.4, although to a lesser extent. It can, nevertheless, be seen in Figure 5.5 that the coefficient transformation method

gives the most accurate results, with errors in the same order of magnitude as the errors in Jekeli's ellipsoidal harmonics conversion method (relative accuracy of the order of 10^{-5}). The latter two methods show an error degree variance signal that is almost parallel to the EGM96 signal, which means that they do not give the larger relative errors in the higher degrees that occur in the other methods.

Only for degrees very close to the maximum of 360 do the errors in the coefficient transformation method increase. This is an artefact that all weighted summation methods possess, because the degrees above 360 that are needed in the summation, are absent. This is not problematic, because the SHCs can be computed to a slightly higher degree to avoid this effect, which is in some cases done in practice anyway. For example, in the construction of EGM96, numerical quadrature solutions to degree and order $n = 500$ were computed, even though only the solution to degree and order $n = 460$ was used in further analysis and only the solution to degree and order $n = 360$ was released (Lemoine et al., 1998).

5.5.3 Comparisons in the space domain

In order to examine how the errors in the recovered geopotential coefficients propagate into geoid heights, a spherical harmonic synthesis is performed and Bruns's formula (Equation 2.45) is applied. A more detailed explanation of the computation of geoid heights from a geopotential model is provided in Section 6.1.1. In Figures 5.6 to 5.11 the errors in the geoid heights, i.e., the difference from the EGM96 geoid heights, are shown for the various methods. Naturally, these errors also include numerical errors in the spherical harmonic synthesis and analysis that are performed during the simulation. Based on simulations by Novák et al. (2001) and Huang et al. (2003), the maximum errors in the synthesis-analysis procedure can be expected to be in the order of several centimetres. Global statistics of each method are provided in Table 5.1.

Since analytic methods are mainly of interest in the higher degrees of the geopotential spectrum, the biggest challenge of any method to compute spherical harmonic geopotential coefficients is to properly model the high-frequency error caused by the constant

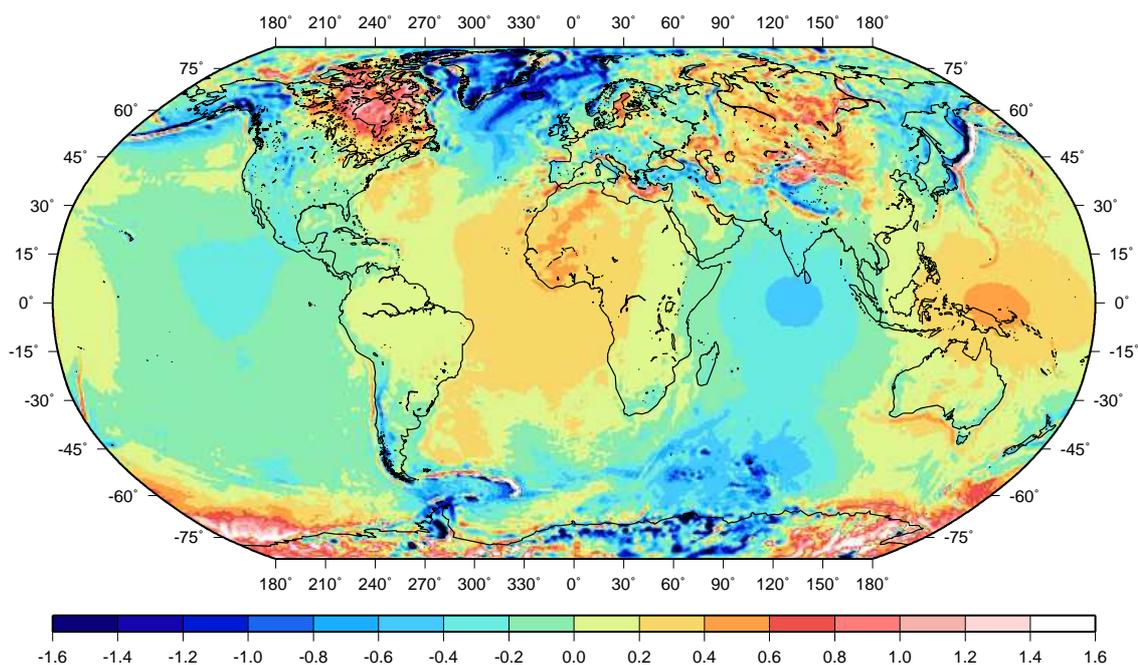


Figure 5.6: Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the spherical approximation for $0 \leq n \leq 360$ (units in m; Robinson projection)

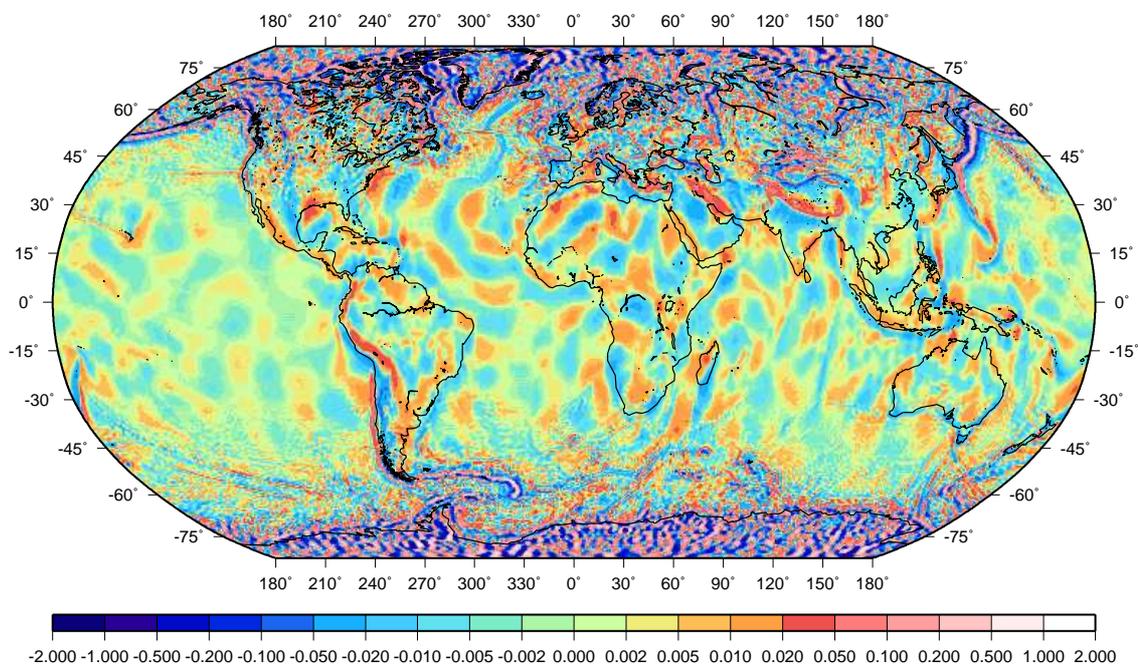


Figure 5.7: Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the method of Cruz (1986) for $20 \leq n \leq 340$ (units in m; Robinson projection)

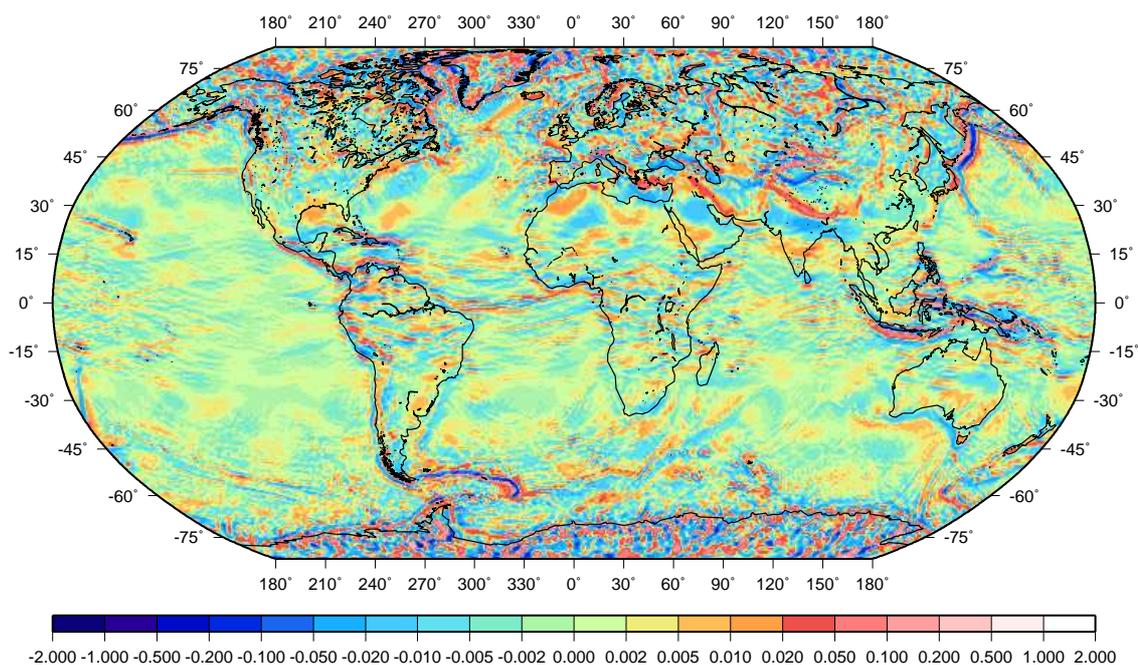


Figure 5.8: Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the method of Petrovskaya et al. (2001) for $20 \leq n \leq 340$ (units in m; Robinson projection)

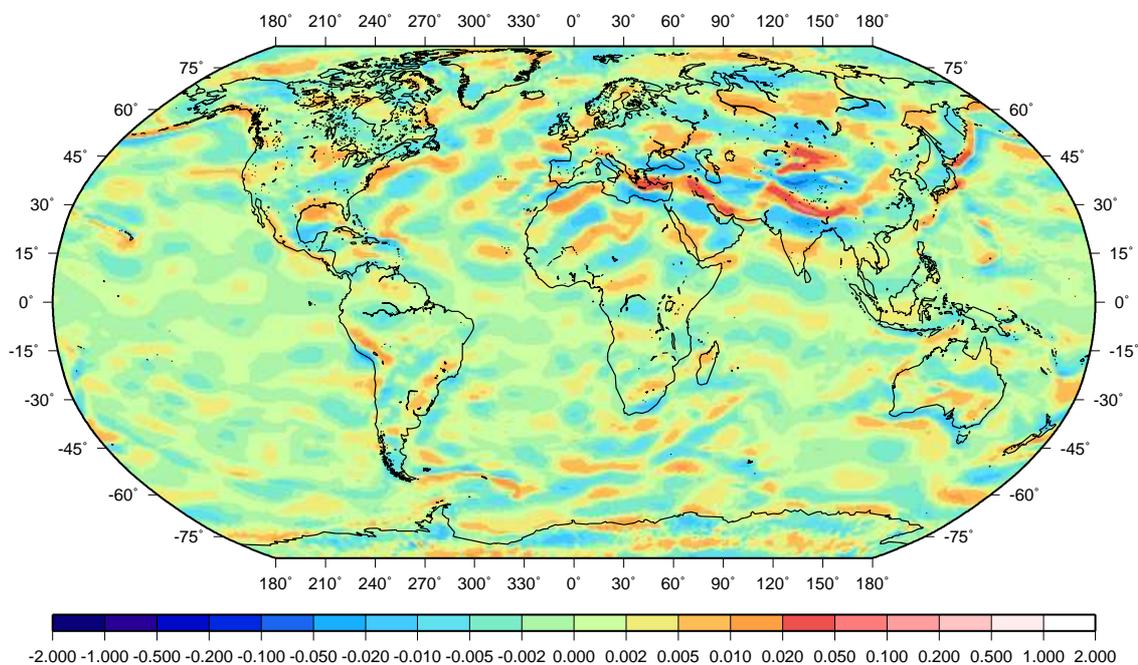


Figure 5.9: Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the ellipsoidal integration method for $20 \leq n \leq 340$ (units in m; Robinson projection)

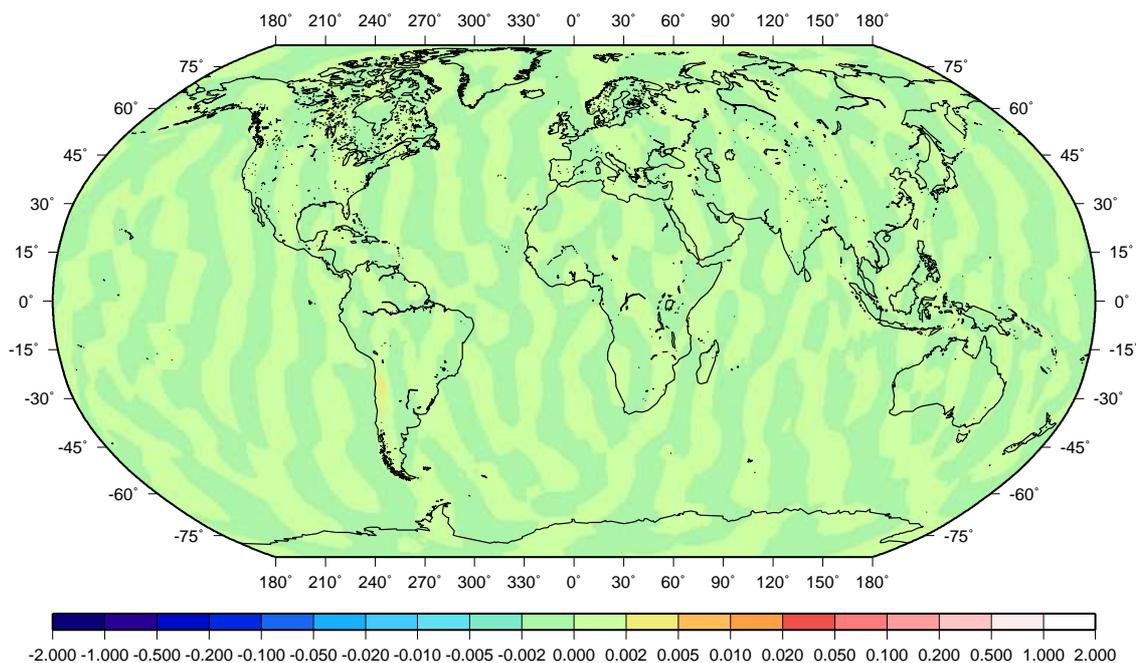


Figure 5.10: Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the coefficient transformation method for $20 \leq n \leq 340$ (units in m; Robinson projection)

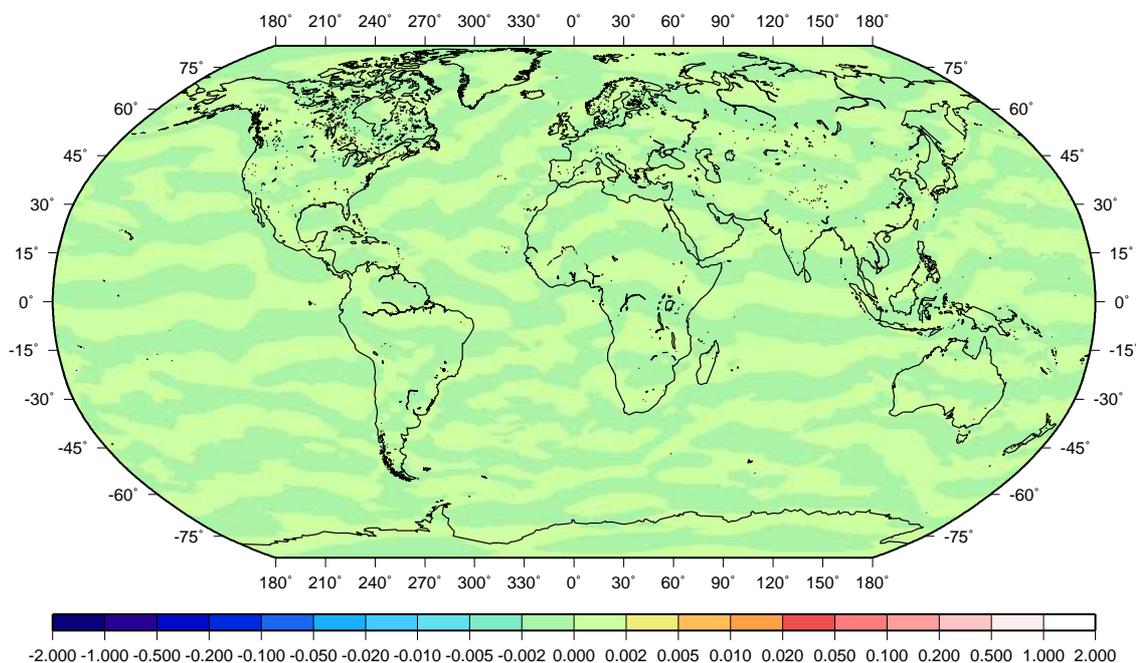


Figure 5.11: Differences in geoid height computed from the EGM96 geopotential coefficients and the coefficients recovered after a closed-loop simulation using the ellipsoidal harmonics method by Jekeli (1988) for $20 \leq n \leq 340$ (solution provided by Holmes (2004, pers. comm.)) (units in m; Robinson projection)

	minimum	maximum	mean	abs. mean
Spherical approximation	-7.85	5.84	$-2.77 \cdot 10^{-3}$	$5.41 \cdot 10^{-1}$
Cruz (1986)	-3.19	2.60	$7.01 \cdot 10^{-4}$	$1.46 \cdot 10^{-1}$
Petrovskaya et al. (2001)	$-6.46 \cdot 10^{-1}$	$8.13 \cdot 10^{-1}$	$-4.20 \cdot 10^{-4}$	$4.28 \cdot 10^{-2}$
Ellipsoidal integration	$-6.51 \cdot 10^{-2}$	$6.98 \cdot 10^{-2}$	$-4.70 \cdot 10^{-5}$	$4.97 \cdot 10^{-3}$
Coefficient transformation	$-2.12 \cdot 10^{-3}$	$2.61 \cdot 10^{-3}$	$6.11 \cdot 10^{-8}$	$2.96 \cdot 10^{-4}$
Jekeli (1988)	$-1.80 \cdot 10^{-4}$	$1.64 \cdot 10^{-4}$	$9.05 \cdot 10^{-8}$	$8.19 \cdot 10^{-6}$

Table 5.1: Statistics of the errors in the geoid height from solid SHCs of the disturbing potential computed using various methods for $20 \leq n \leq 340$ (units in m), where *abs. mean* is the average of the absolute values of the errors

radius approximation. Figures 5.7 to 5.11 and Table 5.1 therefore do not include the degrees $0 \leq n < 20$, so that any errors in the longest wavelengths are filtered out. This serves to clarify the comparison of the medium and high degrees. In addition, the maximum degree in these numerical comparisons is set to 340, in order to filter out the inaccuracies in the close vicinity of the maximum degree that the weighted summation methods suffer from (see Section 5.5.2).

The spherical approximation in Figure 5.6 contains a long-wavelength signal with a magnitude of several decimetres, and this is caused by the approximation of the fundamental Equation (2.52). This long-wavelength pattern, with lows over the Indian Ocean and the East Pacific and highs over West Africa and Papua New Guinea, can be found in many recent publications on ellipsoidal corrections to geoid heights (e.g., Heck and Seitz, 2003; Huang et al., 2003; Hipkin, 2004; Ellmann, 2005).

In this case, however, there is also a high-frequency error signal around the poles, which is caused by the fact that the term $(R/r_e)^{n+2}$ in Equation (4.66) is neglected in the constant radius approximation. The magnitude of this short-wavelength signal is much larger, and the overall error reaches up to an absolute maximum of 7.85 m at 78° S, but can possibly be larger if a geopotential model to higher degree and order is used. The fact that high-frequency errors do not occur around the equator can be

explained by the choice of the radius R equal to the semi-major axis of the ellipsoid a . If, for example, the radius R was alternatively set equal to the semi-minor axis of the ellipsoid b , the high-frequency errors would occur around the equator instead of the poles, because the errors are largest where the ellipsoidal radius r_e and the spherical radius R are furthest apart.

It can be seen in Figures 5.7 and 5.8 that in the methods by Cruz (1986) and Petrovskaya et al. (2001), the high-frequency errors are still dominant. The absolute maximum errors in these methods are 3.19 m and 0.81 m respectively (see Table 5.1), in both cases close to the poles. Around the equator the error is only in the order of a decimetre for Cruz's method and several centimetres for Petrovskaya's method.

In the ellipsoidal integration method (Figure 5.9), the short-wavelength errors are less dominant, but a long-wavelength error of up to 6.9 cm (with $n \geq 20$) is still present. As could be expected from Figure 5.5, the coefficient transformation method and Jekeli's (1988) ellipsoidal harmonics conversion method perform best. Here, the errors are mainly of long-wavelength nature, due to the smoothness of the geoid, and do not exceed 3 mm (for $20 \leq n \leq 340$). It is striking, however, that the coefficient transformation method in Figure 5.10 shows a pattern of stripes in the latitudinal direction, whereas Jekeli's (1988) method in Figure 5.11 shows a pattern of stripes in the longitudinal direction. This indicates that the coefficient transformation method models the zonal harmonics more accurately, whereas Jekeli's method models the sectorial harmonics with higher accuracy.

The relative accuracy of coefficients of various degrees n is shown in Figure 5.12. It can be seen in Figure 5.12 that Jekeli's (1988) method is most accurate for coefficients of absolute order $|m|$ close to degree n , whereas the coefficient transformation method (Section 5.4) is most accurate for coefficients of low order $|m|$. The relative accuracies of the coefficient transformation method range from approximately 10^{-8} for $|m|$ close to zero to approximately 10^{-4} for $|m|$ close to n . Optimal results can be obtained when both methods are combined. Based on Figure 5.12, the coefficient transformation method can, for example, be used for coefficients of $|m| \leq n/6$, and Jekeli's method

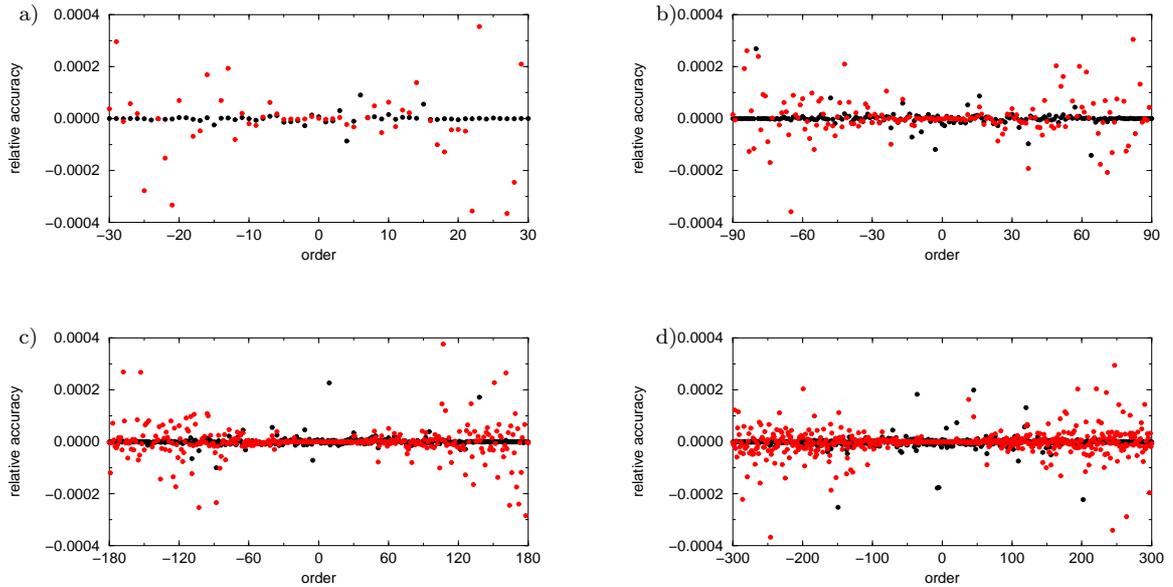


Figure 5.12: Relative accuracy of SHCs \overline{T}_{nm}^R computed from gravity anomalies using Jekeli's (1988) method (black) and the coefficient transformation method (red) compared to the SHCs resulting from a synthesis-analysis procedure for (a) $n = 30$, (b) $n = 90$, (c) $n = 180$ and (d) $n = 300$

for all other coefficients.

Overall, the magnitudes of the errors in the coefficient transformation method are slightly larger than those of Jekeli's method. However, the accuracy of both methods will generally be sufficient, given the sub-centimetre accuracy in geoid height shown in Figures 5.10 and 5.11. Importantly, Figures 5.5 and 5.12 have shown that the relative accuracy of both methods is largely independent of the degree n for all degrees up to 340. This means that coefficients computed using either one of the two methods can also accurately recover gravity functionals with more power in the higher degrees, such as vertical deflections (cf. Jekeli, 1999). The other methods that were tested show lower accuracy for high degrees n (see Figure 5.5), which means that synthesis of gravity field functionals with high power in the higher degrees will show errors that are even more significant than those in the geoid heights shown in Figures 5.7 to 5.9.

Despite the fact that both the coefficient transformation method and Jekeli's (1988) method yield highly accurate coefficients, the formulas to compute them are completely different. It should be noted that the coefficient transformation method is very easily

applicable, because it simply computes the coefficients from a weighted summation over spherically approximated coefficients and avoids the use of ellipsoidal harmonics or hypergeometric series, which are embedded in Jekeli's solution. This simplicity is further illustrated by its limited required computation time. The coefficient transformation up to degree and order 360 takes less than one minute on a Pentium IV 2.4GHz PC.

5.5.4 Computation of very-high-degree and -order geopotential coefficients

Admittedly, the coefficient transformation approach does not perform well in the computation of solid SHCs above degree and order ~ 520 . This is due to the divergence of the iteration process above this degree and order, which also hampers the transformation from surface SHCs on the ellipsoid to solid SHCs. The reason for this was uncovered in Section 3.4, and lies in the loss of diagonal dominance of the weight functions λ_{nmi} . The same principle applies in this application.

The ellipsoidal integration method does not include any iterative procedure to acquire the solution, and it therefore does not encounter this problem. However, in a closed-loop simulation with a very-high-degree expansion of the Earth's gravity field (to $n = 1800$) the ellipsoidal integration approach also shows an inability to accurately recover solid SHCs of the disturbing potential for degrees $n > 720$. This is presumably caused by numerical errors in the computation of the weight functions $\lambda_{nmi}^{\text{ei}}$, and an indication for this is provided in Figure 5.13.

Figure 5.13 shows the weights up to $|i| = 3$ for the ellipsoidal integration method and for the coefficient transformation method. Weights of $|i| > 3$ are not shown for illustrative purposes. The weights for the latter method were derived from a rigorous matrix inversion, in the way indicated after Equation (5.66). The strong correspondence between both methods that was found in Figure 5.4 does not hold beyond degree $n \approx 720$.

It can be seen that in the ellipsoidal integration method, the weights of positive and

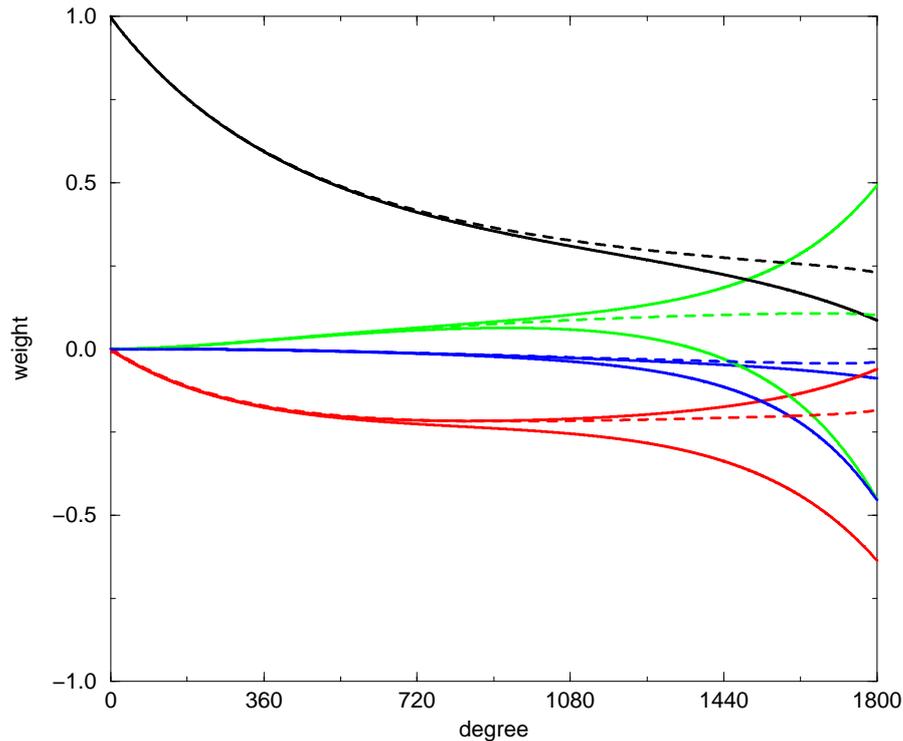


Figure 5.13: Weight functions $\lambda_{nm_i}(T, \tilde{T})$ for zonal harmonics ($m = 0$) with $i = 0$ (black), $|i| = 1$ (red), $|i| = 2$ (green) and $|i| = 3$ (blue) for the ellipsoidal integration method (solid lines) and the coefficient transformation method (dashed lines)

negative index i start to diverge, whereas the weights in the coefficient transformation method do not. Another observation from Figure 5.13 is that the sum of the absolute weights is close to one for all degrees n in the coefficient transformation method, but for the ellipsoidal integration method this is no longer the case for high degrees n .

It appears that numerical errors degrade the accuracy of the ellipsoidal integration method at high degrees. Unfortunately, the coefficient transformation with matrix inversion takes more than 10^2 times as long (for degree n to 2160) to process, and is, in fact, at least as time-consuming as the block-diagonal least-squares estimation procedure. Further investigation of the analytic computation of high-degree solid SHCs of the disturbing potential from gravity anomalies, in particular a more detailed study of the numerical errors that occur, is a possible topic of future research.

5.6 Summary

In this Chapter, the analytic computation of spherical harmonic geopotential coefficients from gravity anomalies on the ellipsoid was discussed. Three alternatives to the ellipsoidal harmonics method of Jekeli (1988) were derived, all of which express the geopotential coefficients as a weighted summation over spherically approximated coefficients. These three weighted summation methods are here called the upward-continuation method (Section 5.2), the ellipsoidal integration method (Section 5.3) and the coefficient transformation method (Section 5.4). The first two methods are based on methods proposed by Cruz (1986) and Sjöberg (2003c), but were here derived to a higher level of accuracy, and the latter method is novel. The three derivations are performed based on the rigorous fundamental equation of physical geodesy (Equation 2.52), whereas all existing methodologies are based on an approximation of this equation to the level of the square of the first numerical eccentricity of the ellipsoid e^2 .

It is shown numerically that this level of approximation is insufficient in the weighted summation methods and a closed-loop simulation revealed that the newly derived rigorous formulas allow us to model the Earth's gravity field up to two orders of magnitude more accurately than the existing weighted summation methods. The coefficient transformation method provides the most accurate results, with an average absolute error in geoid height of 0.3 mm (for $20 \leq n \leq 340$). The coefficients of low-order m are more accurate than those where the order $|m|$ is close to degree n , with relative accuracies ranging from approximately 10^{-8} to 10^{-4} on average. The coefficient transformation method performs better than the ellipsoidal harmonics method of Jekeli (1988) for coefficients of low order $|m|$, and optimal results can therefore be obtained from a combination of both methods. However, the computation of geopotential coefficients for degree $n > 360$ has proven problematic due to numerical inaccuracies, and remains an open area for future research, especially in view of recent efforts to compute a global geopotential model (EGM06) to degree and order 2160 (Kenyon et al., 2005).

6. ELLIPSOIDAL CORRECTIONS TO GRAVIMETRIC GEOID COMPUTATIONS

Ellipsoidal corrections to the spherical approximations in regional geoid computation have received a lot of attention in recent years (e.g., Grafarend and Ardalan, 1997; Fei and Sideris, 2000; Brovar et al., 2001; Heck and Seitz, 2003; Huang et al., 2003; Sjöberg, 2003c, 2004b; Hipkin, 2004). These corrections account for the errors introduced by the spherical and constant radius approximations that are applied in the classical solution of Stokes (1849). Many different approaches can be taken to derive the ellipsoidal corrections. The solutions to ellipsoidal BVPs derived in Chapters 3 and 4 can be utilised for this purpose, and in this Chapter it is shown how these solutions can yield equivalent corrections to Stokes's integration kernel and geoid heights.

In Section 6.1, the spherical approximation is outlined, which is necessary for a sound understanding of the definition of the ellipsoidal corrections applied here. Section 6.2 contains an overview of the various existing methods to compute ellipsoidal corrections and discusses the utility and efficiency of each. A new approach based on the transformation between solid and surface SHCs is derived in Section 6.3. In addition, numerical results are provided, also in view of the often used remove-compute-restore (RCR) approach to geoid computation (Section 6.4).

6.1 The spherical approximation to geoid computation

Ellipsoidal corrections to geoid heights are naturally related to the definition of the spherical approximation that is applied. Huang et al. (2003) note that various authors apply different definitions of the spherical approximation, which makes comparison of ellipsoidal corrections difficult. For example, Martinec and Grafarend (1997a) and Ardestani and Martinec (2000) define the spherical approximation in terms of ellip-

soidal coordinates to accommodate for the use of an ellipsoidal harmonic expansion. A more practical and more commonly used definition sets the spherical approximation equivalent to Stokes's (1849) solution for the computation of geoid heights in spherical polar coordinates, and this is the definition used here.

6.1.1 Definition of spherical approximation

Geoid heights N can be computed from geopotential coefficients using a spherical harmonic synthesis in combination with Bruns's formula (Equation 2.45)

$$N(\theta, \lambda) = \frac{1}{\gamma} \sum_{n=0}^{\infty} \left(\frac{R}{r_e} \right)^{n+1} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (6.1)$$

but it should be noted that some assumptions are embedded in Equation (6.1), as follows. As explained in Chapter 3, the creation of the SHCs \bar{T}_{nm}^R from a geopotential model such as, for example, EGM96 requires the definition of a reference gravity field. The GRS80 reference field (see Table 2.5) was chosen here. Secondly, the use of the simple Bruns formula (Equation 2.45) implies that the potential of the geoid W_0 is chosen equal to the potential of the reference ellipsoid U_Q .

In addition, it is assumed that the geopotential model and the reference model refer to the same tidal model, and that the geopotential model is capable of providing the potential inside the masses of the Earth, which strictly speaking it is not (e.g., Nsombo and Sjöberg, 1996; Rapp, 1997b). However, since only ellipsoidal corrections to the geoid heights are of interest here, Equation (6.1) is used, even though the terminology "geoid height" for N computed from Equation (6.1) can be disputed (Rapp, 1997b; Smith, 1998).

In many publications, the summation over n in Equation (6.1) is shown to run from $n = 2$ to ∞ , but this assumes the choice of a reference field model with a zero-degree term that equals the terrestrial gravitational constant and a coordinate system with its origin in the Earth's centre of mass. Here, generality is maintained by starting the summation at $n = 0$. The geoid computed via Equation (6.1) from the EGM96 geopotential model is shown in Figure 6.1.

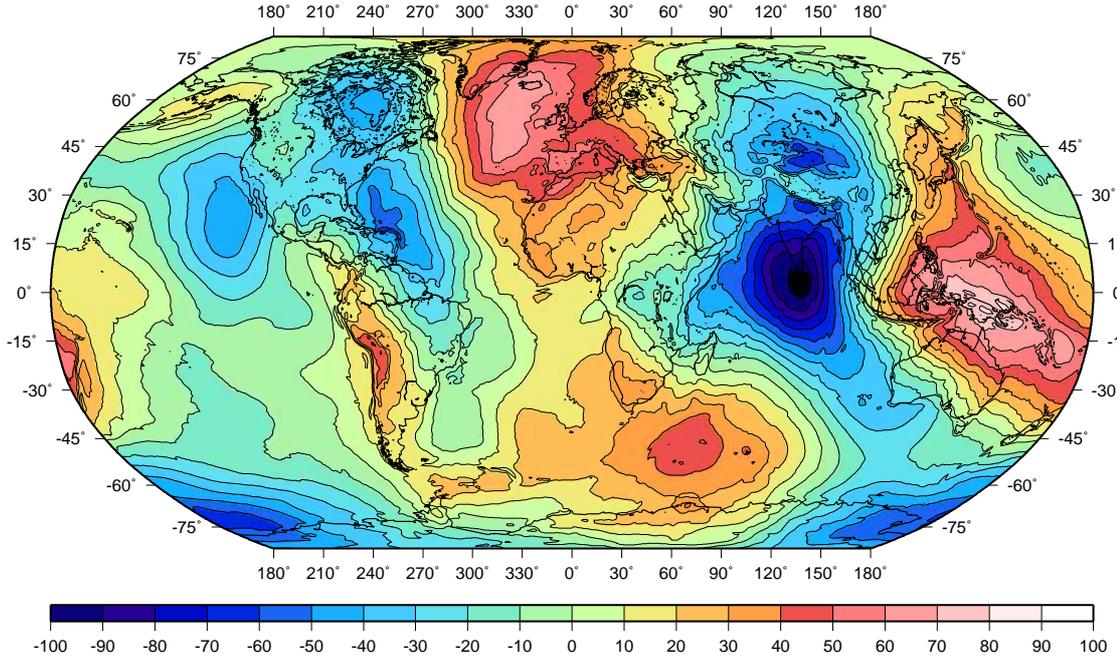


Figure 6.1: Geoid surface computed from Equation (6.1) using the EGM96 global geopotential model and GRS80 reference ellipsoid (units in m; Robinson projection)

Alternatively to Equation (6.1), a surface spherical harmonic expansion of the disturbing potential at the ellipsoid (Equation 2.84) can be utilised in combination with Bruns's formula (Equation 2.45)

$$N(\theta, \lambda) = \frac{1}{\gamma} \sum_{n=0}^{\infty} \sum_{m=-n}^n \bar{T}_{nm}^{r_e} \bar{Y}_{nm}(\theta, \lambda) \quad (6.2)$$

Stokes (1849) derives the geoid heights as a global integration of gravity anomalies in spherical approximation. This integral formula can be obtained by inserting the spectral relation between the disturbing potential and gravity anomalies in spherical and constant radius approximation (Equation 4.68) into Equation (6.1), and additionally approximating the ellipsoidal radius r_e by the spherical radius R so that the term $(R/r_e)^{n+1}$ vanishes

$$\tilde{N}(\theta, \lambda) = \frac{1}{\gamma} \sum_{n=0}^{\infty} \frac{R}{n-1} \sum_{m=-n}^n \overline{\Delta g}_{nm}^{r_e} \bar{Y}_{nm}(\theta, \lambda) \quad (6.3)$$

The spherical approximations introduced by this step are here indicated by a tilde over the geoid height N , and \tilde{N} is generally called the geoid height in spherical approximation.

On comparison of Equations (6.2) and (6.3), it can be seen that the approximation in

the spherical computation thus comes down to

$$\overline{T}_{nm}^{r_e} \approx \frac{R}{n-1} \overline{\Delta g}_{nm}^{r_e} \quad (6.4)$$

It is important to note the difference between the spherical approximation in geoid computation defined by Equation (6.4) and the spherical approximation in the computation of geopotential coefficients, which is defined by Equation (4.68)

$$\overline{T}_{nm}^R \approx \frac{R}{n-1} \overline{\Delta g}_{nm}^{r_e} \quad (6.5)$$

The reason for the difference is the additional constant radius approximation applied to Equation (6.1) to obtain Equation (6.3). It means that the spherical approximation to the geoid height only equals geoid heights computed from ‘spherically approximated’ geopotential coefficients defined by Equation (6.5) when the coefficients are synthesised on the sphere. The differences between geoid heights computed from spherically approximated geopotential coefficients and rigorously computed geopotential coefficients derived in Chapter 5 (Figure 5.6) do therefore not coincide with the differences between the geoid heights in Equations (6.1) and (6.3), since in Chapter 5 the spherical harmonic synthesis was correctly performed on the ellipsoid.

6.1.2 Stokes’s formula

Equation (6.3) requires the gravity anomalies to be given in the form of a set of surface spherical harmonic coefficients, but gravity anomalies are observed as point values. To account for this, the surface spherical harmonic analysis formula (Equation 2.85) can be inserted into Equation (6.3)

$$\tilde{N}(\theta, \lambda) = \frac{1}{4\pi\gamma} \sum_{n=0}^{\infty} \frac{R}{n-1} \sum_{m=-n}^n \int_{\sigma'} \Delta g(\theta', \lambda') \overline{Y}_{nm}(\theta', \lambda') d\sigma' \overline{Y}_{nm}(\theta, \lambda) \quad (6.6)$$

where primes were added to the co-latitude θ and longitude λ inside the integration to avoid confusion with the computation position of the geoid height \tilde{N} . If the order of summation and integration is reversed, a simplification can be obtained using the decomposition formula of spherical harmonics (e.g., Heiskanen and Moritz, 1967, p. 33)

$$P_n(\cos \psi) = \frac{1}{2n+1} \sum_{m=-n}^n \overline{Y}_{nm}(\theta, \lambda) \overline{Y}_{nm}(\theta', \lambda') \quad (6.7)$$

where ψ is the spherical distance between the points (θ, λ) and (θ', λ')

$$\cos \psi = \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos(\lambda' - \lambda) \quad (6.8)$$

Upon insertion of Equation (6.7) into Equation (6.6), Stokes's formula appears

$$\tilde{N}(\theta, \lambda) = \frac{R}{4\pi\gamma} \int_{\sigma'} S(\psi) \Delta g(\theta', \lambda') d\sigma' \quad (6.9)$$

where the spherical Stokes kernel $S(\psi)$ is a function of the spherical distance ψ only

$$S(\psi) = \sum_{n=0}^{\infty} \frac{2n+1}{n-1} P_n(\cos \psi) \quad (6.10)$$

A geoid computation using Stokes's formula is thus homogeneous, i.e. independent of the location of the point of interest, and isotropic, i.e. independent of the azimuth of the point of interest and the location of the gravity anomaly. This is of great importance in view of its practicality, as will be discussed in Section 6.2.2.

The infinite summation in Equation (6.10) converges very slowly, but a more efficient analytical closed solution can also be derived (e.g., Heiskanen and Moritz, 1967, p. 94)

$$S(\psi) = \frac{1}{\sin \frac{\psi}{2}} - 6 \sin \frac{\psi}{2} + 1 - 5 \cos \psi - 3 \cos \psi \ln\left(\sin \frac{\psi}{2} + \sin^2 \frac{\psi}{2}\right) \quad (6.11)$$

Equation (6.9) is used in the vast majority of local and regional geoid computations, where a global geopotential model is often used to model the long-wavelength part of the geoid, and the integration is limited to a spherical cap around the point of interest, whereby the Stokes kernel is sometimes modified to account for the resulting truncation error (e.g., Vaníček and Sjöberg, 1991; Sjöberg, 1991; Featherstone et al., 1998). However, this kernel modification does generally not account for the error caused by the spherical and constant radius approximations that yield Equation (6.4).

6.2 Existing methods for the computation of ellipsoidal corrections

Many authors have derived formulas for the computation of geoid heights in the rigorous ellipsoidal approximation (e.g., Rapp, 1981; Moritz, 1989; Martinec and Grafarend, 1997b; Fei and Sideris, 2000; Heck and Seitz, 2003; Sjöberg, 2003c). Practically without exception, these contributions present the results as ellipsoidal corrections to the

formulation in spherical approximation. These ellipsoidal corrections differ in various ways, giving rise to many methods to compute them, and somewhat disturbingly not all results are equivalent (Huang et al., 2003).

6.2.1 Types of ellipsoidal corrections

The ellipsoidal corrections in geoid computation can be presented in three distinctly different ways. The most common approach is to apply a correction to the geoid height (e.g., Heck and Seitz, 2003; Sjöberg, 2003c)

$$N(\theta, \lambda) = \frac{R}{4\pi\gamma} \int_{\sigma'} S(\psi) \Delta g(\theta', \lambda') d\sigma' + \delta N(\theta, \lambda) \quad (6.12)$$

Alternatively, the Stokes kernel can be modified (e.g., Martinec and Grafarend, 1997a; Fei and Sideris, 2000), noting that this is different to the kernel modification to reduce the truncation error

$$N(\theta, \lambda) = \frac{R}{4\pi\gamma} \int_{\sigma'} (S(\psi) + \delta S(\theta, \lambda, \theta', \lambda')) \Delta g(\theta', \lambda') d\sigma' \quad (6.13)$$

or ellipsoidal corrections to the gravity anomalies can be applied prior to computation

$$N(\theta, \lambda) = \frac{R}{4\pi\gamma} \int_{\sigma'} S(\psi) (\Delta g(\theta', \lambda') + \delta \Delta g(\theta', \lambda')) d\sigma' \quad (6.14)$$

Naturally, combinations of these three options are also possible, and have been derived by many authors (e.g., Molodenskii et al., 1962; Rapp, 1981; Moritz, 1989; Brovar et al., 2001), but this does generally not enhance the clarity of the method. An advantage of Equation (6.12) is that it gives an immediate idea of the magnitude, and thereby the relevance, of the ellipsoidal corrections, whereas use of Equations (6.13) and (6.14) require a global integration over the sphere to properly assess the effect of the corrections on the geoid height.

6.2.2 Ellipsoidal corrections from a global integration

The corrections to the Stokes kernel $\delta S(\theta, \lambda, \theta', \lambda')$ and/or to the gravity anomalies $\delta \Delta g(\theta', \lambda')$ have received much attention in literature. Formulations can be found by

application of Green's second identity (e.g., Molodenskii et al., 1962; Mather, 1973; Zhu, 1981; Fei, 2000; Fei and Sideris, 2000; Brovar et al., 2001), by application of an ellipsoidal harmonic expansion (e.g., Martinec and Grafarend, 1997b; Martinec, 1998; Ardestani and Martinec, 2000), or by application of a spherical harmonic expansion (e.g., Sjöberg, 2003c).

The ellipsoidal Stokes kernel can also be obtained in terms of spherical harmonics, using the theories developed in Chapters 3 and 4. The derivation is very similar to that in Section 6.1 (Equations 6.1 to 6.10), but without the spherical and constant radius approximations. Thus, instead of implementing the spherical spectral relation between the disturbing potential and the gravity anomalies, the rigorous spectral relation derived in Section 4.3.3 (Equation 4.86) is inverted and subsequently inserted into Equation (6.1)

$$N(\theta, \lambda) = \frac{1}{\gamma} \sum_{n=0}^{\infty} \left(\frac{R}{r_e}\right)^{n+1} \sum_{m=-n}^n \sum_{i=-\infty}^{\infty} \lambda_{nmi}(T, \Delta g) \overline{\Delta g}_{n+2i,m}^{r_e} \overline{Y}_{nm}(\theta, \lambda) \quad (6.15)$$

The surface spherical harmonic analysis (Equation 2.85) can be inserted to obtain the geoid height as an integration over gravity anomalies

$$N(\theta, \lambda) = \frac{1}{4\pi\gamma} \sum_{n=0}^{\infty} \left(\frac{R}{r_e}\right)^{n+1} \sum_{m=-n}^n \sum_{i=-\infty}^{\infty} \lambda_{nmi}(T, \Delta g) \times \int_{\sigma'} \Delta g(\theta', \lambda') \overline{Y}_{n+2i,m}(\theta', \lambda') d\sigma' \overline{Y}_{nm}(\theta, \lambda) \quad (6.16)$$

As for the spherical and constant radius approximation, the order of summation and integration can be reversed. However, the decomposition formula (Equation 6.7) cannot be applied, for two different reasons. Firstly, the weights $\lambda_{nmi}(T, \Delta g)$ depend on the order m , and can therefore not be placed outside the summation over m , and secondly, the degree of the spherical harmonics $\overline{Y}_{n+2i,m}(\theta', \lambda')$ depends on the summation index i and is therefore not always equal to that of the other spherical harmonics $\overline{Y}_{nm}(\theta, \lambda)$. Equation (6.16) can thus be written as

$$N(\theta, \lambda) = \frac{R}{4\pi\gamma} \int_{\sigma'} S_e(\theta, \lambda, \theta', \lambda') \Delta g(\theta', \lambda') d\sigma' \quad (6.17)$$

where $S_e(\theta, \lambda, \theta', \lambda')$ is the ellipsoidal generalisation of Stokes's kernel

$$S_e(\theta, \lambda, \theta', \lambda') = \frac{1}{R} \sum_{n=0}^{\infty} \left(\frac{R}{r_e}\right)^{n+1} \sum_{m=-n}^n \overline{Y}_{nm}(\theta, \lambda) \sum_{i=-\infty}^{\infty} \lambda_{nmi}(T, \Delta g) \overline{Y}_{n+2i,m}(\theta', \lambda') \quad (6.18)$$

It is immediately obvious that this equation is far less efficient than the closed form of the spherical approximation in Equation (6.11), for three main reasons.

1. The weights $\lambda_{nmi}(T, \Delta g)$ do not follow directly from the derivations in Section 4.3.3. Equation (4.87) provides the weights $\lambda_{nmi}(\Delta g, T)$ that transform the surface spherical harmonic coefficients of the gravity anomalies into the solid spherical harmonic coefficients of the disturbing potential. However, the inverse relation is needed here, which requires a rigorous matrix inversion. Since the weights increase with increasing degree n , due to the constant radius approximation, weights up to very high degree and order need to be taken into account, and the matrix that needs to be inverted will thus be fairly large and non-diagonally dominant, as explained in Section 4.3.
2. Equation (6.18) involves slowly converging series of weighted products of spherical harmonic base functions, and the derivation of a closed solution is not as straightforward as it is in the case of the spherical and constant radius approximation. Martinec (1998) and Fei (2000) present approximations of the ellipsoidal Stokes kernel based on an ellipsoidal harmonic expansion and Green's second identity, respectively.
3. Solutions to points 1 and 2 above cannot overcome the third impracticality of the ellipsoidal Stokes kernel in Equation (6.18): the ellipsoidal form of Stokes kernel cannot be expressed as a function of the spherical distance ψ . This can be seen from the fact that on an ellipsoid, a line of points at equal surface distance from a certain computation point will not form a circle as it does on the sphere, nor will it be of equal shape for computation points at different latitudes. In other words, the process becomes non-isotropic and non-homogeneous. Therefore, the ellipsoidal Stokes kernel will always depend upon four parameters, i.e., longitude and latitude of the computation point and longitude and latitude of the integration point, whereas the spherical Stokes kernel only depends on one parameter (the spherical distance), and is thus homogeneous and isotropic. The ellipsoidal kernel will thus be less efficient.

The ellipsoidal correction to the geoid height δN in Equation (6.12) is more practically useful than the modifications to the Stokes kernel or gravity anomalies, since it can simply be applied after the spherical geoid computation has been performed. Secondly, it avoids singularity in the vicinity of the computation point that the Stokesian integration suffers from, although the spherical Stokes integral naturally still contains this singularity. Ellipsoidal corrections to the geoid height can be computed in an efficient way, avoiding the need for integration, by evaluation of a global spherical harmonic geopotential model (e.g., Heck and Seitz, 2003; Sjöberg, 2003c).

6.2.3 Ellipsoidal corrections from spherical harmonic geopotential coefficients

The computation of ellipsoidal corrections from a global set of geopotential coefficients has been proposed by many authors (e.g., Rapp, 1981; Moritz, 1989; Heck and Seitz, 2003; Sjöberg, 2003c). Lelgemann (1970) had already realised that the low frequencies are dominant in the ellipsoidal corrections, which permits the truncation of high-degree coefficients. The level of error introduced by various truncation degrees is discussed in Section 6.4. The dominance of the low frequencies has been confirmed by recent numerical studies into the ellipsoidal corrections (e.g., Heck and Seitz, 2003; Huang et al., 2003; Hipkin, 2004; Ellmann, 2005). Huang et al. (2003) state that the contribution of coefficients beyond degree 20 only contributes 10% of the total ellipsoidal correction.

Heck and Seitz (2003, Equation 59) and Sjöberg (2003c, Equation 24) have shown that the ellipsoidal corrections to geoid heights can be presented as a summation over solid SHCs of the disturbing potential, roughly of the form

$$\delta N(\theta, \lambda) \doteq \sum_{n=0}^{\infty} \sum_{m=-n}^n \sum_{i=-\infty}^{\infty} \tau_{nmi} \bar{T}_{n+2i,m} \bar{Y}_{nm}(\theta, \lambda) \quad (6.19)$$

where τ_{nmi} depend on degree n , order m and the parameters of the reference gravity field. In both Heck and Seitz (2003) and Sjöberg (2003c), the boundary condition is approximated to the order of e^2 and the summation over i in Equation (6.19) is truncated after the first terms of $|i| = 1$. It was shown in Section 3.4 that this is highly inaccurate for high-degree coefficients in the transformation between surface and solid coefficients. This is due to the fact that the terms of index i have an impact of order

$\mathcal{O}(e^{2i}n^i)$. However, Heck and Seitz (2003) show that the terms of index i in Equation (6.19) is of order $\mathcal{O}(e^{2i}n^0)$, which means that disregarding all terms of $|i| > 1$ provides ellipsoidal corrections with an accuracy of order e^2 .

Using the spectral relations derived in Chapters 3 and 4, τ_{nmi} can be derived rigorously for all indices i . Moreover, the approximation of surface SHCs by solid SHCs that is required in the method of Heck and Seitz (2003) can be avoided, and a rigorous solution can be obtained. This is here called the surface harmonics method to ellipsoidal corrections, described next.

6.3 The surface harmonics method to ellipsoidal corrections

This method provides an easy and efficient way to compute ellipsoidal corrections to geoid heights from geopotential coefficients. The derivations here are made for any arbitrary reference sphere, contrary to many other ellipsoidal corrections which were only derived for a certain set reference sphere with radius R . For example, Moritz (1989) derives ellipsoidal corrections for the ‘mean’ radius $R = (a^2b)^{1/3}$ (which is the radius of the sphere with the same volume as the ellipsoid), Fei and Sideris (2000) derive corrections for $R = b$ and $R = (a^2b)^{1/3}$, and Heck and Seitz (2003) derive corrections for $R = a$. The choice of the radius of the reference sphere proves to have a large impact on the magnitude and spectral sensitivity of the corrections, and this is discussed in Section 6.3.3 and 6.3.4.

6.3.1 Derivation of the correction formulas

The ellipsoidal correction to the geoid height can be computed from a set of solid spherical harmonic geopotential coefficients from the difference between Equations (6.1) and (6.3)

$$\delta N = \frac{1}{\gamma} \sum_{n=0}^{\infty} \sum_{m=-n}^n \left[\left(\frac{R}{r_e} \right)^{n+1} \bar{T}_{nm}^R - \frac{R}{n-1} \overline{\Delta g}_{nm}^{r_e} \right] \bar{Y}_{nm}, \quad n \neq 1 \quad (6.20)$$

where the surface harmonic coefficients of the gravity anomalies $\overline{\Delta g_{nm}^{r_e}}$ can be computed from solid spherical harmonic geopotential coefficients $\overline{T_{nm}^R}$ via the forward transformation in Equation (4.86). Equation (6.20) does not hold for the case $n = 1$, due to the division over $n - 1$, and this special case is discussed in Section 6.3.2. An alternative for Equation (6.20) is possible using a surface spherical harmonic expansion of the disturbing potential (Equation 6.2)

$$\delta N = \frac{1}{\gamma} \sum_{n=0}^{\infty} \sum_{m=-n}^n \left(\overline{T_{nm}^{r_e}} - \frac{R}{n-1} \overline{\Delta g_{nm}^{r_e}} \right) \overline{Y}_{nm}, \quad n \neq 1 \quad (6.21)$$

Equation (6.21) is much more practical than Equation (6.20), because the term between the brackets is independent of position. It therefore provides the surface SHCs of the ellipsoidal corrections to the disturbing potential

$$\delta T = \sum_{n=0}^{\infty} \sum_{m=-n}^n \overline{\delta T_{nm}^{r_e}} \overline{Y}_{nm} \quad (6.22)$$

where the surface SHCs $\overline{\delta T_{nm}^{r_e}}$ are given by

$$\overline{\delta T_{nm}^{r_e}} = \overline{T_{nm}^{r_e}} - \frac{R}{n-1} \overline{\Delta g_{nm}^{r_e}}, \quad n \neq 1 \quad (6.23)$$

and the ellipsoidal corrections to geoid heights follow from Bruns's formula

$$\delta N = \frac{\delta T}{\gamma} \quad (6.24)$$

Thus, all that is required to obtain a global set of ellipsoidal corrections are surface spherical harmonic expansions of the disturbing potential and the gravity anomalies. These can be obtained from a set of geopotential coefficients. The surface SHCs of the disturbing potential and gravity anomalies can be computed via a synthesis-analysis procedure, or alternatively and much more efficiently via the transformations derived in Chapters 3 and 4 (Equations 3.81 and 4.86).

This new approach to the computation of ellipsoidal corrections to geoid heights is arguably of great simplicity. Importantly, it only requires forward transformations from solid SHCs $\overline{T_{nm}^R}$ to surface coefficients $\overline{T_{nm}^{r_e}}$ and $\overline{\Delta g_{nm}^{r_e}}$ that are numerically stable to at least degree and order 360 (see Sections 3.4.1 and 5.5.2), and the problems encountered with the inverse transformation for coefficients of high degrees in Chapters 3 and 5 are avoided. No iterations or approximations are required, making this method for the

computation of ellipsoidal corrections theoretically exact, unlike any existing method, because all existing methods involve approximations to the order of e^2 .

Naturally, the spherical harmonic expansions involved will in practice always be truncated, and so will the summations in the transformation formulas. However, the errors resulting from this truncation will always be small, since the ellipsoidal corrections are dominated by the long wavelengths. This is shown for various degrees of truncation in Section 6.4.

6.3.2 The zero- and first-degree coefficients

The solid SHCs \bar{T}_{nm}^R of degree $n = 0$ and $n = 1$ are commonly set to zero by the choice of the reference ellipsoid. The zero-degree solid SHC of the Earth's gravitational potential equals the terrestrial gravitational constant μ (e.g., Heiskanen and Moritz, 1967), and $\bar{T}_{0,0}^R$ will thus become zero if the terrestrial gravitational constant of the reference gravity field equals the terrestrial gravitational constant of the geopotential model.

The first-degree solid SHCs $\bar{T}_{1,-1}^R$, $\bar{T}_{1,0}^R$ and $\bar{T}_{1,1}^R$ will be zero if the origin of the coordinate system coincides with the centre of mass of the Earth. They are sometimes called the forbidden or inadmissible harmonic coefficients, which according to Heiskanen and Moritz (1967, p. 62) "must vanish in any spherical harmonic expansion of the Earth's potential". However, this does not hold for a surface spherical harmonic expansion of the Earth's gravity potential from function values on any other surface than a sphere, because a one-to-one relation between solid SHCs and surface SHCs from a non-spherical surface does not exist. This is acknowledged by Heck (1991) for surface SHCs from an ellipsoid, and this case is described next.

It can be seen from the transformation between surface and solid SHCs (Equation 3.73) that the surface SHCs $\bar{T}_{1,m}^e$ will not vanish if the solid SHCs $\bar{T}_{1,m}^R$ are equal to zero,

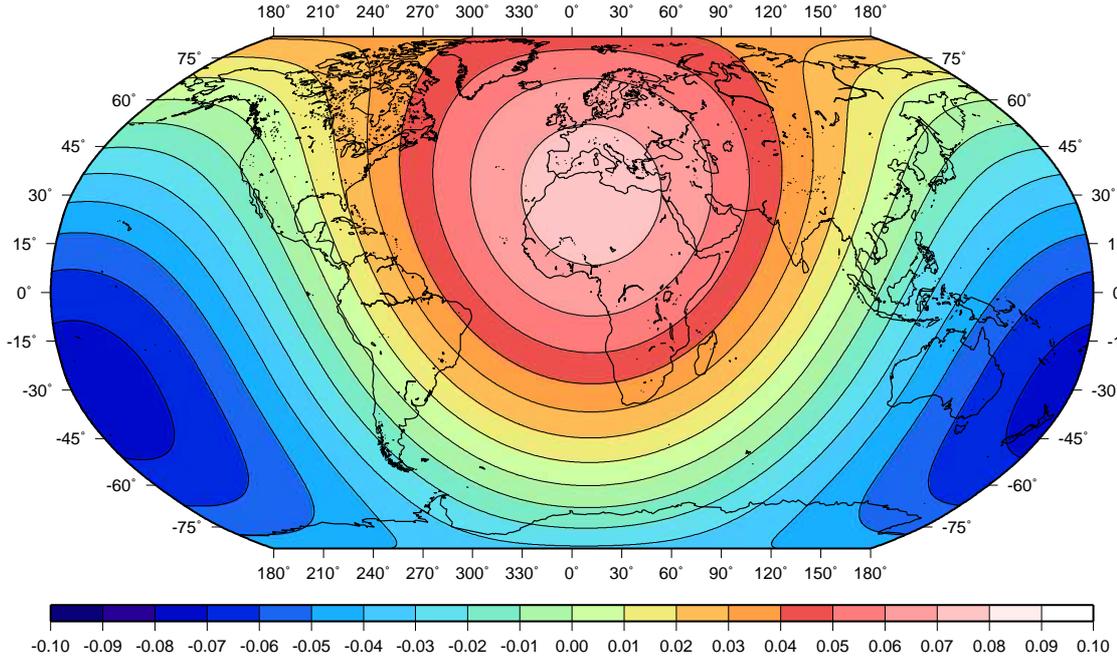


Figure 6.2: Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 1$ (units in m; Robinson projection)

but depend on all solid SHCs of equal order m and odd degree n

$$\bar{T}_{1,m}^{r_e} = \sum_{i=1}^{\infty} \lambda_{1,m,-i} \bar{T}_{1+2i,m}^R \quad (6.25)$$

Similarly, according to Equation (4.86), the first-degree coefficients in the surface spherical harmonic expansion of gravity anomalies on the ellipsoid will not vanish either

$$\bar{\Delta g}_{1,m} = \sum_{i=1}^{\infty} \lambda_{1,m,-i} (\Delta g, T) \bar{T}_{1+2i,m}^R \quad (6.26)$$

even though they do vanish in spherical approximation (see Equation 4.68). The inverse of Equation (6.26), which can be computed iteratively or via a matrix inversion, reads

$$\bar{T}_{1,m}^R = \sum_{i=1}^{\infty} \lambda_{1,m,-i} (T, \Delta g) \bar{\Delta g}_{1+2i,m} \quad (6.27)$$

which shows that the gravity data should fulfill three consistency conditions (Heck, 1991) (the cases $m = -1, 0, 1$), when $\bar{T}_{1,m}^R$ are indeed equal to zero. This will generally not be the case for empirical data, but it can be expected that the effects on practical computations are negligible (Heck, 1991).

Since the first-degree terms disappear in spherical approximation, the first-degree surface SHCs of the ellipsoidal correction simply equal the first-degree surface SHCs of

the disturbing potential

$$\overline{\delta T}_{1,m}^{r_e} = \overline{T}_{1,m}^{r_e} \quad (6.28)$$

Huang et al. (2003) reveal that the methods for the computation of ellipsoidal corrections by Molodenskii et al. (1962), Moritz (1989) and Fei (2000) fail to reproduce the first-degree term. Later tests with synthetic data by Huang, equal to the tests described in Huang et al. (2003), have revealed that the method by Heck and Seitz (2003) correctly computes the first-degree term, as well as all other terms, with an accuracy within the computation noise (± 1 cm). Comparison of the results of Equation (6.28) to the results of Heck and Seitz (2003) shows a very strong resemblance, as can be concluded from comparison of Figure 6.2 with Heck and Seitz (2003, Figure 5). The small differences between both methods, which amount to less than 1mm for any point on Earth, can possibly be explained by the fact that the coefficients in Heck and Seitz (2003) are only accurate up to the order of e^2 .

6.3.3 Ellipsoidal corrections for various reference spheres

It can be seen from Equation (6.23) that the ellipsoidal corrections to geoid heights do not only depend on the surface SHCs of disturbing potential and gravity anomalies, but also on the radius of the reference sphere R . The radius R effectively works as a scale factor to the geoid heights computed by Stokes's integral (Equation 6.9). The choice of R has a substantial effect on the ellipsoidal corrections. For example, the difference between the ellipsoidal correction, where R is chosen equal to the semi-major axis of the ellipsoid a and the semi-minor axis of the ellipsoid b , respectively, is of the order of the flattening of the ellipsoid. This effect is as large as the corrections themselves, since ellipsoidal corrections are of the order of the flattening (e.g., Rummel, 1985).

Some existing formulas for ellipsoidal corrections, such as those by Heck and Seitz (2003), only hold for a reference sphere equal to the semi-major axis a , and if another reference sphere is used, all formulas should be re-derived. In the derivations of Sjöberg (2003c, 2004b) a special scale factor is introduced to allow computation for other reference spheres, and this scaling is separated from the ellipsoidal corrections. Here, the

scaling was included in the ellipsoidal corrections by introduction of the scale factor c (Equation 3.64) for reasons of convenience.

The ellipsoidal corrections for $R = a$ and $R = b$ are shown in Figures 6.3 and 6.4 respectively, and the corrections for the ‘mean’ spherical radius $R = (a^2b)^{1/3}$ ($R = 6,371,005$ m) (Moritz, 1980) are shown in Figure 6.5, where the GRS80 values of a and b (Table 2.5) were used. The spherical harmonic expansion was evaluated for degrees up to 340, to avoid the error that is introduced in the SHCs close to the maximum degree, which occurs due to the absence of SHCs of higher degrees (see Section 5.5.2).

The corrections are based on the EGM96 geopotential model (Lemoine et al., 1998) that was also used in the numerical experiments in Chapters 3 and 5. The choice for EGM96 in this application is based on the fact that most earlier numerical studies into the ellipsoidal corrections (e.g., Heck and Seitz, 2003; Huang et al., 2003; Ellmann, 2005) also use EGM96, which makes the results easy to compare. For example, Figure 6.3 is very similar to Heck and Seitz (2003, Figure 3), which increases the probability that both methods are valid. Moreover, the validity of the transformations from solid SHCs of the disturbing potential to surface SHCs of the disturbing potential and gravity anomaly was already confirmed by the numerical experiments in Chapters 3 and 5 respectively, and therefore no additional closed-loop simulation is needed here.

It can be seen from Figures 6.3 to 6.5 that the ellipsoidal corrections show some similarities with the geoid computed from EGM96 (Figure 6.1). In particular, the lowest point of the geoid appears in the vicinity of Sri Lanka, and this corresponds with the largest negative ellipsoidal correction in Figures 6.3 to 6.5. Likewise, there is a high in the geoid over Papua New Guinea, and this is also reflected in the ellipsoidal corrections. This indicates that the spherical approximation underestimates the absolute values of the geoid heights. For $R = b$, this effect is more profound than for $R = a$, which is understandable given that R scales the geoid heights. Since even $R = a$ does apparently not scale the geoid heights by a high enough factor, it follows that the ellipsoidal corrections are smaller when $R > a$.

Figure 6.6 shows the ellipsoidal corrections for $R = a + 20\text{km}$. Comparison with Figure

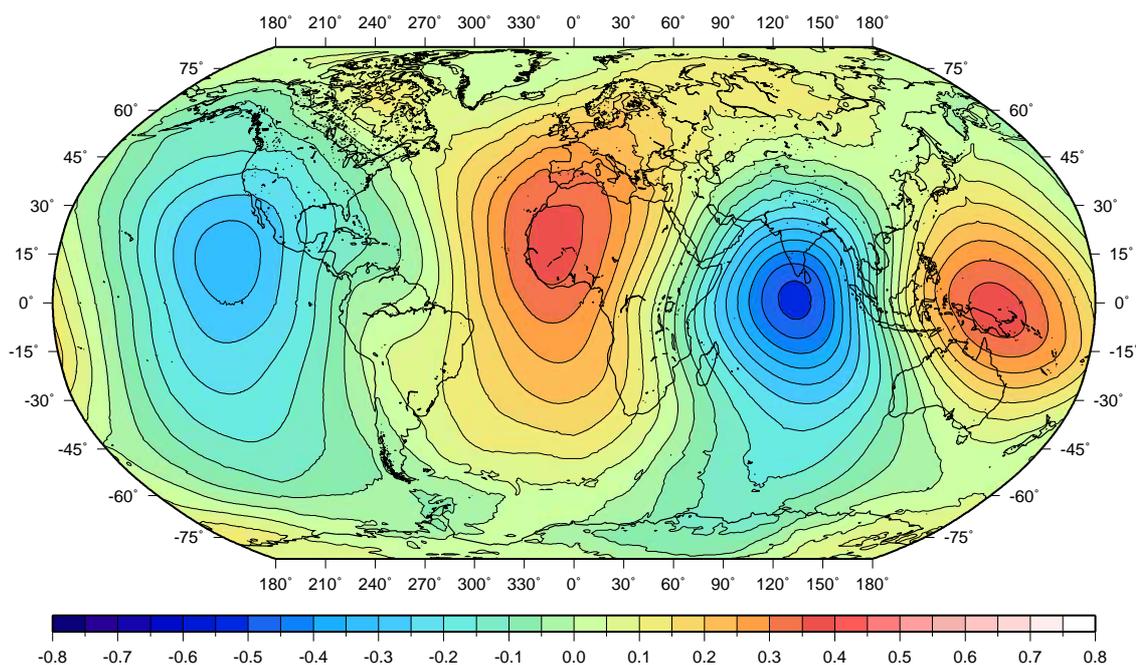


Figure 6.3: Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 340$ and a reference sphere with radius $R = a$ (units in m; Robinson projection)

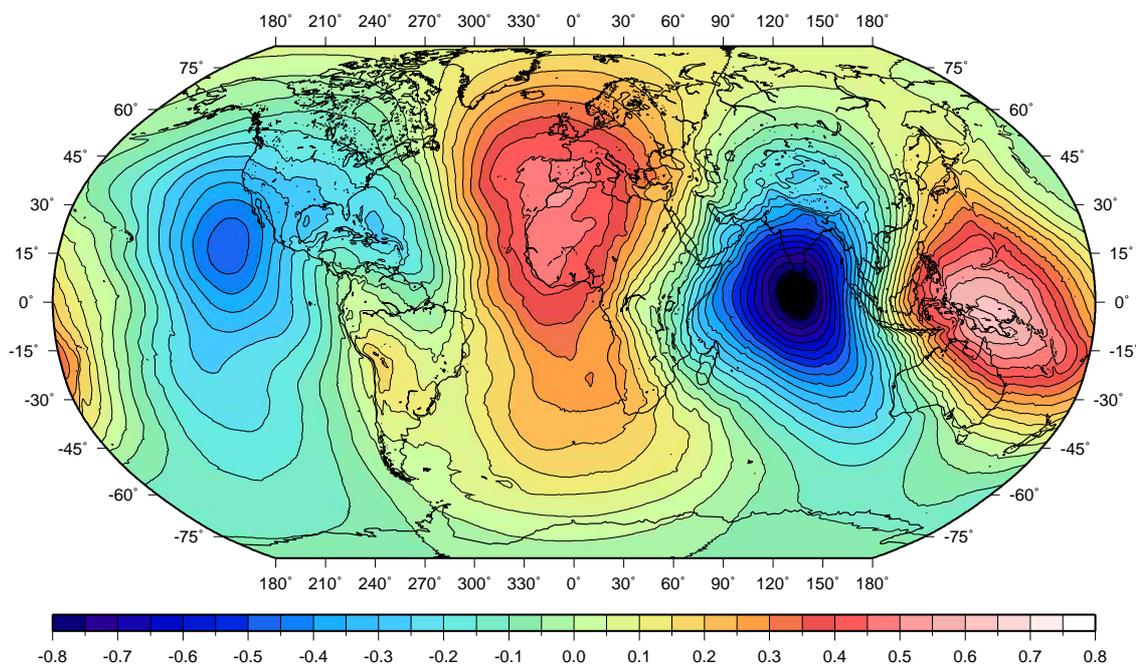


Figure 6.4: Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 340$ and a reference sphere with radius $R = b$ (units in m; Robinson projection)

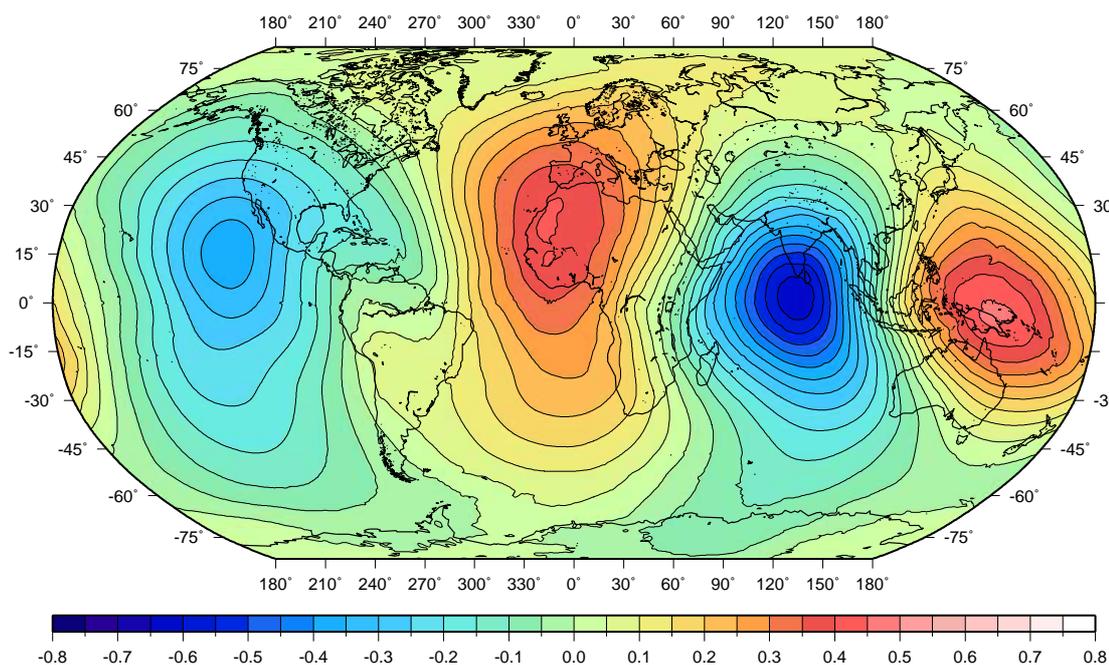


Figure 6.5: Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 340$ and a reference sphere with radius $R = (a^2b)^{1/3}$ (units in m; Robinson projection)

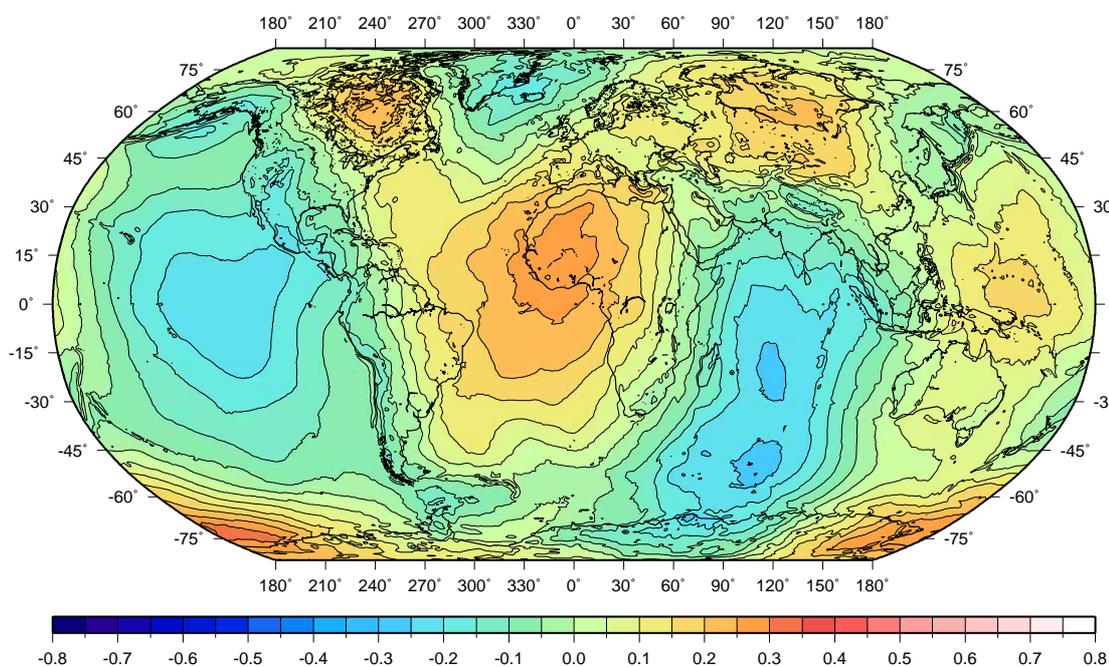


Figure 6.6: Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 340$ and a reference sphere with radius $R = a + 20\text{km}$ (units in m; Robinson projection)

	$R = b$	$R = (a^2b)^{1/3}$	$R = a$	$R = a + 20\text{km}$
minimum	-0.863	-0.637	-0.524	-0.304
maximum	0.655	0.469	0.380	0.333
mean	0.006	0.008	0.009	0.011
absolute mean	0.164	0.125	0.114	0.106
standard deviation	0.219	0.170	0.150	0.127

Table 6.1: Statistics of the ellipsoidal corrections for a reference sphere with radius R equal to semi-minor axis b , ‘mean’ radius $(a^2b)^{1/3}$, semi-major axis a and semi-major axis $a + 20\text{km}$, obtained from geopotential model EGM96 and degrees $0 \leq n \leq 340$ synthesised on a $30'$ grid (units in m)

6.3 shows that the ellipsoidal corrections are in this case increased around the poles, but attenuated around the equator, where the largest corrections appeared. It can be seen from the statistics in Table 6.1 that the minimum and maximum values of the ellipsoidal corrections are substantially lower for $R = a + 20\text{km}$ than in the cases where R is chosen smaller, as are the absolute mean and standard deviation.

This leads to the conclusion that the spherical approximation obtains the ‘best’ results, in a least-squares sense, when the gravity anomalies are assumed to lie on a sphere with a radius that is larger than the equatorial radius. However, when increasing the radius R further, the ellipsoidal corrections become larger again, and errors up to several decimetres will always remain in the spherical approximation for any choice of R . Moreover, when $R > a$ the high frequencies are amplified, which makes it more difficult to model the ellipsoidal corrections accurately from a global geopotential model. Ideally, the ellipsoidal corrections should contain as little power in the high degree coefficients as possible, to allow highly-accurate modelling from low-degree coefficients only. This can be achieved with an optimal choice of the reference sphere R , and this is shown next.

6.3.4 Ellipsoidal corrections with varying reference sphere radius

As stated, the vast majority of numerical studies into ellipsoidal effects either choose the radius of the reference sphere $R = a$ (e.g., Huang et al., 2003; Heck and Seitz, 2003) or $R = (a^2b)^{1/3}$ (e.g., Ellmann, 2005). Interestingly, it can be observed from Figures 6.3 to 6.5 that the ellipsoidal corrections show the least small-scale variations where R equals or is close to the ellipsoidal radius r_e : in Figure 6.3, where $R = a$, the least small-scale variations occur around the equator; in Figure 6.4, where $R = b$, the least small-scale variations occur around the poles; and in Figure 6.5, where $R = (a^2b)^{1/3}$, the least small-scale variations occur along the mid-latitudes ($\theta \approx 45^\circ$ and $\theta \approx 135^\circ$). This indicates that the short-wavelength effects are mainly caused by upward- or downward-continuation from the reference sphere to the ellipsoid.

It can therefore be deduced that ellipsoidal corrections with relatively more power in the low frequencies can be obtained if the radius of the reference sphere R is chosen equal to the ellipsoidal radius r_e at the point of computation. Admittedly, the radius of the reference sphere R is required to be constant for all points on the sphere in order to derive the spherical Stokes formula (Equation 6.9). However, from a pragmatic point of view, in the computation of the geoid height at a certain point, the choice $R = r_e$ simply means that the global set of gravity anomalies are all assumed to lie on the sphere that intersects the ellipsoid at that computation point and has the same origin as the ellipsoid. Another point at a different latitude can thus be computed with a different constant radius. It should be noted that the choice $R = r_e$ is applied in the computation of, for example, geoid models for the British Isles (Featherstone and Olliver, 1994) and Australia (Featherstone et al., 2001), although the choice of the reference sphere is seldomly stated explicitly in publications.

The computation of ellipsoidal corrections with $R = r_e$ is slightly more complicated than in case of a constant radius. The correction to the geoid height becomes, according to Equation (6.21)

$$\delta N = \frac{1}{\gamma} \sum_{n=0}^{\infty} \sum_{m=-n}^n \left(\bar{T}_{nm}^{r_e} - \frac{r_e}{n-1} \bar{\Delta g}_{nm}^{r_e} \right) \bar{Y}_{nm} \quad (6.29)$$

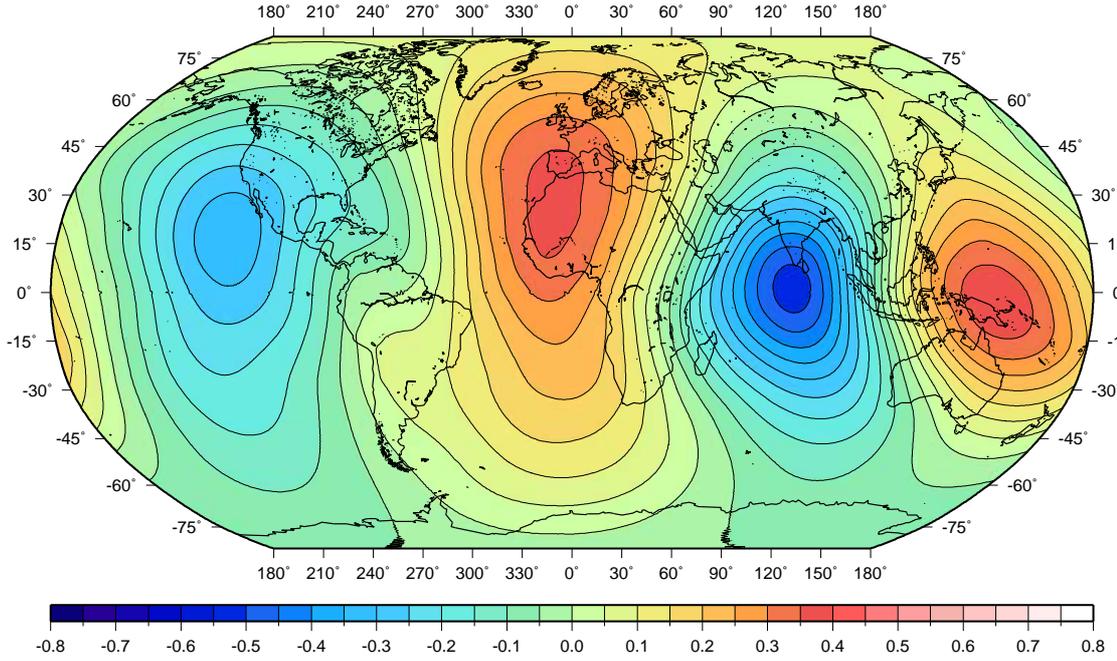


Figure 6.7: Ellipsoidal corrections to the geoid heights computed by Equations (6.22) to (6.24) using the EGM96 geopotential model with degrees $0 \leq n \leq 340$ and $R = r_e$ (units in m; Robinson projection)

Since the ellipsoidal radius r_e depends on latitude, the surface SHCs of the correction to the disturbing potential $\overline{\delta T}_{nm}^{r_e}$ cannot be recovered straight away by comparing Equation (6.29) to Equation (6.22). However, it is possible to remove the dependence of latitude in r_e using the transformation derived in Section 3.3. The coefficient transformation for SHCs that are multiplied by the term $(R/r_e)^{n+1}$ is given in Equation (3.73), and the special case $(R/r_e)^{-1}$ can be applied to obtain an expression for the surface SHC of the correction to the disturbing potential

$$\overline{\delta T}_{nm}^{r_e} = \overline{T}_{nm}^{r_e} - \frac{R}{n-1} \sum_{i=-\infty}^{\infty} \sum_{k=|i|}^{\infty} \alpha_{-2-2i,k} \overline{K}_{n-2i,m}^{2i,2k} \overline{\Delta g}_{n+2i,m}^{r_e} \quad (6.30)$$

The ellipsoidal corrections can then be computed via Equations (6.22) and (6.24), and the results are shown in Figure 6.7. It can be seen that the corrections in the case of $R = r_e$ contain less power in the high degrees than in the cases where R was chosen constant (Figures 6.3 to 6.6). The strong dominance of the low-degree coefficients can also be derived from the degree variance spectra of $\overline{\delta T}_{nm}^{r_e}$ for various values of R that are shown in Figure 6.8. The degree variances for the case $R = r_e$ decrease very rapidly, but at about the level of 10^{-7} they start to oscillate heavily. This is caused by numerical inaccuracies that only affect coefficients of even degree n .

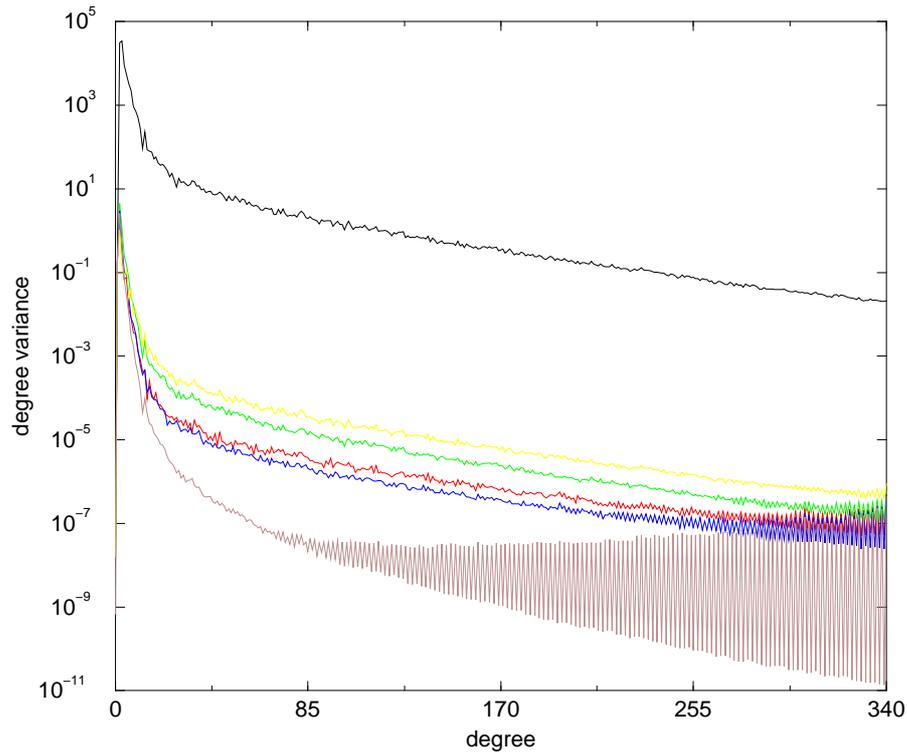


Figure 6.8: From top to bottom: Degree variances of the solid SHCs \overline{T}_{nm}^R (black) and of the surface SHCs $\overline{\delta T}_{nm}^{r_e}$ for $R = a + 20\text{km}$ (yellow), $R = b$ (green), $R = a$ (red), $R = (a^2b)^{1/3}$ (blue) and $R = r_e$ (brown)

The same can be seen in Figure 6.8 for the degree variances of the surface SHCs $\overline{\delta T}_{nm}^{r_e}$ for other values of R , although it is less obvious there because the degree variances are larger. Without the numerical inaccuracies, the degree variances of even degrees would be of the same order of magnitude as those of odd degrees, and the contribution of degrees $n > 70$ is less than a millimetre for any point on the ellipsoid, as will be shown in Section 6.4. The surface spherical harmonic series with SHCs computed via Equation (6.30) can therefore safely be truncated at degree 70, avoiding the influence of the numerical inaccuracies.

6.4 Ellipsoidal corrections in view of the remove-compute-restore (RCR) approach to geoid computation

In local or regional geoid computation, the so-called remove-compute-restore (RCR) approach (e.g., Torge, 1991), which combines the local gravity data with long-wavelength information from a global geopotential model, is often applied, either with or without modifications to the spherical Stokes kernel (e.g., Vaníček and Sjöberg, 1991; Featherstone et al., 1998, 2004). The use of the RCR approach has important implications for the application of ellipsoidal corrections, as is shown by the numerical investigations of, e.g., De Min (1995), Ardestani and Martinec (2003b) and Ellmann (2005).

6.4.1 The remove-compute-restore approach

In the classical RCR approach, i.e., without a modified kernel (e.g., Vincent and Marsh, 1973; Rummel and Rapp, 1976), the long-wavelength part of the geoid is provided by a global geopotential model, and local gravity data are used to provide the short-wavelength part. The geoid height is thus split into two parts

$$N(\theta, \lambda) = N_1(\theta, \lambda) + N_2(\theta, \lambda) \quad (6.31)$$

where N_1 is the contribution from the global geopotential model and N_2 is the high-frequency contribution from a Stokesian integration over the local gravity anomalies. The spherical harmonic geopotential model is evaluated until a certain maximum degree and order n_{\max} . The optimal value of n_{\max} depends mainly on the quality of the local gravity data and of the global geopotential model, but is commonly simply chosen equal to the maximum degree of the global model used. Very roughly, n_{\max} can be as low as 20 for satellite-only models (Vaníček and Krakiwsky, 1986), but is usually higher (up to or even beyond 360) when a combined model, such as EGM96, is used.

The main advantage of the RCR approach is that a global set of gravity anomalies is not required and local data can be used to refine the coarse geoid resulting from the global model, although small errors will still be present because the short-wavelength

part of gravity anomalies outside the region is neglected, the truncation error. However, the approach also introduces some complications, and the most important of these is that the local gravity data also contain a long-wavelength contribution. Therefore, to avoid accounting for the long wavelengths twice, the low-frequency part of the local gravity anomalies needs to be removed beforehand. This can be done by computing gravity anomalies from the global model until degree and order n_{\max} and subtracting this from the local gravity anomalies.

This step is often performed in spherical approximation using Equation (4.67) or a slightly better approximation that only includes the spherical approximation to the boundary condition, but not the constant radius approximation

$$\Delta g_{n_{\max}}(r_e, \theta, \lambda) = \sum_{n=0}^{n_{\max}} \left(\frac{R}{r_e}\right)^{n+1} \frac{n-1}{r_e} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (6.32)$$

where the subscript n_{\max} indicates that this gravity anomaly is only evaluated for $0 \leq n \leq n_{\max}$. Equation (6.32) was, for example, used in the construction of a geoid model for Australia (Featherstone et al., 2001). It is, however, reasonably straightforward to compute the band-limited gravity anomalies in a more exact way, using the rigorous boundary condition (Equation 2.52). The expansion of the normal derivative of the disturbing potential in Equation (4.16) can be inserted into the boundary condition to obtain a rigorous expression for the gravity anomalies in terms of spherical harmonic geopotential coefficients

$$\Delta g_{n_{\max}} = \frac{1}{R} \sum_{n=0}^{n_{\max}} \left(\frac{R}{r_e}\right)^{n+2} \sum_{m=-n}^n \bar{T}_{nm}^R \left[\frac{r_e}{\gamma} \frac{\partial \gamma}{\partial h} + (n+1) \cos \phi + \sin \phi \frac{\partial}{\partial \theta} \right] \bar{Y}_{nm} \quad (6.33)$$

where it should be noted that R here needs to be equal to the reference sphere that the SHCs of the disturbing potential refer to, and cannot be chosen freely (unless of course the SHCs \bar{T}_{nm}^R are also transformed to the new radius R using Equation (3.7)). The application of Equation (6.33) instead of the spherical approximation in Equation (4.67), can be seen as an ellipsoidal correction to the long-wavelength term of the gravity anomalies.

The derivation of $\Delta g_{n_{\max}}$ on the ellipsoid is performed in an alternative way from the method described here by Hipkin (2004), who expresses the long-wavelength effect of

the gravity anomaly on the ellipsoid (in the current notation) as

$$\Delta g_{n_{\max}} = \frac{1}{R} \sum_{n=0}^{n_{\max}} \left(\frac{R}{r_e}\right)^{n+2} \sum_{m=-n}^n \bar{T}_{nm}^R \left[(n-1) \cos \phi - \sin \phi \frac{\partial}{\partial \theta} \right] \bar{Y}_{nm} - E(\theta) N(\theta, \lambda) \quad (6.34)$$

where

$$E(\theta) = \frac{\gamma(2 - e^2 - e^2 \cos^2 \vartheta) \sqrt{1 - e^2 \cos^2 \vartheta}}{a(1 - e^2)} - \frac{2\gamma \cos \phi}{r_e} + 2\omega^2 \quad (6.35)$$

The main difference between Hipkin's (2004) solution (Equation 6.34) and Equation (6.33) is that Hipkin applies a correction including the geoid height N separate from the summation over geopotential coefficients, and suggests that the required geoid height can be evaluated from geopotential coefficients beforehand. The term $E(\theta)$ in Equation (6.35) mainly contains an explicit expression for the normal derivative of the reference gravity $\partial\gamma/\partial h$. Despite the difference in notation, Equations (6.33) and (6.34) are the same, except for one small difference. In Equation (6.33), the term containing the latitudinal derivative of the SHF \bar{Y}_{nm} is added to the other terms, whereas in Equation (6.34) it is subtracted. This can presumably be attributed to a sign error in Hipkin's (2004) publication.

Equation (6.33) is, in principle, equivalent to the weighted summation derived in Equation (4.84), but it is for this purpose not strictly necessary to bring all dependence on latitude inside the SHFs. Instead, the ellipsoidal radius r_e and ellipsoidal deflection angle ϕ , as well as the reference gravity γ and its normal derivative $\partial\gamma/\partial h$, can be computed for every desired latitude separately. However, Equation (4.84) is computationally more efficient and removes the singularity at the poles in Equation (6.33), which is caused by the appearance of the first latitudinal derivative of the SHF.

In conclusion, the long-wavelength term N_1 reads

$$N_1(\theta, \lambda) = \frac{1}{\gamma} \sum_{n=0}^{n_{\max}} \left(\frac{R}{r_e}\right)^{n+1} \sum_{m=-n}^n \bar{T}_{nm}^R \bar{Y}_{nm}(\theta, \lambda) \quad (6.36)$$

and the short-wavelength term N_2 follows from an integration

$$N_2(\theta, \lambda) = \frac{R}{4\pi\gamma} \int_{\sigma'} S(\psi) [\Delta g(\theta', \lambda') - \Delta g_{n_{\max}}(\theta', \lambda')] d\sigma' \quad (6.37)$$

which can in principle be confined to a spherical cap around the computation point due to the absence of the long wavelengths in the gravity anomalies. To reduce long-wavelength information entering in the computation of N_2 , the spherical radius of the

cap ψ_0 can be chosen according to the rule of thumb $\psi_0 \approx \pi/n_{\max}$. However, the restriction of the integration limits results in the omission of far-zone short-wavelength gravity data, causing a truncation error (e.g., Rummel and Rapp, 1976). Sjöberg (2005) shows that this truncation error can reach 3 cm for a spherical cap with a radius of 2 degrees, which according to the rule of thumb corresponds to $n_{\max} = 90$.

Alternatively, or in addition, to the removal of the low-frequency part of gravity anomalies, the spherical Stokes kernel can be modified, primarily to reduce the truncation error (e.g., Wong and Gore, 1969; Vaníček and Sjöberg, 1991; Sjöberg, 1991; Featherstone et al., 1998). This approach has, for example, been applied in the computation of geoid models for the British Isles (Featherstone and Olliver, 1994), Canada (Vaníček et al., 1995), and Australia (Featherstone et al., 2001, 2004). The ellipsoidal correction to one of the modified Stokes kernels has been studied by Sjöberg (2003c, 2004b). However, the simple unmodified RCR approach has also been applied in numerous geoid computations, such as, for example, in the computation of geoid models for Canada (Sideris and She, 1995) and the United States (Smith and Roman, 2001). Comparisons of the unmodified and the modified RCR approach (Vaníček and Sjöberg, 1991) are provided in Vaníček and Featherstone (1998), Featherstone et al. (2004) and Sjöberg (2005).

The ellipsoidal corrections in the unmodified RCR approach were investigated by Ardestani and Martinec (2003b) for the ellipsoidal harmonics method of Martinec and Grafarend (1997a), and are here shown for the surface spherical harmonics method derived in Section 6.3. The effect of kernel modification on the ellipsoidal corrections using the spherical harmonics method is a possible topic for further research.

6.4.2 High-frequency ellipsoidal corrections

In the computation of ellipsoidal corrections in view of the RCR approach, it is usually assumed that the geopotential coefficients in the global model do not contain any spherical approximation. Recent satellite-only geopotential models that provide highly accurate low-degree coefficients, such as GGM02 (Tapley et al., 2005) or EIGEN-02S

(Reigber et al., 2005), can be used to compute the term N_1 in Equation (6.31). According to, e.g., Rummel et al. (2002), Koch (2005) and Mayer-Gürr et al. (2005), these models provide accurate coefficients up to approximately degree 30 – 60. Since satellite-only models are not computed from data on an ellipsoid, additional ellipsoidal corrections to N_1 are not necessary. Note, however, that geoid heights computed from other recent geopotential models that also include terrestrial gravity data, such as EGM96 (Lemoine et al., 1998), do not require any ellipsoidal corrections either, because in their computation the ellipticity of the Earth is already taken into consideration. EGM96 can therefore also be used to compute N_1 without the need for ellipsoidal corrections to this term.

The term N_2 in Equation (6.31) still requires attention. This term is computed from local gravity data from which the long-wavelength effect obtained from Equation (6.33) is subtracted. Although a geopotential model does generally not provide as much detail about the high-frequency part of the Earth’s gravity field as the local gravity data, it can still provide an accurate estimate of the magnitude of the ellipsoidal corrections at these high frequencies. The ellipsoidal correction to term N_2 can be obtained from a spherical harmonic series that starts at degree n_{\max}

$$\delta N_2 = \frac{1}{\gamma} \sum_{n=n_{\max}}^{N_{\max}} \sum_{m=-n}^n \overline{\delta T}_{nm}^{r_e} \overline{Y}_{nm} \quad (6.38)$$

where N_{\max} is the maximum degree and order of the geopotential model.

The absolute error in δN_2 due to truncation at degree N_{\max} will be very small independent of the choice of n_{\max} , since the ellipsoidal corrections are predominantly of long-wavelength nature, especially if the radius of the reference sphere R is chosen equal to the ellipsoidal radius r_e (see Section 6.3.4). The degree variance spectra of the ellipsoidal corrections in Figure 6.8 are roughly parallel to the degree variance spectrum of the disturbing potential itself, and approximately five orders of magnitude lower. Therefore, the absolute truncation error in the computation of δN_2 will be of the order of five times smaller than the truncation error in the geoid height. Based on the investigation into the spectral properties of the geoid height by Schwarz (1985), which shows that the contribution of SHCs beyond degree 360 is less than 0.1%, the

truncation error in δN_2 can be approximated to be well below 1 mm for any choice of reference sphere R .

6.4.3 Magnitude of ellipsoidal corrections in various bandwidths

The magnitude of ellipsoidal corrections in the RCR approach are investigated for the region of Australia, which was chosen as the test area with a view to the determination of a new AUSGeoid model (cf. Featherstone et al., 2001). In any geoid computation, the optimal cut-off degree of the global geopotential model n_{\max} will depend on the quality of the global geopotential model, and also on the quality of the local gravity data, which means that the optimal choice of n_{\max} is different for every region. However, most people choose $n_{\max} = N_{\max}$, the appropriateness of which will not be discussed here.

The choice of n_{\max} has a great impact on the magnitude of the ellipsoidal correction term δN_2 . As an example, computations of the term δN_2 based on a choice of $n_{\max} = 90$ and $N_{\max} = 340$ are shown in Figures 6.9 to 6.11 for $R = a$, $R = (a^2b)^{1/3}$ and $R = r_e$, respectively, based on the EGM96 global geopotential model. Global statistics based on this choice of parameters are provided in Table 6.2. The parameter N_{\max} was chosen slightly smaller than 360 (the maximum degree of EGM96), to avoid errors in the computation of the ellipsoidal corrections in the vicinity of the maximum degree. The contribution of SHCs beyond degree 340 is negligible, judging from the degree variance spectra in Figure 6.8.

As expected, it can be seen from Figures 6.9 and 6.10, respectively, that the largest corrections appear in the south (at the largest distance from the equator) for $R = a$, and in the north (at the largest distance from the mid-latitudes) for $R = (a^2b)^{1/3}$. The magnitude of the corrections reaches a maximum of approximately 5 mm in both cases. However, for $R = r_e$, the magnitude of the corrections remains well below 0.5 mm over the whole region, as can be seen in Figure 6.11. This size of correction is close to the level of the numerical errors, which cause the pattern of longitudinal stripes, specifically in the south (see Figure 6.11).

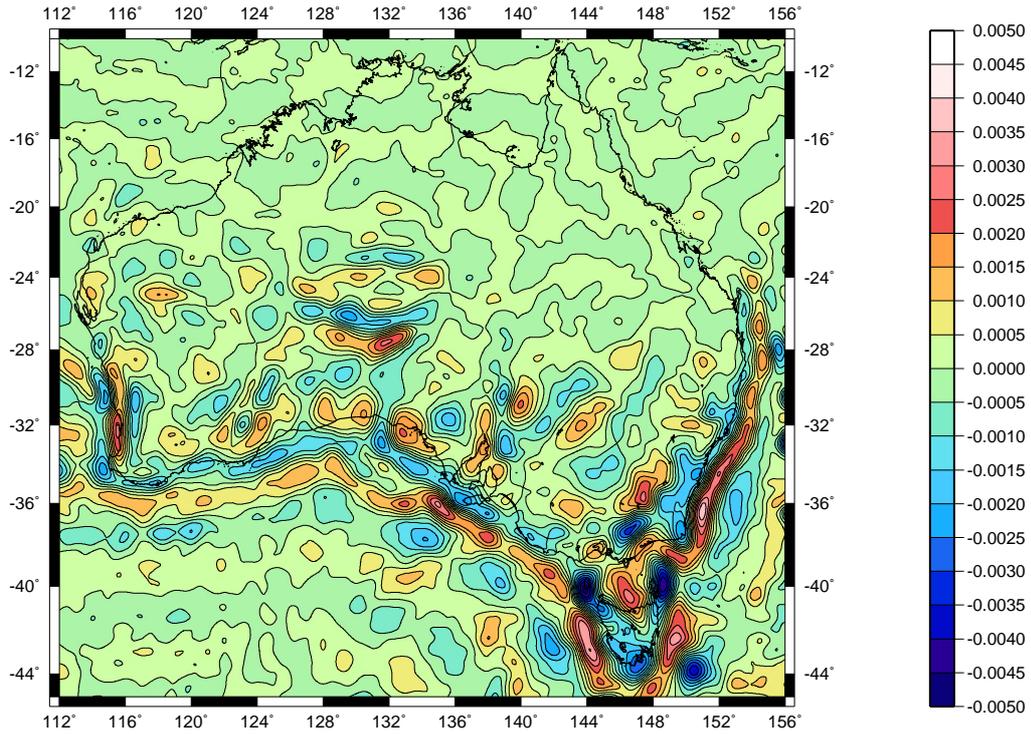


Figure 6.9: Ellipsoidal correction term δN_2 (Equation 6.38) from the EGM96 geopotential model with degrees $90 \leq n \leq 340$ and a reference sphere with radius $R = a$ (units in m; Mercator projection)

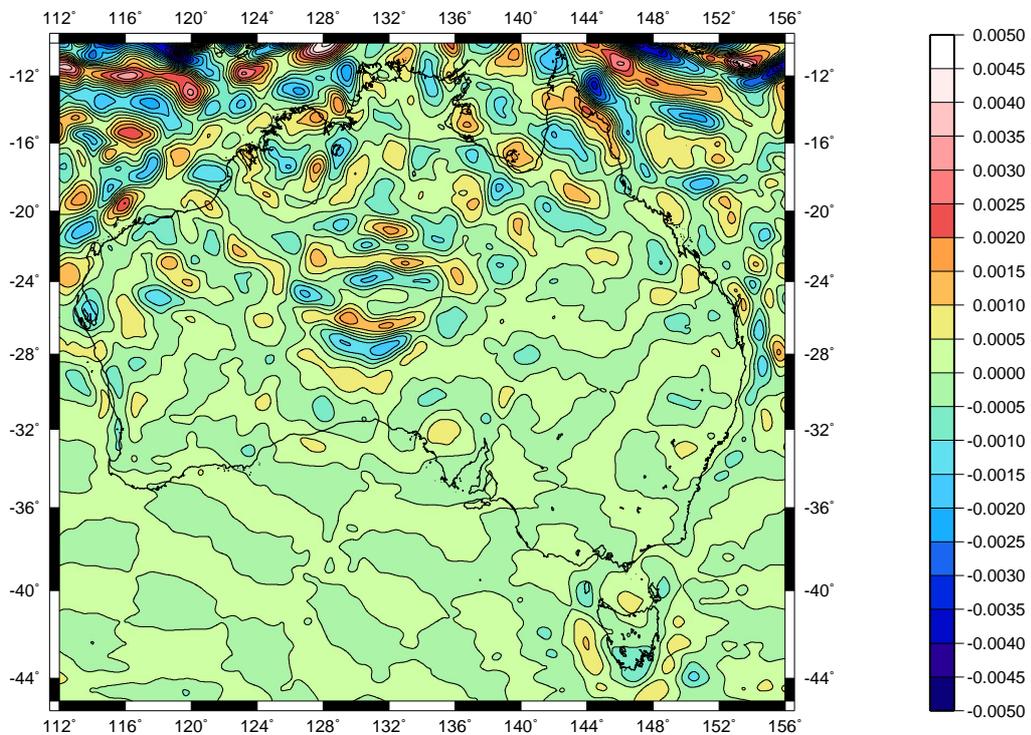


Figure 6.10: Ellipsoidal correction term δN_2 (Equation 6.38) from the EGM96 geopotential model with degrees $90 \leq n \leq 340$ and a reference sphere with radius $R = (a^2b)^{1/3}$ (units in m; Mercator projection)

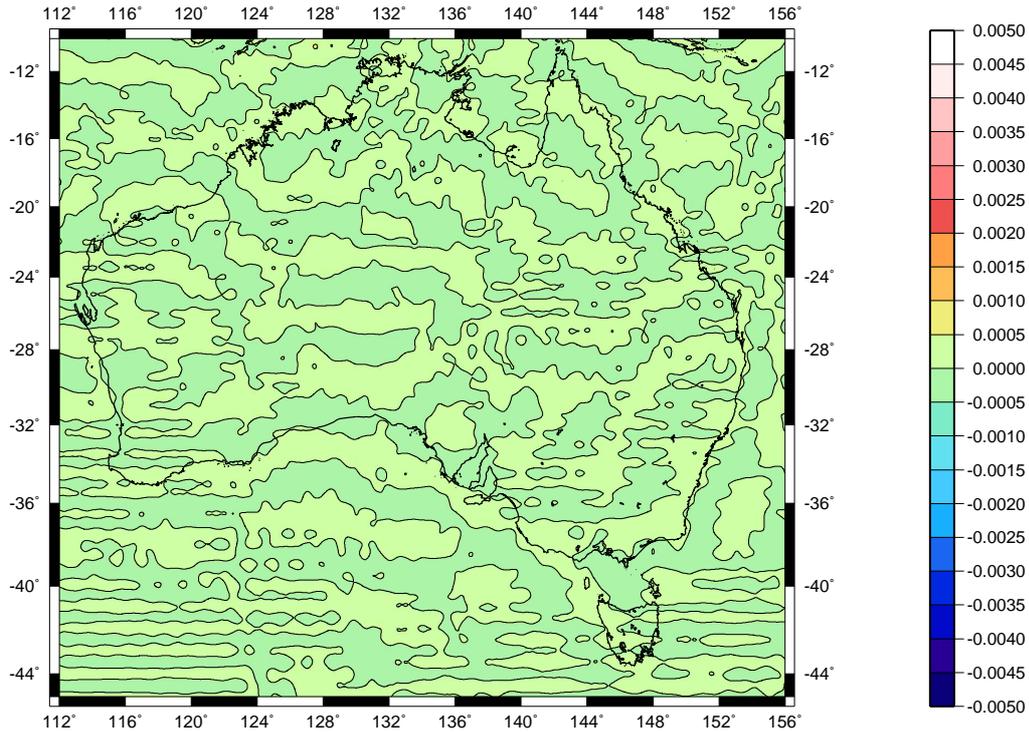


Figure 6.11: Ellipsoidal correction term δN_2 (Equation 6.38) from the EGM96 geopotential model with degrees $90 \leq n \leq 340$ and a reference sphere with radius $R = r_e$ (units in m; Mercator projection)

This leads to the conclusion that truncation at degree $n = 90$ will only lead to sub-millimetre errors. In fact, this does not only hold for Australia, but for the entire Earth. The global extremes of close to 1 cm that can be seen in Table 6.2 occur at the poles and are caused by the numerical errors in the coefficients of even degree, but without these numerical errors, the ellipsoidal corrections are below 1 mm for even the most mountainous regions.

Therefore, evaluation of a global geopotential model using Equations (3.81), (4.86) and (6.22) to (6.24) up to degree 90 provides very accurate ellipsoidal corrections. In addition, when a global geopotential model is used to account for the effect up to degree $n = 90$, the ellipsoidal correction to δN_2 can be ignored. This may even be the case for $n < 90$. Figure 6.12 shows the ellipsoidal correction δN_2 for $R = r_e$ and $n_{\max} = 20$. Also in this case, the corrections are very small, with an absolute maximum of 2 mm in Australia. The strong correlation between the ellipsoidal corrections and the geoid heights can be observed from comparison of Figures 6.12 and 6.13.

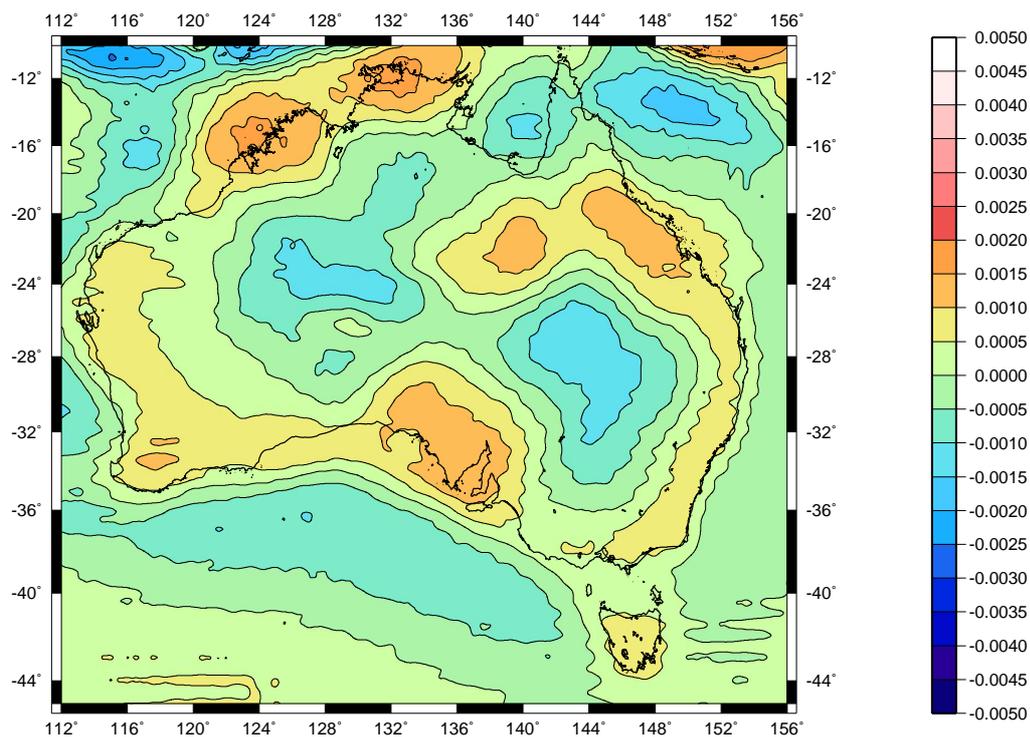


Figure 6.12: Ellipsoidal correction term δN_2 (Equation 6.38) from the EGM96 geopotential model with degrees $20 \leq n \leq 340$ and a reference sphere with radius $R = r_e$ (units in m; Mercator projection)

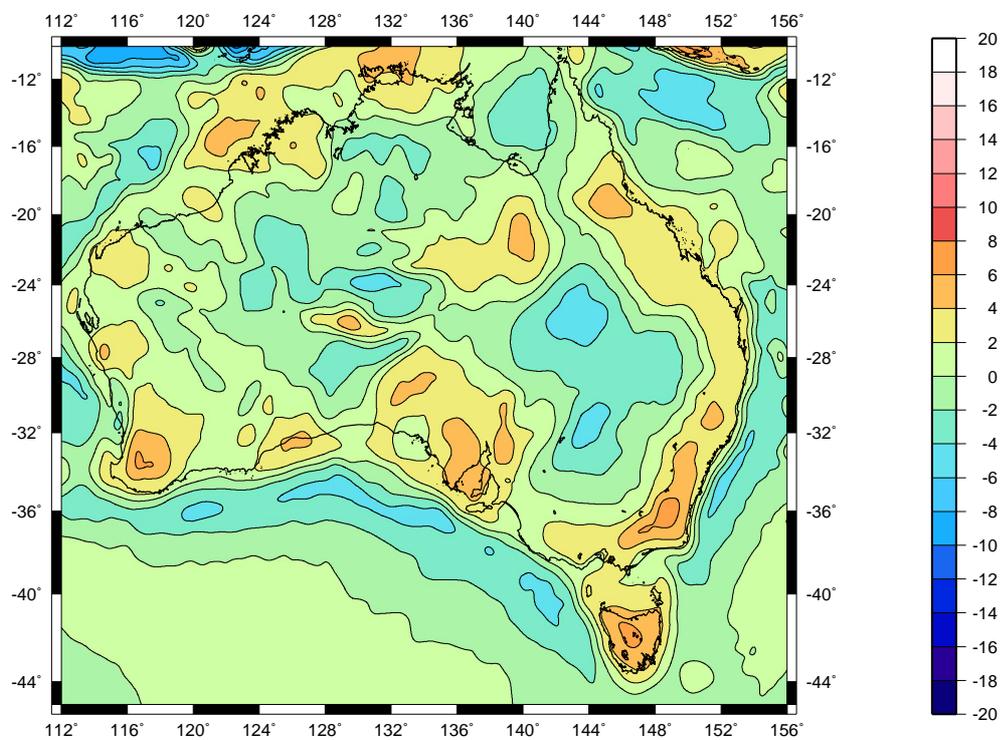


Figure 6.13: Geoid computed from the EGM96 geopotential model with degrees $20 \leq n \leq 340$ (units in m; Mercator projection)

	$R = b$	$R = a$	$R = (a^2b)^{1/3}$	$R = r_e$
minimum	-35.8	-44.4	-29.5	-9.2
maximum	55.2	39.7	25.6	9.9
mean	0.007	0.024	0.018	0.007
absolute mean	1.34	1.54	1.08	0.32

Table 6.2: Global statistics of the ellipsoidal corrections for a reference sphere with radius R equal to semi-minor axis b , ‘mean’ radius $(a^2b)^{1/3}$, semi-major axis a and ellipsoidal radius r_e , obtained from geopotential model EGM96 and degrees $90 \leq n \leq 340$ synthesised on a 30’ grid (units in mm)

	$R = b$	$R = a$	$R = (a^2b)^{1/3}$	$R = r_e$
minimum	-83.0	-64.4	-41.5	-14.2
maximum	85.4	49.9	33.2	11.1
mean	-0.035	0.195	0.118	0.006
absolute mean	3.72	3.37	2.53	0.62

Table 6.3: Global statistics of the ellipsoidal corrections for a reference sphere with radius R equal to semi-minor axis b , ‘mean’ radius $(a^2b)^{1/3}$, semi-major axis a and ellipsoidal radius r_e , obtained from geopotential model EGM96 and degrees $20 \leq n \leq 340$ synthesised on a 30’ grid (units in mm)

The global statistics of δN_2 for $n_{\max} = 20$ are shown in Table 6.3. The ellipsoidal correction beyond degree 20 amounts to less than 1 cm in almost every place on Earth, with an absolute maximum of 1.4 cm. The statement by Huang et al. (2003) that “the first 20 degrees of the geopotential harmonic series contribute approximately 90% of the ellipsoidal correction” only holds for reference spheres of constant radius globally. For $R = r_e$, the first 20 degrees of the geopotential harmonic series contribute approximately 98% of the ellipsoidal correction.

Since practically every modern geoid computation uses a global geopotential model at least up to degree $n = 20$, the errors introduced by the spherical approximation to δN_2 will be sub-centimetre, which is negligible in most cases, provided that: 1) the reference

sphere R in Equation (6.9) is set to the ellipsoidal radius r_e for each computation point, and 2) the long-wavelength contribution to the gravity anomaly is computed via the rigorous ellipsoidal formula in Equation (6.33).

6.5 Summary

In this Chapter, new formulas for ellipsoidal corrections to regional geoid heights were derived based on the spectral relationships found in Chapters 3 and 4. These formulas express the ellipsoidal corrections as a function of global geopotential coefficients, which is more efficient than evaluation of the ellipsoidal Stokes's integral formula. The formulas derived here are not restricted to the order of the square of the first numerical eccentricity of the ellipsoid e^2 , unlike all of the existing methods. The numerical results show a very close agreement with the ellipsoidal corrections computed by Heck and Seitz (2003), which was shown to be within 1 cm accuracy with the synthetic test of Huang et al. (2003).

An important aspect of ellipsoidal corrections to geoid heights is the dependence of the solution on the reference sphere used in the spherical approximation, which appears largely unrecognised in the literature on the subject. It can be concluded that the dominance of the short-wavelength part of the ellipsoidal corrections can be attenuated significantly if the radius of the reference sphere in the spherical approximation is set equal to the ellipsoidal radius of the computation point ($R = r_e$). This reduces the impact of spherical harmonic degrees beyond 20 on the ellipsoidal corrections from $\sim 10\%$ to $\sim 2\%$, contributing a maximum of 1 cm to the geoid height only. This means that the ellipsoidal corrections can be modelled more easily and efficiently, and more importantly that the ellipsoidal corrections to the Stokes integration can be ignored if an RCR procedure to geoid computation is applied, which is often the case in practice.

7. SUMMARY, CONCLUSIONS AND OUTLOOK

In this thesis, new solutions to various ellipsoidal boundary-value problems (BVPs) have been derived, and their application in gravity field modelling and geoid determination was investigated. Section 7.1 contains a short summary of the approach and derived solutions. The main conclusions resulting from this research are summarised in Section 7.2, and based on these conclusions, an outlook to possible future research subjects is provided in Section 7.3.

7.1 Summary

The main objective of this research was to derive highly accurate solutions to ellipsoidal BVPs in the framework of spherical harmonics for application in gravity field modelling. Based on new relations among spherical harmonic functions (SHFs) derived in Section 3.2, a solution to the general Dirichlet BVP was formulated (Section 3.3). Along similar lines, solutions to the Neumann BVP and second-order BVP were derived in Chapter 4, and this has led to the formulation of an ellipsoidal Meissl scheme that connects surface spherical harmonic coefficients (SHCs) of the disturbing potential and its first- and second-order derivatives on ellipsoids of various shapes and sizes.

Based on these derivations, solutions to the ellipsoidal second and third geodetic BVPs were also derived. The application of these solutions in two areas – the computation of solid SHCs of the Earth’s gravity field, and ellipsoidal gravimetric geoid computation – was investigated in Chapters 5 and 6 respectively. The necessity for high theoretical accuracy in the computation of solid SHCs of the disturbing potential, avoiding approximations to the rigorous fundamental equation of physical geodesy (Equation 2.52), was proven. In addition, new insights into the nature of ellipsoidal corrections to gravimetric geoid computation were obtained.

7.2 Conclusions

The most important conclusions that can be drawn from the work presented in this thesis are itemised below, and a short explanation is appended to each conclusion.

1. *The product of associated Legendre functions (ALFs) or spherical harmonic functions (SHFs) with an arbitrary power of the sine or cosine of co-latitude θ can be expressed as a weighted summation over ALFs or SHFs of the same order m (Equations 3.28 and 3.31).*

The weight functions in these expressions are called sinusoidal or cosinusoidal Legendre weight functions (LWFs), respectively. They can be computed from four different iterative algorithms (Equations 3.33, 3.38, 3.39 and 3.40) and the numerical accuracy of the relations is better than 10^{-12} for degree n and order m up to 360, and for $\sin \theta$ or $\cos \theta$ up to at least a power of 16 (with $\theta = 0^\circ, 30^\circ, 45^\circ, 60^\circ$ and 90°). Similar relations for the first- and second-order derivatives of ALFs and SHFs with respect to θ were also derived (Equations 3.54 and 3.61).

2. *The new relations among SHFs allow for a transformation between solid SHCs and surface SHCs of a function defined on an ellipsoid.*

A harmonic function can be expressed as a solid spherical harmonic expansion, which can be used to synthesise function values on the surface of an ellipsoid. Moreover, function values on an ellipsoid can be expressed as a surface spherical harmonic expansion. The relation between the solid SHCs and the surface SHCs from the ellipsoid was derived (Equation 3.73), based on new relations among SHFs (Equation 3.31). It expresses the surface SHCs as a weighted summation over solid SHCs of equal order m , where the weights only depend on degree n , order m , and the shape and size of the ellipsoid used.

3. *Theoretically exact solutions to many ellipsoidal BVPs can be found in the spherical harmonic framework using a transformation from surface SHCs to solid SHCs.*

The transformation of surface SHCs from an ellipsoid into solid SHCs of the harmonic function under consideration provides a solution to an ellipsoidal BVP, since solid SHCs represent the harmonic function anywhere in space where harmonicity holds. Such a coefficient transformation was derived for the case that the function values or its first- or second-order derivatives are given on the surface of the ellipsoid (the Dirichlet, Neumann and second-order BVPs; Equations 3.73, 4.33 and 4.57). In addition, it was shown that such a transformation can also be derived for the ellipsoidal second and third geodetic BVPs (Equations 4.33 and 4.86).

4. *Solutions to ellipsoidal BVPs from a transformation between surface and solid SHCs yield accurate results (with average relative error in degree variances of 10^{-12}) efficiently up to spherical harmonic degree and order ~ 520 for an ellipsoid with the Earth's eccentricity.*

All solutions to ellipsoidal BVPs that use a transformation from surface to solid SHCs require the inversion of a linear system of equations, where the size of the system depends on the maximum degree of the spherical harmonic expansions used. This inversion can be performed efficiently via a Jacobi iteration (or a related iterative procedure), but convergence of such an iterative procedure can only be guaranteed as long as the design matrix is diagonally dominant. It was shown in Section 3.3.2 that for an ellipsoid with the eccentricity of the Earth ($e^2 = 0.0067$), diagonal dominance only occurs for spherical harmonic expansions up to degree $n \approx 520$. This value was found to be almost identical for the Dirichlet (Section 3.3.2), Neumann (Section 4.1.3), and free-geodetic BVPs (Section 4.3.2), due to the fact that non-diagonal dominance at the higher degrees is primarily caused by the continuation of data from the ellipsoid to a sphere, and is thus mainly independent of the boundary condition.

5. *Using spectral relations between surface SHCs of the disturbing potential and its first- and second-order derivatives on an ellipsoid, an ellipsoidal Meissl scheme can be constructed.*

Based on the relations between surface and solid SHCs, an ellipsoidal Meissl scheme (Figure 4.4) was constructed, as an extension of the spherical Meissl scheme (Rummel and Van Gelderen, 1992). In this spherical scheme, all surface SHCs are connected by one-to-one relations, but this does not hold for functions on the ellipsoid, which makes the ellipsoidal Meissl scheme more complicated than the spherical one. Instead, the relations are presented as the product of a block-diagonal matrix with a vector containing surface SHCs of one quantity, yielding a vector containing surface SHCs of a second quantity. The transformation from a quantity on an ellipsoid to a quantity on another surface requires the inversion of a matrix that becomes non-diagonally dominant for spherical harmonic expansions beyond degree ~ 520 , but diagonal dominance is assured for transformations between different quantities on the same ellipsoid (see Section 4.4.2).

6. *Solid SHCs of the disturbing potential can be computed from gravity anomalies on the ellipsoid, using a weighted summation over spherically approximated solid SHCs, by three different methods: the upward-continuation method, the ellipsoidal integration method and the coefficient transformation method.*

The upward-continuation and ellipsoidal integration methods that were introduced by Cruz (1985) and Sjöberg (2003c), respectively, form a simpler and more efficient alternative to the ellipsoidal harmonics method by Jekeli (1988). It was shown in Sections 5.2 and 5.3 that these two methods can be derived more rigorously, avoiding approximations to the order of e^2 that hinder the accuracy of existing methods. In addition, a third and novel approach was introduced in Section 5.4: the coefficient transformation method. This third approach relies entirely on the spectral relationship between solid SHCs of the disturbing potential and surface SHCs of gravity anomalies on the ellipsoid, and is also void of approximations to the order of e^2 .

7. *The coefficient transformation and ellipsoidal integration methods for computation of solid SHCs of the disturbing potential significantly improve on the accuracy of existing weighted summation methods.*

A numerical comparison of the ellipsoidal integration method and the coefficient transformation method with several existing methods has shown the superiority of the newly derived solutions over other weighted summation methods (see Figure 5.5). The accuracy of the coefficient transformation method was the highest, leading to a mean error in geoid height of 0.3 mm and a maximum of 2.6 mm (for $20 \leq n \leq 340$). This is slightly worse than Jekeli's (1988) ellipsoidal harmonics method, which is primarily caused by numerical errors in the coefficients of order $|m|$ close to degree n , where the relative accuracy is in the order of 10^{-4} . For low orders $|m|$, the accuracy of the coefficient transformation method is in the order of 10^{-8} , which is slightly better than that of Jekeli's method (see Figure 5.12). This means that optimal results can be obtained via a combination of the two methods. Both methods show no degradation of accuracy with increasing degree up to $n = 340$.

8. *Solid SHCs of very-high-degree and order ($n > 720$) computed from a weighted summation method contain numerical inaccuracies that increase with increasing degree n .*

The upward-continuation and coefficient transformation methods (Sections 5.2 and 5.4) rely on the solution of a linear system of equations via a Jacobi iteration (or a related iterative procedure), for which convergence can only be guaranteed up to degree and order ~ 520 . This causes the accuracy of the solution to deteriorate with increasing degree n . The alternative rigorous matrix inversion is very time-consuming and therefore inefficient. The ellipsoidal integration method (Section 5.3) does not require an iterative procedure, but the computation of the weights suffers from numerical inaccuracies that become especially large above degree $n \approx 1000$.

9. *Ellipsoidal corrections to a Stokesian geoid computation can be evaluated from surface SHCs of the disturbing potential and of gravity anomalies, which can both be obtained from a global geopotential model using a coefficient transformation.*

Many methodologies exist to compute ellipsoidal corrections to geoid heights via a global integration over gravity anomalies, using a generalisation of the spherical Stokes kernel and/or a priori corrections to the gravity anomalies (e.g., Martinec and Grafarend, 1997a; Fei and Sideris, 2000). However, due to the dominance of the long-wavelengths, ellipsoidal corrections can more efficiently be evaluated with the use of a global geopotential model. Heck and Seitz (2003) and Sjöberg (2003c) provide formulas to do so using approximations of the order of e^2 , but it is shown in Equation (6.21) that the ellipsoidal corrections can also be evaluated from surface SHCs of the disturbing potential and of the gravity anomalies, which can be computed via a coefficient transformation (Equations 3.73 and 4.86). This formulation contains no approximations other than those already embedded in the rigorous fundamental equation of physical geodesy (Equation 2.52).

10. *The magnitude of ellipsoidal corrections to geoid heights strongly depends on the reference sphere chosen in the spherical approximation.*

The geoid heights resulting from a spherical approximation depend on the choice of the radius of the reference sphere, and the magnitude of the ellipsoidal corrections in turn depend on this. It was shown theoretically and numerically, that the influence of the choice of the reference sphere radius on the magnitude of the ellipsoidal corrections is as large as the magnitude of the corrections themselves (see Figures 6.3 to 6.6). The ellipsoidal corrections are smallest, in a least-squares sense, for a reference sphere radius that is larger than the semi-major axis of the ellipsoid.

11. *When the reference sphere radius is chosen equal to the ellipsoidal radius for each computation point ($R = r_e$), the short wavelengths of the ellipsoidal corrections are strongly attenuated, reducing the impact of degrees $n > 20$ from 10% to 2% (contributing a maximum of 1 cm to the geoid height).*

In the literature, ellipsoidal corrections are almost exclusively computed using a constant reference sphere radius (e.g., Huang et al., 2003; Ellmann, 2005), even though the reference sphere radius can be chosen independently for each computation point. It was shown that the choice of $R = r_e$ makes the dominance of the long-wavelengths of the ellipsoidal corrections even more profound (see Figure 6.8), which means that they can more easily and more accurately be modelled from a global geopotential model. In modern geoid computations, a remove-compute-restore (RCR) approach is widely applied, whereby the long-wavelengths of the geoid signal are provided by a global geopotential model up to at least degree 20. In this case, ellipsoidal corrections are not necessary for centimetre accuracy, provided that the long-wavelength contribution is removed from the gravity anomalies using the rigorous ellipsoidal formula (Equation 6.33).

7.3 Outlook

The new solutions to ellipsoidal BVPs presented in this thesis were proven to provide very high numerical accuracy (relative errors $< 10^{-4}$) for low and medium degree and order ($n < 360$). They can therefore successfully be applied in this range of degrees for either spherical harmonic analysis or computation of ellipsoidal corrections to gravimetric geoid computation. Secondly, the demonstrated importance of the choice of the reference sphere radius in a Stokesian geoid computation can be of use in geoid computations around the world and may avoid the improper or inconsistent use of ellipsoidal corrections.

The main point of concern in the solutions to ellipsoidal BVPs is the decrease in accuracy with increasing degree beyond $n = 360$. Contrary to the majority of existing

solutions to ellipsoidal geodetic BVPs, the cause does not lie in the approximation of the boundary condition to the level of e^2 , since such approximations were not introduced in the derivations here. The inaccuracies are instead caused by numerical inaccuracies and/or divergence of iterative procedures.

In the case of the coefficient transformation method, the problems mainly occur due to divergence of the iterative inversion of a large system of linear equations. A different inversion strategy may yield more accurate solutions, but given the large size of the system of equations, the inversion must be highly efficient. Due to the sparseness of the design matrix (despite its non-diagonal dominance), an efficient and accurate inversion scheme may exist, but the search for and application of an optimal inversion procedure would require further research.

The need for inversion of a linear system of equations can be avoided by the ellipsoidal integration method, which can be used in the computation of solid SHCs of the disturbing potential from gravity anomalies on the ellipsoid. However, also in this method, degradation of accuracy for very high degrees ($n > 720$) was observed, presumably caused by numerical inaccuracies. The exact reason for the occurrence of these inaccuracies is as yet unknown, but if they can be circumvented, the ellipsoidal integration method would be a highly efficient alternative to the ellipsoidal harmonics method.

Independent of the possibilities to improve the numerical performance of the solutions to ellipsoidal BVPs derived in this thesis, the question can be posed whether the spherical harmonic framework is indeed the most suitable for a global gravity field model of very-high-degree and order. The inaccuracies in spherical harmonic analysis encountered here are not the only difficulties that arise with very-high-degree and order spherical harmonic expansions. For example, Holmes (2003) points out that a spherical harmonic synthesis will exhibit increasingly large rounding errors at the poles with increasing degree n .

Exclusive use of ellipsoidal harmonic expansions in gravity field modelling would avoid the numerical difficulties described above. However, this would require that all users of global geopotential models have stable algorithms for very-high-degree and order

ellipsoidal harmonic synthesis at their disposal. Such algorithms are unfortunately not available today, and their development would presumably still require considerable research effort (cf. Sona, 1995). Overall, a shift from spherical to ellipsoidal harmonics as the primary framework for global gravity field models would present a vast challenge to the global geodetic community, but may ultimately be beneficial.

REFERENCES

- Abramowitz, M. and Stegun, I. A. (1972) *Handbook of Mathematical Functions*, Dover Publications, New York, USA, 1006 pp.
- Andersen, O. B., Knudsen, P. and Tscherning, C. C. (1996) Investigation of methods for global gravity field recovery from the dense ERS-1 geodetic mission altimetry, in: *Proceedings of the International Association of Geodesy: Global Gravity Field and Its Temporal Variations*, Rapp, R.H., Cazenave, A.A. and Nerem, R.S. (eds.), IAG Symposia Series, Vol. 116, Springer, Berlin, Germany, pp. 218–226.
- Ardalan, A. A. and Grafarend, E. W. (2001a) Ellipsoidal geoidal undulations (ellipsoidal Bruns formula): case studies, *Journal of Geodesy*, Vol. 75, No. 9-10, pp. 544–552, DOI: 10.1007/s001900100212.
- Ardalan, A. A. and Grafarend, E. W. (2001b) *Spheroidal harmonic analysis to degree/order 1000/1000 in the World Geodetic Datum 2000*, paper presented to the 15th IAG Scientific Assembly, Budapest, Hungary, September.
- Ardalan, A. A. and Grafarend, E. W. (2001c) *Towards a Spheroidal Gravitational Earth Model (SEGM2000), spheroidal harmonic analysis on the international reference ellipsoid*, paper presented to the 26th General Assembly of the European Geophysical Society, Nice, France, March.
- Ardalan, A. A. and Safari, A. (2004) Ellipsoidal terrain correction based on multi-cylindrical equal-area map projection of the reference ellipsoid, *Journal of Geodesy*, Vol. 78, No. 1-2, pp. 114–123, DOI: 10.1007/s00190-004-0381-6.
- Ardestani, V. E. and Martinec, Z. (2000) Geoid determination through ellipsoidal Stokes boundary-value problem, *Studia Geophysica et Geodaetica*, Vol. 44, No. 3, pp. 353–363, DOI: 10.1023/A:1022156303247.
- Ardestani, V. E. and Martinec, Z. (2001) Ellipsoidal Stokes boundary-value problem with ellipsoidal corrections in the boundary condition, *Studia Geophysica et Geodaetica*, Vol. 45, No. 2, pp. 109–126, DOI: 10.1023/A:1021808110012.
- Ardestani, V. E. and Martinec, Z. (2003a) Far zone contribution in ellipsoidal Stokes boundary value problem, *Studia Geophysica et Geodaetica*, Vol. 47, No. 4, pp. 719–723, DOI: 10.1023/A:1026321917446.
- Ardestani, V. E. and Martinec, Z. (2003b) Geoid determination through ellipsoidal Stokes boundary-value problem by splitting its solution to the low-degree and the high-degree parts, *Studia Geophysica et Geodaetica*, Vol. 47, No. 1, pp. 73–82, DOI: 10.1023/A:1022299521920.
- Arnold, K. (1986) Geodetic boundary value problems I, *Report 84*, Zentralinstitut für Physik der Erde, Potsdam, Germany, 221 pp.

- Arnold, K. (1987) Geodetic boundary value problems II, *Report 89*, Zentralinstitut für Physik der Erde, Potsdam, Germany, 229 pp.
- Arnold, K. (1989) Geodetic boundary value problems III, *Report 97*, Zentralinstitut für Physik der Erde, Potsdam, Germany, 196 pp.
- Arnold, K. (1990) Geodetic boundary value problems IV, *Report 114*, Zentralinstitut für Physik der Erde, Potsdam, Germany, 188 pp.
- Biancale, R., Balmino, G., Lemoine, J. M., Marty, J. C., Moynot, B., Barlier, F., Exertier, P., Laurain, O., Gegout, P., Schwintzer, P., Reigber, C., Bode, A., König, R., Massmann, F. H., Raimondo, J. C., Schmidt, R. and Zhu, S. Y. (2000) A new global Earth's gravity field model from satellite orbit perturbations: GRIM5-S1, *Geophysical Research Letters*, Vol. 27, No. 22, pp. 3611–3614, DOI: 10.1029/2000GL011721.
- Bjerhammar, A. (1962) On an explicit solution of the gravimetric boundary value problem for an ellipsoidal surface of reference, *Technical report*, The Royal Institute of Technology, Division of Geodesy, Stockholm, Sweden, 92 pp.
- Bjerhammar, A. (1966) On the determination of the shape of the geoid and the shape of the Earth from an ellipsoidal surface of reference, *Bulletin Géodésique*, Vol. 81, pp. 235–265.
- Bjerhammar, A. and Svensson, L. (1983) On the geodetic boundary value problem for a fixed boundary surface - a satellite approach, *Bulletin Géodésique*, Vol. 57, No. 4, pp. 382–393.
- Bölling, K. and Grafarend, E. W. (2005) Ellipsoidal spectral properties of the Earth's gravitational potential and its first and second derivatives, *Journal of Geodesy*, Vol. 79, No. 6-7, pp. 300–330, DOI: 10.1007/s00190-005-0465-y.
- Bomford, G. (1971) *Geodesy* (3rd Edition), Oxford University Press, Oxford, UK, 731 pp.
- Borkowski, K. M. (1987) Transformation of geocentric to geodetic coordinates without approximations, *Astrophysics and Space Science*, Vol. 139, pp. 1–4.
- Bowring, B. R. (1976) Transformation from spatial to geographical coordinates, *Survey Review*, Vol. 23, No. 181, pp. 323–327.
- Brovar, V. V. (1964). On the solution of Molodensky's boundary value problem, *Bulletin Géodésique*, No. 72, pp. 169-173.
- Brovar, V. V., Kopeikina, Z. S. and Pavlova, M. V. (2001) Solution of the Dirichlet and Stokes exterior boundary problems for the Earth's ellipsoid, *Journal of Geodesy*, Vol. 74, No. 11-12, pp. 767–772, DOI: 10.1007/s001900000100.
- Buchdahl, H. A., Buchdahl, N. P. and Stiles, P. J. (1977) On a relation between spherical and spheroidal harmonics, *Journal of Physics A: Mathematical and General*, Vol. 10, No. 11, pp. 1833–1836, DOI: 10.1088/0305-4470/10/11/011.

- Burša, M. (1995) Geoidal potential free of zero-frequency tidal distortion, *Earth, Moon and Planets*, Vol. 71, No. 1-2, pp. 59–64, DOI: 10.1007/BF00612869.
- Claessens, S. J. (2003) A Synthetic Earth Model, *DUP Science Series*, Delft University Press, Delft, The Netherlands, 61 pp.
- Claessens, S. J. (2005) New relations among associated Legendre functions and spherical harmonics, *Journal of Geodesy*, Vol. 79, No. 6-7, pp. 398–406, DOI: 10.1007/s00190-005-0483-9.
- Claessens, S. J. and Featherstone, W. E. (2005) Computation of geopotential coefficients from gravity anomalies on the ellipsoid, in: *Proceedings of the International Association of Geodesy: A Window on the Future of Geodesy*, Sansò, F. (ed.), IAG Symposia Series, Vol. 128, Springer, Berlin, Germany, pp. 459–464.
- Colombo, O. L. (1981) Numerical methods for harmonic analysis on the sphere, *Report 310*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 140 pp.
- Colombo, O. L. (1982) Convergence of the external expansion of the gravity field inside the bounding sphere, *manuscripta geodaetica*, Vol. 7, No. 3, pp. 209–246.
- Cruz, J. Y. (1985) Disturbance vector in space from surface gravity anomalies using complementary models, *Report 366*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 142 pp.
- Cruz, J. Y. (1986) Ellipsoidal corrections to potential coefficients obtained from gravity anomaly data on the ellipsoid, *Report 371*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 33 pp.
- De Min, E. (1995) Using Rapp's global geopotential models OSU89 and OSU91 in local geoid combination solutions, in: *Proceedings of the European Geophysical Society: Latest Developments in the Computation of Regional Geoids*, Vermeer, M. (ed.), Report No. 95:7, Finnish Geodetic Institute, Masala, Finland, pp. 53–59.
- Dechambre, D. and Scheeres, D. J. (2002) Transformation of spherical harmonic coefficients to ellipsoidal harmonic coefficients, *Astronomy and Astrophysics*, Vol. 387, No. 3, pp. 1114–1122, DOI: 10.1051/0004-6361:20020466.
- Dragomir, V. C., Ghitău, D. N., Mihăilescu, M. S. and Rotaru, M. G. (1982) *Theory of the Earth's Shape*, Elsevier, Amsterdam, The Netherlands, 694 pp.
- Drinkwater, M. R., Floberghagen, R., Haagmans, R., Muzi, D. and Popescu, A. (2003) GOCE: ESA's first Earth explorer core mission, in: *Earth Gravity Field From Space - from Sensors to Earth Sciences*, Beutler, G. B., Drinkwater, M. R., Rummel, R. and Von Steiger, R. (eds.), Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 419–432.
- Ellmann, A. (2005) A numerical comparison of different ellipsoidal corrections to Stokes' formula, in: *Proceedings of the International Association of Geodesy: A Window on the Future of Geodesy*, Sansò, F. (ed.), IAG Symposia Series, Vol. 128, Springer, Berlin, Germany, pp. 409–414.

- Ellmann, A. and Sjöberg, L. E. (2004) Ellipsoidal correction for the modified Stokes formula, *Bollettino di Geodesia e Scienze Affini*, Vol. 63, No. 3, pp. 153–172.
- Faller, J. E. (2002) Thirty years of progress in absolute gravimetry: a scientific capability implemented by technological advances, *Metrologia*, Vol. 39, No. 5, pp. 425–428, DOI:10.1088/0026–1394/39/5/3.
- Fan, H. (1989) Geoid determination by global geopotential models and integral formulas, *Technical report*, The Royal Institute of Technology, Division of Geodesy, Stockholm, Sweden.
- Featherstone, W. E. and Olliver, J. G. (1994) A new gravimetric determination of the geoid of the British Isles, *Survey Review*, Vol. 32, No. 254, pp. 464–478.
- Featherstone, W. E., Evans, J. D. and Olliver, J. G. (1998) A Meissl-modified Vaníček and Kleusberg kernel to reduce the truncation error in gravimetric geoid computations, *Journal of Geodesy*, Vol. 72, No. 3, pp. 154–160, DOI: 10.1007/s001900050157.
- Featherstone, W. E., Holmes, S. A., Kirby, J. F. and Kuhn, M. (2004) Comparison of remove-compute-restore and University of New Brunswick techniques to geoid determination over Australia, and inclusion of Wiener-type filters in reference field contribution, *Journal of Surveying Engineering*, Vol. 130, No. 1, pp. 40–47, DOI: 10.1061/(ASCE)0733–9453(2004)130:1(40).
- Featherstone, W. E., Kirby, J. F., Kearsley, A. H. W., Gilliland, J. R., Johnston, G. M., Steed, J., Forsberg, R. and Sideris, M. G. (2001) AUSGeoid98 geoid model of Australia: data treatment, computations and comparisons with GPS-levelling data, *Journal of Geodesy*, Vol. 75, No. 5/6, pp. 313–330, DOI: 10.1007/s001900100177.
- Fei, Z. L. (2000) Refinements of Geodetic Boundary Value Problem Solutions, *PhD thesis*, Department of Geomatics Engineering, University of Calgary, Calgary, Canada, 198 pp.
- Fei, Z. L. and Sideris, M. G. (2000) A new method for computing the ellipsoidal correction for Stokes’s formula, *Journal of Geodesy*, Vol. 74, No. 2, pp. 223–231, DOI: 10.1007/s001900050280 (erratum in Vol. 74, No. 9, p. 671, DOI: 10.1007/s001900000131).
- Finn, G., Ardalán, A. and Grafarend, E. (2002) *The modeling of the ellipsoidal gravitational terrain effect. Case study: Ellipsoidal vertical deflections*, paper presented to the 27th General Assembly of the European Geophysical Society, Nice, France, April.
- Fu, L. L. and Cazenave, A. (2001) *Satellite Altimetry and Earth Sciences*, International Geophysics Series, Vol. 69, Academic Press, San Diego, USA, 463 pp.
- Garnier, R. and Barriot, J. P. (2001) Ellipsoidal harmonic expansion of the gravitational potential: Theory and application, *Celestial Mechanics and Dynamical Astronomy*, Vol. 79, No. 4, pp. 235–275, DOI: 10.1023/A:1017555515763.

- Giacaglia, G. E. O. and Burša, M. (1980) Transformations of spherical harmonics and applications to geodesy and satellite theory, *Studia Geophysica et Geodaetica*, Vol. 24, No. 1, pp. 1–11, DOI: 10.1007/BF01628375.
- Gleason, D. M. (1988) Comparing ellipsoidal corrections to the transformation between the geopotential's spherical and ellipsoidal spectrums, *manuscripta geodaetica*, Vol. 13, No. 2, pp. 114–129.
- Gleason, D. M. (1989a) Obtaining minimally aliased geopotential coefficients from discrete data forms, *manuscripta geodaetica*, Vol. 14, No. 3, pp. 149–162.
- Gleason, D. M. (1989b) Some notes on the evaluation of ellipsoidal and spherical harmonic coefficients from surface data, *manuscripta geodaetica*, Vol. 14, No. 2, pp. 110–116.
- Grafarend, E. W. (2001) The spherical horizontal and spherical vertical boundary value problem - vertical deflections and geoidal undulations - the completed Meissl diagram, *Journal of Geodesy*, Vol. 75, No. 7-8, pp. 363–390, DOI: 10.1007/s001900100186.
- Grafarend, E. W. and Ardalan, A. A. (1997) W_0 - an estimate in the Finnish height datum N60, epoch 1993.4, from twenty-five GPS points of the Baltic Sea level project, *Journal of Geodesy*, Vol. 71, No. 11, pp. 673–679, DOI: 10.1007/s001900050134.
- Grafarend, E. W. and Engels, J. (1994) The convergent series expansion of the gravity field of a starshaped body, *manuscripta geodaetica*, Vol. 19, No. 1, pp. 18–30.
- Grafarend, E. W. and Niemeier, W. (1971). The free nonlinear boundary value problem of physical geodesy, *Bulletin Géodésique*, No. 101, pp. 243–262.
- Grafarend, E. W., Ardalan, A. and Sideris, M. G. (1999) The spheroidal fixed-free two-boundary-value problem for geoid determination (the spheroidal Bruns' transform), *Journal of Geodesy*, Vol. 73, No. 10, pp. 513–533, DOI: 10.1007/s001900050263.
- Grafarend, E. W., Heck, B. and Knickmeyer, E. H. (1985) The free versus fixed geodetic boundary value problem for different combinations of geodetic observations, *Bulletin Géodésique*, Vol. 59, No. 1, pp. 11–32.
- Grafarend, E. W., Krarup, T. and Syffus, R. (1996) An algorithm for the inverse of a multivariate homogeneous polynomial of degree n , *Journal of Geodesy*, Vol. 70, No. 5, pp. 276–286, DOI: 10.1007/BF00867348.
- Grebenitcharsky, R. and Sideris, M. G. (2005) The compatibility conditions in altimetry-gravimetry boundary value problems, *Journal of Geodesy*, Vol. 78, No. 10, pp. 626–636, DOI: 10.1007/s00190-004-0429-7.
- Gross, R. S. (2001) A combined length-of-day series spanning 1832-1997: LUNAR97, *Physics of the Earth and Planetary Interiors*, Vol. 123, No. 1, pp. 65–76.

- Groten, E. (1979) *Geodesy and the Earth's Gravity Field*, Dümmler Verlag, Bonn, Germany, 723 pp.
- Gruber, T. (2001) High-resolution gravity field modeling with full variance-covariance matrices, *Journal of Geodesy*, Vol. 75, No. 9-10, pp. 505–514, DOI: 10.1007/s001900100202.
- Heck, B. (1989) A contribution to the scalar free boundary value problem of physical geodesy, *manuscripta geodaetica*, Vol. 14, No. 2, pp. 87–99.
- Heck, B. (1991) On the linearized boundary value problems of physical geodesy, *Report 407*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 63 pp.
- Heck, B. and Seitz, K. (1995) Non-linear effects in the geodetic version of the free GBVP based on higher order reference fields, in: *Proceedings of the International Association of Geodesy: Geodetic Theory Today*, Sansò, F. (ed.), IAG Symposia Series, Vol. 114, Springer, Berlin, Germany, pp. 332–339.
- Heck, B. and Seitz, K. (2003) Solutions of the linearized geodetic boundary value problem for an ellipsoidal boundary to order e^3 , *Journal of Geodesy*, Vol. 77, No. 3-4, pp. 182–192, DOI: 10.1007/s00190-002-0309-y.
- Heiskanen, W. and Moritz, H. (1967) *Physical Geodesy*, W.H. Freeman & Co., San Francisco, USA, 364 pp.
- Hipkin, R. G. (2004) Ellipsoidal geoid computation, *Journal of Geodesy*, Vol. 78, No. 3, pp. 167–179, DOI: 10.1007/s00190-004-0389-y.
- Hobson, E. W. (1931) *The theory of spherical and ellipsoidal harmonics*, Cambridge University Press, Cambridge, UK, 500 pp.
- Holmes, S. A. (2003) High Degree Spherical Harmonic Synthesis for Simulated Earth Gravity Modelling, *PhD thesis*, Department of Spatial Sciences, Curtin University of Technology, Perth, Australia, 176 pp.
- Holmes, S. A. (2004). *Personal Communication*, Raytheon ITSS, Lanham, Maryland, USA.
- 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 normalised associated Legendre functions, *Journal of Geodesy*, Vol. 76, No. 5, pp. 279–299, DOI: 10.1007/s00190-002-0216-2.
- Hotine, M. (1969) *Mathematical Geodesy*, *ESSA Monograph 2*, U.S. Department of Commerce, Washington, USA, 416 pp.
- Huang, J., Véronneau, M. and Pagiatakis, S. D. (2003) On the ellipsoidal correction to the spherical Stokes solution of the gravimetric geoid, *Journal of Geodesy*, Vol. 77, No. 3-4, pp. 171–181, DOI: 10.1007/s00190-003-0317-6.

- Hwang, C. (1995) A method for computing the coefficients in the product-sum formula of associated Legendre functions, *Journal of Geodesy*, Vol. 70, No. 1-2, pp. 110–116, DOI: 10.1007/BF00863422.
- Hwang, C. and Parsons, B. E. (1996) An optimal procedure for deriving marine gravity from multi-satellite altimetry, *Geophysical Journal International*, Vol. 125, No. 3, pp. 705–718.
- Jekeli, C. (1981) The downward continuation to the Earth's surface of truncated spherical and ellipsoidal harmonic series of the gravity and height anomalies, *Report 323*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 140 pp.
- Jekeli, C. (1982) A numerical study of the divergence of spherical harmonic series of the gravity and height anomalies at the Earth's surface, *Bulletin Géodésique*, Vol. 57, No. 1, pp. 10–28.
- Jekeli, C. (1988) The exact transformation between ellipsoidal and spherical harmonic expansions, *manuscripta geodaetica*, Vol. 13, No. 2, pp. 106–113.
- Jekeli, C. (1996) Spherical harmonic analysis, aliasing and filtering, *Journal of Geodesy*, Vol. 70, No. 4, pp. 214–223, DOI: 10.1007/s001900050010.
- Jekeli, C. (1999) An analysis of vertical deflections derived from high-degree spherical harmonic models, *Journal of Geodesy*, Vol. 73, No. 1, pp. 10–22, DOI: 10.1007/s001900050213.
- Jinghai, Y. and Xiaoping, W. (1997) The solution of mixed boundary value problems with the reference ellipsoid as boundary, *Journal of Geodesy*, Vol. 71, No. 8, pp. 454–460, DOI: 10.1007/s001900050113.
- Jones, G. C. (2002) New solutions for the geodetic coordinate transformation, *Journal of Geodesy*, Vol. 76, No. 8, pp. 437–446, DOI: 10.1007/s00190-002-0267-4.
- Keller, W. (1996) On a scalar fixed altimetry-gravimetry boundary value problem, *Journal of Geodesy*, Vol. 70, No. 8, pp. 459–469, DOI: 10.1007/s001900050034.
- Keller, W. (2002) Geodetic pseudodifferential operators and the Meissl scheme, in: *Geodesy - the Challenge of the Third Millennium*, Grafarend, E.W., Krumm, F.W. and Schwarze, V.S. (eds.), Springer, Berlin, Germany, pp. 207–212.
- Kellogg, O. D. (1929) *Foundations of Potential Theory*, Springer, Berlin, Germany, 384 pp.
- Kenyon, S., Factor, J., Pavlis, N., Saleh, J. and Holmes, S. (2005) *Towards the next Earth Gravitational Model to degree 2160: status and progress*, paper presented to the Dynamic Planet 2005 Conference, Cairns, Australia, August.
- Kern, M., Schwarz, K. P. and Sneeuw, N. (2003) A study on the combination of satellite, airborne, and terrestrial gravity data, *Journal of Geodesy*, Vol. 77, No. 3-4, pp. 217–225, DOI: 10.1007/s00190-003-0313-x.

- Klees, R. (1997) Topics on boundary element methods, in: *Geodetic Boundary Value Problems in View of the One Centimeter Geoid*, Sansò, F. and Rummel, R. (eds.), Springer, Berlin, Germany, pp. 482–531.
- Klees, R., Koop, R., Visser, P. and Van den IJssel, J. (2000) Efficient gravity field recovery from GOCE gravity gradient observations, *Journal of Geodesy*, Vol. 74, No. 7-8, pp. 561–571, DOI: 10.1007/s001900000118.
- Koch, K. R. (1968) Solution of the geodetic boundary value problem in case of a reference ellipsoid, *Report 104*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA.
- Koch, K. R. (1969) Solution of the geodetic boundary value problem for a reference ellipsoid, *Journal of Geophysical Research*, Vol. 74, No. 15, pp. 3796–3803.
- Koch, K. R. (2005) Determining the maximum degree of harmonic coefficients in geopotential models by Monte Carlo methods, *Studia Geophysica et Geodaetica*, Vol. 49, No. 3, pp. 259–275, DOI: 10.1007/s11200-005-0009-1.
- Koop, R. and Stelpstra, D. (1989) On the computation of the gravitational potential and its first and second order derivatives, *manuscripta geodaetica*, Vol. 14, No. 6, pp. 373–382.
- Krarup, T. (1969) A contribution to the mathematical foundation of physical geodesy, *Publication No. 44*, Danish Geodetic Institute, Copenhagen, Denmark.
- Kuhn, M. and Featherstone, W. E. (2003) On the optimal spatial resolution of crustal mass distributions for forward gravity field modelling, in: *Proceedings of the 3rd Meeting of the International Gravity and Geoid Commission: Gravity and Geoid 2002*, Tziavos, I.N. (ed.), Ziti Editions, Thessaloniki, Greece, pp. 189–194.
- Kusche, J. and Klees, R. (2002) Regularization of gravity field estimation from satellite gravity gradients, *Journal of Geodesy*, Vol. 76, No. 6-7, pp. 359–368, DOI: 10.1007/s00190-002-0257-6.
- Ladyzhenskaya, O. A. (1985) *The Boundary Value Problems of Mathematical Physics*, Springer, New York, USA, 322 pp.
- Lambeck, K. (1980) *The Earth's Variable Rotation*, Cambridge University Press, Cambridge, UK, 449 pp.
- Ledin, J. (2001) *Simulation Engineering*, CMP Books, Lawrence, USA, 400 pp.
- Lehmann, R. (1999a) Boundary-value problems in the complex world of geodetic measurements, *Journal of Geodesy*, Vol. 73, No. 9, pp. 491–500.
- Lehmann, R. (1999b) Studies on the altimetry-gravimetry problems, *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy*, Vol. 24, No. 1, pp. 47–52, DOI:10.1016/S1464-1895(98)00009-X.
- Lehmann, R. and Klees, R. (1999) Numerical solution of geodetic boundary value problems using a global reference field, *Journal of Geodesy*, Vol. 73, No. 10, pp. 543–554, DOI: 10.1007/s001900050265.

- Lelgemann, D. (1970) Untersuchungen zu einer genaueren Lösung des Problems von Stokes, *Report C 155*, Veröffentlichungen Deutsche Geodätische Kommission, Potsdam, Germany.
- Lemoine, F. G., Kenyon, S. C., Factor, J. K., Trimmer, R. G., Pavlis, N. K., Chinn, D. S., Cox, C. M., Klosko, S. M., Luthcke, S. B., Torrence, M. H., Wang, Y. M., Williamson, R. G., Pavlis, E. C., Rapp, R. H. and Olson, T. R. (1998) The development of the joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) geopotential model EGM96, *Technical Report NASA/TP-1998-206861*, National Aeronautics and Space Administration, Washington, USA, 575 pp.
- Mackie, A. G. (1965) *Boundary Value Problems*, Oliver and Boyd Ltd., Edinburgh, UK, 249 pp.
- MacMillan, W. D. (1958) *The Theory of the Potential*, Dover Publications, New York, USA, 469 pp.
- MacRobert, T. M. (1967) *Spherical Harmonics* (3rd Edition), Pergamon Press, Oxford, UK, 349 pp.
- Martensen, E. and Ritter, S. (1997) Potential theory, in: *Geodetic Boundary Value Problems in View of the One Centimeter Geoid*, Sansò, F. and Rummel, R. (eds.), Springer, Berlin, Germany, pp. 19–66.
- Martinec, Z. (1998) *Boundary-Value Problems for Gravimetric Determination of a Precise Geoid*, Lecture Notes in Earth Sciences, Vol. 73, Springer, Berlin, Germany, 223 pp.
- Martinec, Z. and Grafarend, E. W. (1997a) Construction of Green's function to the external Dirichlet boundary-value problem for the Laplace equation on an ellipsoid of revolution, *Journal of Geodesy*, Vol. 71, No. 9, pp. 562–570, DOI: 10.1007/s001900050124.
- Martinec, Z. and Grafarend, E. W. (1997b) Solution to the Stokes boundary value problem on an ellipsoid of revolution, *Studia Geophysica et Geodaetica*, Vol. 41, No. 2, pp. 103–129, DOI: 10.1023/A:1023380427166.
- Martinec, Z. and Vaníček, P. (1994) Direct topographical effect of Helmert's second condensation method, *manuscripta geodaetica*, Vol. 19, No. 5, pp. 257–268.
- Masters, G. and Richards-Dinger, K. (1998) On the efficient calculation of ordinary and generalized spherical harmonics, *Geophysical Journal International*, Vol. 135, No. 1, pp. 307–309.
- Mather, R. S. (1973) *A Solution of the Geodetic Boundary Value Problem to Order e^3* , Goddard Space Flight Center, Greenbelt, USA, 127 pp.
- Mayer-Gürr, T., Ilk, K. H., Eicker, A. and Feuchtinger, M. (2005) ITG-CHAMP01: A CHAMP gravity field model from short kinematical arcs of a one-year observation period, *Journal of Geodesy*, Vol. 78, No. 7-8, pp. 462–480, DOI: 10.1007/s00190-004-0413-2.

- Meissl, P. (1971) A study of the covariance functions related to the Earth's disturbing potential, *Report 151*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 81 pp.
- Meissl, P. (1985) The use of finite elements in physical geodesy, *Report 313*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 201 pp.
- Mohr, P. J. and Taylor, B. N. (2005) CODATA recommended values of the fundamental physical constants: 2002, *Reviews of Modern Physics*, Vol. 77, No. 1, pp. 1–107.
- Molodenskii, M. S., Eremeev, V. F. and Yurkina, M. I. (1962) *Methods for study of the external gravity field and figure of the Earth*, Israel Program for Scientific Translations, Jerusalem, Israel, 248 pp.
- Moritz, H. (1980) Geodetic Reference System 1980, *Bulletin Géodésique*, Vol. 54, No. 4, pp. 395–405.
- Moritz, H. (1989) *Advanced Physical Geodesy* (2nd Edition), Wichmann, Karlsruhe, Germany, 500 pp.
- Moritz, H. (1990) *The Figure of the Earth*, Wichmann, Karlsruhe, Germany, 279 pp.
- Nahavandchi, H. and Sjöberg, L. E. (1998) Terrain corrections to power H^3 in gravimetric geoid determination, *Journal of Geodesy*, Vol. 72, No. 3, pp. 124–135, DOI: 10.1007/s001900050154.
- Newton, I. (1729) *The Mathematical Principles of Natural Philosophy*, trans. A. Motte, Middle-Temple-Gate, London, UK, 393 pp.
- Nord, T. (1992) A numerical study of different corrections and transformations to spherical geopotential coefficients recovered on an ellipsoid, *manuscripta geodaetica*, Vol. 17, No. 6, pp. 373–380.
- Novák, P. and Grafarend, E. W. (2005) Ellipsoidal representation of the topographical potential and its vertical gradient, *Journal of Geodesy*, Vol. 78, pp. 691–706, DOI: 10.1007/s00190-005-0435-4.
- Novák, P., Vaníček, P., Véronneau, M., Holmes, S. A. and Featherstone, W. E. (2001) On the accuracy of modified Stokes's integration in high-frequency gravimetric geoid determination, *Journal of Geodesy*, Vol. 74, No. 9, pp. 644–654, DOI: 10.1007/s001900000126.
- Nsombo, P. and Sjöberg, L. E. (1996). Numerical studies on the downward continuation error at the geoid and satellite derived potential models, *Geomatics Research Australasia*, No. 65, pp. 27–41.
- Nutz, H. (2002) A Unified Setup of Gravitational Field Variables, *PhD thesis*, Department of Mathematics, University of Kaiserslautern, Kaiserslautern, Germany, 304 pp.

- Pail, R. (1999) Synthetic Global Gravity Model for Planetary Bodies and Applications in Satellite Gravity Gradiometry, *PhD thesis*, Technical University of Graz, Graz, Austria, 138 pp.
- Paul, M. K. (1973). A note on computation of geodetic coordinates from geocentric (Cartesian) coordinates, *Bulletin Géodésique*, No. 108, pp. 135-139.
- Pavlis, N. K. (1988) Modelling and estimation of a low degree geopotential model from terrestrial gravity data, *Report 386*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 173 pp.
- Pavlis, N. K., Chan, J. C. and Lerch, F. J. (1996) Alternative Estimation Techniques for Global High-degree Gravity Modeling, in: *Proceedings of the International Association of Geodesy: Global Gravity Field and Its Temporal Variations*, Rapp, R.H., Cazenave, A.A. and Nerem, R.S. (eds.), IAG Symposia Series, Vol. 116, Springer, Berlin, Germany, pp. 111–120.
- Pellinen, L. P. (1966) A method for expanding the gravity potential of the Earth in spherical functions, in: *Trans. Central Scientific Res. Inst. Geodesy, Aerial Survey and Cartography*, No. 171, Nedra, Moscow, Russia.
- Pellinen, L. P. (1982) Effects of the Earth's ellipticity on solving geodetic boundary value problem, *Bollettino di Geodesia e Scienze Affini*, Vol. 41, No. 2, pp. 89–103.
- Petrovskaya, M. S. (1995) Solution of the geodetic boundary value problem in spectral form, in: *Proceedings of the International Association of Geodesy: Geodetic Theory Today*, Sansò, F. (ed.), IAG Symposia Series, Vol. 114, Springer, Berlin, Germany, pp. 294–303.
- Petrovskaya, M. S., Vershkov, A. N. and Pavlis, N. K. (2001) New analytical and numerical approaches for geopotential modeling, *Journal of Geodesy*, Vol. 75, No. 12, pp. 661–672, DOI: 10.1007/s001900100215.
- Protter, M. H. and Morrey, C. B. (1964) *Modern Mathematical Analysis*, Addison-Wesley, Reading, USA, 790 pp.
- Pugh, C. C. (2002) *Real Mathematical Analysis*, Springer, New York, USA, 437 pp.
- Rapp, R. H. (1977) The use of gravity anomalies on a bounding sphere to improve potential coefficient determinations, *Report 254*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 18 pp.
- Rapp, R. H. (1981) Ellipsoidal corrections for geoid undulation computations using gravity anomalies in a cap, *Journal of Geophysical Research*, Vol. 86, No. B11, pp. 10843–10848.
- Rapp, R. H. (1986) Global geopotential solutions, in: *Mathematical and Numerical Techniques in Physical Geodesy*, Süinkel, H. (ed.), Springer, Berlin, Germany, pp. 365–415.
- Rapp, R. H. (1997a) Global models for the 1 cm geoid - present status and near term prospects, in: *Geodetic Boundary Value Problems in View of the One Centimeter Geoid*, Sansò, F. and Rummel, R. (eds.), Springer, Berlin, Germany, pp. 273–311.

- Rapp, R. H. (1997b) Use of potential coefficient models for geoid undulation determinations using a spherical harmonic representation of the height anomaly/geoid undulation difference, *Journal of Geodesy*, Vol. 71, No. 5, pp. 282–289, DOI: 10.1007/s001900050096.
- Rapp, R. H. and Bašić, T. (1992) Oceanwide gravity anomalies from GEOS-3, SEASAT and GEOSAT altimeter data, *Geophysical Research Letters*, Vol. 19, No. 19, pp. 1979–1982.
- Rapp, R. H. and Cruz, J. Y. (1986) Spherical harmonic expansions of the Earth's gravitational potential to degree 360 using 30' mean anomalies, *Report 376*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 22 pp.
- Rapp, R. H. and Pavlis, N. K. (1990) The development and analysis of geopotential coefficient models to spherical harmonic degree 360, *Journal of Geophysical Research*, Vol. 95, No. B13, pp. 21585–21911.
- Reigber, C., Balmino, G., Schwintzer, P., Biancale, R., Bode, A., Lemoine, J. M., Koenig, R., Loyer, S., Neumayer, H., Marty, J. C., Barthelmes, F., Perosanz, F. and Zhu, S. Y. (2002a) A high-quality global gravity field model from CHAMP GPS tracking data and accelerometry (EIGEN-1S), *Geophysical Research Letters*, Vol. 29, No. 14, DOI:10.1029/2002GL015064.
- Reigber, C., Luehr, H. and Schwintzer, P. (2002b) CHAMP mission status, *Advances in Space Research*, Vol. 30, No. 2, pp. 129–134, DOI: 10.1016/S0273-1177(02)00276-4.
- Reigber, C., Schmidt, R., Flechtner, F., Koenig, R., Meyer, U., Neumayer, K. H., Schwintzer, P. and Zhu, S. Y. (2005) An Earth gravity field model complete to degree and order 150 from GRACE: EIGEN-GRACE02S, *Journal of Geodynamics*, Vol. 39, No. 1, pp. 1–10, DOI: 10.1016/j.jog.2004.07.001.
- Ritter, S. (1998) The nullfield method for the ellipsoidal Stokes problem, *Journal of Geodesy*, Vol. 72, No. 2, pp. 101–106, DOI: 10.1007/s001900050151.
- Rizos, C. (1979). An efficient computer technique for the evaluation of geopotential from spherical harmonic models, *Australian Journal of Geodesy, Photogrammetry and Surveying*, No. 31, pp. 161-169.
- Rummel, R. (1985) From the Observational Model to Gravity Parameter Estimation, in: *Proceedings of the Beijing International Summer School: Local Gravity Field Approximation*, Schwarz, K.P. (ed.), University of Calgary, Calgary, Canada, pp. 67–106.
- Rummel, R. (1997) Spherical spectral properties of the Earth's gravitational potential and its first and second derivatives, in: *Geodetic Boundary Value Problems in View of the One Centimeter Geoid*, Sansò, F. and Rummel, R. (eds.), Springer, Berlin, Germany, pp. 359–404.

- Rummel, R. and Rapp, R. (1976) The influence of the atmosphere on geoid and potential coefficient determinations from gravity data, *Journal of Geophysical Research*, Vol. 81, No. 2, pp. 5639–5642.
- Rummel, R. and Van Gelderen, M. (1992) Spectral analysis of the full gravity tensor, *Geophysical Journal International*, Vol. 111, No. 1, pp. 159–169.
- Rummel, R. and Van Gelderen, M. (1995) Meissl scheme - spectral characteristics of physical geodesy, *manuscripta geodaetica*, Vol. 20, No. 5, pp. 379–385.
- Rummel, R., Balmino, G., Johannesen, J. A., Visser, P. and Woodworth, P. (2002) Dedicated gravity field missions - principles and aims, *Journal of Geodynamics*, Vol. 33, No. 1-2, pp. 3–20, DOI: 10.1016/S0264-3707(01)00050-3.
- Rummel, R., Teunissen, P. and Van Gelderen, M. (1989) Uniquely and overdetermined geodetic boundary value problems by least squares, *Bulletin Géodésique*, Vol. 63, No. 1, pp. 1–33.
- Sacerdote, F. and Sansò, F. (1985) Overdetermined boundary value problems in physical geodesy, *manuscripta geodaetica*, Vol. 10, No. 3, pp. 195–207.
- Sacerdote, F. and Sansò, F. (1986) The scalar boundary value problem of physical geodesy, *manuscripta geodaetica*, Vol. 11, No. 1, pp. 15–28.
- Safari, A., Ardalan, A. A. and Grafarend, E. W. (2005) A new ellipsoidal gravimetric, satellite altimetry and astronomic boundary value problem, a case study: the geoid of Iran, *Journal of Geodynamics*, Vol. 39, No. 5, pp. 545–568, DOI: 10.1016/j.jog.2005.04.009.
- Sagrebin, D. W. (1956) Die theorie des regularisierten geoids, *Report No. 9*, Veröffentlichungen Geodätisches Institut Potsdam, Potsdam, Germany.
- Sansò, F. (1981) Recent advances in the theory of the geodetic boundary value problem, *Reviews of Geophysics and Space Physics*, Vol. 19, No. 3, pp. 437–449.
- Sansò, F. (1995) The long road from measurements to boundary value problems in physical geodesy, *manuscripta geodaetica*, Vol. 20, No. 5, pp. 326–344.
- Sansò, F. and Sona, G. (2001) ELGRAM: an ellipsoidal gravity model manipulator, *Bollettino di Geodesia e Scienze Affini*, Vol. 60, No. 3, pp. 215–226.
- Sansò, F. and Tscherning, C. C. (2003) Fast spherical collocation: Theory and examples, *Journal of Geodesy*, Vol. 77, No. 1-2, pp. 101–112, DOI: 10.1007/s00190-002-0310-5.
- Schwarz, K. P. (1985) Data types and their spectral properties, in: *Proceedings of the Beijing International Summer School: Local Gravity Field Approximation*, Schwarz, K.P. (ed.), University of Calgary, Calgary, Canada, pp. 1–66.
- Schwarz, K. P. (1996) *Fundamentals of Geodesy*, Department of Geomatics Engineering, University of Calgary, Calgary, Canada, 187 pp.

- Seitz, K., Schramm, B. and Heck, B. (1994) Non-linear effects in the scalar free geodetic boundary value problem based on reference fields of various degrees, *manuscripta geodaetica*, Vol. 19, No. 6, pp. 327–338.
- Sideris, M. G. and She, B. B. (1995) A new, high-resolution geoid for Canada and part of the US by the 1D-FFT method, *Bulletin Géodésique*, Vol. 69, No. 2, pp. 92–108, DOI: 10.1007/BF00819555.
- Sideris, M. G., Fei, Z. L. and Blair, J. A. R. (2002). Ellipsoidal corrections for the inverse Hotine/Stokes formulas, in: *Geodesy - the Challenge of the Third Millennium*, Grafarend, E.W., Krumm, F.W. and Schwarze, V.S. (eds.), Springer, Berlin, Germany, pp. 247–256.
- Sigl, R. (1985) *Introduction to Potential Theory*, Abacus Press, Tunbridge Wells, UK, 229 pp.
- Sjöberg, L. E. (1980) On the convergence problem for the spherical harmonic expansion of the geopotential at the surface of the earth, *Bollettino di Geodesia e Scienze Affini*, Vol. 39, No. 3, pp. 261–271.
- Sjöberg, L. E. (1984) On the convergence problem of the satellite derived spherical harmonic expansion of the geopotential at the Earth's surface, *Proceedings of the 5th International Symposium "Geodesy and Physics of the Earth"*, Magdeburg, German Democratic Republic, pp. 108–114.
- Sjöberg, L. E. (1988) A new integral approach to geopotential coefficient determinations from terrestrial gravity or satellite altimetry data, *Bulletin Géodésique*, Vol. 62, No. 2, pp. 93–101.
- Sjöberg, L. E. (1991) Refined least squares modification of Stokes' formula, *manuscripta geodaetica*, Vol. 16, No. 6, pp. 367–375.
- Sjöberg, L. E. (2003a) The ellipsoidal correction to Stokes's formula, in: *Proceedings of the 3rd Meeting of the International Gravity and Geoid Commission: Gravity and Geoid 2002*, Tziavos, I.N. (ed.), Ziti Editions, Thessaloniki, Greece, pp. 97–101.
- Sjöberg, L. E. (2003b) The correction to the modified Stokes formula for an ellipsoidal Earth, in: *Honoring the academic life of Petr Vaníček*, UNB Technical Report No. 218, Santos, M. (ed.), University of New Brunswick, Fredericton, Canada, pp. 99–110.
- Sjöberg, L. E. (2003c) The ellipsoidal corrections to order e^2 of geopotential coefficients and Stokes' formula, *Journal of Geodesy*, Vol. 77, No. 3-4, pp. 139–147, DOI: 10.1007/s00190-003-0321-x.
- Sjöberg, L. E. (2004a) The ellipsoidal corrections to the topographic geoid effects, *Journal of Geodesy*, Vol. 77, No. 12, pp. 804–808, DOI: 10.1007/s00190-004-0377-2.
- Sjöberg, L. E. (2004b) A spherical harmonic representation of the ellipsoidal correction to the modified Stokes formula, *Journal of Geodesy*, Vol. 78, No. 3, pp. 180–186, DOI: 10.1007/s00190-004-0378-1.

- Sjöberg, L. E. (2005) A discussion on the approximations made in the practical implementation of the remove-compute-restore technique in regional geoid modelling, *Journal of Geodesy*, Vol. 78, No. 11-12, pp. 645–653, DOI: 10.1007/s00190-004-0430-1.
- Smith, D. A. (1998). There is no such thing as “the” EGM96 geoid: Subtle points on the use of a global geopotential model, *IGeS Bulletin*, No. 8, pp. 17-28.
- Smith, D. A. and Roman, D. R. (2001) GEOID99 and G99SSS: 1-arc-minute geoid models for the United States, *Journal of Geodesy*, Vol. 75, No. 9-10, pp. 469–490, DOI: 10.1007/s001900100200.
- Sneeuw, N. (1994) Global spherical harmonic analysis by least-squares and numerical quadrature methods in historical perspective, *Geophysical Journal International*, Vol. 118, No. 3, pp. 707–716.
- Sneeuw, N. (2000) A Semi-analytical Approach to Gravity Field Analysis from Satellite Observations, *PhD thesis*, Institut für Astronomische und Physikalische Geodäsie, Technische Universität München, Munich, Germany, 112 pp.
- Sneeuw, N., Gerlach, C., Földvary, L., Gruber, T., Peters, T., Rummel, R. and Švehla, D. (2005) One year of time-variable CHAMP-only gravity field models using kinematic orbits, in: *Proceedings of the International Association of Geodesy: A Window on the Future of Geodesy*, Sansò, F. (ed.), IAG Symposia Series, Vol. 128, Springer, Berlin, Germany, pp. 288–293.
- Sona, G. (1995) Numerical problems in the computation of ellipsoidal harmonics, *Journal of Geodesy*, Vol. 70, No. 1-2, pp. 117–126, DOI: 10.1007/BF00863423.
- Stewart, J. (1995) *Calculus* (3rd Edition), Brooks/Cole, Pacific Grove, USA, 1130 pp.
- Stokes, G. G. (1849). On the variation of gravity and the surface of the Earth: Cambridge Philosophical Society, Transactions Vol. 8, pp. 672-695.
- Strang, G. (1986) *Linear Algebra and its Applications* (3rd Edition), Harcourt Brace and Company, Fort Worth, USA, 505 pp.
- Svensson, S. L. (1988) Some remarks on the altimetry-gravimetry problem, *manuscripta geodaetica*, Vol. 13, No. 2, pp. 63–74.
- Tapley, B. D., Bettadpur, S., Watkins, M. and Reigber, C. (2004) The gravity recovery and climate experiment: Mission overview and early results, *Geophysical Research Letters*, Vol. 31, No. 9, DOI:10.1029/2004GL019920.
- Tapley, B. D., Chambers, D. P., Bettadpur, S. and Ries, J. C. (2003) Large scale ocean circulation from the GRACE GGM01 geoid, *Geophysical Research Letters*, Vol. 30, No. 22, DOI:10.1029/2003GL018622.
- Tapley, B., Ries, J., Bettadpur, S., Chambers, D., Cheng, M., Condi, F., Gunter, B., Kang, Z., Nagel, P., Pastor, R., Pekker, T., Poole, S. and Wang, F. (2005) GGM02 - An improved Earth gravity field model from GRACE, *Journal of Geodesy*, Vol. 79, No. 8, pp. 467–478, DOI: 10.1007/s00190-005-0480-z.

- Tenzer, R., Vaníček, P., Santos, M., Featherstone, W. E. and Kuhn, M. (2005) The rigorous determination of orthometric heights, *Journal of Geodesy*, Vol. 79, No. 1-3, pp. 82–92, DOI: 10.1007/s00190-005-0445-2.
- Thông, N. C. (1996) Explicit expression and regularization of the harmonic reproducing kernels for the Earth's ellipsoid, *Journal of Geodesy*, Vol. 70, No. 9, pp. 533–538, DOI: 10.1007/s001900050041.
- Thông, N. C. and Grafarend, E. W. (1989) A spheroidal harmonic model of the terrestrial gravitational field, *manuscripta geodaetica*, Vol. 14, No. 5, pp. 285–304.
- Tikhonov, A. N. and Arsenin, V. Y. (1977) *Solutions to Ill-Posed Problems*, Winston, New York, USA, 258 pp.
- Torge, W. (1991) *Geodesy*, De Gruyter, Berlin, Germany, 264 pp.
- Tscherning, C. C. (2004) A discussion on the use of spherical approximation or no approximation in gravity field modelling with emphasis on unsolved problems in least squares collocation, in: *Proceedings of the 5th Hotine-Marussi Symposium on Mathematical Geodesy*, Sansò, F. (ed.), IAG Symposia Series, Vol. 127, Springer, Berlin, Germany, pp. 184–188.
- Tziavos, I. N. (1996) Comparisons of spectral techniques for geoid computations over large regions, *Journal of Geodesy*, Vol. 70, No. 6, pp. 357–373, DOI: 10.1007/BF00868188.
- Van Gelderen, M. and Rummel, R. (2001) The solution of the general geodetic boundary value problem by least squares, *Journal of Geodesy*, Vol. 75, pp. 1–11, DOI: 10.1007/s001900000146.
- Vaníček, P. and Featherstone, W. E. (1998) Performance of three types of Stokes's kernel in the combined solution for the geoid, *Journal of Geodesy*, Vol. 72, No. 12, pp. 684–697, DOI: 10.1007/s001900050209.
- Vaníček, P. and Krakiwsky, E. (1986) *Geodesy, the Concepts* (2nd Edition), Elsevier Science, Amsterdam, The Netherlands, 697 pp.
- Vaníček, P. and Sjöberg, L. E. (1991) Reformulation of Stokes's theory for higher than second degree reference field and modification of integral kernels, *Journal of Geophysical Research*, Vol. 96, No. B4, pp. 6529–6540.
- Vaníček, P., Huang, J., Novák, P., Pagiatakis, S. D., Véronneau, M., Martinec, Z. and Featherstone, W. E. (1999) Determination of the boundary values for the Stokes-Helmert problem, *Journal of Geodesy*, Vol. 73, pp. 180–192.
- Vaníček, P., Kleusberg, A., Martinec, Z., Sun, W., Ong, P., Najafi, M., Vajda, P., Harrie, L. and Ter Horst, B. (1995) Compilation of a precise regional geoid, DSS Contract No. 23244-1-4405/01-SS, *Report for Geodetic Survey Division*, Ottawa, 45 pp.

- Vaníček, P., Kleusberg, A., Martinec, Z., Sun, W., Ong, P., Najafi, M., Vajda, P., Harrie, L., Tomášek, P. and Ter Horst, B. (1996) Downward continuation of Helmert's gravity, *Journal of Geodesy*, Vol. 71, No. 1, pp. 21–34, DOI: 10.1007/s001900050072.
- Vaníček, P., Tenzer, R., Sjöberg, L. E., Martinec, Z. and Featherstone, W. E. (2004) New views of the spherical Bouguer gravity anomaly, *Geophysical Journal International*, Vol. 159, No. 2, pp. 460–472, DOI: 10.1111/j.1365–246X.2004.02435.x.
- Vergos, G. S., Tziavos, I. N. and Andritsanos, V. D. (2005) On the determination of marine geoid models by least-squares collocation and spectral methods using heterogeneous data, in: *Proceedings of the International Association of Geodesy: A Window on the Future of Geodesy*, Sansò, F. (ed.), IAG Symposia Series, Vol. 128, Springer, Berlin, Germany, pp. 332–337.
- Vermeille, H. (2002) Direct transformation from geocentric to geodetic coordinates, *Journal of Geodesy*, Vol. 76, No. 8, pp. 451–454, DOI: 10.1007/s00190–002–0273–6.
- Vincent, S. and Marsh, J. G. (1973) Global detailed gravimetric geoid, in: *Proceedings of the International Symposium on the use of Artificial Satellites for Geodesy and Geodynamics*, Veis, G. (ed.), National Technical University, Athens, Greece, pp. 825–855.
- Wahr, J., Svenson, S., Zlotnicki, V. and Velicogna, I. (2004) Time-variable gravity from GRACE: First results, *Geophysical Research Letters*, Vol. 31, L11501, DOI: 10.1029/2004GL019779.
- Wang, Y. M. (1999) On the ellipsoidal corrections to gravity anomalies computed using the inverse Stokes integral, *Journal of Geodesy*, Vol. 73, No. 1, pp. 29–34, DOI: 10.1007/s001900050215.
- Wenzel, H. G. (1998) Ultra-high degree geopotential models GPM98A and GPM98B to degree 1800, in: *Proceedings of the Second Continental Workshop on the Geoid in Europe*, Vermeer, M and Ádám, J. (eds.), Report 98:4, Finnish Geodetic Institute, Masala, Finland, pp. 71–80.
- Wong, L. and Gore, R. (1969) Accuracy of geoid heights from modified Stokes' kernels, *Geophysical Journal of the Royal Astronomical Society*, No. 18, pp. 81–91.
- Yu, J. H. and Cao, H. S. (1996) Elliptical harmonic series and the original Stokes problem with the boundary of the reference ellipsoid, *Journal of Geodesy*, Vol. 70, No. 7, pp. 431–439, DOI: 10.1007/BF01090818.
- Yu, J., Jekeli, C. and Zhu, M. (2003) Analytical solutions of the Dirichlet and Neumann boundary-value problems with an ellipsoidal boundary, *Journal of Geodesy*, Vol. 76, No. 11–12, pp. 653–667, DOI: 10.1007/s00190–002–0271–8.
- Yurkina, M. I. (2002) A solution of Stokes' problem for the ellipsoidal Earth by means of Green's function, in: *Geodesy - the Challenge of the Third Millennium*, Grafarend, E.W., Krumm, F.W. and Schwarze, V.S. (eds.), Springer, Berlin, Germany, pp. 261–268.

- Zhang, C. D., Hsu, H. T., Wu, X. P., Li, S. S., Wang, Q. B., Chai, H. Z. and Du, L. (2005) An alternative algebraic algorithm to transform Cartesian to geodetic coordinates, *Journal of Geodesy*, Vol. 79, No. 8, pp. 413–420, DOI: 10.1007/s00190-005-0487-5.
- Zhu, Z. (1981) The Stokes problem for the ellipsoid using ellipsoidal kernels, *Report 319*, Department of Geodetic Science and Surveying, Ohio State University, Columbus, USA, 19 pp.

APPENDIX A: RELATION AMONG FULLY NORMALISED LEGENDRE WEIGHT FUNCTIONS

A proof for the identity among fully normalised cosinusoidal Legendre weight functions (LWFs) \overline{F}_{nm}^{ij} given in Equation (3.42) can be found by mathematical induction, where Equation (3.42) forms the proposition

$$\overline{F}_{nm}^{ij} = \overline{F}_{n+i,m}^{-i,j} \quad (\text{A-1})$$

It can easily be seen from Equations (3.19) and (3.20) that this proposition holds for the base case $i = j = 1$. Assuming that the inductive hypothesis

$$\overline{F}_{nm}^{pq} = \overline{F}_{n+p,m}^{-p,q} \quad (\text{A-2})$$

is true for all p and q , the two propositions of the inductive step remain to be proven, i.e.,

$$\overline{F}_{nm}^{p+1,q} = \overline{F}_{n+p+1,m}^{-p-1,q} \quad (\text{A-3})$$

$$\overline{F}_{nm}^{p,q+1} = \overline{F}_{n+p,m}^{-p,q+1} \quad (\text{A-4})$$

Inserting Equation (3.38) in the left-hand side of Equation (A-3), and subsequently applying the inductive hypothesis in Equation (A-2), gives an equation, which – according to Equation (3.36) – is exactly the same as the right-hand side of Equation (A-3)

$$\begin{aligned} \overline{F}_{nm}^{p+1,q} &= \overline{F}_{nm}^{p+2,q-1} \overline{F}_{n+p+2,m}^{-1,1} + \overline{F}_{nm}^{p,q-1} \overline{F}_{n+p,m}^{1,1} \\ &= \overline{F}_{n+p+2,m}^{-p-2,q-1} \overline{F}_{n+p+1,m}^{1,1} + \overline{F}_{n+p,m}^{-p,q-1} \overline{F}_{n+p+1,m}^{-1,1} \\ &= \overline{F}_{n+p+1,m}^{-p-1,q} \quad \text{Q.E.D.} \end{aligned} \quad (\text{A-5})$$

A similar procedure can be followed to prove Equation (A-4)

$$\begin{aligned} \overline{F}_{nm}^{p,q+1} &= \overline{F}_{nm}^{-1,1} \overline{F}_{n-1,m}^{p+1,q} + \overline{F}_{nm}^{1,1} \overline{F}_{n+1,m}^{-p-1,q} \\ &= \overline{F}_{n-1,m}^{1,1} \overline{F}_{n+p,m}^{-p-1,q} + \overline{F}_{n+1,m}^{-1,1} \overline{F}_{n+p,m}^{-p+1,q} \\ &= \overline{F}_{n+p,m}^{-p,q+1} \quad \text{Q.E.D.} \end{aligned} \quad (\text{A-6})$$

Since the base case and the inductive step are true, the proposition in Equation (A-1) is true for all i and j . Moreover, according to Equation (3.30), the base case $i = j = 1$ of Equation (A-1) also holds for the sinusoidal Legendre weight functions \bar{K}_{nm}^{ij} . Because these sinusoidal Legendre weight functions also obey the recursive relationships (Equations 3.33 and 3.38), it can be proven along the same lines that Equation (A-1) holds for all sinusoidal Legendre weight functions as well.

APPENDIX B: REARRANGEMENT OF SUMMATION ORDER

When two summations are nested, the order of the summations can be rearranged. In the situation where the summed quantities can be split into one part depending on both summation indices and one part depending on the sum of both summation indices, the summation order can be rearranged so that the second part only depends on one of the summation indices

$$\sum_{i=i_{min}}^{i_{max}} \sum_{j=j_{min}}^{j_{max}} a_{ij} x_{i+j} = \sum_{k=i_{min}+j_{min}}^{i_{max}+j_{max}} b_k x_k \quad (\text{B-1})$$

where

$$b_k = \sum_{l=l_{min}}^{l_{max}} a_{l,k-l} \quad (\text{B-2})$$

and

$$l_{min} = \max(i_{min}, k - j_{max}) \quad (\text{B-3})$$

$$l_{max} = \min(i_{max}, k - j_{min}) \quad (\text{B-4})$$

The boundaries i_{min} , i_{max} , j_{min} and j_{max} can be any integer or $\pm\infty$. Equations (B-1) - (B-4) can be obtained from investigation of Figure B-1. A 2D summation over a rectangular grid of points is commonly presented as a summation over the two axes i and j , since the summation limits are in this case all constants. It is, however, also possible to transform the summation indices to k and l , where the k -axis is perpendicular to the lines of constant $i + j$. The summation limits over l then depend on index k and are given by Equations (B-3) and (B-4), as can be verified by inspection of Figure B-1.

A special case of rearrangement formula (Equation B-1) is the so-called Cauchy multiplication (e.g., Protter and Morrey, 1964), which is a discrete version of convolution. This involves the product of two power series and enables to combine these into one power series. Given two power series

$$A = \sum_{i=i_{min}}^{i_{max}} a_i x^i \quad (\text{B-5})$$

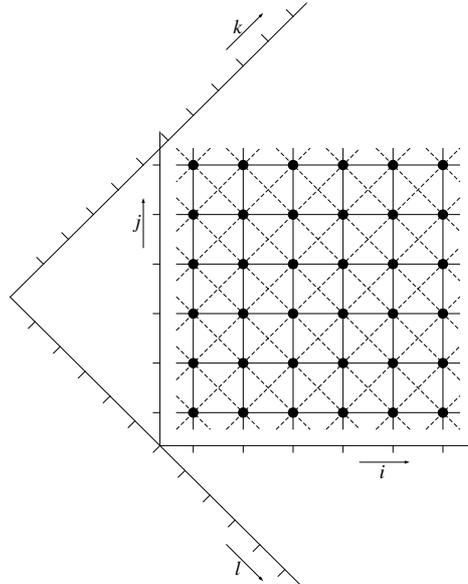


Figure B-1: Schematic overview of a rearrangement of summation order from indices i and j to indices k and l

and

$$B = \sum_{j=j_{\min}}^{j_{\max}} b_j x^j \quad (\text{B-6})$$

the product of both series is of the form of the left-hand side of Equation (B-1)

$$AB = \sum_{i=i_{\min}}^{i_{\max}} \sum_{j=j_{\min}}^{j_{\max}} a_i b_j x^{i+j} \quad (\text{B-7})$$

Thus, the product AB can be written as a summation where the parameter x only depends on one summation index

$$AB = \sum_{k=i_{\min}+j_{\min}}^{i_{\max}+j_{\max}} \sum_{l=l_{\min}}^{l_{\max}} a_l b_{k-l} x^k \quad (\text{B-8})$$

This is a generalisation of the commonly presented Cauchy multiplication of two power series running from zero to infinity (e.g., Protter and Morrey, 1964; Pugh, 2002), in which case k also runs from zero to infinity and l runs from zero to k .

In the case of infinite series, it is important to note that convergence of the resulting series can only be guaranteed if one of the series converges and the other series converges absolutely. The scalar to which the series in Equation (B-7) will converge is in this case always equal to the product of the infinite sums of A and B .

When the limits of the summation over j in the left-hand side of Equation (B-1) are dependent on index i , Equations (B-1) and (B-8) do not hold. A case that frequently appears in this thesis is the situation where i runs from 0 to ∞ , and j runs from $-i$ to i

$$C = \sum_{i=0}^{\infty} \sum_{j=-i}^i c_{ij} \quad (\text{B-9})$$

This also allows for a rearrangement of the summation order, but one of a whole different type, where the indices of the variable are not affected. The summations over i and j can simply be reversed, changing the summation limits as follows

$$C = \sum_{j=-\infty}^{\infty} \sum_{i=|j|}^{\infty} c_{ij} \quad (\text{B-10})$$

Naturally, such a reversal of summations can also be applied in the case that the range of the second summation is independent of the first summation index, in which case the summation limits do not change at all.

APPENDIX C: BINOMIAL SERIES EXPANSION

A function of the form

$$f(x) = (1 + x)^k \quad (\text{C-1})$$

where x and k are real numbers, can be expressed as a binomial series

$$f(x) = \sum_{j=0}^{\infty} \binom{k}{j} x^j \quad (\text{C-2})$$

where $\binom{k}{j}$ is the binomial coefficient, which can be defined as

$$\binom{k}{j} = \frac{k(k-1)(k-2)\cdots(k-j+1)}{j!} \quad \text{for } j \geq 1, \quad \text{and} \quad \binom{k}{0} = 1 \quad (\text{C-3})$$

or as a closed expression

$$\binom{k}{j} = \frac{\Gamma(k+1)}{j!\Gamma(k-j+1)} \quad (\text{C-4})$$

where Γ is the gamma-function, which is an extension of the factorial into the domain of real and complex numbers. When k is a natural number, the gamma-function is defined as

$$\Gamma(k+1) = k! \quad \text{for } k \in \mathbb{N} \quad (\text{C-5})$$

When k is half of an integer, an explicit expression for the gamma-function can also be found

$$\Gamma(k+1) = \frac{(2k)!\sqrt{\pi}}{2^{2k}(k-\frac{1}{2})!} \quad \text{for } k \in \mathbb{H} \quad (\text{C-6})$$

where \mathbb{H} is the set of half-integers. Thus, implementing Equations (C-5) and (C-6) in Equation (C-4) gives explicit expressions for the binomial coefficient

$$\binom{k}{j} = \frac{k!}{j!(k-j)!} \quad \text{for } k \in \mathbb{N} \quad (\text{C-7})$$

$$\binom{k}{j} = \frac{(2k)!(k-j-\frac{1}{2})!}{2^{2k}j!(k-\frac{1}{2})!(2k-2j)!} \quad \text{for } k \in \mathbb{H} \quad (\text{C-8})$$

The convergence of a binomial series can be investigated using the ratio test (Stewart, 1995). The absolute value of the ratio of two consecutive terms in the binomial

expansion, a_j and a_{j+1} , is

$$\left| \frac{a_{j+1}}{a_j} \right| = \frac{|k-j|}{j+1} |x| \quad (\text{C-9})$$

This value tends to zero for higher terms for $|x| < 1$, which means that absolute convergence is assured in this region

$$\lim_{j \rightarrow \infty} \frac{|k-j|}{j+1} |x| = 0 \quad \text{for } |x| < 1 \quad (\text{C-10})$$

The rate of convergence depends strongly on the values of x and k . For negative x , the series is alternating, in which case the error in the truncated series is always smaller than the absolute value of the first truncated term, provided that the absolute values of the terms decrease. This is only the case if the ratio in Equation (C-9) is less than one, which occurs for

$$j > \frac{k|x| - 1}{1 + |x|} \quad (\text{C-11})$$

If this condition is met, the error ς of a truncation of the binomial series after the j -th term will be less than the term $j + 1$. Thus, an upper limit for the truncation error is the following

$$\varsigma < \binom{k}{j} |x|^j < \frac{k^j}{j!} |x|^j, \quad \text{for } j > 1 \quad (\text{C-12})$$

which is a close approximation for high values of j and k . Equations (C-11) and (C-12) show that, in general, the convergence will be slower for higher values of k and for higher absolute values of x .