NUMERICAL METHODS FOR HARMONIC ANALYSIS
ON THE SPHERE

by

Oscar L. Colombo

The Ohio State University
Department of Geodetic Science
1958 Neil Avenue
Columbus, Ohio 43210

March 1981
Foreword

This report was prepared by Dr. Oscar L. Colombo, Post Doctoral Researcher, Department of Geodetic Science, The Ohio State University, under Air Force Contract No. F19628-79-C-0027, The Ohio State University Research Foundation Project No. 711664, Project Supervisor Richard H. Rapp. The contract covering this research is administered by the Air Force Geophysics Laboratory, Hanscom Air Force Base, Massachusetts, with Mr. Bela Szabo, Contract Monitor.
Foreword

This report was prepared by Dr. Oscar L. Colombo, Post Doctoral Researcher, Department of Geodetic Science, The Ohio State University, under Air Force Contract No. F19628-79-C-0027, The Ohio State University Research Foundation Project No. 711664, Project Supervisor Richard H. Rapp. The contract covering this research is administered by the Air Force Geophysics Laboratory, Hanscom Air Force Base, Massachusetts, with Mr. Bela Szabo, Contract Monitor.

The reproduction and distribution of this report was carried out through funds supplied by the Department of Geodetic Science. This report was also distributed by the Air Force Geophysics Laboratory as document AFGL-TR-80-0294, Scientific Report No. 7 under Contract No. F19628-79-C-0027.
Acknowledgements

I am indebted to Dr. R. H. Rapp for his support during the long months when this report was gestating; to K. Katsambalos and C. Jekeli for comments, particularly on Section 1; to Pamela Hager for typing the beginning of section 1, and section 3; and to Susan L. Carroll for typing all the rest, a hard task that she has done with care and good cheer.

The ideas presented in the pages that follow began to develop while I was working for the late Dr. R. S. Mather at the University of New South Wales, Australia. They continued to grow, out of discussions with him, his students, and many others, first in Australia and then in the USA. Those discussions made clear to me that there was (and still is, of course) real need to find efficient ways for handling the large amounts of data that are becoming available to Earth scientists on a global basis, chiefly through the use of space technology. To all of them, my sincere thanks for their beneficial influence.
Table of Contents

Foreword .............................................................. iii

Acknowledgements ...................................................... iv

1. Introduction ......................................................... 1

1.1 Spherical Harmonic Analysis and Synthesis: Definitions ...... 2

1.2 Relationships between Spherical Harmonics and 2-D Fourier Series ........................................ 6

1.3 Sampling Errors ................................................... 9

(a) Number of equations and number of unknowns .......... 11

(b) 100% aliasing ......................................................... 12

(c) Orthogonality ......................................................... 12

1.4 Number of Operations in Analysis and in Synthesis .......... 13

1.5 Algorithm for the Analysis of Point Values .................. 15

1.6 Algorithm for the Synthesis of Point Values ................. 17

1.7 Algorithms for the Analysis and Synthesis of Area Means ... 18

1.8 Duality between Analysis and Synthesis ...................... 18

1.9 Usefulness of the Fast Fourier Transform Method ............ 19

1.10 Functions Harmonic in Space and their Gradients ........... 21

2. Error Measure and Optimal Quadrature Formulas .............. 23

2.1 The Isotropic Covariance ......................................... 23

2.2 Some Additional Notation ......................................... 26

2.3 Estimation Errors, Sampling Errors, and Propagated Noise .29

2.4 The Quadratic Error Measure. .................................. 30

(a) Propagated noise measure ...................................... 30

(b) Sampling Error Measure ......................................... 31
2.5 The Meaning of the Error Measure ......................... 32
2.6 Simple Formulas for Area Means .......................... 33
2.7 Optimum de-Smoothing Factors ........................... 35
2.8 Least Squares Collocation ................................. 36
2.9 The Best Quadrature Formula for non-Uniform, Uncorrelated Noise ........................................ 38
2.10 The Structure of the Covariance Matrix and its Consequences .... 40
2.11 Setting up and Inverting the Covariance Matrix ............... 43
2.12 Optimal Formulas for non-Uniform, Correlated Noise ...... 46
2.13 Least Squares Adjustment and Least Squares Collocation .... 47
   (a) Band-limited signal .................................. 47
   (b) Infinite bandwidth ................................... 50
2.14 Ridge Regression and Least Squares Collocation .............. 51
2.15 Structure of the Normal Matrix ............................. 53
2.16 Global Adjustment and Collocation with Scattered Data ..... 56
   (a) Full region bound by lines of latitude and longitude ..... 56
   (b) Arbitrarily scattered data ............................. 57
2.17 The Error Matrix in the Band Limited Case .................. 63
2.18 The Use of a Priori Information of the Coefficients ......... 63
   (a) Simple weighted averages ............................. 64
   (b) A priori values included as data ..................... 65
   (c) The method of Kaula and Rapp ....................... 65
2.19 Optimal Estimation over a Band of Spatial Frequencies ... 66
3. Numerical Examples .................................................. 68
   3.1 Generation and Analysis of Simulated Data .................. 68
   3.2 Agreement between the Actual and the Theoretical Measures of the Sampling Errors .............................................. 71
   3.3 Accuracies of Various Quadratures Formulas .................. 71
   3.4 The Analysis of a Global Data Set of $5^\circ \times 5^\circ$ Mean Anomalies . 80
4. Covariances between Area Means ...................................... 82
   4.1 Derivation of an Approximate Formula for the Covariance . 82
   4.2 Choosing $N_{max}$ ..................................................... 84
   4.3 Numerical Examples .................................................. 85
   4.4 Covariances between Mean Values and Point Values .......... 90
5. Conclusions ...................................................................... 92

References ........................................................................ 94

Appendix A: Term by Term Integration of Formula (4.5) (Proof of) . 97

Appendix B: Computer Programs (Descriptions and Listings) .... 100

   B.1 General Programming Considerations .............................. 100
   B.2 Subroutine SSYNTH .................................................... 101
   B.3 Subroutine HARMIN ................................................. 106
   B.4 Subroutine NORMAL .................................................. 112
   B.5 Subroutine NORMAX .................................................. 124
   B.6 Subroutine LEGFDN ................................................... 131
   B.7 Subroutine NVAR ....................................................... 134
   B.8 Subroutine COVBLK ................................................... 138
1. Introduction

Spherical harmonics are closely associated with the basic theory of gravitational and magnetic fields, such as those of the Earth and planets; for this reason they are important both in geodesy and in Earth and planetary physics.

The present work considers the numerical aspects of the reduction of global data sets to spherical harmonic coefficients, so the emphasis has been laid on the algorithms for this purpose. The procedures for harmonic analysis (and synthesis) given here are general enough to be used in the study of magnetic or electric fields, but most conclusions regarding their accuracy are restricted to the gravity field of our planet and to fields with the same power spectrum. The accuracy of these methods cannot be separated from the type of signal being used.

Modern instrumentation has provided scientists and engineers with vast amounts of information, and modern computers have made the processing of it possible, and even routine, thanks to constant improvements in both hardware and software. In the mid sixties, those branches of applied mathematics, physics, and engineering concerned with the sifting of data, or with the study of very large regular structures, were greatly affected by the advent of the Fast Fourier Transform (FFT). In spite of the fact that spherical harmonics are members of the family of Fourier transforms, closely related to two dimensional Fourier series, geodesy has lagged behind in the development of techniques similar to the FFT, partly because of the rather wicked nature of the sphere on which data are usually given, partly because there have not been enough data to make the development of powerful techniques a general concern. The topological differences between the Euclidean plane and the surface of the sphere may very well prevent the finding of algorithms as efficient as the FFT for the latter (certainly none seems to have been reported to date) but such algorithms should be regarded, nonetheless, as the desideratum for all those who wish to put their time and work into developing good numerical methods for spherical harmonic analysis.

The increasing use of artificial satellites for surveying the gravity field, particularly by radar altimetry and by the projected tracking of one satellite by another, are making the use of very efficient and accurate techniques for handling the resulting data imprescindible: even a casual review of the literature of the last few years will show that serious efforts to provide such techniques are getting under way. The days of scarce, scattered, unreliable data are just about over.

The remainder of this section defines the basic problem and the associated notation, shows the relationships between spherical harmonics and 2D-Fourier series, presents some of the similitudes and differences between both, and explains some efficient algorithms for harmonic analysis and synthesis that are common to a number of different problems. Section 2 begins by defining a quadratic measure for the accuracy of the estimated harmonic coefficients based
on the covariance functions of the signal and the noise; a discussion on the optimization of this measure follows, leading to the application of least squares collocation to harmonic analysis. The use and implementation of least squares adjustment follows, and then a discussion of the connection between least squares and least squares collocation, shown as alternative and efficient techniques for solving the same problem. The section closes with an algorithm for the case when data are irregularly distributed. Section 3 illustrates with several numerical examples some of the methods presented earlier. Section 4 introduces an efficient formula for computing the covariances between block averages when analyzing mean values by collocation. This formula is much more efficient than others based on the numerical quadratures of the "point" covariance function.

1.1. Spherical Harmonic Analysis and Synthesis: Definitions

A square integrable, analytical function \( f(\theta, \lambda) \) defined on the unit sphere \( 0 \leq \theta \leq \pi \) and \( 0 \leq \lambda \leq 2\pi \) can be expanded in a series of surface spherical harmonics

\[
f(\theta, \lambda) = \sum_{n=0}^{\infty} \sum_{s=-n}^{n} \overline{P}_{ns}(\cos \theta) \left( \overline{C}_{ns} \cos m\lambda + \overline{S}_{ns} \sin m\lambda \right)
\]

(1.1)

where: \( \overline{P}_{ns} \) are the associated Legendre functions of the first kind, fully normalized so \( \frac{1}{4\pi} \int_{-1}^{1} \overline{P}_{nm}(\cos \theta)^{2} \left( \frac{\cos \theta}{\sin \theta} \right)^{2} \sin \theta \, d\theta = 1 \) (here \( \int_{-1}^{1} d\theta \) indicates integration on the unit sphere);

\( \overline{C}_{ns}, \overline{S}_{ns} \) are the fully normalized spherical harmonic coefficients.

For the sake of brevity, the following alternative notation shall be used when possible:

\[
\overline{Y}_{ns}^{\alpha}(\theta, \lambda) = \begin{cases} \overline{P}_{ns}(\cos \theta) \cos m\lambda & \text{if } \alpha = 0 \\ \overline{P}_{ns}(\cos \theta) \sin m\lambda & \text{if } \alpha = 1 \\
\end{cases}
\]

and

\[
\overline{C}_{ns}^{\alpha} = \begin{cases} \overline{C}_{ns} & \text{if } \alpha = 0 \\ \overline{S}_{ns} & \text{if } \alpha = 1 \\
\end{cases}
\]

The purpose of spherical harmonic analysis is to estimate the coefficients \( \overline{C}_{ns}^{\alpha} \) from measurements of the signal \( f(\theta, \lambda) \). These measurements, which may be corrupted by some noise or error signal "n", and which are assumed in what follows to be finite in number, constitute the data. The individual samples are called \( z_{ij} \), so \( z_{ij} = f(\theta_{i}, \lambda_{j}) + n_{ij} \). The subscripts \( i \) and \( j \) are used only to designate the position of the sample in a two-dimensional array, or grid, covering in some more or less regular way the sphere; \( i \) corresponds always to (co)latitude, and \( j \) to longitude. While, as in the last paragraph of section 3, some places in the grid may be empty, the grid itself is defined by a set of complete parallels and meridians. Unless otherwise specified, the separation between the lines of
latitude can be variable, but that between meridians is always constant and equal to \( \Delta \lambda = \pi / N \), where \( N \) is an integer. The grid most often considered in this work is the equal angular grid, also called the regular grid; in this case \( \Delta \theta \) is also constant and equal to \( \Delta \lambda \).

For equal angular grids \( i \) and \( j \) take values \( 0 \leq i \leq N - 1 \) and \( 0 \leq j \leq 2N - 1 \), \( i \) increasing from North to South, and \( j \) from West to East. Formulas where the \( i,j \) subscripts appear are in the form appropriate to the regular grid, though most of them can be extended in a very simple way to other partitions.

Data may consist of values determined at the intersections of the grid, in which case they are referred to as "point data", or they may be averages over the blocks defined by the lines of the grid, and then they are called "area means" or "block means". In the equal angular grid the northernmost and southernmost blocks reach to the respective poles; there are \( N \) rows of blocks (i.e. blocks between the same parallels), there are \( 2N \) blocks per row, and \( i,j \) identify blocks according to the row and column they are in. Grids for equal angular point data may be "center point" grids, data being measured at the center of each block, so \( i = 0 \) corresponds to \( \theta = \Delta \theta / 2 = \Delta \lambda / 2 \).

No blocks straddle the equator, so regular grids are symmetrical with respect to it; in other words: \( N \) is always even (extension to \( N \) odd is trivial).

Area means are identified here by overbars; they can be expanded in series simply by integrating (1.1) term by term, which can be done because the spherical harmonic series is always uniformly convergent for \( 0 \leq \theta \leq \pi \) and \( 0 \leq \lambda \leq 2\pi \):

\[
\overline{f}_{ij} = \frac{1}{\Delta_{ij}} \sum_{s=0}^{\infty} \sum_{z=0}^{n} \sum_{\alpha=0}^{1} C_{sn}^\alpha \int_{\sigma_{ij}} \overline{Y}_{zn}^\alpha(\theta, \lambda) \, d\sigma
\]

\[
= \frac{1}{\Delta_{ij}} \sum_{s=0}^{\infty} \sum_{z=0}^{n} \sum_{\alpha=0}^{1} \int_{\theta_{ij}}^{\theta_{ij} + \Delta \theta} \int_{\lambda_{ij}}^{\lambda_{ij} + \Delta \lambda} \overline{P}_{zn}(\cos \theta) \sin \theta \, d\theta \, d\lambda \left[ C_{sn}^\alpha \right]_{\lambda_{ij}^{\Delta \lambda}}^{\lambda_{ij}^{\Delta \lambda}} \cos m \lambda \\
+ \overline{\xi}_{zn} \int_{\lambda_{ij}^{\Delta \lambda}}^{\lambda_{ij}^{\Delta \lambda}} \sin m \lambda \, d\lambda \]  \quad (1.2)

Here \( \overline{f}_{ij} \) is the area mean of \( f(\theta, \lambda) \) on the block \( \sigma_{ij} \) whose area is \( \Delta_{ij} = \Delta \lambda (\cos \theta_{ij} - \cos (\theta_{ij} + \Delta \theta)) \).

A basic property of spherical harmonics is their orthogonality:

\[
\frac{1}{4\pi} \int_{\sigma} \overline{Y}_{zn}^\alpha(\theta, \lambda) \overline{Y}_{zm}^\beta(\theta, \lambda) \, d\sigma = \begin{cases} 1 & \text{if } \alpha = \beta, \quad m = p \quad \text{and} \quad n = k \\ 0 & \text{otherwise} \end{cases} \quad (1.3)
\]

as a consequence of which

\[
\overline{c}_{zn}^\alpha = \frac{1}{4\pi} \int_{\sigma} \overline{Y}_{zn}^\alpha(\theta, \lambda) f(\theta, \lambda) \, d\sigma \quad (1.4)
\]
Expression (1.4) is the inverse of (1.1); both constitute the basis of spherical harmonic analysis. In general, (1.4) cannot be calculated analytically, because what is available is not the function \( f(\theta,\lambda) \), but a finite set of noisy measurements in the form of point values \( z_{ij} \) or their area means \( \bar{z}_{ij} \). Discretizing (1.4) on an equal angular grid results, for instance, in the following numerical quadrature formula:

\[
\hat{C}_{n^2} = \frac{1}{4\pi} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \tilde{Y}_{n^2}(\theta_i,\lambda_j) f(\theta_i,\lambda_j) \Delta \theta \Delta \lambda
\]  
(1.5)

where \( \hat{C}_{n^2} \) indicates the estimate of \( \bar{C}_{n^2} \), as this type of formula is usually only an approximation. Formulas resembling (1.5) can be called "point values-type quadrature formulas" and can be handled by algorithms that are all identical from a structural point of view and whose prototype is that of paragraph 1.5.

If the data are block averages, several simple approximations have been proposed that take the form

\[
\hat{C}_{n^2} = \mu_n \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \tilde{f}_{ij} \int_{\sigma_{ij}} \tilde{Y}_{n^2}(\theta, \lambda) \, d\sigma
\]

where the \( \mu_n \) are scale factors. This kind of formula shall be studied further in section 2. Writing the above expression in full

\[
\hat{C}_{n^2} = \mu_n \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \tilde{f}_{ij} \int_{\sigma_{ij}} \tilde{f}(\theta + \Delta \theta) \bar{P}_{n^2}(\cos \theta) \sin \theta \, d\theta \int_{\lambda_j}^{\lambda_j + \Delta \lambda} \cos^m \lambda \, m \, d\lambda
\]  
(1.6)

This belongs to the general type

\[
\hat{C}_{n^2} = K \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \tilde{f}_{ij} \left[ \sum_{m=0}^{m_{\text{max}}} \left( \begin{array}{c} \text{A} \,(m) \end{array} \right) \cos m \lambda \Delta \lambda + \left( \begin{array}{c} \text{B} \,(m) \end{array} \right) \sin m \lambda \Delta \lambda \right]
\]  
(1.7)

where \( \text{B} \,(m) = \left\{ \begin{array}{ll} (\cos m \Delta \lambda - 1)/m & \text{if } m \neq 0 \\ 0 & \text{if } m = 0 \end{array} \right. \)

\( \text{A} \,(m) = \left\{ \begin{array}{ll} (\sin m \Delta \lambda)/m & \text{if } m \neq 0 \\ \Delta \lambda & \text{if } m = 0 \end{array} \right. \)

Formulas resembling (1.6) or (1.7) appear several times in this work, and are called here area means-type quadrature formulas.

If all \( \bar{C}_{n^2} \) with \( 0 \leq n \leq N_{\text{max}} \) are known, they can be used to compute the \( (N_{\text{max}} + 1)^2 \) terms in

\[
f(\theta,\lambda)_{N_{\text{max}}} = \sum_{n=0}^{N_{\text{max}}} \sum_{m=0}^{n} \sum_{\alpha=0}^{1} \bar{C}_{n^2} \hat{Y}_{n^2}(\theta,\lambda)
\]

which can be regarded as an approximation to \( f(\theta,\lambda) \) at the point \((\theta,\lambda)\). Expression (1.8), together with the truncation at \( N_{\text{max}} \) of (1.2) (area means), defines the object of spherical harmonic synthesis: given the coefficients, estimate the function. As shown later, analysis and synthesis are related by a simple duality, so they require about the same number of operations when performed on a certain grid and with a given number of coefficients.
The set of all degree variances
\[ \sigma_n^2 = \sum_{n=0}^{N_{\text{max}}} \overline{C}_{nm}^2 + \overline{S}_{nm}^2 \] (1.9)
constitutes the power spectrum of \( f(\theta, \lambda) \). If all coefficients of degree \( n > N_{\text{max}} \) are zero, so are the corresponding \( \sigma_n \); in such case the function \( f(\theta, \lambda) \) is said to be band limited.

Closely associated with the power spectrum is the isotropic covariance function
\[ \text{cov}(f(P), f(Q)) = \sum_{n=0}^{\infty} \sigma_n^2 P_n(\cos \psi_{PQ}) \] (1.10)
Here \( P \) and \( Q \) are two points on the sphere, \( \psi_{PQ} = \cos^{-1}[\cos \theta_P \cos \theta_Q + \sin \theta_P \sin \theta_Q \cos(\lambda_P - \lambda_Q)] \) is the geocentric angle or spherical distance between both, and \( P_n(\cos \psi) \) is the unnormalized Legendre polynomial of degree \( n \). The power spectrum and the covariance of a function on the sphere, like those of functions defined on the real line, on the plane, and on Euclidean spaces of higher dimensions, can be used to obtain estimators related to discrete Wiener filters and predictors, i.e., satisfying a minimum variance condition. The spherical version of such technique is known as least squares collocation (Moritz, 1972). The inverse of expression (1.10) is
\[ \sigma_n^2 = (2n+1) \int_0^{\pi} P_n(\cos \psi_{PQ}) \text{cov}(f(P), f(Q)) \sin \psi_{PQ} \, d\psi_{PQ} \] (1.11)
Equations (1.10) and (1.11) show that, as usual, power spectrum and covariance function are linear transforms of each other. A formula such as (1.11) is sometimes called a Legendre transform.

In addition to the autocovariance (1.10), one can define, more generally,
\[ \text{cov}(u(P), v(Q)) = \sum_{n=0}^{\infty} c_n^{u,v} P_n(\cos \psi_{PQ}) \] (1.10*)
as the covariance of two functions \( u(\theta_P, \lambda_P) \) and \( v(\theta_Q, \lambda_Q) \) on the sphere, where the
\[ c_n^{u,v} = \sum_{m=0}^{n} \overline{C}_{nm}^{u} \overline{C}_{nm}^{v} + \overline{S}_{nm}^{u} \overline{S}_{nm}^{v} , \quad n = 0, 1, \ldots \] (1.9*)
constitute the "power crossspectrum".

A relationship similar to (1.11) applies to the \( c_n^{u,v} \) and to \( \text{cov}(u(P), v(Q)) \).

A very important property of spherical harmonics is Parseval's theorem:
\[ \frac{1}{4\pi} \int_0^\pi f(\theta, \lambda)^2 \, d\sigma = \sum_{n=0}^{\infty} \sigma_n^2 \] (1.12)
The orthogonality of spherical harmonics, and the fact that they form a complete orthogonal set of functions on the sphere, are among the reasons why they are used so widely in theory and in practice, but there is more to them than orthogonality.

The solid spherical harmonics

\[ \frac{1}{r^{n+1}} Y_{n}^{\alpha}(\theta, \lambda) \quad \text{and} \quad r^{n} Y_{n}^{\alpha}(\theta, \lambda) \]

where \( r \) is the distance to the origin of coordinates, are all solutions of Laplace's equation, which in Cartesian coordinates is

\[ \nabla^{2} W = \frac{\partial^{2} W}{\partial x^{2}} + \frac{\partial^{2} W}{\partial y^{2}} + \frac{\partial^{2} W}{\partial z^{2}} = 0 \]  
(1.13)

This makes them appropriate for the study of harmonic functions, such as the gravitational potential, in spherical coordinates. Another property they have, unique among functions that are orthogonal on the sphere, is the relationship

\[ \frac{1}{4\pi} \int_{\sigma} f(P) \sum_{n=0}^{\infty} \frac{(2n+1)}{4^n} P_{n}(\psi_{Q}) \, d\sigma = \sum_{n=0}^{\infty} \sum_{\alpha=0}^{n} \frac{1}{\alpha} a_{n}^{\alpha} \alpha_{n}^{\alpha} Y_{n}^{\alpha}(Q) \]  
(1.14)

which corresponds to the Convolution Theorem for ordinary Fourier series, and is the basis of such fundamental formulas as Stokes' in gravimetric geodesy.

1.2. Relationship Between Spherical Harmonics and 2-D Fourier Series

As indicated in the previous paragraph, spherical harmonics share important properties with ordinary (trigonometric) Fourier series in one or more dimensions. There is also a very immediate relationship with ordinary two-dimensional (2-D) series that will be explained here. From Hobson (1931, Ch. III, formula (7)) we know that

\[ P_{nm}(\cos \theta) = \frac{(-1)^{n} (2n)!}{2^{n} n! (n-m)!} \sin^{n} \theta \left\{ \cos^{n-m} \theta - \frac{(n-m)(n-m-1)}{2(n-1)} \cos^{n-m-2} \theta + \right. \]

\[ \left. + \frac{(n-m)\ldots(n-m-3)}{2\cdot4\cdot(2n-1)(2n-3)} \cos^{n-m-4} \theta - \ldots \right\} \]  
(1.15)

so the normalized Legendre function is

\[ \overline{P}_{nm}(\cos \theta) = \frac{(-1)^{n} (2n)!}{2^{n} n! (n-m)!} \sqrt{\frac{2(2n+1)(n-m)!}{(m+n)!}} \sin^{n} \theta \sum_{k=0}^{(n+1)} \frac{L_{n}^{(2k+1)}}{a_{n}(n,m) \cos^{n-m-2k} \theta} \]

where

\[ L_{n}(n,m) = \begin{cases} (n-m)/2 & \text{if } n-m \text{ is even} \\ (n-m-1)/2 & \text{if } n-m \text{ is odd} \end{cases} \]

while

\[ a_{k}(n,m) = (-1)^{k} (n-m)(n-m-1)\ldots(n-m-2k+1)[2\cdot4\ldots2k(2n-1)\ldots(2n-2k+1)]^{2} \]

for \( k > 0 \).
In the interval $-\pi \leq \theta \leq \pi$ \(\sin^m \theta\) is even when \(m\) is even, and odd when \(m\) is odd. In the interval $-\pi \leq \theta \leq \pi$ the sum
\[
\sum_{k=0}^{\lfloor \frac{m}{2} \rfloor} a_k (n, m) \cos^{m-2k} \theta
\]
is always even (it is a sum of even powers of \(\cos \theta\)), so the parity of
\(P_n (\cos \theta)\) is the same as that of \(m\) if $-\pi \leq \theta \leq \pi$. An even function can be expanded into a sum of \textbf{cosines}, and an odd function into a sum of \textbf{sines}. The highest frequency term will correspond to the highest frequency in the expansion of \(\sin^m \theta \cos^{m-n} \theta\), so this term will be of the form \(a_n \cos \theta\) or \(b_n \sin \theta\). Therefore, the Legendre function satisfies one of the following equations in $-\pi \leq \theta \leq \pi$:

\[
\overline{P}_n (\cos \theta) = \sum_{t=0}^{n} C_t^n \cos t \theta
\]
(a) if \(m\) is even;

\[
\overline{P}_n (\cos \theta) = \sum_{t=0}^{n} S_t^n \sin t \theta
\]
(b) if \(m\) is odd

A spherical harmonic $Y_n^\alpha (\theta, \lambda)$ can have one of four possible forms:

for \(m\) even:
\[
\overline{Y}_n^\alpha (\theta, \lambda) = \begin{cases} 
\sum_{t=0}^{n} C_t^n \cos t \theta \cos m \lambda & \text{for } \alpha = 0 \\
\sum_{t=0}^{n} C_t^n \cos t \theta \sin m \lambda & \text{for } \alpha = 1
\end{cases}
\]

(1.16a)

for \(m\) odd:
\[
\overline{Y}_n^\alpha (\theta, \lambda) = \begin{cases} 
\sum_{t=0}^{n} C_t^n \sin t \theta \cos m \lambda & \text{for } \alpha = 0 \\
\sum_{t=0}^{n} C_t^n \sin t \theta \sin m \lambda & \text{for } \alpha = 1
\end{cases}
\]

(1.16b)

A sum of spherical harmonics such as (1.8) is equivalent to a sum of terms of the form
\(C_t^n \sin t \theta \cos m \lambda, C_t^n \sin t \theta \sin m \lambda, C_t^n \cos t \theta \cos m \lambda\) and
\(C_t^n \cos t \theta \sin m \lambda\), which are also the basic functions of 2-D Fourier series.

The highest \(m\) and \(n\) in the spherical harmonic expansion
\[
f (\theta, \lambda) = \sum_{n=0}^{N_{\text{max}}} \sum_{\alpha=0}^{\lfloor \frac{n}{2} \rfloor} \sum_{m=0}^{n-\alpha} -\overline{\alpha} \overline{Y}_{n}^{\alpha} (\theta, \lambda)
\]
are equal to \(N_{\text{max}}\), so the highest degree and order (or spatial frequencies) in the Fourier series are also equal to \(N_{\text{max}}\). In conclusion: every surface spherical harmonic expansion where the highest degree is \(N_{\text{max}}\) is identical to a 2-D Fourier series (where the highest \(m\) and \(n\) are also \(N_{\text{max}}\)), in the domain $-\pi \leq \theta \leq \pi$, $0 \leq \lambda \leq 2\pi$. The converse is not true, because continuous functions on a sphere, such as the $\overline{Y}_{n}^{\alpha}$, must satisfy certain conditions at the poles that ordinary functions on the $-\pi \leq \theta \leq \pi$, $0 \leq \lambda \leq 2\pi$ domain do not have to. Spherical harmonics correspond to a subclass (linear subspace) of 2-D Fourier series.
For example: (Heiskanen and Moritz, Chapter 1, 1967)

\[ \tilde{P}_{11} (\cos \theta) = \sqrt{3} \sin \theta \]

\[ \tilde{P}_{31} (\cos \theta) = \sqrt{30} \sin \theta \cos \theta = \frac{1}{2} \sqrt{30} \sin 2\theta \]

So

\[ \tilde{Y}_{11}^0 = \sqrt{3} \sin \theta \cos \lambda, \quad \tilde{Y}_{11}^1 = \sqrt{3} \sin \theta \sin \lambda \]

\[ \tilde{Y}_{11}^0 = \frac{1}{2} \sqrt{30} \sin 2\theta \cos 2\lambda, \quad \tilde{Y}_{11}^1 = \frac{1}{2} \sqrt{30} \sin 2\theta \sin 2\lambda \]

Calling the 2-D Fourier coefficients "\( \alpha_{p,n}^\beta \)" where

\( \alpha_{p,n}^0 \) correspond to terms of the form \( \cos p \theta \cos m \lambda \)

\( \alpha_{p,n}^1 \) correspond to terms of the form \( \cos p \theta \sin m \lambda \)

\( \alpha_{p,n}^2 \) correspond to terms of the form \( \sin p \theta \cos m \lambda \)

\( \alpha_{p,n}^3 \) correspond to terms of the form \( \sin p \theta \sin m \lambda \)

these can be related to the respective \( \tilde{C}_{p,n}^\alpha \) by expressions of the form

\[
\tilde{C}_{p,n}^\alpha = \frac{1}{4} \sqrt{\frac{2(2n+1)(n-m)!}{(n+m)!}} \sum_{p=0}^{n-1} \int_{0}^{\pi} \cos p\theta P_n(\cos \theta) \sin \theta \ d\theta \ a_{p,n}^\alpha \quad n = m, m + 1, m + 2, \ldots \]

(1.17)

where \( \beta = \alpha \) if \( m \) is even, and \( \beta = \alpha + 2 \) if \( m \) is odd. The \( I_{n,p}^\beta \) are defined as

\[
I_{n,p}^\beta = \int_{0}^{\pi} \cos p\theta \cos \theta P_n(\cos \theta) \sin \theta \ d\theta \quad \text{if } m \text{ is even}
\]

\[
I_{n,p}^\beta = \int_{0}^{\pi} \sin p\theta \cos \theta P_n(\cos \theta) \sin \theta \ d\theta \quad \text{if } m \text{ is odd}
\]

and can be computed recursively using the formula

\[
I_{n,p}^\beta = \frac{2n-1}{2(n-m)} \left\{ I_{n-1,p+1}^\beta + I_{n-1,p-1}^\beta \right\} - \frac{n+m-1}{n-m} I_{n-2,p}^\beta \quad (1.18a)
\]

with the following starting values

\[
I_{n,p}^\beta = \begin{cases} 
0 & \text{if } (m+p) \text{ is odd} \\
\frac{2(m+1)(2m)!}{2^m \left[ (m+1)^2 - p^2 \right] \left[ (m-1)^2 - p^2 \right] \cdots \left[ 3^2 - p^2 \right] \left[ 1^2 - p^2 \right]} & \text{if } m \text{ is even, } p \text{ even,} \\
\frac{2^m \left[ (m+1)^2 - p^2 \right] \left[ (m-1)^2 - p^2 \right] \cdots \left[ 2^2 - p^2 \right] \left[ -p \right]} {2^{m+1}(2m)!} & \text{if } m \text{ is odd, } p \text{ odd.}
\end{cases}
\]

(1.18b)

The \( I_{n,p}^\beta \) are zero for alternate values of both \( p \) and \( m \). These equations were reported by Ricardi and Burrows (1972), and show how to obtain the
Once the $a_{n}^{\theta}$ have been computed from the data by means of the 2-D discrete Fourier transform, normally this would be impossible, because the data only exists in the upper half of the $-\pi \leq \theta \leq \pi$, $0 \leq \lambda \leq 2\pi$ interval, and the 2-D algorithm requires information on all of it. However, the fact that the $P_{n}^{\lambda} (\cos \theta)$ are sums of sines only or cosines only makes the calculation possible.

While the maximum degree and order in a 2-D Fourier series do not reach the Nyquist frequency\(^1\) ($m < N = \frac{\pi}{\Delta \lambda}$) in an equal angular grid, all the coefficients can be recovered exactly by solving $(2N)^2$ equations such as

$$f (\theta, \lambda) = \sum_{n=-N}^{N} \sum_{s=-N}^{N} \sum_{\theta=0}^{\pi} a_{ns}^{\theta} \{ \cos \} n \sin \{ \cos \} m \lambda$$  \hspace{1cm} (1.19)

When $n$ or $m$ exceed $N_{\text{max}}$, the matrix of the system of equations becomes singular, and the discrete Fourier transform consists of coefficients that "fit" the data, but differ from the true coefficients. The estimated coefficients are said to have been aliased with those that exceed the Nyquist frequency. In the case of spherical harmonics, which are a special case of 2-D Fourier series, a similar situation must arise: the harmonics in the data with $n \geq N$ are going to be aliased with those of lower degree, so the information available is not enough to recover all coefficients because the sampling is too coarse.

The aliasing of spherical harmonics sampled on regular grids is a consequence of the aliasing of the respective 2-D Fourier series, so it makes sense to talk of a "Nyquist frequency" in the case of those functions. Having established the connection between aliasing in both types of series, it is time to point out also some important differences.

### 1.3 Sampling Errors

Expressions (1.16a-b) shows that spherical harmonics are finite sums of 2-D Fourier harmonics, which is not the same as being each a Fourier harmonic. From this simple fact follow some important distinctions.

To understand them better, let us begin by stating some basic properties of Fourier series in one dimension, which carry over to higher dimensions but are easier to explain in one dimension.

If sampled at a constant interval $\Delta \lambda = \frac{\pi}{N}$, the following is always true of sines and of cosines:

---

\(^1\)Named after the Nyquist Theorem: the Fourier coefficients of a function of period $2N\Delta \lambda$ can be recovered only if $N_{\text{max}} < N$. 

---
\[
\sum_{k=0}^{2N-1} \cos mk\Delta \lambda \cos pk\Delta \lambda = 0 \quad \text{if m} \neq p < N
\]
\[
\sum_{k=0}^{2N-1} \sin mk\Delta \lambda \sin pk\Delta \lambda = 0 \quad \text{if m} \neq p < N
\]
\[
\sum_{k=0}^{2N-1} \cos mk\Delta \lambda \sin pk\Delta \lambda = 0 \quad \text{for all m and all p.}
\]

These expressions are discrete counterparts of
\[
\int_0^{2\pi} \cos m\lambda \cos p\lambda \, d\lambda = \int_0^{2\pi} \sin m\lambda \sin p\lambda \, d\lambda = 0 \quad \text{when m} \neq p
\]
\[
\int_0^{2\pi} \cos m\lambda \sin p\lambda \, d\lambda = 0 \quad \text{for all m and p.}
\]

and show that the orthogonal properties of sines and cosines are maintained when these are sampled regularly, provided the Nyquist frequency is not exceeded. From this follows that
\[
a^m_n = H \sum_{k=0}^{2N-1} (\cos) m k \Delta \lambda f(k\Delta \lambda) = h \int_0^{2\pi} (\cos) m \lambda f(\lambda) \, d\lambda
\]

where \( H = \begin{cases} \frac{1}{2N} & \text{if m = 0} \\ \frac{1}{N} & \text{otherwise} \end{cases} \) and \( h = \begin{cases} \frac{i}{2\pi} & \text{if m = 0} \\ \frac{1}{2\pi} & \text{otherwise} \end{cases} \), so the "Fourier" counterpart of (1.5), (i.e. (1.21)) is an exact "numerical quadratures" formula for Fourier series. When the Nyquist frequency is exceeded, the trigonometric relationships
\[
\cos mk\Delta \lambda = \cos (2N \frac{m}{2} \pm m) k \Delta \lambda
\]
\[
\sin mk\Delta \lambda = \sin (2N \frac{m}{2} \pm m) k \Delta \lambda
\]

imply that (1.21) will give, not the true coefficients, but the aliased estimates
\[
\begin{align*}
a^0_n & = a^0_n + \sum_{h=0}^{K} a^0_{2hN+n} + \sum_{h=1}^{K} a^0_{2hN-n} \\
\tilde{a}^1_n & = a^1_n + \sum_{h=0}^{K} a^1_{2hN+n} - \sum_{h=1}^{K} a^1_{2hN-n}
\end{align*}
\]

with \( K \) such that \( 2KN \) is below, and \( 2(K+1)N \) is above the highest frequency present in \( f(\lambda) \). Expression (1.20 a-b), (1.21), and (1.22) are the foundations of discrete Fourier analysis (also known as the computation of the discrete Fourier transform, or D.F.T.), and so well known that they are almost second nature to many engineers and scientists. Unfortunately, none of these discrete formulas has exact counterparts in spherical harmonic analysis, and this fact has been the cause of considerable confusion. The most common
misunderstandings seem to occur in the following areas:

(a) **Number of equations and number of unknowns:**

At each point in the grid it is possible to write an equation such as

\[ f(\theta, \lambda) = \sum_{n=0}^{N} \sum_{m=0}^{M} \sum_{\alpha=0}^{\infty} \bar{C}_{n,m} \bar{Y}_{n,m}(\theta, \lambda) \]  

(1.23)

similar to (1.19). The maximum number of coefficients that can be recovered without exceeding the Nyquist frequency (i.e., those with \( n < N \)) is \( N^2 \). Coefficients of degree or order equal or greater than \( N \) shall not be, in general, free from aliasing. Since there are \( 2N^2 \) points in an equiangular grid, it follows that the maximum number of fully recoverable coefficients is also half the number of points (equations) in the grid. By contrast, in 2-D Fourier analysis the number of coefficients equals the number of points in the grid (in the interval \(-\pi \leq \theta \leq \pi, 0 \leq \lambda \leq 2\pi\)). One could have a grid with only \( N^2 \) points, as proposed by Giacaglia and Lundquist (1972), but such a grid would not be equal angular on the sphere.

The author had discussed this problem elsewhere (Colombo, 1979a, paragraph (5.2)), showing that the system of equations (1.23) becomes singular when all coefficients of degree and order \( 0 \leq (n, m) \leq M \) are included among the unknowns, and \( M \geq N \). In other words: it is not possible to solve for a complete set of coefficients to degree and order \( M \geq N \). The relevant part of that argument can be summarized as follows: the columns of \( \bar{A} \), the matrix of the system of equations (1.23), consist of successions of values \( \bar{Y}_{n,m}(\theta, \lambda) \) of the harmonics corresponding to the unknowns \( \bar{C}_{n,m} \) at the points \( (\theta, \lambda) \) in the grid. The scalar product of two such columns is,

\[
\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \bar{Y}_{n,m} \bar{Y}_{n,q} = \sum_{i=0}^{N} \bar{P}_{n,q} (\cos \theta) \bar{P}_{p,q} (\cos \theta)
\]

\[
x \sum_{i=0}^{N-1} \{ \cos \theta \} \{ \cos \theta \} m j \Delta \lambda = 0 \quad \text{if} \quad \alpha \neq \beta
\]

\[
x \sum_{i=0}^{N-1} \{ \sin \theta \} \{ \sin \theta \} q j \Delta \lambda = 0 \quad \text{if} \quad m \neq q
\]

\[\neq 0 \quad \text{otherwise}\]

according to (1.20-a). Therefore, if two columns correspond to unknowns of different orders \( m \) and \( q \), they must be orthogonal and, thus, independent. For the whole matrix not to be singular, all columns of the same order \( m \) must form sub-matrices \( A(m) \) that have full rank. Otherwise, there will be columns in those \( A(m) \), and consequently in \( A \), that are linearly dependent, so \( A \) cannot be inverted. Consider \( A(0) \), corresponding to all unknowns of order 0. This is a \( 2N^2 \times (M + 1) \) matrix, and the elements of the columns of \( A(0) \) have the same values as \( \bar{P}_{n,q}(\cos \theta) \), \( 0 \leq n \leq M, 0 \leq i \leq N-1 \). The \( \bar{P}_{n,q} \) are functions of \( \theta \) only, and there are \( N \) parallels in the grid, so there are \( Nc \) more than \( N \) independent rows in \( A(0) \). Because the \( \bar{P}_{n,q}(\cos \theta) \) are polynomials of degree \( n \) in \( \cos \theta \), these rows form a sub-matrix \( S(0) \) of \( A(0) \) that has \( M + 1 \) independent columns,
as long as \( M + 1 \leq N \) or \( M < N \). If \( M \geq N \) the extra columns will turn \( S(0) \) into a rectangular matrix with more columns than rows; in other words: there will be no square submatrix in \( A(0) \) of rank \( M + 1 \) if \( M \geq N \), so \( A(0) \) shall be rank deficient, and from this follows that \( A \) must be singular.

Emphasis must be placed on the word complete when referring to the set of "solvable" coefficients: it is possible, by removing some coefficients with \( n < N \) from the unknowns, to introduce others in their stead with \( n \geq N \), but then the solution will not be a complete set of coefficients.

(b) 100% aliasing:

If a term in a Fourier series has a frequency \( n > N \), then it will be aliased with lower frequency terms and become impossible to discriminate from them. For most functions of practical interest, the higher the frequency, the smaller the term, so the coefficient estimated using the sum in (1.21) will be dominated by the lower frequency terms: the estimation error is thus likely to exceed 100%. Estimates above the Nyquist frequency are usually regarded as meaningless and the closer a term is to that frequency with increasing \( n \), the less reliance is placed on its estimate.

In the case of spherical harmonics, expressions (1.16a-b) show clearly that the harmonic \( Y^{\alpha}_{n \beta} \) consists of several Fourier terms of frequencies ranging from 0 to \( n \). When \( n > N \), only that part of the Fourier expansion of \( Y^{\alpha}_{n \beta} \) above the Nyquist frequency will become scrambled beyond recovery; part of the harmonic is left intact: the low frequency "tail", which means that the effect of aliasing on the recovered coefficients does not necessarily reach 100\% (or even 70\%) at the Nyquist frequency, as shown in the examples of section 3.

(c) Orthogonality

From (1.20a) follows that the matrix of equations (1.19) for the Fourier series is orthogonal, so the coefficients estimated according to (1.21) are independent from each other.

In the case of the \( \bar{Y}^{\alpha}_{n \beta} \), orthogonality does not carry over to all the sampled harmonics, unless special "quadratures' weights" are introduced in (1.5) or (1.6). This lack of orthogonality affects, for instance, the formulas for mean values discussed in section 2. The method of Gaussian quadratures is an example of "quadratures with weights" that gives exact coefficients when the Nyquist frequency is not exceeded, though it requires a special grid where the parallels are situated at the same latitudes as the zeroes of \( P_{n \beta} \). The use of this method is possible because the product \( [ \tilde{P}_{n \beta} (\cos \theta) \tilde{P}_{p \gamma} (\cos \theta) ] \) is a polynomial in \( \cos \theta \) of degree \( n + p \leq 2N \); the grid, however, is an unusual one. Details of the application of Gaussian quadratures to spherical harmonic analysis are given in a report by Payne.
(1971). Other examples of methods that recover the coefficients exactly when the data is noiseless and \( N_{s x} < N \) are: least squares collocation, least squares adjustment, and the algorithm developed by Rice and Burrows from expressions (1.17) and (1.18 a-b).

In general, not all the discretised harmonics retain the orthogonality property, and the estimated coefficients are affected by the values of many of the other \( C_{n}^{\alpha} \), in addition to those whose degrees exceed the Nyquist frequency. More about this will be said in section 2, when discussing least squares collocation and adjustment.

Summarising: there are enough differences between the aliasing of Fourier series and that of spherical harmonics, in spite of their being so closely related, to require a great deal of caution before using the intuition gained from one type of analysis when attempting the other. For this reason, the expression "aliasing error" will be replaced by the just as appropriate "sampling error", which is perhaps less charged with misleading connotations because it has not been applied almost exclusively to Fourier series.

It should be noticed here that Gaposchkin (1980) has published formulas for the sampling error on a type of equal area grid (i.e., all blocks have the same area as the equatorial ones). His formulas are the equivalent, for such a grid, of expression (1.22) for Fourier analysis, but much more complicated; they are made tractable numerically by the use of certain recursive expressions that he provides, thus showing an interesting new approach to the study of the problem.

1.4 Number of Operations in Analysis and in Synthesis

Expressions (1.17) and (1.18 a-b) can be used to calculate the \( C_{n}^{\alpha} \) once the 2-D Fourier coefficients of \( f(\theta_{1}, \lambda_{1}) \), or \( a_{\lambda_{1}}^{\alpha} \), have been estimated by means of the 2-D Discrete Fourier Transform (DFT). Using the Fast Fourier Transform (FFT) algorithm (Cooley and Tuckey, 1965), the number of operations required\(^1\) is proportional to the number of data points, which is the order of \( N^{3} \), or \( O(N^{3}) \) for short. The FFT is discussed further in paragraph 1.9.

Having obtained the \( a_{\lambda_{1}}^{\alpha} \), calculation of (1.17) requires \( N \) operations per \( C_{n}^{\alpha} \), or \( N x N^{2} = N^{3} \) for all of them. Finding the coefficients

\(^{1}\) Most of these calculations have the form of scalar products

\[ p = \sum_{k=0}^{p-1} a_{k} b_{k}, \]  

so the basic operation of finding \( p^{k} = a_{k} b_{k} + p^{k-1} \) \((p^{0} = 0, p^{n-1} = p)\)

consists of one sum and one product.
by means of the recursives (1.18 a-b) adds another \( O(N^3) \) operations that can be obviated by computing the \( R_{sp}^a \) once, and then storing them on magnetic tape or disk. One way or the other, \( O(N^2) + O(N^3) \) operations are needed altogether, or \( O(N^3) \) when \( N \) is large (say, \( N \geq 180 \)).

As already mentioned, this procedure was first described by Ricardi and Burrows in 1972; more recently (1977) Goldstein developed a very similar idea and formulated a similar algorithm for synthesis. Goldstein's method uses recursive formulas for the \( R_{sp}^a \) different from (1.18 a-b). The synthesis algorithm also requires \( O(N^3) \) operations.

As explained in paragraphs 1.5 through 1.7, the procedures presented there also require \( O(N^3) \) operations for analysis, and as many for synthesis, though they are formally different from Ricardi and Burrows'. In paragraph 1.8 it will be shown that synthesis requires as many operations as analysis, because one is the dual of the other.

The fact that two rather different approaches (Ricardi and Burrows' and the one described in this report) require essentially the same number of operations suggests that "\( O(N^3) \)" might be a property of all analysis and synthesis algorithms on regular spherical grids due, somehow, to the nature of the sphere itself. This is speculation, of course, but if not, are there other ways of partitioning the sphere for which faster methods exist? The author has discussed this possibility before (Colombo, 1979, paragraph 4.6). It is interesting to notice that in all these procedures the \( O(N^3) \) operations are those associated with \( \theta_i \), or "column operations"; "row operations" are only \( O(N) \). In the case of the Euclidean plane, the 2-D Fourier transform requires the same number of operations per row than per column, \( O(N) \), thus the total is only \( O(N^2) \), or \( O(N) \) times faster than its spherical "counter-part."

While not as efficient as the 2-D FFT, the algorithms for the sphere considered here can be much faster than the straightforward implementation of expressions (1.5), (1.6), (1.2), or (1.8). The latter has been the approach of many scientists who have developed their own software, but whose main interest has generally been far removed from the study of numerical techniques. In 1976, while working at the University of New South Wales (Australia), the author developed the two algorithms of paragraphs 1.5 and 1.6, and C. Rizos programmed them. Subsequently they were used at Goddard Space Flight Center, in Maryland. To everybody's surprise, Rizo's programs turned out to be more than 100 times faster than those in use at the time, when run under the same conditions. More recently, this author has written the subroutines HARMIN and SSYNTH described in appendix B. SSYNTH has been used, after the fashion of the numerical experiments described in Section 2, to generate 64000 \( 1^\circ \times 1^\circ \) mean values (simulated averaged gravity anomalies), each the sum of the 90000 terms of an expansion.
complete to degree and order 300. This took less than 50 central processor
unit seconds in the AMDHAL 470 V/6-II owned by The Ohio State University
(OSU). All calculations were in double precision (32 bits words), using
FORTRAN H EXTENDED. Precomputed values of the Legendre functions
(or their integrals) were read from tape, and all operations involving trigo-
nometric functions were carried out by a fast Fourier Transform subroutine.
These two characteristics, plus a generally tighter coding, are the reasons
for the greater speed of this program, compared to the older versions men-
tioned before.

In all these methods all operations along a given row or parallel (con-
stant \( \theta_i \)) are independent from those for any other row, so a parallel processing
computer with \( N \) processors (arithmetic and control units) could analyse or
synthesize a full grid of \( N \) rows as fast as an ordinary computer with one
central processor can do a single row. This \( N \)-fold increase in speed can be
obtained with the same type of basic hardware (gates, registers) that is
currently used in conventional "general purpose" main-frame machines. The
full power of the algorithms presented in this work will be realized when
computers of parallel structure become more widely available for scientific
applications than they are today.

1.5 Algorithm for the Analysis of Point Values

Expression (1.5) written in full becomes

\[
\alpha_{n}^n = \frac{1}{4\pi} \sum_{i=0}^{N-1} \sum_{j=0}^{2n-1} \bar{P}_{n} (\cos \theta_i) \left\{ \cos \left( \sin \left( mj \Delta \lambda \right) f(\theta_i, \lambda_j) \right) \Delta \lambda \right\}
\]

which corresponds to the general type

\[
\alpha_{n}^n = \sum_{i=0}^{N-1} \sum_{j=0}^{2n-1} \chi_{i}^{m} \sin \left( \sin \left( mj \Delta \lambda \right) f(\theta_i, \lambda_j) \right) \Delta \lambda \quad (1.24)
\]

where \( \chi_{i}^{m} \) could be \( \frac{1}{4\pi} \bar{P}_{n} (\cos \theta_i) \Delta \lambda \) as above, or \( \bar{P}_{n} (\cos \theta_i) \omega_i \) in
the case of quadrature with weights \( \omega_i \), etc.

To simplify the discussion, the grid is supposed to be equal angular
and \( N_{zax} = N-1 \). This and the algorithms that follow can be easily adapted
for the equal area grids current today. Subroutines HARMIN AND SSYNTH
(Appendix B) can handle the cases \( N_{zax} < N-1 \) and \( N_{zax} > N-1 \) as well as
\( N_{zax} = N-1 \).

The equal angular grid is symmetrical with respect to the Equator,
and assuming that (the same as \( \bar{P}_{n} (\cos \theta_i) \) or \( \bar{P}_{n} (\cos \theta_i) \sin \theta_i \))
\( \chi_{i}^{m} = \chi_{N-2-i} \) if \( n-m \) is even, and \( \chi_{i}^{m} = -\chi_{N-1-i} \) if \( n-m \) is odd, one
can write
\[ C_{n_m}^{n_n} = \sum_{i=0}^{2^{n_n-1}} \left( \chi_i^{n_n} \left[ \sum_{j=0}^{2^{n_n-1}} \frac{\cos}{\sin} mj \Delta \lambda f(\theta_j, \lambda_i) \right] + (-1)^{n_m} \chi_i^{\bar{n_m}} \left[ \sum_{j=0}^{2^{n_n-1}} \frac{\cos}{\sin} mj \Delta \lambda f(\theta_{n-1-i}, \lambda_i) \right] \right) \]  

(1.25)

This formula suggests the following procedure in two nested loops:

**START:** set \( i = -1, \ C_{n_m}^{n_m} = 0 \) for all \( 0 \leq n, m \leq N; \)

**Outer loop:**
(a) increment \( i \) by 1 unless \( i = \frac{1}{2} N - 1 \), in which case **STOP**

**Inner loop:**
(b) compute all
\[
\begin{align*}
\begin{bmatrix}
a_i^{n_n} \\
b_i^{n_n}
\end{bmatrix} &= \sum_{j=0}^{2^{n_n-1}} \frac{\cos}{\sin} mj \Delta \lambda f(\theta_j, \lambda_i) \\
\begin{bmatrix}
a_{n-1-i}^{n_n} \\
b_{n-1-i}^{n_n}
\end{bmatrix} &= \sum_{j=0}^{2^{n_n-1}} \frac{\cos}{\sin} mj \Delta \lambda f(\theta_{n-1-i}, \lambda_i)
\end{align*}
\]
(c) find all
\[
C_{n_m}^{n_n} = C_{n_m}^{n_n(1)} + K \left[ \begin{bmatrix} a_i^{n_n} \\ b_i^{n_n} \end{bmatrix} + (-1)^{n_m} \begin{bmatrix} a_{n-1-i}^{n_n} \\ b_{n-1-i}^{n_n} \end{bmatrix} \right] \chi_i^{n_n}
\]

(1.26)

for \( 0 \leq n, m \leq N \) (where "\((-1)^{n_m} \{ \} \)" merely indicates that \( \{ \} \) is to be added or substracted according to the parity of \( n-m \)); **GO BACK TO** (a).

At the end of the outer loop, \( C_{n_m}^{n_n(\frac{1}{2}N-1)} = C_{n_m}^{n_n}. \)

The \( a_i^{n_n} \) and \( b_i^{n_n} \) in (1.26) can be computed by taking the Fourier transform along row \( i \) and row \( N-1-i \) of the values of \( f(\theta_j, \lambda) \). This involves \( O(N) \) operations. There are half as many \( \chi_i^{n_n} \) as \( \chi_i^{n_n(1)} \), therefore it takes \( 2(N+1)^2 \) products, and just as many sums, to form the \( C_{n_m}^{n_n(1)} \) for the pair of rows \( i \) and \( N-1-i \). Consequently, there are \( O(N) + O(N^2) \) operations per pair of rows or \( O(N^2) + O(N^3) \) for the grid as a whole. This is the same as with the Ricardi and Burrows' algorithm, quickly approaching \( O(N^3) \) as \( N \) increases.

Subroutine HARMIN (Appendix B) implements this technique.
1.6 Algorithms for the Synthesis of Point Values

Expression (1.8) is of the general form

\[ f(\theta_i, \lambda_j) = \sum_{n=0}^{N-1} \sum_{s=0}^{n} \chi_i^n \left[ C_{a_s} \cos m\lambda_j + S_{a_s} \sin m\lambda_j \right] \]  \hspace{1cm} (1.27)

where \( \chi_i^n \) can be, for instance, \( K \bar{P}_{mm} \cos \theta_d \), etc., with \( K \) being a proportionality constant. Rearranging terms and considering the parity of \( \chi_i^n \) leads to

\[
\begin{bmatrix}
  f(\theta_i, \lambda_j) \\
  f(\theta_{N-1-i}, \lambda_j)
\end{bmatrix} = \sum_{a=0}^{N-1} \left( \sum_{n=0}^{N-1} \chi_i^n (-1)^{n-a} \chi_i^{n-a} \begin{bmatrix} C_{a_s} \\ S_{a_s} \end{bmatrix} \right) \cos m j \Delta \lambda + \\
\begin{bmatrix}
  \sum_{n=0}^{N-1} \chi_i^n (-1)^{n-a} \chi_i^{n-a} \end{bmatrix} \begin{bmatrix} C_{a_s} \\ S_{a_s} \end{bmatrix} \sin m j \Delta \lambda
\]  \hspace{1cm} (1.28)

which suggests a procedure in two nested loops:

START: set \( i = -1 \)

Outer loop:
(a) increment \( i \) by 1, unless \( i = \frac{1}{2} N - 1 \), in which case STOP

Inner loop:
(b) compute all

\[
\begin{bmatrix}
  \alpha_i^n \\
  \beta_i^n
\end{bmatrix} = \sum_{s=n}^{N-1} \chi_i^s \begin{bmatrix} C_{a_s} \\ S_{a_s} \end{bmatrix}
\]

\[
\begin{bmatrix}
  \alpha_i^{N-1-i} \\
  \beta_i^{N-1-i}
\end{bmatrix} = \sum_{n=i}^{N-1} \chi_i^n (-1)^{n-a} \begin{bmatrix} C_{a_s} \\ S_{a_s} \end{bmatrix}
\]

for \( 0 \leq m \leq N \);

(c) find all

\[
\begin{bmatrix}
  f(\theta_i, \lambda_j) \\
  f(\theta_{N-1-i}, \lambda_j)
\end{bmatrix} = K \sum_{a=0}^{N-1} \begin{bmatrix} \alpha_i^a \\ \beta_i^a \end{bmatrix} \cos m j \Delta \lambda + \\
\begin{bmatrix}
  \cos m j \Delta \lambda \\
  \sin m j \Delta \lambda
\end{bmatrix}
\]  \hspace{1cm} (1.29)

for \( 0 \leq j \leq 2N - 1 \) (where \((-1)^{n-a} \{ -\} \) means the same as in the previous paragraph); GO BACK TO (a).

At the end of the outer loop all \( f(\theta_i, \lambda_j) \) in the grid are known.
Expression (1.29) is computed by applying the FFT to the \(2N\) \(\alpha^i_n, \beta^i_n\) (per row) and the \(2N\) \(\alpha^{n-1}_n, \beta^{n-1}_n\), taking \(O(N)\) operations. The first part of the inner loop (forming the \(\alpha^i_n, \beta^i_n\) and \(\alpha^{n-1}_n, \beta^{n-1}_n\)) involves \(O(N^3)\) operations; for all \(\frac{1}{2}N\) pairs of rows the total is \(O(N^3) + O(N^3)\), and this tends to \(O(N^3)\) as \(N\) increases.

Subroutine SSYNTM (Appendix B) implements this technique.

1.7 Algorithms for the Analysis and Synthesis of Area Means

(I) Analysis

Rearranging (1.7)

\[
\frac{\Delta^i_n}{\Delta_{n-1}} = \sum_{i=0}^{i=n-1} \chi^i_n \begin{bmatrix} A(m)\left(\alpha^i_n + (-1)^{\alpha^i_n-\beta^i_n}\right) + B(m)\left(\beta^i_n + (-1)^{\alpha^i_n-\beta^i_n}\right) \\
A(m)\left(\alpha^i_n - (-1)^{\alpha^i_n-\beta^i_n}\right) + B(m)\left(\beta^i_n - (-1)^{\alpha^i_n-\beta^i_n}\right) \end{bmatrix}
\]

(1.30)

were \(\alpha^i_n, \beta^i_n\) are the same as in paragraph (1.5). A procedure similar to that for the analysis of point values can be obtained directly by replacing the bracket in (1.26) with that in (1.30), and then proceeding as in the algorithm for point values. The total number of operations is, once more, \(O(N^3) + O(N^3)\), or \(O(N^3)\) for large \(N\). Subroutine HARMON also implements this algorithm.

(II) Synthesis

Truncating the series in equation (1.2), replacing

\[
\int_{\theta_1}^{\theta_1 + \Delta \theta} \bar{p}_n (\cos \theta) \sin \theta d \theta
\]

with \(\chi^i_n\), rearranging terms, considering the parity of \(\chi^i_n\), and using \(\alpha^i_n, \beta^i_n\) as defined in paragraph (1.6), leads to the expression

\[
\begin{bmatrix} T_{i,i} \\ T_{n-1,i} \end{bmatrix} = \sum_{n=0}^{n=N-1} \begin{bmatrix} \alpha^i_n \\ \alpha^{n-1}_n \end{bmatrix} A(m) + \begin{bmatrix} \beta^i_n \\ \beta^{n-1}_n \end{bmatrix} B(m) \cos mj\lambda + \begin{bmatrix} \alpha^i_n \\ \alpha^{n-1}_n \end{bmatrix} B(m) + \begin{bmatrix} \beta^i_n \\ \beta^{n-1}_n \end{bmatrix} A(m) \sin mj\lambda
\]

(1.31)

The algorithm for the synthesis of the \(T_{i,i}\) is a direct extension of that for point values. The number of operations, once more, is \(O(N^3) + O(N^3)\) for large \(N\). Subroutine SSYNTM implements this algorithm as well.

1.8 Duality between Analysis and Synthesis

Pairs of direct and inverse linear transforms, such as Fourier transforms, possess dual characteristics: certain words and mathematical
expressions can be arranged in pairs \((a, b)\) such that, if every "a" is replaced by its "b" in any statement or equation valid for a function \(f\), then the modified statement is valid for the transform \(F\) of \(f\), and vice versa.

Analysis and synthesis of spherical harmonics are reciprocal linear transformations of data into coefficients and of coefficients into "data" closely akin to Fourier transforms, so they can be expected to exhibit some dual properties. Comparing the formulas and algorithms in paragraphs (1.5), (1.6) and (1.7) shows many similarities, among them the number of operations. This can be understood as being a consequence of duality. To make this point clear, consider the following pairs:

\[
\begin{pmatrix}
  f(\theta_i, \lambda_i) \\
  f(\theta_{n-1,i}, \lambda_i)
\end{pmatrix}
\begin{pmatrix}
  C_{n,m} \\
  S_{n,m}
\end{pmatrix}
\quad \begin{pmatrix}
  A_{n,m} \\
  B_{n,m}
\end{pmatrix}
\]

\(\chi_i^n (-1)^{-n}, \sin m\Delta\lambda)\); \((i, m); (j, n); (\chi_j^n, \cos m\Delta\lambda)\)

\[
\sum_{i=0}^{2^{N-1}} \sum_{m=0}^{N-1} \sum_{j=0}^{N-1} \sum_{n=0}^{N-1}
\]

From these we can derive the following pairs:

\[
\begin{pmatrix}
  a_i^n \\
  a_{n-1,i}
\end{pmatrix}
\begin{pmatrix}
  \alpha_i^n \\
  \beta_i^n
\end{pmatrix}
\quad \begin{pmatrix}
  b_i^n \\
  b_{n-1,i}
\end{pmatrix}
\begin{pmatrix}
  \beta_i^n \\
  \alpha_i^n
\end{pmatrix}
\]

\[
\begin{pmatrix}
  A(m) \\
  B(m)
\end{pmatrix}
\begin{pmatrix}
  \alpha_i^n \\
  \beta_i^n
\end{pmatrix}
\quad \begin{pmatrix}
  B(m) \\
  A(m)
\end{pmatrix}
\begin{pmatrix}
  \beta_i^n \\
  \alpha_i^n
\end{pmatrix}
\]

\[
\begin{pmatrix}
  a_i^n, A(m); (b_i^n, B(m); (a_{n-1,i}, -B(m); (b_{n-1,i}, A(m)); etc.,
\end{pmatrix}
\]

and, in conclusion:

"ANALYSIS", "SYNTHESIS"

Each one of the analysis algorithms becomes its synthesis counterpart by a simple replacement of terms. Once an algorithm for analysis (synthesis) is defined, the corresponding algorithm for synthesis (analysis) follows. For instance, one can easily apply the principle of duality to the Ricardi and Burrows' method of paragraph (1.4) to obtain a synthesis technique.

1.9 Usefulness of the Fast Fourier Transform Method

The excellent book by Brigham (1974) gives a thorough presentation of the 1-D discrete Fourier transform and its applications, and explains in detail the method known as FFT for computing such transform. The Fourier transform in 2 and higher dimensions can be found simply as follows:
first, get the 1-D transform of each row, then that of each column of the modified array... etc., until all dimensions have been exhausted in this way. Understanding the workings of the 1-D transform is enough to understand those of the N-dimensional transform as well. The FFT requires $O$(number of points) operations for each row along the nth dimension, so the number for all points in a regular, euclidean array is always of the order of the total number of points in that array.

Before the mid-sixties — when the FFT came along — the best techniques available for the analysis of data on regular arrays required $O$(number of points)$^2$ operations. The increase in speed of $O$(number of points) brought about a true revolution in data processing: work that had been long regarded as impossible became feasible overnight, the field of industrial and scientific applications for numerical Fourier transforms expanded tremendously; the impact in areas as diverse as crystallography and communications engineering was remarkable.

Having mentioned the positive side of the FFT, which is used in the algorithms described so far (at least in principle) and in the programs HARMIN and SSYNT4, it is only proper to say something about its alternatives. The FFT calculates all $2N$ Fourier coefficients $a_8, b_8$ very efficiently, but takes just about as many operations to get only a few coefficients as it takes to get all: for $N_{max}$ small compared to $N$ there may be a real disadvantage in using the FFT. The FFT is most efficient when the grid is such that $N$ is an integer power of 2. The grids used in geodesy are usually based on the division of the circle in 360°, and many on the sexagesimal division of 1° as well. In all of these $N$ contains factors other than 2, so a less efficient version of the FFT, known as the mix-radix FFT (Singleton’s algorithm) must be used.

Finally, the mix-radix algorithm is rather convoluted, so it is best to take ready available subroutines from software libraries (as it is done in HARMIN and SSYNT4) rather than to incorporate the FFT "on line" in the program one is writing. This means that the program is going to be less self-contained.

The "pre-FFT" methods can be more efficient than the FFT when $N_{max} \ll N$; they are also very easy to program. For the sake of completeness, the outline of a method this author has used quite often will be given here.

Consider the trigonometric relationships

$$\cos(\alpha \beta) = 2\cos \beta \cos(\alpha - 1) \beta - \cos(\alpha - 2) \beta$$  \hspace{1cm} (1.32)  

$$\sin(\alpha \beta) = 2\cos \beta \sin(\alpha - 1) \beta - \sin(\alpha - 2) \beta$$  \hspace{1cm} (1.33)
with \( \alpha, \beta \) real. If all the values of the trigonometric functions could be obtained with one operation each, the number of operations involved in finding

\[
\begin{bmatrix}
    a_n \\
    b_n
\end{bmatrix} = \sum_{j=0}^{2N-1} \begin{bmatrix}
    \cos m j \Delta \lambda (\lambda_j) \\
    \sin m j \Delta \lambda (\lambda_j)
\end{bmatrix} f(\theta_j, \lambda_j)
\]

would be \( 2(2N) \), or \( 4N^2 \) for all \( 0 \leq m \leq N \). In other words: \( O(N^2) \), as expected. This is precisely what can be done using (1.32), (1.33) as recursive expressions for \( \cos m(j \Delta \lambda) \) and \( \sin m(j \Delta \lambda) \) with \( m \) integer and \( 0 \leq j \leq 2N - 1 \). The values of \( \cos m(-\Delta \lambda) = \cos m \Delta \lambda \) and \( \sin m(-\Delta \lambda) = -\sin m \Delta \lambda \), are needed to start the recursion; they can be calculated with standard trigonometric subroutines. The use of such subroutines increases the number of operations slightly over \( 4N^2 \), but if \( N \) is large enough, this is negligible.

With all calculations carried out in double precision, this method gives values of cosine and sine that coincide to better than 5 significant figures with those provided by the standard FORTRAN functions, when \( N \) is as large as 1800 (0.1° x 0.1° grid). By taking advantage of half-wave symmetries in the sine and cosine, and by ingenious programming, the number of operations can be reduced by a factor of 4 or more rather easily.

1.10 Functions Harmonic in Space and their Gradients

If \( f(\theta, \lambda, r) \) satisfies Laplace's equation \( \nabla^2 f = 0 \) in the space outside a sphere of radius \( a \), then it can be represented, in that space, by the solid spherical harmonic expansion

\[
f(\theta, \lambda, r) = \sum_{n=0}^{\infty} \sum_{\alpha=0}^{n} \frac{a^n}{r^{n+1}} \mathcal{C}_n \mathcal{Y}_n (\theta, \lambda)
\]

(1.34)

If we consider, at a point \( P = (\theta, \lambda, r) \) in space, the local triad \( \vec{r}, \vec{h}, \vec{t} \) oriented downwards to the origin, West to East along the tangent to the local parallel, and North-South along the local meridian, the components of the gradient of \( f(\theta, \lambda, r) \) along this three axes are

\[
\frac{\partial f}{\partial r} (\theta, \lambda, r) = \sum_{n=0}^{\infty} \sum_{\alpha=0}^{n} \frac{a^n}{r^{n+2}} (n + 1) \mathcal{C}_n \mathcal{Y}_n (\theta, \lambda) \]

(1.35)

\[
\frac{\partial f}{\partial h} (\theta, \lambda, r) = \sum_{n=0}^{\infty} \sum_{\alpha=0}^{n} \frac{a^n}{r^{n+2}} \mathcal{P}_n (\cos \theta) \csc (\theta) [m\mathcal{S}_n \cos m \lambda - m \mathcal{C}_n \sin m \lambda]
\]

(1.36)

\[
\frac{\partial f}{\partial t} (\theta, \lambda, r) = \sum_{n=0}^{\infty} \sum_{\alpha=0}^{n} \frac{a^n}{r^{n+2}} \frac{d\mathcal{P}_n}{d\theta} (\cos \theta) [\mathcal{C}_n \cos m \lambda + \mathcal{S}_n \sin m \lambda]
\]

(1.37)

1Here one "operation", as mentioned in paragraph (1.4), consists of one sum and one product.
Expressions (1.34), (1.35), (1.36), and (1.37) appear often in the discussion of geodetic problems, and their calculation from a set of coefficients is an important problem. If all the values of any of them (with $r$ constant, and on a regular grid) are required, the methods discussed so far can be used after a few minor additions.

The synthesis algorithms can be thought of, for the purpose of this discussion, as "black boxes" with the coefficients $C_{n}^{\alpha}$, the Legendre functions, $N_{n,\alpha}$, and $N$ as inputs, and all the $2N^2$ values of $f(\theta, \lambda)$ on the corresponding regular grid as outputs. To compute the expressions given above, only the part of the input consisting in the coefficients and/or the Legendre functions has to be modified before they enter the "box", which remains untouched. For instance: to compute (1.34) one should replace $C_{n}^{\alpha}$ with $\frac{a_n}{r^{n+2}} C_{n}^{\alpha}$ in the "input"; the others are equally obvious and will not be explained further. The following recursive formulas can be used to obtain the derivatives of the Legendre functions:

$$
\frac{d}{d\theta} \bar{P}_{n} = \frac{(\sin \theta)^{-1}}{n} \bar{P}_{n} (\cos \theta) \cos \theta - \left[ \frac{(n^2-m^2)(2n+1)}{2n-1} \right]^{-\frac{1}{2}} P_{n-1} \frac{m(\cos \theta)}{m+1}
$$

(1.38a)

$$
\frac{d}{d\theta} \bar{P}_{n} = \left[ \frac{(2n+1)}{(2n)} \right]^{\frac{1}{2}} \left\{ \sin \theta \frac{d}{d\theta} \bar{P}_{n-1} + \cos \theta \bar{P}_{n+1} \right\}
$$

(1.38b)

with the starting value

$$
\frac{d}{d\theta} \bar{P}_{\infty} = 0
$$

These recursives follow from the unnormalized formula

$$(\cos^2 \theta - 1) \frac{d}{d(\cos \theta)} \bar{P}_{n} (\cos \theta) = n \cos \theta \bar{P}_{n} (\cos \theta) - (n + m) \bar{P}_{n-1} (\cos \theta)$$

(N.N. Lebedev, "Special Functions", Dover, 1972, Ch. 7, equation 7.12.16), and from

$$
\bar{P}_{n} = \left[ \frac{(2n+1)}{(2n-1)} \right]^{\frac{1}{2}} \sin \theta \bar{P}_{n-1}
$$

(see paragraph (4.4))

and

$$
\bar{P}_{n} = \left[ \frac{(n + m)!}{2(n + 1)(n - m)!} \right]^{-\frac{1}{2}} P_{n} \; ; \; \bar{P}_{\infty} = \sqrt{2n+1} \bar{P}_{\infty}
$$

The complete recursive expressions for the $\bar{P}_{n}$ are given in paragraph (4.4).

The expansion for the area means defined by (1.2) can be differentiated term by term because it converges uniformly. The expressions for area means gradients, equivalent to those given here for point values, are immediate. They can be computed after simple modifications to the $C_{n}^{\alpha}$ and/or the $\chi_{n}^{\alpha}$, and using the same programs for computing the area means.

Subroutine "LEGFDN", listed on Appendix B, can compute both the normalized $\bar{P}_{n} (\cos \theta)$ and their derivatives $\frac{d}{d\theta} \bar{P}_{n} (\cos \theta)$. 

-22-
2. Error Measure and Optimal Quadrature Formulas

This section introduces a criterion for quantifying the errors of numerical quadratures formulas that is based on the statistical properties of the data. Three qualities are highly desirable in an error measure: (a) it should be easy to determine; (b) it should be mathematically tractable; (c) it should provide a good idea of the likely size of the actual errors. Point (a) is taken into account by choosing a quadratic measure, because the numerical formulas are linear estimators of the $\sum_{n} \frac{\partial^2}{\partial n^2} \tilde{C}_{n}^{\alpha}$, and the linear, quadratic estimation problem is fairly simple, with its mathematical side very well understood and developed today, which takes care also of (b). Regarding (c), the reader will have to wait till section three, where certain evidence, obtained from numerical experiments, supports the assertion that, though statistical in nature, the error measure adopted represents the actual errors very closely.

Having defined the error measure, the notion of optimal or best formula according to such measure is investigated, leading to the application of least squares collocation and least squares adjustment to spherical harmonic analysis.

2.1 The Isotropic Covariance

The isotropic covariance (expression (1.10)) between two functions $u(\theta, \lambda)$, $v(\theta, \lambda)$ on the unit sphere, both expandable in spherical harmonic series

$$u(\theta, \lambda) = K_u \sum_{n=0}^{\infty} \sum_{m=0}^{n} \sum_{\alpha=0}^{m} \tilde{C}_{n}^{\alpha} \tilde{V}_{n}^{\alpha}(\theta, \lambda);$$

$$v(\theta, \lambda) = K_v \sum_{n=0}^{\infty} \sum_{m=0}^{n} \sum_{\alpha=0}^{m} \tilde{C}_{n}^{\alpha} \tilde{V}_{n}^{\alpha}(\theta, \lambda)$$

can be formally defined as follows

$$\text{cov}(u(P), v(Q)) = M \left\{ u(P), v(Q) \right\}$$

(2.1)

where $M \{ \}$ is the isotropic averaging operator and $P$ and $Q$ are two points on the sphere separated by the spherical distance $\psi_{PQ}$. The operator $M \{ \}$ symbolizes the average of its argument (in the present case the product $u(P) v(Q)$) over all rotations of the sphere. This can be visualized if one thinks of the points $P$ and $Q$ as given in a fixed system of coordinates, while the sphere, on which $u$ and $v$ are defined, rotates in all possible ways. After all the (infinitely many) possible rotations, the average product $u(P) v(Q)$ will be identical to $\text{cov}(u(P), v(Q))$. This kind of covariance, though purely geometrical, resembles closely that of stochastic processes such as time series.
The importance of the isotropic operator and the isotropic covariance function in spherical harmonic analysis stems from the fact that the latter can be described as the estimation of certain parameters of a function \( f(\theta, \lambda) \), the \( C_{\alpha}^{\lambda} \), from data sampled on a sphere. From paragraph (2.7) on, this report deals with optimal estimators for the \( C_{\alpha}^{\lambda} \) based on the theory of least squares collocation. Such optimal estimators minimize a quadratic measure of the error that is defined in terms of the operator \( M \{ \} \), this measure being introduced in paragraph (2.4).

The idea of least squares collocation is related to the basic principles of such linear, minimum variance estimators for time series as the Wiener and Kalman filters, which have found wide application in the physical sciences and in engineering over the last thirty years, and have been generalized to deal with both continuous and discrete time processes, and also "processes" in more than one dimension, such as are found in pattern recognition and in digital image enhancement. Two-dimensional Wiener filtering, of which the reader can find several fine descriptions in the special issue of the Proceedings of the IEEE, Vol 65, No. 6, 1977, is also applicable to "flat-Earth" geodetic calculations; least squares collocation can be regarded as the extension of this type of filtering to calculations on the sphere. Isotropic average operators are not the only ones that could be used in the "statistical" approach, though they are probably the easiest to work with and, perhaps, the best for the sort of application considered here. For a description of other likely operators, the reader is referred to the paper by Rumml and Schwarz (1978). Probably the most didactic introduction to the method of collocation remains Heiskanen and Moritz, (Ch. 7, 1967).

Reasoning as in Heiskanen and Moritz (ibid), one can show that

\[
\text{cov}(u(P), v(Q)) = \sum_{n=0}^{\infty} C_{2n}^{\alpha} R_{n}(\cos \psi_{PQ})
\]

which is, in fact, expression (1.10*), the definition of the isotropic covariance given in section 1 without any reference to \( M \{ \} \). Similarly,

\[
\text{cov}(u(P), u(Q)) = \sum_{n=0}^{\infty} \sigma_{2n}^{\alpha} R_{n}(\cos \psi_{PQ})
\]

usually known as "the covariance of \( u \)" (expression 1.10), while (1.10*) represents the "covariance between \( u \) and \( v \)", or "the crosscovariance of \( u \) and \( v \)". The one-to-one relationship between covariance and power spectrum (or crosscovariance and crossspectrum) should be clear from these expressions.

To apply the notions introduced above to the \( C_{\alpha}^{\lambda} \), it is necessary to think of them as functions rather than fixed values. This is possible if one considers changes in the coordinates \( \theta, \lambda \) brought about by rotation. Each
such change results in different coefficients, though the function they describe is the same, only rotated with respect to the old system. The new system can be related to the first by three angles: the coordinates $\theta, \lambda$ of the shifted north pole, and the azimuth $A$ of the zero meridian. Therefore, the $C_{n,m}^\alpha$ are functions of $\theta, \lambda, A$ and this is enough to define the average over all rotations of the product of each $C_{n,m}^\alpha$ by itself or by another function, in a meaningful way. Two important properties of spherical harmonics are:

\[
M \left\{ \sum_{n,m} \overline{C_{n,m}^\alpha} \overline{S_{n,m}} \right\} = M \left\{ \sigma_n^2 \right\} = \sigma_n^2 \tag{2.2}
\]

i.e., the power spectrum is invariant with respect to rotations. This follows from the plain fact that the integral $\int_\sigma f^2(\theta, \lambda) d\sigma$ is invariant over rotations, and from Parseval's identity (1.12); (2.2) implies that the isotropic covariance function (1.10) is likewise invariant.

\[
M \left\{ \overline{Y_n^\alpha(p)} \overline{Y_p^\beta(q)} \right\} = 0 \text{ if } \begin{cases} \alpha \neq \beta \\ n \neq p \\ m \neq q \end{cases} \tag{2.3}
\]

i.e., the orthogonality properties of spherical harmonics with respect to integration on the sphere are also true with respect to averaging over rotations.

As a consequence of (2.2) and (2.3) above, the following relationships are also true:

\[
M \left\{ \overline{C_{n,m}^2} \right\} = M \left\{ \overline{S_{n,m}^2} \right\} = \frac{\sigma_n^2}{2n+1} \tag{2.4-a}
\]

\[
M \left\{ \overline{C_{n,m}^\alpha C_p^\beta} \right\} = 0 \text{ if } \begin{cases} \alpha \neq \beta \\ n \neq p \\ m \neq q \end{cases} \tag{2.4-b}
\]

\[
M \left\{ \overline{C_{n,m}^\alpha} \right\} = 0 \text{ if } n \neq 0 \tag{2.5}
\]

\[
M \left\{ \overline{C_{n,m}^\alpha f(\theta, \lambda)} \right\} = \frac{\sigma_n^2}{2n+1} \overline{Y_n^\alpha(\theta, \lambda)} \tag{2.6}
\]

\[
M \left\{ \overline{C_{n,m}^\alpha \overline{f_{1,1}}} \right\} = \frac{\sigma_n^2}{(2n+1)A_{1,1}} \int_{\sigma_{1,1}} \overline{Y_n^\alpha(\theta, \lambda)} d\sigma \tag{2.7}
\]
For the derivation of (2.6) the reader can see (Rummel, 1976), and (Sjoberg, 1978) for (2.7). As for (2.4a-b) and (2.5), the proof is given now.

According to (1.4):

\[
16 \pi^3 \mathcal{M} \left[ \overline{C}_{n_{\alpha}^\beta} \right]^2 = \mathcal{M} \left[ \int_{\sigma} \int_{\sigma'} \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) f(\theta',\lambda') \ d\sigma \int_{\sigma'} \overline{Y}_{n_{\alpha}^\beta}(\theta',\lambda') f(\theta,\lambda') \ d\sigma' \right] =
\]

\[
= \mathcal{M} \left[ \int_{\sigma} \int_{\sigma'} \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) f(\theta,\lambda) \ f(\theta',\lambda') \overline{Y}_{n_{\alpha}^\beta}(\theta',\lambda') \ d\sigma \ d\sigma' \right] =
\]

\[
\int_{\sigma} \int_{\sigma'} \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) \overline{Y}_{n_{\alpha}^\beta}(\theta',\lambda') \ M \left[ f(\theta,\lambda) f(\theta',\lambda') \right] \ d\sigma d\sigma' =
\]

\[
\int_{\sigma} \int_{\sigma'} \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) \overline{Y}_{n_{\alpha}^\beta}(\theta',\lambda') \ \text{cov} \ (f(P), f(Q)) \ d\sigma d\sigma'
\]

where \( P = (\theta,\lambda) \) and \( Q = (\theta',\lambda') \). According to (1.10):

\[
16 \pi^2 \mathcal{M} \left[ \overline{C}_{n_{\alpha}^\beta} \right]^2 = \int_{\sigma} \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) \ d\sigma \int_{\sigma'} \overline{Y}_{n_{\alpha}^\beta}(\theta',\lambda') \sum_{n=0}^{\infty} \sigma_n^2 \ P_n(\cos \psi_q) \ d\sigma' =
\]

\[
= 4 \pi \int_{\sigma} \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) \ \frac{\sigma_n^2}{2n+1} \ d\sigma = \frac{\sigma_n^2}{2n+1} \ 16 \pi^2 \ \text{because of (1.14)}
\]

Similarly,

\[
\mathcal{M} \left[ \overline{C}_{n_{\alpha}^\beta} \overline{C}_{p_{\alpha}^\beta} \right] = \frac{\sigma_n^2}{2n+1} \cdot \frac{1}{4 \pi} \int_{\sigma} \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) \overline{Y}_{p_{\alpha}^\beta}(\theta,\lambda) \ d\sigma = 0 \ \text{if } \begin{cases} \alpha \neq \beta \\ n \neq p \\ m \neq q \end{cases}
\]

Finally, recalling that \( Y_{\alpha 0}(\theta,\lambda) = P_{\alpha 0}(\cos \theta) = 1 \) for all \(- \pi \leq \theta \leq \pi\) and all \( \lambda \),

\[
\mathcal{M} \left[ \overline{C}_{n_{\alpha}^\beta} \right] = \frac{1}{4 \pi} \int_{\sigma} \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) \ M \left[ f(\theta,\lambda) \right] \ d\sigma = \frac{M}{4 \pi} \left\{ f(\theta,\lambda) \right\} \int_{\sigma} \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) \ d\sigma
\]

\[
= \frac{M}{4 \pi} \left\{ f(\theta,\lambda) \right\} \int_{\sigma} \overline{Y}_{n_{\alpha}^\beta}(\theta,\lambda) \overline{Y}_{00}(\theta,\lambda) \ d\sigma = 0 \ \text{if } n \neq 0,
\]

which completes the proof.

2.2 Some Additional Notation

So far, data points on the sphere have been identified by the subscripts \( i \) and \( j \). Alternatively, they could be arranged according to a single subscript \( k = 2Ni + j \) (where \( N \) is the number of parallels in a grid and \( 2N \) the number of meridians), so the points in the "0" row, ordered by increasing \( j \), are followed by those in the "1" row, in the same order, etc., the last element in the "N - 1" row closing the sequence. Based on this convention, the set of all
values of \( f(\theta_i, \lambda_j) \) (or \( \bar{f}_{ij} \)) can be arranged in \( N_p \) vector form according to \( k \):

\[
\underline{z} = [c_0 \; z_1 \ldots c_0 \; z_{N_p-1}]^T
\]

(2.8)

where \( z_k = f(\theta_i, \lambda_j) \) or \( z_k = \bar{f}_{ij} \) with \( k = 2Ni + j \)

and where \( N_p \) is the number of data points, or \( 2N^2 \) for equal angular grids.

In a similar way, the coefficients \( \underline{C}_{\alpha} \) can be ordered according to a single subscript \( p = n^2 + \alpha n + m + 1 \) (with the understanding that the meaningless \( S_{<0} \) are not included) defining the following \( N_s \) vector:

\[
\underline{c} = [c_0 \; c_1 \ldots c_p \ldots c_{N_s-1}]^T
\]

(2.9)

\[
c_p = \underline{C}_{\alpha}^\alpha \quad p = n^2 + \alpha n + m + 1
\]

where \( N_c = (N_{max} + 1)^2 \). Using this notation, expression (1.27) for point data quadratures can be written

\[
\underline{C}_{\alpha}^\alpha = \underline{f}_{\alpha}^\alpha \underline{z} \quad (2.10)
\]

where the estimator vector \( \underline{f}_{\alpha}^\alpha \), of dimension \( N_p \), has elements of the form

\[
f_{\alpha} = \chi_{\lambda}^a \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} \quad m \lambda \Delta \lambda
\]

under the convention given above relating \( k \), \( i \) and \( j \).

Grouping all the estimates \( \underline{C}_{\alpha}^\alpha \) in a vector \( \underline{c} \) ordered in the same way as \( \underline{z} \), the relationship between the \( \underline{C}_{\alpha}^\alpha \) and \( \underline{z} \) can be written, in matrix form,

\[
\underline{c} = F \underline{z} \quad (2.11)
\]

where \( F \) is the estimator matrix implied by (2.10). It is a \( N_p \times N_c \) matrix (where \( N_p \) is the number of data points in the grid), each row being formed by the coefficients of the quadrature formula for the corresponding \( \underline{C}_{\alpha}^\alpha \). Such row is also the transpose of the estimator vector of this \( \underline{C}_{\alpha}^\alpha \), designated \( \underline{f}_{\alpha}^\alpha \) in (2.10).

In the same way as the covariance function between scalars, the covariance between vector functions can be defined in terms of \( M \left\{ z, \; z^T \right\} \):

\[
M \left\{ z, \; z^T \right\} = C_{zz} \quad (2.12)
\]

where \( C_{zz} \) is the covariance matrix of \( z \), of dimension \( N_p \times N_p \). This matrix is a function of the relative positions of the points in the grid on which \( z \) has been determined, in the same way as the scalar covariance depends only on the distance between two points. The elements of \( C_{zz} \) are
\[ c_{zz}^r = M \left\{ z_{r} \ z_{s} \right\} = M \left\{ f(P_{r}) \ f(P_{s}) \right\}, \]

i.e., the values of the scalar covariance corresponding to pairs of points in the grid.

In the same way

\[ M \left\{ c \ c^\top \right\} = C \]

(2.13)

is a \( N_x \times N_x \) diagonal matrix according to (2.4a–b). \( C \) is the covariance matrix of the coefficients. Similarly, the covariance between \( z \) and \( \bar{z} \) is

\[ M \left\{ z \ \bar{z}^\top \right\} = C_{zz} = \left[ M \left\{ z \ c^\top \right\} \right]^\top = C_{zz}^\top \]

(2.14)

where \( C_{zz} \) is a \( N_x \times N_x \) matrix, the elements of which are

\[ c_{zz}^{pk} = M \left\{ c_p \ z_k \right\} = M \left\{ \overline{f^{\alpha}} \ f(\theta_i, \lambda_j) \right\} \]

where the right hand side is given by (2.6).

Finally, when estimating the \( \overline{f^{\alpha}} \), not from samples of \( f(\theta, \lambda) \), but from measurements corrupted by noise

\[ m(\theta_i, \lambda_j) = f(\theta_i, \lambda_j) + n_i \]

(2.15)

the measurement errors can be grouped in a \( N_p \)-vector \( n \) with the same ordering as \( z \), and the sum of both will be, then, the \( N_p \)-vector of observed values

\[ m = z + n \]

(2.16)

The measurement errors are values that occur in time, as successive observations are carried out: they constitute a time series. The average operator appropriate to them is the usual statistical expectation operator \( E \left\{ \right\} \). The measurements are supposed to be unbiased, so \( E \left\{ n_k \right\} = 0 \) for all \( k \). The covariance implied by this operator is the usual statistical covariance: \( E \left\{ n_k^2 \right\} = \sigma_k^2 \) and \( E \left\{ n_k \ n_r \right\} = \sigma_{kr} \). This can be generalized for the noise vector \( n \):

\[ E \left\{ n \ n^\top \right\} = D \]

(2.17)

where \( D \) is a \( N_p \times N_p \) matrix of elements

\[ d_{kr} = E \left\{ n_k \ n_r \right\} = \sigma_{kr}^2 \]

(2.18)

Both \( C_{zz} \) and \( D \) have in common a very important property:
\[
\begin{align*}
\mathbf{x}^T \mathbf{C}_{zz} \mathbf{x} & \geq 0 \quad \text{if } \mathbf{x}^T \mathbf{x} \neq 0 , \text{ any } \mathbf{N}_p \text{ vector,} \\
\mathbf{x}^T \mathbf{D} \mathbf{x} & \geq 0
\end{align*}
\]

i.e., they are always positive matrices, moreover, in all the cases considered here, at least \( \mathbf{D} \) is positive definite:

\[
\mathbf{x}^T \mathbf{D} \mathbf{x} > 0 \quad \text{for all } \mathbf{x}
\]

Positiveness can be inferred readily from the definitions of \( \mathbf{C}_{zz} \) and \( \mathbf{D} \):

i.e.,

\[
\mathbf{x}^T \mathbf{D} \mathbf{x} = \mathbf{x}^T \mathbf{E} \left( \mathbf{u} \mathbf{n}^T \right) \mathbf{x} = \mathbf{x}^T \mathbf{E} \left( \mathbf{u} \mathbf{n}^T \mathbf{x} \mathbf{n} \mathbf{x}^T \right) = \mathbf{E} \left( \mathbf{h}^2 \right) \geq 0
\]

(where \( \mathbf{h} = \mathbf{x}^T \mathbf{n} \)), and similarly for \( \mathbf{x}^T \mathbf{C}_{zz} \mathbf{x} \) (with \( M \{ \} \)).

2.3 Estimation Errors, Sampling Errors, and Propagated Noise

A linear estimator is of the general form

\[
\mathbf{s} = \mathbf{F} \mathbf{m}
\]

where \( \mathbf{m} \) is the vector of measurements defined by (2.16), and \( \mathbf{s} \) is the vector of estimates, made up in our case of the \( \hat{\mathbf{c}}_{\mathbf{z}}^{\mathbf{z}} \). According to (2.11) and (2.16)

\[
\mathbf{s} = \hat{\mathbf{c}} = \mathbf{F} (\mathbf{z} + \mathbf{n})
\]

In general, the estimates will not be exactly equal to that which is estimated, the difference being the estimation error. In matrix notation

\[
\mathbf{e} = \mathbf{c} - \hat{\mathbf{c}} = (\mathbf{c} - \mathbf{F} \mathbf{z}) - (\mathbf{F} \mathbf{n}) \tag{2.19}
\]

\( \mathbf{e} \) being the estimation error vector. The two terms in the expression above can be defined as the components of this error:

\[
\mathbf{e}_s = \mathbf{c} - \mathbf{F} \mathbf{z}
\]

which is the estimation error in the case of noiseless (perfect) data; and

\[
\mathbf{e}_\eta = \mathbf{F} \mathbf{n}
\]

which is the error due to the noise, or propagated noise.

The error \( \mathbf{e}_{sp} \) may be due to a number of reasons.\(^1\) If it is zero for

\(^1\)Using the relationship \( p = \mathbf{n}^2 + \alpha \mathbf{n} + \mathbf{m} + 1 \) of paragraph (2.2), \( \mathbf{e}_{sp} \) stands for the sampling error in \( \hat{\mathbf{c}}_{\mathbf{z}}^{\mathbf{z}} \), and \( \mathbf{e}_\eta \) for the propagated noise.
some estimator, then its presence in other estimators could be blamed on them being somewhat inadequate. For instance, if the estimator was chosen by taking the elements of $F$ from a set of random numbers, then the estimation error is likely to be always high, as the estimator has nothing to do with the actual problem. In particular, the addition of extra measurements to the vector $\mathbf{m}$ is not going to bring any general improvement on the estimates. On the other hand, if attention is paid to the nature of the problem when selecting $F$, one would expect the error to decrease as more data is introduced. If, as the number $N_p$ of samples in $\mathbf{m}$ tends to infinity, $e_{sp}$ tends to zero, one could say that the error is due to the incomplete sampling of the signal $f(\theta, \lambda)$, and call it the sampling error. This is precisely the case with any of the quadrature formulas to be studied here, all of which can be written formally as linear estimators according to (2.11), and for all of which the error $e_{sp}$ vanishes as the number of samples tends to infinity, because the sums become identical with the integrals defined by (1.4). In this sense it is quite suitable to call $e_{sp}$ the sampling error, as in paragraph (1.3).

2.4 The Quadratic Error Measure

The overall error measure will be defined here as the sum of two quadratic terms: one for the propagated noise, the other for the sampling error.

(a) Propagated Noise Measure

This measure is the same as in least squares adjustment, i.e., the variance of the error defined in terms of the usual statistical expectation operator

$$\sigma_{\eta_{mn}}^2 = E\left\{ (f_{mn}^{\alpha^\top} n)^2 \right\}$$  \hspace{1cm} (2.20)

according to (2.10). This variance represents the scatter in the value of $\eta_{mn}$ due to the uncertainty in the values of the data. In matrix form

$$E\eta = E\left\{ \eta \eta^\top \right\} = E\left\{ F \eta n \eta^\top F^\top \right\} = F E\left\{ n n^\top \right\} F^\top = F D F^\top$$ \hspace{1cm} (2.21)

where $E\eta$ is a $N_s \times N_s$ matrix, while $D$ was already presented in paragraph (2.2).

In the special case where the measurement errors are uncorrelated, $D$ is diagonal, and (2.20) becomes

$$\sigma_{\eta_{mn}}^2 = E\left\{ f_{mn}^{\alpha^\top} n n^\top f_{mn}^{\alpha^\top} \right\} = f_{mn}^{\alpha^\top} D f_{mn}^{\alpha^\top} = \sum_{i=0}^{N-1} (x_{i}^{n})^2 \sum_{j=0}^{N-1} \left\{ \cos^2 \right\} m j \Delta \lambda E\left\{ n_{i,j}^2 \right\}$$ \hspace{1cm} (2.22)

which is the usual formula for propagating the covariance of the noise.
(b) Sampling Error Measure

This measure is defined in terms of the isotropic averaging operator of paragraph (2.1)

\[ \sigma^2_{\alpha}\alpha = M \left\{ (c_p - \frac{f}{F} z)^2 \right\} = M \left\{ (\bar{c}_{z}^\alpha - \frac{f}{F} z)^2 \right\} \]  

\[ (2.23) \]

or, in matrix form

\[ E_s = M \left\{ e_s e_s^T \right\} = M \left\{ (c - F z)(c - F z)^T \right\} \]

\[ = C - 2 C_z F^T + F C_z F^T \]

\[ (2.24) \]

where \( E_s \) is a \( N_x \times N_x \) matrix, and \( C, C_z, C_{zz} \) were introduced in paragraph (2.2).

(c) Total Error Measure

The total measure is the sum of (a) and (b)

\[ \sigma^2_{\alpha}\alpha = \sigma^2_{\eta}\eta + \sigma^2_{s}\alpha \]

\[ (2.25) \]

or, in matrix form,

\[ E_T = E_s + E_\eta = C - 2 C_z F^T + F C_{zz} F^T + F D F^T \]

\[ = C - 2 C_z F^T + F (C_{zz} + D) F^T \]

\[ (2.26) \]

where \( E_T \) is the \( N_x \times N_x \) error matrix associated with \( F \) and with the covariances that define \( C, C_z \), and \( (C_{zz} + D) \). Expression (2.26) is a special case of the formula for "\( E_s \)" in least squares collocation (for instance, Moritz (1978), Ch. 3, eqn. (3.20)); moreover, it belongs to a family of formulas also found in the minimum variance estimation and filtering of time series and of processes sampled on the euclidean plane.

The total measure has been chosen simply as the sum of \( \sigma^2_{\alpha}\alpha + \sigma^2_{\eta}\eta \) by making the basic assumption that the sampling error and the propagated noise are due to completely independent causes. The first depends on the values \( f(\theta, \lambda) \), while the second depends on the measurement errors of instruments that, at least ideally, operate with accuracies unaffected by the quantities measured, or in such way that any interactions can be eliminated by simple corrections.

The columns of \( F \) are defined by the quadrature formula used, and such formulas either satisfy, or tend to satisfy, orthogonality conditions (paragraph (1.3(c)). For this reason, provided that \( C_{zz} \) and \( D \) belong to the type to be described in paragraph (2.9), matrices \( E_\eta, E_s \), and, thus, \( E_T \), are either diagonal or diagonal dominant, and in the latter case tend to become diagonal as
the sampling intervals decrease, or \( N_p \to \infty \). For this reason the correlations among the errors for individual coefficients are, or "tend to be", very small.

The diagonal elements of the error matrices \( E_\delta \), \( E_\eta \) and \( E_\gamma \) are the variances of the errors in the respective coefficients, as defined by (2.20), (2.23), and (2.25), respectively.

2.5 The Meaning of the Error Measure

The treatment of the propagated noise is the same as in least squares adjustment, so this part of the error measure should be easily understood. The sampling error measure, on the other hand, is a geometrical measure: \( M \{ \} \) belongs, as a concept, in the field of integral geometry, or the study of "geometric probabilities". This is a branch of mathematics closely related to integration and to measure theory, and also to statistical mechanics. In geodesy, this type of idea is relatively new (Kaula, 1959), Moritz (1965), but it has been used already extensively enough to show its considerable worth.

From expressions (1.10) and (1.11), the covariance and the power spectrum are functions of each other. Since either of them, and the sampling grid, define matrices \( C \), \( C_{az} \) and \( C_{zz} \) in expression (2.24), it follows that a statement on \( \sigma_{zz}^2 \) is, somehow, also a statement on the performance of \( F \) for all the functions that have the same power spectrum that determines the diagonal elements of \( C \). To put this more precisely, consider a function \( f_1 (\theta, \lambda) \) having the given power spectrum. If \( C_{az} \) were estimated for \( f_1 \) and also, at least ideally, for all its rotations, then the mean square of the sampling error \( e_{sp} \) in \( C_{az} \) for all these functions would be, by definition of \( M \{ \} \), the measure \( \sigma_{az}^2 \). If a second function \( f_2 \) (perhaps not a rotation of \( f_1 \)) and all its rotations were then analysed in the same way, the average of \( e_{sp}^2 \) for all these functions would be, once more, \( \sigma_{az}^2 \), as long as \( f_2 \) has the same spectrum as \( f_1 \). Moreover, the average of \( e_{sp}^2 \) for \( f_1 \), \( f_2 \), and their rotations put together, would also be \( \sigma_{az}^2 \). In fact, if we had a finite set of functions \( f_1, f_2, \ldots, f_n \), with arbitrary \( n \), all with the same power spectrum (or covariance), then \( e_{sp}^2 \) would average \( \sigma_{az}^2 \) for all the \( f_i \) and their rotations.

It appears, from the preceding discussion, that one could take a simple step and say "\( \sigma_{az}^2 \) is the mean of the sampling error squared of the estimator \( \hat{\alpha}_{az} = \hat{\alpha}_{az} \) over all possible functions with the given power spectrum." Unfortunately, as mentioned in the Introduction, the sphere is a rather wicked surface. There is a theorem by Lauritzen (1973) that states the impossibility of having the same average \( \sigma_{az}^2 \) for all functions as for every function, when the distribution of the ensemble happens to be gaussian. Moritz (1978) has endeavoured to show that this is no problem if the ensemble of functions
is not gaussian, but using his conclusions here would force the introduction of a rather strange requirement of "non-gaussian" on the ensemble of the signals analysed that is best left out, if possible.

Perhaps there is a way out in going back to the idea of a finite set of functions \( f_1 \), where the problem does not exist, by saying:

"the error measure \( \sigma_{n_3}^2 \), for a certain estimator and a certain type of signal power spectrum, is the average of the squares of the sampling error in \( \sigma_{n_3}^2 \) for all functions with the given power spectrumEVER TO BE ANALYSED with that estimator, and for all their rotations".

After all, accuracy is what geodesists are always interested in, not perfection.

2.6 Simple Formulas for Area Means

The numerical studies of section 3 concentrate in area mean type formulas, because area means are preferred for collating information, particularly on a global basis, at present. The formulas to be studied here and in that section can be divided into "simple" and "optimal". The name "simple" is given here to expressions of the type

\[
\tilde{C}_{n_3}^{\alpha} = \mu_3 \sum_{i=0}^{N-1} \sum_{j=0}^{2^{N-1}-1} \tau_{1, i} \int_{\sigma_{1, i}} \overline{Y}_{n_3}^{\alpha} (\theta, \lambda) \, d\sigma
\]

(2.27)

where \( \mu_3 \) in a scale factor affecting the \( n \)th harmonic as a whole. Expressions of this type have been developed more or less intuitively, along the lines of the following reasoning:

If the signal were constant on each block, it will equal its mean value there, and the coefficients of such a function would be precisely

\[
\tilde{C}_{n_3}^{\alpha} = \frac{1}{4\pi} \sum_{i=0}^{N-1} \sum_{j=0}^{2^{N-1}-1} \tau_{1, i} \int_{\sigma_{1, i}} \overline{Y}_{n_3}^{\alpha} (\theta, \lambda) \, d\sigma
\]

(2.28)

according to (1.4). In general, most signals are not equal to their mean value over whole blocks, so the expression would not be exact. In most cases, the signal would have fluctuations in each block, and it would be less smooth than a function that is constant over each block, so using the formula above with \( \tau_{1, i} \) as data may result in the \( \tilde{C}_{n_3}^{\alpha} \) of a smoothed function. As a refinement, one could try to de-smooth the \( \tilde{C}_{n_3}^{\alpha} \). If the blocks were circular, the relationship between "true" and "smooth" \( \tilde{C}_{n_3}^{\alpha} \) would be
\[ C_{an} = \frac{1}{\beta_a} \tilde{C}_{an} \]  

(2.29)

where \( \beta_a \) is known as the Pellinen smoothing factor of degree \( n \).

The relationship between \( \beta_a \) and the radius of the circular blocks is given in paragraph (4.3). For small blocks, experience shows that there is little difference between the area means of geodetic data on circular or on square blocks, so the error is small if one assumes that they are the same; in such case the modified expression

\[ \hat{C}_{an} = \frac{1}{4\pi} \beta_a \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \int_{\sigma_{ij}} \overline{\gamma}_{n}(\theta, \lambda) \ d\sigma = \tilde{C}_{an} \]  

(2.30)

could be used; in practice, this is only an approximation, though a good one, as showed by Katsambalos (1979), who tested this expression extensively.

In addition to (2.28) and (2.30), Lowes (1978) has proposed using

\[ \hat{C}_{an} = \frac{1}{4\pi} \beta_a \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \int_{\sigma_{ij}} \overline{\gamma}_{n}(\theta, \lambda) \ d\sigma \]  

(2.31)

to estimate the harmonic coefficients. All these expressions have the property that, because \( \beta_a \rightarrow 1 \), and \( \int \int T_{ij} \int_{\sigma_{ij}} \overline{\gamma}_{n}(\theta, \lambda) \ d\sigma \rightarrow \int_{\sigma_{ij}} f(\theta, \lambda) \overline{\gamma}_{n}(\theta, \lambda) \ d\sigma \)

as \( \Delta_{ij} \rightarrow 0 \) (or \( N_{p} \rightarrow \infty \)), it is true that the error \( e_{ap} = C_{an} - C_{an} \) is properly called a sampling error in the sense given to this term in paragraph (2.3).

Comparing (2.28), (2.30), and (2.31) it is easy to see that they all belong to a class of expressions of the form (2.27), with \( \mu_a = \frac{1}{4\pi} \), \( \mu_a = \frac{1}{4\pi \beta_a} \), and \( \mu_a = \frac{1}{4\pi \beta_a^2} \), respectively.

The scaling factor \( \mu \) can also be regarded as a de-smoothing factor, if one wishes to retain the intuitive meaning of these formulas. In the notation of paragraph (2.2), these expressions can be written, according to (2.27), as

\[ \hat{C}_{an} = \mu_a (h_{an}) \frac{\alpha}{a} \]  

(2.32)

with \( \mu_a = \overline{f}_{an} \).}

Replacing (2.32) in the definition of the sampling error measure, (2.23), and adding with respect to \( m \) and \( \alpha \) to obtain the total error in the nth harmonic:

\[ \overline{\sigma}_{n}^2 = \sigma_{n}^2 - \left[ 2 \sum_{\alpha=0}^{n} \overline{C}_{a} \frac{\alpha}{a} \overline{h}_{an} \right] \mu_a \]

\[ \left[ \sum_{\alpha=0}^{n} \left( h_{an} \frac{\alpha}{a} (C_{a}) \overline{h}_{an} \right) \right] \mu_a^2 \]  

(2.33)

-34-
(where $c_{m,n}^{r} r_{n}$ is a row of $C_{m,n}$). This is the sum of certain diagonal elements of $E_{n}$, according to (2.24), when the estimator has the form (2.27). Clearly, (2.33) is a quadratic function of the scalar $\mu_{n}$, and as such it can have either a maximum or a minimum. If $C_{zz}$ is positive definite, it must be a minimum. Finding the corresponding value of $\mu_{n}$ is the same as finding the formula of type (2.27) that has the smallest sampling error per harmonic for signals with the covariance (power spectrum) specified by $C_{zz}$. In addition to the sampling error, the measure of the propagated noise can be added to obtain

$$
\sum_{\alpha=0}^{n} \sum_{n=0}^{2} \sigma_{nn}^{2} \alpha^{2} \left[ \sum_{\alpha=0}^{n} \sum_{n=0}^{2} c_{m,n}^{r} \alpha_{n} \ h_{n}^{\alpha} \right] \mu_{n} + \left[ \sum_{\alpha=0}^{n} \sum_{n=0}^{2} \left( h_{n}^{\alpha} \right)^{r} \left( C_{zz} + D \right) h_{n}^{\alpha} \right] \mu_{n}^{2}
$$

(2.34)

This is also quadratic and has a minimum, and finding the optimum $\mu_{n}$ is the subject of the next paragraph.

2.7 Optimum de-Smoothing Factors

The coefficients of $\mu_{n}^{2}$ and $\mu_{n}$ in (2.33) and (2.34) are both real scalars, and so is the independent term $\sigma_{n}^{2}$. The expressions represent parabolas, and because both $C_{zz}$ and $D$ are positive, if the further (and likely) assumption is made that they are also definite, then $\frac{d^{2}}{d\mu_{n}^{2}} \sum_{\alpha=0}^{n} \sigma_{nn}^{2} \alpha^{2} > 0$, and the parabola has a minimum where $\mu_{n}$ satisfies the condition

$$
\frac{1}{2} \frac{d}{d\mu_{n}} \sum_{\alpha=0}^{n} \sum_{n=0}^{2} \sigma_{nn}^{2} \alpha^{2} \left[ \sum_{\alpha=0}^{n} \sum_{n=0}^{2} c_{m,n}^{r} \alpha_{n} \ h_{n}^{\alpha} \right] \mu_{n} + \left[ \sum_{\alpha=0}^{n} \sum_{n=0}^{2} \left( h_{n}^{\alpha} \right)^{r} \left( C_{zz} + D \right) h_{n}^{\alpha} \right] \mu_{n}^{2}
$$

for the sampling error (2.33)

i.e., at

$$
\mu_{n}^{*} = \frac{\sum_{\alpha=0}^{n} \sum_{n=0}^{2} c_{m,n}^{r} \alpha_{n} \ h_{n}^{\alpha}}{\sum_{\alpha=0}^{n} \sum_{n=0}^{2} \left( h_{n}^{\alpha} \right)^{r} \left( C_{zz} + D \right) h_{n}^{\alpha}}
$$

(2.35)

or at

$$
\mu_{n}^{*} = \frac{\sum_{\alpha=0}^{n} \sum_{n=0}^{2} c_{m,n}^{r} \alpha_{n} \ h_{n}^{\alpha}}{\sum_{\alpha=0}^{n} \sum_{n=0}^{2} \left( h_{n}^{\alpha} \right)^{r} \left( C_{zz} + D \right) h_{n}^{\alpha}}
$$

(2.36)

for the total error (2.34).

Expressions (2.28), (2.30), (2.31), and (2.36) will be studied further, by means of computed examples, in section 3.
2.8 Least Squares Collocation

In the notation of Paragraph (2.2), optimizing \( \mu_n \) is the same as obtaining the optimal vectors \( \hat{f}_n^\alpha \) of the form

\[
\hat{f}_n^\alpha = \mu_n \hat{h}_n^\alpha
\]

for the estimator

\[
\hat{c} = Fm
\]

where \( m = z + n \), and the \( (f_n^\alpha)^\top \) are the rows of \( F \). If no restriction is placed on the form the rows of \( F \) can take, then a reasoning similar to that in the preceding paragraph leads to the best possible linear estimator for \( C_{nn}^\alpha \).

Considering the total measure of error for \( 0 \leq n \leq N \):

\[
\sigma_c^2 = \sum_{\alpha=0}^{N-1} \sum_{n=0}^{n} \sum_{n=0}^{n} C_{nn}^\alpha - 2 \sum_{\alpha=0}^{N-1} \sum_{n=0}^{n} \sum_{s=0}^{s} C_{ns}^\alpha \hat{f}_n^\alpha + \sum_{\alpha=0}^{N-1} \sum_{n=0}^{n} \sum_{s=0}^{s} (f_n^\alpha)^\top (C_{zz} + D) f_n^\alpha
\]

(2.37)

it is not difficult to see that, because all \( C_{nn}^\alpha \) are non-negative, finding the \( F \) that minimizes their sum is the same as finding the \( F \) that minimizes them individually. The sum of the mean squared errors of all coefficients is the trace of the error matrix \( E_f \) of (2.26):

\[
\sigma_c^2 = \sum_{\alpha=0}^{N-1} \sum_{n=0}^{n} \sum_{s=0}^{s} C_{nn}^\alpha = \text{tr}[E_f]
\]

To obtain the condition for a minimum, one must differentiate (2.37) so, according to (2.26),

\[
\frac{1}{2} \frac{\delta \text{tr}[E_f]}{\delta F} = -C_{oz}^\top + (C_{zz} + D) F^\top
\]

(2.38)

as found using well-known matrix analysis formulas.

From this follows that

\[
F = C_{oz} (C_{zz} + D)^{-1}
\]

(2.39)

is the \( F \) that minimizes (2.37), provided that \( (C_{zz} + D) \) is positive definite. As already explained, both matrices are always positive and their sum is usually definite. The expression for the optimal estimator for \( C_{nn}^\alpha \) is

\[
C_{nn}^\alpha = (f_n^\alpha)^\top m = C_{nn}^\alpha \alpha \alpha^\top (C_{zz} + D)^{-1} m
\]

(2.40)
where \( \alpha^* \) is the row of the optimal estimator matrix \( F \) corresponding to \( \alpha \).

The use of expression (2.40) is, in brief, least squares collocation applied to spherical harmonic analysis.

When the optimal \( F \) is used, the error matrix becomes, according to (2.26) and (2.39),

\[
E_t = C - C_{zz} (C_{zz} + D)^{-1} C_{zz}^T
\]  

(2.41)

and the total error measure is the trace of this matrix: by definition, the smallest for all possible \( F \).

Clearly, whether one is interested in estimating coefficients or in determining the likely accuracy of such estimates, using expressions (2.40) or (2.41) require a knowledge of either \((C_{zz} + D)^{-1}\) (inversion) or, at least, of \( C_{zz} (C_{zz} + D)^{-1} \) (solution). Because the matrix \( C_{zz} + D \) has dimension \( N_p \times N_p \), obtaining either requires, by usual linear algebra methods, \( O(N_p^3) \) (or \( O(N_p^2) \)) operations. In the case of a \( 1^\circ \times 1^\circ \) grid, \( N_p = 64800 \), so, at some 200000 products and sums (double precision) per second, a modern computer like the one at OSU would need about one century to obtain all \( C_{zz}^\alpha \) to degree and order 180 from data on such a grid. Fortunately, as explained in paragraph (2.9), if the covariance functions of signal and noise both satisfy certain conditions, and if \( \Delta \lambda \) is constant for the whole grid, then both \( C_{zz} \) and \( D \) (consequently their sum) can be inverted in much fewer operations than by conventional methods, because they possess a particularly strong structure. Moreover, the optimal estimator \( C_{zz}^\alpha = (C_{zz})^{\alpha}_{m} \) turns out to be of the form (1.24) or (1.7), depending on the kind of data \( m \), so, under rather general conditions, the optimal estimator of \( C_{zz}^\alpha \) is also the best quadrature type formula for point data or for area means, as the case may be.

The conditions mentioned above are satisfied, for instance, when both the geometrical covariance \( \text{cov}(f(P), f(Q)) \) and the stochastic covariance \( E[n_{z1} n_{z2}] \) (\( P \equiv (\theta_1, \lambda_1) \), \( Q \equiv (\theta_2, \lambda_2) \) are isotropic, i.e., functions only of the separation between the points \( P \) and \( Q \). By definition of \( M \), the geometrical covariance obtained using this operator is isotropic, so \( C_{zz} \) has the desired structure. A common assumption regarding good instruments is that the \( n_{z1} \) are uncorrelated, so \( D \) is diagonal. If the errors are stationary, so their variances are constant, or at least constant along parallels, then matrix \( D \) has the required structure, and inverting \( C_{zz} + D \) can be greatly expedited. In practice, however, this is not likely to be the case, as the number and quality of measurements will vary from region to region, resulting in different \( \sigma_1^2 \) both globally and along parallels. As a result, the best linear estimator in terms of the chosen error measure will not have the
quadratures form, and it will be very difficult to compute when the number of data values is very large. Nevertheless, as shown in section 3, quadrature formulas can give reasonable estimates of the $\bar{C}_{\alpha\beta}$ with noisy data sets where the noise is uneven, so it would be interesting to get the best quadrature formula for a particular combination of signal and noise, provided that such a formula can be obtained without undue effort.

2.9 The Best Quadrature Formula for non-Uniform, Uncorrelated Noise

If the variance of the noise fluctuates along parallels, matrix $D$, though diagonal, is such that the minimization of the error measure (2.37)

$$\sigma_{\xi}^2 = \text{tr} \left[ C - 2C_{zz} F^T + F(C_{zz} + D) F^T \right] = \xi(D)$$

(see also (2.26))
can be very difficult with large $N_p$, and the optimal estimator is not of the quadrature type. Introducing a "modified noise matrix" $L$, also diagonal and where the diagonal elements are

$$\lambda_{kk} = \frac{1}{2N} \sum_{j=0}^{2N-1} \sigma_{i,j}^2 \quad \text{(with } k = 2NI + j)$$

(2.43)

the following modified error measure $\xi(L)$ can be defined:

$$\xi(L) = \text{tr} \left[ C - 2C_{zz} F^T + F(C_{zz} + L) F^T \right]$$

(2.44)

The optimal estimator for this measure is easy to obtain, and is of the quadratures type.

The parts of $\xi(D)$ and $\xi(L)$ that measure the sampling errors are identical, so any difference between the overall measures must come from the "noise propagation" parts $\text{tr} \left[ FF^T \right]$ and $\text{tr} \left[ FLF^T \right]$. If the estimator (not necessarily optimal) happens to be of the quadratures type, i.e., for point data:

$$\hat{C}_{\alpha\beta}^{\alpha} = \sum_{i=0}^{N-1} \sum_{j=0}^{2N-1} \chi_{i,j} \left[ \cos \right] m j \Delta \lambda \left[ f(\theta_1, \lambda_1) + n_1 \right]$$

(2.45)

then the propagated noise is, assuming the $n_{1i}$ to be uncorrelated,

$$\sigma_{n_{1i}}^2 = \sum_{i=0}^{N-1} \sum_{j=0}^{2N-1} \left( \chi_{i,j}^2 \right)^2 \left[ \cos \right] m j \Delta \lambda \left[ E_{1i}^2 \right]$$

so

$$\tilde{\xi}_\eta(D) = \text{tr} \left[ FF^T \right] = \sum_{\alpha=0}^{N-1} \sum_{\beta=0}^{2N-1} \sum_{i=0}^{n} \sigma_{\alpha\beta}^{\alpha} \sum_{i=0}^{N-1} \sum_{j=0}^{2N-1} \sum_{i=0}^{N-1} \sum_{j=0}^{2N-1} \sigma_{i,j}$$

(2.46)

The "modified noise" on the other hand, is, according to (2.43) and (2.44):
\[ \tilde{\eta}(L) = \text{tr} \left[ F L F^T \right] = \sum_{n=0}^{N-1} \sum_{a=0}^{n} \sum_{i=0}^{a} (\chi_i^a)^2 \sum_{j=0}^{2N-1} \left( \cos^2 m j \Delta \lambda + \sin^2 m j \Delta \lambda \right) \cdot \frac{1}{2N} \sum_{j=0}^{2N-1} \sigma_{ij}^2 = \sum_{n=0}^{N-1} \sum_{a=0}^{n} \sum_{i=0}^{a} (\chi_i^a)^2 \sum_{j=0}^{2N-1} \sigma_{ij}^2 \]  

(2.47)

Comparing the expressions for \( [F L F^T] \) and for \( [F D F^T] \), it follows that they are identical, and since the "sampling" parts are also identical in (2.42) and (2.44), then

\[ \text{tr} \left[ C - 2C_{zz} F^T + F(C_{zz} + L) F^T \right] = \text{tr} \left[ C - 2C_{zz} F_l^T + F(C_{zz} + D) F_l^T \right] \]

(2.48)

This means that the actual and the modified error measures must coincide if the estimator is of the quadratures type.

Replacing \( D \) with \( L \) in equation (2.38) and solving for the estimator matrix, one gets

\[ F_l = C_{zz} (C_{zz} + L)^{-1} \]

(2.49)

where \( F_l \) is the estimator matrix that minimizes the modified error measure (2.44). Because of the way \( L \) has been defined, this estimator is of the quadratures type, so the modified and the actual error measures coincide, as just shown.

Assume that there is an estimator, different from \( \hat{\xi} = F_l m \) but also of the quadratures type, the estimator matrix of which is \( \tilde{F} \), and such that:

\[ \text{tr} \left[ C - 2C_{zz} \tilde{F}^T + \tilde{F}(C_{zz} + D) \tilde{F}^T \right] < \text{tr} \left[ C - 2C_{zz} F_l^T + F_l(C_{zz} + D) F_l^T \right] \]

Then, according to (2.48),

\[ \text{tr} \left[ C - 2C_{zz} \tilde{F}^T + \tilde{F}(C_{zz} + L) \tilde{F}^T \right] < \text{tr} \left[ C - 2C_{zz} F_l^T + F_l(C_{zz} + L) F_l^T \right] \]

(2.50)

which contradicts the fact that \( F_l \) minimizes the modified error measure. Therefore, (2.50) cannot be true, and \( F_l \) must be the matrix of the optimal quadratures type estimator that minimizes the actual error measure \( \sigma_{\xi}^2 \).

The optimal quadratures type estimator, as the name indicates, is the best of a certain kind, not the absolute best. The best estimator, when no conditions as to its form are imposed, will not be (in general) of the quadratures type, unless \( D \) happens to have the "right form" specified before, i.e., unless \( D = L \).

When \( D = L \), the optimal estimator and the best quadrature formula coincide. Regardless of this, the quadrature formula obtained from (2.49) is the best, so its error measure is a lower bound for those of all other quadratures formulas with the given signal and noise.
When \( D \neq L \), while minimizing the sum of all error variances \( \sigma_{\alpha}^2 \), i.e., \( \text{tr} (E_\tau ) \), the optimal quadrature formula does not minimize each individual variance \( \sigma_{\alpha}^2 \). To show this, consider the propagated noise measure for \( C_{\alpha}^2 \) when the \( n_\tau \) are uncorrelated (to simplify the argument):

\[
\sigma_{\eta_{\alpha \alpha}} = \sum_{t \in = 0}^{n_\eta} \left( \chi_{i}^{(m)} \right)^{2 n_\eta - 2} \sum_{j = 0}^{m_\eta} \left[ \frac{\cos^2}{\sin^2} \right] m_\eta j \Delta \lambda \quad \sigma_{\eta_{\alpha \eta}} = \sum_{t \in = 0}^{n_\eta} \left( \chi_{i}^{(m)} \right)^{2 n_\eta - 2} \sum_{j = 0}^{m_\eta} \left[ \frac{\cos^2}{\sin^2} \right] m_\eta j \Delta \lambda \quad \sigma_{\eta_{\xi \xi}} = \sum_{t \in = 0}^{n_\eta} \left( \chi_{i}^{(m)} \right)^{-1} \sum_{j = 0}^{m_\eta} \sigma_{\xi_{\xi}}^2 \]

The modified error measure, minimized by the formula, is

\[
\sum_{t \in = 0}^{n_\eta} \left( \chi_{i}^{(m)} \right)^{2 n_\eta - 2} \sum_{j = 0}^{m_\eta} \left[ \frac{\cos^2}{\sin^2} \right] m_\eta j \Delta \lambda \quad \frac{1}{2 N} \sum_{j = 0}^{m_\eta} \sigma_{\xi_{\xi}}^2 = \sum_{t \in = 0}^{n_\eta} \left( \chi_{i}^{(m)} \right)^{-1} \sum_{j = 0}^{m_\eta} \sigma_{\xi_{\xi}}^2 \]

Clearly, both are not the same, unless

\[
\sum_{t \in = 0}^{n_\eta} \left[ \frac{\cos}{\sin} \right] m_\eta j \Delta \lambda \quad \sigma_{\xi_{\xi}}^2 = 0
\]

which is not likely to be fulfilled for arbitrary \( \sigma_{\xi_{\xi}}^2 \). However, looking at the reasoning which leads to (2.47), one can see that the sums of the modified and the actual error measures for pairs \( (C_{\alpha}, S_{\alpha}) \), and also for the individual \( C_{\alpha} \), are already identical. From this follows that the variance of the error per degree

\[
\sum_{\alpha \in \eta} \sum_{\eta \in \xi} \sigma_{\alpha \alpha}^2 = \sigma_{\xi_{\xi}}^2
\]

and per average coefficient per degree:

\[
\delta_{\xi_{\xi}}^2 = \frac{\sigma_{\xi_{\xi}}^2}{2n + 1}
\]

are also identical to the modified measure. So, while nothing can be predicated of individual harmonics, the error for each harmonic as a whole and that for the "average coefficient" in it are going to be minimum. By Parseval's identity (1.12), if the coefficients were used to calculate, say, geoidal undulations, the mean squared error of the computed geoid, globally, would be the same as the sum of the error squared of the normalized coefficients, so individual coefficient variances are of little interest in this and similar applications, while the \( \sigma_{\xi_{\xi}}^2 \) are very important. This shows that the optimal quadratures formula when \( D \neq L \) can be just as useful as when \( D = L \).

The discussion in this paragraph has been centered on point value type formulas; the conclusions apply equally well to area mean type formulas, the extension of the reasoning being quite straightforward. 

### 2.10 The Structure of the Covariance Matrix and its Consequences

The following discussion summarizes some results presented by this author in a previous report (Colombo, 1979a). In order to be able to calculate the variance of the error \( \sigma_{\alpha \alpha}^2 \) with expressions such as (2.26), and also to be able to obtain the optimal estimator according to collocation theory, it is necessary to create and invert the \( N_{p} \times N_{p} \) matrix \( (C_{\alpha} + D) \), which can be very
large if the number of data \( N_p \) is large. In the case of regularly sampled data this two problems can be greatly simplified if the covariances and the grid has certain symmetries. The most important of these are: (a) the sampling in longitude must be at constant intervals and along parallels (or parallel bands, i.e., rows of blocks); (b) for given \( i \) and \( p \) the covariances \( \text{cov}(\theta_i, \lambda_j), v(\theta_i, \lambda_j) \) (or \( \text{cov}(\bar{u}_i, \bar{v}_q) \) and \( E\{u_i, v_q\} \), must depend only on \(|j-q|\). It is also very advantageous, though not essential, that the grid be symmetrical with respect to the Equator.

In what follows \( N_e \) is the number of parallels and \( N_l \) the number of meridians \( (N_e = N, N_l = 2N \) when the grid is equal angular).

Under this set of conditions, if the data vector \( \mathbf{m} \) is ordered according to (2.8) and is subdivided into partitions \( \mathbf{m}_i \), where

\[
\mathbf{m}_i = [m_{i0} m_{i1} \ldots m_{iN_l-1}]^T
\]

includes all data values in the same parallel or row of blocks, then the matrix \( (C_{xz} + D) \) can be partitioned into \( N_e^2 \) blocks \( C_{ip} \), each of dimension \( N_l \times N_l \), containing the covariances between the data along rows \( i \) and \( p \).

Each block \( C_{ip} \) has a Toeplitz circulant structure, because its elements satisfy the relationships

\[
C_{ip} = C_{ip+i+1} ; \quad C_{ip} = C_{p-N_i} \quad \text{when} \ j > 0
\]

which follow from the fact that parallels are circular, and that the covariance between points in parallels \( i \) and \( p \) is a function of \(|j-q|\). Moreover, the elements in the first row or column (the \( C_{ip} \) are symmetrical) also satisfy

\[
C_{ip} = C_{pN_l-i} \quad \text{when} \ q > 0
\]

Therefore, the first row can be represented exactly as a sum of \( \frac{3}{2}N_l - 1 \) cosines:

\[
C_{ip} = \sum_{n=-N_l}^{N_l} a_n \cos \frac{2\pi}{N_l} q
\]

(2.53)

The \( a_n \) form the discrete Fourier transform of the sequence

\[
C_{00} , C_{01} , \ldots C_{0N_l-1}
\]
If

$$r_n^p = H a_n^p$$

(2.54)

where

$$H = \begin{cases} N1 & \text{if } m = 0 \\ N1 & \text{if } m \neq 0 \end{cases}$$

If \( R(m) \) is the matrix where each \( r_{ip} \) element equals \( r_n^p \), then (as explained by Colombo, (op. cit.)) inverting \( (C_{zz} + D) \) is equivalent to inverting the \( Nr \times Nr \) matrices \( R(m) \) for \( m = 0, 1, \ldots, \frac{N}{2} \). Isotropic covariances satisfy the "\( |j - q| \) condition" mentioned above, so, for fixed \( \Delta \lambda \), the covariance matrix always has this regular structure.

Let

$$\alpha_n = [\cos \ 0 \Delta \lambda, \ \cos \ m \Delta \lambda, \ldots, \cos \ m(Nr - 1) \Delta \lambda]^T$$

A vector of the type

$$\alpha = [v_0 \ \alpha_0^T, \ v_1 \ \alpha_1^T, \ldots, v_{N-1} \ \alpha_{N-1}^T]$$

(2.55)

shall be called, for convenience, a vector "of frequency \( m \)."

Under the conditions described before all the eigenvectors of \( (C_{zz} + D) \) are vectors of frequency \( m \), with \( m = 0, 1, \ldots, \frac{N}{2} \). Moreover, if \( \lambda_{t_n} (t = 1, 2, \ldots, Nr) \) is one of the \( Nr \) eigenvalues of \( R(m) \), and if

$$\xi_n = [s_0^t \ \ldots \ s_{N-1}^t]^T$$

is the corresponding eigenvector of \( R(m) \), then \( \lambda_{t_n} \) is also an eigenvalue of \( (C_{zz} + D) \), and the pair

$$\alpha_n = [s_0^t \ \alpha_0^T, \ldots, s_{N-1}^t \ \alpha_{N-1}^T]^T$$

the two corresponding eigenvectors of \( (C_{zz} + D) \). Therefore, to decompose the large covariance matrix in eigenvectors and eigenvalues is equivalent to decomposing the \( \frac{N}{2} \) \( N \times 1 \) matrices \( R(m) \), and this is why the latter are relevant to the inversion of \( (C_{zz} + D) \): the eigenvalues of the inverse are the reciprocal of the \( \lambda_{t_n} \), while its eigenvectors are the same as the \( \xi_n \). Further, this implies that \( (C_{zz} + D)^{-1} \) has the same structure as the covariance matrix, i.e., it consists of Toeplitz-circulant blocks.

Since \( (C_{zz} + D)^{-1} \) has eigenvectors of frequency \( m \), then, if \( h \) is a linear combination of vectors of a given frequency, \( z = (C_{zz} + D)^{-1} h \) is also a linear combination of vectors of that frequency. In the case of point data, from expression (2.6) follows that the cross-covariances vector in (2.40) is
\[ c_{n_k}^T \alpha_z = \frac{\sigma_n}{2n+1} \left[ \bar{p}_{n_k} (\cos \theta_0) \ c_{n_k}^T \ldots \bar{p}_{n_k} (\cos \theta_{n_k-1}) \ c_{n_k}^T \right] \]
\[ = [k_0 \ c_{n_k}^T \ldots k_{n_k-1} \ c_{n_k}^T] \]

Define
\[ k_{n_k} = [k_{n_k}^T \ldots k_{n_k-1}^T]^T \]
and
\[ \chi_{n_k} = [\chi_0^T \ldots \chi_{n_k-1}^T]^T \]

Then \[ \chi_{n_k}^T = (C_{n_k} + D)^{-1} \ c_{n_k} \alpha_z \] must be of the form
\[ \chi_{n_k}^T = [\chi_0^T \ldots \chi_{n_k-1}^T] \]

where, according to Colombo (ibid),
\[ \chi_{n_k} = R(m)^{-1} k_{n_k} \]

Similarly, for area means,
\[ c_{n_k} \alpha_z = \frac{\sigma_n^2}{(2n+1) \Delta_1} \left[ \int_{\theta_1}^{\theta_1+\Delta \theta} \bar{p}_{n_k} (\cos \theta) \sin \theta d\theta \int_{\lambda_1}^{\lambda_1+\Delta \lambda} c_{n_k}^T \ d\lambda \ldots \right]^T \]
\[ = \frac{\sigma_n^2}{(2n+1) \Delta_1} \left[ \int_{\theta_1}^{\theta_1+\Delta \theta} \bar{p}_{n_k} (\cos \theta) \sin \theta d\theta \left( \left\{ \begin{array}{c} A(m) \\ B(m) \end{array} \right\} c_{n_k}^T + \left\{ \begin{array}{c} A(m) \\ B(m) \end{array} \right\} \right)^T \right] \ldots \] (2.59)

(Where \( \Delta_1 \) is supposed to be independent of \( j \)) according to expressions (1.7) and (2.7), so
\[ \chi_{n_k} = [\ldots \chi_{n_k}^T (\left\{ \begin{array}{c} A(m) \\ B(m) \end{array} \right\} c_{n_k}^T + \left\{ \begin{array}{c} A(m) \\ B(m) \end{array} \right\}^T \ldots ]^T \]

In conclusion, the optimal estimator for point values has the form
\[ \hat{\alpha}_{n_k} = \chi_{n_k}^T m = \sum_{i=0}^{N_k-1} \sum_{j=0}^{N_k-1} \chi_{n_k}^T \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} m_j \Delta \lambda m_{i,j} \]
(2.60)

While that for area means is of the type
\[ \hat{\alpha}_{n_k} = \chi_{n_k}^T m = \sum_{i=0}^{N_k-1} \sum_{j=0}^{N_k-1} \chi_{n_k}^T \begin{bmatrix} A(m) \\ B(m) \end{bmatrix} \cos m_j \Delta \lambda + \begin{bmatrix} A(m) \\ B(m) \end{bmatrix} \sin m_j \Delta \lambda \] m_{i,j} \]
(2.61)

So they are both of the quadratures kind, as anticipated in the preceding paragraph.

2.11 Setting up and Inverting the Covariance Matrix

Each block \( C_{n_k} \) of \( (C_{n_k} + D) \) is wholly determined by the 1st \( \frac{1}{2} N_k + 1 \) elements in its first row; if the number of operations required to compute any
element of $C^{\oplus}$ is $k$, then only $(2Nl + 1)k$ operations are needed per block, instead of $Nn_{1}k$, as would be the case if $C^{\oplus}$ did not have the Toeplitz structure described previously. This is a reduction of the number of operations by a factor of $2Nl$, and clearly applies not only to $C^{\oplus}$ but to the whole covariance matrix as well. So $(C_{zz} + D)$ can be set up about $2Nl$ times faster than an ordinary matrix of the same size.

The total number of elements to be computed is $\frac{1}{2}Nl \times N^{2}$, or $N^{3}$ in the case of equal angular grids. If the grid is a fine one, this can still be a very large number of covariances. This is particularly serious in the case of area means, because the area mean covariances are given by expressions of the form

$$\text{cov}(\bar{u}_{1j}, \bar{u}_{pq}) = M \left\{ \int_{\sigma_{1j}} u d\sigma \int_{\sigma_{pq}} u d\sigma \right\} = \int_{\sigma_{1j}} \int_{\sigma_{pq}} M \left\{ u(\theta, \lambda) \ u(\theta', \lambda') \right\} d\sigma d\sigma'$$

(2.62)

involving double area integrals of the covariance function. Numerical quadratures methods, such as the one described in paragraph (4.3), have been used in the past to obtain $\text{cov}(\bar{u}_{1j}, \bar{u}_{pq})$ (see, for example, Rapp (1977)). These methods take so much time in the case of fine equal angular grids for instance, that it may be practically impossible to use them to set up the covariance matrix of a global data set, in spite of the reduction by $2Nl$ in the number of operations. Fortunately, the coefficients $a_{n}^{\oplus}$ in the Fourier expansion of the elements

$$c_{1j}^{l} = \text{cov}(\bar{u}_{1j}, \bar{u}_{pq}) + E \left\{ \bar{n}_{1j} \bar{n}_{pq} \right\}$$

(expression (2.53)) can be obtained by means of a series expansion (truncated to a conveniently high degree $N_{m_{x}}$) according to expression (4.14) in paragraph (4.1). These coefficients are

$$a_{n}^{\oplus} = \left[ \sum_{h=0}^{K} \sum_{n=-N_{m_{x}}}^{N_{m_{x}}} I_{n, 2n h + n, h} + \sum_{h=0}^{K} \sum_{n=-N_{m_{x}}}^{N_{m_{x}}} I_{n, 2Nl n h + n, 0} \right] F(m)$$

where

$$I_{n, z, 1} = \int_{\sigma_{1j}} \int_{\sigma_{pq}} \frac{\cos \theta \sin \theta d\theta}{\sqrt{2n + 1}} \frac{\sigma_{n}}{\Delta \lambda} \frac{1}{\cos(\theta_{i} + \Delta \theta)}$$

(2.63)

$$F(m) = \begin{cases} \frac{\Delta \lambda^{2}}{(2m^{2})} & \text{if } m = 0 \\ \frac{1}{2m^{2}} & \text{if } (2h + 1) \ N \leq N_{m_{x}} \end{cases}$$

A similar reasoning to that for area means leads to an analogous formula for point values:

$$a_{n}^{\oplus} = \left[ \sum_{h=0}^{K} \sum_{n=-N_{m_{x}}}^{N_{m_{x}}} (\bar{P}_{n2n h + n} \cos \theta_{i}) + \sum_{h=0}^{K} \sum_{n=-N_{m_{x}}}^{N_{m_{x}}} (\bar{P}_{n2Nl n h + n} \cos \theta_{p}) \right] \frac{\sigma_{n}^{2}}{2n + 1}$$

(2.64)
The importance of (2.63) and (2.64) is that, if the signal and noise are such that the number of terms in the summations is not too large (\(N_{\text{max}}\) is a "manageable" number), they allow the direct determination of the elements of the \(R(m)\) matrices according to (2.54). In this way, the \(R(m)\) can be created without first having to set up the whole covariance matrix and then to obtain the discrete Fourier transform of the first row of each \(C^b\). This advantage further increases in the case when the grid is symmetrical with respect to the equator, a situation that applies to all equal angular grids. Then each \(R(m)\) is persymmetrical, i.e., symmetrical with respect both to the main diagonal and the main antidiagonal, provided \(D\) is also persymmetrical (for instance, uniform noise). This means that only approximately \(\frac{1}{4}N_r^3\) elements in each \(R(m)\) are different and have to be calculated individually.

Having set up the \(R(m)\) without first creating the covariance matrix, the inverse of \((C_{zz} + D)\) can be found by the equivalent operation of obtaining all \(R(m)^{-1}\). The number of operations in a matrix inversion is usually \(O(\text{dimension}^3)\), or \(O(N^6)\) for a covariance matrix of an equal angular data set. The number of operations per \(R(m)\) is \(O(N_r^2)\), or \(O(N^3)\) for the equal angular grid. In fact, as explained in (Colombo, 1979a), the inversion of a persymmetrical \(R(m)\) is equivalent to that of the two matrices of half its dimension, one related to vectors of frequency \(m\) of the cosine type, and the other to vectors of the same frequency of the sine type. This further reduces calculation by a factor of \(\frac{1}{4}\). With \(O(N)\) \(R(m)\) matrices to be inverted, the total comes to \(O(N^4)\) operations, or \(O(N^3)\) times less than for the inversion of \((C_{zz} + D)\) by ordinary techniques (Choleskii factorization, Gauss-Jordan elimination, etc.). \(O(N^2)\) is also the order of the number of data points in the grid, so in the case of a \(1^\circ \times 1^\circ\) equal angular grid with 64800 elements the reduction in computing time is \(O(64800)\).

The numerical examples in section 3 all involve \(5^\circ \times 5^\circ\) data sets with 2592 elements, so \((C_{zz} + D)\) is of dimension 2592. Setting up and inverting such a matrix is a large exercise, even with a modern digital computer such as the AMDHAL 470 at Ohio State, unless the matrix has a strong structure that can be exploited to simplify the work. As such is indeed the case here, the subroutine NORMAL described in Appendix B has been able to do the whole setting up and inversion in only 20 seconds.

The inversion of \((C_{zz} + D)\) requires \(O(\text{dimension}^2)\) operations \(O(N^4)\) instead of \(O(\text{dimension}^3)\) because of the Toeplitz circulant structure of the \(C^b\) blocks. This "\(O(\text{dimension}^2)\)" property is common to other algorithms for inverting Toeplitz-type matrices, such as the famous Trench algorithm (Trench, 1965), and the Justice algorithm (Justice, 1977), the first for data sampled on the real line and the second for data sampled on the plane. So, in spite of its "rather wicked" nature, the sphere allows this very convenient property of regular grids to apply also on its surface. In fact, not only on the sphere, but also on any body of revolution (cone, oblate and prolate spheroid, hyperboloids and paraboloids of revolution, etc.), regular sampling and
covariances that satisfy the "|j-q| condition" will result in covariance matrices of the type described here, and this is also true of other matrices based on symmetrical kernels, such as the normal matrices of point mass models, when the points belong to a regular grid, etc. Finally, the optimal estimator \( \hat{C}_{\alpha} = \frac{1}{m} \sum_{m} \) for this type of covariance matrix is, as shown in the previous paragraph, of the quadratures type, so the optimal \( \hat{C}_{\alpha} \) can be obtained using the same efficient algorithms described in paragraphs (1.5)-(1.7). Altogether, the powerful structure of the covariance matrix for regular global data sets is most remarkable. One of its many advantageous features is that, because the creation and inversion of each \( R(m) \) can be done quite independently from those of the others, the algorithms developed for this type of matrices are eminently suited for implementation in parallel processing computers.

The separation of the algorithm according to orders also means that, although setting up and inverting all the \( R(m) \) may require a large number of operations, only a fraction of those actually correspond to the recovery of the \( \hat{C}_{\alpha} \) of any given \( m \), so the numerical errors due to rounding or truncation are not likely to accumulate to any great extent in the results.

2.12 Optimal Formulas for non-Uniform, Correlated Noise

Irregular noise, already discussed in paragraph (2.9), may be due not only to the varying quality of the measurements, but also to the way the data is "gridded", i.e., the way the value attributed to a node (or block) \( ij \) is obtained by interpolation from actual measurements nearby, as usually data is not sampled regularly on a global basis. As the number, disposition, and quality of the measurements used will vary from point to point in the grid, so will the accuracy of the interpolated values. Furthermore, even if the measurements themselves are not correlated, the gridded values may be correlated because some of the data may be used for more than one interpolated value. This brings about the question of what can be done when \( D \) is neither diagonal, nor are the \( D^{tr} \) blocks in \( D \), corresponding to the \( C^{tr} \) blocks in \( C_{zz} \), all Toeplitz circulant. The answer is a simple extension of the results already obtained for the uncorrelated case.

When the noise is both non-stationary and correlated, replacing the covariances \( E \{ n_{1}, n_{s} \} \) with

\[
\tilde{\sigma}^{tr}[h] = \frac{1}{4N} \left[ \sum_{h=0}^{N-1} (E \{ n_{1}, n_{r+h} \} + E \{ n_{1}, n_{r-j+h} \} ) \right], \quad \text{where} \quad h = j - s,
\]

will result in a modified "noise matrix" \( L \) where the \( L^{tr} \) (corresponding to the \( D^{tr} \) and the \( C^{tr} \)) will be all Toeplitz circulant, because the "covariance" \( \sigma^{tr}[h] \) satisfies the condition that, for a given \( i \) and \( r \), it is a function of \(|j-s|\) alone. The optimal estimator for the modified measure

\[
\hat{\mathcal{J}}(L) = \text{tr}[C - 2C_{zz} F^{tr} + F(C_{zz} + L) F]
\]

-46-
must be of the quadratures type, because of the structure of \( L \). To show
that is also the best estimator of this kind in terms of the original norm

\[
\hat{\delta}(D) = \text{tr} \left[ C - 2C_{\eta z} F^T + F(C_{\eta z} + D) F^T \right],
\]

the proof will proceed much as in the case of paragraph (2.9).

The propagated error measure for \( \hat{\Delta}_{n_{1r} n_{r_{1r}}}^{n_{1r} n_{r_{1r}}} \) is

\[
\sigma_{n_{1r} n_{r_{1r}}}^{n_{1r} n_{r_{1r}}} = E \left\{ \sum_{i=0}^{n_{1r}} \sum_{j=0}^{N_{1r}} \sum_{r=0}^{n_{r_{1r}}} \sum_{s=0}^{N_{r_{1r}}} \chi_i^{n_{1r}} \chi_j^{n_{1r}} \chi_r^{n_{r_{1r}}} \chi_s^{n_{r_{1r}}} \cos (j-s) \Delta \lambda \right\}
\]

so, for \( \hat{\Delta}_{n_{1r} n_{r_{1r}}}^{n_{1r} n_{r_{1r}}} \) and \( \hat{\Delta}_{n_{1r} n_{r_{1r}}}^{n_{1r} n_{r_{1r}}} \) combined,

\[
\sum_{n_{1r} n_{r_{1r}}} \sigma_{n_{1r} n_{r_{1r}}}^{n_{1r} n_{r_{1r}}} = E \left\{ \sum_{n_{1r} n_{r_{1r}}} \sum_{r=0}^{n_{r_{1r}}} \sum_{s=0}^{N_{r_{1r}}} \chi_i^{n_{1r}} \chi_j^{n_{1r}} \chi_r^{n_{r_{1r}}} \chi_s^{n_{r_{1r}}} \cos (j-s) \Delta \lambda \right\}
\]

thus

\[
\hat{\delta}_{\eta}(D) = \text{tr} \left[ FDF^T \right] = \sum_{n_{1r} n_{r_{1r}}} \sum_{n_{1r} n_{r_{1r}}} \sum_{r=0}^{n_{r_{1r}}} \sum_{s=0}^{N_{r_{1r}}} \chi_i^{n_{1r}} \chi_j^{n_{1r}} \chi_r^{n_{r_{1r}}} \chi_s^{n_{r_{1r}}} \cos (j-s) \Delta \lambda \right\}
\]

Replacing both \( E \{ n_{1r} n_{r_{1r}} \} \) and \( E \{ n_{1r} n_{r_{1r}} \} \) in the last expression with

\[
\hat{\sigma}_{\eta}^{1 \times |n|}
\]

is the same as replacing \( D \) with \( L \), so

\[
\hat{\delta}_{\eta}(L) = \text{tr} \left[ FLF^T \right] = \sum_{n_{1r} n_{r_{1r}}} \sum_{n_{1r} n_{r_{1r}}} \sum_{r=0}^{n_{r_{1r}}} \sum_{s=0}^{N_{r_{1r}}} \chi_i^{n_{1r}} \chi_j^{n_{1r}} \chi_r^{n_{r_{1r}}} \chi_s^{n_{r_{1r}}} \cos (j-s) \Delta \lambda \right\}
\]

because of the definition of \( \hat{\sigma}_{\eta}^{1 \times |n|} \). Comparing the expressions for \( \hat{\delta}_{\eta}(L) \)
and for \( \hat{\delta}_{\eta}(D) \) it is clear that they are identical. From this follows that the
modified error measure \( \hat{\delta}_{\eta}(L) \) coincides with \( \hat{\delta}(D) \) when the estimator is the
optimal estimator of the quadratures type for \( \hat{\delta}_{\eta}(L) \), and that this must be the
optimal estimator of the quadratures type for \( \hat{\delta}(D) \) as well. The other con-
clusions arrived at in paragraph (2.9) for the uncorrelated case apply equally
well here.

2.13 Least Squares Adjustment, and Least Squares Collocation

(a) Band-Limited Signal

If there is a degree \( N_{\text{ax}} \) above which the degree variances \( \sigma^2_{\eta} \) are all
negligible or zero, then the signal can be said to be band limited, and the data
will satisfy equations of the type
\[ m_{ij} = \sum_{n=0}^{N_{max}} \sum_{s=0}^{n} \sum_{\lambda=0}^{C_{nn} \lambda_{ij}} \frac{C_{nn} \alpha_{ij}}{\bar{y}_{nn}} (\theta_{ij}, \lambda_{ij}) + n_{ij} \]  (2.65)

(the treatment here is for point values; the extension to area means is trivial)

With one equation such as (2.65) per point in the grid, the result is a system of equations

\[ m + y = Ac \quad (y = -n) \]  (2.66)

where \( A \) is a \( N_p \times N_o \) matrix (\( N_p \) is the number of points in the grid, \( N_o \) the number of coefficients). The columns of \( A \) consist of elements of the type

\[ a_{ij}^{nn} = \bar{y}_{nn} (\theta_{ij}, \lambda_{ij}) \]  (2.67)

According to the discussion in paragraph (1.3), if the grid is equal angular, \( A \) has full rank when \( N_{max} < N \). \( N_o = N^2 \) and the upper limit in the summations is \( N - 1 \) in what follows.

Least squares adjustment is a method for solving for the \( C_{nn}^{\alpha} \) while minimizing the propagated noise defined in paragraph (2.4). The least squares solution is

\[ \hat{c} = (A^T D^{-1} A)^{-1} A^T D^{-1} m \]  (2.68)

where

\[ G = A^T D^{-1} A \]  (2.69)

is the \( N_c \times N_c \) normal matrix, while \( D = E \{ n n^T \} \) is the same noise matrix considered before. Clearly, the least squares estimator matrix is

\[ F_{ls} = (A^T D^{-1} A) A^T D^{-1} \]

When the noise has zero mean \( (E \{ n \} = 0) \), the estimator of (2.68) is the best linear unbiased estimator, because it minimizes

\[ \text{tr} \left[ E \{ F n n^T F^T \} \right] = \text{tr} \left[ FD F^T \right] \]

and \( E \{ F (z + n) \} = 0 \). If, in addition to all this, the probability distribution of the noise is Gaussian, then (2.68) corresponds to the maximum likelihood estimator as well. In many scientific applications the noise has approximately zero mean and near-Gaussian distribution, while \( D \) is known reasonably well;
for this reason, methods based on expression (2.68) are used quite often. The linearity of the resulting estimators is helpful, because this avoids the use of methods based on non-linear formulas that are usually difficult both from a theoretical and from a practical point of view. The variances of the estimates are given by the corresponding diagonal elements of the a posteriori variance-covariance matrix

$$E_{\delta} = (A^T D^{-1} A)^{-1} = G^{-1}$$

Therefore, to obtain both the estimates and their variances it is necessary to know $G^{-1}$. Sometimes, because of the nature of $A$ and $D$, $G$ can be seriously ill-conditioned, the inversion suffering from strong numerical instabilities. To reduce this problem, a simple device known as regularization is often used (see, for instance, Tikhonov and Arsenin, 1977). Generally speaking, regularization is the introduction of a slight change in a problem, so the solution virtually remains the same, but the modified problem has better numerical properties. In least squares methods regularization usually implies adding a small positive definite-matrix $K$ (diagonal, as a rule) to $G$ before attempting to invert it.

The regularized optimal estimator would be

$$\hat{c} = (A^T D^{-2} A + K)^{-1} A^T D^{-2} m$$

The inverse of the covariance matrix of the harmonic coefficients $C$ is a positive, diagonal matrix which could be used to regularize the normal matrix:

$$\hat{\delta} = (A^T D^{-3} A + C^{-1}) A^T D^{-3} m$$

(2.71)

It is easy to see that this expression minimizes the quadratic form

$$Q = \hat{c}^T C^{-1} \hat{c} + \hat{v}^T D^{-1} \hat{v}$$

(2.72)

subject to the constraint

$$m = A \hat{c} + \hat{v}$$

(2.73)

Moreover, (2.72) is the equivalent of the least squares collocation error measure when then signal is band-limited (Moritz, 1980). This idea has been used, among others, by Schwarz (1975) for the determination of low degree zonal coefficients of the geopotential, and by Lerch et al. (1979), who employed it to stabilize the adjustment of the GEM-9 gravity field model with remarkable success. The equivalence of (2.71) to the collocation estimator is true only for band-limited signals; in the "real world" the gravity field has infinite bandwidth, so (2.71) is no more than an approximation. The band-limited assumption is a reasonable one, however, as the $\sigma_n$ eventually become negligible for large $n$. This is particularly true at satellite altitudes; in any case, geodesy is science of wise approximations. Moritz (1980) has provided a very clear and concise explanation of the use of collocation in general, and expression (2.71) in particular, in spherical harmonic analysis. -49-
An alternative derivation of (2.71) follows from the matrix equation

\[ CA^T (ACA^T + D)^{-1} = (A^T D^{-1} A + C^{-1})^{-1} A^T D^{-1} \]  \hspace{1cm} (2.74)

(see, for instance, Utillia (1978), equation (29)), which is valid for symmetrical matrices, provided the inverses or pseudo-inverses of \( D, C \), and \( (ACA^T + D) \) do exist. According to the definitions in paragraph (2.2):

\[ C_{zz} = M \left\{ C \begin{bmatrix} z \end{bmatrix}^T \right\} = M \left\{ C \begin{bmatrix} c \end{bmatrix} A^T \right\} = M \left\{ C \begin{bmatrix} c \end{bmatrix} \right\} A^T = CA^T \]  \hspace{1cm} (2.75)

\[ C_{zz} = M \left\{ A \begin{bmatrix} c \end{bmatrix} A^T \right\} = A M \left\{ A \begin{bmatrix} c \end{bmatrix} \right\} A^T = ACA^T \]  \hspace{1cm} (2.76)

Replacing \( C_{zz} \) and \( C_{zz} \) in the expression of the collocation estimator matrix (2.39) with their equivalents given by (2.75) and (2.76):

\[ F = C_{oz} (C_{zz} + D)^{-1} = CA^T (ACA^T + D)^{-1} \]

\[ = (A^T D^{-1} A + C^{-1})^{-1} A^T D^{-1} \]  \hspace{1cm} (2.77)

according to equation (2.74). This shows that the "regularized" estimator matrix \( (A^T D^{-1} A + C^{-1})^{-1} A^T D^{-1} \) is indeed the same as the collocation estimator matrix, so (2.71) represents an alternative form of collocation when the data is band-limited.

(b) **Infinite Bandwidth**

In this case the "observation equations" are

\[ m_{ij} = \sum_{n=0}^{\infty} \sum_{s=0}^{n} \sum_{\alpha=0}^{1} C_{\alpha n}^{\alpha} \bar{y}_{\alpha n}^{\alpha} (\theta_1, \lambda_1) + n_{ij} \]

Calling

\[ w_{1j} = \sum_{n=N+1}^{\infty} \sum_{s=0}^{n} \sum_{\alpha=0}^{1} C_{\alpha n}^{\alpha} \bar{y}_{\alpha n}^{\alpha} (\theta_1, \lambda_1) \]  \hspace{1cm} (2.78a)

and

\[ \underline{w} = [w_1, \ldots, w_k, \ldots, w_{N+1}]^T \]  \hspace{1cm} (k = 2N+1+j)  \hspace{1cm} (2.78b)

then

\[ m_{ij} = \sum_{n=0}^{\infty} \sum_{s=0}^{n} \sum_{\alpha=0}^{1} C_{\alpha n}^{\alpha} \bar{y}_{\alpha n}^{\alpha} (\theta_1, \lambda_1) + w_{1j} + n_{ij} \]  \hspace{1cm} (2.79)

and, regarding this expression as a modified observation equation, and replacing \( D \) with \( D + M \left\{ \underline{w} \underline{w}^T \right\} \) in (2.71), the linear estimator that minimizes the quadratic form

\[ \bar{Q} = \bar{c}^T C^{-1} \bar{c} + v^T (D + M \left\{ \underline{w} \underline{w}^T \right\} )^{-1} v \]  \hspace{1cm} (2.80)

is

\[ \hat{\theta} = (A^T (D + M \left\{ \underline{w} \underline{w}^T \right\} )^{-1} A + C^{-1})^{-1} A^T (D + M \left\{ \underline{w} \underline{w}^T \right\} )^{-1} m \]  \hspace{1cm} (2.81)

It is easy to show, either following the lines of Moritz (1972), or going
back once more to the matrix identify (2.74), that expression (2.81) is identical with the estimator of least squares collocation. Rigorously speaking, expression (2.81) should be used whenever $\sigma_n^2 \neq 0$ for $n > N$, even if $\sigma_n^2 = 0$ for $n > N_{\text{max}}$, for some finite $N_{\text{max}} > N$.

2.14 Ridge Regression and Least Squares Collocation

Consider once more the estimator

$$\hat{\theta} = F_{\ell_2} m = (A^T D^{-1} A)^{-1} A^T D^{-1} m$$

If there is no noise, so $m = z = A \hat{c}$, and if $n < N$, then

$$\hat{c} = (A^T D^{-1} A)^{-1} A^T D^{-1} A \hat{c} = \hat{c}$$

According to the definition given in paragraph (2.4), the sampling error of this estimator is zero, so the measure of this error must be also zero. If the noise has zero mean, it follows that

$$E\{\hat{c}\} = E\{F_{\ell_2} m\} = F_{\ell_2} E\{A \hat{c} + n\} = F_{\ell_2} A \hat{c} + E\{n\} = \hat{c}$$

or, as it is usually said, the estimator is unbiased. Moreover, by a simple extension of the Gauss-Markov theorem to the case of a general symmetrical positive matrix $D$ (see, for instance, Bibby and Toutenburg, 1977), $F_{\ell_2} m$, of all linear unbiased estimators, has the least propagated error measure

$$\text{tr}\{F_{\ell_2} D F_{\ell_2}^T\} = \text{tr}\{(A^T D^{-1} A)^{-1}\}$$

as mentioned previously.

The estimator of expression (2.77) does not, in general, give perfect estimates of $\hat{c}$ in the absence of noise: it is a biased estimator, and the measure of the bias is $\text{tr}\left\{C - 2 C_{zz} F' + F C_{zz} F'\right\}$ (this term can no longer be regarded as the measure of the sampling error, as it is the presence of $C^{-1}$ inside the parenthesis in (2.71) and not the sampling that brings about this error). According to (2.39), $F$ is the estimator that minimizes the total error measure, so

$$\sigma_\varepsilon^2 = \text{tr}\left\{C - 2 C_{zz} F' + F(C_{zz} + D) F'\right\}$$

$$\leq \text{tr}\left\{C - 2 C_{zz} F_{\ell_2}^T + F_{\ell_2} (C_{zz} + D) F_{\ell_2}^T\right\}$$

$$= \text{tr}\left\{F_{\ell_2} D F_{\ell_2}^T\right\}$$

If the covariance matrix is positive definite (which only requires that all $\sigma_n^2 \neq 0$ for $0 < n < N$) then the last expression applies with strict inequality

$$\text{tr}\left\{C - 2 C_{zz} F' + F(C_{zz} + D) F'\right\} < \text{tr}\left\{F_{\ell_2} D F_{\ell_2}^T\right\}$$

$$= \text{tr}\left\{(A^T D^{-1} A)^{-1}\right\}$$

1 In the band-limited case.
Some may find this result rather surprising: the best estimator with zero bias is in fact worse than the biased estimator of (2.71)! The difficulty is only apparent: \( F_{k}^{*} \) is the best estimator with no bias; once the "no bias" condition is removed, the expression above merely indicates that there is an estimator in the larger class of the estimators that have a bias (including those with zero bias) such that the sum of the bias and the propagated noise is smaller. From this, it is clear that

\[
\text{tr} \left\{ FDF^{\dagger} \right\} < \text{tr} \left\{ (A^{\dagger}D^{-1}A) \right\} \quad \text{(as } \text{tr} \left\{ C - 2C_{zz}F^{\dagger} + FC_{zz}F^{\dagger} \right\} \geq 0 )
\]

which is indeed possible when the condition \( C - 2C_{zz}F^{\dagger} + FC_{zz}F^{\dagger} = 0 \) is removed. In fact, there is nothing very new about all this: the use of biased estimators to obtain estimates with small variances is a reasonably well-established practice in applied statistics. In particular, the technique known as **ridge regression** consists in using the biased estimator

\[
\hat{c} = (X^{\dagger}X + K)^{-1} X^{\dagger} m
\]

with a suitable choice of \( K \) (Bibby, 1972). Clearly, this expression is the same as (2.71) when \( X = A, D = I, \) and \( K = C^{-1} \). With some obvious modifications suggested by (2.81), this argument can be extended to the estimation of \( c \) when \( n \geq N \), so it can be said that within the scope of spherical harmonic analysis least squares collocation is a form of ridge regression.

This brings up the question of just how realistic the error measure is; after all the best of all possible estimators in terms of a given norm could be a very bad one for some specific problem where that norm is not suitable. To answer this question, one must start by defining the meaning of "realistic". If one is interested in minimizing the actual error variance of the coefficients per degree, i.e., the expression

\[
\delta_{n}^{2} = \sum_{\alpha = 0}^{L} \sum_{n = 0}^{N} (C_{\alpha n} - \overline{C}_{\alpha n})^{2} (2n + 1)^{-1}
\]

which may be of interest because this corresponds to, say, the global mean square of the error of representing the continuous function \( (\theta, \lambda) = \sum_{\alpha = 0}^{\infty} \sum_{n = 0}^{N} C_{\alpha n} \bar{X}_{\alpha n} (\theta, \lambda) \) with \( f(\theta, \lambda) = \sum_{\alpha = 0}^{\infty} \sum_{n = 0}^{N} \bar{C}_{\alpha n} \bar{X}_{\alpha n} (\theta, \lambda) \), according to Parseval's identity (exp. (1.12)), then one could say that a realistic measure is one that gives close estimates of the actual error variances. The "actual error" measure \( \delta_{n}^{2} \) corresponds to one of infinitely many "events" over which the collocation measure is an average. The proof of a pudding being in the eating, the reader can judge just how realistic the collocation measure is by looking at the numerical results in section 3, where the...
value of the measure turns out to differ only by a small percentage from the actual variance in each one of a number of simulated "events", i.e., the recovery of the \( \hat{C}^\alpha_{mn} \) from "simulated data", where the \( \hat{C}^\alpha_{mn} \) are known random numbers scaled to have the desired power spectrum.

2.15 Structure of the Normal Matrix

The elements of the normal matrix \( G \) (in the case of point data) are of the form

\[
g_{\alpha n \beta m} = \sum_{i=0}^{n-1} \int_{2 \pi}^{2 \pi} \hat{P}_{n \alpha} (\cos \theta_i) \hat{P}_{m \beta} (\cos \theta_i) \sum_{j=0}^{2N-1} \left[ \begin{array}{c} \cos \\ \sin \end{array} \right] \begin{array}{c} m j \Delta \lambda \\ q j \Delta \lambda \end{array} \sigma_i^2 \bar{\sigma}_i^2 \]

where \( \sigma_i^2 = \bar{\sigma}_i^2 \) for all \( 0 \leq j \leq 2N \) (i.e., "regular" noise as defined in paragraph 2.8). If \( n, m < N \), then the following equations apply:

\[
\sum_{i=0}^{2N-1} \left[ \begin{array}{c} \cos \\ \sin \end{array} \right] mj \Delta \lambda \begin{array}{c} \cos \\ \sin \end{array} q j \Delta \lambda = 0 \quad \text{if} \quad \begin{cases} \alpha \neq \beta \\ \text{or} \\ m \neq q \end{cases}
\]

(2.83)

Moreover, if the grid is symmetrical with respect to the Equator (as equal angular grids are), and if \( \sigma_i^2 = \sigma_{m-l-1}^2 \) (for instance, if \( \sigma_i^2 \) is independent also of \( i \) the relationships

\[
\sigma_i^2 \hat{P}_{n \alpha} (\cos \theta_i) = \hat{P}_{n \alpha} (\cos \theta_i) (-1)^{m-n} \sigma_{n-l-1}^2
\]

\[
\sigma_i^2 \hat{P}_{n \alpha} (\cos \theta_i) = \hat{P}_{n \alpha} (\cos \theta_i) (-1)^{n-m} \sigma_{n-l-1}^2
\]

\[
\theta_i' = \pi - \theta_i
\]

must apply, according to par. (1.2), and from these follows

\[
\sum_{i=0}^{n-1} \sigma_i^2 \hat{P}_{n \alpha} (\cos \theta_i) \hat{P}_{n \alpha} (\cos \theta_i) = 0 \quad \text{if} \quad n - p \text{ is odd.}
\]

(2.84)

In brief: if \( n, m < N \), and the grid is regularly spaced in longitude, and symmetrical with respect to the equator, then

\[
g_{\alpha n \beta m} = 0 \quad \text{if} \quad \begin{cases} \alpha \neq \beta, m \neq q \\ \text{or} \\ n - p \text{ is odd} \end{cases}
\]

(2.85)

If neither of the conditions listed above apply, then \( g_{\alpha n \beta m} \) may or may not be zero. If the coefficients \( \hat{C}_{mn}^\alpha \) are ordered in \( \alpha \) so that all those of the same \( m \) are grouped together, and for a given \( m \) all \( \hat{C}_{mn}^\alpha \) are separated from all \( \hat{C}_{mn}^\beta \), and further more all \( \hat{C}_{mn}^\alpha \) with \( n - m \) even are separated from all those with \( n - m \) odd, then the normal matrix becomes arranged in such a way that all potentially non-zero elements (i.e., not satisfying (2.85)) are also grouped together forming a series of diagonal blocks \( G_{n \delta}^\alpha \), where \( \delta \) signals...
the parity of $n-p$. Each one of these diagonal blocks is made exclusively of one of the following types of elements

$$g_{z,n-p}^{\alpha,\delta} = \begin{cases} n-p \text{ even (}\delta = 0) \\ n-p \text{ odd (}\delta = 1) \end{cases}$$

If the grid is not symmetrical with respect to the equator, groups of type $(\alpha=1, \delta=0)$ and $(\alpha=1, \delta=1)$, or $(\alpha=0, \delta=1)$, become included into larger non-zero blocks, as there are more non-zero elements in that case. However, here the discussion will cover only the symmetrical case.

The largest blocks are $G_{z}^{\alpha,\delta}$, and their dimension is $N_{z}^{2}$; the smallest blocks have dimension $1 \times 1$, for example $g_{z,1}^{\alpha,0}$.

The inverse of any block-diagonal matrix such as $G$ is another block-diagonal matrix made up of the inverse of the blocks of $G$. There are $4N-2$ diagonal blocks $g_{z}^{\alpha,\delta}$ in $G$, and as many in $G^{-1}$.

The eigenvalues of $G$ are those of the diagonal blocks, and the eigenvectors of $G$, those of the same blocks "expanded" with zeroes at both ends, so as to reach the dimension $N_{p}$ of $G$.

The estimates' vector

$$\vec{c} = G^{-1} A^{T} D^{-1} m$$

can be partitioned in the same way as $c$, each partition $\vec{c}_{z,\alpha,\delta}$ including all coefficients' estimates of the same $m$, $\alpha$, and parity $\delta$ of $n-m$:

$$\vec{c}_{z,\alpha,\delta} = (G_{z}^{\alpha,\delta})^{-1} A_{z}^{\alpha,\delta} D^{-1} m$$  \hspace{1cm} (2.86)

where $A_{z}^{\alpha,\delta}$ is a $(N-m) \times N_{p}$ matrix with rows that are $N_{p}$ vectors of the type

$$\vec{a}_{z,n} = \left[ \ldots \bar{P}_{nm} (\cos \theta_{i}) \left\{ \begin{array}{c} \cos \vspace{0.1cm} \\ \sin \end{array} \right\} m j \Delta \lambda \ldots \right]$$  \hspace{1cm} (2.87)

($n-m$ even if $\delta = 0$, odd if $\delta = 1$) So the rows of $(G_{z}^{\alpha,\delta})^{-1} A_{z}^{\alpha,\delta}$ are linear combinations of vectors of the same frequency $m$ and, therefore, also vectors of the same frequency:

$$\vec{h}_{z,n} = \left[ \ldots \chi_{i}^{m} \left\{ \begin{array}{c} \cos \vspace{0.1cm} \\ \sin \end{array} \right\} m j \Delta \lambda \ldots \right]$$  \hspace{1cm} (2.88)

Consequently, the estimate of a given $\vec{c}_{z,\alpha}$ is

$$\vec{c}_{z,\alpha} = \vec{h}_{z,\alpha} D^{-1} m = \sum_{i=0}^{\frac{N_{p}}{2}} \sum_{j=0}^{\frac{N_{p}}{2}} \chi_{i}^{m} \left\{ \begin{array}{c} \cos \vspace{0.1cm} \\ \sin \end{array} \right\} m j \Delta \lambda \sigma_{i,j}^{2} m_{ij}$$  \hspace{1cm} (2.89)

and this is a quadratures type estimator.
Setting-up and inverting the \( G_{x\delta} \) blocks is tantamount to setting up and inverting the whole matrix \( G \). Since all operations related to one of the blocks are independent from those for the others, the inversion of \( G \) is ideally suited for parallel-processing computing. On average, each \( G_{x\delta} \) requires \( O(N^3) \) operations to invert, or \( O(N^3) \) altogether. Inverting \( G \) by ordinary techniques would involve \( O((N_j^3) \alpha^3) = O(N^5) \), so there is an increase in efficiency of \( O(N^2) \).

Finally, it is quite simple to show that these properties carry over both to the case of area means, and to problems where the surface being studied is not a sphere, but a surface of revolution symmetrical about a plane perpendicular to the axis of rotation, provided that the longitude increments be constant and the grid symmetrical with respect to the "equator". In this latter case, the expansion of the signal in solid spherical harmonics is

\[
z_{ij} = \sum_{\ell=0}^{N-1} \sum_{m=0}^{m=0} \sum_{r=0}^{a_n \lambda_i} \frac{a_n \lambda_i}{r_{ij}} Y_{nm}^\ell (\theta_i, \lambda_j)
\]

and the factors \( \frac{a_n \lambda_i}{r_{ij}} \) are symmetrical about the equator, from which all the properties already mentioned for \( G \) follow.

Clearly, the structure of \( G \) possesses many properties similar or identical to those of \((C_{xz} + D)\) when the data is regularly sampled on a surface of revolution. These similitudes underline the intimate relationship between least squares adjustment and least squares collocation shown in the preceding paragraph. In fact, as least squares, regularized least squares, and collocation differ only in the diagonal matrix \((K, \text{or } C^{-1})\) being added to \( G \) in expressions (2.68), (2.70), and (2.77), all the properties mentioned here for \( G \) apply to the normal matrices in each of the three methods equally well. The important consideration, in the case of collocation, is that the data be band-limited. Otherwise, expression (2.81) indicates that \( (C + M \{w w^T\})^{-1} \) and not \( C^{-1} \) must be added to \( G \). Matrix \( C + M \{w w^T\} \) has the same Toeplitz-type structure of \((C_{xz} + D)\) discussed in paragraph (2.10). Therefore, creating and inverting the normal matrix requires: (a) creating and inverting \( C + M \{w w^T\} \), and (b) creating and inverting \( A^T D^{-1} A + (M \{w w^T\} + C)^{-1} \), which can be shown to have the same block diagonal structure discussed here. This is twice the work needed to set up and invert \((C_{xz} + D)\) using the approach of paragraph (2.11), so, in the case of infinite bandwidth, that approach is more economical in computing and, therefore, more practical.

There may be one important point in favor of using formula (2.77) or (2.81) rather than formula (2.39) for obtaining the optimal estimator matrix \( F \) at least in the band-limited case: as the density of the grid increases, matrix \((C_{xz} + D)\) becomes increasingly more ill-conditioned, because the closer distance between data points results in covariances that have much the same values in consecutive rows or columns. On the other hand, the non-zero diagonal blocks in \( G \) are likely to become more and more diagonal-dominant as \( \Delta \theta, \Delta \lambda \rightarrow 0 \). This will depend on \( D \); for instance, if the variances of the noise were of the form
\[ \sigma_{ij} = \sin \theta_i \]

then

\[ L_{ij} \frac{\Delta \theta \Delta \lambda}{4\pi} \frac{\sigma_{ij}}{\sin \theta_i} = L_{ij} \sum_{\Delta \theta, \Delta \lambda} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \frac{\alpha}{\xi_{xz}(\theta_i, \lambda_j) \xi_{yz}(\theta_j, \lambda_j)} \sin \theta_i \frac{\Delta \theta \Delta \lambda}{4\pi} \]

because of the orthogonality relationships. In general, the variances of the noise are not going to follow a sinusoidal law, but one may reasonably expect (at least with more or less homogeneous noise) that the stability of the normal equations will not deteriorate with \( \Delta \theta, \Delta \lambda \to 0 \).

2.16 Global Adjustment and Collocation with Scattered Data

The efficient set up and inversion of the covariance matrix \((C_{zz} + D)\), or of the normal matrix \(G\), depend on the regular nature of the grid. If not all nodes or blocks in the grid have data associated with them, the data is said to be scattered. The blanks or "holes" in the grid destroy the orderly structure of the matrices, making the application of the techniques previously discussed impossible. Yet so strong is this structure that, even in fragments, still it can be dealt with more efficiently than in the case of ordinary matrices of the same size.

(a) Full Region Bound by Lines of Latitude and Longitude

In the case when there is data at every point or block inside a "square" region limited by parallels and meridians, the partitioning of the data vector along the arcs or parallels inside the zone reveals a strong structure in the \((C_{zz} + D)\) matrix, if all the other assumptions made in paragraph (2.10) still apply.

If \(N_r\) is the number of rows and \(N_s\) the number of meridians that cross the region, then the covariance matrix will consist of \(N_r^2\) blocks \(C^{rs}\) of dimension \(N_s\), both persymmetrical and Toeplitz, though not circulant (i.e., the relationship \(c_{rs}^{pq} = c_{rs}^{q+1}\) is fulfilled, but not \(c_{rs}^{pq} = c_{rs}^{pq+1}\)); moreover the first row in each block does not have the property that \(c_{0q}^{pq} = c_{0q}^{pq+1}\). Clearly, though weaker than in the case of a global grid, there is a definite structure here that can be exploited to make both setting up and inverting the matrix more efficient.

Because each block \(C^{rs}\) is Toeplitz, only the \(N_s\) elements in its first row have to be computed, or about \(\frac{2}{3}N_r N_s^2\) for the matrix as a whole, instead of \(\frac{5}{3}N_r^2N_s^2\); this amounts to a reduction in operations by a factor of \(N_r\).
The solution of the equation \( \dot{\mathbf{f}}_{\mathbf{z}} = (\mathbf{C}_{\mathbf{zz}} + \mathbf{D})^{-1} \mathbf{C}_{\mathbf{z} \mathbf{z}} \mathbf{s} \), for the optimal estimator vector \( \mathbf{f}_{\mathbf{z}} \), can be obtained by a technique such as conjugate gradients, or similar, in which a finite number \( K \) of matrix-vector multiplications \((\mathbf{C}_{\mathbf{zz}} + \mathbf{D}) \mathbf{v}_k \) (where \( \mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_k \), \( \ldots \) \( \mathbf{v}_k \) are \( K \) intermediate \( \mathbf{N}_{p} \) - vectors created during the solution) constitute the bulk of the computing effort. There is no need to go into the details of any specific technique, as the reader will find excellent descriptions in the literature (Householder, 1964, Luenberger, 1969). A discussion of the matrix-vector operation is sufficient here.

Let \( \mathbf{m}^i \) be the \( \mathbf{N}_{r} \) - vector partition of the \( \mathbf{N}_{r} \mathbf{N}_{r} \) data vector \( \mathbf{m} \), containing the measurements along the \( i \)th parallel in the region, and let \( \mathbf{v}_k^i \) be the corresponding partition in any of the \( \mathbf{v}_k \) vectors. The product \((\mathbf{C}_{\mathbf{zz}} + \mathbf{D}) \mathbf{v}_k = \mathbf{p}_k \) is, under such partition,

\[
\mathbf{p}_k = [\mathbf{P}_0^i \ldots \mathbf{P}_{\mathbf{N}_{r}-1}^i]^T
\]

with

\[
\mathbf{P}_k^i = \sum_{p=0}^{\mathbf{N}_{r}-1} \mathbf{C}_{\mathbf{p}}^i \mathbf{v}_k^i
\]

so the whole matrix-vector multiplication can be broken up into \( \mathbf{N}_{r} \mathbf{N}_{r} \) products \( \mathbf{C}_{\mathbf{p}}^i \mathbf{v}_k^i = \mathbf{h}_{k}^i_0 \). Because \( \mathbf{C}_{\mathbf{p}}^i \) is a Toeplitz matrix, the \( \mathbf{N}_{r} \) components of \( \mathbf{h}_{k}^i_0 \) can be obtained by "weighted running averages" or discrete convolution of the elements of \( \mathbf{v}_k^i \) with those in the first row of \( \mathbf{C}_{\mathbf{p}}^i \). Such convolution can be calculated efficiently using the Fast Fourier Transform algorithm (see, for instance, Brigham, 1974, Ch. 13). Therefore, all \( \mathbf{N}_{r} \mathbf{N}_{r} \) products involve \( \mathcal{O}(\mathbf{N}_{r} \mathbf{N}_{r}^{2}) \) operations, and since there are \( K \) matrix-vector multiplications in the whole procedure, the total number of operations needed to obtain \( \mathbf{f}_{\mathbf{z}} \) amounts to \( \mathcal{O}(K \mathbf{N}_{r} \mathbf{N}_{r}^{2}) \). For conjugate gradients, \( K \) does not exceed (in theory) \( \mathbf{N}_{r} = \mathbf{N}, \mathbf{N}_{r}^{2} \), so there should be \( \mathcal{O}(\mathbf{N}_{r}^{2} \mathbf{N}_{r}^{3}) \) operations altogether. If \((\mathbf{C}_{\mathbf{zz}} + \mathbf{D}) \) were handled by conventional techniques, disregarding its well defined structure, the number would be \( \mathcal{O}(\mathbf{N}_{r}^{3} \mathbf{N}_{r}^{2}) \), so the increase in efficiency is \( \mathcal{O}(\mathbf{N}_{r}) \), the same as for the setting up.

(b) Arbitrarily Scattered Data

It is common in geodesy and in geophysics to have a set of measurements scattered throughout the globe, without the data being on the nodes of a regular grid or without all \( \sigma_{i}^{2} \), being equal along parallels (nonhomogenous noise). If the set is dense enough, however, it is possible to interpolate the data quite reliably on the closest nodes of a conveniently chosen grid. Assuming that this is done, and that the accuracies of the interpolated values are known well

-57-
enough, then the problem can be dealt with by conventional least squares or by collocation. In general, there will be blanks or "holes" irregularly distributed over the sphere, the actual data points falling among them in no precise pattern. This problem will be considered here as a least squares adjustment problem. If the data has little or no power above the Nyquist frequency of the grid on which it has been interpolated, then the extension of the ideas that follow to collocation is quite simple, according to paragraph (2.13).

Consider the element \( g_{a_{x}b_{y}}^{\alpha\beta} \) of the matrix \( G = A^T D^\alpha A \)

\[
g_{a_{x}b_{y}}^{\alpha\beta} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \sum_{p=0}^{m-1} \sum_{q=0}^{m-1} \left\{ \begin{array}{c} \cos \\ \sin \end{array} \right\} m_j \Delta \lambda \left\{ \begin{array}{c} \cos \\ \sin \end{array} \right\} p_j \Delta \lambda W_{ij} \sigma_{ij}^2
\]

(2.91)

where \( W_{ij} = 0 \) if the point \( ij \) is blank, otherwise \( W_{ij} = 1 \); \( \sigma_{ij}^2 \) is the variance of the noise \( n_{ij} \). In general \( \sigma_{ij}^2 \) is a function both of \( i \) and of \( j \). Clearly, matrix \( D \) is taken to be diagonal (i.e., uncorrelated noise).

From the relationships

\[
\begin{align*}
\cos mj \Delta \lambda \cos pj \Delta \lambda &= \frac{1}{2} \left[ \cos (m+p)j \Delta \lambda \cos (m-p)j \Delta \lambda \right] \\
\cos mj \Delta \lambda \sin pj \Delta \lambda &= \frac{1}{2} \left[ \sin (m+p)j \Delta \lambda \cos (m-p)j \Delta \lambda \right] \\
\sin mj \Delta \lambda \sin pj \Delta \lambda &= -\frac{1}{2} \left[ \cos (m+p)j \Delta \lambda - \cos (m-p)j \Delta \lambda \right]
\end{align*}
\]

and calling

\[
C_r^\alpha \equiv \begin{cases} 
C_r^{\alpha} = \frac{1}{2} \sum_{i=0}^{N-1} \cos r j \Delta \lambda \sigma_{ij}^2 W_{ij} \\
C_r^{\beta} = \frac{1}{2} \sum_{i=0}^{N-1} \sin r j \Delta \lambda \sigma_{ij}^2 W_{ij}
\end{cases}
\]

(2.92)

where \(-N < r < 2N\), follows

\[
\sum_{i=0}^{n-1} \sigma_{ij}^2 W_{ij} \left[ \begin{array}{c} \cos mj \Delta \lambda \\ \sin mj \Delta \lambda \end{array} \right] \left[ \begin{array}{c} \cos pj \Delta \lambda \\ \sin pj \Delta \lambda \end{array} \right] = [C_r^{\alpha\beta}] + (-1)^{a+\beta} C_r^{\alpha \beta},
\]

(2.93)

Moreover, because

\[
\begin{align*}
\cos (-r) j \Delta \lambda &= \cos r j \Delta \lambda \\
\sin (-r) j \Delta \lambda &= -\sin r j \Delta \lambda \\
\cos r j \Delta \lambda &= \cos (2N - r) j \Delta \lambda \\
\sin r j \Delta \lambda &= -\sin (2N - r) j \Delta \lambda
\end{align*}
\]

(\(2N \Delta \lambda = 2\pi\))

follows

\[
\begin{align*}
C_{-r}^{\alpha} &= C_r^{\alpha} (-1)^{\alpha} \\
C_{-r}^{\alpha} &= C_r^{\alpha} (-1)^{\alpha}
\end{align*}
\]

-58-
where the last one is of special interest when \( r \geq N \). So only the \( C_r^l \alpha \) with \( 0 < r < N \) are needed. Finally, calling

\[
g_{n_1, n_2}^{i \alpha, \beta} = P_n^m (\cos \theta_1) \overline{P}_{n_2} (\cos \theta_1) [C_{n_1}^{i} |\alpha - \beta|] + (-1)^{2\alpha + \beta} C_{n_2}^{i} |\alpha - \beta| \ (2.94)
\]

(where \(|\alpha - \beta|\) is the absolute value of \(\alpha - \beta\))

it is

\[
g_{n_1, n_2}^{i \alpha, \beta} = \sum_{l=0}^{n-1} g_{n_1, n_2}^{i \alpha, \beta} \ (2.95)
\]

Assume that out of \(N\) rows the grid has only \(I\) with any data in them, so at least one \(W_{1l} \neq 0\) in each. Once the corresponding \(C_r^l \alpha \) have been obtained by computing the discrete Fourier transform of \(\sigma_{r}^{2} W_{1l} \) along each row with data \(O(N^2)\) operations for the whole grid, what remains is to get the \(I\) non-zero terms \(g_{n_1, n_2}^{i \alpha, \beta}\) that form each element \(g_{n_1, n_2}^{i \alpha, \beta}\) of \(G\). As there are about \(\frac{2}{2}N^4\) such elements that are different, and \(I\) is \(O(N)\) (except for very sparse data sets), the total number of operations needed to create \(G\) is \(O(N^2) + O(N^5)\), or virtually \(O(N^5)\). If the \(g_{n_1, n_2}^{i \alpha, \beta}\) were computed according to (2.91) as it is written, instead of according to its reduced version. (2.95), the number of operations would be \(O(2N^2 (N^4)\), or \(O(N^6)\), so the gain in efficiency allowed by this approach is \(O(N)\), which is the same as in the case studied in the first part of this paragraph, where the data completely filled a "square" sector of the sphere. This count does not include the time needed to obtain the \(P_n^m (\cos \theta_1)\), as these can be pre-computed once and kept on disk or tape for repeated use.

The number of operations can be reduced further, almost by half, by taking advantage of the fact that \(P_n^m (\cos \theta_1)\) is a common factor in all \(g_{n_1, n_2}^{i \alpha, \beta}\) with the same \(n\) and the same \(m\). Furthermore, if there are pairs of rows \(i\) and \(N-i-1\) (i.e., symmetrical with respect to the Equator) where both rows contain some data, then

\[
g_{n_1, n_2}^{i \alpha, \beta} + g_{n_1, n_2}^{i-1 \alpha, \beta} = P_n^m (\cos \theta_1) \overline{P}_{n_2} (\cos \theta_1)
\]

\[
[(C_{n_1}^{i}|\alpha - \beta|) + (-1)^{2\alpha + \beta} C_{n_2}^{i} |\alpha - \beta|] + (-1)^{2\alpha + \beta} C_{n_2}^{i} |\alpha - \beta|
\]

(2.96)

which leads to further savings in computing at the cost of additional programming complexity. These economies are important, but they will not bring the number of operations much below \(O(N^5)\) unless the grid is so sparse that \(I\) is much smaller than \(N\).

Notice that the normal matrix \(G = A^T D^{-1} A\) is created here without actually forming the observation equations matrix \(A\). This means considerable savings in computing and in storage requirements. As for the right hand side of the normals \(A^T D^{-1}m = b\), the elements of \(b\) are given by the formula

-59-
\[ b_k = \sum_{j=0}^{N-1} \sum_{l=0}^{N-1} \mathbf{p}_{n_{ij}} (\cos \theta_{ij}) \begin{bmatrix} \cos \lambda_m \sigma_{ij}^m \end{bmatrix} \begin{bmatrix} \sin \lambda_m \sigma_{ij}^m \end{bmatrix} W_{ij} m_{ij} \] (2.97)

where \( k = n^2 + \alpha n + m + 1 \) and \( W_{ij} = 0 \) if there is no data at the point \( ij \), as before. Such expression is of the quadratures type, and can be computed efficiently by the corresponding algorithm of section 1, also without first creating \( A \). Finally, the residuals vector \( \mathbf{v} = \mathbf{m} - A \hat{\mathbf{c}} \), usually of interest, \( \alpha \) can be obtained as the difference between the data \( \mathbf{m} \) and the values of

\[ \hat{z} (\theta_1, \lambda_1) = \sum_{\alpha = 0}^{\alpha_n} \sum_{\beta = 0}^{\beta_n} \frac{A \alpha \beta}{m_{ij}} \mathbf{p}_{n_{ij}} (\cos \theta_{ij}) \begin{bmatrix} \cos \lambda_m \sigma_{ij}^m \end{bmatrix} \begin{bmatrix} \sin \lambda_m \sigma_{ij}^m \end{bmatrix} \]

computed by means of the appropriate synthesis procedure given in section 1. Thus \( \mathbf{v} \) can be found without knowing \( A \) explicitly.

The normal equations can be solved by means of conjugate gradients or a similar method involving \( M \) matrix-vector products \( \mathbf{x}_t = \mathbf{G} \mathbf{h}_t \), where \( \mathbf{h}_t \) is part of a sequence of intermediate vectors \( \mathbf{h}_0, \mathbf{h}_1, \ldots, \mathbf{h}_t, \ldots \)

Introducing the notation

\[ \mathbf{h}_{n_{ij}} = \mathbf{h}_t \]

where \( k = n^2 + \alpha n + m + 1 \), and calling \( \mathbf{G}^t \) to the matrix of all \( g_{n_{pq}}^t \alpha, \beta \), so

\[ \mathbf{G} = \sum_{t=0}^{N-1} \mathbf{G}^t \]

\[ \mathbf{x}_t = \sum_{t=0}^{N-1} \mathbf{G}^t \mathbf{h}_t = \sum_{t=0}^{N-1} \mathbf{x}_t \]

then each element \( x_{n_{ij}}^t \alpha, \beta \) of the product vector \( \mathbf{x}_t \) is of the form

\[ x_{n_{ij}}^t \alpha, \beta = \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} \sum_{\beta = 0}^{\beta_n} \frac{g_{n_{pq}}^t \alpha, \beta}{m_{ij}} \mathbf{p}_{n_{pq}} (\cos \theta_{ij}) \sum_{\rho = 0}^{\rho_n} \begin{bmatrix} \cos \lambda_m \sigma_{ij}^m \end{bmatrix} \begin{bmatrix} \sin \lambda_m \sigma_{ij}^m \end{bmatrix} \]

\[ \cdot \mathbf{p}_{pq} (\cos \theta_{ij}) \mathbf{h}_{pq}^\beta \] (2.99)

Every \( \mathbf{h}_{pq}^\beta \) above multiplies always the same \( \mathbf{p}_{pq} (\cos \theta_{ij}) \); calling

\[ d_{pq}^\beta = \mathbf{p}_{pq} (\cos \theta_{ij}) \mathbf{h}_{pq}^\beta \]

expression (2.99) becomes

\[ x_{n_{ij}}^t \alpha, \beta = \mathbf{p}_{n_{ij}} (\cos \theta_{ij}) \sum_{p=0}^{\rho_n} \sum_{q=0}^{N-1} \sum_{\beta = 0}^{\beta_n} \begin{bmatrix} \cos \lambda_m \sigma_{ij}^m \end{bmatrix} \begin{bmatrix} \sin \lambda_m \sigma_{ij}^m \end{bmatrix} d_{pq}^\beta \]

All quantities inside the square brackets are constant if \( \alpha, \beta, m \), and \( q \) are the same. Grouping equal factors together

\[ x_{n_{ij}}^t \alpha, \beta = \mathbf{p}_{n_{ij}} (\cos \theta_{ij}) \sum_{q=0}^{N-1} \sum_{\beta = 0}^{\beta_n} \begin{bmatrix} \cos \lambda_m \sigma_{ij}^m \end{bmatrix} \begin{bmatrix} \sin \lambda_m \sigma_{ij}^m \end{bmatrix} \sum_{p=0}^{\rho_n} d_{pq}^\beta \]

\[ = \mathbf{p}_{n_{ij}} (\cos \theta_{ij}) \sum_{q=0}^{N-1} \sum_{\beta = 0}^{\beta_n} \frac{g_{n_{pq}}^t \alpha, \beta}{m_{ij}} \mathbf{p}_{n_{pq}} (\cos \theta_{ij}) \sum_{\rho = 0}^{\rho_n} \begin{bmatrix} \cos \lambda_m \sigma_{ij}^m \end{bmatrix} \begin{bmatrix} \sin \lambda_m \sigma_{ij}^m \end{bmatrix} d_{pq}^\beta \] (2.100)

where \( D_{pq} d_{pq}^\beta = \sum_{p=0}^{\rho_n} d_{pq}^\beta \)
There are \((N^3)\) products needed to form all \(d_{ij}^{\beta}\); \((N^3)\) sums to compute the \(2N, D_{i\beta}^j\), and \((N^3)\) further multiplications by the \(P_n (\cos \theta)\) to form all \(\times \alpha_{i\beta}^{\alpha}\). As there are 1 non-zero \(G^1\), there are, per matrix-vector product, \(1xO(N^3)\) operations, or \(O(N^3)\) if the data is not too sparse. \(M\), the number of matrix-vector products in the solution, is \(O(N^2)\), so the total comes to \(O(N^5)\). Inverting \(G\) by the usual methods and without having regard for its structure involves \(O(N^5)\) operations. The gain in efficiency is, again, \(O(N)\).

Besides providing a convenient way of demonstrating how the properties of \(G\) can be exploited to make its inversion more efficient (or the solution of the normal equations, both approaches are equivalent), conjugate gradients is interesting on its own right. Sparse data sets with a poor distribution will result in ill-conditioned normals, so the inversion of \(G\) may be numerically impossible. So-called iterative methods, such as conjugate gradients, usually improve the initial guess (represented by \(h_0\)) of the correct values of the unknowns, at least for the first few iterations. Improvement here means a reduction in the quadratic form being minimized, such as the mean square value of the residuals. If the initial guess is a good one, and present day spherical harmonic coefficients of the gravity field are reasonable good for degrees up to 30 or so, then a few iterations are likely to improve this guess, and to produce reasonable estimates of those coefficients that, being wholly unknown, are taken to be zero at the start. If a "few" iterations are much less than the maximum \(N^3\), then a reduction in computing time of \(O(N^2)\) takes place. This might allow scientists to "extend" existing models to much higher degree and order than at present, simply by obtaining approximate solutions of this type.

Clearly, all that has been said here regarding \(G\) and least squares adjustment applies equally well to \((G + C^{-1})\) and least squares collocation. Though the formulas have been developed on the basis of a point values’ formulation, their extension to area means is not difficult.
While the descriptions of the methods for estimating the $\tilde{C}_{n \alpha}^\alpha$ discussed here have been confined to the case where all the data are of one kind (i.e., gravity anomalies only, or magnetic anomalies only, or geopotential numbers, etc.), their extension to mixed data sets is immediate, provided that all values are globally distributed according to grids of the same $\Delta \lambda$, although the latitudes need not be the same as well.

Note on the Accuracies of the $\tilde{C}_{n \alpha}^\alpha$ when Using Conjugate Gradients:

Besides the estimates of the coefficients, one usually wants to know the accuracies of those estimates. When a few iterations of conjugate gradients are used to update some initial estimates in the efficient way described above, what are the accuracies of the improved coefficients? In the course of the conjugate gradients procedure (see references), the conjugate directions $v_k$ of $G = (A^T D^{-1} A + C^{-1} + Q_s^{-1})$ are generated ($Q_s^{-1}$ is the variance covariance matrix of the initial estimates, as explained in paragraph (2.18) part (b)), together with the scalars $v_k^T G v_k = \alpha_k$. The conjugate directions have the property

$\quad v_k^T G v_p = 0$

for $k \neq p$. The estimator implied by $R$ iterations of this procedure is

$\quad \tilde{\alpha} = \tilde{G}(A^T D^{-1} m + Q_s^{-1} \tilde{c}_s) = \tilde{G} A^T D^{-1} (z + n) + \tilde{G} Q_s^{-1} (c + \Delta c_s)$

where $\tilde{c}_s$ is the vector of initial estimates, $\Delta c_s$ the errors in this vector, and

$\quad \tilde{G} = \sum_{k=1}^{R} \alpha_k^{-1} v_k v_k^T$

The variance-covariance matrix of the updated errors, for ordinary least squares, is (assuming that $E\{n \Delta c_s\} = 0$)

$\quad E\{(\tilde{G} A^T D^{-1} n) (\tilde{G} A^T D^{-1} n)^T\} + E\{(\tilde{G} Q_s^{-1} \Delta c_s)(\tilde{G} Q_s^{-1} \Delta c_s)^T\} = \tilde{G} A^T D^{-1} D D^{-1} A \tilde{G}$

$\quad + \tilde{G} Q_s^{-1} Q_s \tilde{G} = \tilde{G} (A^T D^{-1} A + Q_s^{-1}) \tilde{G} = \tilde{G} G \tilde{G}$

$\quad = \sum_{k=1}^{R} \alpha_k^{-1} v_k \alpha_k v_k^T = \sum_{k=1}^{R} \alpha_k^{-1} v_k v_k^T = \tilde{G}$

Therefore, the variances of the errors in the $\tilde{C}_{n \alpha}^\alpha$ are equal to the corresponding diagonal elements of $\tilde{G}$:

$\quad \sigma_{\eta_{n \alpha}^\alpha}^2 = \sum_{k=1}^{R} \alpha_k^{-1} v_k \eta_{n \alpha}^\alpha$

where $v_{n \alpha}^\alpha$ is the element of $v$ corresponding to $\tilde{C}_{n \alpha}^\alpha$ in $\tilde{c}$ ($v$ and $\tilde{c}$ have the same dimension). The same result applies to least squares collocation. All that is needed to obtain the $\sigma_{\eta_{n \alpha}^\alpha}^2$, according to the formula
above, is knowledge of the $\alpha$'s and of the $\nu$'s, which can be saved as they are created, during the $k$th iteration of the procedure. For the efficient algorithm to be applicable, $Q_{k}^{-1}$ must have a suitable structure. One case is when $Q_{k}^{-1}$ is diagonal, or can be satisfactorily approximated by a diagonal matrix.

2.17 The Error Matrix in the Band-Limited Case

From (2.41) and (2.77) results

$$E_{T} = C - C_{\alpha \alpha} (C_{\nu \nu} + D)^{-1} C_{\nu \alpha} = C - F C_{\nu \nu} = C - (A^{T} D^{-1} A + C^{-1})^{-1} A^{T} D^{-1} A C$$

$$= [I - (A^{T} D^{-1} A + C^{-1})^{-1} A^{T} D^{-1} A] C = (A^{T} D^{-1} A + C^{-1})^{-1} (A^{T} D^{-1} A + C^{-1}) - A^{T} D^{-1} A] C$$

$$= (A^{T} D^{-1} A + C^{-1})^{-1}$$

(2.101)

In the case of the best unbiased estimator $C^{-1}$ is not present in the normal matrix, so (2.101) becomes

$$E_{T} = (A^{T} D^{-1} A)^{-1} = G^{-1}$$

(2.102)

which is the well known expression of the error matrix for ordinary least squares.

Because of the block-diagonal structure of the variance-covariance matrix$^{(1)}$ the estimates of $C_{\alpha \alpha}$ of different orders are uncorrelated. The diagonal elements of the variance-covariance matrix are the variances of the estimated coefficients total errors (i.e., sampling plus propagated noise), in (2.101).

Obtaining the variances of the $C_{\alpha \alpha}$ in band-limited collocation is formally identical to getting the variances of the estimates in ordinary least squares, according to (2.101) and to (2.102)

2.18 The Use of a priori Information on the Coefficients

Assume that all coefficients up to some degree and order $M$ are approximately known, and that $Q_{*}$ is the variance-covariance matrix of their errors. This could be the case where a model of the gravity field has been obtained from data gathered using artificial satellites, complete to degree $M$, and terrestrial data is to be used to improve the existing coefficients and obtain new ones beyond degree $M$.

Three possible approaches to this question will be discussed here, using a "point values" formulation in the first two cases for simplicity.

---

$^{(1)}$ When the data set is not sparse.
(a) Simple weighted averages:

The terrestrial data can be used separately to obtain a model to degree and order \( M_T > M \), together with the variance-covariance matrix \( Q_T \) of the coefficients' errors. To combine the satellite and the terrestrial coefficients one can set up the following observation equations

\[
\begin{bmatrix}
\mathbf{c}_s \\
\mathbf{c}_T
\end{bmatrix} + \mathbf{s} = \begin{bmatrix}
I_s \\
I_T
\end{bmatrix}
\]

(2.103)

where \( \mathbf{c} \) is the vector potential coefficients and \( I_s \) is the \((M + 1)^2 \times (M + 1)^2\) unit matrix augmented with zeroes on the right, and \( I_T \) is the \((M_T + 1)^2 \times (M_T + 1)^2\) unit matrix, while \( \mathbf{c}_s \) is the vector of satellite coefficients and \( \mathbf{c}_T \) the vector of terrestrial coefficients; \( \mathbf{s} \) and \( \mathbf{t} \) are the corresponding vectors of residuals. The best linear unbiased estimator for the combined system of observations is

\[
\hat{\mathbf{c}} = (I_s^T I_T) \begin{bmatrix}
(Q_s^{-1} O I_s) \\
O Q_T^{-1} I_T
\end{bmatrix}^{-1} \begin{bmatrix}
I_s^T I_T \\
Q_s^{-1} O I_s
\end{bmatrix} \mathbf{c}_s
\]

(2.104)

\[
= (Q_s^{-1} + Q_T^{-1})^{-1} (Q_s^{-1} \mathbf{c}_s + Q_T^{-1} \mathbf{c}_T)
\]

If \( Q_s \) and \( Q_T \) are the inverses of ordinary least squares normal matrices \( G_s \) and \( G_T \), then the error matrix of the combined solution

\[
E = (Q_s^{-2} + Q_T^{-2})^{-1} = (G_s + G_T)^{-1}
\]

corresponds to the propagated noise only. If they are "collocation" matrices of the type \((A^T D^{-1} A + \mathbf{C}^{-1})^{-1}\) (see expression (2.77)), then the error matrix includes the effect of the sampling error as well. Most satellite models are, to date, "least squares-type" and it would be incongruous to combine them with "collocation-type" models, terrestrial or otherwise. The problem need not be a serious one, because geodetic spacecrafts so far have orbited at altitudes of 300 km or more, where the field is much smoother than at the surface, so the sampling errors are bound to be small compared to the propagated data errors reflected by \( Q_s \).

In the case where the terrestrial model has been derived from a regularly sampled data set using the "band-limited approach" of previous paragraphs, \( Q_T \) is block-diagonal, and those blocks corresponding to orders \( m > M \) are identical to the corresponding blocks in \((Q_s^{-1} + Q_T^{-1})^{-1}\). From this it is not difficult to conclude that the coefficients in the combined model up to order \( M \) will be somewhat different (and presumably better) than those in either the satellite or the terrestrial sets, while those above \( M \) will be identical to the corresponding terrestrial coefficients.
(b) A priori values included as data in the adjustment:

Consider the system of observation equations

\[
\begin{bmatrix}
M \\\nC_s
\end{bmatrix} + \begin{bmatrix}
V \\
S
\end{bmatrix} = \begin{bmatrix}
\hat{A} \\
I_s
\end{bmatrix} \mathbf{c}
\]

(2.105)

which is the system (2.66) augmented with equations of the type

\[
\hat{C}_{\alpha n}(s) - e_{\alpha n}(s) = C_{\alpha n}(s)
\]

where \( e_{\alpha n}(s) \) is the error in the "satellite model" coefficient \( \hat{C}_{\alpha n}(s) \).

The normal matrix of collocation is

\[
E_T^{-1} = \left( [A^T I_s] [\begin{array}{cc}
D^{-1} & O \\
O & Q_s^{-1}
\end{array}] [A] + C^{-1} \right)
\]

(2.106)

and the optimal estimator of the band-limited type is

\[
\hat{\mathbf{c}} = (A^T D^{-1} A + Q_s^{-1} + C^{-1})^{-1} (A^T D^{-1} m + Q_s^{-1} c_s)
\]

(2.107)

Once more, if the data in \( m \) has been sampled regularly on the sphere, the estimated coefficients will be affected by the existence of a priori values only if their order is no higher than \( M \). Naturally, this is a desirable situation. Also it is important that the error measures corresponding to \( E_T \) and \( Q_s \) be congruous, though this is probably not very important in the case of satellite models obtained with high-orbiting spacecraft. Notice that (a) and (b) are equivalent when \( Q_T = (A^T D^{-1} A)^{-1} \) and when \( C^{-1} \) is excluded from (2.106)-(2.107). In other words: these first two approaches are equivalent for ordinary least squares.

(c) The method of Kaula and Rapp:

W. Kaula (1966) proposed a technique for simultaneously filtering errors out of a terrestrial data set and improving the coefficients of a satellite model. This method was later developed by R. Rapp (1968), who more recently (1978) used it to improve a global data set of mean 1° x 1° anomalies by combining it with the potential coefficients of the GEM-9 model. This adjusted data set was used by the author of this report to create the 5° x 5° mean anomalies analysed in one of the numerical experiments of section 3.

The idea is to satisfy condition equations of the type

\[
\hat{C}_{\alpha n}(s) + d_{\alpha n} = (4 \pi \gamma (n - 1) \beta_n S^n)^{-1} \sum_{i=0}^{N-1} \sum_{j=0}^{2n-1} \int \int \gamma n(\phi, \lambda) d \sigma (\Delta g_{ij} + v_{ij}) = 0
\]

(2.108)
while minimizing the quadratic form

$$g(d, v) = d^T Q_s^{-1} d + v^T D^{-1} v$$  \hspace{1cm} (2.109)$$

where \( v \) is the vector of corrections \( v_{ij} \) to the mean gravity anomalies \( \Delta g_{ij} \), and \( d \) is the vector of corrections \( d_{\alpha n}^\alpha \) to the satellite potential coefficients \( \bar{C}_{\alpha n}^\alpha(n) \); \( \beta_n \) is the \( n \)th degree Pellenin factor discussed in paragraph (4.3), \( \gamma \) is the mean value of equatorial gravity, and \( S \) is the ratio between the radius of the Earth's largest inner geocentric sphere and the mean Earth radius. In matrix form, the condition equations (2.108) are

$$c^2 + d - A^T \Delta g - A^T v = 0$$  \hspace{1cm} (2.110)$$

where \( A \) is a \( 2N^2 \times (M + 1)^2 \) matrix having columns \( a_{\alpha n}^\alpha \) of the form

$$a_{\alpha n}^\alpha = (4\pi(n - 1) \gamma S^{m\alpha} \beta_n)^{-1} \left[ \int_{\sigma_{00}} \overline{Y}_{n, s}^\alpha(\theta, \lambda) d\sigma \cdots \int_{\sigma_{NN}} \overline{Y}_{n, s}^\alpha(\theta, \lambda) d\sigma \right]^T$$

so, except for the factor \((4\pi(n - 1) \gamma S^{m\alpha} \beta_n)^{-1}\), \( A \) is the "area means version" of the matrix of system (2.66), and has the same properties as the "A" matrices considered so far.

The optimal estimates are given by the expressions

$$\hat{A} = c_s + d$$  \hspace{1cm} (2.111-a)$$

$$\hat{\Delta g} = \Delta g + v$$  \hspace{1cm} (2.111-b)$$

where

$$d = -(A^T D A)^{-1} + Q_s^{-1} (A^T D A)^{-1} (c_s - A^T \Delta g)$$  \hspace{1cm} (2.112)$$

and

$$v = D A Q_s d$$  \hspace{1cm} (2.113)$$

If the data set is both complete and of uniform quality, the matrix \( A^T D A \) has the block structure first discussed in paragraph (2.15). The presence of \( D \) instead of \( D^{-1} \) makes no difference to the calculations needed to set up and invert the matrix; the procedures are those already explained. Terrestrial data, however, is usually both scattered and of varying quality (i.e., different noise variances). For this type of data, therefore, the methods for scattered measurements given in paragraph (2.16) could be used.

2.19 Optimal Estimation over a Band of Spatial Frequencies

Assume that the signal is of the type

$$m = A c + n$$
where $A$ and $c$ may now be infinite (i.e., all degrees from 0 to $\infty$ may be present). If $c'$ is a sub-vector of $c$ comprising, say, the first $N^2$ coefficients, and if $\hat{s}$ is a vector of estimates of a function $s$ at a given set of points on the sphere, such that the values of $s$ depend only on those of $c'$ according to the relationship

$$s = Bc'$$  \hspace{1cm} (2.114)

where $B$ is some matrix of appropriate size, not necessarily of the same type as $A$, then the optimal estimator for $\hat{s}$ is, according to $(2.39), (2.75),

$$\hat{s} = C_{xx}(C_{xx} + D)^{-1} m = M\left\{Bc'c'^TA^\top(C_{xx} + D)^{-1}\right\} = Bc'c'^TA^\top(C_{xx} + D)^{-1} m$$

or

$$\hat{s} = Bc'$$  \hspace{1cm} (2.115)

according to $(2.40)$.

Expression $(2.115)$ indicates that the optimal estimates of a band-limited function $s$ from data $m$ are identical to the values of $s$ obtained from the optimal estimates of the coefficients $c'$ by means of the relationship $\hat{s} = Bc'$.
3. Numerical Examples

This section presents several computed examples to illustrate some of the ideas and methods discussed earlier on. The question posed here is, basically, that of the accuracy of the various procedures, and is answered by means of error analysis carried out with the formulas for error variance developed in section 2, and also by analyzing simulated data and comparing the recovered coefficients to the original ones in order to find the actual errors. A comparison of the rms of these errors with the theoretical rms (i.e., the square root of the variance) provides both a check on each set of results and, more important, shows just how adequate an error measure the theoretical rms can be. Besides error analysis and simulations, this section shows, in the last paragraph, the results of the harmonic analysis of a real data set: a $5^\circ \times 5^\circ$ equal angular set of mean gravity anomalies covering the whole Earth, from which the coefficients of the disturbing potential have been recovered to degree and order 36 by means of least squares collocation, using the "Toeplitz matrix" approach of paragraphs (2.10) and (2.11).

3.1 Generation and Analysis of Simulated Data

As explained in the preceding section, the variance of the error in the estimate of $\bar{C}_{mm}$ depends on the power spectrum (or covariance function) of the signal, and on the variance-covariance matrix of the noise. The propagation of the noise is quite straightforward, and anybody who has had any practical experience with adjustments of geodetic networks and the like already has enough "feeling" for this part of the error measure, and is capable of understanding its significance when its value is given to him. The part corresponding to the sampling error is somewhat different; it involves a rather unusual geometric average over rotations, and this type of error measure, while not exactly new (collocation, based on this measure, has been around since the mid-sixties) is not so familiar to geodesists yet, and its use in harmonic analysis in particular, far from common practice. For this reason, it is probably fair to the readers to provide some illustration of how "close" this part of the error measure is to the actual sampling error that occurs when data of the assumed power spectrum is analyzed in any of the ways discussed so far to recover spherical harmonic coefficients. By "close" one means that the actual numbers measuring the theoretical and the actual variances (or rms) should differ from each other by a small percentage, or some equally clear-cut criterion.

The theoretical variance considered here is the variance of the estimation errors per degree $\sigma_{e^2}$, defined in terms of the error measure of section 2 as follows (see paragraph (2.8))
\[
\sigma_{\epsilon_n}^2 = \sum_{n=0}^{\infty} \sum_{\alpha=0}^{\infty} \sigma_{n\alpha}^2 = \sum_{n=0}^{\infty} \sum_{\alpha=0}^{\infty} \left( \frac{\sigma_n^2}{2n+1} - 2 \epsilon_{n\alpha} \frac{\alpha}{\epsilon_{n\alpha}} + \left( \frac{\alpha}{\epsilon_{n\alpha}} \right)^{\frac{3}{2}} \right)
\]

(3.1)

The rms of the error is the square root of this variance, and the ratio of this rms to the rms per coefficient:

\[
\frac{\sigma_{\epsilon_n}}{\sigma_n} = \left( 1 + \sigma_{n}^2 \right)^{\frac{1}{2}} \sum_{n=0}^{\infty} \sum_{\alpha=0}^{\infty} \left( \frac{\epsilon_{n\alpha}^\alpha}{\sigma_{n\alpha}^2} - \left( \epsilon_{n\alpha}^\alpha \right)^{\frac{3}{2}} \right)^{\frac{1}{2}}
\]

(3.2)

multiplied by \(100\), or the percentage rms error per degree, is the theoretical quantity to be compared to the "actual" percentage rms error per degree derived from the analysis of simulated data with the same statistical characteristics (i.e., \(\sigma_{n\alpha}^2\), \(C_{n\alpha}\), and \(D\)) in formula (3.2).

The \(\sigma_{n\alpha}^2\) were computed with subroutine NORMAX (Appendix B).

To obtain the actual percentage rms error per degree, sets of simulated data were created on full regular grids as follows: the artificial data consisted of area means computed globally, using the algorithm outlined in paragraph (1.7) and subroutine SSYNTH (Appendix B), on the basis of expression (1.2). The \(\overline{C}_{n\alpha}\), complete to degree and order \(N_{\max} \geq N = \frac{27}{3\Lambda_n}\), came from sequences of random numbers. The random numbers, obtained using the IMSL subroutine "GGNOR" with generating seeds of the order of \(10^8\), were scaled to give them the desired degree variances \(\sigma_n^2\). For each simulation, a sequence of \((N_{\max} + 1)^2\) numbers was obtained, the first corresponding to \(\overline{C}_{00}\), the second, third, and fourth to \(\overline{C}_{10}, \overline{C}_{11}\), and \(\overline{C}_{11}\), respectively, and so forth. If \(r_{n0}, r_{n1}, \ldots, r_{n\alpha}\) were the \((2n+1)\) numbers corresponding to degree \(n\), then the scaling that resulted in the corresponding \(\overline{C}_{n\alpha}\) was

\[
\overline{C}_{n\alpha} = \left[ r_{n\alpha} \frac{\sigma_n}{\sqrt{\sum_{n=0}^{\infty} \sum_{\alpha=0}^{\infty} (r_{n\alpha})^2}} \right]^{\frac{3}{2}}
\]

(3.3)

The harmonic coefficients \(\overline{C}_{n\alpha}\) obtained in this way were the "actual" coefficients to which the \(\tilde{C}_{n\alpha}\), recovered by some of the procedures described in section 2, were then compared to obtain the actual percentage rms errors per degree

\[
\zeta_n = \left[ \sum_{n=0}^{\infty} \sum_{\alpha=0}^{\infty} (\tilde{C}_{n\alpha} - \overline{C}_{n\alpha})^2 \right]^{\frac{1}{2}} \sigma_n \times 100
\]

(3.4)

The analysis of the simulated data was done with subroutine HARMIN (Appendix B). This type of numerical experiment was carried out three or more times in each case, varying only the seed used to generate the random numbers, much as a Monte Carlotype of analysis is conducted. The seeds were chosen widely apart, to ensure that the correlation between "trials" would be virtually nil.
The maximum degree and order in the set of artificial coefficients, \( N_{\text{max}} \), was chosen so that the power in the mean values above degree \( N_{\text{max}} \)

\[
P_{N_{\text{max}}} = \sum_{n=N_{\text{max}}+1}^{2000} \beta_n^2 \sigma_n^2 \tag{3.5}
\]

were less than 1% of the power between degrees 0 and 2000. The \( \beta_n \) are the Pellinen coefficients, discussed in paragraph (4.3), corresponding to 5° x 5° area means. The values used for the degree variances \( \sigma_n^2 \) had been empirically obtained from terrestrial data in the manner described below. The data were supposed to be noise-free, as only the sampling part of the error was studied in this way, for the reasons given at the beginning of this paragraph. The estimators being linear, the propagated noise and the sampling error merely add arithmetically to each other, and can therefore be studied separately, if so desired. Summing up, it can be said that the simulated data consisted in global data sets of artificial gravity anomalies, averaged over equal angular grids.

The empirical degree variances were obtained as follows: up to degree 100 they were those implied by a set of coefficients, complete to degree 180, obtained by R. Rapp and associates at O.S.U. from a global data set of 1° x 1° mean anomalies. Above degree 100, the \( \sigma_n^2 (\text{Ag}) \) were obtained from a model of the form

\[
\sigma_n^2 (\text{Ag}) = (n-1) \left( \frac{\alpha_1}{(n+A)} S_1^{n-2} + \frac{\alpha_2}{(n+B)} S_2^{n-2} \right) \text{ [m gal] } \tag{3.6}
\]

(Moritz, 1976), where the parameters \( \alpha_1, \alpha_2, S_1, S_2, A, \) and \( B \) have been adjusted to fit existing gravimetric data, satellite altimetry, satellite field models, and other geophysical data. The parameters used in most examples were

\[
\begin{align*}
\alpha_1 &= 3.4050 & S_1 &= 0.998006 & A &= 1. \\
\alpha_2 &= 140.03 & S_2 &= 0.914232 & B &= 2.
\end{align*}
\]

corresponding to the best model of this type given in a report by R. Rapp (1979) who, in the same work, discusses also the empirical degree variances obtained from his 180, 180 field model. The degree variances implied by (3-6) with the parameters listed above are also very similar to those obtained by quite different means by Wagner and Colombo (1979), who analyzed the (Fourier) power spectrum of short arcs of GEOS-3 altimetry, and converted their average to a spherical harmonics spectrum using formulas that follow from the relationship between spherical harmonics and Fourier series. The empirical variances for \( n \leq 100 \) are included in the listing of subroutine NVAR, in Appendix B.

In order to understand how critical the choice of empirical degree variances is to the theoretical and actual errors, a different two-term model obtained by C. Jekeli (1978) was used as well. This model has the following parameter
values

\[ \alpha_1 = 18.3906 \quad S_1 = 0.9943667 \quad A = 140. \]
\[ \alpha_2 = 658.6132 \quad S_2 = 0.9048949 \quad B = 10. \]

All the examples considered in this section refer to complete equal angular sets of mean values with the same statistical properties of terrestrial gravity anomalies (as far as such properties are known). The analysis of actual mean gravity anomalies is shown in the last paragraph. Besides being important in geodetic studies, gravity anomalies constitute a type of geophysical data with reasonably well known statistical properties, and their study here is meant to give the reader some idea of how effective are the ideas presented earlier when it comes to handling "real data" (or something resembling it).

3.2 Agreement between the Actual and the Theoretical Measures of the Sampling Errors

Table (3.1) lists side by side the theoretical percentage rms per degree of the sampling error according to (3.2) and the actual value of this percentage for two different sets of coefficients (i.e., from random sequences with different seeds). The coefficients were recovered using the quadratures formula

\[ C_{n}^{\alpha} = \frac{1}{4\pi} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \tilde{g}_{ij} \int_{\sigma_{ij}} Y_{n}^{\alpha} (\theta, \lambda) \, d\sigma \]

This type of formulas has been discussed in paragraph (2.6). The simulated data consisted in full sets of 5° × 5° mean anomalies obtained from harmonic coefficients complete to degree and order \( N_{ax} = 140 \). The power above degree 140 in 5° × 5° anomalies is negligible, according to the empirical power spectrum model that was used. The results shown here are fairly typical of similar tests conducted with other quadrature formulas, so the conclusions that can be drawn are likely to be valid for the analysis of area means by numerical quadratures in general. There is clear agreement between the theoretical and the actual rms of the errors, and not just the average rms of actual errors, but the actual rms of each trial as well. The agreement is close, and the reader will probably agree that to use a theoretical error measure that can predict the actual error so well is a meaningful way of quantifying the error.

3.3 Accuracies of Various Quadratures Formulas

Five quadratures formulas for area means have been studied: the first four of the type

\[ C_{n}^{\alpha} = \mu_{n} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \int_{\sigma_{ij}} Y_{n}^{\alpha} (\theta, \lambda) \, d\sigma \, \tilde{g}_{ij} \]

-71-
Table 3.1

Comparison of actual versus theoretical percentage rms error per degree.
5° x 5° mean anomalies, \( N_{\text{max}} = 140 \), 0. mgal rms noise.

<table>
<thead>
<tr>
<th>( n )</th>
<th>Actual, No 1</th>
<th>Actual, No 2</th>
<th>Average</th>
<th>Theoretical((\text{expression (3.2) x 100)})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(seed=53218)</td>
<td>(seed=31765)</td>
<td>1 and 2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.50</td>
<td>0.62</td>
<td>0.56</td>
<td>0.51</td>
</tr>
<tr>
<td>5</td>
<td>1.06</td>
<td>1.10</td>
<td>1.08</td>
<td>1.23</td>
</tr>
<tr>
<td>10</td>
<td>4.99</td>
<td>5.81</td>
<td>5.40</td>
<td>4.89</td>
</tr>
<tr>
<td>15</td>
<td>13.06</td>
<td>11.03</td>
<td>12.05</td>
<td>12.75</td>
</tr>
<tr>
<td>20</td>
<td>23.48</td>
<td>25.97</td>
<td>24.72</td>
<td>24.77</td>
</tr>
<tr>
<td>25</td>
<td>23.04</td>
<td>30.82</td>
<td>26.93</td>
<td>30.03</td>
</tr>
<tr>
<td>30</td>
<td>45.84</td>
<td>46.26</td>
<td>46.05</td>
<td>43.28</td>
</tr>
<tr>
<td>*36 (N)</td>
<td>55.33</td>
<td>61.10</td>
<td>58.22</td>
<td>60.35</td>
</tr>
<tr>
<td>40</td>
<td>70.61</td>
<td>75.82</td>
<td>73.22</td>
<td>75.14</td>
</tr>
<tr>
<td>45</td>
<td>85.91</td>
<td>86.41</td>
<td>86.16</td>
<td>87.03</td>
</tr>
<tr>
<td>50</td>
<td>101.51</td>
<td>101.43</td>
<td>101.47</td>
<td>101.47</td>
</tr>
</tbody>
</table>

differ only in \( \mu_a \):

(a) \( \mu_a = \frac{\mu_a^*}{\mu_a} \) the optimal de-smoothing factor given by expression (2.36) in paragraph (2.7):

(b) \( \mu_a = \frac{1}{4\pi \beta_a^2} \)

(c) \( \mu_a = \frac{1}{4\pi \beta_a} \)

(d) \( \mu_a = \frac{1}{4\pi} \)

(e) The optimal quadratures - type formula, in the least squares collocation sense (i.e., minimum combined error measure) for the given grid; signal and noise (paragraph (2.8)). The grid was equal angular in all five cases. Table (3.2) compares the percentage rms of the errors per degree for a 30° x 30° grid; table (3.3) corresponds to a 10° x 10° grid; and table (3.4) to a 5° x 5° grid. \( N_{\text{max}} \) was 100 for the first two tables, and 140 for the last. All these values are theoretical, computed in accordance to the formulas in paragraph (3.1). In all three cases noise is not present, so these errors are purely sampling errors.

Table (3.5) corresponds to a 5° x 5° grid and an uniform noise of 5 mgal rms. Here the effect of the noise has been included in the results. Table (3.6) compare the degree correlation coefficients for the various methods, for a noise of 0 mgal. The correlation coefficient for the nth degree is defined as

* (N): "Nyquist frequency" of the grid.
\[ \rho_n = \left[ \sum_{n=0}^{\infty} \sum_{\alpha=0}^{\infty} C_{mn}^\alpha L_{mn}^\alpha \left( \sum_{n=0}^{\infty} \sum_{\alpha=0}^{\infty} C_{mn}^{\alpha 2}\right)^{-1} \left( \sum_{n=0}^{\infty} \sum_{\alpha=0}^{\infty} L_{mn}^{\alpha 2}\right)^{-1} \right]^{1/2} \] (3.7)

and it is also equal to

\[ \rho_n = \left[ \frac{\int_\sigma \Delta g_n \Delta g_n \, d\sigma}{\int_\sigma \Delta g_n^2 \, d\sigma \int_\sigma \Delta g_n^2 \, d\sigma} \right]^{1/2} \] (3.8)

This coefficient can be regarded either as a measure of the agreement between the actual and the recovered coefficients of the nth harmonic, or of between the nth harmonic \( \Delta g_n \) in the signal and the harmonic \( \Delta g_n^\alpha \) that can be computed from the recovered coefficients. If the \( C_{mn}^\alpha \) could be seen as random variables with Gaussian distribution, the interpretation of \( \rho_n \) would move along well-worn paths; however, as it was mentioned in paragraph (2.5), there are some unexpected problems when extending the idea of a Gaussian random process to the sphere, so is better to choose another approach. One could regard the coefficients of the nth harmonic as the coordinates of a vector in \((2n + 1)\)-dimensional Euclidean space. The actual coefficients will define thus one vector, and the recovered coefficients another. Expression (3.7) then merely defines \( \rho_n \) as the scalar product of these two vectors. Likewise, expression (3.8) is that of the scalar product of two elements of a function space. The angle formed by these vectors is 0° when correlation is (maximum), and 90° for 0 correlation; a minimum correlation of -1 corresponds to the case when the vectors are equal but of opposite sense. The scalar product is independent of any scale factors that may multiply the vectors; it depends only on their mutual orientation. For this reason, the correlation coefficients are the same for the four quadratures formulas of the type

\[ C_{mn}^\alpha = \mu_n \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \Delta E_{ij} \int_\sigma \mathcal{Y}_{mn}^\alpha (\theta, \lambda) \, d\sigma \]

because the difference between the errors for the same harmonic, predicted with two different formulas of this kind, consists in a scale factor \( \mu_n^{(1)} / \mu_n^{(2)} \).

Clearly, as the rms of the error increases with \( n \), \( \rho_n \) decreases from almost 1 where the error is smallest (very low degrees) to below 0.5 where the error exceeds 90% (highest degrees analyzed).

Observing the percentual rms of the errors in the first four tables, it is easy to see that they by no means reach 100% as soon as the Nyquist frequency is reached \( n = N \), but that they remain substantially smaller than 100% even at degrees considerably higher than \( N \); this is in line with the conclusions in paragraph (1.3). The optimal estimator itself cannot have an error larger than 100%, be it due to sampling, noise, or both. Otherwise, a null estimator (one that predicts only zeroes) would be better than the optimal, which is not possible.

Table (3.7), compares the theoretical errors with zero noise (i.e., the sampling errors) of the collocation estimator obtained, first according to
\( \sigma_n^2 \) implied by R. Rapp's model (used in all the other tables), and then according to the "2L" model of C. Jegely, both described in paragraph (3.1). This table is included here to give the reader an idea of how sensitive the theoretical error variances are to the empirical degree variances used to compute them. The "2L" model has considerably more power than Rapp's at high degrees, and this may be reflected in the somewhat larger errors in the corresponding column of the table.

### Table 3.2
Theoretical percentage rms error per degree.
30° x 30° mean anomalies, \( N_{max} = 100, \) 0. mgal rms noise.

<table>
<thead>
<tr>
<th>n</th>
<th>( \mu_n )</th>
<th>( \frac{1}{4\pi\beta_n} )</th>
<th>( \frac{1}{4\pi} )</th>
<th>( \frac{1}{4\pi\beta_n^2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>12.51</td>
<td>13.17</td>
<td>13.95</td>
<td>15.86</td>
</tr>
<tr>
<td>4</td>
<td>25.09</td>
<td>25.60</td>
<td>26.91</td>
<td>36.12</td>
</tr>
<tr>
<td>6</td>
<td>58.77</td>
<td>65.13</td>
<td>68.03</td>
<td>67.18</td>
</tr>
<tr>
<td></td>
<td>92.82</td>
<td>96.01</td>
<td>194.08</td>
<td>97.13</td>
</tr>
<tr>
<td>12</td>
<td>98.44</td>
<td>99.27</td>
<td>915.49</td>
<td>105.44</td>
</tr>
</tbody>
</table>

### Table 3.3
Theoretical percentage rms error per degree.
10° x 10° mean anomalies, \( N_{max} = 100, \) 0. mgal rms noise.

<table>
<thead>
<tr>
<th>n</th>
<th>( \mu_n )</th>
<th>( \frac{1}{4\pi\beta_n} )</th>
<th>( \frac{1}{4\pi} )</th>
<th>( \frac{1}{4\pi\beta_n^2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.52</td>
<td>1.63</td>
<td>1.71</td>
<td>2.04</td>
</tr>
<tr>
<td>4</td>
<td>2.24</td>
<td>2.46</td>
<td>3.00</td>
<td>4.76</td>
</tr>
<tr>
<td>6</td>
<td>4.27</td>
<td>4.68</td>
<td>5.83</td>
<td>9.51</td>
</tr>
<tr>
<td>8</td>
<td>8.74</td>
<td>9.60</td>
<td>10.99</td>
<td>16.39</td>
</tr>
<tr>
<td>10</td>
<td>13.19</td>
<td>14.02</td>
<td>15.89</td>
<td>23.68</td>
</tr>
<tr>
<td>12</td>
<td>31.60</td>
<td>33.03</td>
<td>33.10</td>
<td>37.84</td>
</tr>
<tr>
<td>14</td>
<td>40.43</td>
<td>41.41</td>
<td>41.41</td>
<td>46.45</td>
</tr>
<tr>
<td>16</td>
<td>41.41</td>
<td>42.18</td>
<td>42.36</td>
<td>51.41</td>
</tr>
<tr>
<td>18</td>
<td>56.81</td>
<td>59.94</td>
<td>61.26</td>
<td>63.41</td>
</tr>
<tr>
<td>20</td>
<td>76.50</td>
<td>78.99</td>
<td>94.69</td>
<td>79.07</td>
</tr>
<tr>
<td>22</td>
<td>76.02</td>
<td>78.54</td>
<td>89.75</td>
<td>79.05</td>
</tr>
<tr>
<td>24</td>
<td>76.02</td>
<td>78.54</td>
<td>116.57</td>
<td>87.59</td>
</tr>
<tr>
<td>26</td>
<td>90.40</td>
<td>94.03</td>
<td>164.50</td>
<td>97.85</td>
</tr>
<tr>
<td>28</td>
<td>92.21</td>
<td>95.98</td>
<td>193.17</td>
<td>100.35</td>
</tr>
<tr>
<td>30</td>
<td>93.31</td>
<td>98.30</td>
<td>232.24</td>
<td>101.88</td>
</tr>
<tr>
<td>32</td>
<td>95.80</td>
<td>98.30</td>
<td>303.89</td>
<td>103.49</td>
</tr>
<tr>
<td>34</td>
<td>94.95</td>
<td>97.40</td>
<td>281.73</td>
<td>97.77</td>
</tr>
<tr>
<td>36</td>
<td>96.87</td>
<td>98.08</td>
<td>485.72</td>
<td>98.45</td>
</tr>
</tbody>
</table>
### Table 3.4
Theoretical percentage rms error per degree
$5^\circ \times 5^\circ$ mean anomalies, $N_{max} = 140$, 0. mgal rms noise.

<table>
<thead>
<tr>
<th>n</th>
<th>Optimal Estim.</th>
<th>$\mu_n$</th>
<th>$\frac{1}{4\pi \beta_n}$</th>
<th>$\frac{1}{4\pi}$</th>
<th>$\frac{1}{4\pi \beta_n^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.47</td>
<td>0.49</td>
<td>0.51</td>
<td>0.59</td>
<td>0.50</td>
</tr>
<tr>
<td>6</td>
<td>1.36</td>
<td>1.44</td>
<td>1.71</td>
<td>2.61</td>
<td>1.48</td>
</tr>
<tr>
<td>12</td>
<td>9.74</td>
<td>10.11</td>
<td>10.40</td>
<td>12.27</td>
<td>10.37</td>
</tr>
<tr>
<td>18</td>
<td>15.14</td>
<td>15.55</td>
<td>16.35</td>
<td>21.20</td>
<td>16.46</td>
</tr>
<tr>
<td>24</td>
<td>30.58</td>
<td>31.06</td>
<td>31.21</td>
<td>36.30</td>
<td>34.99</td>
</tr>
<tr>
<td>30</td>
<td>42.88</td>
<td>43.28</td>
<td>43.28</td>
<td>49.04</td>
<td>53.32</td>
</tr>
<tr>
<td>36 (N)</td>
<td>57.62</td>
<td>59.15</td>
<td>60.34</td>
<td>62.67</td>
<td>85.90</td>
</tr>
<tr>
<td>42</td>
<td>72.16</td>
<td>73.51</td>
<td>80.25</td>
<td>74.72</td>
<td>137.19</td>
</tr>
<tr>
<td>48</td>
<td>82.15</td>
<td>83.62</td>
<td>101.19</td>
<td>83.82</td>
<td>216.98</td>
</tr>
</tbody>
</table>

### Table 3.5
Theoretical percentage rms per degree
$5^\circ \times 5^\circ$ mean anomalies, $N_{max} = 140$, 5. mgal rms noise.

<table>
<thead>
<tr>
<th>n</th>
<th>Optimal Estim.</th>
<th>$\mu_n$</th>
<th>$\frac{1}{4\pi \beta_n}$</th>
<th>$\frac{1}{4\pi}$</th>
<th>$\frac{1}{4\pi \beta_n^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8.74</td>
<td>8.81</td>
<td>8.84</td>
<td>8.83</td>
<td>8.85</td>
</tr>
<tr>
<td>6</td>
<td>9.33</td>
<td>9.34</td>
<td>9.34</td>
<td>9.43</td>
<td>9.41</td>
</tr>
<tr>
<td>12</td>
<td>32.80</td>
<td>32.88</td>
<td>33.79</td>
<td>33.92</td>
<td>35.28</td>
</tr>
<tr>
<td>18</td>
<td>35.94</td>
<td>36.04</td>
<td>36.54</td>
<td>39.25</td>
<td>39.86</td>
</tr>
<tr>
<td>24</td>
<td>53.32</td>
<td>53.47</td>
<td>56.73</td>
<td>66.49</td>
<td>66.99</td>
</tr>
<tr>
<td>30</td>
<td>61.97</td>
<td>62.14</td>
<td>67.15</td>
<td>87.22</td>
<td>87.72</td>
</tr>
<tr>
<td>36 (N)</td>
<td>71.98</td>
<td>72.62</td>
<td>82.81</td>
<td>122.67</td>
<td>123.19</td>
</tr>
<tr>
<td>42</td>
<td>81.55</td>
<td>82.17</td>
<td>102.91</td>
<td>181.19</td>
<td>181.77</td>
</tr>
<tr>
<td>48</td>
<td>88.00</td>
<td>88.72</td>
<td>124.09</td>
<td>271.87</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3.6
Correlation factor per degree.
$5^\circ \times 5^\circ$ mean anomalies, $N_{max} = 140$, 0. mgal noise.

<table>
<thead>
<tr>
<th>n</th>
<th>Optimal Estim.</th>
<th>Simple quadratures</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>6</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>12</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>18</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>24</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>30</td>
<td>0.90</td>
<td>0.89</td>
</tr>
<tr>
<td>36 (N)</td>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td>42</td>
<td>0.72</td>
<td>0.71</td>
</tr>
<tr>
<td>48</td>
<td>0.59</td>
<td>0.57</td>
</tr>
</tbody>
</table>
The results in the tables show that the optimal estimator errors are the smallest, as expected, and that the quadratures formula with the optimal de-smoothing factors is the best of the four simple quadratures formulas compared here, though it is not quite as good as the optimal estimator, also as expected. In the case of zero noise each of the three non-optimal quadratures formulas has errors that, in some region of the spectrum, are smaller than those for the other two; concretely, the de-smoothing factor \( \mu_n = \frac{1}{4\pi \beta_n} \) works better for low degree harmonics (or \( n \leq \frac{1}{3} N \)), i.e., the percentage rms of the error is less than for the other two formulas there, the factor \( \mu_n = \frac{1}{4\pi \beta_n} \) is best for middle harmonics (approximately \( 1/3N < n \leq N \)), and \( \mu_n = \frac{1}{4\pi} \) is best above the Nyquist frequency \( N \). It is possible, therefore, to obtain a simple "composite" quadratures formula that combines the good properties of all three formulas, by defining its de-smoothing factor as follows:

\[
\mu_n = \frac{1}{4\pi \eta_n} \quad \text{where} \quad \eta_n = \begin{cases} 
\beta_n^2 & \text{if } 0 \leq n \leq 1/3N \\
\beta_n & \text{if } 1/3N < n \leq N \\
1 & \text{if } n > N
\end{cases} \tag{3.9}
\]

This composite formula has been implemented in the version of subroutine HARMIN listed in Appendix B, through the subroutine can be easily changed to compute other quadrature formulas.

It is clear from the tables that, while better than all the others, the optimal formula is only marginally so: from a practical point of view, the simple "composite" formula (3.9) above is virtually as good, but it is much easier to implement and compute. Therefore, when analysing data of the type considered here (resembling mean anomalies with uniform noise on an equal angular grid), the quadratures formulas discussed the composite in particular, are about as good as any linear technique for estimating the coefficients, and also very easy to program and very efficient.
There may be cases, however, when this is not true. If the data set were very noisy and/or incomplete, or if the coefficients to be recovered were not those of the signal, but those of some complicated transformation of it (as in the case of satellite-to-satellite tracking data, where the signal depends on a combination of differences of radial and horizontal derivatives of the gravitational potential) so no simple integral formula like (1.4) exists that can be readily discretized into summations like (1.5) or (1.6), then simple quadrature formulas like those in section 1 would be of no real use, and the optimal estimator could provide the only practical way of obtaining the coefficients. Both collocation and least squares adjustment are very similar, as shown in section 2, and both can be implemented with reasonable efficiency if nothing simpler is available. More important, the optimal formulas provide a theoretical background on which one can build a coherent and comprehensive understanding of the other linear techniques for spherical harmonic analysis. It is, after all, because of the theory developed in the previous section that it has been possible to obtain the results shown in the preceding tables, results that constitute the factual basis for these considerations.

The techniques for setting up and inverting the variance-covariance matrix of the data are of interest in a number of estimation and filtering problems, besides harmonic analysis. The author hopes that the examples provided in this section will encourage the wider use of least squares collocation for the processing of large, global sets of data, both point values and area means. Creating the simulated mean anomalies with subroutine SSYTH, obtaining the "weights" $\chi_i^m$ of the optimal quadratures formula with NORMAL (Appendix B), and recovering the $\hat{\lambda}_m^\alpha$ for comparison with the original $C_{22}^{\alpha}$ up to degree and order 72 (same 5000 coefficients) on a $5^\circ \times 5^\circ$ degree grid (about 2600 "data" values) took less than 20 seconds, central processor unit time, in the AMDHAL computer at Ohio State. To recover the harmonic coefficients up to degree and order 180 from a complete equal angular set of $1^\circ \times 1^\circ$ mean values would require less than two hours, using the same machine. Most of this time would be dedicated to creating and inverting $(C_{22}^{\alpha} + D)$.

But in the optimal estimation and filtering of geoidal undulations, deflections of the vertical, and any other function of the gravity field estimable from (say) gravity anomalies, the fact that $(C_{22}^{\alpha} + D)$ can be set up and inverted efficiently transcends harmonic analysis. A major implication is that such extremely large global adjustments require a computational effort that is already within the reach of most researchers.

As mentioned already in section 2, matrix $(C_{22}^{\alpha} + D)$ may become poorer conditioned, i.e., its numerical inversion less stable, as the data distribution becomes denser. This tendency towards instability was noticed: when there was no noise ($D = 0$) the $R(m)$ matrices had to be regularized by adding a small positive constant $k$ to each diagonal term (Colombo, 1979, par. (4.5)) before they could be inverted successfully. This constant, which in most cases was much smaller than the diagonal elements it was being added to, was $10^{-6}$ for $30^\circ \times 30^\circ$ and $10^\circ \times 10^\circ$, but had to be increased to $10^{-5}$ in the case of a $5^\circ \times 5^\circ$ grid. When noise was present, the nonzero diagonal elements in $D$
were sufficient to provide stability, and no regularization was needed. Because of this tendency to instability, the "band-limited" approach of paragraphs (2.13) to (2.17) may be preferable, whenever it can be properly applied.

As shown in Table (3.1), the theoretical and the actual sampling errors are almost the same in most cases. The propagated noise measure is very easy to compute in the case of uncorrelated noise, and can be added to the actual sampling error (variance) to get an estimate of the total error, both actual and theoretical. This estimate, where the sampling part is the result of a Monte Carlo-like approach, is much easier to obtain than the theoretical one that involves setting up matrix \( C_{xx} + D \), or at least the \( R(m) \) matrices. In the case of equal angular data sets like those considered here, this empirical estimate is likely to be just as accurate, when it comes to judge the performance of any given type of harmonic analysis. Such estimate has been used to evaluate the likely errors in the potential coefficients obtained from \( 1^\circ \times 1^\circ \) mean anomalies using the quadratures formula

\[
\alpha_{\alpha} = \frac{1}{4\pi \beta_{\alpha} \sum \frac{1}{\xi}} \int_{\alpha} \frac{1}{\xi} \sum \int_{\gamma} \left( \phi, \lambda \right) d\sigma \int_{\delta} \, d\gamma \, d\phi \, d\lambda \, d\xi \, d\eta .
\]

The Monte Carlo method described in paragraph (3.1) was implemented with the help of subroutines SSYNT and HARMINE, the error variance being equated to its average over three "trials" (three sets of coefficients created from different random sequences), C. Jekell, also at O.S.U., who undertook this work as part of his own research, fitted a quartic to the percentage rms per degree thus obtained. When this was expressed as a function of the "normalized degree" \( n/N \), the quartic fitted equally well the theoretical results for \( 30^\circ \times 30^\circ \), \( 10^\circ \times 10^\circ \), and \( 5^\circ \times 5^\circ \) presented in this section. Jekell's quartic expression for the truncation error is (private communication):

\[
\frac{\sigma_{\varepsilon}}{C_{\varepsilon}} \times 100 = \left[ \left( -16.19570 \left( \frac{n}{N} \right) + 30.34506 \left( \frac{n}{N} \right) + 40.29588 \left( \frac{n}{N} \right)^2 \right) \right]
\]

It is quite remarkable that such a complex phenomenon can be described satisfactorily by such a simple law.

The expansion of the simulated \( 1^\circ \times 1^\circ \) mean anomalies was complete up to degree and order \( N_{max} = 300 \). Creating (or analysing) area mean values up to degree and order \( N_{max} = 300 \) required about 50 seconds c.p.u. time using double precision arithmetic in the AMDHAL computer at O.S.U.

Table 3.8 shows the actual percentage rms sampling error per degree as computed in one of the trials, and the percentage rms propagated noise (theoretical) corresponding to a 1 mgal rms noise in the data. Clearly, the errors are much smaller than for any of the cases considered previously; this improvement is due to the finer sampling (the sampling error tends to zero as the area of the blocks tends to zero). The data, however, tends to be noisier when averaged on smaller blocks, so the propagated noise may increase. For a given rms error in the data, multiply the number in the "propagated noise" column by this rms (in mgals) to obtain the corresponding percentage. These numbers are only valid for the estimator where \( \mu_n = \frac{1}{4\pi \beta} \). Repeated trials with different random coefficients resulted in much the same percentages for the sampling errors, so these values are probably fairly typical.
Table 3.8
Percentage rms errors per degree for a $1^\circ \times 1^\circ$
equal angular grid. $\mu_n = \frac{1}{4\pi \beta_n}, N_{nax} = 300,$
$\epsilon \Delta g = 1$ mgal

<table>
<thead>
<tr>
<th>n</th>
<th>actual sampling</th>
<th>propagated error (theor.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.01</td>
<td>0.47</td>
</tr>
<tr>
<td>10</td>
<td>0.10</td>
<td>0.72</td>
</tr>
<tr>
<td>20</td>
<td>0.46</td>
<td>1.39</td>
</tr>
<tr>
<td>30</td>
<td>0.98</td>
<td>1.72</td>
</tr>
<tr>
<td>40</td>
<td>1.55</td>
<td>1.93</td>
</tr>
<tr>
<td>50</td>
<td>2.68</td>
<td>2.07</td>
</tr>
<tr>
<td>60</td>
<td>4.71</td>
<td>2.20</td>
</tr>
<tr>
<td>70</td>
<td>6.94</td>
<td>2.34</td>
</tr>
<tr>
<td>80</td>
<td>9.43</td>
<td>2.47</td>
</tr>
<tr>
<td>90</td>
<td>13.51</td>
<td>2.62</td>
</tr>
<tr>
<td>100</td>
<td>14.79</td>
<td>2.77</td>
</tr>
<tr>
<td>110</td>
<td>17.98</td>
<td>2.96</td>
</tr>
<tr>
<td>120</td>
<td>25.06</td>
<td>3.08</td>
</tr>
<tr>
<td>130</td>
<td>25.33</td>
<td>3.26</td>
</tr>
<tr>
<td>140</td>
<td>30.87</td>
<td>3.43</td>
</tr>
<tr>
<td>150</td>
<td>37.02</td>
<td>3.62</td>
</tr>
<tr>
<td>160</td>
<td>44.17</td>
<td>3.81</td>
</tr>
<tr>
<td>170</td>
<td>44.43</td>
<td>4.03</td>
</tr>
<tr>
<td>180 (N)</td>
<td>53.07</td>
<td>4.24</td>
</tr>
<tr>
<td>190</td>
<td>61.87</td>
<td>4.47</td>
</tr>
<tr>
<td>200</td>
<td>65.90</td>
<td>4.73</td>
</tr>
</tbody>
</table>

A possible way of bringing the sampling error down is to use weighted
area means of the form

$$\Delta \bar{g}_{11} = \sum_{k=1}^{N_{11}} \Delta g_{11}^{(k)} W_k$$

where $\Delta g_{11}^{(k)}$ is the kth measurement inside the block $\sigma_{11}$, and where the
$W_k$ are functions of the distance of $\Delta g_{11}^{(k)}$ to the center of the block. If the
$W_k$ decay gently towards the border of the area element, the resulting
weighted means will be smoother than the ordinary area means considered so
far (all $W_k = N_{11}^{-1}$) and their harmonic content above the Nyquist frequency
will be attenuated. Consequently, the harmonics below $N$ can be recovered
with less sampling error. This idea certainly deserves further study. Ob-
viously, it is applicable only in those cases where the original measurements
$\Delta g_{11}^{(k)}$ are available.
3.4 The Analysis of a Global Data Set of $5^\circ \times 5^\circ$ Mean Anomalies

As a final demonstration of the use of optimal estimators, this paragraph presents some results obtained by analysing a real data set, consisting of $5^\circ \times 5^\circ$ mean gravity anomalies. These anomalies were obtained from a global set of $1^\circ \times 1^\circ$ mean values created by R. Rapp and associates (Rapp, 1978) from the combination of land and gravity measurements, satellite altimetry over the oceans, and the GEM-9 satellite model (Lerch et al., 1979).

The $5^\circ \times 5^\circ$ values, $\Delta \bar{g}(\varepsilon)$, were obtained from the $1^\circ \times 1^\circ$ values, $\Delta \bar{g}(\varepsilon)$, using the formula

$$\Delta \bar{g}(\varepsilon) = \frac{1}{25} \sum_{i=1}^{10} \Delta \bar{g}(\varepsilon)_{i}$$

where $\sum_{i}$ represents summation over all $1^\circ \times 1^\circ$ blocks inside a $5^\circ \times 5^\circ$ block. The variances of the $\Delta \bar{g}(\varepsilon)$ were obtained from the formula

$$\sigma^2 \Delta \bar{g}(\varepsilon) = \frac{1}{(25)^2} \sum_{i=1}^{10} \sigma^2 \Delta \bar{g}(\varepsilon)_{i}$$

where the assumption has been made that the errors in the $1^\circ \times 1^\circ$ values were uncorrelated, which is not quite true, as the values used were the product of an adjustment. The variances $\sigma^2 \Delta \bar{g}(\varepsilon)$ were different from $5^\circ$ block to $5^\circ$ block, but they were "homogenized" as described in paragraph (2.9), in order to obtain a $(C_{22} + D)$ matrix that was easily invertible and resulted in a quadrature-type estimator inexpensive to implement.

Figure (3.1) shows a comparison between the collocation solution, complete to degree and order 36, and the same coefficients obtained by R. Rapp by numerical quadratures from the original $1^\circ \times 1^\circ$ values. The figure shows the rms per degree ($\sqrt{\sigma^2/(2n+1)}$) for potential coefficients $(\sigma_2^0(T)) = \sigma_2^0(\Delta \bar{g}) \gamma^{n}(n-1)^{2n}, \gamma = 979.800 \text{ mgal})$. The circles correspond to Rapp's results, and the triangles to those obtained using collocation as already explained. Because the grid used by Rapp is much finer, the corresponding results are likely to be less affected by sampling errors, at least up to degree 36, than those obtained from the $5^\circ \times 5^\circ$ anomalies; for this reason Rapp's rms values are regarded here as the "true" ones. The solid line corresponds to C. Jelekli's "2L" model for the $\sigma_2^0$, used here to obtain the optimal estimator. It is interesting to notice that the triangles follow the circles (or "true" values) rather than the line. A common concern among those using this type of estimators is to what extent the "a priori" power spectrum or covariance function used to set up the estimator may "bias" the results by forcing the spectrum of the output to resemble the "a priori" spectrum. Here there is little evidence of such a "bias".

In addition to Rapp's coefficients, those of GEM-10B (Lerch, 1980) were also used for comparison. The "collocation" values follow them very closely too (the three sets of results agree, in fact, very well with each other). The
GEM values are not shown in the figure, to have a cleaner picture. In any event, the "collocation" values compare very well with the other two, and in the higher degrees (31 to 36) where the divergence between Rapp's model and collocation is largest, GEM-10B fits right in between them. The results shown here were presented in a previous paper by the author (Colombo, 1979-b).

Figure 3.1

- Rapp 1978
- Collocation (5°×5°)
- C. Jekeli's "2L" Covariance Function
  (O.S.U. Report 275)
4. Covariances Between Area Means

In what follows, the expression "point" covariance refers to

\[
\text{cov}(u(P), v(P')) = \sum_{n=0}^{\infty} \sum_{\lambda} c_n^u c_n^v P_n(\psi_{\theta \theta'}) = M[u(P), v(P')] \quad (4.1)
\]

The covariance between the area means of two functions \( u \) and \( v \) is related to the "point" covariance function by the following integral relationship (Sjöberg, 1978):

\[
\text{cov}(\bar{u}_{ij}, \bar{v}_{kl}) = \frac{1}{\Delta_{ij} \Delta_{kl}} \int_{\sigma_{ij}} \int_{\sigma_{kl}} \text{cov}(u(P), v(P')) d\sigma' \quad (4.2)
\]

where \( P \) belongs to the area block \( \sigma_{ij} \), and \( P' \) to \( \sigma_{kl} \), while

\[
\bar{u}_{ij} = \frac{1}{\Delta_{ij}} \int_{\sigma_{ij}} u(P) d\sigma
\]

is the average of \( u \) over the block \( \sigma_{ij} \) of area

\[
\Delta_{ij} = \Delta \lambda (\cos(\theta_i) - \cos(\theta_i + \Delta \theta_i)) \quad (4.4)
\]

\( \Delta \theta_i, \Delta \lambda \) are the colatitude and longitude spans of the blocks in the "row" between colatitudes \( \theta_i \) and \( \theta_i + \Delta \theta_i \). To simplify the discussion, \( \Delta \theta_i = \Delta \theta \) is assumed constant; extension to the more general case where \( \Delta \theta_i \) varies from row to row is straightforward.

4.1. Derivation of an Approximate Formula for the Covariance \(^1\)

Expression (4.2) can be computed by numerical quadratures (Rapp, 1977). Because of the double integral, this is a very laborious process, and it is not practical if many covariance values have to be found, as in the case of large data sets. A more efficient alternative is needed.

Replacing the covariance function in the integrand of (4.2) with its Legendre expansion

\[
\text{cov}(\bar{u}_{ij}, \bar{v}_{kl}) = \frac{1}{\Delta_{ij} \Delta_{kl}} \int_{\sigma_{ij}} \int_{\sigma_{kl}} \sum_{n=0}^{\infty} c_n^u c_n^v P_n(\psi_{\theta \theta'}) \quad (4.5)
\]

Here \( \psi_{\theta \theta'} = \cos^{-1} (\cos \theta \cos \theta' + \sin \theta \sin \theta' \cos(\lambda - \lambda')) \) is the spherical distance between points \( P \equiv (\theta, \lambda) \), and \( P' \equiv (\theta', \lambda') \) in the unit sphere, while

\[
c_n^{uv} = \sum_{n=0}^{\infty} C_n^u C_n^v + S_n^u S_n^v
\]

is the nth degree variance of the cross-spectrum of \( u \) and \( v \). As shown in the Appendix, the order of summation and integration can be reversed.

\(^1\) In this section, expression \( P_n(\psi_{\theta \theta'}) \) is shorthand notation for \( P_n(\cos \psi_{\theta \theta'}) \).
\[
\text{cov}(\vec{u}_{ij}, \vec{v}_{kl}) = \frac{1}{\Delta_{ij} \Delta_{kl}} \sum_{n=0}^{\infty} \sum_{l=0}^{n} c_n^{ij} \int_{\sigma_{ij}} d\sigma \int_{\sigma_{kl}} P_n(\psi_{ij}) P_l(\psi_{kl}) d\sigma' \]
\[
(4.6)
\]

Let \( P_n(\cos \theta) \) be the fully normalized associate Legendre function of the first kind of degree \( n \) and order \( m \). Applying the Summation Theorem for such functions, the \( P_n(\psi_{ij}) \), can be replaced in the integrand as shown below:

\[
\text{cov}(\vec{u}_{ij}, \vec{v}_{kl}) = \frac{1}{\Delta_{ij} \Delta_{kl}} \sum_{n=0}^{\infty} \frac{c_n^{ij}}{2n+1} \int_{\sigma_{ij}} d\sigma \int_{\sigma_{kl}} P_n(\cos \theta) P_n(\cos \theta') \cos m(\lambda - \lambda') d\sigma'
\]
\[
(4.7)
\]

The sum in the general term of the series above has a finite number of terms \((n+1)\), so it is valid to exchange summation and integration once more:

\[
\text{cov}(\vec{u}_{ij}, \vec{v}_{kl}) = \frac{1}{\Delta_{ij} \Delta_{kl}} \sum_{n=0}^{\infty} \frac{c_n^{ij}}{2n+1} \int_{\sigma_{ij}} d\sigma \int_{\sigma_{kl}} P_n(\cos \theta) P_n(\cos \theta') \cos m(\lambda - \lambda') d\sigma' \int_{\sigma_{ij}} d\sigma
\]
\[
(4.8)
\]

Writing out the area integrals as colatitude and longitude integrals:

\[
\text{cov}(\vec{u}_{ij}, \vec{v}_{kl}) = \frac{1}{\Delta_{ij} \Delta_{kl}} \sum_{n=0}^{\infty} \frac{c_n^{ij}}{2n+1} \sum_{l=0}^{n} \int_{\sigma_{ij}} d\sigma \int_{\sigma_{kl}} P_n(\cos \theta) P_n(\cos \theta') \cos m(\lambda - \lambda') d\sigma' \int_{\sigma_{ij}} d\sigma
\]
\[
(4.9)
\]

Calling \( \Delta_1 \equiv \Delta_{ij} \), as all blocks in the same "row" have the same area, we introduce two functions,

\[
F(m) = \begin{cases} 
\Delta \lambda^2 & \text{if } m = 0 \\
\frac{1}{2n+1} \frac{1}{(2m^2)(1-\cos m\Delta \lambda)} & \text{otherwise,}
\end{cases}
\]

and

\[
I_{m,n} = \left( \frac{c_n^{ij}}{2n+1} \right)^{\frac{1}{2}} \Delta_1^{\frac{1}{2}} \int_{\sigma_{ij}} d\sigma \int_{\sigma_{kl}} P_n(\cos \theta) \sin \theta' d\theta' \int_{\sigma_{ij}} d\sigma \int_{\sigma_{kl}} P_n(\cos \theta) \sin \theta \cos m(\lambda - \lambda') d\sigma' d\theta
\]
\[
(4.10-a)
\]

Then expression (4.9) can be written, after integrating and reordering terms, as follows:

\[
\text{cov}(\vec{u}_{ij}, \vec{v}_{kl}) = \sum_{n=0}^{\infty} F(m) \sum_{m=0}^{\infty} I_{m,n} I_{m,k} \cos m(\lambda_1 - \lambda_2)
\]
\[
(4.11-a)
\]

This regrouping of the series is valid because, as shown in the Appendix, the series is absolutely convergent. The integrals in the definition of \( I_{m,n} \) (expression (4.10-b)) can be calculated very accurately and efficiently with recursive formulae obtained by M. K. Paul (1979). Regarding \( j \) as being fixed, the last expression is also that of the Fourier series of cosines of \( \text{cov}(u_{ij}, v_{kl}) \), with amplitudes

\[
a_{m}^{ij,k} = F(m) \sum_{n=0}^{\infty} I_{m,n} I_{m,k}
\]
\[
(4.11-b)
\]

and phases \( \varphi_\lambda = -m \lambda_1 \).

-88-
While (4.11-a, b) are valid for all values of \( \lambda_j \) in the interval \( 0 \leq \lambda_j \leq 2\pi \), this expression has to be calculated only for \( \lambda_j = 0, \Delta \lambda, 2\Delta \lambda, \ldots, j\Delta \lambda \). According to the sampling theorem for ordinary Fourier Series, at such regularly spaced \( \lambda_j \) expression (4.11-a) takes precisely the same values as the finite sum

\[
\text{cov} \left( \bar{u}_{1j}, \bar{v}_{kl} \right) = \sum_{n=0}^{N} \hat{a}_n \cos (\lambda_j - \lambda_l)
\]

where \( N = \pi / \Delta \lambda \)

\[
\hat{a}_n = \sum_{h=0}^{\infty} a_{2nh + n} + \sum_{h=1}^{\infty} a_{2nh - n}
\]

(4.13)

To calculate the covariances, (4.11-a, b) and (4.13) have to be truncated, excluding all harmonics above some degree \( N_{\text{max}} \); (4.12) becomes:

\[
\text{cov} \left( \bar{u}_{1j}, \bar{v}_{kl} \right) \approx \text{cov} \left( \bar{u}_{1j}, \bar{v}_{kl} \right)_{\text{max}} = \sum_{n=0}^{N_{\text{max}}} \sum_{h=0}^{K} \sum_{m=0}^{N_{\text{max}}} \left( I_{n, 2nh + n} \hat{a}_{n} \right) \left( I_{m, 2nh - n} \hat{a}_{m} \right) F(m) \cos (\lambda_j - \lambda_l)
\]

(4.14)

where \((2K + 1)N \leq N_{\text{max}}\).

4.2. Choosing \( N_{\text{max}} \)

To calculate the values of covariances between equal angular mean anomalies, \( \bar{\Delta g} \), the value of \( N_{\text{max}} \) could be chosen so that the percentage error

\[
\nu = \left| \frac{\text{cov} \left( \bar{\Delta g}_{1j}, \bar{\Delta g}_{kl} \right) - \text{cov} \left( \bar{\Delta g}_{1j}, \bar{\Delta g}_{kl} \right)_{\text{max}}}{\text{cov} \left( \bar{\Delta g}_{1j}, \bar{\Delta g}_{kl} \right)} \right| \times 100
\]

(\( \text{cov} \left( \bar{\Delta g}_{1j}, \bar{\Delta g}_{kl} \right) \) being computed by numerical quadratures) does not exceed a prescribed upper bound \( \epsilon \). The smallest \( N_{\text{max}} \) that meets this condition increases towards the poles, because the decrease in area of equal angular blocks with latitude means that the averages have a high frequency content that increases accordingly. On the other hand, the absolute values of the integrals of the \( \tilde{P}_{nm} \), and therefore the Fourier coefficients \( a_n \) in (4.12), decrease quite fast with increasing order \( m \) near the poles, so their contribution soon becomes insignificant. This is fortunate, because the need for lengthy calculations for each Fourier coefficient nearer to the poles can be offset by the existence of fewer coefficients there. In fact, because of the finite arithmetic of digital computers, all \( \hat{a}_n \) for blocks less than 30° from the poles are rounded off to zero for \( m \) considerably less than \( N \) in the cases presented here. Because of this, calculations near the poles can be less laborious than close to the equator, in spite of the larger \( N_{\text{max}} \). To take advantage of this in the programming of (4.14), the \( \hat{a}_n \) coefficients were not computed above the first \( m \) for which the following condition was met:
\[
\frac{\Delta a^2 - \Delta a_{n-1}^2}{\sum_{L=0}^{L=n-2} \Delta a_L^2} \leq \delta \quad \text{and} \quad \frac{\Delta a^2}{\sum_{L=0}^{L=n-2} \Delta a_L^2} \leq \delta,
\]
where \( \delta \leq 10^{-13} \). This ensured no change in the first six significant figures of the result when compared to the case where no coefficients were ignored, but there were great savings in computing.

4.3. Numerical Examples

To verify the accuracy of the truncated series in expression (4.14), mean gravity anomaly covariances were computed both with this formula and by numerical quadratures. All calculations were carried out in double precision (32 bytes), some of those for expression (4.14) being repeated in extended precision (64 bytes). The agreement between both sets of results was better than 6 significant figures, suggesting that both expression (4.14) and Paul’s recursives (used to obtain the \( I_{n,1} \)) are quite stable numerically at all latitudes.

The covariances computed were between a mean anomaly in a fixed block and all the anomalies in the same row of blocks (i.e., all blocks bounded by the same parallels), calculations being done for several rows, at close intervals from equator to pole. Results for only a few of those will be shown here, because they are typical of the rest. Both a \( 5^\circ \) and a \( 1^\circ \) grid (equal angular) were studied. The results were compared to those obtained by numerical quadratures, and by the approximate formula

\[
\text{cov} (\Delta g_{1,1}, \Delta g_{k,1}) \approx \sum_{n=0}^{N_{\text{max}}} b_{n,1} b_{n,k} c_n \Delta \xi_n \Delta \xi P_n (\psi_{11})
\]

(4.15)

where \( Y \) and \( Y' \) are the center points of \( \sigma_{1,1} \) and \( \sigma_{k,1} \), while

\[
b_{n,1} = \frac{1}{1 - \cos \phi} \frac{1}{2n+1} [ P_{n-1} (\cos \phi) - P_{n+1} (\cos \phi) ]
\]

is the \( n \)th degree Pellinen smoothing factor (the formula is Meissner’s), and

\[
\psi_{11} = \cos^{-1} \left[ \frac{\Delta \lambda}{2 \pi} (\cos (\phi_1 + 1) - \cos \phi_1) + 1 \right]
\]

This formula gives the covariances of averages on circular blocks of the same area as that of the equal angular blocks in the row between colatitudes \( \theta_1 \) and \( \theta_{1+1} \). These covariances are used sometimes as approximations to the equal angular covariances between mean values.

The numerical quadratures technique consists in the following: (a) subdividing each block with a grid of \( k \) equally spaced latitude and \( k \) equally spaced longitude lines; (b) computing the covariances between point gravity anomalies at the nodes of each subdivision (there are \( k^2 \) different pairs of nodes to be considered); (c) obtaining the approximate value of the covariance between mean anomalies as
\[ \text{cov}(\Delta g_{ij}, \Delta g_{kl}) = \sum_{n=0}^{k-1} \sum_{m=0}^{k-1} \sum_{r=0}^{k-1} \sum_{s=0}^{k-1} \text{cov}(\Delta g_{ij}, \Delta g_{kl}) \]  

(4.16)

where \( m, n, r, \) and \( s \) are indexes that identify the elements of each pair of nodes in the subdivisions of \( \sigma_{ij} \) and \( \sigma_{kl} \). The covariances between point anomalies such as \( \Delta g_{ij} \) and \( \Delta g_{kl} \) were obtained using a two term model for the degree variances of the anomalies:

\[ c_n^{\Delta\psi} = 18.3906 \frac{(n-1)}{(n+100)} \frac{0.9943667^{n+2} + 658.6132 \cdot (n-1)}{(n+20)(n-2)} \frac{0.9043949^{n+2}}{n} \]

This is the "2L" model of C. Jekeli (1978), and has the advantage that the value of the point covariance function can be computed using finite recursion formulas (Moritz, 1977). The same degree variances were used in (4.14) and (4.15). The number of point covariances being very large \( (k^2) \), a table with entries spaced at \( \Delta \psi = 0.05^\circ \) intervals was created first, the required values being obtained by linear interpolation from this table. Numerical tests showed that \( k = 10 \) was large enough for both \( 5^\circ \) and \( 1^\circ \) blocks, because doubling this number resulted in a change of less than 0.2% in the values given by (4.16). Reducing the interval \( \Delta \psi \) from \( 0.05^\circ \) to \( 0.005^\circ \) had a negligible effect also, therefore the values obtained with (4.16) are probably accurate enough to test those given by (4.14). The only exception was the "polar" row, where the equal angular blocks are, in fact, triangles with a common vertex at the pole. Both with \( 5^\circ \) and \( 1^\circ \) blocks the discrepancies between (4.14) and (4.16) were large (more than 30%), regardless of how large a \( k \), how small a \( \Delta \psi \), or how big a \( N_{\text{max}} \), were chosen. The probable explanation is that the pole is given undue weight in (4.16) because it is treated as a whole row, instead of as a single point. For this reason, the numerical examples presented here stop at the row immediately below the pole.

The first two tables show the covariances between \( 5^\circ \) mean anomalies in the row between latitudes \( 0^\circ \) and \( 5^\circ \) (just above the equator) and in the row between latitudes \( 80^\circ \) and \( 85^\circ \) (one below the pole). The error is at most 3%, though much less in most cases, and \( N_{\text{max}} = 180 \) in each table. Under "Pellinen" one finds the values obtained using (4.15) with due regard for the change in block areas with latitude. While there is very good agreement near the equator, there is no resemblance at all close to the pole to the other values listed.
Table 4.1. Comparison between covariances of $5^\circ$ mean anomalies $\Delta g$ computed with expressions (4.14), (4.16) and (4.15), respectively. $N_{max} = 180$. Row between $0^\circ$ and $5^\circ$.

<table>
<thead>
<tr>
<th>Expression (4.14)</th>
<th>Numerical</th>
<th>Pellinen</th>
<th>Block No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>250.53</td>
<td>253.77</td>
<td>251.78</td>
<td>1</td>
</tr>
<tr>
<td>148.80</td>
<td>149.46</td>
<td>151.03</td>
<td>2</td>
</tr>
<tr>
<td>193.93</td>
<td>93.95</td>
<td>93.90</td>
<td>3</td>
</tr>
<tr>
<td>57.16</td>
<td>57.15</td>
<td>57.10</td>
<td>4</td>
</tr>
<tr>
<td>31.80</td>
<td>31.76</td>
<td>31.67</td>
<td>5</td>
</tr>
<tr>
<td>13.94</td>
<td>13.93</td>
<td>13.86</td>
<td>6</td>
</tr>
<tr>
<td>-18.09</td>
<td>-18.09</td>
<td>-18.08</td>
<td>12</td>
</tr>
<tr>
<td>9.12</td>
<td>9.12</td>
<td>9.11</td>
<td>24</td>
</tr>
</tbody>
</table>

Table 4.2. Comparison between covariances of $5^\circ$ mean anomalies $\Delta g$ computed with expressions (4.14), (4.16) and (4.15), respectively. $N_{max} = 180$. Row between $80^\circ$ and $85^\circ$.

<table>
<thead>
<tr>
<th>Expression (4.14)</th>
<th>Numerical</th>
<th>Pellinen</th>
<th>Block No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>418.60</td>
<td>437.23</td>
<td>835.53</td>
<td>1</td>
</tr>
<tr>
<td>329.03</td>
<td>318.27</td>
<td>800.09</td>
<td>2</td>
</tr>
<tr>
<td>220.91</td>
<td>229.17</td>
<td>709.03</td>
<td>3</td>
</tr>
<tr>
<td>191.75</td>
<td>196.40</td>
<td>592.19</td>
<td>4</td>
</tr>
<tr>
<td>179.79</td>
<td>179.70</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>165.58</td>
<td>168.46</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>120.60</td>
<td>123.73</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>69.89</td>
<td>73.27</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>54.66</td>
<td>58.01</td>
<td>149.40</td>
<td>36</td>
</tr>
</tbody>
</table>

The next three tables show results for $1^\circ$ blocks. Here the numerical method was conducted with the same $\Delta \psi$ and $k$ as in the case of the $5^\circ$ grid. Results for rows between latitudes $0^\circ$ and $1^\circ$, $45^\circ$ and $46^\circ$, and $88^\circ$ and $89^\circ$ are shown. In tables 4.3 and 4.4 the discrepancies between (4.14) and (4.16) stay below 1%; this increases to about 5% near the pole (table 4.5). $N_{max}$ is 300 for the equatorial row, and rises to 400 from $45^\circ$ on. As with $5^\circ$-blocks, the "Pellinen" values are quite close to these of (4.14) and (4.16) near the equator, but become very different near the pole.
Table 4.3. Comparison between covariances of 1° mean gravity anomalies computed with (4.14) and (4.16). $N_{\text{max}} = 300$. Row between 0° and 1°.

<table>
<thead>
<tr>
<th>Expression (4.14)</th>
<th>Numerical</th>
<th>Block No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>849.68</td>
<td>855.16</td>
<td>1</td>
</tr>
<tr>
<td>411.32</td>
<td>410.30</td>
<td>2</td>
</tr>
<tr>
<td>220.35</td>
<td>219.63</td>
<td>3</td>
</tr>
<tr>
<td>181.84</td>
<td>181.50</td>
<td>4</td>
</tr>
<tr>
<td>163.33</td>
<td>163.21</td>
<td>5</td>
</tr>
<tr>
<td>149.11</td>
<td>149.15</td>
<td>6</td>
</tr>
<tr>
<td>1.07</td>
<td>1.08</td>
<td>31</td>
</tr>
<tr>
<td>-16.79</td>
<td>-16.79</td>
<td>61</td>
</tr>
<tr>
<td>1.99</td>
<td>1.99</td>
<td>91</td>
</tr>
<tr>
<td>8.54</td>
<td>8.53</td>
<td>121</td>
</tr>
<tr>
<td>-4.79</td>
<td>-4.79</td>
<td>151</td>
</tr>
<tr>
<td>-14.39</td>
<td>-14.38</td>
<td>181</td>
</tr>
</tbody>
</table>

Table 4.4. Comparison between covariances of 1° mean gravity anomalies computed with (4.14) and (4.16). $N_{\text{max}} = 400$. Row between 45° and 46°.

<table>
<thead>
<tr>
<th>Expression (4.14)</th>
<th>Numerical</th>
<th>Block No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>952.25</td>
<td>959.17</td>
<td>1</td>
</tr>
<tr>
<td>531.15</td>
<td>531.97</td>
<td>2</td>
</tr>
<tr>
<td>275.55</td>
<td>274.65</td>
<td>3</td>
</tr>
<tr>
<td>210.81</td>
<td>210.46</td>
<td>4</td>
</tr>
<tr>
<td>185.36</td>
<td>185.57</td>
<td>5</td>
</tr>
<tr>
<td>170.82</td>
<td>171.08</td>
<td>6</td>
</tr>
<tr>
<td>27.79</td>
<td>27.79</td>
<td>31</td>
</tr>
<tr>
<td>-14.37</td>
<td>-14.37</td>
<td>61</td>
</tr>
<tr>
<td>-18.01</td>
<td>-17.01</td>
<td>91</td>
</tr>
<tr>
<td>-8.27</td>
<td>-8.28</td>
<td>121</td>
</tr>
<tr>
<td>-1.01</td>
<td>-2.92</td>
<td>151</td>
</tr>
<tr>
<td>1.39</td>
<td>1.39</td>
<td>181</td>
</tr>
</tbody>
</table>
Table 4.5. Comparison between covariances of 1° mean gravity anomalies computed with (4.14), (4.16), and (4.15), respectively. \( N_{\text{max}} = 400 \). Row between 88° and 89°.

<table>
<thead>
<tr>
<th>Expression (4.14)</th>
<th>Numerical</th>
<th>Pellinen</th>
<th>Block No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1137.97</td>
<td>1187.94</td>
<td>1795.88</td>
<td>1</td>
</tr>
<tr>
<td>1135.79</td>
<td>1184.15</td>
<td>1795.02</td>
<td>2</td>
</tr>
<tr>
<td>1129.28</td>
<td>1177.99</td>
<td>1792.45</td>
<td>3</td>
</tr>
<tr>
<td>1118.55</td>
<td>1155.08</td>
<td>1788.18</td>
<td>4</td>
</tr>
<tr>
<td>1103.76</td>
<td>1131.34</td>
<td>1782.25</td>
<td>5</td>
</tr>
<tr>
<td>1085.15</td>
<td>1102.85</td>
<td>1774.68</td>
<td>6</td>
</tr>
<tr>
<td>408.59</td>
<td>443.12</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>248.27</td>
<td>257.88</td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>207.19</td>
<td>210.64</td>
<td></td>
<td>91</td>
</tr>
<tr>
<td>189.65</td>
<td>192.55</td>
<td></td>
<td>121</td>
</tr>
<tr>
<td>181.93</td>
<td>184.53</td>
<td></td>
<td>151</td>
</tr>
<tr>
<td>179.65</td>
<td>179.86</td>
<td>368.60</td>
<td>181</td>
</tr>
</tbody>
</table>

Calculations were carried out in the Ohio State University's AMDHAL 470 V/6-II computer, using the FORTRAN H EXTENDED compiler, and double precision. The computing times for obtaining all 37 different covariances between the elements of a row of 5° mean anomalies, and all 181 covariances in a row of 1° mean anomalies, using (4.14) and (4.16) are listed in Table 4.6 for comparison. The integrals of the Legendre functions required by (4.14), and the table of point anomalies covariances needed in (4.16), were precomputed and stored in core memory arrays, so the times given here do not include the determination of those auxiliary values. In most ordinary applications of these formulas, those quantities can be read from disk or tape whenever needed, because they are the same for a whole variety of problems. Clearly, using (4.14) can be orders of magnitude more efficient than using (4.16), while the accuracy is much the same. In fact, accuracy is probably better at the polar rows with (4.14) than with (4.16), because the latter seems to have problems handling triangular blocks. Finally, not all of the time-saving properties of (4.14) were exploited in the computer program used to calculate it, so there is scope for some improvement in efficiency beyond that shown in the table. Notice the time saved in the 88°-89° row thanks to the neglect of terms in (4.14) that become too small near the poles, as explained in paragraph (4.2).
Table 4.6. Comparative efficiencies of algorithms based on expressions (4.14) and (4.16).

<table>
<thead>
<tr>
<th>Block Size</th>
<th>Row</th>
<th>$N_{\text{max}}$</th>
<th>Expression (4.14)</th>
<th>Numerical Quadratures</th>
</tr>
</thead>
<tbody>
<tr>
<td>5°</td>
<td>0° - 5°</td>
<td>180</td>
<td>0.1 sec.</td>
<td>23.3 sec.</td>
</tr>
<tr>
<td>5°</td>
<td>80° - 85°</td>
<td>180</td>
<td>0.1 sec.</td>
<td>23.3 sec.</td>
</tr>
<tr>
<td>1°</td>
<td>0° - 1°</td>
<td>300</td>
<td>0.35 sec.</td>
<td>113.3 sec.</td>
</tr>
<tr>
<td>1°</td>
<td>45° - 46°</td>
<td>400</td>
<td>0.43 sec.</td>
<td>113.3 sec.</td>
</tr>
<tr>
<td>1°</td>
<td>88° - 89°</td>
<td>400</td>
<td>0.06 sec.</td>
<td>113.3 sec.</td>
</tr>
</tbody>
</table>

4.4. Covariances Between Mean Values and Point Values

The prediction by least squares collocation of mean values from point values, or vice versa, requires finding the corresponding "mixed" covariances. In such a case, formula (4.2) becomes

\[
\text{cov}(\bar{u}_{ij}, v(\theta, \lambda)) = \frac{1}{\hat{A}_{ij}} \int_{\sigma_{ij}} \text{cov}(u(P'), v(P)) \, d\sigma
\]  

(4.17)

where $P' = (\theta', \lambda') \in \sigma_{ij}$, $P = (\theta, \lambda)$, $v(P)$ is the "point" value of $v$, and $\bar{u}_{ij}$ is the average over the $i,j$ block of the grid, as before. Following a similar reasoning, one arrives at a formula that corresponds to (4.14), except that only one area integral has to be considered. The new expression is:

\[
\text{cov}(\bar{u}_{ij}, v(\theta, \lambda)) = \frac{1}{\hat{A}_{ij}} \sum_{z=0}^{N_{\text{max}}} \left[ \sum_{k=0}^{K} \sum_{n=-\infty}^{N_{\text{max}}} \frac{c_{n}^{z} \chi_{n}^{z}}{2n+1} \left( I_{n,2Nh+z,1} \bar{P}_{n,2Nh+z}(\cos \theta) \right) \right. 
\]

\[
+ \sum_{k=1}^{K} \sum_{n=-\infty}^{N_{\text{max}}} \frac{c_{n}^{z} \chi_{n}^{z}}{2n+1} \left( I_{n,2Nh-z,1} \bar{P}_{n,2Nh-z}(\cos \theta) \right) \left[ (A(m) \cos m(\lambda_{i}-\lambda)) 
\right.
\]

\[
+ B(m) \sin m(\lambda_{i}-\lambda) \right] \]  

(4.18)

where

\[ A(m) = \begin{cases} \Delta \lambda & \text{if } m = 0 \\ (\cos m\Delta \lambda - 1)/m & \text{otherwise} \end{cases} \]

\[ B(m) = \begin{cases} 0 & \text{if } m = 0 \\ (\sin m\Delta \lambda)/m & \text{otherwise} \end{cases} \]

and the $I_{n,2Nh+1}$, the $c_{n}^{z} \chi_{n}^{z}$, and the $N_{i}, K$, and $N_{\text{max}}$ are as in (4.14). Expression (4.18) assumes that the area means belong to a grid with constant $\Delta \lambda$. If the point values are also on a grid, and if this grid is congruent with that of the mean anomalies, implementation of (4.18) is quite efficient. In fact, the speed of a good algorithm for doing this should be much the same as that of one for implementing (4.14). On the other hand, computing the same covariances by numerical quadratures is $k^{2}$ times faster in this "mixed" case than it was in the previous one, because there is only one area integration involved. Assuming $k = 10$, as in the previous examples, then expression (4.18) should be only 2-3 times faster than numerical quadratures for 5° mean values, and from 4 to 15 times faster in the case of 1° averages.
Both the normalized Legendre function and their integrals are needed for (4.18). They can be precomputed and stored on disk or tape until needed. In this study the following recursive relationships were used to generate their values:

\[
\bar{P}_{n-1,n-1}(\cos \theta) = \left[ \frac{(2n-1)(2n-3)}{(n-m)(n+m-2)} \right]^{\frac{1}{2}} \cos \theta \bar{P}_{n-2,n-1}(\cos \theta)
- \left[ \frac{(2n-1)(n+m-3)(n-m-1)}{(2n-5)(n+m-2)(n-m)} \right]^{\frac{1}{2}} \bar{P}_{n-3,n-1}(\cos \theta), \quad m \neq n \quad (4.19a)
\]

\[
\bar{P}_{n-1,n-1}(\cos \theta) = \left[ \frac{(2n-1)^{\frac{1}{2}}}{(2n-2)^{\frac{1}{2}}} \right] \sin \theta \bar{P}_{n-2,n-3}(\cos \theta) \quad (4.19b)
\]

\[
\int_{\theta_1}^{\theta_2} \bar{P}_{n-1,n-1}(\cos \theta) \sin \theta \, d\theta = -\frac{1}{n} \left[ \frac{(2n-1)(2n-3)}{(n-m)(n+m-2)} \right]^{\frac{1}{2}} \sin^2 \theta \bar{P}_{n-2,n-1}(\cos \theta) \bigg|_{\theta_1}^{\theta_2} +
+ \frac{(n-3)}{n} \left[ \frac{(2n-1)(n+m-3)(n-m-1)}{(2n-5)(n+m-2)(n-m)} \right]^{\frac{1}{2}} \int_{\theta_1}^{\theta_2} \bar{P}_{n-3,n-1}(\cos \theta) \sin \theta \, d\theta \quad (4.20a)
\]

\[
\int_{\theta_1}^{\theta_2} \bar{P}_{n-1,n-1}(\cos \theta) \sin \theta \, d\theta = \frac{1}{2n} \left[ \frac{(2n-1)}{(n-1)(n-2)} \right]^{\frac{1}{2}} \sin^2 \theta \bar{P}_{n-2,n-3}(\cos \theta) \bigg|_{\theta_1}^{\theta_2} +
+ \frac{(n-1)(2n-1)(2n-3)}{2n(n-2)}^{\frac{1}{2}} \int_{\theta_1}^{\theta_2} \bar{P}_{n-3,n-3}(\cos \theta) \sin \theta \, d\theta \quad (4.20b)
\]

if \( \theta \) is not very small, or

\[
\int_{\theta_1}^{\theta_2} \bar{P}_{n-1,n-1}(\cos \theta) \sin \theta \, d\theta = -\left[ \frac{(2n-1)(2n-3) \ldots 3}{(2n-2)(2n-4) \ldots 4} \right]^{\frac{1}{2}} \sin^{n+1} \theta \cdot \left[ \frac{1}{n+1} + \frac{1}{2} \frac{\sin^2 \theta}{n+3} + \frac{1}{2 \cdot 4} \frac{\sin^4 \theta}{n+5} + \frac{1}{2 \cdot 4 \cdot 6} \frac{\sin^6 \theta}{n+7} + \ldots \right] \bigg|_{\theta_1}^{\theta_2} \quad (4.20c)
\]

if \( \theta \) is very small. Here \( y(\theta) \bigg|_{\theta_1}^{\theta_2} = y(\theta_2) - y(\theta_1) \). The recursive formulas for the integrals of the Legendre functions were derived by M.K. Paul (1978); the author has been fortunate enough to have available a FORTRAN subroutine programmed by Paul, and kindly sent by him to Professor R.H. Rapp of the Department of Geodetic Science at O.S.U. The results reported here have been made possible by, and bear witness to, the great numerical stability of Paul's formulas.
5. Conclusions

The relationships between spherical harmonic series and Fourier series, coupled to the symmetries of spherical grids, permit the development of efficient algorithms for numerical analysis of data regularly sampled on the sphere.

The algorithms presented in section 1 for implementing numerical quadratures are efficient enough to allow the analysis of 64800 1° x 1° mean values through degree 180 in less than 20 seconds, and the summation of the 90000 terms of a harmonic series complete through degree 300, on a full 1° x 1° equal angular grid, in less than 1 minute. The analysis of large global data sets to a very high resolution is a relatively trivial operation with modern digital computers.

The principles and ideas behind the optimal estimators of section 2 provide a rational basis for the study of linear techniques for spherical harmonic analysis, both optimal and non-optimal. The error measure introduced in this section is shown to be a very reasonable way of evaluating the estimation error, as illustrated by the results listed in table (3.1).

The optimal estimators themselves are reasonably easy to compute and use, particularly when they are of the quadratures type, which happens under the fairly general conditions discussed in section 2. Even when such conditions are not present, as in the case of scattered data, the problem still has a structure strong enough to allow efficient algorithms for creating and inverting the normal matrix.

The separation of the problem of estimating the coefficients (by least squares collocation or by least squares adjustment) according to the order m of the coefficients, allows both for efficiency and for numerical stability. Even if the total number of unknown coefficients is very large, the largest matrix to be inverted is of dimension N, instead of O(N^2), as it would be if the problem could not be separated in this way.

All the algorithms presented here, when the grid is complete and regular, are well suited to parallel processing.

In the case of full grids of mean values with even noise, the results of section 3 suggest that the optimal "collocation" estimator can be approximated very closely by a much simpler quadratures-type formula, the "composite formula" (3.9). The search for simple, near-optimal estimators is just as important, from a practical point of view, as the search for efficient algorithms for obtaining the optimal estimators themselves. This is a topic that certainly deserves further research.

The methods for creating and inverting the normal matrix, that make possible to find optimal estimators for large data sets, have application outside
spherical harmonic analysis, in all areas of estimation, filtering and prediction on the sphere. This has been the subject of a previous report (Colombo, 1979). It must be added that the principles presented here can be generalized to bodies of revolution other than the sphere; to the case where the data are not homogeneous (i.e., a mixture of, say, gravity anomalies, satellite altimetry, etc.); to the case where the coefficients to be estimated are not those of the signal as given, but of some more or less complex linear transformation of this signal, satellite-satellite tracking data being a good example. In fact, the author is at present considering the error analysis of the determination, by least squares collocation, of the potential coefficients from satellite-satellite tracking, following some of the principles of section 2. This will be the subject of another report.

The method developed in section 4 to calculate the covariance between two area means without employing cumbersome numerical integrations is of interest, not only in spherical harmonic analysis, but more generally in filtering, prediction and estimation from mean values on the sphere.

The computer programs described and listed in Appendix B should help the interested reader to implement some of the techniques discussed in this report. The author sincerely hopes that this will be done by workers concerned with improving and further developing such methods.

Above all, the author hopes that he has conveyed to those who had read this far, the idea that the detailed analysis of very large sets of global, regularly sampled data can be done within the computing resources available today to most scientists who work at universities and research institutions everywhere. The processing, be it by numerical quadratures, or by simultaneous adjustment, of "all the data in the world" is not a fanciful thought, but a practical possibility.
References


Colombo, O. L., Optimal Estimation from Data Regularly Sampled on a Sphere with Applications in Geodesy, Department of Geodetic Science Report No. 291, The Ohio State University, Columbus, Ohio, 1979 (a).


-94-


Moritz, H., Advanced Least Squares Methods, Dept. of Geodetic Science, Report No. 175, The Ohio State University, Columbus, 1972.

Moritz, H., Covariance Functions in Least Squares Collocation, Dept. of Geodetic Science Report No. 240, The Ohio State University, Columbus, 1976.


Rapp, R.H., Comparison of Two Methods for the Combination of Satellite and Gravimetric Data. Dept. of Geodetic Science Report No. 113, The Ohio State University, Columbus, 1968.

Rapp, R.H., Mean Gravity Anomalies and Sea Surface Heights Derived from Geos-3 Altimeter Data. Dept. of Geodetic Science Report No. 268, The Ohio State University, Columbus, 1977.


Appendix A. Term by Term Integration of Formula (4.5), and Rearrangement of Terms in Formula (4.9) to Arrive at Formulas (4.11-a) and (4.14)

The area integrals in expression (4.5) can be split into integrals in $\theta$ and $\lambda$, and the summation theorem for fully normalized spherical harmonics can be used to replace $P_n(\psi_{p'})$ with an equivalent expression in the general term of the series:

$$\text{cov} (\bar{u}_{11}, \bar{v}_{11}) = \frac{1}{\Delta_{\lambda \lambda} \Delta_{\lambda \lambda}} \int_{\theta_{\lambda}}^{\theta_{\lambda} + \Delta \theta} \sin \theta d\theta \int_{\theta_{\lambda}}^{\theta_{\lambda} + \Delta \theta} \sin \theta' d\theta' \int_{\lambda_{\lambda}}^{\lambda_{\lambda} + \Delta \lambda} d\lambda \int_{\lambda_{\lambda}}^{\lambda_{\lambda} + \Delta \lambda} d\lambda' \sum_{n=0}^{\infty} c_n^{u_{p'}} \cdot$$

$$\cdot \sum_{n=0}^{\infty} P_{m}(\cos \theta) P_{m}(\cos \theta') \cos m(\lambda - \lambda')(2n+1)^{-1} \quad (A.1)$$

Starting with (A.1), the proof proceeds in four steps, each justifying in turn the taking of one of the four integrals inside the summation symbols. Each time, three theorems are invoked:

The first is the "M-Test" theorem, due to Weierstrass (see, for instance, Carslaw, 1950):

The series $f(x) = \sum_{n=0}^{\infty} u(x)_n$ will converge uniformly in $a \leq x \leq b$ if there is a convergent series of positive constants $M_0, M_1, ..., M_n, ...$ such that, for all $x$ in $a \leq x \leq b$, $|u(x)_n| \leq M_n$ for every positive integer $n$.

The second theorem is (also according to Carslaw):

If the general term $u(x)_n$ of the series $\sum_{n=0}^{\infty} u(x)_n$ is continuous, and if the series converges uniformly to some function $f(x)$ in the interval $a \leq x \leq b$, then

$$\int_{a}^{b} f(x) \, dx = \int_{a}^{b} \sum_{n=0}^{\infty} u(x)_n \, dx = \sum_{n=0}^{\infty} \int_{a}^{b} u(x)_n \, dx$$

This is a sufficient condition for term by term integration. The third theorem is the mean value theorem for integrals

If $f(x)$ is analytic in $a \leq x \leq b$ then $\int_{a}^{b} f(x) \, dx = (b-a)f(c)$ for some $c$ such that $a \leq c \leq b$.

Proof: The series in the left hand side of (A.1), if all variables but $\lambda'$ are kept fixed, is uniformly convergent in the interval $\lambda_1 \leq \lambda' \leq \lambda_1 + \Delta \lambda$ because

$$(2n+1)^{-1} c_n^{u_{p'}} \sum_{n=0}^{\infty} P_{m}(\cos \theta) P_{m}(\cos \theta') \cos m(\lambda - \lambda') = c_n^{u_{p'}} P_{n}(\psi_{p'})$$

and $|c_n^{u_{p'}} P_{n}(\psi_{p'})| \leq c_n^{u_{p'}}$ because $\max_{0 \leq \psi \leq \pi} |P_{n}(\psi)| = 1$ (the argument of $P_{n}$ is real) for
all \( n \), while \( \sum_{n=0}^{\infty} c_n^{u,v} = \text{cov}(u(P), v(P)) \) is always finite (equals the value of the "point" covariance when \( \psi_{u,v} = 0 \)). The "M-Test" condition is satisfied so the series converges uniformly; therefore, term by term integration with respect to \( \lambda' \) is valid:

\[
\text{cov}(\bar{u}_{1j}, \bar{v}_{k1}) = \frac{1}{\Delta_{1j} \Delta_{k1}} \sum_{n=0}^{\infty} c_n^{u,v} \sum_{n=0}^{\infty} \bar{P}_{n\alpha}(\cos \theta) \bar{P}_{n\alpha}(\cos \theta') \int_{\lambda_1}^{\lambda_1 + \Delta \lambda} \cos m(\lambda' - \lambda) \, d\lambda (2n+1)^{-1}
\]

Applying the mean value theorem to the last expression:

\[
\text{cov}(\bar{u}_{1j}, \bar{v}_{k1}) = \frac{1}{\Delta_{1j} \Delta_{k1}} \sum_{n=0}^{\infty} c_n^{u,v} \sum_{n=0}^{\infty} \Delta \lambda \bar{P}_{n\alpha}(\cos \theta) \bar{P}_{n\alpha}(\cos \theta') \cos m(\lambda' - \lambda_q) (2n+1)^{-1}
\]

where \( \lambda_1 \leq \lambda_q \leq \lambda_1 + \Delta \lambda \). Removing the common factor \( \Delta \lambda \) from the summation:

\[
\text{cov}(\bar{u}_{1j}, \bar{v}_{k1}) = \frac{\Delta \lambda}{\Delta_{1j} \Delta_{k1}} \sum_{n=0}^{\infty} c_n^{u,v} \sum_{n=0}^{\infty} \bar{P}_{n\alpha}(\cos \theta) \bar{P}_{n\alpha}(\cos \theta') \cos m(\lambda' - \lambda_q)(2n+1)^{-1} \tag{A. 2}
\]

The general term in the partially integrated series is

\[
(2n+1)^{-1} c_n^{u,v} \sum_{n=0}^{\infty} \bar{P}_{n\alpha}(\cos \theta) \bar{P}_{n\alpha}(\cos \theta') \cos m(\lambda' - \lambda_q) = c_n^{u,v} \bar{P}_{\alpha}(\psi_{Q})
\]

where \( Q \equiv (\theta', \lambda_q) \); now \( |c_n^{u,v} \bar{P}_{\alpha}(\psi_{Q})| \leq c_n^{u,v} \) for all \( n \), and for all \( \lambda \) in the interval \( \lambda_1 \leq \lambda \leq \lambda_1 + \Delta \lambda \), so the "M-Test" is satisfied again and the series is uniformly convergent, and thus integrable, with respect to \( \lambda \). Therefore

\[
\text{cov}(\bar{u}_{1j}, \bar{v}_{k1}) = \frac{\Delta \lambda}{\Delta_{1j} \Delta_{k1}} \sum_{n=0}^{\infty} c_n^{u,v} \sum_{n=0}^{\infty} \bar{P}_{n\alpha}(\cos \theta) \bar{P}_{n\alpha}(\cos \theta') \cos m(\lambda - \lambda_q)(2n+1)^{-1} \tag{A. 3}
\]

where \( \lambda_1 \leq \lambda_q \leq \lambda_1 + \Delta \lambda \). Once more the general term of the twice integrated series satisfies the "M-Test", because

\[
|(2n+1)^{-1} c_n^{u,v} \sum_{n=0}^{\infty} \bar{P}_{n\alpha}(\cos \theta) \bar{P}_{n\alpha}(\cos \theta') \cos m(\lambda - \lambda_q)| = |c_n^{u,v} \bar{P}_{\alpha}(\psi_{R})| \leq c_n^{u,v}
\]

(where \( R \equiv (\theta, \lambda_q) \)) for all \( n \) and, in particular, for \( \theta' \) in the interval

-98-
\[ \theta_k \leq \theta' \leq \theta_k + \Delta \theta. \] In consequence

\[
\text{cov}(\bar{u}_{ij}, \bar{v}_{k1}) = \frac{\Delta \lambda^2 \sin \theta \Delta \theta}{\Delta_{ij} \Delta_{k1}} \int_{\theta_i}^{\theta_i + \Delta \theta} \sin \theta \, d\theta \sum_{n=0}^{\infty} c_n^{u,v} \cdot \sum_{n=0}^{\infty} \frac{\bar{p}_{nn}(\cos \theta) \bar{p}_{nn}(\cos \theta_5) \cos m(\lambda_8 - \lambda_9)}{(2n+1)^{-1}} \quad (A.4)
\]

Finally,

\[
|\frac{1}{(2n+1)^{-1}} c_n^{u,v} \sum_{n=0}^{\infty} \bar{p}_{nn}(\cos \theta) \bar{p}_{nn}(\cos \theta_5) \cos m(\lambda_8 - \lambda_9)| \leq |c_n^{u,v} \rho_k(\psi_k)| \leq c_n^{u,v}
\]

(where \( \psi \equiv (\theta_5, \lambda_9) \)) for all \( n \) and, in particular, for \( \theta \) in the interval \( \theta_i \leq \theta \leq \theta_i + \Delta \theta \) so the last integral can be put inside of the summations, and the proof of the term by term integration of (4.5) is complete:

\[
\text{cov}(\bar{u}_{ij}, \bar{v}_{k1}) = h \sum_{n=0}^{\infty} c_n^{u,v} \sum_{n=0}^{\infty} \frac{\bar{p}_{nn}(\cos \theta_7) \bar{p}_{nn}(\cos \theta_5) \cos m(\lambda_8 - \lambda_9)}{(2n+1)^{-1}} \quad (A.5)
\]

where \( \theta_i \leq \theta' \leq \theta_i + \Delta \theta \) and \( h = \Delta \lambda^2 \sin \theta \Delta \theta \sin \theta' \Delta \theta / \Delta_{ij} \Delta_{k1} \)

The general term in (A.5) satisfies the "M-Test":

\[
|z_n| = \left| \frac{1}{(2n+1)^{-1}} c_n^{u,v} \sum_{n=0}^{\infty} \bar{p}_{nn}(\cos \theta_7) \bar{p}_{nn}(\cos \theta_5) \cos m(\lambda_8 - \lambda_9) \right| \leq c_n^{u,v}
\]

Since the series \( \sum_{n=0}^{\infty} c_n^{u,v} \) converges, any series of positive terms \( |z_n| \) satisfying \( |z_n| \leq c_n \) must converge also; the "M-Test" condition implies the absolute convergence of \( \sum_{n=0}^{\infty} z_n \). Absolutely convergent series can have the order of their terms changed arbitrarily, without changing the value of their limit sums. This justifies the reordering of (4.9) that leads to expression (4.11-a).
Appendix B: Computer Programs (Descriptions and Listings)

This Appendix contains the description and listing of each of the major subroutines developed in the course of this research, together with their own auxiliary routines. The listings of the main subroutines contain explanatory comments that the author hopes will be of help to those who may use them. The description accompanying each main program defines arguments, gives the dimensions of the memory arrays required, mentions the relevant formulas from preceding sections, and gives a brief explanation of the various segments of the program.

B.1 General Programming Considerations

Only one-dimensional arrays are used in the software described here. The language used for all of them is FORTRAN IV; some subroutines from the International Mathematical and Statistical Libraries Inc. (IMSL) are called by the main subroutines. Implicit DO loops in READ or WRITE statements have been avoided as much as possible, because their execution may be rather slow, depending on the compiler; instead, subroutines FREAD, FWRITE and REWIND are used for all input/output operations involving large files on tape or disk, in some of the subroutines. All operations involving real arithmetic have been coded in double precision (8 bytes, or 32 bits), which is equivalent to retaining the first 7 significant figures in all arithmetic operations.

The arrays containing the associated Legendre functions or their integrals, as the case may be, are arranged first by degree, and then by order: 00, 10, 11, 20, 21, 22, ... (Nmax, Nmax). The Cmn are arranged accordingly, always in two separated arrays: one for the Cmn^2 and another for the Cmn^2. In order to get the value of the element 'mn' from one of these arrays, the following formula is used:

\[ k = \frac{1}{2} n (n + 1) + m + 1 \]

where \( k \) is the position of this element in the one-dimensional array. When the elements are recovered sequentially from the beginning (00), the following type of DO loop is used:

\begin{verbatim}
KOUNT = 1
DO XX   N1 = 1, Nmax
   DO XX   M1 = N1, Nmax
      LEGEND (KOUNT) = ARRAY (KOUNT)**2
   END XX
XX KOUNT = KOUNT + 1
\end{verbatim}

where, in this particular example, the \( mn \) \( (n=N1-1, m=M1-1) \) element in array \texttt{LEGEND} is equated to the square of the \( mn \) element in \texttt{ARRAY}. Avoidance of two-dimensional memory arrays results in considerable improvements in efficiency.
This subroutine computes the sum of a spherical
harmonic series complete to degree and order NMAX at each one of the \(2N^2\)
points or blocks in an equal angular grid. The subroutine can calculate point
values (IFLAG = 0) or area means (IFLAG = 1). The number of rows or parallels \(N\) (Nyquist frequency) must be even. Subroutines FFTP from the IMSL Double
Precision Library is used to calculate the sum of the series along rows by means
of the Mix Radix Fast Fourier Transform algorithm.

The procedure used is that described in paragraphs (1.6) and (1.7). The
symmetry of the grid with respect to the equator, and the corresponding even-
odd symmetry of the values of the Legendre functions or their integrals, (i.e.,
the \(\chi_i^{2n}\) of section 1) are exploited. The values of those functions, or of their
integrals, are read from mass storage (disk or tape) into array ROW, in the
order described in the previous paragraph. All the values for one latitude, or
row, are read at once, so the dimension of ROW is \(\frac{3}{2}(N_{ax} + 1)(N_{ax} + 2)\). All the coe-
cfficients \(C_{ax}^\alpha\) are also stored in core, the corresponding RCNM and RSNM ar-
rays (for \(C_{an}\) and \(S_{an}\) respectively) have the same dimension (the \(S_{ax}\) are
included, though they are all zero). The output consists of \(2N^2\) values in array DATA. This array is organized in rows, from North Pole to South Pole.
The rows, of \(2N\) points or blocks each, have their values written consecutively.
The following is the list of arrays, and their dimensions:

<table>
<thead>
<tr>
<th>NAME</th>
<th>DIMENSION</th>
<th>TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROW</td>
<td>(\frac{3}{2}(N_{ax} + 1)(N_{ax} + 2))</td>
<td>REAL * 8</td>
</tr>
<tr>
<td>RCNM</td>
<td>RD</td>
<td>&quot;</td>
</tr>
<tr>
<td>RSNM</td>
<td>RD</td>
<td>&quot;</td>
</tr>
<tr>
<td>X</td>
<td>RD</td>
<td>INTEGER * 2</td>
</tr>
<tr>
<td>DATA</td>
<td>(2N^2) ((N = 180/\text{BLOCK}))</td>
<td>REAL * 8</td>
</tr>
<tr>
<td>CR1</td>
<td>(N + 1)</td>
<td>&quot;</td>
</tr>
<tr>
<td>CR2</td>
<td>(N + 1)</td>
<td>&quot;</td>
</tr>
<tr>
<td>SR1</td>
<td>(N + 1)</td>
<td>&quot;</td>
</tr>
<tr>
<td>SR2</td>
<td>(N + 1)</td>
<td>&quot;</td>
</tr>
<tr>
<td>AM</td>
<td>(N + 1)</td>
<td>&quot;</td>
</tr>
<tr>
<td>BM</td>
<td>(N + 1)</td>
<td>&quot;</td>
</tr>
<tr>
<td>F AUX1</td>
<td>(4N)</td>
<td>&quot;</td>
</tr>
<tr>
<td>F AUX2</td>
<td>(4N)</td>
<td>&quot;</td>
</tr>
<tr>
<td>F IWK</td>
<td>see IMSL Handbook</td>
<td>INTEGER * 4</td>
</tr>
<tr>
<td>F LL</td>
<td>&quot; &quot; &quot;</td>
<td>LOGICAL * 4</td>
</tr>
<tr>
<td>F A</td>
<td>&quot; &quot; &quot;</td>
<td>REAL * 8</td>
</tr>
<tr>
<td>IV</td>
<td>(N + 1)</td>
<td>&quot;</td>
</tr>
</tbody>
</table>

("F" designates those arrays required by the IMSL subroutine FFTP). In ad-
dition, the size (in degrees) of the blocks is defined by BLOCK; \(NPD = \frac{3}{2}(N_{ax} + 1)
(N_{ax} + 2)\); IU is the number of the unit (disk, tape) from which the Legendre
functions or their integrals are to be read.

Array X contains information on whether a given \(\chi_i^{2n}\) is even or odd;

-101-
SUBROUTINE SSYNTH(BLOCK, DATA, NMAX, BGNM, BSNM, AN, DM, X, ROW, CR1, SR1,
2 CR2, SR2, AUX1, AUX2, IVK, IV, IU, NPP, A, IFLAG, LL)
C
C THIS SUBROUTINE CREATES CENTER VALUES (IFLAG = 0) OR
C AREA MEANS (IFLAG = 1) ON A
C (DEGREES X BLOCK (DEGREES)) EQUAL AREA GRID, BY
C ADDING UP A SPHERICAL HARMONIC EXPANSION (WITH COEFFS.
C CIN, SMN IN ARRAYS BGNM AND BSNM, RESPECTIVELY)
C TO DEGREE AND ORDER MMAX (LIMITED ONLY BY STORAGE CAPACITY
C AND AVAILABLE COMPUTER TIME). THIS SUBROUTINE USES
C A FILE IN UNIT IU CONTAINING THE LEGENDRE FUNCTIONS
C OR THEIR INTEGRALS, AS THE CASE MAY BE, FOR ALL DISCRETE
C LATITUDE BANDS (*RADI*), FROM N. POLE TO THE EQUATOR.
C IT USES THE FAST FOURIER SUBROUTINE (NON BINARY NUMBER
C OF DATA POINTS, ON *RARRI*) *FFT* FROM THE
C IMSL LIBRARY. THE VALUES OF THE HARMONIC SUM ARE STORED
C IN THE NC X NC/2 ARRAY *DATA*, TO FORM AN EQUAL
C ANGULAR ARRAY OF SYNTHETIC DATA. TWO LATITUDE "ROWS" ARE
C CREATED TOGETHER, TAKING ADVANTAGE OF THE FACT THAT ALL
C LEGENDRE FUNCTIONS AND THEIR INTEGRALS ARE EITHER EVEN
C OR ODD WITH RESPECT TO THE EQUATOR. THE SIZE OF THE
C BLOCK IS SUPPOSED TO BE AN EXACT DIVIDER OF 90 DEGREES,
C AND EQUAL TO THE BLOCK = NC/2
C
C PROGRAMMED BY OSCAR L. COLOMBO. GEODETIC SCIENCE,
C OHIO STATE UNIVERSITY, COLUMBUS, FEBRUARY 1990
.

IMPLICIT REAL*4(A-H,O-Z)
INTEGER*4 X
LOGICAL*4 LL
DIMENSION DATA(1, BGNM, 1), BSNM(1), AN(1), DM(1), X(1), ROW(NPP),
2 CR1(1), SR1(1), CR2(1), SR2(1), AUX1(1), AUX2(1), IVK(1), IU(1),
3 LL(1), IU(1)
REWIND IU
PI = 3.141592653589793D0
DIGCONV = PI/180. D0
BLK = BLOCK*DIGCONV
BLK2 = BLK*90. D0
NPI = NMAX*4
NC = 360. D0/BLOCK+6. D0000001D0
NC2 = NC/2
NC2P = NC2+1
NBNDM = NC/4
NGP = NC+1
NDC = 1
DO = 2
M1 = 1, NPI
M1 = 1, NPI
XI(NG3) = 1
IF(CR1(NI-MI)/2. NE. (NI-MI)) X(NDC) = -1
2 NDC = NDC+1
IF( (IFLAG .NE. 0) ) GO TO 5
DO = 4
MI = 2, NPI
K = MI-1
ARK(M1) = BSNM(NI+BLK)/M
4 ARK(M1) = -(1, DO=DO0000001+BLK)) / M
ARK(1) = BLK

102
FORTRAN IV GI RELEASE 2.0  SSYNTH  DATE * 00242  12/29/34  PAGE 002

0029  IBM(1) = 0.0
0030  GO TO 8
0031  5 DO 6  MI = 1, NP1
0032  N = MI-1
0033  BL2M = BL2*N
0034  AM(MI) = DCOS(BL2M)
0035  6 BM(MI) = -DSIN(BL2M)
0036  0 NOT = NC+NC2
0037  DO 9  I = 1, NP1
0038  9 IV(I) = 1*(1-1)/2
0039  ISH = -NC
0040  ISHF = NOT
0041  PHI1 = 90.0D0*DSINV
0042  PHI2 = PHI1-BLK

C MAIN LOOP, WHERE *DATA* IS CREATED, TWO EQUATORIALLY
C SYMMETRICAL ROWS AT A TIMES.
C (OUTER LOOP)

0043  DO 40  NR = 1, NROW2
0044  40 PROWIN = 0.5D0
0045  IF(FLAG, NR, 1) GO TO 12

C computing the inverse of the area of a block in
C row nr.

C PROWIN = 0.5D0/*DSIN(PHI1)-DSIN(PHI2)*BLK
0046  PHI1 = PHI1-BLK
0047  PHI2 = PHI1-BLK
0048  IF(ISH, ISH, 1)
0049  12 ISH = ISH+NC
0050  ISHF = ISHF-NC

C FORM THE COS AND SIN FOURIER COEFFICIENTS OF EACH "ROW"
C UP TO FREQL. N = NMAX, IN ARRAYS CH1, SRL AND
C CH2, SRL, RESPECTIVELY.
C (INNER LOOP)

C READ THE (INTEGRALS OF) NORMALIZED LEGENDRE FUNCTIONS
C CORRESPONDING TO THE NORTHERN "ROW", FROM UNIT *10*
C (UNFORMATTED RECORDS).

0051  READ(10) ROW
0052  DO 20  M1 = 1, NP1
0053  20 CI = 0.0
0054  C2 = 0.0
0055  NI = 0.0
0056  S2 = 0.0

C THIS IS THE BEGINNING OF THE INNERMOST DO LOOP.

0057  DO 15  M1 = 1, NP1
0058  NUC = IV(M1)+MI
0059  R = BOS(RUC)
0060  S = R+RSN(MUC)
0061  C = R+RSN(MUC)
0062  IF(NUC, NUC, 1) GO TO 14
0063  14 CI = CI+C
C2 = C2+C
S1 = S1+S
S2 = S2+S
GO TO 15
C = C1+C
C2 = C2-C
S1 = S1+S
S2 = S2+S
CONTINUE
END OF INNERMOST LOOP.

CR1(M1) = C1*AM(M1)-S1*BM(M1)
CR2(M1) = C2*AM(M1)-S2*BM(M1)
SR1(M1) = S1*AM(M1)+C1*BM(M1)
SR2(M1) = S2*AM(M1)+C2*BM(M1)

20 CONTINUE
CR1(I) = 2.00*CR1(I)
CR2(I) = CR2(I)*2.00

DO 25 M1 = 1,NC2P
NA = M1
MB = -M1+2
DO 25 K = 1,MP1
NA = NA+NC
MB = MB+NC
IF(MA,CT,NP1) GO TO 25
IF(MA,CT,NP1) GO TO 21
CR1(M1) = CR1(M1)+CR1(MA)+CR2(MB)
CR2(M1) = CR2(M1)+CR2(MA)+CR2(MB)
SR1(M1) = SR1(M1)+SR1(MA)+SR2(MB)
SR2(M1) = SR2(M1)+SR2(MA)+SR2(MB)
GO TO 25
CR1(M1) = CR1(M1)+CR1(MB)
CR2(M1) = CR2(M1)+CR2(MB)
SR1(M1) = SR1(M1)+SR1(MB)
SR2(M1) = SR2(M1)+SR2(MB)

25 CONTINUE
FORM THE COMPLEX FOURIER COEFFICIENTS.

K1 = -1
K2 = 0
DO 30 M1 = 1,NC
K1 = K1+2
K2 = K2+2
IF(M1,CT,NC2P) GO TO 26
AUX1(K1) = CR1(M1)
AUX2(K1) = CR2(M1)
AUX1(K2) = -SR1(M1)
AUX2(K2) = -SR2(M1)
GO TO 30

26 NX = NC*M+1
AUX1(K1) = CR1(MK)
AUX2(K1) = CR2(MK)
AUX1(K2) = SR1(M5)
AUX2(K2) = SR2(M5)

30 CONTINUE

C OBTAIN THE CONTENTS OF EACH ROW IN "DATA" AS FOURIER
TRANSFORMS OF AUX1 AND AUX2.

CALL FFT5(AUX1, NC, 1WK, A, LL)
CALL FFT5(AUX2, NC, 1WK, A, LL)

K1 = -1

C END OF OUTER AND INNER LOOPS.

DO 40 I = 1, NC
K1 = K1 + 2
DATA(ISH+1) = AUX1(K1)*FR0VIN
DATA(ISH+1) = AUX2(K1)*FR0VIN

40 CONTINUE

REWIND IU
RETURN
END
arrays CR1, SR1, and CR2, SR2, respectively, contain the Fourier coefficients $\alpha_1^i (a^i_1)$ and $\beta_1^i (b^i_1)$ of rows $i$ and $N-1-1$; arrays AM and BM contain the values of $A(m)$ and $B(m)$, as defined in expression (1.7); the auxiliary array IV contains the numbers $\frac{1}{2} n(n+1)$ needed to locate individual elements within arrays ROW, RCNM, and RSNM, when they are not addressed sequentially.

The comments in the listing are probably enough to understand most of it on close inspection; one point however may be worth explaining further: the "aliasing" of the Fourier coefficients has been incorporated to take care of the case when $N_{max} > N - 1$. In such situation the $\alpha_1^i (a^i_1)$ and $\beta_1^i (b^i_1)$ become aliased, as Fourier coefficients must, and it is their aliased values that the FFT subroutine requires to compute the values of the spherical harmonic expansion along parallels. The formulas for the aliased coefficients are

$$\alpha_1^i = \alpha_1^i + \sum_{k=1}^{KM} (\alpha_{1k} - \alpha_1^i)$$
$$\beta_1^i = \beta_1^i + \sum_{k=1}^{KM} (\beta_{1k} - \beta_1^i)$$

where $KM$ is a large enough integer. A similar expression applies to $a_1^i$ and to $b_1^i$.

The arrangement of the output in latitude corresponds, in the case of area means, to the intervals on which the $P_z$ are integrated; for point values, it is defined by the latitudes $\theta_i$ at which the $P_z (\cos \theta_i)$ have been precomputed. As regards longitude, the grid starts from the zero meridian used for defining the coefficients. In the case of point values it is usual to compute all values at the center of each block. To do this, the $P_z$ must be precomputed at the latitudes of the center points, while the longitudes are taken care by modifying the coefficients as follows

$$C_{1z} = C_z \cos m \frac{\Delta \lambda}{2} + S_z \sin m \frac{\Delta \lambda}{2}$$
$$S_{1z} = S_z \cos m \frac{\Delta \lambda}{2} - C_z \sin m \frac{\Delta \lambda}{2}$$

This is equivalent to rotating the grid eastwards from the zero meridian by $\Delta \lambda / 2$.

**B.3 Subroutine HARMIN**

This subroutine implements either the algorithm of paragraph (1.5) for the harmonic analysis of area means, or that of paragraph (1.7) for the analysis of point values.

The subroutine calls IMSL's FFCSIN to calculate the $a_1^i, b_1^i$ or the $\alpha_1^i, \beta_1^i$ by means of the Fast Fourier Transform (Mix Radix) algorithm. It also calls subroutine QUADFS, that returns in array $A$ the de-smoothing
SUBROUTINE HARMX BLOCK, DATA, NN, NCM, BSNM, AM, DM, X, ROW, CR1,
2 SR1, C62, SR2, AUX1, AUX2, IWK, IV, IU, NDD, A, IFLAG)

C
C THIS SUBROUTINE USES THE INTEGRALS OF LEGENDRE FUNCTIONS
C STORED IN UNIT "IU" (CREATED USING PAUL'S RECURSIVES).
C OR, ALTERNATIVELY, POINT VALUES OF LEGENDRE FUNCTIONS, TO
C ANALYZE REAL VALUES ON A COMPLETE, EQUAL ANGULAR GRID
C OR POINT DATA (IFLAG = 0) ON THE SAME TYPE OF GRID.
C THE INPUT, OR DATA, OCCUPIES ARRAY "DATA", WHILE OUTPUT
C APPEARS IN ARRAYS "RCNI" AND "ISNI". MAXIMUM DEGREE AND
C ORDER CONSIDERED IS NN.
C THE NUMBER OF BLOCKS PER ROW (NC) MUST BE EVEN.
C IT USES THE FAST FOURIER SUBROUTING
C "FCSIN" FROM THE UBLD (DOUBLE) LIBRARY.
C USES SUBROUTINE "QUADS" PROVIDES THE "WEIGHTS" FOR THE
C QUADRATURES' FORMULA OF INTEREST.
C FOR ADDITIONAL INFORMATION ON THE VARIOUS ARGUMENTS, SEE
C SUBROUTINE "SSYNTH".
C PROGRAMMED BY OSCAR L. COLONAO, GEOMETIC SCIENCE,
C OHIO STATE UNIVERSITY, COLUMBUS, FEBRUARY 1980.

C
C IMPLICIT REAL*8(A-H,O-Z)
C INTEGER*2 K
C DIMENSION BSNM(1), AM(1), DM(1), NCM(1), CR1(1), SR2(1),
2 AUX1(1), AUX2(1), IWK(1), IU(1), NDD(1), A(1)

KREWING IU

F1 = 3.1415926535897938D0
FT = 1.0D0/(0.30+F1)
HRONG = F1/100.0D0
BLK = BLOCK+HRONG
IL2 = BLKX4.500D0
IVY (IFLAG, NIC, 1) FT = FT x BLK**2
NN1 = NN+1
NC = 360.0D0/BLK
NCC = NC/2
NHS = NC2
Nhws = NHS/2
NUG = 1
DO 2 MI = 1, NN1
2 DO 4 MI = 1, NN1
4 XNUC = 1
1 IF ((N1-M1)/2)**2 x (M1+1) x NUCC = -1
DO 4 IFY (IFLAG, NIC, 1) GO TO 4
4 DO 5 MI = 2, NN1
5 N = MI-1
6 AM (MI) = D3IN(MMHIK+M)
7 BM (MI) = -(1.0D0-BCOS(N+BLK))/M
8 AM (1) = BLK
9 NN1 = 0.30
GO TO 7
10 DO 6 MI = 1, NN1
11 N = MI-1
12 BM2 = BLK+M
13 AM (MI) = BCOS(BM2)
FORTRAN IV G1 RELEASE 2.0

0026    6 BM(1) = -DNSN(DL2M)
0037    7 CALL QUADFP(A, NM1, BLOCK, NC)
0038         DO R = 1, NM1
0039             RM(R) = 0.10
0040         DO 5 RSRH(R) = 0.10
0041         DO 9 R = 1, NM1
0042             IF(FLAG.EQ.1) A(I) = A(I)*FT
0043         9 IV(I) = 2*IC(1-1)/I"

C C C C C
0011   RCNM(NUC) = RCNM(NUC) + R*CP
0012   IRSN(NUC) = IRSN(NUC) + R*SP
0013   GO TO 20
0014  10   RCNM(NUC) = RCNM(NUC) + R*CM
0015   IRSN(NUC) = IRSN(NUC) + R*SM
0016   20   CONTINUE
               END OF OUTER AND INNER LOOPS.
007    C     IF(1FLAG.EQ.0) GO TO 999
008    C     IN THE CASE OF AREA MEANS, MULTIPLY COEFFICIENTS BY
009    C     DE-SMOOTHING FACTORS (IN ARRAY A).
0010   MAC = 1
0011   DO 25  N1 = 1, NM1
0012      DO 25  N1 = 1, NM1
0013  25   RCNM(NAC) = RCNM(NAC) * A(N1)
0014   RCNM(NAC) = RCNM(NAC) * A(N1)
0093   25   RCNM(NAC) = RCNM(NAC) * A(N1)
0094   CONTINUE
0095   999   CONTINUE
0096   RCMBR IU
0096   RETURN
0097   END
SUBROUTINE QUADFS(A, NN1, BLOCK, NC)
C
C    THIS SUBROUTINE COMPUTES THE VECTOR OF INTEGRATION WEIGHTS.
C
IMPLICIT REAL*8(A-H, O-Z)
DIMENSION A(N), POLS(N)
COMMON /AN/ ANM1(N+1), A2NM1(N)
NN = NN1-1
PI = 3.1415926535897932
DCMOV = PI/100.0D0
NC2 = NC/2
NC3 = NC2/3
BLK = BLOCK*DCMOV
BLCK = DSQRT(DSIN(BLK)*BLK/PI)/DCMOV
F = 1.00/((1.00-DCOS(BLCK*DCMOV))
DO 5 N = 2, NN
ANM(N) = -(N-1)*1.00/N
5 CALL LEGP00P(POLS, BLCK, NN1)
A(1) = 0.00
A(2) = 0.00
DO 10 N = 2, NN
A(N+1) = 1.00*(F/(2*N+1)*(POLS(N-1)-POLS(N+1))
10 IF(N.EQ.NC1) A(N+1) = A(N+1)**2
IF(N.EQ.NC2) A(N+1) = 1.00
10 CONTINUE
WRITE (6, 170) (1, A(I), I=1, NN1)
170 FORMAT (IX, 4(2X, I3, 2X, G20.10))
RETURN
END
SUBROUTINE LEGPOL(POLS, PSI, NMAX)
IMPLICIT REAL*8(A-H, O-Z)
COMMON /AR1/ AR1(101), A2N1(301)
DIMENSION POLS(1)
PI = 3.14159265358979326
DMCONV = PI / 180.

10 T = DCOS(PI*DMCONV)
00001 PM1 = 1.0
00009 PN = T
00100 POLS(1) = T
00110 DO 20 N = 2, NMAX
00120 POLS(N) = AR1(N) + PM1 + A2N1(N) * PM1 + T
00130 PM1 = PR
00200 PN = POLS(N)
00210 20 CONTINUE
0016 RETURN
END

000242 12/29/94
0026400
0026600
0026700
0026800
0026900
0027000
0027100
0027200
0027300
0027400
0027500
0027600
factors \( \mu_n \). QUADFS calls LEGPOL, a subroutine that computes the Legendre polynomials up to degree \( NN + 1 \) needed for the \( \beta_n^1 \) in \( \mu_n \).

The data is arranged as in SSYNTH, in array DATA, before the subroutine is called. Afterwards, the contents of DATA are destroyed, as the \( a_1^r \), \( b_1^r \) or the \( \alpha_1^r \), \( \beta_1^r \) are formed in place of them, row by row, by FFC3SIN. The resulting coefficients' estimates are put into arrays RCNM and RSNM, in the same order as for HARMIN. The other arrays, with the exception of A, are as in SSYNTH. This is the true of the scalars, with the exception of \( NN \). NN is the highest degree and order to be estimated. NDD is the total number of Legendre functions, or their integrals, to be read from unit IU, per \( \theta_1 \). This number is \((NN+1)+(NN+2)/2\). A is a REAL*8 array of dimension \( NN + 1 \). The dimension in QUADFS allows for a maximum \( NN = 300 \); for larger solutions, the dimensions there and in LEGPOL must be increased accordingly.

In the case of point data, the estimated coefficients are computed using a center point formula that assumes that the data are situated at the centers of the blocks; the resulting coefficients are referred, nonetheless, to a grid starting at the zero meridian (the "rotation" of the coefficients takes place between statements 0071 and 0073, when IFLAG = 0). When IFLAG = 1, the area means formula (1.30) is computed; the \( \chi_n^a = \mu_n \int_0^{\pi/2} \phi \sin \theta \, d\theta \), and the \( \mu_n \) are those produced by QUADFS, as already mentioned. The integrals of the Legendre functions are read from unit "IU", as in SSYNTH (same format), and the size of the blocks is specified by BLOCK (in degrees). The version of QUADFS listed here implements the "composite" estimator of paragraph (3.3). If another is desired, this can be achieved simply by replacing lines 0021 through 0024 in QUADFS.

B.4 Subroutine NORMAL

This subroutine creates the optimal estimators for the \( \bar{C}_{zz}^\alpha \) based on the formation and inversion of the \( R(m) \) matrices described in paragraph (2.10). The algorithm exploits the fact that \( (C_{zz} + D) \) is a block matrix of Toeplitz circulant sub-matrices. This subroutine is meant only for mean values.

The grid is as in SSYNTH and HARMIN. The symmetry with respect to the equator is only partly exploited: matrix \( D \) may not be persymmetric, so the total matrix \((C_{zz} + D)\) may not be so either. \( C_{zz} \) however, is always persymmetric, and this is taken into account to save computing and storage. A general diagonal matrix \( D \) corresponds to a rather broad class of actual problems, such as the analysis of the \( 5^\circ \times 5^\circ \) real gravity anomalies described in paragraph (3.4).

This subroutine requires four input/output units: 8 (read only) contains the values of the integrals of the Legendre functions, row by row, arranged as in SSYNTH or HARMIN; 10 contains the right hand sides of the "reduced normals" \( \chi_n^a = R(m) \chi_n^\alpha \) (expression (2.58)); 15 contains the \( R(m) \) matrices, ordered by
SUBROUTINE NORMAL(MAX, MM, DGRID, REGUL, IGEO, NNUN, ROWP, RNPO, RHS, S, A, 00000100
2, UL, R2, W, DWAN, FC, PN, SS)

THIS SUBROUTINE IMPLEMENTS THE FAST COLLOCATION
ALGORITHM. IT CREATES THE OPTIMAL FILTER FOR ESTIMATING
SPHERICAL HARMONIC COEFFICIENTS OF AREA MEANS (EQUAL
ANGULAR) ON A SPHERE. A CHOICE CAN BE MADE OF EITHER
GEODETIC OR GEODETIC COORDINATES.

THE GRID MUST BE BOTH COMPLETE AND SYMMETRICAL RESPECT
TO THE EQUATOR
BECAUSE OF THE GRID'S SYMMETRY, THE "SIGNAL" PART OF EACH
RCM MATRIX IS POSITIVE, AND FULL ADVANTAGE IS TAKEN
OF THIS FACT WHEN SETTING UP THE RCM'S. THE "NOISE" PART,
ON THE OTHER HAND, CAN ABE ANY DIAGONAL MATRIX (ASYMMETRIC).

INTEGRALS RPM ARE READ IN FROM UNIT 8, STORED, AFTER
MULTIPLYING THEM BY CERTAIN FACTORS, IN UNIT 10.
THE RCM'S MATRICES ARE STORED IN UNIT 15.
OPTIMAL "QUADRATURES" WEIGHTS" IN UNIT 30.

THIS PROGRAM USES SUBROUTINES "LUDENT" AND "LIPMLP" FROM
THE ISSL DOUBLE LIBRARY. IN TURN, THESE USE OTHER
SUBROUTINES FROM THE ISSL SINGLE LIBRARY. THIS PROGRAM
ALSO CALLS SUBROUTINE "FUR" TO COMPUTE THE FOURIER
COEFFICIENTS OF THE BLOCK COVARIANCES, AND SUBS. "FREAD",
"FWRITE", AND "FREND" FOR FAST INPUT-OUTPUT. THE USER
CAN REPLACE THESE WITH HIS OWN, OR WITH THE STANDARD "READ",
"WRITE" AND "REWIND" COMMANDS. RNPO, ROWP, ETC ARE THE
LENGTHS#1 (I.E., IN BYTES) OF THE ARRAYS TO BE READ
ON WRITER.
IF THE RCM MATRIX OF A PARTICULAR ORDER CANNOT BE
INVERTED, NULL QUADRATURES "WEIGHTS" FOR THAT ORDER ARE
STORED IN UNIT 30.

PROGRAMMED BY OSCAR L. COLONNO, GEOD. SC., OHIO STATE U., SEPT. 1979.

IMPLICIT REAL(A-H, O-Z)
00001500
DIMENSION ROWP(1, ROWP), RHS(1), R2(1), A(1), UL(1)
00001600
DIMENSION ARLIT(400), S1(1), X(200), Y(200), Z(200)
00001700
DIMENSION ROWP(400), S2(1), T(200), U(200), FC(1), PN(1), W(1), PH(200), DWAN(1), FIMP
3 (400)
00001800
COMMON/HR, HT(310)
00001900
BASIC CONSTANTS
00002000
P1 = 3.1415926535897930
00002100
DIAG = P1/100.0
00002200
F = 1.80/246.25700
00002300
PROBLEM DEFINITION
00002400
FORTRAN IV G1 RELEASE 2.0

NORMAL

DATE = 00242 12/29/34

PAGE 0002

DCRD = BLOCK SIZE IN DEGREES. MINIMUM: 1 DEGREE.

IGEO = 2 IF GEODETIC LATITUDES ARE TO BE CONVERTEO0000300

TO GEOCENTRIC.

MAX. DEG. AND OBLON TO BE ANALYZED (LESS THAN

NC NO. OF "COLUMNS", ALWAYS EVEN.

MR = NO. OF ROWS. THERE IS NO POLAR Rows.

RMAX MAX. DEGREE IN TRUNCATED EXPANSION OF BLOCK CO-

VARIANCE FUNCTIONS. NOT GREATER THAN NC.

MRUN = 0 IF RCMD'S ARE CREATED AND STORED IN

UNIT 15.

= 999 IF RCMD'S ARE READ FROM UNIT 15 INSTEADO0000400

OF BEING CREATED.

REGUL REGULARIZATION PARAMETER. ADDED TO ALL DIAGONALO00004280

TERM OF THE RCM MATRICES.

0009

DCRD = DCRD1+DCRVY

0010

NC = 360.00/DCRD+0.0000100

0011

NC4 = NC4+0.25+0.0000100

0012

NC2 = NC2+2

0013

NR = NC2

0014

N1 = NR+1

0015

NF = 1

0016

NP = 1+MAX+1

0017

NPP = (NP1*(NP1+1))/2

0018

NLL = (NL1*(NL1+1))/2

0019

NTR = NR1

0020

NLD = NR1

0021

NC2 = NC2+1

0022

NC2 = NC2+1

0023

NPH = NPH+2

0024

DVAY0 = 0.10

0025

DO 5 J = 1, NC

0026

PH(J) = 0.10

0027

5 X(J) = 0.10

0028

FROM ARRAY OF GEOCENTRIC OR GEODETIC GRID'S GEOCENTRIC LATSO00006100

0029

EX = F/(2.10-F)

0030

ET = 2.10*EX/(1.00*EX**2)

0031

NC2 = NC2+1

0032

AN = 90.00

0033

DO 6 J = 1, NC2

0034

PH(J) = ANG+RHCONV

0035

IF(NC2.EQ.1) PH(J) = PH(1)-EM*DSIN(2.00*PH(1))

0036

AN = ANG-NCRH1

0037

DO 15 J = 1, 2, EP1

0038

1 = J-1

0039

FIMG(J) = DSIGN(DVAY(J)-(2.10*J+1))

0040

CONTINUE

0041

FIMG(J) = DSIGN(DVAY)

0042

IRF = -NN1

00001000

00001000
FORTAN IV C1 RELEASE 2.0  
NORMAL  DATE = 00242  
12/29/34  PAGE 0003

0043   DO 25  NT = 1, NC2
0044   INS = INS+1
0045   FOWN = 1.00/((DSIN(PHI(NT-1)+DSIN(PHI(NT+1)))*DCR)
0046   CALL FREAD ROW(1), B, MPPB, 0999, 8999)
0047   NIS = 1
0048   DO 20 I = 1, NP1
0049   FX = F1MIP(I) * FOWN
0050   DO 20 IL = 1, I-1
0051   HOWN(I) = HOWN(I) * FX
0052   20 NIS = NIS + 1
0053   CALL WRITE(ROWP, 10, MPPB, 0999, 8999)
0054   25 CONTINUE
0055   NB = NB+1
0056   NBI = NBI+1
0057   WRITE(6, 555) DCRI1, NMAX, NN, NR, NNUN
0058   555 FORMAT(,' BLOCK SIZE ', FT = 3, MAX. DEGREE IN COVARIANCE ', 15, ,'
2. MAX. DEGREE TO BE ANALYZED ', 15, , NUMBER OF ROWS ', 15, , NNUN ,
3 ')
0059   WRITE(6, 556) NC
0060   556 FORMAT(,' NUMBER OF COLUMNS ', 15, ,)
0061   IF (ACOB.EQ.1) WRITE (6, 557)
0062   IF (ACOB.EQ.2) WRITE (6, 558)
0063   557 FORMAT(,' DATA ON A SPHERE ', A)'
0064   558 FORMAT(,' DATA ON AN ELLIPSOID ', A)
0065   DO 99 MI = 1, NP1
0066   99 SNMP(M) = ((M-1)*M)/2
0067   DO 100 MI = 2, NP1
0068   100 FF(M) = 2.00*1.00-DCOS((M-1)*DCR)/(M-1)**2
0069   C READ THE R.H.S.'S OF THE REDUCED 'NORMALS' FROM UNIT 10
0070   C INTO VECTOR BNS .
0071   CALL RWHI(10)
0072   CALL RWHI(10)
0073   ISH = -NIL
0074   DO 210 I = 1, NC4
0075   CALL FREAD ROW(1), 10, NLLA, 8999, 5999)
0076   ISH = ISH+1
0077   DO 210 J = 1, NIL
0078   210 BNS(I) = ROWP(J)
0079   IF (SNAN. OR. 999) GO TO 1500
0080   C MAIN LOOP FOR THE CREATION OF THE RCM'S MATRICES.
0081   CALL RWHI(10)
0082   1000 CONTINUE
0083   DO 101 I = 1, ND
0084   101 S(I) = 0.00
0085   C POSITION THE INMP'S TAPE TO CURRENT ROW NF .
0086   C DO 102 NF = 1, NP
0087   102 CALL FREADROW(1), 10, MPPB, 0999, 8999)
0088   IS = (NP-2)*NN1
0089   MCV = NG2-NP1
0090   DO 120 SQ = NP, 0
0091   IF (SQ.NE.RP) CALL FREADROW(1), 10, MPPB, 0999, 8999)
0092   120 IF (SQ.NE.RP) GO TO 105
0093   0093
BEGIN MAIN LOOP WHERE ERRORS ARE ESTIMATED.

DO 500 M = 2,1,1

CALL REW1(W(4))

IFS = F(I,M)*NC

N = M-1

IF(M.EQ.0) GO TO 214

DO 213 I = 1,M

213 CALL FREACS(1,15,NC20,0999,0999)

READ IN MATRIX RCM FROM UNIT 15, COLUMN BY COLUMN.

AND PUT IT IN SYMMETRIC STORAGE FORM IN VECTOR A.

214 DO 220 J = 1,NC4

NC211 = NC2-J+1

IF(J.EQ.1) GO TO 222

DO 215 I = 1,NC4

215 CALL FREACS(1,15,NC20,0999,0999)

222 CALL FREACS(1,15,NC20,0999,0999)

DO 220 I = J,NC2

220 CONTINUE

FORM THE CHOLESKY FACTOR UL.

CALL LUBDCV(A,UL,NC2,NI,12,15)

WRITE(6,225) M

221 FORMAT(//'### MATRIX R(N,N) CORRESPONDING TO ORDER ',I5,' IS"

SINGULAR ###'/)

225 WRITE(2,220) M

220 IF(IER.EQ.125) GO TO 296

SOLVE THE REDUCED EQUATIONS RCM X(M,N) = R(M,N) FOR ALL

N AND ALL M; AND SAVE THE OPTIMAL "QUADRATURE WEIGHTS",

OR SOLUTIONS, IN UNIT 10.

FIRST FORM R(M,N)

DO 240 M = 1,M

SMN = 1.00

IF((M-N)/2)*2.0 .LT. (M-N) SMN = -1.00

ISH = N-1

K1 = ODBCNI+1

DO 230 I = 1,NC4

230 IF(SMN.EQ.-1) B(MC21) = -B(I1)

SOLVE THE (N,N) NORMAL EQUATION.
CALL LKLN(PUL, B, NC2, X)

COMPUTE THE A POSTERIORI STANDARD ERRORS OF THE COEFFICIENTS PER DEGREE.

FM0ISE = 0.00
FCSD = 0.00
DO 232 I = 1, NB
FCSD = FCSD+B(I)*X(I)
232 FM0ISE = FM0ISE+X(I)**2*(W(I)-REGUL)
END
FCSD = FCSD+FF
FM0ISE = FM0ISE+FF
FC(1) = FC(1)+FCSD
FM(1) = FM(1)+FM0ISE
CALL WRITE(X(I), 30, 8E6, 8999, 8999)
CONTINUE

CHECK FOR AGREEMENT BETWEEN ORIGINAL AND RECONSTRUCTED R.H.S.'S
BY TESTING THE (N,X) CASES.

SO = 0.00
SAVE = 0.00
DO 244 J = 1, NC2
BT(J) = 0.00
DO 244 J = 1, NC2
KT = IBN(J) + J
IF(J.GT.1) KT = IBN(J) + 1
244 BT(1) = BT(1) + KT*(A(KT))
SAVE = SAVE+BT(J)**2
245 SO = SO+(S(J)**2/BT(J))**2
SO = DSQRT(SO/NC2)
IF(SAVE.EQ.0) SO = DSQRT(SO/SAVE)
WRITE(6, 247) N, SO
247 FORMAT('ACCUACY OF INVERSION OF R(1,1) IS', 16, 8.2)
GO TO 300
291 DO 299 J = 1, NM1
290 CALL WRITE(300(1), N, 500)
299 CONTINUE
300 CONTINUE
CALL REWIND(30)
FINISH COMPUTING THE A POSTERIORI COEFFICIENTS' ERRORS PER DEGREE.

DO 310 I = 2, NM1
210 E2IM1 = 2*1
220 ALIAS = E2IM1+1
230 CONTINUE
310 CONTINUE
999 CONTINUE
RETURN
SUBROUTINE FUB(AN,M1,HP,BQ,IDD,FF,NP1)

C THIS SUBROUTINE COMPUTES THE FOURIER COEFFICIENTS
C OF THE BLOCK COVARIANCES.

C

IMPLICIT REAL*8 (A-H,O-Z)

DIMENSION HP(1),BQ(1),IDD(1),FF(1)

DIMENSION AHP(310)

COMMON/MM/M(310)

IF(MM=M1.EQ.1) GO TO 15

AR = 0.0

DO 10 M1 = M1,NP1

10 AR = AR+HP(IX)*BQ(IX)

AHP(M1) = AR

M1 = M1+1

RETURN

END
FORTRAN IV C1 RELEASE 2.0 ANALYSIS DATE = 08242 12/29/74 PAGE 0001

0001 SUBROUTINE ANALYSIS X, BLOCK, DATA, NN, RCNM, RSMH, AM, BM, BROWS, 00056100
      2 ROW, CR, SH, IW, CAA, CBB, SAA, SBI

C
C THIS SUBROUTINE USES THE OUTPUT OF SUBR. "NORMAL", 00056400
C READ FROM UNIT 10, TO COMPUTE THE SPHERICAL HARMONIC 00056400
C COEFFICIENTS OF THE DATA STORED IN ARRAY "DATA" AS 00056400
C (BLOCK) DEGREES X (BLOCK) DEGREES AREA MEANS, USING 00056400
C AN ALGORITHM FOR FAST COLLOCATION, THIS SUBR. CALLS SUBR. 00056400
C "FFCSIN" FROM THE IMSL DOUBLE LIBRARY. 00057400
C THE ORIGINAL CONTENTS OF ARRAY "DATA" ARE DESTROYED. 00057400
C THE GRID IS SUPPOSED TO BE EQUAL ANGULAR. 00057400
C THE NUMBER OF BLOCKS PER ROW (NC) MUST BE EVEN. 00057400
C PROGRAMEHD BY OSCAR L. COLOMBO, DEPT. OF GEODET. SC., OHIO S.U., 00057400
C 00057600

0002 IMPLICIT REAL(A-H,O-Z)

0003 DIMENSION RNRH (1), RSMH(1), DATA(1), AM(1), BM(1), XI(1), FR(1), CR(1)
      2, ROW(1), CRH(1), CAA(1), CBB(1), SAA(1), SBI(1)

0004 DIMENSION ROW(1)

0005 DIMENSION D1(4000)

0006 PI = 3.14159265358979320

0007 DCONV = PI/100, Do

0008 BLK = BLOCK + BLOCK

0009 M1 = M+1

0010 M0 + M1(M1+1)/2

0011 NC = 360, Do BLOCK

0012 NC2 = NC/2

0013 NC2 = NC2

0014 NC2F = NC2+1

0015 NKB + RKB890

0016 DO 5 M = 2, M1

0017 H = H+1

0018 AM(M1) = DSIN(M+BLK)/M

0019 5 IPH(M1) = (1.0-DCOS(M+BLK))/M

0020 AM(M) = BLK

0021 H = H+1, 10 DO 

0022 B0 H = 1, M0

0023 RCNM(1) = 0.0, 10 DO

0024 0 RCNM(1) = 0.0

0025 0 ISHH = -NC

0026 ISHP = -NC

0027 ISHP = -NC

0028 DO 10 NKB = 1, NBROWS

0029 ISHP = ISHP+NC

0030 10 DO H = 1, NC

0031 10 DO 1 = DATA(I HS)

0032 CALL FFCS1M ROW, CR, SH, IW, CAA, CBB, SAA, SBI

0033 LH = LH+NC2F

0034 LH = LH+NC2F

0035 LH = LH+NC2F

0036 LH = LH+NC2F

0037 LH = LH+NC2F

0038 LH = LH+NC2F

0039 LH = LH+NC2F

0040 LH = LH+NC2F

0041 LH = LH+NC2F

0042 LH = LH+NC2F

0043 LH = LH+NC2F
NOFFC = NOFFC+NC
NOFFF = NOFFS+RCP
CA(A(I)) = B(NOFFC)*AM(M1)
CRH(I) = B(NOFFC)*BM(M1)
SAA(I) = B(NOFFS)*AM(M1)
SBB(I) = B(NOFFS)*BM(M1)

166 CONTINUE
DO 20 MI = M1,M1
20 CONTINUE
DO 20 I = 1,NROWS
CA = CAA(I)
CR = CRH(I)
SA = SAA(I)
SB = SBB(I)
HR = I
IF(M.LE.NC2) GO TO 16
SA = -SA
SB = -SB
16 RCMF(RUC) = RCMF(RUC)+X(NR)*(CA+SB)
20 CONTINUE
END
increasing \( m \), stored in vector array form column by column; \( 30 \) stores the \( \chi_1^{a} \) of the optimal quadratures-type estimator. The \( \chi_i^{a} \) are stored from \( N \) Pole to \( S \) Pole, and according to \( n \), as the Legendre integrals and the coefficients. In some circumstances the grid may be geocentric rather than geodetic and a change of coordinates might be desirable; this can be achieved by setting the parameter IGEO to \( 2 \). The flattening assumed for this transformation is \( F = 1/298.257 \).

After the \( R(m) \) matrices have been created, they are inverted by IMSL subroutine LUDECP, that performs a Choleski factorization. IMSL subroutine LUELM solves the equations resulting in the \( \chi_1^{a} \); if during the inversion LUDECP detects an ill-conditioned (or a singular) matrix, the solution part is avoided, and a set of null \( \chi_1^{a} \) is stored for that particular \( m \). As an additional check for the stability of the solution, the relative residuals.

\[
\begin{align*}
    r &= \sum_{i=0}^{N} \frac{v_i^2}{(k_i^{a})^2} \\
    \text{where } v &= [v_0, v_1, \ldots v_{M-1}]^T \text{ is}
\end{align*}
\]

\[
    v = k^{a} - R(m) \chi^{a} \text{ (computed)}
\]

are computed and printed. In all the cases studied here these residuals indicated an agreement of at least 9 significant figures. To improve the stability of the solution, a regularizing constant REGUL is added to the diagonal elements of the \( R(m) \) (Paragraph (3.3)).

Arrays \( PN \), \( SS \), and \( FC \) contain the propagated noise, sampling, and total error measure (variance) per degree. \( W \) contains the averaged row variances (expression (2.43)) arranged from North to South.

The scalar arguments, \( NMAX \), \( NN \), \( DGRID \), \( IGEO \), \( REGUL \), \( NRUN \), and \( NC \), are described in the comments inserted between statements 0006 and 0010. The arrays are as follows:

<table>
<thead>
<tr>
<th>NAME</th>
<th>DIMENSION</th>
<th>TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROWP</td>
<td>( \frac{1}{2}(NMAX+1)(NMAX+2) )</td>
<td>REAL * 8</td>
</tr>
<tr>
<td>ROWQ</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>RHS</td>
<td>( \frac{1}{4}NN \ (NN \ + \ 1) \ N \ (N \ = \ \text{Nyquist freq.}) )</td>
<td>&quot;</td>
</tr>
<tr>
<td>S</td>
<td>( (NN \ + \ 1) \ N )</td>
<td>&quot;</td>
</tr>
<tr>
<td>A</td>
<td>( \frac{1}{2}N(N-1)+N )</td>
<td>&quot;</td>
</tr>
<tr>
<td>UL</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>W</td>
<td>N</td>
<td>&quot;</td>
</tr>
<tr>
<td>DVAR</td>
<td>( NMAX + 1 )</td>
<td>&quot;</td>
</tr>
<tr>
<td>FC</td>
<td>( NN + 1 )</td>
<td>&quot;</td>
</tr>
<tr>
<td>PN</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
<tr>
<td>SS</td>
<td>&quot;</td>
<td>&quot;</td>
</tr>
</tbody>
</table>

Arrays \( ANMPQ \), \( FF \), \( XO \), \( B \), \( X \), \( BT \), \( FINMP \) (all REAL * 8), and \( IDD \) (REAL * 4),
all dimensioned 200 or 400 in the subroutine itself, are large enough for
problems where \( N < 200 \). For finer grids, the size of these arrays should be
increased in the same proportion as that of \( N \).

The \( R(m) \) matrices are formed according to expression (2.63). Subroutine
FUR computes the "aliased" Fourier coefficients of the covariance functions
that are, in fact, the elements of the \( R(m) \), scaled by \( N \) or \( 2N \), depending on
\( m \). Common MM and array MT are part of a logic set up to ensure that the
Fourier coefficients are not computed more than once each.

Subroutine ANALYS uses the \( \chi_{i}^{2} \) stored in unit 30 to analyze the data
in array DATA. The \( \alpha_{i}, \beta_{i} \) are formed in place, as in HARMIN, so the
original values in DATA are destroyed. Arrays CR, SR, CAA, CBB, SAA, and SBB
have all the same description as CR1, SR1, etc., in HARMIN.
IMSL subroutine FFCSIN (double precision) is used to obtain the \( \alpha_{i}, \beta_{i} \). The
reason why ANALYS is used instead of HARMIN, is the arrangement of the
\( \chi_{i}^{2} \) "columnwise", or by increasing latitudes, rather than "rowwise" (i.e., all
the \( \chi_{i}^{2} \) for the same \( i \) stored together) as HARMIN would require.

The listings of FUR, ANALYS, and those of the fast input/output sub-
renoutines FREAD, FWRITE, REWIND, are given after that of NORMAL.
The input/output subroutines have dummy arguments, because originally NORMAL
was written to work with certain subroutines available at O.S.U. that may not
be in the software libraries of other institutions.

B.5 Subroutine NORMAX

A modified version of NORMAL, this subroutine was created to compute
the variance of the estimation errors in ordinary quadratures formulas according
to the theory in section 2. Essentially, it computes

\[
\sigma^{2}_\epsilon = \sigma^{2}_\epsilon - 2 \sum_{z=0}^{n} \sum_{\alpha=0}^{r} C_{z} \Sigma_{\alpha} \cdot f_{\alpha}^{\alpha} + \sum_{z=0}^{n} \sum_{\alpha=0}^{r} (f_{z}^{\alpha} \cdot (C_{z} + D) \cdot f_{\alpha}^{\alpha}
\]

by forming and using the \( R(m) \). Since no inversion or solution of the normal
equations is required, the corresponding segments have been removed from
NORMAL, and a new final segment added for the computation of the various
accuracies.

The theory behind the calculation of the \( \sigma^{2}_\epsilon \) using the \( R(m) \) matrices is
as follows:

In the case of ordinary formulas of the type (see expression (1.7))
SUBROUTINE NORMAX(MMAX, NN, DCR1D, ICED, NRN, ROPT, ROWQ, RBS, S, A,  
2 UL, NC2, K, DVAR, QUADS, XNS, FC, PM, SS)

SUBROUTINE FOR COMPAREING NUMERICAL QUADRATURES FORMULAS USING
COLLOCATION THEORY AND "FAST COLLOCATION"
FOR DATA AVERAGED OVER EQUAL ANGULAR BLOCKS
ON A SURFACE OF REVOLUTION.

INTEGALS IRNP ARE READ IN FROM UNIT 6, STORED, AFTER
MULTIPLYING THEM BY SUITABLE FACTORS, IN UNITS 10 AND 12.
THE RCM'S MATRICES ARE STORED IN UNIT 15.
OPTIMAL "QUADRATURES" WEIGHTS IN UNIT 30.

PROGRMRED BY OSCAR L. COLOMBO, GEOS. SC., OHIO STATE U., SEPT. 1979.

IMPLICIT REAL*8(A-H, O-Z)

DIMENSION ROMP(1), ROWQ(1), RBS(1), S(1), A(1), UL(1)
DIMENSION ARMP(200), SS(1, RI2, RI1, XU, UC), PP(200)
2 IDK(200), AIC(200), FIC(200), PM(1), W(1)
3 PP(1), FVAPR(200), FVBAR(200), DVAR(1)
DIMENSION QUADS(1)
DIMENSION XNS(1)
COMMON/FR/ HTG(10)

BASIC CONSTANTS
PI = 3.1415926535897932D0
F = 1.00-294.2579D0

PROBLEM DEFINITION
DCR1D: BLOCK SIZE IN DEGREES, MINIMUM: 5 DEGREES.
ICED = 2 IF GEODETIC LATITUDES ARE TO BE CONVERTED TO EQUATORI
TO EQUATORIAL.

NN MAX. DEC. AND ORDER TO BE ANALYZED (LESS THAN NRS)
NC NO. OF "COLUMNS", ALWAYS EVEN.
NR NO. OF ROWS. THERE IS NO EQUATORIAL ROW.

RMAX MAX. DEGREE IN BAND LIMITED COVARIANCE FUNCTION.
NRN = 0 IF RCM'S ARE CREATED AND STORED IN
UNIT 15.
RKU = 999 IF RCM'S ARE READ FROM UNIT 15 INSTEAD.
REGU REGULARIZATION PARAMETER, ADDED TO ALL DIAGONALS.

DCR = DCR1D+DCRNV
NC = 360.16/DCR1D+0.000010D0
NC4 = NC0.25+0.000010D0
NC2 = NC4/2
NR = NC2
RR1 = RR+1
NF = 1
NP1 = NMAX+1
M1 = (NN1+1)/(N1+1)
M2 = (NN1+1)/(N1+1)
MNP = DNP
MLH = DMLH

PAGE 0001
C FORM ARRAY OF GEOCENTRIC OR GEODETIC GRID'S GEOCENTRIC LATS

C

0835    EX = F((2.00-F)
0836    XM = 2.00+M/(1.00+MX*2)
0837    YC2P = NC2+1
0838    ANG = 90.00
0839    DO 6 I = 1, NC2P
0840    PHI(I) = ANG*DGONV
0841    IF(DEO, EQ, 2) PHI(I) = PHI(I) - EM*DSIN(2.00*PHI(I))
0842    6 ANG = ANG-DCRD
0843    DO 15 I = 2, NF1
0844     1 = 11-1
0845     FINP(I) = DSQRT(DVAR(I)/(2.00*11))
0846    15 CONTINUE

C READ TAPE OF INTEGRALS OF LEGENDRE FUNCTIONS FROM UNIT 6

C AND MULTIPLY THEM BY THE SQUARE ROOT OF THE DEGREE VARIANCE
C FOR COEFFICIENT, TIMES THE INVERSE OF THE BLOCK AREA
C (FROM RN), STORING THEM IN UNIT 10.

C

0848    IBS = -N51
0849    DO 25 NT = 1, NC2
0850    IMS = IMS+MS1
0851    FINN1 = 1.00/((DSIN(PHI(NT)) - DSIN(PHI(NT+1)))*DCR)
0852    CALL FREADEH(RHS, NT, NPP, 8999, 8999)
0853    IF(NOT, END, R) GO TO 19
0854    NUS = 1
0855    DO 30 NI = 1, N1
0856    DO 10 MI = 1, NI
0857    BIGN = NSI*NORTH, NORT=RHS, QUADE(IMS+MI)
0858    DO 84 NUS = N1-1
0859    CALL FWRITE(NORTH, 10, MI, NI, 8999, 8999)
0860    18 NUS = N1+1
0861    19 CONTINUE

C FORM ARRAY OF GEOCENTRIC OR GEODETIC GRID'S GEOCENTRIC LATS

C

0862    NIS = 1
0863    DO 20 I = 1, NI
0864    FX = FINP(I)*FROWN
0865    DO 20 11 = 1, 1
0866    ROWN = NIS, NIS*FX
0867    20 NIS = NIS+1
0868    CALL FWRITE(ROWN, 10, NF1, 8999, 8999)
0869    23 CONTINUE

C FORM ARRAY OF GEOCENTRIC OR GEODETIC GRID'S GEOCENTRIC LATS

C

0870    NIS = NIS+1
0871    WRITE(5, 553) BGRID, NMAX, NN, NR, NRW
FORTAN IV C1 RELEASE 2.0

0072  555 FORMAT(3,F7.2)     BLOCK SIZE ',x,F7.2,' MAX. DEGREE IN COVARIANCE ',x,5,'  
0073  2 MAX. DEGREE TO BE ANALYZED ',x,5,' NUMBER OF ROWS ',x,5,' NUMBER 
0074  3 15/11)                   WRITE(6,556) NC                    000391000
0075  556 FORMAT(' NUMBER OF COLUMNS ',I5/)                   000391100
0076  IF (ICEO.EQ.1) WRITE(6,557)                           000391200
0077  IF (ICEO.EQ.2) WRITE(6,558)                           000391300
0078  557 FORMAT(' DATA ON A SPHERE',/)                    000391400
0079  558 FORMAT(' DATA ON AN ELLIPSOID',/)                000391500
0079  DO 99 MI = 1, M1                                   000391600
0080  99 DO 100 M1 = 1, M1                                 000391700
0081  DO 100 MI = 2, M1                                   000391800
0082  100 FF(MI) = 2.0D0*(1.0D0-BDOS((MI-1)*BDGR)/(MI-1)**2  000391900
0083  C READ THE R.H.S. 'S OF THE REDUCED 'NORMALS' FROM UNIT 10 . 000392000
0084  C INTO VECTOR RHS .                                000392100
0085  C                                              000392200
0086  0083 CALL REWINDA(10)                               000392300
0087  CALL REWINDA(30)                                   000392400
0088  ISH = MLL                                         000392500
0089  DO 210 J = 1, NC4                                   000392600
0090  CALL FREAD(Rowi(1),10,MLLB,9999,9999)              000392700
0091  CALL FREAD(Rowi(1),30,MLLB,9999,9999)              000392800
0092  ISH = ISH+MLL                                      000392900
0093  DO 210 J = 1, NLL                                   000393000
0094  XBTE(ISH+J) = ROWP(J)                              000393100
0095  210 XBTE(ISH+J) = ROWP(J)                           000393200
0096  IF (RNBN.EQ.999) GO TO 1500                         000393300
0097  C MAIN LOOP FOR THE CREATION OF THE R(M)'S MATRICES. 000393400
0098  C                                                000393500
0099  0094 CALL REWINDA(10)                               000393600
0100  CALL REWINDA(30)                                   000393700
0101  CONTINUE                                           000393800
0102  DO 101 J = 1, ND                                   000393900
0103  101 S(I) = 0.0                                      000394000
0104  C POSITION THE R(M)'S TAPE TO CURRENT ROW NP .    000394100
0105  DO 102 JX = 1, NP                                   000394200
0106  102 CALL FREAD(Rowi(1),10,NPPB,9999,9999)          000394300
0107  103 CALL FREAD(Rowi(1),30,NPPB,9999,9999)          000394400
0108  NC = RC2-NP+1                                      000394500
0109  DO 120 JQ = NP, NC                                  000394600
0110  120 IF (SQ.NE.NP) CALL FREAD(Rowi(1),10,NPPB,9999,9999) 000394700
0111  IF (SQ.NE.NP) GO TO 105                            000394800
0112  DO 105 J = 1, M1                                   000394900
0113  105 BOXQ(I) = BOXP(I)                              000395000
0114  C OBTAIN THE FOURIER COEFFICIENTS OF THE BLOCK COVARIANCES 000395100
0115  C CORRESPONDING TO HOMS P AND Q IN EBD.             000395200
0116  C AND ALIAS THEM TO FORM THE ELEMENTS OF THE R(M)'S MATRICES, 000395300
0117  C                                                  000395400
0118  105 CONTINUE                                         000395500
0119  DO 103 J = 1, 1310                                  000395600
0120  103 M1(I) = 0.0                                     000395700
0121  1 M1(I) = 0                                        000395800
0122  1 CONTINUE                                         000395900
0160 CALL FREAD(SB(1), 15, NC20, 00999, 9999)
0161 DO 220 J = 1, NC2
0162 IF(1,CT,NC2JJ) GO TO 220  
0163 K = 100(J) + J
0164 NC21 = NC21 - 1
0165 KT = IDB(KC2JJ)+NC211  
0166 A(K) = SS(J)
0167 A(KP) = SS(tt)
0168 IF(1,RE,JC) GO TO 220  
0169 A(K) = A(K)+WC(J)
0170 A(KP) = A(KP)+WC(NC211)

220 CONTINUE

C C FORM B(M,N), X(N,M), AND COMPUTE THE THEORETICAL 
C ERRORS PER DEGREE .
C
0172 DO 240 M = 1, MM1  
0173 BM = 1.00  
0174 IF((M-M1)/2)*2*NE.(M-M1) MM = 1.00  
0175 IR = M1
0176 KT = IDB(K) + M1
0177 DO 230 J = 1, NC4  
0178 IS = 1SH+M1
0179 X(J) = IRS(1SH+KT)*1XMEP(M1)  
0180 X(J) = IRS(1SH+KT)
0181 NC21 = NC21 - 1
0182 X(NC211) = X(J)
0183 X(NC211) = X(J)
0184 IF(SMM.EQ.-1.00) X(NC211) = -X(J)
0185 IF(SMM.FA.X.-1.00) X(NC211) = -X(MC211)

230 CONTINUE

C C FPROISE = 0.00  
0187 FSBD = 0.00  
0188 DO 223 J = 1, NR
0189 FSBD = FSBD*XC*(J)*SH(J)
0190 F(t) = F(t)+FSD*XC*(J)*SH(J)

223 CONTINUE  

C X = 0.00  
0192 DO 245 J = 1, NC2
0193 HT(J) = FSO.J+J
0194 DO 244 J = 1, NC2
0195 KT = IDB(K) + J
0196 IF(J, CT, J) KT = IDB(K) + J
0197 HT(J) = FR(J, CT, J) = IDB(K) + J
0198 244 HT(J) = HT(J)+K(J)*SH(J)

245 J = 100(J) + J
0199 NC(S1) = NC(S1)+FSD+(2,00+FSBD)+SS(J)
0200 FN(S1) = FN(S1)+FSD+FSD+FPROISE
0201 ULSN1 = ULC*(1)+SOS+FDDY
0202 CALL WRITE(X, M, S, MM1, 9999, 9999)

240 CONTINUE

300 CONTINUE  
0204 DO 240 J = 1, MM1
0205 FSO(J) = FSBD*J
0206 IF(J, CT, J) FSBD*J
0207 ULSN1 = ULC*(1)+SOS+FDDY
0208 ALIAS = E50F(J)+FN(J)
0209 FG(J) = E5OF(J)+FN(J)
0210 S(S1) = ALIAS
2111 310 CONTINUE
0212 CALL DEWIND(4)
0213 CALL DEWIND(13)
\[ C_{\alpha \alpha} = (f_{\alpha \alpha})^T \mu \]

\[ = \mu_n \sum_{i=1}^{N-1} \sum_{j=0}^{N-1} \int_{\sigma_{ij}} \nabla^\alpha \sum_{t=0}^{N-1} \left\{ \begin{array}{l}
A(m) \cos mj \Delta \lambda + \{B(m) \sin mj \Delta \lambda \}
\end{array} \right\} m_{ij} \]

the estimator vector is a combination of a "sine" and a "cosine" vector of
"frequency" m (terminology introduced in paragraph (2.10)). The product of
such vector by \((C_{zz} + D)\) is, because of the structure of this matrix, another
combination of a "sine" and a "cosine" vector of the same "frequency". From
the properties of the normal matrix follows that

\[(C_{zz} + D) f_{\alpha \alpha} = \begin{bmatrix} 2N & \text{if } m = 0 \\
\text{or } m = N \\
N & \text{if } m \neq 0 
\end{bmatrix} \begin{bmatrix} \ldots \nu_{1}^{\alpha} \\
A(m) \cos mj \Delta \lambda + \{B(m) \sin mj \Delta \lambda \} \\
A(m) \end{bmatrix} \]

where, calling

\[ \nu_{\alpha}^{m} = [\nu_{0}^{\alpha}, \nu_{1}^{\alpha}, \ldots, \nu_{N-1}^{\alpha}]^T = R(m) [x_{\alpha,0}, x_{\alpha,1}, \ldots, x_{\alpha,N-1}]^T \]

and

\[ F(m) = \begin{cases}
2N \Delta \lambda^2 & \text{if } m = 0 \text{ or } m = N \\
4N(1 - \cos m \Delta \lambda) & \text{if } m \neq 0, m \neq N
\end{cases} \]

is

\[ (f_{\alpha \alpha})^T (C_{zz} + D) f_{\alpha \alpha} = F(m) \sum_{t=0}^{N-1} \nu_{t}^{\alpha} \]

(B4-a)

Regarding the scalar product \(2C_{\alpha \alpha}^\dagger \mathbf{f}_{\alpha \alpha}^\dagger \), it is easy to show that

\[ 2C_{\alpha \alpha}^\dagger \mathbf{f}_{\alpha \alpha}^\dagger = 2 \frac{\sigma_{\alpha}^{2}}{2n + 1} \sum_{t=0}^{N-1} (x_{t}^{\alpha})^2 F(m) \]

(B4-b)

Expressions (B4-a) and (B4-b) are implemented by NORMAX to obtain the
error variances. This subroutine also uses FUR. Subroutine QUADS is
also called, to obtain the de-smoothing factors. In the case listed here, this
factor is \( \mu_n = \frac{1}{4n^2 \mu_{\alpha}} \). Array QUADS has been added to the list of arguments,
and it contains the \( \mu_n \) after the call to QUADS.

B.6 Subroutine LEGFDN

This subroutine computes the normalized Legendre functions and, if so
desired, their derivatives at a given \( \theta_1 \). All values correspond to the same
order M; if more than one order at a time is needed, a DO loop, where the
subroutine is called once for each order, can be set up. The subroutine is
based on formulas (4.19 a-b) and (1.38 a-b). The use of this subroutine is
explained by the comments inserted in the listing. The stability of the recursive
formulas was tested by computing \( \mathbf{R}_{\alpha} \) (cos \( \theta \)) and \( (d\mathbf{R}_{\alpha}/d\theta)(\cos \theta) \) for \( m = 350 
and 350 \leq n \leq 400 \), \( 2.5^\circ \leq \theta \leq 90^\circ \), at \( 5^\circ \) intervals. Calculations were done
first in double precision (8 byte words) and then repeated in extended precision.
SUBROUTINE LEGMN(M,THETA,RLEG,DLLEG,NNX,IR,RLMN,IFLAG)

C THIS SUBROUTINE COMPUTES ALL NORMALIZED LEGENDRE FUNCTIONS
C IN 'RLLEG' AND THEIR DERIVATIVES IN 'DLLEG'. ORDER IS ALWAYS
C N, AND COLATITUDE IS ALWAYS THETA (RADIANS). MAXIMUM DEGREE
C IS NNX. ALL CALCULATIONS IN DOUBLE PRECISION.
C IR MUST BE SET TO ZERO BEFORE THE FIRST CALL TO THIS SUB.
C THE DIMENSIONS OF ARRAYS RLEG, DLEG, AND RLMN MUST BE
C AT LEAST EQUAL TO NNX+1.
C IF THIS SUBROUTINE IS TO BE USED TO COMPUTE FUNCTIONS
C AND THEIR DERIVATIVES FOR MORE THAN ONE ORDER M, THEN
C THE HIGHEST ORDER SHOULD BE COMPUTED IN THE FIRST CALL.
C
C THIS PROGRAM DOES NOT COMPUTE DERIVATIVES AT THE POLES.
C
C IF IFLAG = 1, ONLY THE LEGENDRE FUNCTIONS ARE
C COMPUTED.

C PROGRAMMER: OSCAR L. COLOMBO, DEPT. OF GEODETIC SCIENCE,
C THE OHIO STATE UNIVERSITY, AUGUST 1969. *****************************

C
C IMPLICIT REAL(A-H,O-Z)
C DIMENSION RLEG(1),DLEG(1),RLMN(1)
C 2, DINTN(100), DIINTN(100)
C 0004 NNX1 = NNX+1
C 0005 NNX2P = 2*NNX+1
C 0006 N1 = M+1
C 0007 N2 = M+2
C 0008 N3 = M+3
C 0009 IF(IR.EQ.1) GO TO 10
C 0010 IR = 1
C 0011 DO 3 N = 1,NNX2P
C 0012 DIINTN(N) = DIINTN(N)+1,DIINTN(1)
C 0013 3 DIINTN(N) = 1.0/DIINTN(N)
C 0014 10 COTHET = DIINTN(N)
C 0015 STHET = RSIND(THETA)
C 0016 STH1 = 1.0/STHET
C
C COMPUTE THE LEGENDRE FUNCTIONS.
C
C 0017 RLMN(1) = 1.0
C 0018 RLMN(2) = STHET*STHET
C 0019 DO 15 N = 3,M
C 0020 N = N-1
C 0021 N2 = 2*N
C 0022 15 RLMN(N) = DIINTN(N2+1)*DIINTN(N2)*STHET*RLMN(N-1)
C 0023 IF(N.EQ.0) GO TO 20
C 0024 IF(N.EQ.9) GO TO 16
C 0025 RLEG(2) = RLMN(2)
C 0026 DLEG(3) = DIINTN(3)*COTHET*RLEG(2)
C 0027 GO TO 20
C 0028 16 RLEG(1) = 1.0
C 0029 RLEG(2) = COTHET*DIINTN(3)
C 0030 CONTINUE
C 0031 DLEG(1) = RLMN(M)
C 0032 DLEG(2) = DIINTN(M+2)*COTHET*RLEG(M+1)
C 0033 DO 30 N = M+1,NNX
C 0034 N = N-1
IF(M.EQ.0. AND.M.LT.2.OR.M.EQ.1.AND.M.LT.3) GO TO 30
N2 = 2*M
2 BLEG(N1) = DRTS(N2+1)*DIRT(N+M)*DIRT(N-M)*(DIRS(N2-1)*COTHET
2 BLEG(N1-1)*DIRT(N+M-1)*DIRT(N-M-1)*DIRT(N2-3)*BLEG(N1-2))
GO TO 30
CONTINUE
IF(IFLAG.EQ.1) RETURN
V(SITHET.EQ.0.D0) WRITE(6,99)
99 FORMAT(// *** LEGFDN DOES NOT COMPUTE DERIVATIVES AT THE POLES
2 1V(SITHET.EQ.0.D0) RETURN
C
C COMPUTE ALL THE DERIVATIVES OF THE LEGENDRE FUNCTIONS.
C
1
1
1
1
1
(16 byte words). Both sets of results agreed with each other to better than 6 significant figures.

B.7 Subroutine NVAR

This subroutine computes the degree variances of the gravity anomalies (point values), up to degree 100, according to the coefficients obtained by R. Rapp from a complete, equal angular set of mean 1° x 1° anomalies, as mentioned in par. (3.1). Above \( n = 100 \), the subroutine uses a two-term model to calculate the \( \sigma^2_1(\Delta g) \). The resulting degree variances are stored in array DVAR. The first element in DVAR is \( \sigma^2_1(\Delta g) \).
SUBROUTINE NVAR(DVAR, NMAX, DVAR0)

THIS SUBROUTINE COMPUTES ALL DEGREE VARIANCES UP TO DEGREE "NMAX" USING SOME GIVEN MODEL.

IMPLICIT REAL*8 (A-H, O-Z)

DIMENSION DVAR(I)

DVAR(1) = 0.0

DVAR(2) = 7.639038
DVAR(3) = 33.942063
DVAR(4) = 19.743429
DVAR(5) = 19.694103
DVAR(6) = 10.393514
DVAR(7) = 16.455112
DVAR(8) = 16.399562
DVAR(9) = 8.6391739
DVAR(10) = 9.1839627
DVAR(11) = 6.369778
DVAR(12) = 2.8991999
DVAR(13) = 6.7296386
DVAR(14) = 3.2833850
DVAR(15) = 3.819187
DVAR(16) = 4.9444726
DVAR(17) = 3.7168758
DVAR(18) = 4.1951569
DVAR(19) = 3.456849
DVAR(20) = 2.2831356
DVAR(21) = 3.003872
DVAR(22) = 3.5602909
DVAR(23) = 2.6127612
DVAR(24) = 2.6800878
DVAR(25) = 3.4741027
DVAR(26) = 2.2678667
DVAR(27) = 2.1498928
DVAR(28) = 2.0683299
DVAR(29) = 2.0560112
DVAR(30) = 2.0287232
DVAR(31) = 2.4776694
DVAR(32) = 2.4971826
DVAR(33) = 3.2904088
DVAR(34) = 3.1679239
DVAR(35) = 3.1679239
DVAR(36) = 2.9680243
DVAR(37) = 3.0251384
DVAR(38) = 2.1799143
DVAR(39) = 2.0229179
DVAR(40) = 2.064616
DVAR(41) = 3.8649245
DVAR(42) = 3.0534486
DVAR(43) = 2.2442727
DVAR(44) = 3.0297999
DVAR(45) = 2.9169444
DVAR(46) = 2.7235179
DVAR(47) = 3.4543886
DVAR(48) = 2.6100541
DVAR(49) = 2.2049252
DVAR(50) = 3.4272534
<table>
<thead>
<tr>
<th>FORTRAN IV CI RELEASE 2.0</th>
<th>NVAR</th>
<th>DATE  =  09242</th>
<th>12/29/34</th>
<th>PAGE 0002</th>
</tr>
</thead>
<tbody>
<tr>
<td>0055</td>
<td>DVAR(51) = 2.8176826</td>
<td></td>
<td>00002600</td>
<td></td>
</tr>
<tr>
<td>0056</td>
<td>DVAR(52) = 2.6578377</td>
<td></td>
<td>00007700</td>
<td></td>
</tr>
<tr>
<td>0057</td>
<td>DVAR(53) = 3.4227993</td>
<td></td>
<td>00007900</td>
<td></td>
</tr>
<tr>
<td>0058</td>
<td>DVAR(54) = 3.0583736</td>
<td></td>
<td>00007900</td>
<td></td>
</tr>
<tr>
<td>0059</td>
<td>DVAR(55) = 3.1630691</td>
<td></td>
<td>00008000</td>
<td></td>
</tr>
<tr>
<td>0060</td>
<td>DVAR(56) = 2.8089749</td>
<td></td>
<td>00011000</td>
<td></td>
</tr>
<tr>
<td>0061</td>
<td>DVAR(57) = 3.2199365</td>
<td></td>
<td>00012000</td>
<td></td>
</tr>
<tr>
<td>0062</td>
<td>DVAR(58) = 2.6270499</td>
<td></td>
<td>00013000</td>
<td></td>
</tr>
<tr>
<td>0063</td>
<td>DVAR(59) = 2.9265646</td>
<td></td>
<td>00014000</td>
<td></td>
</tr>
<tr>
<td>0064</td>
<td>DVAR(60) = 3.2459338</td>
<td></td>
<td>00015000</td>
<td></td>
</tr>
<tr>
<td>0065</td>
<td>DVAR(61) = 2.6905177</td>
<td></td>
<td>00016000</td>
<td></td>
</tr>
<tr>
<td>0066</td>
<td>DVAR(62) = 3.1935159</td>
<td></td>
<td>00017000</td>
<td></td>
</tr>
<tr>
<td>0067</td>
<td>DVAR(63) = 2.6971074</td>
<td></td>
<td>00018000</td>
<td></td>
</tr>
<tr>
<td>0068</td>
<td>DVAR(64) = 2.7474096</td>
<td></td>
<td>00019000</td>
<td></td>
</tr>
<tr>
<td>0069</td>
<td>DVAR(65) = 2.3063364</td>
<td></td>
<td>00020000</td>
<td></td>
</tr>
<tr>
<td>0070</td>
<td>DVAR(66) = 3.4103538</td>
<td></td>
<td>00021000</td>
<td></td>
</tr>
<tr>
<td>0071</td>
<td>DVAR(67) = 2.6969258</td>
<td></td>
<td>00022000</td>
<td></td>
</tr>
<tr>
<td>0072</td>
<td>DVAR(68) = 2.5621252</td>
<td></td>
<td>00023000</td>
<td></td>
</tr>
<tr>
<td>0073</td>
<td>DVAR(69) = 3.1792376</td>
<td></td>
<td>00024000</td>
<td></td>
</tr>
<tr>
<td>0074</td>
<td>DVAR(70) = 2.4727642</td>
<td></td>
<td>00025000</td>
<td></td>
</tr>
<tr>
<td>0075</td>
<td>DVAR(71) = 2.3711426</td>
<td></td>
<td>00026000</td>
<td></td>
</tr>
<tr>
<td>0076</td>
<td>DVAR(72) = 2.7993256</td>
<td></td>
<td>00027000</td>
<td></td>
</tr>
<tr>
<td>0077</td>
<td>DVAR(73) = 2.8756998</td>
<td></td>
<td>00028000</td>
<td></td>
</tr>
<tr>
<td>0078</td>
<td>DVAR(74) = 2.9569934</td>
<td></td>
<td>00029000</td>
<td></td>
</tr>
<tr>
<td>0079</td>
<td>DVAR(75) = 2.9771508</td>
<td></td>
<td>00030000</td>
<td></td>
</tr>
<tr>
<td>0080</td>
<td>DVAR(76) = 2.5442199</td>
<td></td>
<td>00031000</td>
<td></td>
</tr>
<tr>
<td>0081</td>
<td>DVAR(77) = 2.3192014</td>
<td></td>
<td>00032000</td>
<td></td>
</tr>
<tr>
<td>0082</td>
<td>DVAR(78) = 2.6358035</td>
<td></td>
<td>00033000</td>
<td></td>
</tr>
<tr>
<td>0083</td>
<td>DVAR(79) = 2.5151899</td>
<td></td>
<td>00034000</td>
<td></td>
</tr>
<tr>
<td>0084</td>
<td>DVAR(80) = 2.0430023</td>
<td></td>
<td>00035000</td>
<td></td>
</tr>
<tr>
<td>0085</td>
<td>DVAR(81) = 2.0864117</td>
<td></td>
<td>00036000</td>
<td></td>
</tr>
<tr>
<td>0086</td>
<td>DVAR(82) = 3.1918396</td>
<td></td>
<td>00037000</td>
<td></td>
</tr>
<tr>
<td>0087</td>
<td>DVAR(83) = 2.7446152</td>
<td></td>
<td>00038000</td>
<td></td>
</tr>
<tr>
<td>0088</td>
<td>DVAR(84) = 3.0000133</td>
<td></td>
<td>00039000</td>
<td></td>
</tr>
<tr>
<td>0089</td>
<td>DVAR(85) = 2.3413319</td>
<td></td>
<td>00040000</td>
<td></td>
</tr>
<tr>
<td>0090</td>
<td>DVAR(86) = 2.6756262</td>
<td></td>
<td>00041000</td>
<td></td>
</tr>
<tr>
<td>0091</td>
<td>DVAR(87) = 2.6775225</td>
<td></td>
<td>00042000</td>
<td></td>
</tr>
<tr>
<td>0092</td>
<td>DVAR(88) = 2.5171941</td>
<td></td>
<td>00043000</td>
<td></td>
</tr>
<tr>
<td>0093</td>
<td>DVAR(89) = 2.4930952</td>
<td></td>
<td>00044000</td>
<td></td>
</tr>
<tr>
<td>0094</td>
<td>DVAR(90) = 2.4403774</td>
<td></td>
<td>00045000</td>
<td></td>
</tr>
<tr>
<td>0095</td>
<td>DVAR(91) = 2.7512343</td>
<td></td>
<td>00046000</td>
<td></td>
</tr>
<tr>
<td>0096</td>
<td>DVAR(92) = 2.3106676</td>
<td></td>
<td>00047000</td>
<td></td>
</tr>
<tr>
<td>0097</td>
<td>DVAR(93) = 2.5410976</td>
<td></td>
<td>00048000</td>
<td></td>
</tr>
<tr>
<td>0098</td>
<td>DVAR(94) = 2.5721650</td>
<td></td>
<td>00049000</td>
<td></td>
</tr>
<tr>
<td>0099</td>
<td>DVAR(95) = 2.2715446</td>
<td></td>
<td>00050000</td>
<td></td>
</tr>
<tr>
<td>0100</td>
<td>DVAR(96) = 2.3346230</td>
<td></td>
<td>00051000</td>
<td></td>
</tr>
<tr>
<td>0101</td>
<td>DVAR(97) = 2.3560790</td>
<td></td>
<td>00052000</td>
<td></td>
</tr>
<tr>
<td>0102</td>
<td>DVAR(98) = 2.6100727</td>
<td></td>
<td>00053000</td>
<td></td>
</tr>
<tr>
<td>0103</td>
<td>DVAR(99) = 2.2631529</td>
<td></td>
<td>00054000</td>
<td></td>
</tr>
<tr>
<td>0104</td>
<td>DVAR(100) = 2.3901125</td>
<td></td>
<td>00055000</td>
<td></td>
</tr>
<tr>
<td>0105</td>
<td>IF (NMAX.I.E. 100) GO TO 20</td>
<td></td>
<td>00056000</td>
<td></td>
</tr>
<tr>
<td>0106</td>
<td>A  =  1.0000</td>
<td></td>
<td>00057000</td>
<td></td>
</tr>
<tr>
<td>0107</td>
<td>B  =  2.0000</td>
<td></td>
<td>00058000</td>
<td></td>
</tr>
<tr>
<td>0108</td>
<td>AI  =  3.4085000</td>
<td></td>
<td>00060000</td>
<td></td>
</tr>
<tr>
<td>0109</td>
<td>A2  =  1.4000000</td>
<td></td>
<td>00061000</td>
<td></td>
</tr>
<tr>
<td>0110</td>
<td>N1  =  0.9900000</td>
<td></td>
<td>00062000</td>
<td></td>
</tr>
<tr>
<td>0111</td>
<td>N2  =  0.914423300</td>
<td></td>
<td>00063000</td>
<td></td>
</tr>
<tr>
<td>0112</td>
<td>D0  =  10</td>
<td>N  =  101,NMAX</td>
<td></td>
<td>00063000</td>
</tr>
</tbody>
</table>
B. 8 SUBROUTINE COVBLK

Subroutine COVBLK calculates covariances between area means according to expression (4.14). This subroutine is listed in the following pages. The arguments are explained in the comments at the beginning of the listing. In addition, the following things have to be born in mind: the dimension of the array DVAR is Nmax; the dimension of both RINS and RINN is \( \frac{N_{\text{max}}+1}{2}(N_{\text{max}}+2) \); the dimension of COVS is 360/BLOCK. The values returned in DVAR are the original degree variances \( c_0^{uv} \), each divided by \( 2n+1 \). If LB is less than 360/BLOCK, the LB+1, LB+2, ..., (360/BLOCK - LB - 1) elements in COVS are returned as zeroes, the remainder contain the first (and last) LB covariances. The dimension of \( F \) is 180/BLOCK (Nyquist frequency). To use this subroutine with \( N_{\text{max}} > 400 \), the arrays FF and IDD (whose dimension should be no less than \( N_{\text{max}} \)), should be redimensioned.

The subroutine does not take advantage of the "aliasing" of Fourier coefficients built into expression (4.14). Implementing this aspect should lead to some additional improvement in efficiency. The Fourier series is computed, once the coefficients have been determined, by multiplying each coefficient by the corresponding cosine of \( m\lambda_1 \) and adding the products together. The values of \( \cos m\lambda_1 \) are computed using the following recursive formula:

\[
\cos m\lambda_1 = 2 \cos \lambda_1 \cos (m-1)\lambda_1 - \cos (m-2)\lambda_1
\]

which avoids repeated calculation of the FORTRAN COS function (only \( \cos \lambda_1 \) is required to start the recursion). Actual calculation of the Fourier series requires about 0.04 seconds in the most time consuming case: the grid of 1° blocks. The greater part of the time taken by this subroutine goes into finding the Fourier coefficients of the mean value covariances. For this reason, there is not much difference between computing all covariances in a certain row, or just a few of them, using this procedure.
SUBROUTINE COVBLK(DVAR,RINS,RINN,MAX,BLOCK,DLATS,DLATN,COVS,F,K,2,LB)

THIS SUBROUTINE COMPUTES COVARIANCES BETWEEN AN AREA MEAN
IN ROW "N" AND LB CONTIGUOUS AREA MEANS IN ROW "S",
BEGINNING AT THE BLOCK WITH THE SAME LONGITUDE LIMITS AS
THE ONE IN ROW "N".

DVAR : VECTOR OF DEGREE VARIANCES (PRECOMPUTED)
RINS : VECTOR OF INTEGRAL OF LEGENDRE FUNCTIONS, S. ROW
(PRECOMPUTED)
RINN : VECTOR OF INTEGRALS OF LEGENDRE FUNCTIONS, N. ROW
(PRECOMPUTED)
MAX : MAXIMUM DEGREE CONSIDERED IN THE EXPANSION
OF THE COVARIANCE FUNCTION.
BLOCK : SIZE OF BLOCKS IN THE GRID (DEGREES).
LB : NUMBER OF CONTIGUOUS BLOCKS IN ROW "S",
DLATN : SOUTHWEST LAT. IN ROW "N" (NORTH) (DEGREES)
DLATN : SOUTHWEST LAT. IN ROW "S" (SOUTH) (DEGREES)
COVS : VECTOR OF COMPUTED COVARIANCES BETWEEN BLOCK MEANS
F : VECTOR OF FOURIER COSINE COEFFICIENTS OF THE
TRIGONOMETRIC EXPANSION OF THE BLOCK MEANS' COVARIANCE FUNCTION.
K : INTEGER#4 CONSTANT SET TO 0 BEFORE, FIRST CALL TO
THIS SUBROUTINE.

PROGRAMMED BY OSCAR L. COLONIO, DEPT. OF GEODETIC SC.,
OHIO STATE UNIVERSITY, COLUMBUS, DECEMBER 1978 . 
********** 

IMPLICIT REAL*4(A-H,O-Z)
DIMENSION DVAR(1),RINS(1),RINN(1),COVS(1),F(1),FF(400),IDDM(400),C

2 (400)
IF(K,NE.0) GO TO 10
20
RPI = RPI*1
40
I = 3.14159265358979326
60
IDDD = F/1400.0
80
NB = 360.0/IDDD +0.000000010
90
RHI = RHI/2
110
RPIH = RPI +1

BLK = BLOCK+IDDD
100
DO 1 = 2, NP1
120
FP(1) = 0.20*DCOS((1-1)*BLK)/(1-1)**2
2
DO 2 = 1, NP1
140
2 DVAR(1) = DVAR(1)/2.0*DO1-1.0
160
FP(1) = BLK**2
180
DO 3 = 1, NP1
200
3 IDDM(1) = ((1-1)*1+1)/2
220
DO 4 = 1, NB
240
4 CX(1) = DCOS(BLK*(1-1))
260
K = 1
280
10 BLIS = DLATN+IDDD
300
BLIS = DLATP+IDDD
320
BLIS = DLATN+BLK
340
BLIS = BL2P+BLK
360
FP = 1.00/((DSIN(RHIN)-DSIN(RLIS)) *BLK)*((DSIN(RL2N)-DSIN(RL2S))
2 )*BLK)

COMPUTE DE COSINE FOURIER COEFFICIENTS OF THE BLOCK

C
C COVARIANCE FUNCTION.
C
0027 RA = 0.D0
0028 DO 20 MI = 1,NP1
0029 NM = M1
0030 DO 10 N1 = M1,NP1
0031 IX = IDB(M1)+M1
0032 10 F(M1) = F(M1)+RINS(IX)*RINS(IX)*DVAR(M1)
0033 F(M1) = F(M1)+FF(M1)*FR
0035 FZ = F(M1)**2
0036 RA = RA+FZ
0037 RAG = RA*1.D-12
0038 IF(M1.GT.1.AND.F(M1-1)**2.LT.RAG.AND.FZ.LT.RAG) GO TO 21
0039 20 CONTINUE
C
C FIND THE COVARIANCES BETWEEN BLOCKS BY FOURIER SYNTHESIS.
C
0040 21 DO 40 I = 1,NBHP
0041 COVS(1) = 0.D0
0042 IF(1.GT.1B) GO TO 40
0043 CMH = C(I)
0044 C2 = 2.D0*C(I)
0045 CN = 1.D0
0046 COVS(1) = F(I)
0047 DO 30 N1 = 2,NXX
0048 CNF1 = C2*CMH-CMH
0049 CMH = CN
0050 CN = CNF1
0051 30 COVS(I) = COVS(I)+CNF1*F(N1)
0052 40 CONTINUE
0053 DO 43 I = 1,NBH
0054 DOVS(NB-1+I) = COVS(I+1)
0055 43 CONTINUE
0056 RETURN
0057 END