Stokes's two-boundary-value-problem

Zdeněk Martinec

It is a common belief that, after removing the first-degree spherical harmonics from the gravitational potential, only a regularization of the downward continuation of a high frequency part of the gravity is necessary to guarantee the existence of a unique solution to the Stokes boundary-value problem for gravimetric determination of the geoid. In this paper, we will deal with the original formulation of the problem prior to the downward continuation of gravity. We intend to demonstrate numerically that, besides the spherical harmonics of degree one, the existence of the solution is not also guaranteed for higher-degree harmonics. This lack of guaranty is due to the fact that the input data – the surface gravity and the potential of the geoid – are prescribed on different boundaries.

1. Formulation of the Stokes two-boundary-value problem

Let the geocentric radius of the geoid S_g be described by an angularly dependent function $r = r_g(\Omega)$, where (r, Ω) are the geocentric spherical coordinates, i.e., $(r_g(\Omega), \Omega)$ are points lying on the geoid. We will assume that the function $r_g(\Omega)$ is not known. Let $H(\Omega)$ be the height of the Earth's surface above the geoid reckoned along the geocentric radius. Unlike the geocentric radius of the geoid, we will assume that $H(\Omega)$ is a known function. Finally, let the following quantities be given: the gravity $g_S(\Omega)$ measured on the Earth's surface, the density $\varrho(r, \Omega)$ of the topographical masses (the masses between the geoid and the Earth's surface), and the gauge value W_0 of the gravity potential on the geoid.

The question we pose is: how to determine the gravity potential $W(r, \Omega)$ **inside** and **outside** the topographical masses and the radius $r_g(\Omega)$ of the geoid? The problem is governed by the Poisson equation with the boundary conditions given on the free boundaries S_t and S_g coupled by means of height $H(\Omega)$:

$$\nabla^2 W = -4\pi G\varrho + 2\omega^2 \qquad \text{outside } S_g \ , \tag{1}$$

$$|\mathrm{grad}\, W| = g_S \qquad \text{on } S_t \ , \tag{2}$$

$$W = W_0 \qquad \text{on } S_g \ , \tag{3}$$

$$W = \frac{1}{2}\omega^2 r^2 \sin^2\vartheta + \frac{GM}{r} + O\left(\frac{1}{r^3}\right) \qquad r \to \infty \ , \tag{4}$$

where G is the gravitational constant, M is the mass of the Earth, and ϱ is equal to zero outside the Earth. The first-degree harmonics are left out from the potential W because of the geocentric coordinate system.

Martinec and Matyska (1997) have shown that the boundary-value problem (1)–(4) can be linearized with respect to the anomalous potential T^h such that

$$\nabla^2 T^h = 0 \qquad \text{outside } S_g, \tag{5}$$

$$\left.\frac{\partial T^h}{\partial r}\right|_P + \frac{2}{r_Q} T^h_{P_g} - \epsilon_h(T^h_P) - \epsilon_\gamma(T^h_{P_g}) = -\Delta g^h + \sum_{m=-1}^{1} a_{1m} Y_{1m}(\Omega) \, , \tag{6}$$

$$T^h = \frac{c}{r} + O\left(\frac{1}{r^3}\right) \qquad \text{for } r \to \infty \, , \tag{7}$$

where P, P_g and Q are the points on the Earth's surface, the geoid and the level ellipsoid, respectively, ϵ_h and ϵ_γ are ellipsoidal corrections (e.g., Jekeli, 1981), Δg^h is the Helmert gravity anomaly and a_{1m} are constants to be determined.

2. Numerical investigations

The original problem (1)–(4) as well as the problem described by eqns.(5)–(7) are scalar non-linear free boundary-value problems since the radial coordinate of the geoid is one of the unknowns to be determined. Having some approximation of geoid, it is easy to transform the latter free boundary-value problem to a problem with fixed boundaries. For example, replacing P_g by r_Q, r_Q being the radius of the normal point Q, and P by $r_Q + H(\Omega)$ in eqn.(6) yields the ellipsoidal approximation of the Stokes two-boundary-value problem, where eqns.(5)–(7) serve to determine T^h; Bruns's formula then gives the geoidal height N. Another possibility, most often used in geoid height computations, is to approximate the geoid in the boundary condition (6) by a mean sphere with radius $R = 6371$ km. This means the radius of the point P_g is replaced by R and radius of the point P by $R + H(\Omega)$. The relative error introduced by this spherical approximation is of the order of 3×10^{-3} in the classical problems (Heiskanen and Moritz, 1967, sect.2-14), which then causes a long-wavelength error of at most 0.5 metres in geoidal heights. In regional problems, where only shorter wavelengths are to be determined, this approximation is often reasonable. In the following numerical tests we will employ the spherical approximation of boundary condition (6) for its simplicity. We intend to concentrate on the effects connected with the 'two-boundary nature' of this condition that appear only in a very short wavelength part of the solution.

The solution to the Laplace equation (5) with the condition (7) can be represented as a series of solid spherical harmonics $r^{-j-1} Y_{jm}(\Omega)$,

$$T^h(r, \Omega) = \sum_{\substack{j = j_{min} \\ j \neq 1}}^{j_{max}} \sum_{m=-j}^{j} T_{jm} \left(\frac{R}{r}\right)^{j+1} Y_{jm}(\Omega) \, , \tag{8}$$

where $j_{min}(\geq 0)$ and j_{max} are the respective minimum and maximum cut-off degrees, $Y_{jm}(\Omega)$ are spherical harmonics of degree j and order m, and T_{jm} are the coefficients of potential T^h to be determined. In order to normalize the potential coefficients T_{jm}, we have introduced the mean Earth's radius R into the expansion (8). Equation (6) in the spherical approximation then becomes

$$\frac{1}{R} \sum_{\substack{j = j_{min} \\ j \neq 1}}^{j_{max}} \sum_{m=-j}^{j} \left[(j+1) \left(\frac{R}{R + H(\Omega)} \right)^{j+2} - 2 + e_0^2 (3 \cos^2 \vartheta - 2) \right] Y_{jm}(\Omega) T_{jm} +$$

$$+ \frac{e_0^2}{R} \sum_{\substack{j = j_{min} \\ j \neq 1}}^{j_{max}} \sum_{m=-j}^{j} \sin\vartheta \cos\vartheta \, \frac{\partial Y_{jm}(\Omega)}{\partial \vartheta} T_{jm} = \Delta g^h - \sum_{m=-1}^{1} a_{1m} Y_{1m}(\Omega) \, . \tag{9}$$

This boundary condition must hold in any direction Ω. In order to ensure it, we will employ the Galerkin method in which eqn.(9) can be rewritten as a system of linear algebraic equations for coefficients T_{jm}:

$$\boldsymbol{A}\boldsymbol{m} = \boldsymbol{d} \ , \tag{10}$$

where $\boldsymbol{m}$ is a column vector composed of potential coefficients T_{jm}, i.e.,

$$\boldsymbol{m} := \{T_{jm} | j = j_{min}, ..., j_{max}, j \neq 1, m = -j, ..., j\} \ , \tag{11}$$

$\boldsymbol{A}$ is the matrix composed of the weighted left-hand side of eqn.(9),

$$A_{j_1 m_1, jm} :=$$

$$:= \int_{\Omega_0} \left[(j+1) \left(\frac{R}{R + H(\Omega)} \right)^{j+2} - 2 + e_0^2 (3 \cos^2 \vartheta - 2) \right] Y_{jm}(\Omega) Y^*_{j_1 m_1}(\Omega) d\Omega + \tag{12}$$

$$+ e_0^2 \int_{\Omega_0} \sin \vartheta \cos \vartheta \, \frac{\partial Y_{jm}(\Omega)}{\partial \vartheta} Y^*_{j_1 m_1}(\Omega) d\Omega \ ,$$

and $\boldsymbol{d}$ is a column vector of weighted right-hand side of eqn.(9),

$$d_{j_1 m_1} := R \int_{\Omega_0} \Delta g^h(\Omega) Y^*_{j_1 m_1}(\Omega) d\Omega \ , \tag{13}$$

where $j_1 = j_{min}, ..., j_{max}, j_1 \neq 1$, and $m_1 = -j_1, ..., j_1$.

2.1. An example: constant height

Let us first consider a simple, but illustrative, case when $H = H_0 = const.$ over the Earth, and $e_0^2 = 0$. Introducing function

$$K_j(H_0) := (j+1) \left(\frac{R}{R + H_0} \right)^{j+2} - 2, \quad \text{for} \ \ j \geq 2 \ , \tag{14}$$

the transfer matrix $A_{j_1 m_1, jm}$ between unknown parameters T_{jm} and the gravity anomalies Δg^h on the right-hand side of eqn.(9) becomes $A_{j_1 m_1, jm} = K_j(H_0)\, \delta_{jj_1} \delta_{mm_1}$ and thus

$$T_{jm} = \frac{R}{K_j(H_0)} \int_{\Omega_0} \Delta g^h(\Omega) Y^*_{jm}(\Omega) d\Omega \ . \tag{15}$$

Since $0.998 < R/(R + H_0) < 1$ for the Earth, it is clear that $\lim_{j \to \infty} K_j = -2$ for any fixed $H_0 > 0$. On the other hand, $K_j > 0$ for low degrees j because $0.976 < K_2 < 1$. This means that there is a range of j's in which K_j is zero or near zero. For those j's the solution of eqn.(10) is unstable or even does not exist once $K_j = 0$.

Let us estimate the range of j's for which the solution of eqs.(10) becomes unstable for this simple example. Figure 1 plots the values of K_j for height H_0 equal to 1 km, 5 km and 10 km. We can see that the increase of K_j with increasing j is confined to low degrees j and then K_j starts to decrease to its limiting value -2. That is why, the determination of disturbing potential T^h is stable only in some part of the spectral domain. The width of the stable part grows with decreasing H_0.

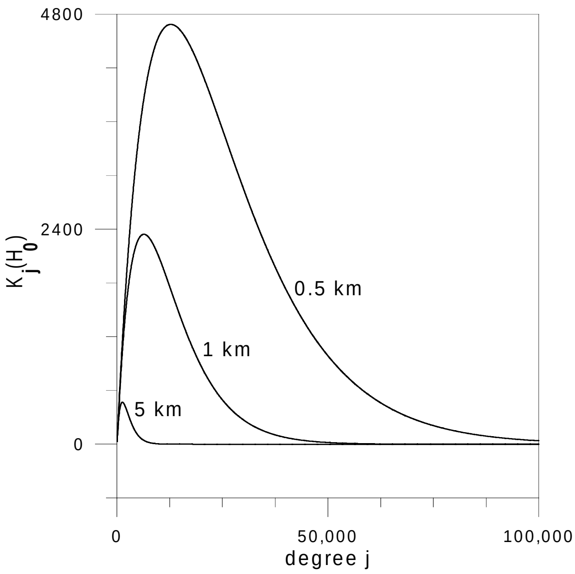

Figure 1: Transfer function $K_j(H_0)$ between unknown coefficients T_{jm} and gravity anomalies Δg^h for $H_0 = 1$ km, 5 km, and 10 km.

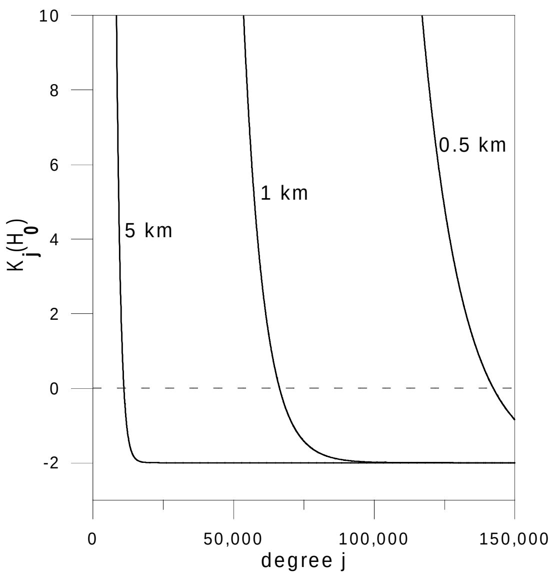

Figure 1a: A detail of Figure 1.

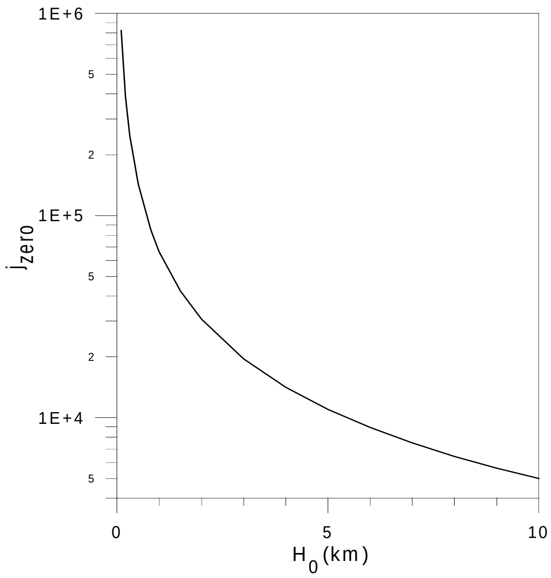

Figure 2: The roots j_{zero} of function $K_j(H_0)$ for $H_0 \in (100\,\mathrm{m}, 10^4\,\mathrm{m})$.

Figure 2 plots those j_{zero} for which function $K_j(H_0)$ vanishes. For such degrees matrix $\boldsymbol{A}$ is singular and the solution of system of equations (10) does not exist. Since spherical degree j corresponds to a given resolution $\Delta\Omega$ in a spatial domain, $\Delta\Omega = \pi/j$, we may also convert critical degree j_{zero} to a critical spatial resolution size $\Delta\Omega_{zero}$, $\Delta\Omega_{zero} = \pi/j_{zero}$, for which the solution to our problem does not exist. Figure 2 shows that, for instance, $j_{zero} = 10980$, for $H_0 = 5$ km, and the critical spatial resolution size is $\Delta\Omega_{zero} \doteq 1$ arcmin. To interpret the result in other words, let us imagine that the Earth's topography is a Bouguer spherical shell with a constant height of 5 km above the geoid and the Stokes two-boundary-value problem is solved in a spatial domain such that the potential $T^h(R,\Omega)$ is parameterized by discrete values $T^h(R,\Omega_i)$ in a regular angular grid with grid step size $\Delta\Omega$. Then the solution to the Stokes two-boundary-value problem will not exist if the grid step size $\Delta\Omega$ of the parameterization of T^h is less than or equal to the critical step size $\Delta\Omega_{zero}$, i.e., of about 1 arcmin in our example, even though the surface gravity data would be known continuously on the Earth's surface.
To map the non-existence of the solution for regional geoid determination and for a more realistic model of the Earth's topography, we need to set up and to solve the system of eqn.(10) for high degrees and orders ($j_{max} = 10^4 - 10^5$). This leads to computational difficulties because of huge consummation of computational time and memory; with today's computer equipment it is impossible to carry out the analysis of the existence for such a general case. Thus, we are forced to approximate the Earth's surface by a simplified model of axisymmetric geometry. By making use of the analysis of this simplified case, we will attempt to estimate the range of critical spectral degrees j_{zero} for the actual case.

2.2. Axisymmetric geometry

Let the height $H(\vartheta,\lambda)$ of the Earth's surface above the geoid is modelled by zonal as well as tesseral and sectoral spherical harmonics of the global digital terrain model TUG87 (Wieser, 1987) cut at degree 180. To create a rotational symmetric body, axisymmetric height $H(\vartheta)$ will be generated by height $H(\vartheta,\lambda)$ taken along a fixed meridian $\lambda = \lambda_0$. In the case of an axisymmetric surface, the elements $A_{j_1 m_1, jm}$ of matrix $\boldsymbol{A}$ do not depend on angular orders m and m_1; they can be written as

$$A_{j_1 j} = \int_{\vartheta=0}^{\pi} \left[(j+1) \left(\frac{R}{R+H(\vartheta)} \right)^{j+2} - 2 + \right.$$

$$+e_0^2(3\cos^2\vartheta - 2)\Bigg] P_j(\cos\vartheta) P_{j_1}(\cos\vartheta)\sin\vartheta d\vartheta + \tag{16}$$

$$+e_0^2 \int_{\vartheta=0}^{\pi} \sin\vartheta\cos\vartheta \frac{dP_j(\cos\vartheta)}{d\vartheta} P_{j_1}(\cos\vartheta)\sin\vartheta d\vartheta \ .$$

Note that the elements $A_{j_1 j}$ can only be evaluated by a method of numerical quadrature.
To analyse the posedness of the Stokes two-boundary-value problem, we will employ the eigenvalue analysis of matrix $\boldsymbol{A}$. According to this method, a non-symmetric matrix $\boldsymbol{A}$ can be decomposed to the product of three matrices,

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{-1} \ , \tag{17}$$

where the columns of matrix $\boldsymbol{U}$ are formed from the right eigenvectors of $\boldsymbol{A}$, the rows of $\boldsymbol{U}^{-1}$ are formed from the left eigenvectors of $\boldsymbol{A}$, and the diagonal matrix $\boldsymbol{\Lambda}$ consists of eigenvalues of $\boldsymbol{A}$. We have employed subroutines BALANC, ELMHES and HQR (Press et al., 1992, sect.11.5 and 11.6) to find the eigenvalues of a non-symmetric matrix $\boldsymbol{A}$.

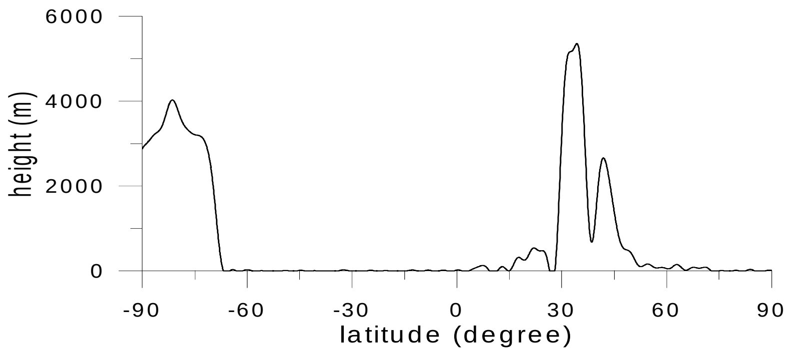

Figure 3: The meridian profile $\lambda = 80^o$ of topographical height $H(\vartheta, \lambda)$ generated by the global digital terrain model TUG87 (Wieser, 1987) cut at degree 180. This profile is used to create a body with the axisymmetric geometry of external surface.

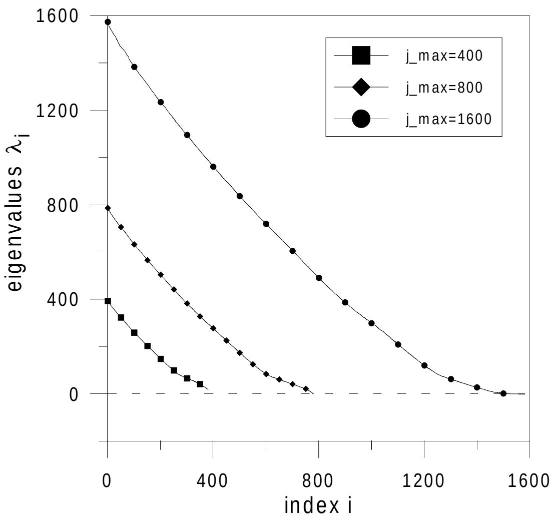

Figure 4: The eigenvalue spectra of matrix $\boldsymbol{A}$ for various cut-off degrees j_{max} and a body with axisymmetric surface generated by height $H(\vartheta, \lambda = 80^o)$ multiplied by 10 ($j_{min} = 21$). The ellipsoidal corrections ϵ_h and ϵ_γ are equal to zero.

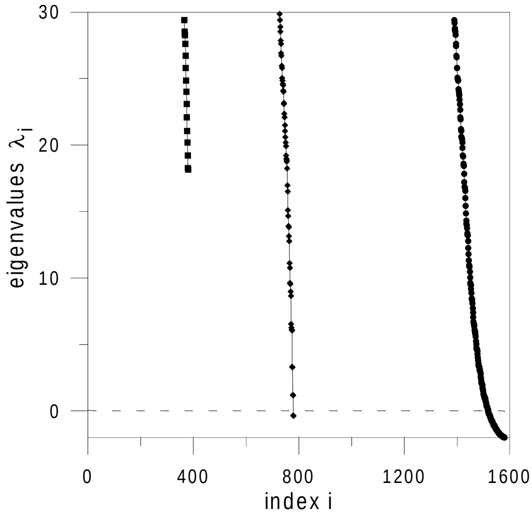

Figure 4a: A detail of Figure 4.

Figure 3 shows the topographical height $H(\vartheta, \lambda_0)$ along the meridian profiles $\lambda_0 = 80^o$ reaching value $H_{max} = 5353$ metres. The consequent Figure 4 shows a plot of the eigenvalues of matrix $\boldsymbol{A}$ for an axisymmetric body with the outer surface generated by this meridian profile. In order to avoid high degrees j, and thus, be able to perform the eigenvalue analysis in real CPU time, we multiply function $H(\vartheta)$ by a factor of 10. The minimum spherical degrees j_{min} of the potential series (8) is $j_{min} = 21$, which models the situation when low-degree harmonics of potential T^h are determined by another approach, e.g., when considering a satellite gravitational model. In Figure 4, where we further put the eccentricity of the level ellipsoid equal to zero, $e_0 = 0$, we change the maximum cut-off degree j_{max} of the disturbing potential T^h and plot eigenvalues of matrix $\boldsymbol{A}$ ordered according to their size (note that the eigenvalues are real numbers in this particular case). Inspecting Figure 4 we can observe that the eigenvalue spectrum of matrix $\boldsymbol{A}$ intersects the zero level starting from degree $j_{zero} \doteq 800$. Once the cut-off degree j_{max} of the spherical harmonic expansion (8) of potential T^h is greater or equal to j_{zero}, the eigenvalue spectrum of $\boldsymbol{A}$ contains a null eigenvalue or an eigenvalues of a very small size. The matrix $\boldsymbol{A}$ becomes ill-conditioned or even singular and the inverse $\boldsymbol{A}^{-1}$ may be distorted by large round-off errors or may not exist at all; in such a case the Stokes two-boundary-value problem does not have a unique and stable solution. As in the preceeding section, the critical degree j_{zero} can again be converted to the critical spatial discretization size $\Delta\Omega_{zero}$ for a case when the Stokes two-boundary-value problem is solved in a spatial domain.

The next test investigates the influence of the ellipsoidal corrections terms ϵ_h and ϵ_γ on the posedness of matrix $\boldsymbol{A}$. We choose the same body as in the preceeding example together with $j_{min} = 21$ and $j_{max} = 1600$ and compute the eigenvalues of matrix $\boldsymbol{A}$ putting $e_0^2 = 0$ and $e_0^2 = 0.006694$, respectively. Figure 5 shows those eigenvalues the magnitudes of which are smaller than 3. (Note that the eigenvalues of $\boldsymbol{A}$ for the case $e_0^2 = 0.006694$ are complex numbers.) We can observe that the eigenvalue spectrum of $\boldsymbol{A}$ changes significantly when e_0^2 differs from zero: there is no null eigenvalue and the magnitude of the smallest eigenvalue is larger than 1. In other words, the ellipsoidal corrections ϵ_γ and ϵ_h

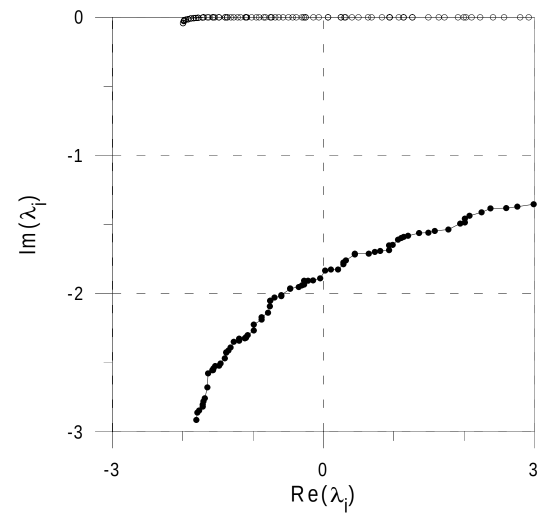

Figure 5: The real vs. imaginary parts of the eigenvalues of matrix $\boldsymbol{A}$ with (lower branch) and without (upper branch) the ellipsoidal corrections ϵ_h and ϵ_γ. The axisymmetric body is the same as that considered in Figure 4 ($j_{min} = 21$, and $j_{max} = 1600$).

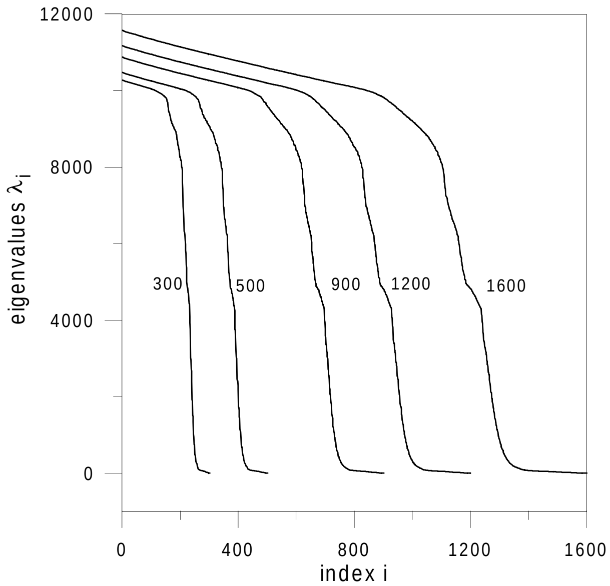

Figure 6: The eigenvalue spectra of matrix $\boldsymbol{A}$ for various cut-off degrees $j_{max} = j_{min} + \Delta j, \Delta j = 300, 500, ..., 1600$, and a body with axisymmetric surface generated by height $H(\vartheta, \lambda = 80^o)$ ($e_0^2 = 0$ and $j_{min} = 10000$).

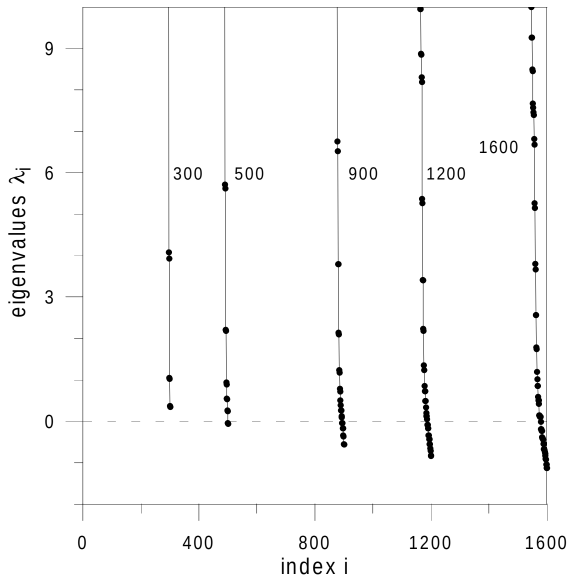

Figure 6a: A detail of Figure 6.

act as regularization factors removing the ill-posedness of matrix $\boldsymbol{A}$. It also means that, in this particular case, ϵ_γ and ϵ_h cannot be subtracted from the right-hand side of eqn.(10) as known quantities determined a priorily by using a known global gravitational model of the Earth; such usage of ellipsoidal corrections is often recommended in real geoid computations.

In order to create a more realistic example, we use the same profile of topographical height as plotted in Figure 3, but now, in contrast with preceeding example, we will not multiplied height $H(\vartheta)$ by 10. In this case, it is not possible to carry out the eigenvalue analysis of matrix $\boldsymbol{A}$ starting from degree $j_{min} = 21$ and going up to degrees $j_{max} \approx 10^4 - 10^5$ due to a huge consummation of computer time and memory. We have to confine ourselves to a smaller range of sought spherical harmonics. That is why we choose $j_{min} = 10000$ and j_{max} in the range between 10300 and 11600. The results for the case $e_0^2 = 0$ are shown in Figure 6. We can again observe that eigenvalue spectra intersect the zero-level starting at degree $j_{zero} \doteq 10500$. It means that whenever $j_{max} \geq j_{zero}$, the spectrum of matrix $\boldsymbol{A}$ contains an eigenvalue which is very close or equal to zero. Consequently, matrix $\boldsymbol{A}$ becomes ill-conditioned or even singular. Putting $e_0^2 = 0.006694$ (this case is not plotted here) has a similar stabilization effect as in the case shown in Figure 5.

To carry out the eigenvalue analysis of matrix $\boldsymbol{A}$ needs a lot of computer time. However, the critical spherical degree j_{zero} for which the existence of the solution to the Stokes two-boundary-value problem is not guaranteed can be estimated by analysing the existence of a solution for a model with a constant topographical height over the world. If we replace H_0 in the example in section 2.1. with the maximum topographical height H_{max}, then such an estimate j_{const} obviously underestimates the actual j_{zero}, i.e., it is too pessimistic, and hence it holds

$$j_{zero} \geq j_{const} \; , \tag{18}$$

where j_{const} is determined by the roots of function $K_j(H_{max})$ given by eqn.(14), i.e., j_{const} satisfies the equation

$$(j_{const} + 1) \left(\frac{R}{R + H_{max}} \right)^{j_{const}+2} - 2 = 0 \; . \tag{19}$$

For the examples in Figures 4 and 6, we obtain $j_{const} \doteq 698$ when $H_{max} = 53530$ metres, and $j_{const} \doteq 10158$ when $H_{max} = 5353$ metres. We have already learnt that the actual critical numbers are $j_{zero} \doteq 800$ and $j_{zero} \doteq 10500$, respectively. So, the criterion (18) estimates j_{zero} quite well.

3. Conclusion

This paper formulated and discussed the existence of a solution to the Stokes two-boundary-problem for geoid determination. We considered the boundary condition (6) relating to this problem without assuming that the surface gravity data had been continued from the Earth's surface to the geoid. The boundary condition (6) has not a usual form, because it contains the unknown anomalous potential referred to both the Earth's surface and the geoid coupled by the known topographical height. The numerical analysis of the 'two-boundary' condition (6) performed for a simplified model of the Earth's surface has revealed that the transfer matrix between the unknown potential on the geoid and the surface gravity anomalies may become ill-conditioned or even singular at a certain critical wavelength of a **finite** length. The existence of solution is not guaranteed for this critical geoidal wavelength. Once this ill-posed case occurs, to obtain a bounded and non-oscillating solution, the Stokes two-boundary-value problem must be regularized in such a way that this critical geoidal wavelength and its vicinity are excluded from the solution. We have given an estimate of critical geoidal wavelength; for the highest part of the Earth's surface, the critical geoidal wavelength is about 1 arcmin.

References

Heiskanen, W.H. and H. Moritz (1967). *Physical Geodesy.* W.H.Freeman and Co., San Francisco.

Jekeli, C. (1981). The downward continuation to the Earth's surface of truncated spherical and ellipsoidal harmonic series of the gravity and height anomalies. Rep. **323**, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus.

Martinec, Z. and C. Matyska (1997). On the solvability of the Stokes pseudo-boundary-value problem for geoid determination. *J. Geod.*, **71**, 103-112.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling (1992). *Numerical Recipes. The Art of Scientific Computing.* Cambridge Univ.Press, Cambridge.

Wieser, M. (1987). The global digital terrain model TUG87. Internal Report on Set-up, Origin and Characteristics, Inst. of Math. Geodesy, Technical Univ. of Graz, Austria.