GEODESY The Concept

DEFINITION OF GEODESY

Geodesy is the science concerned with the study of the shape and size of the earth in the geometric sense as well as with the form of the equipotential surfaces of the gravity potential.

Friedrich R. Helmert (1880)
One of the main founder of Geodesy

The following second definition represents a more up-to-date description of geodesy given by the Committee on Geodesy of the U. S. National Academy of Sciences in 1978.

- Establishment and maintenance of national and global three dimensional geodetic networks
- Measurement and analyses of geodynamic phenomena (earth rotation, earth tides, crustal movements, etc.)
- Determination of the earth's gravity field - Items 1-3 include also changes with time

Why has geodesy to deal with the gravity field ?
First of all, every geodetic measurement is a function of the gravity field (Example: by putting an instrument into the horizontal plane (by using spirit bubbles), it aligns its vertical axis with the local plumb line (local gravity vector) which, unfortunately, may vary from point to point.

Secondly, in defining heights we have to use an equipotential surface of the gravity field as vertical reference (Where does water flow?).

FIGURE OF THE EARTH AND REFERENCE SURFACES
The figure of the earth was approximated first by a sphere and later by an ellipsoid. Whereas these approximations are of geometrical character, the geoid represents a dynamical reference surface, a certain equipotential surface of the earth's gravity field.

The Earth as a Sphere
Various opinions on the form of the earth prevailed in the past, e. g., the notion of an earth disk encircled by Oceanus (Homer's Illiad, ~ 800 BC). Pythagoras (~ 580-500 BC) and his school as well as Aristotle (384-322 BC) among others expressed themselves for the spherical shape.

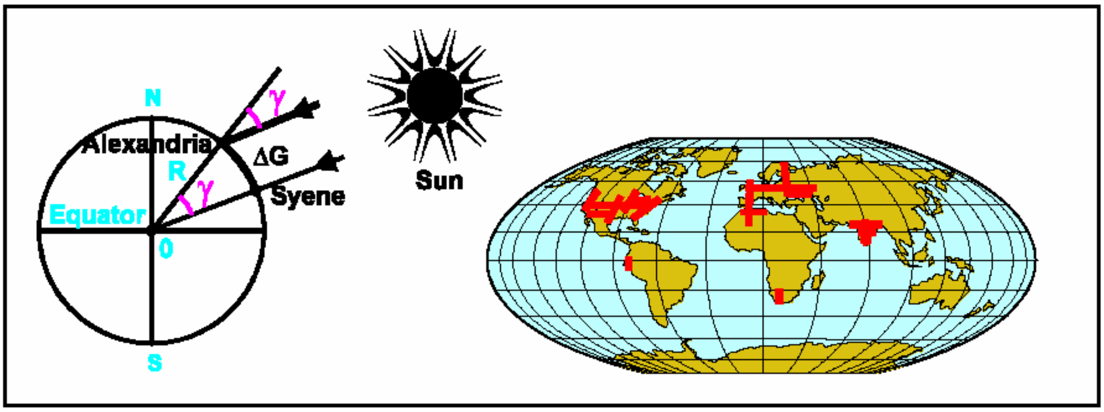

Fig. B-1. *The Earth as a sphere, derived from arc measurements*

The founder of scientific geodesy was Eratosthenes (276-195 BC.) of Alexandria who, assuming the earth was spherical, deduced from measurements a radius for the earth with an uncertainty of 2% (left hand side of Fig. B-1).
The principle of *arc measurements* developed by him was still applied in modern ages: From geodetic measurements the length ΔG of a meridian arc can be determined. Astronomical observations furnish the associated central angle γ . The radius of the earth is then given by $R = \Delta G/\gamma$.
The arc measurements in middle ages were characterized by fundamental advances in instrumentation technology. Arc measurements and early triangulations are shown on the world map of Fig. B-1.

The Earth as an Ellipsoid

Towards the end of the seventeenth century, Newton demonstrated that the concept of a truly spherical earth was inadequate as an explanation of the equilibrium of the ocean surface. He argued that because the earth is a rotating planet, the forces created by its own rotation would tend to force any liquids on the surface to the equator. He showed, by means of a simple theoretical model, that hydrostatic equilibrium would be maintained if the equatorial axis of the earth were longer than the polar axis. This is equivalent to the statement that the body is flattened towards the poles.
Flattening is defined by

$$f = (a - b)/a \qquad \text{(B-1)}$$

Where a is the semimajor, and b is the semiminor axis of the ellipsoid.

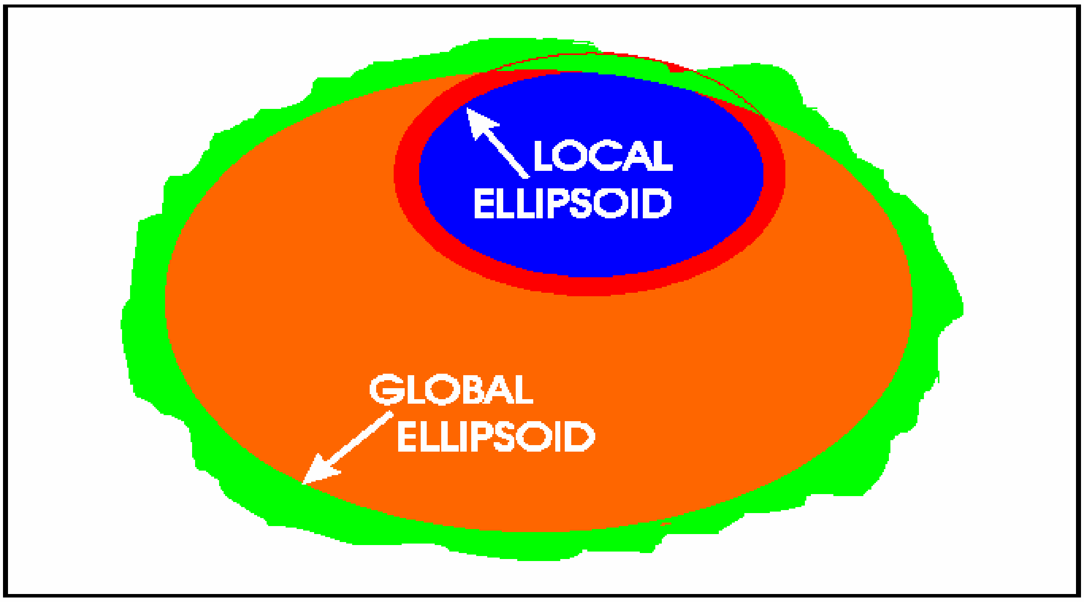

Fig. B-2. *Local ellipsoids are best-fitted to the specific country*

In the eighteenth and nineteenth century, ellipsoids were defined which were fitted best to a certain region of the earth (Fig. B-2). These local ellipsoids still provide the geometrical reference for the horizontal coordinates of various national geodetic (triangulation) networks.

Tab. B-1 shows examples of the ellipsoidal parameters of various ellipsoids. Note that the East Europeans (former Soviet Union) based their horizontal coordinates on a triaxial ellipsoid (Krassowsky).

THE EARTH AS A GEOID

Laplace (1802), Gauss (1828), Bessel (1837) and others had already recognized that the assumption of an ellipsoidal earth model was not tenable when compared against high accuracy observations. One could no longer ignore the deflection of the physical plumb line, to which measurements refer, from the ellipsoidal normal (deviation of the vertical, see

Fig. B-3). By an adjustment of several arc measurements for the determination of the ellipsoidal parameters a and f, contradictions arose which exceeded by far the observational accuracy.

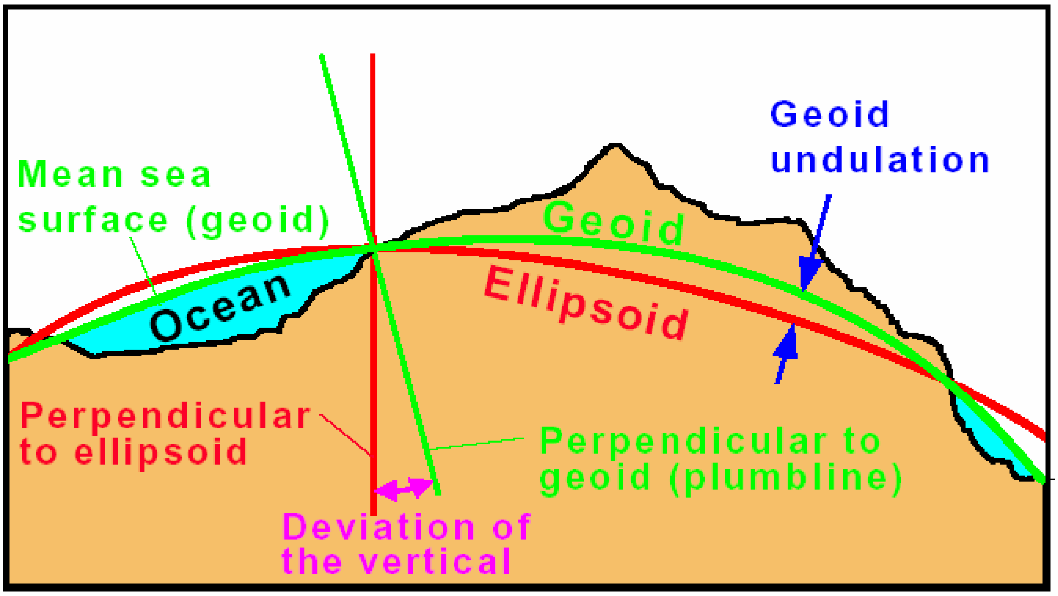

Fig. B-3. *The Earth as a geoid*

The equipotential surface of the earth's gravity field which would coincide with the ocean surface, if the earth were undisturbed and without topography. Listing (1873)

Listing (1873) had given the name ‘geoid’, Helmert (1880,1884) made the transition to the current concept of the figure of the earth. Here the deflections of the vertical are also taken into account in the computation of the ellipsoidal parameters.

The determination of the geoid has been, for the last hundred years, a major goal of geodesy. Its importance increased recently by the new concept of replacing the measurements of spirit levelling by GPS space observations and the use of precise geoid heights. Other global considerations require a unified vertical reference, i.e. a geoid determination with centimeter or even millimeter accuracy. This remains a challenge for geodesy in the coming years.

There are difficulties in defining a geoid: Sea-surface topography, sea-level rise (melting of the polar ice caps), density changes (earthquakes, etc.), ...

COORDINATE SYSTEMS AND REFERENCE ELLIPSOIDS

The general geodetic definitions:

Coordinate

One of a set of N numbers designating the location of a point in N dimensional space

Coordinate system

A set of rules for specifying how coordinates are to be assigned to points

$\rightarrow$ origin, set of axes

Type of coordinate system

- Local
- Geocentric, earth-fixed
- Cartesian / Ellipsoidal

Local Coordinate System

In the past, national survey departments computed ellipsoids best-fitted to their country to provide the basis for mapping. Origin and orientation of coordinate system is arbitrary, but often “the ball under the cross on top of the tower of a specific church” served as the zero-point (or origin) of a national coordinate system (Example: Soldner´s coordinate system in Bavaria with the Munich cathedral "Liebfrauendom" as origin.). The national ellipsoids are the geometric reference surfaces *only* for *horizontal* coordinates.

Geocentric Earth-Fixed Cartesian System (X, Y, Z)

As a fundamental terrestrial coordinate system, one introduces an earth-fixed spatial Cartesian system (X, Y, Z) whose origin is the earth´s centre of mass S (geocentre, centre of mass including the mass of the atmosphere, see Fig. B-4. Earth-fixed spatial Cartesian system (X, Y, Z)). The Z-axis coincides with the *mean rotational axis of the earth (Polar motion, CIO Pole).*

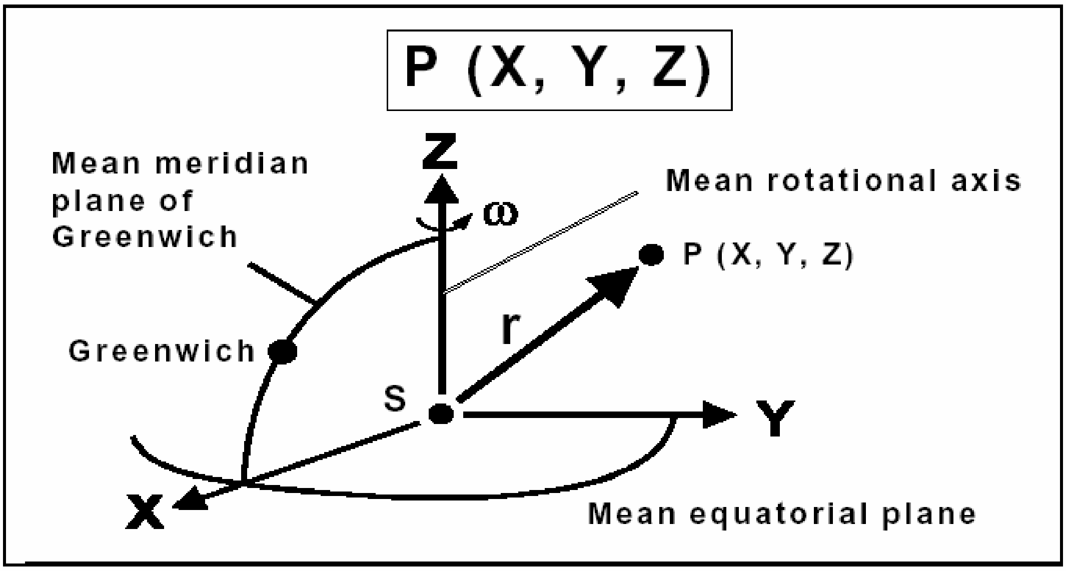

Fig. B-4. *Earth-fixed spatial Cartesian system (X, Y, Z)*

The mean equatorial plane perpendicular to this axis forms the (X-Y) plane. The (X-Z) plane is generated by the mean meridian plane of Greenwich. The latter is defined by the mean rotational axis and the zero meridian of the BIH (Bureau International de l´Heure) adopted longitudes ("mean" observatory of Greenwich). The Y-axis is directed so as to obtain a right handed system. The introduction of a mean rotational axis is necessary because in the course of time, the rotation changes with respect to the earth’s body. This applies to the position of the earth´s rotation axis (polar motion) and to the angular velocity of the rotation.

Ellipsoidal Geographic Coordinates

As Fig. B-5 shows, the earth´s surface may be closely approximated by a rotational ellipsoid with flattened poles (height deviation from the geoid <100 m). As a result geometrically defined ellipsoidal systems are frequently used instead of the spatial Cartesian coordinate system.

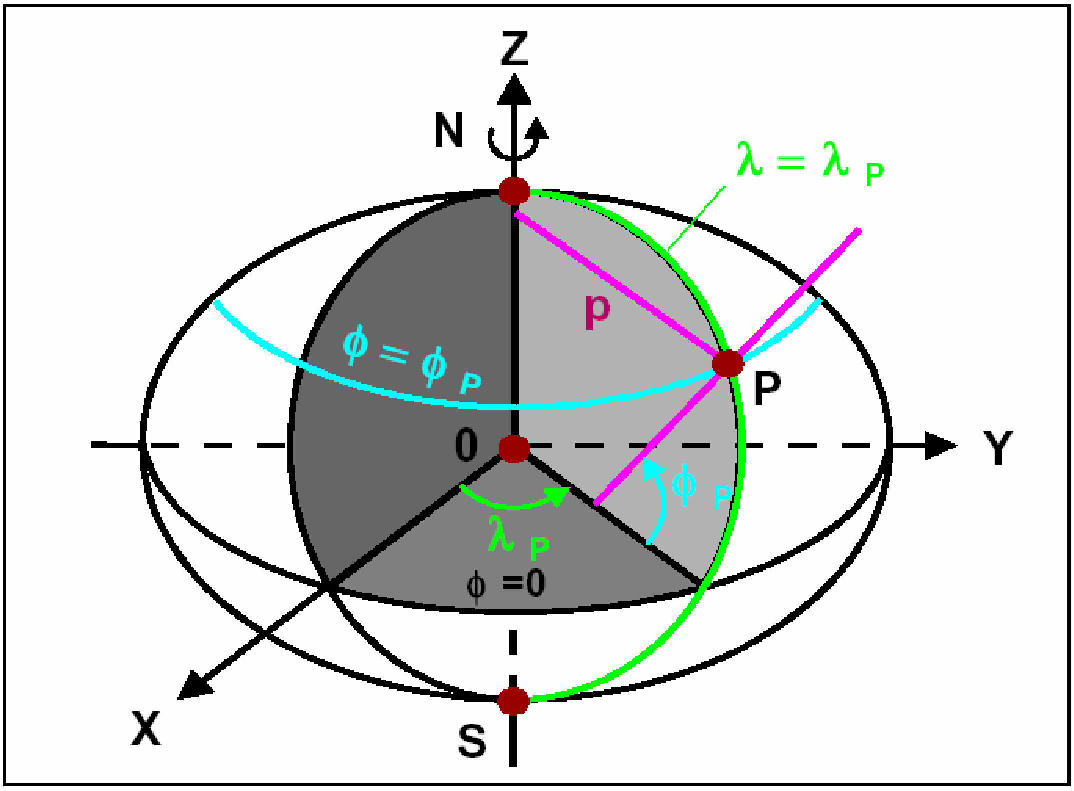

Fig. B-5. *Ellipsoidal geographic coordinates*

The rotational ellipsoid is created by rotating the meridian ellipse about its minor axis. The shape of the ellipsoid is therefore described by two geometric parameters, the semimajor axis a and the semiminor axis b. Generally, b is replaced by a smaller parameter which is more suitable: the (geometrical) flattening f.

$$f = (a - b)/a \qquad \text{(B-2)}$$

Further definitions:

- ***Origin***

Earth's centre of mass

- ***Geographic (geodetic) latitude f***

Angle measured in the meridian plane between the equatorial (x, y)-plane and the surface normal at P

- ***Geographic (geodetic) longitude l***

Angle measured in the equatorial plane between the zero meridian (X-axis) and the meridian plane of P

Spatial Ellipsoidal Coordinate System

For the spatial determination of points on the physical surface of the earth (or in space) with respect to the rotational ellipsoid, the height h above the ellipsoid is introduced in addition to the geographic coordinates ϕ, λ. The ellipsoidal height 'h' is measured along the surface normal (Fig. B-6).

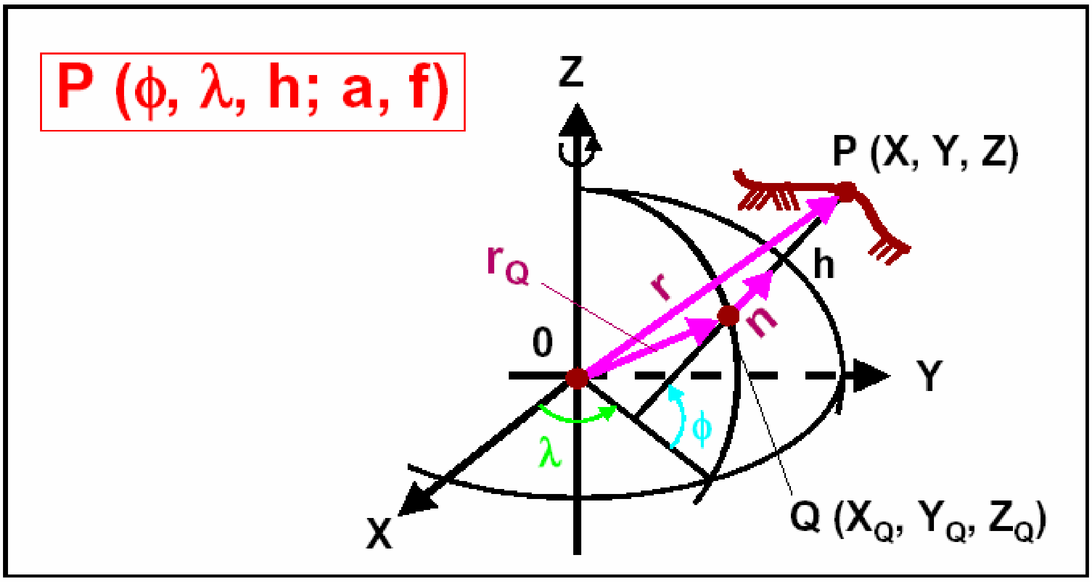

Fig. B-6. *Spatial ellipsoidal coordinate system*

The *spatial ellipsoidal coordinates* ϕ, λ, h are designated as *geodetic coordinates*. The point Q on the ellipsoid is obtained by projecting the surface (or space) point P along the ellipsoidal normal. A point in space is defined by (ϕ, λ, h) *and* the shape of the ellipsoid (a, f).
A standard earth model as a geodetic reference body should guarantee a good fit to the earth's surface and to the external gravity field; but also, it should possess a simple principle of formation.
In this respect, the *rotational ellipsoid*, already introduced as a geometric reference surface, is well suited. In addition to the semimajor axis a and the flattening f as geometric parameters, the total mass M and the rotational angular velocity ω as physical parameters are introduced. The gravity field is then formed as a result of gravitation and rotation.
If we now require the surface of this ellipsoid to be a level surface of its own gravity field then, according to Stokes Theorem, the gravity field is uniquely defined in the space exterior to this surface. This body is known as a *level*

(or *equipotential*) *ellipsoid*. Additionally, the geocentric gravitational constant GM and the dynamic flattening C2 0 , (2nd order zonal harmonic of an earth gravity model) are given. If the ellipsoidal parameters are given those values which correspond to the real earth, then this yields the optimum approximation to the geometry of the geoid and to the external gravity field: *mean earth ellipsoid.*

GEODETIC DATUM

Definitions

The terminology required to describe the geodetic datum problem is rather complex and has developed over more than 100 years. In order to avoid confusion and misunderstanding care will be taken to use the various terms precisely.

The following definitions are adopted by the international geodetic community:

Geodetic reference system (GRS)

- Conceptual idea of an earth-fixed Cartesian system (X, Y, Z)

Geodetic reference frame

- Practical realization of a geodetic reference system by observations It is important to make a difference between a reference system and a reference frame.

A *reference system* is the conceptual idea of a particular coordinate system; (theoretical definition).

A *reference frame* is the practical realization of a reference system by observations and measurements (which have errors). In practical surveying we are only concerned with reference frames, but the underlying concepts of a specific reference frame are of fundamental importance.

Global GRS

- Origin: Earth’s centre of mass
- Z-axis: Coincides with mean rotational axis of Earth
- X-axis: Mean meridian plane of Greenwich and ± to Z-axis
- Y-axis: Orthogonal

Local GRS

- Origin and orientation of axes is "arbitrary"

Geodetic datum

- Minimum set of parameters required to define location and orientation of the local system with respect to the global reference system/frame

Furthermore, it is important to distinguish between global and local reference frames. Looking at the entire set of possible reference frames located in the body of the Earth there is only one truly global reference system. The origin of a global reference system coincides with the centre of the Earth, the Z-axis should coincide with the mean rotational axis of Earth and the X-axis is contained in the mean meridian plane of Greenwich and is perpendicular to the Z-axis. The Y-axis is orthogonal to both the X- and Z-axis (right hand system).

A geodetic datum is expressed in terms of the set of transformation parameters which are required to define the location and orientation of the local frame with respect to the global one.

Note: The term “datum” is often used when one actually means “reference frame”.

What is a Geodetic Datum ?

We have to distinguish between a *Cartesian datum* and an *ellipsoidal datum*.

A Cartesian datum is defined by a set of:

- 3 shifts: ΔX, ΔY, ΔZ
- 3 rotations: α, β, γ
- A scale factor parameter: μ

These 7 parameters are needed to relate two Cartesian 3-d reference frames.

Because the Earth is a curved surface, approximated by an ellipsoid, navigators usually work in geographical coordinates (latitude, longitude). In order to define geographical coordinates the shape of the so called reference ellipsoid has also to be considered. The shape of an ellipsoid is defined by its semi-major and semi-minor axes, i.e. two additional parameters are required. These two additional parameters constitute the difference between a Cartesian and an ellipsoidal datum. Thus, an ellipsoidal datum is defined by 9 transformation parameters.

Rule of thumb:

Ellipsoidal Datum = Cartesian Datum + Shape of Earth Ellipsoid

TRANSFORMATIONS

A geodetic datum transformation is a mathematical rule used to transform surveyed coordinates given in a Reference Frame 1 into coordinates given in Reference Frame 2. The mathematical rule is a function of the set of necessary datum transformation parameters.

The nine parameters

- translation of the origin ΔX, ΔY, ΔZ,
- rotation angles εx, εy, εz,
- Scale factor μ,
- change in ellipsoidal semimajor axis Δa and flattening Δf define the location and orientation of a (local) coordinate system with respect to a global reference frame.

These parameters are needed for a computational coordinate transformation using Helmet's formula.

THE HEIGHT PROBLEM

What is a "Height"

Usually, the implicit imagination behind the term "height" is the answer to the question: Where does water flow? Physically, we consider a lake where water is in rest as a surface of equal heights. More specific, it is an earth's gravity potential surface. Moving on such a surface means no work is carried out, no forces are acting on it. Thus, the definition of a height with such a physical meaning cannot be defined geometrically nor can the reference surface (zero surface) be the ellipsoid's geometrical surface.

Geodetic Networks

The application of differential GPS satellite observations delivers

- Horizontal WGS 84 coordinates: ellipsoidal latitude ϕ and longitude λ,
- Vertical WGS 84 coordinates: ellipsoidal height h.

The ellipsoidal height does not have a physical meaning; it is a geometric quantity which does not indicate a level surface (i.e. it does not indicate the direction of flow of water).

Geodetic networks consist, in general, of geometrically defined and referred ellipsoidal latitude and longitude, whereas national heights refer to the geoid ("mean sea level") as zero surface.

The Geoid as Reference Surface for Heights

The geoid can be considered as an idealized ocean extending under the continents (Fig. B-7). It is a particular equipotential gravity surface of the earth coinciding with approximately two thirds of the worlds surface. There is only one geoid.

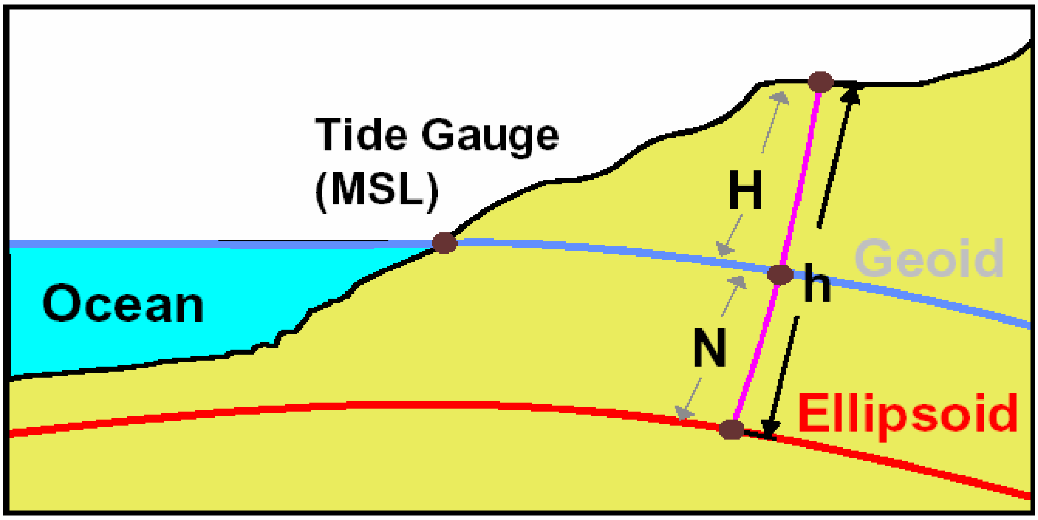

Fig. B-7. The geoid as reference surface for heights

The geoid is realized in practice by observing "mean sea level" (MSL) at tide gauges at the coasts over a certain time period. However, there are certain complications brought about by wind, salinity, currents, etc. producing deviations from the geoid of up to 2 m ("sea surface topography"). This means, that the zero point, and consequently the heights, of different national networks may differ by similar magnitudes. Heights above the geoid are called "*orthometric heights H*". The relation between an ellipsoidal height h and H is given by

$$H = h - N \qquad \text{(B-3)}$$

Where N is the geoid height.

Vertical Datum Problem

WGS 84 is a 3-dimensional reference frame coordinated in X, Y, Z or in ϕ, λ, h. The parameter h is the (geometric) height above the WGS 84 ellipsoid.

In aviation, heights (flight level) are defined by atmospheric pressure. All aircraft are therefore equipped with baro-altimeters. For this reason the ICAO approach has been to initially use only WGS 84 geographical coordinates (ϕ,λ) and to exclude the geometric height (h) from consideration. However, because ICAO is considering the technical issues, surveyors are advised to measure and report the heights of navigation facilities, if the necessary field work and computation can be incorporated with any re-surveys of plan positions.

One has to be very careful when dealing with heights. The differences between the different zero points of national vertical networks may vary up to 3 m!

Presently, there is worldwide effort to come up with an unified height system. It is hoped that this zero surface (namely the geoid) can be determined world-wide to an accuracy < 20 cm by using satellite altimetry.