3-1-2005

What Does Height Really Mean? Part II: Physics and Gravity

Thomas H. Meyer
University of Connecticut, thomas.meyer@uconn.edu

Daniel R. Roman
National Geodetic Survey

David B. Zilkoski
National Geodetic Survey

What Does *Height* Really Mean?
Part II: Physics and Gravity[1]

Thomas H. Meyer, Daniel R. Roman, and David B. Zilkoski

ABSTRACT: This is the second paper in a four-part series considering the fundamental question, "what does the word height really mean?" The first paper in this series explained that a change in National Geodetic Survey's policy, coupled with the modern realities of GPS surveying, have essentially forced practicing surveyors to come to grips with the myriad of height definitions that previously were the sole concern of geodesists. The distinctions between local and equipotential ellipsoids were considered, along with an introduction to mean sea level. This paper brings these ideas forward by explaining mean sea level and, more importantly, the geoid. The discussion is grounded in physics from which gravitational force and potential energy will be considered, leading to a simple derivation of the shape of the Earth's gravity field. This lays the foundation for a simplistic model of the geoid near Mt. Everest, which will be used to explain the undulations in the geoid across the entire Earth. The terms *geoid, plumb line, potential, equipotential surface, geopotential number,* and *mean sea level* will be explained, including a discussion of why mean sea level is not everywhere the same height; why it is not a level surface.

Introduction: Why Care About Gravity?

Any instrument that needs to be leveled in order to properly measure horizontal and vertical angles depends on gravity for orientation. Surveying instruments that measure gravity-referenced heights depend upon gravity to define their datum. Thus, many surveying measurements depend upon and are affected by gravity. This second paper in the series will develop the physics of gravity, leading to an explanation of the geoid and geopotential numbers.

The direction of the Earth's gravity field stems from the Earth's rotation and the mass distribution of the planet. The inhomogeneous distribution of that mass causes what are known as *geoid undulations,* the geoid being defined by the National Geodetic Survey (1986) as "The equipotential surface of the Earth's gravity field which best fits, in a least squares sense, global mean sea level." The geoid is also called the "figure of the Earth." Quoting Shalowitz (1938, p. 10), "The true figure of the Earth, as distinguished from its topographic surface, is taken to be that surface which is everywhere perpendicular to the direction of the force of gravity and which coincides with the mean surface of the oceans." The direction of gravity varies in a complicated way from place to place. Local vertical remains perpendicular to this undulating surface, whereas local normal remains perpendicular to the ellipsoid reference surface. The angular difference of these two is the *deflection of the vertical*.

The deflection of the vertical causes angular traverse loop misclosures, as do instrument setup errors, the Earth's curvature, and environmental factors introducing errors into measurements. The practical consequence of the deflection of the vertical is that observed angles differ from the angles that result from the pure geometry of the stations. It is as if the

Thomas H. Meyer, Department of Natural Resources Management and Engineering, University of ConnecticutStorrs, CT 06269-4087. Tel: (860) 486-2840; Fax: (860) 486-5480. E-mail:<thomas.meyer@uconn.edu>. **Daniel R. Roman,** National Geodetic Survey, 1315 East-West Highway, Silver Springs, MD 20910. E-mail: <Dan.Roman@noaa.gov>. **David B. Zilkoski,** National Geodetic Survey, 1315 East-West Highway, Silver Springs, MD 20910. E-mail: <Dave.Zilkoski@noaa.gov>.

[1]Throughout the series we will enumerate figures, tables, and equations with a Roman numeral indicating the paper in the series from which it came. For example, the third figure in the second paper will be numbered, "Figure II.3".

observing instrument were misleveled, resulting in traverses that do not close. This is true for both plane and geodetic surveying, although the effect for local surveys is seldom measurable because geoid undulations are smooth and do not vary quickly over small distances. Even so, it should be noted that the deflection of the vertical can cause unacceptable misclosures even over short distances. For example, Shalowitz (1938, p. 13, 14) reported deflections of the vertical created discrepancies between astronomic coordinates and geodetic (computed) coordinates up to a minute of latitude in Wyoming. In all cases, control networks for large regions cannot ignore these discrepancies, and remain geometrically consistent, especially in and around regions of great topographic relief. Measurements made using a gravitational reference frame are reduced to the surface of a reference ellipsoid to remove the effects of the deflection of the vertical, skew of the normals, topographic enlargement of distances, and other environmental effects (Meyer 2002).

The first article in this series introduced the idea that mean sea level is not at the same height in all places. This fact led geodesists to a search for a better surface than mean sea level to serve as the datum for vertical measurements, and that surface is the *geoid*. Coming to a deep understanding of the geoid requires a serious inquiry (Blakely 1995; Bomford 1980; Heiskanen and Moritz 1967; Kellogg 1953; Ramsey 1981; Torge 1997; Vanicek and Krakiwsky 1996), but the concepts behind the geoid can be developed without having to examine all the details. The heart of the matter lies in the relationship between gravitational force and gravitational potential. Therefore, we review the concepts of force, work, and energy so as to develop the framework to consider this relationship.

Physics

Force, Work, and Energy

Force is what makes things go. This is apparent from Newton's law, $\mathbf{F} = m\,\mathbf{a}$, which gives that the acceleration of an object is caused by, and is in the direction of, a force **F** and is inversely proportional to the object's mass m. Force has magnitude (i.e., strength) and direction. Therefore, a force is represented mathematically as a vector whose length and direction are set equal to those of the force. We denote vectors in bold face, either upper or lower case, e.g., **F** or **f,** and scalars in standard face, e.g., the speed of light is commonly denoted as c. Force has units of mass times length per second squared and is named the "newton," abbreviated N, in the meter-kilogram-second (mks) system.

There is a complete algebra and calculus of vectors (e.g., see Davis and Snider (1979) or Marsden and Tromba (1988)), which will not be reviewed here. However, we remind the reader of certain key concepts. Vectors are ordered sets of scalar components, e.g., (x,y,z) or $\mathbf{F} = (F_1,F_2,F_3)$, and we take the magnitude of a vector, which we denote as $|\mathbf{F}|$, to be the square root of the sum of the components: $|\mathbf{F}| = \sqrt{F_1^2 + F_2^2 + F_3^2}$.

For example,
if $\mathbf{F} = (1,-4,2)$, then $|\mathbf{F}| = \sqrt{1^2 + (-4)^2 + 2^2} = \sqrt{21}$.

Vectors can be multiplied by scalars (e.g., $c\,\mathbf{A}$) and, in particular, the negative of a vector is defined as the scalar product of minus one with the vector: $-\mathbf{A} = -1\,\mathbf{A}$. It is easy to show that $-\mathbf{A}$ is a vector of magnitude equal to **A** but oriented in the opposite direction. Division of vectors by scalars is simply scalar multiplication by a reciprocal: $\mathbf{F}/c = 1/c\ \mathbf{F}$. A vector **F** divided by its own length results in a **unit vector**, being a vector in the same direction as **F** but having unit length—a length of exactly one. We denote a unit vector with a hat: $\hat{\mathbf{F}} = \mathbf{F}/|\mathbf{F}|$.

Vectors can be added (e.g., $\mathbf{A} + \mathbf{B}$) and subtracted, although subtraction is defined in terms of scalar multiplication by -1 and vector addition (i.e., $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$). The result of adding/subtracting two vectors is another vector; likewise with scalar multiplication. By virtue of vector addition (the law of superposition), any vector can be a composite of any finite number of vectors:

$$\mathbf{F} = \sum_{i=1}^{n} \mathbf{f}_i,\ n < \infty.$$

The *inner* or *scalar* product of two vectors $\mathbf{a} \cdot \mathbf{b}$ is defined as:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| \times |\mathbf{b}| \cos\theta \qquad \text{(II.1)}$$

where θ is the angle between **a** and **b** in the plane that contains them. In particular, note that if **a** is perpendicular to **b**, then $\mathbf{a} \cdot \mathbf{b} = 0$ because cos 90° = 0. We will make use of the fact that the inner product of a force vector with a unit vector is a scalar equal to the magnitude of the component of the force that is applied in the direction of the unit vector.

Newton's law of gravity specifies that the gravitational force exerted by a mass M on a mass m is:

$$\mathbf{F}_g = -\frac{GMm\hat{\mathbf{r}}}{|\mathbf{r}|^2} \quad \text{(II.2)}$$

where:

G = universal gravitational constant; and

$\mathbf{r}$ = a vector from M's center of mass to m's center of mass.

The negative sign accounts for gravity being an attractive force by orienting $\mathbf{F}_g$ in the direction opposite of $\hat{\mathbf{r}}$ (since $\hat{\mathbf{r}}$ is the unit vector from M to m, $\mathbf{F}_g$ needs to be directed from m to M). In light of the discussion above about vectors, Equation (II.2) is understood to indicate that the magnitude of gravitational force is in proportion to the masses of the two objects, inversely proportional to the square of the distance separating them, and is directed along the straight line joining their centroids.

In geodesy, M usually denotes the mass of the Earth and, consequently, the product $G\ M$ arises frequently. Although the values for G and M are known independently (G has a value of approximately 6.67259×10^{-11} m^3 s^{-2} kg^{-1} and M is approximately 5.9737×10^{-24} kg), their product can be measured as a single quantity and its value has been determined to have several, nearly identical values, such as GM=398600441.5±0.8x10^6 m^3 s^2 (Groten 2004).

Gravity is a force field, meaning that the gravity created by any mass permeates all of space. One consequence of superposition is that gravity fields created by different masses are independent of one another. Therefore, it is reasonable and convenient to consider the gravitational field created by a single mass without taking into consideration any objects within that field. Equation (II.2) can be modified to describe a gravitational field simply by omitting m. We can compute the strength of the Earth's gravitational field at a distance equal to the Earth's equatorial radius (6,378,137 m) from the center of M by:

$$\mathbf{E}_g = -\frac{GM\hat{\mathbf{r}}}{|\mathbf{r}|^2} \quad \text{(II.3)}$$

$$= -\frac{398600441.5\ \mathrm{m^3 s^2} \cdot \hat{\mathbf{r}}}{(6378137\ \mathrm{m/s})^2}$$

$$= 9.79829\ \mathrm{m/s^2}(-\hat{\mathbf{r}}) \quad \text{(II.4)}$$

This value is slightly larger than the well-known value of 9.78033 m/s^2 because the latter includes the effect of the Earth's rotation.[2] We draw attention to the fact that Equation (II.3) has units of acceleration, not a force, by virtue of having omitted m.

It is possible to use Equation (II.3) to draw a picture that captures, to some degree, the shape of the Earth's gravitational field (see Figure II.1). The vectors in the figure indicate the magnitude and direction of force that would be experienced by unit mass located at that point in space. The vectors decrease in length as distance increases away from the Earth and are directly radially towards the Earth's center, as expected. However, we emphasize that the Earth's gravitational field pervades all of space; it is not discrete as the figure suggests. Furthermore, it is important to realize that, in general, any two points in space experience a different gravitational force, if perhaps only in direction.

We remind the reader that the current discussion is concerned with finding a more suitable vertical datum than mean sea level, which is, in some sense, the same thing as finding a better way to measure heights. Equation (II.3) suggests that height might be inferred by measuring gravitational force because Equation (II.3) can be solved for the magnitude of $\mathbf{r}$, which would be a height measured using the Earth's center of gravity as its datum. At first, this approach might seem to hold promise because the acceleration due to gravity can be measured with instruments that carefully measure the acceleration of a standard mass, either as a pendulum or free falling (Faller and Vitouchkine 2003). It seems such a strategy would deduce height in a way that stems from the physics that give rise to water's downhill motion and, therefore, would capture the primary motivating concept behind height very well. Regrettably, this is not the case and we will now explain why.

Suppose we use gravitational acceleration as a means of measuring height. This implies that surfaces of equal acceleration must also be level surfaces, meaning a surface across which water does not run without external impetus. Thus, our mean sea level surrogate is that set of places

[2] The gravity experienced on and around the Earth is a combination of the gravitation produced by the Earth's mass and the centrifugal force created by its rotation. The force due solely to the Earth's mass is called gravitational and the combined force is called gravity. For the most part, it will not be necessary for the purposes of this paper to draw a distinction between the two. The distinction will be emphasized where necessary.

that experience some particular gravitational acceleration; perhaps the acceleration of the normal gravity model, g_0, would be a suitable value. The fallacy in this logic comes from the inconsideration of gravity as a vector; it is not just a scalar. In fact, the heart of the matter lies not in the *magnitude* of gravity but, rather, in its *direction*.

If a surface is level, then water will not flow across it due to the influence of gravity alone. Therefore, a level surface must be situated such that all gravity force vectors at the surface are perpendicular to it; none of the force vectors can have any component directed across the surface. Figure II.2 depicts a collection of force vectors that are mutually perpendicular to a horizontal surface, so the horizontal surface is level, but the vectors have differing magnitudes. Therefore, it is apparent that choosing a surface of equal gravitational acceleration (i.e., magnitude) does not guarantee that the surface will be level. Of course, we have not shown that this approach necessarily would not produce level surfaces. It might be the case that it happens that the magnitude of gravity acceleration vectors just happen to be equal on level surfaces. However, as we will show below, this is not the case due to the inhomogeneous distribution of mass within the Earth.

We can use this idea to explain why the surface of the oceans is not everywhere the same distance to the Earth's center of gravity. The first article in this series noted several reasons for this, but we will discuss only one here. It is known that the salinity in the oceans is not constant. Consequently, the density of the water in the oceans is not constant, either, because it depends on the salinity. Suppose we consider

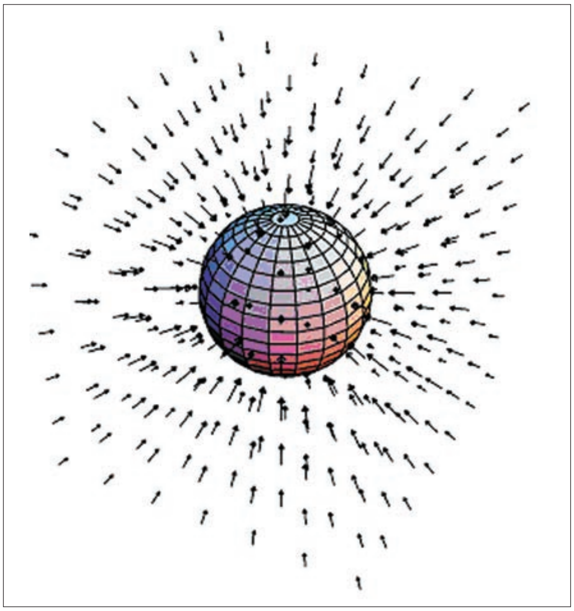

Figure II.1. The gravitational force field of a spherical Earth. Note that the magnitude of the force decreases with separation from the Earth.

columns of water along a coast line and suppose that gravitational acceleration is constant along the coasts (see Figure II.3). In particular, consider the columns A and B. Suppose the water in column A is less dense than in column B; perhaps a river empties into the ocean at that place. We have assumed or know that:

- The force of gravity is constant,
- The columns of water must have the same weight in order to not flow, and
- The water in column A is less dense than that in column B.

It takes more water of lesser density to have the same mass as the amount of water needed of greater density. Water is nearly incompressible, so the water column at A must be taller than the column of water at B. Therefore, a mean sea level station at A would not be at the same distance from the Earth's center of gravity as a mean sea level station at B.

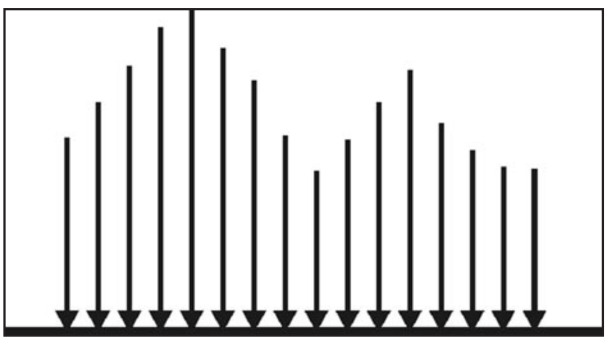

Figure II.2. A collection of force vectors that are all normal to a surface (indicated by the horizontal line) but of differing magnitudes. The horizontal line is a level surface because all the vectors are normal to it; they have no component directed across the surface.

As another example showing why gravitational force is not an acceptable way to define level surfaces, Figure II.4 shows the force field generated by two point–unit masses located at (0,1) and (0,-1). Note the lines of symmetry along the x and y axes. All forces for places on the x-axis are parallel to the axis and directed towards (0,0). Above or below the x-axis, all force lines ultimately lead to the mass also located on that side. Figure II.5 shows a plot of the magnitude of the vectors of Figure II.4. Note the local maxima around $x \pm 1$ and the local minima at the origin. Figure II.6 is

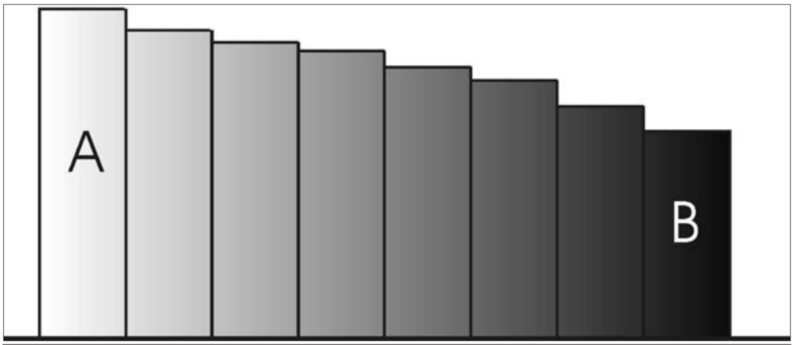

Figure II.3. A collection water columns whose salinity, and therefore density, has a gradient from left to right. The water in column A is least dense. Under constant gravity, the height of column A must be greater than B so that the mass of column A equals that of column B.

a plot of the "north-east" corner of the force vectors superimposed on top of an isoforce plot of their magnitudes (i.e., a "contour plot" of Figure II.5). Note that the vectors are not perpendicular to the isolines. If one were to place a drop of

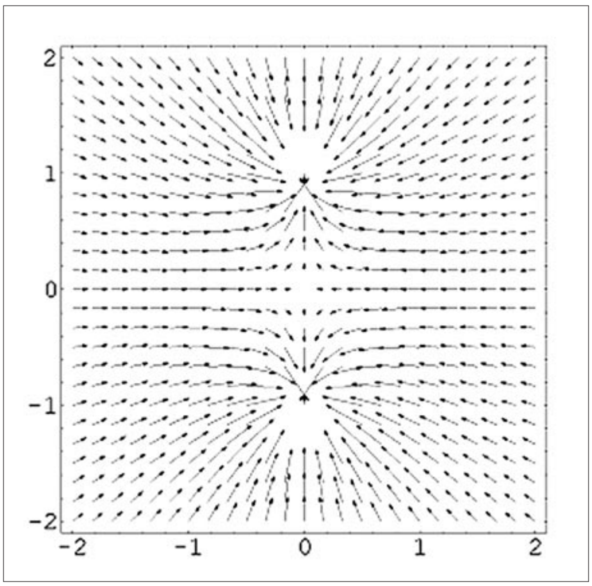

Figure II.4. The force field created by two point masses.

water anywhere in the space illustrated by the figure, the water would follow the vectors to the peak and would both follow and cross isoforce lines, which is nonsensical if we take isoforce lines to correspond to level surfaces. This confirms that equiforce surfaces are not level.

These three examples explain why gravitational acceleration does not lead to a suitable vertical datum, but they also provide a hint where to look. We require that water not flow between two points of equal height. We know from the first example that level surfaces have gravity force vectors that are normal to them. The second example illustrated that the key to finding a level surface pertains to *energy* rather than force, because the level surface in Figure II.3 was created by equalizing the weight of the water columns. This is related to potential energy, which we will now discuss.

Work and Gravitational Potential Energy

Work plays a direct role in the definition of the geoid because it causes a change in the potential energy state of an object. In particular, when work is applied against the force of gravity causing an object to move against the force of gravity, that object's potential energy is increased, and this is an important concept in understanding the geoid. Therefore, we now consider the physics of work.

Work is what happens when a force is applied to an object causing it to move. It is a scalar quantity with units of distance squared times mass per second squared, and it is called the "joule," abbreviated J, in the mks system. Work is computed as force multiplied by distance, but only the force that is applied in the direction of motion contributes to the work done on the object.

Suppose we move an object in a straight line. If we denote a constant force by **F** and the displacement of the object by a vector **s**, then the work done on the object is W=F·s (Equation (II.1)). This same expression would be correct even if **F** is not directed exactly along the path of motion, because the inner product extracts from **F** only that portion that is directed parallel to **s**. Of course, in general, force can vary with position, and the path of motion might not be a straight line. Let *C* denote a curve that has been parameterized by arc length *s*, meaning that **p** = *C* (*s*) is a point on *C* that is *s* units from *C*'s starting point. Let $\hat{\mathbf{t}}(s)$ denote a unit vector tangent to *C* at *s*. Since we want to allow force to vary along *C*, we adopt a notion that the force is a function of position **F**(*s*). Then, by application of the calculus, the work expended by the application of a possibly varying force along a possibly curving path *C* from $s = s_0$ to $s = s_1$ is:

$$W = \int_{s_0}^{s_1} \mathbf{F}(s) \cdot \hat{\mathbf{t}}(s)\, ds \qquad \text{(II.5)}$$

Equation (II.5) is general so we will use it as we turn our attention to motion within a gravitational force field. Suppose we were to move some object in the presence of a gravitational force field. What would be the effect? Let us first suppose that we move the object on a level surface, which implies that the direction of the gravitational force vector is everywhere normal to that surface and, thus, perpendicular to $\hat{\mathbf{t}}(s)$, as well. Since by assumption $\mathbf{F}_g$ is perpendicular to $\hat{\mathbf{t}}(s)$, $\mathbf{F}_g$ plays no part in the work being done because $\mathbf{F}_g(s)\cdot\hat{\mathbf{t}}(s)=0$. Therefore, moving an object over a level surface in a gravity field is identical to moving it in the absence of the field altogether, as far as the work done against gravity is concerned.

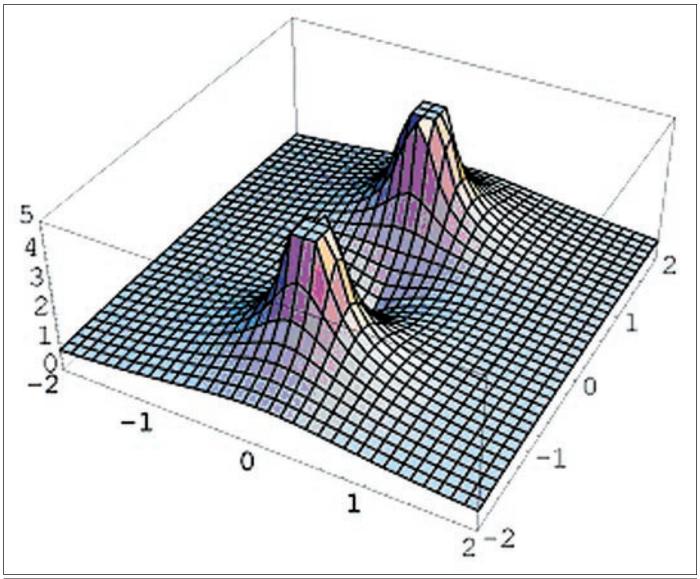

Figure II.5. The magnitude of the force field created by two point masses.

Now, suppose that we move the object along a path such that the gravitational force is not everywhere normal to the direction of motion. From Equation (II.5) it is evident that either more or less work will be needed due to the force of gravity, depending on whether the motion is against or with gravity, respectively. The gravity force will simply be accounted for by adding it to force we apply; the object can make no distinction between them. Indeed, we can use superposition to separate the work done in the same direction as gravity from the work done to move laterally through the gravity field; they are orthogonal. We now state, without proof, a critical result from vector calculus: the work done by gravity on a moving body does not depend on the path of motion, apart from the starting and ending points. This is a consequence of gravity being a conservative field (Blakely 1995; Schey 1992). As a result, the work integral along the curve defining the path of motion can be simplified to consider work only in the direction of gravity. This path is called a *plumb line* and, over short distances, can be considered to be a straight line, although the force field lines shown in Figure II.6 show that plumb lines are not straight, in general. Therefore, from Equation (II.5), the work needed to, say, move some object vertically through a gravity field is given by:

$$W=\int_{h_0}^{h_1}\mathbf{F}_g(h)\cdot\hat{\mathbf{t}}(h)\,dh \qquad \text{(II.6)}$$

where:

h = height (distance along the plumbline); and

$\hat{\mathbf{t}}(h)$ = the direction of gravity.

However, $\mathbf{F}_g(h)$ is always parallel to $\hat{\mathbf{t}}(h)$, so

$$\mathbf{F}_g(h)\cdot\hat{\mathbf{t}}(h)=\pm F_g(h),$$

depending on whether the motion is with or against gravity. If we assume $F_g(h)$ is constant, Equation (II.6) can be simplified as:

$$W=\int_{h_0}^{h_1}\mathbf{F}_g(h)\cdot\hat{\mathbf{t}}(h)dh \qquad \text{(II.6)}$$

$$=\int_{h_0}^{h_1}m\mathbf{E}_g(h)\cdot\hat{\mathbf{t}}(h)dh \qquad \text{(II.3)}$$

$$=mE_g\int_{h_0}^{h_1}dh \qquad \text{assuming } E_g \text{ is constant}$$

$$=m\,g\,\Delta h \qquad \text{(II.7)}$$

where we denote the assumed constant magnitude of gravitational acceleration at the Earth's surface by g, as is customary. The quantity $m\,g\,h$ is called *potential energy*, so Equation (II.7) indicates that the release of potential energy will do work if the object moves along gravity force lines. The linear dependence of Equation (II.7) on height (h) is a key concept.

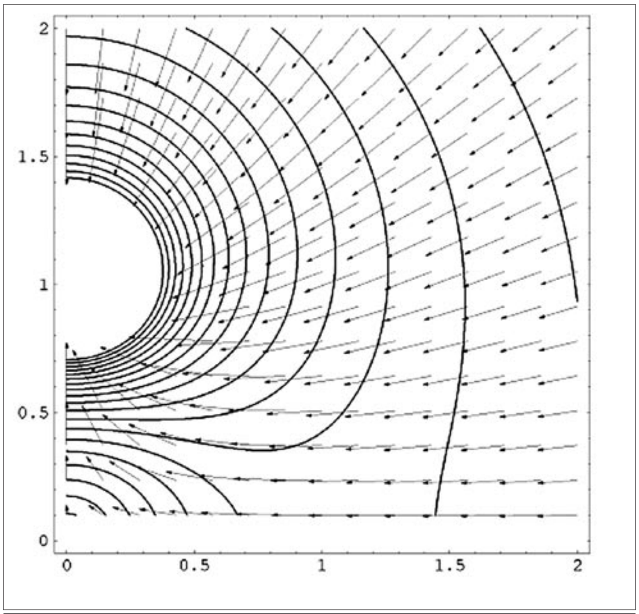

Figure II.6. The force field vectors shown with the isoforce lines of the field. Note that the vectors are not perpendicular to the isolines thus illustrating that equiforce surfaces are not level.

The Geoid

What is the Geoid?

Although Equation (II.7) indicates a fundamental relationship between work and potential energy, we do not use this relationship directly because it is not convenient to measure work to find potential. Therefore, we rely on a direct relationship between the Earth's potential field and its gravity field that we state without justification: (II.8)

$$E_g = \nabla U$$

where:

U = the Earth's potential field; and

∇ = gradient operator.[3] Written out in Cartesian coordinates, Equation (II.8) becomes:

$$\mathbf{E}_g = \frac{\partial U}{\partial x}\hat{i} + \frac{\partial U}{\partial y}\hat{j} + \frac{\partial U}{\partial z}\hat{k}$$

where $\hat{i}, \hat{j}$, and $\hat{k}$ are unit vectors in the *x, y,* and *z* directions, respectively. In spherical coordinates, Equation (II.8) becomes:

$$\mathbf{E}_g = \frac{\partial U}{\partial r}\hat{\mathbf{r}} \qquad \text{(II.9)}$$

Equation (II.8) means that the gravity field is the gradient of the potential field. For full details, the reader is referred to the standard literature, including (Blakely 1995; Heiskanen and Moritz 1967; Ramsey 1981; Torge 1997; Vanicek and Krakiwsky 1996). Although Equation (II.8) can be proven easily (Heiskanen and Moritz 1967, p.2), the intuition behind the equation does not seem to be so easy to grasp.

We will attempt to clarify the situation by asking the reader to consider the following, odd, question: why do air bubbles go upwards towards the surface of the water? The answer that is usually given is because air is lighter than water. This is surely so but $\mathbf{F} = m\,\mathbf{a}$, so if bubbles are moving, then there must be a force involved. Consider Figure II.7, which shows a bubble, represented by a circle, which is immersed in a water column. The horizontal lines indicate water pressure. The pressure exerted by a column of water increases nearly linearly with depth (because water is nearly incompressible). The water exerts a force inwards on the bubble from all directions, which are depicted by the force vectors. If the forces were balanced, no motion would occur. It would be like a rope in a tug-of-war in which both teams are equally matched. Both teams are pulling the rope but the rope is not moving: equal and opposite forces cause no motion.

However, the bubble has some finite height: the depth of the top of the bubble is less than the depth of the bottom of the bubble. Therefore, the pressure at the top of the bubble is less than the pressure at the bottom, so the force on the top of the bubble is less than that at the bottom. This pressure gradient creates an excess of force from below that drives the bubble upwards. Carrying the thought further, the difference in magnitude between any two lines of pressure is the gradient of the force field; it is the potential energy of the force field. The situation with gravity is exactly analogous to the situation with water pressure. Any surface below the water at which the pressure is constant might be called an "equipressure" surface. Any surface in or around the Earth upon which the gravity potential is constant is called an *equipotential surface*. Thus, a gravity field is caused by the difference in

[3] Other authors write Equation (II.8) as $\mathbf{E}_g = -\nabla U$, but the choice of the negative sign is essentially one of perspective: if the negative sign is included, the equation describes work done to overcome gravity. We prefer the opposite perspective because Equation (II.8) follows directly from Equation (II.3), in which the negative sign is necessary to capture the attractive nature of gravitational force.

the gravity potential of two infinitely close gravity equipotential surfaces.

By assuming a spherical, homogeneous, non-rotating Earth, we can derive its potential field from Equation (II.9), and denoting $|\mathbf{r}|$ by r:

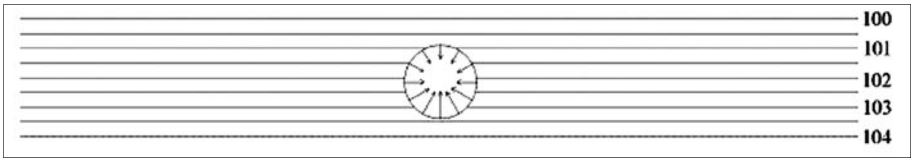

Figure II.7. The force experienced by a bubble due to water pressure. Horizontal lines indicate surfaces of constant pressure, with sample values indicated on the side.

$$\frac{\partial U}{\partial r}\hat{\mathbf{r}} = \mathbf{E}_g$$
$$\int dU = -\int \frac{GM}{r^2}dr$$
$$U = \frac{GM}{r} + c \qquad \text{(II.10)}$$

The constant of integration in Equation (II.10) can be chosen so that zero potential resides either infinity far away or at the center of M. We choose the former convention. Consequently, potential increases in the direction that gravity force vectors point and the absolute potential of an object of mass m located a distance h from M is:

$$\begin{aligned} U &= -\int_{\infty}^{h} \frac{GMm}{r^2}dr \\ &= \frac{GMm}{r}\Big|_{\infty}^{h} \\ &= \frac{GMm}{h} - \frac{GMm}{\infty} \\ &= \frac{GMm}{h} \end{aligned} \qquad \text{(II.11)}$$

We now reconsider the definition of the geoid, being the equipotential surface of the Earth's gravity field that nominally defines mean sea level. From Equation (II.10), the geoid is some particular value of U and, furthermore, if the Earth were spherical, homogeneous, and not spinning, the geoid would also be located at some constant distance from the Earth's center of gravity. However, none of these assumptions are correct, so the geoid occurs at various distances from the Earth's center—it undulates.

One can prove mathematically that $\mathbf{E}_g$ is perpendicular to U. To illustrate this, see Figure II.8. The figure shows the force vectors as seen in Figure II.6 but superimposed over the potential field computed using Equation (II.10) instead of the magnitude of the force field. Notice that the vectors are perpendicular to the isopotential lines. Water would not flow along the isopotential lines; only across them. In three dimensions, the isopotential lines would be equipotential surfaces, such as the geoid.

The Shape of the Geoid

We now consider the shape of the geoid as it occurs for the real Earth. It is evident from Equation (II.10) that the equipotential surfaces of a spherical, homogeneous, non-rotating mass would be concentric, spherical shells—much like layers of an onion. If the sphere is very large, such as the size of the Earth, and we examined a relatively small region near the surface of the sphere, the equipotential surfaces would almost be parallel planes.

Now, suppose we add some mass to the sphere in the form of a point mass roughly equal to that of Mt. Everest positioned on the surface of the sphere. The resulting gravity force field and isopotential lines are shown in Figure II.9. The angles and magnitudes are exaggerated for clarity; the deflection of the vertical is very apparent. In particular, we draw attention to the

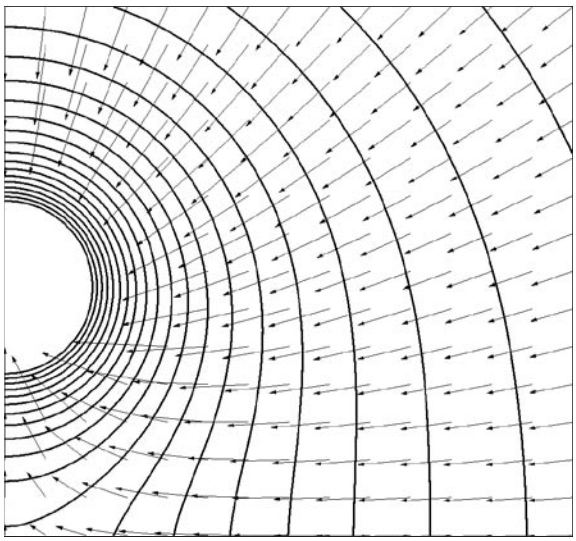

Figure II.8. The gravity force vectors created by a unit mass and the corresponding isopotential field lines. Note that the vectors are perpendicular to the field lines. Thus, the field lines extended into three dimensions constitute level surfaces.

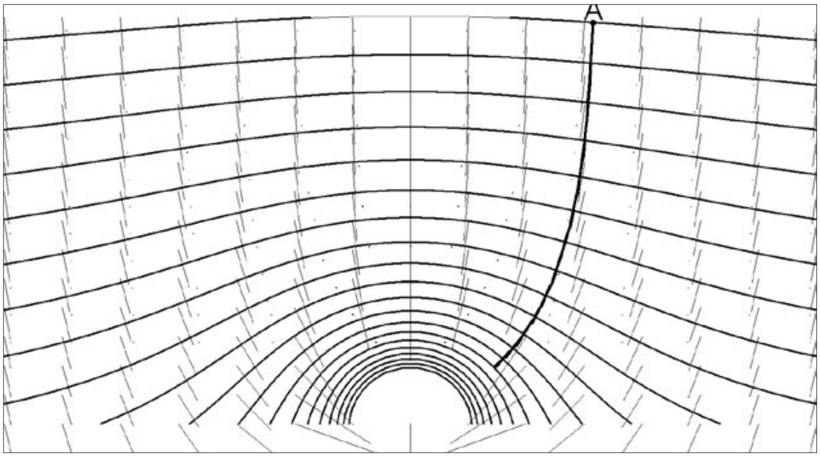

Figure II.9. The gravity force vectors and isopotential lines created at the Earth's surface by a point with mass roughly equal to that of Mt. Everest. The single heavy line is a plumb line.

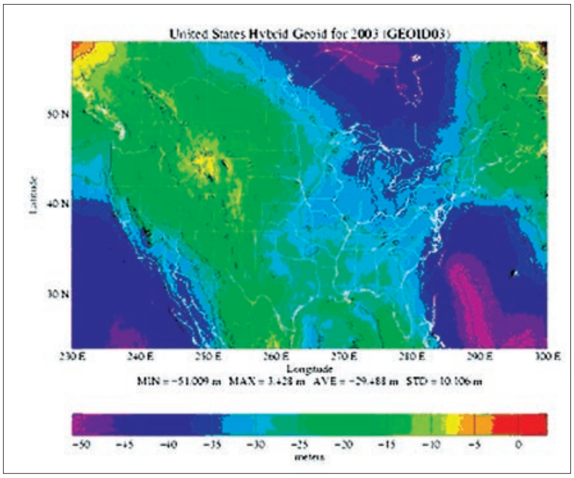

Figure II.10. GEOID03 local geoid model for the conterminous United States. From Roman et. al (2004).

shape of the isopotential lines which run more-or-less horizontally across the figure. Notice how they bulge up over the mountain. This is true in general: the equipotential surfaces roughly follow the topographic shape of the Earth in that they bow up over mountains and dip down into valleys. Also, any one of the geopotential lines shown in Figure II.9 can be thought of as representing the surface of the ocean above an underwater seamount. Water piles up over the top of subsurface topography to exactly the degree that the mass of the additional water exactly balances the excess of gravity caused by the seamount. Thus, one can indirectly observe seafloor topography by measuring the departure of the ocean's surface from nominal gravity (Hall 1992). The geoid, of course, surrounds the Earth, and Figure II.10 shows the ellipsoid height of the geoid with respect to NAD 83 over the conterminous United States as modeled by GEOID03 (Roman et al. 2004). At first glance, one could mistake the image for a topographic map. However, closer examination reveals numerous differences.

Geopotential Numbers

The geoid is usually considered the proper surface from which to reckon geodetic heights because it honors the flow of water and nominally resides at mean sea level. Sea level, itself, does not exactly match the geoid because of the various physical factors mentioned before. Therefore, actually finding the geoid in order to realize a usable vertical datum is currently not possible from mean sea level measurements. Ideally, one would measure potential directly in some fashion analogous to measuring gravity acceleration directly. If this were possible, the resulting number would be a *geopotential number*. In other words, a geopotential number is the potential of the Earth's gravity field at any point in space. Using geopotential numbers as heights is appealing for several reasons:

- Geopotential defines hydraulic head. Therefore, if two points are at the same geopotential number, water will not flow between them due to gravity alone. Conversely, if two points are not at the same geopotential number, gravity will cause the water to flow between them if the waterway is unobstructed (ignoring friction).
- Geopotential decreases linearly with distance from the center of the Earth (Equation (II.10)). This makes it a natural measure of distance.
- Geopotential does not depend on the path taken from the Earth's center to the point of interest. This makes a geopotential number stable.
- The magnitude of a geopotential number is less important than the relative values between two places. Therefore, one can scale geopotential numbers to any desirable values, such as defining the geoid to have a geopotential number of zero.

Equation (II.11) gives hope of determining height by measuring a gravity-related quantity, namely, absolute potential. Regrettably, potential cannot be measured directly. This is understandable because the manifestation of potential (the force of gravity) is created by potential differences, not in the potential itself. That is, two pairs of potential energies, say (150, 140) and (1000, 990) result in a force of the same magnitude. This is true because the difference of the two pairs is the same, namely, 10 newtons. In light of this, one might ask how images of the geoid, such as Figure (II.10), came into being. The image in Figure (II.10) is the result of a sophisticated mathematical model based on Stokes' formula, which we take from Heiskanen and Moritz' (1967, p. 94) equation 2-163b, and present here for completeness:

$$N = \frac{R}{4\pi G} \int_{\sigma} \Delta g \, S(\psi) d\sigma \qquad \text{(II.12)}$$

where:

N = geoid height at a point of interest;
R = mean radius of the Earth;
G = the universal gravitational constant;
σ = the surface of the Earth;
Δg = the reduced, observed gravity measurements around the Earth;
ψ = the spherical distance from each surface element $d\sigma$ to the point of interest, and
$S(\psi)$, which is known as Stokes' function, given by Heiskanen and Moritz' (1967, p. 94) equation 2-164:

$$S(\psi) = \frac{1}{\sin(\psi/2)} - 6\sin(\psi/2) + 1 - 5\cos\psi - 3\cos\psi \ln\left(\sin(\psi/2) + \sin^2(\psi/2)\right)$$

The model is calibrated with, and has boundary conditions provided by, reduced gravity measurements taken in the field—the Δg's in Equation (II.12). These measurements together with Stokes' formula permit the deduction of the potential field that must have given rise to the observed gravity measurements.

In summary, in spite of their natural suitability, geopotential numbers are not practical to use as heights because practicing surveyors cannot easily measure them in the field.[4] They are, however, the essence of what the word *height* really means, and subsequent papers in this series will come to grips with how orthometric and ellipsoid heights are related to geopotential numbers by introducing *Helmert orthometric heights and dynamic heights*.

Summary

This second paper in a four-part series that reviews the fundamental concept of *height* presented simple derivations of the physics concepts needed to understand the force of gravity, since mean sea level and the Earth's gravity field are strongly interrelated. It was shown that one cannot use the magnitude of the force of gravity to define a vertical datum because equiforce surfaces are not level surfaces. However, it was observed that gravity potential gives rise to gravity force and, furthermore, gravity force is normal to equipotential surfaces. The practical consequence of this is that water will not flow along an equipotential surface due to the force of gravity alone. Therefore, equipotential surfaces are level surfaces and suitable to define a vertical datum. In particular, although there is an infinite number of equipotential surfaces, the geoid is often chosen to be the equipotential surface of the Earth's gravity field that best fits mean sea level in a least squares sense, and the geoid has thus become the fundamental vertical datum for mapping. It was shown that mean sea level itself is not a level surface, therefore, one cannot deduce the location of the geoid by measuring the location of mean sea level alone. Furthermore, one cannot measure gravity potential directly. Therefore, we model the geoid mathematically, based on gravity observations.

A geopotential number was defined to be a number proportional to the gravity potential at that place. Geopotential numbers capture the notion of height exactly because they vary linearly with vertical distance and define level surfaces. However, they are usually unsuitable for use as distances themselves because they cannot be measured directly and have units of energy rather than length.

REFERENCES

Blakely, R.J. 1995. *Potential theory in gravity and magnetic applications*. Cambridge, U.K.: Cambridge University Press. 441p.

Bomford, G. 1980. *Geodesy*, 4th ed. Oxford, U.K.: Clarendon Press. 561p.

[4] Geopotential numbers have units of energy, not length. We suspect that most practicing surveyors would object to using heights that don't have length units, as well.

Davis, H.F., and A. D. Snider. 1979. *Introduction to vector analysis*. Boston, Massachusetts: Allyn and Bacon, Inc. 340p.

Faller, J.E., and A. L. Vitouchkine. 2003. Prospects for a truly portable absolute gravimeter. *Journal of Geodynamics* 35(4-5): 567-72.

Groten, E. 2004. Fundamental parameters and current (2004) best estimates of the parameters of common relevance to astronomy, geodesy, and geodynamics. *Journal of Geodesy* 77(10-11): 724-97.

Hall, S.S. 1992. *Mapping the next millennium*. New York, New York: Random House. 477p.

Heiskanen, W.A., and H. Moritz. 1967. *Physical geodesy*. New York, New York: W. H. Freeman & Co. 364p.

Kellogg, O.D. 1953. *Foundations of potential theory*. New York, New York: Dover Publications. 384p.

Marsden, J.E., and A. J. Tromba. 1988. *Vector calculus*, 3rd ed. New York, New York: W.H. Freeman and Company. 655p.

Meyer, T. H. 2002. Grid, ground, and globe: Distances in the GPS era. *Surveying and Land Information Science* 62(3): 179-202.

National Geodetic Survey. 1986. *Geodetic glossary*. Rockville, Maryland: National Oceanic and Atmospheric Administration. 274p.

Ramsey, A.S. 1981. *Newtonian attraction*. Cambridge, U.K.: Cambridge University Press. 184p.

Roman, D. R., Y. M. Wang, W. Henning, and J. Hamilton. 2004. Assessment of the new National Geoid Height Model, GEOID03. Presented at the 2004 Conference of the American Congress on Surveying and Mapping, Nashville, Tennessee.

Schey, H.M. 1992. *Div, grad, curl and all that: An informal text on vector calculus*. New York, New York: W.W. Norton & Company. 164p.

Shalowitz, A.L. 1938. The geographic datums of the Coast and Geodetic Survey. *US Coast and Geodetic Survey Field Engineers Bulletin* 12: 10-31.

Torge, W. 1997. *Geodesy*, 2nd ed. New York, New York: Walter du Gruyter. 264p.

Vanicek, P., and E. J. Krakiwsky. 1996. *Geodesy: The concepts*, 2nd ed. Amsterdam, Holland: Elsevier Scientific Publishing Company. 697p. ■

TRIG-STAR

A High School Trigonometry Skill Awards Program
Administered by the National Society of Professional Surveyors

««PURPOSE OF THE CONTEST««

- To promote excellence in the mastery of mathematics in high school;
- To honor high school students who have demonstrated their superior skill among classmates at the local, state and national levels;
- To acquaint high school students with the use and practical applications of mathematics in the surveying professions;
- To build an awareness of surveying and mapping as a profession among mathematically skilled high school students, career guidance counselors, and high school math teachers.

««SPONSORSHIP APPLICATION««

A local sponsor is needed for each participating High School. Only one TRIG-STAR will be allowed per high school per year. The local sponsor must apply to NSPS on this application form and list the name of the High School where the contest will be held. A contest packet containing contest instructions, a test packet, an award plaque (if ordered), and a sample news release will be forwarded by NSPS to the local sponsor. The local sponsor will be responsible for making all the local arrangements. The winner will be eligible to compete in the NSPS National Trig-Star Contest and contest information will be supplied with the contest packet.

High School Name (required) ______________________________

City and State (required) ______________________________

Local Sponsor ______________________________

Mailing Address ______________________________

City______________ **State**______________ **Zip**__________

Phone Number ( ) ______________ **Fax Number ( )** ______________

If email address is provided you will receive the contest packet by email.
Visit www.trig-star.info for resources to help your prepare for you presentation.

Email Address______________________________

Sponsorship fee: $45.00 (with plaque) and $35.00 (without plaque) Date:______________

Enclosed is $________ for _____ Packets. Makes checks payable to NSPS TRIG-STAR
Visa /MC # ______________________________Expiration Date _____/_____
Name on Card ______________________ Signature ______________________

Packets are mailed standard USPS. If you wish to receive packet Next Day by Federal Express or UPS please add $25.00 for shipping costs.

Send To: TRIG-STAR, 6 Montgomery Village Avenue, Suite #403, Gaithersburg, MD 20879

Phone: 240-632-9716 ext. 103 or 105 Fax: 240-632-1321 E-mail: trisha.milburn@acsm.net; www.trig-star.info