PHYSICAL GEODESY

PETR VANÍČEK

January 1976

LECTURE NOTES
NO. 43

University of New Brunswick

Department of Surveying Engineering

PHYSICAL GEODESY

BY

Petr Vaníček

January 1976

PREFACE

In order to make our extensive series of lecture notes more readily available, we have scanned the old master copies and produced electronic versions in Portable Document Format. The quality of the images varies depending on the quality of the originals. The images have not been converted to searchable text.

Content

1) The topic of physical geodesy 1

2) Elements of the theory of potential 3

2.1) Concept of a field of force 3

2.2) Newton's gravitation 4

2.3) Gravitation field of a point with mass M 5

2.4) Gravitation field of a physical body 6

2.5) Field of force on and above the surface of a rotating body. 7

2.6) Notion of potential 9

2.7) Potential of an attracting point 10

2.8) Potential of an attracting body 11

2.9) Potential of gravity of a rotating body 12

2.10) Potential as a solution to Poisson's or Laplace's equation 13

2.11) Harmonic functions and their properties 16

2.12) Boundary-value problems 18

2.13) Some methods for solving the boundary-value problems 20

2.14) Eigenvalues and eigenfunctions 22

2.15) Laplacean in curvilinear coordinates, Lamé's coefficients 24

2.16) Fourier method as applied to the Laplacean in spherical coordinates 29

2.17) Eigenfunctions of the Laplacean in spherical coordinates, spherical harmonics 31

2.18) Orthogonality of spherical harmonics and development in spherical harmonics 35

2.19) Complete solution of the Laplace's equation in spherical coordinates 37

2.20) Complete solution of the Laplace's equation in ellipsoidal coordinates 40

2.21) Solution to the spherical boundary-value problems using spherical harmonics 41

2.22) Connection between the coefficients of the spherical harmonics and the attracting body 44

2.23) Physical interpretation of the coefficients by lower degree harmonics 47

2.24) Equipotential surfaces, lines of force 50

3) The gravity field of the earth and its approximations

3.1) Geoid 52

3.2) Remarks on Spheroid 54

3.3) "Normal" and disturbing potential 56

3.4) Sphere as a "normal" reference surface 57

3.5) Rotational ellipsoid as a "normal" reference surface..... 59

3.6) "Normal" gravity related to ellipsoid reference surface.. 62

3.7) Clairaut's theorem for gravity and geometrical flattennings 65

3.8) Somigliana's formulae for normal gravity 67

3.9) Cassinis' formulae for normal gravity, international formulae 69

3.10) Definitions of gravity anomaly, gravity disturbance, geoidal height, deflection of vertical 70

3.11) Relation between disturbing potential and geoidal heights, 2nd Bruns' formulae 73

3.12) Fundamental gravimetric equation 75

3.13) Discussion of the fundamental gravimetric equation, mixed boundary-value problem of geodesy 76

3.14) Vertical gradient of gravity 78

3.15) Solution of the mixed boundary-value problem of physical geodesy 82

3.16) Stokes' integral 84

3.17) Stokes' formula, gravimetric determination of the geoid.. 86

3.18) A few remarks about the Stokes' formula 87

3.19) Vening-Meinesz' formulae 89

3.20) Outline of numerical solution of Stokes' and Vening-Meinesz' formulae 91

4) Fundamentals of gravimetry 94

4.1) Gravity observations 94

4.2) Instruments used in gravimetry 95

4.3) Gravimetric networks 98

4.4) Processing of the observed gravity 100

4.5) Free-air correction and anomaly 101

4.6) Bouguer correction and anomaly 102

4.7) Terrain correction and refined Bouguer anomaly 107

4.8) Isostatic correction and anomaly 111

4.8.1) Isostasy 111

4.8.2) Pratt's model and theory 111

4.8.3) Airy's model and theory 113

4.8.4) Vening-Meinesz's model and theory....................114

4.8.5) Isostatic correction and anomaly.....................116

4.9) Other gravity corrections and anomalies.....................117

4.10) Indirect effect, cogeoid...118

4.11) Discussion of the individual gravity anomalies................119

4.12) Gravity maps, gravity data banks............................120

5) Heights...121

5.1) Observed heights..121

5.2) Geopotential numbers...123

5.3) Dynamic heights...124

5.4) Orthometric heights..125

5.5) Normal heights...130

5.6) Heights based on normal gravity.............................132

5.7) Discussion of the individual height systems...................133

5.8) Mean sea level as a height reference..........................135

6) Use of astronomic observations in geodesy..........................137

6.1) Geocentric Cartesian, Geodetic and Astronomic coordinate systems and their transformations; astro-deflections..................137

6.2) Definition of a "corresponding point" on the reference ellipsoid to a surface point; Helmert's and Pizzeti's projections...140

6.3) Relation between astro-deflection and gravimetric deflection...142

6.4) Astronomic and geodetic azimuths; Laplace's equation..........146

6.5) Astro-geodetic determination of the geoid (astronomic levelling)152

6.6) Astro-gravimetric determination of the geoid (astro-gravimetric levelling).......................................154

6.7) Determination of an optimum reference ellipsoid from the deflections of vertical and geoidal heights....................157

6.7.1) Relation between the change of the reference ellipsoid and the deflections and geoidal heights......157

6.7.2) Determination of the local optimum reference ellipsoid..161

6.7.3) Determination of the mean earth ellipsoid..............163

7) Miscellaneous topics..164

7.1) Reduction of observed horizontal angles to the reference ellipsoid...164

7.2) Reduction of observed distances to the reference ellipsoid.....166

7.3) Triangulated heights...170

7.4) "Three-dimensional" adjustment of geodetic networks............172

7.5) Discussion of various methods for geoid determination..........173

Acknowledgements...176

References...176

1) The topic of physical geodesy

In surveying, we are concerned with the determination of mutual position of points. When we work in small areas, we are able to get away with the measured relations. Hence the relation between two points can be expressed as

$$P_1 \rightleftarrows P_2 .$$

In larger areas, we cannot do the same. We are not able to measure the direct relations between the points and have to refer them to a common framework that interrelates the two points. Hence then we really speak about the relation

$$P_1 \rightleftarrows \text{framework} \rightleftarrows P_2 .$$

The description of such a framework and the relations between the points and the framework is one of the main concerns to geodesy.

In geodesy usually some kind of surface, "close" to the topographic surface of the earth, is chosen as a reference surface that plays the role of the framework. It is, of course, desirable that the reference surface be as close to the topographic surface as possible so that the individual points (whose position towards the topographic surface can be measured) can be related to the reference surface in a simple manner.

On the other hand, for the computational convenience, we want the reference surface to have the simplest possible geometrical shape. It is, for instance, conceivable that the topographical surface would not be a good reference from this point of view.

When measuring the positions and relations between the points on the surface of the earth (and above and below the surface point as well) we are subject to all kinds of physical influences from the physical environment. Our instruments obey some physical "laws" and "rules" which we have to try to understand in order to be able to interpret our measurements. We are all aware of the gravity force, Coriolis' force, air refraction, influences of temperature variations to name but a few.

For the static processes - as the geodetic observations are - the two most important physical influences are the refraction and the gravity. They both change the geometry of the space we are working in and have to be therefore studied and understood as clearly as possible. While we shall leave the study of refraction alone completely - this being one of the topics of surveying courses - we are going to devote our attention almost completely to the gravity.

The theoretical understanding of the gravity field, its determination and its relation (relevance) to the geometrical investigations (that constitute the main topic of surveying) is the field of physical geodesy. Hence the first semester of our course will be devoted to two main subjects. First we shall try to get some understanding and grasp of the mathematical model of a gravity field. This subject is known as the theory of potential. The second subject will be the earth gravity field and its approximations used in geodesy.

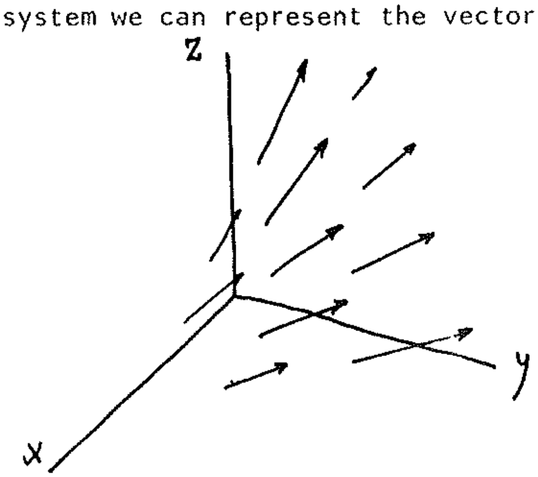
In this first half of the forthcoming semester, we should learn something about the mathematical tools used in physical geodesy.

The knowledge of these tools should enable us to follow, in the second half, the development of the classical concepts as how to determine the relation between the gravity field and some of the reference surfaces used in geodesy.

2) Elements of the theory of potential

2.1) Concept of a field of force

Where, in a certain area of our time-space, we have some physical forces acting, we often describe the area of interest by a vector-field, instead of dealing with the forces, their sources and the objects of the forces directly. By a vector field we understand a triplet of real numbers attributed to each point,(given by a four-tuple of real numbers) of our time space. Using Cartesian coordinate system we can represent the vector field graphically thus:



at any point of time.

To make things easier, we in physical geodesy, consider all the vector fields we work with as stationary, i.e., not changing in time.

Hence any such stationary field can be fully described by a three-valued function f, usually denoted as

$$\vec{f}\ ,\ \vec{f}(\vec{r})\ \in R_3\ ,\ \vec{r}\ \in R_3$$

(to describe the "three valuedness") of the arguments - the coordinates of the point in space. These three coordinates, real numbers, can be regarded as coordinates of the radius-vector of the point in question.

2.2) Newton's gravitation

At the beginning of all were the experimental results (astronomical observations) by a Danish astronomer Tycho-de-Brahe made in the second half of 16th century. These observations constituted the foundations on which a German astronomer-mathematician Johannes Kepler based the formulation of his famous three laws governing the motion of planets around the sun (beginning of 17th century). From these three experimental laws the English mathematician and physicist Isaac Newton derived his principle of gravitation (Philosophiae naturalis principia mathematica, 1687) which remained until our day a corner stone of the Newtonian mechanics.

The classical formulation of this principle is - "the force of mutual attraction of two masses m_1, m_2 is proportional to their product and inversely proportional to the square of their distance". In vector notation

$$\vec{f}_1 = \varkappa \frac{m_1 m_2}{\rho^3} \vec{\rho}_1\ ,\ \ \vec{f}_2 = \varkappa \frac{m_1 m_2}{\rho^3} \vec{\rho}_2$$

where $\vec{\rho}_1 = -\vec{\rho}_2$ are the vectors connecting the two masses (do not confuse with accelerations) and directed against the forces $\vec{f}_1, \vec{f}_2$ and κ is the constant of proportionality, called gravitation (Newton's) constant. From a multitude of measurements, the value of κ was determined as

$$\kappa = 6.67\ 10^{-8} \pm 0.007\ 10^{-8}\ [cm^3 g^{-1}\ sec^{-2}]$$

and accepted by a number of international scientific organizations as the best approximation known at this time. The argument is still going on as to whether the value of κ is varying with time.

Note the physical units of κ.

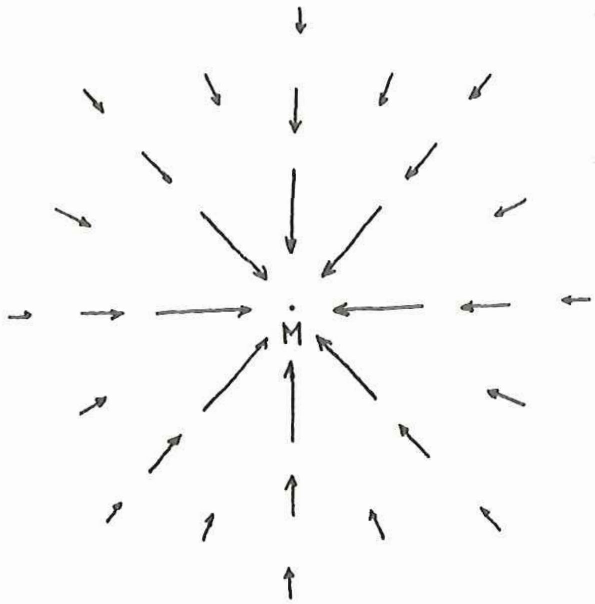
2.3) Gravitiation field of a point with mass M

We can see, that the Newton's principle of gravitation is completely symmetrical; i.e., there is no preference as to the masses. However, it is a matter of convenience to call one of the masses attracting and the other attracted. This allows us to reformulate the principle in terms of a force-field (vector field) as

$$\boxed{-\vec{f} = \kappa \frac{M}{\rho^3}\vec{\rho}}$$

understanding that the vector $\vec{f}$ represents a force exerted by the mass M on a unit mass m. The vector $\vec{\rho}$ is "directed" from M to m and is nothing but the radius vector of m if M is located in the ~~center~~ origin of coordinate system.

This is an example of a radial (or central) vector field where all the vectors point ~~from outside~~ towards one point (M). In any two-dimensional projection it looks thus:



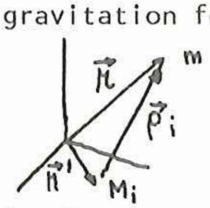
It is called <u>gravitation field of a point</u>.

Note that in this case, we are <u>not interested</u> in the effect that m exerts on M.

2.4) <u>Gravitation field of a physical body</u>

It was established through experiments that gravitation forces can be added in the same way as the ordinary three-dimensional vectors in Eucleidian space E_3. Hence if we have two masses M_1, M_2 acting on a unit mass m, we can write for the compound gravitation force:

$$\vec{f} = \vec{f}_1 + \vec{f}_2 = \varkappa \left(-\frac{M_1}{\rho_1^3}\vec{\rho}_1 - \frac{M_2}{\rho_2^3}\vec{\rho}_2\right),$$



the vectors ρ_1, ρ_2 being the respective vectors connecting M_1, M_2 with m.

We can similarly write for a whole system of masses M_1, $M_2, \ldots,$ M_n:

$$\vec{f} = \sum_{i=1}^{n} \vec{f}_i = -\varkappa \sum_{i=1}^{n} \frac{M_i}{\rho_i^3} \vec{\rho}_i \; , \; \vec{\rho} = \vec{r} - \vec{r}\,' \; , \; \vec{r}\,' \text{ is the rad. vector of } M_i .$$

Here, we are again not interested in the gravitation acting between the individual masses M_i, neither are we interested in the effect of m on the M's.

If we imagine a physical body as a~~n~~ region area B of E_3, with a density $\sigma(\vec{r})$ attributed to each point from the area, then the mass ΔM of a differential part ΔB of the body will be given by the product $\Delta M = \Delta B \sigma\ (\vec{r})$

Where $\sigma(\vec{r}\,')$ is the value of density in a representative point of ΔB. alligen

Then we can write for the gravitation field of the body B:

$$\boxed{\vec{f} = -\kappa \int_B \frac{\sigma}{\rho^3}\ \vec{\rho}\ dB \qquad \vec{\rho} = \vec{r} - \vec{r}\,'} \; ,$$

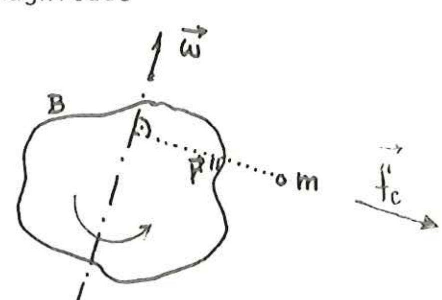
$\vec{r}\,'$being the rad. vector of the element dM., and thus the dummy variable in the integration.

Note that here σ is a function of the position of the element dB and $\vec{\rho}$ is a function of the position of both the element and the point we are investigating the field for.

2.5) Field of force on and above the surface of a rotating body (when you rotate with it)

It is again known from experience that a forced rotation of a mass m with rotational (angular) velocity ω at a distance r'' from the axis of rotation forces ~~pushes~~ the mass in the direction away from the axis of rotation. This pressure (force) has a magnitude

$$f_c = r''\omega^2 m \; .$$



The expression for the centrifugal force, as it is known, in vectorial form is

$$\boxed{\vec{f}_c = \omega^2 m \vec{r}'' .}$$

Let us imagine now a situation when a unit mass is forced to rotate on or above a body B. It is first attracted by the gravitation force of the body and second ~~pushed away~~ repudiated by the centrifugal force. The combined force, known as gravity is hence given by

$$\vec{f} = \vec{f}_g + \vec{f}_c = -\varkappa \int_B \frac{\sigma}{\rho^3} \vec{\rho}\, dB + \omega^2 \vec{r}'' .$$

Note the difference between $\vec{r}''$ and $\vec{\rho}$!

These are the two forces we are experiencing on the surface of the earth acting on a stationary object.

Note that if $f_g > f_c$ the object is attracted towards the body, if $f_g < f_c$ the object is pushed from the body.

2.6) Notion of potential

The field of force is a very useful representation of a physical environment. However, the necessity of having to know three real numbers (coordinates of the force vector) for each point in space is inconvenient. For this reason, it is better to adopt a simpler tool to depict the physical framework. One of such simpler tools is the potential.

The relation of potential (scalar field) to the field of force (vector field) is very much the same as the relation of the primitive function to the original function in the analysis of real variable. There the primitive function F (if it exists) is related to the original function f by relations

$$F(x) = \int f\ (x)\ dx, \quad \frac{dF(x)}{dx} = f(x) .$$

Here the potential V (if it exists) is related to the force $\vec{f}$ by similar equations:

$$V(\vec{r}) = \int \vec{f}(\vec{r})\ d\vec{r} \ , \quad \nabla(V(\vec{r})) = \text{grad } V(\vec{r}) = \vec{f}(\vec{r})$$

where ∇ (or grad) operator is the vectorial equivalent of $\frac{d}{dx}$ operator in the ordinary analysis. We speak of V as of the potential of $\vec{f}$ and of $\vec{f}$ as of the gradient of V.

Note that $\vec{r}$ here means the radius (position) vector of the point we are computing the potential (force) for. In E_3 $\vec{r}$ is simply (x,y,z) or as it is sometimes denoted

$$\vec{r} = x\vec{i} + y\vec{j} + z\vec{k} \ ,$$

$\vec{i}$, $\vec{j}$, $\vec{k}$ being the unit vectors in the coordinates axes.

It is usually not easy to integrate the vector field to get its potential even if it exists. It leads to the integral equations - difficult subject on its own. Thus we usually try to bypass this difficulty somehow. If the potential exists it sufices to show that its gradient is the original vector field. In other words if we find a scalar field the gradient of which is identical with the original vector field we have found the potential.

Potential is the most important notion used in physical geodesy.

2.7) Potential of an attracting point

We can show that the potential of an attracting point of a mass M is given by:

$$\boxed{V(\vec{r}) = \varkappa \frac{M}{r}},$$ assuming M located again in the centre of the coordinate system.

We have:

$$\nabla(V) = \frac{\partial V}{\partial x}\vec{i} + \frac{\partial V}{\partial y}\vec{j} + \frac{\partial V}{\partial z}\vec{k} = \frac{\partial V}{\partial r}\left(\frac{\partial r}{\partial x}\vec{i} + \frac{\partial r}{\partial y}\vec{j} + \frac{\partial r}{\partial z}\vec{k}\right)$$

$$= \frac{\partial V}{\partial r}\nabla(r)$$

$$r = (x^2+y^2+z^2)^{1/2} \Rightarrow \frac{\partial r}{\partial x} = \frac{1}{2} r^{-1} 2x = xr^{-1}, \frac{\partial r}{\partial y} = yr^{-1},$$

$$\frac{\partial r}{\partial z} = zr^{-1}, \nabla(\mathbf{r}) = \mathbf{r}/|\mathbf{r}|, \frac{\partial V}{\partial r} = -\kappa M r^{-2}$$

Hence

$$\nabla(V) = -\varkappa \frac{M}{r^3}\vec{r}$$

which is the expression for gravitation of a mass M as shown in 2.3 .

Thus

$$\nabla(V) = \vec{f}$$

which is the sufficient and necessary condition for V to be the potential of $\vec{f}$.

Notice the sign of V!

2.8) Potential of an attracting body

Similar to 2.7) it can be shown that

$$\boxed{V(\vec{r}) = \kappa \int_B \frac{\sigma}{\rho}\, dB}$$

where $\vec{\rho} = \vec{r} - \vec{r}\,'$, $\vec{r}\,'$ being the radius vector of the element $dM = \sigma dB$, is the potential of the attracting body B. We have

$$\nabla(V) = \nabla\left(\kappa \int_B \frac{\sigma}{\rho}\, dB\right) = \kappa \int_B \sigma\, \nabla\left(\frac{1}{\rho}\right) dB \; .$$

Since $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$, $\vec{r}\,' = \xi\vec{i} + \eta\vec{j} + \zeta\vec{k}$

we have $\vec{\rho} = (x - \xi)\,\vec{i} + (y - \eta)\,\vec{j} + (z - \zeta)\,\vec{k}$

and $\rho = ((x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2)^{1/2}$.

Hence $\nabla(1/\rho) = \nabla(\rho^{-1}) = \frac{\partial}{\partial \rho}(\rho^{-1})\left(\frac{\partial \rho}{\partial x}\vec{i} + \frac{\partial \rho}{\partial y}\vec{j} + \frac{\partial \rho}{\partial z}\vec{k}\right)$,

where $\frac{\partial \rho^{-1}}{\partial \rho} = -\rho^{-2}$ and $\frac{\partial \rho}{\partial x} \Leftarrow 2\rho d\rho = 2(x - \xi)\, dx$.

Therefore $\frac{\partial \rho}{\partial x} = (x - \xi)\rho^{-1}$, and cyclically:

$$\frac{\partial \rho}{\partial y} = (y - \eta)\rho^{-1}, \quad \frac{\partial \rho}{\partial z} = (z - \zeta)\,\rho^{-1} \; .$$ (Note the similarity with 2.7)

We get finally

$$\nabla(1/\rho) = -\rho^{-2}\,\rho^{-1}\,((x - \xi)\,\vec{i} + (y - \eta)\,\vec{j} + (z - \zeta)\,\vec{k}\,)$$

$$= -\frac{\vec{\rho}}{\rho^3}$$

and $\quad \nabla(V) = -\varkappa\int_B \frac{\sigma}{\rho^3}\vec{\rho}\, dB = \vec{f} \quad$ (see 2.4)

which is the necessary and sufficient condition for V to be the potential of $\vec{f}$. Notice again the sign of V.

2.9) Potential of gravity of a rotating body

The gravity force $\vec{f}$ is given as a sum of the gravitation force $\vec{f}_g$ and the centrifugal force $\vec{f}_c$. Since ∇ is a linear operation, i.e. $\nabla(A + B) = \nabla(A) + \nabla(B)$ for any two scalars A,B, we can try to find the potential of gravity in terms of a sum of two potentials - potential of gravitation and potential of the centrifugal force. Denoting the first by V and second by W we can write

$$\nabla(V + W) = \nabla(V) + \nabla(W) = \vec{f}_g + \vec{f}_c .$$

We know already V from 2.8 so that the problem is solved upon finding W.

It can be shown that

$$\boxed{W(\vec{r}) = \frac{1}{2} r''^2 \omega^2} \qquad (\text{note } \vec{r}'' = \vec{r}''(\vec{r}))$$

is the potential of the centrifugal force. We have

$$r'' = r \cos \alpha,$$

$\vec{r}''$ is the projection of $\vec{r}$ in the plane perpendicular to the axis of rotation. Let us put, for convenience sake, z axis into the axis of rotation. (This is not detrimental to the generality of the treatment). We get

$$\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$$

$$\vec{r}'' = x\vec{i} + y\vec{j} + o\vec{k}$$

and $\nabla(W) = \frac{1}{2}\omega^2 \nabla(r''^2)$

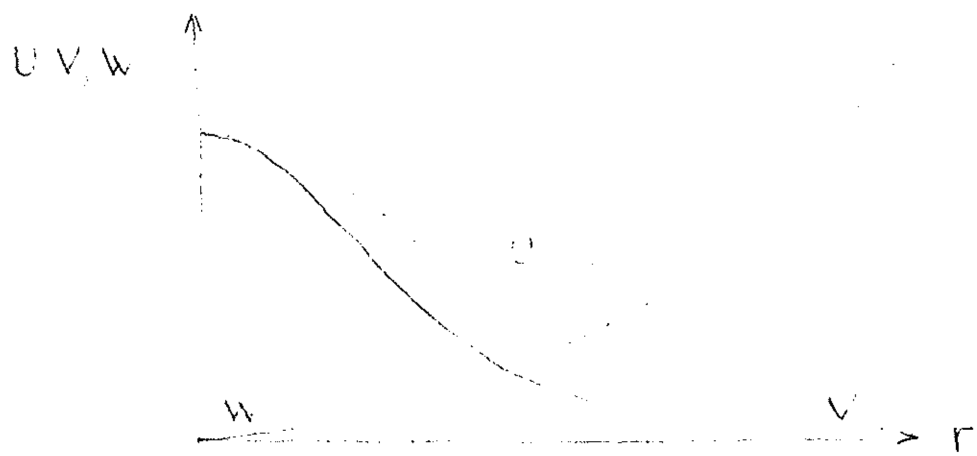
$$\nabla(r''^2) = 2r'' \left(\frac{x}{r''} i + \frac{y}{r''} j\right) = 2\vec{r}''.$$

Hence $\nabla(W) = \omega^2 \vec{r}'' = \vec{f}_c$ (see 2.5) .

Therefore the potential of gravity of a rotating body B is

$$\boxed{U = V + W = \kappa \int_B \frac{\sigma}{\rho}\, dB + \frac{1}{2} r''^2 \omega^2 .}$$

A brief look on the last formula will convince us that the first term decreases with ρ as we go away from the surface of B while the second term increases with r''^2. Hence there must be a locus where $\vec{f}_c = -\vec{f}_g$. The following diagram shows the situation in the terms of potentials.



Therefore there is a minimum of U (that coincides with $\vec{f}c = -\vec{f}g$) where ∇U does not have a radial component, i.e. the gradient of U is directed in the tangential direction. This is actually the place where the stationary satellites are placed.

Note that the integral has got a singularity at the point $\vec{r}$ if this is within or on the surface of the body B. Then $\vec{\rho}$ of the same point becomes zero and the integrated function goes to infinity. We may notice the same phenomenon with the gravitation force as well. This is a rather unfortunate property.

We shall now show that the problem of finding the appropriate potential can be transformed to a boundary value problem in partial derivatives.

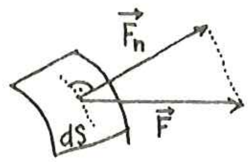
The above integral cannot be evaluated because $\sigma(\vec{r})$ is not known. Hence we have to look for another way how to evaluate U = > BVP

2.10) Potential as a solution to Poisson's or Laplace's equation

As we know from vector analysis the "first derivative" of a vector field $\vec{F}$, called also divergence of $\vec{F}$, is expressed as

$$\nabla(\vec{F}) = \operatorname{div} \vec{F} = \lim_{V_o \to 0} \frac{\oint\!\!\oint_S F_n dS}{V_o}$$

where V_o is the volume of the space embraced by the surface S and F_n is the magnitude of the vector $\vec{F}_n$ which in turn is the projection of $\vec{F}$ onto the normal to S.



The integration $\oiint_S$ is carried out over the whole enclosed surface S.

The term $F_n dS$ is often called <u>flux</u> of $\vec{F}$ through dS. If $\nabla(\vec{F})$ is in $\vec{r}$ positive, we speak about a "<u>source</u>" in $\vec{r}$, negative is called "<u>sink</u>".

Let us take now for simplicity the <u>S to be a sphere</u> .
(It can be shown that the shape of S in irrelevant and in limit we get always the same answer) and ask "what will be the divergence of a gravitation field?". The mass embraced by S will obviously be

$$M = \sigma V_o$$

with σ being the density of the mass within S. This mass will "radiate" (or rather "absorb") gravitation force

$$\vec{F} = - (\varkappa M/r^3)\ \vec{r}$$

when we locate the S in the centre of the coordinate system for simplicity. This force will obviously be normal to S everywhere so that

$$F_n = - \sigma V_o/r^2 \ .$$

Hence $\oiint_S F_n dS = - \dfrac{\varkappa\sigma V_o}{r^2} \oiint_S dS = - \dfrac{\varkappa\sigma V_o}{r^2} 4\pi r^2 = - 4\pi\varkappa\sigma V_o$

and $\qquad \nabla(\vec{F}) = \lim\limits_{V_o \to o} (-4\pi\varkappa\sigma) = -4\pi\varkappa\sigma \qquad$ (sink).

This <u>equation for divergence is valid for all the points throughout the space</u>. We may notice that $\nabla(\vec{F})$ depends on the value of σ. If we hence take a physical body B with density σ in a space with density 0 , we get

$$\nabla(\vec{F}(\vec{r})) = \begin{cases} -4\pi\varkappa\sigma & \vec{r} \in B \text{ except the surface of } B \\ -4\pi\varkappa\frac{\sigma}{2} = -2\pi\varkappa\sigma & \vec{r} \text{ on the surface of } B \\ 0 & \vec{r} \notin B. \end{cases}$$

Let us now have a look on $\nabla(\vec{F})$ itself. We can write

$$\nabla(\vec{F}) = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = \nabla\cdot F \quad \text{(scalar product)}.$$

But we have learnt that

$$\vec{F} = \nabla(V)$$

V being the potential of $\vec{F}$. Hence

$$\nabla(\vec{F}) = \nabla\cdot(\nabla(V)) = \frac{\partial^2 V}{\partial_x^2} + \frac{\partial^2 V}{\partial_y^2} + \frac{\partial^2 V}{\partial_z^2} = \Delta(V) \ .$$

Here $\nabla\cdot\nabla\ (\ \ ) = \nabla^2(\ \ ) = \Delta(\ \ )$ is a differential operator of second order. It is known as the Laplace's operator.

Putting these two results together we end up with the partial differential equations for V:

$$\Delta(V) = \begin{cases} -4\pi\varkappa\sigma & \text{in } B \\ -2\pi\varkappa\sigma & \text{on the surface of } B \\ 0 & \text{outside } B. \end{cases}$$

The first two equations are known as Poisson's **equation,** the general formula being

$$\Delta(V) = h(\vec{r}) \neq 0 \text{ ,where } h \text{ is a known function.}$$

The last equation

$$\Delta(V) = 0$$

is known as the Laplace's equation. They are the two fundamental differential equations of the theory of potential

We have thus learnt that the potential of gravitation of a physical body must satisfy the Poisson's equation in and on the surface of the body and the Laplace's equation outside the body.

Note that Δ is again a linear operator so that

$$\Delta(A + B) = \Delta(A) + \Delta(B)$$

$$\Delta(KA) = K\ \Delta(A)$$

for any two functions A, B and a constant K

Let us see, what differential equation is valid for the potential of the centifugal force. Puting z-axis into the axis of rotation of the body, we can write

$$r''^2 = x^2 + y^2.$$

On the other hand

$$\Delta W = \Delta(\tfrac{1}{2} r''^2 \omega^2) = \tfrac{1}{2}\omega^2\ \Delta r''^2.$$

But

$$\begin{aligned}\Delta r''^2 &= \frac{\partial^2}{\partial x^2}(x^2+y^2) + \frac{\partial^2}{\partial y^2}(x^2+y^2) + \frac{\partial^2}{\partial z^2}(x^2+y^2)\\ &= 2 + 2 + 0\\ &= 4.\end{aligned}$$

We conclude that since $\Delta r''^2 = 4$ in this particular coordinate system, it equals 4 in any coordinate system and we have

$$\Delta W = 2\omega^2,$$

We shall just notice that potential of gravitation outside the attracting body satisfies the Laplace's equation, the potential of gravity does not. The potential of gravity has to satisfy the following Poisson's equation:

$$\boxed{\Delta(V + W) = \Delta U = -4\pi\varkappa\sigma + 2\omega^2}$$

where only σ is a function of the position. Hence throughout the space U has discontinuous second derivatives only at the points (surfaces) where $\sigma(\vec{r})$ is discontinuous, i.e., on the surface of the body or inside the body if it has discontinuous density (Regions, layers, points). U itself is continuous throughout the space.

go to page 29

2.11) Harmonic functions and their properties

Function satisfying the Laplace's equation in region $\mathcal{A}$ are called harmonic in $\mathcal{A}$. For example, the gravitation potential of the attracting body is a harmonic function outside the body. Any harmonic function has got the following properties:

i) Attains both maximum and minimum values on the boundary of any enclosed region $\mathcal{B} \subset \mathcal{A}$, the values inside $\mathcal{B}$ being smaller than the maximum and larger than the minimum.

ii) Is analytic in all the points of $\mathcal{A}$, i.e. has got derivatives of any order.

iii) It lends itself to spherical inversion. This means that if $V(\vec{r})$ is harmonic inside (outside) a unit sphere, $\frac{1}{r} V(\vec{R})$, where $\vec{R} = \vec{r}/r^2$, is harmonic outside (inside) the same sphere, while the sphere transforms to itself (i.e. we talk about unit sphere centered upon the origin of the coordinate system). This property can be generalized for any $\mathcal{B} \subset \mathcal{A}$ with the consequence that $\mathcal{B}$ gets also inverted.

iv) The value that V, harmonic inside a sphere, attains in the center of the sphere is equal to the mean of all the values on the sphere. Hence, if the sphere is centered upon the origin and has a radius R we have

$$V(\vec{0}) = \frac{1}{4\pi R^2} \oiint_S V(\vec{r})\, dS.$$

v) The most important property has been established by a Frenchmathematician Dirichlet who proved that the values of a harmonic function on closed boundary surface determine one and only one harmonic function within the boundary. This became consequently known as the Dirichlet's principle. It has been shown that the corresponding harmonic function always exists if the boundary is "sufficiently smooth", i.e. has a continuously varying tangent plane and if the harmonic function is allowed to disappear in the infinity (when the region is considered infinite).

Let us just state the proof of the Dirichlet's principle, which is the easiest to prove: let $\mathcal{B}$ be a finite region with boundary S. Let us suppose that there are two harmonic functions, V, W that can attain the same values on the boundary S. Hence U = V - W, also a harmonic function due to the linearity of Δ operator, must have its boundary-value zero. But, according to the first property of harmonic functions, both maximum and minimum lay on the boundary. Since the extremes for U equal both to zero, U must be zero even within the whole $\mathcal{B}$. Hence V - W = 0 and V = W which concludes the proof.

Let us mention here that the function

$$\gamma(\vec{r}) = 1/(4\pi\vec{\rho})$$

is known as the fundamental harmonic function or fundamental solution to the Laplace's equation. The use of it will be shown later.

2.12) Boundary-value problems

The Dirichlet's principle insures that there exists a solution to the Laplace's equation if we know the values of the sought function on the boundary of a certain region.The problem of finding the harmonic function that would satisfy certain boundary conditions is called a boundary-value problem.

There are basically three types of boundary-value problems: The first, due to Dirichlet, whose name it usually bears is the one expressed in Dirichlet's principle. It can be stated as - "given the region of interest and the values of a harmonic function V on the boundaries of the region, find the harmonic function V in the region". This means that we have to solve the Laplace's equation ($\Delta V = 0$) knowing the value $V(\vec{r})$ $\vec{r} \in S$, where S is the closed boundary surface of the region of interest. The problem has a solution if and only if the assumptions of the Dirichlet's principle are satisfied.

Second boundary-value problem, bearing the name of Neuman, differs from the first only so far that we do not know the value of V on the boundary but instead we know the derivative

$$\frac{\partial V}{\partial n}(\vec{r}), \ \vec{r} \in S$$

of the sought function along the normal n towards the boundary S. For the second problem to have a solution inside a region it is necessary that

$$\oiint_S \frac{\partial V}{\partial n} \, dS = 0.$$

This condition follows immediately from the assumption that V is harmonic within the region, hence the flux of its gradient through the whole surface S has to be nil. This, together with the conditions for the Dirichlet's principle are all the sufficient conditions to ensure that the Newman's problem has a solution.

We speak about the third boundary-value problem when a linear combination of the first two boundary values is given on the surface S:

$$f(\vec{r}) = c_1 V(\vec{r}) + c_2 \frac{\partial V(\vec{r})}{\partial n}, \quad \vec{r} \in S,$$

where f is a function. Note that together with the assumptions for Dirichlet's principle, the equation

$$\oiint_S f(\vec{r}) \, dS = c_1 \oiint_S V(\vec{r}) \, dS$$

must be satisfied $(c_2 \oiint_S \frac{\partial V}{\partial n} dS = 0)$.

2.13) Some methods for solving the boundary-value problems

There are many different ways how to solve a boundary-value problem. We can use the operator calculus (Laplace's, Fourier and other transformations), functional analysis methods, transition to integral equations, Green's method, Fourier's method, numerical methods, to name but a few. They are all interrelated one way or the other and to venture into these would take a considerable amount of time. We shall briefly touch only the Green's method and devote our time mostly to the Fourier's method which is the best suited and therefore the most popular in physical geodesy.

The Green's method as applied to Dirichlet's problem for the interior of a region $\mathcal{A}$ consists of two steps:

i) first, we try to find the Green's function in the form

$$G(\vec{r}\,',\vec{r}) = \gamma(\vec{r}) + \mathbf{v}(\vec{r}\,',\vec{r})$$

where $\vec{r}\,', \vec{r} \in \mathcal{A}$, γ is the fundamental solution of the Laplace's equation (see 2.11) and $\mathbf{v}$ is, for any fixed $\vec{r}$, harmonic in $\mathcal{A}$, function of $\vec{r}\,'$. In addition, G on the boundary must be identically zero, i.e. $\mathbf{v} = -\gamma$ on the boundary. We can hence see, that G is a function of the shape of $\mathcal{A}$ only and generally it is a difficult problem to find the G for a specific $\mathcal{A}$.

ii) Once the Green's function is known, the solution to the Dirichlet's internal problem is given explicitly by:

$$V(\vec{r}\,') = -\oiint_S \frac{\partial G}{\partial n} V_S(\vec{r})\, dS$$

where $\frac{\partial G}{\partial n}$ is the outer gradient of G on the boundary S and V_S is the boundary value of V. Note that $\vec{r}$ becomes dummy variable in the integration.

For external problem, we have to use the spherical inversion. The Green's method in a slightly different form can be used for solving the boundary-value problem for Poisson's equation as well.

The special case of the Green's solution for the sphere of radius R is known as Poisson's integral:

$$V(\vec{r}\,') = \frac{R(r'^2 - R^2)}{4\pi} \oiint_S \frac{V(R,\theta,\lambda)}{(r^2+R^2-2Rr\cos\psi)^{3/2}}\, dS$$

where Ψ is the spherical angle of $\vec{r}\,'$ and $\vec{r}$.

The Fourier's method is based on an entirely different principle. It seeks the solution $V(x\ y\ z)$ in terms of a product of three independent functions:

$$V(x,y,z) = X(x)\cdot Y(y)\cdot Z(z)$$

or, as we sometimes call it, it seeks the separation of variables. The development leads to three separate ordinary differential equations of second order:

i) we suppose first that

$$V(x,y,z) = X(x) \cdot \Phi(y,z) .$$

Hence

$$\Delta V = \Delta(X\Phi) = \Phi\frac{\partial^2 X}{\partial x^2} + X\left(\frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2}\right) = 0$$

and

$$X^{-1}\frac{\partial^2 X}{\partial x^2} = -\Phi^{-1}\left(\frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2}\right) .$$

ii) Since the left hand side is a function of x only while the other side is a function of y,z, both sides must be constant (they obviously cannot vary because if they did they would vary independently and the equation cannot be satisfied). Hence

$$X^{-1}\frac{\partial^2 X}{\partial x^2} = C_1, \quad \Phi^{-1}\left(\frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2}\right) = -C_1$$

or, as we usually write it;

$$X'' - C_1 X = 0, \quad \Phi''_{yy} + \Phi''_{zz} + C_1\Phi = 0 .$$

Thus we have separated the first variable x.

iii) Let

$$\Phi(y,z) = Y(y)\cdot Z(z).$$

Then the second equation becomes

$$ZY'' + YZ'' + C_1 YZ = 0 .$$

Applying the same argument as in ii) we get

$$Y'' - C_2 Y = 0, \quad Z'' + (C_1 + C_2)\, Z = 0 .$$

The original partial differential equation is thus split into 3 ordinary differential equations that are related through the constants. Any solution of these three equations (for any value of C_1, C_2) that satisfies the boundary conditions is at the same time the solution of the boundary-value problem.

2.14) Eigenvalues and eigenfunctions

The three ordinary differential equations we end up with in Fourier's method may or may not have solutions for arbitrary values of the two constants C_1, C_2 under the prescribed boundary conditions. As a matter of fact, the ordinary differential equations we have to deal with are mostly of the Sturm-Luiville type (special case of self-adjoint differential equations) that have a solution only for some particular values of the constant. These values are known as eigenvalues of the equation in question.

The Sturm-Luiville equation is usually written as

$$(Ky')' - qy + \lambda\rho y = 0$$

where y is the sought function of x, K, ρ are some known, positive functions of x, q is a known, non-negative function of x and λ is a real number. In addition, ρ, known as the weight function, is required to be bounded. It can be shown that such an equation has got a solution for $x \in [a, b]$ for infinitity many values of λ (eigenvalues), all of them non-negative.

Every particular value of λ, say λ_i, gives one and only one particular solution y_i to the equation. These solutions (functions) are called eigenfunctions of the equation. There are, therefore, infinitely many, distinctly different, eigenfunctions for any Sturm-Luiville equation. It can be proved, that they create an orthogonal system of functions on [a,b] with weight ρ. We have hence:

$$\int_a^b y_i(x)\, y_j(x)\, \rho(x)\, dx = N_i\, \delta_{ij} \, ,$$

where $N_i = \int_a^b y_i^2(x)\, \rho(x)\, dx$ is know as the norm of y_i and δ_{ij} is the Kronecker's δ.

Example: For $K(x) = 1$, $q(x) = 0$, $\rho(x) = 1$ on [a,b], the Sturm-Luiville equation represents the equation of harmonic motion. We get, for its eigenvalues:

$$\lambda_i = \frac{4\pi^2}{(b-a)^2}\, i^2 \, , \quad i = 0,1,2,\ldots.$$

Its eigenfunctions are

$$\cos\sqrt{\lambda_i}\,x \, , \ \sin\sqrt{\lambda_i}\,x \, , \ i = 0,1,2,\ldots$$

As we know, any linear combination of the particular solutions that satisfies the boundary conditions is the solution of our boundary-value problem (one dimensional). We are going to show that this concept has basic importance in physical geodesy so that we shall be working with the eigenfunctions extensively. Note that all three ordinary differential equations derived in 2.13 are of the Sturm-Luiville type.

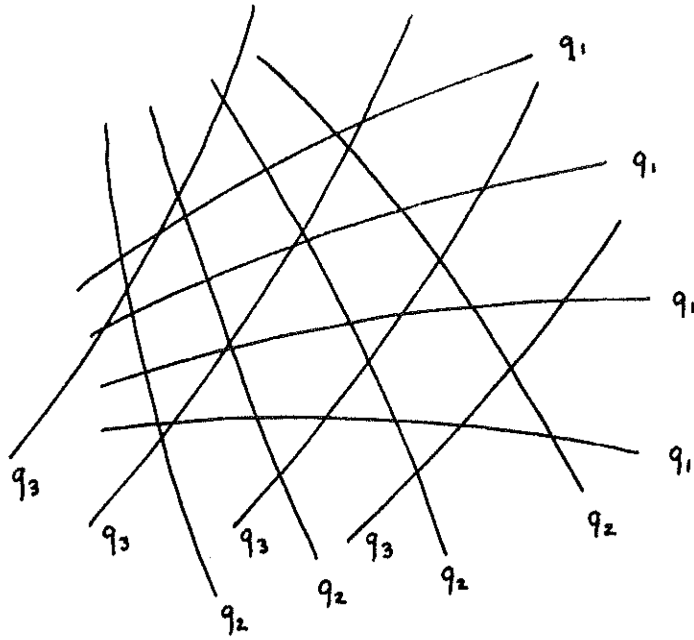
So far, we have been working with the common Cartesian coordinates x,y,z. However, they are not the best suited for geodetic purposes where we deal with the body of the earth which is roughly spherical or ellipsoidal. Hence we find the spherical or ellipsoidal coordinates handier. The transition to these systems will furnish the topic for the next few chapters.

2.15) Laplacean in curvilinear coordinates, Lamé's coefficients

Curvilinear coordinates: we say that we have defined a curvilinear coordinate system $(q_1 q_2 q_3)$ in E_3 if for every point (x,y,z) we have

$$(q_1, q_2, q_3) \rightleftarrows (x,y,z)$$

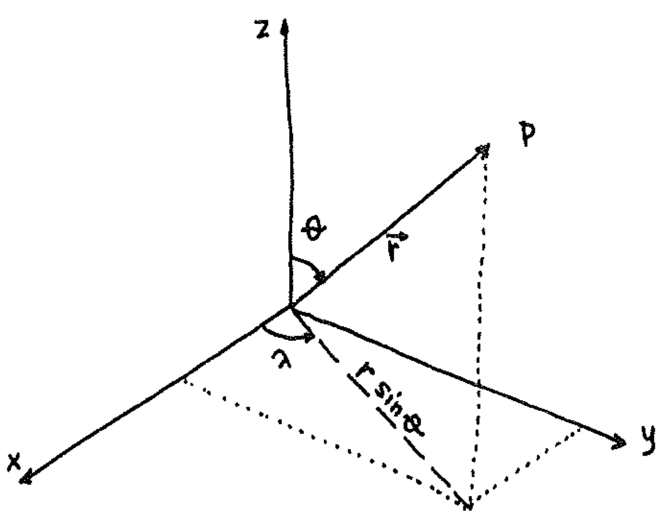
i.e. if we can express each q_i as a function of (x,y,z) and vice versa all the x,y,z as functions of q_i. Hence the one-to-one relationship is required.



Above is one example of coordinate lines of a curvilinear system.

Examples

1) Spherical coordinates:



$$x = r \sin \Theta \cos \lambda,$$

$$y = r \sin \Theta \sin \lambda,$$

$$z = r \cos \Theta,$$

and

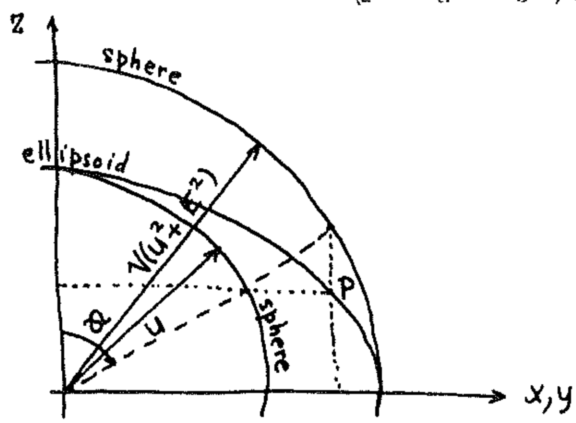
$$r = \sqrt{(x^2+y^2+z^2)},$$

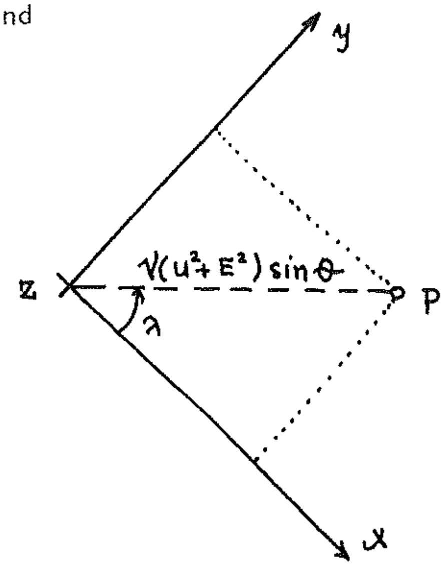
$$\Theta = \operatorname{arctg} \left(\sqrt{(x^2+y^2)}/z\right),$$

$$\lambda = \operatorname{arctg} (y/x) .$$

2) Ellipsoidal coordinates:

this system requires the focal length E to be given beforehand ($E^2 = a^2 - b^2$).





Then:

$$x = \sqrt{(u^2+E^2)}\ \sin\Theta \cos\lambda ,$$

$$y = \sqrt{(u^2+E^2)}\ \sin\Theta \sin\lambda ,$$

$$z = u \cos\Theta ,$$

and u given by equation

$$u^4 - u^2(x^2+y^2+z^2-E^2) - z^2E^2 = 0 ,$$

$$\lambda = \mathrm{arctg}\ (y/x) ,$$

$$\Theta = \arccos z/u .$$

Both these systems are locally orthogonal.

Lamé's coefficients. The functions H_i of q_1,q_2,q_3 defined as

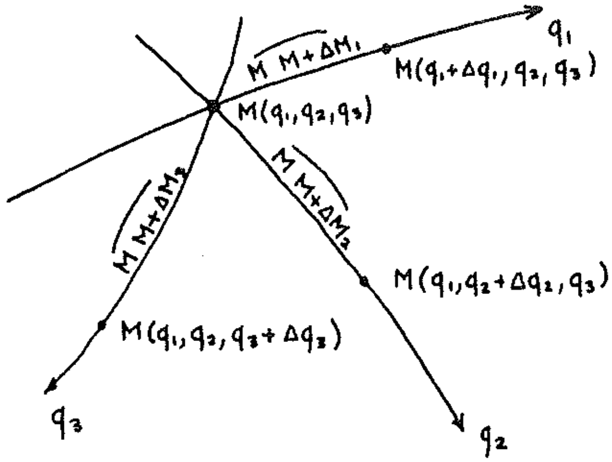
$$H_i = \lim_{\Delta q_i \to 0} \frac{\overset{\frown}{M\ M+\Delta M_i}}{\Delta q_i} \qquad i = 1,2,3$$

where $\widehat{M\,M+\Delta M_i}$ is the length of the q_i line connecting the two points

$$M = M(q_1,q_2,q_3),\quad M + \Delta M_1 = M(q_1+\Delta q_1,q_2,q_3) \text{ or}$$

$$M + \Delta M_2 = M(q_1,q_2+\Delta q_2,q_3) \text{ or}$$

$$M + \Delta M_3 = M(q_1,q_2,q_3+\Delta q_3) \ :$$



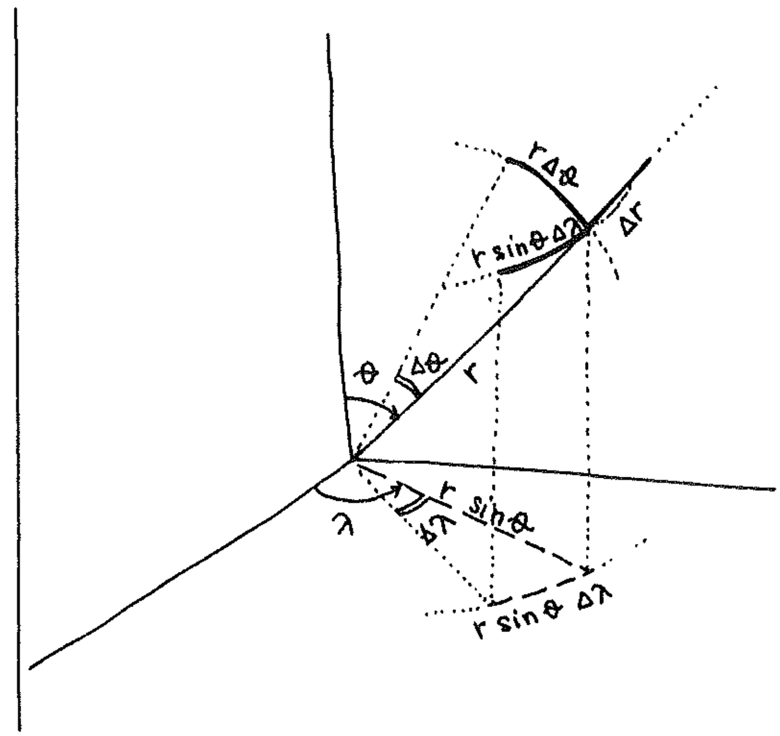
are known as Lamé's coefficients.

Ex. Spherical coordinates

$$H_r = \lim_{\Delta r\to 0} \frac{M(r,\Theta,\lambda)\ M(r+\Delta r,\Theta,\lambda)}{\Delta r} = \lim_{\Delta r\to 0} \frac{\Delta r}{\Delta r} = 1,$$

$$H_\Theta = \lim_{\Delta\Theta\to 0} \frac{M(r,\Theta,\lambda)\ M(r,\Theta+\Delta\Theta,\lambda)}{\Delta\Theta} = \lim_{\Delta\Theta\to 0} \frac{r\Delta\Theta}{\Delta\Theta} = r,$$

$$H_\lambda = \lim_{\Delta\lambda\to 0} \frac{M(r,\Theta,\lambda)\ M(r,\Theta,\lambda+\Delta\lambda)}{\Delta\lambda} = \lim_{\Delta\lambda\to 0} \frac{r\ \sin\theta\Delta\lambda}{\Delta\lambda} = r\ \sin\Theta$$



Problem: Derive the Lame's coefficients for ellipsoidal coordinates.

Note that for the x,y,z system, all the H_i's equal to 1!

It is evident that using the Lame's coefficients we can express the differential length increment dS_i along individual coordinate lines q_i as

$$dS_i = \widehat{MM + dM_i} = H_i dq_i \ .$$

The derivatives of any scalar field along these lines are then given by

$$\frac{\partial f}{\partial S_i} = \frac{1}{H_i}\frac{\partial f}{\partial q_i}$$

so that we can write for the gradient of f in curvilinear coordinates:

$$\nabla f = \sum_{i=1}^{3} \frac{1}{H_i}\frac{\partial f}{\partial q_i}\vec{e}_i \ .$$

Taking the differential volume $dV = \prod_i dS_i = \prod_i H_i dq_i$ we can similarly derive the expression for divergence of a vector field

$$\nabla \vec{f} = (\prod_i H_i^{-1}) \sum_i \frac{\partial (f_i H_i^{-1} \prod_j H_j)}{\partial q_i} \quad .$$

Realizing that $\Delta f = \nabla \cdot \nabla f$ we get for the Laplacean:

$$\Delta f = (\prod_i H_i^{-1}) \sum_i (\frac{\partial}{\partial q_i} (H_i^{-1} (\prod_j H_j) H_i^{-1} \frac{\partial f}{\partial q_i}))$$

$$= (\prod_i H_i^{-1}) \sum_i (\frac{\partial}{\partial q_i} ((\prod_j H_j) H_i^{-2} \frac{\partial f}{\partial q_i})) \quad .$$

Ex. Laplacean in spherical coordinates will be simply obtained by substituting for H_i from the earlier formulae

$$\Delta f = \frac{1}{r^2 \sin\Theta} (\frac{\partial}{\partial r} (\frac{r^2 \sin\Theta}{1} \frac{\partial f}{\partial r}) + \frac{\partial}{\partial \Theta} (\frac{r^2 \sin\Theta}{r^2} \frac{\partial f}{\partial \Theta}) + \frac{\partial}{\partial \lambda} (\frac{r^2 \sin\Theta}{r^2 \sin^2\Theta} \frac{\partial f}{\partial \lambda}))$$

$$= \frac{1}{r^2 \sin\Theta} (2r \sin\Theta \frac{\partial f}{\partial r} + r^2 \sin\Theta \frac{\partial^2 f}{\partial r^2} + \cos\Theta \frac{\partial f}{\partial \theta} + \sin\Theta \frac{\partial^2 f}{\partial \theta^2} + \frac{1}{\sin\Theta} \frac{\partial^2 f}{\partial \lambda^2})$$

$$\boxed{= \frac{2}{r} \frac{\partial f}{\partial r} + \frac{\partial^2 f}{\partial r^2} + \frac{\cot g\Theta}{r^2} \frac{\partial f}{\partial \Theta} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \Theta^2} + \frac{1}{r^2 \sin^2\Theta} \frac{\partial^2 f}{\partial \lambda^2}} \quad .$$

Problem: Derive the Laplacean in ellipsoidal coordinates.

2.16) Fourier method as applied to the Laplacean in spherical coordinates

Let us take the Laplacean in spherical coordinates as derived in 2.15 and seek the solution f in the following form:

$$f(r, \Theta, \lambda) = R(r) \cdot Y(\Theta, \lambda) \quad .$$

We get:

$$\frac{\partial f}{\partial r} = \frac{dR}{dr} Y = R'Y , \quad \frac{\partial^2 f}{\partial r^2} = R''Y , \quad \frac{\partial f}{\partial \Theta} = R \frac{\partial Y}{\partial \Theta} , \quad \frac{\partial^2 f}{\partial \Theta^2} = R\frac{\partial^2 Y}{\partial \Theta^2} ,$$

$$\frac{\partial f}{\partial \lambda} = R \frac{\partial Y}{\partial \lambda} , \quad \frac{\partial^2 f}{\partial \lambda^2} = R \frac{\partial^2 Y}{\partial \lambda^2} .$$ Substituting these into the

Laplace's equation, we get:

$$\Delta f = \frac{2}{r} R'Y + R''Y + \frac{\cot g\Theta}{r^2} R \frac{\partial Y}{\partial \Theta} + \frac{1}{r^2} R \frac{\partial^2 Y}{\partial \Theta^2} + \frac{1}{r^2 \sin^2\Theta} R \frac{\partial^2 Y}{\partial \lambda^2} = 0 .$$

Let us multiply the equation by $r^2/(RY)$. We obtain:

$$\Delta f = 2r \frac{R'}{R} + r^2 \frac{R''}{R} + \cot g\Theta \frac{\partial Y}{\partial \Theta}/Y + \frac{\partial^2 Y}{\partial \Theta^2}/Y + \frac{1}{\sin^2\Theta} \frac{\partial^2 Y}{\partial \lambda^2}/Y = 0 .$$

Hence

$$\frac{1}{R} (2rR' + r^2R'') = -\frac{1}{Y} \left(\cot g\Theta \frac{\partial Y}{\partial \Theta} + \frac{\partial^2 Y}{\partial \Theta^2} + \sin^{-2}\Theta \frac{\partial^2 Y}{\partial \lambda^2}\right) = \text{const.} = C_1$$

and we have separated the first variable r into the equation

$$r^2R'' + 2rR' - C_1R = 0 . \qquad \text{I}$$

The remaining two variables Θ, λ must satisfy the equation

$$\cot g\Theta \frac{\partial Y}{\partial \Theta} + \frac{\partial^2 Y}{\partial \Theta^2} + \sin^{-2}\Theta \frac{\partial^2 Y}{\partial \lambda^2} + C_1 Y = 0. \qquad \text{II}$$

Let us again seek the solution of II in terms of a product of two independent functions of T and L :

$$Y(\Theta, \lambda) = T(\Theta) \cdot L(\lambda).$$

We have:

$$\frac{\partial Y}{\partial \Theta} = \frac{dT}{d\Theta} L = T' \quad , \frac{\partial^2 Y}{\partial \Theta^2} = T'' \quad , \frac{\partial Y}{\partial \lambda} = TL', \quad \frac{\partial^2 Y}{\partial \lambda^2} = TL'' \text{ and after}$$

substitution to II we get:

$$\cot g\Theta\; T'L + T''L + \sin^{-2}\Theta\; TL'' + C_1 T\; L = 0.$$

Let us multiply this equation by $\sin^2\Theta/(TL)$ and we obtain:

$$\sin\Theta \cos\Theta\; T'/T + \sin^2\Theta\; T''/T + L''/L + C_1 \sin^2\Theta = 0 .$$

Hence:

$$\frac{1}{T}\left(\sin^2\Theta \;\; T'' + \sin\Theta \cos\Theta\; T'\right) + C_1 \sin^2\Theta = -\frac{L''}{L} = \text{const.} = C_2$$

and we end up with

$$\sin^2\Theta\; T'' + \sin\Theta \cos\Theta\; T' + (C_1 \sin^2\Theta - C_2)\, T = 0 \qquad \text{II}'$$

$$L'' + C_2 L = 0 . \qquad \text{II}''$$

Any function of r, Θ, λ that would satisfy the three equations (I, II', II'') and satisfy the boundary conditions as well is the solution of our boundary-value problem (formulated in spherical coordinates).

2.17) Eigenfunctions of the Laplacean in spherical coordinates, spherical harmonics

In order to see for which values of C_1 and C_2 the three equations have a solution, let us take the last equation first. The equation II'' is obviously the "equation of harmonic motion". Hence, according to 2.14, the eigenfunctions of II'' are

$$\cos\lambda\sqrt{\mu_m},\ \sin\lambda\sqrt{\mu_m},\ m = 0,1,2,\ldots\ .$$

The definition interval of λ is $[0,2\pi]$ or $[-\pi,\pi]$. Hence the eigenvalues are:

$$(C_2 =)\ \mu_m = m^2 \frac{4\pi^2}{4\pi 2} = m^2, \quad m = 0,1,2,\ldots$$

and the eigenfunctions can be written as

$$\cos m\lambda\ ,\ \sin m\lambda\ ,\ m = 0,1,2,\ldots\ .$$

Hence any linear combination of these trig. functions satisfies the equation II''.

The equation II' is slightly more difficult to deal with. It can be solved, for instance, by substitution

$$\tau = \cos\Theta.$$

Then we get: $\Theta = \arccos\tau$ and $\tau\in[-1,1]$. Further, we obtain:

$$T(\Theta) = T\ (\arccos\tau)\ ,$$

$$T' = \frac{dT}{d\Theta} = \frac{dT}{d\tau}\frac{d\tau}{d\Theta}\ ,$$

$$T'' = \frac{d}{d\Theta}\left(\frac{dT}{d\tau}\frac{d\tau}{d\Theta}\right) = \frac{d}{d\Theta}\left(\frac{dT}{d\tau}\right)\frac{d\tau}{d\Theta} + \frac{dT}{d\tau}\frac{d}{d\Theta}\left(\frac{d\tau}{d\Theta}\right) = \frac{d^2T}{d\tau^2}\left(\frac{d\tau}{d\Theta}\right)^2 + \frac{dT}{d\tau}\frac{d^2\tau}{d\Theta^2}\ .$$

Denoting $\frac{dT}{d\tau}$ by T'_τ and $\frac{d^2T}{d\tau^2}$ by $T''_{\tau\tau}$ and establishing that $\frac{d\tau}{d\Theta} = -\sin\Theta$, $\frac{d^2\tau}{d\Theta^2} = -\cos\Theta$ we get:

$$\sin^2\Theta(T''_{\tau\tau}\sin^2\Theta - T'_\tau\cos\Theta) + \sin\Theta\cos\Theta\, T'_\tau\,(-\sin\Theta) + (C_1\sin^2\Theta - C_2)\, T = 0.$$

This can be rewritten as

$$\sin^4\Theta\, T''_{\tau\tau} - 2\sin^2\Theta\cos\Theta\, T'_\tau + (C_1\sin^2\Theta - C_2)\, T = 0\ .$$

Substituting τ for $\cos\theta$ and $\sqrt{(1-\tau^2)}$ for $\sin\theta$ we get:

$$(1-\tau^2)^2\, T''_{\tau\tau} - 2\tau(1-\tau^2)T'_{\tau} + (C_1(1-\tau^2) - C_2)\, T = 0$$

or, as we usually write

$$(1-\tau^2)\, T''_{\tau\tau} - 2\tau T'_{\tau} + (C_1 - \frac{C_2}{1-\tau^2})\, T = 0 \ .$$

This equation is known as <u>Legendre's equation of $\sqrt{C_2}$-th order.</u>

It makes sense to try to find a solution only for such values of C_2 for which even II" has a solution; i.e., for

$$C_2 = m^2 \ .$$

Hence

$$(1-\tau^2)\, T''_{\tau\tau} - 2\tau T'_{\tau} + (C_1 - \frac{m^2}{1-\tau^2})\, T = 0 \ .$$

It can be seen that the Legendre's equation is again of the Sturm-Liuville type, particularly when we write:

$$((1-\tau^2)\, T'_{\tau})'_{\tau} - \frac{m^2}{1-\tau^2} T + C_1 T = 0 \ .$$

Here obviously $(1-\tau^2) = K$, $\frac{m^2}{1-\tau^2} = \sigma$, $1 = \rho$ which for $\tau \in (-1,1)$ satisfies the requirements for S.-L. equation.

It can be shown that its eigenvalues are:

$$(C_1 =)\ \ddot{\mu}_n = n(n+1) \ , \quad n \geq m$$

and the corresponding eigenfunctions

$$P_{nm}(\tau) = (1-\tau^2)^{\frac{m}{2}} \frac{d^m}{d\tau^m} P_n(\tau)$$

where

$$P_n(\tau) = \frac{1}{n!\cdot 2^n} \frac{d^n}{d\tau^n} (\tau^2-1)^n .$$

The functions P_{nm} are known as Legendre's associated functions (polynomials) of n-th order and m-th degree while P_n are known as Legendre's polynomials (functions). The latter are only a special case of the former for zero-degree.

Thus any linear combination of the associated Legendre's functions is a solution of equation II ! Hence any linear combination of the trig. functions and the associated functions is a solution to equation II (2.16). We can thus write:

$$\boxed{Y(\Theta,\lambda) = \sum_{\substack{m=o\\ n>m}}^{\infty} [(A_{nm} \cos m\lambda + B_{nm} \sin m\lambda) P_{nm} (\cos\Theta)] ,}$$

where A_{nm}, B_{nm} are arbitrary constants. The above expression can be written also in following ways:

$$Y = \sum_{n=o}^{\infty} Y_n = \sum_{n=o}^{\infty} \sum_{m=o}^{n} Y_{nm} = \sum_{n=o}^{\infty} \sum_{m=o}^{n} (A_{nm} C_{nm} + B_{nm} S_{nm}) .$$

The functions Y_n, Y_{nm}, C_{nm}, S_{nm} are all called (surface) spherical harmonics.

It is not difficult to see that on a sphere of radius $r = a$ (we have $R(a) = \text{const} = k$) we have

$$f(a,\Theta,\lambda) = k \sum_{n=o}^{\infty} Y_n = \sum_{n=o}^{\infty} \tilde{Y}_n$$

where $\tilde{Y}_n = kY_n$. Hence the solution of the Laplace's equation on any sphere is given by a linear combination of spherical harmonics. Therefore, the spherical harmonics are eigenfunctions of Laplacean on any sphere.

Reading on the Legendre's associated functions as well as the Legendre's functions of second kind is left on the students. The recommended source: W. A. Heiskanen, H. Moritz: Physical Geodesy.

2.18) Orthogonality of spherical harmonics and development in spherical harmonics

We have seen in 2.14 that any two eigenfunctions of a Sturm-Luiville equation are orthogonal on the appropriate interval with the weight ρ. Hence the functions $\cos m\lambda$, $\sin m\lambda$ are orthogonal on $[-\pi,\pi]$ with the weight 1. We have:

$$\int_{-\pi}^{\pi} \phi_i(\lambda)\ \phi_j\ (\lambda)\ d\lambda = N_i\ \delta_{ij}$$

for ϕ either sin or cos and

$$N_i = \begin{cases} 2\pi & i=0 \\ \pi & i\neq 0 \end{cases} .$$

The integral is, of course, always zero if ϕ_i, ϕ_j are not both either cos or sin.

On the other hand, the functions P_{nm} are orthogonal on $[0,\pi]$ for θ, ($[-1,1]$ for τ) with the weight 1. It can again be shown that

$$\int_{-1}^{1} P_{nm}(\tau)\ P_{km}(\tau)\ d\tau \quad =$$

$$= \int_{0}^{\pi} P_{nm}\ (\cos\theta)\ P_{km}\ (\cos\theta)\ \sin\theta\ d\theta \ = \ M_{nm}\ \delta_{nk} \qquad ,$$

where

$$M_{nm} = \frac{2}{2n+1}\ \frac{(n+m)!}{(n-m)!} \ .$$

Hence even any two functions

$$\phi_{nm}(\theta,\lambda) = \phi_m(\lambda)\; P_{nm}(\cos\theta)$$

where again $\phi_m(\lambda)$ is either $\cos m\lambda$ or $\sin m\lambda$, are orthogonal in the area

$$\mathcal{A} \equiv (0 \le \theta \le \pi,\; -\pi \le \lambda \le \pi)$$

with weight 1. We have

$$\int_{\mathcal{A}} \phi_{nm}(\theta,\lambda)\; \phi_{kl}(\theta,\lambda)\; d\mathcal{A} = \int_0^{\pi} \int_{-\pi}^{\pi} {}_m(\lambda)\; P_{nm}(\cos\theta)\; \phi_l(\lambda)$$

$$P_{kl}(\cos\theta)\sin\theta\; d\lambda\;\; d\theta \;=\; \int_0^{\pi} P_{nm}(\cos\theta)\; P_{kl}(\cos\theta)\sin\theta\; d\theta\;.$$

$$\int_{-\pi}^{\pi} \phi_m(\lambda)\; \phi_l(\lambda)\; d\lambda = M_{nm}\,\delta_{nk} \;\cdot\; N_m\,\delta_{ml} \;=$$

$$= M_{nm}\, N_n\, \delta_{nk}\, \delta_{ml} = \begin{cases} \dfrac{4\pi}{2n+1}\,\dfrac{(n+m)!}{(n-m)!}\,\delta_{nk}\,\delta_{ml} & m = 0 \\[2ex] \dfrac{2\pi}{2n+1}\,\dfrac{(n+m)!}{(n-m)!}\,\delta_{nk}\,\delta_{ml} & m \neq 0\,. \end{cases}$$

Note that by dividing the individual functions ϕ_{nm} by $\sqrt{(M_{nm}\, N_{nm})} =$ $= \sqrt{0_{nm}}$ the system becomes <u>orthonormal</u>. Functions

$$\phi_{nm}(\theta,\lambda) \;=\; \phi_{nm}(\theta,\lambda)/\sqrt{0_{nm}}$$

are orthonormal; i.e.,

$$\int_{\mathcal{A}} \phi_{nm}\, \phi_{kl}\; d\mathcal{A} = \delta_{nk}\, \delta_{ml}\;.$$

Given any integrable function $h(\theta,\lambda)$ defined on $\mathcal{A}$ we can develop it into generalized two-dimensional Fourier series:

$$h(\theta,\lambda) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} C_{nm}\, \phi_{nm}(\theta,\lambda) = \sum_{n=0}^{\infty} Y_n \; ,$$

where the coefficients C_{nm} are given by

$$C_{nm} = O_{nm}^{-1} \int_{\mathcal{A}} h(\theta,\lambda)\, \phi_{nm}(\theta,\lambda)\, d\mathcal{A} .$$

We are purposely not using $\oiint_{\mathcal{A}}$ for $\int_{\mathcal{A}}$ since $\mathcal{A}$ does not have to be closed in this development.

Note that a spherical surface is one of such areas $\mathcal{A}$ and that any function defined on the spherical surface can be thus developed into the series of spherical harmonics, without any connection with the Laplace's equation. If the function h happens to be the boundary-value of a boundary-value problem, then

$$R(r)\, h(\theta,\lambda)$$

is the solution of the boundary-value problem outside or inside the sphere for which the h is known.

2.19) Complete solution of the Laplace's equation in spherical coordinates

So far we have established that any linear combination of the surface spherical harmonics is a solution to equation II from 2.16. To complete the discussion of the Fourier method applied on Laplacean in spherical coordinates we have to find the solution of equation I from 2.16.

We have learnt that equation II has a solution only for

$$C_1 = n(n+1) \quad , \quad n = m,\, m+1\, , \quad \ldots \quad .$$

This must be born in mind when solving equation I, that changes to

$$r^2 R'' + 2rR' - n(n+1)\, R = 0 \; .$$

This equation is known as Euler's and can be solved by substitution $r = \exp(t)$. We can write:

$$t = \ln r \; ; \quad R'_r = R'_t \frac{dt}{dr} = R'_t \frac{1}{r} = R'_t e^{-t} \; ;$$

$$R''_{rr} = \frac{d}{dr}\left(R'_t \frac{dt}{dr}\right) = R''_{tt} \left(\frac{dt}{dr}\right)^2 + R'_t \frac{d^2t}{dr^2} = R'_{tt} e^{-2t} + R'_t \left(-\frac{1}{r^2}\right)$$

$$= R''_{tt} e^{-2t} - R'_t e^{-2t} \; .$$

Hence the Euler's equation becomes:

$$e^{2t}\left(R''_{tt} e^{-2t} - R'_t e^{-2t}\right) + 2e^t R'_t e^{-t} - n(n+1)R = 0$$

or

$$R''_{tt} + R'_t - n(n+1)R = 0 \; .$$

This is a linear equation of second order whose characteristic equation is

$$\alpha^2 + \alpha - n(n+1) = 0$$

and

$$\alpha_{1,2} = -\frac{1}{2} \pm \sqrt{\left(\frac{1}{4} + n(n+1)\right)} = -\frac{1}{2} \pm \sqrt{\left(n^2+n+\frac{1}{4}\right)} = -\frac{1}{2} \pm \sqrt{\left(n+\frac{1}{2}\right)^2}$$

$$= -\frac{1}{2} \pm \left(n+\frac{1}{2}\right) = \left\langle \begin{array}{l} n \\ -n-1 \; . \end{array} \right.$$

There are thus two sets of functions satisfying the Euler's equation:

$$R_n^{(1)} = e^{nt} = r^n \; , \quad R_n^{(2)} = e^{-(n+1)t} = r^{-(n+1)} \; .$$

We know already that for the boundary-value problem to have a solution outside a sphere, the disappearance of the solution in the infinity is a prerequisite. Hence $R^{(1)}$ cannot supply the solution to a boundary-value problem outside a sphere. On the ohter hand, $R^{(2)}$ cannot give a solution to a boundary-value problem inside a sphere because it grows beyond all limits for $r \to 0$ which contradicts the first and the fourth property of harmonic functions (see 2.11). Thus $R^{(1)}$ gives the solution to the Laplace's equation inside and $R^{(2)}$ outside a sphere.

We may, of course, have a sphere for which a function would be harmonic outside and inside (apart from a certain region or point because no function can be harmonic throughout the space !) in which case we require that both external and interval solution have the same value on the sphere. It is evident that this can happen only on a sphere with radius $r = 1$, the unit sphere. Really, one can see that the two solutions to Laplace's equations

$$f_i = \sum_{n=0}^{\infty} r^n Y_n \quad , \quad f_e = \sum_{n=0}^{\infty} r^{-(n+1)} Y_n$$

lend themselves to spherical inversion (see 2.11) if and only if one is a solution inside the unit sphere and the other is a solution outside the unit sphere.

In practice, though we seldom want to solve the boundary-value problem for the unit sphere. If we wish to solve the problem for a sphere of radius 'a', all we have to do is to scale the solution in such a way as to make them both agree on this new sphere. This is easily done and we can see that

$$\tilde{f}_i = \sum_{n=0}^{\infty} \left(\frac{r}{a}\right)^n Y_n \quad , \quad \tilde{f}_e = \sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^{n+1} Y_n$$

are the complete solutions for the inside of the sphere of radius 'a' or for the outside respectively.

2.20) Complete solution of the Laplace's equation in ellipsoidal coordinates

We shall just mention here that the complete solutions to the Laplace's equation in spherical coordinates has an analogy in ellipsoidal coordinates:

$$\tilde{f}_i = \sum_{n=0}^{\infty} \sum_{m=0}^{n} [p_{nm}(u,E,b)\ P_{nm}(\cos\Theta)\ (A_{nm}\cos m\lambda + B_{nm}\sin m\lambda)]$$

$$\tilde{f}_e = \sum_{n=0}^{\infty} \sum_{m=0}^{n} [q_{nm}(u,E,b)\ P_{nm}(\cos\Theta)\ (A_{nm}\cos m\lambda + B_{nm}\sin m\lambda)]$$

where
$$p_{nm}(u,E,b) = P_{nm}(i\,\frac{u}{E}) \,/\, P_{nm}(i\,\frac{b}{E})$$

$$q_{nm}(u,E,b) = Q_{nm}(i\,\frac{u}{E}) \,/\, Q_{nm}(i\,\frac{b}{E}) \quad .$$

Here "i" is the imaginary unit, Q_{nm} are the Legendre's functions of second kind and "b" is the semiminor axis of the ellipsoid (defined by b and E) towards which f_i is the internal and f_e the external solutions. This ellipsoid hence plays the same role as the sphere of radius 'a' has played in 2.19.

Note the similarity in structure of these formulae with the spherical ones. If it was not for the indeces m by p and q we would be able to write them in the same manner. Here because of the asymmetry of ellipsoidal coordinates with respect to Θ , the "radial" functions p, q depend on the order as well as degree of the surface spherical harmonic with which it is combined.

Further reading on this topic is left to the reader. (Use Heiskanen and Moritz: Physical Geodesy).

If we denote $Y_n = \sum_{m=0}^{n} Y_{nm}$ we can write

$$f_i = \sum_{n=0}^{\infty} \sum_{m=0}^{n} p_{nm} Y_{nm}$$

$$f_e = \sum_{n=0}^{\infty} \sum_{m=0}^{n} q_{nm} Y_{nm}$$

2.21) <u>Solution to the spherical boundary-value problems using spherical harmonics</u>

In 2.18 we have shown that R'h is the solution to spherical Dirichlet's problem if $h(\Theta,\lambda)$ is the boundary value on the spherical surface of radius 'a'. Hence the solution of the spherical Dirichlet's problem can be written as

$$\boxed{f_i = \sum_{n=0}^{\infty} \left(\frac{r}{a}\right)^n Y_n \ , \quad f_e = \sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^{n+1} Y_n}$$

where the coefficients A_{nm}, B_{nm} by the surface spherical harmonics

$$\boxed{Y_n = \sum_{m=0}^{n} [P_{nm} (\cos \Theta) \ (A_{nm} \cos m\lambda + B_{nm} \sin m\lambda \ )]}$$

are determined by the integrals (developed in 2.18):

$$\boxed{\begin{aligned} A_{nm} &= \frac{2n+1}{2\pi} \frac{(n-m)!}{(n+m)!} \oint\!\!\oint_S h(\Theta,\lambda) \ P_{nm} \ (\cos \Theta) \ \cos m\lambda \ dS \ , \\ B_{nm} &= \frac{2n+1}{2\pi} \frac{(n-m)!}{(n+m)!} \oint\!\!\oint_S h(\Theta,\lambda) \ P_{nm} \ (\cos \Theta) \ \sin m\lambda \ dS \ . \end{aligned}}$$

(for m = 0 the term 2π is replaced by 4π). Hence $h(\Theta,\lambda) = \sum_{n=0}^{\infty} Y_n(\Theta,\lambda)$.

Note that $h(\Theta,\lambda) = f(a,\Theta,\lambda)$. The integration is carried over the whole sphere.

For the spherical Neuman's problem, when the boundary-value

$$h\ (\theta,\lambda) = \frac{\partial}{\partial n} f\ (r,\theta,\lambda)\Big|_{r=a} = \frac{\partial}{\partial r} f\ (r,\theta,\lambda)\Big|_{r=a}$$

is given we try again to get a solution in the form (we shall be for simplicity interested in the outside of the sphere 'a' only):

$$f_e = \sum_{n=0}^{\infty} R'_n\ Y'_n$$

where Y'_n is the spherical harmonic of n^{-th} degree determined from the same formulae as for Dirichlet's problem (but $h\ (\theta,\lambda)$ is now the derivative of the sought function with respect to the outer normal to the sphere), R'_n equals $\alpha_n R_n$ and α_n is for a particular n constant.

We shall show that such a solution really exists and

$$\alpha_n = -\ a/(n+1)\ .$$

To prove it, let us take the assumed solution

$$f_e = \sum_{n=0}^{\infty} -\frac{a}{n+1}\ (\frac{a}{r})^{n+1}\ Y'_n \qquad (*)$$

and differentiate it with respect to r. We obtain

$$\frac{\partial f}{\partial r} = \sum_{n=0}^{\infty} -\frac{a^{n+2}}{n+1}\ Y'_n \frac{\partial}{\partial r} r^{-n-1} = \sum_{n=0}^{\infty} (\frac{a}{r})^{n+2}\ Y'_n$$

and for r = a

$$\frac{\partial f}{\partial r}\Big|_{r=a} = \sum_{n=0}^{\infty} Y'_n$$

which is the boundary value again. Hence if

$$\oiint_S h\ dS = 0$$

and the Neuman's problem has a solution; the assumed solution (*) is the only one. It is usually written as

$$\boxed{f_e = -a \sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^{n+1} \frac{Y'_n}{n+1} .}$$

The most important spherical boundary-value problem in physical geodesy is the third. . We speak about the third boundary-value problem when a function

$$h(\Theta,\lambda) = \left(c_1 f(r,\Theta,\lambda) + c_2 \frac{\partial f}{\partial n}(r,\Theta,\lambda)\right)\Big|_{r=a} = \left(c_1 f(r,\Theta,\lambda) + c_2 \frac{\partial f(r,\Theta,\lambda)}{\partial r}\right)\Big|_{r=a}$$

is given on the sphere. We shall be again interested in the external case only and seek the solution in the form:

$$f_e = \sum_{n=0}^{\infty} R''_n Y''_n .$$

Here we assume again $h = \sum_{n=0}^{\infty} Y''_n$ and $R''_n = \beta_n R_n$. We can show that such a solution exists if we take

$$\beta_n = 1/(c_1 - c_2/a\ (n+1)) .$$

To prove it, we adopt the same approach as for Neuman's problem. The proof is left to the reader. We just recapitulate by stating that

$$\boxed{f_e = \sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^{n+1} \frac{Y''_n}{c_1 - c_2/a\ (n+1)}}$$

solves the third type spherical boundary-value problem for the outside of the sphere of radius r=a.

Note that <u>any</u> truncation of the series of spherical harmonics supplies a <u>precise</u> solution of the Laplace's equation, i.e. is a harmonic function From this point of view it does not matter how well the truncated series approximates the boundary value. Therefore any truncated series of spherical harmonics represents always a <u>potential</u> of some force. The degree of approximation of the actual potential depends on the degree of approximation of the boundary value. This is the main advantage of using spherical harmonics.

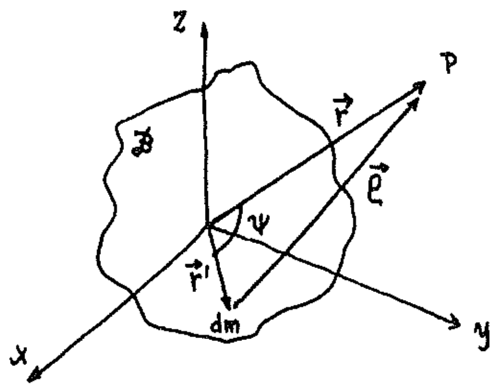
2.22) Connection between the coefficients of the spherical harmonics and the attracting body

Suppose, we are interested in solving the Dirichlet's problem for outside of a sphere of radius 'a' encompassing fully an attracting body $\mathcal{B}$. We are interested in what relation there will be between the coefficients of the spherical harmonics and the attracting body, i.e., can we say anything about the body $\mathcal{B}$ when we know its potential?

To establish this, let us take the potential of the attracting body (see 2.8)

$$V(\vec{r}) = \kappa \int_{\mathcal{B}} \frac{\sigma}{\rho}\, d\mathcal{B} .$$

Here $\vec{\rho} = \vec{r} - \vec{r}\,'$. Thus $\rho^2 = r^2 - 2\vec{r}\vec{r}\,' + r'^2$.



The scalar product $\vec{r}\vec{r}\,'$ can be rewritten as $rr' \cos\Psi$ so that

$$\rho = \sqrt{(r^2 - 2rr'\cos\Psi + r'^2)}$$

$$= r\sqrt{(1 - 2\frac{r'}{r}\cos\Psi + \frac{r'^2}{r^2})} .$$

It is known from the theory of Legendre's functions that $y = (1 - 2xt + t^2)^{-1/2}$ for $|x| \le 1$ and $|t| < 1$ is the "generating function" for the Legendre's polynomials that can be expressed as

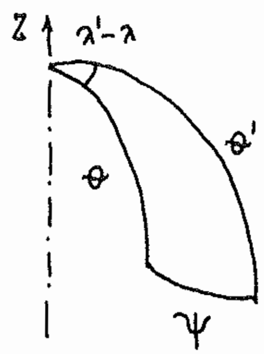
$$y = \sum_{n=0}^{\infty} P_n(x)\, t^n$$

where P_n are Legendre's polynomials (of zero degree).

It is easily seen that in our case $r > r'$ hence $|r'/r| < 1$ and $|\cos\Psi| \le 1$ so that we can write:

$$1/\rho = 1/r \sum_{n=0}^{\infty} P_n(\cos\Psi) \left(\frac{r'}{r}\right)^n .$$

Here $\cos\Psi$ can be expressed from the spherical triangle:



$$\cos\Psi = \cos\Theta\cos\Theta' + \sin\Theta\sin\Theta' \cos(\lambda'-\lambda) .$$

It can be shown by tedious computations that

$$P_n(\cos\Psi) = P_{n0}(\cos\Theta)\, P_{n0}(\cos\Theta') + $$

$$+ 2 \sum_{m=1}^{n} \Big[\frac{(n-m)!}{(n+m)!} P_{nm}(\cos\Theta)\, P_{nm}(\cos\Theta')$$

$$(\cos m\lambda \cos m\lambda' + \sin m\lambda \sin m\lambda')\Big] .$$

This formula is known as <u>decompositon formula</u> and we may notice its complete symmetry in Θ,Θ' and λ,λ' .

Substituting this result back into the expression for $1/\rho$ and that again back into the formula for V one gets:

$$V(\vec{r}) = \kappa\int_{\mathcal{B}} \sigma/r \sum_{n=0}^{\infty} \Big(\frac{r'}{r}\Big)^n \Big[P_{n0}(\cos\Theta)\, P_{n0}(\cos\Theta') + 2 \sum_{m=1}^{n} \Big(\frac{(n-m)!}{(n+m)!} P_{nm}(\cos\Theta)$$

$$P_{nm}(\cos\Theta')\, (\cos m\lambda \cos m\,\lambda' + \sin m\lambda \sin m\lambda')\Big)\Big]\, d\mathcal{B} .$$

Here the integration is carried over the whole body $\mathcal{B}$, i.e. all the points with coordinates r',Θ',λ' . Thus we can write:

$$V(\vec{r}) = \sum_{n=0}^{\infty} \Big(\frac{1}{r}\Big)^{n+1} \Big\{P_{n0}(\cos\Theta) \int_{\mathcal{B}} \kappa\sigma r'^n\, P_{n0}(\cos\Theta')\, d\mathcal{B} +$$

$$+ \sum_{m=0}^{n} P_{nm}(\cos\Theta) \Big(\int_{\mathcal{B}} 2\frac{(n-m)!}{(n+m)!} \kappa\sigma r'^n\, P_{nm}(\cos\Theta') \cos m\lambda'\, d\mathcal{B} \cos m\lambda$$

$$+ \int_{\mathcal{B}} 2\frac{(n-m)!}{(n+m)!} \kappa\sigma r'^n\, P_{nm}(\cos\Theta') \sin m\lambda'\, d\mathcal{B} \sin m\lambda\Big)\Big\} .$$

Multiplying and dividing each term by a^{n+1} we obtain

$$V(\vec{r}) = \sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^{n+1} \{P_{n0}(\cos\theta)\ \frac{\kappa}{a}\int_{\mathcal{B}} \sigma\left(\frac{r'}{a}\right)^n P_{n0}\ (\cos\theta')\ d\mathcal{B}\ +$$

$$\sum_{m=1}^{n} P_{nm}\ (\cos\theta)\ [\frac{2\kappa}{a}\ \frac{(n-m)!}{(n+m)!}\int_{\mathcal{B}} \sigma\left(\frac{r'}{a}\right)^n P_{nm}\ (\cos\theta')\ \cos m'\lambda\ d\mathcal{B}\ \cos m\lambda\ +$$

$$\frac{2\kappa}{a}\ \frac{(n-m)!}{(n+m)!}\int_{\mathcal{B}} \sigma\left(\frac{r'}{a}\right)^n P_{nm}\ (\cos\theta')\ \sin m'\lambda\ d\mathcal{B}\ \sin m\lambda\ ]\}.$$

Denoting

$$\left.\begin{aligned} &\frac{\kappa}{a}\int_{\mathcal{B}} \sigma\left(\frac{r'}{a}\right)^n P_{n0}\ (\cos\theta')\ d\mathcal{B} = A_{n0} \\ &\frac{2\kappa}{a}\ \frac{(n-m)!}{(n+m)!}\int_{\mathcal{B}} \sigma\left(\frac{r'}{a}\right)^n P_{nm}\ (\cos\theta')\ \cos m\lambda'\ d\mathcal{B} = A_{nm} \\ &\frac{2\kappa}{a}\ \frac{(n-m)!}{(n+m)!}\int_{\mathcal{B}} \sigma\left(\frac{r'}{a}\right)^n P_{nm}\ (\cos\theta')\ \sin m\lambda'\ d\mathcal{B} = B_{nm} \end{aligned}\right\}\ (*)$$

we can write:

$$V\ (\vec{r}) = \sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^{n+1} Y_n$$

which is our well-known formula for the external Dirichlet's solution from 2.21. The above equations (*) determine the relationship between the coefficients A_{nm}, B_{nm} and the attracting body $\mathcal{B}$.

Notice the structure of the integrand which is a product of a harmonic (inside the sphere r=a) function (with unit coefficients by the harmonics) with the density.

We can see that the formulae (*) do not give us much information about the body $\mathcal{B}$. However, we shall show in the next paragraph that some information can be gained from the lower degree harmonics.

2.23) Physical interpretation of the coefficients by lower degree harmonics

The formulae developed in 2.22 allow us to interpret the coefficients by lower degree harmonics physically. To do so we have to evaluate, the terms

$$C'_{nm} = P_{nm}\ (\cos \Theta')\ \cos m\lambda',\ S'_{nm} = P_{nm}\ (\cos \Theta')\ \sin m\lambda' \ .$$

The associated functions are given by:

$$P_{00} = 1$$
$$P_{10} = \cos \Theta'$$
$$P_{11} = \sin \Theta'$$
$$P_{20} = \frac{3}{2} \cos^2 \Theta' - \frac{1}{2}$$
$$P_{21} = 3 \sin \Theta' \cos \Theta'$$
$$P_{22} = 3 \sin^2 \Theta' \ .$$

Hence the functions C'_{nm}, S'_{nm} can be written as follows

$$C'_{00} = 1 \qquad S'_{00} = 0$$
$$C'_{10} = \cos \Theta' \qquad S'_{10} = 0$$
$$C'_{11} = \sin \Theta' \cos \lambda' \qquad S'_{11} = \sin \Theta' \sin \lambda'$$
$$C'_{20} = \frac{3}{2} \cos^2 \Theta' - \frac{1}{2} \qquad S'_{20} = 0$$
$$C'_{21} = 3 \sin \Theta' \cos \Theta' \cos \lambda' \qquad S'_{21} = 3 \sin \Theta' \cos \Theta' \sin \lambda'$$
$$C'_{22} = 3 \sin^2 \Theta' \cos 2\lambda' \qquad S'_{22} = 3 \sin^2 \Theta' \sin 2\lambda' \ .$$

Changing over to **Cartesian** coordinates x,y,z using the transformation from 2.15:

$$x = r' \sin \Theta' \cos \lambda' \ , \ y = r' \sin \Theta' \sin \lambda' \ , \ z = r' \cos \Theta'$$

we get:

$$
\begin{aligned}
C'_{00} &= 1 & S'_{00} &= 0\\
r'C'_{10} &= z & r'S'_{10} &= 0\\
r'C'_{11} &= x & r'S'_{11} &= y\\
r'^2C'_{20} &= \tfrac{1}{2}x^2 - \tfrac{1}{2}y^2 + z^2 & r'^2S'_{20} &= 0\\
r'^2C'_{21} &= 3xz & r'^2S'_{21} &= 3yz\\
r'^2C'_{22} &= 3x^2-3y^2 & r'^2S'_{22} &= 6xy\,.
\end{aligned}
$$

The proof of the latter formulae is left to the reader.

Substituting these results into the expressions (*) for A_{nm}, B_{nm}, we obtain

$$
\begin{aligned}
A_{00} &= \frac{\kappa}{a}\int_{\mathfrak{B}} \sigma \, d\mathfrak{B} & B_{00} &= 0\\
A_{10} &= \frac{\kappa}{a^2}\int_{\mathfrak{B}} \sigma z \, d\mathfrak{B} & B_{10} &= 0\\
A_{11} &= \frac{\kappa}{a^2}\int_{\mathfrak{B}} \sigma x \, d\mathfrak{B} & B_{11} &= \frac{\kappa}{a^2}\int_{\mathfrak{B}} \sigma y \, d\mathfrak{B}\\
A_{20} &= \frac{\kappa}{a^3}\int_{\mathfrak{B}} \sigma (2z^2-x^2-y^2) \, d\mathfrak{B} & B_{20} &= 0\\
A_{22} &= \frac{\kappa}{a^3}\int_{\mathfrak{B}} \sigma xz \, d\mathfrak{B} & B_{21} &= \frac{\kappa}{a^3}\int_{\mathfrak{B}} \sigma yz \, d\mathfrak{B}\\
A_{22} &= \frac{\kappa}{4a^3}\int_{\mathfrak{B}} (x^2-y^2) \, d\mathfrak{B} & B_{22} &= \frac{\kappa}{2a^3}\int_{\mathfrak{B}} \sigma xy \, d\mathfrak{B}\,.
\end{aligned}
$$

On the other hand, the coordinates of the centre of gravity of $\mathfrak{B}$ are given by (M is the mass of $\mathfrak{B}$):

$$\xi = \frac{1}{M}\int_{\mathfrak{B}} \sigma x d\mathfrak{B}$$

$$\eta = \frac{1}{M}\int_{\mathfrak{B}} \sigma y d\mathfrak{B}$$

$$\zeta = \frac{1}{M}\int_{\mathfrak{B}} \sigma z d\mathfrak{B}.$$

Introducing, in addition, the matrix of the tensor of inertia of $\mathfrak{B}$ in the origin of the coordinate system:

$$\mathcal{J} = \left\| \begin{matrix} \int_{\mathfrak{B}}(y^2+z^2)\sigma d\mathfrak{B}, & -\int_{\mathfrak{B}} xy\sigma d\mathfrak{B}, & -\int_{\mathfrak{B}} xz\sigma d\mathfrak{B} \\ -\int_{\mathfrak{B}} xy\sigma d\mathfrak{B}, & \int_{\mathfrak{B}}(x^2+z^2)\sigma d\mathfrak{B}, & -\int_{\mathfrak{B}} yz\sigma d\mathfrak{B} \\ -\int_{\mathfrak{B}} xz\sigma d\mathfrak{B}, & -\int_{\mathfrak{B}} yz\sigma d\mathfrak{B}, & \int_{\mathfrak{B}}(x^2+y^2)\sigma d\mathfrak{B} \end{matrix} \right\| =$$

$$= \left\| \begin{matrix} A & -D & -E \\ -D & B & -F \\ -E & -F & C \end{matrix} \right\|$$

where A,B,C are the principal moments of inertia in the coordinate system's origin and D,E,F are the products of inertia (deviation moments), we get

$$A_{00} = \frac{\kappa}{a} M \qquad B_{00} = 0$$

$$A_{10} = \frac{\kappa}{a^2} M \zeta \qquad B_{10} = 0$$

$$A_{11} = \frac{\kappa}{a^2} M \xi \qquad B_{11} = \frac{\kappa}{a^2} M \eta$$

$$A_{20} = \frac{\kappa}{a^3}\left(\frac{A+B}{2} - C\right) \qquad B_{20} = 0$$

$$A_{21} = \frac{\kappa}{a^3} E \qquad B_{21} = \frac{\kappa}{a^3} F$$

$$A_{22} = \frac{\kappa}{4a^3} (B-A) \qquad B_{22} \frac{\kappa}{2a^3} D ,$$

Hence, the coefficients by the lower degree harmonics have quite definite physical meaning. This discovery will help us at later stages to get some insight into the formulae used in physical geodesy. Here, we just notice that whatever shape the attracting body may have and whatever distribution of density, the first few spherical harmonics of its gravitation potential depend only on its principal moments of inertia and its products of inertia.

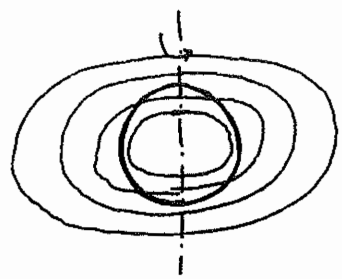
2.24) Equipotential surfaces, lines of force

The loci of equal potential

$$V(\vec{r}) = \text{const.}$$

are called equipotential surfaces of the potential V. For various values of the constant we get various equipotential surfaces. Because the potential is continuous throughout the space, anayltical in the area where it is harmonic and has only discontinuous second derivatives on the boundaries of validity of Laplace's and Poisson's equation, the equipotential surfaces of a gravity potential (gravity field) are smooth. Their curvature varies smoothly apart from the places where density changes suddenly, i.e., their curvature changes as suddenly as the density does. The equipotential surfaces never cross each other and would look very much like the onion peals.

Example a crossection of a gravity potential of a rotating rigid sphere with homogeneous distribution of density may look thus

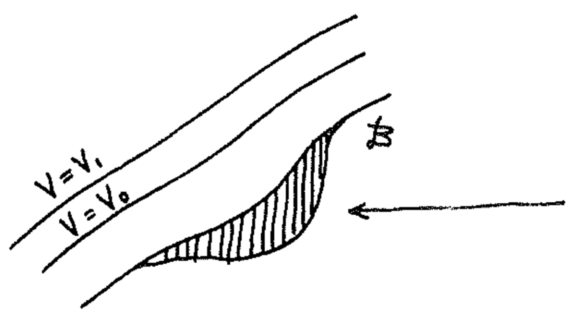


The lines of force are the curves to which the gradient of the potential; i.e., the field of force, is tangent in every point. They are always perpendicular to the equipotential surfaces as can be proved by a simple computation. We get for the total differential of the potential V:

$$dV = \frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy + \frac{\partial V}{\partial z} dz = \nabla\vec{V} \cdot \vec{da} \quad .$$

This formula provides us with the tool to determine what happens with dV when we point $\vec{da}$ in various directions. It is clear that if $\vec{da}$ lies in the tangent plane to the equipotential V = const., dV must be zero; there is no increment of the potential if we move on the equipotential surface. But for $\vec{\nabla}V\cdot\vec{da}$ to be zero it is necessary that $\vec{\nabla}V$ be perpendicular to $\vec{da}$; i.e., $\nabla\vec{V}$ must coincide with normal to V = const., which was to be proved.

We can also see that there is no force acting in the equipotential surface. This is the reason why a homogeneous elastic body tries to reach a shape that conforms with one of the equipotential surfaces. In such a state there are no tangential forces (strains) acting on the surface and the body is in equilibrium. For a rigid body though the tangential forces are always present. If the elastic body is not homogeneous, it does not follow the shape of an equipotential surface. Denser elements are "pulled into the body" more than the lighter ones, pushing the lighter elements aside. This "additional" force contributes towards the balance of forces making it different from the equilibrium of homogeneous body. The equipotential surfaces then reflect, to a certain extent, the "immersion" of the denser elements



area of higher density

3) The gravity field of the earth and its approximations

3.1) Geoid

The earth as a whole behaves as an unhomogeneous elastic body. It has reached a certain equilibrium so that it does not depart "too far" from one of its equipotential surfaces. Whenever it does depart, it is due to

i) locally rigid earth crust (with its topography that, of course, cannot conform with the equipotential surface;

ii) unevenly distributed density of its masses.

If the oceans were homogeneous; i.e., if the density of water were the same throughout (salinity, temperature, mineral content, etc.), and if there were no dynamic effects (currents, sheer stresses, river discharges etc.) the ocean surface would follow an equipotential surface. Unfortunately, the oceans do not behave reasonably and their surface depart from the equipotential surface at places by alegedly some $\pm$ 2m. In addition, they do not all have the same level and are probably even "sloping" away from north and south, due to the continuous melting of the polar caps.

The equipotential surface going through the ocean surfaces in average is called geoid. Mathematically, the geoid can be again written as

$$U(r,\phi,\lambda) = \sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^{n+1} Y_n + \frac{1}{2}\omega^2 r''^2 = U_o = \text{const.} ,$$

where 'a' is the radius of a sphere encompassing all the masses of the earth; i.e., sphere outside which the earth gravitational potential harmonic. Such a sphere is generally known as reference sphere. In practice, the reference sphere is not required to encompass all the earth. The excess of the masses outside the sphere can be eliminated computationally. We can see that if we knew the value of gravitational potential (or for that matter if we knew the normal derivative of this potential or alternatively a linear combination of the potential and its normal derivative) on the reference sphere, we would be able to determine the geoid. The determination would involve the evaluation of $U(\vec{r})$ in a certain area and tracing of the geoidal surface $U(\vec{r}) = U_o$.

Another way to express the geoid is

$$U(u,\phi,\lambda) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} q_{nm} (u,E,b)\, Y_{nm} + \frac{1}{2}\omega^2 r''^2 = U_o = \text{const.}$$

using the ellipsoidal harmonics. Here, the ellipsoid (b,E) is the reference surface, called reference ellipsoid. The geoid expressed one way or the other is, of course, identical. The spherical harmonics Y_n in the spherical solution are exactly the same as Y_{nm} used in the ellipsoidal solution. The reference ellipsoid should contain all the earth. In practice it does not and the deficiency is accounted for artificially.

Unfortunately, though, we do not know the values of the gravitational potential on the reference surface. We cannot, therefore, use the described approach. In the forthcoming chapters, we are going to see how we go about solving the problem.

The question may remain in a reader's mind, why are we interested in the equipotential surfaces rather than any other characteristic of the gravity field. The explanation is simple -- the equipotential surfaces have an immediate application in geodesy. They define the local "horizontal plane" (tangent plane to the equipotential surface in a point) to which we align our instruments when setting them up. Hence they define the geometry of the space we work in, in the most obvious way. They represent the framework to which all our surveying is attached.

3.2) Remarks on Spheroid

By spheroid (in the non-English literature) we understand a simplified geoid (in English literature, spheroid coincides with rotational ellipsoid). Bruns' spheroid takes the potential of the gravitation force of the earth developed into spherical harmonics up to the 2nd degree plus the potential of the centrifugal force. At the same time, it takes the origin of the coordinate system to coincide with the center of the earth. In addition, it assumes that the axis of rotation coincides with the main axis of inertia (i.e., the products of inertia D,E,F = 0) and let its z-axis coincide with these two.

We can write for the potential of the earth:

$$U(\vec{r}) = \sum_{n=0}^{2} \left(\frac{a}{r}\right)^{n+1} Y_n + \frac{\omega^2}{2} r''^2$$

and substituting for the coefficients by the spherical harmonics from 2.23 we get under the above assumptions:

$$U_B(\vec{r}) = \frac{\kappa}{r} M + \frac{\kappa}{r^3} \left(\frac{A+B}{2} - C\right) C_{20} + \frac{\kappa}{4r^3} (B-A)\ C_{22} + \frac{\omega^2}{2} r''^2 = \text{const.}$$

Substituting for C_{20}, C_{22} the expressions involving x,y,z (2-23) we get

$$U_B(\vec{r}) = \frac{\kappa M}{r} + \frac{\kappa}{r^5} \left(\left(\frac{A+B}{2} - C\right) \left(z^2 - \frac{1}{2}x^2 - \frac{1}{2}y^2\right) + \frac{3}{4}(B-A)\ (x^2-y^2)\right) + \frac{\omega^2}{2}(x^2+y^2)$$

$$= \frac{\kappa M}{r} + \frac{\kappa}{2r^5} \left[(B+C-2A)\ x^2 + (A+C-2B)\ y^2 + (A+B-2C)\ z^2\right] + \frac{\omega^2}{2}(x^2+y^2) = \text{const.},$$

allegedly a surface of 14^{th} order.

Helmert's spheroid is based on the same assumptions although it uses spherical harmonics up to the 4^{th} degree. The result is a surface of 22nd degree. The expressions for the spheroid can be further simplified by assuming a rotational symmetry of the earth; i.e., A = B. If we do that we discover that the spheroid departs from an ellipsoid of rotation by only a very little. This is the reason why we are not much interested in the spheroid as an approximation to the geoid.

Note that the expression for spheroid contains the following unknown quantities: M, A, B, C, ω .

3.3) "Normal" and disturbing potential

One way how we try to bypass the difficulties mentioned at the end of 3.1 is to define "normal" potential and the corresponding "normal" gravity. The idea behind doing this is to split the actual potential U of the earth into two parts:

$$U = U_N + T$$

where U_N is a potential whose one of the equipotential surfaces coincides with the reference surface and T is the difference between the actual and assumed potential. The assumed potential U_N is usually called <u>normal potential</u>, T is then taken as <u>disturbing potential</u>. The reference surface is hence called <u>normal reference surface</u>.

The reference surface, towards which the normal potential can be related, is usually either a sphere or an ellipsoid of rotation.

Notice that if we succeeded to split the actual potential

$$U = V + \frac{1}{2}\omega^2 r''^2$$

in such a way as to make

$$U_N = V_N + \frac{1}{2}\omega^2 r''^2$$

we end up with

$$T = V - V_N$$

and the disturbing potential satifies the Laplace's equation outside the attracting body -- the earth. We are going to show that this can be done for both the sphere and the ellipsoid.

3.4) Sphere as a "normal" reference surface

We can write for the sphere:

$$U = V + \frac{1}{2}\omega^2 r''^2 = U_N^{(S)} + T^{(S)} = V_N^{(S)} + \frac{1}{2}\omega^2 r''^2 + T^{(S)}$$

where $U_N^{(S)}$ is constant on the sphere of radius a.

Hence

$$U_N^{(S)}\Big|_{r=a} = U_{NO}^{(S)} = \underbrace{\sum_{n=0}^{\infty} \left(\frac{a}{r}\right)^{n+1} Y_n^{(S)}}_{V_N^{(S)}}\Bigg|_{r=a} + \frac{1}{2}\omega^2 r''^2\Bigg|_{r=a} = \text{const.}$$

where $Y_n^{(S)} = A_{no}P_{no}$ (cos Θ) because it does not vary with λ (rotational symmetry of the sphere). The components in Θ are present because they have to compensate the assymmetry in Θ of the centrifugal term.

Here

$$r''\Big|_{r=a} = a \sin\Theta .$$

Hence $\frac{1}{2}\omega^2 r''^2 = \frac{1}{2}\omega^2 a^2 \sin^2\Theta$. In terms of Legendre's functions

$$\sin^2\Theta = 1 - \cos^2\Theta = \frac{2}{3}\left(\frac{3}{2} - \frac{3}{2}\cos^2\Theta\right) = \frac{2}{3}\left(1 - \frac{3}{2}\cos^2\Theta + \frac{1}{2}\right)$$

$$= \frac{2}{3}\left(P_{oo}(\cos\Theta) - P_{20}(\cos\Theta)\right)$$

and the potential of the centrifugal force is

$$\frac{1}{3}\omega^2 a^2 \left(P_{oo}(\cos\Theta) - P_{20}(\cos\Theta)\right) .$$

We can write for $U_{NO}^{(S)}$:

$$U_{NO}^{(S)} = A_{00}P_{00}+A_{10}P_{10}+A_{20}P_{20}+\frac{1}{3}\omega^2a^2(P_{00}-P_{20}) + \sum_{n=3}^{\infty} A_{n0}P_{n0} .$$

Therefore:

$$(A_{00}+\frac{1}{3}\omega^2a^2-U_{NO}^{(S)})\ P_{00}+A_{10}P_{10} + (A_{20}-\frac{1}{3}\omega^2a^2)\ P_{20} + \sum_{n=3}^{\infty} A_{n0}P_{n0} = 0 .$$

In order to be zero for all θ , all the coefficients by P_{i0} must be zero. Hence

$$(A_{00}+\frac{1}{3}\omega^2a^2 - U_{NO}^{(S)}) = 0, \quad (A_{20}-\frac{1}{3}\omega^2a^2) = 0,\ A_{10},\ A_{30},\ A_{40},\ \ldots = 0.$$

Substituting these results back into the original formula for $V_N^{(S)}$ we get:

$$V_N^{(S)} = \frac{a}{r}\ (U_{NO}^{(S)}-\frac{1}{3}\ \omega^2a^2)\ P_{00} + \frac{a^3}{r^3}\ \frac{a^2\omega^2}{3}\ P_{20} .$$

The value $U_{NO}^{(S)}$ (constant) should be selected in such a way as to correspond to the attractive force of the actual earth. For the geoid, we have (3.1)

$$U(\vec{r}) = \sum_{n=0}^{\infty} (\frac{a}{r})^{n+1} Y_n + \frac{1}{2}\ \omega^2 r''^2 = V(\vec{r}) + \frac{1}{2}\ \omega^2 r''^2$$

where $V(\vec{r})$ is the attractive potential of the earth. It can be written (3.2) as

$$V(\vec{r}) = \frac{\kappa}{r} M + O(r^{-3})$$

where $O(r^{-3})$ are the terms of lower than r^{-2} order. Hence we may write approximately, comparing the first terms of the two potentials:

$$\frac{a}{r}\ (U_{NO}^{(S)} - \frac{1}{3}\ \omega^2a^2) \simeq \frac{\kappa}{r} M$$

or

$$U_{NO}^{(S)} \simeq \frac{\kappa M}{a} + \frac{\omega^2a^2}{3} .$$

Substituting this result back into the equation for $U_N^{(S)}$ we end up with the expression for normal potential using a sphere as a reference:

$$U_N^{(S)} \simeq \frac{\kappa M}{r} + \frac{a^5}{r^3}\frac{\omega^2}{3} P_{20} + \frac{1}{3}\omega^2 a^2 (1 - P_{20})$$

$$= \frac{\kappa M}{r} + \frac{\omega^2 a^2}{3} + \frac{a^2\omega^2}{3}\left(\frac{a^3}{r^3} - 1\right) P_{20} \quad .$$

Note that the normal potential is known up to three unknown quantities: κM , a, ω.

Note that the equipotential surface of the normal gravity are not generally spherical. Only for $r=a$ do we get spherical equipotential surface -- the reference sphere.

3.5) Ellipsoid of rotation as a "normal" reference surface

Ellipsoid of rotation is the normal reference surface almost exclusively used in geodesy. This is because of its closeness to geoid or spheroid. The normal potential related to it can be developed similarly to the case of sphere. We again write

$$U = V + \frac{1}{2}\omega^2 r''^2 = U_N^{(E)} + T^{(E)} = V_N^{(E)} + \frac{1}{2}\omega^2 r''^2 + T^{(E)}$$

where we require $U_N^{(E)}$ to be constant on the reference ellipsoid (b, E) as yet not specified. Hence, using ellipsoidal harmonics:

$$U_N^{(E)}\Big|_{u=b} = U_{N0}^{(E)} = \sum_{n=0}^{\infty} q_{n0}(u,E,b)\Big|_{u=b} A_{n0} P_{n0}(\cos\theta) + \frac{1}{2}\omega^2 (u^2+E^2)\Big|_{u=b} \sin^2\theta = \text{const.}$$

because the terms contianing λ must all vanish, the (b,E) being rotational in the λ-plane.

Here evidently

$$q_{n0}(u,E,b)\Big|_{u=b} = Q_{n0}(i\frac{b}{E})/Q_{n0}(i\frac{b}{E}) = 1, \quad u^2+E^2\Big|_{u=b} = a^2 \quad \text{(see 2.15)}$$

so that we can write:

$$U_{NO}^{(E)} = \sum_{n=0}^{\infty} A_{n0}P_{n0}(\cos\Theta) + \frac{\omega^2 a^2}{2}\sin^2\Theta$$

$$= \sum_{n=0}^{\infty} A_{n0}P_{n0}(\cos\Theta) + \frac{\omega^2 a^2}{3}(P_{00}(\cos\Theta) - P_{20}(\cos\Theta)) \ .$$

Analogous to 3.4, it can be satisfied if and only if all the coefficients by P_{n0} are zero. Thus:

$$(A_{00} + \frac{\omega^2 a^2}{3} - U_{NO}^{(E)}) = A_{10} = (A_{20} - \frac{\omega^2 a^2}{3}) = A_{30} = \ldots = 0$$

and

$$A_{00} = U_{NO}^{(E)} - \frac{\omega^2 a^2}{3}, \quad A_{20} = \frac{\omega^2 a^2}{3} \ .$$

$V_N^{(E)}$ therefore becomes:

$$V_N^{(E)} = q_{00}(u,E,b)\ (U_{NO}^{(E)} - \frac{\omega^2 a^2}{3}) + q_{20}\ (u,E,b)\ \frac{\omega^2 a^2}{3} P_{20}\ (\cos\Theta) \ .$$

Since

$$q_{00}(u,E,b) = \text{arctg}(E/u)/\ \text{arctg}(E/b)$$

we can write, developing arctg E/u and again u into a series using r:

$$q_{00}(u,E,b) = E/(r\ \text{arctg}(E/b)) + o(r^{-3}) \ .$$

To establish the value of $U_{NO}^{(E)}$, we again compare the first terms of $V_N^{(E)}$ with V deduced for the geoid asking for such a value $U_{NO}^{(E)}$ that would correspond to the attractive force of the actual

3.6) "Normal" gravity related to ellipsoid reference surface

Taking the formula for the normal potential, we can compute the force -- gravity force -- corresponding to this potential on and above the reference surface (reference ellipsoid). We know that the force belonging to the potential can be obtained as the gradient of the potential. Hence.

$$\vec{\gamma} = \text{grad}\ (U_N^{(E)}) = \nabla U_N^{(E)} \ .$$

∇ in ellipsoidal coordinates is given by (see 2.15)

$$\nabla = \sum_{i=1}^{3} \frac{\vec{e}_i}{H_i} \frac{\partial}{\partial q_i}$$

where the Lame's coefficients equal:

$$H_u = \sqrt{\frac{u^2+E^2\cos^2\theta}{u^2+E^2}} \ , \quad H_\theta = \sqrt{(u^2+E^2\cos^2\theta)} \ , \quad H_\lambda = \sqrt{(u^2+E^2)}\ \sin\theta \ .$$

Since $U_N^{(E)}$ does not depend on λ (symmetrical around z axis) we have

$$\frac{\partial U_N^{(E)}}{\partial \lambda} = 0 \ .$$

Differentiating $U_N^{(E)}$ with respect to the other two coordinates we find:

i) $$\frac{\partial U_N^{(E)}}{\partial u} \simeq \frac{\kappa M}{E} \frac{\partial}{\partial u}\text{arctg}(E/u) + \frac{2}{3}\,\omega^2 u + \left(\frac{\partial}{\partial u}\, q_{20}\, \frac{\omega^2 a^2}{3} - \frac{2}{3}\,\omega^2 u\right) P_{20} .$$

Here $\frac{\partial}{\partial u}\ \text{arctg}(E/u) = \frac{1}{1+E^2/u^2}\ (-E/u^2) = -\frac{E}{u^2+E^2}$.

$\frac{\partial q_{20}}{\partial u} = \frac{\partial}{\partial u}\ (Q_{20}(i\ \frac{u}{E})/Q_{20}(i\ \frac{b}{E}))$ can be evaluated approximately by

developing the Legendre's functions of second kind into the power series in u/E or b/E respectively. Denoting

$$q_{20} = f(u)/f(b)$$

where $f(x) = (3(\frac{x}{E})^2 + 1)\ \mathrm{arctg}(E/x) - 3\frac{x}{E}$ we get

$$\mathrm{arctg}(E/x) = E/x - \frac{1}{3}(E/x)^3 + \frac{1}{5}(E/x)^5 - \ldots$$

and

$$f(x) = (3(E/x)^{-2} + 1)\ ((E/x) - \frac{1}{3}(E/x)^3 + \frac{1}{5}(E/x)^5 + \ldots) - 3(E/x)^{-1}$$

$$= 3(E/x)^{-1} - (E/x) + \frac{3}{5}(E/x)^3 - \frac{3}{7}(E/x)^5 + \ldots + (E/x) - \frac{1}{3}(E/x)^3 +$$

$$+ \frac{1}{5}(E/x)^5 - \ldots - 3(E/x)^{-1}$$

$$= (\frac{3}{5} - \frac{1}{3})\ (E/x)^3 + (\frac{1}{5} - \frac{3}{7})\ (E/x)^5 + \ldots$$

$$= \frac{4}{15}(E/x)^3 - \frac{8}{35}(E/x)^5 + \ldots$$

$$= \frac{4}{15}(E/x)^3\ (1 - \mathcal{O}((E/x)^2))\ .$$

Hence

$$q_{20} \simeq \frac{\frac{4}{15}(E/u)^3\ (1 - 0((E/u)^2))}{\frac{4}{15}(E/b)^3\ (1 - 0((E/b)^2))}$$

$$\simeq \frac{b^3}{u^3}\ (1 - 0((E/u)^2))\ (1 + 0((E/b)^2))$$

$$\simeq \frac{b^3}{u^3}\ (1 - 0((E/u)^4))$$

and finally

$$\frac{\partial q_{20}}{\partial u} \simeq -3\frac{b^3}{u^4}\ (1 - 0((E/u)^4)) \simeq -3\frac{b^3}{u^4}$$

Thus the partial derivative can be written as

$$\frac{\partial U_N^{(E)}}{\partial u} = - \frac{\kappa M}{u^2+E^2} - \frac{\omega^2 a^2 b^3}{u^4} P_{20} + \frac{2}{3} \omega^2 u \, (1 - P_{20}) \, .$$

ii) $$\frac{\partial U_N^{(E)}}{\partial \Theta} = (q_{20}(u,E,b) \, \frac{\omega^2 a^2}{3} - \frac{\omega^2 (u^2+E^2)}{3} \,) \, \frac{d}{d\Theta} P_{20} \, (\cos \Theta)$$

where $\frac{d}{d\Theta} P_{20} \, (\cos \Theta) = \frac{d}{d\Theta} \, (\frac{3}{2} \cos^2 \Theta - \frac{1}{2}) = - 3 \cos\Theta \sin\Theta = - P_{21} \, (\cos \Theta)$.

It is not difficult to see, that on the reference ellipsoid, the normal gravity $\vec{\gamma}$ must be perpendicular (normal) to the ellipsoid. Hence, denoting the normal gravity on the ellipsoid by $\vec{\gamma}_o$ we get:

$$\vec{\gamma}_o = \frac{1}{H_u} \frac{\partial U_N^{(E)}}{\partial u} \vec{e}_u$$

and after substituting for H_u and $\frac{\partial U_N^{(E)}}{\partial u}$ we get

$$\vec{\gamma}_o = - \sqrt{\frac{u^2+E^2}{u^2+E^2\cos^2\Theta}} \, \left(\frac{\kappa M}{u^2+E^2} + \frac{a^2\omega^2 b^3}{u^4} P_{20} - \frac{2}{3} \omega^2 u \, (1 - P_{20})\right)\Big|_{u=b} \, \vec{e}_u$$

or

$$\gamma_o = \sqrt{\frac{a^2}{b^2+E^2\cos^2\Theta}} \, \left(\frac{\kappa M}{a^2} - \frac{2}{3} \omega^2 b + (\frac{E^2\omega^2}{b} + \frac{2}{3} \omega^2 b) \, P_{20}\right)$$

Realizing that $E^2 = a^2 - b^2$ we get

$$b^2+E^2\cos^2\Theta = b^2 + a^2\cos^2\Theta - b^2\cos^2\Theta = a^2\cos^2\Theta + b^2\sin^2\Theta$$

and

$$\gamma_o = \frac{\kappa M}{a\sqrt{(a^2\cos^2\Theta + b^2\sin^2\Theta)}} \, \left(1 - \frac{2}{3} \frac{a^2 b\omega^2}{\kappa M} + (\frac{\omega^2 a^4}{b\kappa M} + \frac{2}{3} \frac{a^2 b\omega^2}{\kappa M}) \, P_{20} \right) \, .$$

The term $a^2b\omega^2/(\kappa M)$ is often denoted by m (being roughly equal to 0.33 10^{-2} for the earth). Using m, the normal gravity on the reference ellipsoid can be expressed as follows:

$$\boxed{\gamma_o \simeq \frac{\kappa M}{a\sqrt{(a^2\cos^2\theta + b^2\sin^2\theta)}} \left(1 - \frac{2}{3}m + \left(\left(\frac{a}{b}\right)^2 m + \frac{2}{3}m\right) P_{20}(\cos\theta)\right) .}$$

3.7) Clairaut's theorem for gravity and geometrical flattennings

Using the normal gravity a very important theorem of physical geodesy can be developed that links the gravity with the geometry of the reference ellipsoid. We can write for the normal gravity on the equator γ_a :

$$\theta = 90°,\ P_{20}(\cos\theta)\Big|_{\theta=90°} = \frac{3}{2}\cos^2\theta\Big|_{\theta=90°} - \frac{1}{2} = -\frac{1}{2} \quad \text{and}$$

$$\gamma_a \simeq \frac{\kappa M}{ab}\left(1 - \frac{2}{3}m - \frac{1}{2}\left(\left(\frac{a}{b}\right)^2 m + \frac{2}{3}m\right)\right) = \frac{\kappa M}{ab}\left(1 - m - \left(\frac{a}{b}\right)^2 \frac{m}{2}\right) .$$

Similarly, for the normal gravity on the poles γ_b:

$$\theta = 0,\ 180°,\ P_{20}(\cos\theta)\Big|_{\theta=0,180°} = 1 \qquad \text{and}$$

$$\gamma_b \simeq \frac{\kappa M}{a^2}\left(1 - \frac{2}{3}m + \left(\frac{a}{b}\right)^2 m + \frac{2}{3}m\right) = \frac{\kappa M}{a^2}\left(1 + \left(\frac{a}{b}\right)^2 m\right) .$$

Hence

$$\frac{\gamma b}{\gamma a} \simeq \frac{b}{a}\left(1 + \left(\frac{a}{b}\right)^2 m\right)\left(1 + m + \left(\frac{a}{b}\right)^2 \frac{m}{2} + \dots\right)$$

$$= \frac{b}{a}\left(1 + m + \frac{3}{2}\left(\frac{a}{b}\right)^2 m + \dots\right) .$$

Realizing that $\gamma_b/\gamma_a - 1 = (\gamma_b - \gamma_a)/\gamma_a$ and $b/a - 1 = (b-a)/a$ we can write

$$\frac{\gamma_b - \gamma_a}{\gamma_a} - \frac{b-a}{a} \simeq \frac{b}{a}\left(m + \frac{3}{2}\left(\frac{a}{b}\right)^2 m + \ldots\right) = \frac{bm}{a}\left(1 + \frac{3}{2}\left(\frac{a}{b}\right)^2 + \ldots\right) .$$

But
$$\frac{bm}{a} = \frac{a\omega^2 b^2}{\kappa M} \simeq \frac{b\omega^2}{\gamma_a}\left(1 - m - \left(\frac{a}{b}\right)^2 \frac{m}{2}\right) .$$

Thus

$$\frac{\gamma_b - \gamma_a}{\gamma_a} + \frac{a-b}{a} \simeq \frac{b\omega^2}{\gamma_a}\left(1 - m - \left(\frac{a}{b}\right)^2 \frac{m}{2}\right)\left(1 + \frac{3}{2}\left(\frac{a}{b}\right)^2 + \ldots\right)$$

$$\simeq \frac{b\omega^2}{\gamma_a}\left(1 + \frac{3}{2}\left(\frac{a}{b}\right)^2 + \ldots\right) .$$

Here, the term $(\gamma_b - \gamma_a)/\gamma_a = f^*$ ($= \alpha$ in older literature) is known as gravity flattenning and $(a-b)/a = f$ ($= i$ in older literature) is the known flattenning (geometrical) of the reference ellipsoid. Hence the formula can be written as

$$\boxed{f^* + f \simeq \frac{b\omega^2}{\gamma_a}\left(1 + \frac{3}{2}\left(\frac{a}{b}\right)^2 + \ldots\right)}$$

which is known as the Clairaut's theorem. It was derived first by a French mathematician Clairaut (1738) in the form:

$$f^* + f \simeq \frac{5}{2}\frac{b\omega^2}{\gamma_a}$$

which is obviously further simplification of the above for $a \simeq b$.

3.8) Somigliana's formulae for normal gravity

Handier formulae for normal gravity are due to the Italian geodesist Somigliana (1929). He has developed the formula for γ_o (from 3.6) along the following lines: let us rewrite $P_{20}(\cos\Theta) = \frac{3}{2}\cos^2\Theta - \frac{1}{2}$, using the identity

$$\frac{1}{2} = \frac{1}{2}(\cos^2\Theta + \sin^2\Theta)$$

as

$$P_{20}(\cos\Theta) = \cos^2\Theta - \frac{1}{2}\sin^2\Theta .$$

Substituting this into the formula for γ_o, we get

$$\gamma_o \simeq \frac{\kappa M}{a\sqrt{(a^2\cos^2\Theta + b^2\sin^2\Theta)}} \left(1 - \frac{2}{3}m + \left(\left(\frac{a}{b}\right)^2 m + \frac{2}{3}m\right)\cos^2\Theta - \right.$$

$$\left. - \frac{1}{2}\left(\left(\frac{a}{b}\right)^2 m + \frac{2}{3}m\right)\sin^2\Theta\right) .$$

Using another identity

$$1 - \frac{2}{3}m = \left(1 - \frac{2}{3}m\right)(\cos^2\Theta + \sin^2\Theta)$$

we can write

$$\gamma_o \simeq \frac{\kappa M}{a\sqrt{(a^2\cos^2\Theta + b^2\sin^2\Theta)}} \left[\left(1 + \left(\frac{a}{b}\right)^2 m\right)\cos^2\Theta + \left(1 - m - \left(\frac{a}{b}\right)^2 \frac{m}{2}\right)\sin^2\Theta\right] .$$

The expressions in the round brackets can be substituted for using the formulae for γ_a, γ_b (3.7):

$$\left(1 + \left(\frac{a}{b}\right)^2 m\right) \simeq \frac{a^2\gamma_b}{\kappa M} , \quad \left(1 - m - \left(\frac{a}{b}\right)^2 \frac{m}{2}\right) \simeq \frac{ab\gamma_a}{\kappa M}$$

so that we get

$$\boxed{\gamma_o \simeq \frac{a\gamma_b \cos^2\Theta + b\gamma_a \sin^2\Theta}{\sqrt{(a^2\cos^2\Theta + b^2\sin^2\Theta)}}} .$$

It is more common in geodesy to work with geodetic latitude ϕ rather than with Θ . As we know from geometric geodesy, ϕ and Θ are related via geocentric latitutde β by following formulae:

$$\beta = 90° - \Theta, \ \mathrm{tg}\beta = \frac{b}{a} \mathrm{tg}\ \phi .$$

Hence $\quad \mathrm{cotg}\Theta = \frac{b}{a} \mathrm{tg}\phi$

and we can write

$$\cos\Theta = cb \sin\phi \quad , \sin\Theta = ca \cos\phi$$

where c is an arbitrary constant. Substituting these expressions into the formula for γ_o we get:

$$\gamma_o \simeq \frac{c^2b^2a\gamma_b \sin^2\phi + c^2a^2b\gamma_a \cos^2\phi}{\sqrt{(a^2b^2c^2 \sin^2\phi + a^2b^2c^2 \cos^2\phi)}} = c\ (b\gamma_b \sin^2\phi + a\gamma_a \cos^2\phi) .$$

On the other hand, we must have $\sin^2\Theta + \cos^2\Theta = 1$. Hence

$$c^2(a^2\cos^2\phi + b^2\sin^2\phi) = 1$$

which yields the value for c. Substituting this into our equation for γ_o we finally obtain:

$$\boxed{\gamma_o \simeq \frac{a\gamma_a \cos^2\phi + b\gamma_b \sin^2\phi}{\sqrt{(a^2\cos^2\phi + b^2\sin^2\phi)}}} .$$

Notice the symmetry of the two Somigliana's formulae for normal gravity.

3.9) Cassinis' formula for normal gravity, international formulae

In geodetic practice it is usual to use yet another formula due also to Somigliana although commonly known as Cassinis' formula. This is because Cassinis was the first to have presented his estimates of the values of the coefficients to the IUGG congress in 1930. The theoretical development goes as follows: expressing $\cos^2\phi$ in the latter formula (in 3.8) as $1 - \sin^2\phi$, we get:

$$\gamma_o \simeq \frac{a\gamma_a + (b\gamma_b - a\gamma_a)\sin^2\phi}{\sqrt{(a^2 + (b^2 - a^2)\sin^2\phi)}} = \gamma_a \frac{1 + (\frac{\gamma_b}{\gamma_a}\frac{b}{a} - 1)\sin^2\phi}{\sqrt{(1 + (\frac{b^2}{a^2} - 1)\sin^2\phi)}} .$$

According to our notation from 3.7

$$\frac{\gamma_b - \gamma_a}{\gamma_a} = \frac{\gamma_b}{\gamma_a} - 1 = f^* , \quad \frac{a-b}{a} = 1 - \frac{b}{a} = f .$$

Hence
$$\frac{\gamma_b}{\gamma_a} = 1 + f^*, \quad \frac{b}{a} = 1 - f$$

and

$$\gamma_o \simeq \gamma_a \frac{1 + [(1+f^*)(1-f) - 1]\sin^2\phi}{\sqrt{(1 + ((1-f)^2 - 1)\sin^2\phi)}} = \gamma_a \frac{1 + (f^* - f - ff^*)\sin^2\phi}{\sqrt{(1 + (f^2 - 2f)\sin^2\phi)}} .$$

Since the flattennings are much smaller than 1 we can develop the denominator into power series:

$$(1 + (f^2 - 2f)\sin^2\phi)^{-1/2} = 1 - {}^1/_2 (f^2 - 2f)\sin^2\phi - \dots .$$

We thus get

$$\gamma_o \approx \gamma_a \, (1+(f^*-f-ff^*) \sin^2\phi) \, (1-\frac{1}{2} (f^2-2f) \sin^2\phi - \ldots)$$

$$= \gamma_a \, [1+(f^*-f-ff^*-\frac{1}{2} f^2+f) \sin^2\phi - \frac{1}{2} (f^2-2f)(f^*-f-ff^*) \sin^4\phi + \ldots] \, .$$

Substituting $\sin^2\phi - \frac{1}{4} \sin^2 2\phi$ for $\sin^4\phi$ we get

$$\gamma_o \approx \gamma_a \, [1 + (f^*-ff^*-\frac{1}{2}f^2-\frac{1}{2} (f^2-2f) \, (f^*-f-ff^*)) \sin^2\phi \; + \frac{1}{8} (f^2-2f)(f^*-f-ff^*) \sin^2 2\phi + \ldots]$$

$$\gamma_c \approx \gamma_a \, (1 + \tilde{\alpha} \sin^2\phi + \tilde{\beta} \sin^2 2\phi)$$

where $\tilde{\alpha} \approx f^* + O(f^{*2})$, $\tilde{\beta} = \frac{f}{4} (f-f^*) + O(f^3)$.

Cassinis' original formula, adopted in 1930 reads

$$\gamma_o \approx 978.0490 \, (1 + 0.0052884 \sin^2\phi - 0.000\,0059 \sin^2 2\phi \,) \text{ gal.}$$

In 1967, the IUGG adopted new values for the coefficients so that now, we have

$$\gamma_o \approx 978.031 \, (1 + 0.0053024 \sin^2\phi - 0.000\,0059 \sin^2 2\phi \,) \text{ gal.}$$

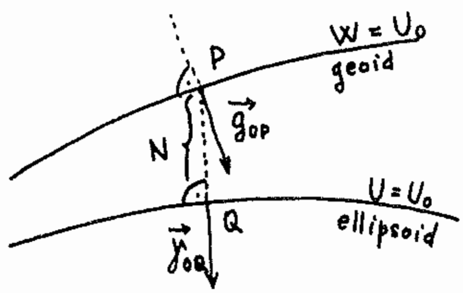
3.10) Definitions of gravity anomaly, gravity distrubance, geoidal height, deflection of vertical

Let us denote by U the normal gravity potential related to a rotational ellipsoid as a reference surface (as yet unknown!) and the actual potential of the earth by W. We have called the difference

$$T = W - U$$

disturbing potential (3.3), it is also known as anomalous potential.

Let us assume, for the sake of the forthcoming definitions, that we know the reference ellipsoid already and let us denote the normal potential (whose one of the equipotential surfaces coincides with the reference ellipsoid!) on the reference ellipsoid by U_o. We can then draw the following cross-section:



We can see, that $U = U_o$ is the reference ellipsoid, $W = U_o$ is the geoid, the surface we would like to determine, $\vec{g}_{0P}$ is the actual gravity on the geoid, $\vec{\gamma}_{0Q}$ is the normal gravity on the reference surface. The distance $N = \overline{PQ}$ is known as <u>geoidal height</u> (geoidal undulation) at point Q. The vector

$$\Delta\vec{g}_{0P} = \vec{g}_{0P} - \vec{\gamma}_{0Q}$$

is called <u>anomaly vector</u> and its absolute value Δg_{0P} is known as <u>gravity anomaly</u> on the reference surface.

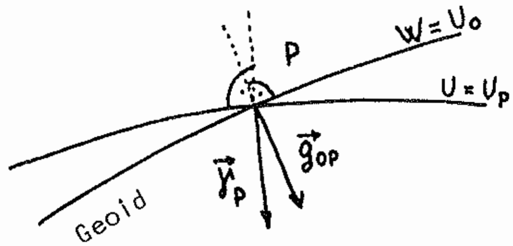
The geodial heights probably do not exceed $\pm$ 100 m anywhere in the world. The angle $\Theta = \sphericalangle\, \vec{g}_{0P}\, \vec{\gamma}_{0Q}$ (do not mix up with the second spherical coordinate), known as the <u>deflection of the vertical</u>, very seldom exceeds 1' and is usually smaller than 5''. ($\Theta \approx 30''$ are considered already large.). Because of this small amplitude of Θ we generally compute the gravity anomaly from

$$\Delta g_{0P} \quad g_{0P} - \gamma_{0Q}$$

instead of

$$\Delta g_{0P} = \left|\vec{g}_{0P} - \vec{\gamma}_{0Q}\right| = g_{0P} - \gamma_{0Q} \cos\Theta .$$

Taking $W=U_o$ and $U=U_P$ in such a way as to let the two surfaces coincide in P on the geoid we get the <u>gravity disturbance vector</u>:



$\vec{\delta g}_{OP} = \vec{g}_{OP} - \vec{\gamma}_P$, <u>gravity disturbance</u> $\delta g_{OP} \simeq g_{OP} - \gamma_P$ and the angle $\sphericalangle\, \vec{g}_{OP}\vec{\gamma}_P$ is for all practical purposes identical with the deflection of vertical. They differ only by the term arising from the curvature of the normal field plumbline.

Realizing that $\vec{g} = \nabla W$ and $\vec{\gamma} = \nabla U$ we get

$\vec{\delta g} = \vec{g} - \vec{\gamma} = \nabla W - \nabla U = \nabla(W-U) = \nabla T.$

Hence the gravity disturbance vector at a point P on the geoid is given by the gradient of the disturbing potential at the point. We can also write

$$g = -\frac{\partial W}{\partial n} \ , \ \gamma = -\frac{\partial U}{\partial n'}$$

where n, n' are local outer normals to the geoid and the ellipsoid respectively. Since the angle between the two normals (deflection of vertical) is small we get:

$$\delta g = |\vec{g}-\vec{\gamma}| \simeq g-\gamma = -\frac{\partial W}{\partial n} + \frac{\partial U}{\partial n'} \simeq -\frac{\partial W}{\partial n} + \frac{\partial U}{\partial n} = -\frac{\partial T}{\partial n} \ .$$

Thus the gravity disturbance is given as negative derivative of the disturbing potential taken with respect to the local outer vertical (or outer ellipsoidal normal for that matter).

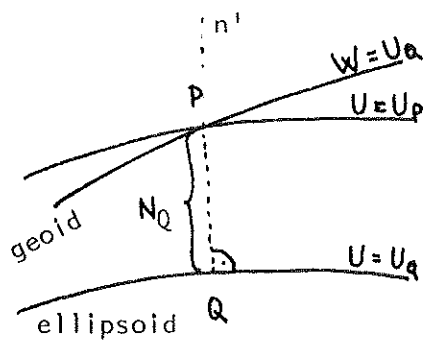
The gravity anomaly is mostly used in the classical terrestrial geodesy, gravity disturbance is widely used in modern theories and satelite geodesy.

Since the gravity disturbance is related to the geoid, i.e., to the point P, rather than the ellipsoid, we shall relate everything systematically to the geoid. Hence even the disturbing potential and geoidal height will be thought of as being related to the geoid.

Note that we are still moving on a superficial level knowing neither the geoid nor the reference ellipsoid. Hence we cannot measure any of the involved quantities.

3.11) Relation between disturbing potential and geoidal heights, 2nd Bruns' formula

Let us take now both cross-sections from 3.10 together:



and let $W(\vec{r}_p)$ be called W_p and similarly U_P, U_Q. We can write

$$U_P - U_Q = \frac{\partial U}{\partial n'}\Big|_Q N_Q = -\gamma_Q N_Q = -\gamma_Q N_P, \quad (N_Q = N_P) \; . \qquad (*)$$

By definition $W_P = U_P + T_P$. Hence

$$W_P = U_Q - \gamma_Q N_P + T_P \; .$$

But W_P equals also to U_o (see 3.5) or U_Q in our notation. Thus we get finally

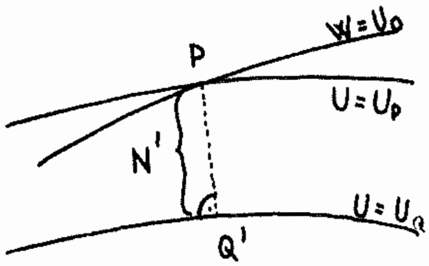
$$T_P = \gamma_Q N_P$$

or

$$N_P = T_P/\gamma_Q$$

which is one of the most important formulae of physical geodesy, due to a German geodesist Bruns (1878). It is known as <u>2nd Bruns' formula</u> and relates the disturbing potential T to the geoidal undulation N.

When we assume a reference ellipsoid, without knowing the proper values of the constants involved (κM, a, E,ω) which is always the case, it is likely to have a wrong value of potential (normal). Let us denote the assumed value of normal potential by U'_Q and the difference $U'_Q - U_o$ by δU



Now $W_p = U_o$ and $\delta U = U'_Q - W_P = U'_Q - (U_P + T_P)$.

But $U_P - U'_Q = -\gamma_{Q'} N'_P$

and we get:

$$\delta U = \gamma_{Q'} N'_P - T_P.$$

Finally, we obtain:

$$\boxed{N'_P = \frac{T_P + \delta U}{\gamma_{Q'}}} \doteq N_P + \frac{U}{\gamma_Q} .$$

This formula is known as generalized Bruns' formula and it relates the disturbing potential (as computed from an <u>assumed</u> normal gravity field) to the <u>geoidal height above the assumed ellipsoid.</u> Note that δU is a function of δM, δa, δE, U, Θ, where by δM, δa, δE, we denote the differences between the correct values M,a,E and the assumed values M',a',E'. In practice δU is assumed constant and interpreted as imprecision in our knowledge of W_p, the value of potential corresponding to the geoid.

3.12) Fundamental Gravimetric equation

Let us assume again that we know the correct size and shape of the reference ellipsoid (a,E) and the other two constants ($\kappa M,\omega$) necessary to determine the normal gravity. Let us differentiate the equation (*) from 3.11 with respect to n'.

We get

$$\frac{\partial U}{\partial n'}\Big|_P - \frac{\partial U}{\partial n'}\Big|_Q = - \frac{\partial \gamma}{\partial n'}\Big|_Q N_P .$$

here $\frac{\partial U}{\partial n'}\Big|_P = - \gamma_P$, $\frac{\partial U}{\partial n'}\Big|_Q = - \gamma_Q$ and we have

$$\gamma_P - \gamma_Q = \frac{\partial \gamma}{\partial n'}\Big|_Q N_P .$$

We have shown in 3.10 that

$$g_P - \gamma_P \simeq - \frac{\partial T}{\partial n'}\Big|_P .$$

Combining these two equations we obtain

$$g_P - \gamma_Q \doteq \frac{\partial \gamma}{\partial n'}\Big|_Q N_P - \frac{\partial T}{\partial n'}\Big|_P .$$

But, according to 3.10, $g_P - \gamma_Q \simeq \Delta g_P$. Hence

$$\Delta g_P \simeq \frac{\partial \gamma}{\partial n'}\Big|_Q N_P - \frac{\partial T}{\partial n'}\Big|_P .$$

(Notice that this equation relates the gravity anomaly to the gravity disturbance, since $- \frac{\partial T}{\partial n'}\Big|_P = \delta g_P$:

$$\Delta g_P \simeq \delta g_P + \frac{\partial \gamma}{\partial n'}\Big|_Q N_P .)$$

Using the 2nd Bruns' formula to substitute for N_P we get

$$\Delta g_P \simeq -\frac{\partial T}{\partial n'}\Big|_P + \frac{1}{\gamma_Q}\frac{\partial \gamma}{\partial n'}\Big|_Q T_P .$$

This fundamental gravimetric equation is usually written as

$$\Delta g \simeq -\frac{\partial T}{\partial h} + \frac{1}{\gamma_Q}\frac{\partial \gamma}{\partial h} T$$

where n' (outer normal to ellipsoid) is replaced by height and all its terms are related to the point Q, i.e. to the reference ellipsoid. We may do that because Δg as well as T are really related to both P and Q and we have earlier decided to denote them by subscript P merely for convenience.

Assuming again an arbitrary (though close to the geoid) reference ellipsoid we end up with the generalized equation

$$\Delta g' \simeq -\frac{\partial (T + \delta U)}{\partial h} + \frac{1}{\gamma'}\frac{\partial \gamma'}{\partial h}(T + \delta U)$$

where all the variables are related to the arbitrary reference ellipsoid.

3.13) Discussion of the fundamental gravimetric equation, mixed boundary-value problem of geodesy

It is not difficult to see that the fundamental gravimetric equation provides us with the boundary values of the mixed type to solve the Laplace's equation

$$\Delta T = 0$$

for the outside of the reference ellipsoid, providing the reference ellipsoid and the values $\kappa M, \omega$ (to compute the normal gravity) are selected properly. There are three difficulties involved in solving the third boundary-value problem using the gravity anomalies (the fundamental gravity equation):

i) We never know, and never will know, the true values of a, E and κM, ω. Hence the unknown term δU will be always present and we cannot apply the earlier developed method (see 2.21) as it is.

ii) The geoid is neither known nor accessible on the continents. Hence the value g_{0P} necessary for determining Δg (see 3.10) is not observable. Therefore even Δg cannot be obtained without introducing further considerations. On the other hand the observations of g_{0P} on the ocean surface are still thin and a matter of concern.

iii) Even the basic requirement for $\Delta T = 0$, i.e. the density $\sigma = 0$ everywhere outside the ellipsoid, is usually not satisfied. The assumed reference ellipsoid is usually approximating the geoid in the mean sense so that it is almost always underneath the terrain on the continent and even underneath the sea level at various places.

The last two difficulties are usually dealt with by means of altering the gravity anomalies in such a way as to neutralize them. We first reduce the gravity observations made on the surface of the earth to the geoid and then account for the masses above the ellipsoid computationally. These gravity reductions, however, do not constitute the topic of the present course and will be dealt with elsewhere. We shall be assuming that the gravity anomalies, used for the mixed boundary-value problem are already corrected in the proper manner.

Before trying to solve the mixed boundary-value problem, let us have a look at the term $\frac{\partial \gamma'}{\partial h}$ and try to find an expression for it. It could be done directly from the normal potential (3.5) if we expressed γ as its gradient and differentiated it with respect to the ellipsoidal normal, but this would be a very tedious job. We shall show a shorter approach using differential geometry.

3.14) Vertical gradient of gravity

Let us take the potential W of the actual earth. The equipotential surface $W(x,y,z) = W_P$ can be regarded as an implicit function of x,y,z. Let us consider a local orthogonal system of coordinates such that $\tilde{z}$-axis coincides with the outer normal to $W = W_P$ (vertical), $\tilde{x}$-axis is in the tangent plane to $W = W_P$ and points to north, $\tilde{y}$-axis points to west.

Total derivative of W with respect to $\tilde{x}$ is

$$\frac{dW}{d\tilde{x}} = \frac{\partial W}{\partial \tilde{x}} + \frac{\partial W}{\partial \tilde{z}} \quad \frac{d\tilde{z}}{d\tilde{x}} \quad ,$$

second total derivative is

$$\frac{d^2W}{d\tilde{x}^2} = \frac{\partial^2 W}{\partial \tilde{x}^2} + \frac{\partial^2 W}{\partial \tilde{x}\partial \tilde{z}} \frac{d\tilde{z}}{d\tilde{x}} + \frac{\partial^2 W}{\partial \tilde{z}\partial \tilde{x}} \quad \frac{d\tilde{z}}{d\tilde{x}} + \frac{\partial^2 W}{\partial \tilde{z}^2} \left(\frac{d\tilde{z}}{d\tilde{x}}\right)^2 + \frac{\partial W}{\partial \tilde{z}} \quad \frac{d^2\tilde{z}}{d\tilde{x}^2}$$

Let us denote $\frac{\partial W}{\partial \tilde{x}} = W'_{\tilde{x}}$ and similarly the other partial derivatives.

Since W = const., we have $\frac{dW}{d\tilde{x}} = \frac{d^2W}{d\tilde{x}^2} = 0$. In addition, $\frac{dz}{d\tilde{x}} = 0$ since z is perpendicular to W = const. The second total derivative then becomes

$$W''_{\tilde{x}\tilde{x}} + W'_{\tilde{z}} \frac{d^2\tilde{z}}{d\tilde{x}^2} = 0 \quad .$$

Similarly, the second total derivative with respect to $\tilde{y}$ yields:

$$W''_{\tilde{y}\tilde{y}} + W'_{\tilde{z}} \frac{d^2\tilde{z}}{d\tilde{y}^2} = 0 \quad .$$

Here $W'_{\tilde{z}} = \frac{\partial W}{\partial \tilde{z}} = \frac{\partial W}{\partial n} = -g$, the absolute value of gravity at the point in question.

From differential geometry we know that the curvature k of the curve $y = y(x)$ is given by

$$k = y'' \, (1 + y'^2)^{-3/2} \quad .$$

In our case then, $\frac{d^2\tilde{z}}{d\tilde{x}^2}$ can be considered the curvature of $z = z(x)$, the N-S profile of the equipotential surface, $\frac{d^2\tilde{z}}{d\tilde{y}^2}$ the curvature in the E-W profile. This is because the first derivatives in both directions are zero.

Denoting by J the negative value of the overall curvature of the equipotential surface, given as the arithmetic mean of the two curvatures of the two perpendicular profiles, we get

$$J = -\frac{1}{2}\,(k_{\tilde{x}} + k_{\tilde{y}}) = -\frac{1}{2g}\,(W''_{\tilde{x}\tilde{x}} + W''_{\tilde{y}\tilde{y}}) \quad .$$

On the other hand, since W is the potential of gravity, it has to satisfy the Poisson's equation

$$\Delta W = -4\pi\kappa\sigma + 2\omega^2 \quad ,$$

and the Laplace's operator in the local coordinates system is given by

$$\Delta W = W''_{\tilde{x}\tilde{x}} + W''_{\tilde{y}\tilde{y}} + W''_{\tilde{z}\tilde{z}} \quad .$$

Combining the last three equations together and realizing that $W''_{\tilde{z}\tilde{z}} = \frac{\partial}{\partial z}\; W'_{\tilde{z}} = -\frac{\partial}{\partial \tilde{z}}\, g = -\frac{\partial g}{\partial h}$ we get finally

$$\frac{\partial g}{\partial h} = -2gJ + 4\pi\kappa\sigma - 2\omega^2 \quad .$$

This is known as the 1-st Bruns' formula and it relates the vertical gradient of gravity to the other parameters determining the potential field. Note that the quantities g, J, σ are related to the point we examine the gradient in. The 1-st Bruns' formula is one of the very few rigorous formulae in physical geodesy.

It is not difficult to see, that for the <u>normal gravity above the ellipsoid</u> we get

$$\frac{\partial \gamma}{\partial h} = -2\gamma J - 2\omega^2 .$$

Here J is as yet unknown. However, dealing with the reference ellipsoid we are able to express J as a function of the meridian radius of curvature M and the radius of curvature of the prime vertical cross-section N (do not mix up these M,N with the mass of the earth and the geoidal undulation!).

We have

$$J = \frac{1}{2}(1/M + 1/N)$$

(note the sign ; the curvature in mathematics is taken positive for a convex surface, in geodesy positive for a concave surface viewed along the outer normal). Here M and N can be expressed, as we know from geometric geodesy,

$$1/M = \frac{b}{a^2}(1 + e'^2 \cos^2\phi)^{3/2} , \quad 1/N = \frac{b}{a^2}(1 + e'^2 \cos^2\phi)^{1/2}$$

where $e' = E/b = \sqrt{(a^2-b^2)}/b$. Hence

$$e'^2 = \frac{a^2-b^2}{b^2} = \frac{a^2}{b^2} - 1 .$$

Using $f = 1 - b/a$ we can write

$$e'^2 = 1/(1-f)^2 - 1 = \frac{1-(1-f)^2}{(1-f)^2} = (2f-f^2)\ (1+2f+\ldots)$$

$$= 2f - f^2 + 4f^2 - 2f^3 + \ldots = 2f + 3f^2 + \ldots \simeq 2f \ .$$

Hence

$$1/M \simeq \frac{b}{a^2}\ (1+2f\ \cos^2\phi)^{3/2} = \frac{b}{a^2}\ (1 + 3f\ \cos^2\phi + \ldots)$$

$$1/N \simeq \frac{b}{a^2}\ (1+2f\ \cos^2\phi)^{1/2} = \frac{b}{a^2}\ (1 + f\ \cos^2\phi + \ldots)$$

and

$$J \simeq \frac{b}{2a^2}\ (2+4f\ \cos^2\phi) = \frac{b}{a^2}\ (1+2f\ \cos^2\phi\ ) \ .$$

We then can write for the vertical gradient of normal gravity.

$$\frac{\partial\gamma}{\partial h} \simeq -\ \frac{2\gamma b}{a^2}\ (1 + 2f\ \cos^2\phi) - 2\omega^2 \ .$$

Here $2\omega^2$ is smaller than the first term and can be considered as a corrective term. It may therefore be approximated by

$$2\omega^2 \simeq \frac{2\gamma m}{a}$$

using the formulae for m from 3.6 and

$$\gamma \simeq \gamma_a \simeq \frac{\kappa M}{ab}$$

from 3.7. Substituting this result back to our original equation, we can write

$$\frac{\partial\gamma}{\partial h} \simeq -\ \frac{2\gamma}{a}\ (\frac{b}{a} + 2\frac{b}{a}\ f\ \cos^2\phi + m) \ .$$

Here $\frac{b}{a} = 1-f$, $\frac{b}{a}$ in the $\cos^2\phi$ term may be equated to 1 (since it is much smaller than the first term) and

$$2f \cos^2\phi = f(1+\cos 2\phi) .$$

Hence we finally end up with

$$\frac{\partial\gamma}{\partial h} \simeq - \frac{2\gamma}{a} (1 + m + f \cos 2\phi)$$

neglecting thus all the higher order terms in m,f.

3.15) Solution to the mixed boundary-value problem of physical geodesy

Substituting the result of 3.14 into the fundamental gravimetric equation we obtain:

$$\Delta g \simeq - \frac{\partial}{\partial h} (T+\delta U) - \frac{2}{a} (1+m+f\cos 2\phi) (T+\delta U)$$

where we understand, from now on, that the equation is valid on an arbitrary reference ellipsoid which is sufficiently close to the geoid. Hence all the involved quantities are computed on the basis of this assumed ellipsoid. This is our boundary value for

$$\Delta(T + \delta U) = 0$$

on the reference ellipsoid.

It can be shown that with an error of the order of 3.10^{-3} the coefficient by $(T+\delta U)$ is on the ellipsoid constant and equal to $- 2/R$ where $R = \sqrt[3]{(a^2b)}$. The solution to the third boundary value problem can be then written approximately (see 2.21)

$$\tilde{T} = T+\delta U \simeq \sum_{n=0}^{\infty} \sum_{m-0}^{n} \frac{q_{nm}(u,E,b)}{-\frac{2}{R} + \frac{(n+1)}{R}} \Delta g_{nm} = \sum_{n} \sum_{m} \frac{R}{n-1} q_{nm} \Delta g_{nm}$$

where Δg_{nm} are the spherical harmonics of Δg. On the reference ellipsoid we get

$$\tilde{T} \simeq \sum_{n=0}^{\infty} \frac{R}{n-1} \Delta g_n \ . \qquad (+)$$

Note that the expression is not defined for n=1. We have to assume that the first-degree harmonic is missing altogether. This is the condition that has to be satisfied for this particular linear combination of boundary-values. It corresponds to the condition

$\oiint_{El} \frac{\partial f}{\partial n} \; dEl = 0$ for the Neuman's problem. Our condition here is

$\oiint_{El} \Delta g \cos\psi \; dEl = 0$. Expressing

$$\tilde{T} = W - (U - \delta U) = W - \tilde{U}$$

we can develop all three potentials into spherical harmonics and find that the coefficients by the first-degree harmonic for $\tilde{T}$ depend on the displacement of the center of the reference ellipsoid from the center of gravity of the earth. Assuming that they coincide, we can bring the first-degree harmonic to zero and write

$$\tilde{T} = \tilde{T}_0 + \sum_{n=2}^{\infty} \frac{R}{n-1} \Delta g_n \ .$$

whereby $\tilde{T}_0$ we denote the zero-degree harmonic of $\tilde{T}$. This is the solution on the surface of the reference ellipsoid.

Since $\tilde{T}_0 \simeq T_0 + \delta U$ and $T_0 \simeq W_0 - U_0 = \kappa\frac{\delta M}{R}$ we can write the above equation as

$$T \simeq \kappa \frac{\delta M}{R} + \sum_{n=2}^{\infty} \frac{R}{n-1} \Delta g_n \, .$$

Here, in dealing with Δg, we assume that it has been corrected for

i) the influence of the masses above the ellipsoid and;

ii) the reduction from the terrain to the geoid of the observed gravity.

3.16) Stokes' integral

Let us develop the gravity anomaly on the reference ellipsoid into spherical (ellipsoidal) harmonics. As we know, we get for Δg as for any arbitary function:

$$\Delta g = \sum_{n=0}^{\infty} \Delta g_n$$

where $$\Delta g_n = \sum_{m=0}^{n} [(A_{nm} \cos m\lambda + B_{nm} \sin m\lambda) \, P_{nm}(\cos \Theta)]$$

with $$A_{nm} = \frac{(n-m)!}{(n+m)!} \, \frac{2n+1}{2\pi} \oiint_{E1} \Delta g(\Theta',\lambda') \cos m\lambda' \, P_{nm} \, (\cos\Theta') \, dE1 \, ,$$

$$B_{nm} = \frac{(n-m)!}{(n+m)!} \, \frac{2n+1}{2\pi} \oiint_{E1} \Delta g(\Theta',\lambda') \sin m\lambda' \, P_{nm} \, (\cos\Theta') \, dE1 \, .$$

(for $m=0$, there will be 4π instead of 2π). By dashes, we denote the dummy variables in the integration. Substituting the expressions for the coefficients back into the expression for Δg_n we get:

$$\Delta g_n = \frac{2n+1}{4\pi} P_{n0} \, (\cos\Theta) \oiint_{E1} \Delta g \, P_{n0} \, (\cos\Theta') \, dE1 + \sum_{m=1}^{n} \{ \frac{(n-m)!}{(n+m)!} \, \frac{2n+1}{2\pi} \, P_{nm} \, (\cos\Theta)$$

$$[\cos m\lambda \oiint_{E1} \Delta g \cos m\lambda' \, P_{nm} \, (\cos\Theta') \, dE1 + \sin m\lambda \oiint_{E1} \Delta g \sin m\lambda' \, P_{nm}(\cos\Theta') \, dE1] \} .$$

Taking the integration sign outside the summation we obtain:

$$\Delta g_n = \frac{2n+1}{4\pi} \oiint_{El} \{\Delta g(\Theta',\lambda') \; [P_{n0}(\cos\Theta) \; P_{n0}(\cos\Theta') + 2 \sum_{m=1}^{n} (\frac{(n-m)!}{(n+m)!}$$

$$P_{nm}(\cos\Theta) \; P_{nm}(\cos\Theta') \; (\cos m\lambda \; \cos m\lambda' + \sin m\lambda \; \sin m\lambda'))]\} \; dEl \; .$$

We can see that the expression in the square brackets equals to $P_n(\cos\psi)$ (see 2.22) with ψ denoting the spherical distance between the point (Θ,λ) for which g_n is computed and the "dummy point" (Θ',λ') involved in the integration. We can thus write

$$\Delta g_n = \frac{2n+1}{4\pi} \oiint_{El} \Delta g \; (\Theta',\lambda') \; P_n(\cos\psi) \; dEl \; .$$

This result can be substituted into the equation (+) in 3.15 and we get

$$\tilde{T} \simeq \sum_{n=0}^{\infty} \; [\frac{R}{n-1} \; \frac{2n+1}{4\pi} \oiint_{El} \Delta g \; P_n(\cos \; \psi) \; dEl]$$

$$= - \frac{R}{4\pi} \oiint_{El} \Delta g \; P_0(\cos\psi) \; dEl + \frac{R}{4\pi} \oiint_{El} \Delta g \sum_{n=2}^{\infty} \frac{2n+1}{n-1} P_n(\cos\psi) \; dEl \; .$$

where the series is known as Stokes' function - $S(\psi)$ - and the first term is nothing but $-R\Delta g_0$. The second term is usually called Stokes' integral and it represents the closed solution to the mixed boundary-value problem on the ellipsoid for $\Delta g_0=0$. It corresponds, to Poisson's integral for the first boundary-value problem on the sphere.

We can see, that in the development we have left out the first-degree harmonic $\tilde{T}_1 = \frac{R}{1-1} \Delta g_1$. This can be done because it was shown that for the center of the ellipsoid coinciding with the center of gravity of the earth T_1 goes to zero. The coincidence of the two centers is to be assumed.

3.17) Stokes' formula, gravimetric determination of the geoid

From the point of view of determination of the geoid the knowledge of $\tilde{T}$ is just an intermittent step; geoid being defined as the equipotential surface $W(r,\theta,\lambda) = W_0 = U_0$. It is easily seen however, that using the generalized Bruns' formula (see 3.11) the geoidal undulations N can be computed from $\tilde{T}$. We can write

$$\tilde{T} = T+\delta U = \gamma N$$

or

$$N \simeq \frac{R}{4\pi\gamma} \oiint_{E1} \Delta g \; S(\psi) \; dE1 - R \; \Delta g_0/\gamma \; .$$

Here we can take a mean gravity G instead of γ with little influence on the already limited accuracy. It remains to be seen whether Δg_0 can be expressed in terms of other parameters.

From the formula for $\tilde{T}$ (3.15) we can see that

$$\tilde{T}_0 = - R \; \Delta g_0$$

where by 0 we denote again the first spherical harmonic in the appropriate development. Recalling the formulae for $\tilde{T}_0$ and T_0 in 3.15, we get

$$- R \; \Delta g_0 \simeq \kappa \frac{\delta M}{R} + \delta U .$$

Writing the formula for N as

$$N + \delta N = \frac{R}{4\pi G} \oiint_{E1} \Delta g \; S(\psi) \; dE1 \; .$$

we get

$$\delta N = - \kappa \frac{\delta M}{RG} - \frac{\delta U}{G} \; .$$

The formula for N is known as Stokes' formula and it supplies us with the undulations of the geoid above the assumed reference surface-ellipsoid.

The correction δN to the computed undulations can be added if we know δM and δU, the errors in the mass and in the potential of the assumed ellipsoid.

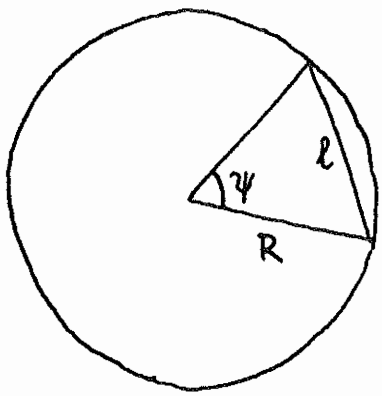
The formula (without the term δN) is due to Sir George Gabriel Stokes (1819-1903) mathematician and physicist in Cambridge. It was first published in 1849 and is, perhaps, the most important formula of physical geodesy. It permits the determination of the geoid (as related to the assumed reference ellipsoid) from the gravimetric data.

The geoid as computed from this formula is, due to our presumptions, always concentric with the reference surface. δN is usually taken as constant and interpreted as correction to one of the ellipsoid's axis. The reason for this is that the parameters E, ω of the reference ellipsoid are known much more accurately than $a, \kappa M$ and δU can be regarded as mostly due to δM and δa. The effect of these two uncertainities is indistinguishable. It is though easier to account for the effect by changing the size of the reference ellipsoid rather than its mass. The reference ellipsoid hence follows more closely the geoid which is a desirable property.

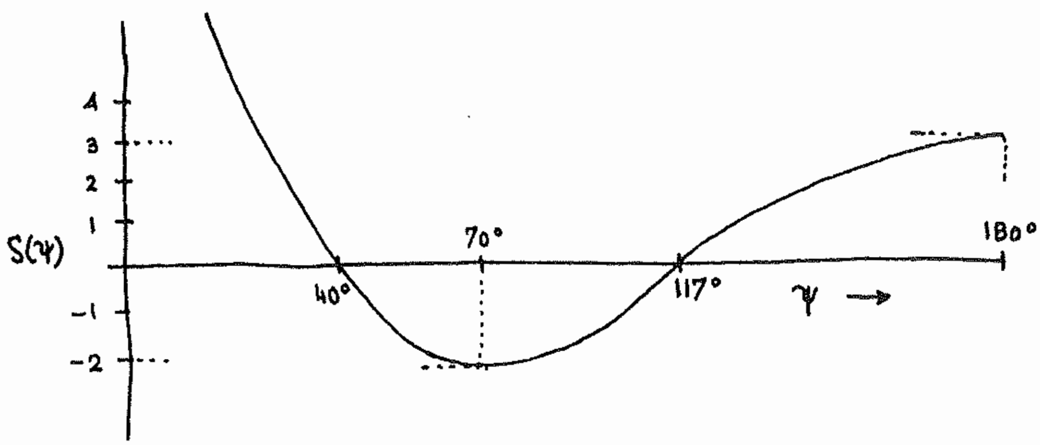
3.18) A few remarks about the Stokes' formula

The Stokes' Function can be expressed without using the infinite series as

$$S(\psi) = 2\,\frac{R}{\ell} + 1 - 3\,\frac{\ell}{R} - \cos\psi\,\left(5 + 3\,\ln\left(\frac{1}{2} - \frac{1}{2}\cos\psi + \frac{\ell}{2R}\right)\right).$$



We can see that for $\ell \to 0$, $\psi \to 0$ and $S(\psi) \to \infty$. Its graph looks thus:



It can be clearly seen how the gravity anomalies over the whole earth contribute to each particular separation N at any place. The closer we go to the point of interest the more the anomaly Δg contributes towards the separation. Therefore, when using the Stokes' formula, we have to know Δg well particularly in the vicinity of the point we are interested in.

The Stokes' formula can be rewritten in various different ways. We are going to mention two of them here. First, we can chose the point of interest as the origin for polar coordinates on the ellipsoid and get

$$N + \delta N = \frac{R}{4\pi G} \int_{\alpha=0}^{2\pi} \int_{\psi=0}^{\pi} \Delta g(\alpha,\psi)\ S(\psi)\ \sin\psi\ d\psi\ d\alpha\ .$$

$$= \frac{R}{2G} \int_{\psi=0}^{\pi} \left[\frac{1}{2\pi} \int_{\alpha=0}^{2\pi} \Delta g(\alpha,\psi)\ d\alpha \right] S(\psi)\ \sin\psi\ d\psi$$

$$= \frac{R}{G} \int_{\psi=0}^{\pi} \bar{\Delta g}(\psi)\ F(\psi)\ d\psi$$

where $F(\psi) = \frac{1}{2} S\ (\psi)\ \sin\psi$ and $\bar{\Delta g}(\psi)$ is the mean anomaly at the angular distance ψ from the point of interest.

Alternatively, $\oiint_{E1}$ can be expressed in terms of geographical coordinates and we get

$$N(\phi,\lambda) + \delta N = \frac{R}{4\pi G} \int_{\lambda'=0}^{2\pi} \int_{\phi'=-\pi/2}^{\pi/2} \Delta g(\phi',\lambda')\ S(\psi)\ \cos\phi'\ d\phi'\ d\lambda'$$

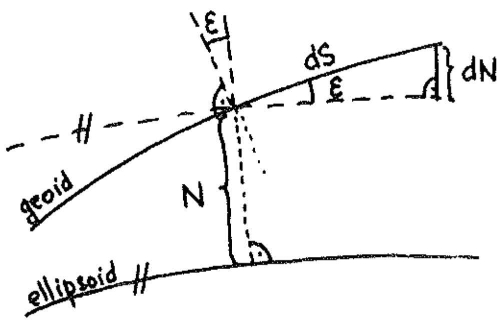
where $\quad \psi = \arccos\ (\sin\phi\ \sin\phi' + \cos\phi\ \cos\phi'\ \cos\ (\lambda'-\lambda))$.

In practice the numerical methods to evaluate these integrals are used almost exclusively.

3.19) Vening-Meinesz' formulae

Another application of the closed solution to the mixed boundary-value problem, i.e., the Stokes' integral, are the formulae permitting the computation of the N-S and W-E components of local deflection of vertical from the gravity anomalies known all over the earth. They can be derived as follows: it can be seen from the cross-section that $dN = \varepsilon dS$. If the cross-section is taken in the plane defined by the two normals, then $\varepsilon = \Theta$. If the cross-section lies in the meridian (prime vertical) plane, ε represents the N-S (W-E) component of Θ called ξ (η).

Signs?



For these two components we get

$$\xi = -\frac{dN}{dS_\phi}\ ,\quad \eta = -\frac{dN}{dS_\lambda}$$

where the negative signs express the convention that for dN positive ξ (η) is taken as decreasing with increasing ϕ (λ). This means that if the geoid increases towards North, ξ decreases and similarly if the geoid increases towards West, η decreases. Some European countries use the reverse convention for η. Denoting by $d\phi$, $d\lambda$ the differential increments of the coordinates in meridian and prime vertical, and using the Lame's coefficients for $r = R$ (on the ellipsoid):

$$dS_\phi = H_\phi d\phi = R d\phi$$

$$dS_\lambda = H_\lambda d\lambda = R \cos\phi \, d\lambda$$

and

$$\xi = -\frac{1}{R}\frac{\partial N}{\partial\phi}\ ,\quad \eta = -\frac{1}{R\cos\phi}\frac{\partial N}{\partial\lambda}\ .$$

Taking N from the Stokes' formula and regarding δN as constant we get (only S is a function of ϕ and λ !):

$$\xi = -\frac{1}{4\pi G}\oiint_{El} \Delta g \frac{\partial S(\psi)}{\partial\phi}\, dEl\ ,$$

$$\eta = -\frac{1}{4\pi G\cos\phi}\oiint_{El} \Delta g \frac{\partial S(\psi)}{\partial\lambda}\, dEl\ .$$

Here we express the partial derivatives as

$$\frac{\partial S(\psi)}{\partial\phi} = \frac{\partial S(\psi)}{\partial\psi}\frac{\partial\psi}{\partial\phi}\ ,\quad \frac{\partial S(\psi)}{\partial\lambda} = \frac{\partial S(\psi)}{\partial\psi}\frac{\partial\psi}{\partial\lambda}$$

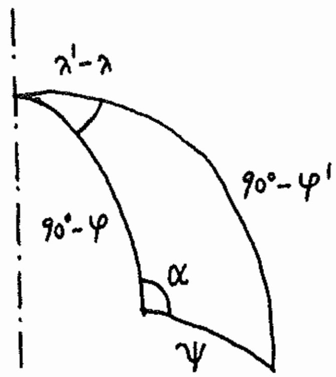
and derive $\frac{\partial\psi}{\partial\phi}$, $\frac{\partial\psi}{\partial\lambda}$ from the formula for cos ψ used in 2.22:

$\cos\psi = \sin\phi\sin\phi' + \cos\phi\cos\phi'\cos(\lambda'-\lambda)$

we get

$$- \sin\psi \frac{\partial\psi}{\partial\phi} = \cos\phi \sin\phi' - \sin\phi \cos\phi' \cos(\lambda'-\lambda)$$

$$- \sin\psi \frac{\partial\psi}{\partial\lambda} = \cos\phi \cos\phi' \sin(\lambda'-\lambda).$$



On the other hand, from the spherical triangle we obtain

$$\sin\psi\cos\alpha=\cos\phi\sin\phi'-\sin\phi\cos\phi'\cos(\lambda'-\lambda)$$

$$\sin\psi \sin\alpha = \cos\phi' \sin(\lambda'-\lambda) .$$

Comparing these two sets of equations we get

$$\frac{\partial\psi}{\partial\phi} = - \cos\alpha \ , \ \frac{\partial\psi}{\partial\lambda} = - \cos\phi \sin\alpha$$

Substituting the results back into the formulae for ξ and η, we finally end up with:

$$\xi = \frac{1}{4\pi G} \oiint_{El} \Delta g \frac{\partial S(\psi)}{\partial\psi} \cos\alpha \, dEl$$

$$\eta = \frac{1}{4\pi G} \oiint_{El} \Delta g \frac{\partial S(\psi)}{\partial\psi} \sin\alpha \, dEl ,$$

the Vening-Meinez' formulae. In these formulae, ψ ,α can be taken again for "polar" coordinates on the ellipsoid or they can be transferred to any other pair of coordinates on the ellipsoid.

3.20) Outline of numerical solution of Stokes' and Vening-Meinesz' formulae

The gravity anomalies Δg in Stokes' and Vening Meinesz' formulae are not available for every point on the reference ellipsoid. They can be computed for a number of discrete points where the values of gravity have been observed on the earth surface. Hence, we cannot integrate them (more precisely integrate the product of the gravity anomalies with the weight function $S(\psi)$ or $\frac{\partial S(\psi)}{\partial\psi}$ respectively) over the ellipsoid and we have to use one of the numerous numerical methods to evaluate the double integrals. All the numerical methods, whichever we use, replace the double integration by double summation over two parameters.

The grid for the summation can be basically of two kinds, either polar or rectangular. The polar grid corresponds to variables α,ψ as described in 3.18. Hence the grid has to be shifted every time so as to make it centered upon the point of interest. Because the weight of Δg varies with ψ (and α as well, in case of Vening Meinesz' formulae) the grid can be designed in such a way as to have larger areas corresponding to smaller weight and vice versa. Evidently, if Δg is weighted "lightly", it can represent larger area without contributing to the result too much and vice versa.

The rectangular grid is generally based on geographical coordinates as mentioned in 3.18. This approach is somewhat preferable to use when we study the whole globe, as opposed to individual points, because the representative values of Δg in individual blocks ($\Delta\phi \times \Delta\lambda$) can be attached to the appropriate blocks once and for all. The grid does not change from one point of interest to another.

The most serious difficulty encountered in the numerical solution is the increasing influence of the gravity anomalies as one approaches to point of interest. A brief look on the weight functions convinces us that the immediate surroundings of the point have considerable effect on the result. This problem can be overcome by two means. First, it is usual to make the grid finer in the immediate environment of the point. Second, various formulae have been designed, to express the influence of the gravity anomalies in the close vicinity via other characteristics of the gravity field, that do not deal with the Stokes' function. Using these formulae, we divide the double integral in two or more parts that reflect the contribution of close and more distant zones. These parts are then evaluated separately.

4) Fundamentals of gravimetry

4.1) Gravity observations

The gravity observations are meant to provide us with the values of gravity acceleration at the desired points. The gravity acceleration, usually denoted by g, is measured in gals (1 gal = 1 cm sec^{-2}) or its decadic fractions (mgal = 10^{-3} gal, μgal = 10^{-6} gal). Obviously the approximate value of g anywhere on the surface of the earth is 980 gals.

From the point of view of the position of the observation points we can divide the gravity observations to:

i) terrain observations;

ii) underwater observations (either observed from the submarine or on the sea bottom);

iii) sea-surface observations (from the ship);

iv) airborne observations (from the aircraft).

From the point of view of the observation technique used we can speak about

i) absolute observations;

ii) relative observations.

The former is based on the idea of observing directly the value of g at a point. The latter observes just the difference in gravity for two stations. For geodetic purposes we would like to know the gravity with an error in absolute value smaller than 0.1 mgal = 10^{-4} gal; i.e., $\simeq 10^{-7}$ times the observed phenomenon g. This is a very high accuracy and is easier achievable using the relative rather than absolute observations.

As we have said earlier , it has not been established yet whether the values of g are subject to any secular changes. So far we regard them as permanent. On the other hand, we know that the value of gravity changes with the position of the Sun and the Moon. This phenomenon, known as gravimetric tide, can account for as much as - 0.16 to +0.08 mgals. Since the tidal variations of gravity are known and predictable they can be corrected for.

The relative measurements are done in such a way that we read the gravimeter reading at one point where gravity is known already. Then take another reading at the unknown point and another one back at the known point. Hence we have two differences in readings that multiplied by a known constant give two differences in gravity. Their discrepancy is attributed to the drift of the instrument and divided linearly with time on both differences.

4.2) Instruments used in gravimetry

There are basically three distinctly different types of devices used for gravity observations (measurement of g)

i) vertical pendulums

ii) gravimeters

iii) free-fall devices.

The first and the third types can be used for absolute measurements, the second cannot.

Pendulums can be either ordinary, reversible, inverted, very long or multiple. Their use is based on the idea that there is a known relationship between the period of swing and the value of g, namely

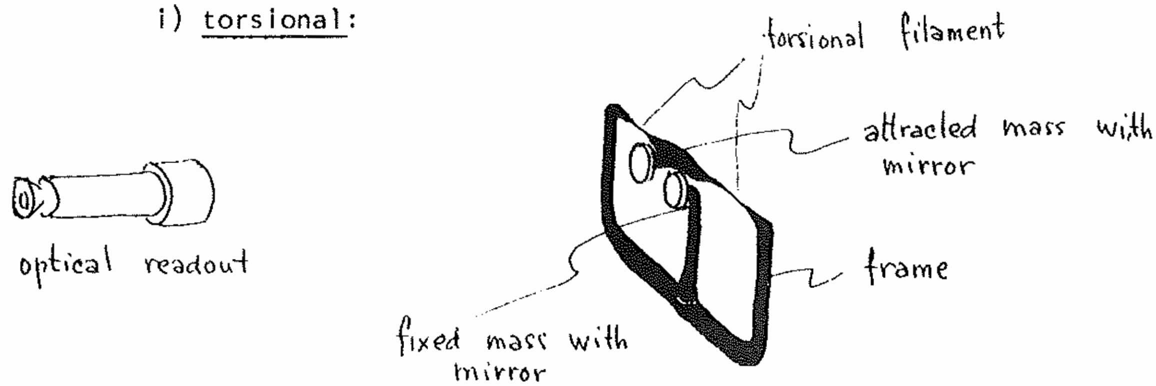
$$T = Cg^{-1/2}$$

where C is the constant related to the mass and the length of the pendulum. The above relation originates in the equation of motion of the pendulum. It is not difficult to see that by observing the period of swing we can deduce the value of g, providing C is known.

The precision attainable with pendulums is of the order of ± 3 to ± 0.1 mgals (for very long pendulums). Vening-Meinesz's submarine pendulum apparatus (three coupled pendulums with photographic registration) had, at best, precision of ± 4 mgals. The major obstacles in achieving any better precision are numerous influences like friction (air and edges), temperature, instability of the fixed construction. Also errors in timing contribute significantly to the relatively low precision.

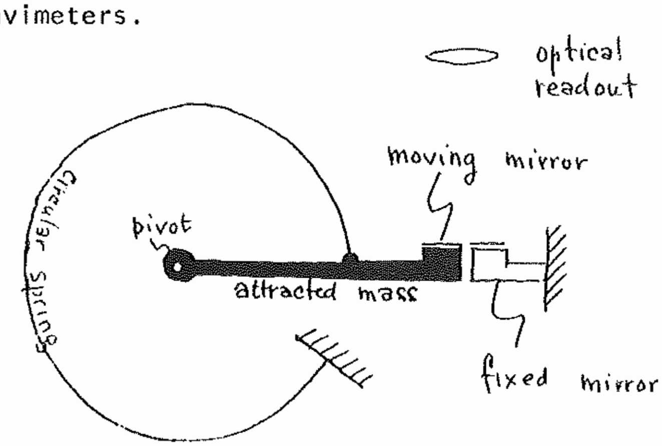
Gravimeters are the widest used devices in gravimetry. All the designs are invariably based on measuring the relative position of a fixed and free masses. There are three distinctly different basic designs in modern gravimeters :

i) torsional:



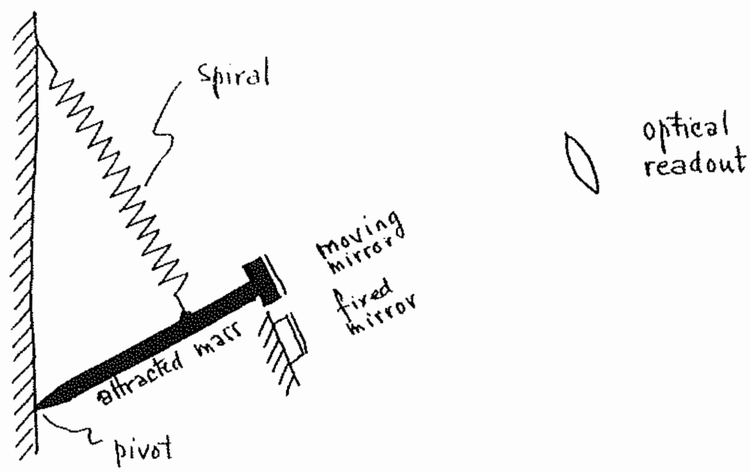
The whole construction is made of one piece of fused quartz. This design was first utlized in a Danish gravimeter designed by Norgaard (precision ± 0.2 mga
Basically the same design, although equipped with various compensational devices and different readout systems, is used by American Worden (precision ± 0.03 mgal) and German Graf gravimeters.

ii) Circular spring



This is the principle used by Molodenskij in his GKA gravimeter (precision $\pm$ 0.3 mgal) and the whole family of subsequent Russian gravimeters. They vary by different compensational and readout system.

iii) Spiral spring



This system with various compensations and readouts is used by North American and Lacoste-Romberg (precision $\pm$ 0.01 mgal) gravimeters and seems to be the most successful one.

Besides these three, there is a number of other designs like gas pressure, vibration or bifilar gravimeters. But they never have gained any wide recognition.

The gravimeters are also used almost exclusively as shipborne or airborne instruments mounted either on gimbals or gyro-stabilized platforms. Their precision is still comparatively low ($\pm$ 0.5 and $\pm$ 10 mgal respectively) mostly because of inadequate accounting for the accelerations of the vessels involved. The same holds true for the sea-bottom gravimeters, where the major flaw is our inaptitude to attach precise coordinates to the observation point.

Generally, the gravimeters have proved more successful than any other gravimetric instrument mainly because of their versatility. They are easy to operate and their theory is well known. The only information one can get from a gravimeter is the reading on a scale. This has to be compared to the proper value of gravity. This comparison is known as calibration. Three major hindrances can be listed against gravimeters in general:

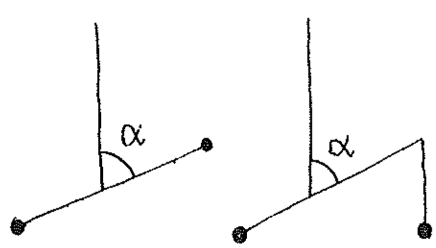
i) the inability to measure the absolute gravity;

ii) the necessity of frequent calibration;

iii) the inevitable presence of drift due to aging or various components and other causes.

Free-fall devices are the newest development in gravimetry. They are based on the timing of a fall of a free body in vacuum. Since the acceleration of a free fall is g, the magnitude of the gravity can be deduced from the free-fall time. The device is still under development and the precision so far within the region of $\pm$ 1 mgal.

One more instrument should be mentioned here even though it does not measure gravity directly. It is the Eötvös's torsion balance (variometer) designed to measure the horizontal gradients of gravity. There are two different types of variometers:



The inclination α changes with the azimuth of the balance arm giving the values of gradients of g in various directions.

4.3) Gravimetric networks

From the point of view of the use of the observed gravity we can distinguish two different gravity surveys:

i) geodetic;

ii) geophysical.

The second, used in geophysical prospecting (location of various mineral deposits) has got a local character. The users are not interested in absolute but relative gravity. It has little interest for a geodesist.

The first, because of the necessity to supply the absolute values of gravity has to be organized on the international level. Hence the national gravimetric works are all connected to one international reference point -- Potsdam. The last adjustment of international gravity networks was carried out in 1971.

The national networks are divided into 3 orders--

i) First order consists of the national reference station and all the "absolute points". They are usually located at the airports so that the access to them is easy. Canadian national reference station is the pier in the basement of the former Dominion Observatory in Ottawa. It was established by relative methods.

ii) Second order consists of points established some 10-20 miles apart within an easy reach by car (along highways, etc.).

iii) Third order has points closer together although their accuracy is lower. This is being currently built in Canada.

Besides the national networks there are some international "calibration lines" stretching across wide areas so as to cover the widest possible range of gravity values. Their points are usually observed very precisely to obtain very precise values of gravity. They are used to calibrate individual gravimeters -- i.e., to derive the one to one correspondence of scale readings with gravity values. The North American calibration line runs from Alaska to South Mexico.

4.4) Processing of the observed gravity

Since there are various users of the gravity information there are also various aspects of the processing of the gravity data. In this and the forthcoming paragraphs of this section we are going to show how the observed gravity can be processed so as to supply us with the gravity anomalies on the reference ellipsoid needed to evaluate the Stokes formula (see Section 3).

We may recall (section 3.15) that in order to get the Δg (gravity anomaly) on the reference ellipsoid we have to

i) know the "actual gravity" g_o on the geoid;

ii) supress the effect of the masses above the reference ellipsoid on g_o. Since all the formulae used for the determination of geoid are only approximate, we can afford to introduce one more highly convenient approximation of the same order. We shall not require that the masses above the ellipsoid are accounted for but replace it by the requirement that all the masses above the geoid are removed. This amounts to the same as if we had formulated the boundary-value problem on the geoid instead of ellipsoid.

However, we do not know the distribution of masses; i.e. the distribution of density σ, within the earth. Neither do we know the variations of density within the surface layers, by which we understand the parts protruding above the geoid. The task to evaluate the Δg on the geoid is hence a formidable one and represents one of the major hindrances of classical geodesy.

The difficulty of the task is reflected in the wide spectrum of ideas, techniques and attitudes displayed by various scholars in geodesy. We are going to show here only the generally accepted and most widely used approaches to the problem.

4.5) Free-air correction and anomaly

The free-air correction originates from the somewhat simplified following imagination: suppose we had taken the observation of g (on the topographical surface of the earth) in absence of all the redundant masses above the geoid. The observation station is thus imagined to be hanging 'free in the air'. Then the only correction (reduction) necessary to obtain the gravity value on the geoid is

$$\delta g_F = \frac{-\partial g}{\partial h} h$$

where h is the height of our station above the geoid. (δg here has nothing to do with the gravity disturbance defined in 3.10).
Considering all the mass of the geoid concentrated in its center of gravity and denoting by r the distance of the geoidal surface from the center of gravity we can write approximately for g_o (on the geoid):

$$g_o \simeq \kappa \frac{M}{r^2} \quad .$$

Differentiating this with respect to r we get

$$\frac{\partial g_o}{\partial r} \simeq - 2\kappa \frac{M}{r^3}$$

that approximates the term $\frac{\partial g}{\partial h}$. In this formula, r can be further approximated by either the length of the radius vector of the reference ellipsoid:

$$\frac{\partial g_o}{\partial h} \simeq \frac{-2\kappa M}{a^3} \left(1 + \frac{3}{2} f \cos 2\phi + \ldots\right)$$

or by the radius $R = \sqrt[3]{(a^2 b)}$ of the reference sphere:

$$\frac{\partial g}{\partial h} \simeq \frac{-2\kappa M}{R^3} .$$

In the majority of cases hence, the simplified formula for the free-air correction

$$\boxed{\delta g_F \simeq \frac{2\kappa M}{R^3} h = 0.3086\ h}$$

(with δg_F in mgal for h in m,) can be used.

The corresponding gravity anomaly

$$\boxed{\Delta g_F = g + \delta g_F - \gamma}$$

is known as the free-air anomaly. Even though the free-air anomaly is based on seemingly quite erroneous assumptions it is very widely used because of some properties we are going to discuss later.

4.6) Bouguer correction and anomaly

Seemingly, the free-air treatment of gravity does not depict the reality well enough. Obviously, when observing the g on the surface of the earth its value is influenced by the masses in between the topographic surface and the geoid as well as by the masses enclosed within the geoid. This influence of the masses above the geoid should be corrected for as well. It is usually done in two steps:

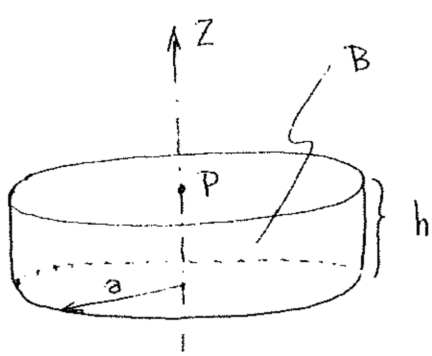
i) removal of the influence of the plate (layer) of uniform thickness h-meters high;

ii) removal of the influence of the irregularities of the topography; i.e., the influence of the masses enclosed between the topographic surface and the flat surface of the plate.

In this section, we are going to deal with the plate -- known as Bouguer's plate according to the French geodesist Bouguer who first used this correction in his gravity survey in Peru in 1749. The second step, considered as a refinement of the first is known as terrain correction and will be dealt with in the next section.

The plate, or the layer, covers the whole of the geoid and has therefore quite a complicated spatial shape. Fortunately, however, it is quite sufficient in the first approximation, to consider the plate as a plane layer. Then a correction for the curvature of the earth can be added to it, as we shall see later.

To derive the correction for the plate, consider a circular cylinder of height h, radius a (do not mix up with the semimajor axis of the reference ellipsoid!) and density σ. What will be its attraction at the point P? We know from the theory of potential that an attracting body B



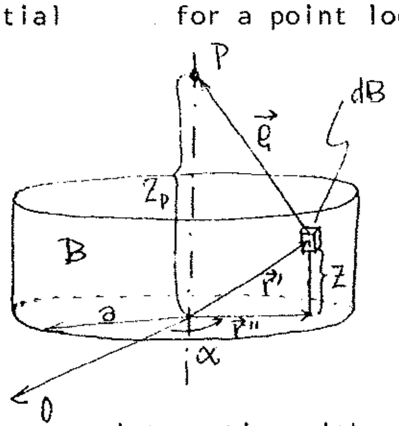
acts on a unit mass with the force (see Physical Geodesy I, § 2.4):

$$\vec{f} = -\kappa \int_B \frac{\sigma}{\rho^3}\vec{\rho}\, dB.$$

The potential of this force is given by (see Physical Geodesy I, § 2.8):

$$V = \kappa \int_B \frac{\sigma}{\rho}\, dB .$$

The easiest way, how to derive the attractive force of our cylinder is to express it in cylindrical coordinates and compute first its potential for a point located on the Z axis:



We get

$$V(P) = \kappa\sigma \int_{\alpha=0}^{2\pi} \int_{Z=0}^{h} \int_{r''=0}^{a} \frac{1}{\rho}\, dB$$

where $dB = r''\, dr''\, dZ\, d\alpha$ and

$$\rho = ((Z_P - Z)^2 + r''^2)^{1/2} .$$

Integration with respect to α yields:

$$V(P) = 2\pi\kappa\sigma \int_{Z=0}^{h} \int_{r''=0}^{a} \frac{r''\,dr''\,dZ}{((Z_p-Z)^2+r''^2)^{1/2}}.$$

The inner integral can be solved by substitution $t^2 = (Z_p-Z)^2+r''^2$.

This yields:

$$\int_{r''=0}^{a} \frac{r''\,dr''\,dZ}{((Z_p-Z)^2+r''^2)^{1/2}} = \sqrt{((Z_p-Z)^2 + a^2)} - Z_p + Z.$$

Hence

$$V(P) = 2\pi\kappa\sigma \int_{Z=0}^{h} [\sqrt{((Z_p-Z)^2 + a^2)} - Z_p + Z]\,dZ$$

$$= 2\pi\kappa\sigma\, [-Z_p h + \frac{1}{2} h^2 + \int_{o}^{h} \sqrt{((Z_p-Z)^2 + a^2)}\,dZ]\ .$$

The integral here can again be solved by another substitution $t = Z_p - Z$ which gives

$$I = \int_{o}^{h} (a^2 + (Z_p-Z)^2)^{1/2} dZ = - \int_{Z_p}^{Z_p-h} (a^2 + t^2)^{1/2}\,d\,t.$$

Here $\int (a^2+t^2)^{1/2}\,dt = \frac{1}{2}\,[t(a^2+t^2)^{1/2} + a^2 \ln\,(t + (a^2 + t^2)^{1/2})]$+const. Hence

$$I = -\frac{1}{2}\,[(Z_p-h)(a^2+(Z_p-h)^2)^{1/2} + a^2 \ln\,((Z_p-h) + (a^2+ (Z_p-h)^2)^{1/2})] +$$

$$+ \frac{1}{2}\,[Z_p \quad (a^2 + Z_p^2)^{1/2} + a^2 \ln\,(Z_p + (a^2 + Z_p^2)^{1/2}]\ .$$

Denoting $(a^2 + (Z_p-h)^2)^{1/2}$ by d and $(a^2 + Z_p^2)^{1/2}$ by b we can write finally

$$\boxed{V(P) = 2\pi\kappa\sigma\,[-Z_p h + \frac{1}{2}h^2 - \frac{1}{2}\,((Z_p-h)d - Z_p b + a^2 \ln \frac{Z_p-h+d}{Z_p+b})].}$$

Since we are interested in getting the vertical gradient of this potential; i.e., the component in Z direction which for all practical purposes coincides with the direction of g, we have to take the first derivative of V (in P) with respect to Z_p. We obtain:

$$\frac{\partial V}{\partial Z_P} = f_P = 2\pi\kappa\sigma \left[-h - \frac{1}{2}\left(d + (Z_P-h)\frac{\partial d}{\partial Z_P} - b - Z_P \frac{\partial b}{\partial Z_P} + a^2 \frac{Z_P+b}{Z_P-h+d} \times \right.\right.$$

$$\left.\left.\frac{\left(1+\frac{\partial d}{\partial Z_P}\right)(Z_P+b) - \left(1+\frac{\partial b}{\partial Z_P}\right)(Z_P-h+d)}{(Z_P+b)^2}\right]\right|_P$$

where

$$\frac{\partial d}{\partial Z_P} = \frac{1}{2} d^{-1}\, 2\,(Z_P-h)\Big|_P = \frac{(Z_P-h)}{2d}\Big|_P, \quad \frac{\partial b}{\partial Z_P} = \frac{1}{2} b^{-1}\, 2Z_P\Big|_P = \frac{Z_P}{b}\Big|_P .$$

Hence f_P can be rewritten as

$$f_P = -2\pi\kappa\sigma \left[h + \frac{d}{2} + \frac{(Z_P-h)^2}{2d} - \frac{b}{2} - \frac{Z_P{}^2}{2b} + \frac{a^2}{2}\frac{Z_P+b}{Z_P-h+d} \times \right.$$

$$\left.\frac{\left(1+\frac{(Z_P-h)}{d}\right)(Z_P+b) - \left(1+\frac{Z_P}{b}\right)(Z_P-h+d)}{(Z_P+b)^2}\right]\Bigg|_P .$$

Evaluating the above formula at P on the cylinder we have $Z_P = h$, $d = a$, $b = \sqrt{(a^2+h^2)}$ and we get:

$$f_P = -\pi\kappa\sigma \left[2h + a - \sqrt{(a^2+h^2)} - \frac{h^2}{\sqrt{(a^2+h^2)}} + \right.$$

$$\left. + a^2 \frac{(h+\sqrt{(a^2+h^2)})(h+\sqrt{(a^2+h^2)}) - \left(1+\frac{h}{\sqrt{(a^2+h^2)}}\right) a}{a\,(h+\sqrt{(a^2+h^2)})^2}\right]$$

$$= -\pi\kappa\sigma \left[2h + a - \sqrt{(a^2+h^2)} - \frac{h^2}{\sqrt{(a^2+h^2)}} + a - \frac{a^2\left(1+\frac{h}{\sqrt{(a^2+h^2)}}\right)}{h+\sqrt{(a^2+h^2)}}\right]$$

$$= -\pi\kappa\sigma\ [2h + 2a - \sqrt{(a^2+h^2)} - \frac{h^2}{\sqrt{(a^2+h^2)}} - \frac{a^2}{\sqrt{(a^2+h^2)}}\ ]$$

$$= -\pi\kappa\sigma\ [2h + 2a - \sqrt{(a^2+h^2)} - \frac{a^2+h^2}{\sqrt{(a^2+h^2)}}\ ]$$

$$\boxed{= -\pi\kappa\sigma\ [2h + 2a - 2\sqrt{(a^2+h^2)}] = - 2\pi\kappa\sigma\ (h+a - \sqrt{(a^2+h^2)}).}$$

Here $\sqrt{(a^2+h^2)}$ can be developed into power series, considering $a > h$ and we obtain

$$f_P = - 2\pi\kappa\sigma\ (h + a - a(1 + \frac{1}{2}\ (\frac{h}{a})^2 + \ldots))$$

$$= - 2\pi\kappa\sigma\ (h - \frac{1}{2}\frac{h^2}{a} - \ldots)$$

$$= - 2\pi\kappa\sigma h\ (1 - \frac{h}{2a} - \ldots)\ .$$

Considering the diameter of the cylinder infinite, i.e., extending the cylinder to Bouguer's plate we get finally:

$$\boxed{\delta g_P = \lim_{a\to\infty} f_P = - 2\pi\kappa\sigma h,}$$

the correction due to the Bouguer's plate. In practice, σ of the upper part of the earth crust is usually assumed to be 2.67 g cm^{-3} giving thus

$$\boxed{\delta\ g_P = - 0.1119\ h}$$

in mgal for h in meters.

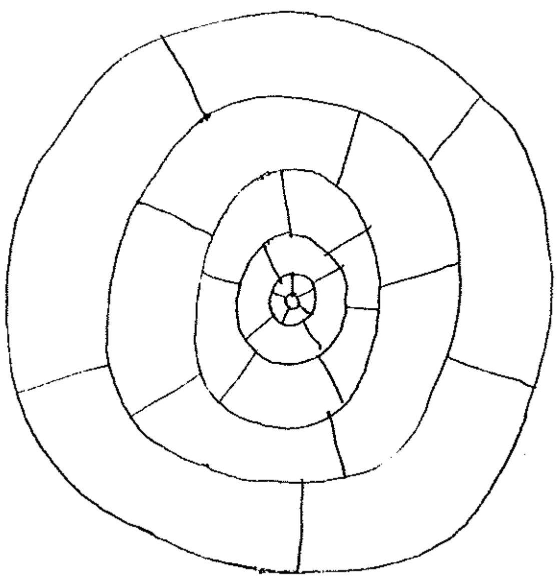
The sum $\delta g_F + \delta g_P = 0.1967$ h became known as <u>incomplete</u> (simple) <u>Bouguer correction</u> and the corresponding anomaly

$$\boxed{\Delta g_P = g + \delta g_F + \delta g_P - \gamma}$$

is called <u>incomplete</u> (simple) <u>Bouguer anomaly</u>.

4.7) Terrain correction and refined Bouguer anomaly

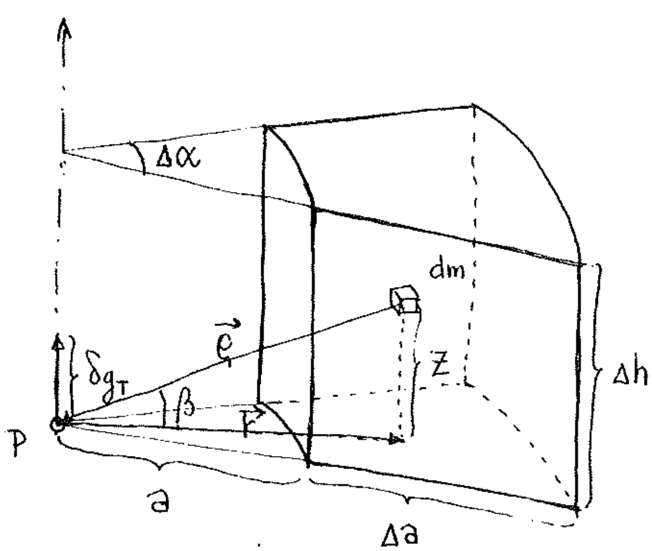
The second step in evaluating the influence of the masses between the geoid and the topographic surface consists of accounting for the masses trapped between the Bouguer plate and the surface. The evaluation is easily done using a template method. We first divide the area surrounding our station into compartments according to a template. One of such possible divisions is shown on the diagram. The contribution of each individual compartment is computed separately and their combined effect then determined.



To determine the contribution of one such compartment we can write, using again cylindrical coordinates r, α, Z for the attracting force in absolute value due to one mass element dm:

$$df_P = \kappa \frac{dm}{\rho^2}$$

where $\rho^2 = r^2 + Z^2$. Note that here, we are omitting the two primes over r. The vertical component of the attractive force, our correction $d\delta g_{T_i}$, due to one mass element in the i-th compartment is given by



$$d\delta g_{T_i} = df_P \sin\beta = df_P \frac{Z}{\rho} .$$

For the whole compartment(Δa by Δh by $\Delta\alpha$)we obtain, considering again homogenous density σ for the whole compartment and an average height Δh above P:

$$\delta g_{T_i} = \kappa\sigma \int_{\alpha=0}^{\Delta\alpha} \int_{Z=0}^{\Delta h} \int_{r=a}^{a+\Delta a} \frac{Z}{\rho^3} r \, dr \, dZ \, d\alpha$$

$$= \kappa\sigma \int_{\alpha=0}^{\Delta\alpha} \int_{Z=0}^{\Delta h} \int_{r=a}^{a+\Delta a} \frac{Zr}{(r^2+Z^2)^{3/2}} \, dr \, dZ \, d\alpha .$$

Integration with respect to α yields:

$$\delta g_{T_i} = \kappa\sigma\Delta\alpha \int_{Z=0}^{\Delta h} \int_{r=a}^{a+\Delta a} \frac{Zr}{(r^2+Z^2)^{3/2}} \, dr \, dZ .$$

Integration with respect to Z can be solved using substitution $t^2 = r^2 + Z^2$. We get:

$$\delta g_{T_i} = \kappa\sigma\Delta\alpha \int_{r=a}^{a+\Delta a} r \left(\int_{r}^{\sqrt{(r^2+\Delta h^2)}} \frac{dt}{t^2} \right) dr ,$$

where the inner integral gives $\left[-\frac{1}{t} \right]_r^{\sqrt{(r^2+\Delta h^2)}} = -\frac{1}{\sqrt{(r^2+\Delta h^2)}} + \frac{1}{r}$.

Hence we can write:

$$\delta g_{T_i} = \kappa\sigma\Delta\alpha \int_{a}^{a+\Delta a} \left(1 - \frac{r}{\sqrt{(r^2+\Delta h^2)}}\right) dr = \kappa\sigma\Delta\alpha \left(\Delta a - \int_{a}^{a+\Delta a} \frac{r}{\sqrt{(r^2+\Delta h^2)}} \, dr \right).$$

Another substitution $t^2 = r^2 + \Delta h^2$ yields

$$\delta g_{T_i} = \kappa\sigma\Delta\alpha \left(\Delta a - \int_{\sqrt{(a^2+\Delta h^2)}}^{\sqrt{((a+\Delta a)^2 + \Delta h^2)}} dt\right) = \kappa\sigma\Delta\alpha \left[\Delta a - \sqrt{((a+\Delta a)^2 + \Delta h^2)} + \sqrt{(a^2+\Delta h^2)}\right] .$$

For $\Delta h \ll a$ we get:

$$\delta g_{T_i} = \kappa\sigma\Delta\alpha\ [\Delta a + a\sqrt{(1+(\frac{\Delta h}{a})^2)} - (a+\Delta a)\sqrt{(1+(\frac{\Delta h}{a+\Delta h})^2)}]$$

or approximately:

$$\delta g_{T_i} \simeq \kappa\sigma\Delta\alpha\ [\Delta a + a(1+\frac{1}{2}(\frac{\Delta h}{a})^2) - (a+\Delta a)(1+\frac{1}{2}(\frac{\Delta h}{a+\Delta a})^2)]$$

$$= \frac{\kappa\sigma\Delta\alpha\Delta h^2}{2}\left(\frac{1}{a}+\frac{1}{a+\Delta a}\right) .$$

The complete terrain correction is given by:

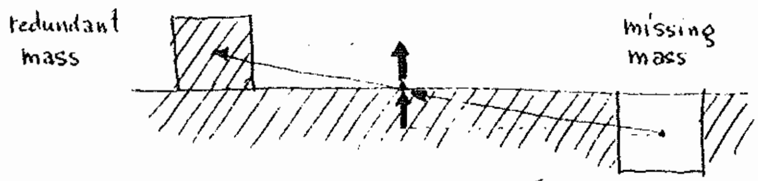
$$\boxed{\delta g_T = \kappa\sigma \sum_j \{\Delta\alpha_j \sum_i [a_{i+1} - a_i + \sqrt{(a_i^2 + \Delta h_{i,j}^2)} - \sqrt{(a_{i+1}^2 + \Delta h_{i,j}^2)}]\} .}$$

Various schemes have been devised for the template to simplify the above equation. With the appearance of computers the templates have lost their significance. However, even on computers one has to use one scheme or another, and one has therefore to understand the way how to set one up.

In all the schemes the determination of the terrain correction involves the determination of an average height Δh for each compartment. The closer compartments contribute more towards the whole correction. The heights there have to be known therefore with a higher accuracy. Hence the grid has to be denser in the vicinity of the gravity station. Sometimes, particularly in mountaineous areas, even a fine grid is not enough to give adequate precision. Then the contribution of the immediate vicinity is to be determined through other methods using gravity gradients or the gravity values themselves. These are to be furnished by additional observations. Obviously, the computations involved are quite tedious and the terrain correction is then applied only when it is absolutely necessary.

The magnitude of terrain correction is usually of the order of a few tenths of a mgal for flat and gently rolling country. It reaches a few milligals in a hilly area and tens of milligals in the mountains. Corrections of about 40 mgal were experienced in Colorado, 70 mgal in the Carpathains, etc.

We notice that the terrain correction is always positive, whether the terrain is lower or higher than the observation station. This fact can be understood from the drawing:



Hence, neglect of the terrain correction introduces a systematic bias.

The sum $\delta g_F + \delta g_P + \delta g_T$ is known as complete (refined) Bouguer correction. Correspondingly,

$$\boxed{\Delta g_B = g + \delta g_F + \delta g_P + \delta g_T - \gamma}$$

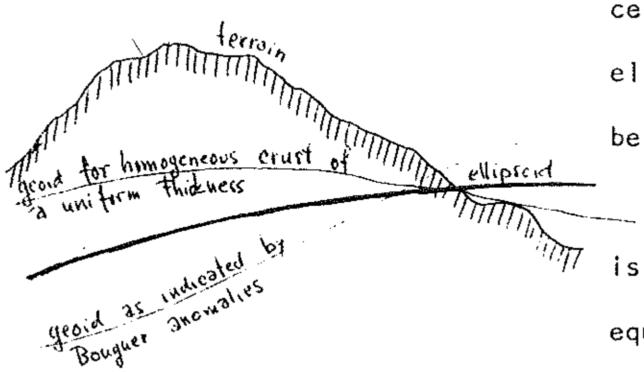
is called complete Bouguer anomaly.

We can note that here, as well as when deriving the incomplete Bouguer correction, we have considered the earth flat. This can be done since the attraction of the masses departing from the horizontal plane of the gravity station diminishes very rapidly. However, for very precise work this "spherical effect" has to be accounted for mathematically. In such a case we usually begin considering the earth curved from the spherical distance of 1.5° (~167 km). This distance corresponds to the radius of the inner area in Hayford template of which more will be said in 4.8.5. From this distance on, the thickness of the topography and even the Bouguer plate may be regarded as zero when compared with the distance. Hence, the Bouguer correction (refined) can be computed from the expression for the potential of a surface layer of variable density.

4.8) Isostatic correction and anomaly

4.8.1) Isostasy

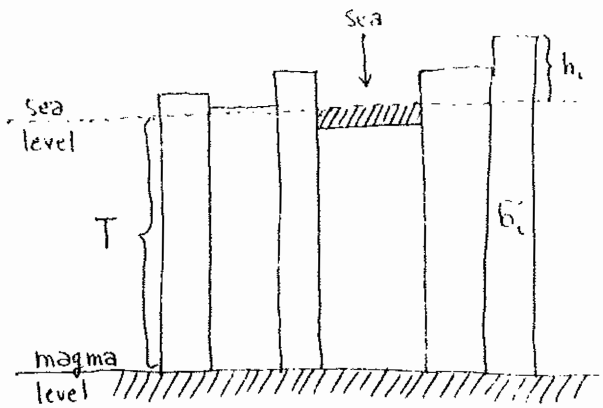
When one computes the Bouguer anomalies in larger areas one discovers that they have systematically large positive values on the oceans and large negative values under the mountains. Such values correspond generally to positive geoidal undulations on the seas and negative undulations under the mountainous regions. This discovery contradicts our expectations based on physical principles. If the earth crust were homogeneous and of a uniform thickness we would expect the geoid to follow the terrain, to a certain degree, to protrude above the ellipsoid under the mountains and lay below the ellipsoid on the seas.



The explanation for this discovery is sought in isostasy, the theory of equilibrium of the earth crust. The idea of isostasy was probably originated already by Leonardo da Vinci. The first mathematical formulations can be found in the theories of J.M. Pratt (1854) and G.B. Airy (1855). More recently (1931) , F. A. Vening-Meinesz produced his theory which is being accepted as the, perhaps, most realistic of all the three today. Since isostasy is important for geodesy we are going to outline all of these theories.

4.8.2) Pratt's model and theory

Pratt's basic idea is that the earth crust is divided into more or less independent blocks of different density. He then considers the blocks as floating on the level of magma that lies in the depth T. This



level is usually called the "compensation level".

For the individual blocks (columns) to exert the same pressure on the magma it is necessary that the product $(T + h_i)\ \sigma_i$ be constant, i.e. be the same for all the columns. Postulating a certain depth T of the compensation level and an average density σ_o, we can express the constant as

$$\text{const.} = T\sigma_o$$

and regard it as a contribution of a column with average height zero. From these two expressions we can determine the density of each individual column, protruding h_i km, in average, above the sea level:

$$\boxed{\sigma_i = \sigma_o \frac{T}{T+h_i}}\ .$$

Here T has to be postulated (usually around 100 km to make the individual densities realistic) and σ_o is generally taken to be 2.67 g cm^{-3}.

Dealing with a block submerged in the ocean, we have to add also the weight of the involved water column. Denoting by σ_W the density of water (usually taken as 1.027 g cm^{-3}) and by D_i the mean depth of the crust block under the sea, we get

$$(T-D_i)\ \sigma_i + D_i\sigma_W = T\sigma_o .$$

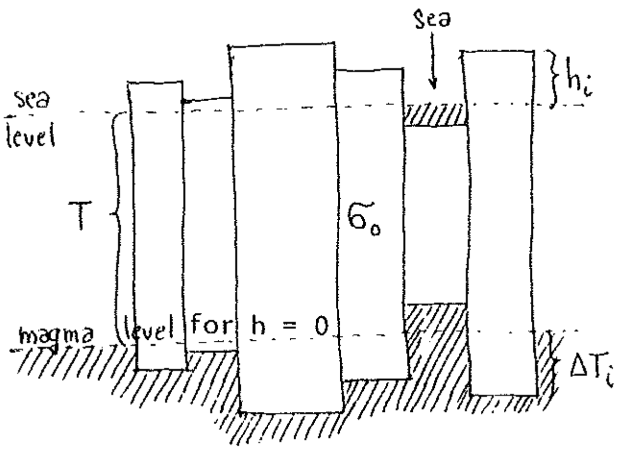
Hence

$$\boxed{\sigma_i = \frac{T\sigma_o - D_i\sigma_W}{T - D_i}}\ .$$

The Pratt's model with certain modifications was used by J.F. Hayford (1912) for smoothing of the gravimetric deflections of vertical which were deployed for the determination of the best-fitting ellipsoid to North America.

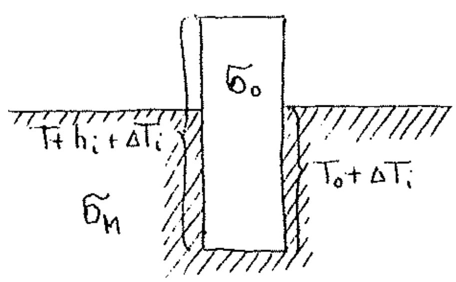
4.8.3) Airy's model and theory

Airy's model is based on the analogy of the earth crust blocks with icebergs. His model assumes a constant density $\sigma_o = 2.67$ g cm^{-3} for all the individual blocks and has, therefore, to conclude that the blocks sink differently into the plastic magma according to their heights.



From the Archimedes law the equilibrium of a column $T + h_i + \Delta T_i$ high, submerged into a liquid $T_o + \Delta T_i$ deep, is achieved when following equality is satisfied:

$$(T + h_i + \Delta T_i)\, \sigma_o = (T_o + \Delta T_i)\, \sigma_M .$$



Here $T_o + \Delta T_i$ is the depth of submersion and σ_M is the density of the liquid, in our case the magma. σ_M is usually assumed to be 3.27 g cm^{-3}.

From the equation for a column of zero height h_i (and therefore zero submersion ΔT_i):

$$T\sigma_o = T_o\sigma_M$$

we can determine

$$T_o = T\ \frac{\sigma_o}{\sigma_M} .$$

Then the first equation yields

$$h_i\sigma_o + \Delta T_i\sigma_o = \Delta T_i\sigma_M$$

and

$$\boxed{\Delta T_i = h_i \, \frac{\sigma_o}{\sigma_M - \sigma_o} \doteq 4.45 \, h_i .}$$

For the blocks under the oceans we can write similarly

$$(T - \Delta T_i - D_i) \, \sigma_o + D_i \, \sigma_W = (T_o - \Delta T_i) \, \sigma_M$$

where again D_i is the average depth of the ocean and σ_W the density of water assumed to be 1.027 g cm^{-3}. Substituting for T_o we get

$$-\Delta T_i \sigma_o - D_i \sigma_o + D_i \sigma_W = - \Delta T_i \sigma_M$$

and

$$\boxed{\Delta T_i = D_i \, \frac{\sigma_o - \sigma_W}{\sigma_M - \sigma_o} = 2.73 \, D_i \; .}$$

The thickness of the crust, according to the Airy's model, is then given by

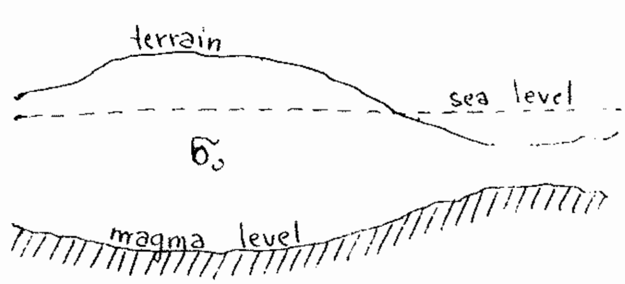
$$t = \begin{cases} T + h + 4.45 \, h = T + 5.45h & \text{on the continents} \\ T - D - 2.73D = T - 3.73 \, D & \text{under the oceans.} \end{cases}$$

T is generally postulated to be somewhere between 30 to 50 km.

The Airy's model was originally used by Heiskanen in his first attempts to compute the isostatically corrected anomalies.

4.8.4) Vening-Meinesz's model and theory

Both the preceding models assume the individual blocks (columns) to move more or less independently of each other. The geophysical investigations have shown some evidence that this is not quite the case and that the models are somewhat oversimplified.



This was the reason why Vening-Meinesz came up with a different model of earth crust. He regards the crust as an elastic homogeneous layer of variable thickness known as regional model. The mean thickness is assumed to be about 30 km. The mathematical description of the model is based on the theory of elasticity and is quite complicated. Because of its complicated mathematics it is rather seldom used in practice.

According to our present knowledge the earth crust behaves as a combination of all of these three models. It has definitely a variable density as well as variable thickness and behaves as a "structured" elastic layer. Until more has been learnt about the crust from other sources (seismology, earth tides, geology, tectonics, etc.) it will be difficult to design a truly realistic model. Besides, the isostatic compensation still goes on and has to be regarded as a continuous dynamic process contributing to the so-called crustal movements.

4.8.5) Isostatic correction and anomaly

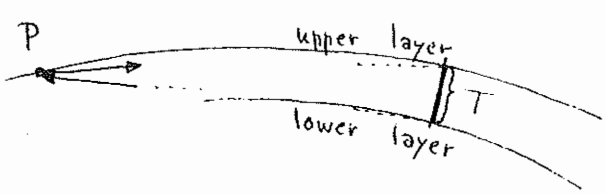
In order to account for the earth crust isostasy (the crust is known to be isostatically compensated over an area covering about 90% of the earth surface) one has to change the Bouguer or the refined Bouguer correction accordingly. This can be done using basically the same technique as we have used for computing the terrain corrections. We divide the earth surface into compartments, using a template, and obtain the corresponding columns of the earth crust. In other words we chose the boundaries of the crustal blocks according to the template disregarding the actual structure of the blocks which (if it exists) is largely unknown anyway.

Then the Bouguer correction (without the free-air component) can be regarded as having accounted for the attraction of some "mean columns" of density σ_0 and height either $T + h_i$, in case of the refined correction, or $T + h$, where h is the height of our gravity station, in case of the simple correction. When we want to account for the isostasy we have to subtract the effect of the difference between these mean columns and the isostatically compensated columns.

Hence, using the Pratt's model, we subtract the effect of the columns of the same sizes but with the densities $\delta_i = \sigma_o - \sigma_i$ computed for σ_i given in 4.8.2. This was what Hayford did in 1912 selecting T = 113.7 km. The template pattern which he devised then, has been serving as a prototype for almost all the subsequent attempts in the field of isostatic corrections. The template used has also proved useful for various other tasks such as the computations of the terrain corrections mentioned already in 4.7. His design was based on the necessity of doing all the compu ations manually and lost somewhat its importance with the introduction of computers.

When we use the Airy's model we have to subtract the effect of the immersed parts of the columns, with density $\sigma_M - \sigma_o$. Various tables assuming different values of T have been published mainly by Finnish geodesists and Heiskanen's name has to be mentioned in this context. Similar tables by Vening-Meinesz are available for his own isostatic model.

When computing the isostatic correction we have to start considering again the earth to be curved from a certain distance. As in the cases of the plate or the terrain corrections this is usually chosen to be 1.5°. From this distance on, the combined correction δg_I (Bouguer without the free-air ~~and~~ compensation corrections) can be computed from plus the expression for the potential of double surface layer T km apart. The upper layer, corresponding to the "condensed topography" (variable density corresponds to the variable heights) is attracting, the lower layer, corresponding to the lack of mass in the column, is repelling. Hence each column of the height T acts as a dipole.



The sum of the combined correction and the free-air correction is known as isostatic correction. The anomaly

$$\boxed{\Delta g_I = g + \delta g_F + \delta g_I - \gamma}$$

is called isostatic (isostatically compensated) anomaly. The combined correction δg_I is generally fairly small in comparison with δg_F and the free air correction is then often regarded as the first approximation to the isostatic correction. This can be understood when we realize that the free air anomaly may be interpreted as assuming the masses above the sea level to be fully compensated for by the lack of masses underneath and therefore having no effect on the gravity station on the surface.

4.9) Other gravity corrections and anomalies

Apart from the corrections and anomalies dealt with in the precedent sections, there are many more corrections and anomalies defined in literature. They are based on hypotheses different from the ones we have met already. To name but a few, let us mention the following:

i) Rudzki's correction based on the principle of spherical inversion with respect to a sphere of a radius $\tilde{R}$ going everywhere underneath the geoid. The masses protruding above the geoid are "shifted" inside the sphere. The correction is mathematically rigorous but does not have any physical interpretation. It is therefore seldom ever used.

ii) Helmert's condensation correction accounts for the redundant masses by means of expressing their effect in terms of a surface layer (on the geoid) of varying density. It is equivalent to the Pratt's isostatic correction computed for $T = 0$. It approximately equals to the free air correction.

iii) Bruns' correction is nothing else but the free-air correction related to the reference ellipsoid rather than geoid. It has, therefore, formally the same shape as the δg_F (see 4.5) with h + N replacing the h. Theoretically, this correction should be preferred against the free-air since it removes the effect of the masses above the ellipsoid and provides us thus with the boundary value on the ellipsoid. Its application is hindered by our lack of knowledge of the geoid.

4.10) Indirect effect, cogeoid

When removing the effect of the Bouguer plate and/or the terrain undulations, we actually disregard the masses above the geoid. In other words, when using the Bouguer anomaly (simple or refined) we mathematically change the real distribution of masses, the potential of the earth and hence even the geoid. Expressing the change of the potential due to the removal of the Bouguer plate and/or the terrain by T_B we may say that the potential W of the earth is changed by $\delta W_B = T_B$. The effect of the Bouguer anomalies on the geoid computed by means of the Stoke's formula is hence given, using the Bruns' formula:

$$\delta N_B = \frac{\delta W_B}{\gamma} = \frac{T_B}{\gamma} \quad .$$

This distortion is usually called the indirect effect of the mass removal and the surface thus distorted is known as cogeoid. δN_B of the Bouguer anomalies can be as large as 440 m, i.e. much larger than N itself.

The cogeoid can be reduced to the geoid by evaluating the quantity T_B and then δN_B all over the earth surface. The template method may be used for this purpose again to determine the T_B's in individual compartments in much the same way as it was used for determination of the terrain corrections. The only difference is that here we would be looking for differences in potential rather than differences of the attractive force.

Similar argument holds true for the isostatic anomalies. However, they yield a different cogeoid that has to be transformed into the geoid using again the above formula, where δW is given as the difference $T_B - T_C$. T_C here denotes the potential of the "anomalous blocks" the effect of which was subtracted from the "regular Bouguer blocks". Since the absolute value of $T_B - T_C$ is much smaller than the absolute value of T_B, even the indirect effect of the isostatic anomalies is much smaller than that of the Bouguer anomalies. It is of the order of $\pm$ 10 m.

It is not difficult to see that there is no indirect effect produced by the free-air anomalies. There we do not manipulate with the masses at all. This is one of the outstanding advantages of using the free-air rather than any other anomalies.

4.11) Discussion of the individual gravity anomalies

In spite of our intuition it is not the Bouguer anomaly that depicts the real distribution of masses the best. Although it seems to account for the visible distribution of masses adequately, the fact that there is the mass deficiency in the lower part of the crust tending to cancel out the mass redundancy in the upper part of the crust, distorts its real meaning. This can be seen from its huge indirect effect. Hence the Bouguer anomaly is not recommended for geoid determination. However, it is still very useful for geophysical prospecting because it varies very smoothly and reflects the local gravity irregularities in the most useful manner.

The isostatically compensated (isostatic) anomalies are obviously the most truthful representation of the nature and would be theoretically the best to use for geoid determination. Their distinct disadvantage is the complicated computation. We have seen that one has to compute the refined Bouguer correction first then determine the compensating correction and finally evaluate the indirect effect after having applied the Stokes' formula. More precisely, the Stokes formula is really evaluated on the cogeoid and therefore the free-air correction should be taken with respect to the cogeoid. Hence the height used in determining the free-air correction should be $h + \delta N_I$ instead of h.

As we have stated in 4.8.5 already, the free-air anomaly can be considered as the first approximation to the isostatically compensated anomaly. In addition, it is very simple to compute and has no indirect effect. These are the two reasons why the free-air anomalies are used almost exclusively for gravimetric determination of geoid.

Many "gravimetric geoids" have been computed by various authors. The best known geoids have been produced by Hirvonen (1934), Jeffreys (1943), Tanin (1948, 49), Heiskanen (1957), Uotila (1962, 64) and Kaula (1961, 1966). The literature is rich with examples. The individual geoids vary quite widely mainly due to different interpolation techniques used by the authors to determine the interpolated values of anomalies in the unsurveyed areas. Even the best geoids are not supposed to have better precision than some $\pm$ 10 m. This is because of the fact that large regions on the surface of the earth still remain unsurveyed and because the Stokes' formula, due to all the approximations used, has its own inherent imprecision. In section 6 we are going to show some more precise techniques for computing geoidal undulations above the local datum.

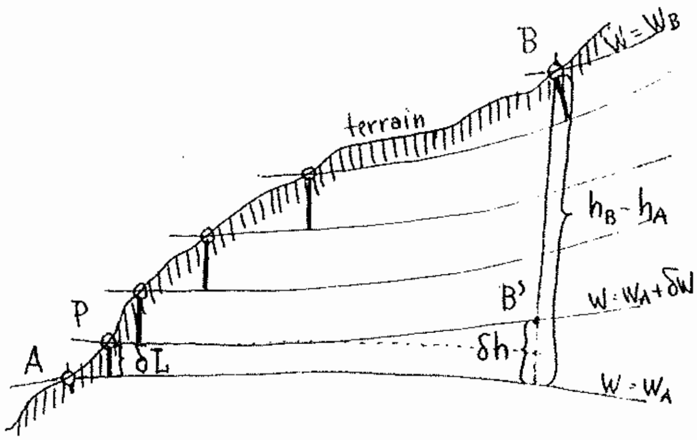
4.12) Gravity maps, gravity data banks

It is usual to depict the results of a gravity survey in the form of a map. The maps may deal with either the observed gravity or with anomalies. They may depict either the individual points or provide us with lines joining the points of equal anomalies -- isoanomales. The maps can have different scales and show therefore regions of different size. For us, the most important are the small scale map of free-air anomalies.

Recently, the tendency has been to replace the maps by other forms of data representation. Namely, the user may be now provided with gravity data from a certain area in a digital form. The data bank of the Gravity Section of the Earth Physics Branch (Department of Energy, Mines and Resources) in Ottawa may serve as an example in this respect. The gravity data can be supplied by them in either the punch-card form or on magnetic tape accompanied by a computer listing and a coarse map in any wanted scale containing the requested anomalies or the observed values. Another example is the U.S. Air Force agency, "Aeronautical Charts and Information Center", dealing with world-wide gravity data that can be obtained from them.

5) Heights

5.1) Observed heights



Let us consider a levelling line A,B with the intermediate stations depicted by circles. When we realize that the gravity field is represented by generally non-parallel equipotential surfaces, it is not difficult to see that the sum of the observed level differences δL is not equal to the sum of the height differences δh. Hence the levelled difference of any two points A,B is not equal to the height difference A and B, $h_B - h_A$. This is due to the non-parallelness of the equipotential surfaces.

The actual relationship between the observed level difference δL and the "corresponding" height difference δh can be expressed through gravity. We can write

$$\left.\frac{dW}{dh}\right|_P = - g_P, \qquad \left.\frac{dW}{dh}\right|_{B'} = - g_{B'},$$

or using differences

$$\delta W = - g_P \ \delta L = - g_{B'} \ \delta h$$

where g_P, $g_{B'}$ is the actual gravity at P and B' respectively. Hence we can write for δh:

$$\boxed{\delta h = \frac{g_P}{g_{B'}} \ \delta L}$$

and if $g_P \neq g_{B'}$ then $\delta h \neq \delta L$.

This unfortunate property results in the fact that various levelling lines connecting the same two points yield different level differences $\Sigma\delta L$. Using differentials instead of differences, i.e., abstracting from the actual levelling, we can write generally

$$\int_A^B dL \neq h_B - h_A$$

and also

$$\underset{\text{Path 1}}{\int_A^B dL} \neq \underset{\text{Path 2}}{\int_A^B dL}.$$

The last property is usually written as

$$\oint dL \neq 0$$

and quoted as integration over a closed circuit.

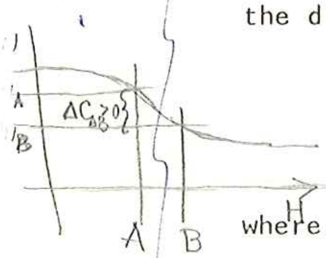
We can conclude that the results of levelling without gravity taken into account do not define the height of every point uniquely. But the unique definition of a height of every point on the surface is a highly desirable feature. In order to achieve it, we have to take the gravity into consideration.

5.2) Geopotential numbers

One way how to define the heights uniquely is to use the equipotential surfaces directly to define the height of a point. We may say that a point laying on equipotential surface $W = C_B$ is above (below) the point laying on equipotential surface $W = C_A$ by

$$\boxed{\Delta C_{AB} = C_{\cancel{B}\,A} - C_{\cancel{A}\,B} .}$$

The numbers $C_0 - C_B$, $C_0 - C_A$, where $C_0 \overset{= W_0}{}$ is the potential of the geoid, are known as geopotential numbers defining the heights of B and A. Evidently, the difference of two equipotential numbers $C_B - C_A$ can be computed from



$$C_{\cancel{B}\,A} - C_{\cancel{A}\,B} = \Delta C_{AB} \overset{-\Delta W_{AB} =}{=} \int_A^B g \, dL \simeq \sum_{i=A}^{B} g_i \, \delta L_i$$

where g (g_i) is the surface (actual) gravity along the levelling line connecting A with B.

It is not difficult to see that the geopotential numbers define the height of every point uniquely. This means that whatever line connecting A with B we take, the ΔC_{AB} will be always the same. We say that the geopotential numbers are not path-dependent and write

$$\oint g \, dL = 0 .$$

The disadvantage of the geopotential numbers is that they are not given in length units but in cm^2/s^2 -- hence the name. Numerically, they depart from the observed heights by some 2% even when we chose the units in the most convenient way, i.e. when we express the gravity in kgal, so that $g \sim 1$. This is far too much for any technical work and it is the reason why geopotential numbers are very seldom used in technical practice.

5.3) Dynamic heights

Dynamic heights are designed to retain the advantage of the geopotential numbers and eliminate ~~its~~ their disadvantage. The dynamic height of a point A is defined as

$$\boxed{h_A^D = C_A/G}$$

where G is a gravity value selected as a reference. G is generally chosen in such a way as to be close to the average value of g for the area in question -- usually a country or a group of countries.

We can see that h^D is expressed in length units and its value does not deviate from the levelled height as much as the value of the corresponding geopotential number does. On the other hand, points laying on one equipotential surface have the same dynamic height. This is usually expressed by ~~a~~ the slogan-- "in dynamic height system a lake surface is flat."

It is a general practice to express the dynamic height difference of two points A,B in terms of a correction to the observed level difference. To derive such a correction let us write:

$$\Delta h_{AB}^D = \frac{\Delta C_{AB}}{G} \doteq \frac{1}{G} \sum_{i=A}^{B} g_i \, \delta L_i \, .$$

Adding and subtracting G to g_i we obtain

$$\Delta h^D_{AB} \doteq \frac{1}{G} \sum_i (g_i + G - G)\delta L_i = \sum_i \delta L_i + \sum_i \frac{g_i - G}{G} \delta L_i .$$

Here

$$\sum_i \delta L_i = \Delta h^M_{AB}$$

is the observed (measured) level difference between A and B and

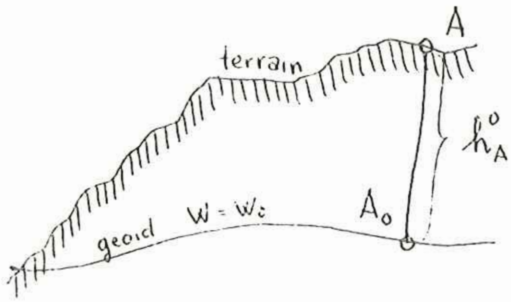
$$\boxed{\sum_i \frac{g_i - G}{G} \delta L_i = \Delta^D_{AB}}$$

with $\frac{g_i - G}{G} \ll 1$

is the dynamic correction we are looking for.

5.4) Orthometric heights

Orthometric height of a point A is defined as the length of the actual plumb-line connecting the point A with the geoid. We have seen in



5.1) that the difference δW can be expressed as

$$\delta W = - g' \delta h$$

where g' is the gravity along the plumb-line. We can write thus for δh

$$\delta h = - \frac{\delta W}{g'}$$

and

$$h^0_A = - \int_{W_0}^{W_A} \frac{dW}{g'} .$$

(i) $h^0_A = \int_0^A dh = -\int_0^A \frac{dW}{g'}$

(ii) "mean value" theorem:

$h^0_A = -\frac{1}{\bar{g}'}(W_A - W_0)$ mean value

(iii) $W_A - W_0 = -C_A$

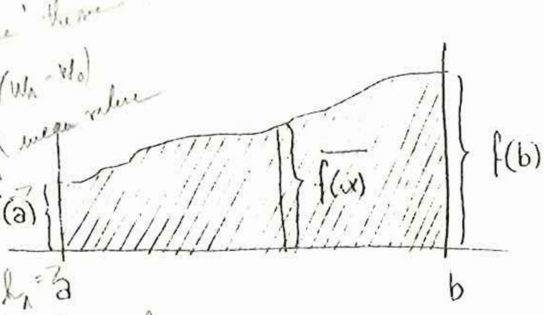
(iv) taking ...

$\frac{\partial g}{\partial h}$ = const $\frac{\partial g}{\partial h} h_A$

$g' = g_A + \frac{\partial g}{\partial h}$

$\bar{g}' = \frac{1}{2}(g_A + g_0) = g_A + \frac{\partial g}{\partial h}\frac{h_A}{2}$

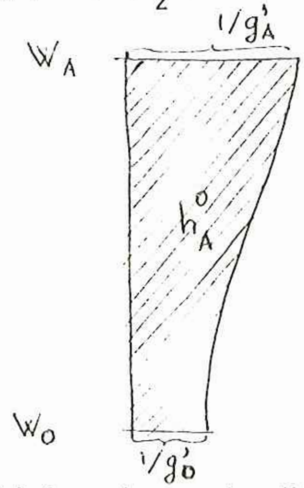
We now recall the theorem about the "mean value" saying that if f is continuous on [a,b] then the definite integral $\int_a^b f(x)\,dx$ can be always expressed as $\overline{f(x)}\,(b-a)$ where $\overline{f(x)}$ is a value between f(a) and f(b).



Obviously, if f is a linear function, then $\overline{f(x)} = f\left(\frac{b-a}{2}\right)$. Applying the theorem on our case, where

$[a,b] = [W_o, W_A]$, $f = 1/g'$,

and the definite integral equals the orthometric height of A, we get:



Feb. 20

$$\boxed{h_A^0 = -\frac{1}{\bar{g}'}(W_A - W_0) = \frac{C_A}{\bar{g}'}\,.}$$

The orthometric height of a point is then defined as the ratio of its geopotential number and a "mean gravity" along its plumb-line in the sense of the theorem quoted earlier.

In order to evaluate the mean gravity, let us write

$$\bar{g}' = -\frac{1}{h_A^0}(W_A - W_0) = -\frac{1}{h_A^0}\int_{W_0}^{W_A} dW = \frac{1}{h_A^0}\int_{h^0=0}^{h_A^0} g'dh\,.$$

In the last integral, g' has to be assumed known along the plumb-line as a function of height. According to the type of assumption (hypothesis) used, we get various definitions of orthometric heights.

The best known of the various definitions is that of <u>Helmert</u> who uses the <u>Poincaré-Pray's hypothesis</u> concerning the gravity along the plumb-line. Their approach is based on the 1st Bruns' formula (see Physical Geodesy I, § 3.14) that reads:

$$\frac{\partial g}{\partial h} = -\,2gJ + 4\pi\kappa\sigma - 2\omega^2\,.$$

The same formula, applied to normal gravity above the mean earth ellipsoid yields

$$\frac{\partial \gamma}{\partial h} = -\,2\gamma J_o - 2\omega^2\,.$$

Here J and J_0 are the mean curvatures of the actual equipotential surface and the corresponding normal equipotential surface respectively. Taking approximately

$$gJ \simeq \gamma J_0$$

we get when subtracting the second from the first formula:

$$\frac{\partial g}{\partial h} \simeq \frac{\partial \gamma}{\partial h} + 4\pi\kappa\sigma .$$

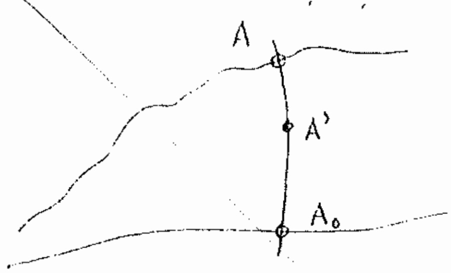
In this formula $\frac{\partial \gamma}{\partial h}$ can be taken approximately equal to the "free air" gradient; i.e., -0.3086 mgal/m. $4\pi\kappa\sigma$, for $\sigma = 2.67$ g cm^{-3}, becomes 0.2238 mgal/m.

Hence

$$\boxed{\frac{\partial g}{\partial h} \simeq -\ 0.0848 \text{ mgal/m.}}$$

This is the Poincaré-Pray gradient of the actual gravity underground.

Now, let us write for any point A' on the plumb-line of A:



$$g_{A'} = g' = g_A + \int_A^{A'} \frac{\partial g}{\partial h}\, dh$$

$$= g_A - 0.0848 \int_A^{A'} dh \quad \text{[mgal]} .$$

Finally, we obtain

$$g' = g_A - 0.0848\ (h^0_{A'} - h^0_A)$$

$$= g_A + 0.0848\ (h^0_A - h^0_{A'}) \quad \text{[mgal]} \qquad (*)$$

for h^0 in meters.

We can now evaluate the mean gravity along the plumb-line. Substituting the last result for g' in the formula for $\overline{g}'$ we get

$$\overline{g}' = \frac{1}{h^0_A} \int_0^{h^0_A} (g_A + 0.0848\ (\ h^0_A - Z))\ dZ$$

$$= g_A + \frac{1}{h_A^0}\, 0.0848\, [h_A^0 z - \frac{z^2}{2}]_0^{h_A^0}$$

$$= g_A + 0.0424\, h_A^0 \quad [\text{mgal for } h_A^0 \text{ in m }].$$

Note that after all the computations we are ending up with the value for $\bar{g}'$ that could have been deduced from (*) right away for g' of the central point of the plumb-line:

$$\bar{g}' = g_A + 0.0848\, (h_A^0 - \frac{h_A^0}{2})\ .$$

This is the consequence of having chosen g to be a linear function of h^0.

Using the determined value of $\bar{g}'$ we can now write for the Helmert's orthometric height:

$$h_A^0 = C_A/(g_A + 0.0424\, h_A^0)$$

$$\bar{g}' = g_A + 0.0848\, \frac{h_A}{2}$$

or, with sufficient precision:

$$\boxed{h^0 = C/(g + 0.0424\, h^M).}$$

Other, more refined variations on the same theme, are definitions due to Niethammer, Mader, Ledersteger, et. al. Many more approximate systems, used extensively in technical practice have been put forward by Baranov, Ramsayer, Ledersteger, et. al.

It is again a general practice, as in the case of dynamic heights, to express the orthometric height (whichever it may be) difference as a sum of two terms -- the observed level difference and the correction. In order to do so let us write for the orthometric height of a point A:

$$h_A^0 = h_A^D\, \frac{G}{\bar{g}_A'}$$

and similarly for B. For the difference of their orthometric heights we get

(i) height diff.
(ii) ortho. → dynamic
(iii) add & subtract 1
(iv) dynamic → orthometric
(v) orthometric → observed

$$\Delta h^O_{AB} = h^O_B - h^O_A = h^D_B \frac{G}{\bar{g}'_B} - h^D_A \frac{G}{\bar{g}'_A}$$

$$= h^D_B + h^D_B \left(\frac{G}{\bar{g}'_B} - 1 \right) - h^D_A - h^D_A \left(\frac{G}{\bar{g}'_A} - 1 \right)$$

$$= \Delta h^D_{AB} + h^D_B \frac{G - \bar{g}'_B}{\bar{g}'_B} - h^D_A \frac{G - \bar{g}'_A}{\bar{g}'_A} .$$

But $\Delta h^D_{AB} = \Delta h^M_{AB} + \Delta^D_{AB}$, $h^D_B = h^O_B \frac{\bar{g}'_B}{G}$ and similarly $h^D_A = h^O_A \frac{\bar{g}'_A}{G}$. Hence

$$\Delta h^O_{AB} = \Delta h^M_{AB} + \Delta^D_{AB} + h^O_A \frac{\bar{g}'_A - G}{G} - h^O_B \frac{\bar{g}'_B - G}{G}$$

$$= \Delta h^M_{AB} + \Delta^O_{AB} .$$

In the corrective terms, the orthometric heights can be replaced by observed heights without any detrimental effect and we finally get for the

<u>orthometric correction</u>

$$\boxed{\Delta^O_{AB} = \Delta^D_{AB} + h^M_A \frac{\bar{g}'_A - G}{G} - h^M_B \frac{\bar{g}'_B - G}{G} .}$$

The interesting thing to note about the orthometric correction is that it can be regarded as being composed of three dynamic corrections for the following three lines AB, A_0A, B_0B. Or to be more precise as a sum of dynamic corrections in the open circuit A_0ABB_0.

To conclude with, let us note that orthometric heights of different points laying on one equipotential surface are generally different. Hence in the orthometric system, a lake surface is not flat and water may "flow uphill." On the other hand, they are usually numerically closer to the levelled heights. Since they have a definite geometrical meaning --

geometrical heights above the sea level (geoid) -- they are also considered as the most appealing intuitively and used in practice almost exclusively.

As an example of the magnitude the orthometric correction can attain is the 23 cm experienced in the levelling of one Alpine road (Biasca-St. Bernardino) on a stretch of 50 km climbing from 300 to 2000 m. This is about 30 times larger than the tolerance limit for precise levelling.

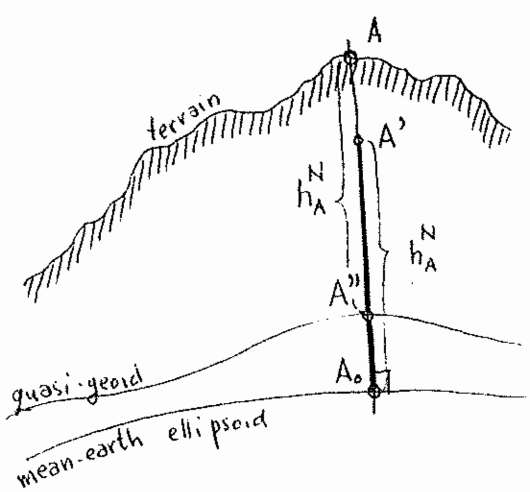
5.5) Normal heights

The normal heights are not supposed to describe the heights above the geoid. They relate the points to another surface known as quasigeoid and are closely tied to modern geodetic theories, Molodenskij and Hirvonen in particular. Quasigeoid is a purely mathematical surface (without any physical meaning) that departs from the geoid by at most a few metres and coincides with it on the seas.

On the other hand, the normal heights can be regarded as approximation to orthometric heights. The normal height of a point A is defined as

$$\boxed{h_A^N = C_A/\bar{\gamma}'}$$

where $\bar{\gamma}'$ is the mean normal gravity along the plumb-line of A taken from the mean earth ellipsoid to a point h_A^N above the ellipsoid. In other words, $\bar{\gamma}'$ is computed as a mean value of normal gravity in A_0A' (see the diagram). Note that $\bar{g}'$ (see the last §) would be a mean value of actual gravity in AA'' and due to the relationship between the geoidal (and therefore to a certain extent even the quasigeoidal) undulations and the gravity field, the two values, $\bar{g}'$, $\bar{\gamma}'$, are not too far apart.



We can again write for the mean value $\bar{\gamma}'$:

$$\bar{\gamma}' = \frac{1}{h_A^N} \int_0^{h_A^N} \gamma' \, dh$$

where h = 0 is the ellipsoid and $\gamma'(h)\Big|_{h=0} = \gamma_0$ is the normal gravity on the ellipsoid. In Physical Geodesy I, § 3.14, we have developed the following formula for the vertical gradient of normal gravity:

$$\frac{\partial\gamma}{\partial h} \simeq -\frac{2\gamma}{a}(1 + m + f \cos 2\phi).$$

Integrating the vertical gradient from the ellipsoid upwards yields

$$\gamma \simeq \gamma_0 - \frac{2\gamma_0}{a}(1 + m + f \cos 2\phi)\, h.$$

Substituting this γ for γ' into the above integral we get finally:

$$\bar{\gamma}' \simeq \gamma_0 - \frac{\gamma_0}{a}(1 + m + f \cos 2\phi)\, h_A^N .$$

(Note that the same result can be obtained from the equation for γ if we take $h = h_A^N/2$). Heights computed for this $\bar{\gamma}'$ are known as Molodenskij heights. They are used exclusively in Russia and Eastern Europe instead of orthometric heights.

We can now have a look at the second term in the above expression. It can evidently be regarded as a corrective term to γ_0. Realizing that

$$\gamma_0 \sim \frac{\kappa M}{R^2}$$

and $a \cong R$ we can write the second term as

$$\frac{\gamma_0}{a}(1 + m + f \cos 2\phi)\, h_A^N \simeq \frac{\kappa M}{R^3}\, h_A^N .$$

Comparison with the free-air correction § 4.5 convinces us that our corrective term can be written in first approximation as negative free-air correction to γ_0 for $h = h_A^N/2$. Numerically, we obtain

$$\bar{\gamma}'^{V} = \gamma_0 - 0.3086\ h_A^N/2 = \gamma_0 - 0.1542\ h_A^N\ .$$

Normal heights using $\bar{\gamma}'^{V}$ are called Vignal heights. They have been adopted in Western Europe for the unified European levelling network.

It is easily seen that the normal correction to the observed level difference Δh_{AB}^M is given by the same formula as the orthometric correction with the exception that instead of $\bar{g}'$ we write $\bar{\gamma}'$. Hence

$$\boxed{\Delta_{AB}^N = \Delta_{AB}^D + h_A^M\ \frac{\bar{\gamma}_A'-G}{G} - h_B^M\ \frac{\bar{\gamma}_B'-G}{G}\ .}$$

Obviously, the normal heights define the vertical position of every point again uniquely.

5.6) Heights based on normal gravity

Wherever the gravity survey is not detailed enough to evaluate the geopotential numbers and the corrections on the basis of observed surface gravity, we can at least take the normal gravity into account. The normal gravity accounts for the effect of the overall convergence of the equipotential surfaces corresponding to the change of γ from the equator to the poles of the order of 5400 mgal. The local irregularities of the actual gravity field, usually less than 200 mgal, are neglected.

If we decide to use the normal instead of observed gravity, we compute the approximate values of geopotential number differences $\tilde{\Delta C}$ from the following formula

$$\boxed{\tilde{\Delta C}_{AB} = \sum_{i=A}^{B} \gamma_i\ \ \delta L_i\ .}$$

The normal gravity here is, of course, a function of both latitude ϕ_i and height h_i of the point for which γ_i is computed. It is quite sufficient to determine γ_i from

$$\gamma_i = \gamma_0\ (\phi_i) - 0.3086\ h_i^M$$

where $\gamma_0(\phi_i)$ is given by one of the international formulae for normal gravity on the mean earth ellipsoid. The error introduced by using ϕ and h on the local ellipsoid instead of the mean earth ellipsoid, is at most a few tens of mgals in Canada.

Orthometric heights are then derived from $\tilde{\Delta C}$ using again the normal gravity in exactly the same manner as the observed gravity was used. The same holds true for dynamic heights as well.

Generally, we may observe that heights based on normal gravity are meant to approximate the proper heights (based on actual, observed, gravity) and may be thus regarded as lower order heights from the point of view of rigour. They must not be mixed up with normal heights in any way. Canada and the U.S.A. are among the countries where these approximate heights are still used exclusively.

5.7) Discussion of the individual height systems

We have seen (§§ 5.3 - 5.5) that geopotential numbers can be regarded as a basis for all the used height systems. According to the value of gravity by which we divide the geopotential number, we get either

i) dynamic height (C divided by a reference gravity constant for one or more countries);

ii) orthometric height (C divided by the mean actual gravity along the plumb-line of the point taken between the geoid and the surface);

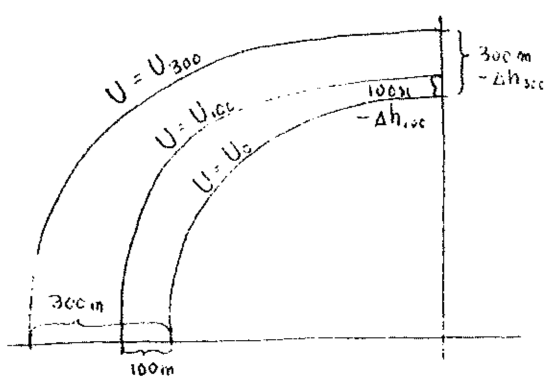
iii) normal height (C divided by the mean normal gravity along the plumb-line of the point, taken between the ellipsoid and a point h_A high above the ellipsoid).

Disregarding the heights based on normal gravity, that can be regarded as a degenerate case of the proper heights, we have thus three distinctly different definitions of heights. They are not, and do not claim to be, equivalent. They represent three different geometrical (physical) quantities.

The dynamic heights are closely related to the concept of equipotential surfaces. One may say that they reflect the geometry of the physical space surrounding us. As we have seen already, the points laying on one equipotential surface have the same dynamic height (for one value of the reference gravity).

The orthometric heights are, what one may call, "the common sense heights". The points that have the same orthometric height have the same vertical distance from the geoid. Because these heights are not meant to depict the physical properties of the space around us they have the unfortunate property that they disregard the water flow; we may have water flowing from a point orthometrically lower to a point orthometrically higher.

In order to see the magnitude of the difference between the orthometric and dynamic heights let us take the normal component of gravity only and draw the following diagram:



The quantities Δh_{100} and Δh_{300} denote the amount of convergence of the normal equipotential surfaces, and therefore also approximately the difference between orthometric and dynamic height. They can be evaluated from following formulae:

$$\Delta h_{100} \doteq (U_{100} - U_0) \left(\frac{1}{\gamma_a} - \frac{1}{\gamma_b}\right)$$

$$\Delta h_{300} \doteq (U_{300} - U_0) \left(\frac{1}{\gamma_a} - \frac{1}{\gamma_b}\right)$$

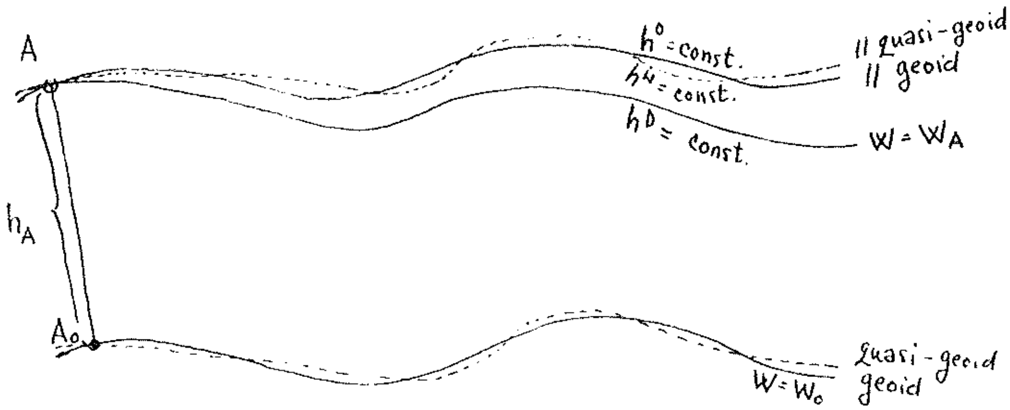
that can be derived from our known relation

$$dU = - \gamma \, dh \ .$$

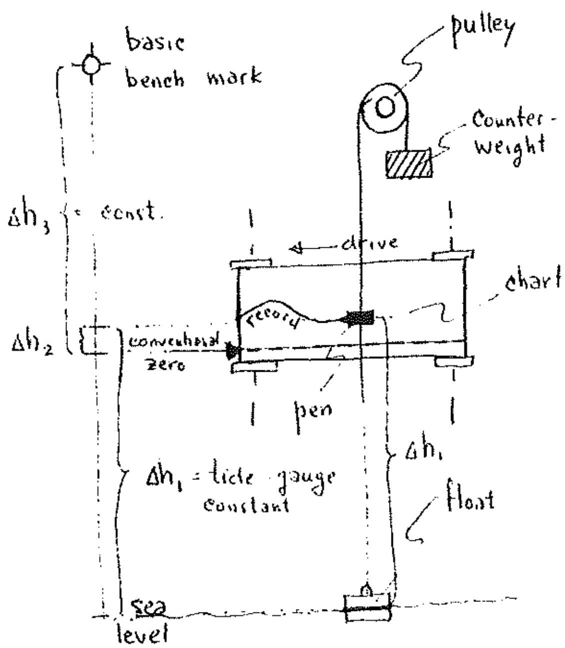
Taking $U_{100} - U_0 \simeq -100$ m $\gamma_a \simeq -97800$ gal·m and $1/\gamma_a - 1/\gamma_b \simeq 5.4 \; 10^{-6}$ gal^{-1}, we get approximately $\boxed{\Delta h_{100} \simeq 53 \text{ cm}, \quad \Delta h_{300} \simeq 160 \text{ cm}.}$

The normal heights are, philosophically speaking, closer to the orthometric heights. The disadvantage though is that the surface they are refered to (quasigeoid) has no physical meaning. On the other hand, there are no hypotheses involved in defining them. For practical purposes the normal heights are just as good as the orthometric.

All three systems define the height of any point uniquely. They may be graphically interpreted as follows:-



5.8) Mean sea level as a height reference



The levelling networks are usually connected to a number of basic bench-marks situated close to tide-gauges. Their heights (generally only a few meters above the sea-level) are determined as $\Delta h_1 - \overline{\Delta h_2} + \Delta h_3$, where Δh_1 is a constant for a tide-gauge, Δh_3 is the level difference between the tide-gauge zero and the bench mark and $\overline{\Delta h_2}$ is the mean chart reading, corresponding to the "mean-sea level". $\overline{\Delta h_2}$ is obtained as a mean of the readings over a period of many years. Then the height of the basic bench-mark can be regarded as directly related to the mean-seal level which is representing the geoid (quasigeoid).

The heights of the individual points of the levelling network are determined by adjustment. First the orthometric (dynamic, normal) height differences of levelling lines are determined, using the levelled differences and the appropriate corrections. Then the network is adjusted, holding the heights of the basic B.M!s fixed.

This procedure of course assumes that

i) the mean chart readings represent the mean sea level that is constant over any period of time;

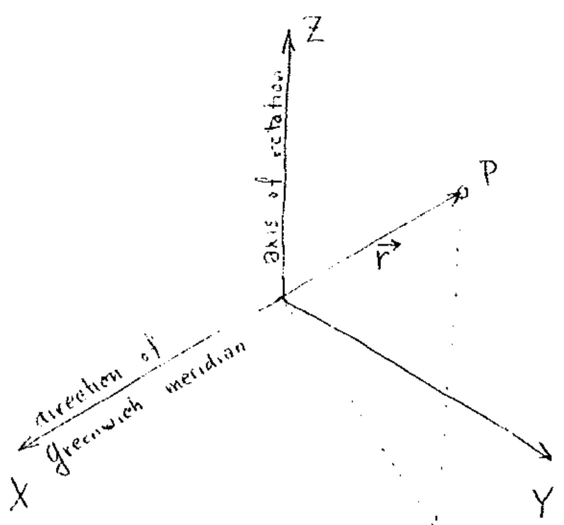
ii) all the tide-gauges refer to the same level -- the geoid (quasigeoid).

Strictly speaking, neither of these two assumptions seem to be valid. The sea level at every point is subject to great many influences. It seems almost certain now that the combination of these influences cause the sea level to rise systematically all over the world. The second assumption appears incorrect because the local conditions (prevailing winds, salinity, temperature, etc.) influence the sea level somewhat permanently.

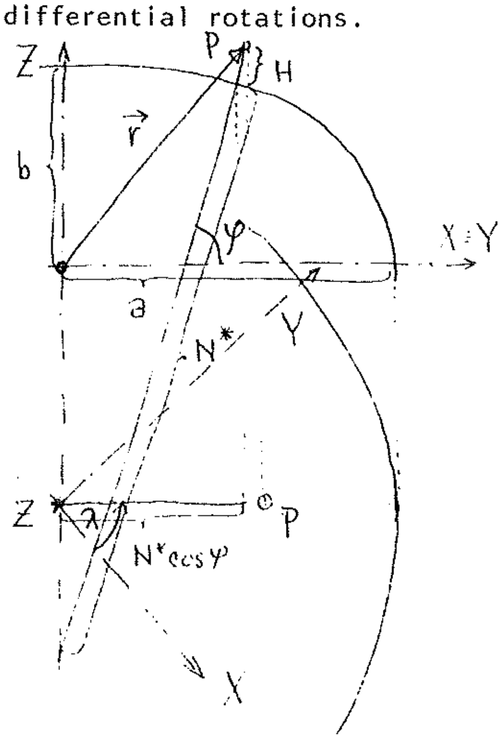
Hence more satisfactory procedure would be to hold only one (reference) point in the network with an assumed height fixed for the adjustment. This reference point should be located close to the center of the net (preferably in a geologically stable area) to achieve the best propagation of errors. After having adjusted the network one can study the heights of the sea levels--as indicated by the tide-gauges--relative to the reference point. From these relative heights, a mean difference between the sea levels and the reference point can be deduced that would be valid for a certain period of time and all the points, including the reference point, can be given the appropriate correction. The height of the reference point then may be declared fixed for a certain period of time.

6) Use of astronomic observations in geodesy

6.1) Geocentric Cartesian, Geodetic and Astronomic coordinate systems and their transformations; astro-deflection



i) The Geocentric Cartesian system can either be average-using the average axis of rotation of the earth--or instantaneous--using the instantaneous axis of rotation for Z axis. It is centered on the center of gravity of the earth and +XZ plane contains the Greenwich observatory. If the axes do not interesect in the center of gravity, we speak about Relative Cartesian system. The transformations between these systems are a matter of translation and differential rotations.



ii) Geodetic Coordinates of a point P are ϕ,λ,H. In order to be able to relate them to the Cartesian coordinates, the reference ellipsoid has to be given as well, usually by its center and the two main axes a and b. The radius vector of the point P is given by

$$\vec{r} = \begin{bmatrix} (N^* + H)\cos\phi\cos\lambda \\ (N^* + H)\cos\phi\sin\lambda \\ (N^* (\frac{b}{a})^2 + H)\sin\phi \end{bmatrix} .$$

Here, the radius of curvature in the prime vertical is

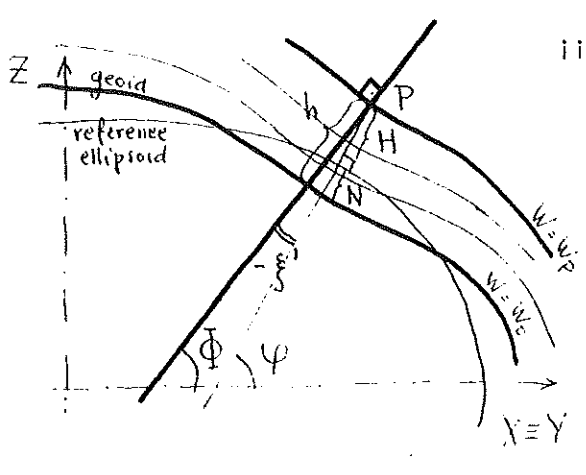
$$N^* = a^2 / \sqrt{(a^2 \cos^2\phi + b^2 \sin^2\phi)}$$

For small height above the reference ellipsoid, H, we get

$$\vec{r} \simeq (N^* + H) \begin{bmatrix} \cos\phi\cos\lambda \\ \cos\phi\sin\lambda \\ (\frac{b}{a})^2 \sin\phi \end{bmatrix} .$$

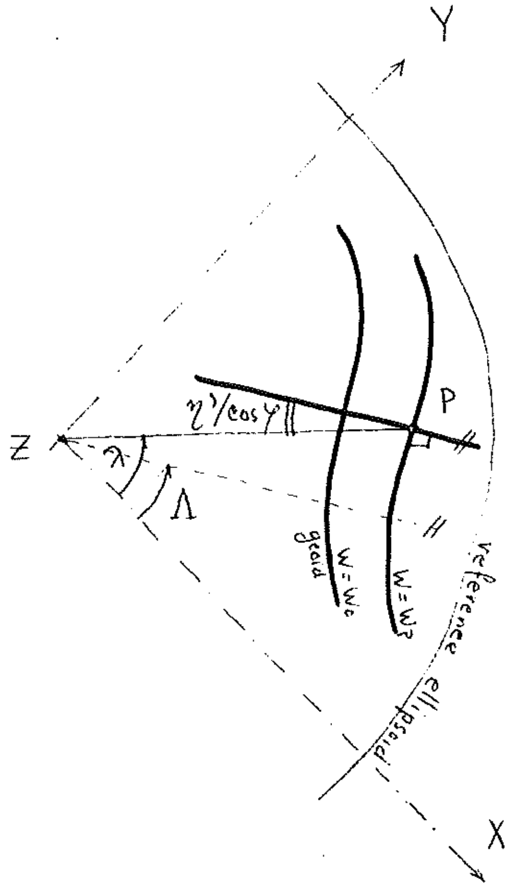
These formulae as well as more details can be found in [Krakiwsky, E. J. & Wells, D. E., 1971: Coordinate Systems in Geodesy, UNB.]

Note that we do not require the center of the reference ellipsoid to coincide with the center of gravity of the earth. On the other hand we usually want its semiminor axis to be parallel with the mean axis of rotation of the earth. The formulae above are, of course, valid only for the two systems (Geodetic and Cartesian) being concentric, and represent the direct transformation of geodetic into Cartesian coordinates. If the two systems are not concentric, then the shift (translation) of the two centres has to be added to the transformation. The inverse transformation is more complicated and is usually solved by an iterative process; see [Krakiwsky & Wells, 1971].



iii) Astronomic coordinates of a point P are Φ, Λ, h. The meaning of these symbols is apparent from the diagrams. We can obviously write

$$\boxed{\begin{aligned} \Phi &= \phi + \xi' \\ h &\doteq H - N. \end{aligned}}$$



The relation between Λ and λ is not so obvious. However, we can write it easily enough realizing that the image of η' in the xy-plane is given by $\eta'/\cos\phi$ and thus

$$\boxed{(\Lambda-\lambda)\cos\phi = \eta'.}$$

The following equation is also valid approximately

$$(\Lambda-\lambda)\cos\Phi \simeq \eta'.$$

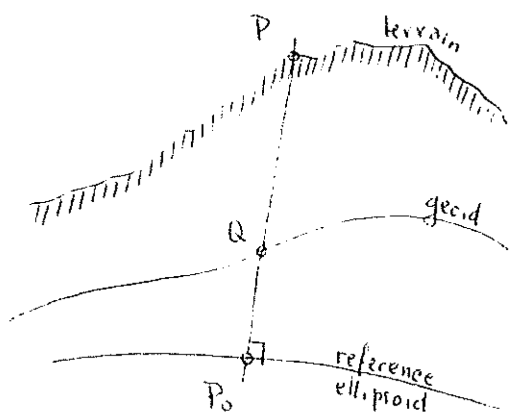
Hence we can transform the local astronomic coordinates to geodetic coordinates only if we know the geoidal undulation N and the two components of the deflection of the vertical ξ', η' <u>at the point P</u>. The geodetic coordinates can subsequently be transformed to the Cartesian coordinates through the reference ellipsoid on which the geodetic coordinates are known. Therefore, we can conclude that in order to transform the astronomic coordinates (that can be regarded as "observed coordinates") to the Cartesian we have to know two surfaces--the reference ellipsoid and the geoid or more precisely the ellipsoid and the gravity field, as given by the equipotential surfaces, between the geoid and the point P.

Note that if Φ is observed astronomically, it may or may not be corrected for the effects of polar wobble. Φ and Λ are in geodetic networks observed on a whole set of points known as <u>deflection points</u>. Obviously, if the astronomic coordinates Φ,Λ of a point are observed and its geodetic coordinates ϕ,λ derived from the terrestrial network (computed on the reference ellipsoid), the relationship of these two pairs of coordinates, can be used to provide us with the components ξ', η' of the deflection of the vertical on the surface of the earth, known as <u>astro-deflection</u> at P. This technique is very widely used in practice.

6.2) Definition of a "corresponding point" on the reference ellipsoid to a surface point; Helmert's and Pizzeti's projections

It is known from geometric geodesy that the reference ellipsoid is the surface on which the adjustment of the horizontal networks is usually carried out. Hence we have to define the projection of the points on the surface of the earth onto the reference ellipsoid or, in other words, we have to define what we mean by a "corresponding point" P_o on the ellipsoid to the point P on the surface.

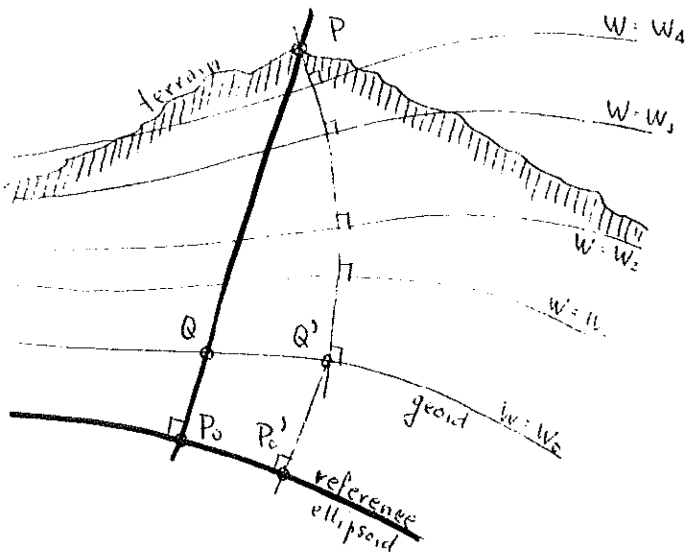
The most obvious way how to define the corresponding point P_o to $P \equiv (\phi, \lambda, H)$ is to take $P_o \equiv (\phi, \lambda, 0)$. This definition is due to the German geodesist Helmert and the projection of P to P_o is hence known as <u>Helmert's projection</u>. Its geometric interpretation is very easy and can be regarded as straightforward generalization of the earlier used definition of corresponding points on the geoid and on the ellipsoid.



If the point P on the surface is determined by its astronomical coordinates Φ, Λ, h, the corresponding point on the ellipsoid is then given by

$$P_o \equiv (\phi = \Phi - \xi', \ \lambda = \Lambda - \eta'/\cos\phi, \ 0)$$

where ξ', η' are the components of the astro-deflection at the point P.



The second definition is due to the Italian geodesist Pizzeti. He suggests to compute first the "corresponding point Q' " on the geoid projecting P along the actual plumb-line. Then the point Q' on the geoid is projected along the ellipsoidal normal to the ellipsoid, to get the corresponding point P_o'. This double projection became known as <u>Pizzeti's projection</u>. Its mathematical expression is evidently more complicated than this of Helmert.

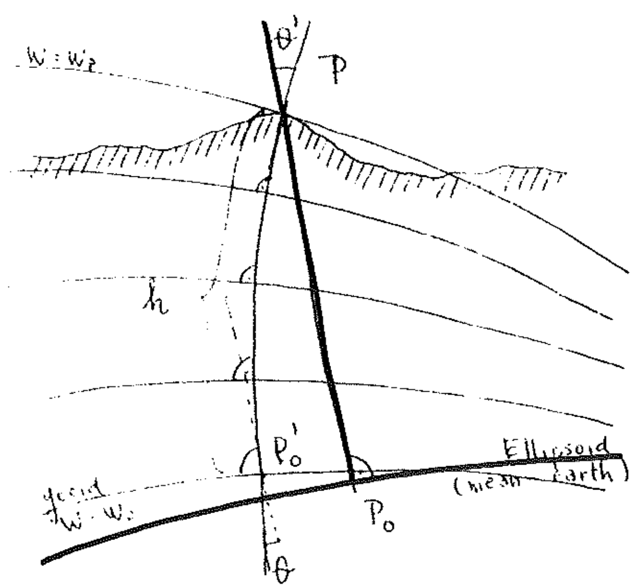
In practice, the Helmert's definition is used almost everywhere mainly because of its simplicity. These two definitions, however, yield almost identical results, the distance between P_o and P_o' being usually no more than a few centimeters. It may occasionally reach the magnitude of about 1 meter.

6.3) Relation between astro-deflection and gravimetric deflection

So far, we have come across two kinds of deflections of vertical. One, introduced in 3.19, described the relationship between the geoid and the mean-earth-ellipsoid. These deflections can be computed from the gravity data and are known as gravimetric or absolute deflections. Their components are denoted here by ξ, η, conforming to the previous notation.

Second kind, the components ξ', η' of which we have met in §6.1, relates the equipotential surface of the point on the terrain with the reference ellipsoid. Since these deflections can be determined from astronomic observations (and geodetic observations and computations) they are usually called astronomic or relative deflections.

If the reference ellipsoid, used for the determination of the astro-deflections happens to be the mean-earth-ellipsoid we would end up with the situation depicted, on the diagram, where θ' denotes the astro-deflection and θ the gravimetric deflection. In such a case the two deflections would differ just by the term $\delta\theta$ due to the curvature of the actual plumb-line. Later, we shall see that when we know the mutual relationship of the mean-earth-ellipsoid and the reference ellipsoid



(used for deriving the astro-deflections) we are able to transform the original astro-deflections to new astro-deflections related to the mean-earth-ellipsoid. In other words, knowing the relationship of the two ellipsoids we are able to talk about the relationship of the two kinds of deflections in the terms used above, since the astro-deflections can be first transformed to the astro-deflections related to the mean ellipsoid. Hence for the rest of this paragraph we shall assume that the astro-deflections are related to the mean-earth ellipsoid.

In order to develop the expressions for the two components of the curvature term let us denote first

$$\Delta\xi = \xi - \xi', \quad \Delta\eta = \eta - \eta', \quad \Delta\theta = \sqrt{(\Delta\xi^2 + \Delta\eta^2)}.$$

Let us take again the local orthogonal system of coordinates x,y,z with z-axis coinciding with the outer normal and x-axis pointing south on the tangent plane to the local equipotential surface (see also

3.14). This particular orientation of x,y,z system is chosen so that it corresponds to the sense of ξ and η. Taking the differential vector increment along the plumb-line

$$\vec{da} = dx\,\vec{i} + dy\,\vec{j} + dz\,\vec{k}$$

parallel to the gravity vector

$$\vec{g} = \nabla W$$

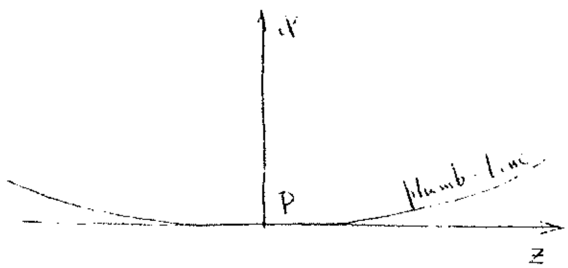
we can write

$$dx/W'_x = dy/W'_y = dz/W'_z$$

where W'_x stands for $\frac{\partial W}{\partial x}$ and similarly W'_y, W'_z. This is the differential equation of the plumb-line.

Since we want to determine the $\Delta\xi, \Delta\eta$; i.e., the curvature terms in the meridian plane and in the plane of the prime vertical, let us investigate the projections of the plumb-line in the xz and yz-planes. Writing for the projection into xz plane

$$x = x(z).$$



Recalling the formula for the curvature of $x = x(z)$:

$$k = x''\,(1 + x'^2)^{-3/2}$$

and realizing that $x' = 0$ due to our choice of the coordinate system (x has got a minimum at P) we can write for the curvature in xz plane

$$k_x = \frac{d^2x}{dz^2}\,.$$

On the other hand, from the differential equation of the plumb-line, we have

$$\frac{dx}{dz} = \frac{W'_x}{W'_z}\,.$$

Taking the derivative of this expression with respect to z we obtain:

$$\frac{d^2x}{dz^2} = \frac{1}{W_z'^2}\left[W'_z\left(W''_{xz} + W''_{xx}\frac{dx}{dz}\right) - W'_x\left(W''_{zz} + W''_{zx}\frac{dx}{dz}\right)\right].$$

But here again $\frac{dx}{dz} = 0$ as above and $W'_x = 0$ because the x-axis is tangent to the equipotential surface W = constant.

Hence

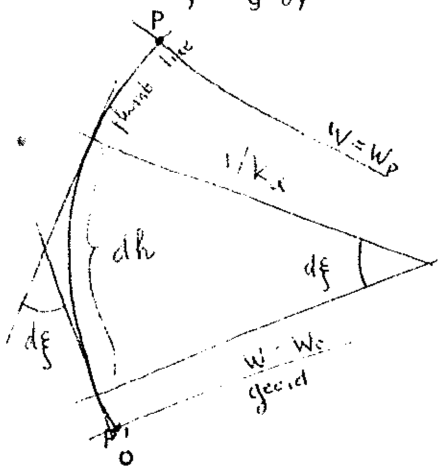
$$k_x = \frac{d^2x}{dz^2} = \frac{W'_z \, W''_{xz}}{W'^2_z} = \frac{W''_{xz}}{W'_z} \ .$$

We know that W is inside the earth analytic so that $W''_{xz} = W''_{zx}$. Realizing that $W'_z = -g$ we can write finally:

$$k_x = \frac{1}{g} \frac{\partial g}{\partial x} \ .$$

Analogously

$$k_y = \frac{1}{g} \frac{\partial g}{\partial y} \ .$$



Denoting by dξ the differential change in the meridian component of the deflection, corresponding to a differential change of height dh (see the diagram), we obtain

$$d\xi = k_x dh = \frac{dh}{g} \frac{\partial g}{\partial x} \ .$$

The total change Δξ in ξ corresponding to the displacement from P to P'_o is then given by

$$\boxed{\Delta\xi = \int_P^{P'_o} \frac{1}{g} \frac{\partial g}{\partial x} \, dh.}$$

Analogously

$$\boxed{\Delta\eta = \int_P^{P'_o} \frac{1}{g} \frac{\partial g}{\partial y} \, dh.}$$

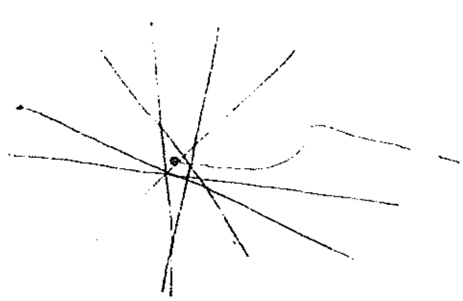
Various formulae suitable for practical evaluation of these quantities may be found in literature. They are based on different hypotheses for the actual gravity inside the earth and we are not going to deal with them here. Let us just recapitulate here that the curvature corrections $\Delta\xi$ and $\Delta\eta$ have to be applied when we want to reduce the "surface deflection" to the "geoidal deflection" or vice versa.

The magnitude of $\Delta\xi, \Delta\eta$ can attain several seconds of arc. It is likely to be higher in mountaineous area, lower in the flat regions. Therefore, in the flat regions, the curvature corrections are often neglected altogether.

6.4) Astronomic and geodetic azimuths; Laplace's equation

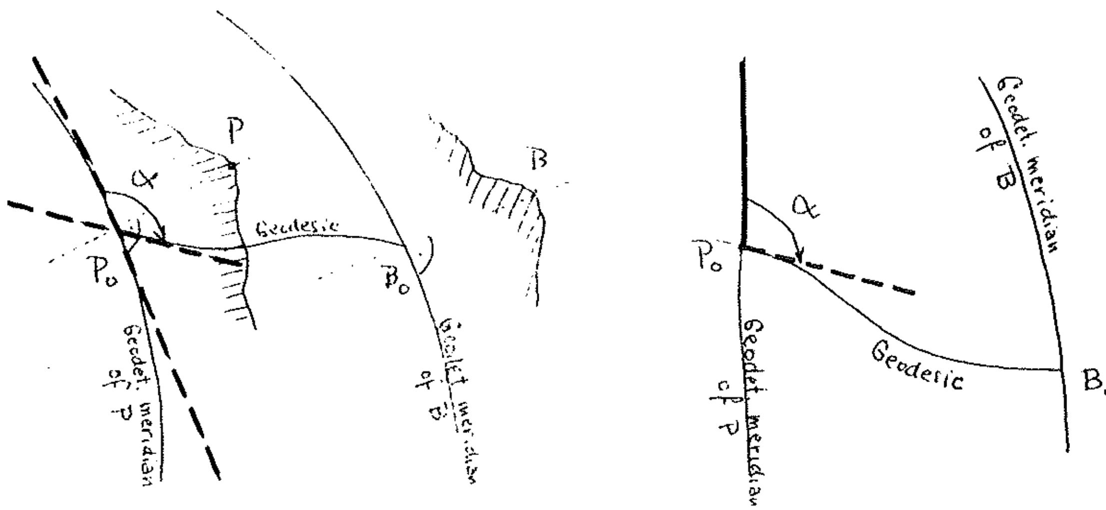
When we deal with a pair of points on the surface of the earth we can define an azimuth of one with respect to the other. In geodesy, we work with two different kinds of azimuths - astronomic and geodetic.

Astronomic azimuth A of a point B with respect to point P is defined as the angle between two planes: the local astronomic meridian plane of P and the vertical plane containing the point B. The local astronomic meridian plane is given by one line-local vertical; i.e., the tangent line to the local plumb-line at P - and one point - the infinitely distant point on the average axis of rotation as defined by the position among the star formations. The vertical plane containing the point B is then defined by the local vertical as above and the point B on the surface of the earth. We may note that the astronomic meridian planes do not generally contain the average pole of rotation on the surface of the earth. The bunch of the astro-meridian planes intersecting a unit sphere centered upon the earth center of gravity may look thus



average axis of rotation.

Geodetic azimuth α of a point B with respect to point P is defined as the angle between two lines: the projection of the geodetic meridian into the tangent plane to the reference ellipsoid at P_o (corresponding to P) and the tangent to the projection of the geodetic curve connecting P_o with B_o (corresponding to B), on the ellipsoid, into the same tangent plane. The geodetic meridian



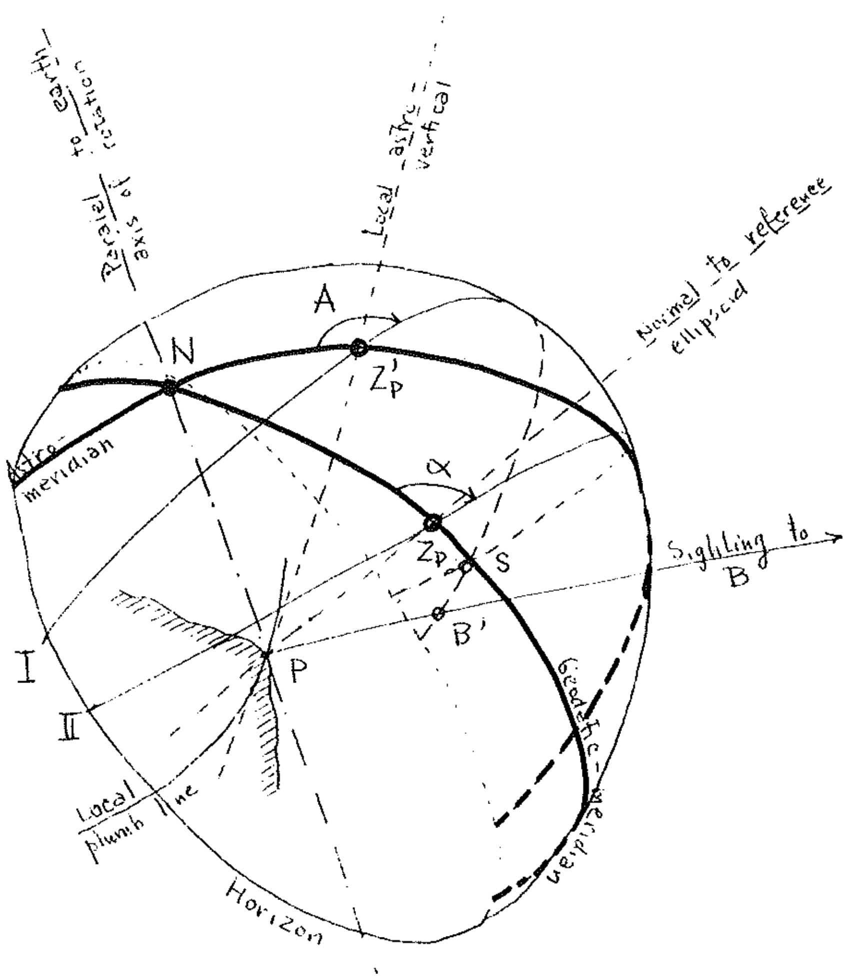
of P (and also P_o) is the geodesic on the reference ellipsoid containing both ellipsoidal poles and the point P_o. Note that the geodetic meridian plane does not, in general, contain the earth axis of rotation, when the reference ellipsoid is not co-axial with the earth.

To establish the relationship between the astronomic and geodetic azimuths, let us first assume that the axis of rotation of the earth is parallel (not coincident) with the semi-minor axis of the reference elliposid. We can then take a unit sphere centered upon the point P. The diagram shows the various points and circles we get on the sphere:

$\underline{Z_p}$ - the local ellipsoidal zenith;

$\underline{Z'_p}$ - the local astronomical zenith;

$\underline{N}$ - image of the North pole; i.e., intersection of the Northern branch of the line parallel with the earth axis of rotation going through P, with the sphere;



$\underline{B'}$ - projection of the point B onto the sphere;

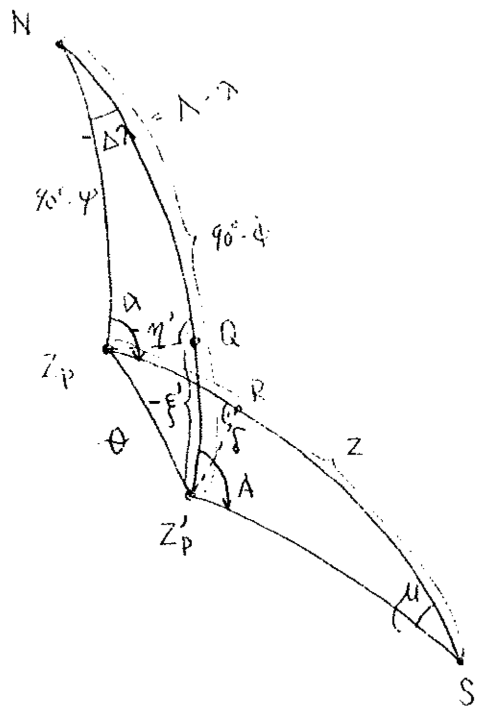
$\underline{S}$ - intersection of the planes I and II on the sphere.

The plane denoted by I contains the local astronomic vertical (tangent to the local plumb-line at P) and the point B. Plane II contains the local normal to the ellipsoid and the tangent to the geodesic P_oB_o.

Redrawing, the two triangles we are interested in, namely $NZ_pZ'_p$ and Z'_pZ_pS, we obtain the following formation. Applying the Napier's rule to the rectangular triangle NZ_pQ we get:

$$\sin\phi = \cos\,(90° - \Phi + \xi')\,\cos\,\eta'$$

$$-\sin\eta' = \cos\,(90° + \Delta\lambda)\,\cos\phi .$$



Realizing that $\xi', \eta', \Delta\lambda$ are very small in absolute value we can write

$\sin\phi \doteq \sin(\Phi - \xi')$

or

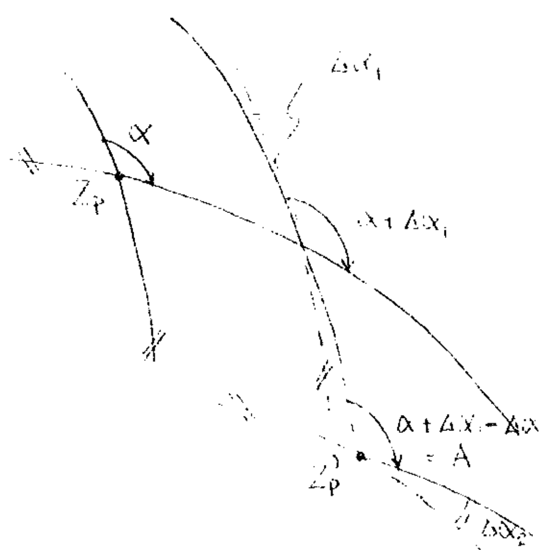
$\phi \simeq \Phi - \xi'$,

and

$\eta' \simeq \sin\Delta\lambda \cos\phi \simeq \Delta\lambda\cos\phi$.

These equations coincide with the equations in §6.1 defining the astronomic deflection which in fact may be regarded as a proof of correctness of the representation of the triangle $NZ_pZ'_p$ in the way we have done it.

On the other hand, the central part of the last diagram can be redrawn and denoted thus.



Then we obtain

$\Delta\alpha_1 \doteq \Delta\lambda \cos(90° - \Phi)$

$\simeq \Delta\lambda \sin\phi$

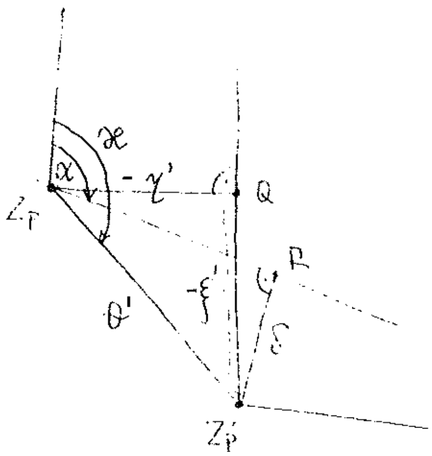
and

$\Delta\alpha_2 \simeq \mu \cos Z$.

We can thus write:

$A - \alpha = \Delta\alpha_1 - \Delta\alpha_2 \simeq \Delta\lambda \sin\phi - \mu\cos Z$.

Taking the triangles $QZ_pZ'_p$, $RZ_pZ'_p$ and projecting them on the plane tangent to the sphere at, say Z_p we get the following diagram. Here the lines can be drawn straight because of the extremely small size of the formation.



From the triangle $QZ_pZ'_p$ we obtain:

$$-\xi' = \theta' \sin(\kappa - 90°)$$

or

$$\xi' = \theta' \cos\kappa$$

and

$$-\eta' = \theta' \sin\kappa.$$

These formulae relate the deflection θ' and its azimuth κ to the deflection components ξ', η'. Note that $\theta' = \sqrt{(\xi'^2 + \eta'^2)}$ and $\operatorname{tg}\kappa = \eta'/\xi'$.

On the other hand, we can write, using the triangle $RZ_pZ'_p$:

$$\delta = \theta' \sin(\kappa - \alpha) = \theta' (\sin\kappa \cos\alpha - \cos\kappa \sin\alpha)$$

$$= -\eta' \cos\alpha - \xi' \sin\alpha.$$

But, from the spherical triangle SRZ'_p we get:

$$\sin\delta = \sin Z \sin\mu$$

or, considering δ and μ very small:

$$\mu \simeq \delta/\sin Z = -(\eta' \cos\alpha + \xi' \sin\alpha)/\sin Z.$$

Thus, we can finally write for $A-\alpha$:

$$\boxed{A - \alpha \simeq (\Lambda - \lambda)\sin\phi + (\xi' \sin\alpha + \eta' \cos\alpha)\operatorname{cotg} Z.}$$

This is the well known Laplace's equation (do not mix it up with the other Laplace's equation $\Delta V = 0$) in its full form. It expresses the relationship between the astronomic and geodetic azimuths through other quantities.

We can note that the second term contains Z, that depends on the zenith distance of the sighting PB as well as the deflection at the station. and the curvature of its plumb-line. Its precise evaluation is very difficult. Z, in geodetic literature, is usually assumed to equal just the mentioned zenith distance, which may cause some anxiety. Application of the right corrections to the observed astronomical azimuths, namely the "skew normal" correction, the correction "to geodesic", as known from geometric geodesy, and the correction for the "curvature of the actual plumbline", eliminates this possible source of errors. These corrections are, however, seldom applied in practice. For astronomical purposes, the second term is usually disregarded altogether.

The azimuth (geodetic) determined from

$$\alpha' = A - (\Lambda - \lambda) \sin \phi$$

is known as Laplace's azimuth. A point, for which the Laplace's azimuth was determined is then called a Laplace's point. The Laplace's points are used for i) orienting the horizontal networks on the reference ellipsoid;

ii) orienting the reference ellipsoid so that its semi-minor axis becomes parallel to the earth axis of rotation. The straightforward least squares solution minimizing the summation of the squares $(\alpha' - \alpha)^2$ is usually deployed for this purpose.

Note that the simplified Laplace's equation allows us to find another formula for η' yet. Realizing that $\Lambda - \lambda = \eta'/\cos \phi$ (see earlier), we derive

$$\eta' \approx (A - \alpha) \operatorname{cotg} \phi$$

relating the E-W component of the astro-deflections to the difference of the two azimuths.

6.5) Astrogeodetic determination of the geoid (astronomic levelling)

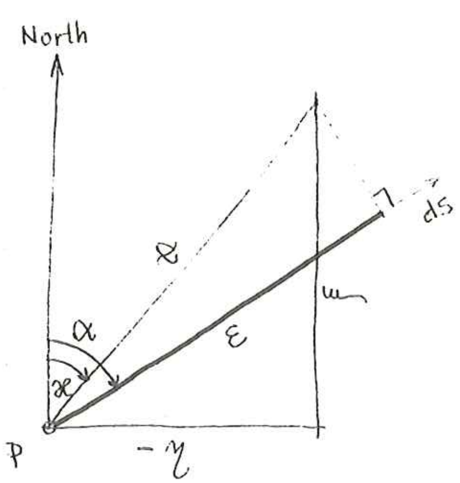
As we have seen in § 3.1~~8~~9, there is a close relationship between the deflections of vertical on the geoid and the geoidal undulations. We have seen that following equation holds:

$$dN = - \varepsilon ds$$

where ε is the deflection component in the azimuth α of ds. The sign is minus because of the sign convention for ξ and η.

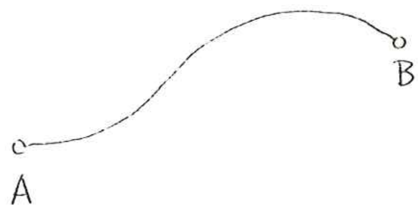
Obviously, if the deflection components on the geoid refer to a reference ellipsoid (mean earth or local) we can use them for determining the dN which are referred to the same reference ellipsoid. Hence the astrodeflections can be used for the purpose of determining the increments of geoidal height with respect to the used reference ellipsoid when corrected for the actual plumb-line curvature between the surface and the geoid. In other words, correcting the astro-deflection components for the curvature term (see 6.3) they can be used in the above formula giving the variations of the geoidal height with respect to the reference ellipsoid.

The quantity ε can be determined from the components ξ,η and the azimuth α of the line segment ds from the diagram as



$$\varepsilon = \xi \cos\alpha + \eta \sin\alpha \text{ .}$$

The idea of the astrogeodetic determination of the geoid, due to Helmert again, is based on the following. Consider the astro-deflection component known along a given line $\overline{AB}$ on the surface. Providing, we can reduce these components to the geoid

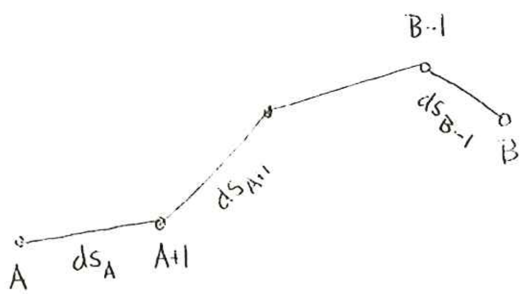


(by applying the curvature corrections) and providing the geoidal height at A, N_A, is known we can determine the geoidal height of B from the evident formula

$$N_B = N_A + \int_A^B dN = N_A - \int_A^B (\xi \cos\alpha + \eta \sin\alpha)\, ds .$$

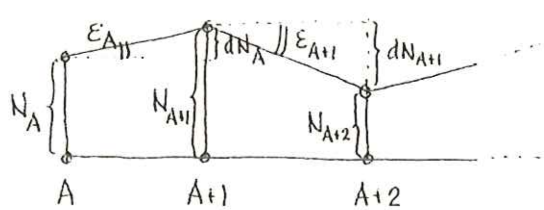
In practice, we do not know the astro-deflections continuously along the profile but providing the points with known deflections are spaced densely enough we can replace the integral by summation and write

$$\boxed{N_B \simeq N_A - \sum_{i=A}^{B-1} (\xi_i \cos\alpha_i + \eta_i \sin \alpha_i)\, ds_i \ .}$$



This formula corresponds to the case shown on the diagrams.

It is usual in practice, to design closed loops consisting of such piece-wise lines. Then the geoidal undulations can be determined for all the involved points and the loops, or network of such loops, adjusted in much the same fashion as levelling loops. The results are the more reliable the closer are the astro-deflection points together.



We may conclude by noting that while the Stokes' formula furnishes the geoidal undulations above the mean earth ellipsoid (once the proper scale is determined), the <u>astrolevelling</u>, as this method became known, provides us with the geoidal heights above the reference ellipsoid; i.e., the ellipsoid used for computing the geodetic coordinates on. The geoidal profiling (or loops) starts usually at the origin of the reference ellipsoid, where the geoidal undulation is assumed (generally assumed to be zero).

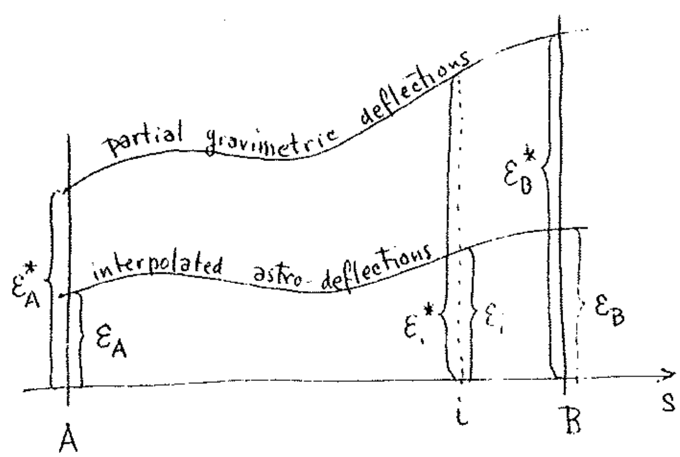
6.6) Astro-gravimetric determination of the geoid (astro-gravimetric levelling)

The astro-geodetic determination of geoidal profiles, viewed mathematically, represents an example of the use of the trapezoidal formula for evaluating the involved finite integral. The weakness of the technique lies in the linear interpolation between any two adjacent astro-deflection points. Hence, as we have mentioned already, it is an imperative there to have the deflection-points spaced as closely as possible. This, of course, represents a very costly requirement.

One way, how to overcome the necessity of having too many deflection-points is to use the knowledge of the local gravity field to supplement the deflection points. As we know (see § 3.13) it is possible to obtain the absolute (gravimetric) deflection at any point on the geoid from the gravity anomalies, using the Vening-Meinesz's formulae. Taking the Vening-Meinesz's formulae for ξ and η, we can split the integration involved there into two terms: i) integration over the vicinity of the point of interest (up to a few degrees or few hundred miles); ii) integration over the rest of the terrestrial globe. The first term when computed for densely spaced points along a profile will reflect the "local effects"; i.e., the effect of the immediate environment of the points. It will therefore vary irregularly. The second term will vary only very slowly and its variations between any two adjacent deflection points may be taken as approximately linear.

Hence, the first term of the Vening-Meinesz's formulae can be used for densification of deflections along a profile connecting two adjacent astro-deflection points. The fact that the Vening-Meinesz's formulae yield gravimetric deflections; i.e., deflections related to the mean-earth ellipsoid, is immaterial. The difference between astronomic and gravimetric deflections can certainly be regarded as practically linear in between any two adjacent deflection points and has therefore approximately the same influence on the results as the second term in the Vening-Meinesz's formulae.

One way how to use the "partial" gravimetric deflections (given by the first term) would be to compute their values ε^* along the profile AB including the points A and B. Assuming, that the difference between ε's and ε^*'s varies linearly we can write for the interpolated astro-deflection ε_i



$$\boxed{\varepsilon_i = \varepsilon_i^* + \varepsilon_A - \varepsilon_A^* + \frac{S_{Ai}}{S_{AB}} (\varepsilon_A^* - \varepsilon_A - \varepsilon_B^* + \varepsilon_B)}$$

which the reader can easily verify. This idea is due to Molodenskij and the method became known as astro-gravimetric determination of the geoid.

It is not difficult to see that quadratic (or higher order) interpolation can be used instead of the linear interpolation. In such a case we would have to consider three (or more) astrodeflection points along the profile. Other sources of information can be used for interpolating between or amongst the astro-deflection points as well. We could, theoretically use, for instance, the values of the horizontal gradient of gravity, zenith distances observed between two adjacent detail points on the profile or just the computed values of isostatically corrected terrain attraction.

We may note that the most serious difficulty with the astro-gravimetric levelling is again the necessity to account for the effect of immediate surroundings (up to a few miles) of the interpolated point. This may require some gravity densification at certain points.

To conclude with, let us point out that both these methods (astro-geodetic and astro-gravimetric) suffer from one unfortunate property. The incertitude in the geoidal height increases as we go away from the point of origin due to the build-up of random and systematic errors inherently contained in the astro-deflections. Hence, perhaps, a least-squares fitting of a surface to the astro-deflections (and/or the densified deflections) might provide a better answer than the profiles.

6.7) Determination of an optimum reference ellipsoid from the deflections of vertical and geoidal heights*

6.7.1) Relation between the change of the reference ellipsoid and the deflections and geoidal heights

When dealing with deflections and geoidal heights we always have to bear in mind the fact that they are related to the specific reference ellipsoid. Hence, if we change the ellipsoid all the deflections and the geoidal heights change as well (since the geoid remains the same for all our work).

Talking about the change of the reference ellipsoid we talk about three distinctly different changes:

i) the change of the center of the ellipsoid with respect to the center of gravity of the earth; i.e., translation;

ii) the change of the orientation given by three rotation angles;

iii) the change of the shape and size of the ellipsoid; i.e., the change of the semi-major axis and the flattenning or any two parameters describing the shape and size of an ellipsoid of rotation.

All in all, there are 8 parameters involved.

Fortunately, the earth is almost spherical and therefore the reference ellipsoids are usually almost spherical too (the flattenning of the order of 1/300 does not flatten the sphere very much). Since the orientation of the reference ellipsoid is generally quite good and because the orientation of an almost spherical ellipsoid plays a minor role anyway, the rotation is usually not taken into account at all. Thus, by a change of the reference ellipsoid we shall understand the change of the remaining 5 parameters -- the three translation components and the two shape and size parameters.

As we know from geometric geodesy [Krakiwsky, and **Wells**, 1971], the assumed reference ellipsoid is usually not given by its center and its shape and size parameters. It is generally defined by the geoidal deflection components ξ_o, η_o and the geodial undulation N_o at the point of origin (Meade's Ranch for the NAD 27, Pulkovo for the Krasovski's ellipsoid, Potsdam for the European datum etc.) and its shape and size parameters. Thus, instead of talking about parameters x_o, y_o, z_o, a, f, we shall talk about ξ_o, η_o, N_o, a, f, taking the proper orientation for granted.

The formula expressing the changes in the geodetic coordinates ϕ, λ, H of a point P in relation to the change of x_o, y_o, z_o, a, f, has been derived in [Krakiwsky & Wells, 1971]. It reads

$$\begin{bmatrix} \delta\phi \\ \delta\lambda \\ \delta H \end{bmatrix} = A \begin{bmatrix} \delta x_o \\ \delta y_o \\ \delta z_o \end{bmatrix} + B \begin{bmatrix} \delta a \\ \delta f \end{bmatrix} \qquad (*)$$

where

$$A = \frac{1}{a \cos\phi} \begin{Vmatrix} \cos\phi \sin\phi \cos\lambda\ , & \cos\phi \sin\phi \sin\lambda\ , & -\cos^2\phi \\ \sin\lambda\ , & -\cos\lambda\ , & 0 \\ -a\cos^2\phi \cos\lambda\ , & -a\cos^2\phi \sin\lambda\ , & -a \sin\phi \cos\phi \end{Vmatrix}$$

and

$$B = \begin{Vmatrix} 0\ , & 2 \sin\phi \cos\phi \\ 0\ , & 0 \\ 1\ , & a \sin^2\phi \end{Vmatrix} .$$

Here, all the elements of the matrices A and B are related to the point P. To derive the relation between the change of ξ_o, η_o, N_o, a, f and the change of ξ, η, N at an arbitrary point, let us first establish the relation between $\delta\phi_o, \delta\lambda_o, \delta H_o$ on one side and $\delta x_o, \delta y_o, \delta z_o, \delta a, \delta f$ on the other side. Obviously, this relationship is just a special case of the relationship introduced above and we can write:

$$\begin{bmatrix} \delta\phi_o \\ \delta\lambda_o \\ \delta H_o \end{bmatrix} = A_o \begin{bmatrix} \delta x_o \\ \delta y_o \\ \delta z_o \end{bmatrix} + B_o \begin{bmatrix} \delta a \\ \delta f \end{bmatrix} . \qquad (**)$$

Here the matrices A_o, B_o are nothing else but A,B related to the point of origin.

Now we can get rid of the translation components $\delta x_o, \delta y_o, \delta z_o$ that we shall not be interested in. For this purpose, let us multiply the equation (*) by A^{-1} from the left, the equation (**) by A_o^{-1} from the left and subtract (**) from (*). We get:

$$A^{-1} \begin{bmatrix} \delta\phi \\ \delta\lambda \\ \delta H \end{bmatrix} - A_o^{-1} \begin{bmatrix} \delta\phi_o \\ \delta\lambda_o \\ \delta H_o \end{bmatrix} = A^{-1} B \begin{bmatrix} \delta a \\ \delta f \end{bmatrix} - A_o^{-1} B_o \begin{bmatrix} \delta a \\ \delta f \end{bmatrix} .$$

This equation can be rewritten as

$$\begin{bmatrix} \delta\phi \\ \delta\lambda \\ \delta H \end{bmatrix} = AA_o^{-1} \begin{bmatrix} \delta\phi_o \\ \delta\lambda_o \\ \delta H_o \end{bmatrix} + (B - AA_o^{-1} B_o) \begin{bmatrix} \delta a \\ \delta f \end{bmatrix}$$

$$= C \begin{bmatrix} \delta\phi_o \\ \delta\lambda_o \\ \delta H_o \end{bmatrix} + D \begin{bmatrix} \delta a \\ \delta f \end{bmatrix} .$$

Let us recall, at this stage, the definition of the geoidal deflection components and the ellipsoidal height H. We know that

$$\xi = \xi' + \Delta\xi = \Phi - \phi + \Delta\xi \ ; \quad \eta = \eta' + \Delta\eta = (\Lambda-\lambda) \cos \phi + \Delta\eta \ ;$$

$$H = N + h.$$

Here the quantities Φ, Λ, h (astronomic coordinates) and $\Delta\xi,\Delta\eta$ (curvature corrections) have nothing to do with the reference ellipsoid and will therefore not be influenced by the change of the ellipsoid. We can therefore write:

$$\delta\xi = -\delta\phi, \quad \delta\eta = -\delta\lambda \cos \phi, \quad \delta N = \delta H.$$

Substituting these results back into the transformation equation above we get finally:

$$\begin{bmatrix} \delta\xi \\ \delta\eta \\ \delta N \end{bmatrix} = \tilde{C} \begin{bmatrix} \delta\xi_o \\ \delta\eta_o \\ \delta N_o \end{bmatrix} + \tilde{D} \begin{bmatrix} \delta a \\ \delta f \end{bmatrix}$$

where $\tilde{C}$ and $\tilde{D}$ can be found to equal:

$$\tilde{C} = \begin{Vmatrix} \cos\phi_o \cos\phi + \sin\phi_o \sin\phi \cos(\lambda-\lambda_o), & -\sin\phi \sin(\lambda-\lambda_o), & -\frac{1}{a}(\sin\phi_o \cos\phi - \cos\phi_o \sin\phi \cos(\lambda-\lambda_o)) \\ \sin\phi_o \sin(\lambda-\lambda_o), & \cos(\lambda-\lambda_o), & \frac{1}{a}(\cos\phi_o \sin(\lambda-\lambda_o)) \\ -a[\cos\phi_o \sin\phi + \sin\phi_o \cos\phi \cos(\lambda-\lambda_o)], & -a\cos\phi \sin(\lambda-\lambda_o), & (\sin\phi_o \sin\phi + \cos\phi_o \cos\phi \cos(\lambda-\lambda_o)) \end{Vmatrix}$$

$$\tilde{D} = \begin{Vmatrix} -\frac{1}{a}(\sin\phi_o \cos\phi - \cos\phi_o \sin\phi \cos(\lambda-\lambda_o)), & -\sin^3\phi_o \cos\phi + \cos\phi_o \sin^2\phi_o \sin\phi \cos(\lambda-\lambda_o) - 2\cos\phi(\sin\phi - \sin\phi_o) \\ \frac{1}{a}\cos\phi_o \sin(\lambda-\lambda_o), & \cos\phi_o \sin(\lambda-\lambda_o) \sin^2\phi_o \\ (\sin\phi_o \sin\phi + \cos\phi_o \cos\phi \cos(\lambda-\lambda_o) - 1), & \sin^3\phi_o \sin\phi + \cos\phi_o \sin^2\phi_o \cos(\lambda-\lambda_o) + \sin^2\phi_o - 2\sin\phi_o \sin\phi \end{Vmatrix}$$

This formula bears the name of Vening-Meinesz, who was among the first to derive it, and it is one of the most important formulae in geodesy. It relates the changes of the deflection components (surface or geoidal) and geoidal undulations to the changes of the reference ellipsoid represented by $\delta\xi_o, \delta\eta_o, \delta N_o, \delta a, \delta f$ as we set originally to derive.

6.7.2) Determination of the local optimum reference ellipsoid

The deflections of vertical (and the geoidal undulations) provide us with the geometric relationship between the two surfaces -- the reference ellipsoid and the geoid. They can therefore be used for two different purposes: i) assuming the reference ellipsoid known, to determine the geoid. We have treated this problem in 6.5 and 6.6 already.

ii) assuming the geoid, to determine the best fitting (optimun) reference ellipsoid.

The optimum reference ellipsoid for a certain area can be defined in a variety of possible ways. The two most widely used definitions of optimum ellipsoids are:

1) the best fitting in the sense that it minimizes the sum of squares of geoidal deflections known in the area;

2) the best fitting in the sense that it minimizes the sum of squares of geoidal undulations determined in the area.

Obviously, in both cases, the formulae developed in 6.7.1 will come useful.

For the first optimum ellipsoid, we can write

$$\begin{bmatrix} \tilde{\xi} \\ \tilde{\eta} \end{bmatrix} = \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} \delta\xi \\ \delta\eta \end{bmatrix} = \tilde{C}_R \begin{bmatrix} \delta\xi_o \\ \delta\eta_o \\ \delta N_o \end{bmatrix} + \tilde{D}_R \begin{bmatrix} \delta a \\ \delta f \end{bmatrix} + \begin{bmatrix} \xi \\ \eta \end{bmatrix}$$

where $\tilde{\xi}, \tilde{\eta}$ are the "new" deflection components obtained from the "old" ξ, η component after changing the "old" ellipsoid $(\xi_o, \eta_o, N_o, a, f)$ to the "new" ellipsoid $(\xi_o+\delta\xi_o, \eta_o+\delta\eta_o, N_o+\delta N_o, a+\delta a,\ f+\delta f)$ and $\tilde{C}_R$, $\tilde{D}_R$ are the reduced matrics $\tilde{C}, \tilde{D}$ from § 6.7.1 containing the first two rows only (we are not dealing with δN in this case!). Denoting the first row of $\tilde{C}$ by $\tilde{C}_1$ and the first row of $\tilde{D}$ by $\tilde{D}_1$ we can write

$$\boxed{\xi + \delta\xi = \tilde{C}_1 \begin{bmatrix} \delta\xi_o \\ \delta\eta_o \\ \delta N_o \end{bmatrix} + \tilde{D}_1 \begin{bmatrix} \delta a \\ \delta f \end{bmatrix} + \xi}$$

and similarly

$$\boxed{\eta + \delta\eta = \tilde{C}_2 \begin{bmatrix} \delta\xi_o \\ \delta\eta_o \\ \delta N_o \end{bmatrix} + \tilde{D}_2 \begin{bmatrix} \delta a \\ \delta f \end{bmatrix} + \eta\ .}$$

From these linear observation equations we can determine the adjusted increments $\delta\xi_o, \delta\eta_o, \delta N_o, \delta a, \delta f$ that render the expression

$$\boxed{\sum_{i=1}^{n} [(\xi_i + \delta\xi_i)^2 + (\eta_i + \delta\eta_i)^2]}$$

minimum. Then the best fitting reference ellipsoid for the area is obtained by adding the adjusted increments to the "old" ellipsoid parameters.

It is not difficult to see that if the optimum reference ellipsoid in the sense of the second definition is wanted we can write immediately

$$\boxed{N + \delta N = \tilde{C}_3 \begin{bmatrix} \delta\xi_o \\ \delta\eta_o \\ \delta N_o \end{bmatrix} + \tilde{D}_3 \begin{bmatrix} \delta a \\ \delta f \end{bmatrix} + N}$$

and proceed to get the adjusted increments minimizing the expression

$$\boxed{\sum_{i=1}^{n} (N + \delta N)^2}\ .$$

It is worth noting that for a limited area the two mentioned optimum ellipsoids will not be the same. For a limited area one has to be also aware of the fact that δN and δa are almost linearly dependent (due to the almost spherical shape of the ellipsoid sought) and the matrix of normal equations for the adjusted increments is likely to be poorly conditioned.

6.7.3) Determination of the mean earth ellipsoid

The same technique; i.e., the minimization of either $\Sigma\ (N_i + \delta N_i)^2$ or $\Sigma\ [(\xi+\delta\xi)^2 + (\eta+\delta\eta)^2]$, can be applied to the whole earth in order to determine the best fitting ellipsoid to the whole globe. Such an ellipsoid is usually called the mean earth or absolute ellipsoid. It has been shown that for the whole earth, the two above conditions are equivalent and yield therefore the same ellipsoid.

The question may arise now as to whether the "gravimetric geoid", given by the Stoke's formula can be used for the purpose of determining the mean earth ellipsoid as well. The answer to this question is positive. In this approach we generally define the mean earth ellipsoid as such an ellipsoid that shares with the geoid the mass M, the potential W_o, angular velocity ω and the "difference" of the principal moments of inertia $C - \frac{1}{2}\ (A + B)$. It has again been shown that these physical quantities not only determine one ellipsoid only but also that this ellipsoid is identical with the one defined by the adjustment process as described above.

An interesting and often used alternative way of determining the mean earth ellipsoid is the use of both the gravimetric and the astro-deflections (reduced to the geoid). They should obviously be the same theoretically, if the astro-deflections were computed on the mean earth ellipsoid. Due to the fact, that they are usually computed on the local reference ellipsoid and due to the inevitable errors, thought random, the two deflections do not coincide. This idea can be thus exploited to formulate the adjustment model seeking the corrections to the parameters of the used reference ellipsoid that would minimize the expression

$$\Sigma\ [(\xi-\xi^*)^2 + (\eta-\eta^*)^2]$$

where ξ,η are the astro-deflection and ξ^*,η^* are the gravimetric deflection components. Theoretically, the ellipsoid arrived at using this technique, should again be the same as the previous two.

7) Miscellaneous topics

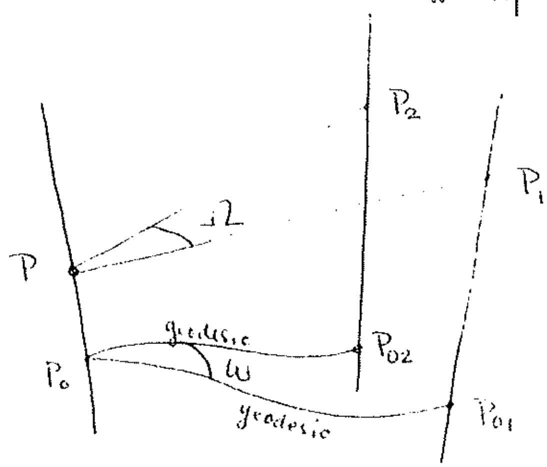
7.1) Reduction of observed horizontal angles to the reference ellipsoid

As we know already, our horizontal and vertical networks are referred to two different surfaces. While the vertical network (heights) are related to the geoid or quasi-geoid, the horizontal network is referred to the used reference ellipsoid. The heights can theoretically be related to the ellipsoid if we know the heights of the geoid (quasi-geoid) above the ellipsoid; i.e., the geoidal (quasi-geoidal) undulations. Hence, in order to be able to talk about relative positions of individual points in three-dimensional space, we have to know the relative position of the geoid with respect to the ellipsoid or vice versa.

In order to be able to treat both the networks on the ellipsoid we have to project the individual points from the surface of the earth to the ellipsoid first. This can be done using one of the projections described in §6.2. At the same time we have to project onto the ellipsoid even the individual observed elements, the horizontal angles and the distances, to be able to compute the positions of the trigonometric points, whose astronomic coordinates are not observed, on the ellipsoid. In this paragraph we shall show, how to project (reduce) the horizontal angles.

To reduce an observed horizontal angle from the terrain to the ellipsoid, we can use the formula developed in §6.4 for the relationship between the astronomic and geodetic azimuths. It is conceivable that the observed angle Ω can be expressed as the difference of two astronomic azimuths

$$\Omega = A_1 - A_2$$



while the reduced angle on the ellipsoid as the difference between the corresponding geodetic azimuths

$$\omega = \alpha_1 - \alpha_2 \ .$$

Hence, denoting by $\Delta\omega$ the difference $\Omega - \omega$ we can write:

$$\begin{aligned}\Delta\omega &= (A_1 - A_2) - (\alpha_1 - \alpha_2)\\ &= (A_1 - \alpha_1) - (A_2 - \alpha_2)\\ &= (\Lambda - \lambda)\sin\phi + (\xi' \sin\alpha_1 + \eta' \cos\alpha_1)\operatorname{cotg} Z_1 - (\Lambda - \lambda)\sin\phi -\\ &\quad - (\xi' \sin\alpha_2 + \eta' \cos\alpha_2)\operatorname{cotg} Z_2 \ .\end{aligned}$$

Substituting here the expressions (§ 6.4) for ξ' and η'

$$\xi' = \theta' \cos\kappa \ , \ -\eta' = \theta' \sin\kappa$$

we get

$$\Delta\omega = \theta' (\cos\kappa\sin\alpha_1 - \sin\kappa\cos\alpha_1)\operatorname{cotg} Z_1 - \theta'(\cos\kappa\sin\alpha_2 - \sin\kappa\cos\alpha_2)\operatorname{cotg} Z_2$$

$$\boxed{\Delta\omega = \theta' [\sin(\alpha_1 - \kappa)\operatorname{cotg} Z_1 - \sin(\alpha_2 - \kappa)\operatorname{cotg} Z_2]}$$

where

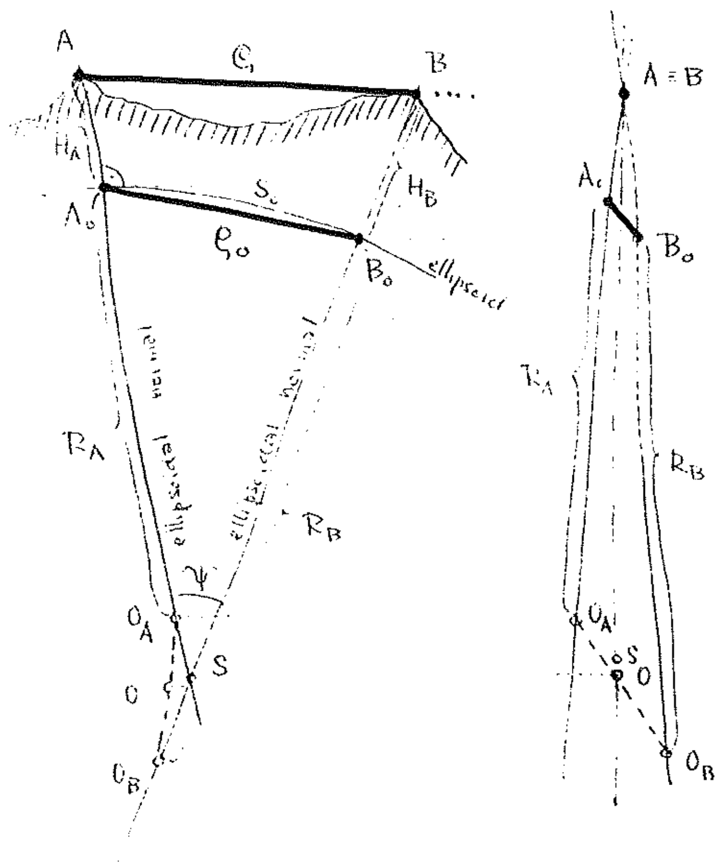
$$-\kappa = \text{arctg}\ \eta'/\xi'$$

is the azimuth of the astro-deflection θ'.

In order to get some feeling for the magnitudes involved let us consider an angle Ω observed in high mountains where we have $\theta' = 20''$, $\alpha_1 = 45°$, $\alpha_2 = -45°$, $\kappa = 0$, $Z_1 = Z_2 = 100°$. We get

$$\Delta\omega = 20'' \cdot 2 \cdot \sin 45° \ \text{cotg}\ 100° \simeq -40'' \cdot 0.707 \cdot 0.176$$
$$\simeq -5'' .$$

7.2) Reduction of observed distances to the reference ellipsoid



In order to be able to derive the reduction formulae, let us draw the diagram depicting the problem we shall be dealing with. In the diagram, ρ is the observed distance and H_A, H_B are the ellipsoidal heights. ρ_o, s_o are the chord and ellipsoidal distances between the two projections A_o, B_o of A and B. ψ is the plane angle of the projections of the two normals into the projection plane, given by three points: A,B,O, O being the center of O_A, O_B. R_A,R_B are the radii of curvature of the ellipsoidal section AB_o given by

$$\frac{1}{R_A} = \frac{\cos^2\alpha_A}{M_A} + \frac{\sin^2\alpha_A}{N_A} \quad , \quad \frac{1}{R_B} = \frac{\cos^2\alpha_B}{M_B} + \frac{\sin^2\alpha_B}{N_B} .$$

Here α_A, α_B are the azimuths of the ellipsoidal section at A and B, M,N are the principal radii of curvature.

Denoting $\overline{A_oS}$ by R'_A and $\overline{B_oS}$ by R'_B we can write from the triangle SAB:

$$\rho^2 = (R'_A + H_A)^2 + (R'_B + H_B)^2 - 2(R'_A + H_A)\ (R'_B + H_B)\cos\psi .$$

Expressing $\cos\psi$ as $1 - 2\sin^2\frac{\psi}{2}$ we obtain

$$\rho^2 = [(R'_A + H_A) - (R'_B + H_B)]^2 + 4\ (R'_A + H_A)\ (R'_B + H_B)\sin^2\frac{\psi}{2} .$$

Applying the same approach to the triangle SA_oB_o we get

$$\rho_o^2 = [R'_A - R'_B]^2 + 4R'_AR'_B\sin^2\frac{\psi}{2} .$$

and

$$\sin^2\frac{\psi}{2} = \frac{\rho_o^2 - [R'_A - R'_B]^2}{4R'_AR'_B} .$$

This result can be substituted back into the formula for ρ^2.

Denoting $R'_A - R'_B$ by $\Delta R'$ and $H_A - H_B$ by ΔH we get

$$\rho^2 = [\Delta R' + \Delta H]^2 + \frac{\rho_o^2 - \Delta R'^2}{R'_AR'_B}(R'_A + H_A)\ (R'_B + H_B)$$

$$= [\Delta R' + \Delta H]^2 + (\rho_o^2 - \Delta R'^2)\ (1 + \frac{H_A}{R'_A})\ (1 + \frac{H_B}{R'_B}) .$$

Denoting the mean of R_A, R_B by R, and $\delta R_A = R_A - R_A' = \overline{O_AS}$, $\delta R_B = R_B - R_B' = \overline{O_BS}$ we obtain

$$R_A + R_B = R_A' + R_B' + \delta R_A + \delta R_B$$

or

$$2R = R_A' + R_A' - \Delta R' + \delta R_A + \delta R_B$$

$$2R = R_B' + \Delta R' + R_B' + \delta R_A + \delta R_B .$$

Hence

$$R_A' = R + \frac{\Delta R'}{2} - \frac{\delta R_A + \delta R_B}{2}$$

$$R_B' = R - \frac{\Delta R'}{2} - \frac{\delta R_A + \delta R_B}{2} .$$

Taking $\frac{\delta R_A + \delta R_B}{2} << \frac{\Delta R'}{2}$

we get

$$R_A' \approx R + \frac{\Delta R'}{2}, \quad R_B' \approx R - \frac{\Delta R'}{2}$$

and

$$\rho^2 \simeq (\Delta R' + \Delta H)^2 + (\rho_o^2 - \Delta R'^2)\left(1 + \frac{H + \frac{1}{2}\Delta H}{R + \Delta R^{\frac{1}{2}}}\right)\left(1 + \frac{H - \frac{1}{2}\Delta H}{R - \Delta R^{\frac{1}{2}}}\right)$$

with H denoting the mean of H_A, H_B and ΔH denoting their difference. After some development, the product of the last two terms becomes approximately

$$1 + 2\frac{H}{R} - \frac{\Delta H \Delta R'}{R^2} + \frac{H^2}{R^2} \doteq 1 + 2\frac{H}{R} \quad .$$

We can see that the same result is obtained for

$$\frac{\delta R_A + \delta R_B}{2} >> \frac{\Delta R'}{2}$$

as well as

$$\frac{\delta R_A + \delta R_B}{2} \approx \frac{\Delta R'}{2} \, .$$

Substituting this approximation back into the expression for ρ^2 we obtain

$$\begin{aligned}\rho^2 &\simeq \Delta R'^2 + 2\Delta R'\Delta H + \Delta H^2 + \rho_o^2\left(1 + 2\frac{H}{R}\right) - \Delta R'^2\left(1 + 2\frac{H}{R}\right) \\ &= \rho_o^2\left(1 + 2\frac{H}{R}\right) + \Delta H^2 + \Delta R'\left(2\Delta H - 2\Delta R'\frac{H}{R}\right) \\ &\simeq \rho_o^2\left(1 + 2\frac{H}{R}\right) + \Delta H^2 + 2\Delta R'\Delta H.\end{aligned}$$

Hence, we can finally write the equation for ρ_o^2

$$\boxed{\rho_o^2 = \frac{\rho^2 - \Delta H^2 - 2\Delta R'\Delta H}{\left(1 + 2\frac{H}{R}\right)}}$$

where $\Delta R'$ for short distance ρ is usually so small that it can be neglected altogether.

We can notice that $\rho^2 - \Delta H^2$ represents the square of the horizontal distance, say ρ_H^2. Disregarding $\Delta R'$, we can write

$$\boxed{\rho_o = \sqrt{\frac{\rho_H^2}{1 + 2\frac{H}{R}}} \simeq \rho_H \left(1 - \frac{H}{R}\right)}$$

which is the approximate formula used in practice for reducing short horizontal distances.

For the ellipsoidal distance s_o, we get approximately

$$s_o \simeq R\psi .$$

But ψ can be expressed from the earlier formula for $\sin^2 \frac{\psi}{2}$ as

$$\psi \doteq 2 \arcsin \sqrt{\frac{\rho_o^2 - \Delta R'^2}{4R^2}} .$$

Therefore

$$s_o \simeq 2R \text{ arc sin } [\frac{1}{2R} \sqrt{} (\rho_o^2 - \Delta R'^2)]$$

$$\boxed{s_o \simeq 2R \text{ arc sin } [\frac{\rho_o}{2R} (1 - \frac{1}{2} \frac{\Delta R'^2}{\rho_o^2})]} .$$

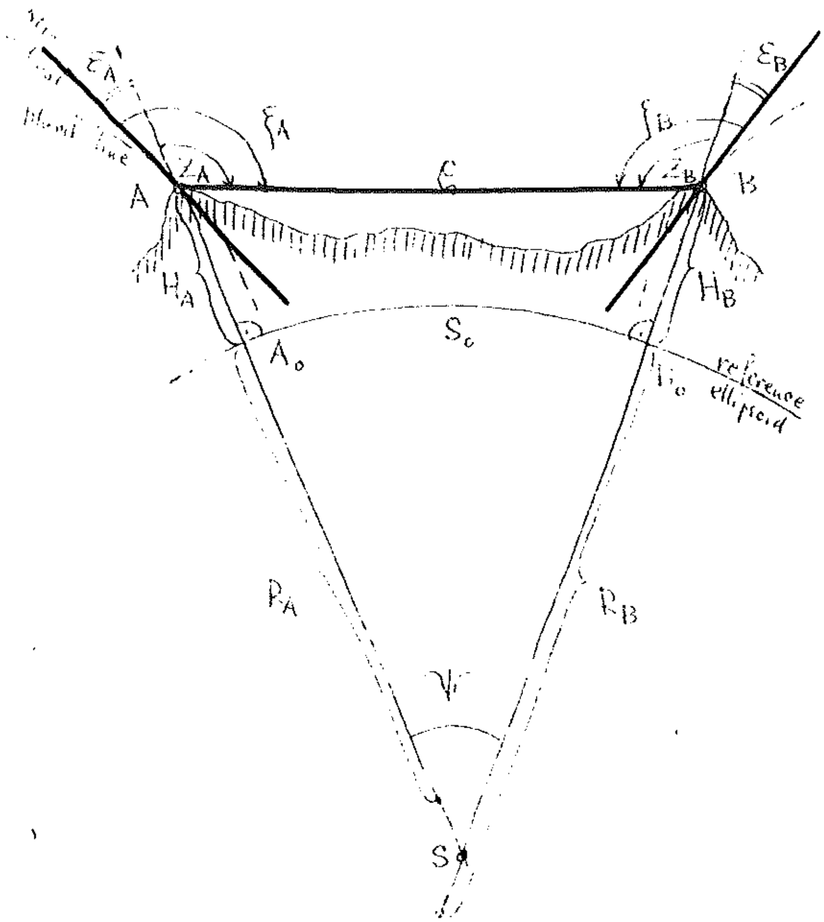
As an example, let us take a distance ρ of 20 km measured by an EDM device between two points on the same meridian on latitutde $\phi \approx 45°$. Let H_A = 200 m, H_B = 800 m. We first neglect $\Delta R'$ and get

$$\rho_o^2 \doteq \frac{(2 \cdot 10^4)^2 - (6 \cdot 10^2)^2}{1 + 2 \frac{5 \cdot 10^2}{6.4 \cdot 10^6}} \quad m^2 \approx 4.10^8 \left(1 - \frac{36.10^4}{4 \cdot 10^8}\right) (1 - 1.56 \cdot 10^{-4})$$

$$\approx \rho^2 (1 - 9\ 10^{-4}) (1 - 1.56\ 10^{-4})\ m^2 .$$

Hence $\rho_o \approx \rho\ (1 - 4.5\ 10^{-4})\ (1 - 0.78\ 10^{-4})$ m and the "slope" and "ellipsoid" corrections come to 9 m and 1.56 m respectively.

7.3) Triangulated heights



Leaving refraction aside, the zenith distances ζ_A, ζ_B are what we get from the observations on the surface of the earth. On the other hand, the angles Z_A, Z_B are the quantities we should like to get that would enable us to determine the difference in ellipsoidal heights of the two points A and B.

Let us deal with short distances s_o only, for which we can write $R_A \simeq R_B \simeq R$ with sufficient precision, and regard s_o as being a spherical arc. First, we can write:

$$Z_A = \zeta_A + \varepsilon_A = \zeta_A + (\xi_A' \cos \alpha_{AB} - \eta_A' \sin \alpha_{AB})$$

where the formula for ε_A was derived in § 6.5, since ε_A is evidently the component of θ_A in the azimuth α_{AB} of the line AB. Completely analogous formula can be written for B. Note the sign of the term in brackets. It originates from the definition of the sense of ξ',η' and in the diagram above both ε_A and ε_B would be negative.

Once the quantities Z_A and Z_B are determined we can evaluate the difference H_B - H_A as a function of both Z_A and Z_B. Applying the tangent law to the triangle SAB and considering $R_A \doteq R_B \doteq R$ we get:

$$(R+H_A-R-H_B)\ tg[\tfrac{1}{2}(180°-Z_A+180°-Z_B)] \simeq (R+H_A+R+H_B)\ tg[\tfrac{1}{2}(180°-Z_A-180°+Z_B)]$$

or

$$(H_A-H_B)\ tg\ [\tfrac{1}{2}(360°-Z_A-Z_B)] \simeq (2R+H_A+H_B)\ tg\ \frac{Z_B-Z_A}{2}\ .$$

Here evidently $180° - (Z_A + Z_B) = \psi = s_o/R$ so that we can rewrite the equation as

$$(H_A - H_B)\ tg\ (90° + \frac{s_o}{2R}) \simeq 2(R + H)\ tg\ \frac{Z_B-Z_A}{2}$$

with H denoting the mean of H_A and H_B.

We can now substitute $-\ cotg\ \frac{s_o}{2R}$ for $tg\ (90° + \frac{s_o}{2R})$ and obtain

$$H_B - H_A \simeq 2(R + H)\ \ tg\ \frac{s_o}{2R}\ \ tg\ \frac{Z_B-Z_A}{2}\ .$$

Since $s_o/(2R)$ is a very small angle, we can develop the first tangent into power series:

$$tg\ \frac{s_o}{2R} = \frac{s_o}{2R} + \frac{1}{3}\ (\frac{s_o}{2R})^3 + \ldots = \frac{s_o}{2R}\ (1 + \frac{s_o^2}{12R^2} + \ldots)\ .$$

Substituting this back into our equation for the height difference we get

$$H_B - H_A \simeq 2R(1 + \frac{H}{R})\ \ \frac{s_o}{2R}\ (1 + \frac{s_o^2}{12R^2})\ tg\ \frac{Z_B - Z_A}{2}$$

$$\boxed{H_B - H_A \simeq s_o\ (1 + \frac{H}{R} + \frac{s_o^2}{12R^2})\ tg\ \frac{Z_B - Z_A}{2}}\ .$$

Hence, using the zenith distances we are theoretically able to determine directly the heights above the ellipsoid. Unfortunately, the observed zenith distances are very sensitive to atmospheric refraction -- the major problem in geodetic observations -- and the precision of triangulated heights is still very low. Vehement research is going on in this area. The newest concept being to get away from the necessity to determine the astro-deflections and the refraction corrections by using redundant observations.

7.4) "Three-dimensional" adjustment of geodetic networks

Recently, some interest of the geodetic community has been directed towards the so-called "three-dimensional" approach to the geodetic problems. This approach was advocated already by Bruns, at the end of 19-th century, and in the fifties of our century revived by Hotine.

The whole idea of the three-dimensional treatment is that of using the three-dimensional Eucleidian space rather than the curved two-dimensional space--the ellipsoid. Using this approach, each point of the network is given by a three-tuple of coordinates, either X,Y,Z (Geocentric Cartesian) or ϕ,λ,H (geodetic). In order to be able to reconcile the astronomic observations with the terrestrial we have to be able to express also the local astronomical verticals (tangent to the local plumb-lines). This is usually done by means of two direction parameters, either directly Φ and Λ or ξ' and η'.

Once this is done, we can formulate the equations for astronomic azimuth A, zenith distance ζ taken with respect to the local astro-vertical and the cord distance ρ between any two points of the network. Providing we have enough observations of the actual quantities A,ζ,ρ and the horizontal angles $\omega_{ij} = A_i - A_j$ we can formulate the parametric adjustment for the unknown increments dx,dy,dz, $d\Phi$, $d\Lambda$ or $d\phi,d\lambda,dH,d\Phi$, $d\Lambda$. Then the spatial positions of the individual points and their plumb-line directions can be determined. The ways to formulate the models and to adjust them are numerous and it was not considered the aim of this write-up to tackle them.

If both adjustments; ie., the "three-dimensional" and the one using the reference ellipsoid, are formulated properly, they should yield the same results. They are mathematically equivalent. From the numerical point of view, however, one might be preferred to the other because of less labor involved. The difference between the two is that they use the reference ellipsoid either explicitly or implicitly or not at all and this difference is reflected in the formulae.

7.5) Discussion of various methods for geoid determination

So far, we have met three different methods for geoid determination:

i) gravimetric, using Stokes' formula dealing with gravity anomalies (see § 3.17);

ii) astro-geodetic, using the astro-deflections of vertical (see § 6.5);

iii) astro-gravimetric, using the astro-deflections and the gravity anomalies (see § 6.6).

We have also seen that the first method furnishes the geoid related to the mean-earth ellipsoid; i.e., ellipsoid determined in this case by normal gravity up to a scale factor. (The scale factor of thus determined ellipsoid has to be supplied from other sources.) On the other hand, the gravimetric geoid is concentric with the mean-earth ellipsoid and known therefore as absolute geoid.

The second and the third methods give the geoid related to the used, usually local, reference ellipsoid and the geoid is hence known as relative geoid. The scale of the relative geoid is evaulated properly. The disadvantage of the relative geoid however is that its precision decreases with the distance from the point of origin, where the relative position of the two surfaces (geoid and reference ellipsoid) is assumed. If we knew the relation of the local ellipsoid with respect to the mean-earth ellipsoid we could transform the relative geoid to the absolute geoid and vice-versa.

The gravimetric (absolute) geoid has the inherent incertitude of the order of $\pm$ 5 to 10 m, in the best case. The incertitude is due to two main factors. First, the gravity anomalies are not yet known uniformly well all over the globe. The gravity observations in various regions like seas and polar areas are very scarce. Second, the Stokes' formula is only a spherical approximation; i.e., it cannot be expected to give any better precision than $\pm N\, 3\, 10^{-3}$.

On the other hand, the astro-levelling and even more so the astro-gravimetric levelling are more precise. The precision is, of course, a function of the coverage of the area by astro-deflection points and the gravity points (in case of the second technique). Theoretically, the precision can be better than $\pm$ 1 m. The major hindrance here is the necessity to correct the astro-deflections for the curvature of the actual plumb-line.

In recent years the gravimetric determination was given a boost by the satellites. The idea of using the satellite information for the determination of the absolute geoid is basedon the possibility to interpret the satellite orbit perturbations in terms of gravity disturbances of the gravity field above the earth. These disturbances can then be transmitted to the mean-earth ellipsoid and converted into disturbing potential and ultimately into geoidal heights. The satellite orbit information can be used in two different ways; either on its own or in conjunction with the terrestrial gravity data. Hence we can have two more distinctly different modes for absolute geoid determination:

iv) <u>satellite</u> determination;

v) <u>combined</u> determination.

From the two last methods, the former suffers from the effect of the altitude of the satellites used. At the altitude of the satellite orbit the "details" of the terrestrial gravity field are becoming undistinguishable and we are not able to resolve the geoid to any detail. It is basically the same problem as trying to determine the shape of an underground deposit from the gravity profiles on the surface, a problem encountered in applied geophysics. Speaking in terms of spherical harmonics, the satellite determination can supply only a few low-degree harmonics.

The combined solution is hence preferable in sofar that the advantage of the global coverage by satellites is supplemented by the terrestrial information on the detailed structure at least in the surveyed areas. It can therefore be regarded as theoretically better.

Acknowledgements

The author wishes to acknowledge the kind assistance of Dr. E. J. Krakiwsky. The endless discussions he has had with Dr. Krakiwsky during the period of writing these two sets of lecture notes helped clarifying a great number of points.

All comments and corrections will be highly appreciated.

References

Interested reader is recommended to further his understanding by referring to the three main comprehensive text-books on geodesy published in English, namely:

Bomford, G., 1971: Geodesy. Clarendon, Press, Oxford.

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

Hotine, M., 1969: Mathematical Geodesy. ESSA monograph 2, U. S. Dept. of Commerce, Washington.

A wealth of references to special publications, scientific papers and text-books in other languages can also be found in the three textbooks.