3.02 Potential Theory and Static Gravity Field of the Earth

C. Jekeli, The Ohio State University, Columbus, OH, USA

3.02.1	**Introduction**	12
3.02.1.1	Historical Notes	12
3.02.1.2	Coordinate Systems	14
3.02.1.3	Preliminary Definitions and Concepts	14
3.02.2	**Newton's Law of Gravitation**	15
3.02.3	**Boundary-Value Problems**	18
3.02.3.1	Green's Identities	18
3.02.3.2	Uniqueness Theorems	20
3.02.3.3	Solutions by Integral Equation	21
3.02.4	**Solutions to the Spherical BVP**	22
3.02.4.1	Spherical Harmonics and Green's Functions	22
3.02.4.2	Inverse Stokes and Hotine Integrals	25
3.02.4.3	Vening-Meinesz Integral and Its Inverse	26
3.02.4.4	Concluding Remarks	27
3.02.5	**Low-Degree Harmonics: Interpretation and Reference**	28
3.02.5.1	Low-Degree Harmonics as Density Moments	28
3.02.5.2	Normal Ellipsoidal Field	30
3.02.6	**Methods of Determination**	32
3.02.6.1	Measurement Systems and Techniques	33
3.02.6.2	Models	37
3.02.7	**The Geoid and Heights**	38
References		41

Glossary

deflection of the vertical Angle between direction of gravity and direction of normal gravity.

density moment Integral over the volume of a body of the product of its density and integer powers of Cartesian coordinates.

disturbing potential The difference between Earth's gravity potential and the normal potential.

eccentricity The ratio of the difference of squares of semimajor and semiminor axes to the square of the semimajor axis of an ellipsoid.

ellipsoid Surface formed by rotating an ellipse about its minor axis.

equipotential surface Surface of constant potential.

flattening The ratio of the difference between semimajor and semiminor axes to the semimajor axis of an ellipsoid.

geodetic reference system Normal ellipsoid with defined parameters adopted for general geodetic and gravimetric referencing.

geoid Surface of constant gravity potential that closely approximates mean sea level.

geoid undulation Vertical distance between the geoid and the normal ellipsoid, positive if the geoid is above the ellipsoid.

geopotential number Difference between gravity potential on the geoid and gravity potential at a point.

gravitation Attractive acceleration due to mass.

gravitational potential Potential due to gravitational acceleration.

gravity Vector sum of gravitation and centrifugal acceleration due to Earth's rotation.

gravity anomaly The difference between Earth's gravity on the geoid and normal gravity on the

ellipsoid, either as a difference in vectors or a difference in magnitudes.
gravity disturbance The difference between Earth's gravity and normal gravity, either as a difference in vectors or a difference in magnitudes.
gravity potential Potential due to gravity acceleration.
harmonic function Function that satisfies Laplace's field equation.
linear eccentricity The distance from the center of an ellipsoid to either of its foci.
mean Earth ellipsoid Normal ellipsoid with parameters closest to actual corresponding parameters for the Earth.
mean tide geoid Geoid with all time-varying tidal effects removed.
multipoles Stokes coefficients.
Newtonian potential Harmonic function that approaches the potential of a point mass at infinity.
non-tide geoid Mean tide geoid with all (direct and indirect deformation) mean tide effects removed.
normal ellipsoid Earth-approximating reference ellipsoid that generates a gravity field in which it is a surface of constant normal gravity potential.
normal gravity Gravity associated with the normal ellipsoid.
normal gravity potential Gravity potential associated with the normal ellipsoid.
orthometric height Distance along the plumb line from the geoid to a point.
potential Potential energy per unit mass due to the gravitational field; always positive and zero at infinity.
sectorial harmonics Surface spherical harmonics that do not change in sign with respect to latitude.
Stokes coefficients Constants in a series expansion of the gravitational potential in terms of spherical harmonic functions.
surface spherical harmonics Basis functions defined on the unit sphere, comprising products of normalized associated Legendre functions and sinusoids.
tesseral harmonics Neither zonal nor sectorial harmonics.
zero-tide geoid Mean tide geoid with just the mean direct tidal effect removed (indirect effect due to Earth's permanent deformation is retained).
zonal harmonics Spherical harmonics that do not depend on longitude.

3.02.1 Introduction

Gravitational potential theory has its roots in the late Renaissance period when the position of the Earth in the cosmos was established on modern scientific (observation-based) grounds. A study of Earth's gravitational field is a study of Earth's mass, its influence on near objects, and lately its redistributing transport in time. It is also fundamentally a geodetic study of Earth's shape, described largely (70%) by the surface of the oceans. This initial section provides a historical backdrop to potential theory and introduces some concepts in physical geodesy that set the stage for later formulations.

3.02.1.1 Historical Notes

Gravitation is a physical phenomenon so pervasive and incidental that humankind generally has taken it for granted with scarcely a second thought. The Greek philosopher Aristotle (384–322 BC) allowed no more than to assert that gravitation is a natural property of material things that causes them to fall (or rise, in the case of some gases), and the more material the greater the tendency to do so. It was enough of a self-evident explanation that it was not yet to receive the scrutiny of the scientific method, the beginnings of which, ironically, are credited to Aristotle. Almost 2000 years later, Galileo Galilei (1564–1642) finally took up the challenge to understand gravitation through observation and scientific investigation. His experimentally derived law of falling bodies corrected the Aristotelian view and divorced the effect of gravitation from the mass of the falling object – all bodies fall with the same acceleration. This truly monumental contribution to physics was, however, only a local explanation of how bodies behaved under gravitational influence. Johannes Kepler's (1571–1630) observations of planetary orbits pointed to other types of laws, principally an inverse-square law according to which bodies are attracted by forces that vary with the inverse square of distance. The genius of Issac Newton (1642–1727) brought it all together in his *Philosophiae Naturalis Principia Mathematica* of 1687 with a single and simple all-embracing law that in one bold stroke explained the dynamics of the entire universe (today there is more to understanding the dynamics of the cosmos, but Newton's law remarkably holds its own). The mass of a body was again an essential aspect, not as a self-attribute as Aristotle had implied, but as the source of attraction for other bodies: each material body attracts every other material body according to a very specific rule (Newton's law of gravitation; see Section 3.02.2). Newton regretted that he could not explain exactly why mass has this property (as one still yearns to know today within the standard models of particle and quantum theories). Even Albert Einstein (1879–1955) in developing his general theory of relativity (i.e., the theory of gravitation) could only improve on Newton's theory by incorporating and explaining action at a distance (gravitational force acts with the speed of light as a fundamental tenet of the theory). What actually mediates the gravitational attraction still intensely occupies modern physicists and cosmologists.

Gravitation since its early scientific formulation initially belonged to the domain of astronomers, at least as far as the observable universe was concerned. Theory successfully predicted the observed perturbations of planetary orbits and even the location of previously unknown new planets (Neptune's discovery in 1846 based on calculations motivated by observed perturbations in Uranus' orbit was a major triumph for Newton's law). However, it was also discovered that gravitational acceleration varies on Earth's surface, with respect to altitude and latitude. Newton's law of gravitation again provided the backdrop for the variations observed with pendulums. An early achievement for his theory came when he successfully predicted the polar flattening in Earth's shape on the basis of hydrostatic equilibrium, which was confirmed finally (after some controversy) with geodetic measurements of long triangulated arcs in 1736–37 by Pierre de Maupertuis and Alexis Clairaut. Gravitation thus also played a dominant role in geodesy, the science of determining the size and shape of the Earth, promulgated in large part by the father of modern geodesy, Friedrich R. Helmert (1843–1917).

Terrestrial gravitation through the twentieth century was considered a geodetic area of research, although, of course, its geophysical exploits should not be overlooked. But the advancement in modeling accuracy and global application was promoted mainly by geodesists who needed a well-defined reference for heights (a level surface) and whose astronomic observations of latitude and longitude needed to be corrected for the irregular direction of gravitation. Today, the modern view of a height reference is changing to that of a geometric, mathematical surface (an ellipsoid) and three-dimensional coordinates (latitude, longitude, and height) of points on the Earth's surface are readily obtained geometrically by ranging to the satellites of the Global Positioning System (GPS). The requirements of gravitation for GPS orbit determination within an Earth-centered coordinate system are now largely met with existing models. Improvements in gravitational models are motivated in geodesy primarily for rapid determination of traditional heights with respect to a level surface. These heights, for example, the orthometric heights, in this sense then become derived attributes of points, rather than their cardinal components.

Navigation and guidance exemplify a further specific niche where gravitation continues to find important relevance. While GPS also dominates this field, the vehicles requiring completely autonomous, self-contained systems must rely on inertial instruments (accelerometers and gyroscopes). These do not sense gravitation (see Section 3.02.6.1), yet gravitation contributes to the total definition of the vehicle trajectory, and thus the output of inertial navigation systems must be compensated for the effect of gravitation. By far the greatest emphasis in gravitation, however, has shifted to the Earth sciences, where detailed knowledge of the configuration of masses (the solid, liquid, and atmospheric components) and their transport and motion leads to improved understanding of the Earth systems (climate, hydrologic cycle, tectonics) and their interactions with life. Oceanography, in particular, also requires a detailed knowledge of a level surface (the geoid) to model surface currents using satellite altimetry. Clearly, there is an essential temporal component in these studies, and, indeed, the temporal gravitational field holds center stage in many new investigations. Moreover, Earth's dynamic behavior influences point coordinates and Earth-fixed coordinate frames, and we come back to fundamental geodetic concerns in which the gravitational field plays an essential role!

This section deals with the static gravitational field. The theory of the potential from the classical Newtonian standpoint provides the foundation for modeling the field and thus deserves the focus of the exposition. The temporal part is a natural extension that is readily achieved with the addition of the time variable (no new laws are needed, if we neglect general relativistic effects) and will not be expounded here. We are primarily concerned with gravitation on and external to the solid and liquid Earth since this is the domain of most applications. The internal field can also be modeled for specialized purposes (such as submarine navigation), but internal geophysical modeling, for example, is done usually in terms of the sources (mass density distribution), rather than the resulting field.

3.02.1.2 Coordinate Systems

Modeling the Earth's gravitational field depends on the choice of coordinate system. Customarily, owing to the Earth's general shape, a spherical polar coordinate system serves for most applications, and virtually all global models use these coordinates. However, the Earth is slightly flattened at the poles, and an ellipsoidal coordinate system has also been advocated for some near-Earth applications. We note that the geodetic coordinates associated with a geodetic datum (based on an ellipsoid) are never used in a foundational sense to model the field since they do not admit to a separation-of-variables solution of Laplace' differential equation (Section 3.02.4.1).

Spherical polar coordinates, described with the aid of **Figure 1**, comprise the spherical colatitude, θ, the longitude, λ, and the radial distance, r. Their relation to Cartesian coordinates is

$$\begin{aligned} x &= r\sin\theta\cos\lambda \\ y &= r\sin\theta\sin\lambda \\ z &= r\cos\theta \end{aligned} \qquad [1]$$

Considering Earth's polar flattening, a better approximation, than a sphere, of its (ocean) surface is an ellipsoid of revolution. Such a surface is

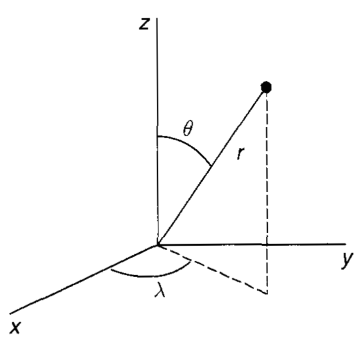

Figure 1 Spherical polar coordinates.

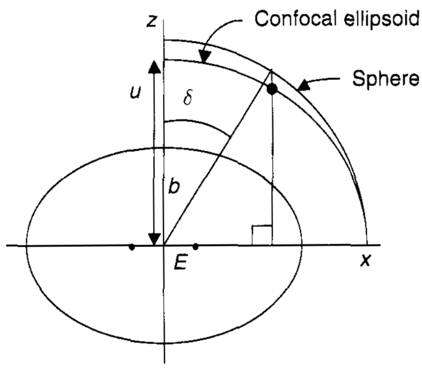

Figure 2 Ellipsoidal coordinates.

generated by rotating an ellipse about its minor axis (polar axis). The two focal points of the best-fitting, Earth-centered ellipsoid (ellipse) are located in the equator about $E = 522$ km from the center of the Earth. A given ellipsoid, with specified semiminor axis, b, and linear eccentricity, E, defines the set of ellipsoidal coordinates, as described in **Figure 2**. The longitude is the same as in the spherical case. The colatitude, δ, is the complement of the so-called reduced latitude; and the distance coordinate, u, is the semiminor axis of the confocal ellipsoid through the point in question. We call (δ, λ, u) ellipsoidal coordinates; they are also known as spheroidal coordinates, or Jacobi ellipsoidal coordinates. Their relation to Cartesian coordinates is given by

$$\begin{aligned} x &= \sqrt{u^2 + E^2}\sin\delta\cos\lambda \\ y &= \sqrt{u^2 + E^2}\sin\delta\sin\lambda \\ z &= u\cos\delta \end{aligned} \qquad [2]$$

Points on the given ellipsoid all have $u = b$; and, all surfaces, $u =$ constant, are confocal ellipsoids (the analogy to the spherical case, when $E = 0$, should be evident).

3.02.1.3 Preliminary Definitions and Concepts

The gravitational potential, V, of the Earth is generated by its total mass density distribution. For applications on the Earth's surface it is useful to include the potential, ϕ, associated with the centrifugal acceleration due to Earth's rotation. The sum, $W = V + \phi$, is then known, in geodetic terminology, as the gravity potential, distinct from gravitational potential. It is further advantageous to define a relatively simple reference potential, or normal potential, that accounts for the bulk of the gravity potential

(Section 3.02.5.2). The normal gravity potential, U, is defined as a gravity potential associated with a best-fitting ellipsoid, the normal ellipsoid, which rotates with the Earth and is also a surface of constant potential in this field. The difference between the actual and the normal gravity potentials is known as the disturbing potential: $T = W - U$; it thus excludes the centrifugal potential. The normal gravity potential accounts for approximately 99.999 5% of the total potential.

The gradient of the potential is an acceleration, gravity or gravitational acceleration, depending on whether or not it includes the centrifugal acceleration. Normal gravity, γ, comprises 99.995% of the total gravity, g, although the difference in magnitudes, the gravity disturbance, δg, can be as large as several parts in 10^4. A special kind of difference, called the gravity anomaly, Δg, is defined as the difference between gravity at a point, P, and normal gravity at a corresponding point, Q, where $W_P = U_Q$, and P and Q are on the same perpendicular to the normal ellipsoid.

The surface of constant gravity potential, W_0, that closely approximates mean sea level is known as the geoid. If the constant normal gravity potential, U_0, on the normal ellipsoid is equal to the constant gravity potential of the geoid, then the gravity anomaly on the geoid is the difference between gravity on the geoid and normal gravity on the ellipsoid at respective points, P_0, Q_0, sharing the same perpendicular to the ellipsoid. The separation between the geoid and the ellipsoid is known as the geoid undulation, N, or also the geoid height (**Figure 3**). A simple Taylor expansion of the normal gravity potential along the ellipsoid perpendicular yields the following important formula:

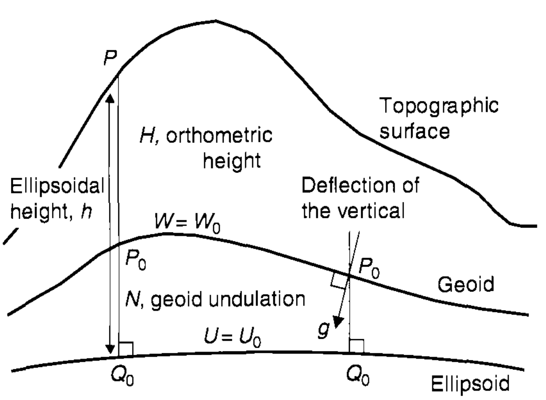

Figure 3 Relative geometry of geoid and ellipsoid.

$$N = \frac{T}{\gamma} \qquad [3]$$

This is Bruns' equation, which is accurate to a few millimeters in N, and which can be extended to $N = T/\gamma - (W_0 - U_0)/\gamma$ for the general case, $W_0 \neq U_0$. The gravity anomaly (on the geoid) is the gravity disturbance corrected for the evaluation of normal gravity on the ellipsoid instead of the geoid. This correction is $N \partial\gamma/\partial h = (\partial\gamma/\partial h)(T/\gamma)$, where h is height along the ellipsoid perpendicular. We have $\delta g = -\partial T/\partial h$, and hence

$$\Delta g = -\frac{\partial T}{\partial h} + \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} T \qquad [4]$$

The slope of the geoid with respect to the ellipsoid is also the angle between the corresponding perpendiculars to these surfaces. This angle is known as the deflection of the vertical, that is, the deflection of the plumb line (perpendicular to the geoid) relative to the perpendicular to the normal ellipsoid. The deflection angle has components, ξ, η, respectively, in the north and east directions. The spherical approximations to the gravity disturbance, anomaly, and deflection of the vertical are given by

$$\begin{aligned} \delta g &= -\frac{\partial T}{\partial r}, \qquad & \Delta g &= -\frac{\partial T}{\partial r} - \frac{2}{r} T \\ \xi &= \frac{1}{\gamma}\frac{\partial T}{r\partial\theta}, \qquad & \eta &= -\frac{1}{\gamma}\frac{\partial T}{r \sin\theta\,\partial\lambda} \end{aligned} \qquad [5]$$

where the signs on the derivatives are a matter of convention.

3.02.2 Newton's Law of Gravitation

In its original form, Newton's law of gravitation applies only to idealized point masses. It describes the force of attraction, **F**, experienced by two such solitary masses as being proportional to the product of the masses, m_1 and m_2; inversely proportional to the distance, ℓ, between them; and directed along the line joining them:

$$\mathbf{F} = G\frac{m_1 m_2}{\ell^2}\mathbf{n} \qquad [6]$$

G is a constant, known as Newton's gravitational constant, that takes care of the units between the left- and right-hand sides of the equation; it can be determined by experiment and the current best value is (Groten, 2004):

$$G = (6.67259 \pm 0.00030) \times 10^{-11}\,\mathrm{m^3\,kg^{-1}\,s^{-2}} \quad [7]$$

The unit vector $\mathbf{n}$ in eqn [6] is directed from either point mass to the other, and thus the gravitational force is attractive and applies equally to one mass as the other. Newton's law of gravitation is universal as far as we know, requiring reformulation only in Einstein's more comprehensive theory of general relativity which describes gravitation as a characteristic curvature of the space–time continuum (Newton's formulation assumes instantaneous action and differs significantly from the general relativistic concept only when very large velocities or masses are involved).

We can ascribe a gravitational acceleration to the gravitational force, which represents the acceleration that one mass undergoes due to the gravitational attraction of the other. Specifically, from the law of gravitation, we have (for point masses) the gravitational acceleration of m_1 due to the gravitational attraction of m_2:

$$\mathbf{g} = G\frac{m_2}{\ell^2}\mathbf{n} \quad [8]$$

The vector $\mathbf{g}$ is independent of the mass, m_1, of the body being accelerated (which Galileo found by experiment).

By the law of superposition, the gravitational force, or the gravitational acceleration, due to many point masses is the vector sum of the forces or accelerations generated by the individual point masses. Manipulating vectors in this way is certainly feasible, but fortunately a more appropriate concept of gravitation as a scalar field simplifies the treatment of arbitrary mass distributions.

This more modern view of gravitation (already adopted by Gauss (1777–1855) and Green (1793–1841)) holds that it is a field having a gravitational potential. Lagrange (1736–1813) fully developed the concept of a field, and the potential, V, of the gravitational field is defined in terms of the gravitational acceleration, $\mathbf{g}$, that a test particle would undergo in the field according to the equation

$$\nabla V = \mathbf{g} \quad [9]$$

where $\boldsymbol{\nabla}$ is the gradient operator (a vector). Further elucidation of gravitation as a field grew from Einstein's attempt to incorporate gravitation into his special theory of relativity where no reference frame has special significance above all others. It was necessary to consider that gravitational force is not a real force (i.e., it is not an applied force, like friction or propulsion) – rather, it is known as a kinematic force, that is, one whose action is proportional to the mass on which it acts (like the centrifugal force; see Martin, 1988). Under this precept, the geometry of space is defined intrinsically by the gravitational fields contained therein.

We continue with the classical Newtonian potential, but interpret gravitation as an acceleration different from the acceleration induced by real, applied forces. This becomes especially important when considering the measurement of gravitation (Section 3.02.6.1). The gravitational potential, V, is a 'scalar' function, and, as defined here, V is derived directly on the basis of Newton's law of gravitation. To make it completely consistent with this law and thus declare it a Newtonian potential, we must impose the following conditions:

$$\lim_{\ell\to\infty} \ell V = Gm \quad \text{and} \quad \lim_{\ell\to\infty} V = 0 \quad [10]$$

Here, m is the attracting mass, and we say that the potential is regular at infinity. It is easy to show that the gravitational potential at any point in space due to a point mass, in order to satisfy eqns [8]–[10], must be

$$V = \frac{Gm}{\ell} \quad [11]$$

where, again, ℓ is the distance between the mass and the point at which the potential is expressed. Note that V for $\ell = 0$ does not exist in this case; that is, the field of a point mass has a singularity. We use here the convention that the potential is always positive (in contrast to physics, where it is usually defined to be negative, conceptually closer to potential energy).

Applying the law of superposition, the gravitational potential of many point masses is the sum of the potentials of the individual points (see **Figure 4**):

$$V_P = G\sum_j \frac{m_j}{\ell_j} \quad [12]$$

And, for infinitely many points in a closed, bounded region with infinitesimally small masses, $\mathrm{d}m$, the summation in eqn [12] changes to an integration,

$$V_P = G\int_{\mathrm{mass}} \frac{\mathrm{d}m}{\ell} \quad [13]$$

or, changing variables (i.e., units), $\mathrm{d}m = \rho\,\mathrm{d}v$, where ρ represents density (mass per volume) and $\mathrm{d}v$ is a volume element, we have (**Figure 5**)

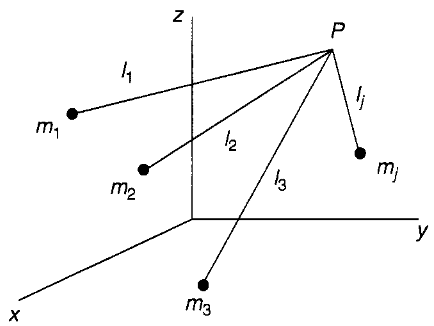

Figure 4 Discrete set of mass points (superposition principle).

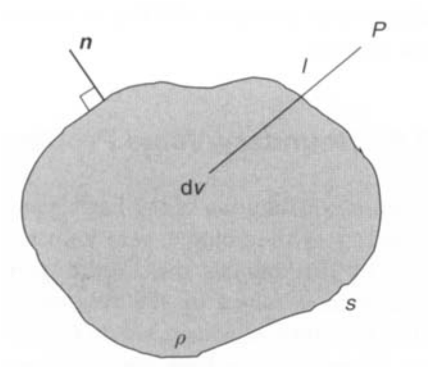

Figure 5 Continuous density distribution.

$$V_P = G\iiint_{\text{volume}} \frac{\rho}{\ell}\mathrm{d}v \qquad [14]$$

In eqn [14], ℓ is the distance between the evaluation point, P, and the point of integration. In spherical polar coordinates (Section 3.02.1.2), these points are (θ, λ, r) and (θ', λ', r'), respectively. The volume element in this case is given by $\mathrm{d}v = r'^2 \sin\theta'\,\mathrm{d}\lambda'\,\mathrm{d}\theta'\,\mathrm{d}r'$. V and its first derivatives are continuous everywhere – even in the case that P is on the bounding surface or inside the mass distribution, where there is the apparent singularity at $\ell = 0$. In this case, by changing to a coordinate system whose origin is at P, the volume element becomes $\mathrm{d}v = \ell^2 \sin\psi\,\mathrm{d}\alpha\,\mathrm{d}\psi\,\mathrm{d}\ell$ (for some different colatitude and longitude ψ and α); and, clearly, the singularity disappears – the integral is said to be weakly singular.

Suppose the density distribution over the volume depends only on radial distance (from the center of mass): $\rho = \rho(r')$, and that P is an exterior evaluation point. The surface bounding the masses necessarily is a sphere (say, of radius, R) and because of the spherically symmetric density we may choose the integration coordinate system so that the polar axis passes through P. Then

$$\ell = \sqrt{r'^2 + r^2 - 2r'r\cos\theta'}, \qquad \mathrm{d}\ell = \frac{1}{\ell}r'r\sin\theta'\mathrm{d}\theta' \qquad [15]$$

It is easy to show that with this change of variables (from θ to ℓ) the integral [14] becomes simply

$$V(\theta, \lambda, r) = \frac{GM}{r}, \quad r \geq R \qquad [16]$$

where M is the total mass bounded by the sphere. This shows that to a very good approximation the external gravitational potential of a planet such as the Earth (with concentrically layered density) is the same as that of a point mass.

Besides volumetric mass (density) distributions, it is of interest to consider surface distributions. Imagine an infinitesimally thin layer of mass on a surface, s, where the units of density in this case are those of mass per area. Then, analogous to eqn [14], the potential is

$$V_P = G\iint_s \frac{\rho}{\ell}\mathrm{d}s \qquad [17]$$

In this case, V is a continuous function everywhere, but its first derivatives are discontinuous at the surface. Or, one can imagine two infinitesimally close density layers (double layer, or layer of mass dipoles), where the units of density are now those of mass per area times length. It turns out that the potential in this case is given by (see Heiskanen and Moritz, 1967, p. 8)

$$V_P = G\iint_s \rho\frac{\partial}{\partial n}\left(\frac{1}{\ell}\right)\mathrm{d}s \qquad [18]$$

where $\partial/\partial n$ is the directional derivative along the perpendicular to the surface (**Figure 5**). Now, V itself is discontinuous at the surface, as are all its derivatives. In all cases, V is a Newtonian potential, being derived from the basic formula [11] for a point mass that follows from Newton's law of gravitation (eqn [6]).

The following properties of the gravitational potential are useful for subsequent expositions. First, consider Stokes's theorem, for a vector function, **f**, defined on a surface, **s**:

$$\iint_s (\nabla \times \mathbf{f})\cdot\mathbf{n}\ \mathrm{d}s = \oint_p \mathbf{f}\cdot\mathrm{d}\mathbf{r} \qquad [19]$$

where p is any closed path in the surface, **n** is the unit vector perpendicular to the surface, and **dr** is a differential displacement along the path. From eqn [9], we find

$$\nabla \times \mathbf{g} = \mathbf{0} \qquad [20]$$

since $\nabla \times \nabla = 0$; hence, applying Stokes's theorem, we find with $\mathrm{F} = m\mathrm{g}$ that

$$w = \oint \mathbf{F} \cdot \mathrm{d}\mathbf{s} = 0 \qquad [21]$$

That is, the gravitational field is conservative: the work, w, expended in moving a mass around a closed path in this field vanishes. In contrast, dissipating forces (real forces!), like friction, expend work or energy, which shows again the special nature of the gravitational force.

It can be shown (Kellogg, 1953, p. 156) that the second partial derivatives of a Newtonian potential, V, satisfy the following differential equation, known as Poisson's equation:

$$\nabla^2 V = -4\pi G\rho \qquad [22]$$

where $\nabla^2 = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}$ formally is the scalar product of two gradient operators and is called the Laplacian operator. In Cartesian coordinates, it is given by

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \qquad [23]$$

Note that the Laplacian is a scalar operator. Eqn [22] is a local characterization of the potential field, as opposed to the global characterization given by eqn [14]. Poisson's equation holds wherever the mass density, ρ, satisfies certain conditions similar to continuity (Hölder conditions; see Kellogg, 1953, pp. 152–153). A special case of eqn [22] applies for those points where the density vanishes (i.e., in free space); then Poisson's equation turns into Laplace' equation,

$$\nabla^2 V = 0 \qquad [24]$$

It is easily verified that the point mass potential satisfies eqn [24], that is,

$$\nabla^2 \left(\frac{1}{\ell}\right) = 0 \qquad [25]$$

where $\ell = \sqrt{(x-x')^2+(y-y')^2+(z-z')^2}$ and the mass point is at (x', y', z').

The solutions to Laplace' equation [24] (that is, functions that satisfy Laplace' equations) are known as harmonic functions (here, we also impose the conditions [10] on the solution, if it is a Newtonian potential and if the mass-free region includes infinity). Hence, every Newtonian potential is a harmonic function in free space. The converse is also true: every harmonic function can be represented as a Newtonian potential of a mass distribution (Section 3.02.3.1).

Whether as a volume or a layer density distribution, the corresponding potential is the sum or integral of the source value multiplied by the inverse distance function (or its normal derivative for the dipole layer). This function depends on both the source points and the computation point and is known as a Green's function. It is also known as the 'kernel' function of the integral representation of the potential. Functions of this type also play a dominant role in representing the potential as solutions to boundary-value problems (BVPs), as shown in subsequent sections.

3.02.3 Boundary-Value Problems

If the density distribution of the Earth's interior and the boundary of the volume were known, then the problem of determining the Earth's gravitational potential field is solved by the volume integral of eqn [14]. In reality, of course, we do not have access to this information, at least not the density, with sufficient detail. (The Preliminary Reference Earth Model (PREM), of Dziewonsky and Anderson (1981), still in use today by geophysicists, represents a good profile of Earth's radial density, but does not attempt to model in detail the lateral density heterogeneities.) In this section, we see how the problem of determining the exterior gravitational potential can be solved in terms of surface integrals, thus making exclusive use of accessible measurements on the surface.

3.02.3.1 Green's Identities

Formally, eqn [24] represents a partial differential equation for V. Solving this equation is the essence of the determination of the Earth's external gravitational potential through potential theory. Like any differential equation, a complete solution is obtained only with the application of boundary conditions, that is, imposing values on the solution that it must assume at a boundary of the region in which it is valid. In our case, the boundary is the Earth's surface and the exterior space is where eqn [24] holds (the atmosphere and other celestial bodies are neglected for the moment). In order to study the solutions to these BVPs (to show that solutions exist and are unique), we take advantage of some very important theorems and identities. It is noted that only a rather elementary introduction to BVPs is offered here with no attempt to address the much larger field of solutions to partial differential equations.

The first, seminal result is Gauss' divergence theorem (analogous to Stokes' theorem, eqn [19]),

$$\iiint_v \nabla \cdot \mathbf{f} \; dv = \iint_s f_n \, ds \qquad [26]$$

where **f** is an arbitrary (differentiable) vector function and $f_n = \mathbf{n} \cdot \mathbf{f}$ is the component of **f** along the outward unit normal vector, **n** (see **Figure 5**). The surface, s, encloses the volume, v. Equation [26] says that the sum of how much **f** changes throughout the volume, that is, the net effect, ultimately, is equivalent to the sum of its values projected orthogonally with respect to the surface. Conceptually, a volume integral thus can be replaced by a surface integral, which is important since the gravitational potential is due to a volume density distribution that we do not know, but we do have access to gravitational quantities on a surface by way of measurements.

Equation [26] applies to general vector functions that have continuous first derivatives. In particular, let U and V be two scalar functions having continuous second derivatives, and consider the vector function $\mathbf{f} = U\nabla V$. Then, since $\mathbf{n} \cdot \nabla = \partial/\partial n$, and

$$\nabla \cdot (U\nabla V) = \nabla U \cdot \nabla V + U\nabla^2 V \qquad [27]$$

we can apply Gauss' divergence theorem to get Green's first identity,

$$\iiint_v (\nabla U \cdot \nabla V + U\nabla^2 V) dv = \iint_s U \frac{\partial V}{\partial n} ds \qquad [28]$$

Interchanging the roles of U and V in eqn [28], one obtains a similar formula, which, when subtracted from eqn [28], yields Green's second identity,

$$\iiint_v (U\nabla^2 V - V\nabla^2 U) dv = \iint_s \left(U\frac{\partial V}{\partial n} - V\frac{\partial U}{\partial n} \right) ds \qquad [29]$$

This is valid for any U and V with continuous derivatives up to second order.

Now let $U = 1/\ell$, where ℓ is the usual distance between an integration point and an evaluation point. And, suppose that the volume, v, is the space exterior to the Earth (i.e., Gauss' divergence theorem applies to any volume, not just volumes containing a mass distribution). With reference to **Figure 6**, consider the evaluation point, P, to be inside the volume (free

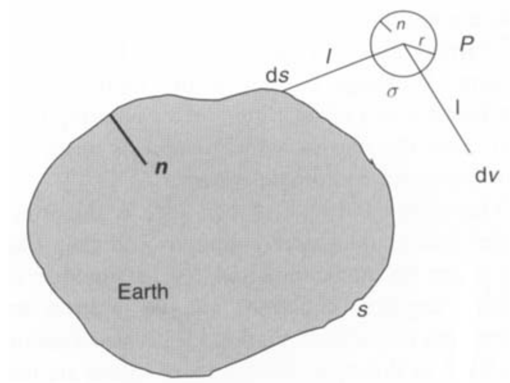

Figure 6 Geometry for special case of Green's third identity.

space) that is bounded by the surface, s; P is thus outside the Earth's surface. Let V be a solution to eqn [24], that is, it is the gravitational potential of the Earth. From the volume, v, exclude the volume bounded by a small sphere, σ, centered at P. This sphere becomes part of the surface that bounds the volume, v. Then, since U, by our definition, is a point mass potential, $\nabla^2 U = 0$ everywhere in v (which excludes the interior of the small sphere around P); and, the second identity [29] gives

$$\iint_s \left(\frac{1}{\ell}\frac{\partial V}{\partial n} - V\frac{\partial}{\partial n}\left(\frac{1}{\ell}\right) \right) ds + \iint_\sigma \left(\frac{1}{\ell}\frac{\partial V}{\partial n} - V\frac{\partial}{\partial n}\left(\frac{1}{\ell}\right) \right) d\sigma = 0 \qquad [30]$$

The unit vector, **n**, represents the perpendicular pointing away from v. On the small sphere, **n** is opposite in direction to $\ell = r$, and the second integral becomes

$$\begin{aligned} \iint_\sigma \left(-\frac{1}{r}\frac{\partial V}{\partial r} + V\frac{\partial}{\partial r}\left(\frac{1}{r}\right) \right) d\sigma &= -\iint_\Omega \frac{1}{r}\frac{\partial V}{\partial r} r^2 \, d\Omega - \iint_\Omega V \, d\Omega \\ &= -\iint_\Omega \frac{\partial V}{\partial r} r \, d\Omega - 4\pi \bar{V} \end{aligned} \qquad [31]$$

where $d\sigma = r^2 d\Omega$, Ω is the solid angle, 4π, and $\bar{V}$ is an average value of V on σ. Now, in the limit as the radius of the small sphere shrinks to zero, the right-hand side of eqn [31] approaches $0 - 4\pi V_P$. Hence, eqn [30] becomes (Kellogg, 1953, p. 219)

$$V_P = \frac{1}{4\pi}\iint_s \left(\frac{1}{\ell}\frac{\partial V}{\partial n} - V\frac{\partial}{\partial n}\left(\frac{1}{\ell}\right) \right) ds \qquad [32]$$

with **n** pointing down (away from the space outside the Earth). This is a special case of Green's third identity. A change in sign of the right-hand side transforms **n** to a normal unit vector pointing into v, away from the masses, which conforms more to an Earth-centered coordinate system.

The right-hand side of eqn [32] is the sum of single- and double-layer potentials and thus shows that every harmonic function (i.e., a function that satisfies Laplace' equation) can be written as a Newtonian potential. Equation [32] is also a solution to a BVP; in this case, the boundary values are independent values of V and of its normal derivative, both on s (Cauchy problem). Below and in Section 3.02.4.1, we encounter another BVP in which the potential and its normal derivative are given in a specified linear combination on s. Using a similar procedure and with some extra care, it can be shown (see also Courant and Hilbert, 1962, vol. II, p. 256 (footnote)) that if P is on the surface, then

$$V_P = \frac{1}{2\pi}\iint_s \left(\frac{1}{\ell}\frac{\partial V}{\partial n} - V\frac{\partial}{\partial n}\left(\frac{1}{\ell}\right)\right)\mathrm{d}s \qquad [33]$$

where n points into the masses. Comparing this to eqn [32], we see that V is discontinuous as one approaches the surface; this is due to the double-layer part (see eqn [18]).

Equation [32] demonstrates that a solution to a particular BVP exists. Specifically, we are able to measure the potential (up to a constant) and its derivatives (the gravitational acceleration) on the surface and thus have a formula to compute the potential anywhere in exterior space, provided we also know the surface, s. Other BVPs also have solutions under appropriate conditions; a discussion of existence theorems is beyond the present scope and may be found in (Kellogg, 1953). Equation [33] has deep geodetic significance. One objective of geodesy is to determine the shape of the Earth's surface. If we have measurements of gravitational quantities on the Earth's surface, then conceptually we are able to determine its shape from eqn [33], where it would be the only unknown quantity. This is the basis behind the work of Molodensky *et al.* (1962), to which we return briefly at the end of this section.

3.02.3.2 Uniqueness Theorems

Often the existence of a solution is proved simply by finding one (as illustrated above). Whether such as solution is the only one depends on a corresponding uniqueness theorem. That is, we wish to know if a certain set of boundary values will yield just one potential in space. Before considering such theorems, we classify the BVPs that are typically encountered when determining the exterior potential from measurements on a boundary. In all cases, it is an exterior BVP; that is, the gravitational potential, V, is harmonic ($\nabla^2 V = 0$) in the space exterior to a closed surface that contains all the masses. The exterior space thus contains infinity. Interior BVPs can be constructed, as well, but are not applicable to our objectives.

- *Dirichlet problem* (*or, BVP of the first kind*). Solve for V in the exterior space, given its values everywhere on the boundary.
- *Neumann problem* (*or, BVP of the second kind*). Solve for V in the exterior space, given values of its normal derivative everywhere on the boundary.
- *Robin problem* (*mixed BVP, or BVP of the third kind*). Solve for V in the exterior space, given a linear combination of it and its normal derivative on the boundary.

Using Green's identities, we prove the following theorems for these exterior problems; similar results hold for the interior problems.

Theorem 1. If V is harmonic (hence continuously differentiable) in a closed region, v, and if V vanishes everywhere on the boundary, s, then V also vanishes everywhere in the region, v.

Proof. Since $V = 0$ on s, Green's first identity (eqn [28]) with $U = V$ gives

$$\iiint_v (\nabla V)^2\,\mathrm{d}v = \iint_s V\frac{\partial V}{\partial n}\,\mathrm{d}s = 0 \qquad [34]$$

The integral on the left side is therefore always zero, and the integrand is always non-negative. Hence, $\boldsymbol{\nabla} V = \mathbf{0}$ everywhere in v, which implies that $V =$ constant in v. Since V is continuous in v and $V = 0$ on s, that constant must be zero; and so $V = 0$ in v.

This theorem solves the Dirichlet problem for the trivial case of zero boundary values and it enables the following uniqueness theorem for the general Dirichlet problem.

Theorem 2 (*Stokes' theorem*). If V is harmonic (hence continuously differentiable) in a closed region, v, then V is uniquely determined in v by its values on the boundary, s.

Proof. Suppose the determination is not unique: that is, suppose there are V_1 and V_2, both harmonic in v and both having the same boundary values on s. Then the function $V = V_2 - V_1$ is harmonic in v with all boundary values equal to zero. Hence, by Theorem 1, $V_2 - V_1 = 0$ identically in v, or $V_2 = V_1$ everywhere, which implies that any determination is unique based on the boundary values.

Theorem 3. If V is harmonic (hence continuously differentiable) in the exterior region, v, with closed boundary, s, then V is uniquely determined by the values of its normal derivative on s.

Proof. We begin with Green's first identity, eqn [28], as in the proof of Theorem 1 to show that if the normal derivative vanishes everywhere on s, then V is a constant in v. Now, suppose there are two harmonic functions in v: V_1 and V_2, with the same normal derivative values on s. Then the normal derivative values of their difference are zero; and, by the above demonstration, $V = V_2 - V_1 =$ constant in v. Since V is a Newtonian potential in the exterior space, that constant is zero, since by eqn [10], $\lim_{\ell \to \infty} V = 0$. Thus, $V_2 = V_1$, and the boundary values determine the potential uniquely.

This is a uniqueness theorem for the exterior Neumann BVP. Solutions to the interior problem are unique only up to an arbitrary constant.

Theorem 4. Suppose V is harmonic (hence continuously differentiable) in the closed region, v, with boundary, s; and, suppose the boundary values are given by

$$g = \alpha V|_s + \beta \frac{\partial V}{\partial n}\bigg|_s \qquad [35]$$

Then V is uniquely determined by these values if $\alpha/\beta > 0$.

Proof. Suppose there are two harmonic functions, V_1 and V_2, with the same boundary values, g, on s. Then $V = V_2 - V_1$ is harmonic with boundary values

$$\alpha(V_2 - V_1)|_s + \beta\left(\frac{\partial V_2}{\partial n} - \frac{\partial V_1}{\partial n}\right)\bigg|_s = 0 \qquad [36]$$

With $U = V = V_2 - V_1$, Green's first identity, eqn [28], gives

$$\iiint_v (\nabla(V_2 - V_1))^2\, \mathrm{d}v = \iint_s (V_2 - V_1)\frac{-\alpha}{\beta}(V_2 - V_1)\, \mathrm{d}s \qquad [37]$$

Then

$$\iiint_v (\nabla(V_2 - V_1))^2\, \mathrm{d}v + \frac{\alpha}{\beta}\iint_s (V_2 - V_1)^2\, \mathrm{d}S = 0 \qquad [38]$$

Since $\alpha/\beta > 0$, eqn [38] implies that $\nabla(V_2 - V_1) = 0$ in v; and $V_2 - V_1 = 0$ on s. Hence $V_2 - V_1 =$ constant in v; and $V_2 = V_1$ on s. By the continuity of V_1 and V_2, the constant must be zero, and the uniqueness is proved.

The solution to the Robin problem is unique only in certain cases. The most famous problem in physical geodesy is the determination of the disturbing potential, T, from gravity anomalies, Δg, on the geoid (Section 3.02.1.3). Suppose T is harmonic outside the geoid; the second of eqns [5] provides an approximate form of boundary condition, showing that this is a type of Robin problem. We find that $\alpha = -2/r$, and, recalling that when v is the exterior space the unit vector n points inward toward the masses, that is, $\partial/\partial n = -\partial/\partial r$, we get $\beta = 1$. Thus, the condition in Theorem 4 on α/β is not fulfilled and the uniqueness is not guaranteed. In fact, we will see that the solution obtained for the spherical boundary is arbitrary with respect to the coordinate origin (Section 3.02.4.1).

3.02.3.3 Solutions by Integral Equation

Green's identities show how a solution to Laplace's equation can be transformed from a volume integral, that is, an integral of source points, to a surface integral of BVPs, as demonstrated by eqn [32]. The uniqueness theorems for the BVPs suggest that the potential due to a volume density distribution can also be represented as due to a generalized density layer on the bounding surface, as long as the result is harmonic in exterior space, satisfies the boundary conditions, and is regular at infinity like a Newtonian potential. Molodensky *et al.* (1962) supposed that the disturbing potential is expressible as

$$T = \iint_s \frac{\mu}{\ell}\, \mathrm{d}s \qquad [39]$$

where μ is a surface density to be solved using the boundary condition. With the spherical approximation for the gravity anomaly, eqn [5], one arrives at the following integral equation

$$2\pi\mu\cos\zeta - \iint_s \left(\frac{\partial}{\partial r}\left(\frac{1}{\ell}\right) + \frac{2}{r\ell}\right)\mu\, \mathrm{d}s = \Delta g \qquad [40]$$

The first term accounts for the discontinuity at the surface of the derivative of the potential of a density layer, where ζ is the deflection angle between the normal to the surface and the direction of the (radial) derivative (Heiskanen and Moritz, 1967, p. 6; Günter, 1967, p. 69). This Fredholm integral equation of the second kind can be simplified with further approximations, and, a solution for the density, μ, ultimately leads to the solution for the disturbing potential (Moritz, 1980).

Other forms of the initial representation have also been investigated, where Green's functions other than $1/\ell$ lead to simplifications of the integral equation (e.g., Petrovskaya, 1979). Nevertheless, most practical solutions rely on approximations, such as the spherical approximation for the boundary condition, and even the formulated solutions are not strictly guaranteed to converge to the true solutions (Moritz, 1980). Further treatments of the BVP in a geodetic/mathematical setting may be found in the volume edited by Sansò and Rummel (1997).

In the next section, we consider solutions for T as surface integrals of boundary values with appropriate Green's functions. In other words, the boundary values (whether of the first, second, or third kind) may be thought of as sources, and the consequent potential is again the sum (integral) of the product of a boundary value and an appropriate Green's function (i.e., a function that depends on both the source point and the computation point in some form of inverse distance in accordance with Newtonian potential theory). Such solutions are readily obtained if the boundary is a sphere.

3.02.4 Solutions to the Spherical BVP

This section develops two types of solutions to standard BVPs when the boundary is a sphere: the spherical harmonic series and an integral with a Green's function. All three types of problems are solved, but emphasis is put on the third BVP since gravity anomalies are the most prevalent boundary values (on land, at least). In addition, it is shown how the Green's function integrals can be inverted to obtain, for example, gravity anomalies from values of the potential, now considered as boundary values. Not all possible inverse relationships are given, but it should be clear at the end that, in principle, virtually any gravitational quantity can be obtained in space from any other quantity on the spherical boundary.

3.02.4.1 Spherical Harmonics and Green's Functions

For simple boundaries, Laplace's equation [24] is relatively easy to solve provided there is an appropriate coordinate system. For the Earth, the solutions commonly rely on approximating the boundary by a sphere. This case is described in detail and a more accurate approximation based on an ellipsoid of revolution is briefly examined in Section 3.02.5.2 for the normal potential. In spherical polar coordinates, (θ, λ, r), the Laplacian operator is given by Hobson (1965, p. 9)

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\lambda^2} \qquad [41]$$

A solution to $\nabla^2 V = 0$ in the space outside a sphere of radius, R, with center at the coordinate origin can be found by the method of separation of variables, whereby one postulates the form of the solution, V, as

$$V(\theta, \lambda, r) = f(\theta)g(\lambda)h(r) \qquad [42]$$

Substituting this and the Laplacian above into eqn [24], the multivariate partial differential equation separates into three univariate ordinary differential equations (Hobson, 1965, p. 9; Morse and Feshbach, 1953, p. 1264). Their solutions are well-known functions, for example,

$$V(\theta, \lambda, r) = P_{nm}(\cos\theta)\sin m\lambda\frac{1}{r^{n+1}} \qquad [43a]$$

or

$$V(\theta, \lambda, r) = P_{nm}(\cos\theta)\cos m\lambda\frac{1}{r^{n+1}} \qquad [43b]$$

where $P_{nm}(t)$ is the associated Legendre function of the first kind and n, m are integers such that $0 \leq m \leq n$, $n \geq 0$. Other solutions are also possible (e.g., $g(\lambda) = e^{a\lambda}$ $(a \in \mathbb{R})$ and $h(r) = r^n$), but only eqns [43] are consistent with the problem at hand: to find a real-valued Newtonian potential for the exterior space of the Earth (regular at infinity and 2π-periodic in longitude). The general solution is a linear combination of solutions of the forms given by eqns [43] for all possible integers, n and m, and can be written compactly as

$$V(\theta, \lambda, r) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left(\frac{R}{r}\right)^{n+1} v_{nm}\bar{Y}_{nm}(\theta, \lambda) \qquad [44]$$

where the $\bar{Y}_{nm}$ are surface spherical harmonic functions defined as

$$\bar{Y}_{nm}(\theta, \lambda) = \bar{P}_{n|m|}(\cos\theta)\begin{cases}\cos m\lambda, & m \geq 0\\ \sin|m|\lambda, & m < 0\end{cases} \qquad [45]$$

and $\bar{P}_{nm}$ is a normalization of P_{nm} so that the orthogonality of the spherical harmonics is simply

$$\frac{1}{4\pi}\iint_{\sigma} \bar{Y}_{nm}(\theta,\lambda)\,\bar{Y}_{n'm'}(\theta,\lambda)\,\mathrm{d}\sigma = \begin{cases} 1, & n = n' \quad \text{and} \quad m = m' \\ 0, & n \neq n' \quad \text{or} \quad m \neq m' \end{cases} \qquad [46]$$

and where $\sigma = \{(\theta,\lambda)|0 \le \theta \le \pi,\ 0 \le \lambda \le 2\pi\}$ represents the unit sphere, with $\mathrm{d}\sigma = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\lambda$. For a complete mathematical treatment of spherical harmonics, one may refer to Müller (1966). The bounding spherical radius, R, is introduced so that all the constant coefficients, v_{nm}, also known as Stokes constants, have identical units of measure. Applying the orthogonality to the general solution [44], these coefficients can be determined if the function, V, is known on the bounding sphere (boundary condition):

$$v_{nm} = \frac{1}{4\pi}\iint_{\sigma} V(\theta,\lambda,R)\,\bar{Y}_{nm}(\theta,\lambda)\mathrm{d}\sigma \qquad [47]$$

Equation [44] is known as a spherical harmonic expansion of V and with eqn [47] it represents a solution to the Dirichlet BVP if the boundary is a sphere. The solution thus exists and is unique in the sense that these boundary values generate no other potential. We will, however, find another equivalent form of the solution.

In a more formal mathematical setting, the solution [46] is an infinite linear combination of orthogonal basis functions (eigenfunctions) and the coefficients, v_{nm}, are the corresponding eigenvalues. One may also interpret the set of coefficients as the spectrum (Legendre spectrum) of the potential on the sphere of radius, R (analogous to the Fourier spectrum of a function on the plane or line). The integers, n, m, correspond to wave numbers, and are called degree (n) and order (m), respectively. The spherical harmonics are further classified as zonal ($m = 0$), meaning that the zeros of $\bar{Y}_{n0}$ divide the sphere into latitudinal zones; sectorial ($m = n$), where the zeros of $\bar{Y}_{nn}$ divide the sphere into longitudinal sectors; and tesseral (the zeros of $\bar{Y}_{nm}$ tessellate the sphere) (**Figure 7**).

While the spherical harmonic series has its advantages in global representations and spectral interpretations of the field, a Green's function representation provides a more local characterization of the field. Changing a boundary value anywhere on the globe changes all coefficients, v_{nm}, according to eqn [47], which poses both a numerical challenge in applications, as well as in keeping a standard model up to date. However, since the Green's function essentially depends on the inverse distance (or higher power thereof), a remote change in boundary value generally does not appreciably affect the local determination of the field.

When the boundary is a sphere, the solutions to the BVPs using a Green's function are easily derived from the spherical harmonic series representation. Moreover, it is possible to derive additional integral relationships (with appropriate Green's functions) among all the derivatives of the potential. To formalize and simultaneously simplify these derivations, consider harmonic functions, f and h, where h depends only on θ and r, and function g, defined on the sphere of radius, R. Thus let

$$f(\theta,\lambda,r) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left(\frac{R}{r}\right)^{n+1} f_{nm}\,\bar{Y}_{nm}(\theta,\lambda) \qquad [48]$$

$$h(\theta,r) = \sum_{n=0}^{\infty}(2n+1)\left(\frac{R}{r}\right)^{n+1} h_n P_n(\cos\theta) \qquad [49]$$

$$g(\theta,\lambda,R) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} g_{nm}\,\bar{Y}_{nm}(\theta,\lambda) \qquad [50]$$

where $P_n(\cos\theta) = \bar{P}_{n0}(\cos\theta)/\sqrt{2n+1}$ is the nth degree Legendre polynomial. Constants f_{nm} and g_{nm} are the respective harmonic coefficients of f and g when these functions are restricted to the sphere of radius, R. Then, using the decomposition formula for Legendre polynomials,

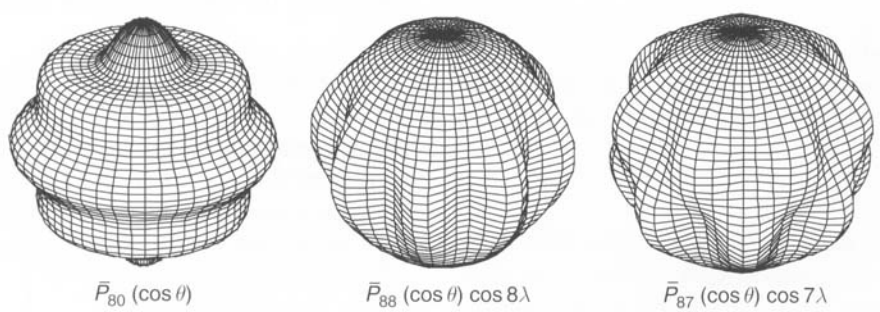

Figure 7 Examples of zonal, sectorial, and tesseral harmonics on the sphere.

$$P_n(\cos\psi) = \frac{1}{2n+1}\sum_{m=-n}^{n} \bar{Y}_{nm}(\theta,\lambda)\bar{Y}_{nm}(\theta',\lambda') \qquad [51]$$

where

$$\cos\psi = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\lambda-\lambda') \qquad [52]$$

it is easy to prove the following theorem.

Theorem (convolution theorem in spectral analysis on the sphere).

$$f(\theta,\lambda,r) = \frac{1}{4\pi}\iint_\sigma g(\theta',\lambda',R)h(\psi,r)\mathrm{d}\sigma \quad \text{if and only if} \quad f_{nm} = g_{nm}h_n \qquad [53]$$

Here, and in the following, $\mathrm{d}\sigma = \sin\theta'\mathrm{d}\theta'\mathrm{d}\lambda'$. The angle, ψ, is the distance on the unit sphere between points (θ, λ) and (θ', λ').

Proof. The forward statement [53] follows directly by substituting eqns [51] and [49] into the first equation [53], together with the spherical harmonic expansion [50] for g. A comparison with the spherical harmonic expansion for f yields the result. All steps in this proof are reversible, and so the reverse statement also holds.

Consider now f to be the potential, V, outside the sphere of radius, R, and its restriction to the sphere to be the function, $g: g(\theta,\lambda) = V(\theta,\lambda,R)$. Then, clearly, $h_n = 1$, for all n; by the theorem above, we have

$$V(\theta,\lambda,r) = \frac{1}{4\pi}\iint_\sigma V(\theta',\lambda',R)U(\psi,r)\mathrm{d}\sigma \qquad [54]$$

where

$$U(\psi,r) = \sum_{n=0}^{\infty}(2n+1)\left(\frac{R}{r}\right)^{n+1}P_n(\cos\psi) \qquad [55]$$

For the distance

$$\ell = \sqrt{r^2 + R^2 - 2rR\cos\psi} \qquad [56]$$

between points (θ,λ,r) and (θ',λ',R), with $r \geq R$, the identity (the Coulomb expansion; Cushing, 1975, p. 155),

$$\frac{1}{\ell} = \frac{1}{R}\sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+1}P_n(\cos\psi) \qquad [57]$$

yields, after some arithmetic (based on taking the derivative on both sides with respect to r),

$$U(\psi,r) = \frac{R(r^2-R^2)}{\ell^3} \qquad [58]$$

Solutions [44] and [54] to the Dirichlet BVP for a spherical boundary are identical (in view of the convolution theorem [53]). The integral in [54] is known as the Poisson integral and the function U is the corresponding Green's function, also known as Poisson's kernel.

For convenience, one separates Earth's gravitational potential into a reference potential (Section 3.02.1.3) and the disturbing potential, T. The disturbing potential is harmonic in free space and satisfies the Poisson integral if the boundary is a sphere. In deference to physical geodesy where relationships between the disturbing potential and its derivatives are routinely applied, the following derivations are developed in terms of T, but hold equally for any exterior Newtonian potential. Let

$$T(\theta,\lambda,r) = \frac{GM}{R}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left(\frac{R}{r}\right)^{n+1}\delta C_{nm}\bar{Y}_{nm}(\theta,\lambda) \qquad [59]$$

where M is the total mass (including the atmosphere) of the Earth and the δC_{nm} are unitless harmonic coefficients, being also the difference between coefficients for the total and reference gravitational potentials (Section 3.02.5.2). The coefficient, δC_{00}, is zero under the assumption that the reference field accounts completely for the central part of the total field. Also note that these coefficients specifically refer to the sphere of radius, R.

The gravity disturbance is defined (in spherical approximation) to be the negative radial derivative of T, the first of eqns [5]. From eqn [59], we have

$$\begin{aligned}\delta g(\theta,\lambda,r) &= -\frac{\partial}{\partial r}T(\theta,\lambda,r)\\ &= \frac{GM}{R^2}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left(\frac{R}{r}\right)^{n+2}(n+1)\delta C_{nm}\bar{Y}_{nm}(\theta,\lambda)\end{aligned} \qquad [60]$$

and, applying the convolution theorem [53], we obtain

$$T(\theta,\lambda,r) = \frac{R}{4\pi}\iint_\sigma \delta g(\theta',\lambda',R)H(\psi,r)\mathrm{d}\sigma \qquad [61]$$

where with $g_{nm} = (n+1)\delta C_{nm}/R$ and $f_{nm} = \delta C_{nm}$, we have $h_n = f_{nm}/g_{nm} = R/(n+1)$, and hence (taking care to keep the Green's function unitless)

$$H(\psi,r) = \sum_{n=0}^{\infty}\frac{2n+1}{n+1}\left(\frac{R}{r}\right)^{n+1}P_n(\cos\psi) \qquad [62]$$

The integral in [61] is known as the Hotine integral, the Green's function, H, is called the Hotine kernel, and with a derivation based on equation [57], it is given by (Hotine 1969, p. 311)

$$H(\psi,r) = \frac{2R}{\ell} - \ln\left(1 + \frac{\ell}{2R\sin^2\psi/2}\right) \qquad [63]$$

Equation [61] solves the Neumann BVP when the boundary is a sphere.

The gravity anomaly (again, in spherical approximation) is defined by eqn [5]

$$\Delta g(\theta, \lambda, r) = \left(-\frac{\partial}{\partial r} - \frac{2}{r}\right) T(\theta, \lambda, r) \qquad [64]$$

or, also,

$$\Delta g(\theta, \lambda, r) = \frac{GM}{R^2} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{R}{r}\right)^{n+2} (n-1)\delta C_{nm} \bar{Y}_{nm}(\theta, \lambda) \qquad [65]$$

In this case, we have $g_{nm} = (n-1)\delta C_{nm}/R$ and $h_n = R/(n-1)$. The convolution theorem in this case leads to the geodetically famous Stokes integral,

$$T(\theta, \lambda, r) = \frac{R}{4\pi} \iint_\sigma \Delta g(\theta', \lambda', R) S(\psi, r) \mathrm{d}\sigma \qquad [66]$$

where we define Green's function to be

$$\begin{aligned} S(\psi, r) &= \sum_{n=2}^{\infty} \frac{2n+1}{n-1} \left(\frac{R}{r}\right)^{n+1} P_n(\cos\psi) \\ &= 2\frac{R}{\ell} + \frac{R}{r} - 3\frac{R\ell}{r^2} - 5\frac{R^2}{r^2}\cos\psi \\ &\quad - 3\frac{R^2}{r^2}\cos\psi \ln\frac{\ell + r - R\cos\psi}{2r} \end{aligned} \qquad [67]$$

more commonly called the Stokes kernel. Equation [66] solves the Robin BVP if the boundary is a sphere, but it includes specific constraints that ensure the solution's uniqueness – the solution by itself is not unique, in this case, as proved in Section 3.02.3.2. Indeed, eqn [65] shows that the gravity anomaly has no first-degree harmonics for the disturbing potential; therefore, they cannot be determined from the boundary values. Conventionally, the Stokes kernel also excludes the zero-degree harmonic, and thus the complete solution for the disturbing potential is given by

$$\begin{aligned} T(\theta, \lambda, r) &= \frac{GM}{r}\delta C_{00} \\ &\quad + \frac{GM}{R} \sum_{m=-1}^{1} \left(\frac{R}{r}\right)^2 \delta C_{1m} \bar{Y}_{1m}(\theta, \lambda) \\ &\quad + \frac{R}{4\pi} \iint_\sigma \Delta g(\theta', \lambda', R) S(\psi, r)\, \mathrm{d}\sigma \end{aligned} \qquad [68]$$

The central term, δC_{00}, is proportional to the difference in GM of the Earth and reference ellipsoid and is zero to high accuracy. The first-degree harmonic coefficients, δC_{1m}, are proportional to the center-of-mass coordinates and can also be set to zero with appropriate definition of the coordinate system (see Section 3.02.5.1). Thus, the Stokes integral [66] is the more common expression for the disturbing potential, but it embodies hidden constraints.

We note that gravity anomalies also serve as boundary values in the harmonic series form of the solution for the disturbing potential. Applying the orthogonality of the spherical harmonics to eqn [65] yields immediately

$$\delta C_{nm} = \frac{R^2}{4\pi(n-1)GM} \iint_\sigma \Delta g(\theta', \lambda', R) \bar{Y}_{nm}(\theta', \lambda') \mathrm{d}\sigma, \quad n \geq 2 \qquad [69]$$

A similar formula holds when gravity disturbances are the boundary values ($n-1$ in the denominator changes to $n+1$). In either case, the boundary values formally are assumed to reside on a sphere of radius, R. An approximation results if they are given on the geoid, as is usually the case.

3.02.4.2 Inverse Stokes and Hotine Integrals

The convolution integrals above can easily be inverted by considering again the spectral relationships. For the gravity anomaly, we note that $f = r\Delta g$ is harmonic with coefficients, $f_{nm} = GM(n-1)\delta C_{nm}/R$. Letting $g_{nm} = GM\ \delta C_{nm}/R$, we find that $h_n = n-1$; from the convolution theorem, we can write

$$\Delta g(\theta, \lambda, r) = \frac{1}{4\pi R} \iint_\sigma T(\theta', \lambda', R) \hat{Z}(\psi, r)\, \mathrm{d}\sigma \qquad [70]$$

where

$$\hat{Z}(\psi, r) = \sum_{n=0}^{\infty} (2n+1)(n-1) \left(\frac{R}{r}\right)^{n+2} P_n(\cos\psi) \qquad [71]$$

The zero- and first-degree terms are included provisionally. Note that

$$\hat{Z}(\psi, r) = -R\left(\frac{\partial}{\partial r} + \frac{2}{r}\right) U(\psi, r) \qquad [72]$$

that is, we could have simply used the Dirichlet solution [54] to obtain the gravity anomaly, as given by [70], from the disturbing potential. It is convenient to separate the kernel function as follows:

$$\hat{Z}(\psi, r) = Z(\psi, r) - \sum_{n=0}^{\infty} (2n+1) \left(\frac{R}{r}\right)^{n+2} P_n(\cos\psi) \qquad [73]$$

where

$$Z(\psi, r) = \sum_{n=1}^{\infty} (2n+1) n \left(\frac{R}{r}\right)^{n+2} P_n(\cos\psi) \qquad [74]$$

Next Page

We find that

$$\Delta g(\theta, \lambda, r) = -\frac{1}{r} T(\theta, \lambda, r) + \frac{1}{4\pi R} \iint_\sigma T(\theta', \lambda', R) Z(\psi, r)\, d\sigma \quad [75]$$

Now, since Z has no zero-degree harmonics, its integral over the sphere vanishes, and one can write the numerically more convenient formula:

$$\Delta g(\theta, \lambda, r) = -\frac{1}{r} T(\theta, \lambda, r) + \frac{1}{4\pi R} \iint_\sigma (T(\theta', \lambda', R) - T(\theta, \lambda, R)) Z(\psi, r)\, d\sigma \quad [76]$$

This is the inverse Stokes formula. Given T on the sphere of radius R (e.g., in the form of geoid undulations, $T = \gamma N$), this form is useful when the gravity anomaly is also desired on this sphere. It is one way to determine gravity anomalies on the ocean surface from satellite altimetry, where the ocean surface is approximated as a sphere. Analogously, from eqns [60] and [64], it is readily seen that the inverse Hotine formula is given by

$$\delta g(\theta, \lambda, r) = \frac{1}{r} T(\theta, \lambda, r) + \frac{1}{4\pi R} \iint_\sigma (T(\theta', \lambda', R) - T(\theta, \lambda, R)) Z(\psi, r)\, d\sigma \quad [77]$$

Note that the difference of eqns [76] and [77] yields the approximate relationship between the gravity disturbance and the gravity anomaly inferred from eqns [5].

Finally, we realize that, for $r = R$, the series for $Z(\psi, R)$ is not uniformly convergent and special numerical procedures (that are outside the present scope) are required to approximate the corresponding integrals.

3.02.4.3 Vening-Meinesz Integral and Its Inverse

Other derivatives of the disturbing potential may also be determined from boundary values. We consider here only gravity anomalies, being the most prevalent data type on land areas. The solution is either in the form of a series – simply the derivative of the series [59] with coefficients given by eqn [69], or an integral with appropriate derivative of the Green's function. The horizontal derivatives of the disturbing potential are often interpreted as the deflections of the vertical to which they are proportional in spherical approximation (eqn [5]):

$$\begin{Bmatrix} \xi(\theta, \lambda, r) \\ \eta(\theta, \lambda, r) \end{Bmatrix} = \frac{1}{\gamma(\theta, r)} \begin{Bmatrix} \dfrac{1}{r}\dfrac{\partial}{\partial\theta} \\ -\dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\lambda} \end{Bmatrix} T(\theta, \lambda, r) \quad [78]$$

where ξ, η are the north and east deflection components, respectively, and γ is the normal gravity. Clearly, the derivatives can be taken directly inside the Stokes integral, and we find

$$\begin{Bmatrix} \xi(\theta, \lambda, r) \\ \eta(\theta, \lambda, r) \end{Bmatrix} = \frac{R}{4\pi r\ \gamma(\theta, r)} \iint_\sigma \Delta g(\theta', \lambda', R) \times \frac{\partial}{\partial\psi} S(\psi, r) \begin{Bmatrix} \cos\alpha \\ \sin\alpha \end{Bmatrix} d\sigma \quad [79]$$

where

$$\begin{Bmatrix} \dfrac{1}{r}\dfrac{\partial}{\partial\theta} \\ -\dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\lambda} \end{Bmatrix} = \frac{1}{r} \begin{Bmatrix} \dfrac{\partial\psi}{\partial\theta} \\ -\dfrac{1}{\sin\theta}\dfrac{\partial\psi}{\partial\lambda} \end{Bmatrix} \frac{\partial}{\partial\psi} \quad [80]$$

and

$$\frac{\partial\psi}{\partial\theta} = \frac{1}{\sin\psi}(\sin\theta\cos\theta' - \cos\theta\sin\theta'\cos(\lambda' - \lambda)) = \cos\alpha$$
$$-\frac{1}{\sin\theta}\frac{\partial\psi}{\partial\lambda} = \frac{1}{\sin\psi}\sin\theta'\sin(\lambda' - \lambda) = \sin\alpha \quad [81]$$

The angle, α, is the azimuth of (θ', λ') at (θ, λ) on the unit sphere. The integrals [79] are known as the Vening-Meinesz integrals. Analogous integrals for the deflections arise when the boundary values are the gravity disturbances (the Green's functions are then derivatives of the Hotine kernel).

For the inverse Vening-Meinesz integrals, we need to make use of Green's first identity for surface functions, f and g:

$$\iint_s f\Delta^*(g)\, ds + \iint_s \nabla f \cdot \nabla g\, ds = \int_b f\ \nabla g \cdot \mathbf{n}\, db \quad [82]$$

where b is the boundary (a line) of surface, s; ∇ and Δ^* are the gradient and Laplace–Beltrami operators, which for the spherical surface are given by

$$\nabla = \begin{pmatrix} \dfrac{\partial}{\partial\theta} \\ \dfrac{1}{\sin\theta}\dfrac{\partial}{\partial\lambda} \end{pmatrix}, \quad \Delta^* = \frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\lambda^2} \quad [83]$$

Previous Page

and where $\mathbf{n}$ is the unit vector normal to b. For a closed surface such as the sphere, the line integral vanishes, and we have

$$\iint_\sigma f\Delta^*(g)\,\mathrm{d}\sigma = -\iint_\sigma \nabla f\cdot\nabla g\ \mathrm{d}\sigma \qquad [84]$$

The surface spherical harmonics, $\bar{Y}_{nm}(\theta,\lambda)$, satisfy the following differential equation:

$$\Delta^*\bar{Y}_{nm}(\theta,\lambda) + n(n+1)\bar{Y}_{nm}(\theta,\lambda) = 0 \qquad [85]$$

Therefore, the harmonic coefficients of $\Delta^* T(\theta,\lambda,r)$ on the sphere of radius, R, are

$$[\Delta^* T(\theta,\lambda,r)]_{nm} = -n(n+1)\frac{GM}{R}\delta C_{nm} \qquad [86]$$

Hence, by the convolution theorem (again, considering the harmonic function, $f = r\Delta g$),

$$\Delta g(\theta,\lambda,r) = -\frac{1}{4\pi R}\iint_\sigma \Delta^* T(\theta',\lambda',R)W(\psi,r)\,\mathrm{d}\sigma \qquad [87]$$

where

$$W(\psi,r) = \sum_{n=2}^{\infty}\frac{(2n+1)(n-1)}{n(n+1)}\left(\frac{R}{r}\right)^{n+2}P_n(\cos\psi) \qquad [88]$$

and where the zero-degree term of the gravity anomaly must be treated separately (e.g., it is set to zero in this case). Using Green's identity [84] and eqns [80] and [81], we have

$$\begin{aligned}\Delta g(\theta,\lambda,r) &= \frac{1}{4\pi R}\iint_\sigma \nabla T(\theta',\lambda',R)\cdot\nabla W(\psi,r)\,\mathrm{d}\sigma\\ &= \frac{R\gamma_0}{4\pi R}\iint_\sigma (\xi(\theta',\lambda',R)\cos\alpha\\ &\quad + \eta(\theta',\lambda',R)\sin\alpha)\frac{\partial}{\partial\psi}W(\psi,r)\,\mathrm{d}\sigma \qquad [89]\end{aligned}$$

where normal gravity on the sphere of radius, R, is approximated as a constant: $\gamma(\theta,R)\simeq\gamma_0$. Equation [89] represents a second way to compute gravity anomalies from satellite altimetry, where the along-track and cross-track altimetric differences are used to approximate the deflection components (with appropriate rotation to north and east components). Employing differences in altimetric measurements benefits the estimation since systematic errors, such as orbit error, cancel out. To speed up the computations, the problem is reformulated in the spectral domain (see, e.g., Sandwell and Smith, 1996).

Clearly, following the same procedure for $f = T$, we also have the following relationship:

$$\begin{aligned}T(\theta,\lambda,r) &= \frac{R\gamma_0}{4\pi}\iint_\sigma (\xi(\theta',\lambda',R)\cos\alpha\\ &\quad + \eta(\theta',\lambda',R)\sin\alpha)\frac{\partial}{\partial\psi}B(\psi,r)\,\mathrm{d}\sigma \qquad [90]\end{aligned}$$

where

$$B(\psi,r) = \sum_{n=2}^{\infty}\frac{(2n+1)}{n(n+1)}\left(\frac{R}{r}\right)^{n+1}P_n(\cos\psi) \qquad [91]$$

It is interesting to note that instead of an integral over the sphere, the inverse relationship between the disturbing potential on the sphere and the deflection of the vertical on the same sphere is also more straightforward in terms of a line integral:

$$\begin{aligned}T(\theta,\lambda,R) &= T(\theta_0,\lambda_0,R) + \frac{\gamma_0}{4\pi}\int_{(\theta_0,\lambda_0)}^{(\theta,\lambda)}(\xi(\theta',\lambda',R)\,\mathrm{d}s_\theta\\ &\quad - \eta(\theta',\lambda',R)\,\mathrm{d}s_\lambda) \qquad [92]\end{aligned}$$

where

$$\mathrm{d}s_\theta = R\,\mathrm{d}\theta,\ \mathrm{d}s_\lambda = R\sin\theta\ \mathrm{d}\lambda \qquad [93]$$

3.02.4.4 Concluding Remarks

The spherical harmonic series, [59], represents the general solution to the exterior potential, regardless of the way the coefficients are determined. We know how to compute those coefficients exactly on the basis of a BVP, if the boundary is a sphere. More complicated boundaries would require corrections or, if these are omitted, would imply an approximation. If the coefficients are determined accurately (e.g., from satellite observations (Section 3.02.6.1), but not according to eqn [69]), then the spherical harmonic series model for the potential is not a spherical approximation. The spherical approximation enters when approximate relations such as eqns [5] are used and when the boundary is approximated as a sphere. However determined, the infinite series converges uniformly for all $r > R_c$, where R_c is the radius of the sphere that encloses all terrestrial masses. It may also converge below this sphere, but would represent the true potential only in free space (above the Earth's surface, where Laplace's equation holds). In practice, though, convergence is not an issue since the series must be truncated at some finite degree. Any trend toward divergence is then part of the overall model error, due mostly to truncation.

Model errors exist in all the Green's function integrals as they depend on spherical approximations in the boundary condition. In addition, the surface of integration in these formulas is formally assumed to be the geoid (where normal derivatives of the potential coincide with gravity magnitude), but it is approximated as a sphere. The spherical approximation results, in the first place, from a neglect of the ellipsoid flattening, which is about 0.3% for the Earth. When working with the disturbing potential, this level of error was easily tolerated in the past (e.g., if the geoid undulation is 30 m, the spherical approximation accounts for about 10 cm error), but today it requires attention as geoid undulation accuracy of 1 cm is pursued.

The Green's functions all have singularities when the evaluation point is on the sphere of radius, R. For points above this sphere, it is easily verified that all the Legendre series for the Green's functions converge uniformly, since $|P_n(\cos\psi)| \leq 1$. When $r = R$, the corresponding singularities of the integrals are either weak (Stokes integral) or strong (e.g., Poisson integral), requiring special definition of the integral as Cauchy principal value.

Finally, it is noted that the BVP solutions also require that no masses reside above the geoid (the boundary approximated as a sphere). To satisfy this condition, the topographic masses must be redistributed appropriately by mathematical reduction and the gravity anomalies, or disturbances measured on the Earth's surface must be reduced to the geoid (*see* Chapter 10.05). The mass redistribution must then be undone (mathematically) in order to obtain the correct potential on or above the geoid. Details of these procedures are found in Heiskanen and Moritz (1967, chapter 3). In addition, the atmosphere, having significant mass, affects gravity anomalies at different elevations. These effects may also be removed prior to using them as boundary values in the integral formulas.

3.02.5 Low-Degree Harmonics: Interpretation and Reference

The low-degree spherical harmonics of the Earth's gravitational potential lend themselves to interpretation with respect to the most elemental distribution of the Earth's density, which also leads to fundamental geometric characterizations, particularly for the second-degree harmonics. Let

$$C_{nm}^{(a)} = \frac{a}{GM}\left(\frac{R}{a}\right)^{n+1} v_{nm} \qquad [94]$$

be unitless coefficients that refer to a sphere of radius, a. (Recall that coefficients, v_{nm}, eqn [44], refer to a sphere of radius, R.) Relative to the central harmonic coefficient, $C_{00}^{(a)} = 1$, the next significant harmonic, $C_{20}^{(a)}$, is more than 3 orders of magnitude smaller; the remaining harmonic coefficients are at least 2–3 orders of magnitude smaller than that. The attenuation after degree 2 is much more gradual (**Table 1**), indicating that the bulk of the potential can be described by an ellipsoidal field. The normal gravitational field is such a field, but it also adheres to a geodetic definition that requires the underlying ellipsoid to be an equipotential surface in the corresponding normal gravity field. This section examines the low-degree harmonics from these two perspectives of interpretation and reference.

3.02.5.1 Low-Degree Harmonics as Density Moments

Returning to the general expression for the gravitational potential in terms of the Newtonian density integral (eqn [14]), and substituting the spherical harmonic series for the reciprocal distance (eqn [57] with [51]),

Table 1 Spherical harmonic coefficients of the total gravitational potential[a]

Degree (n)	Order, (m)	$C_{nm}^{(a)}$	$C_{n,-m}^{(a)}$
2	0	−4.841 70E − 04	0.0
2	1	−2.398 32E − 10	1.424 89E − 09
2	2	2.439 32E − 06	−1.400 28E − 06
3	0	9.571 89E − 07	0.0
3	1	2.030 48E − 06	2.481 72E − 07
3	2	9.048 02E − 07	−6.190 06E − 07
3	3	7.212 94E − 07	1.414 37E − 06
4	0	5.399 92E − 07	0.0
4	1	−5.361 67E − 07	−4.735 73E − 07
4	2	3.505 12E − 07	6.624 45E − 07
4	3	9.908 68E − 07	−2.009 76E − 07
4	4	−1.884 72E − 07	3.088 27E − 07

[a]GRACE model GGM02S (Tapley *et al.*, 2005).

$$V(\theta, \lambda, r) = G \iiint_{\text{volume}} \rho \frac{1}{r'} \sum_{n=0}^{\infty} \left(\frac{r'}{r}\right)^{n-1} \times \left(\frac{1}{2n+1} \sum_{m=-n}^{n} \bar{Y}_{nm}(\theta, \lambda) \bar{Y}_{nm}(\theta', \lambda')\right) \mathrm{d}v$$
$$= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{R}{r}\right)^{n+1} \left(\frac{G}{R^{n+1}(2n+1)} \times \iiint_{\text{volume}} \rho (r')^n \bar{Y}_{nm}(\theta', \lambda') \,\mathrm{d}v\right) \times \bar{Y}_{nm}(\theta, \lambda) \quad [95]$$

yields a multipole expansion (so called from electrostatics) of the potential. The spherical harmonic (Stokes) coefficients are multipoles of the density distribution (cf. eqn [44]),

$$v_{nm} = \frac{G}{R^{n+1}(2n+1)} \iiint_v \rho (r')^n \bar{Y}_{nm}(\theta', \lambda') \mathrm{d}v \quad [96]$$

One may also consider the nth-order moments of density (from the statistics of distributions) defined by

$$\mu_{\alpha,\beta,\gamma}^{(n)} = \iiint_v (x')^{\alpha} (y')^{\beta} (z')^{\gamma} \rho \,\mathrm{d}v, \quad n = \alpha + \beta + \gamma \quad [97]$$

The multipoles of degree n and the moments of order n are related, though not all $(n+1)(n+2)/2$ moments of order n can be determined from the 2n+1 multipoles of degree n, when $n \geq 2$ (clearly risking confusion, we defer to the common nomenclature of order for moments and degree for spherical harmonics). This indeterminacy is directly connected to the inability to determine the density distribution uniquely from external measurements of the potential (Chao, 2005), which is the classic geophysical inverse problem.

The zero-degree Stokes coefficient is coordinate invariant and is proportional to the total mass of the Earth:

$$v_{00} = \frac{G}{R} \iiint_{\text{volume}} \rho \,\mathrm{d}v = \frac{GM}{R} \quad [98]$$

It also represents a mass monopole, and it is proportional to the zeroth moment of the density, M.

The first-degree harmonic coefficients (representing dipoles) are proportional to the coordinates of the center of mass, (x_{cm}, y_{cm}, z_{cm}), which are proportional to the first-order moments of the density, as verified by recalling the definition of the first-degree spherical harmonics:

$$v_{1m} = \frac{G}{\sqrt{3}R^2} \iiint_{\text{volume}} \rho \begin{Bmatrix} r' \sin\theta' \sin\lambda', & m = -1 \\ r' \cos\theta', & m = 0 \\ r' \sin\theta' \cos\lambda', & m = 1 \end{Bmatrix} \mathrm{d}v$$
$$= \frac{GM}{\sqrt{3}R^2} \begin{Bmatrix} y_{cm}, & m = -1 \\ z_{cm}, & m = 0 \\ x_m, & m = 1 \end{Bmatrix} \quad [99]$$

Nowadays, by tracking satellites, we have access to the center of mass of the Earth (including its atmosphere), since it defines the center of their orbits. Ignoring the small motion of the center of mass (annual amplitude of several millimeters) due to the temporal variations in the mass distribution, we may choose the coordinate origin for the geopotential model to coincide with the center of mass, thus annihilating the first-degree coefficients.

The second-order density moments likewise are related to the second-degree harmonic coefficients (quadrupoles). They also define the inertia tensor of the body. The inertia tensor is the proportionality factor in the equation that relates the angular momentum vector, $\mathbf{H}$, and the angular velocity, $\boldsymbol{\omega}$, of a body, like the Earth:

$$\mathbf{H} = I\omega \quad [100]$$

and is given by

$$I = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \quad [101]$$

It comprises the moments of inertia on the diagonal

$$I_{xx} = \iiint_{\text{volume}} \rho(y'^2 + z'^2) \,\mathrm{d}v, \quad I_{yy} = \iiint_{\text{volume}} \rho(z'^2 + x'^2) \,\mathrm{d}v, \quad I_{zz} = \iiint_{\text{volume}} \rho(x'^2 + y'^2) \,\mathrm{d}v \quad [102]$$

and the products of inertia off the diagonal

$$I_{xy} = I_{yx} = -\iiint_{\text{volume}} \rho \; x'y' \,\mathrm{d}v, \quad I_{xz} = I_{zx} = -\iiint_{\text{volume}} \rho \; x'z' \,\mathrm{d}v, \quad I_{yz} = I_{zy} = -\iiint_{\text{volume}} \rho \; y'z' \,\mathrm{d}v \quad [103]$$

Note that

$$I_{xx} = \mu_{020}^{(2)} + \mu_{002}^{(2)}, \quad I_{xy} = -\mu_{110}^{(2)}, \quad \text{etc} \qquad [104]$$

and there are as many (six) independent tensor components as second-order density moments. Using the explicit expressions for the second-degree spherical harmonics, we have from eqn [96] with $n = 2$:

$$\begin{aligned} v_{2,-2} &= -\frac{\sqrt{15}\,G}{5R^3} I_{xy}, \quad v_{2,-1} = -\frac{\sqrt{15}\,G}{5R^3} I_{yz} \\ v_{2,1} &= -\frac{\sqrt{15}\,G}{5R^3} I_{xz} \quad v_{2,0} = \frac{\sqrt{5}\,G}{10R^3}\left(I_{xx} + I_{yy} - 2I_{zz}\right) \\ v_{2,2} &= \frac{\sqrt{15}\,G}{10R^3}\left(I_{yy} - I_{xx}\right) \end{aligned} \qquad [105]$$

These are also known as MacCullagh's formulas. Not all density moments (or, moments of inertia) can be determined from the Stokes coefficients.

If the coordinate axes are chosen so as to diagonalize the inertia tensor (products of inertia are then equal to zero), then they are known as principal axes of inertia, or also 'figure' axes. For the Earth, the z-figure axis is very close to the spin axis (within several meters at the pole); both axes move with respect to each other and the Earth's surface, with combinations of various periods (daily, monthly, annually, etc.), as well as secularly in a wandering fashion. Because of these motions, the figure axis is not useful as a coordinate axis that defines a frame fixed to the (surface of the) Earth. However, because of the proximity of the figure axis to the defined reference z-axis, the second-degree, first-order harmonic coefficients of the geopotential are relatively small (**Table 1**).

The arbitrary choice of the x-axis of our Earth-fixed reference coordinate system certainly did not attempt to eliminate the product of inertia, I_{xy}(the x-axis is defined by the intersection of the Greenwich meridian with the equator, and the y-axis completes a right-handed mutually orthogonal triad). However, it is possible to determine where the x-figure axis is located by combining values of the second-degree, second-order harmonic coefficients. Let u, v, w be the axes that define a coordinate system in which the inertia tensor is diagonal, and assume that $I_{ww} = I_{zz}$. A rotation by the angle, $-\lambda_0$, about the w- (also z-) figure axis brings this ideal coordinate system back to the conventional one in which we calculate the harmonic coefficients. Tensors transform under rotation, defined by matrix, $\mathcal{R}$, according to

$$I_{xyz} = \mathcal{R} I_{uvw} \mathcal{R}^{\mathrm{T}} \qquad [106]$$

With the rotation about the w-axis given by the matrix,

$$\mathcal{R} = \begin{pmatrix} \cos\lambda_0 & -\sin\lambda_0 & 0 \\ \sin\lambda_0 & \cos\lambda_0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad [107]$$

and with eqns [105], it is straightforward to show that

$$\begin{aligned} v_{2,-2} &= -\frac{\sqrt{15}G}{10R^3}\left(I_{uu} - I_{vv}\right)\sin 2\lambda_0 \\ v_{2,2} &= -\frac{\sqrt{15}G}{10R^3}\left(I_{uu} - I_{vv}\right)\cos 2\lambda_0 \end{aligned} \qquad [108]$$

Hence, we have

$$\lambda_0 = \frac{1}{2}\tan^{-1}\frac{v_{2,-2}}{v_{2,2}} \qquad [109]$$

where the quadrant is determined by the signs of the harmonic coefficients. From **Table 1**, we find that $\lambda_0 = -14.929°$; that is, the u-figure axis is in the mid-Atlantic between South America and Africa.

The second-degree, second-order harmonic coefficient, $v_{2,2}$, indicates the asymmetry of the Earth's mass distribution with respect to the equator. Since $v_{2,2} > 0$ (for the Earth), equations [108] show that $I_{vv} > I_{uu}$ and thus the equator 'bulges' more in the direction of the u-figure axis; conversely, the equator is flattened in the direction of the v-figure axis. This flattening is relatively small: 1.1×10^{-5}.

Finally, consider the most important second-degree harmonic coefficient, the second zonal harmonic, $v_{2,0}$. Irrespective of the x-axis definition, it is proportional to the difference between the moment of inertia, I_{zz}, and the average of the equatorial moments, $\left(I_{xx} + I_{yy}\right)/2$. Again, since $v_{2,0} < 0$, the Earth bulges more around the equator and is flattened at the poles. The second zonal harmonic coefficient is roughly 1000 times larger than the other second-degree coefficients and thus indicates a substantial polar flattening (owing to the Earth's early more-fluid state). This flattening is approximately 0.003.

3.02.5.2 Normal Ellipsoidal Field

Because of Earth's dominant polar flattening and the near symmetry of the equator, any meridional section of the Earth is closer to an ellipse than a circle. For this reason, ellipsoidal coordinates have often been advocated in place of the usual spherical coordinates. In fact, for geodetic positioning and geographic mapping, because of this flattening, the conventional (geodetic) latitude and longitude are coordinates that define the direction of the perpendicular to an ellipsoid, not the radial direction from the origin. These geodetic coordinates are not the ellipsoidal coordinates defined in Section 3.02.1 (eqn [2]) and would be rather useless in potential modeling because they do not separate Laplace' differential equation.

Harmonic series in terms of the ellipsoidal coordinates, (δ, λ, u), however, can be developed easily. They have not been adopted in most applications, perhaps in part because of the nonintuitive nature of the coordinates. Nevertheless, it is advantageous to model the normal (or reference) gravity field in terms of these ellipsoidal coordinates since it is based on an ellipsoid. Laplace' equation in these ellipsoidal coordinates can be separated, analogous to spherical coordinates, and the solution is obtained by successively solving three ordinary differential equations. Applied to the exterior gravitational potential, V, the solution is similar to the spherical harmonic series (eqn [44]) and is given by

$$V(\delta, \lambda, u) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{Q_{n|m|}(\mathrm{i}u/E)}{Q_{n|m|}(\mathrm{i}b/E)} v_{nm}^{\mathrm{e}} \bar{Y}_{nm}(\delta, \lambda) \quad [110]$$

where E is the linear eccentricity associated with the coordinate system and Q_{nm} is the associated Legendre function of the second kind. The coefficients of the series, v_{nm}^{e}, refer to an ellipsoid of semiminor axis, b; and, with the series written in this way, they are all real numbers with the same units as V. An exact relationship between these and the spherical harmonic coefficients, v_{nm}, was given by Hotine (1969) and Jekeli (1988).

With this formulation of the potential, Dirichlet's BVP is solved for an ellipsoidal boundary using the orthogonality of the spherical harmonics:

$$v_{nm}^{\mathrm{e}} = \frac{1}{4\pi} \iint_{\sigma} V(\delta, \lambda, b) \bar{Y}_{nm}(\delta, \lambda)\, \mathrm{d}\sigma \quad [111]$$

where $\mathrm{d}\sigma = \sin\delta \mathrm{d}\delta \mathrm{d}\lambda$ and $\sigma = \{(\delta, \lambda) | 0 \le \delta \le \pi, 0 \le \lambda \le 2\pi\}$. Note that while the limits of integration and the differential element, $d\sigma$, are the same as for the unit sphere, the boundary values are on the ellipsoid. Unfortunately, integral solutions with analytic forms of a Green's function do not exist in this case, because the inverse distance now depends on two surface coordinates and there is no corresponding convolution theorem. However, approximations have been formulated for all three types of BVPs (see Yu *et al.*, 2002, and references therein). Forms of ellipsoidal corrections to the classic spherical integrals have also been developed and applied in practice (e.g., Fei and Sideris, 2000).

The simplicity of the boundary values of the normal gravitational potential allows its extension into exterior space to be expressed in closed analytic form. Analogous to the geoid in the actual gravity field, the normal ellipsoid is defined to be a level surface in the normal gravity field. In other words, the sum of the normal gravitational potential and the centrifugal potential due to Earth's rotation is a constant on the ellipsoid:

$$V^{\mathrm{e}}(\delta, \lambda, b) + \phi(\delta, b) = U_0 \quad [112]$$

Hence, the normal gravitational potential on the ellipsoid, $V^e(\delta, \lambda, b)$, depends only on latitude and is symmetric with respect to the equator. Consequently, it consists of only even zonal harmonics, and because the centrifugal potential has only zero- and second-degree zonals, the corresponding ellipsoidal series is finite (up to degree 2). The solution to this Dirichlet problem is given in ellipsoidal coordinates by

$$V^{\mathrm{e}}(\delta, \lambda, u) = \frac{GM}{E} \tan^{-1}\frac{E}{u} + \frac{1}{2}\omega_{\mathrm{e}}^2 a^2 \frac{q}{q_0}\left(\cos^2\delta - \frac{1}{3}\right) \quad [113]$$

where a is the semimajor axis of the ellipsoid, ω_e is Earth's rate of rotation, and

$$q = \frac{1}{2}\left(\left(1 + 3\frac{u^2}{E^2}\right)\tan^{-1}\frac{E}{u} - 3\frac{u}{E}\right), \quad q_0 = q|_{u=b} \quad [114]$$

Heiskanen and Moritz (1967) and Hofmann-Wellenhof and Moritz (2005) provide details of the straightforward derivation of these and the following expressions.

The equivalent form of V^e in spherical harmonics is given by

$$V^{\mathrm{e}}(\theta, \lambda, r) = \frac{GM}{r}\left(1 - \sum_{n=1}^{\infty} J_{2n}\left(\frac{a}{r}\right)^{2n} P_{2n}(\cos\theta)\right) \quad [115]$$

where

$$J_{2n} = (-1)^{n+1}\frac{3e^{2n}}{(2n+1)(2n+3)}\left(1 - n + \frac{5n}{e^2}J_2\right), \quad n \ge 1 \quad [116]$$

and $e = E/a$ is the first eccentricity of the ellipsoid. The second zonal coefficient is given by

$$J_2 = \frac{e^2}{3}\left(1 - \frac{2}{15 q_0}\frac{\omega_e^2 a^2 E}{GM}\right) = \frac{I_{zz}^{\mathrm{e}} - I_{xx}^{\mathrm{e}}}{Ma^2} \quad [117]$$

where the second equality comes directly from the last of eqns [105] and the ellipsoid's rotational symmetry ($I^e_{xx} = I^e_{yy}$). This equation also provides a direct relationship between the geometry (the eccentricity or flattening) and the mass distribution (difference of second-order moments) of the ellipsoid. Therefore, J_2 is also known as the dynamic form factor – the flattening of the ellipsoid can be described either geometrically or dynamically in terms of a difference in density moments.

The normal gravitational potential depends solely on four adopted parameters: Earth's rotation rate, ω_e; the size and shape of the normal ellipsoid, for example, a, J_2; and a potential scale, for example, GM. The mean Earth ellipsoid is the normal ellipsoid with parameters closest to actual corresponding parameters for the Earth. GM and J_2 are determined by observing satellite orbits, a can be calculated by fitting the ellipsoid to mean sea level using satellite altimetry, and Earth's rotation rate comes from astronomical observations. **Table 2** gives current best values (Groten, 2004) and adopted constants for the Geodetic Reference Systems of 1967 and 1980 (GRS67, GRS80) and the World Geodetic System 1984 (WGS84).

In modeling the disturbing potential in terms of spherical harmonics, one naturally uses the form of the normal gravitational potential given by eqn [115]. Here, we have assumed that all harmonic coefficients refer to a sphere of radius, a. Corresponding coefficients for the series of $T = V - V^e$ are, therefore,

$$\delta C^{(a)}_{nm} = \begin{cases} 0, & n = 0 \\ C^{(a)}_{n0} - \dfrac{-J_n}{\sqrt{2n+1}}, & n = 2, 4, 6, \ldots \\ C^{(a)}_{nm}, & \text{otherwise} \end{cases} \qquad [118]$$

For coefficients referring to the sphere of radius, R, as in eqn [59], we have $\delta C_{nm} = (a/R)^n \delta C^{(a)}_{nm}$. The harmonic coefficients, J_{2n}, attenuate rapidly due to the factor, e^{2n}, and only harmonics up to degree 10 are significant. Normal gravity, being the gradient of the normal gravity potential, is used only in applications tied to an Earth-fixed coordinate system.

3.02.6 Methods of Determination

In this section, we briefly explore the basic technologies that yield measurements of the gravitational field. Even though we have reduced the problem of determining the exterior potential from a volume integral to a surface integral (e.g., either eqns [66] or [69]), it is clear that in theory we can never determine the entire field from a finite number of measurements. The integrals will always need to be approximated numerically, and/or the infinite series of spherical harmonics needs to be truncated. However, with enough effort and within the limits of computational capabilities, one can approach the ideal continuum of boundary values as closely as desired, or make the number of coefficients in the series representation as large as possible. The expended computational and measurement effort has to be balanced with the ability to account for inherent model errors (such as the spherical approximation) and the noise of the measuring device. To be useful for geodetic and geodynamic purposes, the instruments must possess a sensitivity of at least a few parts per million, and, in fact, many have a sensitivity of parts per billion. These sensitivities often come at the expense of prolonging the measurements (integration time) in order to average out random noise, thus reducing the achievable temporal resolution. This is particularly critical for moving-base instrumentation such as on an aircraft or satellite where temporal resolution translates into spatial resolution through the velocity of the vehicle.

Table 2 Defining parameters for normal ellipsoids of geodetic reference systems

Reference system	a (m)	J_2	GM ($m^3 s^{-2}$)	ω_e (rad s^{-2})
GRS67	6378160	1.0827E − 03	3.98603E14	7.2921151467E − 05
GRS80	6378137	1.08263E − 03	3.986005E14	7.292115E − 05
WGS84	6378137	1.08262982131E − 03	3.986004418E14	7.2921151467E − 05
Best current values[a]	6378136.7 ± 0.1 (mean-tide system)	(1.0826359 ± 0.0000001)E − 03 (zero-tide system)	(3.986004418 ± 0.000000008)E14 (includes atmosphere)	7.292115E − 5 (mean value)

[a]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) (The Geodesist's Handbook pp. 724–731).

3.02.6.1 Measurement Systems and Techniques

Determining the gravitational field through classical measurements relies on three fundamental laws. The first two are Newton's second law of motion and his law of gravitation. Newton's law of motion states that the rate of change of linear momentum of a particle is equal to the totality of forces, **F**, acting on it. Given more familiarly as $m_i \mathrm{d}^2\mathbf{x}/\mathrm{d}t^2 = \mathbf{F}$, it involves the inertial mass, m_i; and, conceptually, the forces, **F**, should be interpreted as action forces (like propulsion or friction). The gravitational field, which is part of the space we occupy, is due to the presence of masses like the Earth, Sun, Moon, and planets, and induces a different kind of force, the gravitational force. It is related to gravitational acceleration, **g**, through the gravitational mass, m_g, according to the law of gravitation, abbreviated here as $m_g\mathbf{g}=\mathbf{F}_g$ Newton's law of motion must be modified to include $\mathbf{F}_g$ separately. Through the third fundamental law, Einstein's equivalence principle, which states that inertial and gravitational masses are indistinguishable, we finally get

$$\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = \mathbf{a} + \mathbf{g} \qquad [119]$$

where **a** is the specific force ($\mathbf{F}/m_i$), or also the inertial acceleration, due to action forces. This equation holds in a nonrotating, freely falling frame (that is, an inertial frame), and variants of it can be derived in more complicated frames that rotate or have their own dynamic motion. However, one can always assume the existence of an inertial frame and proceed on that basis.

There exists a variety of devices that measure the motion of an inertial mass with respect to the frame of the device, and thus technically they sense **a**; such devices are called accelerometers. Consider the special case that an accelerometer is resting on the Earth's surface with its sensitive axis aligned along the vertical. In an Earth-centered frame (inertial, if we ignore Earth's rotation), the free-fall motion of the accelerometer is impeded by the reaction force of the Earth's surface acting on the accelerometer. In this case, the left-hand side of eqn [119] applied to the motion of the accelerometer is zero, and the accelerometer, sensing the reaction force, indirectly measures (the negative of) gravitational acceleration. This accelerometer is given the special name, gravimeter.

Gravimeters, especially static instruments, are designed to measure acceleration at very low frequencies (i.e., averaged over longer periods of time), whereas accelerometers typically are used in navigation or other motion-sensing applications, where accelerations change rapidly. As such, gravimeters generally are more accurate. Earth-fixed gravimeters actually measure gravity (Section 3.02.1.3), the difference between gravitation and centrifugal acceleration due to Earth's spin (the frame is not inertial in this case). The simplest, though not the first invented, gravimeter utilizes a vertically, freely falling mass, measuring the time it takes to fall a given distance. Applying eqn [119] to the falling mass in the frame of the device, one can solve for g (assuming it is constant): $x(t) = 0.5gt^2$. This free-fall gravimeter is a special case of a more general gravimeter that constrains the fall using an attached spring or the arm of a pendulum, where other (action) forces (the tension in the arm or the spring) thus enter into the equation.

The first gravimeter, in fact, was the pendulum, systematically used for gravimetry as early as the 1730s and 1740s by P. Bouguer on a geodetic expedition to measure the size and shape of the Earth (meridian arc measurement in Peru). The pendulum served well into the twentieth century (until the early 1970s) both as an absolute device, measuring the total gravity at a point, or as a relative device indicating the difference in gravity between two points (Torge, 1989). Today, absolute gravimeters exclusively rely on a freely falling mass, where exquisitely accurate measurements of distance and time are achieved with laser interferometers and atomic clocks (Zumberge *et al.*, 1982; Niebauer *et al.*, 1995). Accurate relative gravimeters are much less expensive, requiring a measurement of distance change only, and because many errors that cancel between measurements need not be addressed. They rely almost exclusively on a spring-suspended test mass (Nettleton, 1976; Torge, 1989). Developed early in the twentieth century in response to oil-exploration requirements, the relative gravimeter has changed little since then. Modern instruments include electronic recording capability, as well as specialized stabilization and damping for deployment on moving vehicles such as ships and aircraft. The accuracy of absolute gravimeters is typically of the order of parts per billion, and relative devices in field deployments may be as good but more typically are at least 1 order of magnitude less precise. Laboratory relative (in time) gravimeters, based on cryogenic instruments that monitor the virtual motion of a test mass that is electromagnetically suspended using superconducting persistent currents (Goodkind, 1999), are as accurate as portable absolute devices (or more), owing to the stability of the currents and the controlled laboratory environment (*see* Chapters 3.03 and 3.04).

On a moving vehicle, particularly an aircraft, the relative gravitational acceleration can be determined only from a combination of gravimeter and kinematic positioning system. The latter is needed to derive the (vertical) kinematic acceleration (the left-hand side of eqn [119]). Today, the GPS best serves that function, yielding centimeter-level precision in relative three-dimensional position. Such combined GPS/gravimeter systems have been used successfully to determine the vertical component of gravitation over large, otherwise-inaccessible areas such as the Arctic Ocean (Kenyon and Forsberg, 2000) and over other areas that are more economically surveyed from the air for oil-exploration purposes (Hammer, 1982; Gumert, 1998). The airborne gravimeter is specially designed to damp out high-frequency noise and is usually stabilized on a level platform. Three-dimensional moving-base gravimetry has also been demonstrated using the triad of accelerometers of a high-accuracy inertial navigation system (the type that are fixed to the aircraft without special stabilizing platforms). The orientation of all accelerometers on the vehicle must be known with respect to inertial space, which is accomplished with precision gyroscopes. Again, the total acceleration vector of the vehicle, d^2x/dt^2, can be ascertained by time differentiation of the kinematic positions (from GPS). One of the most critical errors is due to the cross-coupling of the horizontal orientation error, $\delta\psi$, with the large vertical acceleration (the lift of the aircraft, essentially equal to $-g$). This is a first-order effect ($g \sin \delta\psi$) in the estimation of the horizontal gravitation components, but only a second-order effect ($g(1-\cos \delta\psi)$) on the vertical component. Details of moving-base vector gravimetry may be found in Jekeli (2000a, chapter 10) and Kwon and Jekeli (2001).

The ultimate global gravimeter is the satellite in free fall (i.e., in orbit due to sufficient forward velocity) – the satellite is the inertial mass and the 'device' is a set of reference points with known coordinates (e.g., on the Earth's surface, or another satellite whose orbit is known; *see* Chapter 3.05). The measuring technology is an accurate ranging system (radar or laser) that tracks the satellite as it orbits (falls to) the Earth. Ever since *Sputnik*, the first artificial satellite, launched into the Earth orbit in 1957, Earth's gravitational field could be determined by tracking satellites from precisely known ground stations. Equation [119], with gravitational acceleration expressed as a truncated series of spherical harmonics (gradient of eqn [44]), becomes

$$\frac{d^2x}{dt^2} = \sum_{n=0}^{n_{\max}} \sum_{m=-n}^{n} v_{nm} \nabla \left(\left(\frac{R}{r}\right)^{n+1} \bar{Y}_{nm}(\theta, \lambda + \omega_e t) \right) + \delta \mathbf{R} \qquad [120]$$

where ω_e is Earth's rate of rotation and $\delta\mathbf{R}$ represents residual accelerations due to action forces (solar radiation pressure, atmospheric drag, Earth's albedo, etc.), gravitational tidal accelerations due to other bodies (Moon, Sun, planets), and all other subsequent indirect effects. The x left-hand side of eqn [120] is more explicitly $\mathbf{x}(\mathbf{t}) = \mathbf{x}(\theta(\mathbf{t}), \lambda(\mathbf{t}) + \omega_\mathbf{e}\mathbf{t}, \mathbf{r}(\mathbf{t}))$, and the spatial coordinates on the right-hand side are also functions of time. This makes the equation more conceptual than practical since it is numerically more convenient to transform the satellite position and velocity into Keplerian orbital elements (semi-major axis of the orbital ellipse, its eccentricity, its inclination to the equator, the angle of perigee, the right ascension of the node of the orbit, and the mean motion) all of which also change in time, but most much more slowly. This transformation was derived by Kaula (1966) (see also Seeber, 1993).

In the most general case ($n_{\max} > 2$ and $\delta\mathbf{R} \neq 0$), there is no analytic solution to eqn [120] or its transformations to other types of coordinates. The positions of the satellite are observed by ranging techniques and the unknowns to be solved are the coefficients, v_{nm}. Numerical integration algorithms have been specifically adapted to this problem and extremely sophisticated models for $\delta\mathbf{R}$ are employed with additional unknown parameters to be solved in order to estimate as accurately as possible the gravitational coefficients (e.g., Cappelari *et al.*, 1976; Pavlis *et al.*, 1999). The entire procedure falls under the broad category of dynamic orbit determination, and the corresponding gravitational field modeling may be classified as the 'timewise' approach. The partial derivatives of eqn [120] with respect to unknown parameters, $\mathbf{p} = \{\ldots, v_{nm}, \ldots\}$, are integrated numerically in time, yielding estimates for $H = \partial\mathbf{x}/\partial\mathbf{p} = \{\ldots, \partial\mathbf{x}/\partial v_{nm}, \ldots\}$. These are then used in a least-squares adjustment of the linearized model relating observed positions (e.g., via ranges) to parameters

$$\delta \mathbf{x} = H\delta \mathbf{p} + \mathbf{e} \quad [121]$$

where $\delta\mathbf{x}$ and $\delta\mathbf{p}$ are differences with respect to previous estimates, and $\mathbf{e}$ represents errors. (Tapley, 1973).

A gravimeter (or accelerometer) on a satellite does not sense the presence of a gravitational field. This is evident from the fact that the satellite is in free fall (apart from small accelerations due to action forces such as atmospheric drag) and the inertial test mass and the gravimeter, itself, are equally affected by gravitation (i.e., they are all in free fall). However, two accelerometers fixed on a satellite yield, through the difference in their outputs, a gradient in acceleration that includes the gradient of gravitation. On a nonrotating satellite, the acceleration at an arbitrary point, $\mathbf{b}$, of the satellite is given by

$$\mathbf{a}_b = \frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} - \mathbf{g}(\mathbf{b}) \quad [122]$$

in a coordinate system with origin at the center of mass of the satellite. Taking the difference (differential) of two accelerations in ratio to their separation, we obtain

$$\frac{\delta \mathbf{a}_b}{\delta \mathbf{b}} = -\frac{\delta \mathbf{g}}{\delta \mathbf{b}} \quad [123]$$

where the ratios represent tensors of derivatives in the local satellite coordinate frame. For a rotating satellite, this equation generalizes to

$$\frac{\partial \mathbf{a}_b}{\partial \mathbf{b}} = -\frac{\partial \mathbf{g}_b}{\partial \mathbf{b}} + \Omega^2 + \frac{\mathrm{d}}{\mathrm{d}t}\Omega \quad [124]$$

where Ω is a skew-symmetric matrix whose off-diagonal elements are the components of the vector that defines the rotation rate of the satellite with respect to the inertial frame. Thus, a gradiometer on a satellite (or any moving vehicle) senses a combination of gravitational gradient and angular acceleration (including a centrifugal type). Such a device is scheduled to launch for the first time in 2007 as part of the mission *GOCE* (*Gravity Field and Steady-State Ocean Circulation Explorer*; Rummel *et al.*, 2002). If the entire tensor of gradients is measured, then, because of the symmetry of the gravitational gradient tensor and of Ω^2, and the antisymmetry of $\mathrm{d}\Omega/\mathrm{d}t$, the sum $\delta\mathbf{a}_b/\delta\mathbf{b} + (\delta\mathbf{a}_b/\delta\mathbf{b})^{\mathrm{T}}$ eliminates the latter, while the difference $\delta\mathbf{a}_b/\delta\mathbf{b} - (\delta\mathbf{a}_b/\delta\mathbf{b})^{\mathrm{T}}$ can be used to infer Ω, subject to initial conditions.

When the two ends of the gradiometer are fixed to one frame, the common linear acceleration, $\mathrm{d}^2\mathbf{x}/\mathrm{d}t^2$, cancels, as shown above; but if the two ends are independent, disconnected platforms moving in similar orbits, the gravitational difference depends also on their relative motion in inertial space. This is the concept for satellite-to-satellite tracking, by which one satellite precisely tracks the other and the change in the range rate between them is a consequence of a gravitational difference, a difference in action forces, and a centrifugal acceleration due to the rotation of the baseline of the satellite pair. It can be shown that the line-of-sight acceleration (the measurement) is given by

$$\frac{\mathrm{d}^2\rho}{\mathrm{d}t^2} = \mathbf{e}_\rho^{\mathrm{T}}(\mathbf{g}(\mathbf{x}_2) - \mathbf{g}(\mathbf{x}_1)) + \mathbf{e}_\rho^{\mathrm{T}}(\mathbf{a}_2 - \mathbf{a}_1) + \frac{1}{\rho}\left(\left|\frac{\mathrm{d}}{\mathrm{d}t}\Delta\mathbf{x}\right|^2 - \left(\frac{\mathrm{d}\rho}{\mathrm{d}t}\right)^2\right) \quad [125]$$

where $\mathbf{e}_\rho$ is the unit vector along the instantaneous baseline connecting the two satellites, ρ is the baseline length (the range), and $\Delta\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_1$ is the difference in position vectors of the two satellites. Clearly, only the gravitational difference projected along the baseline can be determined (similar to a single-axis gradiometer), and then only if both satellites carry accelerometers that sense the nongravitational accelerations. Also, the orbits of both satellites need to be known in order to account for the centrifugal term.

Two such satellite systems were launched recently to determine the global gravitational field. One is *CHAMP* (*Challenging Mini-Satellite Payload*) in 2000 (Reigber *et al.*, 2002) and the other is *GRACE* (*Gravity Recovery and Climate Experiment*) in 2002 (Tapley *et al.*, 2004a). *CHAMP* is a single low-orbiting satellite (400–450 km altitude) being tracked by the high-altitude GPS satellites, and it also carries a magnetometer to map the Earth's magnetic field. *GRACE* was more specifically dedicated to determining with extremely high accuracy the long to medium wavelengths of the gravitational field and their temporal variations. With two satellites in virtually identical low Earth orbits, one following the other, the primary data consist of intersatellite ranges observed with K-band radar. The objective is to sense changes in the gravitational field due to mass transfer on the Earth within and among the atmosphere, the hydrosphere/cryosphere, and the oceans (Tapley *et al.*, 2004b).

An Earth-orbiting satellite is the ideal platform on which to measure the gravitational field when seeking global coverage in relatively short time. One simply designs the orbit to be polar and circular;

and, as the satellite orbits, the Earth spins underneath, offering a different section of its surface on each satellite revolution. There are also limitations. First, the satellite must have low altitude to achieve high sensitivity, since the nth-degree harmonics of the field attenuate as $(R/r)^{n+1}$. On the other hand, the lower the altitude, the shorter the life of the satellite due to atmospheric drag, which can only be countered with onboard propulsion systems. Second, because of the inherent speed of lower-orbit satellites (about $7\,\mathrm{km\,s^{-1}}$), the resolution of its measurements is limited by the integration (averaging) time of the sensor (typically 1–10 s). Higher resolution comes only with shorter integration time, which may reduce the accuracy if this depends on averaging out random noise. **Figure 8** shows the corresponding achievable resolution on the Earth's surface for different satellite instrumentation parameters, length of time in polar orbit and along-orbit integration time, or smoothing (Jekeli, 2004). In each case, the indicated level of resolution is warranted only if the noise of the sensor (after smoothing) does not overpower the signal at this resolution.

Both *CHAMP* and *GRACE* have yielded global gravitational models by utilizing traditional satellite-tracking methods and incorporating the range rate appropriately as a tracking observation (timewise approach). However, the immediate application of eqn [125] suggests that gravitational differences can be determined *in situ* and used to determine a model for the global field directly. This is classified as the spacewise approach. In fact, if the orbits are known with sufficient accuracy (from kinematic orbit determination, e.g., by GPS), this procedure utilizes a

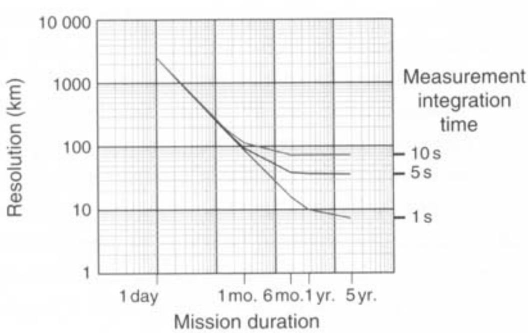

Figure 8 Spatial resolution of satellite measurements vs mission duration and integration time. The satellite altitude is 450 km (after Jekeli, 2004) mo., month. Reproduced by permission of American Geophysical Union.

linear relationship between observations and unknown harmonic coefficients:

$$\left.\frac{\mathrm{d}^2\rho}{\mathrm{d}t^2}\right|_{x_1,x_2} = \sum_{n=0}^{n_{\max}}\sum_{m=-n}^{n} v_{nm}\mathbf{e}_\rho^{\mathrm{T}}\left(\nabla U_{nm}(\theta,\lambda,r)|_{x_2} - \nabla U_{nm}(\theta,\lambda,r)|_{x_1}\right) + \delta a + \delta c \qquad [126]$$

where $U_{nm}(\theta, \lambda, r) = (R/r)^{n+1}\bar{Y}_{nm}(\theta, \lambda)$ and δa, δc are the last two terms in eqn [125]. Given the latter and a set of line-of-sight accelerations, a theoretically straightforward linear least-squares adjustment solves for the coefficients. A similar procedure can be used for gradients observed on an orbiting satellite:

$$\left.\frac{\partial \mathbf{a}}{\partial \mathbf{b}}\right|_{x} = \sum_{n=0}^{n_{\max}}\sum_{m=-n}^{n} v_{nm}\nabla\nabla^{\mathrm{T}} U_{nm}(\theta,\lambda,r)\big|_{x} + \psi \qquad [127]$$

where ψ comprises the rotational acceleration terms in eqn [124].

In situ measurements of line-of-site acceleration or of more local gradients would need to be reduced from the satellite orbit to a well-defined surface (such as a sphere) in order to serve as boundary values in a solution to a BVP. However, the model for the field is already in place, rooted in potential theory (truncated series solution to Laplace' equation), and one may think of the problem more in terms of fitting a three-dimensional model to a discrete set of observations. This operational approach can readily, at least conceptually, be expanded to include observations from many different satellite systems, even airborne and ground-based observations.

Recently, a rather different theory has been considered by several investigators to model the gravitational field from satellite-to-satellite tracking observations. The method, first proposed by Wolff (1969), makes use of yet another fundamental law: the law of conservation of energy. Simply, the range rate between two satellites implies an along-track velocity difference, or a difference in kinetic energy. Observing this difference leads directly, by the conservation law, to the difference in potential energy, that is, the gravitational potential and other potential energies associated with action forces and Earth's rotation. Neglecting the latter two, conservation of energy implies

$$V = \frac{1}{2}\left|\frac{\mathrm{d}}{\mathrm{d}t}x_1\right|^2 - E_0 \qquad [128]$$

where E_0 is a constant. Taking the along-track differential, we have approximately

$$V(\mathbf{x}_2) - V(\mathbf{x}_1) = \left|\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}_1\right| \frac{\mathrm{d}}{\mathrm{d}t}\rho \qquad [129]$$

where $|\mathrm{d}x_2/\mathrm{d}t - \mathrm{d}x_1/\mathrm{d}t| \approx \mathrm{d}\rho/\mathrm{d}t$. This very rough conceptual relationship between the potential difference and the range rate applies to two satellites closely following each other in similar orbits. The precise formulation is given by (Jekeli, 1999) and holds for any pair of satellites, not just two low orbiters. Mapping range rates between two polar orbiting satellites (such as *GRACE*) yields a global distribution of potential difference observations related again linearly to a set of harmonic coefficients:

$$V(\mathbf{x}_2) - V(\mathbf{x}_1) = \sum_{n=0}^{n_{\max}} \sum_{m=-n}^{n} v_{nm} \left(U_{nm}(\theta, \lambda, r)\Big|_{x_2} - U_{nm}(\theta, \lambda, r)\Big|_{x_1} \right) \qquad [130]$$

The energy-based model holds for any two vehicles in motion and equipped with the appropriate ranging and accelerometry instrumentation. For example, the energy conservation principle could also be used to determine geopotential differences between an aircraft and a satellite, such as a GPS satellite (at which location the geopotential is known quite well). The aircraft would require only a GPS receiver (to do the ranging) and a set of accelerometers (and gyros for orientation) to measure the action forces (the same system components as in airborne accelerometry discussed above). Resulting potential differences could be used directly to model the geoid using Bruns' equation.

3.02.6.2 Models

We have already noted the standard solution options to the BVP using terrestrial gravimetry in the form of gravity anomalies: the Green's function approach, Stokes integral, eqn [66], and the harmonic analysis of surface data, either using the integrals [69] or solving a linear system of equations (eqn [65] truncated to finite degree) to obtain the coefficients, δC_{nm}. The integrals must be evaluated using quadratures, and very fast numerical techniques have been developed when the data occupy a regular grid of coordinates on the sphere or ellipsoid (Rapp and Pavlis, 1990). Similar algorithms enable the fast solution of the linear system of eqns [65].

For a global harmonic analysis, the number of coefficients, $(n_{\max} + 1)^2$, must not be greater than the number of data. A general, conservative rule of thumb for the maximum resolution (half-wavelength) of a truncated spherical harmonic series is, in angular degrees on the unit sphere,

$$\Delta\theta = \frac{180^\circ}{n_{\max}}$$

Thus, data on a $1^\circ \times 1^\circ$ angular grid of latitudes and longitudes would imply $n_{\max} = 180$. The number of data (64 800) is amply larger than the number of coefficients (32 761). This majority suggests a least-squares adjustment of the coefficients to the data, in either method, especially because the data have errors (Rapp, 1969). As $n_{\max}$ increases, a rigorous, optimal adjustment usually is feasible, for a given computational capability, only under restrictive assumptions on the correlations among the errors of the data. Also, the obvious should nevertheless be noted that the accuracy of the model in any area depends on the quality of the data in that area. Furthermore, considering that a measurement contains all harmonics (up to the level of measurement error), the estimation of a finite number of harmonics from boundary data on a given grid is corrupted by those harmonics that are in the data but are not estimated. This phenomenon is called aliasing in spectral analysis and can be mitigated by appropriate filtering of the data (Jekeli, 1996).

The optimal spherical harmonic model combines both satellite and terrestrial data. The currently best known model is EGM96 complete to degree and order $n_{\max} = 360$ (Lemoine *et al.*, 1998). It is an updated model for WGS84 based on all available satellite tracking, satellite altimetry, and land gravity (and topographic) data up to the mid-1990s. Scheduled to be revised again for 2006 using more recent data, as well as results from the satellite missions, *CHAMP* and *GRACE*, it will boast a maximum degree and order of 2160 (5' resolution). In constructing combination solutions of this type, great effort is expended to ensure the proper weighting of observations in order to extract the most appropriate information from the diverse data, pertaining to different parts of the spatial gravitational spectrum. The satellite tracking data dominate the estimation of the lower-degree harmonics, whereas the fine resolution of the terrestrial data is most amenable to modeling the higher degrees. It is beyond the present scope to delve into the numerical methodology of combination methods. Furthermore, it is an unfinished story as new *in situ* satellite measurements from *GRACE* and *GOCE* will affect the combination methods of the future.

Stokes integral is used in practice only to take advantage of local or regional data with higher resolution than was used to construct the global models. Even though the integral is a global integral, it can be truncated to a neighborhood of the computation point since the Stokes kernel attenuates rapidly as the reciprocal distance. Moreover, the corresponding truncation error may be reduced if the boundary values exclude the longer-wavelength features of the field. The latter constitute an adequate representation of the remote zone contribution and can be included separately as follows. Let $\Delta g^{(n_{\max})}$ denote the gravity anomaly implied by a spherical harmonic model, such as given by eqn [65], truncated to degree, $n_{\max}$. From the orthogonality of spherical harmonics, eqn [46], it is easy to show that

$$T^{(n_{\max})}(\theta, \lambda, r) = \frac{GM}{R}\sum_{n=2}^{n_{\max}}\sum_{m=-n}^{n}\left(\frac{R}{r}\right)^{n+1}\delta C_{nm}\bar{Y}_{nm}(\theta, \lambda) = \frac{R}{4\pi}\iint_{\sigma}\Delta g^{(n_{\max})}(\theta', \lambda', R)S(\psi, r)\mathrm{d}\sigma \qquad [131]$$

Thus, given a spherical harmonic model $\{\delta C_{nm}|2 \leq n \leq n_{\max}, -n \leq m \leq n\}$, we first remove the model in terms of the gravity anomaly and then restore it in terms of the disturbing potential, changing Stokes formula [66] to

$$T(\theta, \lambda, r) = \frac{R}{4\pi}\iint_{\sigma}\left(\Delta g(\theta', \lambda', R) - \Delta g^{(n_{\max})}(\theta', \lambda', R)\right) \times S(\psi, r)\mathrm{d}\sigma + T^{(n_{\max})}(\theta, \lambda, r) \qquad [132]$$

In theory, if $\Delta g^{(n_{\max})}$ has no errors, then the residual $\Delta g - \Delta g^{(n_{\max})}$ excludes all harmonics of degree $n \leq n_{\max}$, and orthogonality would also allow the exclusion of these harmonics from S. Once the integration is limited to a neighborhood of (θ, λ, R), as it must be in practice, there are a number of ways to modify the kernel so as to minimize the resulting truncation error (Sjöberg, 1991, and references therein). The removal and restoration of a global model, however, is the key aspect in all these methods.

In practical applications, the boundary values are on the geoid, being the surface that satisfies the boundary condition of the Robin BVP (i.e., we require the normal derivative on the boundary and measured gravity is indeed the derivative of the potential along the perpendicular to the geoid). The integral in eqn [132] thus approximates the geoid by a sphere. Furthermore, it is assumed that no masses exist external to the geoid. Part of the reduction to the geoid of data measured on the Earth's surface involves redistributing the topographic masses on or below the geoid. This redistribution is undone outside the solution to the BVP (i.e., Stokes integral) in order to regain the disturbing potential for the actual Earth. Conceptually, we may write

$$T_{\mathrm{P}} = \frac{R}{4\pi}\iint_{\sigma}\left(\Delta g - \Delta g^{(n_{\max})} - \delta c\right)_{P'} S_{P,P'}\mathrm{d}\sigma + T_{\mathrm{P}}^{(n_{\max})} + \delta T_{\mathrm{P}} \qquad [133]$$

where δc is the gravity reduction that brings the gravity anomaly to a geoid with no external masses, and δT_P is the effect (called indirect effect) on the disturbing potential due to the inverse of this reduction. This formula holds for T anywhere on or above the geoid and thus can also be used to determine the geoid undulation according to Bruns' formula [3].

3.02.7 The Geoid and Heights

The traditional reference surface, or datum, for heights is the geoid (Section 3.02.1.3). A point at mean sea level usually serves as starting point (datum origin) and this defines the datum for vertical control over a region or country. The datum (or geoid) is the level continuation of the reference surface under the continents, and the determination of gravity potential differences from the initial point to other points on the Earth's suface, obtained by leveling and gravity measurements, yields heights with respect to that reference (or in that datum). The gravity potential difference, known as the geopotential number, at a point, P, relative to the datum origin, $\bar{P}_0$, is given by (since gravity is the negative vertical derivative of the gravity potential)

$$C_P = W_0 - W_P = \int_{\bar{P}_0}^{P} g\,\mathrm{d}n \qquad [134]$$

where g is gravity (magnitude), dn is a leveling increment along the vertical direction, and W_0 is the gravity potential at $\bar{P}_0$. By the conservative nature of the gravity potential, whatever path is taken for the integral yields a unique geopotential number for P. From these potential differences, one can define various types of height, for example, the orthometric height (**Figure 3**):

$$H_P = \frac{C_P}{\bar{g}_P} \quad [135]$$

where

$$\bar{g}_P = \frac{1}{H_P} \int_{P_0}^{P} g \, \mathrm{d}H \quad [136]$$

is the average value of gravity along the plumb line from the geoid at P_0 to P. Other height systems are also in use (such as the normal and dynamic heights), but they all rely on the geopotential number (see Heiskanen and Moritz, 1967, chapter 4, for details).

For a particular height datum, there is theoretically only one datum surface (the geoid). But access to this surface is far from straightforward (other than at the defined datum origin). If $\bar{P}_0$ is defined at mean sea level, other points at mean sea level are not on the same level (datum) surface, since mean sea level, in fact, is not level. Erroneously assuming that mean sea level is an equipotential surface can cause significant distortions in the vertical control network of larger regions, as much as several decimeters. This was the case, for example, for the National Geodetic Vertical Datum of 1929 in the US for which 26 mean sea level points on the east and west coasts were assumed to lie on the same level surface. Accessibility to the geoid (once defined) at any point is achieved either with precise leveling and gravity (according to eqns [134] and [135]), or with precise geometric vertical positioning and knowledge of the gravity potential. Geometric vertical positioning, today, is obtained very accurately (centimeter accuracy or even better) with differential GPS. Suppose that an accurate gravity potential model is also available in the same coordinate system as used for GPS. Then, determining the GPS position at $\bar{P}_0$ allows the evaluation of the gravity potential, W_0, of the datum. Access to the geoid at any other point, P, or equivalently, determining the orthometric height, H_P, can be done by first determining the ellipsoidal height, h, from GPS. Then, as shown in **Figure 3**,

$$H_P = h_P - N \quad [137]$$

where, with $T = W - U$ evaluated on the geoid, Bruns' extended equation (Section 3.02.1) yields

$$N = T/\gamma - (W_0 - U_0)/\gamma \quad [138]$$

where U_0 is the normal gravity potential of the normal ellipsoid.

In a sense, the latter is a circular problem: determining N requires N in order to locate the point on the geoid where to compute T. However, the computation of T on the geoid can be done with assumptions on the density of the topographic masses above the geoid and a proper reduction to the geoid, using only an approximate height. Indeed, since the vertical gradient of the disturbing potential is the gravity disturbance, of the order of $5 \times 10^{-4}\,\mathrm{m\,s^{-2}}$, a height error of 20 m leads to an error of $10^{-2}\,\mathrm{m^2\,s^{-2}}$ in T, or just 1 mm in the geoid undulation. It should be noted that a model for the disturbing potential as a series of spherical harmonics, for example, derived from satellite observations, satisfies Laplace's equation and, therefore, does not give the correct disturbing potential at the geoid (if it lies below the Earth's surface where Laplace's equation does not hold).

The ability to derive orthometric heights (or other geopotential-related heights) from GPS has great economical advantage over the laborious leveling procedure. This has put great emphasis on obtaining an accurate geoid undulation model for land areas. Section 3.02.6 briefly outlined the essential methods to determine the geoid undulation from a combination of spherical harmonics and an integral of local gravity anomalies. When dealing with a height datum or the geoid, the constant $N_0 = -(W_0 - U_0)/\gamma$ requires careful attention. It can be determined by comparing the geoid undulation computed according to a model that excludes this term (such as eqn [133]) with at least one geoid undulation (usually many) determined from leveling and GPS, according to eqn [137]. Vertical control and the choice of height datum are specific to each country or continent, where a local mean sea level was the adopted datum origin. Thus, height datums around the world are 'local geoids' that have significant differences between them. Investigations and efforts have been under way for more than two decades to define a global vertical datum; however, it is still in the future, awaiting a more accurate global gravity potential model and, perhaps more crucially, a consensus on what level surface the global geoid should be.

On the oceans, the situation is somewhat less complicated. Oceanographers who compute sea-surface topography from satellite altimetry on the basis of eqn [137] depend critically on an accurate geoid undulation, or equivalently on an accurate model of T. However, no reduction of the disturbing potential from mean sea level to the geoid is necessary, the deviation being at most 2 m and causing an error in geoid undulation of less than 0.1 mm. Thus, a spherical harmonic model of T is entirely appropriate. Furthermore, it is reasonable to ensure that the constant, $W_0 - U_0$, vanishes over the oceans. That is, one may choose the geoid such that it best fits mean sea level and choose an ellipsoid that best fits this geoid. It means that the global average value of the geoid undulation should be zero (according to eqn [138]). The latter can be achieved with satellite altimetry and oceanographic models of sea surface topography (Bursa *et al.*, 1997).

Several interesting and important distinctions should be made in regard to the tidal effects on the geoid. The Sun and the Moon generate an appreciable gravitational potential near the Earth (the other planets may be neglected). In an Earth-fixed coordinate system, this extraterrestrial potential varies in time with different periods due to the relative motions of the Moon and Sun, and because of Earth's rotation (Torge, 2001, p. 88). There is also a constant part, the permanent tidal potential, representing the average over time. It is not zero, because the Earth–Sun–Moon system is approximately coplanar. For each body with mass, M_B, and distance, r_B, from the Earth's center, this permanent part is given by

$$V_c^B(\theta, r) = \frac{3}{4} GM_B \frac{r^2}{r_B^2}\left(3\cos^2\theta - 1\right)\left(\frac{1}{2}\sin^2\varepsilon - \frac{1}{3}\right) \qquad [139]$$

where ε is the angle of the ecliptic relative to the equator ($\varepsilon \approx 23°\,44'$). Using nominal parameter values for the Sun and the Moon, we obtain at mean Earth radius, $R = 6371\,\text{km}$,

$$V_c^{s+m}(\theta, R) = -0.97\left(3\cos^2\theta - 1\right)\text{m}^2\,\text{s}^{-2} \qquad [140]$$

The gravitational potential from the Sun and Moon also deforms the quasi-elastic Earth's masses with the same periods and similarly includes a constant part. These mass displacements (both ocean and solid Earth) give rise to an additional indirect change in potential, the tidal deformation potential (there are also secondary indirect effects due to loading of the ocean on the solid Earth, which can be neglected in this discussion; *see* Chapter 3.06). The indirect effect is modeled as a fraction of the direct effect (Lambeck, 1988, p. 254), so that the permanent part of the tidal potential including the indirect effect is given by

$$\bar{V}_c^{s+m}(\theta, R) = (1 + k_2)V_c^{s+m}(\theta, R) \qquad [141]$$

where $k_2 = 0.29$ is Love's number (an empirical number based on observation). This is also called the mean tide potential.

The mean tidal potential is inherent in all our terrestrial observations (the boundary values) and cannot be averaged away; yet, the solutions to the BVP assume no external masses. Therefore, in principle, the effect of the tide potential including its mean, or permanent, part should be removed from the observations prior to applying the BVP solutions. On the other hand, the permanent indirect effect is not that well modeled and arguably should not be removed; after all, it contributes to the Earth's shape as it actually is in the mean. Three types of tidal systems have been defined to distinguish between these corrections. A mean quantity refers to the quantity with the mean tide potential retained (but time-varying parts removed); a nontidal quantity implies that all tidal effects (time-varying, permanent, direct and indirect effects) have been removed computationally; and the zero-tide quantity excludes all time-varying parts and the permanent direct effect, but it retains the indirect permanent effect.

If the geoid (an equipotential surface) is defined solely by its potential, W_0, then a change in the potential due to the tidal potential, V^{tide} (time-varying and constant parts, direct and indirect effects), implies that the W_0-equipotential surface has been displaced. The geoid is now a different surface with the same W_0. This displacement is equivalent to a change in geoid undulation, $\delta N = V^{\text{tide}}/\gamma$, with respect to some predefined ellipsoid. The permanent tidal effect (direct and indirect) on the geoid is given by

$$\delta\bar{N}(\theta) = -0.099(1 + k_2)\left(3\cos^2\theta - 1\right)\text{m} \qquad [142]$$

If N represents the instantaneous geoid, then the geoid without any tidal effects, that is, the nontidal geoid, is given by

$$N_{nt} = N - \delta N \qquad [143]$$

The mean geoid is defined as the geoid with all but the mean tidal effects removed:

$$\bar{N} = N - (\delta N - \delta\bar{N}) \qquad [144]$$

This is the geoid that could be directly observed, for example, using satellite altimetry averaged over time. The zero-tide geoid retains the permanent indirect effect, but no other tidal effects:

$$N_z = N - \left(\delta N + 0.099k_2\left(3\cos^2\theta - 1\right)\right)\text{m} \qquad [145]$$

The difference between the mean and zero-tide geoids is, therefore, the permanent component of the direct tidal potential. We note that, in principle, each of the geoids defined above, has the same potential value, W_0, in its own field. That is, with each correction, we define a new gravity field and the corresponding geoid undulation defines the equipotential surface in that field with potential value given by W_0. This is fundamentally different than what happens in the case when the geoid is defined as a vertical datum with a specified datum origin point. In this case one needs to consider also the vertical displacement of the datum point due to the tidal deformation of the Earth's surface. The potential of the datum then changes because of the direct tidal potential, the indirect effect due to mass changes, and the indirect effect due to the vertical displacement of the datum (for additional details, see Jekeli (2000b)).

References

Bursa M, Radej K, Sima Z, True S, and Vatrt V (1997) Determination of the geopotential scale factor from TOPEX/Poseidon satellite altimetry. *Studia Geophysica et Geodaetica* 41: 203–216.

Cappelari JO, Velez CE, and Fuchs AJ (eds.) (1976) *Mathematical theory of the Goddard Trajectory Determination System*. GSFC Document X-582-76-77. Greenbelt MD: Goddard Space Flight Center.

Chao BF (2005) On inversion for mass distribution form global (time-variable) gravity field. *Journal of Geodynamics* 29: 223–230.

Courant R and Hilbert D (1962) *Methods of Mathematical Physics, vol. II. New York: Wiley.*

Cushing JT (1975) *Applied Analytical Mathematics for Physical Scientists*. New York: Wiley.

Dziewonsky AD and Anderson DL (1981) Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors* 25: 297.

Fei ZL and Sideris MG (2000) A new method for computing the ellipsoidal correction for Stokes's formula. *Journal of Geodesy* 74(2): 223–231.

Goodkind JM (1999) The superconducting gravimeter. *Review of Scientific Instruments* 70: 4131–4152.

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) (The Geodesist's Handbook pp. 724–731).

Gumert WR (1998) An historical review of airborne gravimetry. *The Leading Edge* 17(1): 113–116.

Günter NM (1967) *Potential Theory and Its Applications to Basic Problems of Mathematical Physics*. New York: Frederick Ungar.

Hammer S (1982) Airborne gravity is here!. *Oil and Gas Journal* 80: 2 (January 11, 1982).

Heiskanen WA and Moritz H (1967) *Physical Geodesy*. San Francisco, CA: Freeman.

Helmert FR (1884) *Die Mathematischen und Physikalischen Theorien der Höheren Geodäsie*, vol. 2. Leipzig, Germany: B.G. Teubner, (reprinted in 1962 by Minerva GMBH, Frankfurt/Main, Germany).

Hobson EW (1965) *The Theory of Spherical and Ellipsoidal Harmonics*. New York: Chelsea.

Hofmann-Wellenhof B and Moritz H (2005) *Physical Geodesy*. Berlin, Germany: Springer-Verlag.

Hotine M (1969) *Mathematical Geodesy*. Washington, DC: US Department of Commerce.

Jekeli C (1988) The exact transformation between ellipsoidal and spherical harmonic expansions. *Manuscripta Geodaetica* 14: 106–113.

Jekeli C (1996) Spherical harmonic analysis, aliasing, and filtering. *Journal of Geodesy* 70: 214–223.

Jekeli C (1999) The determination of gravitational potential differences from satellite-to-satellite tracking. *Celestial Mechanics and Dynamical Astronomy* 75(2): 85–100.

Jekeli C (2000a) *Inertial Navigation Systems with Geodetic Applications*. Berlin, Germany: Walter deGruyter.

Jekeli C (2000b) Heights, the geopotential, and vertical Datums. *Report no.459, Department of Civil and Environmental Engineering and Geodetic Science, Ohio State University*. Columbus, OH: Ohio State University.

Jekeli C (2004) High-resolution gravity mapping: The next generation of sensors. In: Sparks RSJ (ed.) *Geophysical Monograph 150, IUGG, vol. 19: The State of the Planet: Frontiers and Challenges in Geophysics*, pp. 135–146. Washington, DC: American Geophysical Union.

Kaula WM (1966) *Theory of Satellite Geodesy*. London: Blaisdell.

Kellogg OD (1953) *Foundations of Potential Theory*. New York: Dover.

Kenyon S and Forsberg R (2000) Arctic gravity project. *SEG Technical Program Expanded Abstracts* 19: 410–413.

Kwon JH and Jekeli C (2001) A new approach for airborne vector gravimetry using GPS/INS. *Journal of Geodesy* 74: 690–700.

Lambeck K (1988) *Geophysical Geodesy*. Oxford: Clarendon Press.

Lemoine FG, Kenyon SC, Factor JK, *et al.* (1998) *The development of the joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) geopotential model EGM96. NASA Technical Paper NASA/TP-1998-206861*, Greenbelt, MD: Goddard Space Flight Center.

Martin JL (1988) *General Relativity: A Guide to Its Consequences for Gravity and Cosmology*. New York: Wiley.

Molodensky MS, Eremeev VG, and Yurkina MI (1962) *Methods for Study of the External Gravitational Field and Figure of the Earth*. Jerusalem, Israel: Israel Program for Scientific Translations (Translation from Russian).

Moritz H (1980) *Advanced Physical Geodesy*. Karlsruhe, Germany: Wichmann, (reprinted in 2001 by Civil and Environmental Engineering and Geodetic Science, Ohio State University, Columbus, OH).

Morse PM and Feshbach H (1953) *Methods of Theoretical Physics, Parts I and II*. New York: McGraw-Hill.

Müller C (1966) *Spherical Harmonics*. Lecture Notes in Mathematics. Berlin, Germany: Springer-Verlag.

Nettleton LL (1976) *Gravity and Magnetics in Oil Prospecting*. New York: McGraw-Hill.

Niebauer T, Sasagawa G, Faller J, Hilt R, and Klopping F (1995) A new generation of absolute gravimeters. *Metrologia* 32(3): 159–180.

Pavlis DE, Moore D, Luo S, McCarthy JJ, and Luthcke SB (1999) *GEODYN Operations Manual, 5 vols. Greenbelt, MD: Raytheon ITSS.*

Petrovskaya MS (1979) Upward and downward continuations in the geopotential determination. *Bulletin Géodésique* 53: 259–271.

Rapp RH (1969) Analytical and numerical differences between two methods for the combination of gravimeter and satellite data. *Bollettino di Geofisica Teorica ed Applicata* XI(41–42): 108–118.

Rapp RH and Pavlis NK (1990) The development and analysis of geopotential coefficient models to spherical harmonic degree 360. *Journal of Geophysical Research* 95(B13): 21885–21911.

Reigber Ch, Balmino G, Schwintzer P, *et al.* (2002) A high quality global gravity field model from CHAMP GPS tracking data and accelerometry (EIGEN-1S). *Geophysical Research Letters* 29(14): 37–41 (doi:10.1029/2002GL015064).

Rummel R, Balmino G, Johannessen J, Visser P, and Woodworth P (2002) Dedicated gravity field missions – Principles and aims. *Journal of Geodynamics* 33: 3–20.

Sandwell DT and Smith WHF (1996) Marine gravity anomaly form GEOSAT and ERS-1 satellite altimetry. *Journal of Geophysical Research* 102(B5): 10039–10054.

Sansò F and Rummel R (eds.) (1997) Lecture Notes in Earth Sciences 65: Geodetic Boundary Value Problems in View of the One Centimeter Geoid. Berlin, Germany: Springer-Verlag.

Seeber G (1993) *Satellite Geodesy, Foundations, Methods, and Applications*. Berlin, Germany: Walter deGruyter.

Sjöberg LE (1991) Refined least squares modifications of Stokes' formula. *Manuscripta Geodaetica* 16: 367–375.

Tapley BD (1973) Statistical orbit determination theory. In: Tapley BD and Szebehely V (eds.) *Recent Advances in Dynamical Astronomy*, pp. 396–425. Dordtecht, The Netherlands: D. Reidel.

Tapley BD, Bettadpur S, Ries JC, Thompson PF, and Watkins M (2004b) GRACE measurements of mass variability in the Earth system. *Science* 305(5683): 503–505.

Tapley BD, Bettadpur S, Watkins M, and Reigber Ch (2004a) The gravity recovery and climate experiment, mission overview and early results. *Geophysical Research Letters* 31(9): 1501 (doi:10.1029/2004GL019920).

Tapley BD, Ries J, Bettadpur S, *et al.* (2005) GGM02 – An improved Earth gravity field model from GRACE. *Journal of Geodesy* 79: 137–139 (doi:10.1007/s00190-005-0480-z).

Torge W (1989) *Gravimetry*. Berlin, Germany: Walter deGruyter.

Torge W (2001): *Geodesy, 3rd edn. Berlin, Germany: Walter de Gruyter.*

Wolff M (1969) Direct measurement of the Earth's gravitational potential using a satellite pair. *Journal of Geophysical Research* 74: 5295–5300.

Yu J, Jekeli C, and Zhu M (2002) The analytical solutions of the Dirichlet and Neumann boundary value problems with ellipsoidal boundary. *Journal of Geodesy* 76(11–12): 653–667.

Zumberge MA, Rinker RL, and Faller JE (1982) A portable apparatus for absolute measurements of the Earth's gravity. *Metrologia* 18(3): 145–152.