GEOPHYSICS, VOL. 66, NO. 6 (NOVEMBER-DECEMBER 2001); P. 1660–1668, 4 FIGS., 3 TABLES.

Tutorial

Ellipsoid, geoid, gravity, geodesy, and geophysics

Xiong Li* and Hans-Jürgen Götze‡

ABSTRACT

Geophysics uses gravity to learn about the density variations of the Earth's interior, whereas classical geodesy uses gravity to define the geoid. This difference in purpose has led to some confusion among geophysicists, and this tutorial attempts to clarify two points of the confusion. First, it is well known now that gravity anomalies after the "free-air" correction are still located at their original positions. However, the "free-air" reduction was thought historically to relocate gravity from its observation position to the geoid (mean sea level). Such an understanding is a geodetic fiction, invalid and unacceptable in geophysics. Second, in gravity corrections and gravity anomalies, the elevation has been used routinely. The main reason is that, before the emergence and widespread use of the Global Positioning System (GPS), height above the geoid was the only height measurement we could make accurately (i.e., by leveling). The GPS delivers a measurement of height above the ellipsoid. In principle, in the geophysical use of gravity, the ellipsoid height rather than the elevation should be used throughout because a combination of the latitude correction estimated by the International Gravity Formula and the height correction is designed to remove the gravity effects due to an ellipsoid of revolution. In practice, for minerals and petroleum exploration, use of the elevation rather than the ellipsoid height hardly introduces significant errors across the region of investigation because the geoid is very smooth. Furthermore, the gravity effects due to an ellipsoid actually can be calculated by a closed-form expression. However, its approximation, by the International Gravity Formula and the height correction including the second-order terms, is typically accurate enough worldwide.

INTRODUCTION

Geophysics has traditionally borrowed concepts of gravity corrections and gravity anomalies from geodesy. Their uncritical use has sometimes had unfortunate results. For example, the "free-air" reduction was historically interpreted by geodesists as reducing gravity from topographic surface to the geoid (mean sea level). This interpretation is a useful fiction for geodetic purposes, but is completely inappropriate for geophysics. In geophysics, gravity is used to learn about the density variations of the Earth's interior. In geodesy, gravity helps define the figure of the Earth, the geoid. This difference in purpose determines a difference in the way to correct observed data and to understand resulting anomalies.

Until a global geodetic datum is fully and formally accepted, used, and implemented worldwide, global geodetic applications require three different surfaces to be clearly defined. They are (Figure 1): the highly irregular topographic surface (the landmass topography as well as the ocean bathymetry), a geometric or mathematical reference surface called the ellipsoid, and the geoid, the equipotential surface that mean sea level follows.

Gravity is closely associated with these three surfaces. Gravity corrections and gravity anomalies have been traditionally defined with respect to the elevation. Before the emergence of satellite technologies and, in particular, the widespread use of the Global Positioning System (GPS), height above the geoid (i.e., the elevation) was the only height measurement we could make accurately, namely by leveling. The GPS delivers a measurement of height above the ellipsoid. Confusion seems to have arisen over which height to use in geophysics.

Published on Geophysics Online May 31, 2001. Manuscript received by the Editor August 9, 2000; Revised manuscript received February 26, 2001.
*Fugro-LCT Inc., 6100 Hillcroft, 5th Floor, Houston, Texas 77081. E-mail: xli@fugro.com.
‡Freie Universität Berlin, Institut für Geologie, Geophysik and Geoinformatik, Malteserstraße 74-100, D-12249 Berlin, Germany. E-mail: hajo@geophysik.fu-berlin.de.

This tutorial explains the concepts of, and relationships among, the ellipsoid, geoid, gravity, geodesy, and geophysics. We attempt to clarify the way to best compute gravity corrections given GPS positioning. In short, h, the ellipsoid height relative to the ellipsoid, is the sum of H, the elevation relative to the geoid, and N, the geoid height (undulation) relative to the ellipsoid (Figure 2):

$$h = H + N. \qquad (1)$$

The geoid undulations, gravity anomalies, and gravity gradient changes all reflect, but are different measures of, the density variations of the Earth. The difference between the geophysical use of gravity and the geodetic use of gravity mirrors the difference between the ellipsoid and the geoid.

ELLIPSOID

As a first approximation, the Earth is a rotating sphere. As a second approximation, it can be regarded as an equipotential ellipsoid of revolution.

According to Moritz (1980), the theory of the equipotential ellipsoid was first given by P. Pizzetti in 1894. It was further elaborated by C. Somigliana in 1929. This theory served as the basis for the International Gravity Formula adopted at the General Assembly of the International Union of Geodesy and Geophysics (IUGG) in Stockholm in 1930. One particular ellipsoid of revolution, also called the "normal Earth," is the one having the same angular velocity and the same mass as the actual Earth, the potential U_0 on the ellipsoid surface equal to the potential W_0 on the geoid, and the center coincident with the center of mass of the Earth. The Geodetic Reference System 1967 (GRS 67), Geodetic Reference System 1980 (GRS 80), and World Geodetic System 1984 (WGS 84) all are "normal Earth."

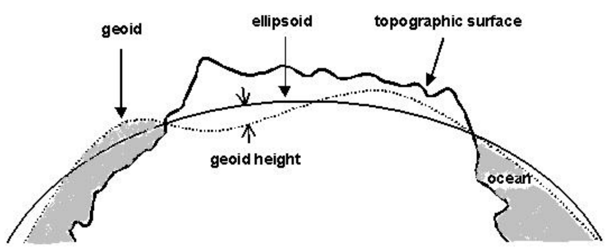

FIG. 1. Cartoon showing the ellipsoid, geoid, and topographic surface (the landmass topography as well as the ocean bathymetry).

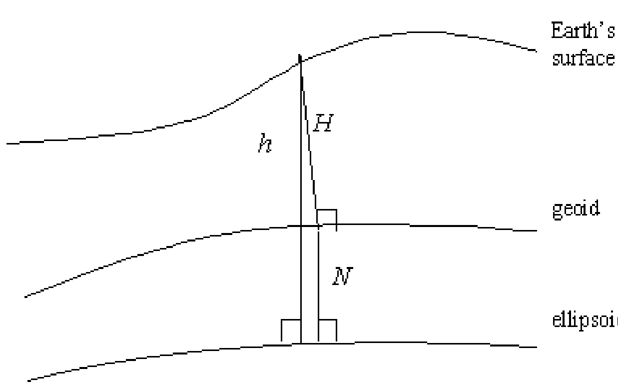

FIG. 2. The elevation H above the geoid, the ellipsoid height h, and the geoid height (undulation) N above the ellipsoid.

Although the Earth is not an exact ellipsoid, the equipotential ellipsoid furnishes a simple, consistent and uniform reference system for all purposes of geodesy as well as geophysics: a reference surface for geometric use such as map projections and satellite navigation, and a normal gravity field on the Earth's surface and in space, defined in terms of closed formulas, as a reference for gravimetry and satellite geodesy. The gravity field of an ellipsoid is of fundamental practical importance because it is easy to handle mathematically, and the deviations of the actual gravity field from the ellipsoidal "theoretical" or "normal" field are small. This splitting of the Earth's gravity field into a "normal" and a remaining small "disturbing" or "anomalous" field considerably simplifies many problems: the determination of the geoid (for geodesists), and the use of gravity anomalies to understand the Earth's interior (for geophysicists).

Although an ellipsoid has many geometric and physical parameters, it can be fully defined by any four independent parameters. All the other parameters can be derived from the four defining parameters. Table 1 lists the two geometric parameters of several representative ellipsoids. Notice how the parameters differ, depending on the choice of ellipsoid.

One of the principal purposes of a world geodetic system is to supersede the local horizontal geodetic datums developed to satisfy mapping and navigation requirements for specific regions of the Earth. A particular reference ellipsoid was used to help define a local datum. For example, the Australian National ellipsoid (Table 1) was used to define the Australian Geodetic Datum 1966. At present, because of a widespread use of GPS, many local datums have been updated using the GRS 80 or WGS 84 ellipsoid.

GRS 80 and WGS 84

Modern satellite technology has greatly improved determination of the Earth's ellipsoid. As shown in Table 1, the semimajor axis of the International 1924 ellipsoid is 251 m larger than for the GRS 80 or WGS 84 ellipsoid, which represents the current best global geodetic reference system for the Earth.

WGS 84 was designed for use as the reference system for the GPS. The WGS 84 Coordinate System is a conventional terrestrial reference system. When selecting the WGS 84 ellipsoid and associated parameters, the original WGS 84 Development Committee decided to adhere closely to the IUGG's approach in establishing and adopting GRS 80.

GRS 80 has four defining parameters: the semimajor axis ($a = 6\,378\,137$ m), the geocentric gravitational constant of the

Table 1. Examples of different reference ellipsoids and their geometric parameters.

Ellipsoid name	Semimajor axis (a in meters)	Reciprocal of flattening ($1/f$)
Airy 1830	6 377 563.396	299.324 964 6
Helmert 1906	6 378 200	298.3
International 1924	6 378 388	297
Australian National	6 378 160	298.25
GRS 1967	6 378 160	298.247 167 427
GRS 1980	6 378 137	298.257 222 101
WGS 1984	6 378 137	298.257 223 563

Earth including the atmosphere ($GM = 3\,986\,005 \times 10^8$ m^3/s^2), the dynamic form factor ($J_2 = 108\,263 \times 10^8$) of the Earth excluding the permanent tidal deformation, and the angular velocity ($\omega = 7\,292\,115 \times 10^{-11}$ rad/s) of the Earth (Moritz, 1980).

Besides the same values of a and ω as GRS 80, the current WGS 84 (National Imagery and Mapping Agency, 2000) uses both an improved determination of the geocentric gravitational constant ($GM = 3\,986\,004.418 \times 10^8$ m^3/s^2) and, as one of the four defining parameters, the reciprocal ($1/f = 298.257\,223\,563$) of flattening instead of J_2. This flattening is derived from the normalized second-degree zonal gravitational coefficient ($C_{2,0}$) through an accepted, rigorous expression, and turned out slightly different from the GRS 80 flattening because the $C_{2,0}$ value is truncated in the normalization process. The small differences between the GRS 80 ellipsoid and the current WGS 84 ellipsoid have virtually no practical consequence.

APPROXIMATE CALCULATION OF THEORETICAL GRAVITY DUE TO AN ELLIPSOID

The theoretical or normal gravity, or gravity reference field, is the gravity effect due to an equipotential ellipsoid of revolution. Approximate formulas are used widely even though we can calculate the exact theoretical gravity analytically. Appendix A gives closed-form expressions as well as approximate ones. In particular, equation (A-2) (see Appendix A) estimates in a closed form the theoretical gravity at any position on, above, or below the ellipsoid.

The International Gravity Formula

The conventionally used International Gravity Formula is obtained by substituting the parameters of the relevant reference ellipsoid into equation (A-3). Helmert's 1901 Gravity Formula, and International Gravity Formulas 1930, 1967, and 1980, correspond respectively to the Helmert 1906, International 1924, GRS 67, and GRS 80 ellipsoids. For example, the 1980 International Gravity Formula is (Moritz, 1980)

$$\gamma_{1980} = 978\,032.7(1 + 0.005\,302\,4 \sin^2\phi - 0.000\,005\,8 \sin^2 2\phi) \text{ mGal}, \qquad (2)$$

where ϕ is the geodetic latitude.

The resulting difference between the 1980 International Gravity Formula and the 1930 International Gravity Formula is

$$\gamma_{1980} - \gamma_{1930} = -16.3 + 13.7 \sin^2\phi \text{ mGal},$$

where the main difference is due to a change from the Potsdam gravity reference datum used in the 1930 formula to the International Gravity Standardization Net 1971 (IGSN71) reference.

The first term of the International Gravity Formula is the value of gravity at the equator on the ellipsoid surface. Unfortunately, in the 1930s, no one really knew what it was. The most reliable estimate at that time was based on absolute gravity measurements made by pendulums at the Geodetic Institute Potsdam in 1906. The Potsdam gravity value served as an absolute datum for worldwide gravity networks from 1909 until 1971. In the 1960s, new measurements across continents made by precise absolute and relative gravity meters became the network of IGSN71 still in use today. A mean difference between the Potsdam datum and the IGSN71 reference has been found to be 14 mGal (Woollard, 1979).

Similarly, we can compare the 1967 formula to the 1980 formula in use today. The difference between the two is relatively small:

$$\gamma_{1980} - \gamma_{1967} = 0.8316 + 0.0782 \sin^2\phi \text{ mGal}.$$

The height correction

The International Gravity Formula estimates the change with latitude on the ellipsoid surface of theoretical gravity due to an ellipsoid. The height correction accounts for the change of theoretical gravity due to the station's being located above or below the ellipsoid at ellipsoid height h. Historically, this height correction has been called the "free-air" correction and thought to be associated with the elevation H, not the ellipsoid height h. In geodesy, the "free-air" correction was interpreted fictitiously as a reduction *to* the geoid of gravity observed on the topographic surface. This has given rise to confusion in geophysics (e.g., Nettleton, 1976, 88).

As a second approximation, the height correction is given in equation (A-4). For the GRS 80 ellipsoid, we have

$$\delta g_{h2} = \gamma_h - \gamma = -(0.308\,769\,1 - 0.000\,439\,8 \sin^2\phi)h + 7.2125 \times 10^{-8} h^2 \text{ mGal}. \qquad (3)$$

However, in exploration geophysics, a first-order formula is widely used, rather than this second approximation.

The famous 0.3086 correction factor

For the International 1924 ellipsoid, the second approximation of the height correction is (Heiskanen and Moritz, 1967, 80)

$$\delta g_{h2} = -(0.308\,77 - 0.000\,45 \sin^2\phi)h + 0.000\,072 h^2.$$

Ignoring the second-order term and setting $\phi = 45^\circ$, we obtain the first approximation of the height correction

$$\delta g_{h1} = -0.3086h \text{ mGal}. \qquad (4)$$

This is just the famous, routinely used, approximate height correction. Again, in exploration geophysics, it is commonly called the (first-order) "free-air" correction and is used with the elevation H rather than the ellipsoid height h.

Errors of approximate formulas

For the GRS 80 ellipsoid, as a first approximation equations (2) and (4) are combined to predict the theoretical gravity at a position above (or below) the ellipsoid. The result is

$$\gamma^1_{1980} = \gamma_{1980} + \delta g_{h1}. \qquad (5)$$

A second approximation is a combination of equations (2) and (3):

$$\gamma^2_{1980} = \gamma_{1980} + \delta g_{h2}. \qquad (6)$$

These two approximate formulas can be compared to the value given by the closed-form formula (A-2). The two differences are denoted as

$$\Delta g_1 = \gamma^1_{1980} - \gamma \tag{7}$$

and

$$\Delta g_2 = \gamma^2_{1980} - \gamma. \tag{8}$$

For an ellipsoid height of 3000 m, differences versus latitudes are given in Table 2. Table 3 shows differences versus ellipsoid heights at 45° latitude.

Because the differences Δg_2 shown in Tables 2 and 3 are smaller than typical exploration survey errors, equation (A-4), together with the International Gravity Formula, produces a sufficiently accurate approximation of the exact theoretical gravity value worldwide. This equation includes the second-order ellipsoid height terms. For the GRS 80 ellipsoid, equation (A-4) becomes equation (3).

GEOID

The geoid is a surface of constant potential energy that coincides with mean sea level over the oceans. This definition is not very rigorous. First, mean sea level is not quite a surface of constant potential due to dynamic processes within the ocean. Second, the actual equipotential surface under continents is warped by the gravitational attraction of the overlying mass. But geodesists define the geoid as though that mass were always underneath the geoid instead of above it. The main function of the geoid in geodesy is to serve as a reference surface for leveling. The elevation measured by leveling is relative to the geoid.

GEODESY: CONVERSION OF GRAVITY TO GEOID

Originally, geodesy was a science solely concerned with global surveying, with the objective of tying local survey nets together by doing careful surveying over long distances. Geodesists tell local surveyors where their positions are with respect to the rest of the world. That includes determining the elevation above sea level.

Why should gravity enter into geodesy?

Many geodetic instruments use gravity as reference. Clearly, mean sea level serves as a reference surface for leveling, and the elevation is relative to mean sea level. In theory, mean sea level could be determined by regular observations at permanent tidal gauge stations. However, one can not very accurately determine the elevation at a location far away from and not tightly tied to an elevation datum defining mean sea level. In practice, the geoid replaces mean sea level as a reference surface for leveling. When we level, what we really measure are the elevations above (or below) the geoid. When geodesists or surveyors say a surface is horizontal, they really mean that it is a surface of constant gravitational potential. So, geodesists have always had to measure gravity—in addition to relative positions—which is why gravity historically was regarded as part of geodesy.

The very early gravity work with pendulum equipment was for geodetic purposes alone. Pierre Bouguer was probably the first to make this kind of observation when he led the expeditions of the French Academy of Sciences to Peru in 1735–1743. Geophysical use of gravity observations started much later. The first use for geological investigation may have been when Hugo de Boeckh, who was at that time the Director of the Geological Survey of Hungary, asked Baron Roland von Eötvös to do a torsion balance survey over the then one-well oil field of Egbell (Gbely) in Slovakia. This survey was carried out in 1915–1916 and showed a clear maximum over the known anticline (Eckhardt, 1940).

Geodesists determine the Earth's figure (i.e., the geoid) in two steps. First, they reduce to the geoid the gravity, observed on the actual Earth's surface. Second, from the reduced gravity, they calculate the geoid undulations (i.e., the deviations from the ellipsoid surface).

The free-air reduction: An historical concept and requirement of classical geodesy

Gravity is measured on the actual surface of the Earth. In order to determine the geoid, the masses outside the geoid must be completely removed or moved inside the geoid by the various gravity corrections, and gravity must be reduced onto the geoid. Geodesists need the elevation H relative to the geoid when they derive the geoid from gravity.

For a reduction of gravity to the geoid, they need the vertical gradient of gravity, $\partial g/\partial H$. Note that $H \ll a$, the semimajor axis of the ellipsoid. If g_s is the observed value on the surface of the Earth, then the value g_g on the geoid may be obtained as a Taylor series expansion. Neglecting all but the linear term, geodesists obtain

$$g_g = g_s + F,$$

Table 2. Differences Δg_1 in equation (7) and Δg_2 in equation (8) of theoretical gravity in equation (A-2) and the two approximations in equations (5) and (6) at an ellipsoid height of 3000 m and different geodetic latitudes.

latitude	0°	15°	30°	45°	60°	75°	90°
Δg_1 (mGal)	−0.114	−0.192	−0.411	−0.728	−1.079	−1.363	−1.474
Δg_2 (mGal)	0.028	0.038	0.061	0.073	0.052	0.009	−0.013

Table 3. Differences Δg_1 in equation (7) and Δg_2 in equation (8) of theoretical gravity in equation (A-2) and the two approximations in equations (5) and (6) at geodetic latitude of 45° and different ellipsoid heights.

height (m)	10	100	500	1000	2000	3000	4000	5000	6000
Δg_1 (mGal)	0.044	0.040	0.006	−0.068	−0.326	−0.728	−1.276	−1.968	−2.805
Δg_2 (mGal)	0.045	0.046	0.050	0.055	0.064	0.073	0.081	0.089	0.096

where

$$F = -\frac{\partial g}{\partial H} H \approx -0.3086H \text{ mGal}. \quad (9)$$

Equation (9) continues to be called the "free-air" effect. Geodesists have assumed that there are no masses above the geoid, or that such masses have been removed beforehand, so that this reduction is as though it were done in "free air". It is so called because, after removal of the topography by the complete Bouguer reduction, the gravity station is left hanging in "free air" (Heiskanen and Moritz, 1967, 131).

In classical geodesy, geodesists employed the fiction that the "free-air" reduction condensed the topographic masses and lowered the station onto the geoid, whereby the geoid became a bounding surface of the terrestrial masses, and gravity g_g was on the geoid (Heiskanen and Moritz, 1967, 145). Therefore, Stokes' formula can be used to calculate geoid undulations from gravity. Unfortunately, geophysicists have often misunderstood and misused this geodetic philosophy.

Calculating geoid undulations from gravity

After estimating g_g, gravity on the geoid, geodesists can then derive the geoid. For a simplification, we take as example the spherically symmetric, rotating Earth. The derivation from gravity to the geoid consists of three substeps (Wahr, 1997, 104–108). First, calculate δg, the angular-dependent component of g_g, by the following relation:

$$g_g \approx \frac{GM}{a^2} - \frac{2}{3}\omega^2 a + \delta g.$$

Then solve

$$\frac{\partial \delta V}{\partial r} + \frac{2}{a}\delta V = -\delta g$$

to find δV, the angular-dependent component of the gravitational potential. Finally, the geoid shape is given by the Bruns formula

$$\delta r = \frac{\delta V}{\gamma},$$

where γ is the theoretical gravity on the surface of the spherical earth and δr is the departure of the geoid from a sphere.

In general and in practice, the geoid undulations are denoted by N. They are the departure from an ellipsoid and can be calculated using Stokes' formula. Details can be found in books on physical geodesy (e.g., Heiskanen and Moritz, 1967; Wahr, 1997).

Geoid model

Equation (1) connects h (the ellipsoid height relative to the ellipsoid), N (the geoid undulation relative to the ellipsoid), and H (the elevation relative to the geoid) (Figure 2).

The geoid undulations range worldwide from −107 m to 85 m relative to the WGS 84 ellipsoid. The primary goal of geodesy is to develop a geoid model, which is then used to connect the three values. Given N, we can compute H or h from the other. For example, when we use GPS as a positioning tool, we measure the ellipsoid height h. The elevation H can be estimated by equation (1) if we have a geoid model.

In general, the global or large-scale features of the geoid are expressed by a spherical harmonic expansion of the gravitational potential. Its higher terms are well defined by the ground gravity data, and the lower terms by the satellite tracking data. The Earth Gravitational Model 1996 (EGM96) is one of the latest global models. It is complete through degree and order 360. The EGM96 global geoid undulations are shown in Figure 3, and have an error range of ±0.5 to ±1.0 m worldwide (Lemoine et al., 1998). The U.S. National Imagery and Mapping Agency recommends that it be used together with the WGS 84 reference ellipsoid (National Imagery and Mapping Agency, 2000).

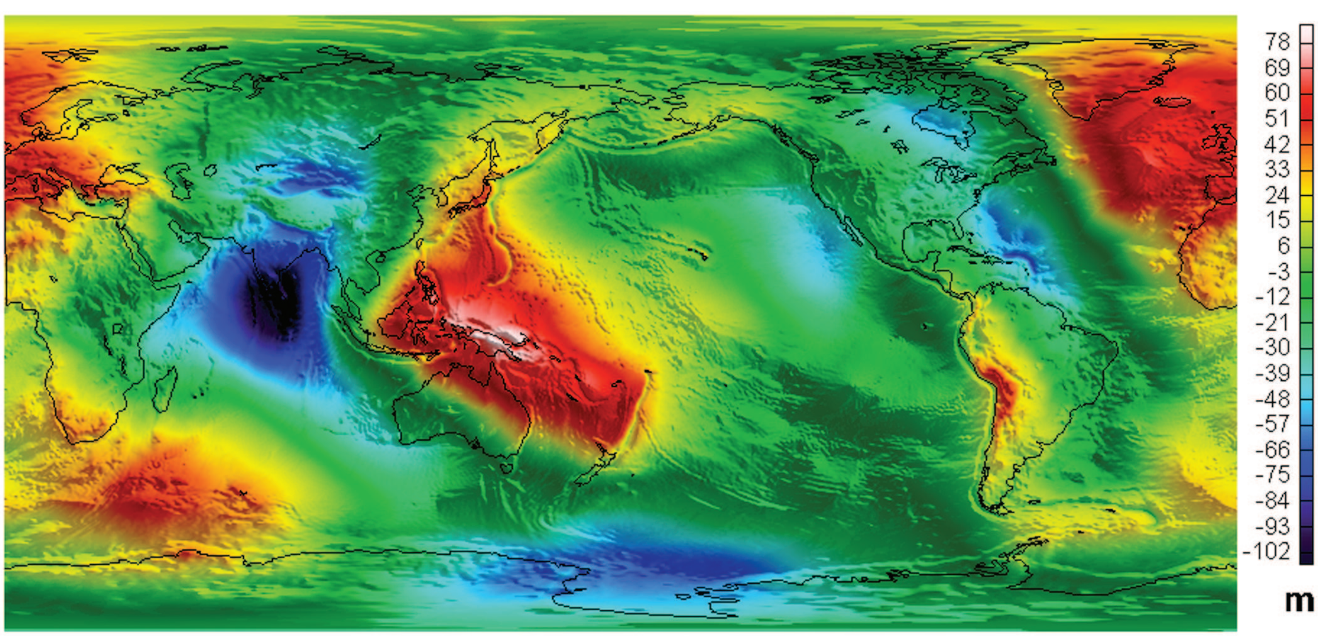

FIG. 3. The 15′ × 15′ global geoid undulations produced by EGM96 (Lemoine et al., 1998). The undulations range from −107 m to 85 m. Black lines indicate coast lines.

Short-wavelength geoid undulations

The relation between spherical harmonic degree n and wavelength λ of geoid undulations is:

$$\lambda = \frac{2\pi R}{n} \approx \frac{40\,000\,000}{n}, \quad (10)$$

where $R = 6\,371\,000$ m is the average radius of the Earth. EGM96 extends to degree and order 360 and thus has the shortest spatial wavelength of 111 km.

At present, there exists no published truly global geoid model that extends beyond degree 360 (i.e., contains a wavelength of shorter than 111 km). Several empirical relations have been established to estimate how the expected power of global gravity and geoid signals drops off with an increase in degree of the spherical harmonic model (Kaula, 1966; Tscherning and Rapp, 1974; Jekeli, 1978). All these relations estimate that the global rms geoid undulation signals are less than 2 cm and 20 cm, when the wavelengths of undulations are 10 km and 100 km, respectively.

In a local area or nationwide, a high-resolution and accurate geoid model may be derived. The GEOID99 model is the latest one for the United States. The geoid grid with a cell size of 1 arcminute (about 2 km) is known as a hybrid geoid model, combining many millions of gravity and elevation points with thousands of control points (i.e., GPS ellipsoid heights on leveled bench marks). For the conterminous United States, when comparing the GEOID99 model back to the same control points, the rms difference is 4.6 cm. Its resolution may be between 10 and 20 km (Smith and Roman, 2001). For most of geophysical exploration purposes, simple height conversions with GEOID99 in the conterminous United States can be sufficient.

CORRECTLY INTERPRETING THE FREE-AIR REDUCTION

Heiskanen and Moritz (1967, chapter 8) defined physical geodesy to be "classical" or "conventional" before M. S. Molodensky proposed his famous theory in the 1940s, and "modern" thereafter. Distinguished from Stokes' formula, Molodensky's theory says that the physical surface of the Earth can be determined without using the density required, for example, by the Bouguer correction. Heiskanen and Moritz (1967, section 8.3 "Molodensky's Problem", 293) clearly wrote:

> The normal gravity on the telluroid [*a variant of the geoid–authors*] is computed from the normal gravity at the ellipsoid by the normal free-air reduction, but now applied upward . . . Therefore the new free-air anomalies have nothing in common with a free-air reduction of actual gravity to sea level, except the name. This distinction should be carefully kept in mind.

And on page 241,

> If, as is usually done, the normal free-air gradient $\partial\gamma/\partial h \approx 0.3086$ mGal/m is used for the free-air reduction, then the free-air anomalies refer, strictly speaking, to the Earth's physical surface (to ground level) rather than to the geoid (to sea level) . . . However, this distinction is insignificant and can be ignored in most cases, so that we may consider Δg as sea-level anomalies.

In geodesy, this distinction is insignificant and can be ignored in most cases because the reduction (downward continuation) to sea level affects relatively short-wavelength anomalies, which are less significant in determination of the geoid. The geoid reflects very-long-wavelength density variations. This is particularly true in a determination of regional or global geoid undulations. But in exploration geophysics, we are interested just in short-wavelength anomalies. The distinction is important for us. Furthermore, it has led to an astonishing level of confusion among geophysicists.

In exploration geophysics, Naudy and Neumann (1965) explicitly noted that the free-air and Bouguer gravity anomalies refer to the observation station. Many algorithms [e.g., the equivalent source technique of Dampney (1969)] have been developed to continue gravity from an undulating observation surface to a horizontal plane. Regardless, even in the 1980s, some publications still referred the free-air anomaly to sea level and incorrectly suggested that the measured vertical gravity gradient should be better used to reduce observed gravity to sea level. For example, Gumert (1985) wrote, "The free-air factor varies significantly with horizontal position and can affect the reduction of observed gravity data. Land gravity measurements made at varying elevation in an area of rugged topography, processed using the standard accepted free-air factor, can produce highly erroneous maps." Again, "Airborne gravity gives the ability to fly multi-level lines in a survey area to compute the free-air factor to apply to the data."

The height (or improperly, "free-air") correction should be made using a consistent, worldwide theoretical standard, that is, one defined by an ellipsoid. The use of local or measured value is inconsistent with the objective of looking for anomalies relative to a universal model of the earth's gravity and is unable to continue observed gravity to any common level.

SATELLITE ALTIMETER GRAVITY: AN EXAMPLE OF CONVERTING GEOID INTO GRAVITY

The primary task of geodesy is to determine the geoid from the observed gravity. However, we can go in the other direction, as well: we can convert the observed geoid into a gravity anomaly. Satellite altimeter gravity (also called satellite-derived gravity) is such a process.

In satellite altimetry, two very precise distance measurements are made so that the topography of the ocean surface (i.e., the geoid) is derived. First, the ellipsoid height h is measured by tracking the satellite from a globally distributed network of lasers and/or Doppler stations. Second, the height of the satellite above the closest ocean surface (i.e., the elevation H) is measured with a microwave radar altimeter. As demonstrated in equation (1), the difference between these two heights is just the geoid undulation N. In practice, altimeter data, collected by different satellites over many years, are combined to achieve a high data density and to average out sea surface disturbing factors such as waves, winds, tides, and currents.

The geoid relatively reflects deeply buried density variations. In order to enhance small-scale features, the high-precision geoid is converted into gravity anomaly. The gravity anomaly can be computed by using inverse Stokes' formula (the geoid-to-gravity method) or by taking the derivatives of the geoid and using Laplace's equation (the slope-to-gravity method; e.g., see Sandwell and Smith, 1997). In the real world, the conversion algorithms are sophisticated, based on laws of physics, geometry, and statistics.

Anyway, there is a simple relationship between gravity anomaly and geoid undulation. For two-dimensional anomalies, an anomaly in geoid with a wavelength λ and amplitude N, the associated gravity anomaly Δg is given by

$$\Delta g = \frac{2\pi \gamma N}{\lambda}, \tag{11}$$

where $\gamma = 980\,000$ mGal, the average gravity of the Earth. This formula can be derived in the Fourier domain by following the work of a determination of gravity anomalies from a grid of geoid undulations (Haxby et al., 1983). Equation (11) says that the bump in the geoid associated with a 10-mGal gravity anomaly and a wavelength of 10 km is just 16 mm. This indicates how precise the geoid must be in order to derive gravity anomalies useful for exploration geophysics. A number of independent studies (Green et al., 1998; Yale et al., 1998) show that satellite altimeter gravity has an accuracy of about 5 mGal and resolution of about 20 km.

GEOPHYSICS: STATION GRAVITY ANOMALY RELATIVE TO THE ELLIPSOID

Equations (4) and (9) appear to be the same. Actually, they have two important differences. First, h in equation (4) is the ellipsoid height, but H in equation (9) the elevation. Second, equation (4) accounts for the change of theoretical gravity due to the ellipsoid with the ellipsoid height, whereas equation (9) represents an historical endeavor of reducing gravity from the Earth's surface to the geoid. *These two differences* [*i.e., equations* (4) *and* (9)] *distinguish geophysics from* (*classical*) *geodesy.* In geophysics, we should follow equation (4) and its implications.

Gravity anomaly is a station anomaly

The geophysical use of gravity is to learn about the Earth's interior. We need to remove the effects of the Earth's irregular (nonellipsoidal) surface. In principle, this means that we should compare the observed gravity to that of ellipsoidally-produced theoretical gravity values at each observation station. Their difference is just the gravity anomaly. The free-air anomaly is the difference between the observed gravity, without terrain-related corrections, and the theoretical gravity. The complete Bouguer anomaly is the difference between the observed gravity with the complete Bouguer correction (the Bouguer slab, curvature, and terrain corrections) and the theoretical gravity. Both the free-air and Bouguer gravity anomalies are located at the gravity station. We must conduct a continuation process in order to obtain the gravity responses on the geoid or another surface/level. As an example, in a continuation to sea level of the ground Bouguer gravity anomaly in the Central Andes, the correction value can reach 30% of the maximum magnitude of the station anomalies (Li and Götze, 1996).

The ellipsoid height, or the elevation plus the geoid

In geophysics, the gravity anomaly is the difference between the observed gravity and the theoretical gravity produced by the ellipsoid. The geophysical gravity anomaly can be calculated simply by using the ellipsoid height h instead of the elevation H in positioning and in all necessary corrections/reductions. In particular, it is not appropriate to estimate the elevation from the ellipsoid height determined by GPS and then use the elevation for corrections/reductions. The extra step produces less reasonable and less significant results.

Traditionally, geophysicists use the elevation as the vertical position of the gravity station and the topographic model. The elevation is used in all the corrections including the height correction and the complete Bouguer correction (the Bouguer slab, curvature, and terrain corrections). Rigorously speaking, in addition we should correct observed gravity for the geoid shape. The gravity effects due to the geoid undulations are called the indirect effects (Chapman and Bordine, 1979). Li and Götze (1996) explained the details of estimating the indirect effects. For example, the indirect effect δg_{ih} caused by the routine height correction is

$$\delta g_{ih} = -0.3086 N \text{ mGal}. \tag{12}$$

Thus, the indirect effect on the free-air gravity anomaly can be up to 30 mGal worldwide.

However, the amplitude of geoid undulations with a wavelength below 10 km is usually smaller than 10 cm, and the amplitude for a wavelength of 100 km is widely smaller than 1 m. Approximately, an elevation change of 10 cm results in a change in computed Bouguer anomaly value of 0.02 mGal, and 1 m results in a change of 0.2 mGal. In practice, at short wavelengths (say, less than 100 km), we don't need to correct for the geoid undulations because the geoid is very smooth, with little power at those wavelengths. In petroleum exploration and in particular in minerals exploration, by ignoring the geoid corrections (i.e., the indirect effects) one is unlikely to introduce any important relative errors across the region of investigation.

Use of geoid, gravity, and gravity gradient

The geoid undulations, gravity anomalies, and gravity gradient changes all are due to the density variations of the Earth's interior, and are transformable from one to another. The wavelengths that the gravity gradient, gravity, and geoid dominate or concentrate range gradually from short (tens of meters) to long (thousands of kilometers). The geoid undulations are used to study the global or very regional problems such as the mantle convection. On the contrary, the gravity gradient is better used to investigate short wavelength effects for engineering, environmental, or mining problems.

SUMMARY

Geodesy uses gravity to determine the geoid. Geodesists must reduce the observed gravity from the actual surface of the Earth to the geoid (mean sea level). In the gravity corrections/reductions, geodesists use the elevation instead of the ellipsoid height.

Geophysics uses gravity to study the Earth's interior. The gravity anomaly is the difference between the observed gravity and the theoretical gravity predicted from the ellipsoid. The gravity anomaly is located at the observation station after the height correction and other routine corrections/reductions are applied. In principle, the ellipsoid height should be used in positioning and in all data corrections/reductions. In practice and in minerals and petroleum exploration, use of the elevation rather than the ellipsoid height hardly introduces significant errors because the geoid is very smooth. However, it should not be recommended as a routine procedure to derive the elevation from the ellipsoid height determined by GPS and then use the elevation for corrections/reductions.

The theoretical gravity on, above, and below the ellipsoid surface can be calculated by a closed-form formula. Its approximation by the International Gravity Formula and the height correction including the second-order terms is typically accurate enough worldwide.

ACKNOWLEDGMENTS

A part of this work was finished when X.L. worked as a research fellow at Freie Universität Berlin, financially supported by the Alexander von Humboldt Stiftung. Writing of this tutorial has been largely stimulated by the discussions at the gravity and magnetic user group *grvmag-l* during May and June 2000. Comments from and discussions with Richard O. Hansen and Maurice D. Craig were helpful in improving the tutorial. We also thank associate editor David A. Chapin and reviewers Alan T. Herring and Dhananjay N. Ravat for their valuable suggestions and comments.

REFERENCES

Chapman, M. E., and Bordine, J. H., 1979, Considerations of the indirect effect in marine gravity modeling: J. Geophys. Res., **84**, 3889–3892.

Dampney, C. N. G., 1969, The equivalent source technique: Geophysics, **34**, 39–53.

Eckhardt, E. A., 1940, A brief history of the gravity method of prospecting for oil: Geophysics, **5**, 231–242.

Green, C. M., Fairhead, J. D., and Maus, S., 1998, Satellite-derived gravity: Where we are and what's next: The Leading Edge, **17**, 77–79.

Gumert, W. R., 1985, Advantages of continuous profiling airborne gravity surveys: Proceedings of the International Meeting on Potential Fields in Rugged Topography, Institut de Géophysique, Université de Lausanne, 16–18.

Haxby, W. F., Karner, G. D., LaBrecque, J. L., and Weissel, J. K., 1983, Digital images of combined oceanic and continental data sets and their use in tectonic studies: EOS, **64**, 995–1004.

Heiskanen, W. A., and Moritz, H., 1967, Physical geodesy: W. H. Freeman and Co.

Jekeli, C., 1978, An investigation of two models for the degree variances of global covariance functions: Ohio State University, Department of Geodetic Science and Surveying, Report 275.

Kaula, W. M., 1966, Theory of satellite geodesy: Blaisdell Publishing Co.

Lakshmanan, J., 1991, The generalized gravity anomaly: Endoscopic microgravity: Geophysics, **56**, 712–723.

Lemoine, F. G., et al., 1998, The development of the joint NASA GSFC and National Imagery and Mapping Agency (NIMA) geopotential model EGM96: Technical Paper NASA/TP-1998-206861.

Li, X., and Götze, H.-J, 1996, Effects of topography and geoid on gravity anomalies in mountainous areas: The Central Andes as an example: Institut für Geologie, Geophysik und Geoinformatik, Freie Universität Berlin.

Moritz, H., 1980, Geodetic Reference System 1980: Bulletin Géodésique, **54**, 395–405.

National Imagery and Mapping Agency, 2000, Department of Defense World Geodetic System 1984: Its definition and relationship with local geodetic systems: Technical Report NIMA TR8350.2, Third Edition.

Naudy, H., and Neumann, R., 1965, Sur la définition de l'anomalie de Bouguer et sec conséquences pratiques: Geophy. Prosp., **13**, 1–11.

Nettleton, L. L., 1976, Gravity and magnetics in oil prospecting: McGraw-Hill Book Co.

Sandwell, D. T., and Smith, W. H. F., 1997, Marine gravity anomaly from Geosat and ERS-1 satellites: J. Geophys. Res., **102**, 10039–10054.

Smith, D. A., and Roman, D. R., 2001, GEOID99 and G99SSS: One arc-minute models for the United States: J. Geodesy, in press.

Tscherning, C. C., and Rapp, R. H., 1974, Closed covariance expression for gravity anomalies, geoid undulations, and deflections of the vertical implied by anomaly degree variance models: Ohio State University, Department of Geodetic Science and Surveying, Report 208.

Wahr, J., 1997, Geodesy and gravity: Class notes: Samizdat Press.

Woollard, G. P., 1979, The new gravity system—Changes in international gravity base values and anomaly values: Geophysics, **44**, 1352–1366.

Yale, M. M., Sandwell, D. T., and Herring, A. T., 1998, What are the limitations of satellite altimetry?: The Leading Edge, **17**, 73–76.

APPENDIX A

THEORETICAL GRAVITY DUE TO AN ELLIPSOID

The theoretical gravity is the gravity effect due to an equipotential ellipsoid of revolution. Approximate formulas are used widely. In fact, we can calculate the theoretical gravity at any position on, above, or below the ellipsoid surface using closed-form expressions.

Closed-form expression: Gravity on the surface of the ellipsoid

The theoretical gravity on the surface of the ellipsoid is given by the formula of Somigliana (Heiskanen and Moritz, 1967, 76):

$$\gamma = \gamma_e \frac{1 + k \sin^2\phi}{\sqrt{1 - e^2 \sin^2\phi}}, \qquad \text{(A-1)}$$

where

$$k = \frac{b\gamma_p}{a\gamma_e} - 1;$$

$$e = \sqrt{\frac{a^2 - b^2}{a^2}} \text{ is the first eccentricity;}$$

and a and b are the semimajor and semiminor axes of the ellipsoid, respectively; γ_e and γ_p are the theoretical gravity at the equator and poles, respectively; and ϕ is the geodetic latitude.

Closed-form formula: Gravity above and below the surface of the ellipsoid

The theoretical gravity at any ellipsoid height h and any geodetic latitude ϕ (Figure A-1) can also be given by a closed-form formula. Starting from the general formula of Heiskanen and Moritz (1967, 67–71), Lakshmanan (1991) derived the formula and published a result containing typographic errors. Li and Götze (1996) repeated the derivation and corrected the errors, obtaining

$$\gamma = \frac{1}{W}\left\{\frac{GM}{b'^2 + E^2} + \frac{\omega^2 a^2 E q'}{(b'^2 + E^2) q_0}\left(\frac{1}{2}\sin^2\beta' - \frac{1}{6}\right) - \omega^2 b' \cos^2\beta'\right\}, \qquad \text{(A-2)}$$

where

$$E = \sqrt{a^2 - b^2} \text{ is linear eccentricity,}$$

$$W = \sqrt{\frac{b'^2 + E^2 \sin^2 \beta'}{b'^2 + E^2}},$$

$$q' = 3\left(1 + \frac{b'^2}{E^2}\right)\left(1 - \frac{b'}{E}\tan^{-1}\frac{E}{b'}\right) - 1,$$

$$q_0 = \frac{1}{2}\left[\left(1 + \frac{3b^2}{E^2}\right)\tan^{-1}\frac{E}{b} - \frac{3b}{E}\right],$$

$$b' = \sqrt{r''^2 - E^2 \cos^2 \beta'},$$

$$\cos \beta' = \sqrt{\frac{1}{2} + \frac{R}{2} - \sqrt{\frac{1}{4} + \frac{R^2}{4} - \frac{D}{2}}},$$

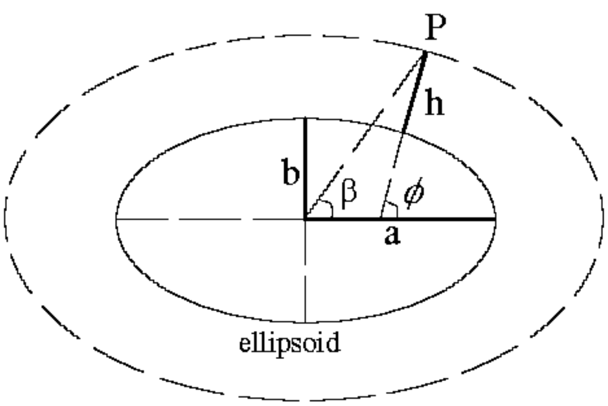

FIG. A-1. A station above an ellipsoid surface. The ellipsoid has the semimajor axis a and semiminor axis b. The position of station P relative to the ellipsoid is defined by ellipsoid height h and geodetic latitude ϕ. The angle β is called the reduced latitude.

See ERRATA for this Figure

and $R = r''^2/E^2$, $D = d''^2/E^2$, $r''^2 = r'^2 + z'^2$, $d''^2 = r'^2 - z'^2$, $r' = a\cos\beta + h\cos\phi$, $z' = b\sin\beta + h\sin\phi$, and $\tan\beta = b/a \tan\phi$.

Approximate formula for the latitude correction

The conventional latitude correction is a second-order series expansion of equation (A-1) (Heiskanen and Moritz, 1967, 77):

$$\gamma = \gamma_e\left(1 + f^* \sin^2\phi - \frac{1}{4} f_4 \sin^2 2\phi\right), \quad \text{(A-3)}$$

with

$$f^* = \frac{\gamma_p - \gamma_e}{\gamma_e} \quad \text{(gravity flattening),}$$

$$f_4 = -\frac{1}{2} f^2 + \frac{5}{2} f m,$$

$$f = \frac{a - b}{a} \quad \text{(flattening of the ellipsoid),}$$

and

$$m = \frac{\omega^2 a^2 b}{GM}.$$

Approximate formula for the height correction

The height correction accounts for the change of theoretical gravity due to the station being located above or below the ellipsoid at ellipsoid height h. As a second approximation (Heiskanen and Moritz, 1967, 79), a Taylor series expansion for the theoretical gravity above the ellipsoid with a positive direction downward along the geodetic normal to the reference ellipsoid is

$$\gamma_h = \gamma\left[1 - \frac{2}{a}(1 + f + m - 2f\sin^2\phi)h + \frac{3}{a^2}h^2\right].$$

The difference $\gamma_h - \gamma$ (i.e., the height correction) is

$$\gamma_h - \gamma = -\frac{2\gamma_e}{a}\left[1 + f + m + \left(\frac{5}{2}m - 3f\right)\sin^2\phi\right]h + \frac{3\gamma_e}{a^2}h^2. \quad \text{(A-4)}$$

GEOPHYSICS, VOL. 67, NO. 3 (MAY-JUNE 2002); DOI 10.1190/1.1489656 031203GPY

To: "Ellipsoid, geoid, gravity, geodesy, and geophysics," X. Li and H. -J. Götze (Geophysics, 66, 1660-1668).

The authors thank Dr. Nico Sneeuw of the University of Calgary for pointing out a graphical error. The reduced latitude β was incorrectly defined in Figure A-1, which should be replaced by the figure shown below. However, this graphical error was not introduced into the derivation of the closed-form expression for the theoretical gravity due to an ellipsoid. All the formulae given in Appendix A are correct.

There is also a typographic error in the dynamic form factor J_2, in the first paragraph on page 1662.

J_2 = 108 263 x 10^{-8} not 108 263 x 10^{8}.

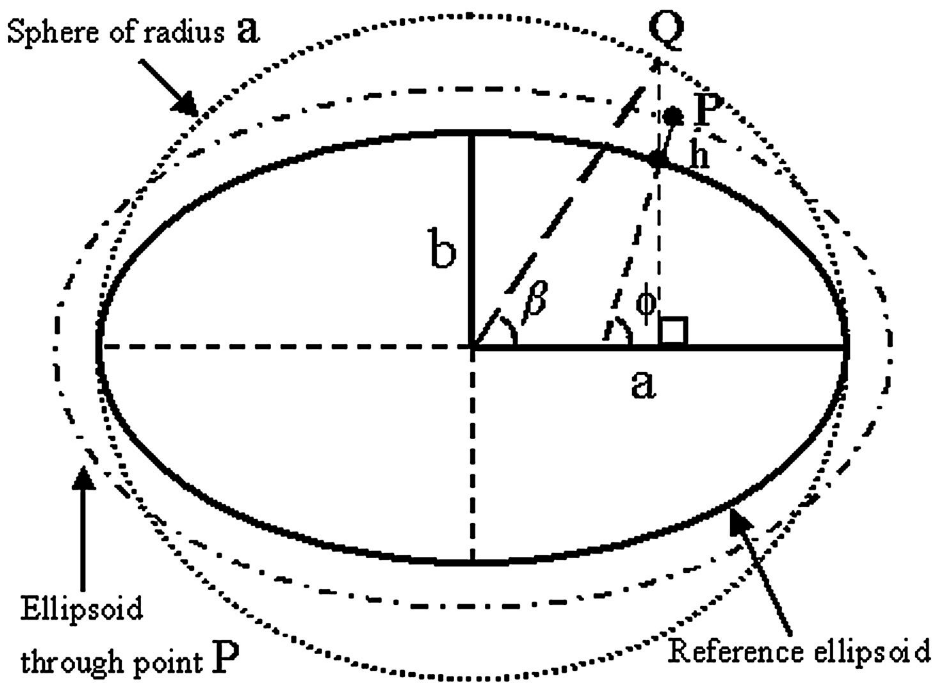

Figure A-1. A station above a reference ellipsoid surface. The reference ellipsoid has a semi-major axis a and a semi-minor axis b. The position of station P relative to the reference ellipsoid is defined by ellipsoid height h and geodetic latitude ϕ. The ellipsoid through point P has the same linear eccentricity as the reference ellipsoid. The reduced latitude β is a geocentric latitude of point Q, which is the vertically projected point, on a sphere of radius a, of station P's normal projection on the reference ellipsoid surface.