A discussion of the use of spherical approximation or no approximation in gravity field modeling with emphasis on unsolved problems in Least-Squares Collocation.

C.C.Tscherning
Department of Geophysics, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen Ø, Denmark

Abstract: Spherical approximation is widely used in gravity field modeling, but it causes errors of the order of the flattening. Different methods for avoiding using this approximation when linearizing observation functionals are possible due to our improved knowledge of the gravity field and of the surface of the Earth. However a number of problems are not solved today, such as those related to the use of a Bjerhammar sphere in Least-Squares Collocation (LSC). If this concept is used, the associated homogeneous /isotropic kernels in general use will be strongly non-homogeneous on the ellipsoid. A possible solution is to use a "Bjerhammar" ellipsoid inside the Earth.

1. Introduction.

Spherical approximation is widely used in gravity field modeling. The best example is the use of the formula

$$\Delta \mathbf{g} = \mathbf{g}(\mathbf{P}) - \gamma(\mathbf{Q}) = -\frac{\mathbf{dT}}{\mathbf{dn}} - \frac{\partial \gamma}{\partial \mathbf{n}}\frac{\mathbf{T}}{\gamma} \approx -\frac{\mathbf{dT}}{\mathbf{dr}} - \frac{2}{\mathbf{r}}\mathbf{T} \quad (1)$$

where Δg is the gravity anomaly, g is gravity measured in the point P and γ is normal gravity in a point Q on the same ellipsoidal normal as P and where the ellipsoidal height of Q is equal to the orthometric height of P. T is the anomalous gravity potential and r is distance to the origin. The derivative in the direction of the gravity vector, $\frac{\partial}{\partial \mathbf{n}}$ is substituted by the radial derivative, the normal gravity is substituted by GM/r^2 and the derivative of gravity with $-2*GM/r^3$ see e.g. Heiskanen and Moritz (1967, section 2-14).

This equation is also the basis for the use of the spherical Stokes equation, solving the boundary problem for gravity data $\Delta \mathbf{g}^*$ on a mean earth sphere with radius R,

$$\mathbf{T}(\mathbf{P}) = \frac{\mathbf{R}}{4\pi}\int_{\sigma} \mathbf{S}(\psi, \mathbf{r})\Delta \mathbf{g}^* \mathbf{d}\sigma \quad (2)$$

where S is Stokes kernel, r is the radial distance of P from the origin and ψ is the spherical distance between P and the integration point on the mean earth surface σ .

Note that there are two approximations in use: the first relating the gravity anomaly to the anomalous potential and the second when using the spherical Stokes equation. A further, and most severe problem, is naturally that we have to estimate gravity anomalies $\Delta \mathbf{g}^*$ on the bounding sphere.

The use of these approximations causes errors at least of the order of the flattening. For local applications this seems to be admissible especially if so-called remove-restore methods are used, see e.g. Torge (2001, p. 286).

Earlier, spherical approximation was used when computing spherical harmonic coefficients from mean surface gravity anomalies. Here the use of an ellipsoidal surface of integration and of ellipsoidal harmonic functions has removed most of the error.

However when treating data from the new gravity missions, where data of very high precision is collected and used, we must try to avoid any approximation. In the following we will discuss how we in many cases may be able to avoid using spherical or even ellipsoidal approximation.

The use of better approximations of the observation functionals due to an improved knowledge of positions of points and of the surface of the Earth are discussed in section 2 and 3.

Spherical approximation is also used in the sense that the surface to which a boundary-value problem is referenced is a sphere. Here, the use of an ellipsoidal surface or of the true Earth surface (Meissl, 1981) will possibly give improvements, which are not discussed here.

The use of an ellipsoidal boundary leads to rather complicated numerical procedures, which makes the use of certain methods like least-squares collocation an extremely heavy task. This is discussed in section 4 and 5.

In this paper, I present problems to which I do not have a solution. Other current problems, which I consider important, are listed at http://www.gfy.ku.dk/~cct/topics_for_research.htm .

2. Using better approximations for the observation functionals.

Spherical approximation is characterized by two approximations: The position of a point in geodetic coordinates, geocentric latitude $\bar{\varphi}$, longitude λ and radial distance r, is put equal to geodetic latitude φ, longitude (unchanged) λ, and r equal to R+h, where R is the mean radius of the Earth (6371 km) and h is the ellipsoidal height. Frequently, one further approximation is made, where h is put equal to the orthometric or normal height.

A second approximation is that the reference ellipsoid is replaced by a sphere with radius R as described in the introduction. Furthermore, observation functionals are used in spherical approximation. Consequently, several approximations are used, so the usual estimate of the error as being of the order of the flattening may not be correct, see e.g. Sansò and Tscherning (2002).

The use of spherical approximation was appropriate earlier, because many other errors (due to data distribution and quality) were also playing a role. But today with modern computers, there is no excuse for using spherical approximation if it is not an absolute necessity. Also, the surface of the Earth is now well known (due to InSAR), and so is the direction of the vertical.

A first improvement is to use "ellipsoidal" approximation; see e.g. Ardalan and Grafarend (2001). We may however get even closer to using no approximation at all.

Using a high degree and order gravity model, contingently enhanced using a regional model, it is possible to come close to making no approximation at all. The gravity vector calculated from the gravity field model must be used when linearizing observation functionals like those associated with the gravity anomaly and the deflection of the vertical. The ellipsoidal height of a point of evaluation must be calculated as the sum of the orthometric height plus the height anomaly calculated using the model. The main point is that the anomalous potential T is now the potential W minus a spherical harmonic model contingently enhanced with a regional model based on local data and topographic information. The lineraized observation equation for a gravity anomaly, eq. (1) then becomes

$$\Delta g = -\frac{\partial T}{\partial n} - \frac{1}{\gamma}\frac{\partial \gamma}{\partial n} T \quad (3)$$

where n is the direction of the gravity vector computed from the model and γ is the magnitude of this gravity vector. The point of evaluation of the partial derivatives and of the gravity vector is a point with altitude computed as the sum of the orthometric height and the height anomaly computed from the model.

3. Consequences for mean anomalies.

Mean gravity and mean height anomalies are generally computed from a set of points interpolated to form a regular grid or calculated using least-squares collocation. The related functional has generally been associated with a surface parallel to the mean earth sphere at the mean orthometric height of the block over which the mean value is defined.

Here we have to be much more precise. However, there is today no generally accepted definition of a mean value. One alternative would be to compute the mean value over the surface with a fixed ellipsoidal height. (This height should be equal to or larger than the maximum ellipsoidal height in the area).

4. Spherical approximation in implementations of Least-Squares Collocation.

The downward continuation needed when applying Stokes formula eq. (2) may be avoided using different approximation methods such as e.g. LSC in a space of functions harmonic down to a surface inside the Earth. Generally a so-called Bjerhammar sphere is used. Hence LSC has the potential of avoiding completely the use of spherical approximation.

Spherical approximation has to a certain extent been used in the GRAVSOFT (Tscherning et al., 1992) implementations of Least-Squares Collocation. In new versions of the software, the use of spherical approximation is optional.

The solutions are expressed as a linear combination of base functions harmonic outside a Bjerhammar sphere totally enclosed in the Earth.

$$\tilde{\mathbf{T}}(\mathbf{P}) = \sum_{\mathbf{i}=1}^{\mathbf{N}} \mathbf{a}_{\mathbf{i}} \mathbf{K}(\mathbf{L}_{\mathbf{i}}, \mathbf{P}) \quad (4)$$

where a_i are constants and K is a reproducing kernel (or covariance function) evaluated in P and with respect to the observation functionals L_i. K is rotationally invariant and homogeneous on spheres concentric with the Bjerhammar sphere, such as on the Earth's surface in spherical approximation, cf. eq. (5).

The use of a mean earth sphere as an approximation of the surface of the Earth enables the representation of a large class of kernels using simple closed expressions, see e.g. Tscherning and Rapp (1974).

If no approximation is used, the rotational invariance and homogeneity is lost. This gives problems in global applications, since the gravity anomaly variances calculated from such a kernel may have a large variation between the Equator and the Poles. As a numerical example we have used a typical kernel

$$K(P,Q) = K(r,r',\psi) = \sum_{i=2}^{\infty} \sigma_i \left(\frac{R^2}{(rr')}\right)^{2i+2} P_i(\cos\psi) \quad (5)$$

where σ_i is equal to the error-degree variances of EGM96 (Lemoine et al., 1998) for i < 361, equal to the degree-variances of GPM98 (Wenzel, 1998) for i> 360 and i < 721 and $\sigma_i = \frac{A}{(i-1)(i-2)(i+4)}$ for i > 720. R = 6349.8 km is the radius of the Bjerhammer sphere and the associated variance of gravity anomalies at the surface of the Earth is 123.4 $mgal^2$ for R=6371000. r and r' are the radial distances of P and Q and ψ is the spherical distance. This kernel was used in numerical experiments reported in Sansò and Tscherning (2002).

In Table 1, values of variances for different functionals are given at different distances r from the origin. The functionals are gravity anomalies, Δg, height anomalies, ς, and vertical gravity gradients, $\mathbf{T_{zz}}$, and their corresponding mean-values over a 0.5° x 0.5° equal area block.

r (km)	Δg	$\overline{\Delta \mathbf{g}}$	ς	$\bar{\varsigma}$	$\mathbf{T_{zz}}$	$\overline{\mathbf{T_{zz}}}$
	$mgal^2$	$mgal^2$	m^2	m^2	EU^2	EU^2
6381	48	31	0.091	0.081	10.8	4.1
6376	75	43	0.112	0.096	21.0	6.8
6371	123	62	0.140	0.116	43.1	11.5
6366	217	93	0.181	0.142	93.6	20.3
6361	409	143	0.246	0.178	214.2	37.2

Table 1. (Signal) variances of different functionals for different distances from the origin, r, corresponding to points on the ellipsoid with latitude varying from 0° to 90° . Mean values are indicated with an over-bar. Note that the mean-values show (as expected) a much smaller dependence on r than the point values. Mean values were evaluated as point values at 5 km altitude.

The variance of the gravity gradients were also calculated at the radial distances (m) r = 340000, 345000, 350000, 355000 and 360000. The variance was in all cases close to 0.00005 EU^2.

If the linear functionals used in eq. (4) are satellite data, and the point of evaluation is on the surface of the Earth, we are implicitly carrying out a downward continuation. The numerical values in Table 1 then show that the results obtained when using LSC for downward continuation will depend on the latitude. One possible method to solve this problem is to use non-isotropic covariance functions or covariance functions based on ellipsoidal harmonics.

The reason for the large variations of the variances is that we use an infinite dimensional space, i.e. the kernels are an infinite series. If we break off the series at a certain, not too low degree, (e.g. i=720), the variances are not changing very much. This may be justified if data at satellite altitude are used to compute spherical harmonic coefficients. However the error-estimates will then not reflect the effect of this. The so-called omission error will be too small.

It would be worthwhile to investigate the magnitude of this type of error.

The use of kernels in a finite dimensional space (degree 720) has been used in numerical experiments with LSC, see Sansò and Tscherning (2002). Data were generated from EGM96, so as to be able to compare estimated spherical harmonic coefficients to the original values. The result of using spherical approximation showed an error much larger than the flattening, up to 10 %. In this experiment, two approximations were used, where in normal practice only one occurs.

In spherical approximation we used r = R+h. The generated gravity anomalies and vertical gravity gradients were evaluated at the correct radial distance, but the derivative was taken in the direction of the radius vector. (In fact the derivative was calculated using the coefficients of degree i multiplied by (i-1) /r). For real observations the derivative is in the direction approximately orthogonal to the bounding surface. This may be the reason why errors of 10 % magnitude normally not are found when using spherical approximation. However, a better understanding of the errors introduced when using spherical approximation is needed.

5. Kernels or covariance functions homogeneous on an ellipsoid.

Kernels/covariance functions may be constructed as product sums of ellipsoidal harmonics.

Let E be the excentric anomaly of a "Bjerhammar" ellipsoid with semi minor axis b, smaller than or equal to the semi-minor axis of the Earth. A point P has the ellipsoidal coordinates β (reduced latitude) and u (the semi-major axis of the confocal ellipse on which the point P is located). The Legendre functions of the first kind are denoted P_{ij} and those of the second kind are denoted Q_{ij}, where i is the degree and j the order. Then a general expression for a positive definite kernel in ellipsoidal harmonics is

$$K(P,Q) = K(u,\beta,\lambda,u',\beta',\lambda') = \sum_{i=2}^{\infty}\sum_{j=0}^{i} \frac{Q_{ij}(i\frac{u}{E})Q_{ij}(i\frac{u'}{E})}{Q_{ij}(i\frac{b}{E})^2} \bullet P_{ij}(\sin\beta)P_{ij}(\sin\beta') \bullet (a_{ij}\cos(j\lambda)\cos(j\lambda') + b_{ij}\sin(j\lambda)\sin(j\lambda')) \quad (6)$$

where a_{ij} and b_{ij} are positive constants. (The "i" occurring in the argument of the associated Legendre function of second order is the imaginary unit).

If the constants all are equal to σ_i for each degree i we have with ψ the "ellipsoidal" distance and u = b

$$K(P,Q) = K(b,b,\psi) = \sum_{i=2}^{\infty}\sum_{j=0}^{i} \frac{\sigma_i}{2i+1} P_{ij}(\cos\psi) \quad .(7)$$

On the "Bjerhammar" ellipsoid we may find closed expressions similar to those found in Tscherning and Rapp (1974) .

Unfortunately general closed expressions have not yet been found for reproducing kernels in infinite dimensional spaces having such functions as kernels. It may however be possible to express the ratios of Legendre function of the second kind as polynomials in u and u' , leading to expressions which are sums of closed expressions.

The use of the ellipsoidal kernels is attractive, but further developments such as the derivations of expressions for the functionals associated with the gradients and the second order derivatives are needed.

6. Conclusion.

Using the known position of points of evaluation and a high resolution gravity field model, it is possible to arrive to virtually approximation-free linearised functionals representing the ground data we have. The geodetic points on the bounding surface, the surface of the Earth, are also known with a very small error.

There are a number of problems associated with abandoning spherical approximation. Here the use of an ellipsoidal "Bjerhammar" surface may be a partial solution for the problems occurring in LSC.

There are other areas where even a planar approximation is used. Such approximations may certainly be valid for smaller areas. But what is the limit? And what are the limits of using spherical or ellipsoidal approximation? Numerical studies may tell us something, but theoretical investigations may be a shortcut to obtaining a better handle on the problems.

Approximations used in geodesy have saved a lot of computer time. But now computers are so fast than one may ask: Why do it wrong, when it can be done right ?

References.

Ardalan, A.A. and E.W.Grafarend: Ellipsoidal geoid undulations (ellipsoidal Bruns formulae): case studies. Journal of Geodesy, Vol. 75, pp. 544-552, 2001.

Heiskanen, W.A. and H. Moritz: Physical Geodesy. W.H. Freeman & Co, San Francisco, 1967.

Lemoine, F.G., S.C. Kenyon, J.K. Factor, R.G. Trimmer, N.K. Pavlis, D.S. Chinn, C.M. Cox, S.M. Klosko, S.B. Luthcke, M.H. Torrence, Y.M. Wang, R.G. Williamson, E.C. Pavlis, R.H. Rapp, and T.R. Olson, The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96, NASA/TP-1998-206861, Goddard Space Flight Center, Greenbelt, MD, July, 1998.

Meissl, P.: The use of finite elements in physical geodesy. Reports of the Dep. of Geodetic Scoence, no. 313, The Ohio State University, 1981.

Sanso' , F. and C.C.Tscherning: Fast spherical collocation: A general implementation. IAG Symposia Vol. 125, pp. 131-137, Springer Verlag, 2002.

Torge, W.: Geodesy. 3. Ed., de Gruyter, 2001.

Tscherning, C.C., R.Forsberg and P.Knudsen: The GRAVSOFT package for geoid determination. Proc. 1. Continental Workshop on the Geoid in Europe, Prague, May 1992, pp. 327-334, Research Institute of Geodesy, Topography and Cartography, Prague, 1992.

Tscherning, C.C. and R.H.Rapp: Closed Covariance Expressions for Gravity Anomalies, Geoid Undulations, and Deflections of the Vertical Implied by Anomaly Degree-Variance Models. Reports of the Department of Geodetic Science No. 208, The Ohio State University, Columbus, Ohio, 1974.

Wenzel, H.G.: Ultra hochaufloesende Kugelfunktionsmodelle GMP98A und GMP98B des Erdschwerefeldes. Proceedings Geodaetische Woche, Kauserslautern, 1998.