UNB

Why do we need a proper geoid

Petr Vaníček
Department of Geodesy and Geomatics Engineering
University of New Brunswick
P.O. Box 4400
Fredericton, N.B. CANADA
E3B 5A3

UNB

My working hypothesis:

"Let us all agree that heights are a necessary component of surveying and mapping and that they should be as useful to the user as possible".

In this talk, I shall emphasize the aspect of geoid as a datum for practical heights leaving the other, scientific aspects alone.

UNB

Practical heights

• Practical heights referred to Mean Sea Level (MSL) as their (vertical) datum ⇒ points on the mean shoreline all have height 0, as practice expects.

•Practical heights must be physically meaningful ⇐ must show what is up and what is down.

Vaníček, 2009 | Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 | 3

UNB

HEIGHTS ARE INTRICATELY INTERTWINED WITH GRAVITY! Consequently, the rigorous vertical datum for practical heights is *not* MSL but an equipotential surface of the Earth's gravity field, the geoid.

Quasigeoid used in the former Soviet Empire and some countries in Europe. Geoid and quasigeoid coincide along the mean shoreline and at sea.

Vaníček, 2009 | Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 | 4

UNB

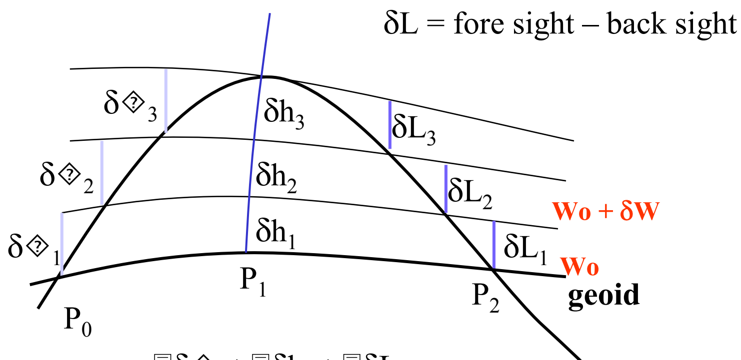
Practical heights are defined by means of real gravity observed at the Earth's surface.

Following 14 slides contain the story of different varieties of practical heights. Need for real gravity in practical height determination should become evident.

UNB

Geodetic Levelling

Levelling is affected by the earth's gravity field



□δ□$_i$ ≠ □δh_i ≠ □δL_i

UNB

$\square\delta\diamond_i \neq \square\delta h_i \neq \square\delta L_i$,

due to the non-parallelism of equipotential surfaces

$$\oint dL \neq 0$$

Consequences $\Rightarrow$

- levelling loops do not close (impossible to adjust)
- non-uniqueness of heights

Must use properly defined heights instead of levelled heights.

UNB

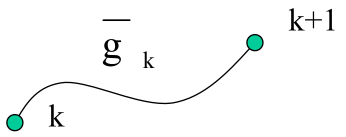
Geopotential Numbers

- The most natural height system, but impractical.
- Use of earth's gravity potential W which defines "height" uniquely (each point can lay on only one equipotential surface!).
- Difference between two close together equipotential surfaces:

$\delta W = -g\ \delta h$

- $\Delta C_{AB} = C_B - C_A = W_B - W_A = -\Delta W_{AB} = \int_A^B g dL \approx \sum_{k=A}^{B} \bar{g}_k \delta L_k$
- $C_A = W_0 - W_A = -\Delta W_{A0}$ (subscript 0 means "at the geoid")
- The higher the point the smaller the potential ➔ geopotential numbers grow with height.

UNB



$$\bar{g}_k = \frac{g^k + g^{k+1}}{2}$$

- Usually g is evaluated at the benchmarks:

  if g has not been observed: interpolate from a map.

  required accuracy for g: 1-2 mGal is good enough

- Geopotential numbers are path independent: $\oint dC = \oint g dL = 0$

  i.e., they are *holonomic* ➔ heights are unique.

- Geopotential numbers are expressed in physical units of the potential (‘length’ times ‘acceleration’), e.g., kGal m (1 kGal m ⇔ 1.02 m). They are not in length units.

UNB

Dynamic Heights

- Geopotential numbers scaled by a *reference gravity* G, to bypass both numerical and unit problems:

$$H^D_A = C_A/G$$

  e.g., G may be normal gravity γ for a reference latitude ϕ_R.

- Points on same equipotential surface have the same H^D and C.
- H^D has physical meaning: water always flows down the hill.
- H^D has no geometrical meaning: $\delta L_1 \neq \delta L_2 \neq \delta L_3$ but $H^D_1 = H^D_2 = H^D_3$.
- Dynamic heights are unique: $\oint dH^D = 0$
- $H^D = 0$ for points on the geoid.

$\Delta H^D_{AB} = H^D_A - H^D_B = \Delta C_{AB} / G$

Replacing ΔC_{AB} and developing the equation:

$$\Delta H^D_{AB} = \sum_{i=A}^{B} \delta L_i + \sum_{i=A}^{B} \frac{\bar{g}_i - G}{G} \delta L_i$$

we get:

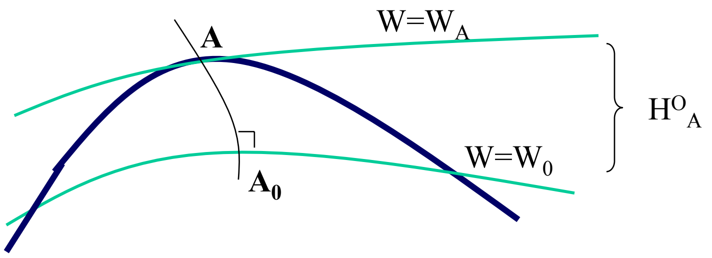
$$\Delta H^D_{AB} = \Delta L_{AB} + DC_{AB}$$

where DC_{AB} is the dynamic correction.

- It is a function of mean observed gravity along the path, hence, path dependent.
- It is small: vary from 0 to 1.3 mm √S[km] for 1st order levelling.
- Must be applied in 1st order levelling

Orthometric Heights

Geometrical concept of height: H^O of a point A corresponds to the distance between the point and the geoid, measured along the plumb line of A.



$$H^O_A = C_A / \tilde{g}_A$$

where $\tilde{g}_A$ is the mean gravity along the plumb line of A.

UNB

•Several prescriptions for evaluating $\tilde{g}_A$ exist. The one most often used in practice is Helmert's orthometric height:

$$H_A^O = \frac{C_A}{g_A + 0.0424\ H_A^O}$$

where g_A is surface gravity at A.

• Recursive equation, one iteration is enough (make H^O_A equal to observed height).

•Because actual value of $\tilde{g}_A$ is unknown, H^O is only an approximation

•Points on same equipotential surface do not generally have same H^O

•H^O has no physical meaning: water may run up the hill.

•Due to its geometric meaning H^O used extensively in technical practice.

•H^O is holonomic and $H^O = 0$ for points on the geoid.

Vaníček, 2009 Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 13

UNB

$\Delta H^O_{AB} = H^O_B - H^O_A$

After some development:

$$\Delta H_{AB}^O = \Delta L_{AB} + \sum_{i=A}^{B} \frac{\bar{g}_i - G}{G}\delta L_i + H_A^O \frac{\tilde{g}_A - G}{G} - H_A^O \frac{\tilde{g}_B - G}{G}$$

$\Delta H^O_{AB} = \Delta L_{AB} + OC_{AB}$ (one iteration is enough)

OC_{AB} is the orthometric correction $\Rightarrow$

•It is larger in mountainous regions.

•Generally |DC| > |OC| ➔ ΔH^O closer to ΔL than ΔH^D is .

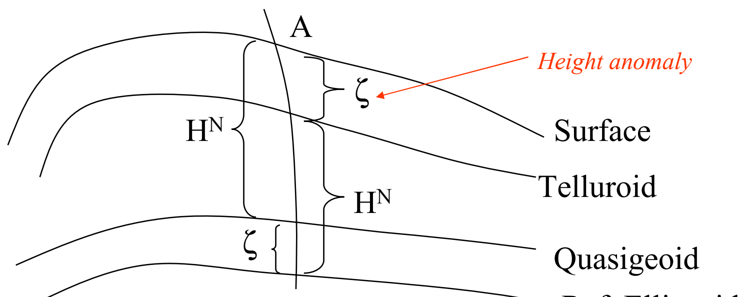
Vaníček, 2009 Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 14

Molodenskij: "gravity along the plumb line, $\tilde{g}_A$, cannot be evaluated accurately enough because the topographical mass density is not known well enough" ⇒ Molodenskij's theory, the quasigeoid and normal heights.

Not the case [Tenzer et al., 2005]: accuracy of orthometric heights is better than 1 cm when lateral topo-density variations are taken into account.

Normal Heights

Describe heights above the quasigeoid (approximation of H^O)



$$H_A^N = C_A / \tilde{\gamma}_A$$

where $\tilde{\gamma}_A$ is the mean normal gravity along the normal plumbline of A.

UNB

Example: Vignal's normal height.

$$H_A^N = \frac{C_A}{\gamma_A - 0.3086 \frac{H_A^N}{2}}$$

γ_A is normal gravity at A.

•Normal height is holonomic.

•**Normal correction**:

$$NC_{AB} = DC_{AB} + H_A^N \frac{\tilde{\gamma}_A - G}{G} - H_B^N \frac{\tilde{\gamma}_B - G}{G}$$

Vaníček, 2009 Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 17

UNB

Real gravity not available? Use normal gravity! Observed gravity, g, is replaced by normal gravity, γ. Use of γ results in heights (of whatever kind) being only *approximate*.

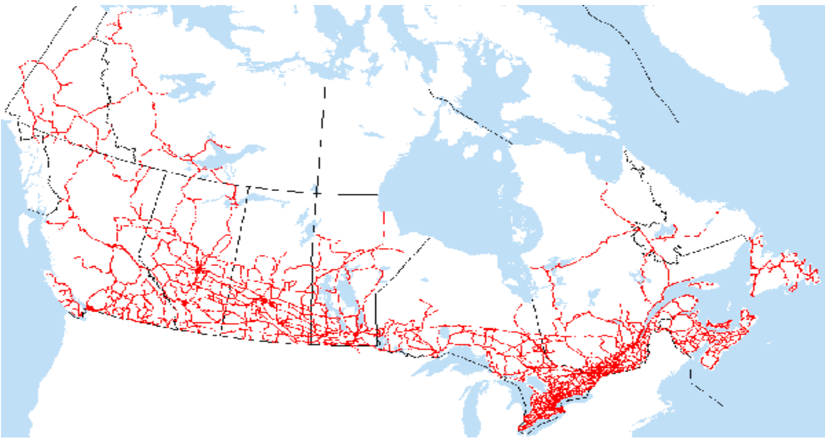
This practice OK 50 years ago. Now, real gravity known almost everywhere. Even freely available gravity from GGMs is much closer to real gravity g than normal gravity γ is.

Vaníček, 2009 Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 18

No excuse for using normal gravity any more!!! Approximate heights not referred to geoid, not physically meaningful.

FORGET THEM!

Canadian Vertical Network



Introduction of space techniques (1960's) opened possibility to measure geodetic heights h (sometimes called, incorrectly, ellipsoidal heights).

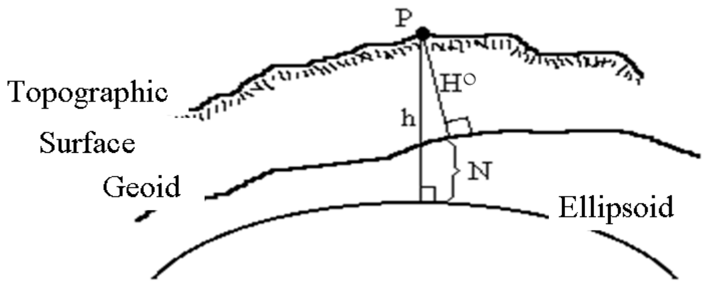
Concept of geodetic heights known and used long before but they were not possible to measure.

NEVER MEANT TO REPLACE ORTHO-METRIC HEIGHTS AS PRACTICAL HEIGHTS.

Geodetic heights are used in parallel with orthometric and other heights.

Orthometric and Geodetic Heights

$H^O = h - N$ (*this expression is quite accurate*)



Measured geodetic heights h are converted to orthometric heights H^O by the equation above. Levelling (getting H^O) expensive and GPS (determination of h) cheap $\Rightarrow$ this approach is used to "replace" levelling by GPS. Results in cheaper but less accurate orthometric heights.

In reality, this is done differentially – see next slide – and zillions of $$ are saved. But there are problems.

Vaníček, 2009 | Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 | 23

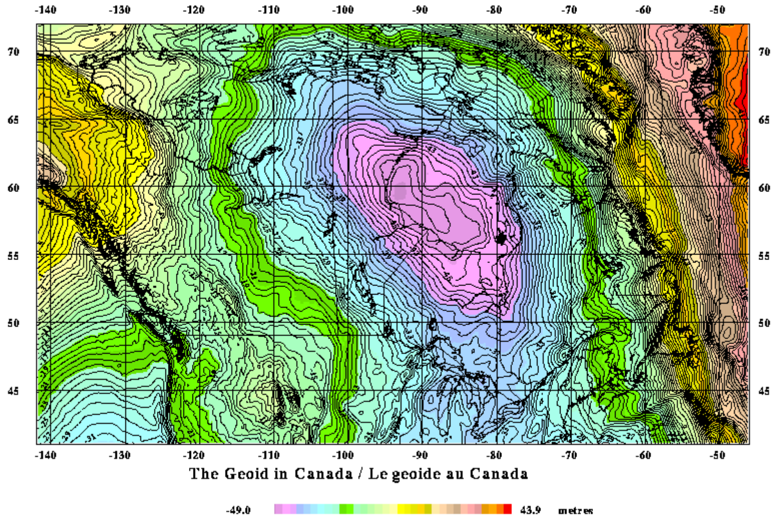
Relation between geodetic and orthometric height differences

- $\Delta H^O = \Delta h - \Delta N$ (*quite accurate*)
- $H^O_B - H^O_A = (h_B - h_A) - (N_B - N_A)$
- ΔH^O from terrestrial observations (geodetic levelling)
- Δh from space observations
- ΔN from geoidal map

Vaníček, 2009 | Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 | 24

UNB

Geoid in Canada



Vaníček, 2009

Geodetic Research Laboratory – University of New Brunswick,
FIG Working Week, Eilat, May 6, 2009

25

UNB

Problems with this approach?

Yes, here they are:

1. Definition of (vertical) datum for h
2. GPS determined Δh is not accurate enough

   (accuracy of h lower still)
3. ΔN from global geoid models not accurate enough

Vaníček, 2009

Geodetic Research Laboratory – University of New Brunswick,
FIG Working Week, Eilat, May 6, 2009

26

UNB

Re1) Defining the vertical datum for h, the geocentric reference ellipsoid, i.e., an equipotential surface of normal gravity γ whose potential has the same value W_0 as the real potential W on the geoid. Entails choosing parameters of normal field. Changed (by IAG) whenever the set is no longer representative of the actual field. Present parameters adopted by IAG in 1981 as Geodetic Reference System of 1980 (GRS80).

UNB

Let's mention also advantages in using Geocentric Reference Ellipsoid:

- Its definition parallel to that of geoid
- Position, orientation, shape, size are exactly and uniquely defined and do not change with time
- Possibility to obtain the geodetic height h of a point exactly from known geocentric coordinates, say x, y, z, of that point.

UNB

Re 2 GPS improving steadily but accuracy of Δh still worse than accuracy of ΔH^O.

Re 3 Accuracy of geoidal height N varies widely.

STATEMENT #3 BRINGS US TO THE MAIN POINT OF THIS CONTRIBUTION!

UNB

Geoidal height N is the link between orthometric and geodetic heights. Also link between two global vertical datums: Geocentric Reference Ellipsoid and the Geoid. It can be obtained in many ways. How? How accurately?

If GPS-derived h are to be used in practice then the geoid (N) must be known as accurately as possible, preferably to within one or two centimetres. This is a stiff order.

Can the artificial surface Z defined by

$$h - H^O = Z,$$

be used instead of the geoid N, as advocated be some practitioners ?

No! Here is why :

- Z known only at some points ⇒interpolation
- H^O are usually referred to MSL ⇒SST missing
- h are much less accurate than Δh.

What is Geoid anyway?

Sea level surface follows closely an equipotential surface ⇒ Gauss's idea: geoid is equivalent to sea level surface.

Formally: geoid is that equipotential surface which "corresponds to within about 1-2 m to the time-averaged ocean surface" [Lambeck, 1988], or "which best fits, in the least squares sense, mean sea level" [US National Geodetic Survey, 1986].

The equivalence is only approximate. In applications of sub-metre accuracy, the difference must be taken into account.

Mean Sea Level (MSL) surface dictates the value W_0 of the potential on the geoid:

$$W = W_0 \text{ (MSL)}$$

MSL is really the surface to which orthometric heights are referred in practice.

Mean Sea Level

How is it determined? We have to study sea tide and other variations at many locations:

- measure by means of tide-gauges,
- analyse these measurements.

World Service for the MSL at Bidston. Study of local tidal behaviour falls under the auspices of *national hydrographic services*.

Sea Surface Topography

Separation of geoid and MSL is called Sea Surface Topography (SST).

What is causing SST? *Temperature, salinity content, suspended particle content.* Met. forces: *wind shear*, *barometric pressure* variations, *river discharge*, and also *steady and turbulent flows*. For a complete list of causes see Montgomery [1937-38], Warren and Wunch [1981]

We usually write:

$$SST(t) = SST_0 + \delta SST(t)$$

where the permanent part SST_0 ranges between -1.5 and + 1.5 metres [Levitus, 1982].

No country accounts for SST_0 when referring its heights to geoid as a vertical datum ⇒ heights referred really to MSL not to geoid.

UNB

It should now be clear what the problems with the artificial surface Z are: inaccurate control through h and potentially large systematic errors due to omitting of SST_0.

FORGET IT!

Vaníček, 2009 Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 37

UNB

What can be done?

To get accurate practical heights:

1. Convert the observed geodetic height differences to orthometric height differences:

$$\Delta H^O = \Delta h - \Delta N$$

2. Correct orthometric heights H^O for the permanent SST_0 effect to have the heights referred to the geoid and not MSL.

Vaníček, 2009 Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 38

Re 1. How accurately are the geoidal heights N known?

- From global geoid models is good to a few metres
- The regional geoid models can be two orders of magnitude better [Vaníček and Martinec, 1994]

Regional model:

- Does not suffer from the cutoff of the high degrees
- It tends to have the latest gravity, satellite altimetry, deflections of the vertical in the solution

Vaníček, 2009 Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 39

Re 2. How can we determine SST_0 ?

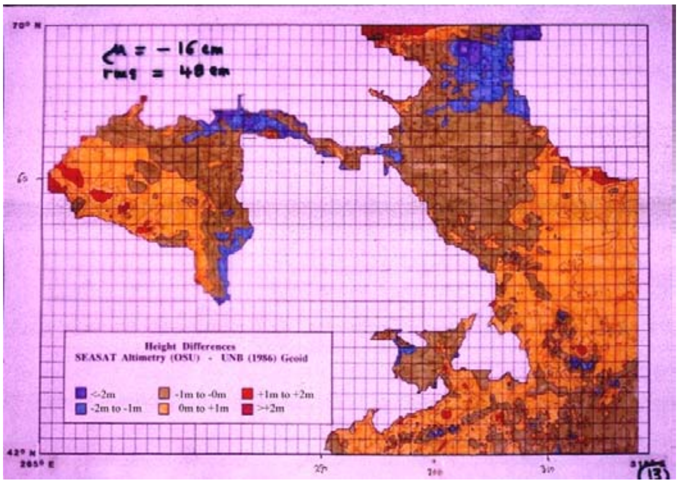
This is the most difficult problem of geodesy. Geoid must be known. Cooperation of physical oceanographers must be secured.

This choice is more difficult then #1.

Vaníček, 2009 Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 40

UNB

Sea Surface Topography – North Atlantic



Vaníček, 2009 | Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 | 41

UNB

The easiest solution: get a good regional geoid. What good would it do? The answer is affirmative and it is twofold

- To enable use of space techniques – explained above
- To unify the height systems - heights on dry land and the heights/depths used in maritime boundary delimitation and in bathymetric charting are from different systems.

Vaníček, 2009 | Geodetic Research Laboratory – University of New Brunswick, FIG Working Week, Eilat, May 6, 2009 | 42

Discontinuity: heights in maps and depths in bathymetric charts. Canadian Hydrographic Service study of unification of vertical datums in Canada (Seamless Vertical Datum).

Additional advantage: gravity data collection and archiving should satisfy even the non-surveying users.