Chapter 4
The Anomalous Potential and Its Determination

4.1 Outline

The knowledge of the normal potential and related ellipsoidal quantities are not enough to properly treat the problem of relating different types of geodetic heights.

To do that we need a much more precise knowledge of the geoid, i.e. of the gravity potential W, than that supplied by the ellipsoid, which leaves out the "last 100 m" of the geoid undulation. To do that we need to learn how to model the difference between W_{P} and U_{P}, namely the anomalous potential T_{P}. How this can be derived by a suitable fusion of different data sources, like surface gravity, satellite tracking, digital terrain models and oceanic mean dynamic heights, is certainly one of the main tasks of Physical Geodesy, requiring a good knowledge of some chapters of mathematics. We shall account here after of one of the main procedures along which the task is performed nowadays. We shall not go deeply into the mathematical background, but for the theorem of Runge-Krarup. The proof of the theorem, even in the simplified form provided here, needs not to be fully understood, however its consequences and implications need to be clearly visualized and kept in mind by the reader.

Although other approaches are present in geodetic literature, all of them need to go through two fundamental steps: the first is linearization of the relations expressing the observables as functionals of the potential, the second is to remove from our unknown T pieces that approximate its long wavelength behaviour as well as its short wavelength behaviour, controlled by the so called topographic signal. Such concepts are properly developed in the chapter. The rest is basically collocation theory as a technique to solve the relevant boundary value problem left for the residual part of T.

4.2 The Anomalous Potential

We define the anomalous potential T (P) as

$$T\,(\mathrm{P}) = W\,(\mathrm{P}) - U\,(\mathrm{P}) \ . \tag{4.1}$$

F. Sansò et al., *Geodetic Heights*, Springer Geophysics,
https://doi.org/10.1007/978-3-030-10454-2_4

Let us immediately observe that, since we have placed the polar axis of $\mathcal{E}$ along the rotation axis of the Earth, the centrifugal potential V_c (P) (see (3.9)) contained in both W (P) and U (P) is the same; therefore (see (3.12) and (3.58))

$$T\,(\mathrm{P}) = V_{\mathrm{N}}\,(\mathrm{P}) - V_{\mathrm{e}}\,(\mathrm{P}) \ . \tag{4.2}$$

Hence, since V_e (P) is harmonic outside $\mathcal{E}$ and even inside, for a depth of thousands of kilometers, from (4.2) and recalling (3.20) we find that T satisfies the Poisson equation

$$\Delta T\,(\mathrm{P}) = -4\pi G\rho\,(\mathrm{P}) \ ; \tag{4.3}$$

in particular T (P) *is harmonic outside the masses*.

Now let us remark as an empirical fact that, at the level of the topographic layer, the following relations of maximum order of magnitude hold

$$\left|\frac{T}{W}\right| \lesssim 2\cdot 10^{-5} \ , \ \frac{|g-\gamma|}{\gamma} \lesssim 10^{-4} \ . \tag{4.4}$$

This implies that T can be usefully considered as a quantity small of the first order, when we have to linearize functionals of W. However we have to underline that, if we try to go inside the masses, the behaviour of W and U (continued as a harmonic function) diverge one from the other (see Sansò and Sideris 2013), so that

$$\frac{|g-\gamma|}{\gamma} \lesssim 4\cdot 10^{-3} \text{ at 20 km depth} \ ,$$
$$\frac{|g-\gamma|}{\gamma} \lesssim 2\cdot 10^{-2} \text{ at 100 km depth} \ .$$

It follows that, some 20/30 km below the Earth surface, the significance of T (P) is lost and one should not use any more the actual normal potential to approximate W (P).

Having characterized the order of magnitude of T close to the masses, let us look now at its behaviour at infinity, i.e. for r tending to ∞. From (4.2) and recalling (3.19) and (3.68), one has

$$\begin{aligned} T\,(\mathrm{P}) &= W\,(\mathrm{P}) - U\,(\mathrm{P}) = V_{\mathrm{N}}\,(\mathrm{P}) - V_{\mathrm{e}}\,(\mathrm{P}) = \\ &= \left[\frac{\mu}{r} + \mathcal{O}\left(\frac{1}{r^3}\right)\right] - \left[\frac{\mu}{r} + \mathcal{O}\left(\frac{1}{r^3}\right)\right] = \mathcal{O}\left(\frac{1}{r^3}\right) \ . \end{aligned} \tag{4.5}$$

Notice that the above asymptotic relation comes from our choice to have the same value of $\mu = GM$ for the actual and normal potential, to put the origin at the barycentre of the masses, also coinciding with the centre of the ellipsoid $\mathcal{E}$, and to make the z axis coinciding with the rotation axis of the Earth as well as with the polar axis of $\mathcal{E}$.

A consequence of (4.5) is that, outside any Brillouin sphere of radius R, one can write the series expansion

$$T\,(\mathrm{P}) = \sum_{n=2}^{+\infty} \sum_{m=-n}^{n} T_{nm} \left(\frac{R}{r}\right)^{n+1} Y_{nm}\,(\sigma)\ . \tag{4.6}$$

Note that T_{nm} have the same dimensions as T, while in the literature, e.g. Sansò and Sideris (2013), we often find non dimensional T_{nm}^{nd}, related to the present coefficients by $T_{nm}^{\mathrm{nd}} = \left(\frac{\mu}{R}\right)^{-1} T_{nm}$. Here R is any radius close to the mean Earth radius.

In (4.6) the first two degrees, $\mathcal{O}\left(\frac{1}{r}\right)$ and $\mathcal{O}\left(\frac{1}{r^2}\right)$, are missing, complying with the asymptotic behaviour (4.5).

Let us recall as well here that, paralleling (3.54), the coefficients T_{nm} are functions of the chosen value for R because

$$T_{nm} = \frac{1}{4\pi} \int T\,(R, \sigma)\, Y_{nm}\,(\sigma)\, \mathrm{d}\sigma\ . \tag{4.7}$$

Now if we take any other sphere with radius $R' > R$, we have obviously

$$T\left(R', \sigma\right) = \sum_{n=2}^{+\infty} \sum_{m=-n}^{n} T_{nm} \left(\frac{R}{R'}\right)^{n+1} Y_{nm}\,(\sigma)\ ; \tag{4.8}$$

on the other hand, $T\left(R', \sigma\right)$ will have as well its own harmonic coefficients T'_{nm} such that

$$T\left(R', \sigma\right) = \sum_{n=2}^{+\infty} \sum_{m=-n}^{n} T'_{nm} Y_{nm}\,(\sigma)\ . \tag{4.9}$$

Comparing (4.8) and (4.9), one finds

$$T'_{nm} = \left(\frac{R}{R'}\right)^{n+1} T_{nm}\ . \tag{4.10}$$

Formula (4.10) represents the upward continuation of the harmonic coefficients of T from the sphere S_R to the sphere $S_{R'}$; as we can see, the upward continued coefficients T'_{nm} become exponentially smaller than the corresponding T_{nm} as the degree increases. This corresponds to a smoothing of T as a function of σ, going from $T\,(R, \sigma)$ to $T\left(R', \sigma\right)$.

However the converse is also true, namely if we start from the outer sphere $S_{R'}$ and its coefficients T'_{nm} and we want to derive the coefficients T_{nm}, i.e. the potential T, we will have an exponential increase of T'_{nm}, namely

$$T_{nm} = \left(\frac{R'}{R}\right)^{n+1} T'_{nm}\ . \tag{4.11}$$

So if we have an imperfect knowledge of T'_{nm}, namely

$$T'_{0nm} = T'_{nm} + \varepsilon_{nm} \ , \tag{4.12}$$

and we try to use the erroneous T'_{0nm} to derive $\widehat{T}_{nm}$ through (4.11), we get

$$\widehat{T}_{nm} = \left(\frac{R'}{R}\right)^{n+1} T'_{nm} + \left(\frac{R'}{R}\right)^{n+1} \varepsilon_{nm} = T_{nm} + \left(\frac{R'}{R}\right)^{n+1} \varepsilon_{nm} \ . \tag{4.13}$$

As we can see, $\widehat{T}_{nm}$ are equal to the true T_{nm} plus an error exponentially amplified. For instance, if ε_{nm} are just random errors, uncorrelated, with constant variance

$$\sigma^2 \left(\varepsilon_{nm}\right) = \sigma_\varepsilon^2 \ ,$$

as it happens if $T_0\left(R', \sigma\right)$ is equal to $T\left(R', \sigma\right)$ plus a white noise on the sphere $S_{R'}$, the error contaminating our estimate $\widehat{T}\left(R', \sigma\right)$ is

$$\delta T\left(R, \sigma\right) = \sum_{n=2}^{N} \sum_{m=-n}^{n} \left(\frac{R'}{R}\right)^{n+1} \varepsilon_{nm} Y_{nm}\left(\sigma\right) \ . \tag{4.14}$$

When the summation in (4.14) goes up to infinity, δT becomes an awkward random variable, with infinite variance, because, recalling (3.42)

$$\sum_{m=-n}^{n} Y_{nm}^2\left(\sigma\right) = \left(2n+1\right) P_n\left(1\right) = \left(2n+1\right) \ ,$$

we find

$$\sigma^2\left(\delta T\right) = \sum_{n=2}^{N} \left(\frac{R'}{R}\right)^{2n+2} \left(2n+1\right) \sigma_\varepsilon^2 \xrightarrow[N\to\infty]{} +\infty \ . \tag{4.15}$$

This shows that, if we try to make a downward continuation from the sphere $S_{R'}$ to the sphere S_R, we can expect a lot of fuzzy numbers because of the increasing variability of errors with the degree. In fact it is well known that, even assuming that we know exactly T'_{nm}, there are potentials that are harmonic outside $S_{R'}$ but not down to S_R, so that formula (4.11) cannot be meaningfully applied (see Moritz 1980, Sansò and Venuti 2010).

Note that the determination of T is an essential tool to be able to perform the transformation between several types of geodetic heights, so we have at least to be aware of how it is done, to handle the necessary calculations involving T.

The determination of T, starting from the historical approach of Stokes (1849), has always been done by building a model $\widehat{T}$ which is harmonic in a domain larger than Ω, i.e. harmonic even inside the masses down to some reference surface S_0, for instance an internal sphere S_{R_0} also called a Bjerhammar sphere. Since then $\widehat{T}$ seems to be a kind of downward continuation of T inside the masses, which is not, it is necessary to clarify the situation by illustrating a cornerstone of Physical Geodesy, namely the Runge-Krarup theorem.

4.3 The Runge-Krarup Theorem: A Mathematical Intermezzo

This is essentially a theorem saying that if we have a closed surface S, with Ω the exterior of S, and another internal surface S_0, with Ω_0 the exterior of S_0, such that $\overline{\Omega} \subset \Omega_0$, then any function harmonic in Ω can be approximated as well as we like by a function harmonic in Ω_0.

When we want to obtain a result of "approximation", we need to specify what this term means for us, i.e. we have to fix some topology for the space of functions harmonic in Ω. This can be done, as it was done by Krarup, in very general terms, but here we shall content ourselves to use the space mostly applied in geodetic literature, namely the space functions harmonic in Ω and such that their trace on S is in $L^2(S)$, i.e.

$$u \in \mathcal{H}(\Omega) \;\Rightarrow\; \Delta u = 0 \text{ in } \Omega\,, \quad \int_S u^2 \mathrm{d}S < +\infty\,. \tag{4.16}$$

Such a space is a Hilbert space with scalar product

$$\langle u, v \rangle_{\mathcal{H}} = \int u\, v\, \mathrm{d}S \tag{4.17}$$

and with the norm derived by (4.17). So $u_N \to u$, i.e. u_N approximates u as well as we like, in $\mathcal{H}$ means

$$\lim_{N \to \infty} \int (u_N - u)^2\, \mathrm{d}S = 0\,. \tag{4.18}$$

Similarly we can define the space $\mathcal{H}_0 = \mathcal{H}(\Omega_0)$ as

$$u_0 \in \mathcal{H}_0 \;\Rightarrow\; \Delta u_0 = 0 \text{ in } \Omega_0\,, \quad \int_{S_0} u_0^2\, \mathrm{d}S < +\infty\,. \tag{4.19}$$

We note that $\forall u_0 \in \mathcal{H}_0$ we can define a function $u_{0\Omega}$ which is the restriction of u_0 to $\overline{\Omega}$ (remember the $\overline{\Omega} \subset \Omega_0$), i.e. we can define a restriction operator $\mathcal{R}_\Omega : \mathcal{H}_0 \to \mathcal{H}$ such that

$$u_{0\Omega} = \mathcal{R}_\Omega u_0 \;\Rightarrow\; u_{0\Omega}(\mathrm{P}) \equiv u_0(\mathrm{P})\;,\; \forall \mathrm{P} \in \Omega\,. \tag{4.20}$$

With the help of $\mathcal{R}_\Omega$ we can give the theorem a synthetic form.

Theorem 4.1 (Runge-Krarup)[1] *The set of functions*

$$U_0 \equiv \{\mathcal{R}_\Omega u_0\,,\ u_0 \in \mathcal{H}_0\} \tag{4.21}$$

is densely embedded in $\mathcal{H}$.

This exactly means that $\forall u \in \mathcal{H}$ we can find $u_{0N} \in \mathcal{H}_0$ such that $u - \mathcal{R}_\Omega u_{0N} \to 0$ in $\mathcal{H}$. Since $\mathcal{H}$ is a Hilbert space, the above is equivalent to saying that there is no element $v \neq 0 \in \mathcal{H}$ which is orthogonal to U_0, i.e.

$$\forall u_0 \in \mathcal{H}\,,\ \langle v, \mathcal{R}_\Omega u_0\rangle_{\mathcal{H}} = 0 \ \Rightarrow\ v = 0\,. \tag{4.22}$$

We sketch here a proof without too many pretenses of rigorousness.

Take

$$u_0\,(\mathrm{P}) = \frac{1}{\ell_{\mathrm{PQ}}}\,,\quad \mathrm{Q} \in B_0 \text{ (interior of } S_0)\,;$$

it is obvious that $u_0 \in \mathcal{H}_0$, $\forall \mathrm{Q} \in B_0$. But then if $v \in \mathcal{H}$ is such that

$$V^v\,(\mathrm{Q}) = \langle v, \frac{1}{\ell_{\mathrm{PQ}}}\rangle_{\mathcal{H}} = \int_S \frac{v\,(\mathrm{P})}{\ell_{\mathrm{PQ}}}\,\mathrm{d}S_{\mathrm{P}} = 0 \qquad \forall \mathrm{Q} \in B_0\,,$$

we have that the single layer potential $V^v\,(\mathrm{Q})$ has to be zero in B_0. Since $V^v\,(\mathrm{Q})$ is harmonic in both B (interior of S) and Ω, and B_0 is an open set contained in B, $V^v\,(\mathrm{Q}) \equiv 0$ in B by the unique continuation property; namely, two functions u, v harmonic in some set B, that are equal in an open subset B_0 of B, have to coincide in the whole of B (Sansò and Sideris 2013). Indeed $V^v\,(\mathrm{Q})$ and 0 are precisely in the above situation.

On the other hand, imposing some regularity hypothesis on the surface S, it is known that a single layer with an $L^2\,(S)$ surface density is continuous throughout all of $\mathcal{R}^3$. This implies that $V^v\,(\mathrm{Q}) \equiv 0$ on S too. But then $V^v\,(\mathrm{Q})$ is harmonic in Ω, continuous in $\overline{\Omega}$ and zero on its boundary S, i.e. it has to be zero everywhere in Ω by the well known maximum principle, i.e. (see Sansò and Sideris 2013)

$$\max_{\mathrm{Q}\in\overline{\Omega}} V^v\,(\mathrm{Q}) = \max_{\mathrm{Q}\in S} V^v\,(\mathrm{Q})\,,\quad \min_{\mathrm{Q}\in\overline{\Omega}} V^v\,(\mathrm{Q}) = \min_{\mathrm{Q}\in S} V^v\,(\mathrm{Q})\,.$$

Since, as for any single layer (MacMillan 1958),

$$v\,(\mathrm{Q}) = \frac{1}{2\pi}\left\{\frac{\partial V^v\,(\mathrm{Q})}{\partial n_+} - \frac{\partial V^v\,(\mathrm{Q})}{\partial n_-}\right\}\,,$$

[1] *Note*: on historical ground Runge proved a similar theorem for analytic functions; the theorem was extended to harmonic functions by T. Krarup.

with $\boldsymbol{n}_{\pm}$ indicating the external/internal limit of the derivative along the normal to S, we find that $v \equiv 0$ and the theorem is proved.

Note that in the theorem S_0 is any closed surface with an open interior domain B_0. Now take a sequence S_k of such surfaces, internal one to the other and shrinking to some point that we take as the origin O, so that

$$B_k \supset \overline{B}_{k+1} \quad \text{or} \quad \Omega_{k+1} \subset \overline{\Omega}_k \ . \tag{4.23}$$

If we consider the corresponding Hilbert spaces $\mathcal{H}_k$, we have indeed

$$u_{k+1} \in \mathcal{H}_{k+1} \ , \quad \mathcal{R}_{\Omega_k} u_{k+1} \in \mathcal{H}_k \ , \quad \mathcal{R}_{\Omega} \mathcal{R}_{\Omega_k} u_{k+1} = \mathcal{R}_{\Omega} u_{k+1} \in \mathcal{H} \ ,$$

so that

$$\mathcal{R}_{\Omega} \mathcal{H}_{k+1} \subset \mathcal{R}_{\Omega} \mathcal{H}_k \subset \cdots \subset \mathcal{H} \ , \tag{4.24}$$

each embedding being dense in $\mathcal{H}$. If we take the intersection

$$\bigcap_{k=0}^{+\infty} \mathcal{R}_{\Omega} \mathcal{H}_k = \mathcal{R}_{\Omega} \dot{\mathcal{H}} \ ,$$

we get the restriction to Ω of all the functions that are harmonic outside the origin,

$$u \in \dot{\mathcal{H}} \ \Rightarrow \ \Delta u = 0 \qquad \forall \mathrm{P} \neq \mathrm{O} \ .$$

$\dot{\mathcal{H}}$ has not a Hilbert space structure, but this is not important to us. More interesting is that, if we take the sequence of solid spherical harmonics

$$S_{nm}(r, \sigma) = \left\{ \left(\frac{R}{r} \right)^{n+1} Y_{nm}(\sigma) \right\} \ ,$$

we have indeed $S_{nm} \in \dot{\mathcal{H}}$ and so any finite linear combination of $\{S_{nm}\}$ is also in $\dot{\mathcal{H}}$, namely

$$u \in \mathcal{H}^F \equiv \left\{ \sum_{n=0}^{N} \sum_{m=-n}^{n} a_{nm} \left(\frac{R}{r} \right)^{n+1} Y_{nm}(\sigma) \right\} \Rightarrow u \in \dot{\mathcal{H}} \ .$$

In particular, what is of utmost importance for us is the following corollary of the Runge-Krarup theorem.

Corollary *The subspace $\mathcal{R}_{\Omega} \mathcal{H}^F$ is densely embedded into $\mathcal{H}(\Omega) \equiv \mathcal{H}$.*

This is rather obvious because writing the elements of $\mathcal{H}^F$ in the form

$$u \in \mathcal{H}^F \Rightarrow u = \sum_{n=0}^{N} \sum_{m=-n}^{n} a'_{nm} \left(\frac{R}{r} \right)^{n+1} Y_{nm}(\sigma) \ , \tag{4.25}$$

for some finite N, we see that, $\forall\varepsilon > 0$,

$$\begin{aligned} &\mathcal{R}_{\Omega_\varepsilon}\mathcal{H}^F \subset \mathcal{H}(\Omega_\varepsilon) \\ &(\Omega_\varepsilon \equiv \{r \geq \varepsilon\}) \ , \end{aligned} \tag{4.26}$$

the embedding being dense, because $\left\{\left(\frac{R}{r}\right)^{n+1} Y_{nm}(\sigma)\right\}$ is an orthogonal, complete sequence in $\mathcal{H}(\Omega_\varepsilon)$. But then, for ε sufficiently small so that $\Omega_\varepsilon \supset \overline{\Omega}$, we get

$$\mathcal{R}_\Omega \mathcal{R}_{\Omega_\varepsilon}\mathcal{H}^F = \mathcal{R}_\Omega\mathcal{H}^F \subset \mathcal{R}_\Omega\mathcal{H}(\Omega_\varepsilon) \subset \mathcal{H} \ , \tag{4.27}$$

each embedding being dense into the next.

Remark The neat result of the above mathematical discussion is that, given any potential $T \in \mathcal{H}(\Omega)$, we can find a $\widehat{T}_M \in \mathcal{H}^F$ that approximates T better than a prefixed level ε, or said in another way

$$\forall T \in \mathcal{H}(\Omega) \ , \ \forall\varepsilon > 0; \ \exists N_\varepsilon, \ \left\{\widehat{T}_{nm}; \ n \leq N_\varepsilon\right\} \ \Rightarrow \ \left\|T - \widehat{T}_M\right\|_{\mathcal{H}} < \varepsilon$$

with

$$\widehat{T}_M = \sum_{n=2}^{N_\varepsilon}\sum_{m=-n}^{n} \widehat{T}_{nm}\left(\frac{R}{r}\right)^{n+1} Y_{nm}(\sigma) \ , \tag{4.28}$$

for some fixed Bjerhammar radius R. A finite sum of the type (4.28) is called a global model of the anomalous potential.

It is clear from the above discussion that a global model $\widehat{T}_M$ is not a downward continuation of T; in addition there are many $\widehat{T}_M$ satisfying the same approximation level.

Actually we use validated models up to degree 2159 (e.g. EGM2008 Pavlis et al. 2012, 2013 or EIGEN-6C4 Förste et al. 2014, Shako et al. 2014), which have a resolution of about 10 km on the Earth surface. The use of these models however does imply calculations with about $4.6 \cdot 10^6$ coefficients T_{nm}, which is feasible but computationally heavy. Even more, if we wanted to reach the resolution of 1 km on S, we should use a model with 100 times coefficients than the above, what seems not particularly economical from the computational point of view.

So in our description on how to represent T, we shall always include a component of the type global model, but we shall leave to other methods a representation of the high resolution particulars of this potential.

4.4 Optimal Degree of Global Models, or Smoothing by Truncation

The decision to represent T by a global model $\widehat{T}_M$ leaves open the question of the degree N up to which $\widehat{T}_M$ should be developed and of which method should be employed to estimate the specific coefficients $\widehat{T}_{nm}$.

As for the second point, we could say that $\widehat{T}_{nm}$ are obtained by solving a specific boundary value problem, as it will be illustrated into the the next sections, while the first point will be discussed here.

In any event we assume that we have a tool that from some data is producing estimates

$$\widehat{T}_{nm} = T_{nm} + \varepsilon_{nm} \ , \tag{4.29}$$

where ε_{nm} are the estimation errors and Eq. (4.29) refers to some suitable radius R.

We call power of the degree n (or full power degree variances) the index

$$C_n\left(T\right) = \sum_{m=-n}^{n} T_{nm}^2 \tag{4.30}$$

and degree variances (we shall explain this term in the next section)

$$\sigma_n^2\left(T\right) = \frac{C_n\left(T\right)}{2n+1} \ . \tag{4.31}$$

Let us note that the quantity

$$\frac{1}{4\pi}\int T^2\left(R,\sigma\right) \mathrm{d}\sigma = \sum_{n=2}^{+\infty} C_n\left(T\right) < +\infty \tag{4.32}$$

has to be finite, implying that $C_n\left(T\right) \to 0$ for $n \to \infty$. Among others, this constitutes a necessary condition to be imposed on R. For instance, for EGM2008 R is close to be equal to a, the equatorial radius. Indeed we do not know exactly $C_n\left(T\right)$, but we can have a guess of them, $\widehat{C}_n\left(T\right) = C_n\left(\widehat{T}\right)$, by using $\widehat{T}_{nm}$. To be more precise, one could observe that the estimator $\widehat{C}_n\left(T\right) = \sum_{m=-n}^{n} \widehat{T}_{nm}^2$ is biased and $E\left\{\widehat{C}_n\left(T\right)\right\} = C_n\left(T\right) + \sum_{m=-n}^{n} \sigma^2\left(\varepsilon_{nm}\right)$; but then, if we assume to know $\sigma^2\left(\varepsilon_{nm}\right)$, we can easily construct the unbiased estimator $\overline{C}_n\left(T\right) = \widehat{C}_n\left(T\right) - \sum_{m=-n}^{n} \sigma^2\left(\varepsilon_{nm}\right)$.

It happens that by inspecting the plot of $\widehat{C}_n\left(T\right)$, for example computed from the EGM2008 coefficients (see Fig. 4.1), one can derive an empirical law for them (see

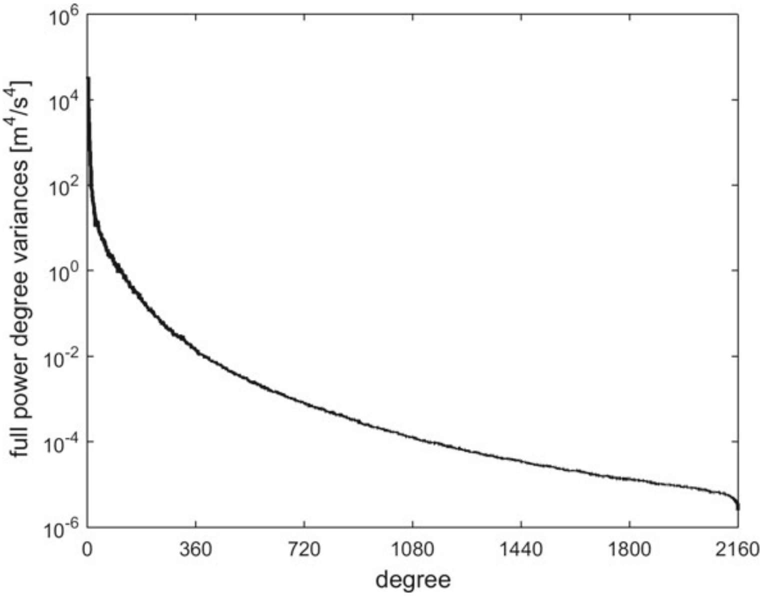

Fig. 4.1 The empirical full power degree variances of EGM2008

discussion in Sansó and Sideris 2013, Sect. 3.8). Older but evergreen models of C_n are also available, like those of Kaula (1966, 2000) and Tscherning and Rapp (1974).

Therefore we could say that, although we do not know the exact T_{nm}, we have a law for $C_n(T)$. This helps us to define the mean square omission error at degree N, i.e.

$$\mathcal{O}\mathcal{E}_N^2 = \sum_{n=N+1}^{+\infty} C_n(T) \ ; \tag{4.33}$$

this is the error that we commit if instead of T we use just its development up to degree N. In fact if we split T (at the level of the sphere S_R) into

$$T = \sum_{n=2}^{N} \sum_{m=-n}^{n} T_{nm} Y_{nm}(\sigma) + \sum_{n=N+1}^{+\infty} \sum_{m=-n}^{n} T_{nm} Y_{nm}(\sigma) = T_{(N)} + T^{(N)} \ , \tag{4.34}$$

we see that, thanks to the orthogonality property of spherical harmonics, see (3.47),

$$\frac{1}{4\pi} \int \left[T - T_{(N)}\right]^2 \mathrm{d}\sigma = \frac{1}{4\pi} \int \left[T^{(N)}\right]^2 \mathrm{d}\sigma = \sum_{n=N+1}^{+\infty} C_n(T) = \mathcal{O}\mathcal{E}_N^2 \ . \tag{4.35}$$

We note as well that $\mathcal{O}\mathcal{E}_N$ is a decreasing function of N and it has to tend to 0 for $N \to \infty$, because of condition (4.32).

Now we observe that, recalling (4.29), one has

$$e(T) = T - \widehat{T}_M = T_{(N)} - \widehat{T}_M + T^{(N)} = \\ = \sum_{n=2}^{N} \sum_{m=-n}^{n} \varepsilon_{nm} Y_{nm}(\sigma) + \sum_{n=N+1}^{+\infty} T_{nm} Y_{nm}(\sigma) \ . \tag{4.36}$$

The mean square error of $e(T)$ over the unit sphere is then

$$\frac{1}{4\pi} \int e^2(T)\, \mathrm{d}\sigma = \sum_{n=2}^{N} \sum_{m=-n}^{n} \varepsilon_{nm}^2 + \mathcal{O}\mathcal{E}_N^2 \ .$$

As we can see, this is still a random variable because it depends on ε_{nm}^2; so we can reasonably define a total error $\mathcal{E}_N^2$ as

$$\mathcal{E}_N^2 = E\left\{\frac{1}{4\pi} \int e^2(T)\, \mathrm{d}\sigma\right\} = \sum_{n=2}^{N} \sum_{m=-n}^{n} \sigma^2(\varepsilon_{nm}) + \mathcal{O}\mathcal{E}_N^2 \ . \tag{4.37}$$

This is the total (mean square) error that we expect by substituting T with $\widehat{T}_M$. As we can see, it is in part due to the propagation of the estimation errors ε_{nm}, in part to the omission of the coefficients by truncating at degree N. The first term in (4.37) is called commission error

$$\mathcal{C}\mathcal{E}_N^2 = \sum_{n=2}^{N} \sum_{m=-n}^{n} \sigma^2(\varepsilon_{nm}) \ . \tag{4.38}$$

As we said, it represents the effect of the estimation errors, up to degree N, which ultimately descend from the presence of measurement noise in the original data that have allowed to estimate the T_{nm} coefficients.

The terms

$$\sigma_n^2(\varepsilon) = \sum_{m=-n}^{n} \sigma^2(\varepsilon_{nm}) \tag{4.39}$$

are called error degree variances and we have

$$\mathcal{C}\mathcal{E}_N^2 = \sum_{n=2}^{N} \sigma_n^2(\varepsilon) \ . \tag{4.40}$$

As it is obvious, $\mathcal{C}\mathcal{E}_N^2$ is an increasing function of N and if for instance $\sigma^2(\varepsilon_{nm}) = \sigma_0^2$, as it happens when ε_{nm} are just white noise, then $\sigma_n^2(\varepsilon) = (2n+1)\,\sigma_0^2$ and indeed $\mathcal{C}\mathcal{E}_N^2 \to \infty$ when $N \to \infty$. This however is not the general case.

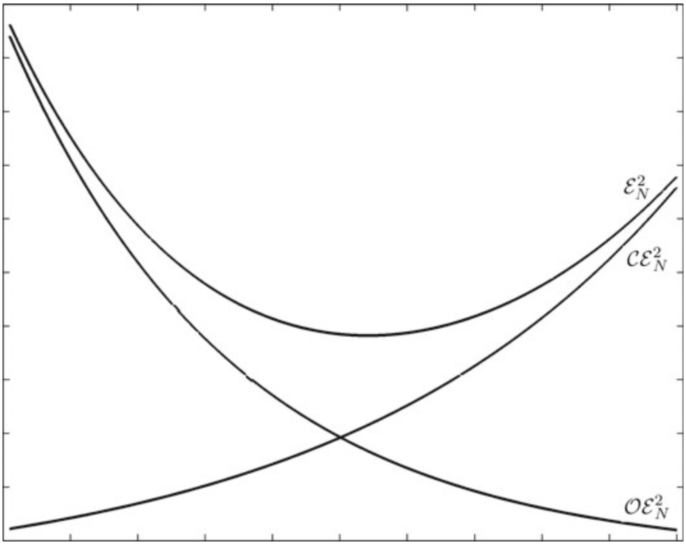

Fig. 4.2 The typical shape of $\mathcal{OE}_N^2$, $\mathcal{CE}_N^2$ and $\mathcal{E}_N^2$

Yet $\mathcal{E}_N^2$ as the sum of $\mathcal{CE}_N^2$ and $\mathcal{OE}_N^2$ will have a typical behaviour as shown in see Fig. 4.2, namely $\mathcal{E}_N^2$ will have a minimum at the degree $\overline{N}$ where the commission and omission errors cross. $\overline{N}$ is indeed our optimal choice for N, because the total error is minimum at this degree.

We note that the above condition implies

$$\sigma^2_{\overline{N}}(\varepsilon) = C_{\overline{N}}(T) \ ; \tag{4.41}$$

for instance, if $\sigma^2(\varepsilon_{nm}) = \sigma_0^2$, then $\sigma^2_{\overline{N}}(\varepsilon) = \left(2\overline{N} + 1\right)\sigma_0^2$ and the optimal criterion is

$$\sigma_0^2 = \frac{C_{\overline{N}}(T)}{\left(2\overline{N} + 1\right)} = \sigma^2_{\overline{N}}(T) \ .$$

This solves the posed problem. As a realistic example in Fig. 4.3 let us display the plot of potential error degree variances of a satellite model, when both the estimate of $\widehat{T}_{nm}$ is unregularized and it is conditioned by using $C_n(T)$ (see next section). As we can see, the optimal $\overline{N}$ in this case is around $\overline{N} = 250$.

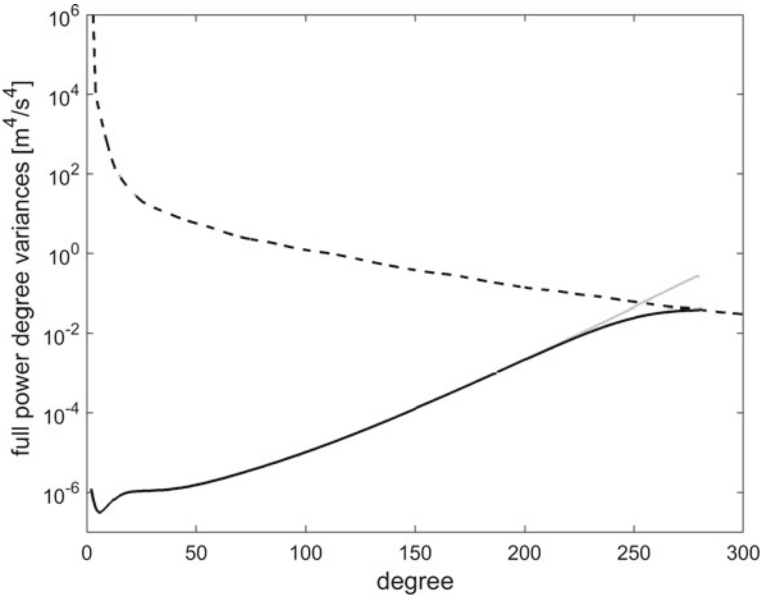

Fig. 4.3 Estimated error degree variances of a model from the ESA-GOCE mission with and without regularization, respectively in black and gray Brockmann et al. (2014), [Brockmann, personal communication, 2015]. The dash line shows the EGM2008 degree variances

4.5 Collocation Theory, or Smoothing by Prior Information

As in the previous section, we assume to know $T_{0nm} = T_{nm} + \varepsilon_{nm}$ up to some degree N, as well as the full power degree variances (4.30) and $\sigma_n^2(T)$. For the moment let us assume further on that ε_{nm} are independent from one another.

We want to state a criterion to estimate $\widehat{T}_{nm}$ that exploits, beyond the "observations" $\{T_{0nm}\}$, also the prior knowledge given by (4.30). In collocation theory this is done by establishing the minimum principle

$$\left\{\widehat{T}_{nm}\right\} = \text{ArgMin}\left\{\sum_{n=2}^{N}\sum_{m=-n}^{n}\frac{\left(T_{0nm}-\widehat{T}_{nm}\right)^2}{\sigma^2\left(\varepsilon_{nm}\right)} + \sum_{n=2}^{N}\sum_{m=-n}^{n}\frac{\widehat{T}_{nm}^2}{\sigma_n^2(T)}\right\} . \tag{4.42}$$

As we can see, this is composed by a first quadratic functional that is essentially the same sum of squares as in least squares theory, while the second part of the functional has the purpose of stabilizing the solution as in Tikhonov theory. We observe also that this second functional would be the natural extension of least squares if we interpreted the prior information in terms of pseudo-observation equations

$$\begin{aligned} &T_{0nm} = 0 + \eta_{nm} \qquad \forall m\,,\ n > N \\ &E\left\{\eta_{nm}\right\} = 0\,,\ \ \sigma^2\left(\eta_{nm}\right) = \sigma_n^2(T)\ . \end{aligned}$$

This is also typical of a Bayesian interpretation in which every variable is stochastic by assumption.

All in all the principle (4.42) has an obvious, but significant, solution

$$\begin{cases} \widehat{T}_{nm} = \dfrac{\sigma_n^2(T)}{\sigma_n^2(T) + \sigma^2(\varepsilon_{nm})} T_{0nm} & (\forall m\ ,\ n \le N) \\ \widehat{T}_{nm} = 0 & (\forall m\ ,\ n > N) \end{cases} . \tag{4.43}$$

As we can see, the analogy to the Wiener-Kolmogorov filter is very strong (Sansó and Sideris 2013, Sect. 5.4).

Also here we truncate the estimated model at degree N, because there is no interaction between $\widehat{T}_{nm}$ $(n > N)$ and the observations. The coefficients of degree $n \le N$ are rescaled and not just equal to T_{0nm}. In particular at low degrees where we expect $\sigma_n^2(T) \gg \sigma^2(\varepsilon_{nm})$, we have $\widehat{T}_{nm} \sim T_{0nm}$, while for high degrees, where $\sigma_n^2(T) \to 0$ and $\sigma^2(\varepsilon_{nm})$ might even tend to a constant or in any way is expected to go zero much slower than $\sigma_n^2(T)$, we have that $\widehat{T}_{nm} \to 0$ much faster than T_{0nm}.

Remark There are significant examples in which T_{0nm} are directly derived from space observations. In these cases a stochastic model with independent estimation errors is too unrealistic; on the contrary the ε_{nm} have a fully populated covariance matrix C_ε. So if we reorganize T_{nm} in a vector $\boldsymbol{T}$ with some ordering and we introduce the diagonal matrix

$$K = \text{diag}\left\{\sigma_n^2(T)\right\} ,$$

meaning that $\sigma^2(T_{nm}) = \sigma_n^2(T)$, $(m = -n, \ldots, 0, \ldots, n)$, the principle (4.42) is extended to

$$\min\left\{\widehat{\boldsymbol{T}}^T K^{-1}\widehat{\boldsymbol{T}} + \left(\boldsymbol{T}_0 - \widehat{\boldsymbol{T}}\right)^T C_\varepsilon^{-1}\left(\boldsymbol{T}_0 - \widehat{\boldsymbol{T}}\right)\right\} . \tag{4.44}$$

The variation equation of (4.44) is

$$\left(K^{-1} + C_\varepsilon^{-1}\right)\widehat{\boldsymbol{T}} = C_\varepsilon^{-1}\boldsymbol{T}_0$$

and its solution is given by

$$\widehat{\boldsymbol{T}} = K\left(K + C_\varepsilon\right)^{-1}\boldsymbol{T}_0 .$$

Such a formula is particularly nice because we do not need to invert two times the large matrix C_ε.

Anyway, what we have done up to now is basically to show how to filter a global model, where the coefficients themselves are considered as observations. On the other hand, we need a more general tool to treat the estimation of $\widehat{T}$ from a general set of observations; this is particularly important because the main sources of information on T come from gravity measurements and not from coefficients.

So to generalize the above discussion, we assume now to have a set of observations

$$m_{0i} = M_i\,(T) + \eta_i \qquad i = 1, 2, \ldots, M\ , \tag{4.45}$$

where $M_i\,(T)$ are linear functionals of T, namely numbers that linearly depend on T. We shall see in the next section how to write $M_i\,(T)$ for the main observables available.

We want to directly estimate $\widehat{T}\,(\mathrm{P})$ at any point P in the harmonicity domain of $\widehat{T}$, recalling that by using the Runge-Krarup theorem $\widehat{T}$ is taken as harmonic down to a Bjerhammar sphere,

$$\begin{aligned}\widehat{T} &= \frac{\mu}{R}\sum_{n=2}^{+\infty}\sum_{m=-n}^{n}\widehat{T}_{nm}\left(\frac{R}{r}\right)^{n+1} Y_{nm}\,(\sigma) = \\ &= \frac{\mu}{R}\sum_{n=2}^{+\infty}\sum_{m=-n}^{n}\widehat{T}_{nm} S_{nm}\,(r, \sigma)\ .\end{aligned}$$

The new optimization principle then becomes

$$\min\left\{\sum_{i=1}^{M}\frac{\left[m_{0i} - M_i\left(\widehat{T}\right)\right]^2}{\sigma^2_{\eta_i}} + \sum_{n=2}^{+\infty}\sum_{m=-n}^{n}\frac{\widehat{T}^2_{nm}}{\sigma^2_n\,(T)}\right\}\ . \tag{4.46}$$

Leaving the proofs e.g. to the text Sansó and Sideris (2013, Sect. 5.5), we directly report here the solution of the principle (4.46). This can be obtained in terms of the so called covariance functions, hereafter defined

$$C\,(\mathrm{P}, \mathrm{Q}) = \sum_{n=2}^{+\infty}\sum_{m=-n}^{n}\sigma^2_n\,(T)\, S_{nm}\,(r_{\mathrm{P}}, \sigma_{\mathrm{P}})\, S_{nm}\left(r_{\mathrm{Q}}, \sigma_{\mathrm{Q}}\right)\ , \tag{4.47}$$

$$C\,(\mathrm{P}, M_i) = \sum_{n=2}^{+\infty}\sum_{m=-n}^{n}\sigma^2_n\,(T)\, S_{nm}\,(r_{\mathrm{P}}, \sigma_{\mathrm{P}})\, M_i\left(S_{nm}\left(r_{\mathrm{Q}}, \sigma_{\mathrm{Q}}\right)\right)\ , \tag{4.48}$$

$$C\,(M_k, M_i) = \sum_{n=2}^{+\infty}\sum_{m=-n}^{n}\sigma^2_n\,(T)\, M_k\,(S_{nm}\,(r_{\mathrm{P}}, \sigma_{\mathrm{P}}))\, M_i\left(S_{nm}\left(r_{\mathrm{Q}}, \sigma_{\mathrm{Q}}\right)\right)\ . \tag{4.49}$$

The optimal solution is then obtained by the formula

$$\widehat{T}\,(\mathrm{P}) = \sum_{i,k=1}^{M} C\,(\mathrm{P}, M_i)\left\{C\,(M_i, M_k) + \sigma^2_{\eta_i}\delta_{ik}\right\}^{(-1)} m_{0k}\ . \tag{4.50}$$

An important feature of the theory is that one can also compute the variance of the estimation error of $\widehat{T}\,(\mathrm{P})$, namely

$$\mathcal{E}^2(\text{P}) = C(\text{P},\text{P}) - \sum_{i,k=1}^{M} C(\text{P}, M_i)\left\{C(M_i, M_k) + \sigma^2_{\eta_i}\delta_{ik}\right\}^{(-1)} C(\text{P}, M_k) \ . \quad (4.51)$$

We cannot go here into the intricacy of the full estimation process and of its numerical implementation. However we shall make some comment on the use of (4.50) in a local area and on the remove-restore principle.

Remark (Collocation in a local refinement environment)

Assume we have global data sets, like satellite tracking or satellite gravity missions or just gravity observations all over the surface S; assume that we have solved the problem of estimating a global model $\widehat{T}_M$ from such global data sets, with a resolution regulated by its maximum degree. Now we have more observations, written as in (4.45), concentrated in a local area and we want to improve our knowledge of T in that area.

As a first operation we can remove the global information putting

$$m_{0i} = M_i(T_M + \delta T) + \eta_i = M_i(T_M) + M_i(\delta T) + \eta_i \ , \quad (4.52)$$

computing $M_i(T_M)$ and removing it from m_{0i}. We are left now with the unknown δT that represents the local behaviour of T. Before estimating δT with a formula like (4.50), it is usually convenient to further smooth the data by exploiting the information coming from a local Digital Terrain Model (DTM). In fact the fine variations of the topography produce a quite significant potential with an important content of high frequency. This is done by what is called the residual terrain correction and its potential δT_{tc}. In fact in general we have a much higher resolution in the knowledge of the topography than for other gravity measurements. This correction is called residual because we know that the long wavelength effect of topography is already captured by the model T_M. So in δT_{tc} we have to put the effect of the masses between the actual terrain and a smoothed version of it. This is usually done by discretizing the masses in prisms (Fig. 4.4).

So we now rewrite (4.52) as

$$\delta m_{0i} = m_{0i} - M_i(T_M) = M_i(\delta T_{\text{tc}}) + M_i\left(\overline{\delta T}\right) + \eta_i \ , \quad (4.53)$$

where $M_i(\delta T_{\text{tc}})$ is computed and removed from the known term (Sansò and Sideris 2013, Sect. 4.4). We are finally left with an unknown $\overline{\delta T}$, where long and short wavelengths have been removed or de-potentiated. It is now to $\overline{\delta T}$ that a collocation solution is applied. At the end we restore all the terms and $\widehat{T}$ is estimated in the area where we have added new measurements by the formula

$$\widehat{T} = \widehat{T}_M + \delta T_{\text{tc}} + \overline{\delta T} \ . \quad (4.54)$$

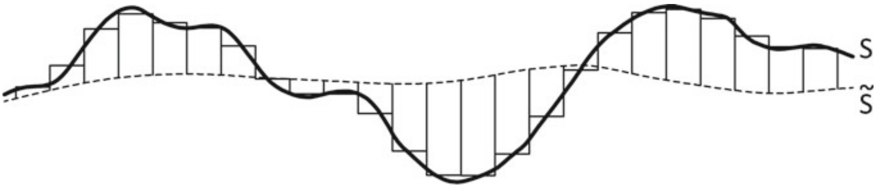

Fig. 4.4 Geometry of the terrain correction, i.e. potential generated by prisms including the masses between the actual topography S and a smoothed topography $\widetilde{S}$

4.6 On the Relation Between Potential and the Surface Gravimetric Observables

Let us remark first of all that W (P), and whence T (P), is related to several spatial observables that we shall not discuss in the present context, because this would require to enter into subjects of satellite dynamics that are far away from the main purpose of the book.

We shall mention however that due to the structure of satellite observation equations and the significant smoothing of T (P) at satellite altitudes, it comes natural that the processing of spatial geodesy data gives as an output the estimate of the harmonic coefficients $\{T_{nm}\}$ of T up to some maximum degree N. At present, with the data of the CHAMP, GRACE and GOCE missions, N can be taken to be as high as $N = 300$.

Given that, we come to the main observables on the surface of the Earth, that provide the major information on the gravity field.

4.6.1 Gravimetry

In principle gravimetry, in its absolute version, provides the measurement of the modulus of the gravity vector on continental areas.

In addition to absolute gravimeters, we have relative gravimeters that can observe the difference of gravity between two points. The old spring gravimeter schematized in Fig. 4.5a is nowadays substituted by superconducting gravimeters, see Fig. 4.5b, which are in principle able of measuring g with an accuracy of 1 μGal (10^{-6} Gal). Such relative measurements are arranged in networks constituted by closed loops, which, thanks to their redundancy, allow to estimate various biases in the observations.

All in all, also correcting the time variable part of g, we end up with a set of points $\{P_i\}$ (gravity stations), where we know

$$g\,(\mathrm{P}) = |\boldsymbol{g}\,(\mathrm{P})| = |\gamma\,(\mathrm{P}) + \nabla T\,(\mathrm{P})|\,. \tag{4.55}$$

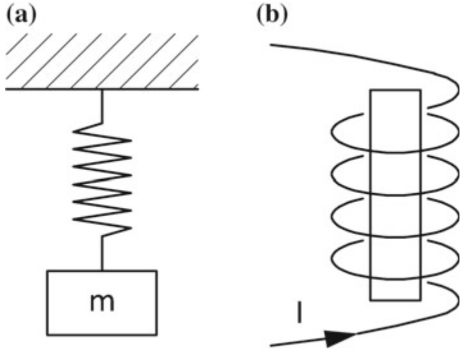

Fig. 4.5 The principle of measurement of gravity, **a** spring gravimeter (not in use any more), **b** superconducting gravimeter

The final accuracy with which we know g (P) can be deemed to be somewhere between 0.1 and 0.01 mGal, which is certainly suitable for geodetic purposes. The actual data set at Bureau Gravimétrique International (BGI) comprises some 10^6 data, on continental areas, with a significant variability of points density. In particular South America, Africa and Antarctica are rather poorly covered by gravity observations.

Let us note explicitly that although nowadays gravity measurements are accompanied by the 3D ellipsoidal coordinates of P, given with sufficient accuracy by GPS observations, this is not the case for the largest part of the data existing in the BGI archives, where P_i have known horizontal coordinates but unknown ellipsoidal height h.

This imposes a particular manipulation of the equations, during linearization, which is characteristic of Physical Geodesy. We only mention that beyond continental gravity measurements, we have a marine gravity data set of direct gravity observations. This however is much less dense than the first and its accuracy is much lower (between 1 and 5 mGal). Furthermore, on oceans we have the more important data set of radar altimetry that we shall discuss hereafter.

Finally we have as well gravity data from aerogravimetry, in part on land and in part on sea; however it is only recently that such data have an accuracy below the mGal level and in any way we can think that they have been processed to provide grids of gravity values on the surface.

4.6.2 Levelling Combined with Gravimetry

Levelling is a kind of classical geodetic measurement, that is schematized in Fig. 4.6 for one of its constitutive steps. As we can see from the figure, the typical reading of a step of levelling is δL, which represents the projection of the vector $\boldsymbol{r}_{AB} = \boldsymbol{r}_B - \boldsymbol{r}_A$ in the vertical direction $\boldsymbol{n}$ at the midpoint M between A and B, namely

$$\delta L = \boldsymbol{n} \cdot \boldsymbol{r}_{AB} = -\frac{\boldsymbol{g}_M}{g_M} \cdot \boldsymbol{r}_{AB} \; . \tag{4.56}$$

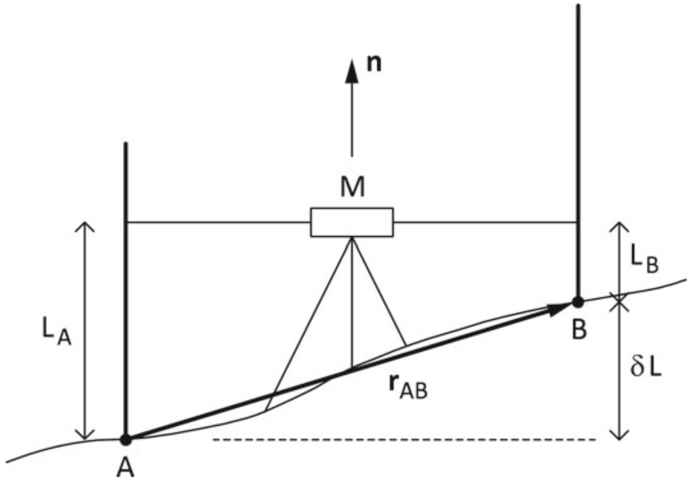

Fig. 4.6 The elementary operation of levelling with two vertical rods and two readings L_A, L_B from the middle in horizontal directions; the observation is $\delta L = L_A - L_B$

Such a measurement can be combined with the value of gravity at M, providing

$$g_M \delta L = -\boldsymbol{g}_M \cdot \boldsymbol{r}_{AB} = -\nabla W(M) \cdot \boldsymbol{r}_{AB} \cong W_A - W_B = -\delta W . \tag{4.57}$$

The last step in (4.57) is justified by the fact that in each levelling station the distance between A and B is typically between 100 and 200 m, so that $\boldsymbol{r}_{AB}$ can be considered to have an infinitesimal length, on a planetary scale.

We shall discuss in a dedicated chapter the levelling operations and their analytical formulation. Here we are interested in the fact that by adding the relation (4.57) along levelling lines, we can arrive to connect all the points of a certain region to an origin point P_0, which ideally could be placed on the geoid. This means that all over the surface of the continents we could arrive to know

$$W(P) = W_0 + \int_{P_0}^{P} dW = W_0 - \int_{P_0}^{P} g\, \delta L . \tag{4.58}$$

For several reasons, including the fact that it is difficult to state that P_0 is on the geoid, even if it is placed at a tide gauge, we could say that $W(P)$ is know, but for an additive constant. Even more such a constant is certainly different for different patches connected to different origins. So for the moment we shall overlook the problem of determining such constants, that will be treated in the last chapter of the book, and we shall assume that we know $W(P)$ at any point on land.

4.6.3 Radar-Altimetry on the Oceans

As already illustrated in Sect. 3.6, a radar-altimeter measures the height of a satellite on the ocean. The position of the radar-altimeter in space is known by GPS tracking at centimetric level; subtracting the former from the latter, we are left with the ellipsoidal height of the sea.

The footprint of the radar beam is regulated with a diameter between 100 and 1000 m, in such a way as to average the wave motion. Tides and barometric effects are modelled and subtracted from the observed height of the sea, so that by averaging in time we arrive at the (quasi) stationary sea surface. This one, in turn, is the sum of the geoid and the mean dynamic ocean topography η, which is related to geostrophic currents and provided by oceanographic models.

All together, one has the observation equation for H_0, i.e.

$$H_0 = h_{\mathrm{R}} - (N + \eta_t) + \nu \ , \tag{4.59}$$

with ν the observation noise, η_t the time dependent dynamic ocean topography, N the geoid undulation, h_{R} the ellipsoidal height of the radar-altimeter, see Fig. 4.7.

All the terms in (4.59), but for the unavoidable measurement error ν and the geoid N, are known or modelled. Hence (4.59) can be used to provide estimates of N over the ocean.

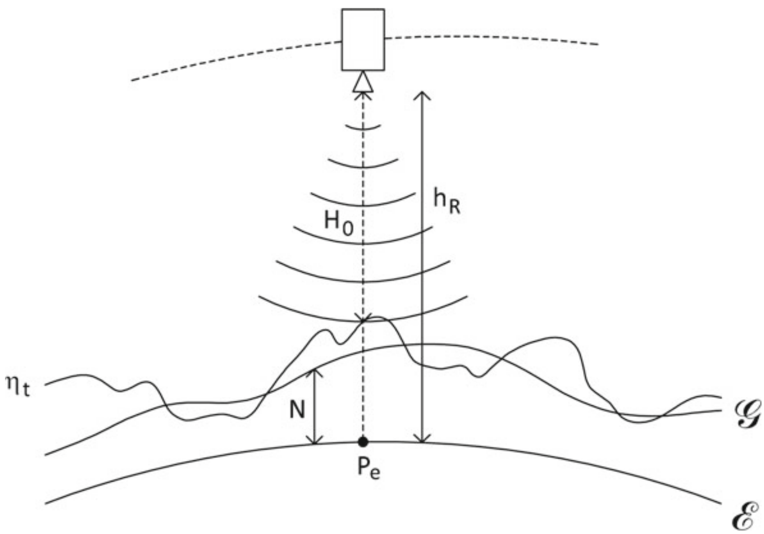

Fig. 4.7 Geometry of radar-altimetric observations: $\mathcal{E}$ ellipsoid, $\mathcal{G}$ geoid, N geoid undulation, η_t dynamic ocean topography, h_{R} ellipsoidal height of the radar-altimeter, H_0 radar-altimeter observation

Summarizing, and with a certain degree of abstraction, we could say that the main observables of Physical Geodesy can provide

$$\text{on continents: } W(\mathrm{P}) \;,\quad g(\mathrm{P}) \tag{4.60}$$

$$\text{on oceans: } \quad N(\mathrm{P_e}) \;. \tag{4.61}$$

We recall that in (4.60) we know the horizontal coordinates (λ, φ) of P, but usually not its ellipsoidal height; in (4.61) $\mathrm{P_e}$ is on the ellipsoid and its (λ, φ) coordinates are known. However, note that from N one has also the third coordinate, namely (λ, φ, N) of the point $\mathrm{P}_{\mathcal{G}}$. So, recalling that on the geoid $W(\mathrm{P})$ has the known value W_0, one could substitute (4.61) with the relation

$$W(\mathrm{P}_{\mathcal{G}}) = W_0 \;, \tag{4.62}$$

where $\mathrm{P}_{\mathcal{G}}$ has known ellipsoidal coordinates.

As a closing remark of the section, we recall again that further important sources of information on the gravity field are space geodetic methods, providing global models up to some degree N (nowadays we have $N \cong 300$), and digital terrain models, basically used at a local level to smooth the gravity field by residual terrain corrections.

To put together all this information is not an easy task; at a conceptual level, this is done by the so called Geodetic Boundary Value Problem theory that we shall review in the next section, especially with the purpose of providing the linearized version of the Eqs. (4.60) and (4.62), where the unknown field is not any more $W(\mathrm{P})$, but the anomalous potential $T(\mathrm{P})$.

4.7 The Geodetic Boundary Value Problem (GBVP)

In principle (4.60) and (4.62) can be put together, to formulate the following BVP:

to find $W(\mathrm{P}) = V(\mathrm{P}) + \frac{1}{2}\omega^2\rho^2$, with $V(\mathrm{P})$ regular harmonic in Ω, the exterior of surface S,

$$\begin{cases} \Delta V = 0 & \text{in } \Omega \\ V = \mathcal{O}\left(\dfrac{1}{r}\right) & r \to \infty \end{cases} \;; \tag{4.63}$$

the surface S is composed by two patches, that we call L and O and correspond respectively to Land and Ocean,

$$S = L \cup O \;.$$

The surface O is geometrically known

$$\boldsymbol{r} \in O \quad \Rightarrow \quad \boldsymbol{r} = (\lambda, \varphi, N) \;,\quad N = N(\lambda, \varphi) \tag{4.64}$$

and on O we know that the potential is constant, i.e.

$$W(\mathrm{P}) = W_0 \qquad (\mathrm{P} \in O) \;; \tag{4.65}$$

on the contrary, the surface L is unknown

$$\boldsymbol{r} \in L \quad \Rightarrow \quad \boldsymbol{r} = (\lambda, \varphi, h) \;, \quad h = h(\lambda, \varphi) \;\; \text{(unknown)} \tag{4.66}$$

but on L both the gravimetric quantities are known, i.e.

$$\begin{cases} W(\mathrm{P}) = W[\lambda, \varphi, h(\lambda, \varphi)] = W_0(\lambda, \varphi) \\ g(\mathrm{P}) = |\nabla W[\lambda, \varphi, h(\lambda, \varphi)]| = g_0(\lambda, \varphi) \end{cases} . \tag{4.67}$$

As such, this BVP can be classified as:

- a BVP for the Laplace operator in a space of regular harmonic functions (see (4.63)),
- a partially fixed boundary (see (4.64)) Dirichlet problem (see (4.65)),
- a partially free boundary (see (4.66)), mixed Dirichlet-Oblique Derivative (see (4.67)), because ∇W is not pointing towards the normal of S, non linear problem, because the second equation in (4.67) is highly non linear in the unknowns W and $h(\lambda, \varphi)$.

This is the GBVP in its most general form, or to be more precise, in its most general scalar form, as opposed to a vector form, previously stated in literature, where on L instead of knowing $(\lambda_\mathrm{P}, \varphi_\mathrm{P})$ it is considered as known the direction of $\boldsymbol{g}$ in an Earth-fixed reference frame. This vector form, though interesting, is certainly less realistic than the scalar one, because the data set of directions

$$\boldsymbol{n}(\mathrm{P}) = -\frac{\boldsymbol{g}(\mathrm{P})}{g(\mathrm{P})}$$

is essentially very poor and globally not very accurate. This is why we have chosen to directly present here the scalar GBVP. To the knowledge of the authors, this problem has never been rigorously analyzed in such a general formulation.

In any event, we shall go here to a linearization and a further simplification of the problem, conducting it to a form which is actually used to derive numerical solutions. We follow here the general approach introduced by Krarup (2006), although we want to mention as well Molodensky et al. (1962), Heiskanen and Moritz (1967) and Heck (1991). To this purpose, we notice that the problem has to be linearized with respect to all its unknowns, which here are the potential $W(\mathrm{P})$ as well as the height h_P of S corresponding to the land L. As for $W(\mathrm{P})$, it is only natural to put

$$W(\mathrm{P}) = U(\mathrm{P}) + T(\mathrm{P}) \;, \tag{4.68}$$

with $T(\mathrm{P})$ the variational unknown, and we shall put as well

$$h(\mathrm{P}) = \widetilde{h}(\mathrm{P}) + \zeta(\mathrm{P}) \;, \qquad \mathrm{P} \in L \;, \tag{4.69}$$

where $\zeta(\mathrm{P})$, the variation of $\widetilde{h}(\mathrm{P})$, is called the generalized height anomaly; generalized because we shall reserve the name of proper height anomaly to a particular choice, that will be made in the sequel, for $\widetilde{h}(\mathrm{P})$.

In any way we recall that $\frac{T}{W} = \mathcal{O}\left(10^{-5}\right)$, so to keep in balance the linearization process we have also to put a constraint on $\widetilde{h}$, in such a way that $\frac{\zeta}{R} = \mathcal{O}\left(10^{-5}\right)$, with R the mean radius of the Earth, say 6371 km. This restricts the a-priori values of $\zeta(\mathrm{P})$ to be of the order of 100 m; such a choice is by the way consistent with the values of $N(\mathrm{P})$, which are the counterparts of $\zeta(\mathrm{P})$ on the oceanic area.

We observe that the problem is indeed already linear for the Laplace equation in Ω, because

$$\Delta T = 0\,, \quad \mathrm{P} \in \Omega\,;$$

however such a relation is of little use because Ω is not yet specified. In fact Ω has to be substituted by an approximate $\widetilde{\Omega}$, with a boundary $\widetilde{S}$ that includes $\left\{h = \widetilde{h}\right\}$ on L. For reasons that will be clearer later, instead of the actual known surface of O, we prefer in any way to make $\widetilde{S}$ to coincide with the ellipsoid $\mathcal{E}$ on the oceanic area. This is consistent with our previous discussion on orders of magnitude. In any way we notice that in doing so we modify the domain of harmonicity of the true $T(\mathrm{P})$, yet, on account of the Runge-Krarup theorem, this does not prevent us from having an excellent approximation of $T(\mathrm{P})$, neglecting only quadratic terms in the range $10^{-9} \div 10^{-10}$ of the potential. So we have an $\widetilde{\Omega}$ that is defined as the exterior of

$$\widetilde{S} \equiv \left\{h = \widetilde{h} \text{ on } L\,;\; h = 0 \text{ on } O\right\} \equiv \widetilde{S}_L \cup \widetilde{S}_O\,. \tag{4.70}$$

Naturally, to guarantee that $\widetilde{S}$ is a closed surface, one has to force $\widetilde{h}$ to go to zero on the coast lines. Therefore on $\widetilde{S}_O$ we can write

$$\begin{aligned} W(\mathrm{P_e}) &\cong W(\mathrm{P}) + g(\mathrm{P_e})\,N \cong W(\mathrm{P}) + \gamma(\mathrm{P_e})\,N = \\ &= W_0 + \gamma N \equiv U(\mathrm{P_e}) + T(\mathrm{P_e}) = U_0 + T(\mathrm{P_e})\,, \end{aligned}$$

with $\mathrm{P_e} \in \widetilde{S}_O$, $\mathrm{P} \in O$. Recalling that $W_0 = U_0$, from the previous relation we derive the boundary condition for $\widetilde{S}_O$

$$T(\mathrm{P_e}) = \gamma(\mathrm{P_e})\,N(\mathrm{P_e}) \qquad \mathrm{P_e} \in \widetilde{S}_O\,, \tag{4.71}$$

where the right hand side is known according to (4.64).

Coming to the land part $\widetilde{S}_L$, we have, considering the two points $\mathrm{P} \in S$ and $\widetilde{\mathrm{P}} \in \widetilde{S}_L$, along the same normal ν, at a distance ζ apart,

$$W(\mathrm{P}) = U(\mathrm{P}) + T(\mathrm{P}) \cong U\left(\widetilde{\mathrm{P}}\right) - \gamma\,\widetilde{\zeta} + T\left(\widetilde{\mathrm{P}}\right)\,. \tag{4.72}$$

Introducing the known potential anomaly

$$\mathrm{D}W = W(\mathrm{P}) - U\left(\widetilde{\mathrm{P}}\right)\,, \tag{4.73}$$

we write (4.72) as

$$\widetilde{\zeta} = \frac{T\left(\widetilde{\mathrm{P}}\right) - \mathrm{D}W}{\gamma} ; \tag{4.74}$$

this is known as the generalized Bruns relation. Notice that we use the non standard notation $\mathrm{D}W$ and (here below) $\mathrm{D}g$ to designate W and g anomalies and to distinguish them from ΔW and Δg that correspond to a particular choice of $\widetilde{S}$ and will be introduced later on.

Moreover, we have

$$\begin{aligned} g\,(\mathrm{P}) &= |\nabla U\,(\mathrm{P}) + \nabla T\,(\mathrm{P})| \cong \\ &\cong \gamma\,(\mathrm{P}) + \boldsymbol{e}_\gamma \cdot \nabla T\left(\widetilde{\mathrm{P}}\right) \cong \\ &\cong \gamma\left(\widetilde{\mathrm{P}}\right) + \gamma' \zeta + \boldsymbol{e}_\gamma \cdot \nabla T\left(\widetilde{\mathrm{P}}\right) ; \end{aligned} \tag{4.75}$$

here we have introduced the notation

$$\boldsymbol{e}_\gamma = \frac{\boldsymbol{\gamma}}{\gamma} , \quad \gamma' = \frac{\partial \gamma}{\partial h} .$$

Considering that on $\widetilde{S}$ (see Sansò and Sideris 2013, Sect. 15.2)

$$\boldsymbol{e}_\gamma \cong -\boldsymbol{\nu}$$

with an accuracy of $5 \cdot 10^{-6}$, and introducing, similarly to (4.73), the gravity anomaly

$$\mathrm{D}g = g\,(\mathrm{P}) - \gamma\left(\widetilde{\mathrm{P}}\right) , \tag{4.76}$$

we can write (4.75) in the form

$$-\boldsymbol{\nu} \cdot \nabla T + \gamma' \widetilde{\zeta} \equiv -T' + \gamma' \widetilde{\zeta} = \mathrm{D}g . \tag{4.77}$$

Finally, using (4.74) in (4.77) and reordering, we get the fundamental equation of Physical Geodesy

$$-T'\left(\widetilde{\mathrm{P}}\right) + \frac{\gamma'}{\gamma} T\left(\widetilde{\mathrm{P}}\right) = \mathrm{D}g + \frac{\gamma'}{\gamma} \mathrm{D}W \qquad \widetilde{\mathrm{P}} \in \widetilde{S}_L . \tag{4.78}$$

Putting everything together, we find the linearized form of the scalar GBVP, namely

$$\begin{cases} \Delta T = 0 & \text{in } \widetilde{\Omega} \\ T = \gamma N & \text{on } \widetilde{S}_O \\ -T' + \dfrac{\gamma'}{\gamma} T = \mathrm{D}g + \dfrac{\gamma'}{\gamma} \mathrm{D}W & \text{on } \widetilde{S}_L \\ T = \mathcal{O}\left(\dfrac{1}{r}\right) \end{cases} . \tag{4.79}$$

A first simplification of (4.79) is to fix explicitly the choice of $\widetilde{h}$. One possible useful choice, though not the only one, is to use the traditional condition

$$\mathrm{D}W = W(\mathrm{P}) - U\left(\widetilde{\mathrm{P}}\right) = 0 .$$

Such a condition gives $\widetilde{h}$ as the solution of the implicit function equation

$$U\left[\sigma, \widetilde{h}(\sigma)\right] = W(\sigma, h_\sigma) \;\Rightarrow\; \widetilde{h} = \widetilde{h}(\sigma) , \qquad \sigma = (\lambda, \varphi) . \tag{4.80}$$

With this choice, we shall denote

$$\widetilde{h} = h^* , \tag{4.81}$$

also called normal height, that we shall study in depth in the next chapter. Under such a choice, the corresponding

$$\zeta = h - h^* = \frac{T(\mathrm{P}^*)}{\gamma} \tag{4.82}$$

is the proper height anomaly and (4.82) is the proper Bruns relation. One can prove empirically that in fact $\mathcal{O}(|\zeta|) = 100$ m, which was one of the a-priori conditions to accept $\widetilde{h}$ as a suitable approximation of h.

We note as well that when P is on the geoid, as it happens in O, then

$$W(\sigma, h_\sigma) = W_0 = U_0 = U(\sigma, 0) ;$$

in other words $h^* = 0$ and

$$\zeta_\sigma \equiv N_\sigma , \quad (\mathrm{P} \in O) .$$

Another quantity that gets fixed by the choice (4.81) is the gravity anomaly that now is denoted as

$$\mathrm{D}g = g(\sigma, h_\sigma) - \gamma\left(\sigma, h^*_\sigma\right) \equiv \Delta g(\sigma) , \tag{4.83}$$

also called free air gravity anomaly. Notice that Δ in (4.83) has no relation with the Laplace operator.

The surface

$$S^* \equiv \left\{h = h^*\right\} = S^*_L \cup S^*_O \tag{4.84}$$

is called the Marussi telluroid (Marussi 1985); as we see, this is naturally a closed surface and this explains why we have chosen to use $S^*_O \subset \mathcal{E}$ as the approximate surface in the O region. In fact if we had chosen $S^*_O \equiv S_O$, which is possible because S_O is known, we would have for S^* a surface broken along the coast lines and this is not acceptable as boundary in a Boundary Value Problem.

In this way (4.79) becomes

$$\begin{cases} \Delta T = 0 & \text{in } \widetilde{\Omega} \\ T = \gamma N & \text{on } S_O^* \\ -T' + \dfrac{\gamma'}{\gamma} T = \Delta g & \text{on } S_L^* \\ T = \mathcal{O}\left(\dfrac{1}{r}\right) \end{cases} . \tag{4.85}$$

The solution of this problem is significantly complicated by the shape of the telluroid S_L^*, which mimics the geometry of the actual Earth surface in land areas, with irregular mountains as high as 10^{-3} R. In addition an important role is played by the geometry of the coasts, that separate S_O^* from S_L^*. So a further simplification is achieved by modifying the boundary condition on O, bringing it to the same form as that on L.

Without going into details, we only mention that, after a model up to some degree 200-300 is subtracted from T (see Rapp 1993), one goes locally from T (P) to Δg (P) by a slight generalization of the collocation theory outlined in Sect. 4.5.

More precisely when we subtract from T (P) a global model, e.g. up to degree 200, we theoretically obtain on O a signal containing only wavelengths below about 100 km. The covariance function of such a signal is decaying much faster than the original one and so a good prediction of Δg from T can be done in O even ignoring land data. By forming block averages, e.g. $5' \times 5'$ and using all available altimetric data, properly manipulated to eliminate biases (cross-over analysis), we finally arrive to determine a Δg field uniformly accurate at the level of about 2 mGal (see Sansò and Sideris 2013, Chaps. 6 and 9).

So the GBVP gets the form

$$\begin{cases} \Delta T = 0 & \text{in } \widetilde{\Omega} \\ -T' + \dfrac{\gamma'}{\gamma} T = \Delta g & \text{on } S^* \end{cases} ; \tag{4.86}$$

in (4.86) T is for the moment just a regular harmonic function in $\widetilde{\Omega}$.

Yet, with the new formulation we have introduced an important structural change into the problem. In fact, in contrast to (4.85), the solution of (4.86) is “almost” non unique. This can be better appreciated passing to the so called spherical approximation of (4.86), which consists in changing the boundary operator (but not the boundary S^*) into

$$-\frac{\partial}{\partial h} + \frac{\frac{\partial \gamma}{\partial h}}{\gamma} \cdot \cong -\frac{\partial}{\partial r} + \frac{\frac{\partial \gamma}{\partial r}}{\gamma} \cdot$$

and taking $\gamma = \frac{\mu}{r^2}$, so that (4.86) becomes

$$\begin{cases} \Delta T = 0 & \text{in } \widetilde{\Omega} \\ -\frac{\partial T}{\partial r} - \frac{2}{r} T = \Delta g & \text{on } S^* \end{cases} . \tag{4.87}$$

This is known as the simple Molodensky problem; would S^* be taken as a sphere, this becomes the Stokes problem, that we shall solve explicitly as an example below.

The theory of the simple Molodensky problem is contained in a few propositions (see Sansò and Sideris 2013, Sect. 15.4):

- first extend the definition of the (spherical approximation of) the gravity anomaly to the whole $\widetilde{\Omega}$, written in the form

$$-r\frac{\partial T}{\partial r} - 2T = r\,\Delta g_{\text{sph}} \equiv u\ ; \tag{4.88}$$

- verify by a direct computation that $r\frac{\partial T}{\partial r} = \boldsymbol{r}\cdot\nabla T$ is harmonic throughout $\widetilde{\Omega}$, so that $u = r\,\Delta g_{\text{sph}}$ is a harmonic function too in $\widetilde{\Omega}$;
- to derive the (regular) harmonic u in $\widetilde{\Omega}$, given its boundary values

$$u_0 = r\,\Delta g_{\text{sph}}\big|_{\widetilde{S}}\ , \tag{4.89}$$

  is to solve the Dirichlet problem; this is very well known (see Sansò and Sideris 2013) to have a unique solution, for instance, if the boundary $\widetilde{S}$ is a Lipschitz surface (basically it admits conical points but not cusps) and $u_0 \in L^2\left(\widetilde{S}\right)$, i.e. $\int_{\widetilde{S}} u_0^2\,(\mathrm{P})\,\mathrm{d}S_{\mathrm{P}} < +\infty$;
- let $\overline{R}$ be any Brillouin radius, so that T and u are both harmonic in $\overline{\Omega} = \left\{r \geq \overline{R}\right\}$; let

$$\begin{aligned} T &= \sum_{n=0}^{+\infty}\sum_{m=-n}^{n} \overline{T}_{nm}\left(\frac{\overline{R}}{r}\right)^{n+1} Y_{nm}\,(\sigma) \\ u &= \sum_{n=0}^{+\infty}\sum_{m=-n}^{n} \overline{u}_{nm}\left(\frac{\overline{R}}{r}\right)^{n+1} Y_{nm}\,(\sigma) \end{aligned} , \tag{4.90}$$

  then by a direct computation of (4.88) one finds the "spatial" relation

$$(n-1)\,\overline{T}_{nm} = \overline{u}_{nm}\ , \quad (|m| \leq n\ ;\ \ n = 0, 1, \ldots)\ ; \tag{4.91}$$

- Equation (4.91) implies that if u is derived from (4.88) then $\overline{u}_{1m} = 0$ $(m = -1, 0, 1)$, i.e. u has no terms of the type $\sum_{m=-1}^{1} c_{1m} \left(\frac{R}{r}\right)^2 Y_{1m}(\sigma)$ in its asymptotic expansion at infinity; we observe as well that if by any chance u is such that $\overline{u}_{00} = 0$, then we would have $\overline{T}_{00} = 0$ too, so that the asymptotic behaviour of T would be
$$T = \mathcal{O}\left(\frac{1}{r^3}\right) , \tag{4.92}$$
as it was in our original definition of the anomalous potential;
- on the other hand, since
$$\delta T = \sum_{m=-1}^{1} c_{1m} \left(\frac{R}{r}\right)^2 Y_{1m}(\sigma) \tag{4.93}$$
is a function of r, homogeneous of degree -2, whatever are constants $c_{1\,-1}$, $c_{1\,0}$, $c_{1\,1}$, we see that δT is such that
$$r\,\delta T' + 2\,\delta T \equiv 0 ;$$
since δT is also obviously harmonic, outside the origin, we have that δT represents a null space of our BVP (4.87); this means that in any way a component like δT of T will never be fixed by the data;
- since in the end we want to find a solution T satisfying the traditional relation (4.92), we decide that the arbitrary δT should be fixed by the condition
$$\delta T \equiv 0 ,$$
that we know to be equivalent to placing the barycentre of T at the origin (or better placing the barycentre of U so as to coincide with that of W); furthermore we shall make some operation on the data $u_0 = r\,\Delta g_{\mathrm{sph}}\big|_{\widetilde{S}}$, so that $\overline{u}_{00} = 0$ implying also that $\overline{T}_{00} = 0$, i.e. (4.92) holds true (see Sansò and Sideris 2013);
- if we do not want to put restrictions directly on u_0, we can change it by introducing four unknown constants, namely substituting the boundary condition $u|_S = u_0$ with
$$u|_S = u_0 + a\,\frac{R}{r}\bigg|_S + \sum_{m=-1}^{1} b_{1m} \left(\frac{R}{r}\right)^2 Y_{1m}(\sigma)\bigg|_S \tag{4.94}$$
and determining a, b_{1m} $(m = -1, 0, 1)$ in such a way that
$$\begin{aligned} &\overline{u}_{00} = 0 \;\; (\text{to imply}\,\overline{T}_{00} = 0) \\ &\overline{u}_{1m} = 0 \;\; (\text{to produce a boundary function } u_0 \\ &\qquad\qquad \text{that is } r \text{ times a spherical gravity anomaly}); \end{aligned}$$

one can prove that such conditions can always be satisfied by suitable constants $\forall u_0 \in L^2(S)$ (see Sansò and Sideris 2013);

- finally we derive $T = T(r, \sigma)$ by integrating radially (4.88) and taking into account that $u = \mathcal{O}\left(\frac{1}{r^3}\right)$, so that the closed expression is found

$$T(r, \sigma) = \frac{1}{r^2} \int_r^{+\infty} s\, u(s, \sigma)\, ds ; \tag{4.95}$$

one can directly prove that such a T satisfies (4.88), that it is a harmonic function and that it satisfies (4.92).

Summarizing, we have recalled the line showing that the simple Molodensky problem, modified as

$$\begin{cases} \Delta T = 0 & \text{in } \widetilde{\Omega} \\ -r\,T' - 2\,T = u_0 - a\,\frac{\overline{R}}{r} - \sum_{m=-1}^{1} b_{1m} \left(\frac{\overline{R}}{r}\right)^2 Y_{1m}(\sigma) & \text{on } S^* \\ T = \mathcal{O}\left(\frac{1}{r^3}\right) \end{cases} , \tag{4.96}$$

has one and only one solution $\{T, a, b_{1\,-1}, b_{1\,0}, b_{1\,1}\}$ whatever is the known term $u_0 \in L^2(S^*)$, i.e. $\Delta g \in L^2(S^*)$.

Once this is achieved, one can return to the original problem (4.86), that now we rewrite as

$$\begin{cases} \Delta T = 0 & \text{in } \widetilde{\Omega} \\ T' + \frac{\gamma'}{\gamma} T = \Delta g + a\,\frac{\overline{R}}{r^2} + \sum_{m=-1}^{1} b_{1m} \left(\frac{\overline{R}^2}{r^3}\right) Y_{1m}(\sigma) & \text{on } S^* \\ T = \mathcal{O}\left(\frac{1}{r^3}\right) \end{cases} , \tag{4.97}$$

and prove, by a perturbative argument, that (4.97) has a unique solution; however we are now obliged to put constraints on the inclination of the normal to S^* with respect to the radial direction $\boldsymbol{e}_r$ to guarantee the convergence of the perturbative process. Yet, a satisfactory result is obtained if we admit to a-priori know a model up to a maximum degree N, so that we can reduce our solution u to satisfy the asymptotic relation

$$u = \mathcal{O}\left(\frac{1}{r^{N+2}}\right) . \tag{4.98}$$

The theorem is the following (see Sansò and Sideris 2013, Sect. 15.4): if we know a model of T complete up to degree and order 20, then a unique solution to the Molodensky problem exists if the inclination of S^* with respect to $\boldsymbol{e}_r$ never exceeds 60°.

Fortunately, nowadays satellite geodesy is able, by analyzing data of low satellites, to provide the knowledge of the first 20 degrees of T with very high accuracy, in fact with an error of the order of 1 mm in terms of geoid. Such a knowledge has been pushed up to degree 200 with an error of about 2 cm, as we shall comment later on in Chap. 7.

As promised, we develop now the explicit solution of (4.96) when S^* is taken as a sphere, i.e. of the Stokes problem.

Example (Stokes theory)

Assume S^* is just a sphere with radius R_0; we want to solve the corresponding B.V.P. (4.96), which is of the simple Molodensky type.

Given our hypothesis, we expect T to be expandable into the spherical harmonic series

$$T = \sum_{n=0}^{+\infty} \sum_{m=-n}^{n} T_{nm} \left(\frac{R_0}{r} \right)^{n+1} Y_{nm}(\sigma) \ ; \tag{4.99}$$

this automatically satisfies the harmonicity condition. On the other hand we have, on the boundary,

$$u_0(\sigma) = R_0 \, \Delta g(\sigma) = R_0 \sum_{n=0}^{+\infty} \sum_{m=-n}^{n} \Delta g_{nm} Y_{nm}(\sigma) \ .$$

Since in this case we can take $\overline{R} = R_0$, we see that the known term in the second equation of (4.96) can be written as

$$\overline{u}_0 = R_0 \sum_{n=2}^{+\infty} \sum_{m=-n}^{n} \Delta g_{nm} Y_{nm}(\sigma) \ ,$$

if we make the choice

$$a = R_0 \, \Delta g_{00} \ , \qquad b_{1m} = R_0 \, \Delta g_{1m} \ ,$$

so that $u_{00} = u_{1m} \equiv 0$, $(m = -1, 0, 1)$. But in this case we know that $T_{00} = 0$ and, also, we can choose $T_{1m} = 0$ to satisfy the third equation of (4.96). Then for $n > 1$, we can use (4.91), i.e.

$$(n-1)\, T_{nm} = u_{nm} = R_0 \, \Delta g_{nm} \ .$$

Returning to the representation of T, we get

$$T\,(\mathrm{P}) = \sum_{n=2}^{+\infty} \sum_{m=-n}^{n} \frac{R_0}{n-1}\,\Delta g_{nm} \left(\frac{R_0}{r_\mathrm{P}}\right)^{n+1} Y_{nm}\,(\sigma_\mathrm{P}) \ ;$$

Now we can remember that

$$\Delta g_{nm} = \frac{1}{4\pi} \int \Delta g\left(\sigma_\mathrm{Q}\right) Y_{nm}\left(\sigma_\mathrm{Q}\right) \mathrm{d}\sigma_\mathrm{Q} \ ,$$

so that the previous relation can be written as

$$\begin{aligned} T\,(\mathrm{P}) &= \frac{1}{4\pi} \int \Delta g\left(\sigma_\mathrm{Q}\right) \left[\sum_{n=2}^{+\infty} \frac{R_0}{n-1} \left(\frac{R_0}{r_\mathrm{P}}\right)^{n+1} \sum_{m=-n}^{n} Y_{nm}\,(\sigma_\mathrm{P})\, Y_{nm}\left(\sigma_\mathrm{Q}\right)\right] \mathrm{d}\sigma_\mathrm{Q} = \\ &= \frac{R_0}{4\pi} \int \Delta g\left(\sigma_\mathrm{Q}\right) \sum_{n=2}^{+\infty} \frac{2n+1}{n-1} P_n\left(\cos\psi_\mathrm{PQ}\right) \left(\frac{R_0}{r_\mathrm{P}}\right)^{n+1} \mathrm{d}\sigma_\mathrm{Q} \ . \end{aligned}$$

The series can be added in a closed form, obtaining the so called Stokes function (see Sansò and Sideris 2013, Sect. 3.4)

$$\begin{aligned} S\left(R_0, r_\mathrm{P}, \psi_\mathrm{PQ}\right) &= \frac{2R_0}{\ell_\mathrm{PQ}} + \frac{R_0}{r_\mathrm{P}} - \frac{3R\,\ell_\mathrm{PQ}}{r_\mathrm{P}^2} - \frac{R^2}{r_\mathrm{P}^2}\cos\psi_\mathrm{PQ} \cdot \\ &\quad \cdot \left[5 + 3\log\frac{r_\mathrm{P} - R_0\cos\psi_\mathrm{PQ} + \ell_\mathrm{PQ}}{2r_\mathrm{P}}\right] , \end{aligned}$$

with

$$\ell_\mathrm{PQ} = \left[R_0^2 + r_\mathrm{P}^2 - 2R_0\, r_\mathrm{P} \cos\psi\right]^{\frac{1}{2}} \ .$$

So the solution of the Stokes problem is written in integral form as

$$T\,(\mathrm{P}) = \frac{R_0}{4\pi} \int S\left(R_0, r_\mathrm{P}, \psi_\mathrm{PQ}\right) \Delta g\left(\sigma_\mathrm{Q}\right) \mathrm{d}\sigma_\mathrm{Q} \ .$$

Let us remark that the GBVP theory, beyond providing a basis for the numerical determination of high degree anomalous models, is in itself one of the foundations of Physical Geodesy because it can specify what is the minimal information that can provide a stable solution T (P), under realistic conditions.

As claimed before, the solution of the GBVP is provided in terms of a finite sum of spherical harmonics of the type (4.99), truncated at a maximum degree N, which is called a global model of the anomalous potential. At present the most important of such models is EGM2008, which is complete up to degree and order 2159. The original data have been processed in such a way as to cover the Earth with a $5' \times 5'$ grid of area mean gravity anomalies; this corresponds to 9,331,200 values from which the model, described by 4,665,595 coefficients, is derived (see Sansò and Sideris 2013, Part II, Chap. 6). Another widely used global model, complete up to degree and order 2159, is EIGEN-6C4 that additionally includes GOCE data, though using the same EGM2008 $5' \times 5'$ grid of area mean gravity anomalies over continents. In 2020, it is foreseen the release of an updated version of EGM2008, called EGM2020, which will benefit from new data sources and procedures.

The overall error of the model, in terms of geoid, evaluated as a mean square estimation error over the whole Earth sphere, is considered to range around 5 cm; however the geographic distribution of the error, reflecting in particular areas of poor coverage of data and mountainous areas, shows that local error r.m.s. can amount up to 1 m.

The resolution of the model indeed cannot be better than the resolution of the input data, which in the average is around 10 km; this is reflected in the maximum degree 2159 chosen.

Indeed one might wonder whether, by using higher resolution data, one could improve the knowledge of the anomalous potential, at least locally. This is the case, although we cannot enter into details in this context; we rather send to literature, e.g. Sansò and Sideris (2013). Here we report only that an improved result can be obtained by first finely tuning the effect of local topographic masses on T (separately accounting for it) and then by applying a kind of local solver operator borrowed from random field prediction theory, for instance a collocation algorithm as recalled in Sect. 4.6, or some other equivalent techniques. What we fix here, about this more complex theory, is that there is a local solving operator S_A that acts on the improved data set $\{\Delta g\}$ in an area A, capable of producing a local anomalous potential

$$T = S_A(\Delta g) \tag{4.100}$$

that provides an approximation of the true T at the level of 2–3 cm in geoid, depending on the data available, the roughness of the surface (telluroid) and the roughness of the field Δg in A.

References

Brockmann J.M., Zehentner N., Höck E., Pail R., Loth I., Mayer-Gürr T., Schuh W.-D. (2014). *EGM TIM RL05: An independent geoid with centimeter accuracy purely based on the GOCE mission*. Geophysical Research Letters, 41(22):8089–8099.

Förste C., Bruinsma S.L., Abrikosov O., Lemoine J.M., Schaller T., Götze H.J., Ebbing J., Marty J.C., Flechtner F., Balmino G., Biancale R. (2014). *EIGEN-6C4: The latest combined global gravity field model including GOCE data up to degree and order 2190 of GFZ Potsdam and GRGS Toulouse*. In: 5th GOCE User Workshop, Paris. 25–28 November 2014. http://icgem.gfz-potsdam.de/ICGEM/documents/Foerste-et-al-EIGEN-6C4.pdf.

Heck B. (1991). *On the linearized Boundary Value Problem of physical geodesy*. Report No. 407, Department of Geodetic Science and Survey, The Ohio State University, Columbus.

Heiskanen W.A. and Moritz H. (1967). *Physical geodesy*. Freeman, San Francisco.

Kaula W.M. (1966). *Tests and combination of satellite determinations of the gravity eld with gravimetry*. Journal of Geophysical Research, 71:5303–5314.

Kaula W.M. (2000). *Theory of satellite geodesy*. Dover, New York.

Krarup T. (2006). *Letters on Molodenskiy's problem, III*. In: Borre K. (Ed.) Mathematical Foundation of Geodesy. Springer Verlag, Berlin, Heidelberg.

MacMillan W.D. (1958). *The theory of the potential*. In: Theoretical mechanics, Vol 2. Dover Publications, New York.

Marussi A. (1985). *Intrinsic geodesy*. Springer, Berlin.

Molodensky M.S., Eremeev V.F., Jurkina M.I. (1962). *Methods for study of the external gravitational field and figure of the Earth*. Translated from Russian, Israel Program for Scientific Translations, Jerusalem.

Moritz H. (1980). *Advanced physical geodesy*, 2nd edn. Wichmann, Karlsruhe.

Pavlis N.K., Holmes S.A., Kenyon S.C., Factor J.K. (2012). *The development and evaluation of the Earth Gravitational Model 2008 (EGM2008)*. Journal of Geophysical Research: Solid Earth, 117(B4):B04406.

Pavlis N.K., Holmes S.A., Kenyon S.C., Factor J.K. (2013). *Correction to "The development and evaluation of the Earth Gravitational Model 2008 (EGM2008)"*. Journal of Geophysical Research: Solid Earth, 118(5):2633.

Rapp, R.H. (1993). *Use of altimeter data in estimating global gravity models*. In: Sansò F., Rummel R. (Eds.), Satellite Altimetry in Geodesy and Oceanography, Lecture Notes in Earth Sciences, 50. Springer-Verlag, Berlin, Heidelberg, pp. 374–417.

Sansò F., Sideris M.G. (2013). *Geoid determination: Theory and methods*. Lecture Notes in Earth System Sciences, Vol. 110. Springer-Verlag, Berlin, Heidelberg.

Sansò F., Venuti G. (2010). *The convergence problem of collocation solutions in the framework of the stochastic interpretation*. Journal of Geodesy, 85(1):51–63.

Shako R., Förste C., Abrykosov O., Bruinsma S., Marty J.-C., Lemoine J.-M., Flechtner F., Neumayer K.-H., Dahle C. (2014). *EIGEN-6C: A High-Resolution Global Gravity Combination Model Including GOCE Data*. In: Flechtner F., Sneeuw N., Schuh W.-D. (Eds.), Observation of the System Earth from Space - CHAMP, GRACE, GOCE and future missions, Advanced Technologies in Earth Sciences. Springer, Berlin, Heidelberg, pp. 155–161.

Stokes G.G. (1849). *On the variation of gravity on the surface of the Earth*. Transactions of the Cambridge Philosophical Society, 8:672–695.

Tscherning C.C. and Rapp R.H. (1974). *Closed covariance expressions for gravity anomalies, geoid undulations, and deections of the vertical implied by anomaly degree variances*. Report No. 208, Department of Geodetic Science and Surveying, The Ohio State University, Columbus.