



Change of boundary: towards a mathematical foundation of global gravity models

Fernando Sansò¹ · Blažej Bucha²

Received: 8 July 2022 / Accepted: 8 May 2023 / Published online: 26 June 2023
© The Author(s) 2023

Abstract

Change of boundary is a method that iteratively downward continues data from a star-shaped boundary to a regular surface, such as the sphere or the ellipsoid of revolution, and then solves a boundary value problem using the downward continued data on the changed boundary. Although the method belongs to concepts involved in the development of recent high-degree Earth's gravity models, it is still unclear whether the iterations converge to the true solution for boundaries as complex as, for instance, the Earth's surface. In this paper, we revisit the method and show that convergence in terms of the Cesaro limit can be achieved under the assumption that the operator performing the iterations is non-expansive. The validity of the assumption is, however, still not proved. Therefore, we examine the hypothesis numerically using boundaries of various complexity. We start with a simple synthetic topography defined by a Legendre polynomial and move to more realistic finite-degree shapes of the asteroids Bennu and Eros. The numerical experiments indicate that the assumption is valid as long as the boundary deviates not too far from a sphere and the truncation degree of the gravity model is not too high (the experiments with the synthetic topography and Bennu). Otherwise, the hypothesis seems to be false (the Eros case). Finding an analytical condition to separate between shapes for which the change of boundary method converges/diverges remains an open issue.

Keywords Geodetic boundary value problem · Change of boundary · Gravity field · Spherical harmonics · Downward continuation

1 Introduction and notation

The determination and use of global gravity models has been the story of a success of physical geodesy in the last decades. A global model of the external anomalous potential of the Earth $T(r, \sigma)$, in this context, is just a harmonic function represented by a truncated series of solid spherical harmonics (Pavlis 2013; Sansò 2013b)

$$u(r, \sigma) = \sum_{n=2}^N \sum_{m=-n}^n u_{nm} \left(\frac{R}{r}\right)^{n+1} Y_{nm}(\sigma), \quad (1)$$

✉ Blažej Bucha
blazej.bucha@stuba.sk

Fernando Sansò
fernandosanso060545@gmail.com

¹ Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy

² Department of Theoretical Geodesy and Geoinformatics, Slovak University of Technology in Bratislava, Radlinského 11, 81005 Bratislava, Slovak Republic

where r is the spherical radius and

$$\sigma = (\vartheta, \lambda) \quad (2)$$

is a pair of spherical angles, colatitude and longitude. $Y_{nm}(\sigma)$ are fully-normalized real spherical harmonics and u_{nm} are spherical harmonic coefficients. N is the maximum (or truncation) degree of the model. R is in principle any radius, essentially defining the length unit of the problem; in fact, it is obvious that, under the transformation

$$R \rightarrow R'; \quad u_{nm}(R) \rightarrow u'_{nm}(R') = \left(\frac{R}{R'}\right)^{n+1} u_{nm}(R), \quad (3)$$

$u(r, \sigma)$ is left invariant.

In the literature, we find the choice $R = a$, the equatorial radius of the ellipsoid, or $R =$ mean radius of the Earth. In this paper, we will use for R any radius of a sphere S_R , internal and separated from the Earth surface $S \equiv \{r = R_\sigma\}$, that we assume to be a star-shaped and Lipschitz surface. Let us notice here too, that we will use the notation Ω for the domain exterior to S . Furthermore, the notation R_σ , which

is a shorthand for a function of σ , should not be confused with R , which is the constant radius of a sphere.

If we put

$$R_0 = \min_{\sigma} R_{\sigma}, \quad \bar{R} = \max_{\sigma} R_{\sigma} \quad (4)$$

(i.e. the sup of the Bjerhammar radius and inf of the Brillouin radius), we will assume that

$$R < R_0. \quad (5)$$

The reason why R is so loosely defined is that, with the up and down transformation (3), we leave $u(r, \sigma)$ invariant and what really matters is that the global model (1) needs to be a “good” approximation of the actual anomalous potential.

On this point one has to be very clear. What we are looking for is a theory that certifies that our global potential (1) is a good approximation of the anomalous potential $T(r, \sigma)$ in Ω (or in some cases on S , too), in the sense that in this domain

$$\lim_{N \rightarrow \infty} u(r, \sigma) = T(r, \sigma) \quad (6)$$

in some topology for a space of functions harmonic in Ω .

In this paper, we will use exclusively the harmonic Hardy space HL^2 with norm defined by

$$\|u\|_{\text{HL}^2}^2 = \lim_{\varepsilon \rightarrow 0} \int_{\sigma} |u(R_{\sigma} + \varepsilon, \sigma)|^2 d\sigma. \quad (7)$$

Let us note that we find it convenient to use in the paper a normalized measure on the unit sphere, that we will denote

$$d\sigma = \frac{1}{4\pi} \sin \vartheta \, d\vartheta \, d\lambda. \quad (8)$$

Fatou’s lemma and Cimmino’s theory (Axler et al. 1992 revisited in Sansò and Sideris 2017 and Sansò and Venuti 2001) guarantee that the limit (7) exists and in fact HL^2 is in a unitary correspondence with $L^2(\sigma)$, namely the Hilbert space of functions of σ with norm

$$\|u\|^2 = \int_{\sigma} |u(R_{\sigma}, \sigma)|^2 d\sigma, \quad (9)$$

of course on condition that (7) be bounded.

The choice of this topology is because it is strongly set in geodetic tradition and, when discretized, it leads to least squares approximation, which is a universal tool in the arsenal of geodetic methods. In addition, the theory of (linearized) geodetic boundary value problems has shown that once a datum of free air gravity anomaly or disturbance is given on S , belonging to $L^2(\sigma)$, then a solution (or anomalous potential T) exists, is unique and continuously dependent on

the data in $L^2(\sigma)$ norm (Sansò 2013a). This fact together with a theorem of completeness of the base functions

$$\varphi_{nm}(\sigma) = \{q_{\sigma}^{n+1} Y_{nm}(\sigma)\}, \quad (10)$$

where

$$q_{\sigma} \equiv \frac{R}{R_{\sigma}}, \quad (11)$$

which is a particular case of Runge–Krarup theorem (Krarup 1975; Sansò 2013c), says that the problem of the approximation of $T(r, \sigma)$ by a global potential of the form (1) is well-posed in $L^2(\sigma)$.

Yet the way in which the most important global models at high resolution ($N \gtrsim 2159$) are actually computed (see for instance Pavlis 2013) is not by the use of least squares directly applied to the original data, but rather the data are moved (“downward continued”) from the surface down to the ellipsoid, which is then perturbatively approximated by a sphere, and finally the global model is computed as the solution of the boundary value problem from this easy to treat “changed boundary”.

Residuals are then computed at the level of S , where data are given, and the process is iterated. Indeed the real way in which a global model is computed includes a number of very important intermediate steps, including the transformation of marine altimetric data into gravity anomalies and the assimilation of satellite only gravity models (see Pavlis 2013; Rapp 1993).

In particular, this last step has a large influence in the estimation of harmonic coefficients up to some intermediate degree (a few hundreds) and has a fundamental role in the correction of biases present in the ground data, due to the non-homogeneous historical measurement process (height datum problem; Sansò and Venuti 2001; Rapp 1993).

Yet for the high-frequency (degree) coefficients the determination process is essentially outlined above. Therefore, the question arises on how good is the result obtained. No doubt the answer has to be given by testing the estimated field on independent data or by cross-validation.

This has been quite successfully done (e.g., Pavlis et al 2012) and therefore it is clear, on a scientific ground, that a mathematical proof of convergence of the iterations is needed to explain/understand such a success. This is how the idea of the change of boundary (CoB) theory was born. Up to now the results obtained are not complete, because of a hard obstacle in proving that a certain operator is non-expansive. Some steps forward have been done by conjecturing that this assumption is correct, based on a perturbative argument (Sansò and Sideris 2017); yet a rigorous proof is still lacking.

In such a situation it was deemed useful by the authors to set up controlled simulations to test at least numerically the

validity of the conjecture. This is what is done in this paper, together with some mathematical refinements, showing the results that can be obtained if the mentioned assumption is correct.

In particular a verifiable sufficient condition is given, to guarantee the convergence if we pose an a priori hypothesis that in real world, or better in our model world, both the anomalous potential and the shape of S are expressed by spherical developments with known maximum degree. Therefore, we want to stress again that the purpose of the paper is not to propose CoB as a new method to estimate global models, but rather to set up a formal frame in which to understand the success of existing methods.

Concluding this section, it is worth mentioning that the item of convergence of global models has been tackled by several authors, also outside the geodetic literature (e.g., Takahashi and Scheeres 2014; Hu and Jekeli 2015; Reimond and Baur 2016; Sebera et al. 2016; Hirt and Kuhn 2017). However, the purpose of this paper is slightly different in that we try to assess the extent at which our present theoretical understanding of the problem is reliable; in particular this will be done by examining the norm of a certain matrix operator, rather than following the sequence of models of increasing degree.

2 Foundation of CoB

In order to set up the formal definition of CoB we need to recall first of all that the linearized geodetic boundary value problem amounts to determining $T(r, \sigma)$ in Ω given, depending on what we consider as known on the boundary,

$$\delta g = -\frac{\partial T}{\partial h} \quad \text{or} \quad \Delta g = -\frac{\partial T}{\partial h} + \frac{\gamma'}{\gamma} T \quad \left(\gamma' = \frac{\partial \gamma}{\partial h} \right) \quad (12)$$

on $S \equiv \{r = R_\sigma\}$; here h is as usual the ellipsoidal height, γ the modulus of the normal gravity vector, δg the gravity disturbance and Δg the gravity anomaly. Let us recall that the CoB has been first defined in Sansò (1993) and its theory more carefully analysed in Sansò and Sideris (2017).

The spherical approximation of (12) (see Sansò and Sideris 2017) is just the so-called *simple Molodensky problem*

$$\delta g \cong -\frac{\partial T}{\partial r} \Big|_S, \quad \Delta g \cong -\frac{\partial T}{\partial r} \Big|_S - \frac{2}{r} T \Big|_S \quad (13)$$

and its solution is the first step for the solution of (12). The point is that in either cases (13) the boundary value problem can be reduced to the solution of a Dirichlet problem; in fact,

one has that

$$u = -r \frac{\partial T}{\partial r} \quad \text{or} \quad u = -r \frac{\partial T}{\partial r} - 2T \quad (14)$$

have to be harmonic too (Heiskanen and Moritz 1967; Sansò and Sideris 2017) in Ω and they satisfy one or the other boundary condition

$$-r \frac{\partial T}{\partial r} \Big|_S = u(R_\sigma, \sigma) = f(\sigma) = R_\sigma \delta g(\sigma), \quad (15)$$

$$-r \frac{\partial T}{\partial r} \Big|_S - 2T \Big|_S = u(R_\sigma, \sigma) = g(\sigma) = R_\sigma \Delta g(\sigma). \quad (16)$$

For the sake of simplicity let us choose one of the two, for instance the first that has no singularity at degree 1, as it happens to the second.

So the problem of finding T harmonic in $\Omega \equiv \{r > R_\sigma\}$, satisfying the boundary condition (15), is equivalent to searching for the solution $u(r, \sigma)$ of the Dirichlet problem with boundary values (15), and then retrieving $T(r, \sigma)$ by integration of $-u(r, \sigma)/r$ along the radius (Sansò 2013a). The geodetic practice is then to “lower” $f(\sigma)$ from S to S_R ; this step can be done in different ways, for instance by collocation or simply by a truncated Taylor series (Martinec 1998; Pavlis 2013). Whatever it is, we call PB (pull back) the operator that lowers $f(\sigma)$ to S_R . After that the Dirichlet problem is solved for the pull back $PB f$ by Poisson formula and the residuals between f and the upward continuation of $PB f$ are computed. Would the residuals be zero, we could say that the problem has been solved. Since this is not the case, we will see that it cannot be, the same procedure is applied to the residuals and the same approach is repeated iteratively.

Then all contributions from the iterations are added to find hopefully a series convergent, in some sense, to the correct solution. The situation is formalized in the scheme of Fig. 1 where PB represents the pull back operator, Π denotes the Poisson operator, harmonically propagating the boundary values on the sphere S_R to the outer domain Ω_R (i.e. the exterior of the sphere S_R) and Γ is the trace on S operator, namely given that $S \equiv \{r = R_\sigma\}$, Γ is the operator that, for any continuous $u(r, \sigma)$ defined in Ω_R , computes

$$\Gamma u = u(R_\sigma, \sigma). \quad (17)$$

Furthermore, we call U the upward continuation operator

$$U = \Gamma \Pi \quad (18)$$

or, more precisely (Heiskanen and Moritz 1967)

$$\forall f(\sigma), \quad Uf \equiv \int_\sigma \frac{R(R_\sigma^2 - R^2)}{\ell_{\sigma\sigma'}^3} f(\sigma') d\sigma' \quad (19)$$

$$S_R \equiv (r = R) \quad \Omega_R \equiv (r > R) \quad S \equiv \{r = R_\sigma\} \quad \Omega \equiv \{r > R_\sigma\}$$

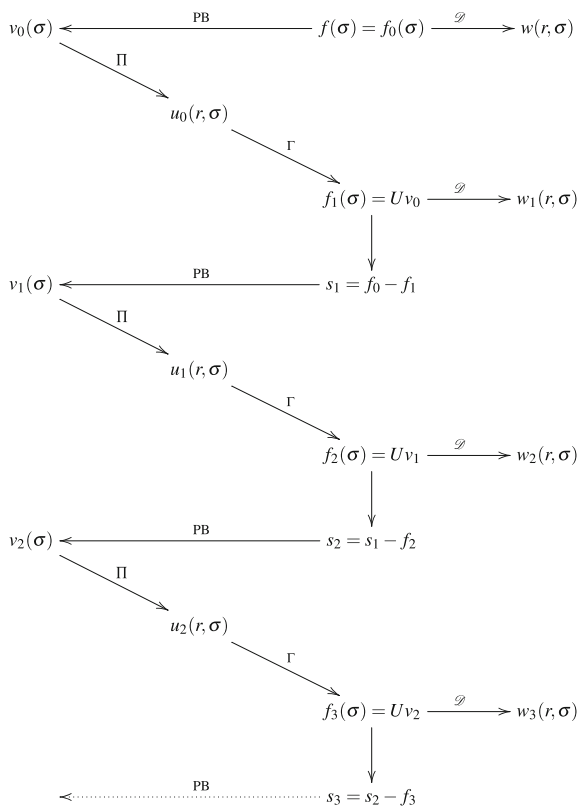


Fig. 1 The iterative scheme of CoB. The first line gives the sets on which functions in the same columns are defined

$$(\ell_{\sigma\sigma'} = [R^2 + R_\sigma^2 - 2 R R_\sigma \cos \psi_{\sigma\sigma'}]^{1/2},$$

$\psi_{\sigma\sigma'}$ = spherical angle between σ and σ').

Notice that the Poisson formula (19) is written without the factor $(4\pi)^{-1}$ which is included by definition in the area element $d\sigma'$ (see 8). Since S is star-shaped by hypothesis, U transforms functions of σ (i.e. defined on the unit sphere) into functions again defined on the unit sphere. Let us further observe that, if $S \subset \Omega_R$ and

$$R_0 = \inf_{\sigma} R_\sigma > R, \tag{20}$$

as we assume, any function harmonic in Ω_R is continuous there, so that (19) is well-defined. Moreover, in Fig. 1 we use \mathcal{D} to denote the Dirichlet solver that, given a function $f(\sigma)$ on S , provides the function $w(r, \sigma)$, harmonic in Ω and agreeing with $f(\sigma)$ on S .

The scheme of Fig. 1 implies the following relations

$$v_\ell(\sigma) = \text{PB } s_\ell(\sigma) \quad (\ell = 1, 2, \dots; v_0(\sigma) = \text{PB } f_0(\sigma)), \tag{21}$$

$$s_\ell(\sigma) = s_{\ell-1}(\sigma) - f_\ell(\sigma) \quad (\ell = 2, 3, \dots; s_1(\sigma) = f_0(\sigma) - f_1(\sigma)), \tag{22}$$

$$f_\ell(\sigma) = Uv_{\ell-1}(\sigma) \quad (\ell = 1, 2, \dots). \tag{23}$$

Combining the three relations one gets:

$$s_\ell(\sigma) = s_{\ell-1}(\sigma) - Uv_{\ell-1}(\sigma) = s_{\ell-1}(\sigma) - U \text{PB } s_{\ell-1}(\sigma) = (I - U \text{PB})s_{\ell-1}(\sigma), \tag{24}$$

$$s_1(\sigma) = f_0(\sigma) - f_1(\sigma) = f(\sigma) - Uv_0(\sigma) = (I - U \text{PB})f(\sigma). \tag{25}$$

So, putting

$$K = I - U \text{PB}, \tag{26}$$

one can write

$$s_\ell(\sigma) = K s_{\ell-1}(\sigma), \tag{27}$$

implying, also recalling (25), that

$$s_\ell(\sigma) = K^{\ell-1} s_1(\sigma) = K^\ell f(\sigma). \tag{28}$$

On the same time, iterating (22) we see that

$$\begin{aligned} s_\ell(\sigma) &= s_{\ell-1}(\sigma) - f_\ell(\sigma) \\ &= s_{\ell-2}(\sigma) - f_{\ell-1}(\sigma) - f_\ell(\sigma) = \dots \\ &= f_0(\sigma) - \sum_{k=1}^{\ell} f_k(\sigma) \equiv f(\sigma) - F_\ell(\sigma) \end{aligned}$$

$$(F_\ell(\sigma) = \sum_{k=1}^{\ell} f_k(\sigma)). \tag{29}$$

Therefore, a first conclusion of our construction is that if in some topology we can claim that

$$s_\ell(\sigma) \rightarrow 0, \tag{30}$$

then we have found a sequence $F_\ell(\sigma)$, defined only by applications of PB and U (i.e. Poisson formula), that in the same sense (topology) tends to $f(\sigma)$.

In addition, still looking at Fig. 1, one realizes that

$$w_\ell(r, \sigma) = \mathcal{D} f_\ell(\sigma), \tag{31}$$

which are the solutions of Dirichlet problem agreeing with $f_\ell(\sigma)$ on S , have to be just the restriction of $u_{\ell-1}(r, \sigma)$ to Ω ; in fact, these are harmonic in $\Omega_R > \Omega$ and they are equal to $f_\ell(\sigma)$ on S , by definition of $f_\ell(\sigma)$. Therefore, if the topologies used are such that

$$\begin{aligned} F_\ell(\sigma) \rightarrow f(\sigma) &\Rightarrow W_\ell(r, \sigma) = \sum_{k=1}^{\ell} w_k(r, \sigma) \\ &= \sum_{k=1}^{\ell} \mathcal{D} f_k(\sigma) = \mathcal{D} F_\ell(\sigma) \rightarrow w(r, \sigma) = \mathcal{D} f(\sigma), \end{aligned} \tag{32}$$

we can conclude too that

$$\sum_{k=0}^{\ell-1} \Pi v_k(r, \sigma) = \sum_{k=0}^{\ell-1} u_k(r, \sigma) \rightarrow w(r, \sigma), \quad (r, \sigma) \in \Omega, \tag{33}$$

namely we have found the approximate solution we are looking for.

By expressing $u_k(r, \sigma)$ in terms of solid spherical harmonics, what is natural since they are given by Poisson formula, we will also find convergence theorems for global models.

Remark 1 It is important to underline again that the operators PB and $U = \Gamma \Pi$ (see 19), which are the constituents of the iterative process, can be interpreted as transforming functions of σ only into functions of σ again so they can be thought of as transformations between some functional space into itself. In this paper this space, i.e. the topology in which we embed the problem, is simply $L^2(\sigma)$, i.e. the Hilbert space with norm

$$\|f\|_{L^2(\sigma)}^2 = \int_{\sigma} f^2(\sigma) \, d\sigma. \tag{34}$$

Moreover, with the $L^2(\sigma)$ topology for $F_{\ell}(\sigma) - f(\sigma)$, the property (31), with $W_{\ell}(r, \sigma) \rightarrow w(r, \sigma)$ interpreted as pointwise convergence or uniform convergence on compact subsets of Ω , holds, in force of Runge–Krarup’s theorem in the simplified form as presented on p. 137 of Sansò (2013b) (see also Krarup 1975, [p. 254]).

Before developing our analysis, we want to recall a first choice done for the operator PB:

$$\text{PB} = I. \tag{35}$$

We justify such a position by examining the two sphere example, first presented in Sansò (1993). We are going to see in fact that any choice of PB as a Taylor formula in radial direction has essentially not very different characteristics than (35) and therefore this is the simplest assumption.

Nevertheless, as we will see in Sect. 6, the alternative

$$(h_{\sigma} = R_{\sigma} - R) \text{ PB} = I - h_{\sigma} \frac{\partial}{\partial r} \Big|_S \tag{36}$$

has been explored too, in an attempt to improve our results. Yet we will find that even more complex gravity field models produce results similar to those obtained by next example.

Since it will be used later on, we give also a spectral representation of PB from (36), when acting on a potential harmonic down to S_R , namely

$$\begin{aligned} \text{PB}u &= \sum_{n,m} u_{nm} c_n(\sigma) q_{\sigma}^{n+1} Y_{nm}(\sigma) \\ &= \sum_{n,m} u_{nm} \left[1 + (n+1) \frac{h_{\sigma}}{R_{\sigma}} \right] q_{\sigma}^{n+1} Y_{nm}(\sigma). \end{aligned} \tag{37}$$

Example 1 (Two spheres S_R, S_{R_0}) In this example we examine the case that $S = S_{R_0}$, a sphere of radius $R_0 > R$. This is because, despite the elementary character of this case, we can work it out analytically and read directly how the CoB scheme works.

We put

$$q_0 = \frac{R}{R_0} (< 1), \quad h = R_0 - R, \quad \eta = \frac{h}{R}; \tag{38}$$

let us notice that in this example the three quantities q_0, h, η are constant. We test the CoB iterations with the two choices

$$\text{PB} = I \quad \text{or} \quad \text{PB} = I - h \frac{\partial}{\partial r} \Big|_{S_{R_0}}.$$

In both cases we can write the spectral representation of PB as

$$\text{PB} \sim c_n \quad (f \sim \{f_{nm}\}, \text{PB}f \sim \{c_n f_{nm}\}) \tag{39}$$

with

$$\begin{cases} c_n = 1 & \text{when PB} = I \\ c_n = 1 + \frac{h}{R_0}(n+1) & \text{when PB} = I - h \frac{\partial}{\partial r} \Big|_{S_{R_0}} \end{cases}. \tag{40}$$

Moreover, the spectral representation of U is just

$$U \sim q_0^{n+1} \tag{41}$$

and so

$$K = I - U \text{PB} \sim 1 - c_n q_0^{n+1} = k_n, \tag{42}$$

with c_n one of the two choices in (40). Clearly if $k_n < 1$, the operator K as well as all K^{ℓ} for any ℓ are continuous operators in $L^2(\sigma)$.

Since $q_0 < 1$, it is obvious that $k_n < 1$ when $c_n = 1$, so we have to look into (42) when $c_n = 1 + \frac{h}{R_0}(n+1)$. Considering that

$$\frac{h}{R_0} = \frac{\eta}{1+\eta}, \quad q_0 = \frac{1}{1+\eta}, \tag{43}$$

we easily get, for the second case in (40),

$$c_n q_0^{n+1} = \frac{1 + (n + 2)\eta}{(1 + \eta)^{n+2}} < 1. \tag{44}$$

Summarising, we have for k_n the values

$$k_n = 1 - q_0^{n+1} = \frac{(1 + \eta)^{n+1} - 1}{(1 + \eta)^{n+1}}, \tag{45}$$

corresponding to the first of (40), and

$$k'_n = \frac{(1 + \eta)^{n+2} - [1 + \eta(n + 2)]}{(1 + \eta)^{n+2}}, \tag{46}$$

corresponding to the second. As we see, we have

$$\forall n \quad 0 < k'_n < k_n < 1; \tag{47}$$

moreover

$$\lim_{n \rightarrow +\infty} k_n = \lim_{n \rightarrow +\infty} k'_n = 1. \tag{48}$$

It is not difficult to verify that the same would happen even if we pushed the Taylor expansion in PB to a higher order. Since the convergence of the iterative scheme critically depends on the behaviour of k_n , which are the eigenvalues of K corresponding to the eigenfunctions $Y_{nm}(\sigma)$, and the relation (47) holds, it is enough here to consider the case of k_n (45) in order to prove convergence of the CoB scheme.

With (45) and recalling (28), one has

$$s_\ell(\sigma) = K^\ell f(\sigma) = \sum_{n,m} [1 - q_0^{n+1}]^\ell f_{nm} Y_{nm}(\sigma), \tag{49}$$

i.e.

$$\|s_\ell(\sigma)\|_{L^2(\sigma)}^2 = \sum_{n,m} [1 - q_0^{n+1}]^{2\ell} f_{nm}^2. \tag{50}$$

Since k_n , satisfying (47), are uniformly bounded, one can take the limit for $\ell \rightarrow \infty$ in (50) under the summation sign, obtaining

$$\lim_{\ell \rightarrow \infty} \|s_\ell(\sigma)\|_{L^2(\sigma)}^2 = \sum_{nm} \lim_{\ell \rightarrow \infty} [1 - q_0^{n+1}]^{2\ell} f_{nm}^2 = 0. \tag{51}$$

This proves the sought convergence, namely $F_\ell(\sigma) = \sum_{k=1}^\ell f_k(\sigma)$ tends to $f(\sigma)$ in $L^2(\sigma)$ and, (32) holding true in this case, we also have $W_\ell(r, \sigma) \rightarrow w(r, \sigma)$ in Ω_{R_0} .

Two important comments are in order.

Remark 2 We have to underline that the forms (45), (46) of the eigenvalues of U , together with relations (47), (48) show that

$$\|K\|_{L^2(\sigma)} = \sup k_n = \sup k'_n = 1, \tag{52}$$

because 1, the spectral radius of K , is also equal to its operator norm in $L^2(\sigma)$.

This will be a hard point when we will try to extend this analysis to a general surface S different from a sphere S_{R_0} . On the other hand, if we truncated the operator K to some maximum degree N , then we would have

$$\|K_N\| = \|I - U_N\|_{L^2(\sigma)} = 1 - q_0^{N+1} < 1 \tag{53}$$

and the convergence will immediately descend from the fact that K becomes a contraction. Also this remark will be used in a more general setting.

Remark 3 Considering Example 1, we have to understand that the coefficients $\{f_{nm}\}$ of $f(\sigma)$ refer to a function, the datum, which is given on the surface S_{R_0} . So the convergence of s_ℓ to 0 leading to the convergence of the method is for functions defined on S_{R_0} and $W_\ell(r, \sigma)$ tend to $w(r, \sigma)$ in Ω_{R_0} .

On the other hand, if we call

$$V_\ell(\sigma) = \sum_{k=0}^{\ell-1} v_k(\sigma), \tag{54}$$

i.e. the sum of the residuals seen at the level of S_R (recall that $R < R_0$), we see from (23) that

$$F_\ell(\sigma) = UV_\ell(\sigma). \tag{55}$$

Now the eigenvalues of U are $c_n = q_0^{n+1}$, so they are positive tending to 0 for $n \rightarrow \infty$. It follows that U is a compact operator and then the convergence of $F_\ell(\sigma)$ to $f(\sigma)$ in $L^2(\sigma)$ does not imply that $V_\ell(\sigma)$ is convergent too to some function in $L^2(\sigma)$. Would this be true, we could always downward continue $f(\sigma)$ from S_{R_0} to S_R , which is known to be false.

A final word of caution, before closing the section, is that although the choice $PB = I$ makes the problem more comfortable to analyse, in practice a radial Taylor expansion should always be applied before going to numbers (see, e.g. Sansò and Sideris 2017; Martinec 1998).

3 Properties of the upward continuation operator U

As we will see in the next section, the success of convergence of CoB depends on whether the statement that $\|K\| = 1$ is

correct or not. Whatever we choose as PB, we have

$$K = I - U \text{ PB}, \tag{56}$$

so it is important to get a quite clear understanding of the properties of U .

They are summarized below in a list from a) to e). Since U can be represented as an integral operator, as in (19), most properties of U are just well-known classical properties of such operators, so they will be only formulated with reference to literature. We would like to underline that such a matter is well-known in mathematics since the times of Fredholm, more than one century ago, and is presented in classical old books of Analysis, like Goursat (1964).

(a) U is a compact operator. According to (19), U is an integral operator with kernel

$$U(\sigma, \sigma') = \Pi(R_\sigma, \psi_{\sigma\sigma'}). \tag{57}$$

It is enough to use the identity

$$\begin{aligned} \ell_{\sigma\sigma'}^2 &= (R_\sigma^2 + R^2 - 2 R R_\sigma \cos \psi_{\sigma\sigma'}) \\ &= (R_\sigma - R)^2 + 2 R_\sigma R (1 - \cos \psi_{\sigma\sigma'}) \end{aligned} \tag{58}$$

to realize that

$$\ell_{\sigma\sigma'} \geq R_\sigma - R \geq R_0 - R = h_0. \tag{59}$$

Similarly, one has

$$\ell_{\sigma\sigma'} \leq R_\sigma + R \leq 2\bar{R}, \tag{60}$$

so that the following (very rough) upper and lower bounds hold for $U(\sigma, \sigma')$:

$$U(\sigma, \sigma') \leq \frac{R(R_\sigma - R)(R_\sigma + R)}{(R_\sigma + R)\ell_{\sigma\sigma'}^2} \leq \frac{2R\bar{R}}{h_0^2} = c, \tag{61}$$

$$U(\sigma, \sigma') \geq \frac{R(R_\sigma - R)(R_\sigma + R)}{(R_\sigma + R)^3} \geq \frac{R h_0}{4\bar{R}^2} = c > 0. \tag{62}$$

So $U(\sigma, \sigma')$ is a bounded strictly positive function. Clearly $U(\sigma, \sigma')$ is also continuous in its arguments, in fact Lipschitz continuous, as it is R_σ by hypothesis.

Whence, since the unit sphere is a compact set with measure 1, by the hypothesis done on $d\sigma$, we have a fortiori

$$\iint U^2(\sigma, \sigma') d\sigma d\sigma' \leq c^2 < +\infty. \tag{63}$$

The first conclusion then is that U is a continuous operator $L^2(\sigma) \rightarrow L^2(\sigma)$. Moreover, U is (sequentially)

completely continuous, namely it transforms any weakly $L^2(\sigma)$ -convergent series into a strongly convergent one.

(b) The kernel $U(\sigma, \sigma')$ admits the representation

$$\begin{aligned} U(\sigma, \sigma') &= \sum_{n,m} q_\sigma^{n+1} Y_{nm}(\sigma) Y_{nm}(\sigma') \\ &= \sum_{n=0}^{+\infty} q_\sigma^{n+1} (2n+1) P_n(\cos \psi_{\sigma\sigma'}), \end{aligned} \tag{64}$$

where q_σ is defined by (11) and P_n are the ordinary Legendre polynomials. In fact, (64) is nothing but the well-known series

$$\Pi(r, \psi_{\sigma\sigma'}) = \sum_{n=0}^{+\infty} \sum_{m=-n}^n \left(\frac{R}{r}\right)^{n+1} Y_{nm}(\sigma) Y_{nm}(\sigma') \tag{65}$$

restricted to $S \equiv \{r = R_\sigma\}$.

An important consequence of the representation (64) is that the range of U in $L^2(\sigma)$, $\mathcal{R}(U)$, is dense in this space. In fact, it is clear that

$$\text{span}\{q_\sigma^{n+1} Y_{nm}(\sigma)\} \subset \mathcal{R}(U), \tag{66}$$

the inclusion being dense and, as it is proved in Theorem 9 of Sansò (2013c), the sequence $\{q_\sigma^{n+1} Y_{nm}(\sigma)\}$ is complete in $L^2(\sigma)$. This is as a matter of fact a by-product of Runge–Krarup’s theorem.

(c) The compact operator U has a representation in terms of singular elements $\{\lambda_i, \varphi_i, \xi_i\}$ such that

$$(i) \lambda_i > 0; \lambda_i \rightarrow 0, i \rightarrow \infty, \tag{67}$$

$$(ii) \{\varphi_i\} \text{ and } \{\xi_i\} \text{ are both orthonormal sequences in } L^2(\sigma),$$

$$(iii) U = \sum_{i=1}^{+\infty} \lambda_i \xi_i \otimes \varphi_i, \tag{68}$$

where

$$(\xi_i \otimes \varphi_i)(u) \equiv \xi_i \langle \varphi_i, u \rangle_{L^2(\sigma)}, \tag{69}$$

$$(iv) U^T \equiv \sum_{i=1}^{+\infty} \lambda_i \varphi_i \otimes \xi_i, \tag{70}$$

$$(v) \begin{cases} U\varphi_i = \lambda_i \xi_i \\ U^T \xi_i = \lambda_i \varphi_i \end{cases} \Rightarrow \begin{cases} U^T U \varphi_i = \lambda_i^2 \varphi_i, \\ U U^T \xi_i = \lambda_i^2 \xi_i. \end{cases} \tag{71}$$

This representation is also standard (see Appendix of Sansò and Sideris 2017). Since the kernel of the operator $\xi_i \otimes \varphi_i$ is

$$\xi_i \otimes \varphi_i \sim \xi_i(\sigma) \varphi_i(\sigma') = \omega_i(\sigma, \sigma'), \tag{72}$$

it is clear that the kernel $U(\sigma, \sigma')$ is given by

$$U(\sigma, \sigma') = \sum_{i=1}^{+\infty} \lambda_i \xi_i(\sigma) \varphi_i(\sigma'). \tag{73}$$

To verify that the series (73) is not only formal one can observe that indeed

$$\iint \omega_i(\sigma, \sigma') \omega_k(\sigma, \sigma') d\sigma d\sigma' = \langle \xi_i, \xi_k \rangle \langle \varphi_i, \varphi_k \rangle = \delta_{ik}, \tag{74}$$

i.e. $\{\omega_i(\sigma, \sigma')\}$ is orthonormal in $L^2(\sigma, \sigma')$, then

$$\sum_{i=1}^{+\infty} \lambda_i^2 \equiv \iint U^2(\sigma, \sigma') d\sigma d\sigma' < c^2, \tag{75}$$

which shows that the series is $L^2(\sigma, \sigma')$ convergent. Further on, by using the orthonormality of $\{\xi_i\}$, we can write, $\forall u \in L^2(\sigma)$

$$\begin{aligned} \|Uu\|_{L^2(\sigma)}^2 &= \left\| \sum_{i=1}^{+\infty} \lambda_i \langle u, \varphi_i \rangle_{L^2(\sigma)} \xi_i \right\|_{L^2(\sigma)}^2 \\ &= \sum_{i=1}^{+\infty} \lambda_i^2 \langle u, \varphi_i \rangle^2. \end{aligned} \tag{76}$$

Since also $\{\varphi_i\}$ is orthonormal in $L^2(\sigma)$, one has

$$\sup_{\|u\|_{L^2(\sigma)} \leq 1} \|Uu\|_{L^2(\sigma)} = \sup_i \lambda_i^2 < c^2. \tag{77}$$

Therefore, if we put λ_i in decreasing order, one can state that

$$\|U\|^2 \doteq \lambda_1^2. \tag{78}$$

Finally, we observe that since by (68)

$$\text{span}\{\xi_i\} \subset \mathcal{R}(U), \tag{79}$$

with dense embedding, and since $\mathcal{R}(U)$ is densely contained in $L^2(\sigma)$ by point b), we have that also $\text{span}\{\xi_i\}$ is dense in $L^2(\sigma)$. Since $\{\xi_i\}$ is orthonormal we conclude that it is an orthonormal basis of $L^2(\sigma)$.

(d) U is an invertible operator, namely the equation

$$v = Uu = 0 \tag{80}$$

has only the solution $u = 0$. In fact, consider the potential

$$v(r, \sigma) = \int \Pi(r, \psi_{\sigma, \sigma'}) u(\sigma') d\sigma', \tag{81}$$

which is harmonic in Ω_R . By (80)

$$v(R_\sigma, \sigma) \equiv 0,$$

so that $v \equiv 0$ in Ω . But then, by the unique continuation property, $v \equiv 0$ in Ω_R , too; so

$$u(\sigma) = \lim_{r \searrow R} v(r, \sigma) \equiv 0, \tag{82}$$

as it was to be proved.

We notice that, from (71), no λ_i can be equal to zero; i.e. λ_i is an infinite bounded sequence tending to zero through positive values. Moreover, since

$$u \in \mathcal{N}(U) \Rightarrow u \perp \text{span}\{\varphi_i\} \tag{83}$$

and since by the above $\mathcal{N}(U) = \emptyset$, we conclude that

$$\text{span}\{\varphi_i\} \subset L^2(\sigma) \text{ (densely);} \tag{84}$$

therefore also $\{\varphi_i\}$ is an orthonormal basis of $L^2(\sigma)$.

(e) As a compact operator we know that U has only a point spectrum of eigenvalues $\{\mu_i\}$ such that

$$\mu_i \rightarrow 0. \tag{85}$$

Since μ_i can never be zero, $\{\mu_i\}$ is an infinite sequence. Since $U(\sigma, \sigma')$ is real and non-symmetric, it is clear that in general μ_i are complex and they come in couples, i.e. if μ_i is an eigenvalue, $\bar{\mu}_i$ is an eigenvalue too, unless they are both real. Moreover, we know that the spectral radius ρ of U is also the norm of U , i.e.

$$\|U\| = \sup_i |\mu_i| = \rho(U). \tag{86}$$

Since $\{\mu_i\}$ is a bounded sequence with zero as unique accumulation point it has to be, by suitably choosing μ_1 ,

$$\sup_i |\mu_i| = \max_i |\mu_i| = |\mu_1|. \tag{87}$$

Therefore, there is an eigenfunction $p_1(\sigma)$ such that

$$\mu_1 p_1(\sigma) = \int \Pi(R_\sigma, \psi_{\sigma \sigma'}) p_1(\sigma') d\sigma'. \tag{88}$$

Now (88) tells us that $p_1(\sigma)$ has to be a continuous function, so that we find

$$|\mu_1| |p_1(\sigma)| \leq \int |\Pi(R_\sigma, \psi_{\sigma\sigma'})| d\sigma' M_1 \tag{89}$$

$$(M_1 = \max_\sigma |p_1(\sigma)|).$$

On the other hand, since $\Pi > 0$,

$$\int |\Pi(R_\sigma, \psi_{\sigma\sigma'})| d\sigma' = \int \Pi(R_\sigma, \psi_{\sigma\sigma'}) d\sigma' \tag{90}$$

$$= q_\sigma = \frac{R}{R_\sigma},$$

as it is clear also from (64). Therefore (89) yields

$$|\mu_1| M_1 \leq M_1 \max_\sigma q_\sigma \leq M_1 q_0, \tag{91}$$

namely

$$|\mu_1| = \|U\| \leq q_0 < 1. \tag{92}$$

So under the geometrical hypotheses done, particularly (5), U is a contraction and $K = I - U$ is an invertible operator.

Remark 4 Let us observe that the operator $K = I - U$, entering into our theory with the choice $PB = I$, has indeed eigenvalues $(1 - \mu_i)$ and its spectral radius is

$$\rho(K) = \sup_i |1 - \mu_i|. \tag{93}$$

On the other hand, since $\mu_i \rightarrow 0$ for $i \rightarrow \infty$, we must have

$$\rho(K) \geq 1. \tag{94}$$

Therefore, considering that, $|\mu_i| \leq |\mu_1| = q_0 < 1$, the only possibility to have $\|K\| = 1$ is that the complex values μ_i belong to the filled area in Fig. 2. In particular, any real negative value of μ_i would imply $\|K\| > 1$.

Remark 5 We underline that the abstract representation of U by (68) and of its kernel $U(\sigma, \sigma')$, (73), is once more an old achievement of Analysis, going back to Schmidt and Picard (Goursat 1964). The form (68) in particular provides an explicit expression for U^{-1} , i.e.

$$U^{-1} = \sum_{i=1}^{+\infty} \frac{1}{\lambda_i} \varphi_i \otimes \xi_i; \tag{95}$$

attention has to be paid to the fact that (72) cannot be represented by a kernel in $L^2(\sigma, \sigma')$, because $\lambda_i \rightarrow 0$. In fact U^{-1} , as the inverse of a compact operator, is an unbounded

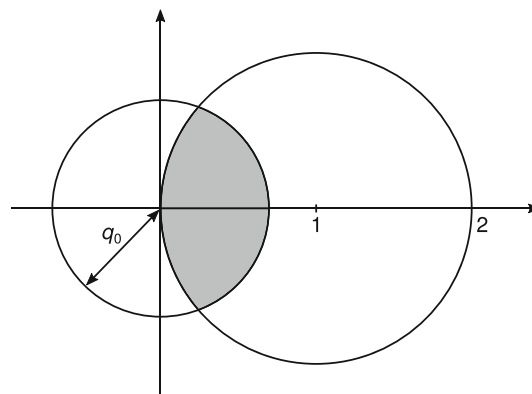


Fig. 2 Area in the complex plain (the grey colour) to which μ_i have to belong to guarantee that $\|K\| = 1$ holds

operator in $L^2(\sigma)$ with domain dense in this space. More explicitly, the equation

$$v = Uu \tag{96}$$

has a solution that in terms of components can be written as

$$\langle u, \varphi_k \rangle = \frac{1}{\lambda_k} \langle v, \xi_k \rangle, \tag{97}$$

showing that $u \in L^2(\sigma)$ only if (Goursat 1964, p. 158)

$$\|u\|_{L^2(\sigma)}^2 = \sum_{k=1}^{+\infty} \langle u, \varphi_k \rangle_{L^2(\sigma)}^2 = \sum_{k=1}^{+\infty} \frac{\langle v, \xi_k \rangle_{L^2(\sigma)}^2}{\lambda_k^2} < +\infty. \tag{98}$$

So it is clear that not for every $v \in L^2(\sigma)$ there is a solution $u \in L^2(\sigma)$ of (96); only if (98) is verified, which is also known as Picard’s condition, can u be “reasonably” computed. This is just another way to claim that the downward continuation operator

$$D = U^{-1}, \tag{99}$$

is unbounded in $L^2(\sigma)$ and therefore the downward continuation operation is improperly posed in $L^2(\sigma)$.

4 The sandwich conjecture and its functional implications

In this section we restrict our choice of PB to $PB = I$ (see 35), because this makes the presentation quite smoother, while, as we will see in Sect. 6, on a numerical ground the other choice (36) gives only small advantages. Then we start from Remark 4 and we observe that when the CoB algorithm is

seen as acting on the whole $L^2(\sigma)$, the operator $K = I - U$ cannot have a norm smaller than 1; therefore, the question is whether we have

$$\|K\| = 1; \tag{100}$$

if on the contrary we would decide that $\|K\| > 1$, the whole CoB concept would be destroyed. Yet in this section we would like to analyse the consequences of (100). In other words, would $\|K\|$ be equal to 1, what happens to the residuals s_ℓ given by (28)?

As already recalled, (100) is a “reasonable” hypothesis descending from the sandwich conjecture, namely that calling h_{nm} the eigenvalues of

$$H = K^T K = (I - U^T)(I - U), \tag{101}$$

one has

$$\left[1 - \left(\frac{R}{R}\right)^{n+1}\right]^2 \geq h_{nm} \geq \left[1 - \left(\frac{R}{R_0}\right)^{n+1}\right]^2; \tag{102}$$

therefore, would (102) be correct, we should have

$$\sup h_{nm} = \|K\|^2 = 1. \tag{103}$$

The relation (102) is only guessed on the basis of a perturbative argument (see Sansò and Sideris 2017, Appendix). We shortly recapitulate the reasoning for the right inequality in (102); a similar reasoning holds for the left inequality.

Proof Let S_R be the sphere completely internal to S , so that $R < R_0 < \bar{R}$ (recall the definition 4) and S is totally inside the layer between S_{R_0} and $S_{\bar{R}}$. Let us put

$$\begin{aligned} U &\sim \sum_{n,m} q_\sigma^{n+1} Y_{nm}(\sigma) Y_{nm}(\sigma'), \\ U_0 &= \sum_{n,m} q_0^{n+1} Y_{nm}(\sigma) Y_{nm}(\sigma'), \end{aligned} \tag{104}$$

where as always $q_\sigma = \frac{R}{R_\sigma}$, $q_0 = \frac{R}{R_0}$ and $q_\sigma < q_0$. Then

$$\delta U = U - U_0, \quad \delta H \cong -\delta U^T (I - U_0) - (I - U_0^T) \delta U. \tag{105}$$

Indeed U_0 is selfadjoint with eigenvalues $h_{0nm} = q_0^{n+1}$ and eigenvectors $\{Y_{nm}(\sigma)\}$.

Accordingly $H_0 = (I - U_0^T)(I - U_0) = (I - U_0)^2$ has eigenvalues $h_{0nm} = \left[1 - q_0^{n+1}\right]^2$ and eigenvectors $\{Y_{nm}(\sigma)\}$.

By first-order perturbative formulas we know that

$$\delta h_{nm} = h_{nm} - h_{0nm} = \langle Y_{nm}, \delta H Y_{nm} \rangle; \tag{106}$$

the right-hand side of (102) is proved if we can show that $\delta h_{nm} > 0$.

But, using (105),

$$\begin{aligned} \delta h_{nm} &= -\langle Y_{nm}, \delta U^T (I - U_0) Y_{nm} \rangle \\ &\quad - \langle Y_{nm}, (I - U_0^T) \delta U Y_{nm} \rangle \\ &= -2\langle (I - U_0) Y_{nm}, \delta U Y_{nm} \rangle \\ &= -2h_{0nm} \langle Y_{nm}, \delta U Y_{nm} \rangle. \end{aligned} \tag{107}$$

On the other hand

$$\delta U Y_{nm} = \left[q_\sigma^{n+1} - q_0^{n+1} \right] Y_{nm}(\sigma), \tag{108}$$

so that

$$\langle Y_{nm}, \delta U Y_{nm} \rangle = \int \left[q_\sigma^{n+1} - q_0^{n+1} \right] Y_{nm}^2(\sigma) d\sigma < 0. \tag{109}$$

Since $h_{0nm} > 0$, (107) and (109) give $\delta h_{nm} > 0$ as we wanted to prove. \square

The authors want to comment once more that the above proof is only heuristic in character, and that it holds under the condition that (105) be satisfied, i.e. that the linearization of U with respect to U_0 gives a good approximation of this operator.

So we assume that (100) is true, namely that K is a non-expansive operator. An immediate consequence is that s_ℓ is a bounded sequence

$$\|s_\ell\| \leq \|f\| \tag{110}$$

and therefore a subsequence of it is weakly convergent; yet the point is whether such a limit is 0 or not and whether 0 is the limit for the whole sequence. Even more we would prefer to have a stronger result, namely that $s_\ell \rightarrow 0$ in the $L^2(\sigma)$ norm. This is not possible in general, yet a close result can be obtained by using the concept of Cesaro limit.

Definition 1 We say that the sequence of real numbers $\{x_n\}$ has the Cesaro limit x if the Cesaro means

$$C_N(\{x_n\}) = \frac{1}{N} \sum_{n=1}^N x_n \tag{111}$$

have x as usual limit, i.e.

$$\lim_{N \rightarrow \infty} C_N(\{x_n\}) = x. \tag{112}$$

In this case we write

$$x_n \xrightarrow{c} x. \tag{113}$$

Remark 6 It is obvious that such a definition can be generalized to any topological space, where the definition of the limit of a sequence is specified by the topology. So, for instance, in a Hilbert space X we say that

$$x_n \xrightarrow{c} x \tag{114}$$

means that

$$\|C_N(\{x_n\}) - x\|_X \rightarrow 0. \tag{115}$$

Similarly, endowing X of the weak topology, we say that

$$x_n \xrightarrow{c^*} x \tag{116}$$

if

$$\langle h, C_N(\{x_n\}) \rangle_X \rightarrow \langle h, x \rangle_X \quad \forall h \in X. \tag{117}$$

Supplied with this definition, we turn our attention to the sequence $\{s_\ell\} \in L^2(\sigma)$ and recalling (28) we see that their Cesaro means can be written as

$$C_N(\{s_n\}) \equiv \frac{1}{N} \sum_{n=1}^N K^n f. \tag{118}$$

It is therefore convenient to define the mean operators

$$M_N = \frac{1}{N} \sum_{n=1}^N K^n; \tag{119}$$

it is obvious that, like K , all the mean operators are non-expansive, namely

$$\|M_N\| \leq 1. \tag{120}$$

Considering the sequence $M_N(f) \equiv C_N(\{s_n\})$, we get an immediate result recurring to the ergodic theory, more specifically to the Mean Ergodic Theorem (Yosida 1995).

Proposition 1 *If $\|K\| \leq 1$ we have*

$$s_n \xrightarrow{c} 0 \text{ in } L^2(\sigma), \tag{121}$$

i.e.

$$\forall f \in L^2(\sigma), \quad M_N(f) \xrightarrow{L^2(\sigma)} 0. \tag{122}$$

Proof We adapt here the very general proof given in Yosida (1995) to our much simpler case.

We first notice that the following identity holds

$$KM_N - M_N = \frac{1}{N}(K^{N+1} - K), \tag{123}$$

so that

$$\|KM_N - M_N\| \leq \frac{2}{N} \rightarrow 0. \tag{124}$$

Now, given (120) we know that $\forall f \in L^2(\sigma)$ fixed, the sequence $M_N(f)$ is bounded by $\|f\|^2$; so there is a subsequence $\{N_j\}$ such that

$$M_{N_j}(f) \xrightarrow{*} g \tag{125}$$

for some $g \in L^2(\sigma)$.

Therefore

$$KM_{N_j}(f) - M_{N_j}(f) \rightarrow 0 \tag{126}$$

by (124) and on the same time (K is continuous)

$$KM_{N_j}(f) - M_{N_j}(f) \xrightarrow{*} Kg - g; \tag{127}$$

so we must have

$$Kg - g = -Ug = 0 \Rightarrow g = 0, \tag{128}$$

because of property d) in Sect. 3. We conclude that

$$M_{N_j}(f) \xrightarrow{*} 0. \tag{129}$$

Since the same reasoning can be repeated for any other subsequence $\{N_k\}$, we can conclude that for the whole sequence

$$M_N(f) \xrightarrow{*} 0. \tag{130}$$

Hence

$$s_n \xrightarrow{c^*} 0. \tag{131}$$

Now take any $w \in \mathcal{B}(U)$, i.e

$$w = Uu, \quad u \in L^2(\sigma) \tag{132}$$

and notice that

$$M_N(w) = M_N(Uu) = U(M_N(u)). \tag{133}$$

Since $M_N(u) \xrightarrow{*} 0$ and U is compact by property a) from Sect. 3, we find

$$M_N(w) \rightarrow 0. \tag{134}$$

On the other hand we know that $\mathcal{R}(U)$ is dense in $L^2(\sigma)$ by property b) from 3; then given any $f \in L^2(\sigma)$ and any $\varepsilon > 0$ we can find $w \in \mathcal{R}(U)$ such that

$$\|f - w\| < \varepsilon. \tag{135}$$

Then we have

$$\begin{aligned} \|M_N(f)\| &= \|M_N(f - w) + M_N(w)\| \\ &\leq \|f - w\| + \|M_N(w)\| \leq \varepsilon + \|M_N(w)\|. \end{aligned} \tag{136}$$

Therefore, by using (134) we get

$$\overline{\lim}_{N \rightarrow \infty} \|M_N(f)\| \leq \varepsilon, \tag{137}$$

i.e.

$$M_N(f) \rightarrow 0. \tag{138}$$

□

Remark 7 Returning to (29), we see that

$$s_\ell \xrightarrow{c} 0 \Rightarrow F_\ell(\sigma) \xrightarrow{c} f(\sigma), \tag{139}$$

namely that

$$\frac{1}{N} \sum_{\ell=1}^N F_\ell(\sigma) \rightarrow f(\sigma). \tag{140}$$

Since, by definition,

$$F_\ell(\sigma) = \sum_{k=1}^{\ell} f_k(\sigma), \tag{141}$$

(140) can be written in the comfortable form

$$\sum_{k=1}^N \left(1 - \frac{k-1}{N}\right) f_k(\sigma) \rightarrow f(\sigma), \tag{142}$$

showing that $f(\sigma)$ can be obtained directly from $f_k(\sigma)$ as limit of the series (142).

5 A modified CoB for global models

A gravity model, as recalled in the Introduction, is just a finite degree potential (see 1). The reason to concentrate on this particular case is twofold: on the one hand, this is what we really want to get in our analysis, namely that CoB is converging for a global model. On the other hand, we will introduce in

the space of global models of degree N a “natural” topology, which however would become more problematic to be extended to the whole $L^2(\sigma)$.

So we first define the space of global models of degree N

$$\mathcal{M}_N = \left\{ u = \sum_{n=0}^N \sum_{m=-n}^n u_{nm} \left(\frac{R}{r}\right)^{n+1} Y_{nm}(\sigma) \right\}, \tag{143}$$

with

$$\{u_{nm}\} \in \mathbb{R}^{(N+1)^2}. \tag{144}$$

Then we observe that, given any smooth star-shaped surface, $S \equiv \{r = R_\sigma, R_\sigma > 0\}$, $u \in \mathcal{M}_N$ is totally identified by its trace on S , i.e. we have the one-to-one correspondences

$$\begin{aligned} v(\sigma) = \sum_{nm}^N u_{nm} Y_{nm}(\sigma) &\longleftrightarrow u \in \mathcal{M}_N \longleftrightarrow f(\sigma) \\ &= \sum_{n,m}^N u_{nm} q_\sigma^{n+1} Y_{nm}(\sigma). \end{aligned} \tag{145}$$

A rapid inspection of the CoB scheme in Fig. 1 shows that (145) can be formally written as

$$f = \Gamma_S u = \Gamma_S \Pi v = \Gamma_S \Pi \Gamma_{S_R} u; \tag{146}$$

each operator in (146) is invertible, whence the above correspondence.

We call

$$\mathcal{H}_N = \text{span}\{q_\sigma^{n+1} Y_{nm}(\sigma)\} \subset L^2(\sigma), \tag{147}$$

$$H_N = \text{span}\{Y_{nm}(\sigma)\} \subset L^2(\sigma) \tag{148}$$

and we notice that $\mathcal{M}_N, H_N, \mathcal{H}_N$ are all $(N+1)^2$ dimensional linear spaces, since their elements depend ultimately on the $(N+1)^2$ real constants u_{nm} . The one-to-one relations (146) can therefore be used to introduce a norm in one of them and then extend it to the others by isometry.

In particular, we will define

$$|u|_{\mathcal{M}_N}^2 = |v|_{H_N}^2 = |f|_{\mathcal{H}_N}^2 = \sum_{n=0}^N \sum_{m=-n}^n u_{nm}^2. \tag{149}$$

Such norms correspond to the usual scalar product in $L^2(\sigma)$ for H_N

$$\langle h, g \rangle_{H_N} = \langle h, g \rangle_{L^2(\sigma)}, \tag{150}$$

such that

$$\{Y_{nm}(\sigma); n \leq N\} \tag{151}$$

is an orthonormal basis of H_N , and to the unusual scalar product

$$\begin{cases} Uh = v \in \mathcal{H}_N, & Ug = w \in \mathcal{H}_N, \\ \langle w, v \rangle_{\mathcal{H}_N} = \langle g, h \rangle_{L^2(\sigma)}. \end{cases} \tag{152}$$

It is clear that in this case

$$\{q_\sigma^{n+1} Y_{nm}(\sigma)\} \tag{153}$$

is an orthonormal basis of \mathcal{H}_N .

It is clear from (149) and (152) that if we tried to go to infinite dimensions, $N \rightarrow \infty$, with this definitions we would end up with functions v defined on S , that in any way can be harmonically continued down to S_R ; at this point it is more handy to work directly with finite degree potentials, what justifies our previous choice.

Now let us return to the CoB scheme of Fig. 1. The iteration is accomplished by going from S to S_R with the pull back operator PB and then going back to S by the upward operator U . Spectrally this means (see 38)

$$PB(q_\sigma^{n+1} Y_{nm}) = c_n q_\sigma^{n+1} Y_{nm}, \tag{154}$$

$$U(Y_{nm}) = q_\sigma^{n+1} Y_{nm}. \tag{155}$$

As we see from (155) we always have

$$U : H_N \rightarrow \mathcal{H}_N; \tag{156}$$

however in general we have that, unless $c_n q_\sigma^{n+1}$ is constant, i.e. S is a sphere, PB sends functions in \mathcal{H}_N to a subspace of $L^2(\sigma)$, spanned by many more harmonics. The ordinary practice at this point would be that we restrict R_σ to be a surface with finite spectrum, i.e. $R_\sigma \in H_p$ for some p . Yet this implies that $q_\sigma = R/R_\sigma$ has full spectrum up to infinity and the same would be true for $f = \Gamma_s u$ (Bucha et al. 2019). So to cut short the discussion, we make here the unconventional, but numerically equivalent, hypothesis that directly $q_\sigma \in H_p$ for some p . Then, returning to our discussion, even assuming that

$$c_n = 1 \quad (PB = I), \quad q_\sigma \in H_p, \tag{157}$$

we have

$$q_\sigma^{n+1} \in H_{(n+1)p}, \quad q_\sigma^{n+1} Y_{nm} \in H_{n+(n+1)p}, \tag{158}$$

namely

$$PB : \mathcal{H}_N \rightarrow H_{N'}, \quad N' = n + (n + 1)p, \tag{159}$$

and

$$U PB : \mathcal{H}_N \rightarrow \mathcal{H}_{N'}. \tag{160}$$

This means that at every cycle in the iteration we are accumulating more and more high degrees in the residuals

$$s = Kf = (I - U PB)f \tag{161}$$

which are not useful in building the approximation to f , which we know a priori to have zero components on all the degrees beyond N . An easy remedy to this anomalous inflation of the norm of the residuals is just to truncate $U PBf$ at the maximum degree N . This can be done in two ways, either by projecting $U PBf$ on \mathcal{H}_N , using the orthonormal basis $\{q_\sigma^{n+1} Y_{nm}\}$ and the scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}_N}$, or by projecting PBf on H_N , using the orthonormal basis $\{Y_{nm}\}$ and the scalar product $\langle \cdot, \cdot \rangle_{H_N} = \langle \cdot, \cdot \rangle_{L^2(\sigma)}$. The two results coincide, so we can choose the second approach and introduce the truncation operator \mathcal{T}_N with kernel

$$\mathcal{T}_N(\sigma, \sigma') = \sum_{n,m}^N Y_{nm}(\sigma) Y_{nm}(\sigma') \tag{162}$$

to be used with the ordinary $L^2(\sigma)$ scalar product.

We find then,

$$f = \sum_{n,m}^N f_{nm} c_n q_\sigma^{n+1} Y_{nm} \in \mathcal{H}_N, \tag{163}$$

$$U \mathcal{T}_N PBf = \sum_{jk} q_\sigma^{j+1} Y_{jk} \left(\sum_{n,m}^N \langle Y_{jk}, c_n q_\sigma^{n+1} Y_{nm} \rangle f_{nm} \right), \tag{164}$$

namely the matrix $K_{jk,nm}$ performing the iteration $f \rightarrow Kf \in \mathcal{H}_N$, is given by

$$K_{jk,nm} = \delta_{jn} \delta_{km} - \langle Y_{jk}, c_n q_\sigma^{n+1} Y_{nm} \rangle. \tag{165}$$

Since the scalar product $\langle \cdot, \cdot \rangle$ is coupled with the norm of f in H_N , we also have that

$$\|K\| = \|I - U \mathcal{T}_N PB\| = \|\{K_{jk,nm}\}\|, \tag{166}$$

where the matrix norm is given by the square root of the maximum eigenvalue, i.e. the spectral radius of the matrix.

Remark 8 It is convenient to elaborate already here the form of the matrix $K_{jk,nm}$ when c_n is not just 1, but it is given by $c_n = 1 + (n + 1) \frac{h_\sigma}{R_\sigma}$. Recalling that $h_\sigma = R_\sigma - R$ and noticing that

$$\frac{h_\sigma}{R_\sigma} = 1 - q_\sigma, \tag{167}$$

we easily find

$$c_n q_\sigma^{n+1} = (n + 2) q_\sigma^{n+1} - (n + 1) q_\sigma^{n+2}, \tag{168}$$

that is the formula we will use in the next section.

Remark 9 It is important to understand that when we pass to finite dimensional potentials and on the same time we restrict the CoB algorithm to a finite dimensional subspace of $L^2(\sigma)$, i.e. \mathcal{H}_N , we represent the operator K by a finite matrix which has only a finite number of singular values. So the argument that necessarily they tend to 1 used to claim that $\|K\| = 1$ (or larger) (see 100) is not valid here and in fact, if the sandwich conjecture (see 101 and 102) holds, we can expect

$$\|K\| < 1. \tag{169}$$

Indeed, in lack of a rigorous proof, the contractive character of the finite dimensional K expressed in (169) is just an expectation, even if the sandwich conjecture should hold true.

6 Numerical experiments

In this section, we numerically examine the conjecture (169), the validity of which is critical for CoB to converge. We start by studying the pull back operator (35), for which $c_n q_\sigma^{n+1} = q_\sigma^{n+1}$ (see 40). The experiments are performed for surfaces S of various complexity: a synthetic topography (Sect. 6.1) and the topographies of the asteroids Benu (Sect. 6.2) and Eros (Sect. 6.3). Finally, in Sect. 6.4, we briefly report the results obtained with the improved pull back operator (36) with $c_n q_\sigma^{n+1}$ being given by (168).

Before moving to the experiments, let us emphasize that numerical investigations of this kind can never prove or disprove the convergence of CoB. Nevertheless, a rigorous analytical study with realistic surfaces S appears to be too difficult, so the numerics is used to shed some light on the behaviour of CoB in practice. Hopefully, the experiments will allow us to draw some general conclusions.

6.1 Synthetic topography

Let S be represented by a zonal function as

$$q_\sigma(\vartheta) = 1 - a + \alpha P_M(\cos \vartheta). \tag{170}$$

The shape of S is dominated by the oscillations of the Legendre polynomial $P_M(\cos \vartheta)$, while the parameters a and α are numerical constants to control the oscillations. In any case, a and α are chosen such that $q_\sigma(\vartheta) < 1$ for all ϑ (see 11). We call this surface a synthetic topography of degree M .

As an advantage, the synthetic topography is independent of the longitude, allowing us to think of CoB as being independent of the longitude, too. This reduces the size of the matrix $K_{jk,nm}$ and simplifies the evaluation of its elements

to (see 165)

$$\begin{aligned} K_{jn} &= \delta_{jn} - \frac{1}{2} \int_0^\pi Y_{j0}(\vartheta) q_\sigma^{n+1}(\vartheta) Y_{n0}(\vartheta) \sin \vartheta \, d\vartheta \\ &= \delta_{jn} - \frac{\sqrt{2j+1} \sqrt{2n+1}}{2} \\ &\quad \times \int_0^\pi P_j(\cos \vartheta) q_\sigma^{n+1}(\vartheta) P_n(\cos \vartheta) \sin \vartheta \, d\vartheta. \end{aligned} \tag{171}$$

Note that the $(2\pi)^{-1}$ factor is missing before the integrals due to the independence on the longitude. For $n = j$ and $q_\sigma = 1$, we therefore have

$$\frac{2n+1}{2} \int_0^\pi P_n(\cos \vartheta) P_n(\cos \vartheta) \sin \vartheta \, d\vartheta = 1. \tag{172}$$

The matrix K_{jk} was computed for two synthetic topographies of degree $M = 180$, one with $a_1 = 0.001$ and $\alpha_1 = 0.0003$ and the other with $a_2 = 0.01$ and $\alpha_2 = 0.003$. On the Earth, the oscillations of the two topographies would roughly correspond to the range of $\max_\sigma R_\sigma - \min_\sigma R_\sigma \approx 9$ km and 90 km, respectively, with the half-wavelength of 110 km. The larger is the range $\max_\sigma R_\sigma - \min_\sigma R_\sigma$, the more non-spherical is the surface S which, in turn, increases the numerical complexity. The $\{K_{jn}\}$ elements were obtained by (171) for $j = 0, 1, \dots, N$ and $n = 0, 1, \dots, N$, where N gradually varies from 0 to 720. The scalar products were computed in Fortran in double and quadruple precision using the numerical integration technique of Fukushima (2017). The relative error tolerance δ , affecting the accuracy of the numerical integration, was set to $\delta = 10^{-16}$ in double precision and to $\delta = 10^{-32}$ in quadruple precision. In double precision, the norms $\|K\|$ were computed for all $N = 0, 1, \dots, 720$ in Python using the `norm` function (the `ord=2` flag) from the `numpy.linalg` module (Harris et al. 2020). In quadruple precision, the norms were computed in MATLAB using the ADVANPIX toolbox (www.advantpix.com).

Figures 3 and 4 show that somewhere around degree 700, the norms in double precision do not satisfy the (169) conjecture, as they are equal to or larger than the threshold of 1. In quadruple precision, the inequality (169) is, however, satisfied for all studied N . Therefore, we attribute the jump near degree 700 to numerical errors caused by the limited accuracy of the double precision environment. A comparison of the two figures also shows that the norms of the second synthetic topography approach the threshold of 1 faster, owing to the stronger oscillations of S .

As a conclusion, we expect the conjecture (169) is satisfied for the studied synthetic topographies, even for the strongly

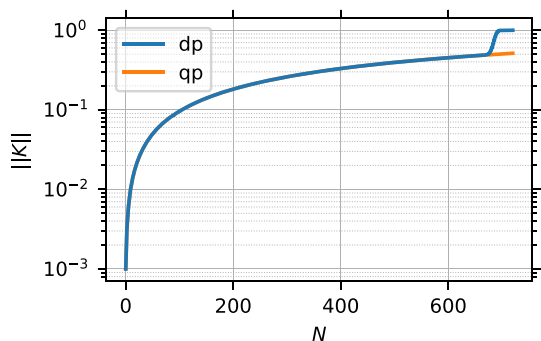


Fig. 3 Norms $\|K\|$ as a function of the maximum harmonic degree of the scalar products N for a synthetic topography (170) defined by $M = 180$, $a_1 = 0.001$ and $\alpha_1 = 0.0003$. The abbreviations dp and qp stand for the double and quadruple precision environments, respectively, in which the scalar products and the norms were computed

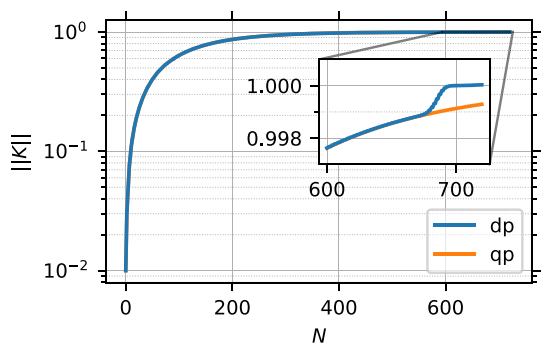


Fig. 4 The same as Fig. 3 but for a more oscillatory synthetic topography defined by $M = 180$, $a_2 = 0.01$ and $\alpha_2 = 0.003$

oscillating one. This conclusion is valid at least up to the studied maximum degree $N = 720$.

6.2 Asteroid Bennu

In this section, the surface S is represented by a realistic shape model of the asteroid Bennu. We define R_σ as

$$R_\sigma = \sum_{n=0}^{15} \sum_{m=-n}^n r_{nm} Y_{nm}(\sigma), \tag{173}$$

where r_{nm} are spherical harmonic coefficients of Bennu’s shape taken from the Bucha and Sansò (2021) study. This scenario is significantly more challenging than that with the synthetic topographies for two reasons. First, R_σ now varies both with the colatitude and longitude, so that all $\{K_{jk,nm}\}$ elements up to some maximum degree N are considered. Second, the ratio between the radii of the minimum Brillouin sphere and the maximum Bjerhammar sphere, the shape ratio \bar{R}/R_0 (see Bucha and Sansò 2021), is about $282.5 \text{ m}/224 \text{ m} \approx 1.261$ for Bennu while it is only about $(6378 \text{ km} + 90 \text{ km})/6378 \text{ km} \approx 1.014$ for the second synthetic topography from Sect. 6.1. The shape ratio reflects the

extent, to which the surface S deviates from a sphere. Large shape ratio therefore implies highly non-spherical surface S , hence increased numerical complexity (see Sect. 6.1).

Due to the complexity of this scenario, we were able to compute $K_{jk,nm}$ only up to degree $N = 65$. To get the scalar products, we used $\delta = 10^{-8}$ in the numerical integration technique at first and executed the computation in double precision. Despite the seemingly low maximum degree of 65, the norms were revealed to exceed the threshold of 1 already at degree $N = 31$ (Fig. 5). To make sure this result is not affected by numerical inaccuracies, we tightened the relative error tolerance to $\delta = 10^{-16}$. With the improved numerical accuracy, it turned out the norms are in fact smaller than 1 for all $N \leq 63$, meaning that the critical degree moved to higher figures (see Fig. 5). It would certainly be of interest to repeat the experiment in quadruple precision with some $\delta < 10^{-16}$ to confirm/disprove that the degree of 63 is not again limited by the numerics, similarly as we did in Sect. 6.1. Unfortunately, this is not possible for us with our hardware and software, as already the integration in double precision with $\delta = 10^{-16}$ required a few weeks of wall-clock time when using about 96 CPU cores. This clearly demonstrates that our numerical experiments cannot provide definite answers on the validity of (169), but at most some hints, though valuable.

To summarize the experiment, we have shown that even with as irregularly shaped body as the asteroid Bennu, the norms stay below the threshold of 1 at least up to degree 63. Using the very same shape of Bennu, Bucha and Sansò (2021) have shown that a relative accuracy of $\sim 10^{-6}$ can be achieved on S in terms of the gravitational potential that is expanded in a single series of external spherical harmonics up to degree 65. This might indicate that CoB could deliver reasonably accurate gravity field models even in advanced conditions.

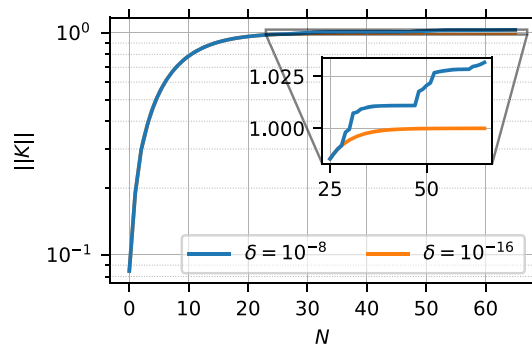


Fig. 5 Norms $\|K\|$ obtained for the asteroid Bennu in double precision with two relative error tolerance parameters δ . The blue line is above the threshold of 1 for $N \geq 31$ and the orange one for $N \geq 64$

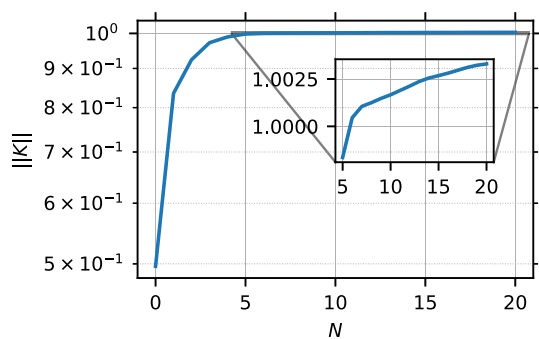


Fig. 6 Norms $\|K\|$ obtained for the asteroid Eros in double precision with the relative error tolerance of $\delta = 10^{-16}$. The norms are larger than 1 for $N \geq 6$

6.3 Asteroid Eros

Finally, we repeated the previous experiment, but this time with a non-spherical asteroid Eros. The shape coefficients r_{nm} of Eros were taken from the Zuber et al. (2000) study up to degree 5. The shape ratio for such defined surface reads roughly $16000 \text{ m}/3250 \text{ m} \approx 4.923$ which demonstrates the non-sphericity of Eros.

The $K_{jk, nm}$ elements were computed up to degree $N = 20$ in double precision with $\delta = 10^{-16}$ and in quadruple precision with $\delta = 10^{-32}$. In both cases, the norms are almost identical and quickly approach the threshold of 1, exceeding it already at degree 6 (Fig. 6). This indicates that there might indeed exist conditions, for which the conjecture (169) is not satisfied. Yet, we cannot reject the conclusion that (169) is satisfied but instead quadruple precision and the relative error tolerance of $\delta = 10^{-32}$ still do not ensure sufficient numerical accuracy for Eros.

6.4 Improved pull back operator

To study the effect of the Taylor order on the downward continuation, we repeated some of the previous experiments with the improved pull back operator (36). We studied the norms for the second synthetic topography in quadruple precision (see Sect. 6.1) and for Eros in double precision (Sect. 6.3). The plots of the norms are not shown here as the curves are visually the same as seen in the previous sections.

With the second synthetic topography, the norms stayed below the threshold of 1 with both variants of the pull back operator. However, the improved pull back operator produced slightly smaller norms (e.g., 0.999 vs. 0.994 at degree 720 in favour of the improved pull back operator). In case of Eros, the norms exceeded the threshold of 1 at degree 7 which is one degree higher than with the pull back operator (35) (see Sect. 6.3). The generally smaller norms indicate that the improved pull back operator (36) has better convergence properties than (35), though the enhancement is probably not significant.

7 Some conclusions and perspectives

In the paper the concept of change of boundary has been revisited and in particular the functional implications of the hypothesis that

$$\|K\| \leq 1 \quad (174)$$

have been explored, showing that even in the case of a non-expansive operator, $\|K\| = 1$, we can have a convergence of the method when the concept of Cesaro limit is used.

The hypothesis (174) has no strict analytical proof for the moment, yet it is a consequence of a first-order perturbative argument, in which the surface $S \equiv \{r = R_\sigma\}$ is supposed to be not too far from a sphere.

In lack of a rigorous proof, our hypothesis has been subject to numerical tests in order to verify its plausibility.

In the first experiment, we have used data from a simulation of the gravity field of objects with various topographic surfaces with a width up to 1.4 % of the radius ($\sim 90 \text{ km}$ in the case of the Earth) and wavelength corresponding to a maximum degree 720.

In the second experiment, data related to the asteroid Bennu (Bucha and Sansò 2021) have been used. In this case the topographic surface is quite smooth but the ratio between minimum Brillouin radius and maximum Bjerhammar radius, \bar{R}/R_0 , is about 1.261 meaning a significant, but not too large, deviation from a spherical shape. In both experiments, the hypothesis (174) has proved to be tenable.

In a third experiment we have been stressing the conditions of validity of hypothesis (174) by using data related to the asteroid Eros which has a shape very far from a sphere, with ratio $\bar{R}/R_0 \approx 4.923$. In this case the $\|K\|$ norm has been estimated to be larger than 1, i.e. the CoB theory is not applicable.

Furthermore two forms of the pull back operator have been tested, arriving at the conclusion that they display a not significantly different behaviour as for what concerns convergence.

Summarizing we may claim that the hypothesis (174), and the consequent convergence of the method, can be considered as effective when working with boundaries S not far from the sphere, as for long wavelength behaviours, while short wavelength perturbations seem not to effect too much convergence.

On the contrary, for a surface S significantly deformed with respect to the sphere, we certainly cannot rely on the CoB method and we expect non-convergence even for low-degree truncated models.

The threshold between the two behaviours still needs to be studied both theoretically and numerically.

Acknowledgements The computations were performed at the HPC centre at the Slovak University of Technology in Bratislava, which is a part of the Slovak Infrastructure of High Performance Computing (SIVVP Project, ITMS code 26230120002, funded by the European region development funds, ERDF). The plots were prepared with the Python's Matplotlib module (Hunter 2007).

Author Contributions Although FS has worked more on the theoretical development of the study and BB on the numerical experiments, a continuous discussion between the authors has made the paper an equal result of theirs.

Funding Open access funding provided by The Ministry of Education, Science, Research and Sport of the Slovak Republic in cooperation with Centre for Scientific and Technical Information of the Slovak Republic. BB was supported by the VEGA 1/0809/21 Project.

Availability of data and materials The spherical harmonic shape models of Bennu and Eros are available at <https://doi.org/10.5281/zenodo.6622624> and <http://sebago.mit.edu/near/nlr.30day.html>, respectively.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Axler S, Bourdon P, Ramey W (1992) Harmonic function theory, 1st edn. Springer, New York
- Bucha B, Sansò F (2021) Gravitational field modelling near irregularly shaped bodies using spherical harmonics: a case study for the asteroid (101955) Bennu. *J Geod* 95:56. <https://doi.org/10.1007/s00190-021-01493-w>
- Bucha B, Hirt C, Kuhn M (2019) Divergence-free spherical harmonic gravity field modelling based on the Runge–Krarup theorem: a case study for the Moon. *J Geod* 93:489–513. <https://doi.org/10.1007/s00190-018-1177-4>
- Fukushima T (2017) Precise and fast computation of the gravitational field of a general finite body and its application to the gravitational study of asteroid Eros. *Astron J* 154(145):15. <https://doi.org/10.3847/1538-3881/aa88b8>
- Goursat E (1964) Integral equations, calculus of variations. Dover, New York
- Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH, Brett M, Haldane A, del Río JF, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C, Oliphant TE (2020) Array programming with NumPy. *Nature* 585:357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Heiskanen WA, Moritz H (1967) Physical geodesy. W. H. Freeman and Company, San Francisco
- Hirt C, Kuhn M (2017) Convergence and divergence in spherical harmonic series of the gravitational field generated by high-resolution planetary topography—a case study for the Moon. *J Geophys Res Planets* 122:1727–1746. <https://doi.org/10.1002/2017JE005298>
- Hu X, Jekeli C (2015) A numerical comparison of spherical, spheroidal and ellipsoidal harmonic gravitational field models for small non-spherical bodies: examples for the Martian moons. *J Geod* 89:159–177. <https://doi.org/10.1007/s00190-014-0769-x>
- Hunter JD (2007) Matplotlib: a 2D graphics environment. *Comput Sci Eng* 9:90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Kravaris T (1975) On potential theory. In: Borre K (ed) Mathematical foundation of geodesy: selected papers of Torben Kravaris, vol 2006. Springer, New York, pp 207–255
- Martinez Z (1998) Boundary-value problems for gravimetric determination of a precise geoid. Springer, Berlin
- Pavlis N (2013) Global gravitational models. In: Sansò F, Sideris MG (eds) Geoid determination: theory and methods. Springer, Heidelberg, pp 261–310
- Pavlis NK, Holmes SA, Kenyon SC, Factor JK (2012) The development and evaluation of the Earth Gravitational Model 2008 (EGM2008). *J Geophys Res* 117(B04406):1–38. <https://doi.org/10.1029/2011JB008916>
- Rapp RH (1993) Use of altimeter data in estimating global gravity models. *Satellite altimetry in geodesy and oceanography*. Springer, Berlin, pp 374–417
- Reimond S, Baur O (2016) Spheroidal and ellipsoidal harmonic expansions of the gravitational potential of small Solar System bodies. Case study: Comet 67P/Churyumov-Gerasimenko. *J Geophys Res Planets* 121:497–515. <https://doi.org/10.1002/2015JE004965>
- Sansò F (1993) Theory of geodetic B.V.P.s applied to the analysis of altimetric data. In: Rummel R, Sansò F (eds) *Satellite altimetry in geodesy and oceanography*. Springer, Berlin, pp 318–371. <https://doi.org/10.1007/BFb0117931>
- Sansò F (2013a) The analysis of geodetic boundary value problems in linear form. In: Sansò F, Sideris MG (eds) *Geoid determination: theory and methods*. Springer, Berlin, pp 663–706
- Sansò F (2013b) Harmonic calculus and global gravity models. In: Sansò F, Sideris MG (eds) *Geoid determination: theory and methods*. Springer, Heidelberg, pp 111–168
- Sansò F (2013c) On potential theory and HS of harmonic functions. In: Sansò F, Sideris MG (eds) *Geoid determination: theory and methods*. Springer, Heidelberg, pp 591–644
- Sansò F, Sideris MG (2017) Geodetic Boundary value problem: the equivalence between Molodensky's and Helmert's solutions. *SpringerBriefs in earth sciences*. Springer, Berlin
- Sansò F, Venuti G (2001) White noise stochastic BVP's and Cimmino's theory. In: Sansò F (ed) *IV Hotine-Marussi symposium on mathematical geodesy*. Springer, Heidelberg, pp 5–13
- Sebera J, Bezděk A, Pešek I, Henych T (2016) Spheroidal models of the exterior gravitational field of asteroids Bennu and Castalia. *Icarus* 272:70–79. <https://doi.org/10.1016/j.icarus.2016.02.038>
- Takahashi Y, Scheeres DJ (2014) Small body surface gravity fields via spherical harmonic expansions. *Celest Mech Dyn Astron* 119:169–206. <https://doi.org/10.1007/s10569-014-9552-9>
- Yosida K (1995) *Functional analysis*, 6th edn. Springer, Berlin
- Zuber MT, Smith DE, Cheng AF, Garvin JB, Aharonson O, Cole TD, Dunn PJ, Guo Y, Lemoine FG, Neumann GA, Rowlands DD, Torrence MH (2000) The shape of 433 Eros from the NEAR-Shoemaker Laser Rangefinder. *Science* 289:2097–2101. <https://doi.org/10.1126/science.289.5487.2097>