Physical Geodesy

WEIKKO A. HEISKANEN

HELMUT MORITZ

Reprint
Institute of Physical Geodesy
Technical University
Graz, Austria
1993

This book is a corrected reprint of the original edition published by W.H. Freeman and Company San Francisco, 1967.
Order from: Institute of Physical Geodesy, Steyrergasse 17, A-8010 Graz.

Preface

Almost every geodetic measurement depends in a fundamental way on the earth's gravity field. Therefore, the study of the physical properties of the gravity field and their geodetic application, which are the subject of physical geodesy, forms an essential part of the geodesist's education.

During the ten years that have passed since the writing of *The Earth and Its Gravity Field* by Heiskanen and Vening Meinesz, geodesy has progressed enormously. To incorporate the results of this progress, which has been theoretical as well as practical, in a new edition of that book became increasingly impossible. It was necessary to write an entirely new textbook, one that is different in both scope and treatment. The great increase in the amount of available information required a strict limitation to geodetic aspects; advances in theory made necessary an increased emphasis on mathematical methods. The outcome is the present book, which is intended to be theoretical in the sense in which the word is used in the term "theoretical physics."

This textbook, intended for graduate students, presupposes the background in mathematics and physics required by geodesy departments of American and European Universities. The necessary fundamentals of potential theory are presented in an introductory chapter.

Chapters 1 through 5 cover the material for a basic course in physical geodesy. Chapters 6 through 8 present a number of more specialized and advanced topics, where current research activity is high. (These chapters are likely to be more subjectively biased than the others.) The reader who has mastered them should be able to begin research of his own. For the sake of completeness we have added a chapter on celestial methods; this material may be included in the basic course.

We have tried hard to make the book self-contained. Detailed derivations are given wherever feasible. Our approach is intuitive: verbal explanations of the principles were felt to be more important than formal mathematical rigor, although the latter is not ignored.

Our general attitude is conservative. We do not believe that the concept of the geoid has become obsolete. This does not mean, however, that we are unaware of the great significance of recent theoretical developments associated mainly with the name of Molodensky: we discuss them in Chapter 8.

Observational techniques such as those used in gravity measurements or astronomical observations are deliberately omitted as being out of place in a theoretically oriented presentation.

Bibliographies of works cited in the text, many of which should be useful for further study, will be found at the end of each chapter; citation in the text is by author's name and year of publication—for example, Kellogg (1929).

We have not attempted to settle questions of priority. Names associated with formulas should be considered primarily as convenient labels. Similarly, the most readily accessible or most comprehensive publication of an author on a particular topic is given rather than his first.

Most of our own research incorporated in this book has been done at The Ohio State University. We wish to thank Dr. Walter D. Lambert for carefully checking parts of the manuscript for correct English.

December 1966

WEIKKO A. HEISKANEN
HELMUT MORITZ

Contents

1

Fundamentals of Potential Theory

2

The Gravity Field of the Earth

3

Gravimetric Methods

4

Heights Above Sea Level

5

Astrogeodetic Methods

6

Gravity Field Outside the Earth

7

Statistical Methods in Physical Geodesy

8

Modern Methods for Determining the Figure of the Earth

9

Celestial Methods

1

Fundamentals of Potential Theory

1-1. Introduction. Attraction and Potential

It is our purpose in this preparatory chapter to present the fundamentals of potential theory, including spherical and ellipsoidal harmonics, in sufficient detail to assure a full understanding of the later chapters. Our intent is to explain the meaning of the theorems and formulas, avoiding long derivations that can be found in any textbook on potential theory (see the references at the end of this chapter). A simple rather than completely rigorous presentation is offered. Still, the reader will probably consider this chapter rather abstract and perhaps more difficult than other parts of the book. Since later practical applications will give concrete meaning to the topics of the present chapter, the reader may wish to read it only cursorily at first and return to it later when necessary.

According to Newton's law of gravitation two points with masses m_1, m_2, separated by a distance l, attract each other with a force

$$F = k\frac{m_1 m_2}{l^2}. \tag{1-1}$$

This force is directed along the line connecting the two points; k is Newton's gravitational constant. In cgs units the gravitational constant has the value

$$k = 66.7 \times 10^{-9}\ \mathrm{cm^3\ g^{-1}\ sec^{-2}}, \tag{1-2}$$

according to measurements made by P. R. Heyl around 1930.

Although the masses m_1, m_2 attract each other in a completely symmetrical way, it is convenient to call one of them the attracting mass and the other the attracted mass. For simplicity we set the attracted mass equal to unity and denote the attracting mass by m. The formula

$$F = k\frac{m}{l^2} \tag{1-3}$$

expresses the force exerted by the mass m on a unit mass located a distance l from m.

We now introduce a rectangular coordinate system xyz, and denote the coordinates of the attracting mass m by ξ, η, ζ and the coordinates of the attracted point P by x, y, z. The force may be represented by a vector $\mathbf{F}$ with magnitude F (Fig. 1-1). The components of $\mathbf{F}$ are given by

$$\begin{aligned} X &= -F\cos\alpha = -\frac{km}{l^2}\frac{x-\xi}{l} = -km\frac{x-\xi}{l^3}, \\ Y &= -F\cos\beta = -\frac{km}{l^2}\frac{y-\eta}{l} = -km\frac{y-\eta}{l^3}, \\ Z &= -F\cos\gamma = -\frac{km}{l^2}\frac{z-\zeta}{l} = -km\frac{z-\zeta}{l^3}, \end{aligned} \tag{1-4}$$

where

$$l = \sqrt{(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2}. \tag{1-5}$$

We next introduce a scalar function

$$V = \frac{km}{l}, \tag{1-6}$$

called the *potential of gravitation*. The components X, Y, Z of the gravitational force $\mathbf{F}$ are then given by

$$X = \frac{\partial V}{\partial x}, \quad Y = \frac{\partial V}{\partial y}, \quad Z = \frac{\partial V}{\partial z}, \tag{1-7}$$

as can be easily verified by differentiating (1-6), since

$$\frac{\partial}{\partial x}\left(\frac{1}{l}\right) = -\frac{1}{l^2}\frac{\partial l}{\partial x} = -\frac{1}{l^2}\frac{x-\xi}{l} = -\frac{x-\xi}{l^3}, \ldots. \tag{1-8}$$

In vector symbolism (1-7) is written

$$\mathbf{F} = (X, Y, Z) = \operatorname{grad} V; \tag{1-7'}$$

that is, the force vector is the *gradient vector* of the scalar function V.

It is of basic importance that according to (1-7) the three components of the vector $\mathbf{F}$ can be replaced by a single function V. Especially when we consider the attraction of systems of point masses or of solid bodies, as we do in geodesy, it is much easier to deal with the potential than with the three components of the force. Even in these complicated cases the relations (1-7) hold; the function V is then simply the sum of the contributions of the respective particles.

Thus if we have a system of several point masses $m_1, m_2, \ldots, m_n$, the potential of the system is the sum of the individual contributions (1-6):

$$V = \frac{km_1}{l_1} + \frac{km_2}{l_2} + \cdots + \frac{km_n}{l_n} = k\sum_{i=1}^{n}\frac{m_i}{l_i}. \tag{1-9}$$

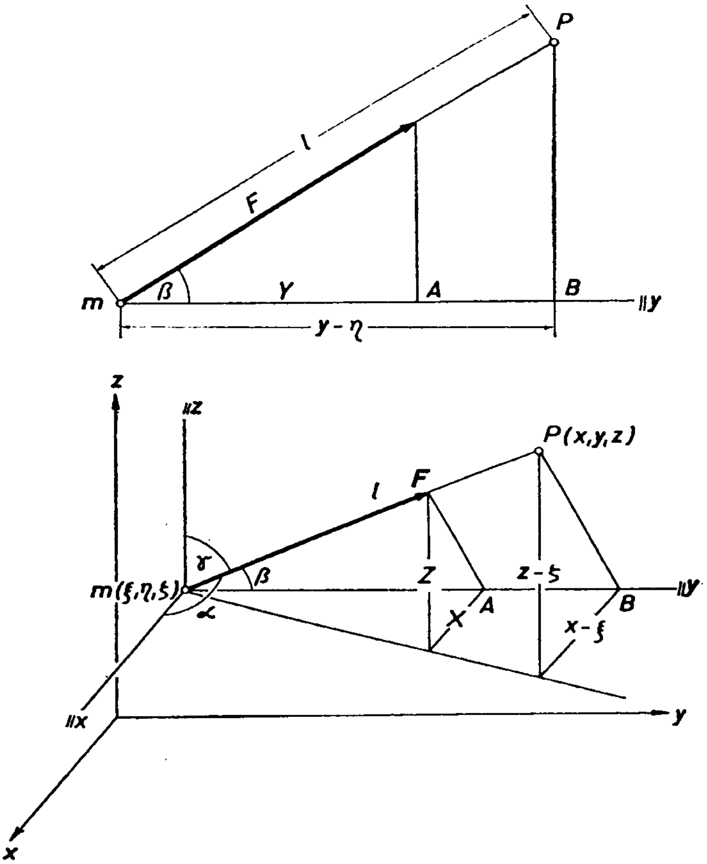

FIGURE 1-1

The components of the gravitational force. The upper figure shows the y-component.

1-2. Potential of a Solid Body

Let us now assume that point masses are distributed continuously over a volume v (Fig. 1-2) with density

$$\rho = \frac{dm}{dv}, \tag{1-10}$$

where dv is an element of volume and dm is an element of mass. Then the sum (1-9) becomes an integral

$$V = k \iiint_v \frac{dm}{l} = k \iiint_v \frac{\rho}{l}\, dv, \tag{1-11}$$

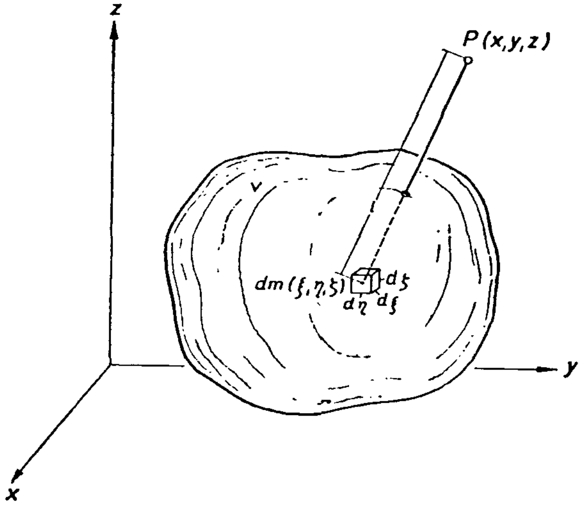

FIGURE 1-2
Potential of a solid body.

where l is the distance between the mass element $dm = \rho\, dv$ and the attracted point P. Denoting the coordinates of the attracted point by (x, y, z) and of the element of mass by (ξ, η, ζ), we see that l is again given by (1-5), and we can write explicitly

$$V(x, y, z) = k \iiint_v \frac{\rho(\xi, \eta, \zeta)}{\sqrt{(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2}}\, d\xi\, d\eta\, d\zeta, \quad \text{(1-11')}$$

since the element of volume is expressed by

$$dv = d\xi\, d\eta\, d\zeta.$$

This is the reason for the triple integrals in (1-11).

The components of the force of attraction are given by (1-7). For instance,

$$X = \frac{\partial V}{\partial x} = k \frac{\partial}{\partial x} \iiint_v \frac{\rho(\xi, \eta, \zeta)}{l}\, d\xi\, d\eta\, d\zeta$$

$$= k \iiint_v \rho(\xi, \eta, \zeta) \frac{\partial}{\partial x}\left(\frac{1}{l}\right) d\xi\, d\eta\, d\zeta.$$

Note that we have interchanged the order of differentiation and integration. Substituting (1-8) into the above expression we finally obtain

$$X = -k \iiint_v \frac{x-\xi}{l^3}\, \rho\, dv. \quad \text{(1-12)}$$

Similar expressions hold for Y and Z.

The potential V is continuous throughout the whole space and vanishes at infinity as $1/l$. This can be seen from the fact that for very great distances l the body acts approximately like a point mass, with the result that its attraction is then approximately given by (1-6). Consequently, in celestial mechanics the planets are usually considered as point masses.

The first derivatives of V, that is, the force components, are also continuous throughout space, but not so the second derivatives. At points where the density changes discontinuously, some second derivatives have a discontinuity. This is evident because the potential V satisfies *Poisson's equation:*

$$\Delta V = -4\pi k\rho, \tag{1-13}$$

where

$$\Delta V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2}. \tag{1-14}$$

The symbol Δ, called the *Laplacian operator*, has the form

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

From (1-13 and 1-14) we sec that at least one of the second derivatives of V must be discontinuous together with ρ.

Outside the attracting bodies, in empty space, the density ρ is zero and (1-13) becomes

$$\Delta V = 0. \tag{1-15}$$

This is *Laplace's equation*. Its solutions are called *harmonic functions*. Hence the potential of gravitation is a harmonic function outside the attracting masses but not inside the masses: there it satisfies Poisson's equation.

1-3. Potential of a Material Surface

Now we assume that the attracting masses form a layer, or coating, on a certain closed surface S, with thickness zero and density

$$\kappa = \frac{dm}{dS},$$

where dS is an element of surface. This case is more or less fictitious but is nevertheless of great theoretical importance.

In exact correspondence to (1-11) the potential is given by

$$V = k \iint_S \frac{dm}{l} = k \iint_S \frac{\kappa}{l}\, dS, \tag{1-16}$$

where l is the distance between the attracted point P and the surface element dS (Fig. 1-3).

On S the potential V is continuous, but there are discontinuities even in the first derivatives. Whereas the tangential derivatives on S (derivatives taken along the tangent plane) are continuous, the normal derivatives differ according to

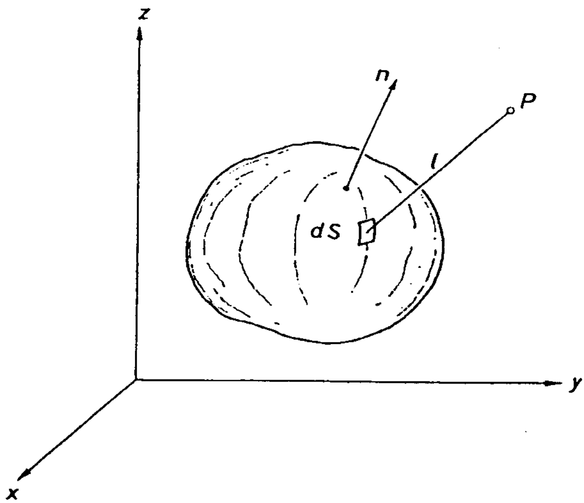

FIGURE 1-3

Potential of a material surface.

whether we approach S from the inner or from the outer side. If from the outside, then the normal derivative has on S the limit

$$\left(\frac{\partial V}{\partial n}\right)_e = -2\pi k\kappa + k \iint_S \kappa \frac{\partial}{\partial n}\left(\frac{1}{l}\right) dS; \tag{1-17a}$$

if from the inside,

$$\left(\frac{\partial V}{\partial n}\right)_i = +2\pi k\kappa + k \iint_S \kappa \frac{\partial}{\partial n}\left(\frac{1}{l}\right) dS. \tag{1-17b}$$

Here and throughout the book, $\partial/\partial n$ will denote the derivative in the direction of the *outer normal n* (Fig. 1-3).

Thus we see that the normal derivative $\partial V/\partial n$ has a discontinuity on S:

$$\left(\frac{\partial V}{\partial n}\right)_e - \left(\frac{\partial V}{\partial n}\right)_i = -4\pi k\kappa. \tag{1-18}$$

The following expressions are generalizations of equations (1-17a, b), and give the discontinuity on S of the derivative of V along an *arbitrary* direction m:

$$\left(\frac{\partial V}{\partial m}\right)_e = -2\pi k\kappa \cos(m, n) + k \iint_S \kappa \frac{\partial}{\partial m}\left(\frac{1}{l}\right) dS, \tag{1-19a}$$

$$\left(\frac{\partial V}{\partial m}\right)_i = +2\pi k\kappa \cos(m, n) + k \iint_S \kappa \frac{\partial}{\partial m}\left(\frac{1}{l}\right) dS, \tag{1-19b}$$

where (m, n) denotes the angle between the direction m and the normal n. These equations are a consequence of (1-17a, b) and of the continuity of the tangential derivatives.

Discontinuities occur only on the surface S; inside and outside S the potential V is everywhere continuous with all its derivatives, satisfying everywhere, except on S itself, Laplace's equation for harmonic functions,

$$\Delta V = 0.$$

At infinity the potential of a surface behaves in the same way as the potential of a solid body, vanishing like $1/l$ for $l \longrightarrow \infty$.

The potential of material surfaces is also called *single-layer potential* in order to distinguish it from the double-layer potential to be considered next.

1-4. Potential of a Double Layer

Imagine a *dipole* consisting of two equal point masses of opposite sign, $+m$ and $-m$, separated by a small distance h (Fig. 1-4). In gravitation such a case is purely fictitious because there are no negative masses, but the notion is nevertheless mathematically useful. In magnetism, however, there are real dipoles. The potential of the positive mass is given by

$$V_+ = \frac{km}{l},$$

the potential of the negative mass by

$$V_- = -\frac{km}{l_1}.$$

The total potential of the dipole is

$$V = V_+ + V_- = km\left(\frac{1}{l} - \frac{1}{l_1}\right).$$

Denoting the direction of the dipole's axis by n, we can expand $1/l_1$ into a Taylor series with respect to h:

$$\frac{1}{l_1} = \frac{1}{l} - \frac{\partial}{\partial n}\left(\frac{1}{l}\right)h + \frac{1}{2}\frac{\partial^2}{\partial n^2}\left(\frac{1}{l}\right)h^2 - \cdots.$$

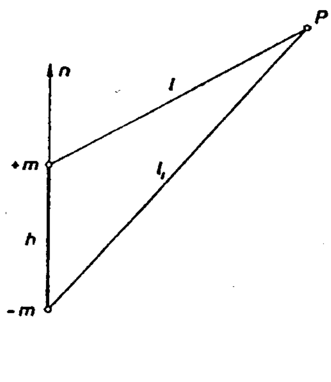

FIGURE 1-4

Potential of a dipole.

Substituting this into the preceding formula we get

$$V = k \cdot mh \cdot \frac{\partial}{\partial n}\left(\frac{1}{l}\right) - k\frac{mh^2}{2}\frac{\partial^2}{\partial n^2}\left(\frac{1}{l}\right) + \cdots$$

or, if we denote the product mh, mass times distance, as M,

$$V = kM\frac{\partial}{\partial n}\left(\frac{1}{l}\right) - k\frac{Mh}{2}\frac{\partial^2}{\partial n^2}\left(\frac{1}{l}\right) + \cdots.$$

The quantity $mh = M$ is called *dipole moment*. Now let the distance h decrease indefinitely and at the same time let the mass m increase so that the dipole moment $M = mh$ remains finite. Then the higher order terms tend toward zero as $h \longrightarrow 0$, and the expression for V reaches a limit:

$$V = kM\frac{\partial}{\partial n}\left(\frac{1}{l}\right). \tag{1-20}$$

This is the *potential of a dipole.*

A *double layer* on a surface S can be pictured as two single layers separated by a small distance h. The surface normal n intersects the two layers at two points P and P' which are very close and have surface densities of equal magnitude κ and opposite sign (Fig. 1-5). Hence, every corresponding pair of points P, P' form a dipole with *dipole density* (density of dipole moment)

$$\mu = \frac{dM}{dS},$$

which in the above figure, is given by $\mu = \kappa h$ (h very small, κ very large).

From (1-20), by summing (integrating) over all dipoles, which are continuously distributed over the surface S, we get

$$V = k\iint_S \frac{\partial}{\partial n}\left(\frac{1}{l}\right) dM = k\iint_S \mu\frac{\partial}{\partial n}\left(\frac{1}{l}\right) dS. \tag{1-21}$$

This is the *potential of the double layer* on the surface S.

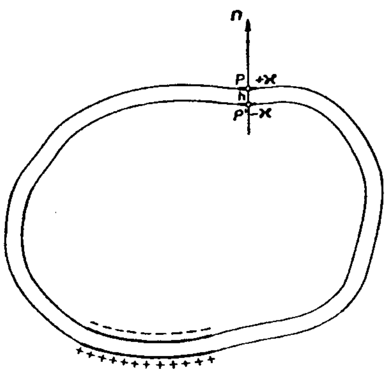

FIGURE 1-5

The double layer potential as the limit of the potential of two single layers on two neighboring parallel surfaces.

It is continuous everywhere except on the surface S; there we obtain two different limits for the potential, depending on the side (inner or outer) from which we approach S:

$$V_e = 2\pi k\mu + k \iint_S \mu \frac{\partial}{\partial n}\left(\frac{1}{l}\right) dS, \tag{1-22a}$$

$$V_i = -2\pi k\mu + k \iint_S \mu \frac{\partial}{\partial n}\left(\frac{1}{l}\right) dS. \tag{1-22b}$$

The difference,

$$V_e - V_i = 4\pi k\mu, \tag{1-23}$$

is the discontinuity to which V is subjected at the surface S as we pass from the outside to the inside.

Although equations (1-22a, b) are similar in appearance to (1-17a, b), they are basically different. In equations (1-17a, b) the differentiation $\partial/\partial n$ refers to the surface normal in the attracted point P if, as a limit, it lies on the surface S itself. In the formulas for the double-layer potential, and consequently in (1-22a, b), the differentiation $\partial/\partial n$ is taken along the surface normal in the variable attracting point which carries the surface element dS. Of course, n is the *outward* direction of the surface normal in both cases.

The double layer must be sharply distinguished from the single layer, or coating, the difference being that between mass dipole and point mass. Common to both is the behavior at infinity (vanishing like $1/l$) and the fact that they are harmonic in both the interior and the exterior of S, satisfying Laplace's equation there. On S itself, however, they have discontinuities of a completely different nature, and it is these very discontinuities that make these fictitious potentials mathematically useful, especially in connection with Green's theorems.

1-5. Gauss' and Green's Integral Formulas

Green's theorems and related integral formulas are among the basic equations of potential theory; they are indispensable tools for certain problems of theoretical geodesy.

Gauss' formula. We start with *Gauss' integral formula*,

$$\iiint_v \operatorname{div} \mathbf{F}\, dv = \iint_S F_n\, dS, \tag{1-24}$$

where v is the volume enclosed by the surface S, F_n is the projection of the vector $\mathbf{F}$ onto the outer surface normal n (i.e., the *normal component* of $\mathbf{F}$), and div $\mathbf{F}$ is the so-called *divergence* of the vector $\mathbf{F}$. If $\mathbf{F}$ has the components X, Y, Z, that is,

$$\mathbf{F} = (X, Y, Z),$$

then

$$\operatorname{div} \mathbf{F} = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z}. \tag{1-25}$$

Since Gauss' formula is well known and can be found in any textbook on engineering mathematics or mathematical physics, we need not derive it here. Instead we shall try to make it understood intuitively.

Formula (1-24) is valid for any vector field, whatever its physical meaning. Especially clear is the case in which **F** is the velocity vector of an incompressible fluid. Inside the surface S there may be sources of flow in which fluid is generated or sinks in which fluid is annihilated. The strength of the sources or sinks is measured by div **F**. The integral on the left-hand side of (1-24) is the amount of fluid generated (or annihilated) in unit time by the combined action of the sources and sinks inside S; the right-hand side is the amount of fluid flowing in unit time across the surface S. Gauss' formula (1-24) expresses the evident fact that both quantities are equal.

In the case where **F** is the vector of the gravitational force, the intuitive interpretation is not so obvious, but the analogy to fluid flow is often useful. In gravitation the components X, Y, Z of the force can be derived from a potential V by equations (1-7):

$$X = \frac{\partial V}{\partial x}, \quad Y = \frac{\partial V}{\partial y}, \quad Z = \frac{\partial V}{\partial z}.$$

Hence

$$\operatorname{div} \mathbf{F} = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = \Delta V,$$

so that by Poisson's equation (1-13)

$$\operatorname{div} \mathbf{F} = -4\pi k\rho.$$

This can be interpreted to mean that the masses are the sources of the gravitational field; the strength of the sources, div **F**, is proportional to the mass density ρ. The right-hand side of (1-24) is called *flux of force*, in our case *gravitational flux*, also in analogy to fluid flow.

For any force whose components can be derived from a potential V according to equations (1-7), Gauss' formula may be expressed in terms of the function V. For the moment we take the positive x-axis in the direction of the outer surface normal n; then the normal component of **F** is the x-component X: $F_n = X$. Then since $\partial V/\partial x = \partial V/\partial n$, the derivative of V in the direction of the outer normal n, we see from (1-7) that

$$F_n = \frac{\partial V}{\partial n}.$$

Inserting this and the relation div $\mathbf{F} = \Delta V$ into (1-24) we get

$$\iiint_v \Delta V \, dv = \iint_S \frac{\partial V}{\partial n} \, dS. \tag{1-26}$$

This is *Gauss' integral formula for the potential.*

In deriving (1-26) from (1-24) we have used only the fact that the force **F** is the gradient of a function V. We need not assume that V satisfies Poisson's equation for the gravitational field. Therefore, Gauss' integral holds for an arbitrary function V which is sufficiently regular and differentiable.

Green's formulas. These formulas are derived from (1-24) by the substitution

$$X = U\frac{\partial V}{\partial x}, \quad Y = U\frac{\partial V}{\partial y}, \quad Z = U\frac{\partial V}{\partial z},$$

where U, V are functions of x, y, z. The normal component of the vector $\mathbf{F} = (X, Y, Z)$ is then given by

$$F_n = U\frac{\partial V}{\partial n}.$$

In order to see this, consider again the x-axis coinciding with the normal n. The divergence is, by (1-25),

$$\operatorname{div}\mathbf{F} = \frac{\partial U}{\partial x}\frac{\partial V}{\partial x} + \frac{\partial U}{\partial y}\frac{\partial V}{\partial y} + \frac{\partial U}{\partial z}\frac{\partial V}{\partial z} + U\,\Delta V.$$

Thus, (1-24) becomes

$$\iiint_v U\,\Delta V\,dv + \iiint_v \left(\frac{\partial U}{\partial x}\frac{\partial V}{\partial x} + \frac{\partial U}{\partial y}\frac{\partial V}{\partial y} + \frac{\partial U}{\partial z}\frac{\partial V}{\partial z}\right) dv = \iint_S U\frac{\partial V}{\partial n}\,dS. \tag{1-27}$$

This is *Green's first identity.*

If in this formula we interchange the functions U and V and subtract the new equation from the original, we obtain

$$\iiint_v (U\,\Delta V - V\,\Delta U)\,dv = \iint_S \left(U\frac{\partial V}{\partial n} - V\frac{\partial U}{\partial n}\right) dS. \tag{1-28}$$

This is *Green's second identity.*

In these formulas we have presupposed the functions U, V to be continuous and finite in the spatial region v (i.e., inside and on the surface S) and to have continuous and finite partial derivatives of the first and second order there.

Of great importance is the case in which

$$U = \frac{1}{l},$$

where l is the distance from a certain fixed point P. If P is outside the surface S, then $1/l$ is regular inside and on S, and U satisfies the conditions mentioned. If, however, P is inside S or on S, then $1/l$ becomes infinite at a point in v and (1-28) cannot be applied directly but must be modified. Omitting the derivation we state only the result:

$$\iiint_v \frac{1}{l}\,\Delta V\,dv = -pV + \iint_S \left[\frac{1}{l}\frac{\partial V}{\partial n} - V\frac{\partial}{\partial n}\left(\frac{1}{l}\right)\right] dS, \tag{1-29}$$

where

$$p = \begin{cases} 4\pi & \text{if } P \text{ inside } S, \\ 2\pi & \text{if } P \text{ on } S, \\ 0 & \text{if } P \text{ outside } S. \end{cases}$$

This is *Green's third identity*. It differs from the second identity (1-28) by the term $-pV$. The reason for the different forms of (1-29), according as point P is inside, on, or outside S, is the term containing $\partial/\partial n(1/l)$, which can be regarded as a double-layer potential having discontinuities on S. If P is outside S, then $1/l$ is regular in v, and equation (1-29), with $p = 0$, is an immediate consequence of (1-28); v is the interior of the surface S (including S itself), and n is the normal to S, directed outward.

Green's third identity (1-29) is also valid if v is the *exterior* of the surface S and the normal n is the *inner* normal of S. If we wish to maintain n as the *outer* normal, then we have to reverse the sign of $\partial/\partial n$, getting

$$\iiint_v \frac{1}{l} \Delta V \, dv = -pV - \iint_S \left[\frac{1}{l}\frac{\partial V}{\partial n} - V\frac{\partial}{\partial n}\left(\frac{1}{l}\right)\right] dS, \qquad (1\text{-}29')$$

where

$$p = \begin{cases} 4\pi & \text{if } P \text{ outside } S, \\ 2\pi & \text{if } P \text{ on } S, \\ 0 & \text{if } P \text{ inside } S. \end{cases}$$

This is *Green's third identity for the exterior of the surface S*. It is valid for functions V that, besides satisfying the general requirements for Green's identities, satisfy certain conditions at infinity, such as vanishing there.

1-6. Applications of Green's Integral Formulas

In order to show the significance and usefulness of Green's identities, we shall now apply them to special cases.

1. In the third identity (1-29) we set $V \equiv 1$. Then

$$\iint_S \frac{\partial}{\partial n}\left(\frac{1}{l}\right) dS = \begin{cases} -4\pi & \text{if } P \text{ inside } S, \\ -2\pi & \text{if } P \text{ on } S, \\ 0 & \text{if } P \text{ outside } S. \end{cases} \qquad (1\text{-}30)$$

These formulas, which are sometimes useful, are also due to Gauss. They may be considered theorems on the potential of a double layer of constant density $k\mu = 1$. Such a potential has a constant value inside the surface and is zero outside, with the characteristic discontinuity (1-23) on S.

2. In this case V is a harmonic function outside S: $\Delta V = 0$. If the point P is also outside S, then the third identity (1-29′) yields ($p = 4\pi$):

$$V = -\frac{1}{4\pi}\iint_S \frac{1}{l}\frac{\partial V}{\partial n} dS + \frac{1}{4\pi}\iint_S V\frac{\partial}{\partial n}\left(\frac{1}{l}\right) dS. \qquad (1\text{-}31)$$

This formula shows that every harmonic function may be represented as the sum of a surface potential (1-16), with density

$$\kappa = -\frac{1}{4\pi k}\frac{\partial V}{\partial n},$$

and a double-layer potential (1-21), with density $\mu = V/(4\pi k)$.

3. Again V is harmonic outside S. We further assume S to be a surface $V = V_0 = \text{const.}$, that is, a surface of constant potential V, or an *equipotential surface*. Then, for a point P outside S, we get from (1-31)

$$V = -\frac{1}{4\pi}\iint_S \frac{1}{l}\frac{\partial V}{\partial n}\,dS + \frac{V_0}{4\pi}\iint_S \frac{\partial}{\partial n}\left(\frac{1}{l}\right) dS.$$

The second integral is zero according to (1-30). Hence

$$V = -\frac{1}{4\pi}\iint_S \frac{1}{l}\frac{\partial V}{\partial n}\,dS. \tag{1-32}$$

This formula, attributed to Chasles, shows that every harmonic function can be represented as a single-layer potential on any of its equipotential surfaces $V = \text{const}$. If V is the Newtonian potential of a solid body which lies inside S, we can say that any solid body can be replaced by a suitable surface layer on one of its outer equipotential surfaces S without changing its potential outside S (Fig. 1-6).

We shall now give two somewhat more elaborate examples which are of basic importance to physical geodesy.

4. In the second identity (1-28) we set $U \equiv 1$. We again get Gauss' formula (1-26):

$$\iiint_v \Delta V\,dv = \iint_S \frac{\partial V}{\partial n}\,dS.$$

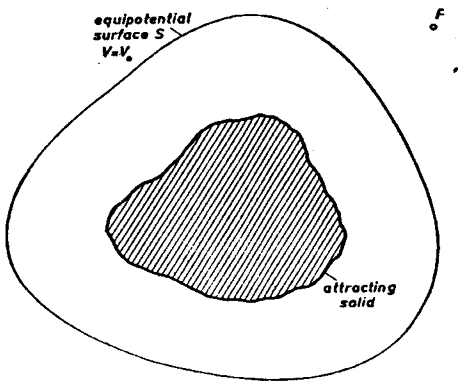

FIGURE 1-6

Theorem of Chasles. The potential, at any point P outside S, of a surface layer of density $\kappa = -(4\pi k)^{-1}\,\partial V/\partial n$ is the same as that of the attracting solid itself.

Apply this formula to the potential W of gravity (gravitation plus centrifugal force; see Section 2-1):

$$\iiint_v \Delta W \, dv = \iint_S \frac{\partial W}{\partial n} \, dS.$$

The function W satisfies an equation (2-6)

$$\Delta W = -4\pi k\rho + 2\omega^2,$$

which is similar to Poisson's equation (1-13); ω is the angular velocity of the earth's rotation. Let v be the earth and S its physical surface. Then, as we shall see,

$$\frac{\partial W}{\partial n} = -g_n,$$

which is the component of gravity normal to the earth's surface S.

Taking these two relations into account we find

$$\iiint_v (-4\pi k\rho + 2\omega^2) \, dv = -\iint_S g_n \, dS$$

or

$$M = \frac{1}{4\pi k} \iint_S g_n \, dS + \frac{\omega^2}{2\pi k} v, \tag{1-33}$$

where

$$M = \iiint_v \rho \, dv$$

is the mass of the earth and v is its volume. Basically, this equation is the reason why it is possible *to determine the mass of the earth from measured gravity*. Note that it is not necessary for this purpose to know the detailed density distribution in the interior of the earth!

5. Consider again the earth and its gravity potential W and apply the third identity (1-29) to a point on the earth's surface. Then $p = 2\pi$, so that we have

$$\iiint_v \frac{1}{l} \Delta W \, dv + 2\pi W - \iint_S \left[\frac{1}{l}\frac{\partial W}{\partial n} - W \frac{\partial}{\partial n}\left(\frac{1}{l}\right)\right] dS = 0.$$

With the same substitutions as before we get

$$\iiint_v \frac{1}{l}(-4\pi k\rho + 2\omega^2) \, dv + 2\pi W + \iint_S \left[W \frac{\partial}{\partial n}\left(\frac{1}{l}\right) + \frac{g_n}{l}\right] dS = 0$$

and, according to (1-11),

$$W = k \iiint_v \frac{\rho}{l} \, dv + \frac{1}{2}\omega^2(x^2 + y^2),$$

we finally obtain

$$-2\pi W + \iint_S \left[W \frac{\partial}{\partial n}\left(\frac{1}{l}\right) + \frac{g_n}{l} \right] dS + 2\pi\omega^2(x^2 + y^2) + 2\omega^2 \iiint_v \frac{dv}{l} = 0. \tag{1-34}$$

All quantities in this equation are referred to the surface S.

Equation (1-34) relates the surface S to the gravity potential W and the gravity g. If W and g are given, it is thus reasonable to assume that one can solve in some way the above equation for the surface S. Actually, this equation may be considered the mathematical basis for the *determination of the physical surface S of the earth from measured potential W and gravity g*, according to the famous theory of Molodensky (see Chapter 8).

1-7. Harmonic Functions. Stokes' Theorem and Dirichlet's Principle

Earlier we have defined the harmonic functions as solutions of Laplace's equation

$$\Delta V = 0.$$

More precisely, a function is called *harmonic in a region v* of space if it satisfies Laplace's equation at every point of v. If the region is the exterior of a certain closed surface S, then it must in addition vanish like $1/l$ for $l \longrightarrow \infty$. It can be shown that *every harmonic function is analytic* (in the region where it satisfies Laplace's equation); that is, it is continuous and has continuous derivatives of any order.

The simplest harmonic function is the reciprocal distance

$$\frac{1}{l} = \frac{1}{\sqrt{(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2}}$$

between two points (ξ, η, ζ) and (x, y, z), considered as a function of x, y, z. It is the potential of a point mass $m = 1/k$, situated at the point (ξ, η, ζ); compare (1-5) and (1-6) for $km = 1$.

It is easy to show that $1/l$ is harmonic. We form the following partial derivatives with respect to x, y, z in the fashion of (1-8):

$$\frac{\partial}{\partial x}\left(\frac{1}{l}\right) = -\frac{x-\xi}{l^3}, \quad \frac{\partial}{\partial y}\left(\frac{1}{l}\right) = -\frac{y-\eta}{l^3}, \quad \frac{\partial}{\partial z}\left(\frac{1}{l}\right) = -\frac{z-\zeta}{l^3};$$

$$\frac{\partial^2}{\partial x^2}\left(\frac{1}{l}\right) = \frac{-l^2 + 3(x-\xi)^2}{l^5}, \quad \frac{\partial^2}{\partial y^2}\left(\frac{1}{l}\right) = \frac{-l^2 + 3(y-\eta)^2}{l^5},$$

$$\frac{\partial^2}{\partial z^2}\left(\frac{1}{l}\right) = \frac{-l^2 + 3(z-\zeta)^2}{l^5}.$$

Adding the last three equations and recalling the definition of Δ we find

$$\Delta\left(\frac{1}{l}\right) = 0; \tag{1-35}$$

that is, $1/l$ is harmonic.

The point (ξ, η, ζ), where l is zero and $1/l$ is infinite, is the only point to which we cannot apply the above derivation; $1/l$ is not harmonic at this singular point.

As a matter of fact, the slightly more general potential (1-6) of an arbitrary point mass m is also harmonic except at (ξ, η, ζ), because (1-35) remains unchanged if both sides are multiplied by km.

Not only the potential of a point mass, but also any other gravitational potential is harmonic outside the attracting masses. Consider the potential (1-11) of an extended body. Interchanging the order of differentiation and integration we find from (1-11)

$$\Delta V = k\Delta\left[\iiint_v \frac{\rho}{l}\,dv\right] = k\iiint_v \rho\Delta\left(\frac{1}{l}\right)dv = 0;$$

that is, the potential of a solid body is also harmonic at any point $P(x, y, z)$ outside the attracting masses.

If P lies inside the attracting body the above derivation breaks down, since $1/l$ becomes infinite for that mass element $dm\ (\xi, \eta, \zeta)$ which coincides with $P(x, y, z)$, and (1-35) does not hold. This is the reason why the potential of a solid body is not harmonic in its interior but instead satisfies Poisson's differential equation (1-13).

In exactly the same manner it may be shown that the potential (1-16) of an attracting layer on a surface S is harmonic at all points except the points of S itself. As a consequence we see that the potential (1-21) of a double layer is also harmonic everywhere except on the surface S, since the double-layer potential may be considered the limit of the combined potential of two neighboring surface layers; compare Fig. 1-5.

Thus the gravitational potential is harmonic at all points where there are no attracting masses, and, consequently, so is the outer potential of the earth if we disregard the atmosphere and the centrifugal force. This is the reason for the basic importance of harmonic functions in physical geodesy.

In general, the same harmonic function can be generated by many different mass distributions. A well-known example is the exterior potential of a homogeneous sphere:

$$V = \frac{kM}{l},$$

where M is the mass of the sphere and l is the distance from its center.[1] Hence all concentric homogeneous spheres of the same total mass M, regardless of their size, generate the same potential. The potential is the same as if the total mass were concentrated at its center, because the potential of a point mass is also given by this formula.

Another example is the theorem of Chasles (1-32). Take any Newtonian

[1] This may be seen immediately from (2-39): in the case of spherical symmetry the J_{nm} and K_{nm} must be zero.

potential V and denote one of its exterior equipotential surfaces by S. Outside S, the potential will be the same as that of a surface layer of density

$$-\frac{1}{4\pi k}\frac{\partial V}{\partial n};$$

see Fig. 1-6.

These are particular instances of *Stokes' theorem*. A function V harmonic outside a surface S is uniquely determined by its values on S. In general, however, there are infinitely many mass distributions which have the given harmonic function V as exterior potential.

It is therefore impossible to determine uniquely the generating masses from the external potential. This *inverse problem of potential theory* has no unique solution (direct problem: determination of the potential from the masses, inverse problem: determination of the masses from the potential). The inverse problem occurs in geophysical prospecting by gravity measurements: invisible masses are inferred from disturbances of the gravity field. To determine the problem more completely, additional information is necessary, which is furnished, for example, by geology or by seismic measurements.

Because of the importance of Stokes' theorem we shall give a simple proof of its first part. Let a given mass distribution generate a potential V and let S be a surface that encloses all masses. Assume that a different mass distribution inside S generates a potential V' which takes on the same values on the surface S. If we denote the difference $V' - V$ by U, then, according to our assumption, $U = 0$ on S. Taking Green's first identity (1-27) and setting both functions equal to each other, we get

$$\iiint_v U\,\Delta U\,dv + \iiint_v \left[\left(\frac{\partial U}{\partial x}\right)^2 + \left(\frac{\partial U}{\partial y}\right)^2 + \left(\frac{\partial U}{\partial z}\right)^2\right] dv = \iint_S U\frac{\partial U}{\partial n}\,dS.$$

We apply this equation to the *exterior* of S, so that v is the region outside S.[1] Because $U = V' - V$, being the difference of two harmonic functions, is also harmonic outside S, we have $\Delta U = 0$ in v; in addition, $U = 0$ on S. Hence the right-hand side and the first integral of the left-hand side vanish, and we get

$$\iiint_v \left[\left(\frac{\partial U}{\partial x}\right)^2 + \left(\frac{\partial U}{\partial y}\right)^2 + \left(\frac{\partial U}{\partial z}\right)^2\right] dv = 0.$$

If only one derivative of U is different from zero, this equation cannot hold, because the integrand is always positive or zero. Thus, all derivatives of U must be zero; that is, U is a constant. Since U, as a harmonic function, must be zero at infinity, the constant must be zero. Hence, $V' - V = 0$ or $V' = V$ throughout v, which is what we set out to demonstrate.

Stokes' theorem states that there is only one harmonic function V that assumes given *boundary values* on a surface S, *provided that such a harmonic*

[1] This is possible if U is harmonic, because the regularity conditions at infinity, mentioned at the end of the preceding sections, are satisfied in this case.

function exists. The assertion that for arbitrarily prescribed boundary values there always exists a harmonic function V that assumes on S the given boundary values is called *Dirichlet's principle.* We have two different cases: V harmonic outside S, and V harmonic inside S.

Dirichlet's principle has been proved for very general cases by the work of many mathematicians, for example, Poincaré and Hilbert; the proof is very difficult.

The problem of computing the harmonic function (inside or outside S) from its boundary values on S is *Dirichlet's problem,* or the *first boundary-value problem of potential theory.* We shall return to it in Section 1-16.

Finally we remark that there is no function that is harmonic throughout the entire space (except the trivial case $V \equiv 0$): at least one singularity always occurs. The potential of a point mass, $V = km/l$, is singular for $l = 0$; the potential of a surface distribution or of a double layer on a surface S is harmonic inside and outside S, but not on S itself.

1-8. Laplace's Equation in Spherical Coordinates

The most important harmonic functions are the so-called *spherical harmonics.* To find them we introduce spherical coordinates: r (radius vector), θ (polar distance), λ (geocentric longitude) (Fig. 1-7). Spherical coordinates are related to rectangular coordinates x, y, z by the equations

$$\begin{aligned} x &= r \sin\theta \cos\lambda, \\ y &= r \sin\theta \sin\lambda, \\ z &= r\cos\theta; \end{aligned} \tag{1-36}$$

or inversely by

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2}, \\ \theta &= \tan^{-1}\frac{\sqrt{x^2 + y^2}}{z}, \\ \lambda &= \tan^{-1}\frac{y}{x}. \end{aligned} \tag{1-37}$$

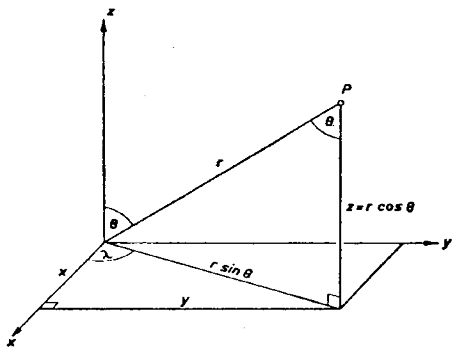

FIGURE 1-7

Spherical and rectangular coordinates.

To get Laplace's equation in spherical coordinates we first determine the element of arc (element of distance) ds in these coordinates. For this purpose we form

$$dx = \frac{\partial x}{\partial r}\,dr + \frac{\partial x}{\partial \theta}\,d\theta + \frac{\partial x}{\partial \lambda}\,d\lambda,$$

$$dy = \frac{\partial y}{\partial r}\,dr + \frac{\partial y}{\partial \theta}\,d\theta + \frac{\partial y}{\partial \lambda}\,d\lambda,$$

$$dz = \frac{\partial z}{\partial r}\,dr + \frac{\partial z}{\partial \theta}\,d\theta + \frac{\partial z}{\partial \lambda}\,d\lambda.$$

By differentiating (1-36) and inserting it into the elementary formula

$$ds^2 = dx^2 + dy^2 + dz^2$$

we obtain

$$ds^2 = dr^2 + r^2\,d\theta^2 + r^2 \sin^2\theta\,d\lambda^2. \tag{1-38}$$

We might have found this well-known formula more simply by geometrical considerations, but the approach used here is more general and can also be applied to ellipsoidal coordinates.

In (1-38) there are no terms with $dr\,d\theta$, $dr\,d\lambda$, and $d\theta\,d\lambda$. This expresses the evident fact that spherical coordinates are orthogonal: the spheres r = const., the cones θ = const., and the planes λ = const. intersect each other orthogonally.

The general form of the element of arc in arbitrary orthogonal coordinates q_1, q_2, q_3 is

$$ds^2 = h_1^2\,dq_1^2 + h_2^2\,dq_2^2 + h_3^2\,dq_3^2. \tag{1-39}$$

It can be shown that Laplace's operator in these coordinates is

$$\Delta V = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial V}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_3 h_1}{h_2}\frac{\partial V}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial V}{\partial q_3}\right)\right]. \tag{1-40}$$

For spherical coordinates we have $q_1 = r$, $q_2 = \theta$, $q_3 = \lambda$. Comparison of (1-38) and (1-39) shows that

$$h_1 = 1, \quad h_2 = r, \quad h_3 = r\sin\theta.$$

Substituting these into (1-40) we get

$$\Delta V = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 V}{\partial\lambda^2}.$$

On performing the differentiations we find

$$\Delta V \equiv \frac{\partial^2 V}{\partial r^2} + \frac{2}{r}\frac{\partial V}{\partial r} + \frac{1}{r^2}\frac{\partial^2 V}{\partial\theta^2} + \frac{\cot\theta}{r^2}\frac{\partial V}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 V}{\partial\lambda^2} = 0, \tag{1-41}$$

which is *Laplace's equation in spherical coordinates*. An alternative expression is obtained by multiplying both sides by r^2:

$$r^2\frac{\partial^2 V}{\partial r^2} + 2r\frac{\partial V}{\partial r} + \frac{\partial^2 V}{\partial\theta^2} + \cot\theta\frac{\partial V}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2 V}{\partial\lambda^2} = 0. \tag{1-41$'$}$$

This form will be somewhat more convenient for our subsequent development.

1-9. Spherical Harmonics

We shall attempt to solve Laplace's equation (1-41) or (1-41′) by separating the variables r, θ, λ by means of the trial substitution

$$V(r, \theta, \lambda) = f(r)\,Y(\theta, \lambda), \tag{1-42}$$

where f is a function of r only and Y is a function of θ and λ only. By making this substitution in (1-41′) and dividing by fY we get

$$\frac{1}{f}(r^2 f'' + 2rf') = -\frac{1}{Y}\left(\frac{\partial^2 Y}{\partial\theta^2} + \cot\theta\,\frac{\partial Y}{\partial\theta} + \frac{1}{\sin^2\theta}\,\frac{\partial^2 Y}{\partial\lambda^2}\right),$$

where the primes denote differentiation with respect to the argument (r, in this case). Since the left-hand side depends only on r and the right-hand side only on θ and λ, both sides must be constant. We can therefore separate the equation into two equations:

$$r^2 f''(r) + 2rf'(r) - n(n+1)f(r) = 0, \tag{1-43}$$

$$\frac{\partial^2 Y}{\partial\theta^2} + \cot\theta\,\frac{\partial Y}{\partial\theta} + \frac{1}{\sin^2\theta}\,\frac{\partial^2 Y}{\partial\lambda^2} + n(n+1)\,Y = 0, \tag{1-44}$$

where we have denoted the constant by $n(n+1)$.

Solutions of (1-43) are given by the functions

$$f(r) = r^n \quad \text{and} \quad f(r) = r^{-(n+1)}; \tag{1-45}$$

this should be verified by substitution. Denoting the as yet unknown solutions of (1-44) by $Y_n(\theta, \lambda)$ we see that Laplace's equation (1-41) is solved by the functions

$$V = r^n\,Y_n(\theta, \lambda) \quad \text{and} \quad V = \frac{Y_n(\theta, \lambda)}{r^{n+1}}. \tag{1-46}$$

These functions are called *solid spherical harmonics*, whereas the functions $Y_n(\theta, \lambda)$ are known as (Laplace's) *surface spherical harmonics*. Both kinds are called *spherical harmonics*; the kind referred to can usually be judged from the context.

Later we shall see that n is not an arbitrary constant but must be an integer $0, 1, 2, \ldots$. If a differential equation is linear, and if we know several solutions, then, as is well known, the sum of these solutions is also a solution. Hence we conclude that

$$V = \sum_{n=0}^{\infty} r^n\,Y_n(\theta, \lambda) \quad \text{and} \quad V = \sum_{n=0}^{\infty} \frac{Y_n(\theta, \lambda)}{r^{n+1}} \tag{1-47}$$

are also solutions of Laplace's equation $\Delta V = 0$; that is, harmonic functions. The important fact is that *every* harmonic function—with certain restrictions—can be expressed in one of the forms (1-47).

1-10. Surface Spherical Harmonics

Now we have to determine Laplace's surface harmonics $Y_n(\theta, \lambda)$. We attempt to solve (1-44) by a new trial substitution

$$Y_n(\theta, \lambda) = g(\theta)h(\lambda), \tag{1-48}$$

where the functions g and h each depend on one variable only. Making this substitution in (1-44) and multiplying by $\sin^2 \theta/gh$ we find

$$\frac{\sin\theta}{g}(\sin\theta\, g'' + \cos\theta\, g' + n(n+1)\sin\theta\, g) = -\frac{h''}{h},$$

where the primes denote differentiation with respect to the argument: θ in g, λ in h. The left-hand side is a function of θ only, and the right-hand side is a function of λ only. Therefore, both sides must again be constant; let the constant be m^2. Thus the partial differential equation (1-44) splits into two ordinary differential equations for the functions $g(\theta)$ and $h(\lambda)$:

$$\sin\theta\, g''(\theta) + \cos\theta\, g'(\theta) + \left[n(n+1)\sin\theta - \frac{m^2}{\sin\theta}\right]g(\theta) = 0; \tag{1-49}$$

$$h''(\lambda) + m^2 h(\lambda) = 0. \tag{1-50}$$

Solutions of the second equation are the functions

$$h(\lambda) = \cos m\lambda \quad \text{and} \quad h(\lambda) = \sin m\lambda, \tag{1-51}$$

as may be verified by substitution. The first equation is more difficult. It can be shown that it has physically meaningful solutions only if n and m are integers 0, 1, 2, . . . and if m is smaller than or equal to n. A solution of (1-49) is the so-called Legendre function $P_{nm}(\cos\theta)$, which will be considered in some detail in the next section. Hence

$$g(\theta) = P_{nm}(\cos\theta), \tag{1-52}$$

and the functions

$$Y_n(\theta, \lambda) = P_{nm}(\cos\theta)\cos m\lambda \quad \text{and} \quad Y_n(\theta, \lambda) = P_{nm}(\cos\theta)\sin m\lambda \tag{1-53}$$

are solutions of the differential equation (1-44) for Laplace's surface harmonics.

Since this equation is linear, any linear combination of the solutions (1-53) is also a solution. Such a linear combination has the general form

$$Y_n(\theta, \lambda) = \sum_{m=0}^{n} [a_{nm}P_{nm}(\cos\theta)\cos m\lambda + b_{nm}P_{nm}(\cos\theta)\sin m\lambda],$$

where a_{nm} and b_{nm} are arbitrary constants. This is the general expression for the surface harmonic Y_n.

Inserting this into equations (1-47) we see that

$$V_i(r, \theta, \lambda) = \sum_{n=0}^{\infty} r^n \sum_{m=0}^{n} [a_{nm}P_{nm}(\cos\theta)\cos m\lambda + b_{nm}P_{nm}(\cos\theta)\sin m\lambda], \tag{1-54a}$$

$$V_e(r, \theta, \lambda) = \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \sum_{m=0}^{n} [a_{nm}P_{nm}(\cos\theta)\cos m\lambda + b_{nm}P_{nm}(\cos\theta)\sin m\lambda] \tag{1-54b}$$

are solutions of Laplace's equation $\Delta V = 0$; that is, harmonic functions. Furthermore, as we have mentioned, they are very general solutions indeed: every function which is harmonic *inside* a certain sphere can be expanded into a series (1-54a), and every function which is harmonic *outside* a certain sphere (such as the earth's gravitational potential) can be expanded into a series (1-54b). Thus we see how spherical harmonics can be useful in geodesy.

1-11. Legendre's Functions

In the preceding section we have introduced Legendre's function $P_{nm}(\cos\theta)$ as a solution of Legendre's differential equation (1-49). The n denotes the *degree* and m the *order* of P_{nm}.

It is convenient to transform Legendre's equation (1-49) by the substitution

$$t = \cos\theta. \tag{1-55}$$

In order to avoid confusion, we use an overbar to denote g as a function of t. Hence

$$g(\theta) = \bar{g}(t),$$
$$g'(\theta) = \frac{dg}{d\theta} = \frac{dg}{dt}\frac{dt}{d\theta} = -\bar{g}'(t)\sin\theta,$$
$$g''(\theta) = \bar{g}''(t)\sin^2\theta - \bar{g}'(t)\cos\theta.$$

Inserting these into (1-49), dividing by $\sin\theta$, and then substituting $\sin^2\theta = 1 - t^2$ we get

$$(1 - t^2)\bar{g}''(t) - 2t\bar{g}'(t) + \left[n(n+1) - \frac{m^2}{1-t^2}\right]\bar{g}(t) = 0. \tag{1-56}$$

The Legendre function $\bar{g}(t) = P_{nm}(t)$, which is defined by

$$P_{nm}(t) = \frac{1}{2^n n!}(1 - t^2)^{m/2}\frac{d^{n+m}}{dt^{n+m}}(t^2 - 1)^n, \tag{1-57}$$

satisfies (1-56). Apart from the factor $(1 - t^2)^{m/2} = \sin^m\theta$ and from a constant, the function P_{nm} is the $(n + m)$th derivative of the polynomial $(t^2 - 1)^n$. It can thus be evaluated without difficulty. For instance,

$$P_{11}(t) = \frac{(1 - t^2)^{1/2}}{2\cdot 1}\frac{d^2}{dt^2}(t^2 - 1) = \frac{1}{2}\sqrt{1 - t^2}\cdot 2 = \sqrt{1 - t^2} = \sin\theta.$$

The case $m = 0$ is of particular importance. The functions $P_{n0}(t)$ are often simply denoted by $P_n(t)$. Then (1-57) gives

$$P_n(t) = P_{n0}(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}(t^2 - 1)^n. \tag{1-57'}$$

Because $m = 0$, there is no square root, that is, no $\sin\theta$. Therefore, the $P_n(t)$ are simply polynomials in t. They are called *Legendre's polynomials*. We give the first few Legendre polynomials for $n = 0$ through $n = 5$:

$$
\begin{aligned}
P_0(t) &= 1, & P_3(t) &= \frac{5}{2}t^3 - \frac{3}{2}t, \\
P_1(t) &= t, & P_4(t) &= \frac{35}{8}t^4 - \frac{15}{4}t^2 + \frac{3}{8}, \\
P_2(t) &= \frac{3}{2}t^2 - \frac{1}{2}, & P_5(t) &= \frac{63}{8}t^5 - \frac{35}{4}t^3 + \frac{15}{8}t.
\end{aligned} \tag{1-58}
$$

Remember that

$$t = \cos\theta.$$

The polynomials may be obtained by means of (1-57′) or more simply by the *recursion formula*

$$P_n(t) = -\frac{n-1}{n}P_{n-2}(t) + \frac{2n-1}{n}\,t\,P_{n-1}(t), \tag{1-59}$$

by which P_2 can be calculated from P_0 and P_1, P_3 from P_1 and P_2, etc. Graphs of the Legendre polynomials are shown in Fig. 1-8.

The powers of $\cos\theta$ can be expressed in terms of the cosines of multiples of θ, such as

$$\cos^2\theta = \frac{1}{2}\cos 2\theta + \frac{1}{2}, \quad \cos^3\theta = \frac{1}{4}\cos 3\theta + \frac{3}{4}\cos\theta.$$

Therefore, we may also express the $P_n(\cos\theta)$ in this way, obtaining

$$
\begin{aligned}
P_2(\cos\theta) &= \frac{3}{4}\cos 2\theta + \frac{1}{4}, \\
P_3(\cos\theta) &= \frac{5}{8}\cos 3\theta + \frac{3}{8}\cos\theta, \\
P_4(\cos\theta) &= \frac{35}{64}\cos 4\theta + \frac{5}{16}\cos 2\theta + \frac{9}{64}, \\
P_5(\cos\theta) &= \frac{63}{128}\cos 5\theta + \frac{35}{128}\cos 3\theta + \frac{15}{64}\cos\theta, \\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned} \tag{1-58′}
$$

If the order m is not zero—that is, for $m = 1, 2, \ldots, n$—Legendre's functions $P_{nm}(\cos\theta)$ are called *associated Legendre functions*. They can easily be reduced to the Legendre polynomials by means of the equation

$$P_{nm}(t) = (1 - t^2)^{m/2}\frac{d^m P_n(t)}{dt^m}, \tag{1-60}$$

which follows from (1-57) and (1-57′). Thus the associated Legendre functions are expressed in terms of the Legendre polynomials of the same degree n. We give some P_{nm}, writing $t = \cos\theta$, $\sqrt{1-t^2} = \sin\theta$:

$$
\begin{aligned}
P_{11}(\cos\theta) &= \sin\theta, & P_{31} &= \sin\theta\left(\frac{15}{2}\cos^2\theta - \frac{3}{2}\right), \\
P_{21}(\cos\theta) &= 3\sin\theta\cos\theta, & P_{32} &= 15\sin^2\theta\cos\theta, \\
P_{22}(\cos\theta) &= 3\sin^2\theta, & P_{33} &= 15\sin^3\theta.
\end{aligned} \tag{1-61}
$$

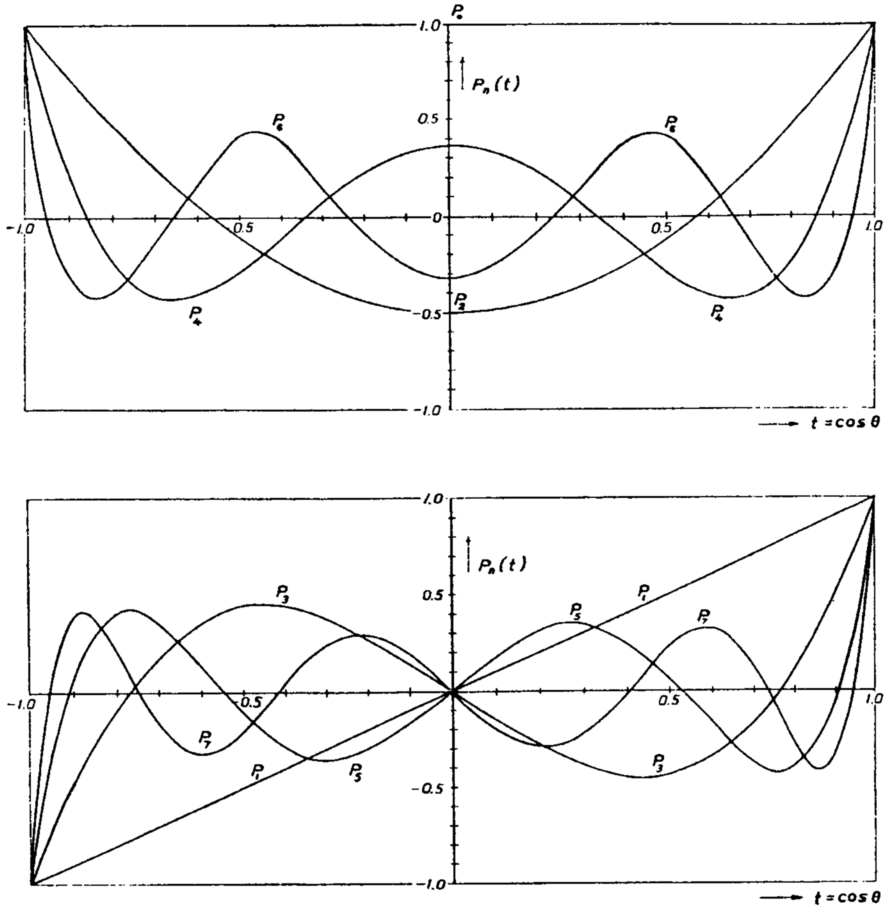

FIGURE 1-8

Legendre's polynomials as functions of $t = \cos\theta$. Top, n even; bottom, n odd.

We also mention an explicit formula for any Legendre function (polynomial or associated function):

$$P_{nm}(t) = 2^{-n}(1 - t^2)^{m/2} \sum_{k=0}^{r} (-1)^k \frac{(2n - 2k)!}{k!(n - k)!(n - m - 2k)!} t^{n-m-2k}, \quad (1\text{-}62)$$

where r is the greatest integer $\leq (n - m)/2$; i.e., r is $(n - m)/2$ or $(n - m - 1)/2$, whichever is an integer. This formula is convenient for use in programming an electronic computer.

As this useful formula is seldom found in the literature, we show the derivation, which is quite straightforward. The necessary information on factorials may be obtained from any collection of mathematical formulas.

The binomial theorem gives

$$(t^2 - 1)^n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} t^{2n-2k} = \sum_{k=0}^{n} (-1)^k \frac{n!}{k!(n-k)!} t^{2n-2k}.$$

Thus (1-57) becomes

$$P_{nm}(t) = \frac{1}{2^n} (1 - t^2)^{m/2} \sum_{k=0}^{n} (-1)^k \frac{1}{k!(n-k)!} \frac{d^{n+m}}{dt^{n+m}} (t^{2n-2k}),$$

the quantity $n!$ having been cancelled out. The rth derivative of the power t^s is

$$\frac{d^r}{dt^r}(t^s) = s(s-1)\cdots(s-r+1)t^{s-r} = \frac{s!}{(s-r)!} t^{s-r}.$$

Setting $r = n + m$ and $s = 2n - 2k$ we have

$$\frac{d^{n+m}}{dt^{n+m}}(t^{2n-2k}) = \frac{(2n-2k)!}{(n-m-2k)!} t^{n-m-2k}.$$

Inserting this into the above expression for $P_{nm}(t)$ and noting that the lowest possible power of t is either t or $t^0 = 1$, we obtain (1-62).

The surface spherical harmonics are Legendre's functions multiplied by $\cos m\lambda$ or $\sin m\lambda$:

degree 0	$P_0(\cos\theta)$;
degree 1	$P_1(\cos\theta)$, $P_{11}(\cos\theta)\cos\lambda$, $P_{11}(\cos\theta)\sin\lambda$;
degree 2	$P_2(\cos\theta)$, $P_{21}(\cos\theta)\cos\lambda$, $P_{21}(\cos\theta)\sin\lambda$, $P_{22}(\cos\theta)\cos 2\lambda$, $P_{22}(\cos\theta)\sin 2\lambda$;

and so on.

The geometrical representation of these spherical harmonics is useful. The harmonics with $m = 0$—that is, Legendre's polynomials—are polynomials of degree n in t, so that they have n zeros. These n zeros are all real and situated in the interval $-1 \leqq t \leqq +1$, that is, $0 \leqq \theta \leqq \pi$ (Fig. 1-8). The harmonics with $m = 0$ thus change their sign n times in this interval; furthermore, they do not depend on λ. Their geometrical representation is therefore similar to Fig. 1-9, case a. Since they divide the sphere into zones, they are also called *zonal harmonics*.

The associated Legendre functions change their sign $n - m$ times in the interval $0 \leqq \theta \leqq \pi$. The functions $\cos m\lambda$ and $\sin m\lambda$ have $2m$ zeros in the interval $0 \leqq \lambda < 2\pi$, so that the geometrical representation of the harmonics for

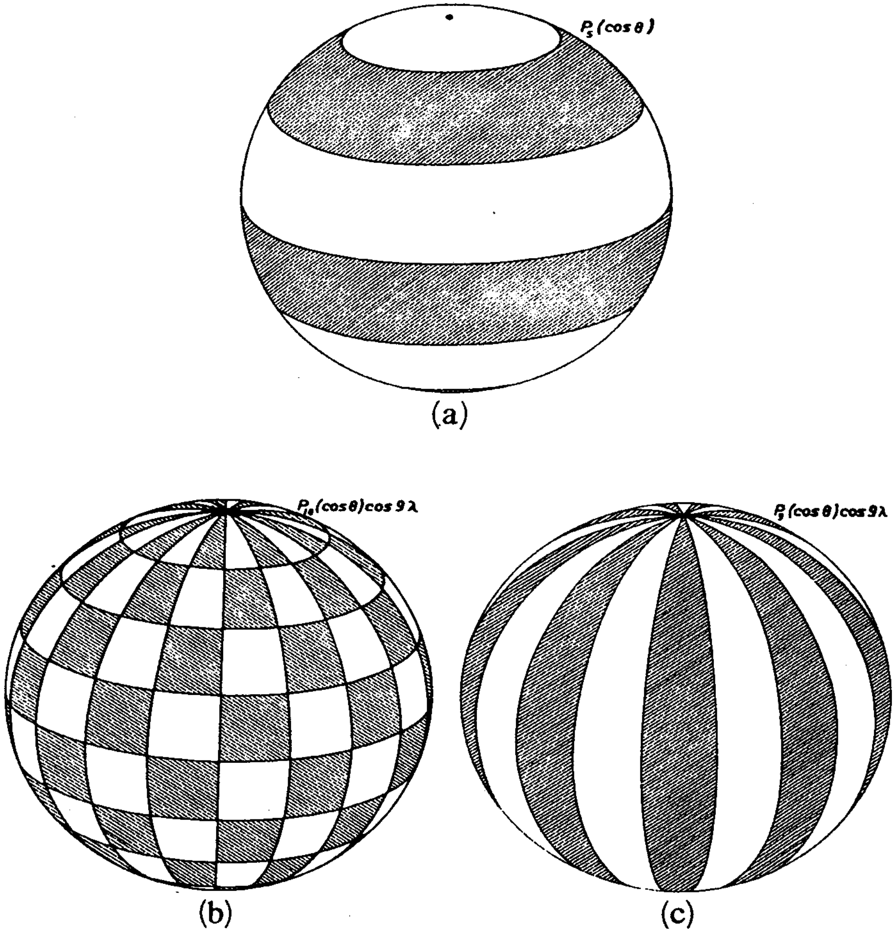

FIGURE 1-9
The different kinds of spherical harmonics: (a) zonal, (b) tesseral, (c) sectorial.

$m \neq 0$ is similar to that of case *b*. They divide the sphere into compartments in which they are alternately positive and negative, somewhat like a chess board, and are called *tesseral harmonics*.[1] In particular, for $n = m$, they degenerate into functions that divide the sphere into positive and negative sectors, in which case they are called *sectorial harmonics* (Fig. 1-9, case *c*).

[1] *Tessera* means a square or rectangle, or also a tile.

1-12. Legendre's Functions of the Second Kind

Legendre's function $P_{nm}(t)$ is not the only solution of Legendre's differential equation (1-56). There is a completely different function which also satisfies this equation. It is called *Legendre's function of the second kind*, of degree n and order m, and denoted by $Q_{nm}(t)$.

Although the $Q_{nm}(t)$ are functions of a completely different nature, they satisfy relationships very similar to those satisfied by the $P_{nm}(t)$.

The "zonal" functions

$$Q_n(t) \equiv Q_{n0}(t)$$

are defined by

$$Q_n(t) = \frac{1}{2} P_n(t) \ln \frac{1+t}{1-t} - \sum_{k=1}^{n} \frac{1}{k} P_{k-1}(t) P_{n-k}(t), \tag{1-63}$$

and the others by

$$Q_{nm}(t) = (1 - t^2)^{m/2} \frac{d^m Q_n(t)}{dt^m}. \tag{1-64}$$

Equation (1-64) is completely analogous to (1-60); furthermore, the functions $Q_n(t)$ satisfy the same recursion formula (1-59) as the functions $P_n(t)$.

If we evaluate the first few Q_n by (1-63) we find

$$\begin{aligned} Q_0(t) &= \frac{1}{2} \ln \frac{1+t}{1-t} = \tanh^{-1} t, \\ Q_1(t) &= \frac{t}{2} \ln \frac{1+t}{1-t} - 1 = t \tanh^{-1} t - 1, \\ Q_2(t) &= \left(\frac{3}{4} t^2 - \frac{1}{4}\right) \ln \frac{1+t}{1-t} - \frac{3}{2} t = \left(\frac{3}{2} t^2 - \frac{1}{2}\right) \tanh^{-1} t - \frac{3}{2} t. \end{aligned} \tag{1-65}$$

These formulas and Fig. 1-10 show that the functions Q_{nm} are really quite different from the functions P_{nm}. From the singularity $\pm\infty$ at $t = \pm 1$ (i.e., $\theta = 0$ or π) we see that it is impossible to substitute $Q_{nm}(\cos\theta)$ for $P_{nm}(\cos\theta)$ if θ means the polar distance, for harmonic functions must be regular.

However, we shall encounter them in the theory of ellipsoidal harmonics (Section 1-20), which is applied to the normal gravity field of the earth (Section 2-7). For this purpose we need Legendre's functions of the second kind as functions of a complex argument. If the argument z is complex we must replace the definition (1-63) by

$$Q_n(z) = \frac{1}{2} P_n(z) \ln \frac{z+1}{z-1} - \sum_{k=1}^{n} \frac{1}{k} P_{k-1}(z) P_{n-k}(z), \tag{1-63'}$$

where Legendre's polynomials $P_n(z)$ are defined by the same formulas as in the case of a real argument t. Thus the only change as compared to (1-63) is the replacement of

$$\frac{1}{2} \ln \frac{1+t}{1-t} = \tanh^{-1} t$$

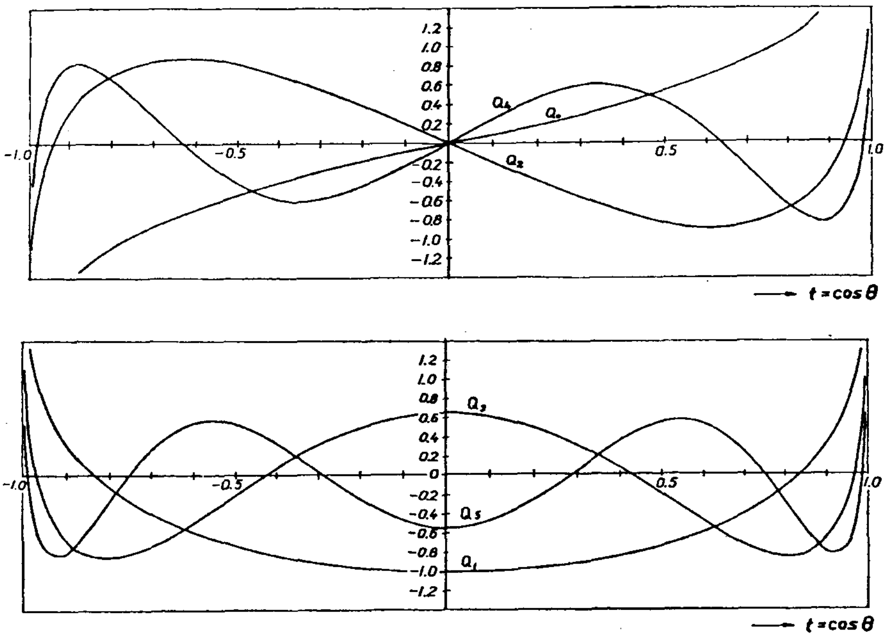

FIGURE 1-10

Legendre's functions of the second kind. Top, n even; bottom, n odd.

by

$$\frac{1}{2}\ln\frac{z+1}{z-1} = \coth^{-1} z.$$

In particular we have

$$Q_0(z) = \frac{1}{2}\ln\frac{z+1}{z-1} = \coth^{-1} z,$$

$$Q_1(z) = \frac{z}{2}\ln\frac{z+1}{z-1} - 1 = z\coth^{-1} z - 1, \tag{1-65'}$$

$$Q_2(z) = \left(\frac{3}{4}z^2 - \frac{1}{4}\right)\ln\frac{z+1}{z-1} - \frac{3}{2}z = \left(\frac{3}{2}z^2 - \frac{1}{2}\right)\coth^{-1} z - \frac{3}{2}z.$$

1-13. Expansion Theorem and Orthogonality Relations

In this section we are concerned with surface spherical harmonics. In (1-54a, b) we have expanded *harmonic* functions in space into series of *solid* spherical harmonics. In a similar way an *arbitrary* (at least in a very general sense) function $f(\theta, \lambda)$ on the surface of the sphere can be expanded into a series of *surface* spherical harmonics:

$$f(\theta, \lambda) = \sum_{n=0}^{\infty} Y_n(\theta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} [a_{nm}R_{nm}(\theta, \lambda) + b_{nm}S_{nm}(\theta, \lambda)], \quad (1\text{-}66)$$

where we have introduced the abbreviations[1]

$$\begin{aligned} R_{nm}(\theta, \lambda) &= P_{nm}(\cos\theta)\cos m\lambda, \\ S_{nm}(\theta, \lambda) &= P_{nm}(\cos\theta)\sin m\lambda. \end{aligned} \quad (1\text{-}67)$$

The symbols a_{nm} and b_{nm} are constant coefficients, which we shall now determine. Essential for this purpose are the so-called *orthogonality relations*. These remarkable relations mean that the integral over the unit sphere of the product of any two *different* functions R_{nm} or S_{nm} is zero:

$$\left.\begin{aligned} \iint_\sigma R_{nm}(\theta, \lambda)R_{sr}(\theta, \lambda)\, d\sigma &= 0 \\ \iint_\sigma S_{nm}(\theta, \lambda)S_{sr}(\theta, \lambda)\, d\sigma &= 0 \end{aligned}\right\} \quad \text{if } s \neq n \text{ or } r \neq m \text{ or both;} \quad (1\text{-}68)$$

$$\iint_\sigma R_{nm}(\theta, \lambda)S_{sr}(\theta, \lambda)\, d\sigma = 0 \qquad \text{in any case.}$$

For the product of two *equal* functions R_{nm} or S_{nm} we have

$$\begin{aligned} &\iint_\sigma [R_{n0}(\theta, \lambda)]^2\, d\sigma = \frac{4\pi}{2n+1}; \\ &\iint_\sigma [R_{nm}(\theta, \lambda)]^2\, d\sigma = \iint_\sigma [S_{nm}(\theta, \lambda)]^2\, d\sigma = \frac{2\pi}{2n+1}\frac{(n+m)!}{(n-m)!} \qquad (m \neq 0). \end{aligned} \quad (1\text{-}69)$$

(There is no S_{n0}, since $\sin 0\lambda = 0$.) In these formulas we have used the abbreviation

$$\iint_\sigma = \int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi}$$

for the integral over the unit sphere. The expression

$$d\sigma = \sin\theta\, d\theta\, d\lambda$$

denotes the surface element of the unit sphere or the element of solid angle, a solid angle being defined as the corresponding area on the unit sphere.

Now the determination of the coefficients a_{nm} and b_{nm} in (1-66) is easy. Multiplying both sides of the equation by a certain $R_{sr}(\theta, \lambda)$ and integrating over the unit sphere gives

$$\iint_\sigma f(\theta, \lambda)R_{sr}(\theta, \lambda)\, d\sigma = a_{sr}\iint_\sigma [R_{sr}(\theta, \lambda)]^2\, d\sigma,$$

[1] We are following MacMillan (1930); he uses the abbreviations $C_{nm}(\theta, \lambda) = P_{nm}(\cos\theta)\cos m\lambda$ and $S_{nm}(\theta, \lambda) = P_{nm}(\cos\theta)\sin m\lambda$.

since in the double integral on the right-hand side all terms except the one with $n = s$, $m = r$ will vanish according to the orthogonality relations (1-68). The integral on the right-hand side has the value given in (1-69), so that a_{sr} is determined. In a similar way we find b_{sr} by multiplying (1-66) by $S_{sr}(\theta, \lambda)$ and integrating over the unit sphere. The result is

$$a_{n0} = \frac{2n+1}{4\pi} \iint_\sigma f(\theta, \lambda) P_n(\cos\theta)\, d\sigma;$$

$$\left.\begin{aligned} a_{nm} &= \frac{2n+1}{2\pi} \frac{(n-m)!}{(n+m)!} \iint_\sigma f(\theta, \lambda) R_{nm}(\theta, \lambda)\, d\sigma \\ b_{nm} &= \frac{2n+1}{2\pi} \frac{(n-m)!}{(n+m)!} \iint_\sigma f(\theta, \lambda) S_{nm}(\theta, \lambda)\, d\sigma \end{aligned}\right\} \quad (m \neq 0). \tag{1-70}$$

The coefficients a_{nm} and b_{nm} can thus be determined by integration.

We note that the Laplace spherical harmonics $Y_n(\theta, \lambda)$ in (1-66) may also be found directly by the formula

$$Y_n(\theta, \lambda) = \frac{2n+1}{4\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} f(\theta', \lambda') P_n(\cos\psi) \sin\theta'\, d\theta'\, d\lambda', \tag{1-71}$$

where ψ is the spherical distance between the points (θ, λ) and (θ', λ'), so that (Fig. 1-11)

$$\cos\psi = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\lambda' - \lambda). \tag{1-72}$$

Equation (1-71) is easily verified by straightforward computation, substituting $P_n(\cos\psi)$ from the decomposition formula (1-82) of Section 1-15.

1-14. Fully Normalized Spherical Harmonics

The formulas of the preceding section for the expansion of a function into a series of surface harmonics are somewhat inconvenient to handle. If we look at equations (1-69) and (1-70) we see that there are different formulas for $m = 0$

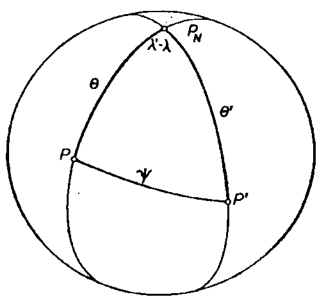

FIGURE 1-11

The spherical distance ψ.

and $m \neq 0$; furthermore, the expressions are rather complicated and difficult to remember.

It has therefore been proposed that the "conventional" harmonics R_{nm} and S_{nm}, defined by (1-67) and (1-57) or (1-62), be replaced by other functions which differ by a constant factor and are easier to handle. We consider here only the *fully normalized*[1] harmonics, which seem to be the most convenient and the most widely used. We denote them by $\bar{R}_{nm}$ and $\bar{S}_{nm}$; they are defined by

$$\begin{gathered}\bar{R}_{n0}(\theta, \lambda) = \sqrt{2n+1}\, R_{n0}(\theta, \lambda) \equiv \sqrt{2n+1}\, P_n(\cos\theta); \\ \left\{\begin{matrix}\bar{R}_{nm}(\theta, \lambda)\\ \bar{S}_{nm}(\theta, \lambda)\end{matrix}\right\} = \sqrt{2(2n+1)\frac{(n-m)!}{(n+m)!}}\left\{\begin{matrix}R_{nm}(\theta, \lambda)\\ S_{nm}(\theta, \lambda)\end{matrix}\right\} \qquad (m \neq 0).\end{gathered} \tag{1-73}$$

The orthogonality relations (1-68) also hold for these fully normalized harmonics, whereas equations (1-69) are thoroughly simplified: they become

$$\frac{1}{4\pi}\iint_\sigma \bar{R}_{nm}^2\, d\sigma = \frac{1}{4\pi}\iint_\sigma \bar{S}_{nm}^2\, d\sigma = 1. \tag{1-74}$$

This means that the *average square of any fully normalized harmonic is unity*, the average being taken over the sphere (average = integral divided by area 4π). This formula now holds for any m, whether it is zero or not.

If we expand an arbitrary function $f(\theta, \lambda)$ into a series of fully normalized harmonics, analogous to (1-66),

$$f(\theta, \lambda) = \sum_{n=0}^{\infty}\sum_{m=0}^{n}[\bar{a}_{nm}\bar{R}_{nm}(\theta, \lambda) + \bar{b}_{nm}\bar{S}_{nm}(\theta, \lambda)], \tag{1-75}$$

then the coefficients $\bar{a}_{nm}$, $\bar{b}_{nm}$ are simply given by

$$\begin{aligned}\bar{a}_{nm} &= \frac{1}{4\pi}\iint_\sigma f(\theta, \lambda)\bar{R}_{nm}(\theta, \lambda)\, d\sigma, \\ \bar{b}_{nm} &= \frac{1}{4\pi}\iint_\sigma f(\theta, \lambda)\bar{S}_{nm}(\theta, \lambda)\, d\sigma;\end{aligned} \tag{1-76}$$

that is, the coefficients are the average products of the function and the corresponding harmonic $\bar{R}_{nm}$ or $\bar{S}_{nm}$.

The simplicity of formulas (1-74) and (1-76) constitutes the main advantage of the fully normalized spherical harmonics, and makes them useful in many respects, even though the functions $\bar{R}_{nm}$ and $\bar{S}_{nm}$ (1-73) are a little more complicated than the conventional R_{nm} and S_{nm}: we have

$$\begin{aligned}\bar{R}_{nm}(\theta, \lambda) &= \bar{P}_{nm}(\cos\theta)\cos m\lambda, \\ \bar{S}_{nm}(\theta, \lambda) &= \bar{P}_{nm}(\cos\theta)\sin m\lambda,\end{aligned}$$

[1] The "fully normalized" harmonics are simply "normalized" in the sense of the theory of real functions; we have to use this clumsy expression because the term "normalized spherical harmonics" has already been used for other functions, unfortunately often for some that are not "normalized" at all in the mathematical sense. A different normalization is used in Jahnke-Emde-Lösch (1960).

where

$$\bar{P}_{n0}(t) = \sqrt{2n+1}\, 2^{-n} \sum_{k=0}^{r} (-1)^k \frac{(2n-2k)!}{k!(n-k)!(n-2k)!} t^{n-2k} \tag{1-77a}$$

for $m = 0$, and

$$\bar{P}_{nm}(t) = \sqrt{2(2n+1)\frac{(n-m)!}{(n+m)!}}\, 2^{-n}(1-t^2)^{m/2}$$

$$\sum_{k=0}^{r} (-1)^k \frac{(2n-2k)!}{k!(n-k)!(n-m-2k)!} t^{n-m-2k} \tag{1-77b}$$

for $m \neq 0$. This corresponds to (1-62); here, as in (1-62), r is the greatest integer $\leq (n-m)/2$.

There are relations between the coefficients $\bar{a}_{nm}$ and $\bar{b}_{nm}$ for fully normalized harmonics and the coefficients a_{nm} and b_{nm} for conventional harmonics that are of course inverse to those in (1-73):

$$\begin{aligned} \bar{a}_{n0} &= \frac{a_{n0}}{\sqrt{2n+1}}; \\ \begin{Bmatrix} \bar{a}_{nm} \\ \bar{b}_{nm} \end{Bmatrix} &= \sqrt{\frac{1}{2(2n+1)}\frac{(n+m)!}{(n-m)!}} \begin{Bmatrix} a_{nm} \\ b_{nm} \end{Bmatrix} \qquad (m \neq 0). \end{aligned} \tag{1-78}$$

1-15. Expansion of the Reciprocal Distance into Zonal Harmonics. Decomposition Formula

The distance l between two points with spherical coordinates

$$P(r, \theta, \lambda), \quad P'(r', \theta', \lambda')$$

is given by

$$l^2 = r^2 + r'^2 - 2rr' \cos\psi, \tag{1-79}$$

where ψ is the angle between the radius vectors r and r' (Fig. 1-12), so that, by (1-72),

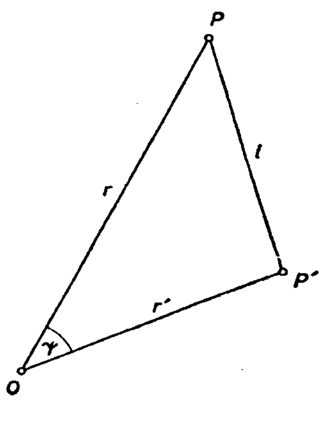

FIGURE 1-12
The spatial distance l.

$$\cos\psi = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\lambda' - \lambda).$$

Assuming $r' < r$ we may write

$$\frac{1}{l} = \frac{1}{\sqrt{r^2 - 2rr'\cos\psi + r'^2}} = \frac{1}{r\sqrt{1 - 2\alpha u + \alpha^2}}$$

where we have put $\alpha = r'/r$ and $u = \cos\psi$. This can be expanded into a power series with respect to α. It is remarkable that the coefficients of α^n are the (conventional) zonal harmonics, or Legendre's polynomials $P_n(u) = P_n(\cos\psi)$:

$$\frac{1}{\sqrt{1 - 2\alpha u + \alpha^2}} = \sum_{n=0}^{\infty} \alpha^n P_n(u) = P_0(u) + \alpha P_1(u) + \alpha^2 P_2(u) + \cdots. \tag{1-80}$$

Hence we obtain

$$\frac{1}{l} = \sum_{n=0}^{\infty} \frac{r'^n}{r^{n+1}} P_n(\cos\psi), \tag{1-81}$$

which is an important formula.

It would still be desirable in this equation to express $P_n(\cos\psi)$ in terms of functions of the spherical coordinates θ, λ and θ', λ' of which ψ is composed according to (1-72). This is achieved by the *decomposition formula*

$$P_n(\cos\psi) = P_n(\cos\theta)P_n(\cos\theta')$$
$$+ 2\sum_{m=1}^{n} \frac{(n-m)!}{(n+m)!}[R_{nm}(\theta,\lambda)R_{nm}(\theta',\lambda') + S_{nm}(\theta,\lambda)S_{nm}(\theta',\lambda')]. \tag{1-82}$$

Substituting this into (1-81) we obtain

$$\frac{1}{l} = \sum_{n=0}^{\infty}\left\{\frac{P_n(\cos\theta)}{r^{n+1}}\cdot r'^n P_n(\cos\theta') + 2\sum_{m=1}^{n}\frac{(n-m)!}{(n+m)!}\right.$$
$$\left.\left[\frac{R_{nm}(\theta,\lambda)}{r^{n+1}}\cdot r'^n R_{nm}(\theta',\lambda') + \frac{S_{nm}(\theta,\lambda)}{r^{n+1}}\cdot r'^n S_{nm}(\theta',\lambda')\right]\right\}. \tag{1-83}$$

The use of fully normalized harmonics simplifies these formulas. Replacing the conventional harmonics in (1-82) and (1-83) by fully normalized harmonics by means of (1-73) we find

$$P_n(\cos\psi) = \frac{1}{2n+1}\sum_{m=0}^{n}[\bar{R}_{nm}(\theta,\lambda)\bar{R}_{nm}(\theta',\lambda') + \bar{S}_{nm}(\theta,\lambda)\bar{S}_{nm}(\theta',\lambda')]; \tag{1-82'}$$

$$\frac{1}{l} = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{1}{2n+1}$$
$$\left[\frac{\bar{R}_{nm}(\theta,\lambda)}{r^{n+1}}\cdot r'^n\bar{R}_{nm}(\theta',\lambda') + \frac{\bar{S}_{nm}(\theta,\lambda)}{r^{n+1}}\cdot r'^n\bar{S}_{nm}(\theta',\lambda')\right]. \tag{1-83'}$$

The last formula will be fundamental for the expansion of the earth's gravitational field in spherical harmonics.

1-16. Solution of Dirichlet's Problem by Means of Spherical Harmonics. Poisson's Integral

In Sec. 1-7 we mentioned *Dirichlet's problem*, or the *first boundary-value problem of potential theory*: given an arbitrary function on a surface S, to determine a function V which is harmonic either inside or outside S and which assumes on S the values of the prescribed function.

If the surface S is a sphere, then Dirichlet's problem can be easily solved by means of spherical harmonics. Let us take first the unit sphere, $r = 1$, and expand the prescribed function, given on the unit sphere and denoted by $V(1, \theta, \lambda)$, into a series of surface harmonics (1-66):

$$V(1, \theta, \lambda) = \sum_{n=0}^{\infty} Y_n(\theta, \lambda), \tag{1-84}$$

the $Y_n(\theta, \lambda)$ being determined by (1-71). The functions

$$V_i(r, \theta, \lambda) = \sum_{n=0}^{\infty} r^n Y_n(\theta, \lambda) \tag{1-85a}$$

and

$$V_e(r, \theta, \lambda) = \sum_{n=0}^{\infty} \frac{Y_n(\theta, \lambda)}{r^{n+1}} \tag{1-85b}$$

assume the given values $V(1, \theta, \lambda)$ on the surface $r = 1$. The series (1-84) converges, and for $r < 1$ we have

$$r^n Y_n < Y_n$$

and for $r > 1$

$$\frac{Y_n}{r^{n+1}} < Y_n.$$

Hence the series (1-85a) converges for $r \leqq 1$, and the series (1-85b) converges for $r \geqq 1$; furthermore, both series have been found to represent harmonic functions. Thus we see that Dirichlet's problem is solved by $V_i(r, \theta, \lambda)$ for the interior of the sphere $r = 1$, and by $V_e(r, \theta, \lambda)$ for its exterior.

For a sphere of arbitrary radius $r = R$ the solution is similar. We expand the given function

$$V(R, \theta, \lambda) = \sum_{n=0}^{\infty} Y_n(\theta, \lambda). \tag{1-86}$$

The surface harmonics Y_n are determined by

$$Y_n(\theta, \lambda) = \frac{2n+1}{4\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} V(R, \theta', \lambda') P_n(\cos\psi) \sin\theta' \, d\theta' \, d\lambda'. \tag{1-86'}$$

Then the series

$$V_i(r, \theta, \lambda) = \sum_{n=0}^{\infty} \left(\frac{r}{R}\right)^n Y_n(\theta, \lambda) \tag{1-87a}$$

solves the first boundary-value problem for the interior, and the series

$$V_e(r, \theta, \lambda) = \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} Y_n(\theta, \lambda) \tag{1-87b}$$

solves it for the exterior of the sphere $r = R$.

Thus we see that Dirichlet's problem can always be solved for the sphere. It is evident that this is closely related to the possibility of expanding an *arbitrary* function on the sphere into a series of *surface* spherical harmonics and a *harmonic* function in space into a series of *solid* spherical harmonics.

Poisson's integral. A more direct solution is obtained as follows. We consider only the exterior problem, which is of greater interest in geodesy. Substituting $Y_n(\theta, \lambda)$ from (1-71) into (1-87b), we obtain

$$V_e(r, \theta, \lambda) = \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} \frac{2n+1}{4\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} V(R, \theta', \lambda') P_n(\cos\psi) \sin\theta' \, d\theta' \, d\lambda'.$$

We can rearrange this as

$$V_e(r, \theta, \lambda) = \frac{1}{4\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} V(R, \theta', \lambda') \left[\sum_{n=0}^{\infty} (2n+1) \left(\frac{R}{r}\right)^{n+1} P_n(\cos\psi)\right] \sin\theta' \, d\theta' \, d\lambda'. \tag{1-88}$$

The sum in the brackets can be evaluated. We denote the spatial distance between the points (r, θ, λ) and (R, θ', λ') by l. Then by (1-81)

$$\frac{1}{l} = \frac{1}{\sqrt{r^2 + R^2 - 2Rr\cos\psi}} = \frac{1}{R} \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} P_n(\cos\psi).$$

Differentiating with respect to r we get

$$-\frac{r - R\cos\psi}{l^3} = -\frac{1}{R} \sum_{n=0}^{\infty} (n+1) \frac{R^{n+1}}{r^{n+2}} P_n(\cos\psi).$$

Multiplying this equation by $-2Rr$, multiplying the expression for $1/l$ by $-R$, and then adding the two equations gives the result

$$\frac{R(r^2 - R^2)}{l^3} = \sum_{n=0}^{\infty} (2n+1) \left(\frac{R}{r}\right)^{n+1} P_n(\cos\psi).$$

The right-hand side is the bracketed expression in (1-88). Substituting the left-hand side, we finally obtain

$$V_e(r, \theta, \lambda) = \frac{R(r^2 - R^2)}{4\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} \frac{V(R, \theta', \lambda')}{l^3} \sin\theta' \, d\theta' \, d\lambda', \tag{1-89}$$

where

$$l = \sqrt{r^2 + R^2 - 2Rr\cos\psi}.$$

This is *Poisson's integral*. It is an explicit solution of Dirichlet's problem for the exterior of the sphere, which has many applications in physical geodesy.

1-17. Other Boundary-value Problems

There are other similar boundary-value problems. In *Neumann's problem*, or the *second boundary-value problem of potential theory*, the normal derivative $\partial V/\partial n$ is given on the surface S, instead of the function V itself. The normal derivative is the derivative along the outward-directed surface normal n to S. In the *third boundary-value problem* a linear combination of V and of its normal derivative

$$hV + k\frac{\partial V}{\partial n}$$

is given on S.

For the sphere the solution of these boundary-value problems is also easily expressed in terms of spherical harmonics. We shall consider the *exterior* problems only, these being of special interest to geodesy.

In *Neumann's problem* we expand the given values of $\partial V/\partial n$ on the sphere $r = R$ into a series of surface harmonics:

$$\left(\frac{\partial V}{\partial n}\right)_{r=R} = \sum_{n=0}^{\infty} Y_n(\theta, \lambda). \tag{1-90}$$

The harmonic function which solves Neumann's problem for the exterior of the sphere is then

$$V_e(r, \theta, \lambda) = -R\sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+1}\frac{Y_n(\theta, \lambda)}{n+1}. \tag{1-91}$$

To verify it we differentiate with respect to r, getting

$$\frac{\partial V_e}{\partial r} = \sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+2} Y_n(\theta, \lambda).$$

Since for the sphere the normal coincides with the radius vector, we have

$$\left(\frac{\partial V}{\partial n}\right)_{r=R} = \left(\frac{\partial V}{\partial r}\right)_{r=R},$$

and we see that (1-90) is satisfied.

The *third boundary-value problem* is particularly relevant to physical geodesy, because the determination of the undulations of the geoid from gravity anomalies is just such a problem. To solve the general case we again expand the function defined by the given boundary values into surface harmonics:

$$hV + k\frac{\partial V}{\partial n} = \sum_{n=0}^{\infty} Y_n(\theta, \lambda).$$

The harmonic function

$$V_e(r, \theta, \lambda) = \sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+1}\frac{Y_n(\theta, \lambda)}{h - (k/R)(n+1)} \tag{1-92}$$

solves the third boundary-value problem for the exterior of the sphere $r = R$. The straightforward verification is completely analogous to the case of (1-91).

In the determination of the geoidal undulations the constants h, k have the values

$$h = -\frac{2}{R}, \quad k = -1,$$

so that

$$V_e(r, \theta, \lambda) = R \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} \frac{Y_n(\theta, \lambda)}{n - 1} \tag{1-92'}$$

solves the so-called *boundary-value problem of physical geodesy*.

As we have seen in the preceding section, the first boundary-value problem can also be solved directly by *Poisson's integral*. Similar integral formulas also exist for the second and third problems. The integral formula that corresponds to (1-92′) for the boundary-value problem of physical geodesy is *Stokes' integral*, which will be considered in detail in Chapter 2.

1-18. The Radial Derivative of a Harmonic Function

For later application to problems related with the vertical gradient of gravity we shall now derive an integral formula for the derivative along the radius vector r of an arbitrary harmonic function which we denote by V. Such a harmonic function satisfies Poisson's integral (1-89):

$$V(r, \theta, \lambda) = \frac{R(r^2 - R^2)}{4\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} \frac{V(R, \theta', \lambda')}{l^3} \sin\theta' \, d\theta' \, d\lambda'.$$

In forming the radial derivative $\partial V/\partial r$ we note that $V(R, \theta', \lambda')$ does not depend on r. Thus we need only to differentiate $(r^2 - R^2)/l^3$, obtaining

$$\frac{\partial V(r, \theta, \lambda)}{\partial r} = \frac{R}{4\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} M(r, \psi) V(R, \theta', \lambda') \sin\theta' \, d\theta' \, d\lambda', \tag{1-93}$$

where

$$M(r, \psi) \equiv \frac{\partial}{\partial r} \frac{r^2 - R^2}{l^3} = \frac{1}{l^5}(5R^2 r - r^3 - Rr^2 \cos\psi - 3R^3 \cos\psi). \tag{1-94}$$

Applying this equation to the special harmonic function

$$V_1(r, \theta, \lambda) = \frac{R}{r}; \quad \frac{\partial V_1}{\partial r} = -\frac{R}{r^2}, \quad V_1(R, \theta', \lambda') = \frac{R}{R} = 1,$$

we obtain

$$-\frac{R}{r^2} = \frac{R}{4\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} M(r, \psi) \sin\theta' \, d\theta' \, d\lambda'.$$

Multiplying both sides of this equation by $V(r, \theta, \lambda)$ and subtracting it from (1-93) gives

$$\frac{\partial V}{\partial r} + \frac{R}{r^2} V_P = \frac{R}{4\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} M(r, \psi)(V - V_P) \sin\theta' \, d\theta' \, d\lambda', \tag{1-95}$$

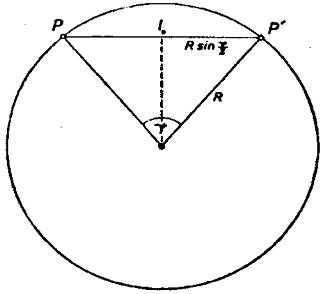

FIGURE 1-13
The spatial distance between two points on a sphere.

where

$$V_P = V(r, \theta, \lambda), \quad V = V(R, \theta', \lambda').$$

In order to find the radial derivative at the surface of the sphere of radius R, we must set $r = R$. Then l becomes (Fig. 1-13)

$$l_0 = 2R \sin \frac{\psi}{2},$$

and the function M takes the simple form

$$M(R, \psi) = \frac{1}{4R^2 \sin^3 \frac{\psi}{2}} = \frac{2R}{l_0^3}. \tag{1-96}$$

For $\psi \longrightarrow 0$ we have $M(R, \psi) \longrightarrow \infty$, and we cannot use the original formula (1-93) at the surface of the sphere $r = R$. In the transformed equation (1-95), however, we have $V - V_P \longrightarrow 0$ for $\psi \longrightarrow 0$, and the singularity of M for $\psi \longrightarrow 0$ will be neutralized.[1] Thus we obtain

$$\frac{\partial V}{\partial r} = -\frac{1}{R} V_P + \frac{R^2}{2\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} \frac{V - V_P}{l_0^3} \sin \theta' \, d\theta' \, d\lambda'. \tag{1-97}$$

This equation expresses $\partial V/\partial r$ on the sphere $r = R$ in terms of V on this sphere; thus we now have

$$V_P = V(R, \theta, \lambda), \quad V = V(R, \theta', \lambda'). \tag{1-98}$$

Solution in terms of spherical harmonics. We may express V_P as

$$V_P = \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} Y_n(\theta, \lambda). \tag{1-99}$$

Differentiation gives

$$\frac{\partial V}{\partial r} = -\sum_{n=0}^{\infty} (n + 1) \frac{R^{n+1}}{r^{n+2}} Y_n(\theta, \lambda).$$

[1] Provided V is differentiable twice at P.

For $r = R$ this becomes

$$\frac{\partial V}{\partial r} = -\frac{1}{R}\sum_{n=0}^{\infty}(n+1)Y_n(\theta, \lambda). \tag{1-100}$$

This is the equivalent of (1-97) in terms of spherical harmonics.

From this equation we get an interesting by-product. Equation (1-100) may be written

$$\frac{\partial V}{\partial r} = -\frac{1}{R}V_P - \frac{1}{R}\sum_{n=0}^{\infty} nY_n(\theta, \lambda).$$

Comparing this with (1-97) we see that if on a sphere of radius R

$$V_P = \sum_{n=0}^{\infty} Y_n(\theta, \lambda), \tag{1-101}$$

then

$$\frac{R^2}{2\pi}\int_{\lambda'=0}^{2\pi}\int_{\theta'=0}^{\pi}\frac{V - V_P}{l_0^3}\sin\theta'\, d\theta'\, d\lambda' = -\frac{1}{R}\sum_{n=0}^{\infty} nY_n(\theta, \lambda). \tag{1-102}$$

This equation is formulated entirely in terms of quantities referred to the spherical surface only. Furthermore, for any function prescribed on the surface of a sphere one can find a function in space that is harmonic outside the sphere and assumes the values of the function prescribed on it. This is done by solving Dirichlet's exterior problem. From these facts we conclude that (1-102) holds for an arbitrary function V defined on the surface of a sphere.

These developments will be used in Secs. 2-23 and 8-8.

1-19. Laplace's Equation in Ellipsoidal Coordinates

Spherical harmonics are most frequently used in geodesy because they are relatively simple and the earth is nearly spherical. Since the earth is more nearly an ellipsoid of revolution it might be expected that ellipsoidal harmonics, which are defined in a way similar to that of the spherical harmonics, would be even more suitable.[1] As they are more complicated, however, they are used only in certain special cases which nevertheless are important, namely, in problems involving rigorous computation of normal gravity.

We introduce *ellipsoidal coordinates* u, θ, λ (Fig. 1-14). In a rectangular system a point P has the coordinates x, y, z. Now we pass through P the surface of an ellipsoid of revolution whose center is the origin O, whose axis coincides with the z-axis, and whose linear eccentricity has the constant value E. The coordinate u is the semiminor axis of this ellipsoid, θ is the complement of the "reduced

[1] The whole matter is a question of mathematical convenience, since both spherical and ellipsoidal harmonics may be used for any attracting body, regardless of its form.

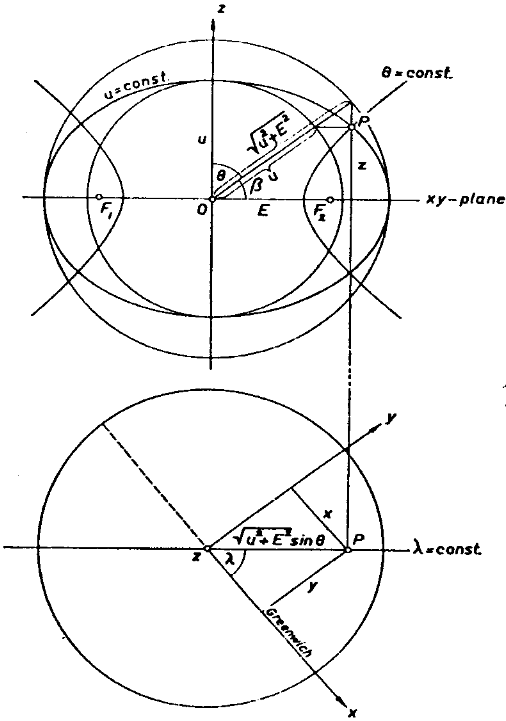

FIGURE 1-14

Ellipsoidal coordinates. Top. View from the front. Bottom. View from above.

latitude" β of P with respect to this ellipsoid (the definition is seen in Fig. 1-14), and λ is the geocentric longitude in the usual sense.[1]

The ellipsoidal coordinates u, θ, λ are related to x, y, z by the equations

$$\begin{aligned} x &= \sqrt{u^2 + E^2} \sin\theta \cos\lambda, \\ y &= \sqrt{u^2 + E^2} \sin\theta \sin\lambda, \\ z &= u \cos\theta, \end{aligned} \tag{1-103}$$

which can be read from the figure, considering that $\sqrt{u^2 + E^2}$ is the semimajor axis of the ellipsoid whose surface passes through P.

If we take u = const. we find

$$\frac{x^2 + y^2}{u^2 + E^2} + \frac{z^2}{u^2} = 1,$$

[1] These coordinates u, θ, λ are specially adapted to an ellipsoid of revolution; they are different from Lamé's ellipsoidal coordinates λ, μ, ν, which refer to an ellipsoid of three different axes. For this reason our ellipsoidal harmonics are different from the ellipsoidal harmonics of Lamé, which are less suited to geodetic problems.

which represents an ellipsoid of revolution. For θ = const. we obtain

$$\frac{x^2 + y^2}{E^2 \sin^2 \theta} - \frac{z^2}{E^2 \cos^2 \theta} = 1,$$

which represents a hyperboloid of one sheet, and for λ = const. we get the meridian plane

$$y = x \tan \lambda.$$

The constant focal length $E = OF_1$, which is the same for *all* ellipsoids u = const., characterizes the coordinate system. For $E = 0$ we have the usual spherical coordinates $u = r, \theta, \lambda$ as a limiting case.

To find the element of arc ds in ellipsoidal coordinates we proceed in the same way as in spherical coordinates, eq. (1-38), and obtain

$$ds^2 = \frac{u^2 + E^2 \cos^2 \theta}{u^2 + E^2} du^2 + (u^2 + E^2 \cos^2 \theta)\, d\theta^2 + (u^2 + E^2) \sin^2 \theta \, d\lambda^2. \quad (1\text{-}104)$$

The coordinate system u, θ, λ is again orthogonal: the products $du\, d\theta$, etc., are missing in the equation for ds. Setting $u = q_1, \theta = q_2, \lambda = q_3$, we have in (1-39)

$$h_1^2 = \frac{u^2 + E^2 \cos^2 \theta}{u^2 + E^2}, \quad h_2^2 = u^2 + E^2 \cos^2 \theta, \quad h_3^2 = (u^2 + E^2) \sin^2 \theta.$$

If we substitute these into (1-40) we obtain

$$\Delta V = \frac{1}{(u^2 + E^2 \cos^2 \theta) \sin \theta} \left\{ \frac{\partial}{\partial u} \left[(u^2 + E^2) \sin \theta \frac{\partial V}{\partial u} \right] + \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial V}{\partial \theta} \right) + \frac{\partial}{\partial \lambda} \left[\frac{u^2 + E^2 \cos^2 \theta}{(u^2 + E^2) \sin \theta} \frac{\partial V}{\partial \lambda} \right] \right\}.$$

Performing the differentiations and cancelling $\sin \theta$ we get

$$\Delta V \equiv \frac{1}{u^2 + E^2 \cos^2 \theta} \left[(u^2 + E^2) \frac{\partial^2 V}{\partial u^2} + 2u \frac{\partial V}{\partial u} + \frac{\partial^2 V}{\partial \theta^2} + \cot \theta \frac{\partial V}{\partial \theta} + \frac{u^2 + E^2 \cos^2 \theta}{(u^2 + E^2) \sin^2 \theta} \frac{\partial^2 V}{\partial \lambda^2} \right] = 0, \quad (1\text{-}105)$$

which is *Laplace's equation in ellipsoidal coordinates.* An alternative expression is obtained by omitting the factor $(u^2 + E^2 \cos^2 \theta)^{-1}$:

$$(u^2 + E^2) \frac{\partial^2 V}{\partial u^2} + 2u \frac{\partial V}{\partial u} + \frac{\partial^2 V}{\partial \theta^2} + \cot \theta \frac{\partial V}{\partial \theta} + \frac{u^2 + E^2 \cos^2 \theta}{(u^2 + E^2) \sin^2 \theta} \frac{\partial^2 V}{\partial \lambda^2} = 0. \quad (1\text{-}105')$$

In the limiting case, $E \longrightarrow 0$, these equations reduce to the spherical expressions (1-41) and (1-41′).

1-20. Ellipsoidal Harmonics

To solve (1-105) or (1-105′) we proceed in a way which is exactly analogous to the method used to solve the corresponding equation (1-41′) in spherical coordinates. What we did there may be summarized as follows. By the trial substitution

$$V(r, \theta, \lambda) = f(r)g(\theta)h(\lambda)$$

we separated the variables r, θ, λ, so that the original partial differential equation (1-41′) was decomposed into three ordinary differential equations (1-43), (1-49), and (1-50).

In order to solve Laplace's equation in ellipsoidal coordinates (1-105′) we correspondingly make the trial substitution

$$V(u, \theta, \lambda) = f(u)g(\theta)h(\lambda). \tag{1-106}$$

Substituting and dividing by fgh we get

$$\frac{1}{f}[(u^2 + E^2)f'' + 2uf'] + \frac{1}{g}(g'' + g' \cot \theta) + \frac{u^2 + E^2 \cos^2 \theta}{(u^2 + E^2)\sin^2 \theta}\frac{h''}{h} = 0.$$

The variable λ occurs only through the quotient h''/h, which consequently must be constant:[1]

$$\frac{h''}{h} = -m^2.$$

The factor by which h''/h is multiplied can be decomposed as follows:

$$\frac{u^2 + E^2 \cos^2 \theta}{(u^2 + E^2)\sin^2 \theta} = \frac{1}{\sin^2 \theta} - \frac{E^2}{u^2 + E^2}.$$

Inserting the last two expressions into the preceding equation and combining functions of the same variable we obtain

$$\frac{1}{f}[(u^2 + E^2)f'' + 2uf'] + \frac{E^2}{u^2 + E^2}m^2 = -\frac{1}{g}(g'' + g' \cot \theta) + \frac{m^2}{\sin^2 \theta}.$$

The two sides are functions of different independent variables and must therefore be constant. Denoting this constant by $n(n + 1)$ we finally get

$$(u^2 + E^2)f''(u) + 2uf'(u) - \left[n(n + 1) - \frac{E^2}{u^2 + E^2}m^2\right]f(u) = 0; \tag{1-107}$$

$$\sin \theta\, g''(\theta) + \cos \theta\, g'(\theta) + \left[n(n + 1)\sin \theta - \frac{m^2}{\sin \theta}\right]g(\theta) = 0; \tag{1-108}$$

$$h''(\lambda) + m^2 h(\lambda) = 0. \tag{1-109}$$

These are the three ordinary differential equations into which the partial differential equation (1-105′) is decomposed by the separation of variables (1-106).

The second and third equations are the same as in the spherical case, equations (1-49) and (1-50); the first equation is different. The substitutions

[1] One sees this more clearly by writing the equation in the form

$$-\frac{(u^2 + E^2)\sin^2 \theta}{u^2 + E^2 \cos^2 \theta}\left\{\frac{1}{f}[(u^2 + E^2)f'' + 2uf'] + \frac{1}{g}(g'' + g' \cot \theta)\right\} = \frac{h''}{h}.$$

The left-hand side depends only on u and θ, the right-hand side only on λ. The two sides cannot be identically equal unless both are equal to the same constant.

$$\tau = i\frac{u}{E} \quad (i = \sqrt{-1}) \quad \text{and} \quad t = \cos\theta$$

transform the first and second equations into

$$(1 - \tau^2)\bar{f}''(\tau) - 2\tau\bar{f}'(\tau) + \left[n(n+1) - \frac{m^2}{1-\tau^2}\right]\bar{f}(\tau) = 0,$$

$$(1 - t^2)\bar{g}''(t) - 2t\bar{g}'(t) + \left[n(n+1) - \frac{m^2}{1-t^2}\right]\bar{g}(t) = 0,$$

where the overbar indicates that the functions f and g are expressed in terms of the new arguments τ and t. From spherical harmonics we are already familiar with the substitution $t = \cos\theta$ and the corresponding equation for $\bar{g}(t)$.

Note that $\bar{f}(\tau)$ satisfies formally the same differential equation as $\bar{g}(t)$, namely, Legendre's equation (1-56). As we have seen, this differential equation has two solutions: Legendre's function P_{nm} and Legendre's function of the second kind Q_{nm}. For $\bar{g}(t)$, where $t = \cos\theta$, the $Q_{nm}(t)$ are ruled out for obvious reasons, as we have seen in Sec. 1-12. For $\bar{f}(\tau)$, however, both sets of functions $P_{nm}(\tau)$ and $Q_{nm}(\tau)$ are possible solutions; they correspond to the two different solutions $f = r^n$ and $f = r^{-(n+1)}$ in the spherical case.

Finally, (1-109) has as before the solutions $\cos m\lambda$ and $\sin m\lambda$.

We summarize all individual solutions:

$$\begin{aligned} f(u) &= P_{nm}\left(i\frac{u}{E}\right) \quad \text{or} \quad Q_{nm}\left(i\frac{u}{E}\right); \\ g(\theta) &= P_{nm}(\cos\theta); \\ h(\lambda) &= \cos m\lambda \quad \text{or} \quad \sin m\lambda. \end{aligned}$$

Here n and $m < n$ are integers $0, 1, 2, \ldots$, as before. Hence the functions

$$\begin{aligned} V(u, \theta, \lambda) &= P_{nm}\left(i\frac{u}{E}\right) P_{nm}(\cos\theta)\begin{Bmatrix}\cos m\lambda \\ \sin m\lambda\end{Bmatrix}, \\ V(u, \theta, \lambda) &= Q_{nm}\left(i\frac{u}{E}\right) P_{nm}(\cos\theta)\begin{Bmatrix}\cos m\lambda \\ \sin m\lambda\end{Bmatrix} \end{aligned} \tag{1-110}$$

are solutions of Laplace's equation $\Delta V = 0$, that is, harmonic functions.

From these functions we may by linear combination form the series

$$V_i(u, \theta, \lambda) = \sum_{n=0}^{\infty}\sum_{m=0}^{n} \frac{P_{nm}\left(i\frac{u}{E}\right)}{P_{nm}\left(i\frac{b}{E}\right)} [a_{nm}P_{nm}(\cos\theta)\cos m\lambda + b_{nm}P_{nm}(\cos\theta)\sin m\lambda]; \tag{1-111a}$$

$$V_e(u, \theta, \lambda) = \sum_{n=0}^{\infty}\sum_{m=0}^{n} \frac{Q_{nm}\left(i\frac{u}{E}\right)}{Q_{nm}\left(i\frac{b}{E}\right)} [a_{nm}P_{nm}(\cos\theta)\cos m\lambda + b_{nm}P_{nm}(\cos\theta)\sin m\lambda]. \tag{1-111b}$$

Here b is the semiminor axis of an arbitrary but fixed ellipsoid which may be called the *reference ellipsoid* (Fig. 1-15). The division by $P_{nm}(ib/E)$ or $Q_{nm}(ib/E)$ is possible because they are constants; its purpose is to simplify the expressions and to make the coefficients a_{nm} and b_{nm} real.

If the eccentricity E reduces to zero, the ellipsoidal coordinates u, θ, λ become spherical coordinates r, θ, λ; the ellipsoid $u = b$ becomes the sphere $r = R$ (because then the semiaxes a and b are equal to the radius R); and we find

$$\lim_{E\to 0} \frac{P_{nm}\left(i\dfrac{u}{E}\right)}{P_{nm}\left(i\dfrac{b}{E}\right)} = \left(\frac{u}{b}\right)^n = \left(\frac{r}{R}\right)^n, \quad \lim_{E\to 0} \frac{Q_{nm}\left(i\dfrac{u}{E}\right)}{Q_{nm}\left(i\dfrac{b}{E}\right)} = \left(\frac{R}{r}\right)^{n+1}, \tag{1-112}$$

so that the series (1-111a) becomes (1-87a), and (1-111b) becomes (1-87b). Thus we see that the function $P_{nm}(iu/E)$ corresponds to r^n and $Q_{nm}(iu/E)$ corresponds to $r^{-(n+1)}$ in spherical harmonics.

Hence the series (1-111a) is harmonic in the interior of the ellipsoid $u = b$, and the series (1-111b) is harmonic in its exterior; this case is relevant to geodesy. For $u = b$, the two series are equal:

$$\begin{aligned} V_i(b, \theta, \lambda) &= V_e(b, \theta, \lambda) \\ &= \sum_{n=0}^{\infty} \sum_{m=0}^{n} [a_{nm} P_{nm}(\cos\theta) \cos m\lambda + b_{nm} P_{nm}(\cos\theta) \sin m\lambda]. \end{aligned} \tag{1-113}$$

Thus the solution of Dirichlet's boundary-value problem for the ellipsoid of revolution is easy. We expand the function $V(b, \theta, \lambda)$, given on the ellipsoid $u = b$, into a series of surface spherical harmonics with the following arguments:

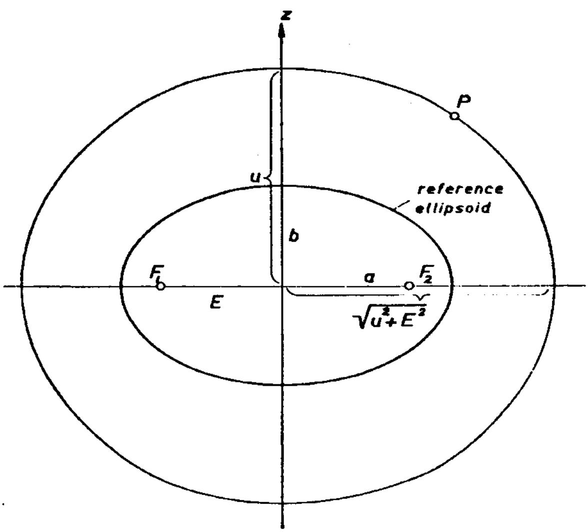

FIGURE 1-15

Reference ellipsoid and ellipsoidal coordinates.

θ = complement of reduced latitude, λ = geocentric longitude. Then (1-111a) is the solution of the interior problem and (1-111b) the solution of the exterior Dirichlet problem.

Formula (1-113) shows that not only functions that are defined on the surface of a sphere can be expanded into a series of surface spherical harmonics. Such an expansion is even possible for rather arbitrary functions defined on a convex surface.

It should be carefully noted that in spherical harmonics θ is the polar distance, which is nothing but the complement of the *geocentric* latitude, whereas in ellipsoidal harmonics θ is the complement of the *reduced* latitude.

References

Byerly, W. E. (1893). An elementary treatise on Fourier's series and spherical, cylindrical, and ellipsoidal harmonics. Boston, Ginn and Co. (New York, Dover Publications, 1959).

Courant, R., and Hilbert, D. (1953-1962). Methods of mathematical physics (2 vols.). New York, Interscience Publishers.

Frank, P., and von Mises, R. (1930-1935). Die Differential- und Integralgleichungen der Mechanik und Physik. Braunschweig, F. Vieweg und Sohn (New York, Dover Publications, 1961).

Hobson, E. W. (1931). The theory of spherical and ellipsoidal harmonics. Cambridge University Press (reprinted 1955).

Jahnke-Emde-Lösch (1960). Tables of higher functions. 6th ed., Stuttgart, B. G. Teubner.

Kellogg, O. D. (1929). Foundations of potential theory. Berlin, J. Springer (New York, Dover Publications, 1953).

MacMillan, W. D. (1930). Theoretical mechanics, vol. 2: The theory of the potential. New York, McGraw-Hill (New York, Dover Publications, 1958).

Wangerin, A. (1921-1922). Theorie des Potentials und der Kugelfunktionen (2 vols.). Berlin, W. de Gruyter & Co.

2

The Gravity Field of the Earth

2-1. Gravity

The force acting on a body at rest on the earth's surface is the resultant of gravitational force and the centrifugal force of the earth's rotation.

Take a rectangular coordinate system whose origin is at the earth's center of gravity and whose z-axis coincides with the earth's mean axis of rotation (Fig. 2-1). The x- and y-axes are so chosen as to obtain a right-handed coordinate

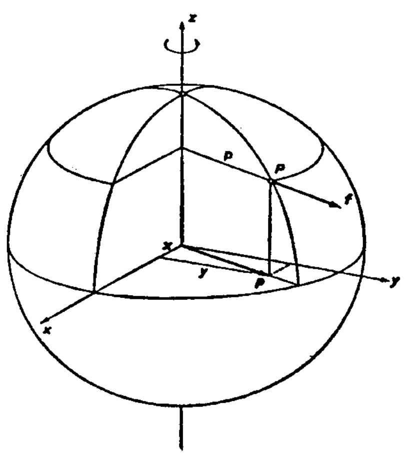

FIGURE 2-1

The centrifugal force.

system; otherwise they are arbitrary. For convenience we may assume an x-axis that is parallel to the meridian plane of Greenwich (see Section 2-4).

The centrifugal force f on a unit mass is given by

$$f = \omega^2 p,$$

where ω is the angular velocity of the earth's rotation and

$$p = \sqrt{x^2 + y^2} \tag{2-1}$$

is the distance from the axis of rotation. The vector $\mathbf{f}$ of this force has the direction of the vector

$$\mathbf{p} = (x, y, 0)$$

and is therefore given by

$$\mathbf{f} = \omega^2 \mathbf{p} = (\omega^2 x, \omega^2 y, 0). \tag{2-2}$$

The centrifugal force can also be derived from a potential

$$\Phi = \frac{1}{2}\omega^2(x^2 + y^2), \tag{2-3}$$

so that

$$\mathbf{f} = \operatorname{grad} \Phi \equiv \left(\frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y}, \frac{\partial \Phi}{\partial z}\right). \tag{2-4}$$

Inserting (2-3) into (2-4) yields (2-2).

The total force, the resultant of gravitational force and centrifugal force, is called *gravity*. The potential of gravity, W, is the sum of the potentials of gravitational force, V (1-11), and centrifugal force, Φ:

$$W = W(x, y, z) = V + \Phi = k \iiint_v \frac{\rho}{l}\, dv + \frac{1}{2}\omega^2(x^2 + y^2), \tag{2-5}$$

where the integration is extended over the earth.

By differentiating (2-3) we find

$$\Delta\Phi \equiv \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = 2\omega^2.$$

If we combine this with Poisson's equation (1-13) for V, we get the *generalized Poisson equation* for the gravity potential W:

$$\Delta W = -4\pi k\rho + 2\omega^2. \tag{2-6}$$

Since Φ is an analytic function, the discontinuities of W are those of V: some second derivatives have jumps at discontinuities of density.

The gradient vector of W,

$$\mathbf{g} = \operatorname{grad} W \equiv \left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}, \frac{\partial W}{\partial z}\right) \tag{2-7}$$

with components

$$g_x = \frac{\partial W}{\partial x} = -k \iiint_v \frac{x - \xi}{l^3} \rho \, dv + \omega^2 x,$$

$$g_y = \frac{\partial W}{\partial y} = -k \iiint_v \frac{y - \eta}{l^3} \rho \, dv + \omega^2 y, \tag{2-8}$$

$$g_z = \frac{\partial W}{\partial z} = -k \iiint_v \frac{z - \zeta}{l^3} \rho \, dv,$$

is called the *gravity vector*; it is the total force (gravitational force plus centrifugal force) acting on a unit mass. As a vector, it has *magnitude* and *direction*.

The magnitude g is called *gravity* in the narrower sense. It has the physical dimension of an acceleration and is measured in gals (1 gal = 1 cm sec^{-2}), the unit being named in honor of Galileo Galilei. The numerical value of g is about 978 gals at the equator, and 983 gals at the poles. In geodesy another unit is often convenient—the milligal, abbreviated mgal (1 mgal = 10^{-3} gal).

The direction of the gravity vector is the direction of the *plumb line*, or the *vertical*; its basic significance for geodetic and astronomic measurements is well known.

In addition to the centrifugal force, another force called the *Coriolis force* acts on a *moving* body. It is proportional to the velocity with respect to the earth, so that it is zero for bodies resting on the earth. Since in geodesy we usually deal with instruments at rest relative to the earth, the Coriolis force plays no part here and need not be considered.

2-2. Level Surfaces and Plumb Lines

The surfaces

$$W(x, y, z) = W_0 = \text{const.}, \tag{2-9}$$

on which the potential W is constant, are called *equipotential surfaces* or *level surfaces*.

Differentiating the gravity potential $W = W(x, y, z)$ we find

$$dW = \frac{\partial W}{\partial x} dx + \frac{\partial W}{\partial y} dy + \frac{\partial W}{\partial z} dz.$$

In vector notation, using the scalar product, this reads

$$dW = \text{grad } W \cdot d\mathbf{x} = \mathbf{g} \cdot d\mathbf{x}, \tag{2-10}$$

where

$$d\mathbf{x} = (dx, dy, dz). \tag{2-11}$$

If the vector $d\mathbf{x}$ is taken along the equipotential surface $W = W_0$, then the potential remains constant and $dW = 0$, so that (2-10) becomes

$$\mathbf{g} \cdot d\mathbf{x} = 0. \tag{2-12}$$

If the scalar product of two vectors is zero, then these vectors are normal to each other. This equation therefore expresses the well-known fact that the *gravity vector is normal to the equipotential surface* passing through the same point.

As the level surfaces are, so to speak, horizontal everywhere, they share the strong intuitive and physical significance of the horizontal; and they share the geodetic importance of the plumb line because they are normal to it. Thus we understand why so much attention is paid to the equipotential surfaces.

The surface of the oceans is, after some slight idealization, part of a certain level surface. This particular equipotential surface was proposed as the "mathematical figure of the earth" by C. F. Gauss, the "Prince of Mathematics," and was later termed the *geoid*. This definition has proved highly suitable, and the geoid is still held by many to be the fundamental surface of physical geodesy.

If we look at equation (2-5) for the gravity potential W, we can see that the equipotential surfaces $W(x, y, z) = W_0$ are rather complicated mathematically. The level surfaces that lie completely outside the earth are at least *analytical* surfaces, although they have no *simple* analytical expression, because the gravity potential W is analytical outside the earth. This is not true of level surfaces that are partly or wholly inside the earth, such as the geoid. They are continuous and "smooth" (i.e., without edges), but they are no longer analytical surfaces; we shall see in the next section that the curvature of the interior level surfaces changes discontinuously with the density.

The lines that intersect all equipotential surfaces normally are not exactly straight but slightly curved (Fig. 2-2). They are called *lines of force*, or *plumb*

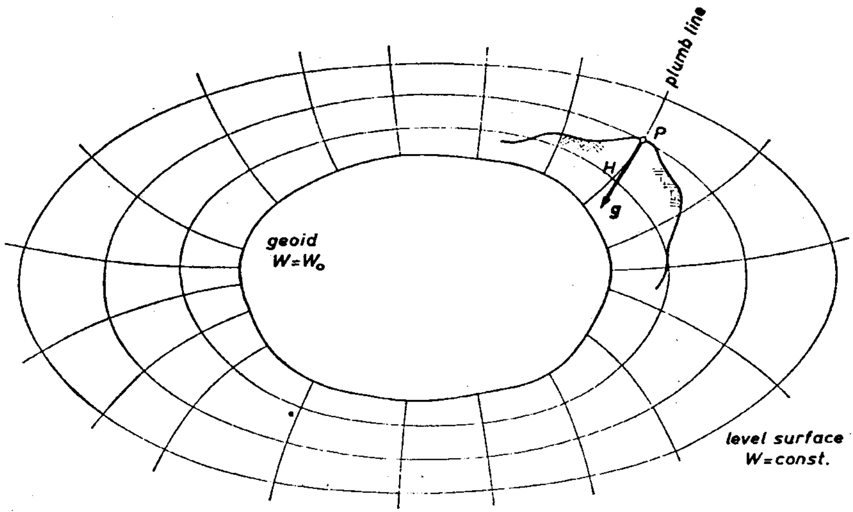

FIGURE 2-2

Level surfaces and plumb lines.

lines. The gravity vector at any point is tangent to the plumb line at that point, hence "direction of the gravity vector," "vertical," and "direction of the plumb line" are synonymous. Sometimes this direction itself is briefly denoted as "plumb line."

The height H of a point above sea level (also called the *orthometric height*) is measured along the curved plumb line, starting from the geoid (Fig. 2-2). If we take the vector $d\mathbf{x}$ along the plumb line, in the direction of increasing height H, then its length will be

$$|d\mathbf{x}| = dH,$$

and its direction is opposite to the gravity vector $\mathbf{g}$, which points downward, so that the angle between $d\mathbf{x}$ and $\mathbf{g}$ is 180°. Since

$$\mathbf{g}\cdot d\mathbf{x} = g\,dH\cos(\mathbf{g}, d\mathbf{x}) = g\,dH\cos 180° = -g\,dH,$$

according to the definition of the scalar product, equation (2-10) becomes

$$dW = -g\,dH. \tag{2-13}$$

This equation relates the height H to the potential W and will be basic for the theory of height determination (Chapter 4). It shows clearly the inseparable interrelation that characterizes geodesy—the interrelation between the geometrical concepts (H) and the dynamic concepts (W).

Another form of equation (2-13) is

$$g = -\frac{\partial W}{\partial H}. \tag{2-14}$$

It shows that gravity is the negative *vertical gradient* of the potential W, or the vertical component of the gradient vector grad W.

Geodetic measurements (theodolite measurements, leveling, etc.) are almost exclusively referred to the system of level surfaces and plumb lines, the geoid playing an essential part. We thus see why the aim of physical geodesy has been formulated as the *determination of the level surfaces of the earth's gravity field*. In a still more abstract but equivalent formulation we may also say that physical geodesy aims at the determination of the potential function $W(x, y, z)$. At first glance the reader is probably perplexed about this definition, which is due to Bruns (1878), but its meaning is easily understood: if the potential W is given as a function of the coordinates x, y, z, then we know all level surfaces including the geoid; they are given by the equation

$$W(x, y, z) = \text{const.}$$

2-3. Curvature of Level Surfaces and Plumb Lines

We recollect the well-known formula for the curvature of a curve $y = f(x)$. It is

$$\kappa = \frac{1}{\rho} = \frac{y''}{(1 + y'^2)^{3/2}},$$

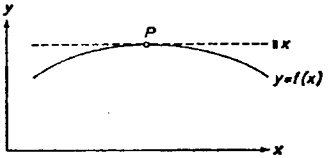

FIGURE 2-3

The curvature of a curve.

where κ is the curvature, ρ is the radius of curvature, and

$$y' = \frac{dy}{dx}, \quad y'' = \frac{d^2y}{dx^2}.$$

For the special case in which a parallel to the x-axis is tangent at the point P under consideration (Fig. 2-3), $y' = 0$, we get simply

$$\kappa = \frac{1}{\rho} = \frac{d^2y}{dx^2}. \tag{2-15}$$

Level surfaces. Consider now a point P on a level surface S. Take a local coordinate system xyz with origin at P whose z-axis is vertical, that is, normal to the surface S (Fig. 2-4). We intersect this level surface

$$W(x, y, z) = W_0$$

with the xz-plane by setting

$$y = 0.$$

Comparing Fig. 2-4 with Fig. 2-3 we see that z now takes the place of y. Therefore, instead of (2-15) we have for the curvature of the intersection of the level surface with the xz-plane:

$$K_1 = \frac{d^2z}{dx^2}. \tag{2-16}$$

If we differentiate $W(x, y, z) = W_0$ with respect to x, considering that y is zero and z is a function of x, we get

$$W_x + W_z \frac{dz}{dx} = 0,$$

$$W_{xx} + 2W_{xz}\frac{dz}{dx} + W_{zz}\left(\frac{dz}{dx}\right)^2 + W_z \frac{d^2z}{dx^2} = 0,$$

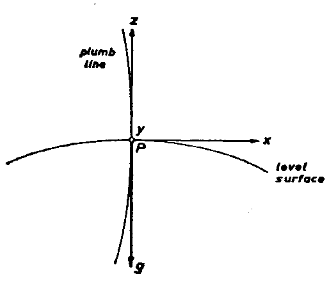

FIGURE 2-4

The local coordinate system.

where the subscripts denote partial differentiation:

$$W_x = \frac{\partial W}{\partial x}, \quad W_{xz} = \frac{\partial^2 W}{\partial x\, \partial z}, \quad \ldots .$$

Since the x-axis is tangent at P, then $dz/dx = 0$ at P, so that

$$\frac{d^2 z}{dx^2} = -\frac{W_{xx}}{W_z}.$$

Since the z-axis is vertical, we have, from (2-14),

$$W_z = \frac{\partial W}{\partial z} = \frac{\partial W}{\partial H} = -g.$$

Thus (2-16) becomes

$$K_1 = \frac{W_{xx}}{g}. \tag{2-17}$$

The curvature of the intersection of the level surface with the yz-plane is found by replacing x with y:

$$K_2 = \frac{W_{yy}}{g}. \tag{2-18}$$

The mean curvature J of a surface at a point P is defined as the arithmetical mean of the curvatures of the curves in which two mutually perpendicular planes through the surface normal intersect the surface (Fig. 2-5). Hence we find

$$J = -\frac{1}{2}(K_1 + K_2) = -\frac{W_{xx} + W_{yy}}{2g}. \tag{2-19}$$

Here the minus sign is only a convention. This is an expression for the *mean curvature of the level surface.*

By the generalized Poisson equation

$$\Delta W \equiv W_{xx} + W_{yy} + W_{zz} = -4\pi k\rho + 2\omega^2$$

we find

$$-2gJ + W_{zz} = -4\pi k\rho + 2\omega^2.$$

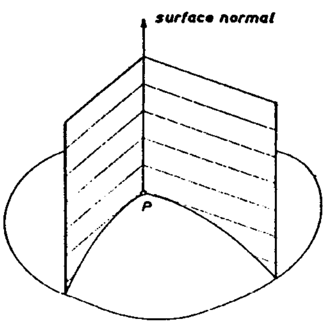

FIGURE 2-5
Definition of mean curvature.

Considering

$$W_z = -g, \quad W_{zz} = -\frac{\partial g}{\partial z} = -\frac{\partial g}{\partial H},$$

we finally obtain

$$\frac{\partial g}{\partial H} = -2gJ + 4\pi k\rho - 2\omega^2. \tag{2-20}$$

This important equation, which relates the *vertical gradient of gravity* $\partial g/\partial H$ to the mean curvature of the level surface, is also due to Bruns (1878). It is another beautiful example of the interrelation between the geometrical and dynamic concepts in geodesy.

Plumb lines. The curvature of the plumb line is needed for the reduction of astronomical observations to the geoid.

A plumb line may be defined as a curve whose line element vector

$$d\mathbf{x} = (dx, dy, dz)$$

has the direction of the gravity vector

$$\mathbf{g} = (W_x, W_y, W_z);$$

that is, $d\mathbf{x}$ and $\mathbf{g}$ differ only by a proportionality factor. This is best expressed in the form

$$\frac{dx}{W_x} = \frac{dy}{W_y} = \frac{dz}{W_z}. \tag{2-21}$$

In the coordinate system of Fig. 2-4 the curvature of the projection of the plumb line onto the xz-plane is given by

$$\kappa_1 = \frac{d^2x}{dz^2};$$

this is equation (2-15) applied to the present case. By (2-21) we have

$$\frac{dx}{dz} = \frac{W_x}{W_z}.$$

We differentiate with respect to z, considering that $y = 0$:

$$\frac{d^2x}{dz^2} = \frac{1}{W_z^2}\left[W_z\left(W_{xz} + W_{xx}\frac{dx}{dz}\right) - W_x\left(W_{zz} + W_{zx}\frac{dx}{dz}\right)\right].$$

In our particular coordinate system the gravity vector coincides with the z-axis, so that its x- and y-components are zero:

$$W_x = W_y = 0.$$

Fig. 2-4 shows that we also have

$$\frac{dx}{dz} = 0.$$

Therefore

$$\frac{d^2x}{dz^2} = \frac{W_zW_{xz}}{W_z^2} = \frac{W_{xz}}{W_z} = \frac{W_{zx}}{W_z}.$$

Considering $W_z = -g$, we finally obtain

$$\kappa_1 = \frac{1}{g}\frac{\partial g}{\partial x}, \tag{2-22a}$$

and similarly,

$$\kappa_2 = \frac{1}{g}\frac{\partial g}{\partial y}. \tag{2-22b}$$

These are the curvatures of the projections of the plumb line onto the xz- and yz-plane, the z-axis being vertical, that is, coinciding with the gravity vector.

The total curvature κ of the plumb line is given, according to differential geometry, by

$$\kappa = \sqrt{\kappa_1^2 + \kappa_2^2} = \frac{1}{g}\sqrt{g_x^2 + g_y^2}. \tag{2-23}$$

For reducing astronomical observations (Sec. 5-6) we shall need only the projection curvatures (2-22a, b).

We mention finally that the various formulas for the curvature of level surfaces and plumb lines are equivalent to the single vector equation

$$\operatorname{grad} g = (-2gJ + 4\pi k\rho - 2\omega^2)\mathbf{n} + g\kappa\mathbf{n}_1, \tag{2-24}$$

where $\mathbf{n}$ is the unit vector along the plumb line (its unit tangent vector) and $\mathbf{n}_1$ is the unit vector along the principal normal to the plumb line.

This may be easily verified. In the local xyz-system used we have

$$\mathbf{n} = (0, 0, 1),$$
$$\mathbf{n}_1 = (\cos\alpha, \sin\alpha, 0),$$

where α is the angle between the principal normal and the x-axis (Fig. 2-6). The z-component of (2-24) yields Bruns' equation (2-20), and the horizontal components yield

$$\frac{\partial g}{\partial x} = g\kappa\cos\alpha, \quad \frac{\partial g}{\partial y} = g\kappa\sin\alpha.$$

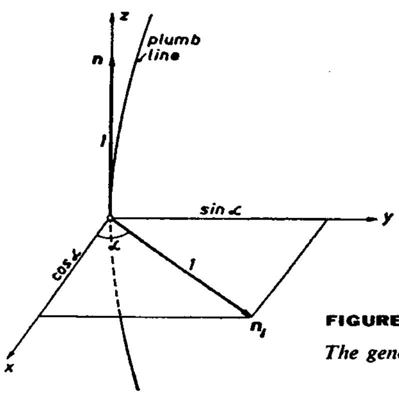

FIGURE 2-6

The generalized Bruns equation.

These are identical to (2-22a, b), since $\kappa_1 = \kappa \cos \alpha$ and $\kappa_2 = \kappa \sin \alpha$, as differential geometry shows. Equation (2-24) is called the *generalized Bruns equation*.

More about the curvature properties and the "inner geometry" of the gravitational field will be found in papers by Marussi (1949) and Hotine (1957).

2-4. Natural Coordinates

The system of level surfaces and plumb lines may be used as a three-dimensional curvilinear coordinate system that is well suited to certain purposes; these coordinates can be measured directly, as opposed to rectangular coordinates x, y, z.

The direction of the earth's axis of rotation and the position of the equatorial plane (normal to the axis) are well defined astronomically. The *geographical latitude* Φ of a point P is the angle between the vertical (direction of the plumb line) at P and the equatorial plane (Fig. 2-7). Consider now a straight line

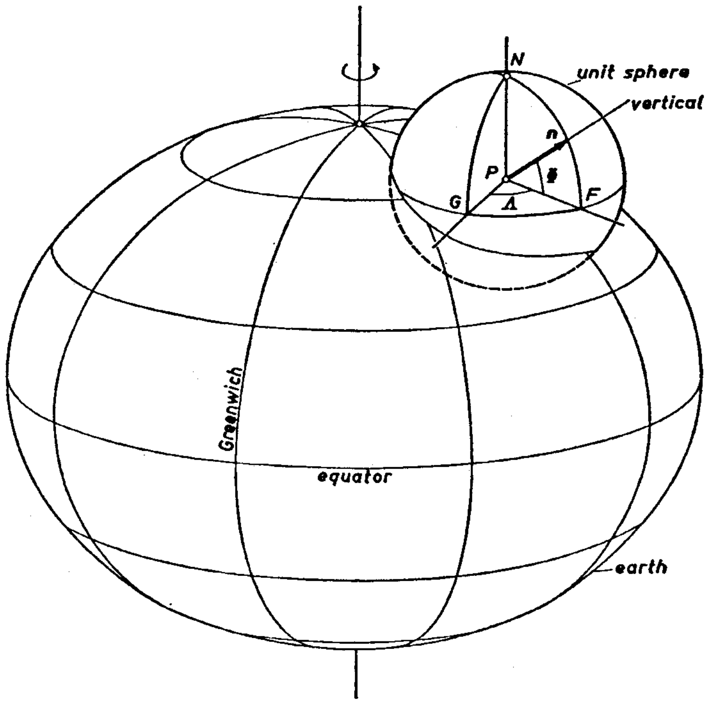

FIGURE 2-7

Definition of the geographical coordinates Φ and Λ of P by means of a unit sphere with center at P. Line PN parallel to rotation axis, plane GPF normal to it, that is, parallel to equatorial plane; $\mathbf{n}$ is the unit vector along the plumb line; plane NPF is the meridian plane of P, and plane NPG is parallel to meridian plane of Greenwich.

through P parallel to the earth's axis. This parallel and the vertical at P together define the meridian plane of P. The angle between this meridian plane and the meridian plane of Greenwich (or some other fixed plane) is the *geographical longitude* Λ of P.

The geographical coordinates, latitude Φ and longitude Λ, form two of the three spatial coordinates of P. As third coordinate we may take the orthometric height H of P or its potential W. Equivalent to W is the *geopotential number* $C = W_0 - W$, where W_0 is the potential of the geoid. The orthometric height H was defined in Sec. 2-2; see also Fig. 2-2. The relations between W, C, and H are given by the equations

$$W = W_0 - \int_0^H g\,dH = W_0 - C,$$

$$C = W_0 - W = \int_0^H g\,dH, \tag{2-25}$$

$$H = -\int_{W_0}^{W} \frac{dW}{g} = \int_0^C \frac{dC}{g},$$

which follow from integrating (2-13). The integral is taken along the plumb line of point P, starting from the geoid ($H = 0$, $W = W_0$); see Fig. 2-8.

The quantities

$$\Phi, \Lambda, W \quad \text{or} \quad \Phi, \Lambda, H$$

are called *natural coordinates*.

They are related in the following way to the geocentric rectangular coordinates x, y, z of Sec. 2-1, the x-axis being parallel to the Greenwich meridian plane. From Fig. 2-7 we read that the unit vector of the vertical $\mathbf{n}$ has the xyz-components

$$\mathbf{n} = (\cos\Phi\cos\Lambda, \cos\Phi\sin\Lambda, \sin\Phi); \tag{2-26}$$

the gravity vector $\mathbf{g}$ is known to be

$$\mathbf{g} = (W_x, W_y, W_z). \tag{2-27}$$

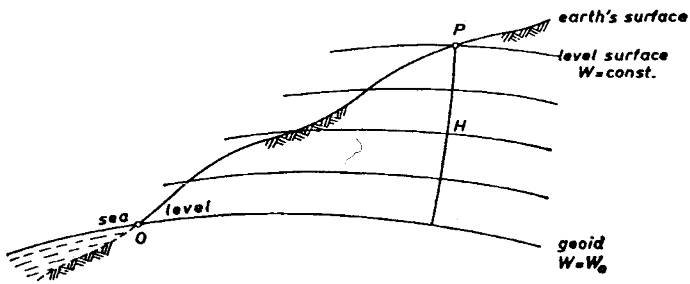

FIGURE 2-8

The orthometric height H.

On the other hand, since **n** is the unit vector corresponding to **g** but of opposite direction, it is given by

$$\mathbf{n} = -\frac{\mathbf{g}}{|\mathbf{g}|} = -\frac{\mathbf{g}}{g},$$

so that

$$\mathbf{g} = -g\mathbf{n}.$$

This equation, together with (2-26) and (2-27), gives

$$\begin{aligned} -W_x &= g \cos \Phi \cos \Lambda, \\ -W_y &= g \cos \Phi \sin \Lambda, \\ -W_z &= g \sin \Phi. \end{aligned} \tag{2-28}$$

Solving for Φ and Λ we finally obtain

$$\begin{aligned} \Phi &= \tan^{-1} \frac{-W_z}{\sqrt{W_x^2 + W_y^2}}, \\ \Lambda &= \tan^{-1} \frac{W_y}{W_x}, \\ W &= W(x, y, z). \end{aligned} \tag{2-29}$$

These three equations relate the natural coordinates Φ, Λ, W to the rectangular coordinates x, y, z, provided the function $W = W(x, y, z)$ is known.

We see that Φ, Λ, H are related to x, y, z in a considerably more complicated way than the spherical coordinates r, θ, λ of Sec. 1-8. Note also the conceptual difference between the geographical longitude Λ and the geocentric longitude λ.

2-5. The Potential of the Earth in Terms of Spherical Harmonics

If we look at the expression (2-5) for the gravity potential W, we see that the part most difficult to handle is the gravitational potential V, the centrifugal potential being a simple analytic function.

The gravitational potential V can be made more manageable for many purposes if we keep in mind the fact that outside the attracting masses it is a harmonic function and can therefore be expanded into a series of spherical harmonics.

We shall now evaluate the coefficients of this series. The gravitational potential V is given by the basic equation (1-11):

$$V = k \iiint_{\text{earth}} \frac{dM}{l}, \tag{2-30}$$

where we now denote the mass element by dM; the integral is extended over the entire earth. Into this integral we insert the expression (1-81):

$$\frac{1}{l} = \sum_{n=0}^{\infty} \frac{r'^n}{r^{n+1}} P_n(\cos \psi),$$

where the P_n are the conventional Legendre polynomials, r is the radius vector of the fixed point P at which V is to be determined, r' is the radius vector of the variable mass element dM, and ψ is the angle between r and r' (Fig. 2-9).

Since r is a constant with respect to the integration over the earth, it can be taken out of the integral. Thus we get

$$V = \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} k \iiint_{\text{earth}} r'^n P_n(\cos\psi)\, dM.$$

If we write this in the usual form as a series of solid spherical harmonics,

$$V = \sum_{n=0}^{\infty} \frac{Y_n(\theta, \lambda)}{r^{n+1}}, \tag{2-31}$$

we see by comparison that the Laplace surface spherical harmonic $Y_n(\theta, \lambda)$ is given by

$$Y_n(\theta, \lambda) = k \iiint_{\text{earth}} r'^n P_n(\cos\psi)\, dM, \tag{2-32}$$

the dependence on θ and λ being effected through the angle ψ since

$$\cos\psi = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\lambda' - \lambda). \tag{2-33}$$

The spherical coordinates θ, λ have been defined in Sec. 1-8.

A more explicit form is obtained by using the decomposition formula (1-83'):

$$\frac{1}{l} = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \frac{1}{2n+1} \left[\frac{\bar{R}_{nm}(\theta, \lambda)}{r^{n+1}} \cdot r'^n \bar{R}_{nm}(\theta', \lambda') + \frac{\bar{S}_{nm}(\theta, \lambda)}{r^{n+1}} \cdot r'^n \bar{S}_{nm}(\theta', \lambda') \right].$$

If we insert this into the integral (2-30) and take the terms that depend on r, θ, λ outside of the integral, we obtain

$$V = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \left[\bar{A}_{nm} \frac{\bar{R}_{nm}(\theta, \lambda)}{r^{n+1}} + \bar{B}_{nm} \frac{\bar{S}_{nm}(\theta, \lambda)}{r^{n+1}} \right], \tag{2-34}$$

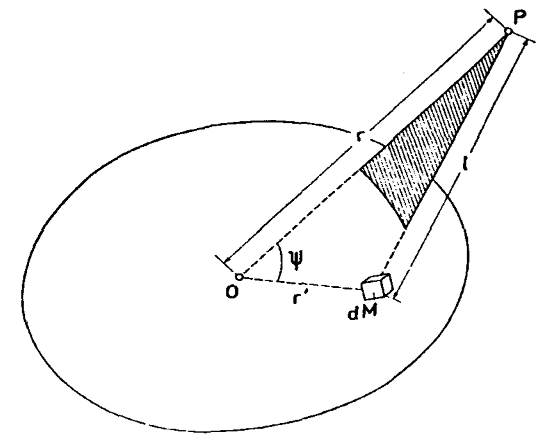

FIGURE 2-9

Expansion into spherical harmonics.

where the constant coefficients $\bar{A}_{nm}$ and $\bar{B}_{nm}$ are given by

$$\begin{aligned}(2n+1)\bar{A}_{nm} &= k \iiint_{\text{earth}} r'^n \bar{R}_{nm}(\theta', \lambda')\, dM,\\ (2n+1)\bar{B}_{nm} &= k \iiint_{\text{earth}} r'^n \bar{S}_{nm}(\theta', \lambda')\, dM.\end{aligned} \tag{2-35}$$

These formulas are very symmetrical and easy to remember: the coefficient, multiplied by $2n + 1$, of the solid harmonic

$$\frac{\bar{R}_{nm}(\theta, \lambda)}{r^{n+1}}$$

is the integral of the solid harmonic

$$r'^n \bar{R}_{nm}(\theta', \lambda').$$

A similar relation holds for $\bar{S}_{nm}$.

Since the mass element is

$$dM = \rho\, dx'\, dy'\, dz' = \rho r'^2 \sin\theta'\, dr'\, d\theta'\, d\lambda', \tag{2-36}$$

the actual evaluation of the integrals requires that the density ρ be expressed as a function of r', θ', λ'. Although no such expression is available at present, this fact does not diminish the theoretical and practical significance of spherical harmonics, since the coefficients A_{nm}, B_{nm} can be determined from the boundary values of gravity at the earth's surface. This is a boundary-value problem that is related to ideas developed in Secs. 1-16 and 1-17 and will be elaborated upon later.

If we recall the relations (1-73) and (1-78) between conventional and fully normalized spherical harmonics, we can also write equations (2-34) and (2-35) in terms of conventional harmonics, readily obtaining

$$V = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \left[A_{nm} \frac{R_{nm}(\theta, \lambda)}{r^{n+1}} + B_{nm} \frac{S_{nm}(\theta, \lambda)}{r^{n+1}} \right], \tag{2-37}$$

where

$$A_{n0} = k \iiint_{\text{earth}} r'^n P_n(\cos\theta')\, dM;$$

$$\left.\begin{aligned} A_{nm} &= 2\frac{(n-m)!}{(n+m)!} k \iiint_{\text{earth}} r'^n R_{nm}(\theta', \lambda')\, dM \\ B_{nm} &= 2\frac{(n-m)!}{(n+m)!} k \iiint_{\text{earth}} r'^n S_{nm}(\theta', \lambda')\, dM \end{aligned}\right\} \quad (m \neq 0). \tag{2-38}$$

These formulas are not as symmetrical as the corresponding formulas (2-35).

In connection with satellite dynamics, the potential V is often written in the form

$$V = \frac{kM}{r} \left\{ 1 - \sum_{n=1}^{\infty} \sum_{m=0}^{n} \left(\frac{a}{r}\right)^n [J_{nm} R_{nm}(\theta, \lambda) + K_{nm} S_{nm}(\theta, \lambda)] \right\}, \tag{2-39}$$

where a is the equatorial radius of the earth, so that

$$\left.\begin{aligned} A_{nm} &= -kMa^n J_{nm} \\ B_{nm} &= -kMa^n K_{nm} \end{aligned}\right\} \qquad (n \neq 0). \tag{2-40}$$

The corresponding fully normalized coefficients

$$\begin{aligned} \bar{J}_{n0} &= \frac{1}{\sqrt{2n+1}} J_{n0}, \\ \left\{\begin{matrix} \bar{J}_{nm} \\ \bar{K}_{nm} \end{matrix}\right\} &= \sqrt{\frac{(n+m)!}{2(2n+1)(n-m)!}} \left\{\begin{matrix} J_{nm} \\ K_{nm} \end{matrix}\right\} \qquad (m \neq 0) \end{aligned} \tag{2-41}$$

are also used.

It is obvious that the nonzonal terms ($m \neq 0$) would be missing in all these expansions if the earth had complete rotational symmetry, since the terms mentioned depend on the longitude λ. In rotationally symmetrical bodies there is no dependence on λ because all longitudes are equivalent. The tesseral and sectorial harmonics will be small, however, since the departures from rotational symmetry are slight.

Finally, we discuss the convergence of (2-34), or of the equivalent series expansions, of the earth's potential. This series is an expansion in powers of $1/r$. Therefore, the larger r is, the better the convergence. For smaller r it is not necessarily convergent. For an arbitrary body, the expansion of V in spherical harmonics can be shown to converge always outside the smallest sphere $r = r_0$ that completely encloses the body (Fig. 2-10). Inside this sphere the series is usually divergent. In certain cases it can converge partly inside the sphere $r = r_0$. If the earth were a homogeneous ellipsoid of about the same dimensions, then the series for V would indeed still converge at the surface of the earth. Owing to the mass irregularities, however, the series of the actual potential V of the earth must be considered *divergent at the surface of the earth* (Moritz, 1961). This impairs the practical significance of the spherical-harmonic expansion of V for terrestrial geodesy; however, besides its theoretical value it has great practical use in satellite dynamics.

It need hardly be pointed out that the spherical-harmonic expansion, always expressing a harmonic function, can represent only the potential *outside the attracting masses*, never inside.

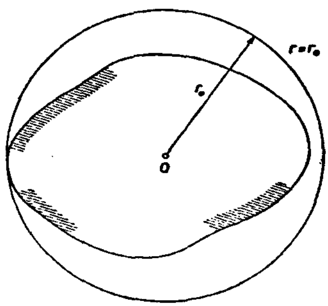

FIGURE 2-10

The spherical-harmonic expansion of V converges outside the sphere $r = r_0$.

2-6. Harmonics of Lower Degree

It is instructive to evaluate the coefficients of the first few spherical harmonics explicitly.

For ready reference we first state some conventional harmonic functions R_{nm} and S_{nm}, using (1-58) and (1-61):

$$\begin{aligned} R_{00} &= 1, & S_{00} &= 0, \\ R_{10} &= \cos\theta, & S_{10} &= 0, \\ R_{11} &= \sin\theta\cos\lambda, & S_{11} &= \sin\theta\sin\lambda, \\ R_{20} &= \tfrac{3}{2}\cos^2\theta - \tfrac{1}{2}, & S_{20} &= 0, \\ R_{21} &= 3\sin\theta\cos\theta\cos\lambda, & S_{21} &= 3\sin\theta\cos\theta\sin\lambda, \\ R_{22} &= 3\sin^2\theta\cos 2\lambda, & S_{22} &= 3\sin^2\theta\sin 2\lambda. \end{aligned} \tag{2-42}$$

The corresponding solid harmonics $r^n R_{nm}$ and $r^n S_{nm}$ are simply homogeneous polynomials in x, y, z. For instance,

$$r^2 S_{22} = 6r^2\sin^2\theta\sin\lambda\cos\lambda = 6(r\sin\theta\cos\lambda)(r\sin\theta\sin\lambda) = 6xy.$$

In this way we find

$$\begin{aligned} R_{00} &= 1, & S_{00} &= 0, \\ rR_{10} &= z, & rS_{10} &= 0, \\ rR_{11} &= x, & rS_{11} &= y, \\ r^2R_{20} &= -\tfrac{1}{2}x^2 - \tfrac{1}{2}y^2 + z^2, & r^2S_{20} &= 0, \\ r^2R_{21} &= 3xz, & r^2S_{21} &= 3yz, \\ r^2R_{22} &= 3x^2 - 3y^2, & r^2S_{22} &= 6xy. \end{aligned} \tag{2-43}$$

Substituting these functions into the expression (2-38) for the coefficients A_{nm} and B_{nm} yields for the zero-degree term

$$A_{00} = k\iiint_{\text{earth}} dM = kM, \tag{2-44a}$$

that is, the product of the mass of the earth times the gravitational constant. For the first-degree coefficients we get

$$A_{10} = k\iiint_{\text{earth}} z'\,dM, \quad A_{11} = k\iiint_{\text{earth}} x'\,dM, \quad B_{11} = k\iiint_{\text{earth}} y'\,dM; \tag{2-44b}$$

and for the second-degree coefficients,

$$\begin{aligned} A_{20} &= \frac{1}{2}k\iiint_{\text{earth}} (-x'^2 - y'^2 + 2z'^2)\,dM, \\ A_{21} &= k\iiint_{\text{earth}} x'z'\,dM, \quad B_{21} = k\iiint_{\text{earth}} y'z'\,dM, \\ A_{22} &= \frac{1}{4}k\iiint_{\text{earth}} (x'^2 - y'^2)\,dM, \quad B_{22} = \frac{1}{2}k\iiint_{\text{earth}} x'y'\,dM. \end{aligned} \tag{2-44c}$$

It is known from mechanics that

$$\xi = \frac{1}{M}\iiint x'\,dM, \quad \eta = \frac{1}{M}\iiint y'\,dM, \quad \zeta = \frac{1}{M}\iiint z'\,dM \tag{2-45}$$

are the rectangular coordinates of the center of gravity. If the origin of the coordinate system coincides with the center of gravity, then these coordinates and hence the integrals (2-44b) are zero. *If the origin $r = 0$ is the center of gravity of the earth, then there will be no first-degree terms in the spherical-harmonic expansion of the potential V.* This is therefore true for our geocentric coordinate system.

The integrals

$$\iiint x'y'\,dM, \quad \iiint y'z'\,dM, \quad \iiint z'x'\,dM$$

are the *products of inertia.* They are zero if the coordinate axes coincide with the principal axes of inertia. Since the z-axis is identical with the mean rotational axis of the earth, which coincides with the axis of maximum inertia, at least the second and third of these products of inertia must vanish. Hence A_{21} and B_{21} will be zero, but not so B_{22}, which is proportional to the first product of inertia; B_{22} would vanish only if the earth had complete rotational symmetry or if a principal axis of inertia happened to fall on the Greenwich meridian.

The five harmonics $A_{10}R_{10}$, $A_{11}R_{11}$, $B_{11}S_{11}$, $A_{21}R_{21}$, and $B_{21}S_{21}$—all first degree harmonics and those of degree 2 and order 1—which must thus vanish in any spherical-harmonic expansion of the earth's potential, are called *forbidden* or *inadmissible harmonics.*

Introducing the *moments of inertia* with respect to the x-, y-, z-axes by the well-known definitions

$$\begin{aligned} A &= \iiint (y'^2 + z'^2)\,dM, \\ B &= \iiint (z'^2 + x'^2)\,dM, \\ C &= \iiint (x'^2 + y'^2)\,dM, \end{aligned} \tag{2-46a}$$

and denoting the xy-product of inertia, which cannot be said to vanish, by

$$D = \iiint x'y'\,dM, \tag{2-46b}$$

we finally have

$$\begin{aligned} A_{00} &= kM, \\ A_{10} &= A_{11} = B_{11} = 0, \\ A_{20} &= k\left(\frac{A+B}{2} - C\right), \\ A_{21} &= B_{21} = 0, \\ A_{22} &= \frac{1}{4}k(B - A), \\ B_{22} &= \frac{1}{2}kD. \end{aligned} \tag{2-47}$$

Now let the x- and y-axes actually coincide with the corresponding principal axes of inertia of the earth. (This is only theoretically possible, since the principal axes of inertia of the earth are only inaccurately known as yet.) Then $B_{22} = 0$, and taking into account (2-42) we may write explicitly

$$V = \frac{kM}{r} + \frac{k}{r^3}\left[\frac{1}{2}\left(C - \frac{A+B}{2}\right)(1 - 3\cos^2\theta) + \frac{3}{4}(B - A)\sin^2\theta\cos 2\lambda\right] + 0\left(\frac{1}{r^4}\right).^{1} \tag{2-48}$$

In rectangular coordinates this assumes the symmetrical form

$$V = \frac{kM}{r} + \frac{k}{2r^5}[(B + C - 2A)x^2 + (C + A - 2B)y^2 + (A + B - 2C)z^2] + 0\left(\frac{1}{r^4}\right), \tag{2-48'}$$

which is easily obtained by taking into account the relations (1-36) between rectangular and spherical coordinates.

Terms of order higher than $1/r^3$ may be neglected for larger distances (say, for the distance to the moon), so that (2-48) or (2-48′), omitting the higher order terms $0(1/r^4)$, are sufficient for many astronomical purposes. For planetary distances even the first term,

$$V = \frac{kM}{r},$$

is generally sufficient; it represents the potential of a point mass. Thus, for very large distances, every body acts like a point mass.

If the form (2-39) of the spherical-harmonic expansion of V is used, then the coefficients of lower degree are obtained from (2-40) and (2-47). We find

$$\begin{aligned} J_{10} &= J_{11} = K_{11} = 0, \\ J_{20} &= \frac{C - \dfrac{A+B}{2}}{Ma^2}, \\ J_{21} &= K_{21} = 0, \\ J_{22} &= \frac{A - B}{4Ma^2}, \\ K_{22} &= -\frac{D}{2Ma^2}. \end{aligned} \tag{2-49}$$

The first of these formulas shows that the summation in (2-39) actually begins with $n = 2$; the others relate the coefficients of second degree to the mass and the moments and products of inertia of the earth.

[1] The notation $0(1/r^4)$ means terms of the order of $1/r^4$.

2-7. The Gravity Field of the Level Ellipsoid

As a first approximation the earth is a sphere; as a second approximation it may be considered an ellipsoid of revolution. Although the earth is not an exact ellipsoid, the gravity field of an ellipsoid is of fundamental practical importance because it is easy to handle mathematically and the deviations of the actual gravity field from the ellipsoidal "normal" field are so small that they can be considered linear. This splitting of the earth's gravity field into a "normal" and a remaining small "disturbing" field considerably simplifies the problem of its determination; the problem could hardly be solved otherwise.

We therefore assume that the normal figure of the earth is a *level ellipsoid*, that is, an ellipsoid of revolution which is an equipotential surface of a *normal gravity field*. This assumption is necessary because the ellipsoid is to be the normal form of the geoid, which is an equipotential surface of the actual gravity field. Denoting the potential of the normal gravity field by

$$U = U(x, y, z),$$

we see that the level ellipsoid, being a surface $U =$ const., exactly corresponds to the geoid, defined as a surface $W =$ const.

The basic point here is that by postulating that the given ellipsoid be an equipotential surface of the normal gravity field, and by prescribing the total mass M, we completely and uniquely determine the normal potential U. The detailed density distribution inside the ellipsoid, which produces the potential U, is quite uninteresting and need not be known at all.[1]

This determination is made possible by Stokes' theorem (Sec. 1-7). Originally it was shown to hold only for the gravitational potential V, but it can as well be applied to the gravity potential

$$U = V + \frac{1}{2}\omega^2(x^2 + y^2) \tag{2-50}$$

if the angular velocity ω is given. The proof follows that of Sec. 1-7, with obvious modifications. Hence, the normal potential function $U(x, y, z)$ is completely determined by

1. the shape of the ellipsoid of revolution, that is, its semiaxes a and b,
2. the total mass M, and
3. the angular velocity ω.

The calculation will now be carried out in detail. The given ellipsoid S_0,

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2} = 1, \tag{2-51}$$

is by definition an equipotential surface

$$U(x, y, z) = U_0. \tag{2-52}$$

[1] In fact, we do not know of any "reasonable" mass distribution for the level ellipsoid. Pizzetti (1894) successfully used a homogeneous density distribution combined with a surface layer of negative density, which of course is quite "unnatural."

It is now convenient to introduce the ellipsoidal coordinates u, $\bar{\theta}$, λ of Sec. 1-19. The ellipsoid S_0 is taken as the reference ellipsoid $u = b$. In this and the following chapters we shall denote the ellipsoidal θ-coordinate by $\bar{\theta}$, the symbol θ being reserved for the spherical polar distance. This distinction is necessary because θ and $\bar{\theta}$ will be used in the same context. Furthermore, we shall introduce

$$\beta = 90^\circ - \bar{\theta},$$

which is the *reduced latitude* much used in geometrical geodesy. See also Figs. 1-14 and 1-15.

Since the gravitational part, V, of the normal potential U will be harmonic outside the ellipsoid S_0, we use the series (1-111b). The field V has rotational symmetry and hence does not depend on the longitude λ. Therefore, all non-zonal terms, which depend on λ, must be zero, and there remains

$$V(u, \beta) = \sum_{n=0}^{\infty} \frac{Q_n\left(i\dfrac{u}{E}\right)}{Q_n\left(i\dfrac{b}{E}\right)} A_n P_n(\sin\beta), \tag{2-53}$$

where

$$E = \sqrt{a^2 - b^2} \tag{2-54}$$

is the linear eccentricity. The centrifugal potential Φ is given by

$$\Phi = \frac{1}{2}\omega^2(x^2 + y^2) = \frac{1}{2}\omega^2(u^2 + E^2)\cos^2\beta. \tag{2-55}$$

Hence the total normal gravity potential may be written

$$U(u, \beta) = \sum_{n=0}^{\infty} \frac{Q_n\left(i\dfrac{u}{E}\right)}{Q_n\left(i\dfrac{b}{E}\right)} A_n P_n(\sin\beta) + \frac{1}{2}\omega^2(u^2 + E^2)\cos^2\beta.$$

On the ellipsoid S_0 we have $u = b$ and $U = U_0$. Hence

$$\sum_{n=0}^{\infty} A_n P_n(\sin\beta) + \frac{1}{2}\omega^2(b^2 + E^2)\cos^2\beta = U_0.$$

This equation must hold for all points of S_0, that is, for all values of β. Since

$$b^2 + E^2 = a^2$$

and

$$\cos^2\beta = \frac{2}{3}[1 - P_2(\sin\beta)],$$

we have

$$\sum_{n=0}^{\infty} A_n P_n(\sin\beta) + \frac{1}{3}\omega^2 a^2 - \frac{1}{3}\omega^2 a^2 P_2(\sin\beta) - U_0 = 0$$

or

$$\left(A_0 + \frac{1}{3}\omega^2a^2 - U_0\right)P_0(\sin\beta) + A_1P_1(\sin\beta)$$
$$+ \left(A_2 - \frac{1}{3}\omega^2a^2\right)P_2(\sin\beta) + \sum_{n=3}^{\infty} A_nP_n(\sin\beta) = 0.$$

This equation will hold for all values of β only if the coefficient of every $P_n(\sin\beta)$ is zero. Thus we get

$$A_0 = U_0 - \frac{1}{3}\omega^2a^2, \quad A_1 = 0,$$

$$A_2 = \frac{1}{3}\omega^2a^2, \quad A_3 = A_4 = \cdots = 0.$$

Inserting these into (2-53) gives

$$V(u,\beta) = \left(U_0 - \frac{1}{3}\omega^2a^2\right)\frac{Q_0\left(i\frac{u}{E}\right)}{Q_0\left(i\frac{b}{E}\right)} + \frac{1}{3}\omega^2a^2\frac{Q_2\left(i\frac{u}{E}\right)}{Q_2\left(i\frac{b}{E}\right)}P_2(\sin\beta). \quad (2\text{-}56)$$

This formula is basically the solution of Dirichlet's problem for the level ellipsoid, but we can give it more convenient forms.

First, we determine the Legendre functions of the second kind, Q_0 and Q_2. As

$$\coth^{-1} ix = \frac{1}{i}\cot^{-1} x = -i\tan^{-1}\frac{1}{x},$$

we find by (1-65′) with $z = iu/E$:

$$Q_0\left(i\frac{u}{E}\right) = -i\tan^{-1}\frac{E}{u},$$

$$Q_2\left(i\frac{u}{E}\right) = \frac{i}{2}\left[\left(1 + 3\frac{u^2}{E^2}\right)\tan^{-1}\frac{E}{u} - 3\frac{u}{E}\right].$$

By introducing the abbreviations

$$q = \frac{1}{2}\left[\left(1 + 3\frac{u^2}{E^2}\right)\tan^{-1}\frac{E}{u} - 3\frac{u}{E}\right], \quad (2\text{-}57)$$

$$q_0 = \frac{1}{2}\left[\left(1 + 3\frac{b^2}{E^2}\right)\tan^{-1}\frac{E}{b} - 3\frac{b}{E}\right], \quad (2\text{-}58)$$

and substituting in equation (2-56), we obtain

$$V(u,\beta) = \left(U_0 - \frac{1}{3}\omega^2a^2\right)\frac{\tan^{-1}\frac{E}{u}}{\tan^{-1}\frac{E}{b}} + \frac{1}{3}\omega^2a^2\frac{q}{q_0}P_2(\sin\beta). \quad (2\text{-}59)$$

Now we can express U_0 in terms of the mass M. For large values of u we have

$$\tan^{-1}\frac{E}{u} = \frac{E}{u} + 0\left(\frac{1}{u^3}\right).$$

From the expressions (1-36) for spherical coordinates and from equations (1-103) for ellipsoidal coordinates we find

$$x^2 + y^2 + z^2 = r^2 = u^2 + E^2 \cos^2 \beta,$$

so that for large values of r we have

$$\frac{1}{u} = \frac{1}{r} + 0\left(\frac{1}{r^3}\right)$$

and

$$\tan^{-1} \frac{E}{u} = \frac{E}{r} + 0\left(\frac{1}{r^3}\right).$$

For very large distances r, the first term in (2-59) is dominant, so that asymptotically

$$V = \left(U_0 - \frac{1}{3}\omega^2 a^2\right) \frac{E}{\tan^{-1}(E/b)} \frac{1}{r} + 0\left(\frac{1}{r^3}\right).$$

We know from the preceding section that

$$V = \frac{kM}{r} + 0\left(\frac{1}{r^3}\right).$$

Comparison of these two expressions shows that

$$kM = \left(U_0 - \frac{1}{3}\omega^2 a^2\right) \frac{E}{\tan^{-1}(E/b)}, \tag{2-60}$$

$$U_0 = \frac{kM}{E} \tan^{-1} \frac{E}{b} + \frac{1}{3}\omega^2 a^2 \tag{2-61}$$

are the desired relations between mass M and potential U_0.

We can substitute these relations into the expression for V, given by (2-59), and express P_2 as

$$P_2(\sin \beta) = \frac{3}{2} \sin^2 \beta - \frac{1}{2}.$$

Finally, if we add the centrifugal potential Φ (2-55), we get the normal gravity potential U as

$$U(u, \beta) = \frac{kM}{E} \tan^{-1} \frac{E}{u} + \frac{1}{2}\omega^2 a^2 \frac{q}{q_0}\left(\sin^2 \beta - \frac{1}{3}\right) + \frac{1}{2}\omega^2(u^2 + E^2) \cos^2 \beta. \tag{2-62}$$

The only constants that occur in this formula are a, b, kM, and ω. This is in complete agreement with Stokes' theorem.

2-8. Normal Gravity

The line element in ellipsoidal coordinates is, according to (1-104), given by

$$ds^2 = w^2\, du^2 + w^2(u^2 + E^2)\, d\beta^2 + (u^2 + E^2) \cos^2 \beta\, d\lambda^2,$$

where

$$w = \sqrt{\frac{u^2 + E^2 \sin^2 \beta}{u^2 + E^2}}. \tag{2-63}$$

Along the coordinate lines we thus have:

$$
\begin{array}{llll}
u = \text{variable}, & \beta = \text{const.}, & \lambda = \text{const.} & ds_u = w\,du, \\
\beta = \text{variable}, & u = \text{const.}, & \lambda = \text{const.} & ds_\beta = w\sqrt{u^2 + E^2}\,d\beta, \\
\lambda = \text{variable}, & u = \text{const.}, & \beta = \text{const.} & ds_\lambda = \sqrt{u^2 + E^2}\cos\beta\,d\lambda.
\end{array}
$$

The components of the normal gravity vector

$$\gamma = \operatorname{grad} U \tag{2-64}$$

along these coordinate lines are accordingly given by

$$
\begin{aligned}
\gamma_u &= \frac{\partial U}{\partial s_u} = \frac{1}{w}\frac{\partial U}{\partial u}, \\
\gamma_\beta &= \frac{\partial U}{\partial s_\beta} = \frac{1}{w\sqrt{u^2 + E^2}}\frac{\partial U}{\partial \beta}, \\
\gamma_\lambda &= \frac{\partial U}{\partial s_\lambda} = \frac{1}{\sqrt{u^2 + E^2}\cos\beta}\frac{\partial U}{\partial \lambda} = 0.
\end{aligned} \tag{2-65}
$$

The component γ_λ is zero because U does not contain λ. This is also evident from the rotational symmetry.

On performing the partial differentiations we find

$$
\begin{aligned}
-w\gamma_u &= \frac{kM}{u^2 + E^2} + \frac{\omega^2 a^2 E}{u^2 + E^2}\frac{q'}{q_0}\left(\frac{1}{2}\sin^2\beta - \frac{1}{6}\right) - \omega^2 u \cos^2\beta, \\
-w\gamma_\beta &= \left(-\frac{\omega^2 a^2}{\sqrt{u^2 + E^2}}\frac{q}{q_0} + \omega^2\sqrt{u^2 + E^2}\right)\sin\beta\cos\beta,
\end{aligned} \tag{2-66}
$$

where we have set

$$q' = -\frac{u^2 + E^2}{E}\frac{dq}{du} = 3\left(1 + \frac{u^2}{E^2}\right)\left(1 - \frac{u}{E}\tan^{-1}\frac{E}{u}\right) - 1. \tag{2-67}$$

Note that q' does *not* mean dq/du; this notation has been borrowed from Hirvonen (1960), where q' is the derivative with respect to another independent variable α, which we are not using here.

For the level ellipsoid S_0 itself we have $u = b$, and we get

$$\gamma_{\beta,0} = 0. \tag{2-68}$$

(We shall often denote quantities referred to S_0 by the subscript 0.) This is also evident because on S_0 the gravity vector is normal to the level surface S_0. Hence, in addition to the λ-component, the β-component is also zero on the reference ellipsoid $u = b$.[1]

Thus the total gravity on the ellipsoid S_0, which we simply denote by γ, is given by

[1] The other coordinate ellipsoids u = const. are *not* equipotential surfaces U = const., so that the β-component will not in general be zero.

$$\gamma = |\gamma_{u,0}| = \frac{kM}{a\sqrt{a^2 \sin^2 \beta + b^2 \cos^2 \beta}}$$

$$\left[1 + \frac{\omega^2 a^2 E}{kM} \frac{q_0'}{q_0}\left(\frac{1}{2}\sin^2 \beta - \frac{1}{6}\right) - \frac{\omega^2 a^2 b}{kM}\cos^2 \beta\right],$$

since the relations

$$\sqrt{u^2 + E^2} = \sqrt{b^2 + E^2} = a,$$

$$w_0 = \frac{1}{a}\sqrt{b^2 + E^2 \sin^2 \beta} = \frac{1}{a}\sqrt{a^2 \sin^2 \beta + b^2 \cos^2 \beta} \qquad (2\text{-}69)$$

hold on S_0.

If we introduce the abbreviation

$$m = \frac{\omega^2 a^2 b}{kM} \qquad (2\text{-}70)$$

and the *second eccentricity*[1]

$$e' = \frac{E}{b} = \frac{\sqrt{a^2 - b^2}}{b}, \qquad (2\text{-}71)$$

and remove the constant terms by noting that

$$1 = \cos^2 \beta + \sin^2 \beta,$$

we obtain

$$\gamma = \frac{kM}{a\sqrt{a^2 \sin^2 \beta + b^2 \cos^2 \beta}}$$

$$\left[\left(1 + \frac{m}{3}\frac{e' q_0'}{q_0}\right)\sin^2 \beta + \left(1 - m - \frac{m}{6}\frac{e' q_0'}{q_0}\right)\cos^2 \beta\right]. \qquad (2\text{-}72)$$

At the equator ($\beta = 0$) we find

$$\gamma_a = \frac{kM}{ab}\left(1 - m - \frac{m}{6}\frac{e' q_0'}{q_0}\right); \qquad (2\text{-}73)$$

at the poles ($\beta = \pm 90°$) normal gravity is given by

$$\gamma_b = \frac{kM}{a^2}\left(1 + \frac{m}{3}\frac{e' q_0'}{q_0}\right). \qquad (2\text{-}74)$$

Normal gravity at the equator, γ_a, and normal gravity at the pole, γ_b, satisfy the relation

$$\frac{a - b}{a} + \frac{\gamma_b - \gamma_a}{\gamma_a} = \frac{\omega^2 b}{\gamma_a}\left(1 + \frac{e' q_0'}{2 q_0}\right), \qquad (2\text{-}75)$$

which should be verified by substitution. This is the rigorous form of an important approximate formula published by Clairaut in 1738. It is therefore

[1] The first eccentricity is $e = E/a$. The prime on e does not denote differentiation, but merely distinguishes the second eccentricity from the first.

called Clairaut's theorem. Its significance will be made clear in Sec. 2-10.

By comparing expression (2-73) for γ_a and expression (2-74) for γ_b with the quantities within parentheses in formula (2-72) we see that γ can be written in the symmetrical form

$$\gamma = \frac{a\gamma_b \sin^2 \beta + b\gamma_a \cos^2 \beta}{\sqrt{a^2 \sin^2 \beta + b^2 \cos^2 \beta}}. \tag{2-76}$$

We finally introduce the geographical latitude on the ellipsoid, ϕ, which is the angle between the normal to the ellipsoid and the equatorial plane (Fig. 2-11). Using the well-known formula from geometrical geodesy,

$$\tan \beta = \frac{b}{a} \tan \phi, \tag{2-77}$$

we obtain

$$\gamma = \frac{a\gamma_a \cos^2 \phi + b\gamma_b \sin^2 \phi}{\sqrt{a^2 \cos^2 \phi + b^2 \sin^2 \phi}}. \tag{2-78}$$

The computation is left as an exercise for the reader. This rigorous formula for normal gravity on the ellipsoid is due to Somigliana (1929).

We shall close this section with a short remark on the vertical gradient of gravity at the reference ellipsoid, $\partial\gamma/\partial s_u = \partial\gamma/\partial h$. Bruns' formula (2-20), applied to the normal gravity field with $\rho = 0$, yields

$$\frac{\partial \gamma}{\partial h} = -2\gamma J - 2\omega^2. \tag{2-79}$$

The mean curvature of the ellipsoid is given by

$$J = \frac{1}{2}\left(\frac{1}{M} + \frac{1}{N}\right), \tag{2-80}$$

where M and N are the principal radii of curvature: M is the radius in the direction of the meridian, and N is the *normal radius of curvature*, taken in the direction of the prime vertical. From geometrical geodesy we borrow the formulas

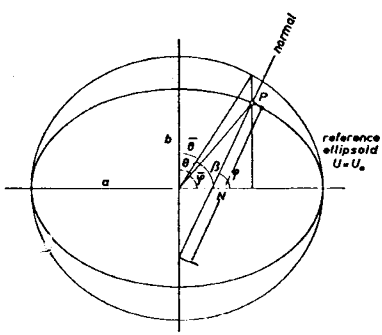

FIGURE 2-11

Geographical (ellipsoidal) latitude ϕ, geocentric latitude $\bar{\phi}$, reduced latitude β, and their complements, for a point P on the ellipsoid.

$$M = \frac{c}{(1 + e'^2 \cos^2 \phi)^{3/2}}, \quad N = \frac{c}{(1 + e'^2 \cos^2 \phi)^{1/2}}, \tag{2-81}$$

where

$$c = \frac{a^2}{b} \tag{2-82}$$

is the radius of curvature at the pole. The normal radius of curvature, N, admits of a simple geometrical interpretation (Fig. 2-11). It is therefore also known as the "normal terminated by the minor axis" (Bomford, 1962, p. 497).

2-9. Expansion of the Normal Potential in Spherical Harmonics

We have found the gravitational potential of the normal figure of the earth in terms of ellipsoidal harmonics to be

$$V = \frac{kM}{E} \tan^{-1} \frac{E}{u} + \frac{1}{3} \omega^2 a^2 \frac{q}{q_0} P_2(\sin \beta). \tag{2-83}$$

Now we wish to express this equation in terms of spherical coordinates r, θ, λ.

We first establish a relation between ellipsoidal and spherical coordinates. By comparing the rectangular coordinates in these two systems according to equations (1-36) and (1-103) we get

$$\begin{aligned} r \sin \theta \cos \lambda &= \sqrt{u^2 + E^2} \cos \beta \cos \lambda, \\ r \sin \theta \sin \lambda &= \sqrt{u^2 + E^2} \cos \beta \sin \lambda, \\ r \cos \theta &= u \sin \beta. \end{aligned}$$

The longitude λ being the same in both systems, we easily find from these equations

$$\begin{aligned} \cot \theta &= \frac{u}{\sqrt{u^2 + E^2}} \tan \beta, \\ r &= \sqrt{u^2 + E^2 \cos^2 \beta}. \end{aligned} \tag{2-84}$$

The direct transformation of (2-83) by expressing u and β in terms of r and θ by means of equations (2-84) is extremely laborious. However, the problem can be solved easily in an indirect way.

We expand $\tan^{-1}(E/u)$ into the well-known power series

$$\tan^{-1} \frac{E}{u} = \frac{E}{u} - \frac{1}{3}\left(\frac{E}{u}\right)^3 + \frac{1}{5}\left(\frac{E}{u}\right)^5 - + \cdots. \tag{2-85}$$

Insertion of this series into the formula (2-57)

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

leads, after simple manipulations, to

$$q = 2\left[\frac{1}{3 \cdot 5}\left(\frac{E}{u}\right)^3 - \frac{2}{5 \cdot 7}\left(\frac{E}{u}\right)^5 + \frac{3}{7 \cdot 9}\left(\frac{E}{u}\right)^7 - + \cdots\right]. \tag{2-86}$$

More concisely we have

$$\tan^{-1}\frac{E}{u} = \frac{E}{u} + \sum_{n=1}^{\infty}(-1)^n \frac{1}{2n+1}\left(\frac{E}{u}\right)^{2n+1},$$

$$q = -\sum_{n=1}^{\infty}(-1)^n \frac{2n}{(2n+1)(2n+3)}\left(\frac{E}{u}\right)^{2n+1}.$$

By inserting this into (2-83) we obtain

$$V = \frac{kM}{u} + \frac{kM}{E}\sum_{n=1}^{\infty}(-1)^n \frac{1}{2n+1}\left(\frac{E}{u}\right)^{2n+1}$$

$$- \frac{\omega^2 a^2}{3q_0}\sum_{n=1}^{\infty}(-1)^n \frac{2n}{(2n+1)(2n+3)}\left(\frac{E}{u}\right)^{2n+1} P_2(\sin\beta).$$

Introducing m, defined by (2-70), and the second eccentricity $e' = E/b$, we find

$$V = \frac{kM}{u} + \sum_{n=1}^{\infty}(-1)^n \frac{kM}{(2n+1)E}\left(\frac{E}{u}\right)^{2n+1}\left[1 - \frac{me'}{3q_0}\frac{2n}{2n+3}P_2(\sin\beta)\right]. \quad (2\text{-}87)$$

We expand the potential V into a series of spherical harmonics. Because of the rotational symmetry there will be only zonal terms, and because of the symmetry with respect to the equatorial plane there will be only *even* zonal harmonics. The zonal harmonics of odd degree change sign for negative latitudes and must therefore be absent. Accordingly, the series has the form

$$V = \frac{kM}{r} + A_2\frac{P_2(\cos\theta)}{r^3} + A_4\frac{P_4(\cos\theta)}{r^5} + \cdots. \quad (2\text{-}88)$$

We next have to determine the coefficients $A_2, A_4, \ldots$. For this purpose we consider a point on the axis of rotation, outside the ellipsoid. For this point we have $\beta = 90°$, $\theta = 0°$, and, by (2-84), $u = r$. Then (2-87) becomes

$$V = \frac{kM}{r} + \sum_{n=1}^{\infty}(-1)^n \frac{kME^{2n}}{2n+1}\left(1 - \frac{2n}{2n+3}\frac{me'}{3q_0}\right)\frac{1}{r^{2n+1}},$$

and (2-88) takes the form

$$V = \frac{kM}{r} + \frac{A_2}{r^3} + \frac{A_4}{r^5} + \cdots = \frac{kM}{r} + \sum_{n=1}^{\infty} A_{2n}\frac{1}{r^{2n+1}}.$$

Here we have used the fact that for all values of n

$$P_n(1) = 1;$$

see also Fig. 1-8. Comparing the coefficients in both expressions for V we find

$$A_{2n} = (-1)^n \frac{kME^{2n}}{2n+1}\left(1 - \frac{2n}{2n+3}\frac{me'}{3q_0}\right). \quad (2\text{-}89)$$

Equations (2-88) and (2-89) give the desired expression for the potential of the level ellipsoid as a series of spherical harmonics.

The second-degree coefficient A_2 is

$$A_2 = k(A - C).$$

This follows from (2-47); we have $A = B$ for reasons of symmetry. The C is the moment of inertia with respect to the axis of rotation, and A is the moment of inertia with respect to any axis in the equatorial plane. By letting $n = 1$ in (2-89) we obtain

$$A_2 = -\frac{1}{3} kME^2 \left(1 - \frac{2}{15}\frac{me'}{q_0}\right).$$

Comparing this with the preceding equation we find

$$k(C - A) = \frac{1}{3} kME^2 \left(1 - \frac{2}{15}\frac{me'}{q_0}\right). \tag{2-90}$$

Thus the difference between the principal moments of inertia is expressed in terms of "Stokes' constants" a, b, M, and ω.

It is possible to eliminate q_0 from equations (2-89) and (2-90), obtaining

$$A_{2n} = (-1)^n \frac{3kME^{2n}}{(2n + 1)(2n + 3)} \left(1 - n + 5n \frac{C - A}{ME^2}\right). \tag{2-91}$$

If we write the potential V in the form

$$V = \frac{kM}{r}\left[1 - J_2 \left(\frac{a}{r}\right)^2 P_2(\cos\theta) - J_4 \left(\frac{a}{r}\right)^4 P_4(\cos\theta) - \cdots\right]$$

$$= \frac{kM}{r}\left[1 - \sum_{n=1}^{\infty} J_{2n} \left(\frac{a}{r}\right)^{2n} P_{2n}(\cos\theta)\right],$$

then the J are given by

$$J_{2n} = (-1)^{n+1} \frac{3e^{2n}}{(2n + 1)(2n + 3)} \left(1 - n + 5n \frac{C - A}{ME^2}\right). \tag{2-92}$$

Here we have introduced the first eccentricity $e = E/a$. For $n = 1$ this gives the important formula

$$J_2 = \frac{C - A}{Ma^2}, \tag{2-92$'$}$$

which is in agreement with equations (2-49).

Finally we note that on eliminating $q_0 = \frac{1}{i} Q_2\left(i\frac{b}{E}\right)$ by using (2-90), and U_0 by using (2-60), we may write the expansion of V in ellipsoidal harmonics, equation (2-56), in the form

$$V(u, \beta) = \frac{i}{E} kM \cdot Q_0\left(i\frac{u}{E}\right)$$

$$+ \frac{15i}{2E^3} k\left(C - A - \frac{1}{3} ME^2\right) \cdot Q_2\left(i\frac{u}{E}\right) P_2(\sin\beta). \tag{2-93}$$

This shows that the coefficients of the ellipsoidal harmonics of degrees zero and two are functions of the mass and of the difference between the two principal moments of inertia. The analogy to the corresponding spherical-harmonic coefficients (2-47) is obvious.

2-10. Series Expansions for the Normal Gravity Field

Since the earth ellipsoid is very nearly a sphere, the quantities

$$\begin{aligned} E &= \sqrt{a^2 - b^2}, \text{ linear eccentricity,} \\ e &= \frac{E}{a}, \text{ first (numerical) eccentricity,} \\ e' &= \frac{E}{b}, \text{ second (numerical) eccentricity,} \\ f &= \frac{a-b}{a}, \text{ flattening,} \end{aligned} \tag{2-94}$$

and similar parameters that characterize the deviation from a sphere, are small. Therefore, series expansions in terms of these or similar parameters will be convenient for numerical calculations.

Linear approximation. In order that the reader may find his way through the subsequent practical formulas we first consider an approximation that is linear in the flattening f. Here we get particularly simple and symmetrical formulas which also exhibit plainly the structure of the higher-order expansions.

It is well known that the radius vector r of an ellipsoid is approximately given by

$$r = a(1 - f \sin^2 \phi). \tag{2-95}$$

As we shall see subsequently, normal gravity may, to the same approximation, be written

$$\gamma = \gamma_a (1 + f^* \sin^2 \phi). \tag{2-96}$$

For $\phi = \pm 90°$, at the poles, we have $r = b$ and $\gamma = \gamma_b$. Hence we may write

$$b = a(1 - f), \quad \gamma_b = \gamma_a(1 + f^*),$$

and solving for f and f^* we obtain

$$f = \frac{a-b}{a}, \tag{2-97}$$

$$f^* = \frac{\gamma_b - \gamma_a}{\gamma_a}, \tag{2-98}$$

so that f is the flattening defined by (2-94), and f^* is an analogous quantity which may be called *gravity flattening*.

To the same approximation, (2-75) becomes

$$f + f^* = \frac{5}{2} m, \tag{2-99}$$

where

$$m \doteq \frac{\omega^2 a}{\gamma_a} = \frac{\text{centrifugal force at equator}}{\text{gravity at equator}}. \tag{2-100}$$

This is *Clairaut's theorem* in its original form. It is one of the most striking formulas of physical geodesy: the (geometrical) flattening f (2-97) can be derived from f^* and m, which are purely dynamical quantities obtained by gravity measurements; that is, *the flattening of the earth can be obtained from gravity measurements.*

Of course, Clairaut's formula is only a first approximation and must be improved, first by the inclusion of higher-order ellipsoidal terms in f, and secondly by taking the deviation of the earth's gravity field from the normal gravity field into account. But the principle remains the same.

Second-order expansion. We shall now expand the closed formulas of the two preceding sections into series in terms of the second eccentricity e' and the flattening f, in general up to and including e'^4 or f^2. Terms of the order of e'^6 or f^3 and higher will usually be neglected.

We start from the series

$$\begin{aligned}
\tan^{-1}\frac{E}{u} &= \frac{E}{u} - \frac{1}{3}\left(\frac{E}{u}\right)^3 + \frac{1}{5}\left(\frac{E}{u}\right)^5 - \frac{1}{7}\left(\frac{E}{u}\right)^7 + - \cdots,\\
q &= 2\left[\frac{1}{3\cdot 5}\left(\frac{E}{u}\right)^3 - \frac{2}{5\cdot 7}\left(\frac{E}{u}\right)^5 + \frac{3}{7\cdot 9}\left(\frac{E}{u}\right)^7 - + \cdots\right],\\
q' &= 6\left[\frac{1}{3\cdot 5}\left(\frac{E}{u}\right)^2 - \frac{1}{5\cdot 7}\left(\frac{E}{u}\right)^4 + \frac{1}{7\cdot 9}\left(\frac{E}{u}\right)^6 - + \cdots\right].
\end{aligned} \tag{2-101}$$

The first two series have already been used in the preceding section; the third is obtained by inserting the $\tan^{-1}$ series into the closed formula (2-67) for q'.

On the reference ellipsoid S_0 we have $u = b$ and

$$\frac{E}{u} = \frac{E}{b} = e',$$

so that

$$\tan^{-1} e' = e' - \frac{1}{3}e'^3 + \frac{1}{5}e'^5 \cdots,$$

$$q_0 = \frac{2}{15}e'^3\left(1 - \frac{6}{7}e'^2 \cdots\right), \tag{2-102}$$

$$q_0' = \frac{2}{5}e'^2\left(1 - \frac{3}{7}e'^2 \cdots\right), \tag{2-103}$$

$$\frac{e' q_0'}{q_0} = 3\left(1 + \frac{3}{7}e'^2 \cdots\right).$$

We shall also need the series

$$b = \frac{a}{\sqrt{1 + e'^2}} = a\left(1 - \frac{1}{2}e'^2 + \frac{3}{8}e'^4 \cdots\right).$$

Potential and gravity. By substituting these expressions into the closed formulas (2-61), (2-73), (2-74), and (2-75) we obtain, up to and including the order e'^4,

potential:

$$U_0 = \frac{kM}{b}\left(1 - \frac{1}{3}e'^2 + \frac{1}{5}e'^4\right) + \frac{1}{3}\omega^2 a^2, \tag{2-104}$$

gravity at the equator and the pole:

$$\gamma_a = \frac{kM}{ab}\left(1 - \frac{3}{2}m - \frac{3}{14}e'^2 m\right), \tag{2-105a}$$

$$\gamma_b = \frac{kM}{a^2}\left(1 + m + \frac{3}{7}e'^2 m\right), \tag{2-105b}$$

Clairaut's theorem:

$$f + f^* = \frac{5}{2}\frac{\omega^2 b}{\gamma_a}\left(1 + \frac{9}{35}e'^2\right). \tag{2-106}$$

The ratio $\omega^2 a/\gamma_a$ may be expressed as

$$\frac{\omega^2 a}{\gamma_a} = m + \frac{3}{2}m^2, \tag{2-107}$$

which is a more accurate version of (2-100).

From equation (2-105a) we find

$$kM = ab\gamma_a\left(1 + \frac{3}{2}m + \frac{3}{14}e'^2 m + \frac{9}{4}m^2\right), \tag{2-108}$$

which gives the mass in terms of equatorial gravity. By means of this equation we can express kM in equation (2-104) in terms of γ_a, obtaining

$$U_0 = a\gamma_a\left(1 - \frac{1}{3}e'^2 + \frac{11}{6}m + \frac{1}{5}e'^4 - \frac{2}{7}e'^2 m + \frac{11}{4}m^2\right). \tag{2-109}$$

Here we have eliminated $\omega^2 a^2$ by replacing it with kMm/b.

Now we can attack equation (2-78) for normal gravity. A simple manipulation yields

$$\gamma = \gamma_a \frac{1 + \dfrac{b\gamma_b - a\gamma_a}{a\gamma_a}\sin^2\phi}{\sqrt{1 - \dfrac{a^2 - b^2}{a^2}\sin^2\phi}}.$$

The denominator is expanded into a binomial series:

$$\frac{1}{\sqrt{1 - x}} = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \cdots.$$

Then the abbreviated series

$$\frac{a^2 - b^2}{a^2} = \frac{e'^2}{1 + e'^2} = e'^2 - e'^4,$$

$$\frac{b\gamma_b - a\gamma_a}{a\gamma_a} = -e'^2 + \frac{5}{2}m + e'^4 - \frac{13}{7}e'^2 m + \frac{15}{4}m^2$$

are introduced and we obtain, upon substitution,

$$\gamma = \gamma_a\left[1 + \left(-\frac{1}{2}e'^2 + \frac{5}{2}m + \frac{1}{2}e'^4 - \frac{13}{7}e'^2m + \frac{15}{4}m^2\right)\sin^2\phi \right.$$
$$\left. + \left(-\frac{1}{8}e'^4 + \frac{5}{4}e'^2m\right)\sin^4\phi\right]. \quad (2\text{-}110)$$

We may also express these quantities in terms of the flattening f by substituting the equation

$$e'^2 = \frac{1}{(1-f)^2} - 1 = 2f + 3f^2 + \cdots.$$

The flattening f is most commonly used; it offers a slight advantage over the second eccentricity e' in that it is of the same order of magnitude as m: it is not immediately apparent that m^2, e'^2m, and e'^4 are quantities of the same order of magnitude. Thus we obtain

$$kM = ab\gamma_a\left(1 + \frac{3}{2}m + \frac{3}{7}fm + \frac{9}{4}m^2\right), \quad (2\text{-}111)$$

$$U_0 = a\gamma_a\left(1 - \frac{2}{3}f + \frac{11}{6}m - \frac{1}{5}f^2 - \frac{4}{7}fm + \frac{11}{4}m^2\right), \quad (2\text{-}112)$$

$$\gamma = \gamma_a\left[1 + \left(-f + \frac{5}{2}m + \frac{1}{2}f^2 - \frac{26}{7}fm + \frac{15}{4}m^2\right)\sin^2\phi \right.$$
$$\left. + \left(-\frac{1}{2}f^2 + \frac{5}{2}fm\right)\sin^4\phi\right]. \quad (2\text{-}113)$$

The last formula is usually abbreviated as

$$\gamma = \gamma_a(1 + f_2\sin^2\phi + f_4\sin^4\phi), \quad (2\text{-}114)$$

so that we have

$$f_2 = -f + \frac{5}{2}m + \frac{1}{2}f^2 - \frac{26}{7}fm + \frac{15}{4}m^2,$$
$$f_4 = -\frac{1}{2}f^2 + \frac{5}{2}fm. \quad (2\text{-}115)$$

By substituting

$$\sin^4\phi = \sin^2\phi - \frac{1}{4}\sin^2 2\phi$$

we finally obtain

$$\gamma = \gamma_a\left(1 + f^*\sin^2\phi - \frac{1}{4}f_4\sin^2 2\phi\right), \quad (2\text{-}116)$$

where

$$f^* = \frac{\gamma_b - \gamma_a}{\gamma_a} = f_2 + f_4 \quad (2\text{-}117)$$

is the "gravity flattening."

Coefficients of spherical harmonics. Equation (2-90) for the principal moments of inertia yields at once

$$\frac{C-A}{ME^2} = \frac{1}{3} - \frac{2}{45}\frac{me'}{q_0}.$$

Expanding by means of (2-102) we find

$$\frac{C-A}{ME^2} = \frac{1}{e'^2}\left(\frac{1}{3}e'^2 - \frac{1}{3}m - \frac{2}{7}e'^2 m\right).$$

Inserting this into (2-92) we obtain

$$\begin{aligned} J_2 = \frac{C-A}{Ma^2} &= \frac{1}{3}e'^2 - \frac{1}{3}m - \frac{1}{3}e'^4 + \frac{1}{21}e'^2 m \\ &= \frac{2}{3}f - \frac{1}{3}m - \frac{1}{3}f^2 + \frac{2}{21}fm, \end{aligned} \tag{2-118}$$

$$J_4 = -\frac{1}{5}e'^4 + \frac{2}{7}e'^2 m = -\frac{4}{5}f^2 + \frac{4}{7}fm. \tag{2-119}$$

The higher J are already of an order of magnitude that we have neglected.

Gravity above the ellipsoid. For a small elevation h above the ellipsoid, normal gravity γ_h at this elevation can be expanded into a series in terms of h:

$$\gamma_h = \gamma + \frac{\partial\gamma}{\partial h}h + \frac{1}{2}\frac{\partial^2\gamma}{\partial h^2}h^2 + \cdots,$$

where γ and its derivatives are referred to the ellipsoid ($h = 0$).

The first derivative $\partial\gamma/\partial h$ is given by Bruns' formula (2-79):

$$\frac{\partial\gamma}{\partial h} = -\gamma\left(\frac{1}{M} + \frac{1}{N}\right) - 2\omega^2, \tag{2-120}$$

where M, N are the principal radii of curvature of the ellipsoid, defined by (2-81). Since

$$\frac{1}{M} = \frac{b}{a^2}(1 + e'^2\cos^2\phi)^{3/2} = \frac{b}{a^2}\left(1 + \frac{3}{2}e'^2\cos^2\phi\cdots\right),$$

$$\frac{1}{N} = \frac{b}{a^2}(1 + e'^2\cos^2\phi)^{1/2} = \frac{b}{a^2}\left(1 + \frac{1}{2}e'^2\cos^2\phi\cdots\right),$$

we have

$$\frac{1}{M} + \frac{1}{N} = \frac{b}{a^2}(2 + 2e'^2\cos^2\phi) = \frac{2b}{a^2}(1 + 2f\cos^2\phi).$$

Here we have limited ourselves to terms linear in f, since the elevation h is already a small quantity. Thus we find from (2-120) after simple manipulations:

$$\frac{\partial\gamma}{\partial h} = -\frac{2\gamma}{a}(1 + f + m - 2f\sin^2\phi). \tag{2-121}$$

The second derivative $\partial^2\gamma/\partial h^2$ may be taken from the spherical approximation, obtained by neglecting e'^2 or f:

$$\gamma = \frac{kM}{a^2},\quad \frac{\partial\gamma}{\partial h} = \frac{\partial\gamma}{\partial a} = -\frac{2kM}{a^3},\quad \frac{\partial^2\gamma}{\partial h^2} = \frac{\partial^2\gamma}{\partial a^2} = \frac{6kM}{a^4},$$

so that

$$\frac{\partial^2\gamma}{\partial h^2} = \frac{6\gamma}{a^2}. \tag{2-122}$$

Thus we obtain

$$\gamma_h = \gamma\left[1 - \frac{2}{a}(1 + f + m - 2f\sin^2\phi)h + \frac{3}{a^2}h^2\right]. \tag{2-123}$$

Using equation (2-113) for γ, we may also write the difference $\gamma_h - \gamma$ in the form

$$\gamma_h - \gamma = -\frac{2\gamma_a}{a}\left[1 + f + m + \left(-3f + \frac{5}{2}m\right)\sin^2\phi\right]h + \frac{3\gamma_a}{a^2}h^2. \tag{2-124}$$

The symbol γ_h denotes the normal gravity for a point at latitude ϕ, situated at height h above the ellipsoid; γ is the gravity at the ellipsoid itself, for the same latitude ϕ, as given by (2-116) or equivalent formulas.

Higher-order series expansions and computational formulas for many quantities of the normal gravity field are given in Hirvonen (1960).

2-11. Numerical Values. The International Ellipsoid

The reference ellipsoid and its gravity field are completely determined by four constants. Usually one takes the following four parameters:

a, semimajor axis;
f, flattening;
γ_a, equatorial gravity; and
ω, angular velocity.

The best-known and most widely used values are those of the *international ellipsoid*:

$$\begin{aligned} a &= 6\,378\,388.000 \text{ meters},\\ f &= 1/297.000,\\ \gamma_a &= 978.049\,000 \text{ gal},\\ \omega &= 0.729\,211\,51 \cdot 10^{-4}\ \text{sec}^{-1}. \end{aligned} \tag{2-125}$$

The geometric parameters a and f were determined by Hayford in 1909 from isostatically reduced astrogeodetic data in the United States. They were adopted for the international ellipsoid by the assembly of the International Association of Geodesy at Madrid in 1924. The equatorial gravity value γ_a was computed by Heiskanen (1928) from isostatically reduced gravity data. The corresponding *international gravity formula*,

$$\gamma = 978.0490(1 + 0.005\,2884\sin^2\phi - 0.000\,0059\sin^2 2\phi)\ \text{gal}, \tag{2-126}$$

whose coefficients were computed from the assumed values for a, f, γ_a, ω by Cassinis (1930) [equations (2-115), (2-116), (2-117)], was adopted by the assembly at Stockholm in 1930.

All parameters of the international ellipsoid and its gravity field can be computed from (2-125) to any desired degree of accuracy, which of course expresses merely the inner consistency. In this way we find

$$\begin{aligned} b &= 6\,356\,911.9 \text{ meters}, \\ E &= 522\,976.1 \text{ meters}, \\ e'^2 &= 0.006\,768\,17, \\ q_0 &= 0.000\,073\,8130, \\ q_0' &= 0.002\,699\,44, \\ m &= 0.003\,449\,86. \end{aligned} \tag{2-127}$$

The potential of the international ellipsoid is

$$U_0 = 6\,263\,978.7 \text{ kgal meters}. \tag{2-128}$$

The product of the earth's mass and the gravitational constant has the value

$$kM = 3.986\,3290 \times 10^{20} \text{ cm}^3 \text{ sec}^{-2}. \tag{2-129}$$

Since the gravitational constant has the value

$$k = 6.67 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1} \text{ sec}^{-2},$$

the earth's mass is

$$M = 5.98 \times 10^{27} \text{ g}.$$

Since k is not very accurate, it would be rather meaningless to give a higher accuracy for M.

For the constants in the spherical-harmonic expansion of the normal gravity field we find the values

$$\begin{aligned} J_2 &= \frac{C - A}{Ma^2} = 0.001\,0920, \\ J_4 &= -0.000\,002\,43. \end{aligned} \tag{2-130}$$

The change of normal gravity with elevation is given by the formula (2-124), which for the international ellipsoid becomes

$$\gamma_h = \gamma - (0.30877 - 0.00045 \sin^2 \phi)h + 0.000\,072h^2, \tag{2-131}$$

where γ_h and γ are measured in gals, and h is the elevation in kilometers.

Although the international ellipsoid can no longer be considered the closest approximation of the earth by an ellipsoid, it may still be used as a reference ellipsoid for geodetic purposes (see Sec. 2-21 for a discussion of this point).

A set of values that probably fit the actual situation more closely has recently been adopted by the assembly of the International Astronomical Union at Hamburg in 1964 (Fricke et al., 1965):

$$\begin{aligned} a &= 6\,378\,160 \text{ meters}, \\ J_2 &= 0.001\,0827, \\ kM &= 3.986\,03 \times 10^{20} \text{ cm}^3 \text{ sec}^{-2}. \end{aligned} \tag{2-132}$$

The corresponding flattening is $f = 1/298.25$. The value of a, which is considerably smaller than that for the international ellipsoid, incorporates recent astrogeodetic determinations; the change in the value of J_2, and consequently of f, is due to the results from artificial satellites.

The eastern countries use the ellipsoid of Krassowsky:

$$a = 6\,378\,245 \text{ meters}, \quad f = 1/298.3. \tag{2-133}$$

In this book we shall continue to use the values (2-125) of the international ellipsoid unless otherwise stated, because most computations, tables, etc., are referred to it; moreover, those values have not as yet been officially changed by the International Union of Geodesy and Geophysics.

2-12. Other Normal Gravity Fields and Reference Surfaces

As we have mentioned, the gravity field of the earth is conveniently split up into a normal and a disturbing field. The normal field comprises the large-scale features, so that the deviations of the actual gravity field from the normal field—the disturbances—are small. The normal field should, furthermore, be mathematically simple. Otherwise it is quite arbitrary.

The use of the ellipsoid as a reference surface for the gravity field is comparatively recent. It was not used officially until 1930, when the assembly of the International Association of Geodesy at Stockholm adopted the theoretical gravity formula (2-126), which is based on an ellipsoid of revolution. Formerly one used the first terms of the spherical-harmonic expansion of W as a normal potential U, that is, the functions

$$U' = \frac{Y_0}{r} + \frac{Y_2(\theta, \lambda)}{r^3} + \frac{1}{2}\omega^2(x^2 + y^2), \tag{2-134a}$$

$$U'' = \frac{Y_0}{r} + \frac{Y_2(\theta, \lambda)}{r^3} + \frac{Y_4(\theta, \lambda)}{r^5} + \frac{1}{2}\omega^2(x^2 + y^2). \tag{2-134b}$$

The first-degree harmonic is missing because the center of the earth is chosen as the origin of coordinates; the third-degree harmonic has been omitted because the normal field is taken to be symmetrical with respect to the equatorial plane. The functions $Y_0 = kM$, Y_2, and Y_4 are assumed to be those of the actual gravity field of the earth.

The corresponding reference surfaces $U = U_0$ are called *earth spheroids:*[1] the surface

$$U'(x, y, z) = U_0 \tag{2-135a}$$

is known as *Bruns' spheroid;* the surface

$$U''(x, y, z) = U_0 \tag{2-135b}$$

is *Helmert's spheroid.*

[1] A spheroid is (1) any surface resembling a sphere; and (2) in particular, an ellipsoid of revolution. In this book we shall use the word "spheroid" in the first, broader meaning rather than in the second, special sense.

According to (2-48′), Bruns' spheroid is given by the equation

$$\frac{kM}{r} + \frac{k}{2r^5}[(B + C - 2A)x^2 + (C + A - 2B)y^2 + (A + B - 2C)z^2] + \frac{1}{2}\omega^2(x^2 + y^2) = U_0. \quad (2\text{-}136)$$

By removing the square root

$$r = \sqrt{x^2 + y^2 + z^2}$$

we find that it is an algebraic surface of the 14th degree. Helmert's spheroid is a surface of the 22nd degree.

Practically, these surfaces closely approximate ellipsoids. However, they are much more complicated mathematically, so that hardly any closed formulas can be obtained for them.

Three reasons given in favor of the ellipsoid as a reference surface in physical geodesy are listed below.

1. Since an ellipsoid is always used as a reference surface for triangulations, etc., the same ellipsoid can be used both as a geometrical and a physical reference surface.
2. The closed formulas for the level ellipsoid permit not only a clear-cut and precise definition of the normal gravity field, but also practical computations of any accuracy.
3. The functions (2-134a) and (2-134b) might be thought of as the natural first approximations of the earth's gravity field. However, the spherical-harmonic expansion of the gravity potential is no more "natural" than, say, an expansion in terms of ellipsoidal harmonics. If we expand W into a series of ellipsoidal harmonics, then the level ellipsoid is the first approximation.

The concept of the reference surface and its gravity field will become still clearer in the following sections, particularly in Sec. 2-21.

2-13. The Anomalous Gravity Field. Geoidal Undulations and Deflections of the Vertical

The small difference between the actual gravity potential W and the normal gravity potential U is denoted by T, so that

$$W(x, y, z) = U(x, y, z) + T(x, y, z); \quad (2\text{-}137)$$

T is called the *anomalous potential*, or *disturbing potential*.

We compare the geoid

$$W(x, y, z) = W_0$$

with a reference ellipsoid

$$U(x, y, z) = W_0$$

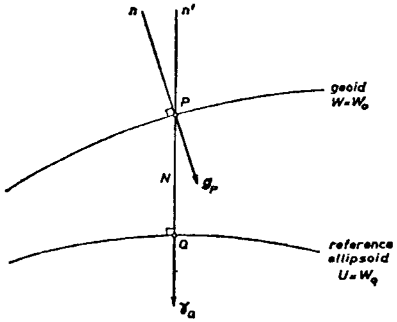

FIGURE 2-12
Geoid and reference ellipsoid.

of the same potential $U_0 = W_0$. A point P of the geoid is projected onto the point Q of the ellipsoid by means of the ellipsoidal normal (Fig. 2-12). The distance PQ between geoid and ellipsoid is called the *geoidal height*, or *geoidal undulation*, and is denoted by N.[1]

Consider now the gravity vector $\mathbf{g}$ at P and the normal gravity vector $\boldsymbol{\gamma}$ at Q. The *gravity anomaly vector* $\Delta\mathbf{g}$ is defined as their difference:

$$\Delta\mathbf{g} = \mathbf{g}_P - \boldsymbol{\gamma}_Q. \tag{2-138}$$

A vector is characterized by *magnitude* and *direction*. The difference in magnitude is the *gravity anomaly*

$$\Delta g = g_P - \gamma_Q; \tag{2-139}$$

the difference in direction is the *deflection of the vertical.*

The deflection of the vertical has two components, a north-south component ξ and an east-west component η (Fig. 2-13). As the direction of the vertical is directly defined by the geographical coordinates latitude and longitude, the components ξ and η can be expressed by them a simple way. The actual geographical coordinates of the geoidal point P, which define the direction of the plumb line n or of the gravity vector $\mathbf{g}$, can be determined by astronomical measurements. They are therefore called *astronomical coordinates* and have been denoted by Φ and Λ. The ellipsoidal geographical coordinates given by the direction of the ellipsoidal normal n' have been denoted by ϕ and λ. It is evident that λ is identical with the geocentric longitude. Thus,

geoidal normal n, astronomical coordinates Φ, Λ;
ellipsoidal normal n', "geodetic" coordinates ϕ, λ.

From Fig. 2-13 we read

$$\begin{aligned} \xi &= \Phi - \phi, \\ \eta &= (\Lambda - \lambda)\cos\phi. \end{aligned} \tag{2-140}$$

[1] Unfortunately there is a conflict of notation here. Denoting both the normal radius of curvature of the ellipsoid and the geoidal height by N is well established in geodetic literature. We shall continue this practice, as there is little chance of confusion.

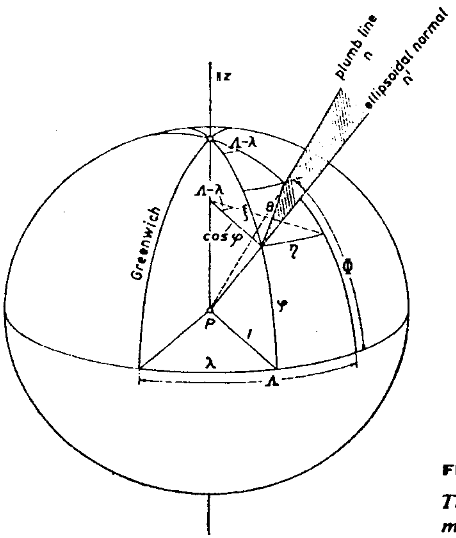

FIGURE 2-13

The deflection of the vertical as illustrated by means of a unit sphere with center at P.

It is also possible to compare the vectors **g** and γ at the same point P. Then we get the *gravity disturbance vector*

$$\boldsymbol{\delta} = \mathbf{g}_P - \boldsymbol{\gamma}_P. \tag{2-141}$$

Accordingly, the difference in magnitude is the *gravity disturbance*

$$\delta g = g_P - \gamma_P. \tag{2-142}$$

The difference in direction—that is, the deflection of the vertical—is the same as before, since the directions of γ_P and γ_Q practically coincide.

The gravity disturbance is conceptually even simpler than the gravity anomaly, but it is not as important in terrestrial geodesy. The significance of the gravity anomaly is that it is given directly: the gravity g is measured on the geoid (or reduced to it, see Chapter 3), and the normal gravity γ is computed for the ellipsoid.

Relations. There are several basic mathematical relations between the quantities just defined. Since

$$U_P = U_Q + \left(\frac{\partial U}{\partial n}\right)_Q N = U_Q - \gamma N,$$

we have

$$W_P = U_P + T_P = U_Q - \gamma N + T.$$

Because

$$W_P = U_Q = W_0,$$

we find

$$T = \gamma N \tag{2-143}$$

or

$$N = \frac{T}{\gamma}. \tag{2-144}$$

This is the famous *Bruns formula*, which relates the geoidal undulation to the disturbing potential.

Next we consider the gravity disturbance. Since

$$\mathbf{g} = \operatorname{grad} W,$$
$$\boldsymbol{\gamma} = \operatorname{grad} U,$$

the gravity disturbance vector (2-141) becomes

$$\boldsymbol{\delta} = \operatorname{grad}(W - U) = \operatorname{grad} T \equiv \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z}\right). \tag{2-145}$$

Then

$$g = -\frac{\partial W}{\partial n}, \quad \gamma = -\frac{\partial U}{\partial n'} \doteq -\frac{\partial U}{\partial n},$$

because the directions of the normals n and n' almost coincide. Therefore, the gravity disturbance is given by

$$\delta g = g_P - \gamma_P = -\left(\frac{\partial W}{\partial n} - \frac{\partial U}{\partial n'}\right) \doteq -\left(\frac{\partial W}{\partial n} - \frac{\partial U}{\partial n}\right)$$

or

$$\delta g = -\frac{\partial T}{\partial n}. \tag{2-146}$$

Since the elevation h is reckoned along the normal, we may also write

$$\delta g = -\frac{\partial T}{\partial h}. \tag{2-146$'$}$$

Comparing (2-146) with (2-145), we see that the gravity disturbance δg, besides being the difference in magnitude of the actual and the normal gravity vector, is also the *normal component of the gravity disturbance vector* $\boldsymbol{\delta}$.

We now turn to the gravity anomaly Δg. Since

$$\gamma_P = \gamma_Q + \frac{\partial \gamma}{\partial h} N,$$

we have

$$-\frac{\partial T}{\partial h} = \delta g = g_P - \gamma_P = g_P - \gamma_Q - \frac{\partial \gamma}{\partial h} N.$$

Remembering the definition (2-139) of the gravity anomaly and taking Bruns' formula (2-144) into account, we find the following equivalent equations:

$$-\frac{\partial T}{\partial h} = \Delta g - \frac{\partial \gamma}{\partial h} N, \tag{2-147a}$$

$$\Delta g = -\frac{\partial T}{\partial h} + \frac{\partial \gamma}{\partial h} N, \tag{2-147b}$$

$$\Delta g = -\frac{\partial T}{\partial h} + \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} T, \tag{2-147c}$$

$$\delta g = \Delta g - \frac{\partial \gamma}{\partial h} N, \tag{2-147d}$$

$$\delta g = \Delta g - \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} T, \tag{2-147e}$$

relating different quantities of the anomalous gravity field.

Another equivalent form is

$$\frac{\partial T}{\partial h} - \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} T + \Delta g = 0. \tag{2-148}$$

This expression is called the *fundamental equation of physical geodesy*, because it relates the measured quantity Δg to the unknown anomalous potential T.

It has the form of a partial differential equation. If Δg were known throughout space, then (2-148) could be discussed and solved as a real partial differential equation. However, since Δg is known only along a surface (the geoid), the fundamental equation (2-148) can be used only as a *boundary condition*, which alone is not sufficient for computing T. Therefore, the name "differential equation of physical geodesy," which is sometimes used for (2-148), is rather misleading.

One usually assumes that *there are no masses outside the geoid*. Of course this is not really true. But neither do we make observations directly on the geoid; we make them on the physical surface of the earth. In reducing the measured gravity to the geoid, the effect of the masses outside the geoid is removed by computation, so that we can indeed assume that all masses are enclosed by the geoid (see Chapters 3 and 8).

In this case, since the density ρ is zero everywhere outside the geoid, the anomalous potential T is harmonic there and satisfies Laplace's equation

$$\Delta T \equiv \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = 0.$$

This is, of course, a true partial differential equation and suffices, if supplemented by the boundary condition (2-148), for determining T at every point outside the geoid.

If we write the boundary condition in the form

$$-\frac{\partial T}{\partial n} + \frac{1}{\gamma}\frac{\partial \gamma}{\partial n} T = \Delta g, \tag{2-148'}$$

where Δg is assumed to be known at every point of the geoid, then we see that a linear combination of T and $\partial T/\partial n$ is given upon that surface. According to Sec. 1-17, the determination of T is therefore a *third boundary-value problem of potential theory*. If it is solved for T, then the geoidal height, which is the most important geometric quantity in physical geodesy, can be computed by Bruns' formula (2-144).

We may therefore say that the basic problem of physical geodesy, the determination of the geoid from gravity measurements, is essentially a third boundary-value problem of potential theory.

2-14. Spherical Approximation. Expansion of the Disturbing Potential in Spherical Harmonics

The reference ellipsoid deviates from a sphere only by quantities of the order of the flattening, $f \doteq 3 \times 10^{-3}$. Therefore, if we treat the reference ellipsoid as a sphere in equations relating quantities of the anomalous field, this may cause a relative error of the order of 3×10^{-3}. This error is usually permissible in N, T, Δg, etc. For instance, the absolute effect of this relative error on the geoidal height is of the order of $3 \times 10^{-3} N$; since N hardly exceeds 100 meters, this error can usually be expected to be less than 1 meter.

As a spherical approximation we have

$$\gamma = \frac{kM}{r^2}, \quad \frac{\partial \gamma}{\partial h} = \frac{\partial \gamma}{\partial r} = -2\frac{kM}{r^3}, \quad \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} = -\frac{2}{r}.$$

We shall introduce a mean radius R of the earth. It is often defined as the radius of a sphere that has the same volume as the earth ellipsoid; from the condition

$$\frac{4}{3}\pi R^3 = \frac{4}{3}\pi a^2 b$$

we get

$$R = \sqrt[3]{a^2 b}.$$

In a similar way we may define a mean value G of gravity over the earth. Numerical values of about

$$R = 6371 \text{ km}, \quad G = 979.8 \text{ gals} \tag{2-149}$$

are usually used. Then

$$\frac{1}{\gamma}\frac{\partial \gamma}{\partial h} = -\frac{2}{R}, \tag{2-150}$$

$$\frac{\partial \gamma}{\partial h} = -\frac{2G}{R}. \tag{2-150'}$$

Since the normal to the sphere is the direction of the radius vector r, we have to the same approximation

$$\frac{\partial}{\partial n} = \frac{\partial}{\partial h} = \frac{\partial}{\partial r}.$$

In Bruns' theorem (2-144) we may replace γ by G, and equations (2-147) and (2-148) become

$$-\frac{\partial T}{\partial h} = \Delta g + \frac{2G}{R}N, \tag{2-151a}$$

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2G}{R} N, \tag{2-151b}$$

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{R} T, \tag{2-151c}$$

$$\delta g = \Delta g + \frac{2G}{R} N, \tag{2-151d}$$

$$\delta g = \Delta g + \frac{2}{R} T; \tag{2-151e}$$

$$\frac{\partial T}{\partial r} + \frac{2}{R} T + \Delta g = 0. \tag{2-151f}$$

The last equation is the spherical approximation of the fundamental boundary condition.

The exact meaning of this spherical approximation should be carefully kept in mind. It is used only in equations relating the small quantities T, N, Δg, etc. The reference surface is never a sphere in any geometrical sense, but always an ellipsoid. As the flattening f is very small, the ellipsoidal formulas can be expanded into power series in terms of f, and then all terms containing f, f^2, etc., are neglected. In this way one obtains formulas that are rigorously valid for the sphere, but approximately valid for the actual reference ellipsoid as well. However, normal gravity γ in the gravity anomaly $\Delta g = g - \gamma$ must be computed for the ellipsoid to a high degree of accuracy.

Since the anomalous potential $T = W - U$ is a harmonic function, it can be expanded into a series of spherical harmonics:

$$T(r, \theta, \lambda) = \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+1} T_n(\theta, \lambda). \tag{2-152}$$

$T_n(\theta, \lambda)$ is Laplace's surface harmonic of degree n. On the geoid, which as a spherical approximation corresponds to the sphere $r = R$, we have formally

$$T = T(R, \theta, \lambda) = \sum_{n=0}^{\infty} T_n(\theta, \lambda) \tag{2-152$'$}$$

(we need not be concerned with questions of convergence here).

Differentiating the series (2-152) with respect to r we find

$$\delta g = -\frac{\partial T}{\partial r} = \frac{1}{r} \sum_{n=0}^{\infty} (n + 1)\left(\frac{R}{r}\right)^{n+1} T_n(\theta, \lambda). \tag{2-153}$$

On the geoid ($r = R$) this becomes

$$\delta g = -\frac{\partial T}{\partial r} = \frac{1}{R} \sum_{n=0}^{\infty} (n + 1) T_n(\theta, \lambda). \tag{2-153$'$}$$

These series express the gravity disturbance in terms of spherical harmonics.

The equivalent of (2-151c) outside the earth is obviously

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{r}T. \tag{2-154}$$

Its exact meaning will be discussed at the end of the following section. The insertion of (2-153) and (2-152) into this equation yields

$$\Delta g = \frac{1}{r}\sum_{n=0}^{\infty}(n-1)\left(\frac{R}{r}\right)^{n+1}T_n(\theta, \lambda). \tag{2-155}$$

On the geoid this becomes

$$\Delta g = \frac{1}{R}\sum_{n=0}^{\infty}(n-1)T_n(\theta, \lambda). \tag{2-155'}$$

This is the spherical-harmonic expansion of the gravity anomaly.

Note that even if the anomalous potential T contains a first-degree spherical term $T_1(\theta, \lambda)$, it will in the expression for Δg *be multiplied by the factor* $1 - 1 = 0$, *so that* Δg *can never have a first-degree spherical harmonic*—even if T has one.

2-15. Gravity Anomalies Outside the Earth

If a harmonic function H is given at the surface of the earth, then, as a spherical approximation, the values of H outside the earth can be computed by Poisson's integral formula (1-89)

$$H_P = \frac{R}{4\pi}\iint_\sigma \frac{r^2 - R^2}{l^3} H\, d\sigma.$$

The symbol $\iint_\sigma$ is the usual abbreviation for an integral extended over the whole unit sphere, or over the full solid angle, which is the same; $d\sigma$ denotes the element of solid angle, defined as the surface element of the unit sphere. Hence the surface element of the terrestrial sphere $r = R$ is $R^2\, d\sigma$. The meaning of the other notations is read from Fig. 2-14. The value of the harmonic function at the

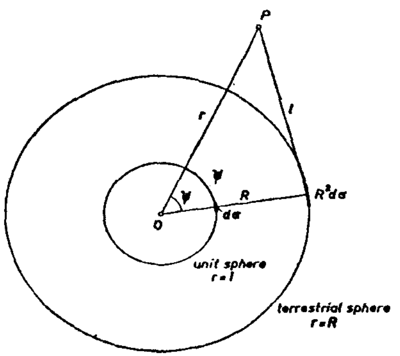

FIGURE 2-14

Notations for Poisson's integral and derived formulas.

variable surface element $R^2\,d\sigma$ is denoted simply by H, whereas H_P refers to the fixed point P. Obviously, then,

$$l = \sqrt{r^2 + R^2 - 2Rr\cos\psi}. \tag{2-156}$$

The harmonic function H can be expanded into a series of spherical harmonics:

$$H = \left(\frac{R}{r}\right)H_0 + \left(\frac{R}{r}\right)^2 H_1 + \sum_{n=2}^{\infty}\left(\frac{R}{r}\right)^{n+1} H_n.$$

By omitting the terms of degrees one and zero we get a new function

$$H' = H - \left(\frac{R}{r}\right)H_0 - \left(\frac{R}{r}\right)^2 H_1 = \sum_{n=2}^{\infty}\left(\frac{R}{r}\right)^{n+1} H_n. \tag{2-157}$$

The surface harmonics are given by

$$H_0 = \frac{1}{4\pi}\iint_\sigma H\,d\sigma, \quad H_1 = \frac{3}{4\pi}\iint_\sigma H\cos\psi\,d\sigma, \tag{2-158}$$

according to equation (1-71). Hence we find from (2-157), on expressing H by Poisson's integral and substituting the integrals (2-158) for H_0 and H_1, the basic formula

$$H'_P = \frac{R}{4\pi}\iint_\sigma \left(\frac{r^2 - R^2}{l^3} - \frac{1}{r} - \frac{3R}{r^2}\cos\psi\right) H\,d\sigma. \tag{2-159}$$

The reason for this modification of Poisson's integral is that the formulas of physical geodesy are simpler if the functions involved do not contain harmonics of degrees zero and one. It is therefore convenient to split these terms off. This is done automatically by the modified Poisson integral (2-159).

We shall now apply these formulas to the gravity anomalies outside the earth. Equation (2-155) yields at once

$$r\,\Delta g = \sum_{n=0}^{\infty}\left(\frac{R}{r}\right)^{n+1}(n-1)T_n(\theta, \lambda).$$

Just as $T_n(\theta, \lambda)$ is a Laplace surface harmonic, so is $(n-1)T_n$. Consequently $r\,\Delta g$, considered as a function in space, can be expanded into a series of spherical harmonics and *is therefore a harmonic function.*

Hence we can apply Poisson's formula to $r\,\Delta g$, getting

$$r\,\Delta g_P = \frac{R}{4\pi}\iint_\sigma \left(\frac{r^2 - R^2}{l^3} - \frac{1}{r} - \frac{3R}{r^2}\cos\psi\right)(R\,\Delta g)\,d\sigma,$$

or

$$\Delta g_P = \frac{R^2}{4\pi r}\iint_\sigma \left(\frac{r^2 - R^2}{l^3} - \frac{1}{r} - \frac{3R}{r^2}\cos\psi\right)\Delta g\,d\sigma. \tag{2-160}$$

This is the formula for the computation of gravity anomalies outside the earth from surface gravity anomalies, or for the *upward continuation of gravity anomalies.*

Finally we shall discuss the exact meaning of the gravity anomaly Δg_P outside the earth. We start with a convenient definition. The level surfaces of the actual gravity potential, the surfaces

$$W = \text{const.},$$

are often called *geopotential surfaces*; the level surfaces of the normal gravity field, the surfaces

$$U = \text{const.},$$

are called *spheropotential surfaces.*

We consider now the point P outside the earth (Fig. 2-15) and denote the geopotential surface passing through it by

$$W = W_P.$$

There is also a spheropotential surface

$$U = W_P$$

of the same constant W_P. The normal plumb line through P intersects this spheropotential surface at the point Q, which is said to correspond to P.

We see that the level surfaces $W = W_P$ and $U = W_P$ are related to each other in exactly the same way as are the geoid $W = W_0$ and the reference ellipsoid $U = W_0$. If, therefore, the gravity anomaly is defined by

$$\Delta g_P = g_P - \gamma_Q,$$

as in Sec. 2-13, then all derivations and formulas of that section also hold for the present situation, the geopotential surface $W = W_P$ replacing the geoid $W = W_0$, and the spheropotential surface $U = W_P$ replacing the ellipsoid $U = W_0$. This is also the reason why (2-154) holds at P as well as at the geoid.

Note that in Sec. 2-13 P is a point at the geoid, which is denoted by P_0 in Fig. 2-15.

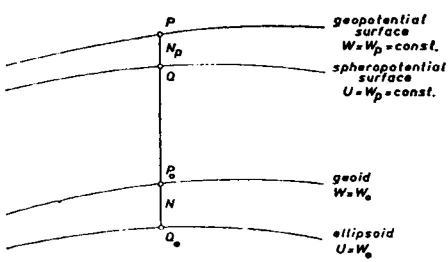

FIGURE 2-15

Geopotential and spheropotential surfaces.

2-16. Stokes' Formula

The basic equation (2-154),

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{r} T,$$

can be considered as a *boundary condition* only, as long as the gravity anomalies Δg are known only at the surface of the earth. However, by the upward continuation integral (2-160) we are now able to compute the gravity anomalies outside the earth. Thus our basic equation changes its meaning radically, becoming a real differential equation that can be integrated with respect to r.[1]

Multiplying by $-r^2$ we get

$$-r^2 \Delta g = r^2 \frac{\partial T}{\partial r} + 2rT = \frac{\partial}{\partial r}(r^2 T).$$

Integrating the formula

$$\frac{\partial}{\partial r}(r^2 T) = -r^2 \Delta g(r)$$

between the limits ∞ and r we find

$$r^2 T\Big|_{\infty}^{r} = -\int_{\infty}^{r} r^2 \Delta g(r)\, dr,$$

where $\Delta g(r)$ indicates that Δg is now a function of r, computed from surface gravity anomalies by means of the formula (2-160). Since this formula automatically removes the spherical harmonics of degrees one and zero from $\Delta g(r)$, the anomalous potential T, as computed from $\Delta g(r)$, cannot contain such terms. Thus we have

$$T = \sum_{n=2}^{\infty} \left(\frac{R}{r}\right)^{n+1} T_n = \frac{R^3}{r^3} T_2 + \frac{R^4}{r^4} T_3 + \cdots.$$

Therefore,

$$\lim_{r\to\infty} (r^2 T) = \lim_{r\to\infty} \left(\frac{R^3}{r} T_2 + \frac{R^4}{r^2} T_3 + \cdots\right) = 0,$$

so that

$$r^2 T\Big|_{\infty}^{r} = r^2 T - \lim_{r\to\infty} (r^2 T) = r^2 T.$$

Hence,[2]

$$r^2 T = -\int_{\infty}^{r} r^2 \Delta g(r)\, dr,$$

[1] Note that this is made possible only because T, in addition to the boundary condition, satisfies Laplace's equation $\Delta T = 0$.

[2] The fact that r is used both as an integration variable and as an upper limit should not cause any difficulty.

and on inserting the upward continuation integral (2-160) we get

$$r^2T = \frac{R^2}{4\pi}\int_\infty^r \left[\iint_\sigma \left(-\frac{r^3 - R^2r}{l^3} + 1 + \frac{3R}{r}\cos\psi\right)\Delta g\, d\sigma\right] dr.$$

Interchanging the order of the integrations gives

$$r^2T = \frac{R^2}{4\pi}\iint_\sigma \left[\int_\infty^r \left(-\frac{r^3 - R^2r}{l^3} + 1 + \frac{3R}{r}\cos\psi\right) dr\right]\Delta g\, d\sigma.$$

The integral in brackets can be evaluated by standard methods. The indefinite integral is[1]

$$\int \left(-\frac{r^3 - R^2r}{l^3} + 1 + \frac{3R}{r}\cos\psi\right) dr$$

$$= \frac{2r^2}{l} - 3l - 3R\cos\psi\ln(r - R\cos\psi + l) + r + 3R\cos\psi\ln r.$$

For large values of r we have

$$l = r\left(1 - \frac{R}{r}\cos\psi\cdots\right) = r - R\cos\psi\cdots,$$

and hence we find that as $r \longrightarrow \infty$, the right-hand side of the above indefinite integral approaches

$$5R\cos\psi - 3R\cos\psi\ln 2.$$

If we subtract this from the indefinite integral we get the definite integral, since infinity is its lower limit of integration. Thus

$$\int_\infty^r \left(-\frac{r^3 - R^2r}{l^3} + 1 + \frac{3R}{r}\cos\psi\right) dr$$

$$= \frac{2r^2}{l} + r - 3l - R\cos\psi\left(5 + 3\ln\frac{r - R\cos\psi + l}{2r}\right).$$

Hence we obtain

$$T(r, \theta, \lambda) = \frac{R}{4\pi}\iint_\sigma S(r, \psi)\,\Delta g\, d\sigma, \tag{2-161}$$

where

$$S(r, \psi) = \frac{2R}{l} + \frac{R}{r} - 3\frac{Rl}{r^2} - \frac{R^2}{r^2}\cos\psi\left(5 + 3\ln\frac{r - R\cos\psi + l}{2r}\right). \tag{2-162}$$

On the geoid itself we have $r = R$, and denoting $T(R, \theta, \lambda)$ simply by T, we find

$$T = \frac{R}{4\pi}\iint_\sigma \Delta g\, S(\psi)\, d\sigma, \tag{2-163a}$$

[1] The reader is advised to perform this integration, taking (2-156) into account, or at least to check the result by differentiating the right-hand side with respect to r.

where

$$S(\psi) = \frac{1}{\sin(\psi/2)} - 6\sin\frac{\psi}{2} + 1 - 5\cos\psi - 3\cos\psi\ln\left(\sin\frac{\psi}{2} + \sin^2\frac{\psi}{2}\right) \quad (2\text{-}164)$$

is obtained from $S(r, \psi)$ by setting

$$r = R \quad \text{and} \quad l = 2R\sin\frac{\psi}{2}.$$

By Bruns' theorem, $N = T/G$, we finally get

$$N = \frac{R}{4\pi G}\iint_\sigma \Delta g\, S(\psi)\, d\sigma. \quad (2\text{-}163\text{b})$$

This formula was published by George Gabriel Stokes in 1849; it is therefore called *Stokes' formula*, or *Stokes' integral*. It is by far the most important formula of physical geodesy because it makes it possible *to determine the geoid from gravity data*. Equation (2-163a) is also called Stokes' formula, and $S(\psi)$ is known as Stokes' function. This and related functions are tabulated in Lambert and Darling (1936).

Using formula (2-161), which was derived by Pizzetti (1911) and later on by Vening Meinesz (1928), we can compute the anomalous potential T at any point outside the earth. On dividing T by the normal gravity at the given point P (Bruns' theorem) we obtain the separation N_P between the geopotential surface $W = W_P$ and the corresponding spheropotential surface $U = W_P$, which, outside the earth, takes the place of the geoidal undulation N. (See Fig. 2-15 and the explanations at the end of the preceding section.)

We mention again that these formulas are based on a spherical approximation; quantities of the order of $3 \times 10^{-3} N$ are neglected. This results in an error of probably less than 1 meter in N, which can be neglected for most practical purposes. Zagrebin, Molodensky, and Bjerhammar have developed higher approximations, which take into account the flattening f of the reference ellipsoid; see Sagrebin (1956), Molodenskii et al. (1962, p. 53), and Bjerhammar (1962).

We next see from the derivation of Stokes' formula by means of the upward continuation integral (2-160) that it automatically suppresses the harmonic terms of degrees one and zero in T and N. The implications of this will be discussed later. We shall see that Stokes' formula in its original form (2-163a, b) holds only for a reference ellipsoid that (1) has the same potential $U_0 = W_0$ as the geoid, (2) encloses a mass that is numerically equal to the earth's mass, and (3) has its center at the center of gravity of the earth. Since the first two conditions are not accurately satisfied by the reference ellipsoids that are in current practical use, and can hardly ever be rigorously fulfilled, Stokes' formula must later be modified for the case of an arbitrary reference ellipsoid.

Finally, T is assumed to be harmonic outside the geoid. This means that the effect of the masses above the geoid must be removed by suitable gravity reductions. This will be discussed in Chapter 3.

2-17. Explicit Forms of Stokes' Integral. Expansion of Stokes' Function in Spherical Harmonics

We shall now write Stokes' formula (2-163b) more explicitly by introducing suitable coordinate systems on the sphere.

The use of spherical *polar coordinates* with origin at P offers the advantage that the angle ψ, which is the argument of Stokes' function, is one coordinate, the *spherical distance*. The other coordinate is the *azimuth* α, reckoned from north. Their definitions are seen in Fig. 2-16. Denoting by P both a fixed point on the sphere $r = R$ (or in space) and its projection on the unit sphere is common practice and will not cause any difficulty.

If P coincides with the north pole, then ψ and α are identical with θ and λ. According to Sec. 1-13 the element of solid angle is then given by

$$d\sigma = \sin\psi\, d\psi\, d\alpha.$$

Since all points of the sphere are equivalent, this relation holds for an arbitrary origin P. In the same way we have

$$\iint_\sigma = \int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\pi}.$$

Hence we find

$$N = \frac{R}{4\pi G}\int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\pi}\Delta g(\psi, \alpha)S(\psi)\sin\psi\, d\psi\, d\alpha \tag{2-165}$$

as an explicit form of (2-163b).

Performing the integration with respect to α first, we obtain

$$N = \frac{R}{2G}\int_{\psi=0}^{\pi}\left[\frac{1}{2\pi}\int_{\alpha=0}^{2\pi}\Delta g(\psi, \alpha)\, d\alpha\right]S(\psi)\sin\psi\, d\psi.$$

The expression in brackets is the *average of* Δg *along a parallel* of spherical radius ψ. We denote this average by $\overline{\Delta g}(\psi)$, so that

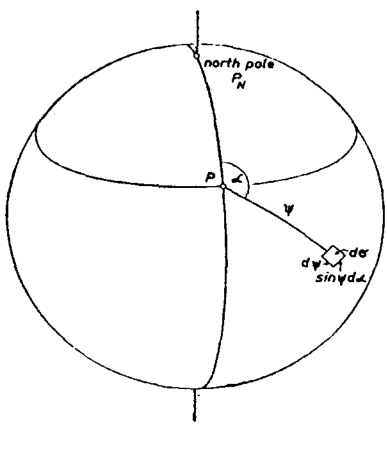

FIGURE 2-16
Polar coordinates on the unit sphere.

$$\overline{\Delta g}(\psi) = \frac{1}{2\pi} \int_{\alpha=0}^{2\pi} \Delta g(\psi, \alpha)\, d\alpha.$$

Hence Stokes' formula may be written

$$N = \frac{R}{G} \int_{\psi=0}^{\pi} \overline{\Delta g}(\psi) F(\psi)\, d\psi, \tag{2-165'}$$

where we have put

$$\frac{1}{2} S(\psi) \sin\psi = F(\psi). \tag{2-166}$$

The functions $S(\psi)$ and $F(\psi)$ are shown in Fig. 2-17.

Alternatively we may use *geographical coordinates* ϕ, λ. As a spherical approximation θ is the complement of geographical latitude:

$$\theta = 90° - \phi, \quad \phi = 90° - \theta.$$

Hence we have

$$\iint_{\sigma} d\sigma = \int_{\lambda=0}^{2\pi} \int_{\phi=-\pi/2}^{\pi/2} \cos\phi\, d\phi\, d\lambda,$$

so that Stokes' formula now becomes

$$N(\phi, \lambda) = \frac{R}{4\pi G} \int_{\lambda'=0}^{2\pi} \int_{\phi'=-\pi/2}^{\pi/2} \Delta g(\phi', \lambda') S(\psi) \cos\phi'\, d\phi'\, d\lambda', \tag{2-167}$$

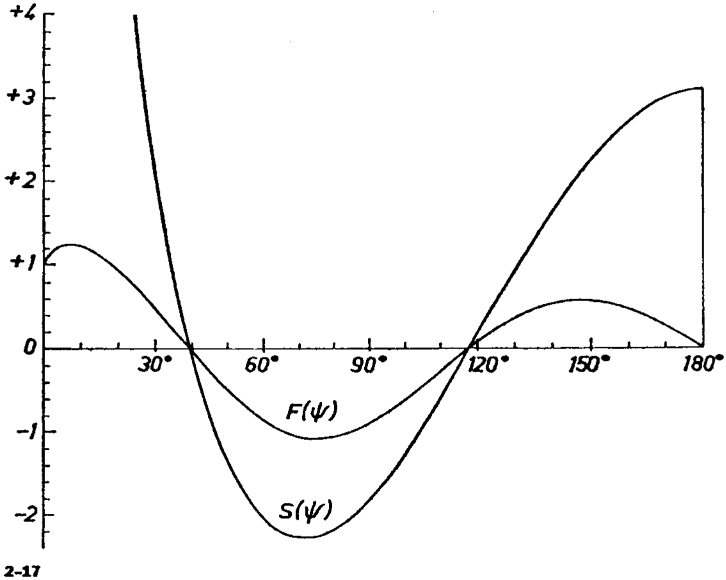

FIGURE 2-17

Stokes' functions $S(\psi)$ and $F(\psi) = \frac{1}{2}S(\psi)\sin\psi$.

where ϕ, λ are the geographical coordinates of the computation point and ϕ', λ' are the coordinates of the variable surface element $d\sigma$. The spherical distance ψ is expressed as a function of these coordinates by

$$\psi = \cos^{-1}[\sin\phi\sin\phi' + \cos\phi\cos\phi'\cos(\lambda' - \lambda)]. \tag{2-168}$$

Stokes' function in terms of spherical harmonics. In Sec. 2-14 we have found

$$\Delta g(\theta, \lambda) = \frac{1}{R}\sum_{n=0}^{\infty}(n-1)T_n(\theta, \lambda).$$

We may also directly express $\Delta g(\theta, \lambda)$ as a series of Laplace surface harmonics:

$$\Delta g(\theta, \lambda) = \sum_{n=0}^{\infty}\Delta g_n(\theta, \lambda).$$

Comparing these two series yields

$$\Delta g_n(\theta, \lambda) = \frac{n-1}{R}T_n(\theta, \lambda), \quad T_n = \frac{R}{n-1}\Delta g_n,$$

so that

$$T = \sum_{n=0}^{\infty}T_n = R\sum_{n=0}^{\infty}\frac{\Delta g_n}{n-1}.$$

This equation shows again that there must be no first-degree term in the spherical-harmonic expansion of Δg; otherwise the term $\Delta g_n/(n-1)$ would be infinite for $n = 1$. As usual, we shall now assume that the harmonics of degrees zero and one are missing. We therefore start the summation with $n = 2$.

Since by equation (1-71)

$$\Delta g_n = \frac{2n+1}{4\pi}\iint_\sigma \Delta g\, P_n(\cos\psi)\, d\sigma,$$

the preceding formula becomes

$$T = \frac{R}{4\pi}\sum_{n=2}^{\infty}\frac{2n+1}{n-1}\iint_\sigma \Delta g\, P_n(\cos\psi)\, d\sigma.$$

By interchanging the order of summation and integration we get

$$T = \frac{R}{4\pi}\iint_\sigma\left[\sum_{n=2}^{\infty}\frac{2n+1}{n-1}P_n(\cos\psi)\right]\Delta g\, d\sigma.$$

Comparing this with Stokes' formula (2-163a) we find the *expression for Stokes' function in terms of Legendre polynomials* (zonal harmonics):

$$S(\psi) = \sum_{n=2}^{\infty}\frac{2n+1}{n-1}P_n(\cos\psi). \tag{2-169}$$

In fact, the analytic expression (2-164) of Stokes' function could have been derived somewhat more simply by direct summation of this series, but we believe that the derivation given in the preceding section is more instructive because it also throws sidelights on important related problems.

2-18. Generalization to an Arbitrary Reference Ellipsoid

As we have seen, Stokes' formula, in its original form, suppresses the spherical harmonics of degrees zero and one in the anomalous potential T and is therefore strictly valid only if these terms are missing. This fact and the condition $U_0 = W_0$ impose on the reference ellipsoid and on its normal gravity field restrictions that are hardly ever fulfilled in practice.

We shall therefore generalize Stokes' formula so that it will apply to an arbitrary ellipsoid of reference, which must satisfy only the condition that it is so close to the geoid that the deviations of the geoid from the ellipsoid can be treated as linear.

Consider the anomalous potential T at the surface of the earth. Its expression in surface spherical harmonics is given by

$$T(\theta, \lambda) = \sum_{n=0}^{\infty} T_n(\theta, \lambda).$$

By separating the terms of degrees zero and one we may write

$$T(\theta, \lambda) = T_0 + T_1(\theta, \lambda) + T'(\theta, \lambda), \tag{2-170}$$

where

$$T'(\theta, \lambda) = \sum_{n=2}^{\infty} T_n(\theta, \lambda). \tag{2-171}$$

In the general case this function T', rather than T itself, is the quantity given by Stokes' formula. It is equal to T only if T_0 and T_1 are missing. Otherwise we have to add T_0 and T_1 in order to get the complete function T.

The zero-degree term in the spherical-harmonic expansion of the potential is equal to

$$\frac{kM}{r},$$

where M is the mass. Hence the zero-degree term of the anomalous potential $T = W - U$ at the surface of the earth ($r = R$) is given by

$$T_0 = \frac{k\delta M}{R}, \tag{2-172}$$

where

$$\delta M = M - M' \tag{2-173}$$

is the difference between the mass M of the earth and the mass M' of the ellipsoid. It would be zero if both masses were equal—but since we do not know the exact mass of the earth how can we make M' equal to M?

Subsequently we shall see that the first-degree harmonic can always be taken to be zero. Assuming this, we can substitute (2-172) into (2-170) and express T' by the conventional Stokes formula (2-163a). Thus we obtain

$$T = \frac{k\delta M}{R} + \frac{R}{4\pi}\iint_\sigma \Delta g \, S(\psi)\, d\sigma. \tag{2-174}$$

This is the generalization of Stokes' formula for T. It holds for an arbitrary reference ellipsoid whose center coincides with the center of the earth.

First-degree terms. The coefficients of the first-degree harmonic in the potential W are, according to (2-44b) and (2-45), given by

$$kM\zeta, \quad kM\xi, \quad kM\eta,$$

where ξ, η, ζ are the rectangular coordinates of the earth's center of gravity. For the normal potential U we have the analogous quantities

$$kM'\zeta', \quad kM'\xi', \quad kM'\eta'.$$

As ξ', η', ζ' are very small in any case, these are practically equal to

$$kM\zeta', \quad kM\xi', \quad kM\eta'.$$

The coefficients of the first-degree harmonic in the anomalous potential $T = W - U$ are therefore equal to

$$kM(\zeta - \zeta'), \quad kM(\xi - \xi'), \quad kM(\eta - \eta'). \tag{2-175}$$

They are zero, and *there is no first-degree harmonic $T_1(\theta, \lambda)$ if the center of the reference ellipsoid coincides with the center of gravity of the earth.* This is usually assumed.

In the general case we find from the first-degree term of (2-37), on putting $r = R$ and using the coefficients (2-44b) together with (2-45),

$$T_1(\theta, \lambda) = \frac{kM}{R^2}\,[(\zeta - \zeta')P_{10}(\cos\theta) + (\xi - \xi')P_{11}(\cos\theta)\cos\lambda$$
$$+ (\eta - \eta')P_{11}(\cos\theta)\sin\lambda].$$

If the origin of the coordinate system is taken to be the center of the reference ellipsoid, then $\xi' = \eta' = \zeta' = 0$. With $P_{10}(\cos\theta) = \cos\theta$, $P_{11}(\cos\theta) = \sin\theta$, and $kM/R^2 \doteq G$ we then obtain the following expression for the first-degree harmonic of T:

$$T_1(\theta, \lambda) = G(\xi\sin\theta\cos\lambda + \eta\sin\theta\sin\lambda + \zeta\cos\theta). \tag{2-176a}$$

By dividing by G we find the first-degree harmonic of the geoidal height:

$$N_1(\theta, \lambda) = \xi\sin\theta\cos\lambda + \eta\sin\theta\sin\lambda + \zeta\cos\theta, \tag{2-176b}$$

where ξ, η, ζ are the rectangular coordinates of the earth's center of gravity, the origin being the center of the reference ellipsoid.

On introducing the vector

$$\boldsymbol{\xi} = (\xi, \eta, \zeta)$$

and the unit vector of the direction (θ, λ),

$$\mathbf{e} = (\sin\theta\cos\lambda,\ \sin\theta\sin\lambda,\ \cos\theta),$$

(2-176b) may be written as

$$N_1(\theta, \lambda) = \boldsymbol{\xi}\cdot\mathbf{e}, \tag{2-177}$$

which is interpreted as the projection of the vector $\boldsymbol{\xi}$ onto the direction (θ, λ).

Hence if the two centers of gravity do not coincide, then we need only add the first-degree terms (2-176a) and (2-176b) to the generalized Stokes formula (2-174) and to its analogue for N [equation (2-181) below], respectively, in order to get the most general solution for Stokes' problem, the computation of T and N from Δg. Equation (2-155′) shows that *any* value of $T_1(\theta, \lambda)$ is compatible with a given Δg field because, for $n = 1$, the quantity $(n - 1)T_1$ is zero and so T_1, whatever be its value, does not at all enter into Δg.

Hence the most general solution for T and N contains three arbitrary constants ξ, η, ζ, which can thus be regarded as the constants of integration for Stokes' problem. In actual practice one always sets $\xi = \eta = \zeta = 0$, thus placing the center of the reference ellipsoid at the center of the earth. This constitutes the enormous advantage of the gravimetric determination of the geoid over the astrogeodetic method, where the position of the reference ellipsoid with respect to the center of the earth remains unknown.

2-19. Generalization of Stokes' Formula for N

Let us first extend Bruns' formula (2-144) to an arbitrary reference ellipsoid. Suppose

$$W(x, y, z) = W^0,$$
$$U(x, y, z) = U^0$$

are the equations of the geoid and the ellipsoid, where in general the constants W^0 and U^0 are different; we have written W^0, U^0 instead of W_0, U_0 in order to avoid confusion with a zero-degree harmonic. As in Section 2-13 we have, using Fig. 2-12,

$$W_P = U_Q - \gamma N + T,$$

$$\text{but now} \quad U_Q = U^0 \neq W^0 = W_P, \quad \text{so that}$$

$$\gamma N = T - (W^0 - U^0).$$

Denoting the difference between the potentials by

$$\delta W = W^0 - U^0$$

we obtain the following simple generalization of Bruns' formula

$$N = \frac{T - \delta W}{\gamma}. \tag{2-178}$$

We shall also need the extension of equations (2-147a–e). Those formulas which contain N instead of T are easily seen to hold for an arbitrary reference ellipsoid as well, but the transition from N to T is now effected by means of (2-178). Hence (2-147b),

$$\Delta g = -\frac{\partial T}{\partial h} + \frac{\partial \gamma}{\partial h} N,$$

remains unchanged, but (2-147c) becomes

$$\Delta g = -\frac{\partial T}{\partial h} + \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} T - \frac{1}{\gamma}\frac{\partial \gamma}{\partial h}\delta W. \tag{2-179}$$

Therefore, the fundamental boundary condition is now

$$-\frac{\partial T}{\partial h} + \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} T = \Delta g + \frac{1}{\gamma}\frac{\partial \gamma}{\partial h}\delta W. \tag{2-180}$$

The spherical approximations of these equations are

$$N = \frac{T - \delta W}{G}, \tag{2-178'}$$

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{R} T + \frac{2}{R}\delta W, \tag{2-179'}$$

$$-\frac{\partial T}{\partial r} - \frac{2}{R} T = \Delta g - \frac{2}{R}\delta W. \tag{2-180'}$$

Various forms of the generalized Stokes formula. By (2-178) we have

$$T = GN + \delta W.$$

Inserting this into (2-174) and dividing by G we obtain

$$N = \frac{k\delta M}{RG} - \frac{\delta W}{G} + \frac{R}{4\pi G}\iint_\sigma \Delta g\, S(\psi)\, d\sigma. \tag{2-181}$$

This is the generalization of Stokes' formula for N. It holds for an arbitrary reference ellipsoid whose center coincides with the center of the earth.

Whereas the formula (2-174) for T contains only the effect of a mass difference δM, the formula (2-181) for N contains, in addition, the potential difference δW. These formulas also show clearly that the simple Stokes integrals (2-163a, b) hold only if $\delta M = \delta W = 0$, that is, *if the reference ellipsoid has the same potential as the geoid and the same mass as the earth.* Otherwise they give N and T only up to additive constants: putting

$$N_0 = \frac{k\delta M}{RG} - \frac{\delta W}{G} \tag{2-182}$$

and taking (2-172) into account, we have

$$T = T_0 + \frac{R}{4\pi}\iint_\sigma \Delta g\, S(\psi)\, d\sigma, \tag{2-183a}$$

$$N = N_0 + \frac{R}{4\pi G} \iint_\sigma \Delta g \, S(\psi) \, d\sigma. \tag{2-183b}$$

Alternative forms of (2-181), which are sometimes useful, are obtained in the following way. Inserting the series (2-152′) and (2-153′) into (2-179′), we get

$$\Delta g(\theta, \lambda) = \frac{1}{R} \sum_{n=0}^{\infty} (n - 1) T_n(\theta, \lambda) + \frac{2}{R} \delta W \tag{2-184a}$$

as the generalization of (2-155′). Expanding the function $\Delta g(\theta, \lambda)$ into the usual series of Laplace surface spherical harmonics,

$$\Delta g(\theta, \lambda) = \sum_{n=0}^{\infty} \Delta g_n(\theta, \lambda), \tag{2-184b}$$

and comparing the constant terms ($n = 0$) of these two equations, we get

$$-\frac{1}{R} T_0 + \frac{2}{R} \delta W = \Delta g_0,$$

where, by (1-71),

$$\Delta g_0 = \frac{1}{4\pi} \iint_\sigma \Delta g \, d\sigma. \tag{2-185}$$

Expressing T_0 by (2-172) in terms of δM, we obtain

$$\Delta g_0 = -\frac{1}{R^2} k\delta M + \frac{2}{R} \delta W. \tag{2-186}$$

The two equations for N_0 (2-182) and for Δg_0 (2-186) can now be solved for δM and δW:

$$k\delta M = R(R \, \Delta g_0 + 2GN_0), \tag{2-187a}$$

$$\delta W = R \, \Delta g_0 + GN_0. \tag{2-187b}$$

The constant N_0 may be expressed by either of these equations:

$$N_0 = -\frac{R}{2G} \Delta g_0 + \frac{k\delta M}{2GR} = -\frac{R}{8\pi G} \iint_\sigma \Delta g \, d\sigma + \frac{k\delta M}{2GR},$$

$$N_0 = -\frac{R}{G} \Delta g_0 + \frac{\delta W}{G} = -\frac{R}{4\pi G} \iint_\sigma \Delta g \, d\sigma + \frac{\delta W}{G}.$$

On inserting these into (2-183b) we obtain

$$N = \frac{R}{4\pi G} \iint_\sigma \Delta g \left[S(\psi) - \frac{1}{2} \right] d\sigma + \frac{k\delta M}{2GR}, \tag{2-188}$$

$$N = \frac{R}{4\pi G} \iint_\sigma \Delta g \, [S(\psi) - 1] \, d\sigma + \frac{\delta W}{G}. \tag{2-189}$$

These formulas are completely equivalent to (2-181); they also hold for an arbitrary reference ellipsoid.

If $M' = M$, even if $U^0 \neq W^0$, we have

$$N = \frac{R}{4\pi G} \iint_\sigma \Delta g \left[S(\psi) - \frac{1}{2} \right] d\sigma, \tag{2-188'}$$

and if $U^0 = W^0$, even if $M' \neq M$, we have

$$N = \frac{R}{4\pi G} \iint_\sigma \Delta g \, [S(\psi) - 1] \, d\sigma. \tag{2-189'}$$

These formulas are slightly more general than the simple Stokes integral, insofar as only *one* of the conditions $M' = M$, $U^0 = W^0$ is presupposed. Equation (2-188') was derived by Pizzetti, and (2-189') by Hirvonen.

Determination of N_0. If the mass M of the earth and the potential W^0 of the geoid were accurately known, then N_0 could be computed by (2-182). The geoidal undulations N could then be calculated accurately by Stokes' formula (2-183b). By applying N to the fixed reference ellipsoid, the geoid would be given absolutely, in its proper scale of length, *without measuring a single distance.*

In practice, of course, we do not know the values of M and W^0 to an accuracy sufficient to enable us to determine N_0. If we evaluate only the original Stokes integral

$$N' = \frac{R}{4\pi G} \iint_\sigma \Delta g \, S(\psi) \, d\sigma, \tag{2-190}$$

then we obtain, instead of the geoid S, merely a surface S' that is parallel to the geoid at the distance N_0 (Fig. 2-18a). Since both surfaces are almost spherical, they are, to a high degree of accuracy, geometrically similar; that is, they differ

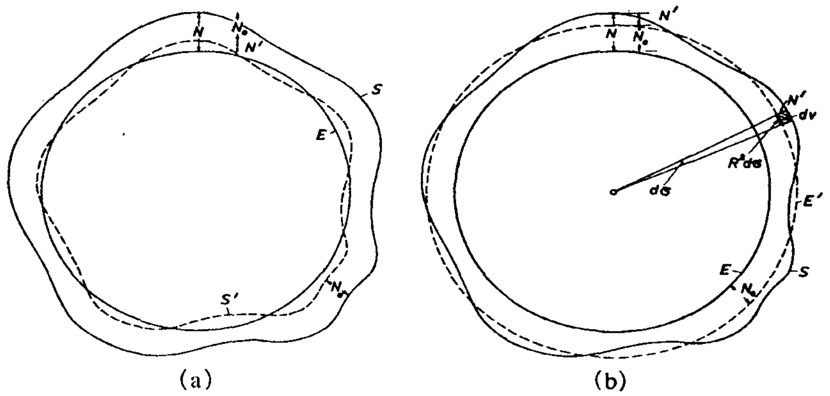

FIGURE 2-18

Two interpretations of Stokes' formula. (a) N' is the height above the ellipsoid E of the surface S' parallel to the geoid. (b) N' is the height of the geoid S above the modified ellipsoid E' parallel to E.

only in the scale. We can thus say that the original Stokes integral (2-190) yields a geoid that lacks only a scale factor. *This scale factor can be determined by a single distance measurement*, whereupon the constant N_0 also is known. This will now be elaborated mathematically.

Let P_1 and P_2 be two geoidal points, and Q_1 and Q_2 be their projections onto the reference ellipsoid (Fig. 2-19); s is the distance between P_1 and P_2 along the geoid, and s' is the distance between Q_1 and Q_2 along the ellipsoid.

We shall derive the relationship between s, s', and N. If we replace the ellipsoidal arc $s' = Q_1Q_2$ by a spherical arc whose radius R is a mean radius of curvature, then Fig. 2-19 shows that

$$\frac{ds \cos \epsilon}{R + N} = \frac{ds'}{R}.$$

Since $\cos \epsilon \doteq 1$ we have

$$ds = ds'\left(1 + \frac{N}{R}\right) = ds' + \frac{N}{R}\,ds' \doteq ds' + \frac{N}{R}\,ds.$$

On integrating we obtain

$$s = s' + \frac{1}{R}\int_{Q_1}^{Q_2} N\,ds, \tag{2-191}$$

which is the desired relation between s, s', and N.

By inserting $N = N_0 + N'$ we find

$$s - s' = \frac{1}{R}\int_{Q_1}^{Q_2} (N_0 + N')\,ds = \frac{1}{R}\int_{Q_1}^{Q_2} N'\,ds + \frac{s}{R}\,N_0,$$

so that

$$N_0 = \frac{R}{s}(s - s') - \frac{1}{s}\int_{Q_1}^{Q_2} N'\,ds. \tag{2-192}$$

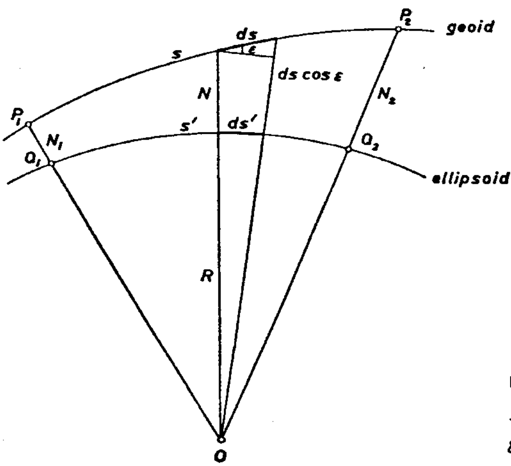

FIGURE 2-19

Determination of the scale of the geoid.

The quantity N' is given by Stokes' integral (2-190). We consider the distance s to be measured on the geoid or reduced to it. The ellipsoidal distance s' can be computed if the coordinates ϕ, λ of its end points Q_1 and Q_2 are given. From equations (2-140) we obtain

$$\begin{aligned}\phi &= \Phi - \xi, \\ \lambda &= \Lambda - \frac{\eta}{\cos\phi}.\end{aligned} \tag{2-193}$$

The astronomical coordinates Φ and Λ are directly measured; the components ξ and η of the deflection of the vertical can be computed from Δg by Vening Meinesz' formula (see Sec. 2-22), so that ϕ and λ will be known.

Thus N_0 can be computed by (2-192). We see that, in principle, one measured distance s is sufficient for this purpose. In practice, of course, many distances, and also angles, will be measured, and N_0 will be obtained from a suitable adjustment (Sec. 5-10).

Interpretation of N_0. We mention finally that N_0, besides being the distance between S and S' (Fig. 2-18a), has another simple geometrical meaning (Fig. 2-18b).

The radius vector r of the geoid, to a sufficient approximation, is obtained by adding the geoidal height N to the ellipsoidal radius vector given by (2-95):

$$r \doteq a(1 - f\sin^2\phi) + N.$$

Now let the semimajor axis a of the reference ellipsoid be changed by δa, the flattening f remaining the same. Since the geocentric radius vector of the geoid is independent of the size of the reference ellipsoid, it is not affected by this change. By differentiating the equation for r we therefore obtain

$$0 = \delta r = \delta a(1 - f\sin^2\phi) + \delta N \doteq \delta a + \delta N,$$

so that the change in the semimajor axis of the reference ellipsoid is compensated by a change in the geoidal undulations of

$$\delta N = -\delta a.$$

If the change is $\delta a = N_0$, then the semimajor axis of the new reference ellipsoid E' is

$$a' = a + N_0,$$

and the new geoidal undulations are

$$N' = N + \delta N = N - N_0.$$

By (2-183b) this is

$$N' = \frac{R}{4\pi G}\iint_\sigma \Delta g\, S(\psi)\, d\sigma.$$

Hence, *on changing the semimajor axis of the reference ellipsoid by N_0, the new geoidal undulations are given by the original Stokes formula.* That is, the values N' obtained by applying the simple Stokes formula refer to an ellipsoid that has the same flattening as the original reference ellipsoid and a semimajor axis $a + N_0$.

Since N' contains no harmonic of degree zero, we have

$$\iint_\sigma N' \, d\sigma = 0. \tag{2-194a}$$

The volume v of the layer between the ellipsoid E' and the geoid is given by

$$v = \iint_\sigma N' \cdot R^2 \, d\sigma,$$

because $R^2 \, d\sigma$ is the surface element of E' as a spherical approximation, so that (Fig. 2-18b)

$$dv = N' \cdot R^2 \, d\sigma.$$

Hence (2-194a) expresses the fact that the total volume of this layer is zero, or that the new ellipsoid E' with $a' = a + N_0$ *encloses the same volume as the geoid.*

Interpretation of Δg_0. The zero-degree harmonic Δg_0 admits of an analogous interpretation.

Gravity g on the geoid is obtained by adding the gravity anomaly Δg to normal gravity given by (2-96):

$$g \doteq \gamma_a(1 + f^* \sin^2 \phi) + \Delta g.$$

Now we let the normal equatorial gravity γ_a be changed by $\delta\gamma_a$, the coefficient f^* remaining the same. Since g is not affected by this change, we find by differentiating this equation

$$0 = \delta g = \delta\gamma_a(1 + f^* \sin^2 \phi) + \delta\Delta g \doteq \delta\gamma_a + \delta\Delta g,$$

so that

$$\delta\Delta g = -\delta\gamma_a.$$

With a change of $\delta\gamma_a = \Delta g_0$, the values become

$$\gamma_a' = \gamma_a + \Delta g_0, \quad \Delta g' = \Delta g - \Delta g_0.$$

Noting the definition (2-185) of Δg_0, we find

$$\iint_\sigma \Delta g' \, d\sigma = 0, \tag{2-194b}$$

which means that the new gravity anomalies $\Delta g'$ contain no zero-degree harmonic.

Since neither N' nor $\Delta g'$ contains a zero-degree harmonic, they must *refer to an ellipsoid enclosing the same mass as the earth and having the same potential as the geoid.* This ellipsoid has the same flattening as the original reference ellipsoid, and its other constants are

$$a' = a + N_0, \quad \gamma_a' = \gamma_a + \Delta g_0.$$

This interpretation is related to the ideas of Ledersteger (1957).

2-20. Determination of the Physical Constants of the Earth

Mass and potential. The fundamental equations for mass and potential were found in the preceding section to be

$$k\delta M = R(R\,\Delta g_0 + 2GN_0),$$
$$\delta W = R\,\Delta g_0 + GN_0.$$

Let us recapitulate how the mass of the earth, M, and the potential of the geoid, W^0, are determined from these equations. We assume an arbitrary, but fixed, reference ellipsoid with constants M' (mass) and U^0 (potential). We then compute the gravity anomalies Δg, referred to this ellipsoid, and calculate Δg_0 by (2-185). By measuring at least one distance s, and also the astronomical latitude Φ and longitude Λ of its end points, we can determine N_0, using formula (2-192). Then the corrections δM and δW are computed from the above equations. Finally, the earth's mass M and the geoidal potential W^0 are found by adding these corrections to the assumed ellipsoidal values M' and U^0:

$$M = M' + \delta M,$$
$$W^0 = U^0 + \delta W.$$

The mass is given in the form kM; that is, the mass is multiplied by the gravitational constant rather than given as M alone because k is not very accurately known.

Note the intimate connection between geometrical and physical constants. Once we know the physical constants kM and W^0, then we also know the linear scale of the earth or, in other words, its size. Conversely, kM and W^0 can be found with the aid of distance measurements. Another significant fact is that since the gravity anomalies on the whole earth are needed in (2-185), *the constants kM and W^0 cannot be determined unless gravity g is known all over the earth.* This again reflects the general principle of the gravimetric method—namely, that it is always required that g be known at every point of the earth's surface.

Higher harmonics. In Sec. 2-5 we found the following expression for the gravitational potential V outside the earth:

$$V = W - \Phi = \frac{kM}{r}\left[1 - \sum_{n=2}^{\infty}\sum_{m=0}^{n}\left(\frac{a}{r}\right)^n (J_{nm}\cos m\lambda + K_{nm}\sin m\lambda)P_{nm}(\cos\theta)\right].$$

Similarly, the normal gravitational potential may be written as

$$U - \Phi = \frac{kM'}{r}\left[1 - \sum_{n=2}^{\infty}\sum_{m=0}^{n}\left(\frac{a}{r}\right)^n (J'_{nm}\cos m\lambda + K'_{nm}\sin m\lambda)P_{nm}(\cos\theta)\right].$$

If we take an ellipsoid of revolution as our reference surface, then all K'_{nm} are zero, and of the J'_{nm} only the J'_{n0} with n even are different from zero (see Sec. 2-9).

Subtracting the above equations and setting $r = a$ we obtain

$$T = W - U = \frac{k\delta M}{a} - \frac{kM}{a} \sum_{n=2}^{\infty} \sum_{m=0}^{n} (\delta J_{nm} \cos m\lambda + \delta K_{nm} \sin m\lambda) P_{nm}(\cos\theta),$$

where

$$\delta J_{nm} = J_{nm} - J'_{nm}, \quad \delta K_{nm} = K_{nm} - K'_{nm} = K_{nm}.$$

This is possible because, for the terms of second degree and higher, the factor kM'/a can be replaced by kM/a.

Comparing this with the expansion (2-152′) of T, we see that the Laplace surface harmonic $T_n(\theta, \lambda)$, for $n \geqq 2$, is given by

$$T_n(\theta, \lambda) = -\frac{kM}{a} \sum_{m=0}^{n} (\delta J_{nm} \cos m\lambda + \delta K_{nm} \sin m\lambda) P_{nm}(\cos\theta).$$

In agreement with the usual spherical approximation we replace a by R, obtaining

$$T_n(\theta, \lambda) = -\frac{kM}{R} \sum_{m=0}^{n} (\delta J_{nm} \cos m\lambda + \delta K_{nm} \sin m\lambda) P_{nm}(\cos\theta).$$

We insert this equation, together with (2-172), into (2-184a) and obtain

$$\Delta g(\theta, \lambda) = -\frac{kM}{R^2} \sum_{n=2}^{\infty} \sum_{m=0}^{n} (n-1)(\delta J_{nm} \cos m\lambda + \delta K_{nm} \sin m\lambda) P_{nm}(\cos\theta) - \frac{k\delta M}{R^2} + \frac{2\delta W}{R}. \tag{2-195a}$$

We can also write the spherical-harmonic expansion of Δg in the usual form (1-66):

$$\Delta g(\theta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} (c_{nm} \cos m\lambda + d_{nm} \sin m\lambda) P_{nm}(\cos\theta), \tag{2-195b}$$

where the coefficients c_{nm} and d_{nm} are given by (1-70):

$$\begin{aligned} c_{n0} &= \frac{2n+1}{4\pi} \iint_\sigma \Delta g\, P_n(\cos\theta)\, d\sigma; \\ \begin{Bmatrix} c_{nm} \\ d_{nm} \end{Bmatrix} &= \frac{2n+1}{2\pi} \frac{(n-m)!}{(n+m)!} \iint_\sigma \Delta g\, P_{nm}(\cos\theta) \begin{Bmatrix} \cos m\lambda \\ \sin m\lambda \end{Bmatrix} d\sigma \qquad (m \neq 0). \end{aligned} \tag{2-196}$$

Equations (2-195a) and (2-195b) are obviously identical to (2-184a) and (2-184b), the Laplace surface harmonics $T_n(\theta, \lambda)$ and $\Delta g_n(\theta, \lambda)$ being written explicitly after the fashion of equation (1-66).

By comparing the coefficients in (2-195a) and (2-195b) we see that

$$\delta J_{nm} = -\frac{R^2}{(n-1)kM} c_{nm}, \quad \delta K_{nm} = -\frac{R^2}{(n-1)kM} d_{nm}.$$

Since $J_{nm} = J'_{nm} + \delta J_{nm}$, then $K_{nm} = \delta K_{nm}$, and since $J'_{nm} = 0$ for $m \neq 0$, we finally obtain

$$J_n = J_n' - \frac{R^2}{(n-1)kM} c_{n0},$$

$$\left.\begin{aligned} J_{nm} &= -\frac{R^2}{(n-1)kM} c_{nm} \\ K_{nm} &= -\frac{R^2}{(n-1)kM} d_{nm} \end{aligned}\right\} \quad (m \neq 0). \tag{2-197}$$

Here we have abbreviated the zonal coefficients J_{n0} by J_n.

Hence the determination of the spherical-harmonic coefficients of the earth's potential can be described as follows. We expand the gravity anomalies Δg, which must be given all over the world, as a series of spherical harmonics, according to (2-195b) and (2-196). We next compute the coefficients J_n' for the reference ellipsoid, for instance by (2-92). Then, formulas (2-197) give the desired result.

Of particular importance is the coefficient

$$J_2 = \frac{C - \bar{A}}{Ma^2}, \tag{2-198}$$

which gives the difference between the principal moments of inertia of the earth: C is the polar moment and

$$\bar{A} = \frac{1}{2}(A + B) \tag{2-199}$$

is the mean equatorial moment of inertia; see (2-49).

2-21. The Mean Earth Ellipsoid

Since a level ellipsoid of revolution and its gravity field are completely determined by four constants, there is one and only one ellipsoid that has the same potential W_0 as the geoid and the same mass M, the same difference of moments of inertia $C - \bar{A}$, and the same angular velocity ω as the earth; $\bar{A}$ is defined by (2-199). By (2-198), this ellipsoid has also the same coefficient J_2. In many respects, it can be considered the best representation of the earth by an ellipsoid; it is therefore called the *mean earth ellipsoid.*

The mean earth ellipsoid, defined by

$$W_0, \quad kM, \quad C - \bar{A}, \quad \omega$$

or, equivalently, by

$$W_0, \quad kM, \quad J_2, \quad \omega,$$

has many desirable properties. As we have seen in Sec. 2-19, it encloses the same volume as the geoid; in Sec. 5-11 we shall see that the sum of the squares of the deviations N of the geoid from the mean earth ellipsoid is a minimum. If the mean earth ellipsoid is in an absolute position, its center coinciding with the earth's center of gravity, then it gives rise to a normal potential U which, for larger distances, is almost exactly equal to the actual potential W of the earth.

The latter property makes the mean earth ellipsoid particularly suited for dynamical astronomy—for instance, with regard to the theory of the motion of the moon or of artificial satellites. The reason is that for larger distances only the harmonics up to the second degree are effective, and these harmonics are equal for W and U because of the equality of kM (zero degree), the absolute position of the ellipsoid (first degree), and the equality of J_2 (second degree, zonal[1]).

This definition of the mean earth ellipsoid enables us to give, in geodesy, precise definitions of the semimajor axis a of the earth, the equatorial gravity γ_a, etc. As a matter of fact, the actual equator of the earth is an irregular curve rather than a circle or radius a, and if we would measure gravity along the equator, we would get many different values, rather than a definite constant γ_a. Something similar is true, for instance, of the flattening $f = (a - b)/a$. These constants, a, f, γ_a, etc., must therefore be considered as derived parameters which refer to an idealized ellipsoid rather than directly to the earth.

To obtain these quantities from given values of W_0, kM, J_2, ω we solve the two equations

$$W_0 = \frac{kM}{E}\tan^{-1} e' + \frac{1}{3}\omega^2 a^2,$$

$$J_2 = \frac{E^2}{3a^2}\left(1 - \frac{2}{15}\frac{me'}{q_0}\right)$$

with respect to a and f and compute γ_a by (2-73). The first of these equations is (2-61); the second is obtained from (2-90) by noting that $J_2 = (C - A)/Ma^2$. In practice, it is more convenient to use the corresponding series expansions (2-104), (2-118), and (2-105a).

Even more convenient is the use of differential formulas. Since $b = a(1 - f)$. we may approximate (2-111) and (2-112) by

$$kM = a^2\gamma_a\left(1 - f + \frac{3}{2}m\right),$$

$$W_0 = a\gamma_a\left(1 - \frac{2}{3}f + \frac{11}{6}m\right).$$

Solving for a and γ_a yields

$$a = \frac{kM}{W_0}\left(1 + \frac{1}{3}f + \frac{1}{3}m\right),$$

$$\gamma_a = \frac{W_0^2}{kM}\left(1 + \frac{1}{3}f - \frac{13}{6}m\right).$$

By differentiating these formulas and neglecting f and m in the coefficients, we find as spherical approximations:

[1] There will also be nonzonal terms of the second degree, because $A \neq B$, but these are much smaller than J_2.

$$\delta a = \frac{1}{a\gamma_a} k\delta M - \frac{1}{\gamma_a} \delta W + \frac{1}{3} a\delta f,$$
$$\delta\gamma_a = -\frac{1}{a^2} k\delta M + \frac{2}{a} \delta W + \frac{1}{3} \gamma_a \delta f. \tag{2-200}$$

By means of (2-182) and (2-186), these can be considerably simplified:

$$\delta a = N_0 + \frac{1}{3} a\delta f,$$
$$\delta\gamma_a = \Delta g_0 + \frac{1}{3} \gamma_a \delta f. \tag{2-200'}$$

From (2-118) we get, approximately,

$$f = \frac{3}{2} J_2 + \frac{1}{2} m.$$

Differentiation finally gives

$$\delta f = \frac{3}{2} \delta J_2. \tag{2-201}$$

This equation expresses the change of flattening in terms of the variation of J_2; the changes in a and γ_a can then be obtained from (2-200) or (2-200′).

It should be remembered, however, that the mean earth ellipsoid, defined in this manner, is not at all the best reference surface for practical geodetic purposes. It is essentially defined empirically by means of empirical determinations of kM, W_0, etc. Its parameters will change with every improvement in the quality or the number of the relevant measurements (gravity, distances, etc.). Since an enormous amount of numerical data is based on an assumed reference ellipsoid, it would be highly impractical to change it very often, for this would involve repeated transformations of all the data. It is much better to use a fixed reference ellipsoid with rigidly assumed parameters, which can be more or less arbitrary if only they give a reasonably good approximation. In this respect even the international ellipsoid would be sufficient, although a change might be desirable for other reasons.

A certain amount of conflict exists between the interests of geodesists and astronomers regarding the earth ellipsoid. The geodesist needs a permanent reference surface, whereas the astronomer wants the best approximation of the earth by an ellipsoid. A good compromise is to use a fixed geodetic reference ellipsoid, but from time to time to compute the "best" corrections to the assumed parameters for astronomical purposes.

2-22. Deflections of the Vertical. Formula of Vening Meinesz

Stokes' formula permits the calculation of the geoidal undulations from gravity anomalies. A similar formula for the computation of the deflections of the vertical from gravity anomalies has been given by Vening Meinesz (1928).

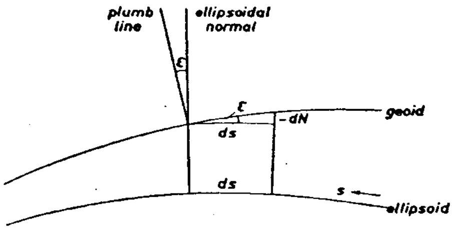

FIGURE 2-20

The relation between the geoidal undulation and the deflection of the vertical.

Fig. 2-20 shows the intersection of geoid and reference ellipsoid with a vertical plane of arbitrary azimuth. If ϵ is the component of the deflection of the vertical in this plane, then

$$dN = -\epsilon\, ds, \tag{2-202}$$

or

$$\epsilon = -\frac{dN}{ds}; \tag{2-203}$$

the minus sign is a convention, the meaning of which will be explained later.

In a north-south direction we have

$$\epsilon = \xi \quad \text{and} \quad ds = ds_\phi = R\, d\phi;$$

in an east-west direction,

$$\epsilon = \eta \quad \text{and} \quad ds = ds_\lambda = R \cos\phi\, d\lambda.$$

In the formulas for ds_ϕ and ds_λ we have again used the spherical approximation; according to (1-38), the linear element on the sphere $r = R$ is given by

$$ds^2 = R^2\, d\phi^2 + R^2 \cos^2\phi\, d\lambda^2.$$

By specializing (2-203) we find

$$\begin{aligned} \xi &= -\frac{dN}{ds_\phi} = -\frac{1}{R}\frac{\partial N}{\partial \phi}, \\ \eta &= -\frac{dN}{ds_\lambda} = -\frac{1}{R\cos\phi}\frac{\partial N}{\partial \lambda}, \end{aligned} \tag{2-204}$$

which gives the connection between the geoidal undulation N and the components ξ and η of the deflection of the vertical.

As N is given by Stokes' integral, our problem is to differentiate this formula with respect to ϕ and λ. For this purpose we use the form (2-167),

$$N(\phi, \lambda) = \frac{R}{4\pi G}\int_{\lambda'=0}^{2\pi}\int_{\phi'=-\pi/2}^{\pi/2} \Delta g(\phi', \lambda') S(\psi) \cos\phi'\, d\phi'\, d\lambda',$$

where ψ is defined as a function of ϕ, λ and ϕ', λ' by (2-168).

The integral on the right-hand side of this formula depends on ϕ and λ only through ψ in $S(\psi)$. Therefore, by differentiating under the integral sign we find

$$\frac{\partial N}{\partial \phi} = \frac{R}{4\pi G}\int_{\lambda'=0}^{2\pi}\int_{\phi'=-\pi/2}^{\pi/2} \Delta g(\phi', \lambda') \frac{\partial S(\psi)}{\partial \phi} \cos\phi'\, d\phi'\, d\lambda' \tag{2-205}$$

and a similar formula for $\partial N/\partial\lambda$. Here we have

$$\frac{\partial S(\psi)}{\partial\phi} = \frac{dS(\psi)}{d\psi}\frac{\partial\psi}{\partial\phi}, \quad \frac{\partial S(\psi)}{\partial\lambda} = \frac{dS(\psi)}{d\psi}\frac{\partial\psi}{\partial\lambda}. \tag{2-206}$$

Writing (2-168) in the form

$$\cos\psi = \sin\phi\sin\phi' + \cos\phi\cos\phi'\cos(\lambda' - \lambda) \tag{2-207}$$

and differentiating with respect to ϕ and λ we obtain

$$-\sin\psi\frac{\partial\psi}{\partial\phi} = \cos\phi\sin\phi' - \sin\phi\cos\phi'\cos(\lambda' - \lambda),$$

$$-\sin\psi\frac{\partial\psi}{\partial\lambda} = \cos\phi\cos\phi'\sin(\lambda' - \lambda).$$

We now introduce the azimuth α, as shown in Fig. 2-16. From the spherical triangle of Fig. 2-21 we get, using well-known formulas of spherical trigonometry,

$$\begin{aligned} \sin\psi\cos\alpha &= \cos\phi\sin\phi' - \sin\phi\cos\phi'\cos(\lambda' - \lambda), \\ \sin\psi\sin\alpha &= \cos\phi'\sin(\lambda' - \lambda). \end{aligned} \tag{2-208}$$

Inserting these into the preceding equations we find the simple expressions

$$\frac{\partial\psi}{\partial\phi} = -\cos\alpha, \quad \frac{\partial\psi}{\partial\lambda} = -\cos\phi\sin\alpha, \tag{2-209}$$

so that

$$\frac{\partial S(\psi)}{\partial\phi} = -\frac{dS(\psi)}{d\psi}\cos\alpha, \quad \frac{\partial S(\psi)}{\partial\lambda} = -\frac{dS(\psi)}{d\psi}\cos\phi\sin\alpha.$$

These are substituted into (2-205) and the corresponding formula for $\partial N/\partial\lambda$, and from equations (2-204) we finally obtain

$$\begin{aligned} \xi(\phi, \lambda) &= \frac{1}{4\pi G}\int_{\lambda'=0}^{2\pi}\int_{\phi'=-\pi/2}^{\pi/2}\Delta g(\phi', \lambda')\frac{dS(\psi)}{d\psi}\cos\alpha\cos\phi'\,d\phi'\,d\lambda', \\ \eta(\phi, \lambda) &= \frac{1}{4\pi G}\int_{\lambda'=0}^{2\pi}\int_{\phi'=-\pi/2}^{\pi/2}\Delta g(\phi', \lambda')\frac{dS(\psi)}{d\psi}\sin\alpha\cos\phi'\,d\phi'\,d\lambda' \end{aligned} \tag{2-210}$$

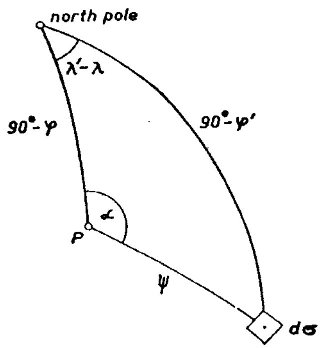

FIGURE 2-21

The relation between geographical and polar coordinates on the sphere.

or, written in the usual abbreviated form,

$$\xi = \frac{1}{4\pi G} \iint_\sigma \Delta g \frac{dS}{d\psi} \cos \alpha \, d\sigma,$$
$$\eta = \frac{1}{4\pi G} \iint_\sigma \Delta g \frac{dS}{d\psi} \sin \alpha \, d\sigma. \tag{2-210'}$$

These are the *formulas of Vening Meinesz*. Differentiating Stokes' function $S(\psi)$, equation (2-164), with respect to ψ we obtain *Vening Meinesz' function*

$$\frac{dS}{d\psi} = -\frac{\cos(\psi/2)}{2\sin^2(\psi/2)} + 8\sin\psi - 6\cos(\psi/2) - 3\frac{1-\sin(\psi/2)}{\sin\psi} + 3\sin\psi \ln[\sin(\psi/2) + \sin^2(\psi/2)]. \tag{2-211}$$

This can be readily verified by using the elementary trigonometric identities. The azimuth α is given by the formula

$$\tan\alpha = \frac{\cos\phi' \sin(\lambda' - \lambda)}{\cos\phi \sin\phi' - \sin\phi\cos\phi'\cos(\lambda' - \lambda)}, \tag{2-212}$$

which is an immediate consequence of (2-208).

The form (2-210) is an expression of (2-210') in terms of geographical coordinates ϕ and λ. As with Stokes' formula (Sec. 2-17) we may also use an expression in terms of spherical polar coordinates ψ and α:

$$\begin{Bmatrix} \xi \\ \eta \end{Bmatrix} = \frac{1}{4\pi G} \int_{\alpha=0}^{2\pi} \int_{\psi=0}^{\pi} \Delta g(\psi, \alpha) \begin{Bmatrix} \cos\alpha \\ \sin\alpha \end{Bmatrix} \frac{dS}{d\psi} \sin\psi \, d\psi \, d\alpha. \tag{2-210''}$$

The reader can easily verify that these equations give the deflection components ξ and η with the correct sign corresponding to the definition (2-140); see also Fig. 2-13. This is the reason why we introduced the minus sign in (2-203).

We note that the formula of Vening Meinesz is valid as it stands for an arbitrary reference ellipsoid, whereas Stokes' formula had to be modified by adding a constant N_0: If we differentiate the modified Stokes formula (2-183b) with respect to ϕ and λ, to get Vening Meinesz' formula, then this constant N_0 drops out and we get equations (2-210').

The practical application of Stokes' and Vening Meinesz' formulas raises many important problems, for which the reader is referred to Sec. 2-24 and to Chapter 3. The function $dS/d\psi$ and related functions are tabulated in Sollins (1947).

2-23. The Vertical Gradient of Gravity. Free-air Reduction to Sea Level

For a theoretically correct reduction of gravity to the geoid we need the vertical gradient of gravity, $\partial g/\partial h$. If g is the observed value at the surface of the earth, then the value g_0 at the geoid may be obtained as a Taylor expansion:

$$g_0 = g - \frac{\partial g}{\partial h} H \cdots,$$

where H is the elevation of the gravity station above the geoid. Neglecting all terms but the linear one, we have

$$g_0 = g + F, \tag{2-213}$$

where

$$F = -\frac{\partial g}{\partial h} H \tag{2-214}$$

is the *free-air reduction* to the geoid. Here, as throughout this chapter, we have assumed that there are no masses above the geoid, or that such masses have been removed beforehand, so that this reduction is indeed carried out "in free air."

Bruns' formula (2-20), with $\rho = 0$,

$$\frac{\partial g}{\partial h} = -2gJ - 2\omega^2,$$

cannot be directly applied for this purpose because the mean curvature J of the level surfaces is unknown. We therefore proceed in the usual way by splitting $\partial g/\partial h$ into a normal and an anomalous part:

$$\frac{\partial g}{\partial h} = \frac{\partial \gamma}{\partial h} + \frac{\partial \Delta g}{\partial h}. \tag{2-215}$$

The normal gradient $\partial\gamma/\partial h$ is given by (2-79) and (2-80), or by (2-121). The anomalous part, $\partial\Delta g/\partial h$, will be considered now.

Expression in terms of Δg. Equation (2-155) may be written as

$$\Delta g(r, \theta, \lambda) = \sum_{n=0}^{\infty} \left(\frac{R}{r}\right)^{n+2} \Delta g_n(\theta, \lambda).$$

By differentiating with respect to r and setting $r = R$, we obtain at sea level:

$$\frac{\partial \Delta g}{\partial r} = -\frac{1}{R}\sum_{n=0}^{\infty}(n+2)\Delta g_n = -\frac{1}{R}\sum_{n=0}^{\infty} n\Delta g_n - \frac{2}{R}\Delta g. \tag{2-216}$$

Now we can apply (1-102), setting $V = \Delta g$ and $Y_n = \Delta g_n$. The result is

$$\frac{\partial \Delta g}{\partial r} = \frac{R^2}{2\pi}\iint_\sigma \frac{\Delta g - \Delta g_P}{l_0^3}\, d\sigma - \frac{2}{R}\Delta g_P. \tag{2-217}$$

In this equation, Δg_P is referred to the fixed point P at which $\partial\Delta g/\partial r$ is to be computed; l_0 is the spatial distance between the fixed point P and the variable surface element $R^2\, d\sigma$, expressed in terms of the angular distance ψ by

$$l_0 = 2R \sin\frac{\psi}{2}.$$

Compare Fig. 1-13 of Sec. 1-18; the element $R^2\, d\sigma$ is at the point P'.

The important integral formula (2-217) expresses the vertical gradient of the gravity anomaly in terms of the gravity anomaly itself. Since the integrand decreases very rapidly with increasing distance l_0, it is sufficient in this formula to extend the integration only over the immediate neighborhood of the point P, as opposed to Stokes' and Vening Meinesz' formulas, where the integration must include the whole earth, if a sufficient accuracy is to be obtained.

Expression in terms of N. By differentiating equation (2-154),

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{r}T,$$

with respect to r we get

$$\frac{\partial \Delta g}{\partial r} = -\frac{\partial^2 T}{\partial r^2} - \frac{2}{r}\frac{\partial T}{\partial r} + \frac{2}{r^2}T.$$

To this formula we add Laplace's equation $\Delta T = 0$, which in spherical coordinates has the form[1]

$$\frac{\partial^2 T}{\partial r^2} + \frac{2}{r}\frac{\partial T}{\partial r} - \frac{\tan\phi}{r^2}\frac{\partial T}{\partial \phi} + \frac{1}{r^2}\frac{\partial^2 T}{\partial \phi^2} + \frac{1}{r^2\cos^2\phi}\frac{\partial^2 T}{\partial \lambda^2} = 0.$$

The result, on setting $r = R$, is

$$\frac{\partial \Delta g}{\partial r} = \frac{2}{R^2}T - \frac{\tan\phi}{R^2}\frac{\partial T}{\partial \phi} + \frac{1}{R^2}\frac{\partial^2 T}{\partial \phi^2} + \frac{1}{R^2\cos^2\phi}\frac{\partial^2 T}{\partial \lambda^2}. \tag{2-218}$$

Since $T = GN$, we may also write

$$\frac{\partial \Delta g}{\partial r} = \frac{2G}{R^2}N - \frac{G}{R^2}\tan\phi\frac{\partial N}{\partial \phi} + \frac{G}{R^2}\frac{\partial^2 N}{\partial \phi^2} + \frac{G}{R^2\cos^2\phi}\frac{\partial^2 N}{\partial \lambda^2}. \tag{2-219}$$

This equation expresses the vertical gradient of the gravity anomaly in terms of the geoidal undulation N and its first and second horizontal derivatives. It can be evaluated by numerical differentiation, using a map of the function N. However, it is less suited for practical application than (2-217) because it requires an extremely accurate and detailed local geoidal map, which is hardly ever available; inaccuracies of N are greatly amplified by forming the second derivatives.

Expression in terms of ξ and η. From equations (2-204) we find

$$\frac{\partial N}{\partial \phi} = -R\xi, \quad \frac{\partial N}{\partial \lambda} = -R\eta\cos\phi,$$

so that

$$\frac{\partial^2 N}{\partial \phi^2} = -R\frac{\partial \xi}{\partial \phi}, \quad \frac{\partial^2 N}{\partial \lambda^2} = -R\frac{\partial \eta}{\partial \lambda}\cos\phi.$$

[1] See equation (1-41); substitute $\theta = 90° - \phi$.

By inserting this into (2-219) we obtain

$$\frac{\partial \Delta g}{\partial r} = \frac{2G}{R^2} N + \frac{G}{R} \xi \tan \phi - \frac{G}{R} \frac{\partial \xi}{\partial \phi} - \frac{G}{R \cos \phi} \frac{\partial \eta}{\partial \lambda}. \tag{2-220}$$

On introducing local rectangular coordinates x, y in the tangent plane we have

$$R\, d\phi = ds_\phi = dx,$$
$$R \cos \phi\, d\lambda = ds_\lambda = dy,$$

so that (2-220) becomes

$$\frac{\partial \Delta g}{\partial r} = \frac{2G}{R^2} N + \frac{G}{R} \xi \tan \phi - G\left(\frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y}\right).$$

The first two terms on the right-hand side can be shown to be very small in comparison to the third term; hence to a sufficient accuracy

$$\frac{\partial \Delta g}{\partial r} = -G\left(\frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y}\right). \tag{2-221}$$

These formulas express the vertical gradient of the gravity anomaly in terms of the horizontal derivatives of the deflection of the vertical. They can again be evaluated by means of numerical differentiation if a map of ξ and η is available. They are better suited for practical application than (2-219) because only first derivatives are required. For a practical computation see Mueller (1961).

These formulas will be used in Sec. 8-8.

2-24. Practical Evaluation of the Integral Formulas

Integral formulas such as Stokes' and Vening Meinesz' integrals must be evaluated approximately by summations. The surface elements $d\sigma$ are replaced by small but finite compartments q, which are obtained by suitably subdividing the surface of the earth. Two different methods of subdivision are used:

1. *Templates* (Fig. 2-22). The subdivision is effected by concentric circles and their radii. The template, which is made of transparent material, is placed on a gravity map of the same scale, so that the center of the template coincides with the computation point P on the map. The natural coordinates for this purpose are *polar coordinates* ψ, α with origin at P.
2. *Grid lines* (Fig. 2-23). The subdivision is effected by the grid lines of some fixed coordinate system, in particular of *geographical coordinates* ϕ, λ. They form rectangular *blocks*—for example, of $10' \times 10'$ or $1° \times 1°$. These blocks are also called *squares*, although they are usually not squares as defined in plane geometry.

As a definite example illustrating the principles of numerical integration consider Stokes' formula

$$N = \frac{R}{4\pi G} \iint_\sigma \Delta g S(\psi)\, d\sigma$$

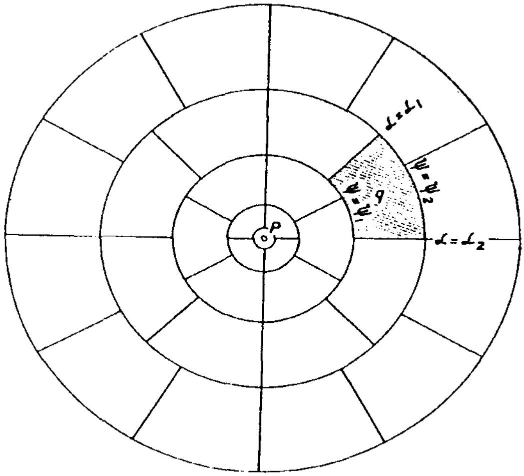

FIGURE 2-22
A template.

with its explicit forms (2-165) for the template method and (2-167) for the method that uses fixed blocks.

For each compartment q_k the gravity anomalies are replaced by their average value $\overline{\Delta g_k}$ in this compartment. Hence the above equation becomes

$$N = \frac{R}{4\pi G} \sum_k \iint_{q_k} \overline{\Delta g_k} S(\psi)\, d\sigma = \frac{R}{4\pi G} \sum_k \overline{\Delta g_k} \cdot \iint_{q_k} S(\psi)\, d\sigma$$

or

$$N = \sum_k c_k \overline{\Delta g_k}, \tag{2-222}$$

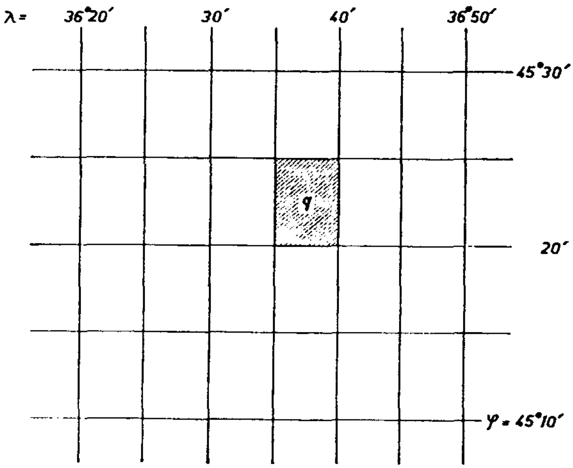

FIGURE 2-23
Blocks formed by a grid of geographical coordinates.

where the coefficients

$$c_k = \frac{R}{4\pi G} \iint_{q_k} S(\psi)\, d\sigma \tag{2-223}$$

are obtained by integration over the compartment q_k; they do not depend on Δg.

If the integrand—in our case, Stokes' function $S(\psi)$—is reasonably constant over the compartment q_k, it may be replaced by its value $S(\psi_k)$ at the center of q_k. Then we have

$$c_k = \frac{R}{4\pi G} S(\psi_k) \iint_{q_k} d\sigma = \frac{S(\psi_k)}{4\pi GR} \iint_{q_k} R^2\, d\sigma.$$

The final integral is simply the area A_k of the compartment. Hence we obtain

$$c_k = \frac{A_k S(\psi_k)}{4\pi GR}. \tag{2-224}$$

This form is simpler, but close to the computation point it may be necessary to use the integrated coefficients (2-223).

If the compartments are formed by lines ϕ = const., λ = const., then the computation of these integrated coefficients is difficult. For the template method, however, where the compartments are formed by lines ψ = const., α = const., it is quite straightforward. We have

$$\begin{aligned} c_k &= \frac{R}{4\pi G} \int_{\alpha=\alpha_1}^{\alpha_2} \int_{\psi=\psi_1}^{\psi_2} S(\psi) \sin\psi\, d\psi\, d\alpha \\ &= \frac{R(\alpha_2 - \alpha_1)}{4\pi G} \int_{\psi_1}^{\psi_2} S(\psi) \sin\psi\, d\psi. \end{aligned}$$

The function

$$J(\psi) \equiv \frac{1}{2} \int_0^{\psi} S(\psi) \sin\psi\, d\psi = \int_0^{\psi} F(\psi)\, d\psi \tag{2-225}$$

(see Sec. 2-17) has been tabulated by Lambert and Darling (1936). Hence we obtain

$$c_k = \frac{R(\alpha_2 - \alpha_1)}{2\pi G} [J(\psi_2) - J(\psi_1)]. \tag{2-226}$$

As another example, consider the formula (2-217) of the preceding section. Here

$$c_k = \frac{R^2}{2\pi} \iint_{q_k} \frac{d\sigma}{l_0^3},$$

where

$$l_0 = 2R \sin\frac{\psi}{2}.$$

We find

$$c_k = \frac{1}{16\pi R}\int_{\alpha=\alpha_1}^{\alpha_2}\int_{\psi=\psi_1}^{\psi_2}\frac{\sin\psi\,d\psi\,d\alpha}{\sin^3(\psi/2)}$$
$$= \frac{\alpha_2-\alpha_1}{16\pi R}\int_{\psi=\psi_1}^{\psi_2}\frac{2\sin(\psi/2)\cos(\psi/2)}{\sin^3(\psi/2)}\,d\psi = \frac{\alpha_2-\alpha_1}{8\pi R}\int_{\psi_1}^{\psi_2}\frac{\cos(\psi/2)}{\sin^2(\psi/2)}\,d\psi.$$

This integral is readily solved by substituting $u = \sin(\psi/2)$; we obtain

$$c_k = \frac{\alpha_2-\alpha_1}{2\pi}\left(\frac{1}{l_{0,1}}-\frac{1}{l_{0,2}}\right). \tag{2-227}$$

The advantage of the template method is its great flexibility. The influence of the compartments near the computation point P is greater than that of the distant ones, and the integrand changes faster in the neighborhood of P. Therefore, a finer subdivision is necessary around P. This can easily be provided by templates. Moreover, the computation of integrated coefficients is easier with the template method.

The advantage of the fixed system of blocks formed by a grid of geographical coordinates lies in the fact that their mean gravity anomalies are needed for many different purposes. These mean anomalies of standard-sized blocks, once they have been determined, can be easily stored and processed by an electronic computer. Also, the same subdivision is used for all computation points, whereas the compartments defined by a template change when the template is moved to the next computation point. The flexibility of the method of standard blocks is of course limited; however, one may use smaller blocks ($5' \times 5'$, for example) in the neighborhood of P and larger ones ($1° \times 1°$, for example) farther away. For electronic computation this method is usually preferred.

It is also possible to combine the two methods, computing the effect of the inner zone by means of a template and using standard blocks outside. This may be advantageous if the integrand changes too rapidly over a $5' \times 5'$ block, which is usually the smallest standard size available.

Effect of the neighborhood. In the innermost zone even the template method may pose difficulties if the integrand becomes infinite as $\psi \longrightarrow 0$. This happens with Stokes' formula, since

$$S(\psi) \doteq \frac{2}{\psi} \tag{2-228}$$

for small ψ. This can be seen from the definition (2-164), because the first term is predominant and for small ψ is given by

$$\frac{1}{\sin(\psi/2)} \doteq \frac{1}{(\psi/2)} = \frac{2}{\psi}.$$

Vening Meinesz' function becomes infinite as well, since to the same approximation,

$$\frac{dS(\psi)}{d\psi} \doteq -\frac{2}{\psi^2}. \tag{2-229}$$

In the gradient formula (2-217) the integrand

$$\frac{1}{l_0^3} \doteq \frac{1}{R^3\psi^3} \tag{2-230}$$

behaves in a similar way.

It is therefore convenient to split off the effect of this innermost zone, which will be assumed to be a circle of radius ψ_0 around the computation point. For instance, Stokes' integral becomes in this way

$$N = N_i + N_e,$$

where

$$N_i = \frac{R}{4\pi G}\int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\psi_0} \Delta g S(\psi)\, d\sigma \quad \text{and} \quad N_e = \frac{R}{4\pi G}\int_{\alpha=0}^{2\pi}\int_{\psi=\psi_0}^{\pi} \Delta g S(\psi)\, d\sigma.$$

The radius ψ_0 of the inner zone corresponds to a linear distance of a few kilometers.

Within this distance we may treat the sphere as a plane, using polar coordinates s, α where

$$s \doteq R\psi \doteq R\sin\psi \doteq 2R\sin\frac{\psi}{2},$$

so that the element of area becomes

$$R^2\, d\sigma = s\, ds\, d\alpha.$$

It is consistent with this approximation to use (2-228) through (2-230), putting

$$S(\psi) \doteq \frac{2R}{s}, \quad \frac{dS}{d\psi} \doteq -\frac{2R^2}{s^2}, \quad \frac{1}{l_0^3} \doteq \frac{1}{s^3}.$$

In both Stokes' and Vening Meinesz' functions the relative error of these approximations is about 1% for $s = 10$ km, and about 3% for $s = 30$ km. In $1/l_0^3$ it is even less. Hence the effect of this inner zone on our integral formulas becomes

$$N_i = \frac{1}{2\pi G}\int_{\alpha=0}^{2\pi}\int_{s=0}^{s_0} \frac{\Delta g}{s}\, s\, ds\, d\alpha, \tag{2-231}$$

$$\begin{Bmatrix}\xi\\ \eta\end{Bmatrix}_i = -\frac{1}{2\pi G}\int_{\alpha=0}^{2\pi}\int_{s=0}^{s_0} \frac{\Delta g}{s^2}\begin{Bmatrix}\cos\alpha\\ \sin\alpha\end{Bmatrix} s\, ds\, d\alpha, \tag{2-232}$$

$$\left(\frac{\partial \Delta g}{\partial h}\right)_i = \frac{1}{2\pi}\int_{\alpha=0}^{2\pi}\int_{s=0}^{s_0} \frac{\Delta g - \Delta g_P}{s^3}\, s\, ds\, d\alpha. \tag{2-233}$$

To evaluate these integrals we expand Δg into a Taylor series at the computation point P:

$$\Delta g = \Delta g_P + x g_x + y g_y + \frac{1}{2!}(x^2 g_{xx} + 2xy g_{xy} + y^2 g_{yy}) + \cdots.$$

The rectangular coordinates x, y are defined by

$$x = s\cos\alpha, \quad y = s\sin\alpha,$$

so that the x-axis points north. We further have

$$g_x = \left(\frac{\partial \Delta g}{\partial x}\right)_P, \quad g_{xx} = \left(\frac{\partial^2 \Delta g}{\partial x^2}\right)_P, \quad \text{etc.}$$

This Taylor series may also be written

$$\Delta g = \Delta g_P + s(g_x \cos\alpha + g_y \sin\alpha) + \frac{s^2}{2}(g_{xx}\cos^2\alpha + 2g_{xy}\cos\alpha\sin\alpha + g_{yy}\sin^2\alpha) + \cdots.$$

On inserting this into the above integrals, we can easily evaluate them. Performing the integration with respect to α first and noting that

$$\int_0^{2\pi} d\alpha = 2\pi,$$

$$\int_0^{2\pi} \sin\alpha\, d\alpha = \int_0^{2\pi} \cos\alpha\, d\alpha = \int_0^{2\pi} \sin\alpha\cos\alpha\, d\alpha = 0,$$

$$\int_0^{2\pi} \sin^2\alpha\, d\alpha = \int_0^{2\pi} \cos^2\alpha\, d\alpha = \pi,$$

we find

$$N_i = \frac{1}{G}\int_0^{s_0}\left[\Delta g_P + \frac{s^2}{4}(g_{xx} + g_{yy}) + \cdots\right] ds,$$

$$\begin{Bmatrix}\xi\\ \eta\end{Bmatrix}_i = -\frac{1}{2G}\int_0^{s_0}\begin{Bmatrix}g_x + \cdots\\ g_y + \cdots\end{Bmatrix} ds,$$

$$\left(\frac{\partial \Delta g}{\partial h}\right)_i = \frac{1}{4}\int_0^{s_0}(g_{xx} + g_{yy} + \cdots)\, ds.$$

We now perform the integration over s, retaining only the lowest nonvanishing terms. The result is

$$N_i = \frac{s_0}{G}\Delta g_P; \tag{2-234}$$

$$\xi_i = -\frac{s_0}{2G}g_x, \quad \eta_i = -\frac{s_0}{2G}g_y; \tag{2-235}$$

$$\left(\frac{\partial \Delta g}{\partial h}\right)_i = \frac{s_0}{4}(g_{xx} + g_{yy}). \tag{2-236}$$

We see that the effect of the innermost circular zone on Stokes' formula depends, to a first approximation, on the value of Δg at P; the effect on Vening Meinesz' formula depends on the first horizontal derivatives of Δg; and the effect on the vertical gradient depends on the second horizontal derivatives.

Note that the contribution of the innermost zone to the total deflection of the vertical has the same direction as the line of steepest inclination of the "gravity anomaly surface" because the plane vector

$$\boldsymbol{\theta}_i = (\xi_i, \eta_i)$$

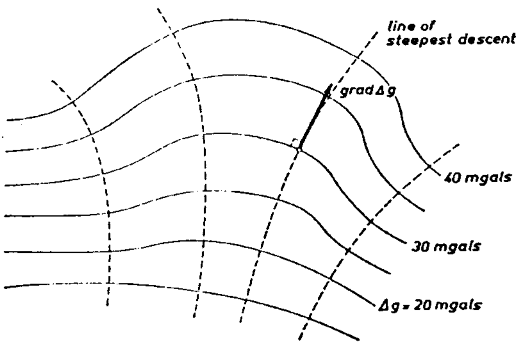

FIGURE 2-24
Lines of constant Δg and lines of steepest descent.

is proportional to the horizontal gradient of Δg,

$$\text{grad}\,\Delta g = (g_x, g_y).$$

The direction of grad Δg defines the line of steepest descent (see Fig. 2-24).

The values of g_x and g_y can be obtained from a gravity map. They are the inclinations of north-south and east-west profiles through P. Values for g_{xx} and g_{yy} may be found by fitting a polynomial in x and y of second degree to the gravity anomaly function in the neighborhood of P.

The influence of distant zones on Stokes' and Vening Meinesz' formulas will be discussed in Sec. 7-4.

For further computational details concerning Stokes' and Vening Meinesz' formulas see Uotila (1960). Applied geophysicists have developed interesting numerical techniques for integration and differentiation which are useful for evaluating formulas like (2-217) and (2-236); see Jung (1961).

References

Baeschlin, C. F. (1948). Lehrbuch der Geodäsie. Zürich, Orell-Füssli.

Bjerhammar, A. (1962). On an explicit solution of the gravimetric boundary value problem for an ellipsoidal surface of reference. Stockholm, Royal Institute of Technology, Geodesy Division.

Bomford, G. (1962). Geodesy. 2nd ed., Oxford University Press.

Bruns, H. (1878). Die Figur der Erde. Berlin, Publ. Preuss. Geod. Inst.

Cassinis, G. (1930). Sur l'adoption d'une formule internationale pour la pesanteur normale. Bull. Géod., No. 26, pp. 40–49.

Fricke, W., D. Brouwer, J. Kovalevsky, A. A. Mikhailov, and G. A. Wilkins (1965). Report to the executive committee of the working group on the system of astronomical constants. Bull. Géod., No. 75, pp. 59–67.

Hayford, J. F. (1909). The figure of the earth and isostasy, from measurements in the United States. U. S. Coast and Geodetic Survey.

Hayford, J. F. (1910). Supplementary investigation in 1909 of the figure of the earth and isostasy. U. S. Coast and Geodetic Survey.

Heiskanen, W. (1924). Untersuchungen über Schwerkraft und Isostasie. Helsinki, Publ. Finn. Geod. Inst., No. 4.

Heiskanen, W. (1928). Ist die Erde ein dreiachsiges Ellipsoid? Gerlands Beitr. Geophys., v. 19, pp. 356–377.

Heiskanen, W. A., and F. A. Vening Meinesz (1958). The earth and its gravity field. New York, McGraw-Hill.

Helmert, F. R. (1884). Die mathematischen und physikalischen Theorien der höheren Geodäsie, vol. 2. Leipzig, B. G. Teubner (reprinted 1962).

Hirvonen, R. A. (1960). New theory of the gravimetric geodesy. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 32.

Hopfner, F. (1933). Physikalische Geodäsie. Leipzig, Akademische Verlagsgesellschaft.

Hotine, M. (1957). Metrical properties of the earth's gravitational field. General Assembly, Int. Union Geod. Geophys., Toronto.

Jeffreys, H. (1962). The earth. 4th ed., Cambridge University Press.

Jung, K. (1956). Figur der Erde. Handbuch der Physik, v. 47, pp. 534–639, Berlin, Springer.

Jung, K. (1961). Schwerkraftverfahren in der Angewandten Geophysik. Leipzig, Akademische Verlagsgesellschaft Geest & Portig.

Lambert, W. D. (1961). The gravity field of an ellipsoid of revolution as a level surface. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 34.

Lambert, W. D., and F. W. Darling (1936). Tables for determining the form of the geoid and its indirect effect on gravity. U. S. Coast and Geodetic Survey, Spec. Publ. No. 199.

Ledersteger, K. (1957). Theoretischer Versuch einer exakten Lösung des gesamten Problems der Erdfigur. Schweizerische Zeitschrift für Vermessung, Kulturtechnik und Photogrammetrie, Sonderheft 2.

Marussi, A. (1949). Fondements de géométrie différentielle absolue du champ potentiel terrestre. Bull. Géod., No. 14, pp. 411–439.

Molodenskii, M. S., V. F. Eremeev, and M. I. Yurkina (1962). Methods for study of the external gravitational field and figure of the earth. Transl. from Russian (1960). Jerusalem, Israel Program for Scientific Translations.

Moritz, H. (1961). Über die Konvergenz der Kugelfunktionsentwicklung für das Außenraumpotential an der Erdoberfläche. Österreichische Zeitschrift für Vermessungswesen, v. 49, pp. 11–15.

Mueller, I. I. (1961). The determination of the regional part of the vertical gradient anomaly by a geodetic method. Geofisica pura e applicata, v. 48, pp. 3–8.

Pizzetti, P. (1894). Sulla espressione della gravità alla superficie del geoide, supposto ellissoidico. Atti R. Accad. Lincei, Ser. V, v. 3, p. 166.

Pizzetti, P. (1911). Sopra il calcolo teorico delle deviazioni del geoide dall' ellissoide. Atti R. Accad. Sci. Torino, v. 46, p. 331.

Sagrebin, D. W. (1956). Die Theorie des regularisierten Geoids. Veröff. Geod. Inst. Potsdam, No. 9.

Sollins, A. D. (1947). Tables for the computation of deflections of the vertical from gravity anomalies. Bull. Géod., No. 6, pp. 279–300.

Somigliana, C. (1929). Teoria generale del campo gravitazionale dell' ellissoide di rotazione. Mem. Soc. Astron. Ital., v. IV.

Stokes, G. G. (1849). On the variation of gravity on the surface of the earth. Trans. Cambridge Phil. Soc., v. 8, pp. 672–695.
Todhunter, I. (1873). A history of the mathematical theories of attraction and of the figure of the earth from the time of Newton to that of Laplace. London, Macmillan (New York, Dover Publications, 1962).
Uotila, U. A. (1960). Investigations on the gravity field and shape of the earth. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 33.
Vening Meinesz, F. A. (1928). A formula expressing the deflection of the plumb-line in the gravity anomalies and some formulae for the gravity field and the gravity potential outside the geoid. Proc. Koninkl. Ned. Akad. Wetenschap., v. 31 (3), pp. 315–331.

3
Gravimetric Methods

3-1. Gravity Reduction

Gravity g measured on the physical surface of the earth is not directly comparable with normal gravity γ referring to the surface of the ellipsoid. A reduction of g to sea level is necessary. Since there are masses above sea level, the reduction methods differ, depending on the way in which these *topographic masses* are dealt with.

Gravity reduction serves as a tool for three main purposes:

1. determination of the geoid,
2. interpolation and extrapolation of gravity,
3. investigation of the earth's crust.

Only the first two purposes are of a direct geodetic nature. The third is of interest to theoretical geophysicists and geologists, who study the general structure of the crust, and to exploration geophysicists, who search for shallow features which might indicate mineral deposits.

The use of Stokes' formula for the determination of the geoid requires that the gravity anomalies Δg represent *boundary values* at the geoid, which implies two conditions: first, gravity g must refer to the geoid; second, there must be no masses outside the geoid (Sec. 2-13). Hence, figuratively speaking, gravity reduction consists of the following steps: the topographic masses outside the geoid are completely removed or shifted below sea level; then the gravity station is lowered from the earth's surface (point P) to the geoid (point P_0, see Fig. 3-1).

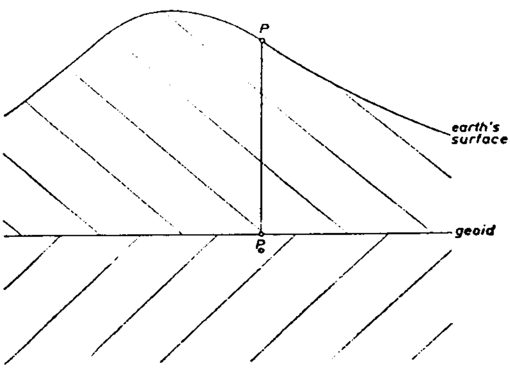

FIGURE 3-1
Gravity reduction.

The first step requires knowledge of the density of the topographic masses, which of course is somewhat problematical.

By such a reduction procedure certain irregularities in gravity due to differences in height of the stations are removed, so that interpolation and even extrapolation to unobserved areas become easier (Sec. 7-10).

3-2. Auxiliary Formulas

Let us compute the potential U and the vertical attraction A of a homogeneous circular cylinder of radius a and height b on a point P situated on its axis at a height c above its base (Fig. 3-2).

P outside cylinder. Assume first that P is above the cylinder, $c > b$. Then the potential is given by the general formula (1-11),

$$U = k \iiint \frac{\rho}{l}\, dv.$$

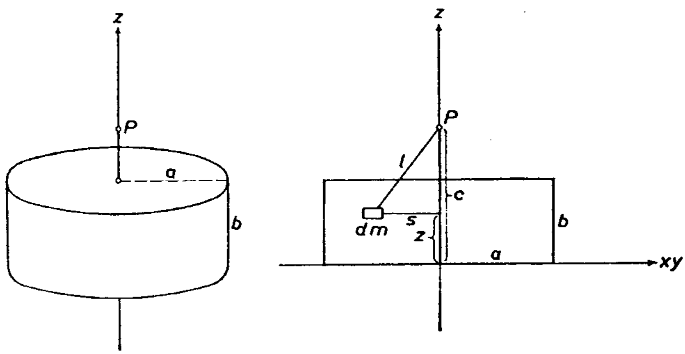

FIGURE 3-2
Potential and attraction of a circular cylinder on an external point.

Introducing polar coordinates s, α in the xy-plane by

$$x = s \cos \alpha, \quad y = s \sin \alpha \tag{3-1}$$

we have

$$l = \sqrt{s^2 + (c - z)^2}$$

and

$$dv = dx\, dy\, dz = s\, ds\, d\alpha\, dz.$$

Hence we find, with the density $\rho = \text{const.}$,

$$U = k\rho \int_{\alpha=0}^{2\pi} \int_{s=0}^{a} \int_{z=0}^{b} \frac{s\, ds\, dz\, d\alpha}{\sqrt{s^2 + (c - z)^2}}$$

$$= 2\pi k\rho \int_{s=0}^{a} \int_{z=0} \frac{s\, ds\, dz}{\sqrt{s^2 + (c - z)^2}}.$$

The integration with respect to s yields

$$\int_0^a \frac{s\, ds}{\sqrt{s^2 + (c - z)^2}} = \sqrt{s^2 + (c - z)^2}\Big|_0^a = \sqrt{a^2 + (c - z)^2} - c + z,$$

so that we have

$$U = 2\pi k\rho \int_0^b (-c + z + \sqrt{a^2 + (c - z)^2})\, dz.$$

The indefinite integral is $2\pi k\rho$ times

$$\frac{1}{2}(c - z)^2 - \frac{1}{2}(c - z)\sqrt{a^2 + (c - z)^2} - \frac{1}{2}a^2 \ln (c - z + \sqrt{a^2 + (c - z)^2}),$$

as may be verified by differentiation. Hence U finally becomes

$$\begin{aligned} U_e = \pi k\rho[(c - b)^2 - c^2 - (c - b)\sqrt{a^2 + (c - b)^2} + c\sqrt{a^2 + c^2} \\ - a^2 \ln (c - b + \sqrt{a^2 + (c - b)^2}) + a^2 \ln (c + \sqrt{a^2 + c^2})], \end{aligned} \tag{3-2}$$

where the subscript e denotes that P is external to the cylinder.

The vertical attraction A is the negative derivative of U with respect to the height c [compare equation (2-14)]:

$$A = -\frac{\partial U}{\partial c}. \tag{3-3}$$

Differentiating (3-2) we obtain

$$A_e = 2\pi k\rho[b + \sqrt{a^2 + (c - b)^2} - \sqrt{a^2 + c^2}]. \tag{3-4}$$

P on cylinder. In this case we have $c = b$ and equations (3-2) and (3-4) become

$$U_0 = \pi k\rho\left(-b^2 + b\sqrt{a^2 + b^2} + a^2 \ln \frac{b + \sqrt{a^2 + b^2}}{a}\right), \tag{3-5}$$

$$A_0 = 2\pi k\rho(a + b - \sqrt{a^2 + b^2}). \tag{3-6}$$

P inside cylinder. We assume that P is now inside the cylinder, $c < b$. By the plane $z = c$ we separate the cylinder into two parts, 1 and 2 (Fig. 3-3), and compute U as the sum of the contributions of these two parts:

$$U_i = U_1 + U_2,$$

where the subscript i denotes that P is now inside the cylinder. The term U_1 is given by (3-5) with b replaced by c, and U_2 by the same formula with b replaced by $b - c$. Their sum is

$$U_i = \pi k\rho \Big[-c^2 - (b-c)^2 + c\sqrt{a^2+c^2} + (b-c)\sqrt{a^2+(b-c)^2} + a^2 \ln \frac{c + \sqrt{a^2+c^2}}{a} + a^2 \ln \frac{b - c + \sqrt{a^2+(b-c)^2}}{a} \Big]. \tag{3-7}$$

It is easily seen that the attraction is the difference $A_1 - A_2$:

$$A_i = 2\pi k\rho[2c - b - \sqrt{a^2+c^2} + \sqrt{a^2+(b-c)^2}]; \tag{3-8}$$

this formula may also be obtained by differentiating (3-7) according to (3-3).

Circular disk. Let the thickness b of the cylinder go to zero, the product

$$\kappa = b\rho$$

remaining finite. The quantity κ may then be considered as the surface density (Sec. 1-3) with which matter is concentrated on the surface of a circle of radius a. We need potential and attraction for an exterior point. By setting

$$\rho = \frac{\kappa}{b}$$

in (3-2) and (3-4) and then letting $b \longrightarrow 0$, we get by well-known methods of the calculus

$$U_e^0 = 2\pi k\kappa(\sqrt{a^2+c^2} - c), \tag{3-9}$$

$$A_e^0 = 2\pi k\kappa\left(1 - \frac{c}{\sqrt{a^2+c^2}}\right). \tag{3-10}$$

Sectors and compartments. The preceding formulas are not used for an entire cylinder or disk, but for sectors and compartments such as those shown in Fig. 2-22. For a sector of radius a and angle

$$\alpha = \frac{2\pi}{n} \tag{3-11}$$

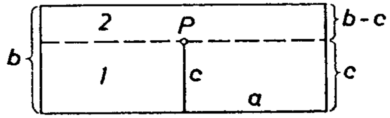

FIGURE 3-3

Potential and attraction on an internal point.

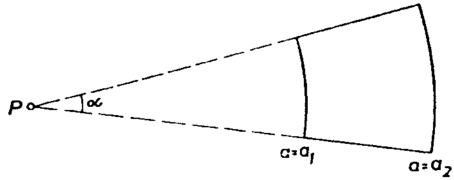

FIGURE 3-4

A template compartment.

we must divide the above formulas by n. For a compartment subtending the same angle and bounded by the radii a_1 and a_2 (Fig. 3-4), we get, in an obvious notation,

$$\Delta U = \frac{1}{n}[U(a_2) - U(a_1)], \tag{3-12}$$

$$\Delta A = \frac{1}{n}[A(a_2) - A(a_1)]. \tag{3-13}$$

Since A_e and A_i differ only by a constant, this constant drops out in (3-13) and we obtain from (3-4) and (3-8)

$$\Delta A_e = \Delta A_i = \frac{2\pi}{n} k\rho[\sqrt{a_2^2 + (c-b)^2} - \sqrt{a_1^2 + (c-b)^2} - \sqrt{a_2^2 + c^2} + \sqrt{a_1^2 + c^2}]. \tag{3-14}$$

On the other hand, $\Delta U_e \neq \Delta U_i$.

3-3. The Bouguer Reduction

The object of the Bouguer reduction of gravity is the complete removal of the topographic masses, that is, the masses outside the geoid.

The Bouguer plate. Assume the area around the gravity station P to be completely flat and horizontal (Fig. 3-5), and let the masses between the geoid and the earth's surface have a constant density ρ. Then the attraction A of this so-called *Bouguer plate* is obtained by letting $a \longrightarrow \infty$ in (3-6), since the plate, considered plane, may be regarded as a circular cylinder of thickness $b = h$ and infinite radius. By well-known rules of the calculus we obtain

$$A_B = 2\pi k\rho h \tag{3-15}$$

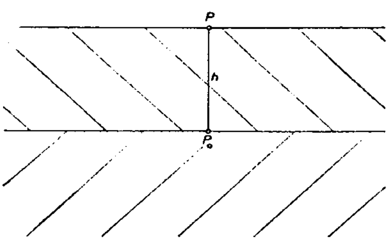

FIGURE 3-5

The Bouguer plate.

as the attraction of an infinite Bouguer plate. With standard density $\rho = 2.67$ g/cm^3 this becomes

$$A_B = 0.1119h \text{ mgal} \tag{3-15'}$$

for h in meters.

Removing the plate is equivalent to subtracting its attraction (3-15) from the observed gravity. This is called *incomplete Bouguer reduction.*

To complete our gravity reduction we must then lower the gravity station from P to the geoid, to P_0. This is done by applying the *free-air reduction*

$$F = -\frac{\partial g}{\partial h}h, \tag{3-16}$$

so called because after removal of the topography the station P is in "free air." For many practical purposes it is sufficient to use the normal gradient of gravity, obtaining

$$F \doteq -\frac{\partial \gamma}{\partial h}h \doteq +0.3086h \text{ mgal} \tag{3-17}$$

for h in meters.

This combined process of removing the topographic masses and applying the free-air reduction is called *complete Bouguer reduction.* Its result is Bouguer gravity at the geoid:

$$g_B = g - A_B + F. \tag{3-18}$$

With the assumed numerical values we have

gravity measured at P	g	
minus Bouguer plate	$-0.1119h$	
plus free-air reduction	$+0.3086h$	
Bouguer gravity at P_0	$g_B = g + 0.1967h.$	(3-18′)

Since g_B now refers to the geoid, we obtain genuine gravity anomalies in the sense of Sec. 2-13 by subtracting normal gravity γ referred to the ellipsoid:

$$\Delta g_B = g_B - \gamma. \tag{3-19}$$

They are called *Bouguer anomalies.*

Terrain correction. This simple procedure is refined by taking into account the deviation of the actual topography from the Bouguer plate of P (Fig. 3-6). This is called *terrain correction.* At A the mass surplus Δm_+, which attracts upward, is removed, causing g at P to increase. At B the mass deficiency Δm_- is made up, causing g at P to increase again. *The terrain correction is always positive.*

The practical determination of the terrain correction A_t is done by means of a template similar to that shown in Fig. 2-22, using (3-14) and adding the effects of the individual compartments:

$$A_t = \sum \Delta A. \tag{3-20}$$

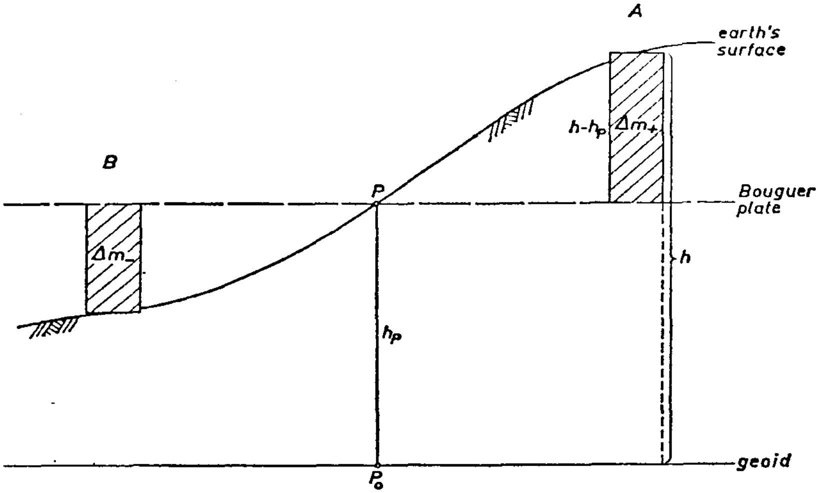

FIGURE 3-6
The terrain correction.

For a surplus mass Δm_+, $h > h_P$, we have

$$b = h - h_P, \quad c = 0;$$

and for a mass deficiency Δm_-, $h < h_P$,

$$b = c = h_P - h.$$

By adding the terrain correction A_t to (3-18) we obtain the *refined Bouguer gravity*

$$g_B = g - A_B + A_t + F. \tag{3-21}$$

The Bouguer reduction and the corresponding Bouguer anomalies Δg_B are called *refined* or *simple*, depending on whether the terrain correction has been applied.

In practice it is convenient to separate the Bouguer reduction into the effect of a Bouguer plate and the terrain correction, because the latter is usually much less. Even for mountains 3000 meters in height the terrain correction is only of the order of 50 mgals (Heiskanen and Vening Meinesz, 1958, p. 154).

Unified procedure. It is also possible to compute the total effect of the topographic masses,

$$A_T = A_B - A_t, \tag{3-22}$$

in one step by using columns with base at sea level (Fig. 3-7), again subdividing the terrain by means of a template. (Note the difference between A_T, the attraction of the topographic masses, and the terrain correction A_t!) Then

$$A_T = \sum \Delta A,$$

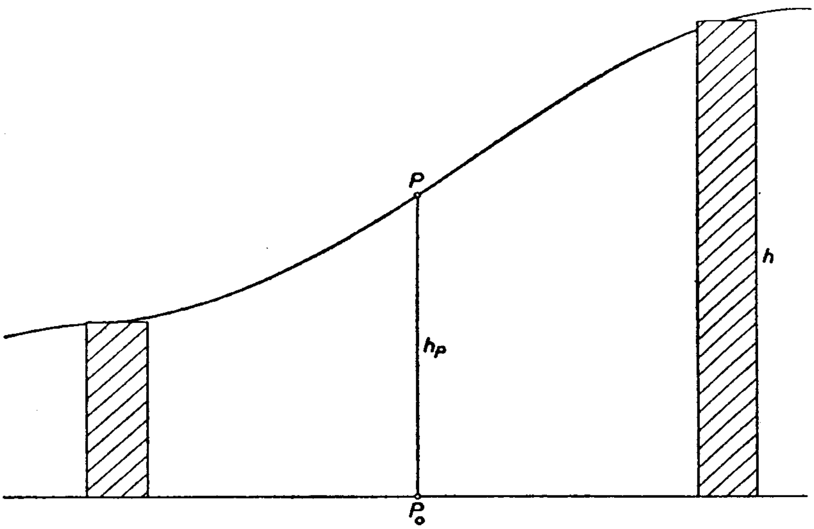

FIGURE 3-7

The Bouguer reduction.

where we now have $b = h$, $c = h_P$. For the innermost circle use (3-6) with $b = h_P$.

Instead of (3-21) we now have

$$g_B = g - A_T + F. \tag{3-21'}$$

The Bouguer reduction may be still further refined by the consideration of density anomalies, anomalies in the free-air gradient of gravity (Sec. 2-23), and spherical effects. More computational formulas may be found in Jung (1961, Sec. 6.4).

3-4. Isostasy

One might be inclined to assume that the topographic masses are simply superposed on an essentially homogeneous crust. If this were the case, the Bouguer reduction would remove the main irregularities of the gravity field, so that the Bouguer anomalies would be very small and would fluctuate randomly around zero. However, just the opposite is true. Bouguer anomalies in mountainous areas are systematically negative and may attain large values, increasing in magnitude on the average by 100 mgals per 1000 meters of elevation. The only explanation possible is that there is some kind of mass deficiency under the mountains. This means that the topographic masses are *compensated* in some way.

There is a similar effect for the deflections of the vertical. The actual deflections are smaller than the visible topographic masses would suggest. In the middle of the nineteenth century J. H. Pratt observed such an effect in the Himalayas. At one station in this area he computed a value of 28″ for the deflection of the vertical from the attraction of the visible masses of the mountains. The value obtained through astrogeodetic measurements was only 5″. Again, some kind of compensation is needed to account for this discrepancy.

Two different theories for such a compensation were developed at almost exactly the same time, by J. H. Pratt in 1854 and 1859, and by G. B. Airy in 1855. According to Pratt the mountains have risen from the underground somewhat like a fermenting dough. According to Airy, the mountains are floating on a fluid lava of higher density, so that the higher the mountain, the deeper it sinks.

Pratt-Hayford system. This system of compensation was outlined by Pratt and put into a mathematical form by J. F. Hayford, who used it systematically for geodetic purposes.

The principle is illustrated by Fig. 3-8. Underneath the level of compensation there is uniform density. Above, the mass of each column of the same cross section is equal. Let D be the depth of the level of compensation, reckoned from sea level, and let ρ_0 be the density of a column of height D. Then the density ρ of a column of height $D + h$ (h representing the height of the topography) satisfies the equation

$$(D + h)\rho = D\rho_0, \tag{3-23}$$

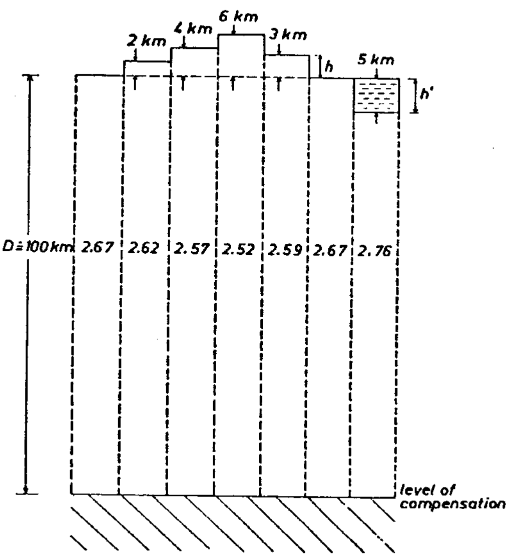

FIGURE 3-8
Isostasy—
Pratt-Hayford model.

which expresses the condition of equal mass. It may be assumed that

$$\rho_0 = 2.67 \text{ g/cm}^3. \tag{3-24}$$

According to (3-23), the actual density ρ is slightly smaller than this normal value ρ_0. Consequently, there is a mass deficiency which, according to (3-23), is given by

$$\Delta\rho = \rho_0 - \rho = \frac{h}{D + h}\rho_0. \tag{3-25}$$

In the oceans the condition of equal mass is expressed as

$$(D - h')\rho + h'\rho_w = D\rho_0, \tag{3-26}$$

where

$$\rho_w = 1.027 \text{ g/cm}^3 \tag{3-27}$$

is the density and h' the depth of the ocean. Hence there is a mass surplus of a suboceanic column given by

$$\rho - \rho_0 = \frac{h'}{D - h'}(\rho_0 - \rho_w). \tag{3-28}$$

As a matter of fact this model of compensation is idealized and schematic. It can be only approximately fulfilled in nature. Values of the depth of compensation around

$$D = 100 \text{ km} \tag{3-29}$$

are assumed.

For a spheroidal earth the columns will converge slightly towards its center, and other refinements may be introduced. We may postulate either equality of mass or equality of pressure; each postulate leads to somewhat different spherical refinements. It may be mentioned that for computational reasons Hayford used still another, slightly different model; for instance, he reckoned the depth of compensation D from the earth's surface instead of from sea level.

Airy-Heiskanen system. Airy proposed this model, and Heiskanen gave it a precise formulation for geodetic purposes and applied it extensively.

Figure 3-9 illustrates the principle. The mountains, of constant density

$$\rho_0 = 2.67 \text{ g/cm}^3, \tag{3-30}$$

float on a denser underlayer of constant density

$$\rho_1 = 3.27 \text{ g/cm}^3. \tag{3-31}$$

The higher they are, the deeper they sink. Thus, *root formations* exist under mountains, and "antiroots" under the oceans.

We denote the density difference $\rho_1 - \rho_0$ by $\Delta\rho$. With the assumed numerical values we have

$$\Delta\rho = \rho_1 - \rho_0 = 0.6 \text{ g/cm}^3. \tag{3-32}$$

If we denote the height of the topography by h and the thickness of the corresponding root by t (Fig. 3-9), then the condition of floating equilibrium is

$$t\Delta\rho = h\rho_0, \tag{3-33}$$

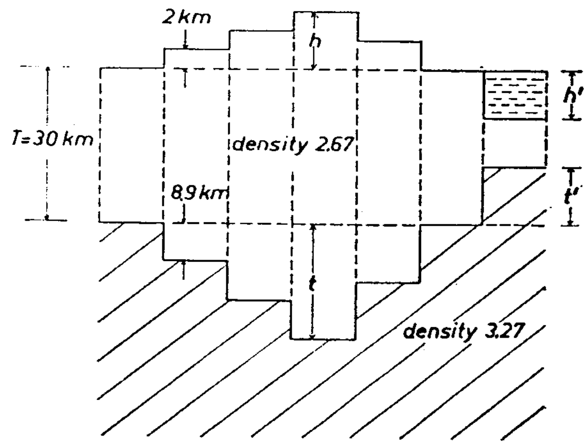

FIGURE 3-9
Isostasy—Airy-Heiskanen model.

so that

$$t = \frac{\rho_0}{\Delta\rho} h = 4.45\text{h}. \tag{3-34}$$

For the oceans the corresponding condition is

$$t' \Delta\rho = h'(\rho_0 - \rho_w), \tag{3-35}$$

where h' and ρ_w are defined as above and t' is the thickness of the antiroot (Fig. 3-9), so that we get

$$t' = \frac{\rho_0 - \rho_w}{\rho_1 - \rho_0} h' = 2.73h' \tag{3-36}$$

for the numerical values assumed.

Again spherical corrections must be applied to these formulas for higher accuracy, and the formulations in terms of equal mass and equal pressure lead to slightly different results.

The normal thickness of the earth's crust is denoted by T (Fig. 3-9). Values of around

$$T = 30\text{ km} \tag{3-37}$$

are assumed. The crustal thickness under mountains is then

$$T + h + t, \tag{3-38}$$

and under the oceans it is

$$T - h' - t'. \tag{3-39}$$

Vening Meinesz regional system. Both systems just discussed are highly idealized in that they assume the compensation to be strictly *local*; that is, they assume that compensation takes place along vertical columns. This presupposes free mobility of the masses to a degree that is obviously unrealistic in this strict form.

For this reason Vening Meinesz modified the Airy floating theory in 1931, introducing regional instead of local compensation. The principal difference between these two kinds of compensation is illustrated by Fig. 3-10. In Vening Meinesz' theory the topography is considered as a load on an unbroken but yielding elastic crust.

Although Vening Meinesz' refinement of Airy's theory is more realistic, it is more complicated and is therefore seldom used by geodesists because, as we shall see, any isostatic system, if consistently applied, serves for geodetic purposes as well.

Geophysical and geodetic evidence shows that the earth is about 90% isostatically compensated, but it is difficult to decide, at least from gravimetric evidence alone, which model best accounts for this compensation. Although seismic results indicate an Airy type of compensation, in some places the compensation seems to follow the Pratt model. Nature will, of course, never conform to any of these models to the degree of precision which we have assumed above. However, a well-defined and consistent mathematical formulation is a prerequisite for the application of isostasy for geodetic purposes.

For more details on isostasy and its geophysical applications see Heiskanen and Vening Meinesz (1958, Chapters 5 and 7).

3-5. Isostatic Reductions

The object of isostatic reduction of gravity is the regularization of the earth's crust according to some model of isostasy. The topographic masses are not completely removed as they are in the Bouguer reduction, but are shifted into the interior of the geoid in order to make up the mass deficiencies that exist under the continents. In the isostatic model of Pratt and Hayford the topographic masses are distributed between the level of compensation and sea level, in order to bring the crustal density from its original value to the constant standard value ρ_0. In the Airy-Heiskanen model the topographic masses are used to fill the roots of the continents, bringing the density from $\rho_0 = 2.67$ to $\rho_1 = 3.27$ g/cm^3.

In other words, the topography is removed together with its compensation, and the final result is ideally a homogeneous crust of density ρ_0 and constant thickness D (Pratt-Hayford) or T (Airy-Heiskanen).

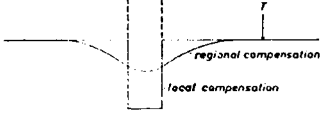

FIGURE 3-10

Local and regional compensation.

Thus we have three steps:

1. removal of topography,
2. removal of compensation,
3. free-air reduction to the geoid.

Steps 1 and 3 are known from Bouguer reduction, so that the techniques of Sec. 3-3 can be applied to them. Step 2 is new and will be discussed now for the two main isostatic systems.

Pratt-Hayford system. The method is the same as for the terrain correction, Sec. 3-3, equation (3-20). The attraction of the (negative) compensation is again computed by

$$A_C = \sum \Delta A,$$

where the attraction of a vertical column representing a compartment is given by (3-14) with

$$b = D, \quad c = D + h_P$$

and ρ replaced by the density defect $\Delta\rho$. If the preceding Bouguer reduction were done with the original density ρ of the column expressed by

$$\rho = \frac{D}{D + h}\rho_0 \tag{3-40}$$

according to (3-23), then $\Delta\rho$ would be given by (3-25).

Usually the Bouguer reduction is performed using the constant density ρ_0; the density defect $\Delta\rho$ must then be computed by

$$\Delta\rho = \frac{h}{D}\rho_0, \tag{3-41}$$

which differs slightly from (3-25), in order to restore equality of mass according to

$$(\rho_0 - \Delta\rho)D + \rho_0 h = \rho_0 D.$$

The first term on the left-hand side represents the mass of the layer between the level of compensation and sea level; the second term represents the mass of the topography, now assumed to have a density ρ_0.

Airy-Heiskanen system. Again we use

$$A_C = \sum \Delta A,$$

where b and c in (3-14) are, according to Fig. 3-11, given by

$$b = t, \quad c = h_P + T + t$$

and ρ is replaced by $\Delta\rho = \rho_1 - \rho_0 = 0.6 \text{ g/cm}^3$.

Total reduction. In analogy with (3-21′) isostatically reduced gravity on the geoid becomes

$$g_I = g - A_T + A_C + F, \tag{3-42}$$

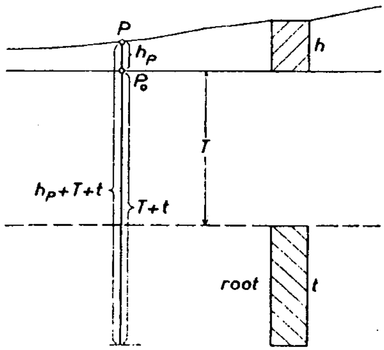

FIGURE 3-11

Topography and compensation—Airy-Heiskanen model.

where $-A_C$ is the attraction of the compensation which is actually negative, so that its removal is equivalent to the term $+A_C$. The quantity A_T is the attraction of topography, to be computed as the effect of a Bouguer plate combined with terrain correction, equation (3-22), or in one step, as described in Sec. 3-3; F is the free-air reduction approximated by (3-17).

Oceanic stations. Here the terms A_T and F of (3-42) are zero, since the station is situated on the geoid, but the term A_C is more complicated.

In the Pratt-Hayford model the procedure is as follows. The mass surplus (3-28) of a suboceanic column of height $D - h'$ (Fig. 3-8) is removed and used to fill the corresponding oceanic column of height h' to the proper density ρ_0. In mathematical terms, this is

$$A_C = -A_1 + A_2, \tag{3-43}$$

where both A_1 and A_2 are of the form (3-20), ΔA being given by (3-14). For A_1 we have

$$b = D - h', \quad c = D$$

and density $\rho - \rho_0$; for A_2 we have

$$b = c = h'$$

and density $\rho_0 - \rho_w$.

In the Airy-Heiskanen model the mass surplus of the antiroot, $\rho_1 - \rho_0$, is used to fill the oceans to the proper density ρ_0. The corresponding value is again given by (3-43) where for A_1 we now have

$$b = t', \quad c = T$$

and density $\rho_1 - \rho_0$; and for A_2 we have, as before,

$$b = c = h'$$

and density $\rho_0 - \rho_w$.

In both models (3-42) reduces for oceanic stations to

$$g_{I,\,\text{ocean}} = g + A_C. \tag{3-44}$$

Isostatic anomalies. The isostatic gravity anomalies are—in analogy to the Bouguer anomalies—defined by

$$\Delta g_I = g_I - \gamma. \tag{3-45}$$

If any of the isostatic systems were rigorously true, then isostatic reduction would fulfil perfectly its object of complete regularization of the earth's crust, which would become level and homogeneous. Then, with a properly chosen reference model for γ, the isostatic anomalies (3-45) would be zero.

The actual isostatic compensation occurring in nature cannot, of course, completely conform to such abstract models. Therefore, nonzero isostatic anomalies (3-45) will be left, but they will be small, smooth, and more or less randomly positive and negative. On account of this smoothness and independence of elevation they are better suited for interpolation or extrapolation than any other type of anomalies; see Chapter 7, particularly Sec. 7-10.

Their computation is relatively laborious; but this fact is rather insignificant in view of the present possibility of using automatic computers. Reduction tables and reduction maps further facilitate the work.

The gravity tables of Hayford and Bowie (1912) have been the prototype of later tables. Hayford's tables give the effects of topography and its isostatic compensation as functions of the height of the compartment, using the isostatic system of Pratt and Hayford with $\rho_0 = 2.67$ g/cm³ and $D = 113.7$ km. Hayford's

Table 3-1

Hayford Zones and Compartments

Zone	*Outer radius (meters)*	*Number of compartments*	*Zone*	*Outer radius*	*Number of compartments*
A	2	1	18	1°41′13″	1
B	68	4	17	1 54 52	1
C	230	4	16	2 11 53	1
D	590	6	15	2 33 46	1
E	1,280	8	14	3 03 05	1
F	2,290	10	13	4 19 13	16
G	3,520	12	12	5 46 34	10
H	5,240	16	11	7 51 30	8
I	8,440	20	10	10 44	6
J	12,400	16	9	14 09	4
K	18,800	20	8	20 41	4
L	28,800	24	7	26 41	2
M	58,800	14	6	35 58	18
N	99,000	16	5	51 04	16
O	166,700	28	4	72 13	12
	(1°29′58″)		3	105 48	10
			2	150 56	6
			1	180 00	1

division into zones and compartments is shown in Table 3-1. The curious values for the radii of the outer zones 1 through 18 are due to the use of the foot as the unit of length. Nevertheless, Hayford's division has become standard for later tables, such as those of Cassinis et al. and of Heiskanen.

The fundamental tables of Cassinis et al. (1937) are essentially a tabulation of the general formula (3-14). They can therefore be used for all kinds of gravity reduction, but for practical application of any reduction method special tables must be computed from them.

Heiskanen (1938) made such special tables for the Airy-Heiskanen isostatic system with $\rho_0 = 2.67$, $\rho_1 = 3.27$ g/cm³, and normal thickness $T = 20$, 30, 40, and 60 km. These tables supersede Heiskanen's (1931) tables for $T = 40$, 60, 80, and 100 km.

Tables for regional isostatic reduction according to the system of Vening Meinesz are given in Vening Meinesz (1939).

Since the combined effect of topography and compensation for the Hayford zones 1 through 18 varies only slowly and smoothly, this effect can be given in the form of an isoanomaly map. This was done by Heiskanen and Nuotio (1938); Niskanen and Kivioja (1951); Heiskanen, Niskanen, and Kärki (1959); and Kärki, Kivioja, and Heiskanen (1961). The last map covers the whole world and uses the Airy-Heiskanen system with $T = 30$ km.

For the Hayford zones 1 through 18 the planar formulas of Sec. 3-2 are no longer sufficient, and spherical formulas must be used. These expressions are highly complicated; we therefore refer the reader to Lambert (1930), Baeschlin (1948, pp. 480–506), and Heiskanen and Vening Meinesz (1958, p. 162). For formulas for automatic computers see Kukkamäki (1955). In Sec. 3-8 we shall give a simplified qualitative treatment.

It may be stressed again that for geodetic purposes the isostatic model used must be mathematically precise and self-consistent, and the same model must be used throughout. Refinements include the consideration of irregularities of density of the topographic masses and the consideration of the anomalous gradient of gravity.

3-6. The Indirect Effect

The removal or shifting of masses which underly the gravity reductions change the gravity potential and hence the geoid. This change of the geoid is an *indirect effect* of the gravity reductions.

Hence the surface computed by Stokes' formula from isostatic anomalies, say, is not the geoid itself but a slightly different surface, the *cogeoid*. To every gravity reduction there corresponds a different cogeoid.

Let the undulation of the cogeoid be N^c. Then the undulation N of the actual geoid is obtained from

$$N = N^c + \delta N, \tag{3-46}$$

by taking the indirect effect on N into account, which is given by

$$\delta N = \frac{\delta W}{\gamma}, \tag{3-47}$$

where δW is the change of potential at the geoid. Equation (3-47) is an application of Bruns' theorem (2-144).

The change of potential δW is for the Bouguer reduction expressed by

$$\delta W_B = U_T, \tag{3-48}$$

and for the isostatic reduction by

$$\delta W_I = U_T - U_C, \tag{3-49}$$

the subscripts of the potential U corresponding to those of the attraction A used in the preceding sections.

For the practical determination of U_T and U_C the template technique, as expressed in (3-20), may again be used:

$$U = \sum \Delta U, \tag{3-50}$$

where the relevant formulas are (3-12), (3-2), (3-5), and (3-7). The point to which U refers is always the point P_0 at sea level (Fig. 3-1).

For U_T we use U_0, (3-5), with $b = h$ and density ρ_0 (see Fig. 3-11). For U_C in the continental case we use U_e, (3-2), with the following values:

Pratt-Hayford:

$$b = c = D, \quad \text{density } \frac{h}{D}\rho_0;$$

Airy-Heiskanen:

$$b = t, \quad c = t + T, \quad \text{density } \rho_1 - \rho_0.$$

The corresponding considerations for the oceanic case are left as an exercise for the reader.

The indirect effect with Bouguer anomalies is very great, of the order of ten times the geoidal undulation itself. See the map (Tafel I) at the end of Helmert (1884), where the maximum value is 440 meters. The reason is, of course, that the earth is in general isostatically compensated. Therefore, the Bouguer anomalies cannot be used for the determination of the geoid.

With isostatic anomalies the indirect effect is smaller than N, of the order of 10 meters, as might be expected. It is necessary, however, to compute the indirect effect δN_I carefully, using exactly the same isostatic model as for the gravity reductions.

Furthermore, before applying Stokes' formula, the isostatic gravity anomalies must be reduced from the geoid to the cogeoid. This is done by a simple free-air reduction, using (3-17), by adding to Δg_I the correction

$$\delta = +0.3086\delta N \text{ mgal}, \tag{3-51}$$

δN in meters. This correction δ is the *indirect effect on gravity*; it is of the order of 3 mgal.

Now the isostatic anomalies refer strictly to the cogeoid. The application of Stokes' formula gives N^c, which according to (3-46) is to be corrected by the indirect effect δN to give the undulation N of the actual geoid.

For the Hayford zones 1 through 18 spherical formulas must be used to compute U. For these formulas we again refer the reader to Lambert (1930) and Baeschlin (1948, pp. 480–506). Tables and maps have been given by Lambert and Darling (1936) and Heiskanen and Niskanen (1941). See also Secs. 3-8 and 8-2.

Deflections of the vertical. The indirect effect on the deflections of the vertical is, in agreement with equations (2-204), given by

$$\delta\xi = -\frac{1}{R}\frac{\partial\delta N}{\partial\phi},$$
$$\delta\eta = -\frac{1}{R\cos\phi}\frac{\partial\delta N}{\partial\lambda}. \tag{3-52}$$

The indirect effect is essentially identical with the so-called *topographic-isostatic deflection of the vertical* (Heiskanen and Vening Meinesz, 1958, pp. 252–255). Spherical formulas and fundamental tables, similar to those of Cassinis et al. (1937) for gravity, were given by Lambert and Darling (1938) and Darling (1949). See also Baeschlin (1948, pp. 336–380).

3-7. Other Gravity Reductions

The inversion reduction of Rudzki. It is possible to find a gravity reduction where the indirect effect is zero. This is done by shifting the topographic masses into the interior of the geoid in such a way that

$$U_C = U_T. \tag{3-53}$$

Then

$$\delta W = U_T - U_C = 0. \tag{3-53'}$$

This procedure was given by M. P. Rudzki in 1905. For the present purpose we may consider the geoid to be a sphere of radius R (Fig. 3-12). Let the mass element dm at Q be replaced by a mass element dm' at a certain point Q' inside the geoid situated on the same radius vector.

The potential due to these mass elements at the geoidal point P_0 is

$$dU_T = k\frac{dm}{l} = \frac{kdm}{\sqrt{r^2 + R^2 - 2Rr\cos\psi}},$$
$$dU_C = k\frac{dm'}{l'} = \frac{kdm'}{\sqrt{r'^2 + R^2 - 2Rr'\cos\psi}}. \tag{3-54}$$

We shall have

$$dU_C = dU_T$$

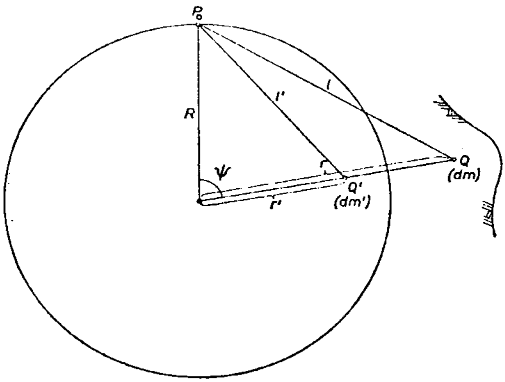

FIGURE 3-12

Rudzki reduction as an inversion in a sphere.

if

$$dm' = \frac{R}{r}\,dm \tag{3-55}$$

and

$$r' = \frac{R^2}{r}. \tag{3-56}$$

This is readily verified by substitution into the second of equations (3-54).

The condition (3-56) means that Q' and Q are related by *inversion in the sphere* of radius R (Kellogg, 1929, p. 231). Therefore, this reduction method is called *inversion reduction* or *Rudzki reduction.*

The condition (3-55) expresses the fact that the compensating mass dm' is not exactly equal to dm, but is slightly smaller. Since this relative decrease of mass is of the order of 10^{-8}, it may be safely neglected by setting

$$dm' = dm. \tag{3-55'}$$

Usually it is even sufficient to replace the sphere by a plane. Then Q' is the ordinary mirror image of Q (Fig. 3-13).

Rudzki gravity at the geoid becomes, in analogy to (3-42),

$$g_R = g - A_T + A_C + F, \tag{3-57}$$

where $A_C = \sum \Delta A$ with $b = h$, $c = h + h_P$, the density being equal to that of topography.

Since the indirect effect is zero, the cogeoid of Rudzki coincides with the actual geoid, but the gravity field outside the earth is changed. In addition,

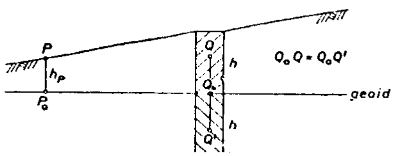

FIGURE 3-13

Rudzki reduction as a plane approximation.

the Rudzki reduction does not correspond to a geophysically meaningful model.

The condensation reduction of Helmert. Here the topography is condensed so as to form a surface layer (Sec. 1-3) on the geoid[1] of density

$$\kappa = \rho h, \tag{3-58}$$

so that the total mass remains unchanged. Again, the mass is shifted along the local vertical (Fig. 3-14).

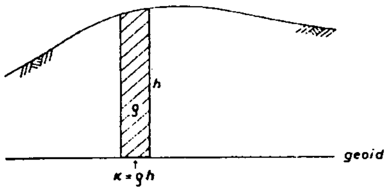

FIGURE 3-14
Helmert's method of condensation.

We may consider Helmert's condensation as a limiting case of an isostatic reduction according to the Pratt-Hayford system as the depth of compensation D goes to zero. This is often useful.[2]

Again we have

$$g_H = g - A_T + A_C + F, \tag{3-59}$$

where $A_C = \sum \Delta A$ is now to be computed using formula (3-10) with $c = h_P$ and $\kappa = \rho h$; h_P is the height of the station P and h the height of the compartment.

The indirect effect is

$$\delta W = U_T - U_C. \tag{3-60}$$

The potential $U_C = \sum \Delta U$ is to be computed using formula (3-9) with $\kappa = \rho h$ as before, but $c = 0$ since it refers to the geoidal point P_0. The corresponding δN is very small, being about 1 meter per 3 km of average elevation. It may therefore be neglected, so that the cogeoid of the condensation reduction practically coincides with the actual geoid.

Even the "direct effect," $-A_T + A_C$, can usually be neglected, as the attraction of the Helmert layer nearly compensates that of the topography. There remains

$$g_H = g + F, \tag{3-61}$$

that is, the simple free-air reduction. In this sense, *the simple free-air reduction may be considered as giving approximate boundary values at the geoid,* to be

[1] This is Helmert's "second method of condensation" (Lambert, 1930). In his first method Helmert (1884, pp. 148–186) condensed the topography on a parallel surface situated 21 km (the difference $a - b$) below the geoid. He did this in order to avoid problems connected with the convergence of the spherical-harmonic series for the potential (Sec. 2-5).

[2] Even the Bouguer reduction may be considered as a limiting case of an isostatic reduction—namely, for the Airy-Heiskanen system as $T \to \infty$, corresponding to a complete removal of the topographic masses.

used in Stokes' formula. To the same degree of approximation, the "free-air cogeoid" coincides with the actual geoid.

Hence the free-air anomalies

$$\Delta g_F = g + F - \gamma \tag{3-62}$$

may be considered as approximations of "condensation anomalies"

$$\Delta g_H = g_H - \gamma. \tag{3-63}$$

The subject of free-air anomalies will be taken up again in Chapter 8, in a rather different context.

The reduction of Poincaré and Prey does not properly belong here because it is intended to provide actual values of gravity inside the earth, and not to give boundary values at the geoid. It cannot be directly used for the determination of the geoid, but is needed to obtain orthometric heights and will therefore be discussed in Sec. 4-3. Actual gravity g_0 at a geoidal point P_0 is related to Bouguer gravity g_B, equation (3-21′), by

$$g_0 = g_B - A_{T,P_0}. \tag{3-64}$$

It is obtained by subtracting from g_B the attraction A_{T,P_0} of the topographic masses on P_0, which corresponds to restoring the topography after the free-air reduction of Bouguer gravity from P to P_0.

These are the main methods that have been proposed for the reduction of gravity.

3-8. Spherical Effects

For the spherical Hayford zones 1 through 18, or at least outside a distance of 400 km from the station, spherical formulas must be used for the computation of gravity reduction and the indirect effect.

We have already mentioned that the exact spherical formulas are very complicated. Therefore, we shall consider here an approximate technique that provides simple qualitative insight and yields results that are accurate enough for many purposes. We shall use surface potentials of simple and double layers on the sphere; see Secs. 1-3 and 1-4.

Surface potentials on a sphere. The potential U_1 at an exterior point P of radius vector

$$r = R + h_P \tag{3-65}$$

due to a simple layer of density κ on a sphere of radius R is given by

$$U_1 = k \iint_\sigma \frac{\kappa}{l} R^2 \, d\sigma = kR^2 \iint_\sigma \frac{\kappa}{l} \, d\sigma. \tag{3-66}$$

The vertical attraction A_1 of this layer at P is the negative radial derivative of U_1:

$$A_1 = -\frac{\partial U_1}{\partial r} = -kR^2 \iint_\sigma \kappa \frac{\partial}{\partial r}\left(\frac{1}{l}\right) d\sigma. \tag{3-67}$$

The potential U_2 of a double layer of density μ on this sphere is expressed by

$$U_2 = k \iint_\sigma \mu \frac{\partial}{\partial n}\left(\frac{1}{l}\right) R^2\, d\sigma = kR^2 \iint_\sigma \mu \frac{\partial}{\partial R}\left(\frac{1}{l}\right) d\sigma, \tag{3-68}$$

since the normal at the surface element $dS = R^2\, d\sigma$ is the radius vector R; compare also Fig. 2-14. The corresponding attraction A_2 is therefore given by

$$A_2 = -\frac{\partial U_2}{\partial r} = -kR^2 \iint_\sigma \mu \frac{\partial^2}{\partial r\, \partial R}\left(\frac{1}{l}\right) d\sigma. \tag{3-69}$$

We now need $1/l$ and its derivatives. We have

$$\frac{1}{l} = (r^2 + R^2 - 2rR\cos\psi)^{-1/2}. \tag{3-70}$$

By differentiation we find

$$\frac{\partial}{\partial r}\left(\frac{1}{l}\right) = -\frac{r - R\cos\psi}{l^3}, \quad \frac{\partial}{\partial R}\left(\frac{1}{l}\right) = -\frac{R - r\cos\psi}{l^3}; \tag{3-71}$$

$$\frac{\partial^2}{\partial r\, \partial R}\left(\frac{1}{l}\right) = \frac{1}{l^5}\left[l^2\cos\psi + 3(R - r\cos\psi)(r - R\cos\psi)\right]. \tag{3-72}$$

Now let P move onto the sphere, $P \longrightarrow P_0$. Then these formulas become

$$\begin{aligned} \frac{1}{l} &= \frac{1}{2R\sin(\psi/2)}, \\ \frac{\partial}{\partial r}\left(\frac{1}{l}\right) &= \frac{\partial}{\partial R}\left(\frac{1}{l}\right) = -\frac{1}{4R^2\sin(\psi/2)}, \\ \frac{\partial^2}{\partial r\, \partial R}\left(\frac{1}{l}\right) &= \frac{1}{8R^3}\left(\frac{1}{\sin^3(\psi/2)} + \frac{1}{\sin(\psi/2)}\right). \end{aligned} \tag{3-73}$$

As a further specialization we consider the case that κ and μ are constant on the spherical zone bounded by the spherical distances ψ_0 and ψ (Fig. 3-15); outside this zone we assume $\kappa = \mu = 0$. We then have

$$\iint_\sigma \kappa\, d\sigma = \kappa \int_{\alpha=0}^{2\pi} \int_{\psi=\psi_0}^{\psi} \sin\psi\, d\psi\, d\alpha = 2\pi\kappa \int_{\psi_0}^{\psi} \sin\psi\, d\psi \tag{3-74}$$

and an analogous expression for μ. By taking this into account and inserting equations (3-73) into equations (3-66) to (3-69), these formulas become

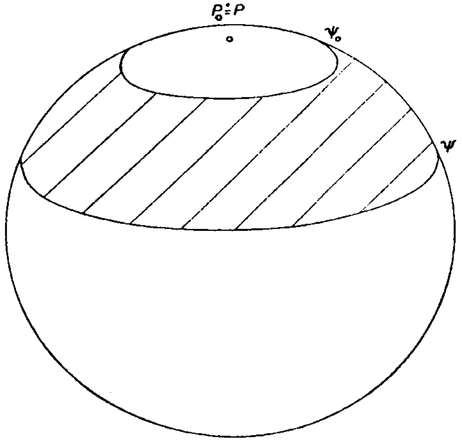

FIGURE 3-15
A spherical zone.

$$U_1 = R\pi k\kappa \int_{\psi_0}^{\psi} \frac{\sin\psi}{\sin(\psi/2)}\,d\psi,$$

$$A_1 = \frac{1}{2}\pi k\kappa \int_{\psi_0}^{\psi} \frac{\sin\psi}{\sin(\psi/2)}\,d\psi;$$

$$U_2 = -\frac{1}{2}\pi k\mu \int_{\psi_0}^{\psi} \frac{\sin\psi}{\sin(\psi/2)}\,d\psi,$$

$$A_2 = -\frac{1}{4R}\pi k\mu \int_{\psi_0}^{\psi} \left(\frac{1}{\sin^3(\psi/2)} + \frac{1}{\sin(\psi/2)}\right)\sin\psi\,d\psi.$$

The integrations are easily performed. For instance,

$$\int_{\psi_0}^{\psi} \frac{\sin\psi}{\sin(\psi/2)}\,d\psi = 2\int_{\psi_0}^{\psi} \cos\frac{\psi}{2}\,d\psi = 4\left(\sin\frac{\psi}{2} - \sin\frac{\psi_0}{2}\right).$$

Since all our formulas will be used as differences for two consecutive radii ψ_2 and ψ_1, the constant term containing $\sin(\psi_0/2)$ can be dropped, so that the result may be considered to be simply

$$4\sin\frac{\psi}{2}.$$

In this way we finally get, apart from a constant depending on ψ_0, the following expressions:

simple layer:

$$U_1 = 4R\pi k\kappa \sin\frac{\psi}{2}, \tag{3-75}$$

$$A_1 = 2\pi k\kappa \sin\frac{\psi}{2}; \tag{3-76}$$

double layer:

$$U_2 = -2\pi k\mu \sin\frac{\psi}{2}, \tag{3-77}$$

$$A_2 = \frac{1}{R}\pi k\mu \frac{\cos^2(\psi/2)}{\sin(\psi/2)}. \tag{3-78}$$

Bouguer reduction. Distant topography may be considered to form a *simple* surface layer on the sphere representing the geoid, and the height of P may be neglected so that P_0 is considered instead of P. The density of the surface layer is, as in (3-58), given by

$$\kappa = \rho h,$$

so that the effect of the attraction on gravity may be computed by (3-76),

$$A_T = 2\pi k\rho h \sin\frac{\psi}{2}. \tag{3-79}$$

For the indirect effect on the potential the formula (3-75) is used:

$$U_T = 4R\pi k\rho h \sin\frac{\psi}{2}. \tag{3-80}$$

Compensating reductions. Here *double* layers are to be used. The dipole moment of a column of unit cross section is, according to Sec. 1-4, given by

$$\mu = h\rho \cdot d, \tag{3-81}$$

since $h\rho$ is the mass of both the topographic column and its compensation; d is the distance between the two centers of mass S_T and S_C of Fig. 3-16 (it is the h of Sec. 1-4). The condition $d \longrightarrow 0$ is very nearly satisfied if d is small compared to the distance from the station.

Hence the combined effect of topography and compensation on gravity and potential are obtained by (3-78) and (3-77) as

$$A_T - A_C = \frac{hd}{R}\pi k\rho \frac{\cos^2(\psi/2)}{\sin(\psi/2)}, \tag{3-82}$$

$$U_T - U_C = -hd \cdot 2\pi k\rho \sin\frac{\psi}{2}. \tag{3-83}$$

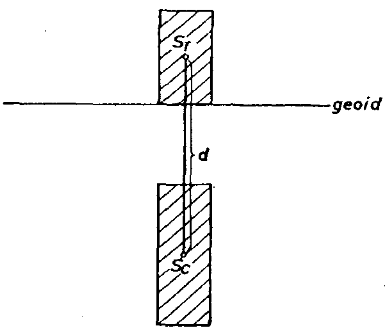

FIGURE 3-16

Topography and compensation as a dipole.

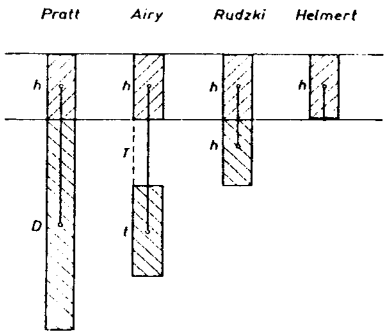

FIGURE 3-17

Topography and compensation for different gravity reductions.

The distance d for different gravity reductions is shown by Fig. 3-17; we have:

Pratt-Hayford:

$$d = \frac{h + D}{2},$$

Airy-Heiskanen:

$$d = T + \frac{h + t}{2},$$

Rudzki (inversion):

$$d = h, \tag{3-84}$$

Helmert (condensation):

$$d = \frac{h}{2}.$$

These hold for the continental case.

For the oceanic case there is no inversion or condensation, but only isostatic compensation. Instead of (3-81) we have

$$\mu = -h'(\rho - \rho_w)\cdot d', \tag{3-85}$$

where ρ_w and h' are density and depth of the ocean, and d' is found in the same way as before to be:

Pratt-Hayford:

$$d' = \frac{D}{2},$$

Airy-Heiskanen:

$$d' = T - \frac{h' + t'}{2}, \tag{3-86}$$

Rudzki and Helmert:

$$d' = 0.$$

The minus sign in (3-85) indicates that mass is shifted in the opposite direction as before.

For isostatic compensation according to Pratt-Hayford (3-82) thus becomes:

continents:

$$A_T - A_C = \frac{h(h + D)}{2R} \pi k \rho \frac{\cos^2 (\psi/2)}{\sin (\psi/2)},$$

oceans:

$$-A_C = -\frac{h' D}{2R} \pi k(\rho - \rho_w) \frac{\cos^2 (\psi/2)}{\sin (\psi/2)}. \tag{3-87}$$

These two formulas were derived by Helmert.

All these formulas are to be used in connection with a system by which the earth is divided into spherical compartments like Hayford's, so that actually

$$\begin{aligned} A &= \sum \Delta A, \\ U &= \sum \Delta U, \end{aligned} \tag{3-88}$$

where the density and elevation can be considered constant in each compartment.

3-9. Practical Determination of the Geoid

Reduction method to be used. In principle, all gravity reductions are equivalent and must lead to the same geoid if they are properly applied, the indirect effect being considered. There are, however, certain requirements that severely restrict the number of practically useful reductions. The main requirements are:

1. The reduction must yield gravity anomalies that are small and smooth, so that they can be easily interpolated and, where necessary, extrapolated. In other words, a single anomaly should be as representative as possible of a whole neighborhood.
2. The reduction must correspond to a geophysically meaningful model, so that the resulting anomalies are also useful for geophysical and geological interpretations.
3. The indirect effect should not be unduly great.

The Bouguer anomalies have good interpolatory properties—they are large but smooth—and are geophysically significant, but the Bouguer reduction must be excluded for the present purpose in view of its excessively large indirect effect (Sec. 3-6).

The Rudzki reduction has no indirect effect on the geoid, but changes the potential outside the earth, which is currently as important as the geoid. The Rudzki anomalies have no geophysical meaning.

The condensation reduction is very simple to compute, since it approximately yields free-air anomalies, and has a negligible indirect effect. It has some geophysical significance, corresponding to an extreme case of isostatic compensation. Free-air anomalies are small but extremely dependent on topography, so that their interpolation is very inaccurate.

Isostatic anomalies fulfil all three requirements. The underlying models correspond best to geological reality. The isostatic anomalies are small, smooth, and independent of topography, so that they are ideally suited for interpolation and extrapolation, and are very representative. The indirect effect is moderate.

Hence the free-air anomalies and the isostatic anomalies must be considered best suited for the present purpose. The main advantage of free-air anomalies is that they are easy to compute; their main disadvantage is that they are difficult to interpolate. With the isostatic reduction, just the opposite holds.

Because of the present possibilities of automatic computation the work involved in isostatic reduction no longer carries great weight. On the other hand, gravity data are still scarce and should be processed in such a way that the maximum amount of information is extracted from them, and that they are made as representative as possible. This speaks strongly in favor of employing isostatic reduction at the present time.

It may be mentioned that isostatic reduction can also be used in connection with the direct gravimetric determination of the physical surface of the earth, a topic which will be discussed in Chapter 8; see Sec. 8-11.

Gravity data. Prerequisites of the gravimetric methods are:

1. Theoretically, gravity anomalies must be given at every point of the earth's surface; practically, a dense gravity net around the computation points and a reasonably uniform distribution of gravity measurements outside are sufficient.
2. All gravity anomalies must be converted to the same system.

Absolute gravity measurements by means of pendulums are very laborious, and it is difficult to achieve the required accuracy of ± 1 mgal. Therefore, relative gravity measurements are preferred, which can be made by pendulums to an accuracy of ± 1 mgal and better, and by gravimeters to around ± 0.1 mgal.

These relative measurements must be tied together in such a way that they refer to a uniform world gravimetric system. One or several stations in each nation form a worldwide gravity base station network (Uotila, 1964a). The present reference datum is the so-called *Potsdam system*, which is based on absolute gravity measurements performed around 1900 at the Geodetic Institute in Potsdam, Germany.

The Potsdam system needs a constant correction which is currently estimated to be about -13 mgals. Several absolute determinations of gravity are in progress. They employ different techniques such as the use of pendulums and the observation of freely falling bodies.

The gravity data are collected and processed at such gravimetric data centers as the Isostatic Institute in Helsinki, The Ohio State University, and the Bureau Gravimetrique Internationale in Paris.

For automatic processing the data are stored as mean values of compartments of standard size, such as $5' \times 5'$, $10' \times 10'$, $1° \times 1°$, $2° \times 2°$, and $5° \times 5°$.

The map of Fig. 3-18 shows the gravity data available in 1959. The distribu-

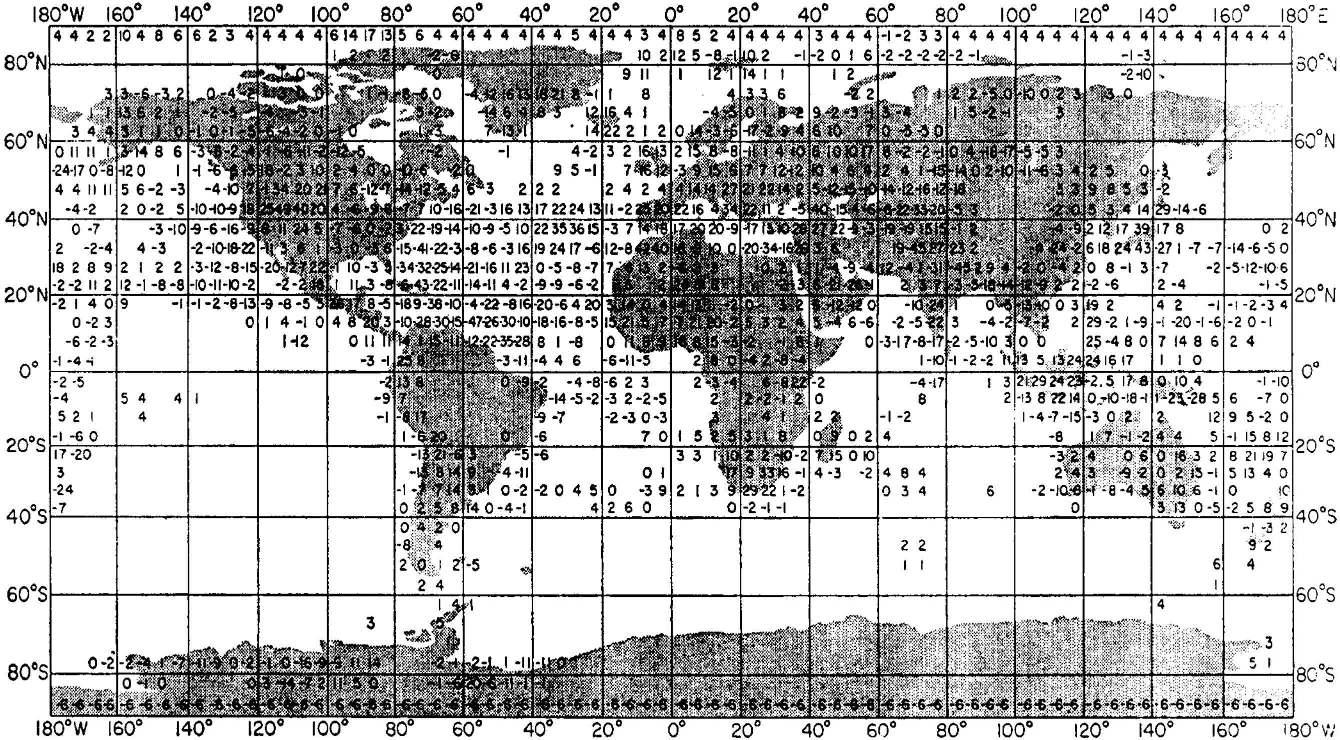

FIGURE 3-18

Mean free-air anomalies of 5° × 5° blocks, unit 1 mgal. Computed at Ohio State University on the basis of gravity data available as of December 31, 1959.

tion is far from being satisfactory. The large unsurveyed areas shown in the oceans can be expected to be filled in with the results of seaborne and airborne gravity measurements in the near future.

In the meantime we must try to fill the gaps with values extrapolated by statistical techniques (Chapter 7) or by means of a geophysical model or with values obtained by using a combination of both methods. Uotila (1964b) computed free-air gravity anomalies representing the effect of topography and its isostatic compensation only, so that they correspond to zero isostatic anomaly, using a spherical-harmonic expansion up to the 37th degree and obtaining 5° × 5° mean values. Part of his results are shown in Fig. 3-19. Whereas Uotila did not use any actual gravity data, Kivioja (1964) tried a combination of meas-

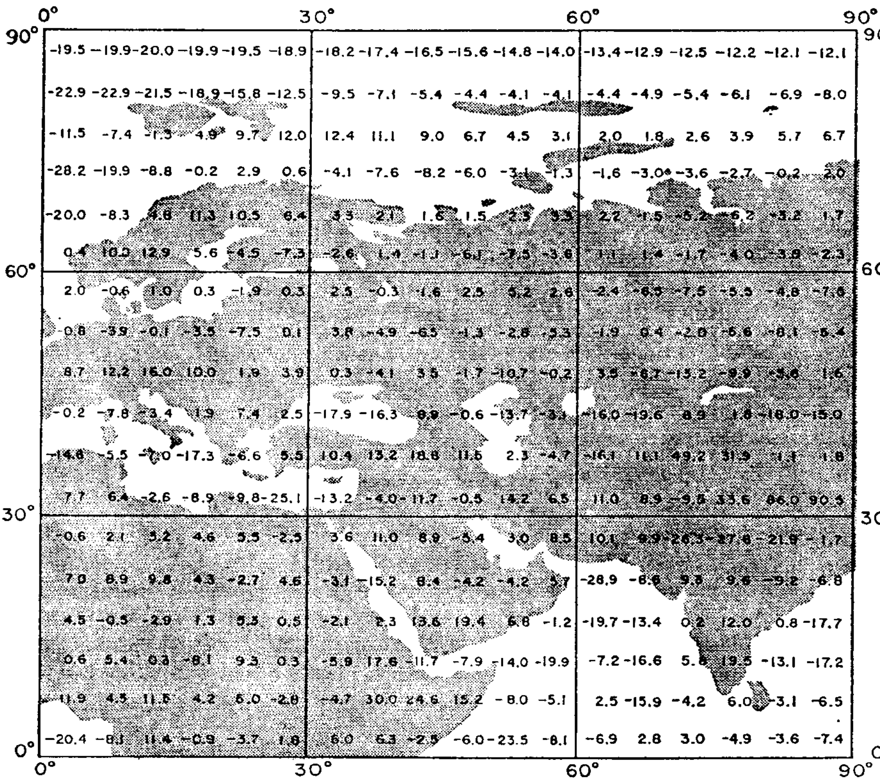

FIGURE 3-19

Mean free air anomalies of 5° × 5° blocks computed by Uotila (1964b) for a mathematical model of the earth in the sector between the longitudes 0° and 90°E from the North Pole to the Equator.

ured gravity data and geophysical extrapolation to the unsurveyed areas, again using an isostatic model.

Geoidal computations. The principles of the computation using Stokes' formula were outlined in Sec. 2-24.

The first practical computation of the geoid on a worldwide scale was carried out by Hirvonen (1934). He calculated the geoidal undulations for 62 points distributed in an east-west band encircling the surface of the earth. He estimated mean free-air anomalies of $5^\circ \times 5^\circ$ blocks in which gravity data were available; in unsurveyed areas he used free-air anomalies corresponding to zero isostatic anomaly.

Tanni (1948, 1949) computed geoidal heights, using Stokes' formula as Hirvonen did, but on the basis of a much larger quantity of gravity data. He employed isostatic reduction by means of both the Pratt-Hayford system with $D = 113.7$ km and the Airy-Heiskanen system with $T = 60$ km. Tanni computed global undulations, using $5^\circ \times 5^\circ$ blocks, and a more detailed geoid for Europe using $1^\circ \times 1^\circ$ blocks.

The most recent detailed gravimetric geoid is the Columbus geoid (Heiskanen, 1957). Five times more gravity data than Tanni had was available then. Free-air anomalies were employed, and the numerical integration in Stokes' formula was performed using an electronic computer. The computational details are described in (Uotila, 1960). Figure 3-20 shows the European geoid.

The large-scale features of the geoid can also be obtained by means of a spherical-harmonic expansion of lower degree, say up to degree four or eight, using methods such as those described in Sec. 2-20. We mention Jeffreys (1943), Zhongolovich (1952), Kaula (1961), Uotila (1962), and Kaula (1966). Figure 3-21 shows Uotila's (1962) geoid, which corresponds to a spherical-harmonic expansion of the fourth degree. The results of the various authors differ with respect to the gravity data available and the methods employed for handling the nonuniform distribution of data.

For more information see, for instance, Heiskanen (1965) and Kaula (1963).

Deflections of the vertical. Vening Meinesz' formula (2-210) for computing the deflections of the vertical is much more sensitive to local gravity anomalies around the computation point than Stokes' formula for the geoidal heights. Therefore, a dense gravity net is needed around the computation point. The effect of distant zones is somewhat less than in Stokes' formula but still considerable (see Sec. 7-4).

Highest accuracy is required, since $\pm 0.3''$ corresponds to about ± 10 meters in position. This is much more difficult to achieve than the corresponding accuracy of ± 10 meters in geoidal height.

For details of the numerical integration the reader is again referred to Sec. 2-24. The effect of the innermost zone involves a careful evaluation of the horizontal gradient of gravity. The radius of this innermost zone varies be-

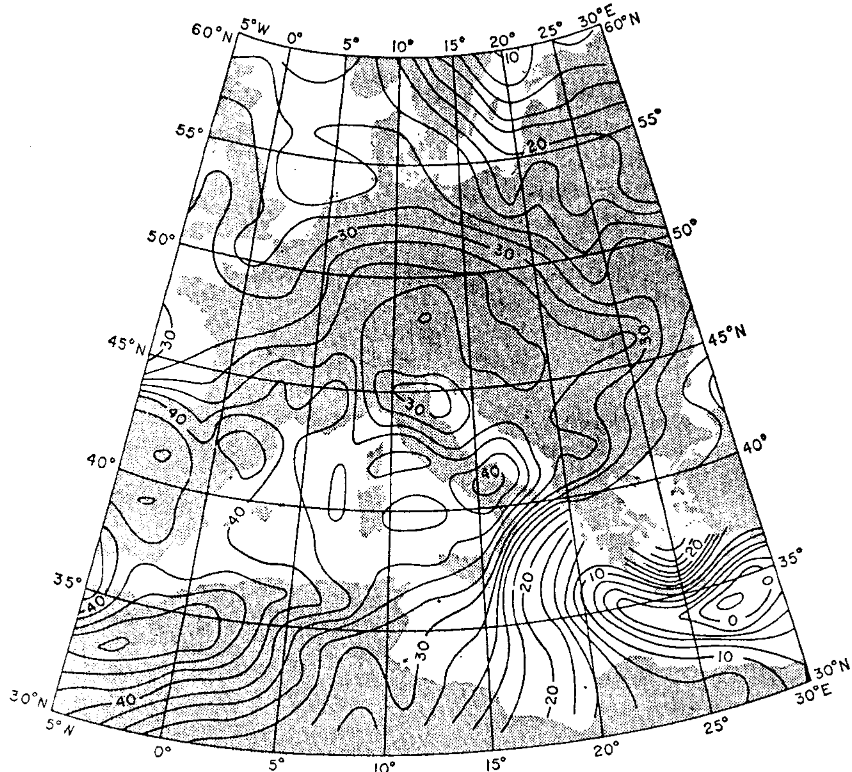

FIGURE 3-20

The Columbus geoid for Europe, referred to the international ellipsoid ($f = 1/297$). Contour interval 2 meters.

tween 0.1 and 10 km according to different authors and also depending on the gravity data available and the accuracy desired. See also Heiskanen and Vening Meinesz (1958, pp. 257–277).

If isostatic anomalies are used, then the indirect effect, which is identical with the topographic-isostatic deflection corresponding to the isostatic model used, must be precisely taken into account (Sec. 3-6). If free-air anomalies are used, then deflections of the vertical at the earth's surface, rather than on the geoid, may be computed utilizing the refinements described in Sec. 8-9.

World geodetic system. Since the gravimetric determination of geoidal heights yields absolute values for a reference ellipsoid coinciding with the center of mass of the earth, it plays a fundamental role in a worldwide geodetic system (Heiskanen, 1951; Heiskanen and Vening Meinesz, 1958, Chapter 9). This requires a combination with astrogeodetic data (see our Chapter 5). During

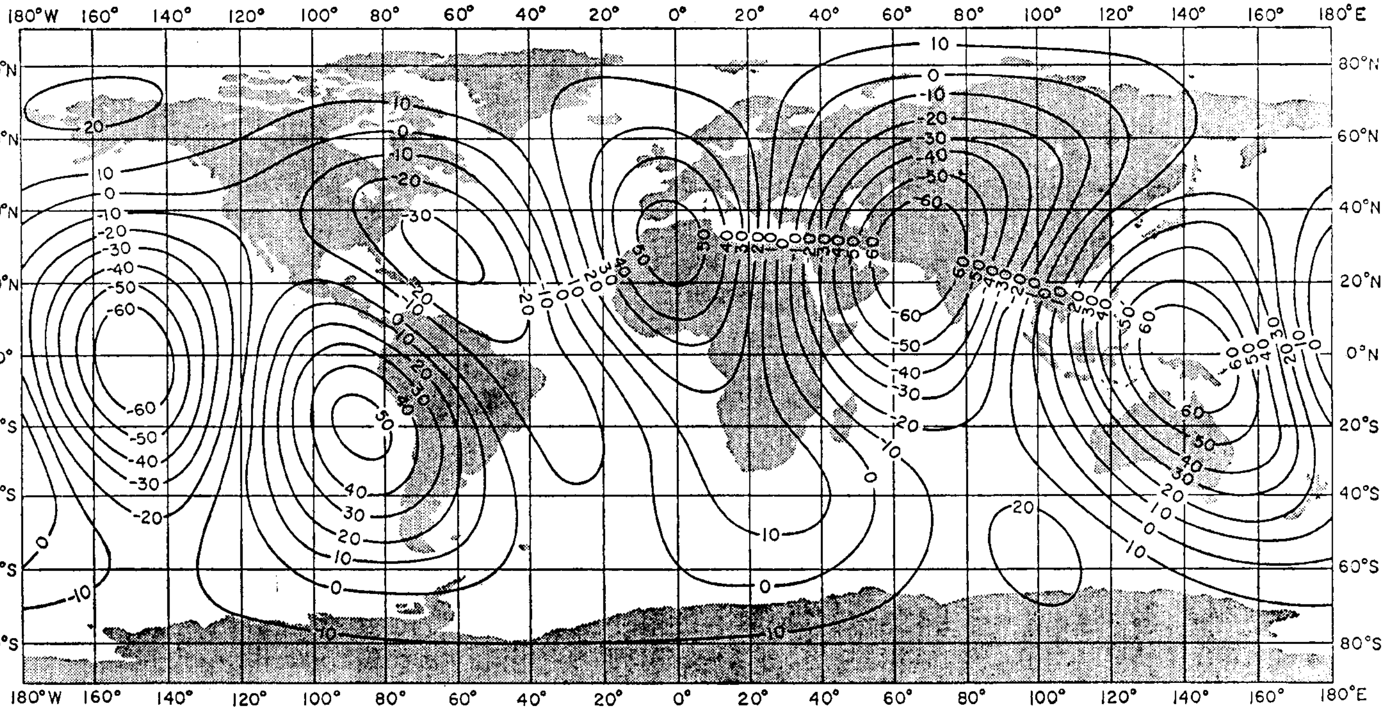

FIGURE 3-21

The generalized geoid of Uotila (1962) computed from a spherical-harmonic expansion of the fourth degree. Unit 1 meter; flattening of the reference ellipsoid $f = 1/298.24$.

the past few years artificial satellites have also been used to gather data for a world geodetic system (see our Chapter 9).

References

Baeschlin, C. F. (1948). Lehrbuch der Geodäsie. Zürich, Orell-Füssli.

Cassinis, G., P. Dore, and S. Ballarin (1937). Fundamental tables for reducing gravity observed values. R. Comm. Geod. Ital. Publ. No. 13.

Darling, F. W. (1949). Fundamental tables for the deflection of the vertical. U. S. Coast and Geodetic Survey, Spec. Publ. No. 243.

Hayford, J. F., and W. Bowie (1912). The effect of topography and isostatic compensation upon the intensity of gravity. U. S. Coast and Geodetic Survey, Spec. Publ. No. 10.

Heiskanen, W. (1924). Untersuchungen über Schwerkraft und Isostasie. Helsinki, Publ. Finn. Geod. Inst., No. 4. Also: Investigations regarding gravity and isostasy. Geol. Survey of India, 1928.

Heiskanen, W. (1931). Tables isostasiques pour le réduction dans l'hypothèse de Airy des intensités de la pesanteur observées. Bull. Géod., No. 30, pp. 87–153.

Heiskanen, W. (1938). New isostatic tables for the reduction of the gravity values calculated on the basis of Airy's hypothesis. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 2.

Heiskanen, W. (1951). On the world geodetic system. *Ibid.*, No. 26.

Heiskanen, W. A. (1957). The Columbus geoid. Trans. Am. Geophys. Union, v. 38, pp. 841–848.

Heiskanen, W. A. (1965). Present problems of physical geodesy. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 49.

Heiskanen, W., and E. Niskanen (1941). World maps for the indirect effect of the undulations of the geoid on gravity anomalies. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 7.

Heiskanen, W., E. Niskanen, and P. Kärki (1959). Topographic-isostatic reduction maps for Europe and North Atlantic in the Hayford zones 18–1 for the Airy-Heiskanen system, $T = 30$ km and 20 km. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 31.

Heiskanen, W., and U. Nuotio (1938). Topographic-isostatic world maps of the effect of the Hayford zones 10, 9, 8 and 7 to 1. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 3.

Heiskanen, W. A., and F. A. Vening Meinesz (1958). The earth and its gravity field. New York, McGraw-Hill.

Helmert, F. R. (1884). Die mathematischen und physikalischen Theorien der höheren Geodäsie, vol. 2. Leipzig, B. G. Teubner (reprinted 1962).

Hirvonen, R. A. (1934). The continental undulations of the geoid. Helsinki, Publ. Finn. Geod. Inst., No. 19.

Jeffreys, H. (1943). The determination of the earth's gravitational field (second paper). M. N. R. A. S. Geo. Supp. 5, pp. 55–66.

Jung, K. (1961). Schwerkraftverfahren in der Angewandten Geophysik. Leipzig, Akademische Verlagsgesellschaft Geest & Portig.

Kärki, P., L. Kivioja, and W. A. Heiskanen (1961). Topographic-isostatic reduction maps for the world for the Hayford zones 18–1, Airy-Heiskanen system, $T = 30$ km, Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 35.

Kaula, W. M. (1961). A geoid and world geodetic system based on a combination of gravimetric, astrogeodetic, and satellite data. J. Geophys. Research, v. 66, pp. 1799–1812.

Kaula, W. M. (1963). Determination of the earth's gravitational field. Reviews of Geophysics, v. 1, pp. 507–551.

Kaula, W. M. (1966). Global harmonic and statistical analysis of gravimetry. Proceedings of the Symposium "Extension of Gravity Anomalies to Unsurveyed Areas" (H. Orlin, ed.), Am. Geophys. Union, Geophys. Monograph No. 9, pp. 58–67.

Kellogg, O. D. (1929). Foundations of potential theory. Berlin, J. Springer (New York, Dover Publications, 1953).

Kivioja, L. (1964). Effect of topographic masses and their isostatic compensation on the mean free-air gravity anomalies of 5° × 5° surface elements. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 46.

Kukkamäki, T. J. (1955). Gravimetric reductions with electronic computors. *Ibid.*, No. 30.

Lambert, W. D. (1930). The reduction of observed values of gravity to sea level. Bull. Géod., No. 26, pp. 107–181.

Lambert, W. D., and F. W. Darling (1936). Tables for determining the form of the geoid and its indirect effect on gravity. U. S. Coast and Geodetic Survey, Spec. Publ. No. 199.

Lambert, W. D., and F. W. Darling (1938). Formulas and tables for the deflection of the vertical. Bull. Géod., No. 57, pp. 29–71.

Niskanen, E., and L. Kivioja (1951). Topographic-isostatic world maps of the effect of the Hayford zones 10 to 1 for the Airy-Heiskanen and Pratt-Hayford systems. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 27.

Rice, D. A. (1952). Deflections of the vertical from gravity anomalies. Bull. Géod., No. 25, pp. 285–312.

Tanni, L. (1948). On the continental undulations of the geoid as determined from the present gravity material. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 18.

Tanni, L. (1949). The regional rise of the geoid in Central Europe. *Ibid.*, No. 22.

Uotila, U. A. (1960). Investigations on the gravity field and shape of the earth. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 33.

Uotila, U. A. (1962). Harmonic analysis of world-wide gravity material. *Ibid.*, No. 39.

Uotila, U. A. (1964a). Adjustment of a world-wide gravity base station network. *Ibid.*, No. 42.

Uotila, U. A. (1964b). Gravity anomalies for a mathematical model of the earth. *Ibid.*, No. 43.

Vening Meinesz, F. A. (1939). Tables fondamentales pour la réduction isostatique régionale. Bull. Géod., No. 63, pp. 711–776.

Zhongolovich, I. D. (1952). The external gravity field of the earth and the fundamental constants connected with it (in Russian). Leningrad, Acad. Sci. Publ. Inst. Teor. Astron.

4

Heights Above Sea Level

4-1. Spirit Leveling

The principle of spirit leveling is well known. To measure the difference in height, δH, between two points A and B, vertical rods are set up at each of these two points and a level (leveling instrument) somewhere between them (Fig. 4-1). Since the line $\bar{A}\bar{B}$ is horizontal, the difference in the rod readings $l_1 = A\bar{A}$ and $l_2 = B\bar{B}$ is the height difference:

$$\delta H_{AB} = l_1 - l_2.$$

For details of the technique of measurement the reader is referred to Bomford (1962).

If we measure a circuit, that is, a closed leveling line, then the algebraic sum of all measured differences in height will not in general be rigorously zero, as one would expect it to be, even if we had been able to observe with perfect precision. This so-called misclosure indicates that leveling is more complicated than it appears at first sight.

Let us look into the matter more closely. Figure 4-2 shows the relevant geometrical principles. Let the points A and B be so far apart that the pro-

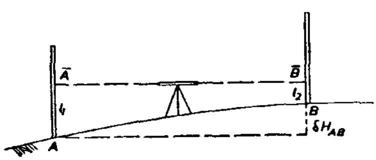

FIGURE 4-1

Spirit leveling.

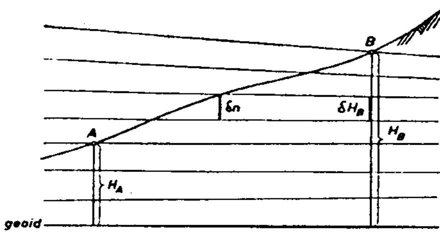

FIGURE 4-2
Leveling and orthometric height.

cedure of Fig. 4-1 must be applied repeatedly. Then the sum of the leveled height differences between A and B will not be equal to the difference in the orthometric heights H_A and H_B. The reason is that the leveling increment δn, as we shall henceforth denote it, is different from the corresponding increment δH_B of H_B (Fig. 4-2), owing to the nonparallelism of the level surfaces. Denoting the corresponding increment of the potential W by δW, we have by (2-13)

$$-\delta W = g\,\delta n = g'\,\delta H_B, \tag{4-1}$$

where g is the gravity at the leveling station and g' is the gravity on the plumb line of B at δH_B. Hence,

$$\delta H_B = \frac{g}{g'}\,\delta n \neq \delta n. \tag{4-2}$$

There is thus no direct geometrical relation between the result of leveling and the orthometric height, since (4-2) expresses a physical relation. What, then, if not height, is directly obtained by leveling? If gravity g is also measured, then

$$\delta W = -g\,\delta n$$

is determined, so that we obtain

$$W_B - W_A = -\sum_A^B g\,\delta n. \tag{4-3}$$

Thus, leveling combined with gravity measurements furnishes *potential differences*, that is, physical quantities.

It is somewhat more rigorous theoretically to replace the sum in (4-3) by an integral, obtaining

$$W_B - W_A = -\int_A^B g\,dn. \tag{4-4}$$

Note that this integral is independent of the path of integration; that is, different leveling lines connecting the points A and B (Fig. 4-3) should give the

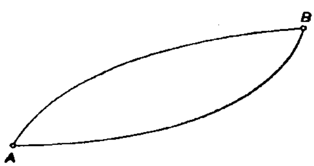

FIGURE 4-3
Two different leveling lines connecting A and B; taken together, they form a circuit.

same result. This is evident because W is a function of position only; thus to every point there corresponds a unique value W. If the leveling line returns to A, then the total integral must be zero:

$$\oint g\,dn = -W_A + W_A = 0. \tag{4-5}$$

The symbol $\oint$ denotes an integral over a circuit.

On the other hand, the measured height difference, that is, the sum of the leveling increments

$$\Delta n_{AB} = \sum_A^B \delta n = \int_A^B dn, \tag{4-6}$$

depends on the path of integration and is thus not in general zero for a circuit:

$$\oint dn = \text{misclosure} \neq 0. \tag{4-7}$$

In mathematical terms, dn is not a perfect differential (the differential of a function of position), whereas $dW = -g\,dn$ is perfect, so that dn becomes a perfect differential when it is multiplied by the integrating factor $(-g)$.

Potential differences are thus the result of leveling combined with gravity measurements. They are basic to the whole theory of heights; even orthometric heights must be considered as quantities derived from potential differences.

Leveling without gravity measurements, although applied in practice, is meaningless from a rigorous point of view, for the use of leveled heights (4-6) as such leads to contradictions (misclosures); it will not be considered here.

4-2. Geopotential Numbers and Dynamic Heights

Let O be a point at sea level, that is, on the geoid; usually a suitable point on the seashore is taken. Let A be another point, connected to O by a leveling line. Then, by formula (4-3), the potential difference between A and O can be determined. The integral

$$\int_O^A g\,dn = W_0 - W_A = C, \tag{4-8}$$

which is the difference between the potential at the geoid and the potential at the point A, has been introduced as the *geopotential number* of A in Sec. 2-4.

As a potential difference, the geopotential number C is independent of the particular leveling line used for relating the point to sea level. It is the same for all points of a level surface; it can thus be considered as a natural measure of height, even if it does not have the dimension of a length.

The geopotential number C is measured in geopotential units (g.p.u.), where

$$1 \text{ g.p.u.} = 1 \text{ kgal meter} = 1000 \text{ gal meter}.$$

Since $g \doteq 0.98$ kgal,

$$C \doteq gH \doteq 0.98H,$$

so that the geopotential numbers in g.p.u. are almost equal to the height above sea level in meters.

The geopotential numbers were adopted at a meeting of a subcommission of the International Association of Geodesy at Florence in 1955. Formerly the *dynamic heights* were used, defined by

$$H^{\text{dyn}} = \frac{C}{\gamma_0}, \tag{4-9}$$

where γ_0 is normal gravity for an arbitrary standard latitude, usually 45°:

$$\gamma_{45^\circ} = 980.6294 \text{ gals}$$

for the international ellipsoid.

Obviously the dynamic height differs from the geopotential number only in the scale or the unit: The division by the constant γ_0 in (4-9) merely converts a geopotential number into a length. However, the dynamic height has no geometrical meaning whatever, so that the division by an arbitrary γ_0 merely obscures the true physical meaning of a potential difference. Hence, the geopotential numbers are in general preferable to the dynamic heights.

Dynamic correction. It is sometimes convenient to convert the measured height difference Δn_{AB} (4-6) into a difference of dynamic height by adding a small correction.

Equation (4-9) gives

$$\Delta H_{AB}^{\text{dyn}} = H_B^{\text{dyn}} - H_A^{\text{dyn}} = \frac{1}{\gamma_0}(C_B - C_A) = \frac{1}{\gamma_0}\int_A^B g\,dn$$

$$= \frac{1}{\gamma_0}\int_A^B (g - \gamma_0 + \gamma_0)\,dn = \int_A^B dn + \int_A^B \frac{g - \gamma_0}{\gamma_0}\,dn,$$

so that

$$\Delta H_{AB}^{\text{dyn}} = \Delta n_{AB} + DC_{AB}, \tag{4-10}$$

where

$$DC_{AB} = \int_A^B \frac{g - \gamma_0}{\gamma_0}\,dn = \sum_A^B \frac{g - \gamma_0}{\gamma_0}\,\delta n \tag{4-11}$$

is the *dynamic correction.*

As a matter of fact, the dynamic correction may also be used for computing differences of geopotential numbers. We at once obtain

$$C_B - C_A = \gamma_0\,\Delta n_{AB} + \gamma_0 DC_{AB}. \tag{4-10'}$$

4-3. The Gravity Reduction of Poincaré and Prey

To convert the results of leveling into orthometric heights, we need from (4-2) the gravity g' inside the earth. Since g' cannot be measured, it must be computed from the surface gravity. This is done by reducing the measured values of gravity according to the method of Poincaré and Prey.

We denote the point at which g' is to be computed by Q, so that $g' = g_Q$. Let P be the corresponding surface point, so that P and Q are situated on the same plumb line (Fig. 4-4). Gravity at P, denoted by g_P, is measured.

The direct way of computing g_Q would be to use the formula

$$g_Q = g_P - \int_Q^P \frac{\partial g}{\partial h}\, dH, \tag{4-12}$$

provided that the actual gravity gradient $\partial g/\partial h$ inside the earth were known. It can be obtained by Bruns' formula (2-20),

$$\frac{\partial g}{\partial h} = -2gJ + 4\pi k\rho - 2\omega^2, \tag{4-13}$$

if the mean curvature J of the geopotential surfaces and the density ρ are known between P and Q.

The normal free-air gradient is given by (2-79):

$$\frac{\partial \gamma}{\partial h} = -2\gamma J_0 - 2\omega^2, \tag{4-14}$$

where J_0 is the mean curvature of the spheropotential surfaces. If the approximation

$$gJ \doteq \gamma J_0$$

is sufficient, then we get from (4-13) and (4-14)

$$\frac{\partial g}{\partial h} = \frac{\partial \gamma}{\partial h} + 4\pi k\rho. \tag{4-15}$$

Numerically, neglecting the variation of $\partial\gamma/\partial h$ with latitude, we find for the density $\rho = 2.67$ g/cm^3 and $k = 66.7 \times 10^{-9}$ c.g.s. units:

$$\frac{\partial g}{\partial h} = -0.3086 + 0.2238 = -0.0848 \text{ gal/km},$$

so that (4-12) becomes

$$g_Q = g_P + 0.0848(H_P - H_Q) \tag{4-16}$$

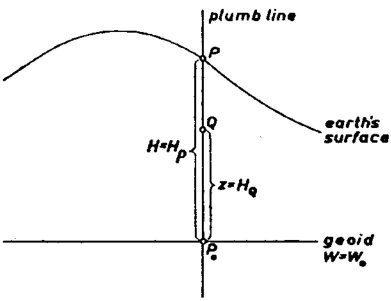

FIGURE 4-4
The Prey reduction.

with g in gals and H in kilometers. This simple formula, although being rather crude, is often applied in practice.

The accurate way to compute g_Q would be to use (4-12) and (4-13) with the actual mean curvature J of the geopotential surfaces, but this would require a knowledge of the detailed shape of these surfaces far beyond what is attainable today.

Another way of computing g_Q, which is more practicable at present, is the following. It is similar to the usual reduction of gravity to sea level (see Chapter 3) and consists of three steps:

1. Remove all masses above the geopotential surface $W = W_Q$, which contains Q, and subtract their attraction from g at P.
2. Since the gravity station P is now "in free air," apply the free-air reduction, thus moving the gravity station from P to Q.
3. Restore the removed masses to their former position, and add algebraically their attraction to g at Q.

The purpose of this somewhat complicated procedure is that in step 2 the *free-air* gradient can be used. If we here replace the actual free-air gradient by the normal gradient $\partial\gamma/\partial h$, the error will presumably be smaller than in using (4-15).[1]

The effect of the masses above Q (steps 1 and 3) may be computed by the methods of Chapter 3—for example, by means of some kind of template. If the terrain correction is neglected and only the infinite Bouguer plate between P and Q of the normal density $\rho = 2.67$ g/cm³ is taken into account, then we obtain simply, with the steps numbered as above:

gravity measured at P	g_P	
1. take away Bouguer plate		$-0.1119\,(H_P - H_Q)$
2. free-air reduction from P to Q		$+0.3086\,(H_P - H_Q)$
3. restore Bouguer plate		$-0.1119\,(H_P - H_Q)$
together: gravity at Q	$g_Q = g_P$	$+\ 0.0848\,(H_P - H_Q)$

This is the same as (4-16), which is thus confirmed independently. We see now that the use of (4-15) or (4-16) amounts to replacing the terrain with a Bouguer plate.

Finally we note that the reduction of Poincaré and Prey, abbreviated as *Prey reduction*, yields the actual gravity which would be measured inside the earth if this were possible. Its purpose is thus completely different from the purpose of the other gravity reductions, to give boundary values at the geoid; see Sec. 3-7.

[1] The free-air gradient can also be accurately computed using (2-217); the gravity anomalies Δg to be used in this formula are the gravity anomalies obtained after performing step 2, that is, Bouguer anomalies referred to the level of point Q.

4-4. Orthometric Heights

We denote the intersection of the geoid and the plumb line through point P by P_0 (Fig. 4-4). Let C be the geopotential number of P, that is,

$$C = W_0 - W,$$

and H its orthometric height, that is, the length of the plumb-line segment between P_0 and P. Perform the integration in (4-8) along the plumb line P_0P. This is permitted because the result is independent of the path. We then get

$$C = \int_0^H g\, dH. \tag{4-17}$$

This equation contains H in an implicit way. It is also possible to get H explicitly. From

$$dC = -dW = g\, dH, \quad dH = -\frac{dW}{g} = \frac{dC}{g}$$

we obtain

$$H = -\int_{W_0}^{W} \frac{dW}{g} = \int_0^C \frac{dC}{g}. \tag{4-18}$$

As before, the integration is extended over the plumb line.

The explicit formula (4-18), however, is of little practical use. It is better to transform (4-17) in a way that at first looks entirely trivial:

$$C = \int_0^H g\, dH = H \cdot \frac{1}{H} \int_0^H g\, dH,$$

so that

$$C = \bar{g} H, \tag{4-19}$$

where

$$\bar{g} = \frac{1}{H} \int_0^H g\, dH \tag{4-20}$$

is the mean value of the gravity along the plumb line between the geoid, point P_0, and the ground, point P. From (4-19) it follows that

$$H = \frac{C}{\bar{g}}, \tag{4-21}$$

which permits H to be computed if the mean gravity $\bar{g}$ is known. Since $\bar{g}$ does not strongly depend on H, equation (4-21) is a practically useful formula and not merely a tautology.

For evaluating (4-21) we need the mean gravity $\bar{g}$. Equation (4-20) may be written

$$\bar{g} = \frac{1}{H} \int_0^H g(z)\, dz, \tag{4-22}$$

where $g(z)$ is the actual gravity at the variable point Q which has the height z (Fig. 4-4).

The simplest approximation is to use the simplified Prey reduction of (4-16):

$$g(z) = g + 0.0848(H - z), \tag{4-23}$$

where g is the gravity measured at the ground point P. The integration (4-22) can now be performed immediately, giving

$$\bar{g} = \frac{1}{H}\int_0^H [g + 0.0848(H - z)]\, dz = g + \frac{1}{H} \cdot 0.0848 \left[Hz - \frac{z^2}{2}\right]_0^H,$$

or

$$\bar{g} = g + 0.0424H \qquad (g \text{ in gals, } H \text{ in km}). \tag{4-24}$$

The factor 0.0424 holds for the normal density $\rho = 2.67$ g/cm^3. The corresponding formula for arbitrary constant density is, by (4-15),

$$\bar{g} = g - \left(\frac{1}{2}\frac{\partial\gamma}{\partial h} + 2\pi k\rho\right) H. \tag{4-25}$$

If we use $\bar{g}$ according to (4-24) or (4-25) in the basic formula (4-21), we obtain the so-called Helmert heights (Helmert, 1890):

$$H = \frac{C}{g + 0.0424H} \tag{4-26}$$

with C in g.p.u., g in gals and H in km.

As we have seen in Sec. 4-3, this approximation replaces the terrain with an infinite Bouguer plate of constant density and of height H. This is often sufficient. Sometimes, in high mountains and for highest precision, it is necessary to apply to g a more rigorous Prey reduction, such as the three steps described in Sec. 4-3. A practical and very accurate method for this purpose has been given by Niethammer (1932). It takes the topography into account, assuming only that the free-air gradient is normal and the density is constant down to the geoid.

It is also sufficient to calculate $\bar{g}$ as the mean of gravity g measured at the surface point P and of gravity g_0 computed at the corresponding geoidal point P_0 by the Prey reduction:

$$\bar{g} = \frac{1}{2}(g + g_0). \tag{4-27}$$

This has been proposed by Mader (1954); it presupposes that gravity g varies linearly along the plumb line. This can usually be assumed with sufficient accuracy, even in extreme cases, as shown by Mader (1954) and by Ledersteger (1955).

Orthometric correction. The orthometric correction is added to the measured height difference, in order to convert it into a difference in orthometric height.

We let the leveling line connect two points A and B (Fig. 4-5). We first apply a simple trick:

$$\begin{aligned}\Delta H_{AB} &= H_B - H_A = H_B - H_A - H_B^{\text{dyn}} + H_A^{\text{dyn}} + (H_B^{\text{dyn}} - H_A^{\text{dyn}}) \\ &= \Delta H_{AB}^{\text{dyn}} + (H_B - H_B^{\text{dyn}}) - (H_A - H_A^{\text{dyn}}).\end{aligned} \tag{4-28}$$

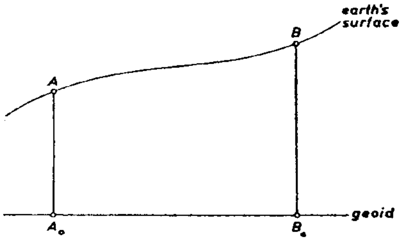

FIGURE 4-5
Orthometric and dynamic correction.

By (4-10) we have

$$\Delta H_{AB}^{\text{dyn}} = \Delta n_{AB} + DC_{AB}. \tag{4-29}$$

Consider now the differences between the orthometric and dynamic heights, $H_A - H_A^{\text{dyn}}$ and $H_B - H_B^{\text{dyn}}$. Imagine a fictitious leveling line leading from the foot A_0 at the geoid to the ground point A along the plumb line. Then, obviously, the measured height difference would be H_A itself: $\Delta n_{A_0A} = H_A$, so that

$$DC_{A_0A} = \Delta H_{A_0A}^{\text{dyn}} - \Delta n_{A_0A} = H_A^{\text{dyn}} - H_A$$

and

$$\begin{aligned} H_A - H_A^{\text{dyn}} &= -DC_{A_0A}, \\ H_B - H_B^{\text{dyn}} &= -DC_{B_0B}. \end{aligned} \tag{4-30}$$

Inserting (4-29) and (4-30) into (4-28) we finally have

$$\Delta H_{AB} = \Delta n_{AB} + DC_{AB} + DC_{A_0A} - DC_{B_0B}$$

or

$$\Delta H_{AB} = \Delta n_{AB} + OC_{AB}, \tag{4-31}$$

where

$$OC_{AB} = DC_{AB} + DC_{A_0A} - DC_{B_0B} \tag{4-32}$$

is the orthometric correction. This is a remarkable relation between the orthometric and dynamic corrections (Ledersteger, 1955).

From (4-11) we find

$$DC_{AB} = \int_A^B \frac{g - \gamma_0}{\gamma_0}\, dn = \sum_A^B \frac{g - \gamma_0}{\gamma_0}\, \delta n,$$

$$DC_{A_0A} = \int_{A_0}^A \frac{g - \gamma_0}{\gamma_0}\, dH = \frac{\bar{g}_A - \gamma_0}{\gamma_0} H_A,$$

$$DC_{B_0B} = \int_{B_0}^B \frac{g - \gamma_0}{\gamma_0}\, dH = \frac{\bar{g}_B - \gamma_0}{\gamma_0} H_B,$$

where $\bar{g}_A$, or $\bar{g}_B$, is the mean value of gravity along the plumb line of A, or B. Thus the orthometric correction (4-32) becomes

$$OC_{AB} = \sum_A^B \frac{g - \gamma_0}{\gamma_0}\, \delta n + \frac{\bar{g}_A - \gamma_0}{\gamma_0} H_A - \frac{\bar{g}_B - \gamma_0}{\gamma_0} H_B. \tag{4-33}$$

Here again we need the mean value of gravity along the plumb line, $\bar{g}_A$ and $\bar{g}_B$; γ_0 is an arbitrary constant—for example, normal gravity at 45° latitude.

Accuracy. Let us first evaluate the effect on H of an error in the mean gravity $\bar{g}$. From $H = C/\bar{g}$ we obtain by differentiation

$$\delta H = -\frac{C}{\bar{g}^2}\,\delta\bar{g} = -\frac{H}{\bar{g}}\,\delta\bar{g}.$$

Since $\bar{g}$ is about 1000 gals we have, neglecting the minus sign, the simple formula

$$\delta H_{\mathrm{mm}} \doteq \delta\bar{g}_{\mathrm{mgal}} H_{\mathrm{km}}, \tag{4-34}$$

the subscripts denoting the units; δH is the error in H, caused by an error $\delta\bar{g}$ in $\bar{g}$.

For $H = 1$ km,

$$\delta H_{\mathrm{mm}} \doteq \delta\bar{g}_{\mathrm{mgal}},$$

which shows that an error $\delta\bar{g}$ as great as 100 mgals falsifies an elevation of 1000 meters by only 10 cm.

Let us now estimate the effect of an error of the density ρ on $\bar{g}$. Differentiating (4-25) and omitting the minus sign we find

$$\delta\bar{g} = 2\pi k H \delta\rho. \tag{4-35}$$

If $\delta\rho = 0.1$ g/cm³ and $H = 1$ km, then

$$\delta\bar{g} = 4.2 \text{ mgals},$$

which causes an error of 4 mm in H. A density error of 0.6 g/cm³, which corresponds to the maximum variation of rock density occurring in practice thus falsifies $H = 1000$ meters by only 25 mm.

Mader (1954) has estimated the difference between the simple computation of mean gravity according to Helmert, equation (4-25), and more accurate methods that take the terrain correction into account. He found for Hochtor, in the Alps, $H = 2504$ meters:

Helmert	(4-25)	$\bar{g} = 980.263$	(Bouguer plate only),
Niethammer		286	(also terrain correction).
$\bar{g} = \frac{1}{2}(g + g_0)$	(4-27)	285	(also terrain correction).

Mean gravity $\bar{g}$ according to (4-27) differs from Niethammer's value by only 1 mgal, which shows the linearity of g along the plumb line even in an extreme case. This corresponds to a difference in H of 3 mm. The simple Helmert height differs by about 6 cm from these more elaborately computed heights.

The differences are thus very small even in this rather extreme case; we see that orthometric heights can be obtained with very high accuracy. This is of great importance for a discussion of the recent theory of Molodensky from a practical point of view. See Chapter 8, particularly the last section.

4-5. Normal Heights

Assume for the moment the gravity field of the earth to be normal, that is, $W = U$, $g = \gamma$, $T = 0$. On this assumption compute "orthometric heights"; they will be called *normal heights* and denoted by H^*. Thus equations (4-17) through (4-20) become

$$W_0 - W = C = \int_0^{H^*} \gamma \, dH^*, \tag{4-36}$$

$$H^* = \int_0^C \frac{dC}{\gamma}, \tag{4-37}$$

$$C = \bar{\gamma} H^*, \tag{4-38}$$

where

$$\bar{\gamma} = \frac{1}{H^*} \int_0^{H^*} \gamma \, dH^* \tag{4-39}$$

is the mean normal gravity along the plumb line.

As the normal potential U is a simple analytic function, these formulas can be evaluated very easily; but since the potential of the earth is evidently not normal, what does all this mean? Consider a point P on the physical surface of the earth. It has a certain potential W_P and also a certain normal potential U_P, but in general $W_P \neq U_P$. However, there is a certain point Q on the plumb line of P, such that $U_Q = W_P$; that is, the normal potential U at Q is equal to the actual potential W at P. The normal height H^* of P is nothing but the geometric height of Q above the ellipsoid, just as the orthometric height of P is the height of P above the geoid.

For more details the reader is referred to Sec. 8-3; Fig. 8-2 illustrates the geometric relations.

We shall now give some practical formulas for the computation of normal heights from geopotential numbers. If we write (4-39) in the form

$$\bar{\gamma} = \frac{1}{H^*} \int_0^{H^*} \gamma(z) \, dz \tag{4-40}$$

corresponding to (4-22), then we can express $\gamma(z)$ by (2-123) as

$$\gamma(z) = \gamma \left[1 - \frac{2}{a}(1 + f + m - 2f \sin^2 \phi) z + \frac{3}{a^2} z^2 \right], \tag{4-41}$$

where γ is the gravity at the ellipsoid, depending on the latitude ϕ but not on z. Thus straightforward integration with respect to z yields

$$\bar{\gamma} = \frac{1}{H^*} \gamma \left[z - \frac{2}{a}(1 + f + m - 2f \sin^2 \phi) \frac{z^2}{2} + \frac{3}{a^2} \frac{z^3}{3} \right]_0^{H^*}$$

$$\frac{1}{H^*} \gamma \left[H^* - \frac{1}{a}(1 + f + m - 2f \sin^2 \phi) H^{*2} + \frac{1}{a^2} H^{*3} \right]$$

or

$$\bar{\gamma} = \gamma \left[1 - (1 + f + m - 2f \sin^2 \phi) \frac{H^*}{a} + \frac{H^{*2}}{a^2} \right]. \tag{4-42}$$

This formula may be used for computing H^* by the formula

$$H^* = \frac{C}{\bar{\gamma}}. \tag{4-43}$$

The mean theoretical gravity itself depends on H^*, by (4-42), but not strongly, so that an iterative solution is very simple.

It is also possible to give a direct expression of H^* in terms of the geopotential number C by inserting (4-42) into (4-43) and expanding into a series of powers of H^*:

$$H^* = \frac{C}{\gamma}\left[1 + \frac{1}{a}(1 + f + m - 2f\sin^2\phi)H^* + 0 \cdot H^{*2} + \cdots\right].$$

Solving this equation for H^* and expanding H^* in powers of C/γ, we obtain

$$H^* = \frac{C}{\gamma}\left[1 + (1 + f + m - 2f\sin^2\phi)\frac{C}{a\gamma} + \left(\frac{C}{a\gamma}\right)^2\right], \tag{4-44}$$

where γ is normal gravity at the ellipsoid, for the same latitude ϕ. The accuracy of this formula will be sufficient for almost all practical purposes; still more accurate expressions are given in Hirvonen (1960).

Corresponding to the dynamic and orthometric corrections there is a *normal correction* NC of the measured height differences. Equation (4-33) immediately yields, on replacing $\bar{g}$ by $\bar{\gamma}$ and H by H^*:

$$NC_{AB} = \sum_A^B \frac{g - \gamma_0}{\gamma_0}\,\delta n + \frac{\bar{\gamma}_A - \gamma_0}{\gamma_0} H_A^* - \frac{\bar{\gamma}_B - \gamma_0}{\gamma_0} H_B^*, \tag{4-45}$$

so that

$$\Delta H_{AB}^* = H_B^* - H_A^* = \Delta n_{AB} + NC_{AB}. \tag{4-46}$$

The normal heights were introduced by Molodensky in connection with his method of determining the physical surface of the earth; see Chapter 8.

4-6. Comparison of Different Height Systems

By means of the geopotential number

$$C = W_0 - W = \int_{\text{geoid}}^{\text{point}} g\,dn$$

we can write the different kinds of height in a common form which is very instructive:

$$\text{height} = \frac{C}{G}, \tag{4-47}$$

where the height systems differ according to how the gravity value G in the denominator is chosen. We have:

dynamic height:

$$G = \gamma_0 = \text{const.},$$

orthometric height:

$$G = \bar{g}, \tag{4-48}$$

normal height:

$$G = \bar{\gamma}.$$

It is seen that one can devise an unlimited number of other height systems by selecting G in a different way.

The geopotential number C is, in a way, the most direct result of leveling and is of great scientific importance. However, it is not a height in a geometrical or practical sense. The dynamic height has at least the dimension of a height, but no geometrical meaning. One advantage is that points of the same level surface have the same dynamic height; this corresponds to the intuitive feeling that if we move horizontally we remain at the same height.[1] The dynamic correction can be very large, because gravity varies from equator to pole by about 5000 mgals. Take, for instance, a leveling line of 1000 m difference of height at the equator, where $g \doteq 978.0$ gals, computed with $\gamma_0 = \gamma_{45°} = 980.6$ gals. Then (4-11) gives a dynamic correction of approximately

$$DC = \frac{978.0 - 980.6}{980.6} \cdot 1000 \text{ meters} = -2.7 \text{ m}.$$

Because of these large corrections, dynamic heights are not suitable as practical heights, and the geopotential numbers are preferable for scientific purposes.

Orthometric heights are the natural "heights above sea level," that is, heights above the geoid. They thus have an unequalled geometrical and physical significance. Their computation is relatively laborious, unless Helmert's simple formula (4-26) is used, which is sufficient in most cases. The orthometric correction is rather small. In the Alpine leveling line of Mader (1954), leading from an elevation of 754 meters to 2505 meters, the orthometric correction is about 15 cm per 1 km of measured height difference.

The physical and geometrical meaning of the normal heights is less obvious; they depend on the reference ellipsoid used. Although they are basic in the new theories of physical geodesy, they have a somewhat artificial character as compared to the orthometric heights. They are, however, easy to compute rigorously; the order of magnitude of the normal corrections is about the same as that of the orthometric corrections. In the countries of the eastern hemisphere they have replaced the orthometric heights in practice.

For estimates of the difference between orthometric height H and normal height H^* we refer the reader to Sec. 8-13.

All these height systems resemble C in being functions of position only.

[1] The orthometric height differs for points of the same level surface because the level surfaces are not parallel. This gives rise to the well-known paradoxes of "water flowing uphill," etc.

There are thus no misclosures, as there are with measured heights. From a purely practical point of view the desired requirements of a height system are that

1. Misclosures be eliminated.
2. Corrections to the measured heights be as small as possible.

Empirical height systems have been devised to give smaller corrections than either the orthometric or the normal heights. They have no clear physical significance, however, and are beyond the scope of the present book.

Accuracy. Leveling is one of the most accurate geodetic measurements. A standard error of ± 0.1 mm per km distance is possible; it increases with the square root of the distance.

If the error of measurement and interpolation, etc., of gravity is negligible, then the differences in the geopotential number C can be determined with an accuracy of ± 0.1 gal·meter per km distance; this corresponds to ± 0.1 mm in measured height. To achieve this, Bomford (1962, p. 206) suggests a distance between gravity stations of 2 to 3 km in level country, 1 to 2 km in moderate hills, and 0.3 to 1.5 km in mountainous areas; however, Ramsayer (1963) found station distances of 15 to 25 km, 10 to 15 km, and 5 to 10 km, to be short enough.

Dynamic heights and normal heights are clearly as accurate as the geopotential numbers, because normal gravity γ is errorless. Orthometric heights, however, are also affected by imperfect knowledge of density, etc., but only slightly; see the end of Sec. 4-4.

4-7. Triangulated Heights

For the sake of completeness we must also deal briefly with the determination of heights by triangulation, that is, by means of zenith distances.

The problem is to determine the differences in the ellipsoidal heights h_1 and h_2 of two points P_1 and P_2 if the horizontal (ellipsoidal) distances, the zenith distances z_1' and z_2', and the deflection components ϵ_1 and ϵ_2 are given (Fig. 4-6). Here ϵ_1 and ϵ_2 are the components of the deflection of the vertical in the direction of the line P_1P_2; as we shall see in Sec. 5-4, equation (5-16), they are computed by

$$\epsilon_1 = \xi_1 \cos\alpha + \eta_1 \sin\alpha,$$
$$\epsilon_2 = \xi_2 \cos\alpha + \eta_2 \sin\alpha,$$

where α is the azimuth of the line P_1P_2, and ξ, η are the components of the deflection of the vertical along the meridian and the prime vertical.

The measured zenith distances refer to the astronomical zenith, that is, to the plumb line. Hence they must be converted to the ellipsoidal zenith, which

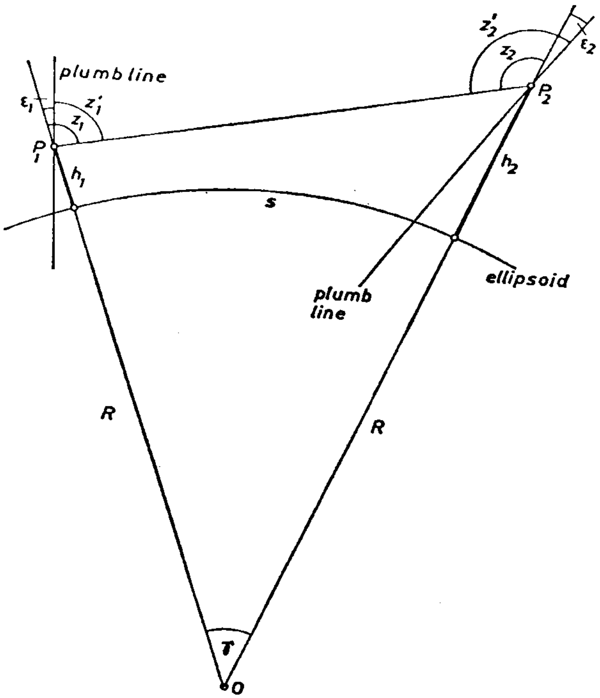

FIGURE 4-6
Triangulated heights.

corresponds to the ellipsoidal normal. These ellipsoidal zenith distances z are computed from the measured zenith distances z' by the formulas

$$\begin{aligned} z_1 &= z_1' + \epsilon_1 = z_1' + \xi_1 \cos\alpha + \eta_1 \sin\alpha, \\ z_2 &= z_2' - \epsilon_2 = z_2' - \xi_2 \cos\alpha - \eta_2 \sin\alpha, \end{aligned} \tag{4-49}$$

which are obtained by inspecting Fig. 4-6.

In this figure we have replaced, with sufficient accuracy, the ellipsoidal arc by a spherical arc of radius

$$R = \frac{1}{2}(R_1 + R_2), \tag{4-50}$$

where R_1 and R_2 are the radii of curvature for the azimuth α at P_1 and P_2:

$$\begin{aligned} \frac{1}{R_1} &= \frac{\cos^2\alpha}{M_1} + \frac{\sin^2\alpha}{N_1}, \\ \frac{1}{R_2} &= \frac{\cos^2\alpha}{M_2} + \frac{\sin^2\alpha}{N_2}; \end{aligned} \tag{4-51}$$

M and N are the principal radii of curvature of the ellipsoid (see Sec. 2-8).

Apply the law of tangents, known from plane trigonometry, to the triangle OP_1P_2, obtaining

$$\frac{R + h_2 - R - h_1}{R + h_2 + R + h_1} = \frac{\tan \frac{1}{2}(180^\circ - z_1 - 180^\circ + z_2)}{\tan \frac{1}{2}(180^\circ - z_1 + 180^\circ - z_2)}$$

or

$$\frac{h_2 - h_1}{2R + h_1 + h_2} = \tan \frac{\gamma}{2} \tan \frac{z_2 - z_1}{2}. \tag{4-52}$$

We read from Fig. 4-6 that

$$\gamma = \frac{s}{R}, \tag{4-53}$$

so that

$$\tan \frac{\gamma}{2} = \frac{\gamma}{2} + \frac{1}{3}\left(\frac{\gamma}{2}\right)^3 \cdots = \frac{s}{2R}\left(1 + \frac{s^2}{12R^2} \cdots\right).$$

On introducing

$$h_m = \frac{1}{2}(h_1 + h_2),$$

the mean elevation of the line P_1P_2, equation (4-52) finally becomes

$$h_2 - h_1 = s\left(1 + \frac{h_m}{R} + \frac{s^2}{12R^2}\right) \tan \frac{z_2 - z_1}{2}. \tag{4-54}$$

If only one zenith distance z_1 has been measured, we can compute γ from (4-53) and z_2 from the condition that the sum of angles of the triangle OP_1P_2 is 180°:

$$180^\circ - z_1 + 180^\circ - z_2 + \gamma = 180^\circ,$$

or

$$z_1 + z_2 - \gamma - 180^\circ = 0. \tag{4-55}$$

Thus,

$$\frac{z_2 - z_1}{2} = 90^\circ + \frac{\gamma}{2} - z_1. \tag{4-56}$$

The main problem in determining heights by triangulation is the effect of atmospheric refraction, which affects the zenith distances much more strongly than the horizontal angles. Owing to uncertainties and variations of refraction an accuracy of $\pm 1''$ in z is at present possible only in exceptional cases, such as in high mountains. To eliminate the effect of refraction as far as possible it is preferable to measure both z_1 and z_2 (reciprocal observations) rather than to measure only z_1 and use (4-56). If $s = 10$ km, the standard error of the elevation difference for reciprocal observations is of the order of ± 10 cm. Thus the accuracy of triangulated heights is much less than that of leveling.

The problem of atmospheric refraction will not be treated here; we have assumed that the observed zenith distances had already been corrected for the effect of refraction.

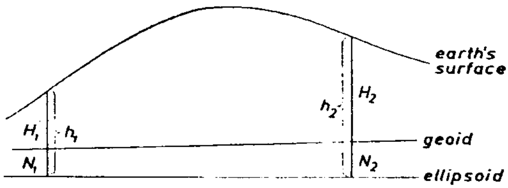

FIGURE 4-7
Heights above the ellipsoid and above the geoid.

Determination of differences in vertical deflections and geoidal undulations. The procedure described in this section presupposes that astrogeodetic deflections of the vertical (see next chapter) are known, at least at certain points. On the other hand, measurements of zenith distances may be used to determine differences in the deflection of the vertical: The insertion of (4-49) into (4-55) gives

$$\epsilon_2 - \epsilon_1 = z_1' + z_2' - \gamma - 180°, \qquad (4\text{-}57)$$

where $\gamma = s/R$ is assumed to be known.

Triangulated heights are heights h above the reference ellipsoid. Orthometric heights H, obtained by leveling, are heights above the geoid. As Fig. 4-7 shows,

$$h = H + N,$$

so that

$$N_2 - N_1 = (h_2 - h_1) - (H_2 - H_1). \qquad (4\text{-}58)$$

Combining the triangulated (geometric) height differences with the leveled (orthometric) height differences we thus obtain differences of geoidal undulation N.

Zenith distance measurements are also important as a means for height determination in a slightly different context, in the so-called three-dimensional geodesy (Sec. 5-12).

References

Baeschlin, C. F. (1960). Rapport spécial sur le nivellement et la pesanteur. Bull. Géod., No. 57, pp. 247–298.

Bodemüller, H. (1957). Beitrag zur Schwerekorrektion geometrischer Nivellements. Deutsche Geodätische Kommission, A, No. 26.

Bomford, G. (1962). Geodesy. 2nd ed., Oxford University Press.

Helmert, F. R. (1884). Die mathematischen und physikalischen Theorien der höheren Geodäsie, vol. 2. Leipzig, B. G. Teubner (reprinted 1962).

Helmert, F. R. (1890). Die Schwerkraft im Hochgebirge, insbesondere in den Tyroler Alpen. Veröff. Königl. Preuss. Geod. Inst., No. 1.

Hirvonen, R. A. (1960). New theory of the gravimetric geodesy. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 32.

Kobold, F., and N. Wunderlin (1963). Die Bestimmung von Lotabweichungen und Meereshöhen im Berner Oberland und im Oberwallis. Commission Géodésique Suisse.

Kukkamäki, T. J. (1958). Rapport Général. Les nivellements de précision, 1954-1956. Travaux, Int. Assoc. Geod., 20 (II).
Ledersteger, K. (1955). Der Schwereverlauf in den Lotlinien und die Berechnung der wahren Geoidschwere. Publication Dedicated to Weikko A. Heiskanen, Publ. Finn. Geod. Inst., No. 46, pp. 109-124.
Mader, K. (1954). Die orthometrische Schwerekorrektion des Präzisions-Nivellements in den Hohen Tauern. Österreichische Zeitschrift für Vermessungswesen, Sonderheft 15.
Niethammer, T. (1932). Nivellement und Schwere als Mittel zur Berechnung wahrer Meereshöhen. Schweizerische Geodätische Kommission.
Ramsayer, K. (1963). Über den zulässigen Abstand der Schwerepunkte bei der Bestimmung geopotentieller Koten im Hochgebirge, Mittelgebirge und Flachland. Deutsche Geodätische Kommission, A, No. 44.

5
Astrogeodetic Methods

5-1. Introduction

Geodesy, as the theory of size and shape of the earth, may appear to be a purely geometrical science. Actually, however, the earth's gravity field, a physical entity, is inextricably involved in most geodetic measurements, even the purely geometric ones. The measurements of geodetic astronomy, triangulation, and leveling all make essential use of the plumb line, which, being the direction of the gravity vector, is no less physically defined than its magnitude, that is, the gravity g.

Thus the astrogeodetic methods, which use the astronomical determinations of latitude, longitude, and azimuth and the geodetic operations of triangulation, base-line measurements, and trilateration may properly be considered as belonging to physical geodesy, fully as much as the gravimetric methods.

As a general distinction, the astrogeodetic methods use the direction of the gravity vector, employing geometrical techniques, whereas the gravimetric methods operate with the magnitude g, making use of potential theory. A sharp demarcation is impossible and there are frequent overlaps. The gravimetric methods are usually considered to constitute physical geodesy in the narrower sense.

In this chapter we shall consider some basic principles of the astrogeodetic methods, which, apart from their intrinsic interest, are also indispensable for a deeper understanding of the gravimetric methods.

A few introductory ideas may help in comprehending this subject. To fix the position of a point in space we need three coordinates. We can use, and have used, a rectangular cartesian coordinate system. For many purposes, however, it is preferable to take what we have called the *natural coordinates*: Φ (geographical latitude), Λ (geographical longitude), and H (height above the geoid), which directly refer to the gravity field of the earth (Sec. 2-4).

The height H is obtained by geometric leveling, combined with gravity measurements, and Φ and Λ are determined by astronomical measurements. As long as the geoid can be identified with an ellipsoid, the use of these coordinates for computations is very simple. Since this identification is sufficient only for results of rather low accuracy, the deviations of the geoid from an ellipsoid must be taken into account. As we have seen, the geoid unfortunately has rather disagreeable mathematical properties. It is a complicated surface with discontinuities of curvature. Thus it is not suitable as a surface on which to perform mathematical computations directly, as on the ellipsoid.

Since the deviations of the geoid from the ellipsoid are small and can be computed, it is convenient to add small reductions to the original coordinates Φ, Λ, H, so as to get values which refer to an ellipsoid. In this way we shall find in Sec. 5-4:

$$\begin{aligned} \phi &= \Phi - \xi, \\ \lambda &= \Lambda - \eta \sec \phi, \\ h &= H + N; \end{aligned}$$

ϕ and λ are the geographical coordinates on the ellipsoid, also called *geodetic latitude* and *geodetic longitude* to distinguish them from the *astronomical latitude* Φ and the *astronomical longitude* Λ. Astronomical and geodetic coordinates differ by the deflection of the vertical (components ξ and η). The quantity h is the *geometric height* above the ellipsoid; it differs from the *orthometric height* H above the geoid by the geoidal undulation N.

Geodetic measurements (angles, distances) are treated similarly. The principle of *triangulation* is well known: distances are obtained indirectly by measuring the angles in a suitable network of triangles; only one base line is necessary to furnish the scale of the network. Triangulation was indispensable in former times, because angles could be measured much more easily than long distances. Nowadays, however, long distances can be measured directly just as easily as angles by means of electronic instruments, so that triangulation, using angular measurements, is often replaced or supplemented by *trilateration*, using distance measurements. The computation of triangulations and trilaterations on the ellipsoid is easy. It is therefore convenient to reduce the measured angles, base lines, and long distances to the ellipsoid, in much the same way as the astronomical coordinates are treated. Then the geodetic (ellipsoidal) coordinates ϕ, λ, obtained (1) by reducing the astronomical coordinates and (2) by computing triangulations or trilaterations on the ellipsoid, can be compared; they should be identical for the same point.

5-2. Projections onto the Ellipsoid

Let us establish the position of a point P by means of the natural coordinates Φ, Λ, H. Then we may project it onto the geoid along the (slightly curved) plumb line. The orthometric height is the distance between P and its projection P_0 onto the geoid, measured along the plumb line (Fig. 5-1). Although this mode of projection is entirely natural, the geoid is not suited for performing computations on it directly; the point P_0 is therefore projected onto the reference ellipsoid by means of the straight ellipsoidal normal, thus getting a point Q_0 on the ellipsoid. In this way, the ground point P and the corresponding point Q_0 on the ellipsoid are connected by a double projection, that is, by two projections which are performed one after the other and which are quite analogous, the orthometric height $H = PP_0$ corresponding to the geoidal undulation $N = P_0Q_0$. This double projection is called *Pizzetti's projection*.

It is simpler to project the point P from the physical surface of the earth directly onto the ellipsoid through the straight ellipsoidal normal, thus obtaining a point Q. The distance $PQ = h$ is the geometrical height above the ellipsoid. The ground point P is then determined by h and the geographical coordinates ϕ, λ of Q on the ellipsoid, so that the so-called *geodetic coordinates* ϕ, λ, h take the place of the *natural coordinates* Φ, Λ, H. This projection is called *Helmert's projection.*

The practical difference between Pizzetti's and Helmert's projection is small. The ellipsoidal height h is equal to $H + N$ within a fraction of a millimeter. The geodetic coordinates ϕ and λ, with respect to the two projections, are related by the equations

$$\phi_{\text{Helmert}} = \phi_{\text{Pizzetti}} + \frac{H}{R}\xi,$$

$$\lambda_{\text{Helmert}} = \lambda_{\text{Pizzetti}} + \frac{H}{R}\eta \sec\phi,$$

which can be read from Figure 5-1, since $QQ_0 \doteq H\epsilon$; $R = 6371$ km is the mean radius of the earth. Even if $\epsilon = 1$ minute of arc and $H = 1000$ meters, the dis-

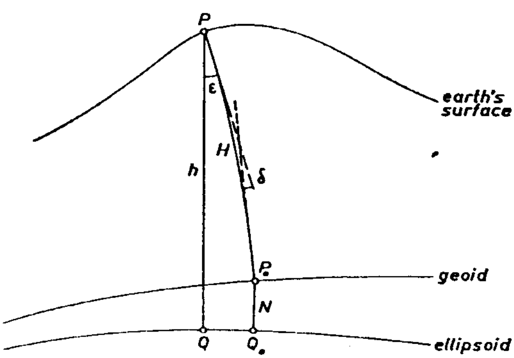

FIGURE 5-1

The projections of Helmert and of Pizzetti.

tance QQ_0 is only about 30 cm and the geodetic coordinates differ by less than 0.01″, which is below the accuracy of astronomical observations. For most purposes we may therefore neglect the difference between the two projections.

Pizzetti's projection is better adapted to the geoid, because there is an exact correspondence between a geoidal point P_0 and an ellipsoidal point Q_0. Helmert's projection has practical advantages, notably the straightforward conversion of the ellipsoidal coordinates ϕ, λ, h into rectangular coordinates x, y, z; it is also simpler in other respects. For this reason we shall henceforth use mainly Helmert's projection, but practically the results hold for both projections.

However, because of the curvature of the plumb line we have to distinguish carefully whether the astronomical coordinates refer to the ground point P or the geoidal point P_0. Even if the angle δ in Fig. 5-1 is only 1 second of arc, a change of 1″ in the geographical latitude means a linear displacement of P_0 by $R\delta \doteq 30$ m. This must be taken into account if we combine astronomical coordinates Φ and Λ, measured at the ground point P, and the gravimetric deflections of the vertical ξ and η, computed by Vening Meinesz' formula for the geoidal point P_0.

5-3. Helmert's Projection. Geodetic and Rectangular Coordinates

We shall now derive the relation between the geodetic coordinates ϕ, λ, h of Helmert's projection and the corresponding rectangular coordinates x, y, z.

The equation of the reference ellipsoid in rectangular coordinates is

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2} = 1. \tag{5-1}$$

The representation of this ellipsoid in terms of geographical coordinates is given by

$$\begin{aligned} x &= N \cos\phi \cos\lambda, \\ y &= N \cos\phi \sin\lambda, \\ z &= \frac{b^2}{a^2} N \sin\phi, \end{aligned} \tag{5-2}$$

where N is the east-west radius of curvature (2-81):

$$N = \frac{a^2}{\sqrt{a^2 \cos^2\phi + b^2 \sin^2\phi}}. \tag{5-3}$$

These equations are known from geometrical geodesy; it may also be verified by direct substitution that a point with xyz-coordinates (5-2) satisfies the equation of the ellipsoid (5-1) and so lies on the ellipsoid.

The components of the unit normal vector $\mathbf{n}$ are

$$\mathbf{n} = (\cos\phi \cos\lambda, \cos\phi \sin\lambda, \sin\phi), \tag{5-4}$$

because ϕ is the angle between the ellipsoidal normal and the xy-plane, which is the equatorial plane (Fig. 5-2).

Now let the coordinates of a point P outside the ellipsoid form the vector

$$\mathbf{X} = (X, Y, Z);$$

similarly we have, for the coordinates of the point Q on the ellipsoid,

$$\mathbf{x} = (x, y, z).$$

From Figure 5-2 we read

$$\mathbf{X} = \mathbf{x} + h\mathbf{n},$$

that is

$$X = x + h \cos \phi \cos \lambda,$$
$$Y = y + h \cos \phi \sin \lambda,$$
$$Z = z + h \sin \phi.$$

By (5-2) this becomes

$$X = (N + h) \cos \phi \cos \lambda, \tag{5-5a}$$

$$Y = (N + h) \cos \phi \sin \lambda, \tag{5-5b}$$

$$Z = \left(\frac{b^2}{a^2} N + h\right) \sin \phi. \tag{5-5c}$$

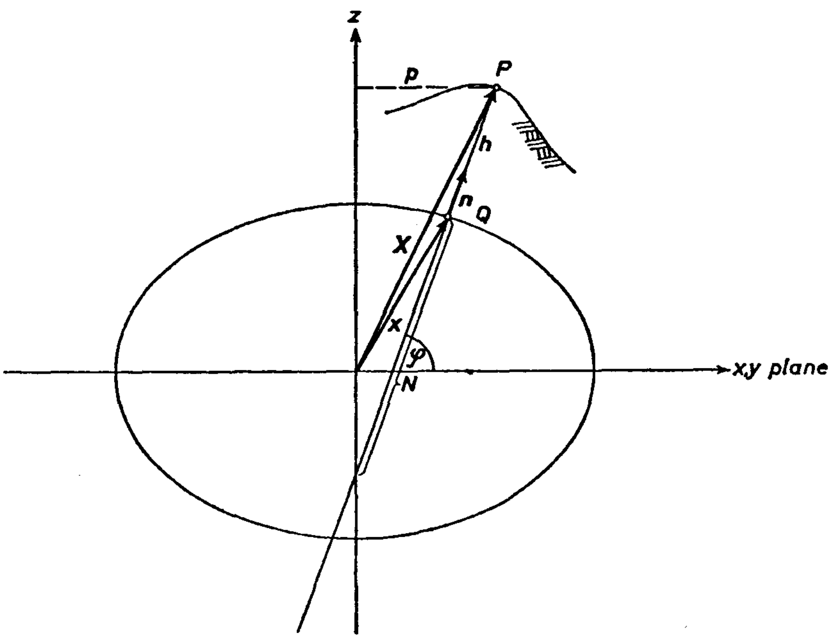

FIGURE 5-2

Geodetic and rectangular coordinates.

These equations are the basic transformation formulas between the geodetic coordinates ϕ, λ, h and the rectangular coordinates X, Y, Z of a point outside the ellipsoid. The origin of the rectangular coordinate system is the center of the ellipsoid, and the z-axis is its axis of rotation; the x-axis has the Greenwich longitude 0° and the y-axis has the longitude 90° east of Greenwich (i.e., $\lambda = +90°$).

A possible source of confusion is that the east-west radius of curvature of the ellipsoid and the geoidal undulation are both denoted by the symbol N; in (5-5) N is, of course, the radius of curvature.

Equations (5-5) permit the computation of rectangular coordinates X, Y, Z from the geodetic coordinates ϕ, λ, h. The inverse procedure, the computation of ϕ, λ, h from given X, Y, Z, is more complicated because (5-5) cannot be solved for ϕ, λ, h in a closed form; accordingly, the computation must be done iteratively (Hirvonen and Moritz, 1963, p. 4).

Denoting $\sqrt{X^2 + Y^2}$ by p, we get from equations (5-5a) and (5-5b) or from Fig. 5-2

$$p = \sqrt{X^2 + Y^2} = (N + h)\cos\phi,$$

whence

$$h = \frac{p}{\cos\phi} - N. \tag{5-6a}$$

Equation (5-5c) may be transformed into

$$Z = \left(N - \frac{a^2 - b^2}{a^2}N + h\right)\sin\phi = (N + h - e^2N)\sin\phi,$$

where $e^2 = (a^2 - b^2)/a^2$. Dividing this equation by the above expression for p we find

$$\frac{Z}{p} = \left(1 - e^2\frac{N}{N + h}\right)\tan\phi,$$

so that

$$\tan\phi = \frac{Z}{p}\left(1 - e^2\frac{N}{N + h}\right)^{-1}. \tag{5-6b}$$

Given X, Y, Z, and hence p, equations (5-6a) and (5-6b) may be solved iteratively for h and ϕ. As a first approximation, we set $h = 0$ in (5-6b), obtaining

$$\tan\phi_{(1)} = \frac{Z}{p}(1 - e^2)^{-1}.$$

Using $\phi_{(1)}$, we compute an approximate value $N_{(1)}$ by means of (5-3). Then (5-6a) gives $h_{(1)}$.

Now, as a second approximation, we set $h = h_{(1)}$ in (5-6b), obtaining

$$\tan\phi_{(2)} = \frac{Z}{p}\left(1 - e^2\frac{N_{(1)}}{N_{(1)} + h_{(1)}}\right).$$

Using $\phi_{(2)}$, improved values for N and h are found, etc. This procedure is repeated until ϕ and h remain practically constant.

It is important not to confuse the geodetic coordinates ϕ, λ, h with the ellip-

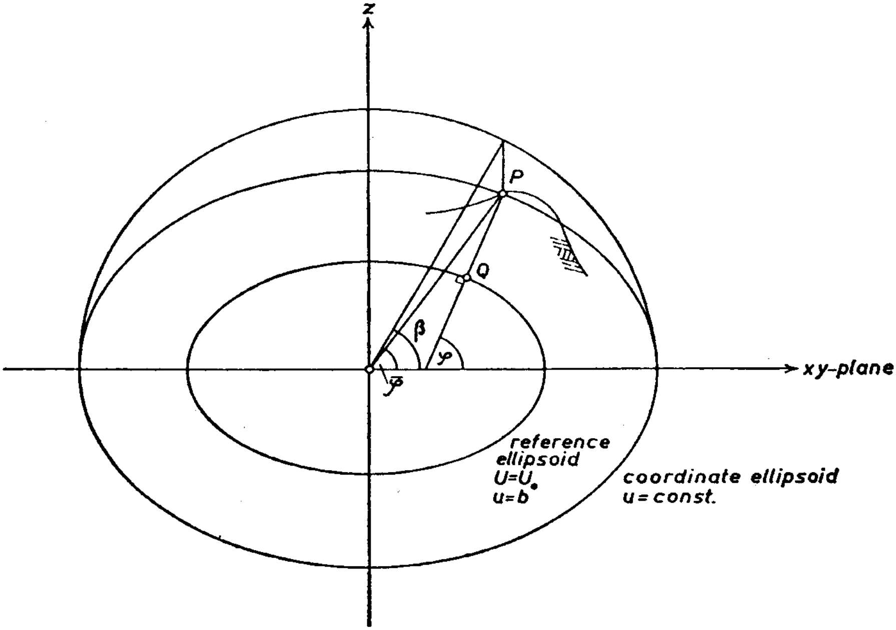

FIGURE 5-3

Geodetic, reduced, and geocentric latitudes.

soidal coordinates β, λ, u, introduced earlier in connection with the normal gravity field, nor with the spherical coordinates $\bar{\phi}$, λ, r. The longitude λ is the same in all three systems; ϕ is the *geodetic* latitude (geographical latitude on the ellipsoid), β is the *reduced* latitude, and $\bar{\phi}$ is the *geocentric* latitude (see Fig. 5-3). The following equations express the rectangular coordinates in these three systems:

$$\begin{aligned} X &= (N + h)\cos\phi\cos\lambda = \sqrt{u^2 + E^2}\cos\beta\cos\lambda = r\cos\bar{\phi}\cos\lambda, \\ Y &= (N + h)\cos\phi\sin\lambda = \sqrt{u^2 + E^2}\cos\beta\sin\lambda = r\cos\bar{\phi}\sin\lambda, \\ Z &= \left(\frac{b^2}{a^2}N + h\right)\sin\phi \quad = u\sin\beta \qquad\qquad\quad = r\sin\bar{\phi}. \end{aligned} \tag{5-7}$$

These relations, which follow from combining equations (1-36), (1-103), and (5-5), can be used if we wish to compute u and β from h and ϕ or from r and $\bar{\phi}$, etc.

5-4. Reduction of Astronomical Observations to the Ellipsoid

Now we shall establish the relation between the natural coordinates Φ, Λ, H and the geodetic coordinates ϕ, λ, h referring to an ellipsoid according to

Helmert's projection. Leaving aside, for the moment, the heights h and H, we may also formulate the problem as the *reduction of the astronomical coordinates Φ and Λ to the ellipsoid.* If we also include astronomical observation of azimuth we have to reduce the astronomical coordinates Φ and Λ and the astronomical azimuth A to the ellipsoid in order to obtain the geodetic coordinates ϕ and λ and the geodetic azimuth α.

Consider a unit sphere (sphere of radius 1) with its center at the observation station P. This sphere has already been used in Sec. 2-13; see also Fig. 2-13. The actual plumb line intersects this sphere at the astronomical zenith Z_a, whereas the ellipsoidal normal intersects it at the geodetic zenith Z_g. Figure 5-4 shows this unit sphere as viewed from above. The line of sight to the target, for which the azimuth A is measured, intersects the unit sphere at the point T and has the zenith distances z' and z with respect to the zeniths Z_a and Z_g. The point P_N corresponds to the direction to the north pole, which has the zenith distances $90° - \Phi$ and $90° - \phi$, with respect to Z_a and Z_g; the angle at P_N is the difference

$$\Delta\lambda = \Lambda - \lambda \tag{5-8}$$

between the astronomical and the geodetic longitude. The angle at Z_a is the astronomical azimuth A, corresponding to the geodetic azimuth α at Z_g. The point F lies on the astronomical meridian, the great circle connecting P_N and

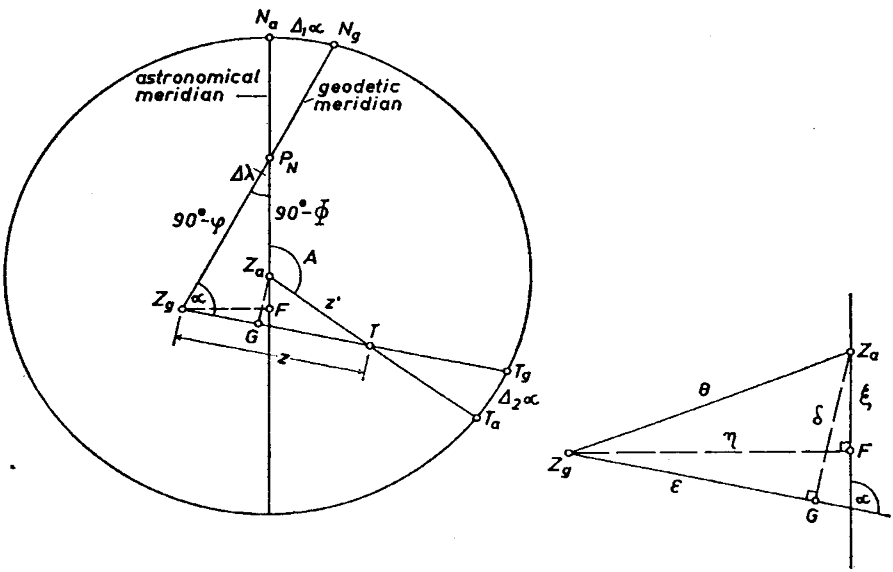

FIGURE 5-4

The unit sphere illustrating the deflection of the vertical, as seen from above, and an enlarged view of the central portion of the figure.

Z_a, so that the angle Z_aFZ_0 is 90°; $\xi = Z_aF$ and $\eta = Z_0F$ are the components of the deflection of the vertical.

Consider first the rectangular spherical triangle with corners Z_0, F, and P_N. By Napier's rules we have

$$\sin \phi = \cos (90° - \Phi + \xi) \cos \eta,$$
$$\sin \eta = \cos (90° - \Delta\lambda) \cos \phi.$$

For the small angles η and $\Delta\lambda$ we may use the approximations

$$\cos \eta \doteq 1, \quad \sin \eta \doteq \eta, \quad \cos (90° - \Delta\lambda) = \sin \Delta\lambda \doteq \Delta\lambda.$$

Thus we find

$$\xi = \Phi - \phi = \Delta\phi, \tag{5-9a}$$

$$\eta = (\Lambda - \lambda) \cos \phi = \Delta\lambda \cos \phi. \tag{5-9b}$$

These are the basic equations that express the components ξ and η of the deflection of the vertical in terms of the geographic (astronomical and geodetic) coordinates, thus linking astronomical and geodetic coordinates. They have already been found in Sec. 2-13 in a different way.

The difference in azimuth

$$\Delta\alpha = A - \alpha \tag{5-10}$$

consists of two parts, $\Delta_1\alpha$ and $\Delta_2\alpha$ (Fig. 5-4):

$$\Delta\alpha = \Delta_1\alpha + \Delta_2\alpha. \tag{5-10'}$$

$\Delta_1\alpha$ is obtained from the spherical triangle $N_0N_aP_N$, which is obviously similar to the triangle Z_0FP_N previously used, $N_0P_N = \phi$ corresponding to $Z_0P_N = 90° - \phi$, and $\Delta_1\alpha$ corresponding to η. The equation corresponding to (5-9b) is thus

$$\Delta_1\alpha = \Delta\lambda \sin \phi; \tag{5-11}$$

together with (5-9b) this becomes

$$\Delta_1\alpha = \eta \tan \phi. \tag{5-12a}$$

On introducing a point G on the great circle connecting Z_0 and T so that the angle Z_aGZ_0 is 90°, and putting $Z_aG = \delta$, we see that the figure $Z_aGTT_0T_a$ has the same geometry as the figure $Z_0FP_NN_aN_0$, so that $\Delta_2\alpha$, δ, z' correspond to $\Delta_1\alpha$, η, $90° - \phi$. The equation corresponding to (5-12a) is thus

$$\Delta_2\alpha = \delta \cot z' \doteq \delta \cot z.$$

Since the small figure Z_aFZ_0G may be considered plane (see the enlarged section of Fig. 5-4), we get by the usual formula of the transformation of plane coordinates

$$\delta = \xi \sin \alpha - \eta \cos \alpha,$$

so that

$$\Delta_2\alpha = (\xi \sin \alpha - \eta \cos \alpha) \cot z, \tag{5-12b}$$

and with (5-12a) we obtain

$$\Delta\alpha = \eta \tan \phi + (\xi \sin \alpha - \eta \cos \alpha) \cot z. \tag{5-13}$$

The first term, $\Delta_1\alpha$, is the same for every target, independent of its azimuth and zenith distance; the second term, $\Delta_2\alpha$, depends on azimuth and zenith distance. The term $\Delta_1\alpha$ results from the astronomical azimuth A being reckoned from astronomical north N_a rather than from geodetic north N_g, as is the geodetic azimuth α. It thus represents a shift of the zero point, which is the same for all targets. The term $\Delta_2\alpha$ arises because the target T is projected from Z_a and Z_g onto different points T_a and T_g of the horizon; the effect is the same as that of an inaccurate leveling of the theodolite.

Usually in first-order triangulation the lines of sight are almost horizontal, so that $z \doteq 90°$, $\cot z \doteq 0$. Therefore, the correction $\Delta_2\alpha$ can in general be neglected and we thus get

$$\Delta\alpha = \eta \tan \phi = \Delta\lambda \sin \phi. \tag{5-14}$$

This is *Laplace's equation* in its usual simplified form. It is remarkable that the differences $\Delta\alpha = A - \alpha$ and $\Delta\lambda = \Lambda - \lambda$ should be related in such a simple way.

For later reference we note that the total deflection of the vertical—that is, the angle θ between the actual plumb line and the ellipsoidal normal—is given by

$$\theta = \sqrt{\xi^2 + \eta^2}, \tag{5-15}$$

and that the deflection component ϵ in the direction of the azimuth α is

$$\epsilon = \xi \cos \alpha + \eta \sin \alpha. \tag{5-16}$$

Both relations can be read immediately from the enlarged section of Fig. 5-4; ϵ and δ are related to ξ and η by a plane coordinate transformation.

Finally, the relationship between the orthometric height H above the geoid and the geometric height h above the ellipsoid can be written down immediately because from Fig. 5-1 we read that, to a sufficient approximation,

$$h = H + N.$$

Thus the conversion formulas from natural to geodetic coordinates are

$$\begin{aligned} \phi &= \Phi - \xi, \\ \lambda &= \Lambda - \eta \sec \phi, \\ h &= H + N, \end{aligned} \tag{5-17}$$

and the corresponding formula for the azimuth is

$$\alpha = A - \eta \tan \phi. \tag{5-18}$$

In the application of these formulas we need the geoidal undulation N and the deflection components ξ and η with respect to the reference ellipsoid used. Two points should be noted:

1. The axis of the reference ellipsoid is parallel to the earth's axis of rotation (otherwise there would be two different poles P_N in Fig. 5-4), but it need not be in an absolute position, its center coinciding with the earth's center of gravity.

2. The deflection components ξ and η refer directly to the point on the ground at which the astronomical observations are made, and not to the geoid.

If components ξ and η of the deflection of the vertical are computed gravimetrically for the geoid by Vening Meinesz' formula, then ϕ, λ, h, and α refer to an ellipsoid in absolute position, but care should be taken because of the curvature of the plumb line; see also the end of Sec. 5-2.

It should also be mentioned that the ellipsoidal azimuth α (5-18) refers to the actual target T, which does not in general lie on the ellipsoid. For the conventional method of computation on the ellipsoid one wishes the azimuth to refer to a target T_0 on the ellipsoid, which is the point at the foot of the normal through T. Furthermore, α refers to what is called a normal section of the ellipsoid, rather than to a geodesic line, which is used in computation. In either case very small azimuth reductions are necessary; since these reductions are purely problems in ellipsoidal geometry, the reader is referred to any textbook on geometrical geodesy or to Bomford (1962).

Effect of polar migration. The direction of the earth's axis of rotation is not rigorously fixed with respect to the earth but undergoes very small, more or less periodic variations. This phenomenon arises from a minute difference between the axes of rotation and of maximum inertia, the angle between these axes being about 0.3″, and is somewhat similar to the precession of a spinning top. This motion of the pole has a main period of about 430 days, the Chandler period, but is rather irregular, presumably because of the movement of masses, atmospheric variations, etc. (Fig. 5-5).

The International Latitude Service, which is maintained by the International Astronomical Union and by the International Union of Geodesy and Geophysics, continuously observes the variation of latitude at several stations and

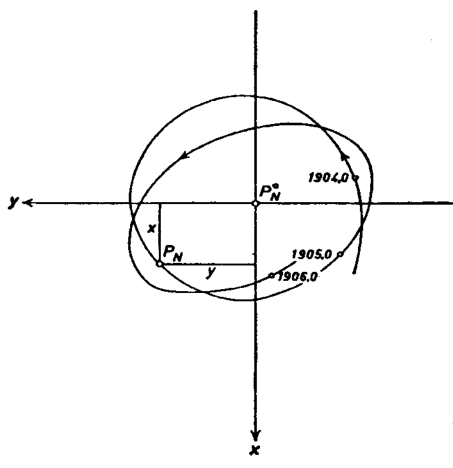

FIGURE 5-5
Polar motion.

thus determines the motion of the pole. The results are published as the rectangular coordinates of the instantaneous pole P_N with respect to a mean pole P'_N (Fig. 5-5). The astronomically observed values of Φ, Λ, and A naturally refer to the instantaneous pole P_N and must therefore be reduced to the mean pole, using the published values of x and y.

This is accomplished by means of the equations

$$\begin{aligned} \Phi &= \Phi_{obs} - x \cos \lambda + y \sin \lambda, \\ \Lambda &= \Lambda_{obs} - (x \sin \lambda + y \cos \lambda) \tan \phi + y \tan \phi_{Gr}, \\ A &= A_{obs} - (x \sin \lambda + y \cos \lambda) \sec \phi. \end{aligned} \tag{5-19}$$

Now Φ, Λ, A are referred to the mean pole; these values are used in geodesy because they do not vary with time. Longitude, throughout this book, is reckoned positive to the east, as is usual in geodesy; it should be mentioned that in the literature these formulas are often written for west longitude, according to the practice of many astronomers. Since the correction terms containing x and y are extremely small (of the order of 0.1″), we may use either the geodetic values ϕ and λ or the astronomical values Φ and Λ in these terms. The term containing ϕ_{Gr} (the latitude of Greenwich) in the formula for Λ is usually omitted, so that the mean meridian of Greenwich remains fixed, rather than the astronomical longitude of Greenwich itself.

The derivation of these formulas is beyond the scope of the present book; it is given in textbooks on spherical astronomy. Nevertheless, it is interesting to note the close similarity between the azimuth reduction (5-13) because of the "zenith variation"—that is, the deflection of the vertical—and the longitude reduction of (5-19) because of the polar variation. Actually, the geometry for both cases is the same. The quantities ξ, η, $90° - z$, ϕ correspond to x, y, ϕ, ϕ_{Gr}; the difference in sign of $\sin \alpha$ and $\sin \lambda$ is due to the fact that, when viewed from the zenith, azimuth is reckoned clockwise and, when viewed from the pole, east longitude is reckoned counterclockwise.

5-5. Reduction of Horizontal and Vertical Angles and of Distances

Horizontal angles. To reduce an observed horizontal angle ω to the ellipsoid we note that every angle may be considered as the difference between two azimuths:

$$\omega = \alpha_2 - \alpha_1.$$

Hence we can apply formula (5-13). In the difference $\alpha_2 - \alpha_1$, the main term $\eta \tan \phi$ drops out, so that for nearly horizontal lines of sight the whole reduction may be neglected.

Vertical angles. The relation between the measured zenith distance z' and the corresponding ellipsoidal zenith distance z was found in Sec. 4-7. Equation (4-49) gives

$$z = z' + \epsilon = z' + \xi \cos \alpha + \eta \sin \alpha, \tag{5-20}$$

where α is the azimuth of the target. This equation may also be obtained by inspecting Fig. 5-4.

Base lines. Figure 5-6 illustrates the reduction of measured base lines to the ellipsoid. Denote an element of the measured distance by dl. It has an inclination β towards the local horizon (the level geopotential surface passing through dl). The deflection component in the direction of the measured line that has the azimuth α is again denoted by ϵ and given by (5-16). The element ds, which is the component of dl parallel to the ellipsoid, is

$$ds = dl \cos(\beta - \epsilon) \doteq dl \cos \beta + \epsilon\, dl \sin \beta.$$

Denoting by dl' the projection of dl onto the local horizon,

$$dl' = dl \cos \beta,$$

and noting that

$$dl \sin \beta \doteq dh,$$

we have

$$ds = dl' + \epsilon\, dh.$$

If R is the local radius of curvature of the azimuth α of the ellipsoid, then it is shown in differential geometry that

$$\frac{1}{R} = \frac{\cos^2 \alpha}{M} + \frac{\sin^2 \alpha}{N}, \tag{5-21}$$

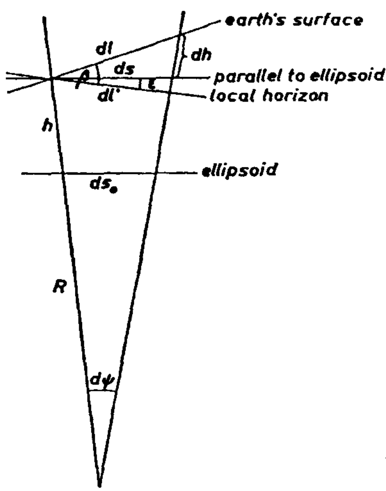

FIGURE 5-6

Reduction of base lines.

where M and N are, respectively, the north-south and east-west radii of curvature. Then, if ds_0 is the projection of dl onto the ellipsoid,

$$\frac{ds}{ds_0} = \frac{R+h}{R} = 1 + \frac{h}{R}$$

or

$$ds_0 = ds - \frac{h}{R}\,ds_0 = dl' + \epsilon\,dh - \frac{h}{R}\,ds_0. \tag{5-22}$$

Setting

$$\frac{ds_0}{R} = d\psi \tag{5-23}$$

we have

$$ds_0 = dl' + \epsilon\,dh - h\,d\psi = dl' + d(\epsilon h) - h\,d(\psi + \epsilon),$$

and on integration between the end points A and B we obtain

$$s_0 = l' + \epsilon_B h_B - \epsilon_A h_A - \int_A^B h\,d(\psi + \epsilon). \tag{5-24}$$

If the elevation h is nearly constant along the line, as occurs almost always in base-line measurements, then the application of a mean-value theorem of integral calculus gives

$$s_0 = l' + \epsilon_B h_B - \epsilon_A h_A - h_m(\epsilon_B - \epsilon_A) - h_m \int_A^B d\psi.$$

Here

$$l' = \int_A^B dl \cos\beta$$

is the sum of the locally reduced dl', and h_m is a mean elevation along the line. On expressing $d\psi$ in terms of ds_0 by (5-23) and integrating we finally obtain

$$s_0 = l' + \epsilon_B(h_B - h_m) - \epsilon_A(h_A - h_m) - \frac{h_m}{R}\,s_0. \tag{5-25}$$

Strictly speaking, R, the local ellipsoidal radius of curvature of azimuth α, is slightly variable along the line from A to B. In practice, however, it is perfectly permissible to replace the local value of R by its average along the line, so that R in (5-23) can be considered constant, which leads to (5-25). This amounts to the approximation of the ellipsoidal arc AB by a circular arc whose radius R is the average along AB of the values given by (5-21).

The terms with ϵ_A and ϵ_B represent the effect of the inclination between the geopotential and spheropotential surfaces; they will often be negligible. The term $s_0 h_m/R$ is due to the convergence of the ellipsoidal normals.

The rigorous reduction of base lines according to (5-25) thus involves the geoidal undulation N, through the height h above the ellipsoid, and the deflection of the vertical ϵ. The base lines are reduced directly to the ellipsoid by means of the straight ellipsoidal normals, in conformity with Helmert's projection.

Spatial distances. Electronic measurement of distance yields straight spatial distances l between two points A and B (Fig. 5-7). These distances may either

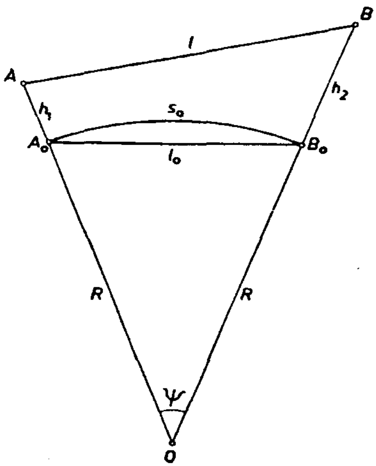

FIGURE 5-7
Reduction of spatial distances.

be used directly for computations in the geodetic coordinate system ϕ, λ, h, as in "three-dimensional geodesy" (see Sec. 5-12), or they may be reduced to the surface of the ellipsoid to obtain chord distances l_0 or geodesic distances s_0.

We shall again approximate the ellipsoidal arc A_0B_0 by a circular arc of radius R that is the mean ellipsoidal radius of curvature along A_0B_0. By applying the law of cosines to the triangle OAB we find

$$l^2 = (R + h_1)^2 + (R + h_2)^2 - 2(R + h_1)(R + h_2)\cos\psi.$$

With

$$\cos\psi = 1 - 2\sin^2\frac{\psi}{2}$$

this is transformed into

$$l^2 = (h_2 - h_1)^2 + 4R^2\left(1 + \frac{h_1}{R}\right)\left(1 + \frac{h_2}{R}\right)\sin^2\frac{\psi}{2};$$

and with

$$l_0 = 2R\sin\frac{\psi}{2}$$

and the abbreviation $\Delta h = h_2 - h_1$, we obtain

$$l^2 = \Delta h^2 + \left(1 + \frac{h_1}{R}\right)\left(1 + \frac{h_2}{R}\right)l_0^2.$$

Hence the chord l_0 and the arc s_0 are expressed by

$$l_0 = \sqrt{\frac{l^2 - \Delta h^2}{\left(1 + \frac{h_1}{R}\right)\left(1 + \frac{h_2}{R}\right)}}; \tag{5-26}$$

$$s_0 = R\psi = 2R\sin^{-1}\frac{l_0}{2R}. \tag{5-27}$$

Ellipsoidal refinements of these formulas may be found in Rinner (1956).

The reason for the great difference between the reduction procedures for base lines and for electronically measured distances is that base lines may be considered as measured along the earth's surface and piecewise reduced to the local horizon, which involves the direction of the vertical, whereas straight spatial distances are independent of the vertical. Therefore, the reduction formula (5-26) does not contain the deflection of the vertical ϵ.

5-6. Reduction of the Astronomical Coordinates for the Curvature of the Plumb Line

The astronomical coordinates Φ and Λ, as observed on the surface of the earth, are not rigorously equal to their corresponding values at the geoid because the plumb line, the line of force, is not straight, or in other words, because the level surfaces are not parallel. Thus if we wish our astronomical coordinates to refer to the geoid, we must reduce our observations accordingly.

Helmert's projection in principle avoids the reduction for plumb-line curvature because it does not use the geoid directly, but we still need this reduction if we want to use or to obtain quantities that are referred to the geoid. Examples of such cases are:

1. The gravimetric deflections are usually computed by Vening Meinesz' formula for the geoid, so that either the gravimetric deflections must be reduced upward to the ground point or the astronomical observations must be reduced downward to the geoid, in order to make the two quantities comparable.
2. If astronomical observations are used for the determination of the geoid, the same reduction must be applied, as explained in the next section.

Consider the projection of the plumb line onto the meridian plane. According to the well-known definition of the curvature of a plane curve, the angle between two neighboring tangents of this projection of the plumb line is

$$d\phi = -\kappa_1\, dh,$$

where the minus sign is conventional and the curvature κ_1 is given by (2-22a):

$$\kappa_1 = \frac{1}{g}\frac{\partial g}{\partial x}.$$

The x-axis is horizontal and points northward. Hence the total change of latitude along the plumb line between a point on the ground, P, and its projection onto the geoid, P_0, is given by

$$\delta\phi = \int_{P_0}^{P} d\phi = -\int_{P_0}^{P} \kappa_1\, dh$$

or

$$\delta\phi = -\int_{P_0}^{P} \frac{1}{g}\frac{\partial g}{\partial x}\,dh. \tag{5-28a}$$

Similarly we find for the change of longitude, κ_2 (2-22b) replacing κ_1,

$$\delta\lambda \cos\phi = -\int_{P_0}^{P} \frac{1}{g}\frac{\partial g}{\partial y}\,dh, \tag{5-28b}$$

where the y-axis is horizontal and points eastwards.

Alternative formulas. There is a close relationship between the curvature reduction of astronomical coordinates and the orthometric reduction of leveling, considered in Sec. 4.4.

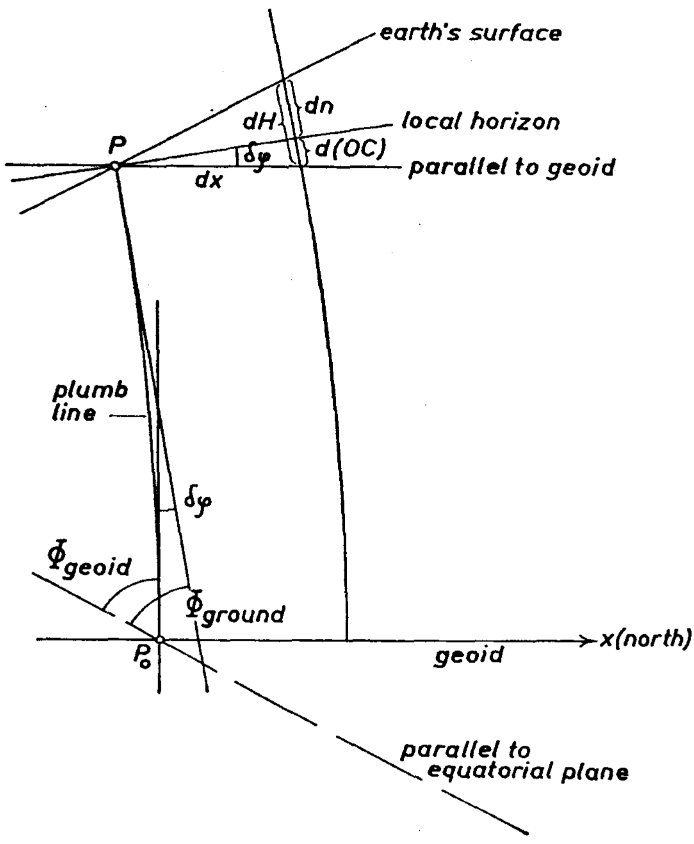

FIGURE 5-1

Plumb-line curvature and orthometric correction.

The orthometric correction $d(OC)$ has been defined as the quantity that must be added to the leveling increment dn in order to convert it into the orthometric height difference dH:

$$d(OC) = dH - dn. \tag{5-29}$$

From Fig. 5-8 we see that, for a north-south profile, the curvature reduction and the orthometric correction are related by the simple formula

$$\delta\phi = \frac{\partial(OC)}{\partial x}. \tag{5-30a}$$

Similarly we find

$$\delta\lambda \cos\phi = \frac{\partial(OC)}{\partial y}. \tag{5-30b}$$

According to Sec. 4-4, we have

$$dC = g\,dn = -dW, \qquad H = \frac{C}{\bar{g}}.$$

Hence (5-29) becomes

$$d(OC) = dH - \frac{1}{g}dC = dH + \frac{1}{g}dW,$$

so that

$$\begin{aligned} \delta\phi &= \frac{\partial H}{\partial x} + \frac{1}{g}\frac{\partial W}{\partial x}, \\ \delta\lambda\cos\phi &= \frac{\partial H}{\partial y} + \frac{1}{g}\frac{\partial W}{\partial y}. \end{aligned} \tag{5-31}$$

These equations relate the reduction for the curvature of the plumb line to the orthometric height H and the potential W. In view of the irregular shape of the plumb lines it is remarkable that such simple general relations as (5-30) and (5-31) exist.

These relations may be used to find computational formulas for the curvature reductions $\delta\phi$ and $\delta\lambda$ (Bodemüller, 1957). We have

$$\begin{aligned} d(OC) &= dH - \frac{dC}{g} = d\left(\frac{C}{\bar{g}}\right) - \frac{dC}{g} \\ &= \frac{dC}{\bar{g}} - \frac{C}{\bar{g}^2}d\bar{g} - \frac{dC}{g} = -\frac{C}{\bar{g}^2}d\bar{g} + \frac{g-\bar{g}}{\bar{g}}\frac{dC}{g} \end{aligned}$$

or

$$d(OC) = -\frac{H}{\bar{g}}d\bar{g} + \frac{g-\bar{g}}{\bar{g}}dn.$$

By substituting this into (5-30a, b) we obtain

$$\begin{aligned} \delta\phi &= -\frac{H}{\bar{g}}\frac{\partial\bar{g}}{\partial x} + \frac{g-\bar{g}}{\bar{g}}\tan\beta_1, \\ \delta\lambda\cos\phi &= -\frac{H}{\bar{g}}\frac{\partial\bar{g}}{\partial y} + \frac{g-\bar{g}}{\bar{g}}\tan\beta_2, \end{aligned} \tag{5-32}$$

where we have set

$$\tan \beta_1 = \frac{\partial n}{\partial x}, \quad \tan \beta_2 = \frac{\partial n}{\partial y},$$

so that β_1 and β_2 are the angles of inclination of the north-south and east-west profiles with respect to the local horizon; $\bar{g}$ is the mean value of gravity between the geoid and the ground. In these formulas we need only this mean value $\bar{g}$, together with its horizontal derivatives, and the ground value g, whereas in (5-28) we must know the horizontal derivatives of gravity all along the plumb line. The detailed shape of the plumb lines does not directly enter into (5-32) as it does into (5-28).

The mean value $\bar{g}$ is found by a Prey reduction of the measured gravity g. In order that the numerical differentiations $\partial\bar{g}/\partial x$ and $\partial\bar{g}/\partial y$ give reliable results, a dense gravity net around the station is necessary, and the Prey reduction must be performed carefully. The inclination angles β_1 and β_2 are taken from a topographical map.

The sign of these corrections may be found in the following way. If g decreases in the x-direction, then formulas (5-28) and (5-32) give $\delta\phi > 0$ and Fig. 5-8 shows that Φ at P_0 is then greater than at P. The same holds for Λ, so that we have

$$\begin{aligned} \Phi_{\text{geoid}} &= \Phi_{\text{ground}} + \delta\phi, \\ \Lambda_{\text{geoid}} &= \Lambda_{\text{ground}} + \delta\lambda. \end{aligned} \tag{5-33}$$

For other methods for determining the curvature of the plumb line see Arnold (1956, Sec. C) and Ledersteger (1955).

Curvature of the normal plumb line. If, instead of the actual gravity g, the normal gravity γ is used for the computation of the plumb line curvature, we find, using

$$\gamma = \gamma_a\left(1 + f^* \sin^2\phi - \frac{2}{a}h \cdots\right),$$

that

$$\frac{\partial\gamma}{\partial x} \doteq \frac{1}{R}\frac{\partial\gamma}{\partial\phi} \doteq \frac{2\gamma_a}{R}f^* \sin\phi\cos\phi \doteq \frac{2\gamma}{R}f^* \sin\phi\cos\phi,$$

$$\frac{\partial\gamma}{\partial y} \doteq \frac{1}{R\cos\phi}\frac{\partial\gamma}{\partial\lambda} = 0.$$

Hence the integrand $(1/\gamma)(\partial\gamma/\partial x)$ in (5-28a) does not depend on h, so that the integration can be performed immediately. We find

$$\begin{aligned} \delta\phi_{\text{normal}} &= -\frac{f^*}{R}h\sin 2\phi = -0.17'' \, h_{\text{km}} \sin 2\phi, \\ \delta\lambda_{\text{normal}} &= 0. \end{aligned} \tag{5-34}$$

The curvature of the normal plumb line in the east-west direction is zero, owing to the rotational symmetry of the ellipsoid of revolution.

The normal reduction (5-34) is often conventionally applied, but this is of little use, for the effect of topographic irregularities on the curvature of the plumb line is often much greater than the "normal" part. In high mountains, the actual reduction may amount to several seconds of arc (Kobold and Hunziker, 1962).

For a rigorous use of the normal reduction (5-34) see Sec. 8-9.

5-7. The Astrogeodetic Determination of the Geoid

The shape of the geoid can be determined if the deflections of the vertical are given. The basic equation is (2-202):

$$dN = -\epsilon\, ds. \tag{5-35}$$

On integrating we get

$$N_B = N_A - \int_A^B \epsilon\, ds, \tag{5-36}$$

where

$$\epsilon = \xi \cos\alpha + \eta \sin\alpha$$

is the component of the deflection of the vertical along the profile AB, whose azimuth is α; see equation (5-16).

This formula expresses the geoidal undulation as an integral of the vertical deflections along a profile. Since N is a function of position, this integral is independent of the form of the line that connects the points A and B. This line need not necessarily be a geodesic on the ellipsoid, and α may in the general case be variable. In practice, north-south profiles ($\epsilon = \xi$) or east-west profiles ($\epsilon = \eta$) are often used. The integral (5-36) is to be evaluated by a numerical or graphical integration. The deflection component ϵ must be given at enough stations along the profile such that the interpolation between these stations can be done reliably. Sometimes a map of ξ and η is available for a certain area. Such a map is constructed by interpolation between well-distributed stations at which ξ and η have been determined. Then the profiles of integration may be suitably selected; loops may be formed to obtain redundancies which must be adjusted.

If the deflection components ξ and η are obtained directly from the equations

$$\xi = \Phi - \phi, \quad \eta = (\Lambda - \lambda)\cos\phi, \tag{5-37}$$

that is, by comparing the astronomic and geodetic coordinates of the same point, then this method is called the *astrogeodetic determination of the geoid.*

The astronomical coordinates are directly observed; the geodetic coordinates are obtained in the following way. In a larger triangulation system a certain "initial point" P_1 is chosen for which the undulation N_1 and the components ξ_1 and η_1 of the deflection of the vertical are prescribed. Here ξ_1, η_1, and N_1 may be assumed arbitrarily in principle; the position of the reference ellipsoid with respect to the earth is thereby fixed. For the sake of definiteness let us consider the case that is of greatest practical importance, that is, the case in which $\xi_1 = \eta_1 = N_1 = 0$. In this case, because $\xi_1 = \eta_1 = 0$, the geoid and the ellipsoid have the same sur-

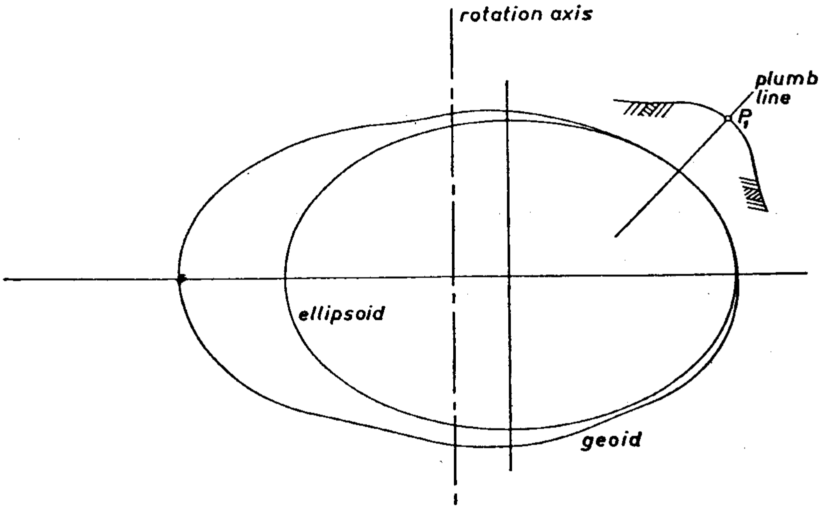

FIGURE 5-9
The reference ellipsoid is tangent to the geoid at P_1.

face normal,[1] so that, because $N_1 = 0$, the ellipsoid is tangent to the geoid below P_1 (Fig. 5-9). The condition that the axis of the reference ellipsoid be parallel to the earth's axis of rotation finally determines the orientation of the triangulation net because Laplace's equation (5-14) then gives $\Delta\alpha_1 = \eta_1 \tan \phi_1 = 0$, so that $\alpha_1 = A_1$; that is, at the initial point the geodetic azimuth is equal to the astronomical azimuth.

Now we can reduce the measured distances and angles to the ellipsoid and compute on it the position of the points of the triangulation net (their geodetic coordinates ϕ and λ) in the usual way. After measuring the coordinates Φ and Λ astronomically at the same points, we can then compute the deflection components ξ and η by (5-37). Starting from the assumed value N_1 at the initial point P_1 (in our case $N_1 = 0$) we can finally compute the geoidal heights N of any point of the triangulation net by repeated application of (5-36). These geoidal heights refer to the ellipsoid that was fixed by prescribing ξ_1, η_1, N_1, and, of course, its semimajor axis a and its flattening f. To employ a frequently used term, they refer to the given *astrogeodetic datum* $(a, f; \xi_1, \eta_1, N_1)$.

By means of N and the orthometric height H, the height h above the ellipsoid is obtained ($h = H + N$), so that the rectangular spatial coordinates X, Y, Z can be computed by (5-5). But unless ξ and η are absolute deflections, the origin of the coordinate system will not be at the center of the earth; see Sec. 5-9.

A flaw in the procedure described above apparently is that N, ξ, η are already

[1] We disregard the curvature of the plumb line.

needed for the reduction of the measured angles and distances to the ellipsoid. However, for this purpose approximate values of N, ξ, η are sufficient. These are obtained by performing the process just explained with unreduced angles and distances. We can also get suitable values for N, ξ, η in other ways, for instance by Stokes' formula.

It should be mentioned that in practice the component η is often obtained from azimuth measurements by (5-18),

$$\eta = (A - \alpha) \cot \phi, \tag{5-38}$$

because astronomical measurements of azimuth are simpler than those of longitude. Moreover, longitude and azimuth are often measured at the same point. Then Laplace's condition

$$\Delta\alpha = \Delta\lambda \sin \phi$$

furnishes an important check on the correct orientation of the net and may be used for adjustment purposes. Astronomical stations with longitude and azimuth observations are therefore called *Laplace stations.*

The astrogeodetic determination of the geoid was known to Helmert (1880); it is also called *astronomical leveling.*

Comparison with the Stokes method. It is quite instructive to compare Helmert's formula

$$N = N_A - \int_A^B \epsilon \, ds$$

for the astrogeodetic method with Stokes' formula

$$N = \frac{R}{4\pi G} \iint_\sigma \Delta g \, S(\psi) \, d\sigma$$

for the gravimetric method. Both methods use the gravity vector **g**. It is compared with a normal gravity vector $\boldsymbol{\gamma}$. The components $\xi = \Delta\phi$ and $\eta = \Delta\lambda \cos \phi$ of the deflection of the vertical represent the differences in *direction*, and the gravity anomaly Δg represents the difference in *magnitude* of the two vectors. Helmert's formula determines the geoidal undulation N from ξ and η, that is, by means of the direction of **g**, and Stokes' formula determines N from Δg, that is, by means of the magnitude of **g**. Both formulas are somewhat similar: they are integrals which contain ϵ, or ξ and η, and Δg in a linear form.

Otherwise the two formulas show marked differences, which are characteristic for the respective methods. In Helmert's formula the integration is extended over part of a profile; thus it is sufficient to know the deflection of the vertical in a limited area. The position of the reference ellipsoid with respect to the earth's center of gravity is unknown, however, and can be determined only by means of the gravimetric method (Sec. 5-10) or the analysis of satellite orbits (Sec. 9-8). Furthermore, the astrogeodetic method can be used only on land, because the necessary measurements are impossible at sea.

In Stokes' formula, however, the integration should be extended over the whole earth. The gravity anomaly Δg must be known all over the earth; however, accurate gravity measurements at sea are possible. The gravimetric method yields, for the whole earth, *absolute* geoidal undulations, the center of the reference ellipsoid coinciding with the center of the earth.

Thus, of the classical geodetic methods, only the gravimetric method makes possible a world-wide geodetic system. The astrogeodetic method is necessary—for instance, to furnish the scale. Thus both methods must be combined, supplemented by such geodetic information as can be obtained in other ways, particularly that obtained from artificial satellites; see Chapter 9.

Correction for the curvature of the plumb line. In formula (5-36), the deflection components ξ and η refer to the geoid. This means that the astronomical observations of Φ and Λ must be reduced to the geoid according to Sec. 5-6.

It is also possible and often more convenient to apply this correction for plumb-line curvature not to the astronomical coordinates Φ and Λ but to the geoidal height differences computed from the unreduced deflection components (Helmert, 1900 and 1901).

These N values, denoted by N', are obtained by using in (5-37) the directly observed Φ and Λ, which define the direction of the plumb line at the station P in Fig. 5-10. The notation N will be reserved for the correct geoidal heights. Then we read from Fig. 5-10:

$$dh = dN + dH = dN' + dn,$$

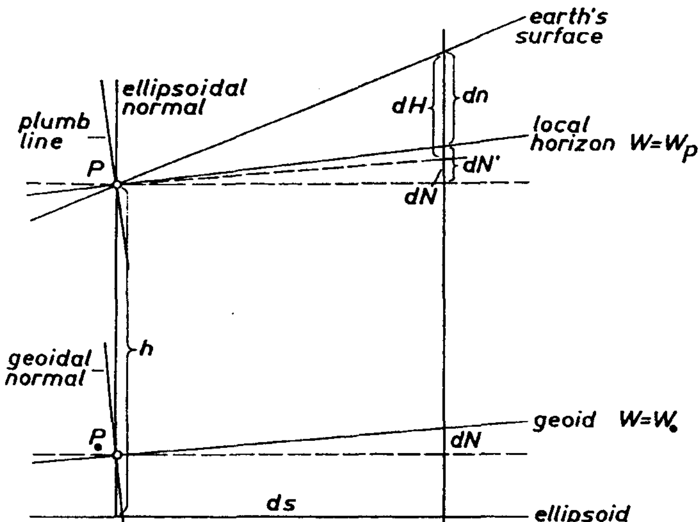

FIGURE 5-10

Reduction of astronomical leveling.

where h is the geometric height above the ellipsoid. Thus we see that the difference between the unreduced and the correct element of geoidal height,

$$dN' - dN = dH - dn = d(OC), \tag{5-39}$$

is equal to the difference between the element dH of orthometric height and the leveling increment dn, which is the orthometric reduction $d(OC)$.

Thus

$$N_B - N_A = N'_B - N'_A - OC_{AB}, \tag{5-40}$$

so that we can immediately apply equation (4-33) of the preceding chapter:

$$N_B - N_A = -\int_A^B \epsilon\, ds - \int_A^B \frac{g - \gamma_0}{\gamma_0}\, dn + \frac{\bar{g}_B - \gamma_0}{\gamma_0} H_B - \frac{\bar{g}_A - \gamma_0}{\gamma_0} H_A, \tag{5-41}$$

where γ_0 is an arbitrary constant value that can be chosen conveniently; the deflection components ϵ are computed from the observed ground values Φ and Λ by (5-37) and (5-16).

The astrogeodetic method has often been applied to the determination of geoidal sections; see, for instance, Bomford (1963), Fischer (1961), Galle (1914), Niethammer (1939), Ölander (1951), Rice (1962), and Wolf (1956). A discussion of the method's practical aspects and accuracy will be found in Bomford (1962, Chap. 5, Sec. 5).

5-8. Interpolation of Deflections of the Vertical. Astrogravimetric Leveling

Helmert's formula (5-36) for the astronomical leveling presupposes that the stations at which the deflections of the vertical are known are very close to one another. Thus a profile for ϵ can be constructed by interpolation, and the integration in (5-36) can be performed numerically or graphically.

If for A and B in (5-36) we take two neighboring astrogeodetic stations, and if they are so close together that the geoidal profile between them can be approximated by the arc of a circle, then this formula becomes

$$N_B - N_A = -\frac{\epsilon_A + \epsilon_B}{2} s, \tag{5-42}$$

where s is the distance between A and B. In this way the interpolation can be avoided; but this is only apparent, since the assumption that the geoid between A and B forms a circular arc is itself equivalent to an interpolation, and not necessarily the best one.

In moderately level areas a station distance of 25 km, say, and the approximation (5-42) are usually satisfactory; but in high mountains a spacing of even 10 km or less may not be sufficient.

Since astronomical observations are time-consuming, more efficient means for interpolation between the astrogeodetic stations have been devised. Such methods are:

measurement of zenith distances;
use of the torsion balance;
astrogravimetric leveling;
use of topographic-isostatic deflections.

We shall now discuss some aspects of these methods.

Zenith distances. Measurements of zenith distances can, at least theoretically, be used to replace astronomical observations (de Graaf-Hunter, 1913).

The principle has already been described in Sec. 4-7. The basic equation is (4-57):

$$\epsilon_2 - \epsilon_1 = z_1' + z_2' - \gamma - 180°, \tag{5-43}$$

where z_1' and z_2' are the measured zenith distances which have been corrected for the effect of atmospheric refraction. The angle γ is given by

$$\gamma = \frac{s}{R}, \tag{5-44}$$

where s is the ellipsoidal distance between the stations 1 and 2, and R is a mean radius of curvature along the arc s. The distance s is obtained by triangulation or trilateration.

The difficulty with this method is, of course, the proper allowance for atmospheric refraction. At present, therefore, its use is limited to high mountains. This method is being applied successfully in the Swiss alps, where deflection differences have been obtained with an accuracy of $\pm 1''$ (Kobold, 1951).

Torsion balance measurements. The torsion balance is an instrument that measures certain combinations

$$\frac{\partial^2 W}{\partial y^2} - \frac{\partial^2 W}{\partial x^2}, \quad \frac{\partial^2 W}{\partial x\,\partial y}, \quad \frac{\partial^2 W}{\partial x\,\partial z}, \quad \frac{\partial^2 W}{\partial y\,\partial z}$$

of the second partial derivatives of the gravity potential with respect to a rectangular coordinate system having a vertical z-axis.

Let us take the x-axis northward and consider the quantity

$$\frac{\partial^2 W}{\partial x\,\partial y}$$

at the geoid. Since the normal potential U is constant along the ellipsoid, and consequently

$$\frac{\partial^2 U}{\partial x\,\partial y} = 0,$$

the xy-plane being tangent to the ellipsoid, we have

$$\frac{\partial^2 T}{\partial x\,\partial y} = \frac{\partial^2 W}{\partial x\,\partial y}.$$

Making use of the basic relations

$$\xi = -\frac{1}{G}\frac{\partial T}{\partial x}, \quad \eta = -\frac{1}{G}\frac{\partial T}{\partial y},$$

where G is a mean value of gravity, we obtain

$$\frac{\partial \xi}{\partial y} = \frac{\partial \eta}{\partial x} = -\frac{1}{G}\frac{\partial^2 W}{\partial x\, \partial y} \tag{5-45}$$

from torsion balance measurements.

We thus know certain horizontal derivatives of the components of the deflection of the vertical. It is clear that we can get differences $\xi_2 - \xi_1$ and $\eta_2 - \eta_1$ of the deflection components by a suitable integration of (5-45). The details are somewhat involved; the reader will find descriptions in Baeschlin (1948) and in Mueller (1963).

This method is very sensitive to topographic irregularities, and the measurement is time-consuming. It is seldom used nowadays, but it is perhaps unduly neglected. Apart from its great theoretical interest it may be of practical importance in level areas where a detailed gravity survey, necessary for astrogravimetric leveling, does not exist or is not feasible—for instance, along coast lines.

Astrogravimetric leveling. If in Vening Meinesz' formula the integration is not extended over the whole earth but only over a neighborhood of the point considered, then an error is introduced because the distant zones are neglected. This error, however, is almost the same for points that are not too far apart, and varies only slowly over the points of a short profile, so that the gravimetric deflections computed in this way can be used for interpolation between astrogeodetic deflections.

From ξ' and η', obtained gravimetrically, the components ϵ' are computed in the usual way:

$$\epsilon' = \xi' \cos\alpha + \eta' \sin\alpha. \tag{5-46}$$

The differences

$$\delta\epsilon = \epsilon - \epsilon' \tag{5-47}$$

between the "correct" astrogeodetic deflections ϵ and the approximate gravimetric values ϵ' vary only slowly and may be assumed to change linearly with distance, so that they can be computed by a linear interpolation

$$\delta\epsilon_P = \delta\epsilon_A + \frac{\delta\epsilon_B - \delta\epsilon_A}{s_{AB}} s_{AP}, \tag{5-48}$$

where P is any point on the profile between the astronomical stations A and B, and s is the distance between the points corresponding to the subscripts.

The procedure is thus as follows. At A and B the astronomical deflections ϵ_A and ϵ_B are given. Compute at these points and at the intermediate points $P_1, P_2, \ldots, P_n$ the gravimetric deflections $\epsilon'_A, \epsilon'_B, \epsilon'_1, \epsilon'_2, \ldots, \epsilon'_n$ and interpolate $\delta\epsilon_i$ at the intermediate points by (5-48). Then the desired deflections of the vertical ϵ at the intermediate points, referred to the astrogeodetic datum, are computed by

$$\epsilon_i = \epsilon_i' + \delta\epsilon_i. \tag{5-49}$$

This combination of astrogeodetic deflections with gravimetrically interpolated values is called *astrogravimetric leveling* (Molodenskii et al., 1962, Chap. 6). It is considered to be the best method of interpolation. If this method is used, then the astrogeodetic stations may be as far apart as 100 to 200 km in level country, but then a sufficiently dense gravity net must extend to at least twice the distance between two stations.

Astrogravimetric leveling shows the great flexibility of the gravimetric method. The Vening Meinesz formula can be applied for two completely different purposes: if we are integrating over the whole earth, it gives absolute deflections of the vertical, thus providing the absolute orientation of the astrogeodetic systems; if we are integrating over a limited area, it helps to interpolate between the relative astrogeodetic deflections.

Use of topographic-isostatic deflections. In (5-49) the vertical deflections ϵ' can also be computed from the effect of topography (Helmert, 1900 and 1901). This method may be refined by considering the effect of the isostatic compensation as well. No gravity information is needed here. This method has been applied successfully for interpolation between Alpine astrogeodetic stations that are not too far apart (Niethammer, 1939). However, it is affected by unknown density anomalies, etc., and is rather laborious. Hence astrogravimetric leveling will be preferred when distances between astrogeodetic stations are large.

5-9. Coordinate Transformations and Datum Shifts

As we have established in Sec. 5-7, a geodetic datum is determined by the dimensions of the reference ellipsoid (semimajor axis a and flattening f) and its position with respect to the earth or the geoid. This relative position is usually given by the geoidal undulation N_1 and the components ξ_1 and η_1 of the deflection of the vertical at an initial point P_1. Instead of ξ_1, η_1, N_1 we might as well use the geodetic coordinates ϕ_1, λ_1, h_1 of P_1 because

$$\begin{aligned} \xi_1 &= \Phi_1 - \phi_1, \\ \eta_1 &= (\Lambda_1 - \lambda_1)\cos\phi, \\ N_1 &= h_1 - H_1. \end{aligned} \tag{5-50}$$

A superficially different but equivalent method is to use the rectangular coordinates x_0, y_0, z_0 of the center of the reference ellipsoid with respect to the center of the earth.

If we vary the geodetic datum—that is, the reference ellipsoid and its position—then the geodetic coordinates ϕ, λ, h and consequently the deflections of the vertical and the undulations of the geoid,

$$\begin{aligned} \xi &= \Phi - \phi, \\ \eta &= (\Lambda - \lambda)\cos\phi, \\ N &= h - H, \end{aligned} \tag{5-51}$$

will also change. Since there are three different ways of fixing the datum, we can formulate these changes in terms of the variation of

$$\xi_1, \eta_1, N_1, \quad \text{or} \quad \phi_1, \lambda_1, h_1, \quad \text{or} \quad x_0, y_0, z_0.$$

Mathematically, the problem is simply a transformation of coordinates, since every geodetic datum corresponds to a different system of geodetic coordinates ϕ, λ, h.

Suppose that the center of the reference ellipsoid does not coincide with the earth's center of gravity, but that the axis of the ellipsoid is parallel to the earth's axis of rotation. Assume a rectangular coordinate system XYZ whose origin is the earth's center of gravity (not the center of the ellipsoid, as before), the axes being directed as usual. Let the coordinates of the center of the ellipsoid with respect to this system be x_0, y_0, z_0, as stated above. Then equations (5-5) must obviously be modified so that they become

$$\begin{aligned} X &= x_0 + (N + h)\cos\phi\cos\lambda, \\ Y &= y_0 + (N + h)\cos\phi\sin\lambda, \\ Z &= z_0 + \left(\frac{b^2}{a^2}N + h\right)\sin\phi. \end{aligned} \tag{5-52}$$

These equations form the starting point for various important differential formulas of coordinate transformation.

First we ask how the rectangular coordinates X, Y, Z change if we vary the geodetic coordinates ϕ, λ, h by small amounts $\delta\phi$, $\delta\lambda$, δh and if we also alter the geodetic datum, namely the reference ellipsoid (a, f) and its position (x_0, y_0, z_0), by δa, δf and δx_0, δy_0, δz_0. Note that δx_0, δy_0, δz_0 correspond to a small translation (parallel displacement) of the ellipsoid, *its axis remaining parallel to the axis of the earth.*

The solution of this problem is found by differentiating (5-52):

$$\begin{aligned} \delta X &= \delta x_0 + \frac{\partial X}{\partial a}\delta a + \frac{\partial X}{\partial f}\delta f + \frac{\partial X}{\partial \phi}\delta\phi + \frac{\partial X}{\partial \lambda}\delta\lambda + \frac{\partial X}{\partial h}\delta h, \\ \delta Y &= \delta y_0 + \frac{\partial Y}{\partial a}\delta a + \frac{\partial Y}{\partial f}\delta f + \frac{\partial Y}{\partial \phi}\delta\phi + \frac{\partial Y}{\partial \lambda}\delta\lambda + \frac{\partial Y}{\partial h}\delta h, \\ \delta Z &= \delta z_0 + \frac{\partial Z}{\partial a}\delta a + \frac{\partial Z}{\partial f}\delta f + \frac{\partial Z}{\partial \phi}\delta\phi + \frac{\partial Z}{\partial \lambda}\delta\lambda + \frac{\partial Z}{\partial h}\delta h, \end{aligned} \tag{5-53}$$

since, according to Taylor's theorem, small changes can be treated as differentials.

In these differential formulas we shall be satisfied with an approximation. Since the flattening f is small, we may expand (2-81) as

$$N = \frac{a^2}{b}(1 + e'^2 \cos^2 \phi)^{-1/2} = \frac{a^2}{b}\left(1 - \frac{1}{2}e'^2 \cos^2 \phi \cdots\right)$$
$$= a(1 + f \cdots)(1 - f\cos^2 \phi \cdots) = a(1 + f - f\cos^2 \phi \cdots);$$
$$N \doteq a(1 + f \sin^2 \phi);$$

and

$$\frac{b^2}{a^2} N = (1 - 2f \cdots)a(1 + f\sin^2 \phi \cdots) \doteq a(1 - 2f + f\sin^2 \phi),$$

since

$$b = a(1 - f), \quad e'^2 = 2f \cdots.$$

Thus equations (5-52) are approximated by

$$\begin{aligned} X &= x_0 + (a + af\sin^2 \phi + h)\cos \phi \cos \lambda, \\ Y &= y_0 + (a + af\sin^2 \phi + h)\cos \phi \sin \lambda, \\ Z &= z_0 + (a - 2af + af\sin^2 \phi + h)\sin \phi. \end{aligned} \tag{5-52'}$$

Now we can form the partial derivatives in (5-53), for instance

$$\frac{\partial X}{\partial a} = (1 + f\sin^2 \phi)\cos \phi \cos \lambda \doteq \cos \phi \cos \lambda,$$

since we may neglect the flattening in these coefficients. This amounts to using for the coefficients, and only for them, a spherical approximation analogous to that of Sec. 2-14. Similarly, all coefficients are easily obtained as partial derivatives, and equations (5-53) become

$$\delta X = \delta x_0 - a \sin \phi \cos \lambda\, \delta\phi - a \cos \phi \sin \lambda\, \delta\lambda$$
$$+ \cos \phi \cos \lambda(\delta h + \delta a + a \sin^2 \phi\, \delta f), \tag{5-54a}$$

$$\delta Y = \delta y_0 - a \sin \phi \sin \lambda\, \delta\phi + a \cos \phi \cos \lambda\, \delta\lambda$$
$$+ \cos \phi \sin \lambda(\delta h + \delta a + a \sin^2 \phi\, \delta f), \tag{5-54b}$$

$$\delta Z = \delta z_0 + a \cos \phi\, \delta\phi + \sin \phi(\delta h + \delta a + a \sin^2 \phi\, \delta f) - 2a \sin \phi\, \delta f. \tag{5-54c}$$

These formulas give the changes in the rectangular coordinates X, Y, Z in terms of the variation in the position (x_0, y_0, z_0) and the dimensions (a, f) of the ellipsoid and in the geodetic coordinates ϕ, λ, h referred to it.

Transformation of the geodetic coordinates. Several important formulas for the transformation of coordinates may be derived from equations (5-54). First, let the position of P in space remain unchanged; that is, let

$$\delta X = \delta Y = \delta Z = 0.$$

Determine the change of the geodetic coordinates ϕ, λ, h if the dimensions of the reference ellipsoid and its position are varied.

The problem is thus to solve equations (5-54) for $\delta\phi$, $\delta\lambda$, δh, the left-hand sides being set equal to zero. To get $\delta\phi$ multiply (5-54a) by $-\sin \phi \cos \lambda$, (5-54b) by $-\sin\phi\sin\lambda$, and (5-54c) by $\cos\phi$, and add all equations obtained in this way. For $\delta\lambda$ the factors are $-\sin\lambda$, $\cos\lambda$, and 0; for δh they are $\cos\phi\cos\lambda$, $\cos\phi\sin\lambda$, and $\sin\phi$. The result is

$$\begin{aligned}
a\,\delta\phi &= \sin\phi\cos\lambda\,\delta x_0 + \sin\phi\sin\lambda\,\delta y_0 - \cos\phi\,\delta z_0 + 2a\sin\phi\cos\phi\,\delta f,\\
a\cos\phi\,\delta\lambda &= \sin\lambda\,\delta x_0 - \cos\lambda\,\delta y_0, \qquad (5\text{-}55)\\
\delta h &= -\cos\phi\cos\lambda\,\delta x_0 - \cos\phi\sin\lambda\,\delta y_0 - \sin\phi\,\delta z_0 - \delta a + a\sin^2\phi\,\delta f.
\end{aligned}$$

We have seen that the translation of the ellipsoid may also be given in terms of the changes in the geodetic coordinates $\delta\phi_1$, $\delta\lambda_1$, δh_1 of an initial point, instead of δx_0, δy_0, δz_0. The problem is then to determine the variations $\delta\phi$, $\delta\lambda$, δh at the other points.

First we express the parallel displacement (δx_0, δy_0, δz_0) of the ellipsoid in terms of the given $\delta\phi_1$, $\delta\lambda_1$, δh_1. In equations (5-54), set $\delta X = \delta Y = \delta Z = 0$ (again because the position of the points in space remains unchanged) and $\phi = \phi_1$, $\lambda = \lambda_1$, $h = h_1$. Then we get

$$\begin{aligned}
\delta x_0 &= a\sin\phi_1\cos\lambda_1\,\delta\phi_1 + a\cos\phi_1\sin\lambda_1\,\delta\lambda_1\\
&\qquad - \cos\phi_1\cos\lambda_1(\delta h_1 + \delta a + a\sin^2\phi_1\,\delta f),\\
\delta y_0 &= a\sin\phi_1\sin\lambda_1\,\delta\phi_1 - a\cos\phi_1\cos\lambda_1\,\delta\lambda_1\\
&\qquad - \cos\phi_1\sin\lambda_1(\delta h_1 + \delta a + a\sin^2\phi_1\,\delta f), \qquad (5\text{-}56)\\
\delta z_0 &= -a\cos\phi_1\,\delta\phi_1 - \sin\phi_1(\delta h_1 + \delta a + a\sin^2\phi_1\,\delta f) + 2a\sin\phi_1\,\delta f.
\end{aligned}$$

These expressions for the shift components δx_0, δy_0, δz_0 are inserted into equations (5-55), so that we finally obtain:

$$\begin{aligned}
\delta\phi &= (\cos\phi_1\cos\phi + \sin\phi_1\sin\phi\cos\Delta\lambda)\,\delta\phi_1 - \sin\phi\sin\Delta\lambda\cdot\cos\phi_1\,\delta\lambda_1\\
&\qquad + (\sin\phi_1\cos\phi - \cos\phi_1\sin\phi\cos\Delta\lambda)\left(\frac{\delta h_1}{a} + \frac{\delta a}{a} + \sin^2\phi_1\,\delta f\right)\\
&\qquad + 2\cos\phi(\sin\phi - \sin\phi_1)\,\delta f,\\
\cos\phi\,\delta\lambda &= \sin\phi_1\sin\Delta\lambda\,\delta\phi_1 + \cos\Delta\lambda\cdot\cos\phi_1\,\delta\lambda_1\\
&\qquad - \cos\phi_1\sin\Delta\lambda\left(\frac{\delta h_1}{a} + \frac{\delta a}{a} + \sin^2\phi_1\,\delta f\right), \qquad (5\text{-}57)\\
\frac{\delta h}{a} &= (\cos\phi_1\sin\phi - \sin\phi_1\cos\phi\cos\Delta\lambda)\,\delta\phi_1 + \cos\phi\sin\Delta\lambda\cdot\cos\phi_1\,\delta\lambda_1\\
&\qquad + (\sin\phi_1\sin\phi + \cos\phi_1\cos\phi\cos\Delta\lambda)\left(\frac{\delta h_1}{a} + \frac{\delta a}{a} + \sin^2\phi_1\,\delta f\right)\\
&\qquad - \frac{\delta a}{a} + (\sin^2\phi - 2\sin\phi_1\sin\phi)\,\delta f,
\end{aligned}$$

where

$$\Delta\lambda = \lambda - \lambda_1.$$

These formulas express the variations $\delta\phi$, $\delta\lambda$, δh at an arbitrary point in terms of the variations $\delta\phi_1$, $\delta\lambda_1$, δh_1 at a given point and the changes δa and δf of the parameters of the reference ellipsoid. They thus relate two different systems of geodetic coordinates, provided these systems are so close to each other that their differences may be considered as linear. Mathematically, equations (5-57) are infinitesimal coordinate transformations; to the geodesist, they give the effect of a change in the geodetic datum. They are equivalent to equations (5-55). Both (5-55) and (5-57) are infinitesimal transformations of geodetic coordinates; they differ only in the parameters used for determining the coordinate system, the geodetic datum; in (5-55) the coordinate system is defined by $(a, f;\, x_0, y_0, z_0)$ and in (5-57) by $(a, f;\, \phi_1, \lambda_1, h_1)$.

Transformation of ξ, η, N. Usually, equations (5-57) are expressed in terms of the variations of the deflection components ξ and η and of the geoidal undulation N. Since the natural coordinates Φ, Λ, H are not affected by a datum shift and remain unchanged, we get from (5-51)

$$\begin{aligned} \delta\phi &= -\delta\xi, \\ \delta\lambda \cos\phi &= -\delta\eta, \\ \delta h &= \delta N, \end{aligned} \tag{5-58}$$

so that equations (5-57) assume the form

$$\begin{aligned}
\delta\xi &= (\cos\phi_1 \cos\phi + \sin\phi_1 \sin\phi \cos\Delta\lambda)\,\delta\xi_1 - \sin\phi \sin\Delta\lambda\,\delta\eta_1 \\
&\qquad - (\sin\phi_1 \cos\phi - \cos\phi_1 \sin\phi \cos\Delta\lambda)\left(\frac{\delta N_1}{a} + \frac{\delta a}{a} + \sin^2\phi_1\,\delta f\right) \\
&\qquad - 2\cos\phi(\sin\phi - \sin\phi_1)\,\delta f, \\
\delta\eta &= \sin\phi_1 \sin\Delta\lambda\,\delta\xi_1 + \cos\Delta\lambda\,\delta\eta_1 \\
&\qquad + \cos\phi_1 \sin\Delta\lambda\left(\frac{\delta N_1}{a} + \frac{\delta a}{a} + \sin^2\phi_1\,\delta f\right), \\
\frac{\delta N}{a} &= -(\cos\phi_1 \sin\phi - \sin\phi_1 \cos\phi \cos\Delta\lambda)\,\delta\xi_1 - \cos\phi \sin\Delta\lambda\,\delta\eta_1 \\
&\qquad + (\sin\phi_1 \sin\phi + \cos\phi_1 \cos\phi \cos\Delta\lambda)\left(\frac{\delta N_1}{a} + \frac{\delta a}{a} + \sin^2\phi_1\,\delta f\right) \\
&\qquad - \frac{\delta a}{a} + (\sin^2\phi - 2\sin\phi_1 \sin\phi)\,\delta f.
\end{aligned} \tag{5-59}$$

These formulas for the effect of a shift of the geodetic datum are among the most important equations of geodesy. They were derived by several scientists,[1] among them Vening Meinesz (1950, 1953), by whose name they are usually known. Superficially similar formulas due to Helmert were in use earlier, but they are based on completely different geometrical principles, and are not suited for the purposes of modern geodesy.[2]

[1] We mention de Graaff-Hunter in 1929, Krassovsky in 1934 and 1942, and Bomford in 1939.

[2] Helmert's idea was the translation of geodesic lines on the ellipsoid, which is essentially a two-dimensional problem, whereas Vening Meinesz' idea is the translation of the ellipsoid in space. Only this corresponds to the essentially three-dimensional character of modern geodesy.

It may be noted that the first two equations (5-59) could also have been derived by differentiating the third of these equations, since (2-204) gives

$$\delta\xi = -\frac{1}{R}\frac{\partial(\delta N)}{\partial\phi} \doteq -\frac{\partial}{\partial\phi}\left(\frac{\delta N}{a}\right),$$

$$\delta\eta = -\frac{1}{R\cos\phi}\frac{\partial(\delta N)}{\partial\lambda} \doteq -\frac{1}{\cos\phi}\frac{\partial}{\partial\lambda}\left(\frac{\delta N}{a}\right)$$

as a spherical approximation.

Applications. As an illustration, we shall apply these formulas to the most important practical case, the absolute orientation of a local geodetic system, or its conversion to a world geodetic system (Heiskanen, 1951). Let us assume that a triangulation or trilateration net has been computed on a local geodetic datum $(a', f'; \xi_1', \eta_1', N_1')$. The quantities referred to this system will be indicated by a prime. Thus ξ_1', η_1', N_1' belong to the fundamental point P_1; they may be assumed to be zero or to have any other values.

Suppose now that at the initial point the absolute geoidal height N_1 and the absolute deflection components ξ_1 and η_1 are known. (Their determination will be discussed in the following section.) The absolute values N, ξ, η will in general refer to a different ellipsoid (a, f), whose center is at the center of gravity of the earth. The quantities a, f; ξ_1, η_1, N_1 determine this "world geodetic system" completely.

It is now very easy to transform the local system $(a', f'; \xi_1', \eta_1', N_1')$ to the world system. Set

$$\begin{aligned} \delta\xi_1 &= \xi_1 - \xi_1', \\ \delta\eta_1 &= \eta_1 - \eta_1', \\ \delta N_1 &= N_1 - N_1'; \end{aligned} \qquad \begin{aligned} \delta a &= a - a', \\ \delta f &= f - f'; \end{aligned} \tag{5-60}$$

and compute, for all points of the local system, the changes $\delta\xi$, $\delta\eta$, δN by equations (5-59). Then the ξ, η, N in the world system are given by

$$\begin{aligned} \xi &= \xi' + \delta\xi, \\ \eta &= \eta' + \delta\eta, \\ N &= N' + \delta N. \end{aligned}$$

The geodetic coordinates in the world geodetic system are obtained from

$$\begin{aligned} \phi &= \phi' - \delta\xi, \\ \lambda &= \lambda' - \delta\eta\sec\phi, \\ h &= h' + \delta N, \end{aligned}$$

and the geocentric rectangular coordinates X, Y, Z can be computed by (5-5).

A related problem is that of determining the coordinates x_0', y_0', z_0' of the center of the original reference ellipsoid defining the local datum $(a', f'; \xi_1', \eta_1', N_1')$. Since the new datum $(a, f; \xi_1, \eta_1, N_1)$, the world datum, is in an absolute position, we have

$$x_0 = y_0 = z_0 = 0,$$

so that

$$\begin{aligned} \delta x_0 &= x_0 - x_0' = -x_0', \\ \delta y_0 &= y_0 - y_0' = -y_0', \\ \delta z_0 &= z_0 - z_0' = -z_0', \end{aligned} \tag{5-61}$$

and

$$x_0' = -\delta x_0, \quad y_0' = -\delta y_0, \quad z_0' = -\delta z_0,$$

where δx_0, δy_0, δz_0 are computed from (5-56). This solves our problem.

5-10. Determination of the Size of the Earth

If we use the gravimetric method with a fixed reference ellipsoid whose center coincides with the earth's center of gravity, then the geoidal undulations are obtained by (2-183b),

$$N = N_0 + \frac{R}{4\pi G} \iint_\sigma \Delta g \, S(\psi) \, d\sigma, \tag{5-62}$$

and the determination of the size of the earth is reduced to the problem of determining the constant N_0 (Sec. 2-19). As we have seen, N_0 has an immediate geometrical meaning: if a is the equatorial radius of the given reference ellipsoid, then

$$a_E = a + N_0 \tag{5-63}$$

is the equatorial radius of an ellipsoid whose normal potential U_0 is equal to the actual potential W_0 of the geoid, and which encloses the same mass as that of the earth, the flattening f remaining the same. If the assumed reference ellipsoid has been chosen so that it has the same value of

$$J_2 = \frac{C - \bar{A}}{Ma^2}$$

as the earth, this quantity now being known accurately from artificial satellites (see Chap. 9), then a_E will be the semimajor axis of the mean earth ellipsoid; see Secs. 2-21 and 5-11.

By the gravimetric method we can determine only the second term on the right-hand side of the above formula, that is, the Stokes integral; to determine N_0 we need to use the astrogeodetic method with at least one measured distance. The principle has already been described in Sec. 2-19; we shall now approach the problem more practically.

The problem may be formulated concisely in the following way. The gravimetric geoid is assumed to be known all over the earth; it is in an absolute position, but since N_0 is not known, its scale is indeterminate. The astrogeodetic geoid is known over part of the earth; it is in a relative position defined by the local geodetic datum, but its scale is correctly known. What we have to do is to fit both geoids together, thus (1) determining the scale of the gravimetric geoid and (2) transforming the local astrogeodetic datum to the world geodetic system.

Suppose that the same reference ellipsoid (a, f) is used in both systems. [If not, we can first transform the astrogeodetic system to the parameters of the gravimetric reference ellipsoid by means of formulas (5-59), setting $\delta\xi_1 = \delta\eta_1 = \delta N_1 = 0$.] We denote the deflection components with respect to the absolute gravimetric datum by ξ, η and those with respect to the local astrogeodetic datum by ξ', η'. Then, because the same reference ellipsoid is used for both systems, we have $\delta a = \delta f = 0$, and with

$$\delta\xi = \xi - \xi', \quad \delta\eta = \eta - \eta'$$

the first two equations of (5-59) yield

$$\begin{aligned} \xi &= \xi' + (\cos\phi_1 \cos\phi + \sin\phi_1 \sin\phi \cos\Delta\lambda)\,\delta\xi_1 - \sin\phi \sin\Delta\lambda\,\delta\eta_1 \\ &\qquad - (\sin\phi_1 \cos\phi - \cos\phi_1 \sin\phi \cos\Delta\lambda)\,\frac{\delta N_1}{a}, \\ \eta &= \eta' + \sin\phi_1 \sin\Delta\lambda\,\delta\xi_1 + \cos\Delta\lambda\,\delta\eta_1 + \cos\phi_1 \sin\Delta\lambda\,\frac{\delta N_1}{a}, \end{aligned} \tag{5-64}$$

If $\delta\xi_1$, $\delta\eta_1$, δN_1 at the initial point are assumed to be known, then we can compute the deflections ξ and η in the world system by means of these formulas and compare them with the corresponding gravimetric deflections, obtained directly by the Vening Meinesz formula. Theoretically, we should obtain the same result. Denoting the transformed astrogeodetic deflections (5-64) by ξ^a, η^a and the direct gravimetric deflections by ξ^g, η^g, we thus should have

$$\xi^a = \xi^g, \quad \eta^a = \eta^g.$$

In practice, $\delta\xi_1$ and $\delta\eta_1$ can be directly computed for the initial point by

$$\delta\xi_1 = \xi_1^g - \xi_1', \quad \delta\eta_1 = \eta_1^g - \eta_1',$$

but δN_1 cannot be determined because N_0 is not at first known. Furthermore, because of inaccuracies in measurement and computation, there would be small nonzero residuals $\xi^a - \xi^g$ and $\eta^a - \eta^g$, even if N_0 were known. It is therefore reasonable to treat $\delta\xi_1$ and $\delta\eta_1$ as unknown, as well as N_1, and to determine them in such a way that the sum of the squares of these residuals at all astronomical stations is a minimum:

$$\sum [(\xi^a - \xi^g)^2 + (\eta^a - \eta^g)^2] = \text{Minimum}. \tag{5-65}$$

This can be done by the usual least-squares adjustment, with $\delta\xi_1$, $\delta\eta_1$, and δN_1 as unknowns and equations (5-64) as the observation equations.[1]

As regards the parameters resulting from the adjustment, the quantities $\delta\xi_1$ and $\delta\eta_1$ help to determine the absolute position of the astrogeodetic system.

[1] The simple condition (5-65) is not quite correct from the theoretical standpoint of adjustment computations, because different weights of observations and correlations between them are not taken into account. For an improved method see Kaula (1959).

Since the sum (5-65) is extended over many astronomical stations, $\delta\xi_1$ and $\delta\eta_1$ from the adjustment will be much more accurate than if they were computed from the gravimetric deflections ξ_1^g and η_1^g at the fundamental point P_1 only, as indicated above.

Of greatest interest, however, is the quantity

$$\delta N_1 = N_1 - N_1'; \tag{5-66}$$

N_1' is the astrogeodetic geoidal height at the initial point, and like ξ_1' and η_1' it has been assumed beforehand to define the local geodetic datum $(a, f; \xi_1', \eta_1', N_1')$. The quantity N_1 is expressed by the generalized Stokes formula (5-62), which contains the constant N_0 as an unknown parameter. But now, since δN_1 has been found, it is possible to determine this parameter by

$$N_0 = N_1' + \delta N_1 - \frac{R}{4\pi G}\iint_\sigma \Delta g\, S(\psi)\, d\sigma, \tag{5-67}$$

as equations (5-62) and (5-66) show.

The parameter δN_1 thus serves two different purposes. It determines the scale of the global gravimetric geoid; and, together with $\delta\xi_1$ and $\delta\eta_1$, it is needed for transforming the local astrogeodetic datum to the global system.

If the flattening of the given reference ellipsoid is in agreement with the coefficient J_2 of the actual earth, then the "equatorial radius of the earth" is computed by adding N_0, obtained from (5-67), to the semimajor axis of the given ellipsoid according to (5-63).

If astrogeodetic geoidal heights N' have been determined at suitable points, it is also possible to express N as a function of $\delta\xi_1$, $\delta\eta_1$, δN_1 by using the third equation of (5-59),

$$\begin{aligned} N = N' &- (\cos\phi_1 \sin\phi - \sin\phi_1 \cos\phi \cos\Delta\lambda) a\,\delta\xi_1 - \cos\phi \sin\Delta\lambda\; a\,\delta\eta_1 \\ &+ (\sin\phi_1 \sin\phi + \cos\phi_1 \cos\phi \cos\Delta\lambda)\,\delta N_1, \end{aligned} \tag{5-68}$$

and to determine $\delta\xi_1$, $\delta\eta_1$, δN_1 from the condition

$$\sum (N^a - N^g)^2 = \text{Minimum}, \tag{5-69}$$

where N^a are the transformed astrogeodetic geoidal heights (5-68) and N^g are those obtained gravimetrically by means of the extended Stokes formula (5-62). The analogy between these equations and (5-64) and (5-65) is obvious.

Actual determinations of the equatorial radius of the earth were made in this way, for example, by Ledersteger (1951) and Fischer (1959).

An alternative approach. We have seen in the preceding section that we may define a geodetic coordinate system or a reference datum in two alternative ways: either by

$$(a, f; x_0, y_0, z_0)$$

or by

$$(a, f; \xi_1, \eta_1, N_1).$$

In the first case the position of the reference ellipsoid is determined by the coordinates x_0, y_0, z_0 of its center with respect to the center of the earth; in the second case, by the assumed geodetic coordinates ϕ_1, λ_1, h_1 of an arbitrary initial point, or by the equivalent quantities ξ_1, η_1, N_1.

The second definition is the more customary of the two, and we have therefore used it so far in this section. The first definition, however, provides an approach to our problem that is in some respects simpler (Molodenskii et al., 1962, pp. 113–117).

We have found, in Sec. 2-18, that a parallel shift of the reference ellipsoid with respect to the earth's center is expressed in the geoidal undulation by adding a first-degree spherical harmonic (2-176b) to the generalized Stokes formula (5-62). An astrogeodetic datum differs by precisely such a shift from a geocentric system to which (5-62) refers. Therefore, the astrogeodetic geoidal undulation must have the form

$$\begin{aligned} N' &= N + \delta x_0 \cos\phi \cos\lambda + \delta y_0 \cos\phi \sin\lambda + \delta z_0 \sin\phi \\ &= N_{St} + N_0 + \delta x_0 \cos\phi \cos\lambda + \delta y_0 \cos\phi \sin\lambda + \delta z_0 \sin\phi, \end{aligned} \tag{5-70}$$

where

$$N_{St} = \frac{R}{4\pi G} \iint_\sigma \Delta g \, S(\psi) \, d\sigma \tag{5-71}$$

is Stokes' integral and δx_0, δy_0, δz_0 are certain constants. In agreement with (5-61) they are the negative coordinates of the center of the ellipsoid with respect to the center of the earth. As compared to (2-176b), we have replaced ξ, η, ζ by δx_0, δy_0, δz_0 and θ by $90° - \phi$.

Since the deflection components are, according to (2-204), defined by

$$\xi = -\frac{1}{R}\frac{\partial N}{\partial \phi}, \quad \eta = -\frac{1}{R\cos\phi}\frac{\partial N}{\partial \lambda},$$

we have

$$\frac{\partial N}{\partial \phi} = -R\xi, \quad \frac{\partial N}{\partial \lambda} = -R\eta\cos\phi.$$

By differentiating (5-70) with respect to ϕ and λ and multiplying both sides of the resulting equations by $-1/R$ we get

$$\begin{aligned} \xi' &= \xi + \frac{1}{R}(\delta x_0 \sin\phi\cos\lambda + \delta y_0 \sin\phi \sin\lambda - \delta z_0 \cos\phi), \\ \eta' &= \eta + \frac{1}{R}(\delta x_0 \sin\lambda - \delta y_0 \cos\lambda), \end{aligned} \tag{5-72}$$

where ξ', η' are the astrogeodetic deflections and ξ, η the gravimetric deflections as obtained by Vening Meinesz' formula.

Equations (5-70) and (5-72) are basic for determining the scale parameter N_0 and the shift parameters δx_0, δy_0, δz_0. The data are: astrogeodetic deflections and undulations, ξ', η', N'; and the gravimetrically computed quantities ξ, η, N_{St}.

Four equations of type (5-70), obtained by setting up (5-70) for four astrogeodetic stations, are the minimum requirement for the determination of all four parameters; three equations of type (5-72) give the shift parameters. It is remarkable that the scale parameter N_0 cannot be determined from deflections of the vertical alone, since equations (5-72) do not contain it.

In practice many equations of types (5-70) and (5-72) will be available, one or more for every astrogeodetic station. This calls for an adjustment. Again we may use the minimum condition (5-65), supplemented by at least one gravimetric determination of N to furnish the scale parameter N_0, or else the condition (5-69). Then

$$\xi^a - \xi^g = \xi' - \xi - \frac{1}{R}(\delta x_0 \sin\phi\cos\lambda + \delta y_0 \sin\phi\sin\lambda - \delta z_0\cos\phi),$$

$$\eta^a - \eta^g = \eta' - \eta - \frac{1}{R}(\delta x_0 \sin\lambda - \delta y_0\cos\lambda), \tag{5-73}$$

$$N^a - N^g = N' - N_{St} - \delta x_0\cos\phi\cos\lambda - \delta y_0\cos\phi\sin\lambda - \delta z_0\sin\phi - N_0.$$

Each of these expressions should ideally be equal to zero.

The only difference between (5-64) and (5-68) on the one hand and (5-72) and (5-70) on the other hand is the choice of parameters: $\delta\xi_1$, $\delta\eta_1$, δN_1 or δx_0, δy_0, δz_0. Therefore, we could also have derived these equations from the formulas of Sec. 5-9, using equations (5-55) with

$$\begin{aligned}\delta\phi &= -\delta\xi = \xi' - \xi,\\ \delta\lambda\cos\phi &= -\delta\eta = \eta' - \eta,\\ \delta h &= \delta N = -(N' - N),\\ \delta a &= \delta f = 0.\end{aligned}$$

5-11. Best-fitting Ellipsoids and the Mean Earth Ellipsoid

In Sec. 2-21 we defined the mean earth ellipsoid physically as that ellipsoid of revolution which shares with the earth the mass M, the potential W_0, the difference between the principal moments of inertia $k(C - \bar{A})$, and the angular velocity ω, where $\bar{A} = (A + B)/2$.

It is also possible to define the mean earth ellipsoid geometrically as that ellipsoid which approximates the geoid most closely. This definition is perhaps more appealing to the geodesist; it may, for instance, be formulated by the condition that the sum of the squares of the deviations N of the geoid from the ellipsoid be a minimum:

$$\iint_\sigma N^2\, d\sigma = \text{Minimum} \tag{5-74}$$

(this integral is to be considered the limit of a sum). The condition of closest approximation may also be expressed in terms of the deflections of the vertical:

$$\iint_\sigma (\xi^2 + \eta^2)\, d\sigma = \text{Minimum}, \tag{5-75}$$

minimizing the sum of the squares of the total deflection of the vertical

$$\theta = \sqrt{\xi^2 + \eta^2}.$$

Many other similar definitions of closest approximation are possible.

The first definition is the most plausible and the most appropriate intuitively, as has been already noted by Helmert; in principle, however, all definitions are more or less conventional and are equivalent theoretically.

The second definition, based on the condition (5-75), uses deflections of the vertical and is thus particularly well adapted to the astrogeodetic method. However, since this method can be applied only over limited areas, at most spanning the continents, the integral (5-75) must be replaced by a sum covering the astronomical stations of a restricted region:

$$\sum (\xi^2 + \eta^2) = \text{Minimum}. \tag{5-75'}$$

In this way we can get only the best-fitting ellipsoid for the region considered, rather than a general earth ellipsoid. As Fig. 5-11 indicates, a locally best-fitting ellipsoid may be quite different from the mean earth ellipsoid, which can be considered a best-fitting ellipsoid for the whole earth.

If a good approximation of the earth ellipsoid by a local best-fitting ellipsoid is desired, it is advisable to subtract the effect of the topography and of its isostatic compensation from the astrogeodetic deflections of the vertical before the minimum condition (5-75′) is applied. The purpose of this procedure is to smooth the irregularities of the geoid. In this way Hayford computed the international ellipsoid as ellipsoid that best fits the isostatically reduced vertical deflections in the United States. See his papers quoted at the end of Chapter 2; Rapp (1963) made an interesting recomputation.

This method is impaired by unknown density anomalies and by the lack of complete isostatic compensation. Therefore, it is better to go still one step

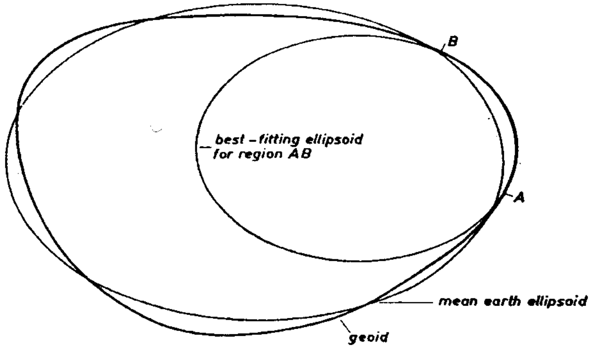

FIGURE 5-11

A local best-fitting ellipsoid and the mean earth ellipsoid.

further and subtract the gravimetrically computed values from the astrogeodetic deflections. Then the minimum condition (5-65) results.

Thus we may say that Hayford's method is equivalent to the use of (5-65), the gravimetric values ξ^g, η^g being approximated by deflections that represent the effect of topography and of its isostatic compensation only. If the isostatic compensation were complete, and if we had perfect knowledge of the density above the geoid, both methods would give exactly the same result if applied properly. It may be noted that these two methods have their exact counterpart in the gravimetrical and the topographic interpolation of the astrogeodetic deflections of the vertical, discussed in Sec. 5-8.

Thus a relation between the two minimum conditions (5-65) and (5-75), or (5-75′), has been found, Hayford's method being a transition. It should be kept in mind, however, that these two conditions are of an essentially different nature. Equation (5-65) involves the residuals $\xi^a - \xi^g$ and $\eta^a - \eta^g$, which are theoretically zero, being nonzero only because of inaccuracies of measurement, whereas the conditions (5-75) or (5-75′) involve the ξ and η themselves, which are essentially nonzero. Hence (5-65) is only a condition of adjustment, whereas (5-75) is an essential definition.

Equivalence of different definitions of the earth ellipsoid. It is quite remarkable that the minimum definitions (5-74) or (5-75) and a similar definition due to Rudzki, using the condition

$$\iint_\sigma (\Delta g)^2 \, d\sigma = \text{Minimum}, \tag{5-76}$$

yield results which, to the usual spherical approximation, are identical with each other and with the physical definition in terms of M, W_0, $C - \bar{A}$, and ω.

This can be seen as follows. We write the spherical-harmonic expansion of the disturbing potential in the form

$$T = \frac{k\delta M}{R} + \sum_{n=1}^{\infty} \sum_{m=0}^{n} [a_{nm}R_{nm}(\theta, \lambda) + b_{nm}S_{nm}(\theta, \lambda)].$$

Then, according to Sec. 2-19, equations (2-178′) and (2-184a), we have

$$N = \frac{k\delta M}{RG} - \frac{\delta W}{G} + \frac{1}{G} \sum_{n=1}^{\infty} \sum_{m=0}^{n} [a_{nm}R_{nm}(\theta, \lambda) + b_{nm}S_{nm}(\theta, \lambda)]$$

and

$$\Delta g = -\frac{k\delta M}{R^2} + \frac{2\delta W}{R} + \frac{1}{R} \sum_{n=1}^{\infty} \sum_{m=0}^{n} [(n-1)a_{nm}R_{nm}(\theta, \lambda) + (n-1)b_{nm}S_{nm}(\theta, \lambda)].$$

The condition of equal masses, $\delta M = 0$, is very natural and will be assumed. If we square the formulas for N and Δg and integrate over the whole earth, then all the integrals of products of different harmonics R_{nm} and S_{nm} will be zero, according to the orthogonality property (1-68), and the remaining integrals will be given by (1-69). Thus we find easily

$$\iint_\sigma N^2\, d\sigma = \frac{4\pi}{G^2}\,\delta W^2 + \frac{4\pi}{G^2}\sum_{n=1}^{\infty}\frac{1}{2n+1}\left[a_{n0}^2 + \sum_{m=1}^{n}\frac{(n+m)!}{2(n-m)!}(a_{nm}^2 + b_{nm}^2)\right], \tag{5-77a}$$

$$\iint_\sigma (\Delta g)^2\, d\sigma = \frac{16\pi}{R^2}\,\delta W^2 + \frac{4\pi}{R^2}\sum_{n=1}^{\infty}\frac{(n-1)^2}{2n+1}\left[a_{n0}^2 + \sum_{m=1}^{n}\frac{(n+m)!}{2(n-m)!}(a_{nm}^2 + b_{nm}^2)\right]. \tag{5-77b}$$

By a more complicated derivation, which we shall omit here but which can be found in Molodenskii et al. (1962, p. 87), one gets the similar formula

$$\iint_\sigma (\xi^2 + \eta^2)\, d\sigma = \frac{4\pi}{R^2G^2}\sum_{n=1}^{\infty}\frac{n(n+1)}{2n+1}\left[a_{n0}^2 + \sum_{m=1}^{n}\frac{(n+m)!}{2(n-m)!}(a_{nm}^2 + b_{nm}^2)\right]. \tag{5-77c}$$

Varying the size and shape of the reference ellipsoid and its position with respect to the earth changes only the coefficients δW, a_{10}, a_{11}, b_{11}, and a_{20}, leaving the other coefficients practically invariant. Thus the minimum of any of the integrals (5-77) is obtained if all these coefficients are equal to zero. Now $\delta W = 0$ means equal potential $U_0 = W_0$; $a_{10} = a_{11} = b_{11} = 0$ means absolute position (coincident centers of gravity); and $a_{20} = 0$ means equality of J_2 or of $C - \frac{A+B}{2}$.

Thus the equivalence of the physical definition by means of M, W_0, $k(C - \bar{A})$, ω and of the condition of closest approximation in any of the forms (5-74), (5-75), or (5-76) has been established. [It may be noted that (5-77b) contains no first-degree term, because of the factor $(n-1)^2$, and that (5-77c) contains no term of degree zero, so that these equations do not determine the missing terms.]

Strictly speaking, the above argument presupposes that there are no masses outside the geoid. Practically, this means that free-air anomalies must be used for Δg; see Sec. 8-13.

5-12. Three-dimensional Geodesy

The idea of a rigorous computation of a triangulation network in space dates back to Bruns (1878).[1] Consider the polyhedron formed by triangulation benchmarks on the surface of the earth and the straight lines of sight connecting them (Fig. 5-12). Another set of straight lines—one through each corner—represents the plumb line at the stations. In order to determine this figure we need five

[1] See also Wolf (1963a).

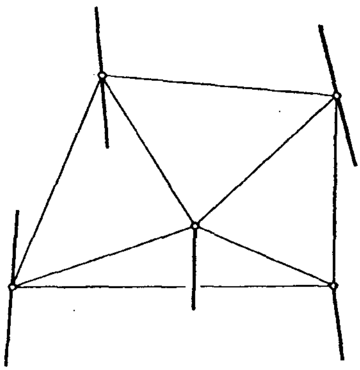

FIGURE 5-12
Bruns' polyhedron.

parameters for each station—three coordinates and two parameters defining the direction of the plumb line. The main observational data for this purpose are:

1. horizontal angles and zenith distances, obtained by theodolite observations;
2. straight spatial distances, obtained, for instance, by electronic distance measurements; and
3. astronomical observations of latitude and longitude to fix the direction of the plumb line, and of azimuth A to determine the orientation of the polyhedron.

We may use a rectangular coordinate system; then the three coordinates to be determined will be X, Y, Z. The parameters defining the direction of the plumb line are conveniently taken to be Φ and Λ, astronomical latitude and longitude. We shall now express the astronomical azimuth A, the measured (astronomical) zenith distance z', and the spatial distance s in terms of these five parameters.

We introduce a local coordinate system u, v, w according to Fig. 5-13. The origin is at the station P under consideration, the w-axis coincides with the plumb line, and the u- and v-axes point northward and eastward, respectively. Then azimuth A and zenith distance z' to a neighboring station Q are given by

$$\tan A = \frac{v}{u},$$
$$\cos z' = \frac{w}{s}. \tag{5-78}$$

We have

$$u = \Delta\mathbf{X}\cdot\mathbf{e}',$$
$$v = \Delta\mathbf{X}\cdot\mathbf{e}'', \tag{5-79}$$
$$w = \Delta\mathbf{X}\cdot\mathbf{n},$$

where $\mathbf{e}'$, $\mathbf{e}''$, $\mathbf{n}$ are the unit coordinate vectors in the *uvw*-system, and $\Delta\mathbf{X}$ is the vector leading from P to Q:

$$\Delta\mathbf{X} = \mathbf{X}_Q - \mathbf{X}_P.$$

In the XYZ-system we have the components

$$\Delta\mathbf{X} = \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} = \begin{pmatrix} X_Q - X_P \\ Y_Q - Y_P \\ Z_Q - Z_P \end{pmatrix} \tag{5-80}$$

and, according to (2-26),

$$\mathbf{n} = \begin{pmatrix} \cos\Phi\cos\Lambda \\ \cos\Phi\sin\Lambda \\ \sin\Phi \end{pmatrix}. \tag{5-81a}$$

In order to express $\mathbf{e}'$ and $\mathbf{e}''$ we inspect Fig. 5-14. The projections of $\mathbf{e}'$ and $\mathbf{e}''$ onto the XY-plane having the lengths $\sin\Phi$ and 1, and the longitudes $\Lambda + 180°$ and $\Lambda + 90°$, we find at once

$$\mathbf{e}' = \begin{pmatrix} -\sin\Phi\cos\Lambda \\ -\sin\Phi\sin\Lambda \\ \cos\Phi \end{pmatrix}, \quad \mathbf{e}'' = \begin{pmatrix} -\sin\Lambda \\ \cos\Lambda \\ 0 \end{pmatrix}. \tag{5-81b}$$

On expressing u, v, w in terms of these components according to (5-79) and inserting them into (5-78) we obtain

$$\begin{aligned} A &= \tan^{-1}\frac{-\Delta X\sin\Lambda + \Delta Y\cos\Lambda}{-\Delta X\sin\Phi\cos\Lambda - \Delta Y\sin\Phi\sin\Lambda + \Delta Z\cos\Phi}, \\ z' &= \cos^{-1}\frac{\Delta X\cos\Phi\cos\Lambda + \Delta Y\cos\Phi\sin\Lambda + \Delta Z\sin\Phi}{\sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2}}, \\ s &= \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2}. \end{aligned} \tag{5-82}$$

For the sake of completeness we have added the obvious equation for the distance s.

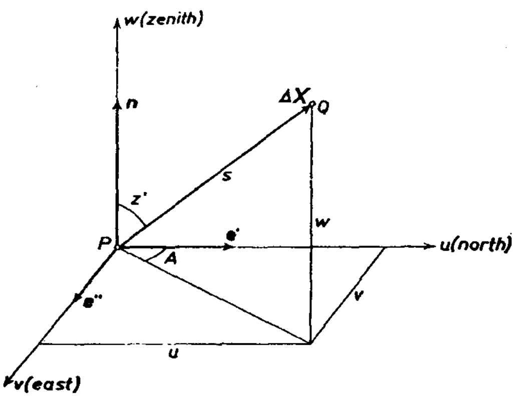

FIGURE 5-13

The local coordinate system uvw.

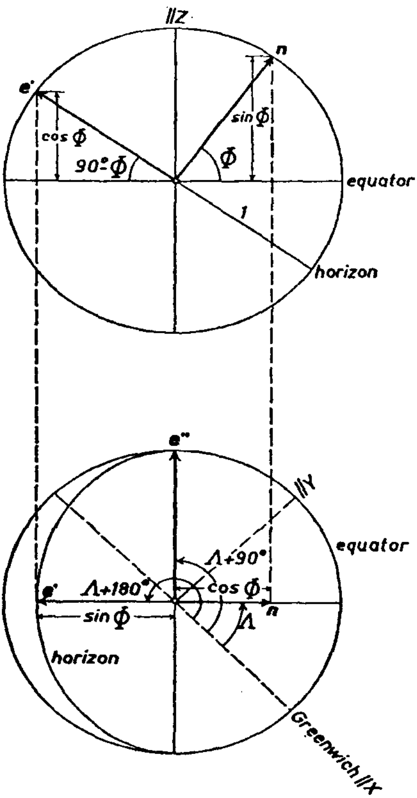

FIGURE 5-14

The unit sphere illustrating the unit vectors **n**, **e'**, **e''**, *as seen from the east (top) and from the north pole (bottom).*

These equations express the observables azimuth, zenith distance, and linear distance in terms of rectangular coordinates and direction parameters:

$$\begin{aligned} A &= A(X_P, Y_P, Z_P;\ X_Q, Y_Q, Z_Q;\ \Phi, \Lambda), \\ z' &= z'(X_P, Y_P, Z_P;\ X_Q, Y_Q, Z_Q;\ \Phi, \Lambda), \\ s &= s(X_P, Y_P, Z_P;\ X_Q, Y_Q, Z_Q). \end{aligned} \tag{5-82'}$$

The first of equations (5-82), or (5-82'), serves this purpose not only for the azimuth A, but also for horizontal angles, since every horizontal angle may be considered as a difference of two azimuths.

Differential formulas. Since equations (5-82) are somewhat complicated, it is often convenient to assume suitable approximate values and to compute corrections by differential formulas. The differential formula for the azimuth has the form

$$\begin{aligned} \delta A = k_1\,\delta X_P + k_2\,\delta Y_P + k_3\,\delta Z_P + k_4\,\delta X_Q + k_5\,\delta Y_Q \\ + k_6\,\delta Z_Q + k_7\,\delta\Phi_P + k_8\,\delta\Lambda_P, \end{aligned} \tag{5-83}$$

where

$$k_1 = \frac{\partial A}{\partial X_P}, \quad \ldots, \quad k_8 = \frac{\partial A}{\partial \Lambda}.$$

It should be noted that Φ and Λ define the direction of the plumb line at P, so that

$$\Phi = \Phi_P, \quad \Lambda = \Lambda_P.$$

Similar expressions hold for $\delta z'$ and δs.

Instead of rectangular coordinates X, Y, Z we may also use geodetic coordinates ϕ, λ, h. The insertion of equations (5-5) into (5-82) gives A, z', and s as functions of ϕ_P, λ_P, h_P; ϕ_Q, λ_Q, h_Q; and Φ_P, Λ_P. These functions, however, are so complicated as to be of little use; we do better to limit ourselves to the corresponding differential formulas from the outset. For this purpose we write equations (5-54) with $\delta x_0 = \delta y_0 = \delta z_0 = \delta a = \delta f = 0$ and $a \doteq R$, obtaining

$$\begin{aligned} \delta X &= -\sin\phi\cos\lambda\cdot R\,\delta\phi - \sin\lambda\cdot R\cos\phi\,\delta\lambda + \cos\phi\cos\lambda\cdot\delta h, \\ \delta Y &= -\sin\phi\sin\lambda\cdot R\,\delta\phi + \cos\lambda\cdot R\cos\phi\,\delta\lambda + \cos\phi\sin\lambda\cdot\delta h, \\ \delta Z &= \cos\phi\cdot R\,\delta\phi + \sin\phi\cdot\delta h, \end{aligned} \tag{5-84}$$

and insert them into (5-83) and the corresponding expressions for $\delta z'$ and δs. Omitting the simple but laborious details we merely state the result:

$$\begin{aligned} \delta A &= a_1\,\delta\phi_P + a_2\,\delta\lambda_P + a_3\,\delta h_P + a_4\,\delta\phi_Q + a_5\,\delta\lambda_Q + a_6\,\delta h_Q \\ &\qquad + a_7\,\delta\Phi_P + a_8\,\delta\Lambda_P, \\ \delta z' &= b_1\,\delta\phi_P + b_2\,\delta\lambda_P + b_3\,\delta h_P + b_4\,\delta\phi_Q + b_5\,\delta\lambda_Q + b_6\,\delta h_Q \\ &\qquad + b_7\,\delta\Phi_P + b_8\,\delta\Lambda_P, \\ \delta s &= c_1\,\delta\phi_P + c_2\,\delta\lambda_P + c_3\,\delta h_P + c_4\,\delta\phi_Q + c_5\,\delta\lambda_Q + c_6\,\delta h_Q, \end{aligned} \tag{5-85}$$

where

$$a_1 = \frac{R\sin\alpha}{s\sin z}, \quad a_2 = -\frac{R\cos\alpha}{s\sin z}\cos\phi, \quad a_3 = 0,$$

$$a_4 = -\frac{R\sin\alpha}{s\sin z}[\cos(\phi_Q - \phi) + \sin\phi_Q\sin\Delta\lambda\cot\alpha],$$

$$a_5 = \frac{R\cos\alpha\cos\phi_Q}{s\sin z}(\cos\Delta\lambda - \sin\phi\sin\Delta\lambda\tan\alpha),$$

$$a_6 = \frac{\cos\alpha\cos\phi_Q}{s\sin z}[\sin\Delta\lambda + (\sin\phi\cos\Delta\lambda - \cos\phi\tan\phi_Q)\tan\alpha],$$

$$a_7 = \sin\alpha\cot z, \quad a_8 = \sin\phi - \cos\alpha\cos\phi\cot z;$$

$$b_1 = -\frac{R}{s}\cos z\cos\alpha, \quad b_2 = -\frac{R}{s}\cos z\sin\alpha\cos\phi, \quad b_3 = \frac{1}{s}\sin z, \tag{5-86}$$

$$b_4 = \frac{R}{s\sin z}(\cos\phi\sin\phi_Q\cos\Delta\lambda - \sin\phi\cos\phi_Q - \cos z\sin z_Q\cos\alpha_Q),$$

$$b_5 = \frac{R\cos\phi_Q}{s\sin z}(\cos\phi\sin\Delta\lambda - \cos z\sin z_Q\sin\alpha_Q),$$

$$b_6 = -\frac{1}{s \sin z}(\cos z \cos z_Q + \sin \phi \sin \phi_Q + \cos \phi \cos \phi_Q \cos \Delta\lambda),$$

$$b_7 = -\cos \alpha, \quad b_8 = -\cos \phi \sin \alpha;$$

$$c_1 = -R \sin z \cos \alpha, \quad c_2 = -R \cos \phi \sin z \sin \alpha, \quad c_3 = -\cos z,$$

$$c_4 = -R \sin z_Q \cos \alpha_Q, \quad c_5 = -R \cos \phi_Q \sin z_Q \sin \alpha_Q, \quad c_6 = -\cos z_Q.$$

Here we have used the abbreviation

$$\Delta\lambda = \lambda_Q - \lambda_P$$

and have omitted the subscript P in quantities referred to point P; α_Q and z_Q are azimuth and zenith distance to P measured in Q.

A complete derivation can be found in Wolf (1963b). It may be mentioned that by following the method outlined above we end up with expressions for the coefficients (5-86) that partly contain the astronomic coordinates Φ, Λ, and partly the geodetic coordinates ϕ, λ. The former arise from the differentiation of (5-82); the latter enter through (5-84). For the sake of consistency we have in (5-86) used the geodetic quantities only, since in these small coefficients we may replace Φ, Λ by ϕ, λ without losing accuracy.

However, the fundamental conceptual difference between the astronomic and the geodetic latitude and longitude should be carefully noted. The astronomical "coordinates" Φ and Λ enter as *direction parameters*, defining the direction of the plumb line [compare equations (5-82)], whereas their geodetic counterparts ϕ and λ, together with h, enter as true *point coordinates* essentially equivalent to X, Y, Z. This is also the reason why the equation for δs does not contain $\delta\Phi$ and $\delta\Lambda$, for the spatial distance s does not depend on the direction of the plumb line.

"Three-dimensional" and "classical" geodesy. The formulas developed so far contain the main principles of *three-dimensional geodesy*. The computation is as follows. Preliminary values of ϕ, λ, h are obtained, for instance, by using the corresponding natural coordinates:

$$\phi^0 = \Phi,$$
$$\lambda^0 = \Lambda,$$
$$h^0 = H.$$

(If available, ϕ^0 and λ^0 may preferably be taken as the geodetic coordinates computed on a local datum.) These preliminary coordinates ϕ^0, λ^0, h^0 are converted into rectangular coordinates X^0, Y^0, Z^0 according to equations (5-5). Then azimuths A^0, zenith distances z^0, and linear distances s^0 are computed from these preliminary coordinates by means of (5-82), as far as needed to compare them with the observed values of A, z', and s, or with measured horizontal angles, which are azimuth differences. Each difference

$$\delta A = A - A^0, \quad \delta z' = z' - z^0, \quad \delta s = s - s^0$$

will furnish one equation of type (5-85). A sufficient number of such equations will permit the solution with respect to the unknowns $\delta\phi$, $\delta\lambda$, δh, $\delta\Phi$, $\delta\Lambda$ for each station, preferably by means of a least-squares adjustment.

This procedure was advocated by Bruns (1878), who stated the form of equations (5-85) without evaluating the coefficients, and by Hotine (1959), who gave a very detailed and complete exposition and initiated extensive recent activity.[1]

It is easy to see the relation between this procedure and the methods described earlier in this chapter. Let us first specialize the first two equations (5-85) for the case

$$\delta\phi = \delta\lambda = \delta h = 0$$

for both points P and Q, and

$$\begin{aligned} \delta\Phi &= \Phi - \phi = \xi, \\ \delta\Lambda &= \Lambda - \lambda = \eta \sec\phi. \end{aligned} \tag{5-87}$$

This means that without changing the ellipsoidal coordinates ϕ, λ, h as such we make the transition from the ellipsoidal normal, for which ϕ, λ are the direction parameters, to the actual plumb line defined by Φ, Λ.

For this special case we shall evidently have

$$\begin{aligned} \delta A &= A - \alpha, \\ \delta z' &= z' - z, \end{aligned}$$

corresponding to the transition from the ellipsoidal azimuth α and zenith distance z (referred to the ellipsoidal normal) to the observed azimuth A and zenith distance z' (referred to the actual plumb line). With these substitutions we obtain from (5-85)

$$\begin{aligned} \alpha &= A - a_7\xi - a_8\eta \sec\phi, \\ z &= z' - b_7\xi - b_8\eta \sec\phi, \end{aligned} \tag{5-88}$$

and with the coefficients (5-86),

$$\begin{aligned} \alpha &= A - \xi \sin\alpha \cot z - \eta(\tan\phi - \cos\alpha \cot z) \\ &= A - \eta \tan\phi - (\xi \sin\alpha - \eta \cos\alpha) \cot z; \\ z &= z' + \xi \cos\alpha + \eta \sin\alpha. \end{aligned} \tag{5-88'}$$

Thus we have again arrived at equations (5-13) and (5-20) for the reduction to the ellipsoid, but in a quite different way.

After this specialization we shall return to the general case of equations (5-85). We write the first two equations (5-85) in the form

$$\begin{aligned} \delta A - a_7\,\delta\Phi - a_8\,\delta\Lambda &= a_1\,\delta\phi_P + a_2\,\delta\lambda_P + a_3\,\delta h_P + a_4\,\delta\phi_Q + a_5\,\delta\lambda_Q + a_6\,\delta h_Q, \\ \delta z' - b_7\,\delta\Phi - b_8\,\delta\Lambda &= b_1\,\delta\phi_P + b_2\,\delta\lambda_P + b_3\,\delta h_P + b_4\,\delta\phi_Q + b_5\,\delta\lambda_Q + b_6\,\delta h_Q. \end{aligned}$$

If we again define $\delta\Phi$ and $\delta\Lambda$ by (5-87), then the left-hand sides become

$$\begin{aligned} \delta A - a_7\xi - a_8\eta \sec\phi &= \delta(A - a_7\xi - a_8\eta \sec\phi) = \delta\alpha, \\ \delta z' - b_7\xi - b_8\eta \sec\phi &= \delta(z' - b_7\xi - b_8\eta \sec\phi) = \delta z. \end{aligned}$$

[1] For simple computational formulas see Hirvonen (1964).

This means that these left-hand sides are now the changes in *ellipsoidal* azimuth α and zenith distance z due to a variation of the geodetic coordinates. In other words, we are left with the purely ellipsoidal equations

$$\delta\alpha = a_1\,\delta\phi_P + a_2\,\delta\lambda_P + a_3\,\delta h_P + a_4\,\delta\phi_Q + a_5\,\delta\lambda_Q + a_6\,\delta h_Q,$$
$$\delta z = b_1\,\delta\phi_P + b_2\,\delta\lambda_P + b_3\,\delta h_P + b_4\,\delta\phi_Q + b_5\,\delta\lambda_Q + b_6\,\delta h_Q.$$

Thus the "classical" method of reduction to the ellipsoid and computation by means of ellipsoidal quantities is seen to be formally equivalent to the "three-dimensional" method as expressed in the original equations (5-85). Both methods, if applied correctly, must give the same result (Levallois, 1963).

Since the deflection components ξ and η, which are needed for a reduction to the ellipsoid, already presuppose good values of the geodetic coordinates ϕ and λ, which can only be computed *after* this reduction, we are again led to an iterative procedure quite analogous to that outlined in Sec. 5-7. The principal difference is that in the astrogeodetic method the geoidal heights N are determined by an integration of the deflections of the vertical, as described in Sec. 5-7, whereas in the Bruns-Hotine method they are obtained through the use of zenith distances, as differences between triangulated and orthometric heights; see Sec. 4-7, equation (4-58). The astrogeodetic determination of N is usually preferred in practice because the measured zenith distances are strongly affected by unknown anomalies of atmospheric refraction.

It goes without saying that the spatial triangulation of Bruns and Hotine and the conventional astrogeodetic method are equivalent also in that they alone are not sufficient to determine the position of the local datums with respect to the center of mass of the earth.

Another aspect of the application of zenith distances, their use for determining deflections of the vertical as outlined in Secs. 4-7 and 5-8, is of course also implicitly contained in the second of equations (5-85).

References

Arnold, K. (1956). Beiträge zur gravimetrischen Geodäsie. Veröff. Geod. Inst. Potsdam, No. 11.

Arnold, K. (1959). Vergleiche zwischen verschiedenen Definitionen des mittleren Erdellipsoids. Gerlands Beitr. Geophys., v. 68, pp. 129–136.

Baeschlin, C. F. (1948). Lehrbuch der Geodäsie. Zürich, Orell-Füssli.

Bodemüller, H. (1957). Beitrag zur Schwerekorrektion geometrischer Nivellements. Deutsche Geodätische Kommission, A, No. 26.

Bomford, G. (1962). Geodesy. 2nd ed., Oxford University Press.

Bomford, G. (1963). The geoid in Europe and connected countries. General Assembly, Int. Union Geod. Geophys., Berkeley.

Bruns, H. (1878). Die Figur der Erde. Berlin, Publ. Preuss. Geod. Inst.

de Graaff-Hunter, J. (1913). Formulae for atmospheric refraction and their application

to terrestrial refraction and geodesy. Survey of India, Professional Paper No. 14.
Fischer, I. (1959). A tentative world datum from geoidal heights based on the Hough ellipsoid and the Columbus geoid. J. Geophys. Research, v. 64, pp. 73–84.
Fischer, I. (1961). The present extent of the astro-geodetic geoid and the geodetic world datum derived from it. Bull. Géod., No. 61, pp. 245–264.
Galle, A. (1914). Das Geoid im Harz. Veröff. Geod. Inst. Potsdam, No. 61.
Heiskanen, W. (1951). On the world geodetic system. Publ. Finn. Geod. Inst., No. 39.
Helmert, F. R. (1880). Die mathematischen und physikalischen Theorien der höheren Geodäsie, vol. 1. Leipzig, B. G. Teubner (reprinted 1962).
Helmert, F. R. (1900 and 1901). Zur Bestimmung kleiner Flächenstücke des Geoides aus Lotabweichungen mit Rücksicht auf Lotkrümmung. Sitzungsberichte Preuss. Akad. Wiss., 2 parts.
Hirvonen, R. A. (1964). Praktische Rechenformeln für die dreidimensionale Geodäsie. Zeitschrift für Vermessungswesen, v. 89, pp. 145–151.
Hirvonen, R. A., and H. Moritz (1963). Practical computation of gravity at high altitudes. Report No. 27, Inst. Geod. Phot. Cart., Ohio State Univ.
Hotine, M. (1959). A primer of non-classical geodesy. Presented to the First Symposium on Three-Dimensional Geodesy, Venice.
Kaula, W. M. (1959). Reconciliation of Stokes' function and astrogeodetic geoid determinations. J. Geophys. Research, v. 64, pp. 61–71.
Kobold, F. (1951). Die Bestimmung der Lotabweichungskomponenten im Meridian des St. Gotthardt aus Höhenwinkelmessungen. Commission Géodésique Suisse.
Kobold, F., and E. Hunziker (1962). Communication sur la courbure de la verticale. Bull. Géod., No. 65, pp. 265–267.
Ledersteger, K. (1951). Die Bestimmung des mittleren Erdellipsoides und der absoluten Lage der Landestriangulationen. Österreichische Zeitschrift für Vermessungswesen, Sonderheft 12.
Ledersteger, K. (1955). Die Reduktion der astronomischen Beobachtungen wegen Lotkrümmung. Schweizerische Zeitschrift für Vermessung, Kulturtechnik und Photogrammetrie, v. 53, pp. 225–239.
Ledersteger, K. (in press). Astronomische und physikalische Geodäsie. Vol. V of Jordan/Eggert/Kneissl, Handbuch der Vermessungskunde, 10th ed. Stuttgart, J. B. Metzler.
Levallois, J. J. (1963). La réhabilitation de la géodésique classique et la géodésie tridimensionelle. Bull. Géod., No. 68, pp. 193–199.
Molodenskii, M. S., V. F. Eremeev, and M. I. Yurkina (1962). Methods for study of the external gravitational field and figure of the earth. Transl. from Russian (1960). Jerusalem, Israel Program for Scientific Translations.
Mueller, I. I. (1963). Geodesy and the torsion balance. Journal of the Surveying and Mapping Division, Proc. Am. Soc. Civil Eng., v. 89, pp. 123–155.
Niethammer, T. (1939). Das astronomische Nivellement im Meridian des St. Gotthardt, Part II. Astronomisch-geodätische Arbeiten in der Schweiz, v. 20, Schweizerische Geodätische Kommission.
Ölander, V. R. (1951). On the geoid in the Baltic area and the orientation of the Baltic ring. Publ. Finn. Geod. Inst., No. 38.
Rapp, R. H. (1963). A consideration of Hayford's best fitting ellipsoid data using the differential change equations of Vening Meinesz. Geofisica pura e applicata, v. 54, pp. 1–5.

Rice, D. A. (1962). A geoidal section in the United States. Bull. Géod., No. 65, pp. 243–252.
Rinner, K. (1956). Über die Reduktion grosser elektronisch gemessener Entfernungen. Zeitschrift für Vermessungswesen, v. 81, pp. 47–55.
Vening Meinesz, F. A. (1950). New formulas for systems of deflections of the plumb-line and Laplace's theorem. Changes of deflections of the plumb-line brought about by a change of the reference ellipsoid. Bull. Géod., No. 15, pp. 33–51.
Vening Meinesz, F. A. (1953). Physical geodesy. Proc. Netherlands Acad. Sci., ser. B, v. 56(1).
Wolf, H. (1956). Versuch einer Geoidbestimmung im mittleren Europa aus astronomisch-geodätischen Lotabweichungen. Deutsche Geodätische Kommission, A, No. 18.
Wolf, H. (1963a). Dreidimensionale Geodäsie, Herkunft, Methodik und Zielsetzung. Zeitschrift für Vermessungswesen, v. 88, pp. 109–116.
Wolf, H. (1963b). Die Grundgleichungen der Dreidimensionalen Geodäsie in elementarer Darstellung. *Ibid.*, v. 88, pp. 225–233.

6

Gravity Field Outside the Earth

6-1. Introduction

The practical interest in the gravity field outside the earth is of a comparatively recent date. The two main purposes of such studies are (1) the evaluation of the effect of gravitational irregularities on motion in the earth's field, and (2) the application of gravity measurements made with airborne instruments.

For computational reasons it is again convenient to split the gravity potential W and the gravity vector

$$\mathbf{g} = \operatorname{grad} W \tag{6-1}$$

into a normal potential U and a normal gravity vector

$$\boldsymbol{\gamma} = \operatorname{grad} U, \tag{6-2}$$

and the disturbing potential $T = W - U$ and the gravity disturbance vector

$$\boldsymbol{\delta} = \operatorname{grad} T = \mathbf{g} - \boldsymbol{\gamma}. \tag{6-3}$$

The normal gravity field is usually taken to be the gravity field of a suitable equipotential ellipsoid. This permits closed formulas and offers other advantages of mathematical simplicity; see Sec. 2-12.

Thus U and $\boldsymbol{\gamma}$ are computed first, and W and $\mathbf{g}$ are then obtained by

$$W = U + T, \tag{6-4}$$

$$\mathbf{g} = \boldsymbol{\gamma} + \boldsymbol{\delta}. \tag{6-5}$$

For some purposes we need the vector of gravitation, grad V (pure attraction without centrifugal force) rather than the vector of gravity. The gravitational vector is computed from the gravity vector by subtracting the vector of centrifugal force:

$$\operatorname{grad} V = \mathbf{g} - \operatorname{grad} \Phi = \mathbf{g} - \begin{pmatrix} \omega^2 x \\ \omega^2 y \\ 0 \end{pmatrix}, \tag{6-6}$$

the notations of Sec. 2-1 being used. The rectangular coordinate system x, y, z will be applied in this chapter in the usual sense: it is geocentric, the x- and y-axes lying in the equatorial plane with Greenwich longitudes 0° and 90° East, respectively, and the z-axis being the rotation axis of the earth.

The sign of the components of $\mathbf{g}$, γ, $\boldsymbol{\delta}$, etc. will always be chosen so that they are positive in the direction of *increasing* coordinates.

6-2. Normal Gravity—Closed Formulas

The gravity field of an equipotential ellipsoid is best expressed in terms of ellipsoidal coordinates u, β, λ, introduced in Secs. 1-19 and 2-7. They are related to rectangular coordinates x, y, z by

$$\begin{aligned} x &= \sqrt{u^2 + E^2} \cos\beta \cos\lambda, \\ y &= \sqrt{u^2 + E^2} \cos\beta \sin\lambda, \\ z &= u \sin\beta. \end{aligned} \tag{6-7}$$

If x, y, z are given, then u, β, λ can be computed by closed formulas. First we find

$$x^2 + y^2 = (u^2 + E^2)\cos^2\beta, \quad z^2 = u^2 \sin^2\beta.$$

On eliminating β between these two equations we obtain a quadratic equation for u^2, whose solution is

$$u^2 = (x^2 + y^2 + z^2 - E^2)\left[\frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{4E^2z^2}{(x^2 + y^2 + z^2 - E^2)^2}}\right]. \tag{6-8a}$$

Then β is given by

$$\tan\beta = \frac{z\sqrt{u^2 + E^2}}{u\sqrt{x^2 + y^2}}, \tag{6-8b}$$

and for λ we simply have

$$\tan\lambda = \frac{y}{x}. \tag{6-8c}$$

The ellipsoidal coordinates now being known, the normal potential U is given by (2-62):

$$U = \frac{kM}{E}\tan^{-1}\frac{E}{u} + \frac{1}{2}\omega^2 a^2 \frac{q}{q_0}\left(\sin^2\beta - \frac{1}{3}\right) + \frac{1}{2}\omega^2(u^2 + E^2)\cos^2\beta. \tag{6-9}$$

The components of γ along the coordinate lines are, by (2-65) and (2-66),

$$\gamma_u = \frac{1}{w}\frac{\partial U}{\partial u} = -\frac{1}{w}\left[\frac{kM}{u^2 + E^2} + \frac{\omega^2 a^2 E}{u^2 + E^2}\frac{q'}{q_0}\left(\frac{1}{2}\sin^2\beta - \frac{1}{6}\right) - \omega^2 u \cos^2\beta\right],$$

$$\gamma_\beta = \frac{1}{w\sqrt{u^2+E^2}}\frac{\partial U}{\partial \beta} = -\frac{1}{w}\left(-\frac{\omega^2 a^2}{\sqrt{u^2+E^2}}\frac{q}{q_0} + \omega^2\sqrt{u^2+E^2}\right)\sin\beta\cos\beta, \quad (6\text{-}10)$$

$$\gamma_\lambda = \frac{1}{\sqrt{u^2+E^2}\cos\beta}\frac{\partial U}{\partial \lambda} = 0.$$

To get the components of γ in the xyz-system, we compute

$$\frac{\partial U}{\partial u} = \frac{\partial U}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial U}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial U}{\partial z}\frac{\partial z}{\partial u}, \quad \text{etc.}$$

The partial derivatives of x, y, z with respect to u, β, λ are obtained by differentiating equations (6-7); we find

$$\frac{\partial U}{\partial u} = \frac{u}{\sqrt{u^2+E^2}}\cos\beta\cos\lambda\frac{\partial U}{\partial x} + \frac{u}{\sqrt{u^2+E^2}}\cos\beta\sin\lambda\frac{\partial U}{\partial y} + \sin\beta\frac{\partial U}{\partial z},$$

$$\frac{\partial U}{\partial \beta} = -\sqrt{u^2+E^2}\sin\beta\cos\lambda\frac{\partial U}{\partial x} - \sqrt{u^2+E^2}\sin\beta\sin\lambda\frac{\partial U}{\partial y} + u\cos\beta\frac{\partial U}{\partial z},$$

$$\frac{\partial U}{\partial \lambda} = -\sqrt{u^2+E^2}\cos\beta\sin\lambda\frac{\partial U}{\partial x} + \sqrt{u^2+E^2}\cos\beta\cos\lambda\frac{\partial U}{\partial y}.$$

On introducing the components

$$\gamma_x = \frac{\partial U}{\partial x}, \ldots; \quad \gamma_u = \frac{1}{w}\frac{\partial U}{\partial u}, \ldots,$$

we obtain

$$\begin{aligned}
\gamma_u &= \frac{u}{w\sqrt{u^2+E^2}}\cos\beta\cos\lambda\,\gamma_x + \frac{u}{w\sqrt{u^2+E^2}}\cos\beta\sin\lambda\,\gamma_y + \frac{1}{w}\sin\beta\,\gamma_z,\\
\gamma_\beta &= -\frac{1}{w}\sin\beta\cos\lambda\,\gamma_x - \frac{1}{w}\sin\beta\sin\lambda\,\gamma_y + \frac{u}{w\sqrt{u^2+E^2}}\cos\beta\,\gamma_z, \qquad (6\text{-}11)\\
\gamma_\lambda &= -\sin\lambda\,\gamma_x + \cos\lambda\,\gamma_y.
\end{aligned}$$

These are the formulas of an orthogonal rectangular coordinate transformation. It is well known that the inverse transformation is obtained by simply interchanging the rows and columns in the matrix of this equation system. Thus we obtain

$$\begin{aligned}
\gamma_x &= \frac{u}{w\sqrt{u^2+E^2}}\cos\beta\cos\lambda\,\gamma_u - \frac{1}{w}\sin\beta\cos\lambda\,\gamma_\beta - \sin\lambda\,\gamma_\lambda,\\
\gamma_y &= \frac{u}{w\sqrt{u^2+E^2}}\cos\beta\sin\lambda\,\gamma_u - \frac{1}{w}\sin\beta\sin\lambda\,\gamma_\beta + \cos\lambda\,\gamma_\lambda, \qquad (6\text{-}12)\\
\gamma_z &= \frac{1}{w}\sin\beta\,\gamma_u + \frac{u}{w\sqrt{u^2+E^2}}\cos\beta\,\gamma_\beta.
\end{aligned}$$

This follows from the definition of these coefficients as direction cosines; equations (6-12) may also be found by solving the linear equations (6-11) with respect to γ_x, γ_y, γ_z in some other way.

The formulas of the present section are completely rigorous. It is possible to expand them into series; more convenient, however, are series expansions in spherical coordinates, which will be considered in the next section.

6-3. Normal Gravity—Series Expansions

In this section we shall use ordinary spherical coordinates r (radius vector), $\bar{\phi}$ (geocentric latitude), and λ (longitude):

$$r = \sqrt{x^2 + y^2 + z^2},$$
$$\bar{\phi} = \tan^{-1} \frac{z}{\sqrt{x^2 + y^2}}, \tag{6-13}$$
$$\lambda = \tan^{-1} \frac{y}{x}.$$

According to Sec. 2-9, the potential of normal gravitation V may be written in the form

$$V = \frac{kM}{r}\left[1 - \sum_{n=1}^{\infty} J_{2n}\left(\frac{a}{r}\right)^{2n} P_{2n}(\sin\bar{\phi})\right]. \tag{6-14}$$

The potential of normal gravity U is then given by

$$U = V + \Phi, \tag{6-15}$$

where Φ is the centrifugal potential. By (2-92), the coefficients J_{2n} are

$$J_{2n} = (-1)^{n+1}\frac{3e^{2n}}{(2n+1)(2n+3)}\left(1 - n + 5n\frac{C-A}{ME^2}\right). \tag{6-16}$$

The components of γ along the coordinate lines are defined by

$$\gamma_r = \frac{\partial U}{\partial r}, \quad \gamma_{\bar{\phi}} = \frac{1}{r}\frac{\partial U}{\partial \bar{\phi}}, \quad \gamma_\lambda = \frac{1}{r\cos\bar{\phi}}\frac{\partial U}{\partial \lambda} = 0. \tag{6-17}$$

These components closely correspond to the components (6-10), because for $E = 0$ we have $u = r, \beta = \bar{\phi}, w = 1$. Thus the rectangular components $\gamma_x, \gamma_y, \gamma_z$ are obtained from (6-12) directly by setting $E = 0$:

$$\begin{aligned}
\gamma_x &= \cos\bar{\phi}\cos\lambda\,\gamma_r - \sin\bar{\phi}\cos\lambda\,\gamma_{\bar{\phi}} - \sin\lambda\,\gamma_\lambda,\\
\gamma_y &= \cos\bar{\phi}\sin\lambda\,\gamma_r - \sin\bar{\phi}\sin\lambda\,\gamma_{\bar{\phi}} + \cos\lambda\,\gamma_\lambda,\\
\gamma_z &= \sin\bar{\phi}\,\gamma_r + \cos\bar{\phi}\,\gamma_{\bar{\phi}}.
\end{aligned} \tag{6-18}$$

These equations also hold when $\gamma_\lambda \neq 0$, but as a matter of fact we have $\gamma_\lambda = 0$ in our case.

It is convenient first to compute the components of the vector of normal gravitation,

$$\boldsymbol{\Gamma} = \text{grad}\, V, \tag{6-19}$$

and then to calculate γ by adding the centrifugal force:

$$\gamma = \boldsymbol{\Gamma} + \text{grad}\,\Phi. \tag{6-20}$$

In xyz-components this equation reads

$$\begin{aligned}\gamma_x &= \Gamma_x + \omega^2 x,\\ \gamma_y &= \Gamma_y + \omega^2 y,\\ \gamma_z &= \Gamma_z.\end{aligned} \tag{6-20'}$$

The vector $\boldsymbol{\Gamma}$ also has considerable interest of its own, because it represents the effect of the normal gravitational attraction of the earth on a satellite.

The components of $\boldsymbol{\Gamma}$ along the coordinate lines are, in analogy to (6-17), given by

$$\Gamma_r = \frac{\partial V}{\partial r}, \quad \Gamma_{\bar\phi} = \frac{1}{r}\frac{\partial V}{\partial \bar\phi}, \quad \Gamma_\lambda = \frac{1}{r\cos\bar\phi}\frac{\partial V}{\partial \lambda} = 0. \tag{6-21}$$

It is easily seen that equation (6-18) also holds if we replace all components of γ by the corresponding components of $\boldsymbol{\Gamma}$.

The components (6-21) are obtained by differentiating (6-14) with respect to r and $\bar\phi$. After elementary manipulations we thus find

$$\begin{aligned}\Gamma_r &= -\frac{kM}{r^2}\left[1 - \sum_{n=1}^{\infty}(2n+1)J_{2n}\left(\frac{a}{r}\right)^{2n} P_{2n}(\sin\bar\phi)\right],\\ \Gamma_{\bar\phi} &= -\frac{kM}{r^2}\sum_{n=1}^{\infty} J_{2n}\left(\frac{a}{r}\right)^{2n}\frac{dP_{2n}}{d\bar\phi}.\end{aligned} \tag{6-22}$$

These equations are suitable for numerical calculations. Since these series converge very rapidly, it is often sufficient to consider terms up to J_4 only.

A slight modification is obtained by introducing

$$G = \frac{kM}{r^2}, \quad G_0 = \frac{kM}{a^2}, \tag{6-23}$$

so that

$$\left(\frac{a}{r}\right)^{2n} = \left(\frac{G}{G_0}\right)^n.$$

Setting

$$\frac{J_{2n}}{G_0^n} = C_{2n} \tag{6-24}$$

we easily obtain

$$\begin{aligned}V &= r\left[G - \sum_{n=1}^{\infty} C_{2n}G^{n+1}P_{2n}(\sin\bar\phi)\right],\\ \Gamma_r &= -G + \sum_{n=1}^{\infty}(2n+1)C_{2n}G^{n+1}P_{2n}(\sin\bar\phi),\\ \Gamma_{\bar\phi} &= -\sum_{n=1}^{\infty} C_{2n}G^{n+1}\frac{dP_{2n}}{d\bar\phi}.\end{aligned} \tag{6-25}$$

These formulas may be used instead of (6-14) and (6-22).

A more explicit form that is directly suited for hand calculations is obtained by expressing the P_{2n} and $dP_{2n}/d\bar\phi$ in powers of $\cos 2\bar\phi$. Substituting

$$t^2 = \sin^2 \bar{\phi} = \frac{1}{2} - \frac{1}{2}\cos 2\bar{\phi}$$

into (1-58) we find

$$P_2 = \frac{1}{4} - \frac{3}{4}\cos 2\bar{\phi}, \quad P_4 = -\frac{13}{32} - \frac{5}{16}\cos 2\bar{\phi} + \frac{35}{32}\cos^2 2\bar{\phi}.$$

These equations are easily differentiated with respect to $\bar{\phi}$, yielding $dP_{2n}/d\bar{\phi}$. On insertion into (6-25) we get, retaining terms up to $n = 2$ only,

$$\begin{aligned} V &= r\Big(G - \frac{1}{4}C_2G^2 + \frac{3}{4}C_2G^2\cos 2\bar{\phi} \\ &\qquad + \frac{13}{32}C_4G^3 + \frac{5}{16}C_4G^3\cos 2\bar{\phi} - \frac{35}{32}C_4G^3\cos^2 2\bar{\phi}\Big), \\ \Gamma_r &= -G + \frac{3}{4}C_2G^2 - \frac{9}{4}C_2G^2\cos 2\bar{\phi} \\ &\qquad - \frac{65}{32}C_4G^3 - \frac{25}{16}C_4G^3\cos 2\bar{\phi} + \frac{175}{32}C_4G^3\cos^2 2\bar{\phi}, \\ \Gamma_{\bar{\phi}} &= \sin\bar{\phi}\cos\bar{\phi}\Big(-3\,C_2G^2 - \frac{5}{4}C_4G^3 + \frac{35}{4}C_4G^3\cos 2\bar{\phi}\Big). \end{aligned} \tag{6-26}$$

With the numerical values of the international ellipsoid (Sec. 2-11), formulas (6-23) and (6-26) become

$$G = \frac{3986\,3290.45}{r^2} \qquad (r \text{ always in kilometers}) \tag{6-27}$$

and

$$\begin{aligned} V &= r(1000G - 0.2786G^2 + 0.8359G^2\cos 2\bar{\phi} \\ &\qquad - 0.0010G^3 - 0.0008G^3\cos 2\bar{\phi} + 0.0028G^3\cos^2 2\bar{\phi}), \\ \Gamma_r &= -1000G + 0.8359G^2 - 2.5077G^2\cos 2\bar{\phi} \\ &\qquad + 0.0051G^3 + 0.0039G^3\cos 2\bar{\phi} - 0.0138G^3\cos^2 2\bar{\phi}, \\ \Gamma_{\bar{\phi}} &= \sin\bar{\phi}\cos\bar{\phi}\,(-3.3436G^2 + 0.0031G^3 - 0.0220G^3\cos 2\bar{\phi}). \end{aligned} \tag{6-28}$$

These expressions give V in geopotential units (1 g.p.u. = 1000 gal·meters) and Γ_r and $\Gamma_{\bar{\phi}}$ in gals, to an accuracy of 1 mgal.[1]

After the computation of Γ_r and $\Gamma_{\bar{\phi}}$, Γ_λ being zero, the rectangular components Γ_x, Γ_y, Γ_z are obtained by means of (6-18), where the components of γ are to be replaced by those of $\boldsymbol{\Gamma}$. If the components of γ are needed, they may be computed from (6-20′).

[1] Series expansions of higher order will be found in Hirvonen and Moritz (1963, p. 12). We have in (6-28) adopted the general notations of this paper, but the derivation is different. It should be noted that Hirvonen defines Γ_r and $\Gamma_{\bar{\phi}}$ with opposite sign and that he denotes the geocentric latitude by ψ.

6-4. Gravity Disturbances—Direct Method

It is convenient to start with the components δ_r, $\delta_{\bar{\phi}}$, δ_λ of the gravity disturbance vector δ, equation (6-3), in the spherical coordinates r, $\bar{\phi}$, λ that were used in the preceding section. In analogy to (6-17) we have[1]

$$\delta_r = \frac{\partial T}{\partial r}, \quad \delta_{\bar{\phi}} = \frac{1}{r}\frac{\partial T}{\partial \bar{\phi}}, \quad \delta_\lambda = \frac{1}{r\cos\bar{\phi}}\frac{\partial T}{\partial \lambda}. \tag{6-29}$$

The disturbing potential T may be expressed in terms of the free-air anomalies at the earth's surface by the formula of Pizzetti, equations (2-161) and (2-162),

$$T_P = T(r, \bar{\phi}, \lambda) = \frac{R}{4\pi}\iint_\sigma \Delta g\, S(r, \psi)\, d\sigma, \tag{6-30}$$

where $S(r, \psi)$ is the extended Stokes function,

$$S(r, \psi) = \frac{2R}{l} + \frac{R}{r} - 3\frac{Rl}{r^2} - \frac{R^2}{r^2}\cos\psi\left(5 + 3\ln\frac{r - R\cos\psi + l}{2r}\right), \tag{6-31}$$

and

$$l = \sqrt{r^2 + R^2 - 2Rr\cos\psi}. \tag{6-32}$$

According to (6-29), we must differentiate (6-30) with respect to r, $\bar{\phi}$, and λ. Here we note that the integral on the right-hand side of (6-30) depends on r, $\bar{\phi}$, λ only through the function $S(r, \psi)$. Thus, Δg being constant with respect to the differentiation, we have

$$\begin{aligned}
\delta_r &= \frac{R}{4\pi}\iint_\sigma \Delta g \frac{\partial S(r, \psi)}{\partial r}\, d\sigma,\\
\delta_{\bar{\phi}} &= \frac{R}{4\pi r}\iint_\sigma \Delta g \frac{\partial S(r, \psi)}{\partial \bar{\phi}}\, d\sigma,\\
\delta_\lambda &= \frac{R}{4\pi r\cos\bar{\phi}}\iint_\sigma \Delta g \frac{\partial S(r, \psi)}{\partial \lambda}\, d\sigma.
\end{aligned} \tag{6-33}$$

The point P at which δ is to be computed has the coordinates $\bar{\phi}$, λ; let the corresponding coordinates of the variable point P', to which Δg and $d\sigma$ refer, be denoted by $\bar{\phi}'$, λ'. Then $d\sigma$ will be expressed by

$$d\sigma = \cos\bar{\phi}'\, d\bar{\phi}'\, d\lambda' \tag{6-34}$$

and ψ, the angular distance between P and P', becomes

$$\psi = \cos^{-1}[\sin\bar{\phi}\sin\bar{\phi}' + \cos\bar{\phi}\cos\bar{\phi}'\cos(\lambda' - \lambda)]. \tag{6-35}$$

We have

$$\frac{\partial S(r, \psi)}{\partial \bar{\phi}} = \frac{\partial S(r, \psi)}{\partial \psi}\frac{\partial \psi}{\partial \bar{\phi}}, \quad \frac{\partial S(r, \psi)}{\partial \lambda} = \frac{\partial S(r, \psi)}{\partial \psi}\frac{\partial \psi}{\partial \lambda}. \tag{6-36}$$

[1] For comparison, we note that in Hirvonen and Moritz (1963) the notations $\delta_n = -\delta_r$, $\delta_m = \delta_{\bar{\phi}}$, $\delta_l = \delta_\lambda$ are used.

Now we recall the corresponding derivations of Sec. 2-22, leading to Vening Meinesz' formula. As a spherical approximation, which is sufficient for T, δ, etc., we may identify the geocentric latitude $\bar{\phi}$ with the geographical latitude ϕ. Thus equations (6-36) and (2-206) are completely analogous, and (2-209) may be borrowed from Sec. 2-22:

$$\frac{\partial\psi}{\partial\bar{\phi}} = -\cos\alpha, \quad \frac{\partial\psi}{\partial\lambda} = -\cos\bar{\phi}\sin\alpha. \tag{6-37}$$

The azimuth α is given by formula (2-212):

$$\tan\alpha = \frac{\cos\bar{\phi}'\sin(\lambda' - \lambda)}{\cos\bar{\phi}\sin\bar{\phi}' - \sin\bar{\phi}\cos\bar{\phi}'\cos(\lambda' - \lambda)}. \tag{6-38}$$

By means of (6-36) and (6-37), equations (6-33) become

$$\delta_r = \frac{R}{4\pi}\iint_\sigma \Delta g \frac{\partial S(r,\psi)}{\partial r}\, d\sigma, \tag{6-39a}$$

$$\begin{aligned} \delta_{\bar{\phi}} &= -\frac{R}{4\pi r}\iint_\sigma \Delta g \frac{\partial S(r,\psi)}{\partial\psi}\cos\alpha\, d\sigma, \\ \delta_\lambda &= -\frac{R}{4\pi r}\iint_\sigma \Delta g \frac{\partial S(r,\psi)}{\partial\psi}\sin\alpha\, d\sigma. \end{aligned} \tag{6-39b}$$

Now we shall form the derivatives of the extended Stokes function (6-31) with respect to r and ψ. By differentiating (6-32) we get

$$\frac{\partial l}{\partial r} = \frac{r - R\cos\psi}{l}, \quad \frac{\partial l}{\partial\psi} = \frac{Rr}{l}\sin\psi. \tag{6-40}$$

By means of these auxiliary relations we find

$$\begin{aligned} \frac{\partial S}{\partial r} = &-\frac{R(r^2 - R^2)}{rl^3} - \frac{4R}{rl} - \frac{R}{r^2} + \frac{6Rl}{r^3} \\ &+ \frac{R^2}{r^3}\cos\psi\left(13 + 6\ln\frac{r - R\cos\psi + l}{2r}\right), \end{aligned} \tag{6-41}$$

$$\begin{aligned} \frac{\partial S}{\partial\psi} = \sin\psi\Bigg[&-\frac{2R^2 r}{l^3} - \frac{6R^2}{rl} + \frac{8R^2}{r^2} \\ &+ \frac{3R^2}{r^2}\left(\frac{r - R\cos\psi - l}{l\sin^2\psi} + \ln\frac{r - R\cos\psi + l}{2r}\right)\Bigg]. \end{aligned} \tag{6-42}$$

Somewhat more convenient expressions are obtained by substituting

$$t = \frac{R}{r}, \tag{6-43}$$

$$D = \frac{l}{r} = \sqrt{1 - 2t\cos\psi + t^2}. \tag{6-44}$$

Then the extended Stokes function (6-31) and its derivatives (6-41) and (6-42) become

$$S(r, \psi) = t\left[\frac{2}{D} + 1 - 3D - t\cos\psi\left(5 + 3\ln\frac{1 - t\cos\psi + D}{2}\right)\right], \tag{6-45}$$

$$\frac{\partial S(r, \psi)}{\partial r} = -\frac{t^2}{R}\left[\frac{1-t^2}{D^3} + \frac{4}{D} + 1 - 6D - t\cos\psi\left(13 + 6\ln\frac{1 - t\cos\psi + D}{2}\right)\right], \tag{6-46a}$$

$$\frac{\partial S(r, \psi)}{\partial \psi} = -t^2 \sin\psi\left(\frac{2}{D^3} + \frac{6}{D} - 8 - 3\,\frac{1 - t\cos\psi - D}{D\sin^2\psi} - 3\ln\frac{1 - t\cos\psi + D}{2}\right). \tag{6-46b}$$

These expressions are used in (6-30) and (6-39) to compute T and δ.

The separation N_P of the geopotential surface through P, $W = W_P$, and the corresponding spheropotential surface $U = W_P$ is according to Bruns' theorem given by

$$N_P = \frac{T_P}{\gamma_Q}; \tag{6-47}$$

see Sec. 2-15 and Fig. 2-15.

The deflection of the vertical, which is the deviation of the actual plumb line from the normal plumb line at P, is represented by its north-south and east-west components,

$$\xi_P = -\frac{1}{r}\frac{\partial N_P}{\partial \bar{\phi}}, \quad \eta_P = -\frac{1}{r\cos\bar{\phi}}\frac{\partial N_P}{\partial \lambda}; \tag{6-48}$$

these equations correspond to (2-204). Since γ varies very little with latitude and is independent of longitude, we have

$$\frac{\partial N_P}{\partial \bar{\phi}} = \frac{\partial}{\partial \bar{\phi}}\left(\frac{T_P}{\gamma_Q}\right) = \frac{1}{\gamma_Q}\frac{\partial T_P}{\partial \bar{\phi}} - \frac{T_P}{\gamma_Q^2}\frac{\partial \gamma_Q}{\partial \bar{\phi}} \doteq \frac{1}{\gamma_Q}\frac{\partial T_P}{\partial \bar{\phi}}$$

and

$$\frac{\partial N_P}{\partial \lambda} = \frac{1}{\gamma_Q}\frac{\partial T_P}{\partial \lambda}.$$

Comparison of (6-29) and (6-48) shows that

$$\xi_P = -\frac{1}{\gamma_Q}\delta_{\bar{\phi}}, \quad \eta_P = -\frac{1}{\gamma_Q}\delta_\lambda. \tag{6-49}$$

We see that N_P, ξ_P, η_P are given by equations (6-30) and (6-39b), apart from the factor $\pm 1/\gamma_Q$. Hence these equations are the extensions of Stokes' and Vening Meinesz' formulas for points outside the earth and reduce to these formulas for $r = R$, $t = 1$.

Writing equations (6-49) in the form

$$\delta_{\bar{\phi}} = -\gamma\xi, \quad \delta_\lambda = -\gamma\eta, \tag{6-49'}$$

we see that the horizontal components of $\boldsymbol{\delta}$ are directly related to the deflection of the vertical, which is the difference *in direction* of the vectors g and γ. The radial component δ_r, however, represents the difference *in magnitude* of these vectors, since as a spherical approximation

$$-\delta_r = \delta g = g_P - \gamma_P, \tag{6-50}$$

which is the scalar gravity disturbance; see Sec. 2-13.

6-5. Gravity Disturbances—Coating Method

An alternative method for computing T and $\boldsymbol{\delta}$ (Orlin, 1959) uses the fact that the disturbing masses may be replaced by a surface layer, or coating, on the reference ellipsoid, without changing the external potential. According to a theorem of potential theory this is rigorously possible if the geoid encloses the total mass of the earth. In the case of the actual earth this is possible to a very good approximation.

In accordance with Sec. 1-3, we represent the disturbing potential in the form (1-16),

$$T_P = T = k \iint_S \frac{\kappa}{l}\, dS. \tag{6-51}$$

The surface S is the reference ellipsoid, which as a spherical approximation is considered as a sphere of radius R. We now have to determine the surface density κ of the coating.

On the ellipsoid S (at sea level) the normal derivative of T is the "outer derivative" (1-17a):

$$\frac{\partial T}{\partial n} = -2\pi k\kappa + k \iint_S \kappa \frac{\partial}{\partial n}\left(\frac{1}{l}\right) dS. \tag{6-52}$$

Generally, according to (6-40),

$$\frac{\partial}{\partial n}\left(\frac{1}{l}\right) = \frac{\partial}{\partial r}\left(\frac{1}{l}\right) = -\frac{r - R\cos\psi}{l^3},$$

and at sea level ($r = R$),

$$\frac{\partial}{\partial n}\left(\frac{1}{l}\right) = -\frac{R(1-\cos\psi)}{8R^3\sin^3(\psi/2)} = -\frac{1}{4R^2\sin(\psi/2)} = -\frac{1}{2Rl}.$$

Hence (6-52) becomes

$$\frac{\partial T}{\partial n} = -2\pi k\kappa - \frac{k}{2R}\iint_S \frac{\kappa}{l}\, dS$$

and, by (6-51),

$$\frac{\partial T}{\partial n} = -2\pi k\kappa - \frac{T}{2R}. \tag{6-53}$$

We now set

$$2\pi k\kappa = \mu, \tag{6-54}$$

so that (6-53) may be written as

$$\mu = -\frac{\partial T}{\partial n} - \frac{T}{2R}. \tag{6-55}$$

Finally we express $\partial T/\partial n$ in terms of the gravity anomaly Δg by the "fundamental equation of physical geodesy" (2-151f)

$$-\frac{\partial T}{\partial n} = \Delta g + \frac{2T}{R}, \tag{6-56}$$

obtaining

$$\mu = \Delta g + \frac{3T}{2R} = \Delta g + \frac{3G}{2R}N; \tag{6-57}$$

G is mean gravity at sea level, and N denotes the geoidal undulation.

Thus the density μ of the coating can be computed if both Δg and N are known.

After expressing κ in terms of μ according to (6-54), the disturbing potential (6-51) becomes

$$T_P = \frac{1}{2\pi}\iint_S \frac{\mu}{l}\,dS = \frac{R^2}{2\pi}\iint_\sigma \frac{\mu}{l}\,d\sigma, \tag{6-58}$$

because as a spherical approximation $dS = R^2\,d\sigma$; the symbols $d\sigma$ and l have the same meaning as in the preceding section.

To form the components (6-29) of the gravity disturbance δ, we must differentiate (6-58) in exactly the same way that we differentiated (6-30) in the preceding section. Instead of

$$\frac{R}{4\pi}S(r, \psi)$$

we now have

$$\frac{R^2}{2\pi}\frac{1}{l},$$

and μ takes the place of Δg. We find the expressions

$$\begin{aligned}
\delta_r &= \frac{R^2}{2\pi}\iint_\sigma \mu\frac{\partial}{\partial r}\left(\frac{1}{l}\right)d\sigma,\\
\delta_{\bar\phi} &= -\frac{R^2}{2\pi r}\iint_\sigma \mu\frac{\partial}{\partial\psi}\left(\frac{1}{l}\right)\cos\alpha\,d\sigma,\\
\delta_\lambda &= -\frac{R^2}{2\pi r}\iint_\sigma \mu\frac{\partial}{\partial\psi}\left(\frac{1}{l}\right)\sin\alpha\,d\sigma,
\end{aligned} \tag{6-59}$$

which are comparable to (6-39). The derivatives with respect to r and ψ are found by using (6-40), so that we have

$$\delta_r = -\frac{R^2}{2\pi}\iint_\sigma \mu \frac{r - R\cos\psi}{l^3}\, d\sigma,$$

$$\begin{Bmatrix}\delta_{\bar\phi}\\ \delta_\lambda\end{Bmatrix} = \frac{R^2}{2\pi}\iint_\sigma \mu \frac{R}{l^3}\sin\psi \begin{Bmatrix}\cos\alpha\\ \sin\alpha\end{Bmatrix} d\sigma. \tag{6-60}$$

On substituting (6-43) and (6-44), equations (6-58) and (6-60) finally become

$$T_P = \frac{Rt}{2\pi}\iint_\sigma \frac{\mu}{D}\, d\sigma, \tag{6-61}$$

$$\delta_r = -\frac{t^2}{2\pi}\iint_\sigma \mu \frac{1 - t\cos\psi}{D^3}\, d\sigma, \tag{6-62a}$$

$$\begin{Bmatrix}\delta_{\bar\phi}\\ \delta_\lambda\end{Bmatrix} = \frac{t^3}{2\pi}\iint_\sigma \mu \frac{\sin\psi}{D^3} \begin{Bmatrix}\cos\alpha\\ \sin\alpha\end{Bmatrix} d\sigma. \tag{6-62b}$$

Again, equations (6-61) and (6-62b) together with (6-47) and (6-49), may be used to compute the separation of corresponding geopotential and spheropotential surfaces, and the deflection of the vertical.

The coating method presupposes the geoidal heights N to be given in addition to the gravity anomalies Δg.

6-6. Gravity Disturbances—Upward Continuation

We apply Poisson's integral formula (1-89) to the harmonic function T:

$$T_P = \frac{R(r^2 - R^2)}{4\pi}\iint_\sigma \frac{T}{l^3}\, d\sigma. \tag{6-63}$$

In the neighborhood of P (Fig. 6-1), the sphere practically coincides with its tangent plane at F. Since the value of the integrand is very small at greater distances from P, we may extend the integration over the tangent plane instead of over the sphere. Then, according to Fig. 6-1,

$$l = \sqrt{s^2 + H^2}. \tag{6-64a}$$

We introduce a rectangular coordinate system x, y, z, the x-axis pointing north and the y-axis pointing east in the tangent plane. Then we may also write

$$l = \sqrt{x^2 + y^2 + H^2}, \tag{6-64b}$$

the surface element becomes

$$R^2\, d\sigma \doteq dx\, dy,$$

and we further have

$$r = R + H,$$

$$r^2 - R^2 = (r + R)(r - R) \doteq 2RH.$$

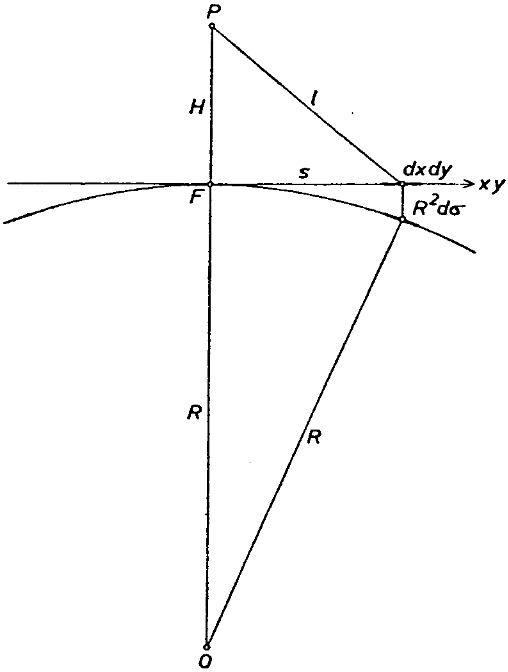

FIGURE 6-1
The plane approximation.

Thus (6-63) becomes the plane formula

$$T_P = \frac{H}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{T}{l^3}\,dx\,dy = \frac{H}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{T}{(x^2+y^2+H^2)^{3/2}}\,dx\,dy. \quad (6\text{-}65)$$

This important formula is called the "*upward continuation integral.*" It permits the computation of the value of the harmonic function T at a point above the xy-plane from the values of T given on the plane, that is, the upward continuation of a harmonic function. Both T and its partial derivatives, $\partial T/\partial x$, $\partial T/\partial y$, $\partial T/\partial z$, are harmonic, because if

$$\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2} = 0,$$

then we also have

$$\frac{\partial^2}{\partial x^2}\left(\frac{\partial T}{\partial x}\right)+\frac{\partial^2}{\partial y^2}\left(\frac{\partial T}{\partial x}\right)+\frac{\partial^2}{\partial z^2}\left(\frac{\partial T}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2}\right) = 0.$$

Thus the upward continuation integral (6-65), which holds for any harmonic function, may also be applied to $\partial T/\partial x$, $\partial T/\partial y$, and $\partial T/\partial z$.

As T is the disturbing potential, its partial derivatives are the components of the gravity disturbance:

$$\frac{\partial T}{\partial x} = \delta_{\bar{\phi}}, \quad \frac{\partial T}{\partial y} = \delta_{\lambda}, \quad \frac{\partial T}{\partial z} = \delta_r.$$

We are not writing δ_x, δ_y, δ_z because we wish to reserve this notation for the components in the geocentric global coordinate system, which should not be confused with the local system introduced in this section. Thus we have in addition to (6-65)

$$\delta_r = \frac{H}{2\pi} \iint_{-\infty}^{\infty} \frac{\delta_r}{l^3}\, dx\, dy, \tag{6-66a}$$

$$\delta_{\bar{\phi}} = \frac{H}{2\pi} \iint_{-\infty}^{\infty} \frac{\delta_{\bar{\phi}}}{l^3}\, dx\, dy,$$

$$\delta_\lambda = \frac{H}{2\pi} \iint_{-\infty}^{\infty} \frac{\delta_\lambda}{l^3}\, dx\, dy. \tag{6-66b}$$

On the left-hand side of these equations, the components of $\boldsymbol{\delta}$ refer to the elevated point P; in the integral on the right-hand side, they are taken at sea level and are to be computed from the expressions

$$\delta_r = -\delta g = -\left(\Delta g + \frac{2G}{R} N\right), \tag{6-67a}$$

$$\delta_{\bar{\phi}} = -G\xi,$$
$$\delta_\lambda = -G\eta, \tag{6-67b}$$

which follow from (6-49′) and (6-50), applied to sea level, together with (2-151d). The symbols R and G denote, as usual, a mean earth radius and a mean value of gravity on the earth's surface.

Hence we may compute T and $\boldsymbol{\delta}$ by means of the upward continuation integral if the geoidal undulations N and the deflection components ξ and η at the earth's surface are given.

The plane approximation is sufficient except for very high altitudes (>250 km, say). Otherwise we must use the spherical formula (6-63) for T. For the radial component δ_r, equations (6-74) or (6-75) below, with δ_r replacing Δg, may be shown to hold. The corresponding spherical formulas for the upward continuation of the horizontal components $\delta_{\bar{\phi}}$ and δ_λ are not known. The reason why the same formula, the upward continuation integral, holds for T and the components of $\boldsymbol{\delta}$ in the planar case only is that the derivatives of T are harmonic only when referred to a cartesian coordinate system.

6-7. Additional Considerations

Reference surface. The preceding formulas for the disturbing potential T and the gravity disturbance vector $\boldsymbol{\delta}$ are rigorously valid if the reference surface is a sphere. In practice the gravity anomalies are referred to an ellipsoid. The above formulas for T and $\boldsymbol{\delta}$ are also valid for an ellipsoidal reference surface if a relative error of the order of the flattening $f \doteq 0.3\%$ is neglected, that is, as a spherical approximation. The reader is reminded that this does not mean that the ellipsoid is replaced by a sphere in any geometrical sense; rather it means that in the originally elliptical formulas the first and higher powers of the flattening are neglected, whereby they formally become spherical formulas.

Since the gravity anomalies, etc., are referred to an ellipsoid, we must be very careful in computing t, which enters into the formulas of Secs. 6-4 and 6-5. If an exact sphere of radius R were used as a reference surface, then we should have $r = R + H$, where H is the elevation of the computation point above the sphere. Actually, we use a reference ellipsoid; then we again have

$$r = R + H, \quad t = \frac{R}{R + H}, \tag{6-68}$$

but H is now the elevation *above the ellipsoid* (or, to a sufficient accuracy, *above sea level*), the constant $R = 6371$ km being the earth's mean radius. Thus r as computed by (6-68) differs from the geocentric radius vector $r = \sqrt{x^2 + y^2 + z^2}$. As a matter of fact, this holds only for Secs. 6-4 and 6-5, but not for the formulas of Sec. 6-3, which rigorously refer to spherical coordinates.

We have already mentioned that we may replace the geocentric latitude $\bar{\phi}$ by the geographic latitude ϕ, as far as T and δ are concerned—for instance, by putting $\bar{\phi} = \phi$ in (6-35) or (6-38).

Data. For all computations dealing with the external gravity field of the earth, *free-air gravity anomalies* must be used for Δg, since all other types of gravity anomalies correspond to some removal or transport of masses whereby the external field is changed. If, in addition to Δg, geoidal undulations N (in the coating method) or deflections of the vertical ξ, η (in the upward continuation) are used, then these quantities should be computed from free-air anomalies.

If, as is usually done, the normal free-air gradient $\partial\gamma/\partial h \doteq 0.3086$ mgal/meter is used for the free-air reduction, then the free-air anomalies refer, strictly speaking, to the earth's physical surface (to ground level) rather than to the geoid (to sea level). The N-values computed from them by Stokes' formula are height anomalies ζ, referring to the ground, rather than heights of the actual geoid. However, this distinction is insignificant and can be ignored in most cases, so that we may consider Δg as sea-level anomalies (see Sec. 8-13).

If we cannot neglect this distinction, aiming at highest accuracy in high and steep mountains for low altitudes H, then we may proceed as follows (see Secs. 8-8 and 8-10). We reduce the free-air anomaly Δg from the ground point A to the corresponding point A_0 at sea level (see Fig. 6-2):

$$\Delta g^* = \Delta g - \frac{\partial \Delta g}{\partial h} h \tag{6-69}$$

and use the sea level anomaly Δg^* so obtained. The vertical gradient $\partial\Delta g/\partial h$ may be computed by formula (2-217) using the ground-level anomalies Δg. Or we may reduce to any other level surface $W = W_1$, for instance to that

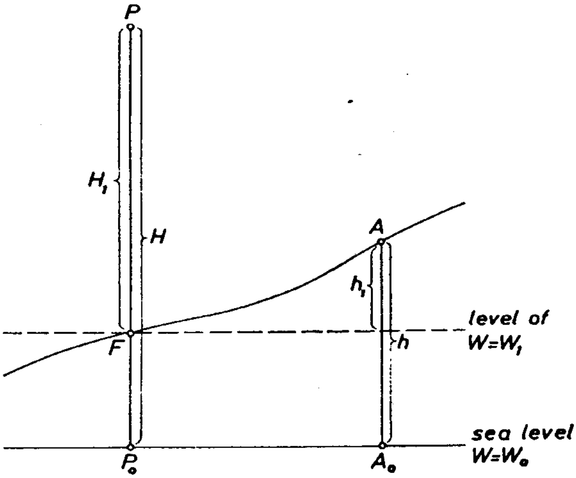

FIGURE 6-2

Reduction to sea level and to the level of F.

passing through F (Fig. 6-2), using h_1 instead of h in (6-69). Then we should also use H_1, rather than H, in (6-68). For large-scale purposes, reduction to sea level appears to be preferable. Probably such a reduction will attain a considerable amount only in exceptional cases, so that it can usually be neglected, and H in the formulas of Secs. 6-4 through 6-6 may be taken as the height of P above sea level or above ground. For other methods of considering the topography see Arnold (1959), Brovar (1963), Levallois (1960), and Moritz (1966).

Comparison of methods. Of all methods described in the three preceding sections, the formulas for the direct method are the most complicated, but they can be handled very well if the required functions have been tabulated or if the formulas are programmed for an automatic computer. Only the gravity anomalies Δg are required here. If the geoidal heights N are known in addition to Δg, then the coating method is preferable because it involves somewhat simpler formulas. Although computation is simplest in the upward continuation method, it requires the largest amount of data: N for T, Δg and N for δ_r, and ξ and η for $\delta_{\bar{\phi}}$ and δ_λ.

To get a better insight into the applicability of these three methods, we shall consider the effect of the distant zones. Table 6-1, taken from Hirvonen and Moritz (1963, p. 63), shows the root mean square influence $\overline{\Delta\delta_r}$, $\overline{\Delta\delta_{\bar{\phi}}} = \overline{\Delta\delta_\lambda}$ of the zones beyond a spherical radius ψ_0 on δ_r, $\delta_{\bar{\phi}}$, and δ_λ. The method by which this table was computed is essentially that to be described in Sec. 7-4. The values in this table hold for all altitudes H from zero up to several hundred kilometers.

We recognize that for $\psi_0 > 20°$ or $30°$ the influence of the distant zones decreases very slowly. Therefore, it seems to be impractical to extend the inte-

Table 6-1

R.M.S. Influence of the Zone beyond a Radius ψ_0 on δ_r, $\delta_{\bar\phi}$, δ_λ.

ψ_0	Direct Method $\overline{\Delta\delta_r}$ (mgal)	Direct Method $\overline{\Delta\delta_{\bar\phi}} = \overline{\Delta\delta_\lambda}$ (mgal)	Coating Method $\overline{\Delta\delta_r}$ (mgal)	Coating Method $\overline{\Delta\delta_{\bar\phi}} = \overline{\Delta\delta_\lambda}$ (mgal)
9.0° (1000 km)	8	8	2	6
13.5° (1500 km)	7	7	2	5
18.0° (2000 km)	6	6	2	5
20°	6	6	2	5
25°	5	5	2	5
30°	5	5	1	5
45°	5	5	1	5
60°	4	5	1	4
90°	3	4	1	3
120°	2	3	1	2
150°	2	3	0	2
180°	0	0	0	0

gration much farther than 20° (coating method) or 30° (direct method), unless it is extended over the *whole* earth.

We further see that the effect of the remote zones on $\delta_{\bar\phi}$ and δ_λ is not very much less in the coating method than in the direct method. The influence on δ_r is less in the coating method; but if we know N in addition to Δg, then we should compute δ_r not by this method but by upward continuation, where the influence of the distant zones is negligibly small.

It is easily understood why this influence is so small in the upward continuation method. If $H = 0$, then the effect of the remote zones in the direct method and in the coating method is still given by Table 6-1. In the method of upward continuation, however, this effect is zero for $H = 0, P = F$, because the "computed" value at P is then identical with the corresponding ground value at F, neighboring values being without influence. If $H \neq 0$, then only the nearest neighborhood of P is of any relevance in this method. We shall see in the next section that it is usually sufficient to go as far as ten times the elevation if upward continuation is used.

This is also the reason why we were able to use the plane approximation in the method of Sec. 6-6, but not in the other methods, which involve larger distances for which this approximation breaks down.

As a summary, the following methods are suitable for practical use: if only Δg is given, the direct method; if Δg and N are given, the coating method for the horizontal components and the upward continuation for the vertical component of δ and for T; if Δg, N, ξ, η are given, then the upward continuation throughout.

The accuracy obtainable is about the same with all three methods if properly applied, in particular if the integration is extended sufficiently far. The standard errors of all three components are approximately proportional to $1/H$ and are very small for large elevations, but the correlation between neighboring values may be appreciable.

Practical integration. The integral formulas of this chapter must be evaluated approximately by summations in precisely the same way as, for instance, Stokes' and Vening Meinesz' formulas. The procedures were described in Sec. 2-24.

Details on the upward continuation method will be found in the following section. As for the direct method and the coating method using standard-sized blocks, we mention that the following sizes are considered appropriate at about 45° latitude. Up to a latitude difference from the computation point of $\Delta\phi = 1.5°$ and a longitude difference of $\Delta\lambda = 2°$ use $5' \times 5'$ blocks; outside this zone, up to $\Delta\phi = 3.5°$ and $\Delta\lambda = 4.5°$, use $20' \times 20'$ blocks; outside this zone, up to $\Delta\phi = 12.5°$ and $\Delta\lambda = 15°$, use $1° \times 1°$ blocks; and outside this zone, $5° \times 5°$ blocks.

For points at elevations of only a few kilometers, even $5' \times 5'$ blocks alone may not be sufficient around the computation point, and additional considerations may be necessary, such as the use of a template for the innermost region or the employment of horizontal gravity gradients analogous to that in Vening Meinesz' formula.

The details of these numerical integrations are thus somewhat intricate; the reader may find more about them in Hirvonen and Moritz (1963).

Computation of the gravity vector. After computing the components δ_r, $\delta_{\bar{\phi}}$, δ_λ by numerical integration, we may transform them into cartesian coordinates δ_x, δ_y, δ_z with respect to the world coordinate system. The transformation equations are (6-18), the components of γ being replaced by the corresponding components of $\boldsymbol{\delta}$; it is easily seen that (6-18) holds for an arbitrary vector.

We may also first form the components of the gravity vector $\mathbf{g}$ in spherical coordinates by

$$g_r = \gamma_r + \delta_r, \quad g_{\bar{\phi}} = \gamma_{\bar{\phi}} + \delta_{\bar{\phi}}, \quad g_\lambda = \delta_\lambda, \tag{6-70}$$

where γ_r, $\gamma_{\bar{\phi}}$, γ_λ are given by the formulas of Sec. 6-3, and then apply (6-18) to $\mathbf{g}$:

$$\begin{aligned} g_x &= \cos\bar{\phi}\cos\lambda\, g_r - \sin\bar{\phi}\cos\lambda\, g_{\bar{\phi}} - \sin\lambda\, g_\lambda, \\ g_y &= \cos\bar{\phi}\sin\lambda\, g_r - \sin\bar{\phi}\sin\lambda\, g_{\bar{\phi}} + \cos\lambda\, g_\lambda, \\ g_z &= \sin\bar{\phi}\, g_r + \cos\bar{\phi}\, g_{\bar{\phi}}. \end{aligned}$$

Another possibility is to use the components in ellipsoidal coordinates according to Sec. 6-2. For the small quantities δ_u, δ_β, δ_λ we may again apply the spherical approximation, neglecting a relative error of the order of the flattening. If the flattening is neglected, then the ellipsoidal coordinates u, β, λ reduce to the spherical coordinates r, $\bar{\phi}$, λ, so that as a spherical approximation

$$\delta_u = \delta_r, \quad \delta_\beta = \delta_{\bar{\phi}}, \tag{6-71}$$

δ_λ being rigorously the same in both systems. Thus δ_r, $\delta_{\bar\phi}$, δ_λ may also be considered as the components of $\boldsymbol{\delta}$ in ellipsoidal coordinates.

Then we have

$$g_u = \gamma_u + \delta_r, \quad g_\beta = \gamma_\beta + \delta_{\bar\phi}, \quad g_\lambda = \delta_\lambda; \tag{6-72}$$

and g_x, g_y, g_z are obtained by (6-12), the components of g replacing the corresponding components of γ. It is evident that the spherical approximation can only be used for $\boldsymbol{\delta}$, so that γ_u and γ_β must be computed by the rigorous formulas (6-10).

The gravity potential W may be computed by (6-4); the gravitational potential V is obtained by subtracting the centrifugal potential $\omega^2(x^2 + y^2)/2$; and the vector of gravitation is given by (6-6).

Spherical harmonics. The anomalous potential T and its derivatives may also be obtained by means of their spherical-harmonic expansions, whose coefficients are computed by harmonic analysis of gravity anomalies (Sec. 2-20). Because of the slow convergence of these series, however, they can be applied for computations at satellite elevations (around 1000 km) only. They are useful for the computation of satellite orbits; see Secs. 9-6 through 9-8.

6-8. Gravity Anomalies Outside the Earth

Suppose gravity g is to be computed at some point P outside the earth (Fig. 6-3); we shall be concerned here only with the *magnitude* of the gravity vector. This is conveniently done by adding a correction to the normal gravity γ. From Sec. 2-13 we recall two different kinds of such a correction, $g - \gamma$:

1. The *gravity disturbance* δg, in which g and γ both refer to the same point P.
2. The *gravity anomaly* Δg. Here g refers to P, but γ refers to the correspond-

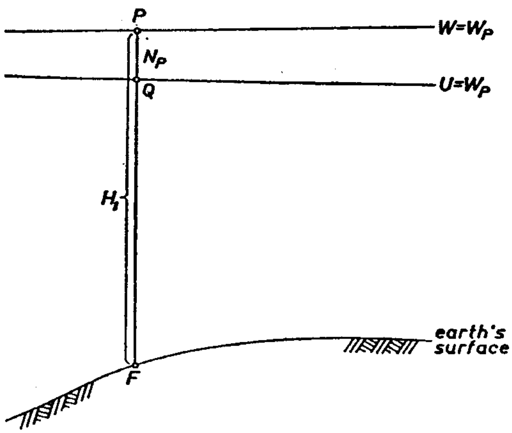

FIGURE 6-3

Gravity anomalies and disturbancies

ing point Q, which is situated on the plumb line of P, and whose normal potential U is the same as the actual potential W of P, that is, $U_Q = W_P$.

These two quantities are connected by

$$\Delta g = \delta g - \frac{2G}{R} N_P;$$

this simple form is sufficient for moderate altitudes.

The *gravity disturbance* is used when the spatial position of P, that is, its geocentric rectangular coordinates x, y, z, is given, such as in the computation of gravity along space trajectories or satellite orbits. Then one usually needs the complete vector **g**, not just its magnitude g, and the computations are done by the methods described in the preceding sections. In Sec. 2-13 we saw that the difference in magnitude δg is practically equal to the vertical component of the gravity disturbance vector:

$$\delta g = -\delta_r.$$

In this section we shall be concerned with the *gravity anomaly* Δg. It is used whenever the natural coordinates (Sec. 2-4), in particular the *potential* W, of P are given. For then we can determine Q as that point whose normal potential is equal to the given value of W; that is, we can compute the height of Q above the ellipsoid by an ellipsoidal formula such as (4-44), with $C = W_0 - W$. Then the normal gravity at Q is given, for instance, by (2-123).

At the earth's surface, the potential W is determined by leveling (Sec. 4-1); this is why gravity anomalies, and not gravity disturbances, are the basic material of gravimetric geodesy. If the height H_1 of P above ground is given, then the potential at P may be obtained by

$$W = W_1 - \bar{g} H_1, \tag{6-73}$$

where W_1 is the potential at the ground point F below P, and $\bar{g}$ is the mean gravity between F and P. Thus even in this case W is given rather than rectangular coordinates x, y, z, and the use of gravity anomalies Δg is appropriate. This is the case, for instance, in airborne gravity measurements, where the height of the aircraft above ground is measured.

Formulas. The basic formula is

$$\Delta g_P = \frac{R^2(r^2 - R^2)}{4\pi r} \iint_\sigma \frac{\Delta g}{l^3} \, d\sigma, \tag{6-74}$$

which differs from (2-160) in that the spherical harmonics of degrees zero and one, which have been excluded there, are left in the present formula. By making the usual substitutions (6-43) and (6-44) we obtain

$$\Delta g_P = \frac{t^2(1 - t^2)}{4\pi} \iint_\sigma \frac{\Delta g}{D^3} \, d\sigma. \tag{6-75}$$

Again

$$t = \frac{R}{R + H},$$

where H is the height above the reference level to which the given anomalies Δg refer; see the corresponding remarks in the preceding section.

Up to flight altitudes we may again use the plane approximation of Sec. 6-6, so that (6-75) reduces to an upward continuation integral of type (6-65):

$$\Delta g_P = \frac{H}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\Delta g}{l^3}\, dx\, dy, \tag{6-76}$$

or in polar coordinates s and α,

$$\Delta g_P = \frac{H}{2\pi} \int_{\alpha=0}^{2\pi} \int_{s=0}^{\infty} \frac{\Delta g}{l^3}\, s\, ds\, d\alpha, \tag{6-76$'$}$$

where

$$l = \sqrt{s^2 + H^2} = \sqrt{x^2 + y^2 + H^2}.$$

Practical integration. We may again use standard blocks ($5' \times 5'$, $10' \times 10'$, or $1° \times 1°$, say), suitable for automatic computation, or we may employ templates.

The integral (6-76) is then replaced by

$$\Delta g_P = \sum_k c_k \overline{\Delta g}_k, \tag{6-77}$$

where $\overline{\Delta g}_k$ is the mean over the kth compartment. If standard blocks of sides $\Delta\phi$ and $\Delta\lambda$ are used, then

$$c_k = \frac{H}{2\pi} \frac{R^2 \Delta\phi\, \Delta\lambda \cos\phi_k}{l_k^3}, \tag{6-78}$$

where ϕ_k and l_k refer to the center of the block. These coefficients are of type (2-224). For a polar template the preferable "integrated" coefficients of type (2-223) also have a simple form. Using (6-76$'$) and the notations of Fig. 2-22 (with ψ replaced by s), we have

$$c_k = \frac{H}{2\pi} \int_{\alpha=\alpha_1}^{\alpha_2} \int_{s=s_1}^{s_2} \frac{s\, ds\, d\alpha}{(s^2 + H^2)^{3/2}} = H \frac{\alpha_2 - \alpha_1}{2\pi} \int_{s_1}^{s_2} \frac{s\, ds}{(s^2 + H^2)^{3/2}},$$

and on performing the integration,

$$c_k = H \frac{\alpha_2 - \alpha_1}{2\pi} \left(\frac{1}{l_1} - \frac{1}{l_2} \right), \tag{6-79}$$

where $l_1 = \sqrt{s_1^2 + H^2}$ and $l_2 = \sqrt{s_2^2 + H^2}$. If the ring between the radii s_1 and s_2 is divided into n compartments, then

$$\frac{\alpha_2 - \alpha_1}{2\pi} = \frac{1}{n},$$

and we obtain

$$c_k = \frac{H}{n}\left(\frac{1}{l_1} - \frac{1}{l_2}\right), \tag{6-79'}$$

where l_1 belongs to the inner and l_2 to the outer radius.

Hirvonen (1962) has made an optimal design for a template. It is constructed so that the error caused by each compartment has the same root mean square size. Table 6-2 gives Hirvonen's coefficients. The radii s_1 and s_2 and the elevation H are to be measured in the same unit.

Table 6-2

Hirvonen's Template Constants

Outer Radius s_2	*Inner Radius* s_1	*Number of Compartments* n	*Coefficients* c_k $H = 1$	$H = 2$	$H = 0.5$
0.4	0	1	0.07152	0.01942	0.21913
1.0	0.4	8	0.02767	0.01077	0.04171
1.8	1.0	12	0.01846	0.01259	0.01496
3.0	1.8	12	0.01412	0.01572	0.00860
4.5	3.0	16	0.00621	0.00928	0.00337
6.7	4.5	16	0.00433	0.00751	0.00225
10.0	6.7	16	0.00301	0.00562	0.00153
15.0	10.0	16	0.00206	0.00400	0.00104
22.0	15.0	16	0.00132	0.00260	0.00066
32.0	22.0	16	0.00089	0.00176	0.00044

As we saw in the preceding section, upward continuation is essentially a local problem. The main contribution to the integrals (6-76) or (6-76′) comes from the area around the computation point P, the influence of distant regions being negligibly small. Let us consider the effect of the zone beyond a certain distance s_0 from P (Fig. 6-4). According to (6-76′), this effect is given by

$$\epsilon = \frac{H}{2\pi}\int_{\alpha=0}^{2\pi}\int_{s=s_0}^{\infty}\frac{\Delta g}{(s^2+H^2)^{3/2}}\, s\, ds\, d\alpha \doteq \frac{H}{2\pi}\int_{\alpha=0}^{2\pi}\int_{s=s_0}^{\infty}\frac{\Delta g}{s^2}\, ds\, d\alpha,$$

because for large s we may replace $l = \sqrt{s^2 + H^2}$ by s. If we introduce a certain average value $\overline{\Delta g}$ of the gravity anomalies in the zone $s > s_0$, then, according to a mean value theorem of integral calculus, we may express the average value of the effect of this zone as

$$\bar{\epsilon} = \frac{H}{2\pi}\overline{\Delta g}\int_0^{2\pi} d\alpha \int_{s_0}^{\infty}\frac{ds}{s^2},$$

which is equal to

$$\bar{\epsilon} = \frac{H}{s_0}\overline{\Delta g}. \tag{6-80}$$

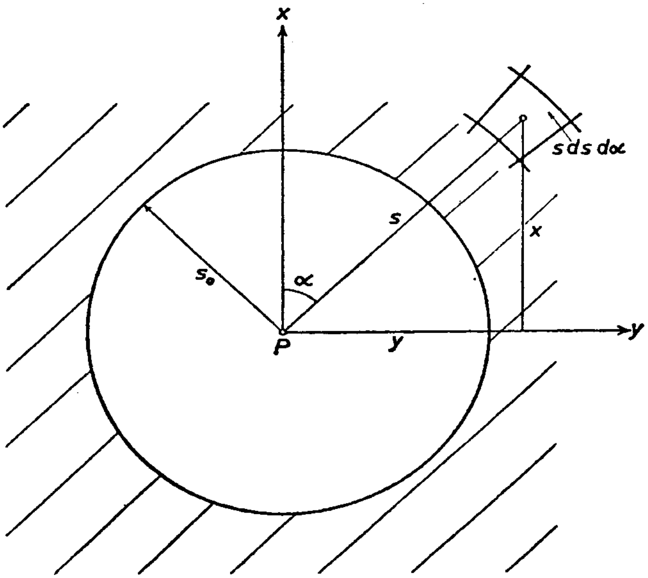

FIGURE 6-4
The zone $s > s_0$.

From this formula we see that s_0 must be roughly proportional to H if we wish to get the same error $\bar{\epsilon}$ for different elevations H. For instance, if $s_0 = 10H$, then $\bar{\epsilon} = 0.1\,\overline{\Delta g}$. If $\overline{\Delta g}$ does not exceed 10 mgals, then $\bar{\epsilon}$ will be smaller than 1 mgal. This can often be assumed, because we may expect that the values of Δg for the zone $s > s_0$ tend to average out for large values of s_0. In such cases it will be sufficient *to extend the integration only as far as 10 times the elevation.*

In many respects the considerations of the preceding section are applicable to the upward continuation of the gravity anomalies as well. Again free-air anomalies referred to ground level or, more accurately, to some level surface, are to be used. If the ground is elevated above sea level, but reasonably flat, it is somewhat better to regard H as elevation above ground rather than above sea level, because the ground may then be considered locally part of a level surface.

For accuracy considerations the reader is referred to Moritz (1962).

The inverse problem, the downward continuation of gravity anomalies, occurs in the reduction of gravity measured on board an aircraft, and also in a certain refined solution of the geodetic boundary-value problem to be described in Sec. 8-10. There is no closed integral formula inverse to (6-75) or (6-76), but the problem of downward continuation may be solved by the iterative method of Sec. 8-10.

Upward and downward continuation are also tools of geophysical exploration, but here the objective is quite different. Several methods have been developed in this connection, some of which are also applicable for geodetic purposes; see, for instance, Jung (1961, Sec. 7.22), Dean (1958), Henderson (1960), and Tsuboi (1961).

References

Arnold, K. (1959). Zur Bestimmung der Geoidundulationen aus Freiluftanomalien. Schwerewerte in grossen Höhen über der Erdoberfläche. Veröff. Geod. Inst. Potsdam, No. 12.

Brovar, V. V. (1963). Computation of the disturbing potential and its derivatives above the earth. English Transl. of the Russian journal "Geodesy and Aerophotography" by Am. Geophys. Union, No. 3, pp. 142–144.

Dean, W. C. (1958). Frequency analysis for gravity and magnetic interpretation. Geophysics, v. XXIII, pp. 97–127.

Henderson, R. G. (1960). A comprehensive system of automatic computation in magnetic and gravity interpretation. Geophysics, v. XXV, pp. 569–585.

Hirvonen, R. A. (1962). On the statistical analysis of gravity anomalies. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 37.

Hirvonen, R. A., and H. Moritz (1963). Practical computation of gravity at high altitudes. Report No. 27, Inst. Geod. Phot. Cart., Ohio State Univ.

Jung, K. (1961). Schwerkraftverfahren in der Angewandten Geophysik. Leipzig, Akademische Verlagsgesellschaft Geest & Portig.

Levallois, J. J. (1960). Sur le problème du champ extérieur de la pesanteur. Paris, Institut Géographique National, No. 26956.

Moritz, H. (1962). Studies on the accuracy of the computation of gravity in high elevations. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 38.

Moritz, H. (1966). The computation of the external gravity field and the geodetic boundary value problem. Proceedings of the Symposium "Extension of Gravity Anomalies to Unsurveyed Areas" (H. Orlin, ed.), Am. Geophys. Union, Geophys. Monograph No. 9, pp. 127–136.

Orlin, H. (1959). The three components of the external anomalous gravity field. J. Geophys. Research, v. 64, pp. 2393–2399.

Tsuboi, C. (1961). Upward continuation of gravity values based on the cylindrical coordinate system. Report No. 16, Inst. Geod. Phot. Cart., Ohio State Univ.

7

Statistical Methods in Physical Geodesy

7-1. Introduction

The most important problems of physical geodesy are formulated and solved in terms of integrals extended over the whole earth. An example is Stokes' formula. Thus, in principle, we need the gravity g at every point of the earth's surface. As a matter of fact, even in the densest gravity net we measure g only at relatively few points, so that we must estimate g at other points by *interpolation*. In large parts of the oceans we have made no observations at all; these gaps must be filled by some kind of *extrapolation*.

Mathematically there is no difference between interpolation and extrapolation; therefore they are denoted by the same term, *prediction*.

Prediction (interpolation or extrapolation) cannot, of course, give exact values; hence, the problem is to estimate the errors that are to be expected in the gravity g or in the gravity anomaly Δg. Since Δg is further used to compute other quantities, such as the geoidal undulation N or the deflection components ξ and η, we must also investigate the influence of the prediction errors of Δg on N, ξ, η, etc. This is called *error propagation*.

It is also important to know which prediction method gives highest accuracy, either in Δg or in derived quantities N, ξ, η, etc. To be able to find these "best" prediction methods, it is obviously necessary to have solved the previous problem, to known the prediction error of Δg and its influence on the derived quantities.

Another question is this. In principle, our integral formulas always involve integrations over the whole earth. In practice, however, integrations are often extended only over a limited area, either because there are no gravity measurements beyond this area, or because there is practically no increase in accuracy if we go farther. The effect of the neglected distant zones is then to be estimated.

Summarizing, we have the following problems:

1. Estimation of interpolation and extrapolation errors of Δg;
2. Estimation of the effect of these errors on derived quantities (N, ξ, η, etc.);
3. Determination of the best prediction method;
4. Estimation of the effect of neglected distant zones.

Since we are interested in the average rather than the individual errors, we are led to a statistical treatment. This will be the topic of the present chapter.

7-2. The Covariance Function

It is quite remarkable that all the problems mentioned above can be solved by means of only one function of one variable, without any other information.[1] This is the *covariance function* of the gravity anomalies.

First we need a measure of the average size of the gravity anomalies Δg. If we form the average of Δg over the whole earth, we get the value zero:

$$M\{\Delta g\} \equiv \frac{1}{4\pi}\iint_\sigma \Delta g \, d\sigma = 0. \tag{7-1}$$

The symbol M stands for the average over the whole earth (over the unit sphere); this average is equal to the integral over the unit sphere divided by its area 4π. The integral is zero if there is no term of degree zero in the expansion of the gravity anomalies Δg into spherical harmonics, that is, if a reference ellipsoid of the same mass as the earth and of the same potential as the geoid is used. This will be assumed throughout this chapter.[2]

Clearly the quantity $M\{\Delta g\}$, which is zero, cannot be used to characterize the average size of the gravity anomalies. Consider then the average square of Δg,

$$\operatorname{var}\{\Delta g\} \equiv M\{\Delta g^2\} = \frac{1}{4\pi}\iint_\sigma \Delta g^2 \, d\sigma. \tag{7-2}$$

It is called the *variance* of the gravity anomalies. Its square root is the *root mean square (r.m.s.) anomaly*:

$$\text{r.m.s.}\{\Delta g\} \equiv \sqrt{\operatorname{var}\{\Delta g\}} = \sqrt{M\{\Delta g^2\}}. \tag{7-3}$$

[1] We are first neglecting the correlation with elevation.

[2] If this is not the case, that is, if $M\{\Delta g\} = m \neq 0$, then we may form new gravity anomalies $\Delta g^* = \Delta g - m$ by subtracting the average value m. Then $M\{\Delta g^*\} = 0$ and all the following developments apply to these "centered" anomalies Δg^*.

The r.m.s. anomaly is a very useful measure of the average size of the gravity anomalies; it is usually given in the form

$$\text{r.m.s. } \{\Delta g\} = \pm 35 \text{ mgals};$$

the plus and minus signs express the ambiguity of the sign of the square root and symbolize that Δg may be either positive or negative. The r.m.s. anomaly is very intuitive; but the variance of Δg is more convenient to handle mathematically and admits of an important generalization.

Instead of the average square of Δg consider the average product of the gravity anomalies $\Delta g\, \Delta g'$ at each pair of points P and P' that are at a constant distance s apart. This average product is called the *covariance* of the gravity anomalies for the distance s and is defined by

$$\text{cov}_s \{\Delta g\} \equiv M \{\Delta g\, \Delta g'\}. \tag{7-4}$$

The average is to be extended over all pairs of points P and P' for which $PP' = s = \text{const}$.

The covariance characterizes the *statistical correlation* of the gravity anomalies Δg and $\Delta g'$, which is their tendency to have about the same size and sign. If the covariance is zero, then the anomalies Δg and $\Delta g'$ are uncorrelated or independent[1] of one another; in other words, the size or sign of Δg has no influence on the size or sign of $\Delta g'$. Gravity anomalies at points that are far apart may be considered uncorrelated or independent, because the local disturbances that cause Δg have almost no influence on $\Delta g'$ and vice versa.

If we consider the covariance as a function of $s = PP'$, then we get the *covariance function* $C(s)$ mentioned at the beginning:

$$C(s) \equiv \text{cov}_s \{\Delta g\} = M \{\Delta g\, \Delta g'\} \qquad (PP' = s). \tag{7-5}$$

For $s = 0$ we have

$$C(0) = M \{\Delta g^2\} = \text{var } \{\Delta g\}, \tag{7-5'}$$

according to (7-2). The covariance for $s = 0$ is the variance.

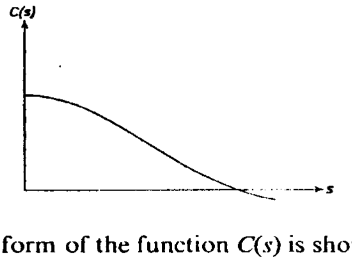

FIGURE 7-1

The covariance function

A typical form of the function $C(s)$ is shown in Fig. 7-1. For small distances s (1 km, say), $\Delta g'$ is almost equal to Δg, so that the covariance is almost equal to the variance; in other words, there is a very strong correlation. The covariance $C(s)$ decreases with increasing s, because then the anomalies Δg and $\Delta g'$

[1] In the precise language of mathematical statistics, zero correlation and independence are not quite the same, but we may neglect the difference here.

become more and more independent. For very large distances the covariance will be very small, but not in general exactly zero because the gravity anomalies are affected not only by local mass disturbances but also by regional factors. So we shall instead expect an oscillation between small positive and negative values.[1]

The practical determination of the covariance function $C(s)$ is somewhat problematical. If we were to determine it exactly, we should have to know gravity at every point of the earth's surface. This we obviously do not know; and if we knew it, then the covariance function would have lost most of its significance, because then we could solve our problems rigorously without needing statistics. As a matter of fact, we can only estimate the covariance function from samples distributed over the whole earth. But even this is not quite possible at present, because of the imperfect or completely missing gravity data over the oceans. For a discussion of sampling and related problems see Kaula (1963, 1966).

The most comprehensive estimate that we have at the present time was made by Kaula (1959). Some of his values are given in Table 7-1. They refer to free-air anomalies. The argument is the spherical distance

$$\psi = \frac{s}{R} \tag{7-6}$$

corresponding to a linear distance s measured on the earth's surface; R is a mean radius of the earth. The r.m.s. free-air anomaly is

$$\text{r.m.s. } \{\Delta g\} = \sqrt{1201} = \pm 35 \text{ mgals.} \tag{7-7}$$

Table 7-1

Estimated Values of the Covariance Function for Free-air Anomalies

Unit 1 mgal²

ψ	$C(\psi)$	ψ	$C(\psi)$	ψ	$C(\psi)$
0.0°	+1201	8°	+124	27°	+18
0.5°	751	9°	104	29°	+ 6
1.0°	468	10°	82	31°	+ 8
1.5°	356	11°	76	33°	+ 5
2.0°	332	13°	54	35°	− 8
2.5°	306	15°	47	40°	−12
3.0°	296	17°	45	50°	−20
4°	272	19°	34	60°	−30
5°	246	21°	35	90°	− 4
6°	214	23°	10	120°	+12
7°	174	25°	20	150°	−21

[1] Positive covariances mean that Δg and $\Delta g'$ tend to have the same size and the *same sign*; negative covariances mean that Δg and $\Delta g'$ tend to have the same size and *opposite* sign. The stronger this tendency, the larger is $C(s)$; the absolute value of $C(s)$ can, however, never exceed the variance $C(0)$.

We see that $C(s)$ decreases with increasing s and that, for $s/R > 30°$, very small values oscillate between plus and minus.

For some purposes we need a *local* covariance function rather than a global one; then the average M is extended over a limited area only, instead of over the whole earth as above. Such a local covariance function is useful for more detailed studies in a limited area—for instance, for interpolation problems. As an example we mention that Hirvonen (1962), investigating the local covariance function of the free-air anomalies in Ohio, found numerical values that are well represented by an analytical expression of the form

$$C(s) = \frac{C_0}{1 + (s/d)^2}, \tag{7-8}$$

where

$$C_0 = 337 \text{ mgals}^2, \quad d = 40 \text{ km}. \tag{7-9}$$

This function is valid for $s < 100$ km.

7-3. Expansion of the Covariance Function in Spherical Harmonics

The more or less complicated integral formulas of physical geodesy usually take on a much simpler form if they are rewritten in terms of spherical harmonics. A good example is Stokes' formula (see Sec. 2-17). Unfortunately this theoretical advantage is in most cases balanced by the practical disadvantage that the relevant series converge very slowly. In certain cases, however, the convergence is good. Then the use of spherical harmonics is very convenient practically; we shall encounter such a case in the next section.

The spherical-harmonic expansion of the gravity anomalies Δg may be written in different ways, such as

$$\Delta g(\theta, \lambda) = \sum_{n=2}^{\infty} \Delta g_n(\theta, \lambda), \tag{7-10}$$

where $\Delta g_n(\theta, \lambda)$ is the Laplace surface harmonic of degree n; or, more explicitly,

$$\Delta g(\theta, \lambda) = \sum_{n=2}^{\infty} \sum_{m=0}^{n} [a_{nm}R_{nm}(\theta, \lambda) + b_{nm}S_{nm}(\theta, \lambda)], \tag{7-11}$$

where

$$\begin{aligned} R_{nm}(\theta, \lambda) &= P_{nm}(\cos\theta)\cos m\lambda, \\ S_{nm}(\theta, \lambda) &= P_{nm}(\cos\theta)\sin m\lambda \end{aligned} \tag{7-12}$$

are the conventional spherical harmonics; or in terms of fully normalized harmonics (see Sec. 1-14):

$$\Delta g(\theta, \lambda) = \sum_{n=2}^{\infty} \sum_{m=0}^{n} [\bar{a}_{nm}\bar{R}_{nm}(\theta, \lambda) + \bar{b}_{nm}\bar{S}_{nm}(\theta, \lambda)]. \tag{7-13}$$

Here θ is the polar distance (complement of geocentric latitude) and λ is the longitude.

Let us now find the average products of two Laplace harmonics

$$\Delta g_n(\theta, \lambda) = \sum_{m=0}^{n} [\bar{a}_{nm}\bar{R}_{nm}(\theta, \lambda) + \bar{b}_{nm}\bar{S}_{nm}(\theta, \lambda)]. \tag{7-14}$$

These average products are

$$M\{\Delta g_n \, \Delta g_{n'}\} = \frac{1}{4\pi}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi} \Delta g_n(\theta, \lambda)\, \Delta g_{n'}(\theta, \lambda) \sin\theta \, d\theta \, d\lambda, \tag{7-15}$$

since the averaging is extended over the whole earth, that is, over the whole unit sphere. Take first $n' = n$, which gives the average square of the Laplace harmonic of degree n:

$$M\{\Delta g_n^2\} = \frac{1}{4\pi}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi} [\Delta g_n(\theta, \lambda)]^2 \sin\theta \, d\theta \, d\lambda. \tag{7-16}$$

Inserting (7-14) and taking into account the orthogonality relations (1-68) and the normalization (1-74), we easily find

$$M\{\Delta g_n^2\} = \sum_{m=0}^{n} (\bar{a}_{nm}^2 + \bar{b}_{nm}^2). \tag{7-17}$$

Consider now the average product (7-15) of two Laplace harmonics of different degree, $n' \neq n$. Owing to the orthogonality of the spherical harmonics the integral in (7-15) is zero:

$$M\{\Delta g_n \, \Delta g_{n'}\} = 0 \qquad \text{if } n' \neq n. \tag{7-18}$$

In statistical terms this means that two Laplace harmonics of different degrees are *uncorrelated* or, broadly speaking, *statistically independent*.

In a way similar to that used for the gravity anomalies we may also expand the covariance function $C(s)$ into a series of spherical harmonics. Let us take an arbitrary, but fixed, point P as the pole of this expansion. Thus spherical polar coordinates ψ (angular distance from P) and α (azimuth) are introduced (Fig. 7-2). The angular distance ψ corresponds to the linear distance s according to (7-6). If we expand the covariance function, with argument ψ, into a series of spherical harmonics with respect to the pole P and coordinates ψ and α, we have

$$C(\psi) = \sum_{n=2}^{\infty}\sum_{m=0}^{n} [c_{nm}R_{nm}(\psi, \alpha) + d_{nm}S_{nm}(\psi, \alpha)],$$

which is of the same type as (7-11). But since C depends only on the distance ψ and not on the azimuth α, the spherical harmonics cannot contain any terms that explicitly depend on α. The only harmonics independent of α are the zonal functions

$$R_{n0}(\psi, \alpha) \equiv P_n(\cos\psi),$$

so that we are left with

$$C(\psi) = \sum_{n=2}^{\infty} c_n P_n(\cos\psi). \tag{7-19}$$

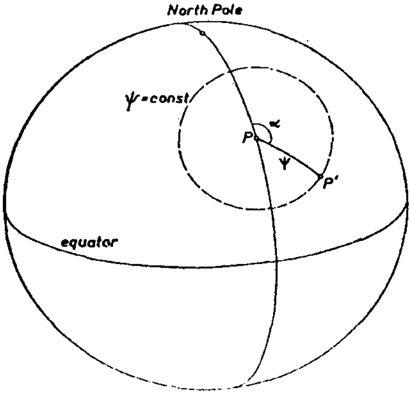

FIGURE 7-2
Spherical coordinates ψ, α.

The $c_n \equiv c_{n0}$ are the only coefficients that are not equal to zero. We shall also use the equivalent expression in terms of fully normalized harmonics:

$$C(\psi) = \sum_{n=2}^{\infty} \bar{c}_n \bar{P}_n(\cos\psi). \tag{7-20}$$

The coefficients in these series, according to Secs. 1-13 and 1-14, are given by

$$\begin{aligned} c_n &= \frac{2n+1}{4\pi} \int_{\alpha=0}^{2\pi} \int_{\psi=0}^{\pi} C(\psi) P_n(\cos\psi) \sin\psi \, d\psi \, d\alpha \\ &= \frac{2n+1}{2} \int_{\psi=0}^{\pi} C(\psi) P_n(\cos\psi) \sin\psi \, d\psi \end{aligned} \tag{7-21}$$

and

$$\bar{c}_n = \frac{c_n}{\sqrt{2n+1}}. \tag{7-22}$$

We shall now determine the relation between the coefficients c_n of $C(\psi)$ in (7-19) and the coefficients $\bar{a}_{nm}$ and $\bar{b}_{nm}$ of Δg in (7-14). For this purpose we need an expression for $C(\psi)$ in terms of Δg, which is easily obtained by writing (7-5) more explicitly. Take the two points $P(\theta, \lambda)$ and $P'(\theta', \lambda')$ of Fig. 7-2. Their spherical distance ψ is given by

$$\cos\psi = \cos\theta \cos\theta' + \sin\theta \sin\theta' \cos(\lambda' - \lambda). \tag{7-23}$$

Here ψ and the azimuth α are the polar coordinates of $P'(\theta', \lambda')$ with respect to the pole $P(\theta, \lambda)$.

The symbol M in (7-5) denotes the average over the unit sphere. To find it requires two steps. First, we average over the spherical circle of radius ψ (denoted in Fig. 7-2 by a broken line), keeping the pole P fixed and letting P' move along the circle, so that the distance PP' remains constant. This gives

$$C^* = \frac{1}{2\pi} \int_{\alpha=0}^{2\pi} \Delta g(\theta, \lambda) \, \Delta g(\theta', \lambda') \, d\alpha,$$

where C^* still depends on the point P chosen as the pole $\psi = 0$. Second, we average C^* over the unit sphere:

$$\frac{1}{4\pi}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi} C^* \sin\theta \, d\theta \, d\lambda$$

$$= \frac{1}{8\pi^2}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi}\int_{\alpha=0}^{2\pi} \Delta g(\theta, \lambda)\,\Delta g(\theta', \lambda') \sin\theta \, d\theta \, d\lambda \, d\alpha.$$

This is equal to the covariance function $C(\psi)$, the symbol M in (7-5) now being written explicitly:

$$C(\psi) = \frac{1}{8\pi^2}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi}\int_{\alpha=0}^{2\pi} \Delta g(\theta, \lambda)\,\Delta g(\theta', \lambda') \sin\theta \, d\theta \, d\lambda \, d\alpha. \tag{7-24}$$

The coordinates θ', λ' in this formula are understood to be related to θ, λ by (7-23) with $\psi = \text{const.}$, but to be arbitrary otherwise; this, of course, expresses the fact that in (7-5) the average is extended over all pairs of points P and P' for which $PP' = \psi = \text{const.}$

To compute the coefficients c_n, insert (7-24) into (7-21), obtaining

$$c_n = \frac{2n+1}{2}\int_{\psi=0}^{\pi} C(\psi)P_n(\cos\psi)\sin\psi \, d\psi$$

$$= \frac{1}{4\pi}\frac{2n+1}{4\pi}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi}\int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\pi} \Delta g(\theta, \lambda)\,\Delta g(\theta', \lambda')$$

$$\cdot P_n(\cos\psi)\sin\psi \, d\psi \, d\alpha \cdot \sin\theta \, d\theta \, d\lambda. \tag{7-25}$$

Consider first the integration with respect to α and ψ. According to (1-71) we have

$$\frac{2n+1}{4\pi}\int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\pi} \Delta g(\theta', \lambda')P_n(\cos\psi)\sin\psi \, d\psi \, d\alpha$$

$$= \frac{2n+1}{4\pi}\int_{\lambda'=0}^{2\pi}\int_{\theta'=0}^{\pi} \Delta g(\theta', \lambda')P_n(\cos\psi)\sin\theta' \, d\theta' \, d\lambda' = \Delta g_n(\theta, \lambda),$$

the change of integration variables being evident. Hence (7-25) becomes

$$c_n = \frac{1}{4\pi}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi} \Delta g(\theta, \lambda)\,\Delta g_n(\theta, \lambda)\sin\theta \, d\theta \, d\lambda. \tag{7-26}$$

This may also be written

$$c_n = M\{\Delta g \, \Delta g_n\}. \tag{7-27}$$

Into this we now insert (7-10), which we write

$$\Delta g(\theta, \lambda) = \sum_{n'=2}^{\infty} \Delta g_{n'}(\theta, \lambda),$$

denoting the summation index by n' instead of n. We get

$$c_n = M\left\{\sum_{n'=2}^{\infty} \Delta g_{n'} \, \Delta g_n\right\} = \sum_{n'=2}^{\infty} M\{\Delta g_n \, \Delta g_{n'}\}.$$

According to (7-18) only the term with $n' = n$ is different from zero, so that by (7-17) we finally obtain

$$c_n = M\{\Delta g_n^2\} = \sum_{m=0}^{n} (\bar{a}_{nm}^2 + \bar{b}_{nm}^2). \tag{7-28}$$

Hence c_n is the average square of the Laplace harmonic $\Delta g_n(\theta, \lambda)$ of degree n, or its variance. For these reasons the c_n are also called *degree variances*. [The "degree covariances" are zero, because of (7-18).]

Equation (7-28) relates the coefficients $\bar{a}_{nm}$ and $\bar{b}_{nm}$ of Δg and c_n of $C(s)$ in the simplest possible way. Note that $\bar{a}_{nm}$ and $\bar{b}_{nm}$ are coefficients of fully normalized harmonics, whereas c_n are coefficients of conventional harmonics. As a matter of fact, we may also use the a_{nm} and b_{nm} (conventional) or the $\bar{c}_n$ (fully normalized); but then (7-28) will obviously become slightly more complicated.[1]

7-4. Influence of Distant Zones on Stokes' and Vening Meinesz' Formulas

The spherical-harmonic expansions of the preceding section will now be used to evaluate the effects of neglecting the distant zones on the computation of the geoidal height and the deflection of the vertical.

Let us split Stokes' integral (2-165) into two parts:

$$N = \frac{R}{4\pi G}\int_{\psi=0}^{\psi_0}\int_{\alpha=0}^{2\pi} \Delta g\, S(\cos\psi) \sin\psi\, d\psi\, d\alpha$$
$$+ \frac{R}{4\pi G}\int_{\psi=\psi_0}^{\pi}\int_{\alpha=0}^{2\pi} \Delta g\, S(\cos\psi) \sin\psi\, d\psi\, d\alpha. \tag{7-29}$$

We are now denoting Stokes' function by $S(\cos\psi)$ instead of $S(\psi)$ in order to have a simple and consistent notation later on in this section.

If the integration is extended not over the whole earth but only up to a spherical distance ψ_0, then only the first integral of (7-29) is considered. The error δN that results from neglecting the zones beyond $\psi = \psi_0$ is therefore given by the second integral in (7-29),

$$\delta N = \frac{R}{4\pi G}\int_{\psi=\psi_0}^{\pi}\int_{\alpha=0}^{2\pi} \Delta g\, S(\cos\psi) \sin\psi\, d\psi\, d\alpha. \tag{7-30}$$

Introducing the (discontinuous) function (Fig. 7-3)

$$\bar{S}(\cos\psi) = \begin{cases} 0 & \text{if } 0 \leqq \psi < \psi_0, \\ S(\cos\psi) & \text{if } \psi_0 \leqq \psi \leqq \pi, \end{cases} \tag{7-31}$$

[1] It should be mentioned that the mathematics behind the statistical description of the gravity anomalies is the theory of *stochastic processes*. The gravity anomaly field is treated as a stationary stochastic process on a sphere; the spherical-harmonic expansions of this section are nothing but the spectral analysis of that process. An elementary introduction to stochastic processes is found in Miller (1956).

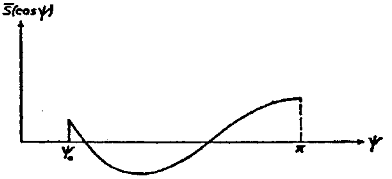

FIGURE 7-3
The function $\bar{S}(\cos\psi)$.

we may write (7-30) in the form

$$\delta N = \frac{R}{4\pi G}\int_{\psi=0}^{\pi}\int_{\alpha=0}^{2\pi} \Delta g\, \bar{S}(\cos\psi)\sin\psi\, d\psi\, d\alpha. \tag{7-32}$$

The integration can now be formally extended over the whole unit sphere because the zones with $\psi < \psi_0$ make no contribution to the value of the integral.

The function $\bar{S}(\cos\psi)$, being piecewise continuous, may be expanded into a series of Legendre polynomials (zonal harmonics):

$$\bar{S}(\cos\psi) = \sum_{n=0}^{\infty} \frac{2n+1}{2} Q_n P_n(\cos\psi). \tag{7-33}$$

For formal reasons we denote the coefficients in this expansion by $(2n+1)Q_n/2$. According to Sec. 1-13, equation (1-70), they are given by

$$\frac{2n+1}{2} Q_n = \frac{2n+1}{4\pi}\int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\pi} \bar{S}(\cos\psi)P_n(\cos\psi)\sin\psi\, d\psi\, d\alpha.$$

The integration with respect to α can be performed immediately, giving

$$\int_{\alpha=0}^{2\pi} d\alpha = 2\pi,$$

so that

$$Q_n = \int_0^{\pi} \bar{S}(\cos\psi)P_n(\cos\psi)\sin\psi\, d\psi.$$

Using (7-31) we finally find

$$Q_n = \int_{\psi_0}^{\pi} S(\cos\psi)P_n(\cos\psi)\sin\psi\, d\psi. \tag{7-34}$$

This equation determines the Q_n as functions of the limiting radius ψ_0. The evaluation of this integral is but a matter of routine; it will be given below.

We now insert (7-33) into (7-32). After interchanging the order of integration and summation we obtain

$$\delta N = \frac{R}{8\pi G}\sum_{n=0}^{\infty}(2n+1)Q_n \int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\pi} \Delta g\, P_n(\cos\psi)\sin\psi\, d\psi\, d\alpha.$$

By (1-71), the double integral is equal to $4\pi\,\Delta g_n/(2n+1)$, so that

$$\delta N(\theta, \lambda) = \frac{R}{2G}\sum_{n=2}^{\infty} Q_n\, \Delta g_n(\theta, \lambda), \tag{7-35}$$

Δg_n being, as before, the nth-degree Laplace harmonic of Δg.

Equation (7-35) gives the error in N at a certain point $P(\theta, \lambda)$ caused by neglecting the gravity anomalies beyond a circle of radius ψ_0 whose center is P. If we want the r.m.s. effect $\overline{\delta N}$, we must form the average M over the unit sphere:

$$\begin{aligned}
\overline{\delta N}^2 \equiv M\{\delta N^2\} &= \frac{R^2}{4G^2} M\left\{\left(\sum_{n=2}^{\infty} Q_n \, \Delta g_n\right)^2\right\} \\
&= \frac{R^2}{4G^2} M\left\{\sum_{n=2}^{\infty} Q_n \, \Delta g_n \sum_{n'=2}^{\infty} Q_{n'} \, \Delta g_{n'}\right\} \\
&= \frac{R^2}{4G^2} M\left\{\sum_{n=2}^{\infty}\sum_{n'=2}^{\infty} Q_n Q_{n'} \, \Delta g_n \, \Delta g_{n'}\right\} \\
&= \frac{R^2}{4G^2} \sum_{n=2}^{\infty}\sum_{n'=2}^{\infty} Q_n Q_{n'} M\{\Delta g_n \, \Delta g_{n'}\}.
\end{aligned}$$

The manipulations performed here are obvious. First we inserted (7-35); then we introduced another summation index n', in order to transform the square of a sum into a double sum; finally we interchanged the order of integration (symbol M) and summation.

According to equation (7-18) of the preceding section all $M\{\Delta g_n \, \Delta g_{n'}\}$ are zero except when $n' = n$. Hence we finally obtain

$$\overline{\delta N}^2 = \frac{R^2}{4G^2} \sum_{n=2}^{\infty} Q_n^2 \, M\{\Delta g_n^2\} = \frac{R^2}{4G^2} \sum_{n=2}^{\infty} Q_n^2 c_n. \tag{7-36}$$

Thus the r.m.s. influence of the remote zones on the geoidal height N may be computed from the degree variances or, what amounts to the same thing, from the covariance function. This is an example of the fundamental role of the covariance function in statistical problems of physical geodesy.

Formulas for the influence of the remote zones on the deflection of the vertical are considerably more difficult to derive. We shall therefore sketch only the main points; a detailed derivation may be found in the paper by Hirvonen and Moritz (1963), referred to in Chapter 6.

By equations (2-204) and (7-35) we have

$$\delta\xi = -\frac{1}{R}\frac{\partial(\delta N)}{\partial\phi} = -\frac{1}{2G}\sum_{n=2}^{\infty} Q_n \frac{\partial \Delta g_n}{\partial\phi},$$

$$\delta\eta = -\frac{1}{R\cos\phi}\frac{\partial(\delta N)}{\partial\lambda} = -\frac{1}{2G}\sum_{n=2}^{\infty} Q_n \frac{1}{\cos\phi}\frac{\partial \Delta g_n}{\partial\lambda}.$$

The total r.m.s. error $\overline{\delta\theta}$ of the deflection of the vertical is thus given by

$$\begin{aligned}
\overline{\delta\theta}^2 &\equiv M\{\delta\xi^2 + \delta\eta^2\} \\
&= \frac{1}{4G^2}\sum_{n=2}^{\infty}\sum_{n'=2}^{\infty} Q_n Q_{n'} M\left\{\frac{\partial\Delta g_n}{\partial\phi}\frac{\partial\Delta g_{n'}}{\partial\phi} + \frac{1}{\cos^2\phi}\frac{\partial\Delta g_n}{\partial\lambda}\frac{\partial\Delta g_{n'}}{\partial\lambda}\right\}.
\end{aligned}$$

It may be shown that for an arbitrary Laplace surface harmonic Y_n of degree n the following relations hold:

$$M\left\{\left(\frac{\partial Y_n}{\partial \phi}\right)^2 + \frac{1}{\cos^2\phi}\left(\frac{\partial Y_n}{\partial \lambda}\right)^2\right\} = n(n+1)\, M\{Y_n^2\},$$
$$M\left\{\frac{\partial Y_n}{\partial \phi}\frac{\partial Y_{n'}}{\partial \phi} + \frac{1}{\cos^2\phi}\frac{\partial Y_n}{\partial \lambda}\frac{\partial Y_{n'}}{\partial \lambda}\right\} = 0 \qquad \text{if } n' \neq n; \tag{7-37}$$

see also Jeffreys (1962, p. 135). Hence, for $Y_n = \Delta g_n$, we obtain

$$\overline{\delta\theta^2} = \frac{1}{4G^2}\sum_{n=2}^{\infty} n(n+1)Q_n^2\, M\{\Delta g_n^2\} = \frac{1}{4G^2}\sum_{n=2}^{\infty} n(n+1)Q_n^2 c_n. \tag{7-38}$$

This formula gives the r.m.s. influence of the remote zones on the total deflection of the vertical θ;[1] it corresponds to equation (7-36) for N.

The coefficients Q_n. To obtain the Q_n explicitly as functions of the radius ψ_0, we must evaluate the integral (7-34). Substituting

$$\sin\frac{\psi}{2} = z, \quad \sin\frac{\psi_0}{2} = t \tag{7-39}$$

we get

$$Q_n = \int_{\psi_0}^{\pi} S(\cos\psi)P_n(\cos\psi)\sin\psi\, d\psi = 4\int_t^1 P_n(1-2z^2)S(1-2z^2)z\, dz,$$

because

$$\cos\psi = 1 - 2\sin^2\frac{\psi}{2} = 1 - 2z^2,$$

$$\sin\psi\, d\psi = 4\sin\frac{\psi}{2}\cdot\cos\frac{\psi}{2}\, d\frac{\psi}{2} = 4z\, dz.$$

By interchanging the limits of integration we finally find

$$Q_n = -4\int_1^t P_n(1-2z^2)S(1-2z^2)z\, dz. \tag{7-40}$$

The $S(1-2z^2)$ means that in the Stokes function $S(\cos\psi)$, we must replace $\cos\psi$ by $1-2z^2$, and $\sin(\psi/2)$ by z:

$$S(1-2z^2) = \frac{1}{z} - 3\ln z(1+z) + 6z^2\ln z(1+z) - 4 - 6z + 10z^2; \tag{7-41}$$

similarly $P_n(1-2z^2)$ means that the argument of the zonal harmonic P_n [t in equations (1-58)] is to be replaced by $1-2z^2$, for instance

$$P_0(1-2z^2) = 1, \quad P_1(1-2z^2) = 1-2z^2, \quad P_2(1-2z^2) = \frac{3}{2}(1-2z^2)^2 - \frac{1}{2}. \tag{7-42}$$

[1] Above, the symbol θ was used to denote the polar distance!

The integral (7-40) can thus be evaluated by the usual methods of integration; we obtain, for instance,

$$
\begin{aligned}
Q_0 &= -4t + 5t^2 + 6t^3 - 7t^4 + (6t^2 - 6t^4)\ln t(1+t),\\
Q_1 &= -2t + 4t^2 + \frac{28}{3}t^3 - 14t^4 - 8t^5 + \frac{32}{3}t^6\\
&\quad + (6t^2 - 12t^4 + 8t^6)\ln t(1+t) - 2\ln(1+t), \qquad (7\text{-}43)\\
Q_2 &= 2 - 4t + 5t^2 + 14t^3 - \frac{53}{2}t^4 - 30t^5 + 47t^6 + 18t^7 - \frac{51}{2}t^8\\
&\quad + (6t^2 - 24t^4 + 36t^6 - 18t^8)\ln t(1+t).
\end{aligned}
$$

Formulas for the Q_n up to $n = 8$ and a table of values may be found in Molodenskii et al. (1962, p. 148–150).

If $\psi_0 = 0$, then the function $\bar{S}(\cos\psi)$ of equation (7-31) reduces to Stokes' function $S(\cos\psi)$ for all values of ψ:

$$\bar{S}(\cos\psi) = \sum_{n=0}^{\infty} \frac{2n+1}{2} Q_n P_n(\cos\psi) = S(\cos\psi) = \sum_{n=2}^{\infty} \frac{2n+1}{n-1} P_n(\cos\psi),$$

so that

$$Q_0 = Q_1 = 0, \quad Q_n = \frac{2}{n-1} \qquad (n \geqq 2) \qquad \text{if } \psi_0 = 0. \qquad (7\text{-}44)$$

Numerical results. Since the size of the Q_n decreases quickly with increasing n, except for small ψ_0, the series (7-36) and (7-38) converge rapidly, so that a few terms are in general sufficient.

Kaula (1959, p. 2419) proposes the following maximum plausible values (mgals²) for the degree variances:

$$c_2 = 15, \quad c_3 = 43, \quad c_4 = 30, \quad c_5 = c_6 = c_7 = c_8 = 25, \qquad (7\text{-}45)$$

which are consistent with the values of the covariance function of Table 7-1. Then the mean effect of the gravity anomalies beyond a spherical radius ψ_0 is given by Table 7-2. The first three values of ψ_0 correspond to linear distances of 1000, 1500, and 2000 km. The summation in (7-36) and (7-38) was extended up to $n = 8$.

Table 7-2

R.M.S. Influence of the Zone beyond the Radius ψ_0 on Geoidal Height N and Deflection of the Vertical θ

ψ_0	$\overline{\Delta N}$	$\overline{\Delta\theta}$	ψ_0	$\overline{\Delta N}$	$\overline{\Delta\theta}$
9.0°	±25 m	±2.4"	60°	±14 m	±1.2"
13.5°	21	2.0	90°	11	1.1
18.0°	18	1.8	135°	8	0.8
30°	14	1.2	180°	0	0.0

Molodenskii et al. (1962, p. 167) give numerical estimates of $\overline{\delta N}$ and $\overline{\delta\theta}$ which are about 70% higher. They are based on values of $c_n = \overline{\Delta g_n^2}$ corresponding to a spherical-harmonic expansion obtained by Zhongolovich in 1952.

7-5. Interpolation and Extrapolation of Gravity Anomalies

As was pointed out in Sec. 7-1, the purpose of prediction (interpolation and extrapolation) is to supplement the gravity observations, which can be made at only a relatively few points, by estimating the values of gravity or of gravity anomalies at all the other points P of the earth's surface.

If P is surrounded by gravity stations, we must interpolate; if the gravity stations are far away from P, we shall extrapolate. Evidently there is no sharp distinction between these two kinds of prediction, and the mathematical formulation is the same in both cases.

In order to predict a gravity anomaly at P, we must have information about the gravity anomaly function. The most important information is, of course, the values observed at certain points. In addition, we need some information on the form of the anomaly function. If the gravity measurements are very dense, then the continuity or "smoothness" of the function is sufficient—for instance, for linear interpolation. Otherwise we may try to use statistical information on the general structure of the gravity anomalies. Here we must consider two kinds of statistical correlation: the *autocorrelation*—the correlation between each other—of gravity anomalies, and the *correlation* of the gravity anomalies *with elevation*.

Correlation with elevation will for the moment be disregarded; Sec. 7-10 will be devoted to this topic. The autocorrelation is characterized by the covariance function considered in Sec. 7-2.

Mathematically, the purpose of prediction is to find a function of the observed gravity anomalies $\Delta g_1, \Delta g_2, \ldots, \Delta g_n$,[1] such that the unknown anomaly Δg_P at P is approximated by the function

$$\Delta g_P \doteq F(\Delta g_1, \Delta g_2, \ldots, \Delta g_n). \tag{7-46}$$

In practice, only *linear* functions of the Δg_i are used. If we denote the predicted value of Δg_P by $\widetilde{\Delta g_P}$, such a linear prediction has the form

$$\widetilde{\Delta g_P} = \alpha_{P1}\,\Delta g_1 + \alpha_{P2}\,\Delta g_2 + \cdots + \alpha_{Pn}\,\Delta g_n \equiv \sum_{i=1}^{n} \alpha_{Pi}\,\Delta g_i. \tag{7-47}$$

The coefficients α_{Pi} depend only on the relative position of P and the gravity stations $1, 2, \ldots, n$; they are independent of the Δg_i. Depending on the way we choose these coefficients, we obtain different interpolation or extrapolation methods. Here are some examples.

Geometrical interpolation. The "gravity anomaly surface," as represented

[1] Here Δg_i denotes the value of Δg at a point i, not a spherical harmonic!

by a gravity anomaly map, may be approximated by a polyhedron by dividing the area into triangles whose corners are formed by the gravity stations, and passing a plane through the three corners of each triangle (see Fig. 7-4). This is approximately what is done in constructing the contour lines of a gravity anomaly map by means of graphical interpolation.

Analytically this interpolation may be formulated as follows. Let point P be situated inside a triangle with corners 1, 2, 3 (Fig. 7-4). To each point we assign its value Δg as its z-coordinate, so that the points 1, 2, and 3 have "spatial" coordinates (x_1, y_1, z_1), (x_2, y_2, z_2), and (x_3, y_3, z_3); x and y are ordinary plane coordinates. The plane through 1, 2, 3 has the equation

$$\begin{aligned} z = {} & \frac{(x_2 - x)(y_3 - y_2) - (y_2 - y)(x_3 - x_2)}{(x_2 - x_1)(y_3 - y_2) - (y_2 - y_1)(x_3 - x_2)} z_1 \\ & + \frac{(x_3 - x)(y_1 - y_3) - (y_3 - y)(x_1 - x_3)}{(x_3 - x_2)(y_1 - y_3) - (y_3 - y_2)(x_1 - x_3)} z_2 \\ & + \frac{(x_1 - x)(y_2 - y_1) - (y_1 - y)(x_2 - x_1)}{(x_1 - x_3)(y_2 - y_1) - (y_1 - y_3)(x_2 - x_1)} z_3. \end{aligned} \tag{7-48}$$

If we replace z_1, z_2, z_3 by Δg_1, Δg_2, Δg_3, then z is the interpolated value $\widetilde{\Delta g}_P$ at point P, which has the plane coordinates x, y. Thus

$$\widetilde{\Delta g}_P = \alpha_{P1}\,\Delta g_1 + \alpha_{P2}\,\Delta g_2 + \alpha_{P3}\,\Delta g_3, \tag{7-49}$$

where the α_{Pi} are the coefficients of z_i in the preceding equation.

Representation. Often the measured anomaly of a gravity station 1 is made to represent the whole neighborhood, so that

$$\widetilde{\Delta g}_P \equiv \Delta g_1 \tag{7-50}$$

as long as P lies within a certain neighborhood of point 1.

Here

$$\alpha_{P1} = 1, \quad \alpha_{P2} = \alpha_{P3} = \cdots = \alpha_{Pn} = 0.$$

This method is rather crude, but simple and accurate enough for many purposes.

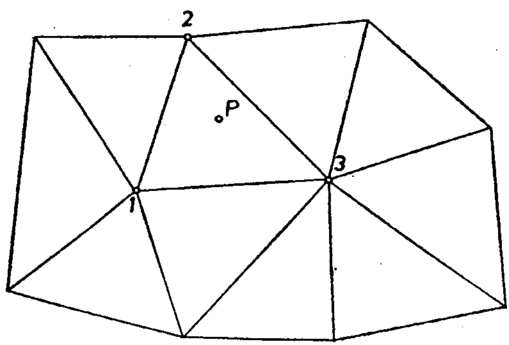

FIGURE 7-4
Geometrical interpolation.

Zero anomaly. If there are no gravity measurements in a large area—for instance, on the oceans—then the estimate

$$\widetilde{\Delta g_P} \equiv 0 \tag{7-51}$$

is used in this area. In this trivial case all α_{Pi}'s are zero.

If all known gravity stations are far away, and if we know of nothing better, then this primitive extrapolation method is applied, although the accuracy is poor. Isostatic anomalies are preferable for this purpose.

None of these three methods gives optimum accuracy. In the next section we shall investigate the accuracy of the general prediction formula (7-47) and find those coefficients α_{iP} that yield the most accurate results.

7-6. Accuracy of Prediction Methods. Least Squares Prediction

In order to compare the various possible methods of prediction, to determine their range of applicability, and to find the most accurate method, we must evaluate their accuracy.

Consider the general case of equation (7-47). The correct gravity anomaly at P is Δg_P, the predicted value is

$$\widetilde{\Delta g_P} = \sum_{i=1}^{n} \alpha_{Pi}\,\Delta g_i.$$

The difference is the error ϵ_P of prediction,

$$\epsilon_P = \Delta g_P - \widetilde{\Delta g_P} = \Delta g_P - \sum_i \alpha_{Pi}\,\Delta g_i. \tag{7-52}$$

By squaring we find

$$\begin{aligned}\epsilon_P^2 &= \Big(\Delta g_P - \sum_i \alpha_{Pi}\,\Delta g_i\Big)\Big(\Delta g_P - \sum_k \alpha_{Pk}\,\Delta g_k\Big)\\ &= \Delta g_P^2 - 2\sum_i \alpha_{Pi}\,\Delta g_P\,\Delta g_i + \sum_i\sum_k \alpha_{Pi}\alpha_{Pk}\,\Delta g_i\,\Delta g_k.\end{aligned} \tag{7-53}$$

Let us now form the average M of this formula over the area considered (either a limited region or the whole earth). Then we have, by (7-5),

$$\begin{aligned} M\{\Delta g_i\,\Delta g_k\} &= C(ik) \equiv C_{ik},\\ M\{\Delta g_P\,\Delta g_i\} &= C(Pi) \equiv C_{Pi},\\ M\{\Delta g_P^2\} &= C(0) \equiv C_0.\end{aligned} \tag{7-54}$$

These are particular values of the covariance function $C(s)$, for $s = ik$, $s = Pi$, and $s = 0$; for instance, ik is the distance between the gravity stations i and k. The abridged notations C_{ik} and C_{Pi} are self-explanatory.

We further set

$$M\{\epsilon_P^2\} = m_P^2. \tag{7-55}$$

Thus m_P is the root mean square error of a predicted gravity anomaly at P, or briefly, the standard *error of prediction* (interpolation or extrapolation).

Taking all these relations into account, we find the average M of (7-53) to be

$$m_P^2 = C_0 - 2\sum_{i=1}^{n} \alpha_{Pi} C_{Pi} + \sum_{i=1}^{n}\sum_{k=1}^{n} \alpha_{Pi}\alpha_{Pk} C_{ik}. \tag{7-56}$$

This is the fundamental formula for the standard error of the general prediction formula (7-47). For the special cases described in the preceding section, the particular values of α_{Pi} are to be inserted.

As an example consider the case of representation, equation (7-50); all α's are zero except one. Here (7-56) yields

$$m_P^2 = C_0 - 2C_{P1} + C_0 = 2C_0 - 2C_{P1}.$$

Often we need not only the standard error m_P of prediction but also the correlation of the prediction errors ϵ_P and ϵ_Q at two different points P and Q, expressed by the "*error covariance*" σ_{PQ}, which is defined by

$$\sigma_{PQ} = M\,\{\epsilon_P \epsilon_Q\}. \tag{7-57}$$

(If the errors ϵ_P and ϵ_Q are uncorrelated, then the error covariance $\sigma_{PQ} = 0$.) By (7-52) we have

$$\begin{aligned}\sigma_{PQ} &= M\,\{(\Delta g_P - \sum_i \alpha_{Pi}\,\Delta g_i)(\Delta g_Q - \sum_k \alpha_{Qk}\,\Delta g_k)\}\\ &= M\,\{\Delta g_P\,\Delta g_Q - \sum_i \alpha_{Pi}\,\Delta g_Q\,\Delta g_i - \sum_k \alpha_{Qk}\,\Delta g_P\,\Delta g_k\\ &\qquad + \sum_i\sum_k \alpha_{Pi}\alpha_{Qk}\,\Delta g_i\,\Delta g_k\},\end{aligned}$$

and finally

$$\sigma_{PQ} = C_{PQ} - \sum_{i=1}^{n} \alpha_{Pi} C_{Qi} - \sum_{i=1}^{n} \alpha_{Qi} C_{Pi} + \sum_{i=1}^{n}\sum_{k=1}^{n} \alpha_{Pi}\alpha_{Qk} C_{ik}. \tag{7-58}$$

The notations are self-explanatory; for instance, $C_{PQ} = C(PQ)$.

The error covariance function. The values of the error covariance σ_{PQ}, for different positions of the points P and Q, form a continuous function of the coordinates of P and Q. This function is called the *error covariance function*, or briefly, the *error function*, and is denoted by $\sigma(x_P, y_P, x_Q, y_Q)$. If P and Q are different, then we simply have

$$\sigma(x_P, y_P, x_Q, y_Q) = \sigma_{PQ}; \tag{7-59a}$$

if P and Q coincide, then (7-58) reduces to (7-56), so that

$$\sigma(x_P, y_P, x_P, y_P) = m_P^2 \tag{7-59b}$$

is the square of the standard prediction error at P.

Thus the error covariances σ_{PQ} may be considered as special values of the error covariance function, just as the covariances C_{PQ} of the gravity anomalies may be considered as special values of the covariance function $C(s)$. To repeat, the error function is the covariance function of the prediction errors, defined as

$$M\,\{\epsilon_P \epsilon_Q\},$$

whereas $C(s)$ is the covariance function of the gravity anomalies, defined as

$$M\,\{\Delta g_P\, \Delta g_Q\}.$$

The term "covariance function" in the narrower sense will be reserved for $C(s)$.

By (7-56) and (7-58) *the error function can be expressed in terms of the covariance function*; we may write more explicitly

$$\sigma(x_P, y_P, x_Q, y_Q) = C(PQ) - \sum_{i=1}^{n} \alpha_{Pi} C(Qi) - \sum_{i=1}^{n} \alpha_{Qi} C(Pi) + \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_{Pi} \alpha_{Qk} C(ik). \tag{7-60}$$

Thus we recognize the basic role of the covariance function in accuracy studies. The error function, on the other hand, is fundamental for problems of error propagation, as we shall see in the following sections.

Least squares prediction. The values of α_{Pi} for the most accurate prediction method are obtained by minimizing the standard prediction error expressed by (7-56) as a function of the α's. The familiar necessary conditions for a minimum are

$$\frac{\partial m_P^2}{\partial \alpha_{Pi}} \equiv -2C_{Pi} + 2 \sum_{k=1}^{n} \alpha_{Pk} C_{ik} = 0 \qquad (i = 1, 2, \ldots, n)$$

or

$$\sum_{k=1}^{n} C_{ik} \alpha_{Pk} = C_{Pi}. \tag{7-61}$$

This is a system of n linear equations in the n unknowns α_{Pk}; the solution is

$$\alpha_{Pk} = \sum_{i=1}^{n} C_{ik}^{(-1)} C_{Pi}, \tag{7-62}$$

where $C_{ik}^{(-1)}$ denote the elements of the inverse of the matrix (C_{ik}).

Inserting (7-62) into (7-47) gives

$$\widetilde{\Delta g_P} = \sum_{k=1}^{n} \alpha_{Pk}\, \Delta g_k = \sum_{i=1}^{n} \sum_{k=1}^{n} C_{ik}^{(-1)} C_{Pi}\, \Delta g_k.$$

In matrix notation this is written

$$\widetilde{\Delta g_P} = (C_{P1}, C_{P2}, \ldots, C_{Pn}) \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix}^{-1} \begin{pmatrix} \Delta g_1 \\ \Delta g_2 \\ \cdot \\ \cdot \\ \cdot \\ \Delta g_n \end{pmatrix}. \tag{7-63}$$

We see that *for optimal prediction we must know the statistical behavior of the gravity anomalies* through the covariance function $C(s)$.

There is a close connection between this optimal prediction method and the method of least squares adjustment. Although they refer to somewhat different problems, both are designed to give most accurate results. The linear equations (7-61) correspond to the "normal equations" of adjustment computations. Prediction by means of formula (7-63) is therefore called "*least squares prediction*." Details will be found in Kaula (1963) and in Moritz (1965).

It is easy to determine the accuracy of least squares prediction. Insert the α's of equation (7-62) into (7-56), after appropriate changes in the indices of summation. This gives

$$\begin{aligned} m_P^2 &= C_0 - 2\sum_k \alpha_{Pk} C_{Pk} + \sum_k \sum_l \alpha_{Pk}\alpha_{Pl} C_{kl} \\ &= C_0 - 2\sum_i \sum_k C_{ik}^{(-1)} C_{Pi} C_{Pk} + \sum_i \sum_j \sum_k \sum_l C_{ik}^{(-1)} C_{Pi} C_{jl}^{(-1)} C_{Pj} C_{kl}. \end{aligned}$$

We have

$$\sum_l C_{jl}^{(-1)} C_{kl} = \delta_{jk} = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

The matrix (δ_{kl}) is the unit matrix. This formula states that the product of a matrix and its inverse is the unit matrix. Thus we further have

$$\sum_k \sum_l C_{ik}^{(-1)} C_{jl}^{(-1)} C_{kl} = \sum_k C_{ik}^{(-1)} \delta_{jk} = C_{ij}^{(-1)},$$

because a matrix remains unchanged on multiplication by the unit matrix. Hence we get

$$\begin{aligned} m_P^2 &= C_0 - 2\sum_i \sum_k C_{ik}^{(-1)} C_{Pi} C_{Pk} + \sum_i \sum_j C_{ij}^{(-1)} C_{Pi} C_{Pj} \\ &= C_0 - 2\sum_i \sum_k C_{ik}^{(-1)} C_{Pi} C_{Pk} + \sum_i \sum_k C_{ik}^{(-1)} C_{Pi} C_{Pk} \\ &= C_0 - \sum_i \sum_k C_{ik}^{(-1)} C_{Pi} C_{Pk}. \end{aligned}$$

Thus the standard error of least squares prediction is given by

$$\begin{aligned} m_P^2 &= C_0 - \sum_{i=1}^{n} \sum_{k=1}^{n} C_{ik}^{(-1)} C_{Pi} C_{Pk} \\ &= C_0 - (C_{P1}, C_{P2}, \ldots, C_{Pn}) \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix}^{-1} \begin{pmatrix} C_{P1} \\ C_{P2} \\ \vdots \\ C_{Pn} \end{pmatrix}. \end{aligned} \qquad (7\text{-}64)$$

In the same way we find the error covariance in the points P and Q:

$$\sigma_{PQ} = C_{PQ} - \sum_{i=1}^{n}\sum_{k=1}^{n} C_{ik}^{(-1)} C_{Pi} C_{Qk}$$

$$= C_{PQ} - (C_{P1}, C_{P2}, \ldots, C_{Pn}) \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix}^{-1} \begin{pmatrix} C_{Q1} \\ C_{Q2} \\ \cdot \\ \cdot \\ \cdot \\ C_{Qn} \end{pmatrix}. \qquad (7\text{-}65)$$

By these two formulas the error covariance function for least squares prediction is given. Both formulas have a form similar to that of (7-63) and are equally well suited for automatic computations, so that $\widetilde{\Delta g}$ and its accuracy can be calculated at the same time.

Practical considerations. Geometrical interpolation (Sec. 7-5) is suited for the interpolation of point anomalies in a dense gravity net, with station distances of 10 km or less. If mean anomalies for blocks of $5' \times 5'$ or larger are needed rather than point anomalies, then some kind of representation, such as that considered in the previous section, may be simpler and hardly less accurate.

Least squares prediction is, of course, more accurate than either geometrical interpolation or representation, but the improvement in accuracy is not striking. The main advantage of least squares prediction is that it permits a systematic, purely numerical processing of gravity data; the construction of gravity anomaly maps is no longer necessary. The same formula applies to both interpolation and extrapolation, so that gaps in the gravity data make no difference in the method of computation, which becomes completely schematic. Because large matrices are involved, an electronic high-speed computer is indispensable. For practical and computational details see Rapp (1964).

For larger station distances, of 50 km or more, prediction of individual point values becomes meaningless. In this case we must work with mean anomalies of, say, $1° \times 1°$ blocks. This will be the subject of Sec. 7-9.

7-7. Error Propagation. Accuracy of Spherical Harmonics

The gravity anomalies are the observational data from which other quantities of geodetic interest, such as geoidal undulations, deflections of the vertical, or the external gravity field, are computed. All these computations are done by means of integral formulas. The problem is now to estimate the accuracy of these derived quantities from the given accuracy of the gravity anomalies.

Conventional error theory does not directly cover this case. It must be slightly modified; this is done by a natural and logical extension of the usual theory of error propagation. Readers interested in the general method are referred to the literature (Moritz 1961, 1964a); here we must limit ourselves to two practical cases, which will be considered in this and in the following section.

The first problem is this. The gravity anomaly field is expanded into a series of fully normalized spherical harmonics (7-13):

$$\Delta g(\theta, \lambda) = \sum_{n=2}^{\infty} \sum_{m=0}^{n} [\bar{a}_{nm}\bar{R}_{nm}(\theta, \lambda) + \bar{b}_{nm}\bar{S}_{nm}(\theta, \lambda)],$$

where

$$\begin{Bmatrix} \bar{a}_{nm} \\ \bar{b}_{nm} \end{Bmatrix} = \frac{1}{4\pi} \int_{\lambda=0}^{2\pi} \int_{\theta=0}^{\pi} \Delta g(\theta, \lambda) \begin{Bmatrix} \bar{R}_{nm}(\theta, \lambda) \\ \bar{S}_{nm}(\theta, \lambda) \end{Bmatrix} \sin\theta \, d\theta \, d\lambda. \tag{7-66}$$

The error covariance function σ of the gravity anomalies is given; we need to determine the accuracy of the coefficients $\bar{a}_{nm}$ and $\bar{b}_{nm}$, that is, their error variances (standard errors) and covariances.

We denote the individual error ϵ_P of the gravity anomaly at a point P with coordinates θ and λ by

$$\epsilon(\theta, \lambda).$$

The totality of these errors at all points of the sphere obviously forms a function of θ and λ. The error covariance function is then, according to (7-59a) and (7-57), given by

$$\sigma(\theta, \lambda, \theta', \lambda') = M\{\epsilon(\theta, \lambda)\, \epsilon(\theta', \lambda')\} \tag{7-67}$$

as the average product of the individual errors at two points with coordinates θ, λ and θ', λ'. The error covariance σ is here considered as a function of spherical coordinates θ, λ rather than a function of plane coordinates x, y.

The effect of these errors $\epsilon(\theta, \lambda)$ on the coefficient $\bar{a}_{nm}$ is, according to (7-66), represented by

$$\eta = \frac{1}{4\pi} \int_{\lambda=0}^{2\pi} \int_{\theta=0}^{\pi} \epsilon(\theta, \lambda)\bar{R}_{nm}(\theta, \lambda) \sin\theta \, d\theta \, d\lambda, \tag{7-68}$$

where η is thus the individual error of $\bar{a}_{nm}$. The error variance of $\bar{a}_{nm}$, the square of its standard error, is evidently given by

$$m^2 \equiv M\{\eta^2\} \tag{7-69}$$

as the average of the individual η^2. Hence, we must first form η^2. We have

$$\begin{aligned}
\eta^2 &= \frac{1}{16\pi^2} \left[\int_{\lambda=0}^{2\pi} \int_{\theta=0}^{\pi} \epsilon(\theta, \lambda)\bar{R}_{nm}(\theta, \lambda) \sin\theta \, d\theta \, d\lambda \right]^2 \\
&= \frac{1}{16\pi^2} \int_{\lambda=0}^{2\pi} \int_{\theta=0}^{\pi} \epsilon(\theta, \lambda)\bar{R}_{nm}(\theta, \lambda) \sin\theta \, d\theta \, d\lambda \\
&\qquad\qquad \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} \epsilon(\theta', \lambda')\bar{R}_{nm}(\theta', \lambda') \sin\theta' \, d\theta' \, d\lambda' \\
&= \frac{1}{16\pi^2} \int_{\lambda=0}^{2\pi} \int_{\theta=0}^{\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} \epsilon(\theta, \lambda)\, \epsilon(\theta', \lambda')\bar{R}_{nm}(\theta, \lambda)\bar{R}_{nm}(\theta', \lambda') \sin\theta \\
&\qquad\qquad \sin\theta' \, d\theta \, d\lambda \, d\theta' \, d\lambda'.
\end{aligned}$$

Here we have used two well-known theorems of integral calculus:

1. The symbols that denote variables of integration in a definite integral are irrelevant; they may be replaced by any other symbols. In our case, θ, λ have been replaced by θ', λ' in the second integral.
2. Products of definite integrals may be written as one multiple integral.

Now we can average the last equation to get the standard error m according to (7-69). We find

$$m^2 = \frac{1}{16\pi^2} \int_{\lambda=0}^{2\pi} \int_{\theta=0}^{\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} M\{\epsilon(\theta, \lambda)\, \epsilon(\theta', \lambda')\} \bar{R}_{nm}(\theta, \lambda) \bar{R}_{nm}(\theta', \lambda') \sin\theta \sin\theta' \, d\theta \, d\lambda \, d\theta' \, d\lambda',$$

We have been able to place the symbol M inside the integral because M, by its definition as the average over the unit sphere, is really a double integral, and the order of integrals with fixed finite limits can be interchanged.

The definition (7-67) finally yields

$$m^2 = \frac{1}{16\pi^2} \int_{\lambda=0}^{2\pi} \int_{\theta=0}^{\pi} \int_{\lambda'=0}^{2\pi} \int_{\theta'=0}^{\pi} \sigma(\theta, \lambda, \theta', \lambda') \bar{R}_{nm}(\theta, \lambda) \bar{R}_{nm}(\theta', \lambda') \sin\theta \sin\theta' \, d\theta \, d\lambda \, d\theta' \, d\lambda'. \quad (7\text{-}70)$$

This is the desired formula for the standard error of the spherical-harmonic coefficient $\bar{a}_{nm}$. If we want the standard error of the coefficient $\bar{b}_{nm}$, we must merely replace $\bar{R}_{nm}$ by the corresponding function $\bar{S}_{nm}$.

This formula thus solves a particular problem of error propagation in gravimetric computations. Like (7-66), it is an integral formula. The error covariance function σ enters essentially; we thus see the fundamental importance of σ for error propagation. If the error function is given, then the evaluation of the integral (7-70) can be effected without theoretical difficulties, for instance, by numerical integration.

A particularly simple result is obtained if we subject the error function to two assumptions:

1. Only errors at neighboring points are noticeably correlated; beyond a certain distance there is no correlation.
2. The accuracy is the same for every point of the earth's surface.

Let us examine what these assumptions mean practically. The principal inaccuracies of the gravity anomalies are caused by interpolation. If other errors are neglected, then the error covariance function may be computed by the formulas of the preceding section. Assumption 1 is natural because, in a reasonably dense gravity net, interpolation errors at points that are at some distance apart are practically uncorrelated. Assumption 2 holds in the idealized case of uniform coverage of the whole earth by gravity measurements. It merely states that the accuracy is the same at every point; the accuracy may, however, be different in different directions, as in the case of profile measurements.

The crucial point, which permits a drastic simplification of the fourfold integral (7-70), is that by assumption 1 the integrand will be noticeably different from zero only if $\theta' \doteq \theta$ and $\lambda' \doteq \lambda$, because the error function for two distant points is zero. Hence we may approximate (7-70) by

$$m^2 \doteq \frac{1}{16\pi^2}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi}\int_{\lambda'=0}^{2\pi}\int_{\theta'=0}^{\pi} \sigma(\theta, \lambda, \theta', \lambda')[\bar{R}_{nm}(\theta, \lambda)]^2 \sin\theta \sin\theta' \, d\theta \, d\lambda \, d\theta' \, d\lambda'$$

and perform the integration over θ' and λ' first. We set

$$\int_{\lambda'=0}^{2\pi}\int_{\theta'=0}^{\pi} \sigma(\theta, \lambda, \theta', \lambda') \sin\theta' \, d\theta' \, d\lambda' = \frac{S}{R^2} \tag{7-71}$$

($R = 6371$ km); by assumption 2 this will be a constant independent of position. The quantity S will be called *error constant;* a practical way of computing it and numerical values will be given in Sec. 7-9.

Then the formula for m^2 becomes

$$m^2 = \frac{S}{16\pi^2 R^2}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi} [\bar{R}_{nm}(\theta, \lambda)]^2 \sin\theta \, d\theta \, d\lambda. \tag{7-72}$$

By equation (1-74), the integral is 4π, so that we finally obtain the simple result

$$m^2 = \frac{S}{4\pi R^2}, \tag{7-73}$$

where m is the standard error of any coefficient $\bar{a}_{nm}$. For $\bar{b}_{nm}$, the function $\bar{R}_{nm}$ must be replaced by $\bar{S}_{nm}$, which obviously gives the same result.

Thus the standard errors of all fully normalized coefficients $\bar{a}_{nm}$ and $\bar{b}_{nm}$ are equal and given by (7-73).

Let us now compute the error covariance of two different spherical-harmonic coefficients $\bar{a}_{nm}$ and $\bar{a}_{pq}$. The individual error η of $\bar{a}_{nm}$ is given by (7-68); the error η^* of $\bar{a}_{pq}$ is

$$\eta^* = \frac{1}{4\pi}\int_{\lambda'=0}^{2\pi}\int_{\theta'=0}^{\pi} \epsilon(\theta', \lambda')\bar{R}_{pq}(\theta', \lambda') \sin\theta' \, d\theta' \, d\lambda'.$$

The error covariance of $\bar{a}_{nm}$ and $\bar{a}_{pq}$ is defined as

$$\sigma(\bar{a}_{nm}, \bar{a}_{pq}) = M\{\eta\eta^*\}.$$

Repeating the procedure leading to (7-70) we find

$$\sigma(\bar{a}_{nm}, \bar{a}_{pq}) = \frac{1}{16\pi^2}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi}\int_{\lambda'=0}^{2\pi}\int_{\theta'=0}^{\pi} \sigma(\theta, \lambda, \theta', \lambda')\bar{R}_{nm}(\theta, \lambda)\bar{R}_{pq}(\theta', \lambda') \sin\theta \sin\theta' \, d\theta \, d\lambda \, d\theta' \, d\lambda'.$$

Instead of (7-72) we now have

$$\sigma(\bar{a}_{nm}, \bar{a}_{pq}) = \frac{S}{16\pi^2 R^2}\int_{\lambda=0}^{2\pi}\int_{\theta=0}^{\pi} \bar{R}_{nm}(\theta, \lambda)\bar{R}_{pq}(\theta, \lambda) \sin\theta \, d\theta \, d\lambda.$$

Because of the orthogonality of two different spherical harmonics this is zero. We would have obtained the same result if we had replaced $\bar{R}_{pq}$ by $\bar{S}_{pq}$ to get the error covariance between the coefficients $\bar{a}_{nm}$ and $\bar{b}_{pq}$. Thus all the coefficients $\bar{a}_{nm}$ and $\bar{b}_{nm}$ are uncorrelated.

As a matter of fact, these simple results hold only as long as the approximate substitution leading from (7-70) to (7-72) is permissible. As one easily recognizes, it breaks down for spherical harmonics of very high degree n, but is valid for the harmonics of lower degree, which are of greatest geodetic interest.

Using these results, one can also easily compute the accuracy of the coefficients J_{nm} and K_{nm} of the gravitational potential V (Moritz, 1964a).

7-8. Accuracy of Geoidal Undulations Computed from Gravity Anomalies

This problem initiated the application of statistical techniques to gravimetric geodesy. Two basic papers (de Graaff-Hunter, 1935; Hirvonen, 1956) have been devoted to it. The second gave rise to an extensive modern development.

We shall again consider an idealized gravity net that is uniform and homogeneous over the whole earth, and study the accuracy in the geoidal undulation N obtainable with such a gravity net. This question is of importance because the result indicates how a gravity survey must be planned in order to achieve a certain prescribed accuracy in N. It is therefore considered in several publications: de Graaff-Hunter (1935), Kaula (1957), Groten and Moritz (1964).

We shall thus study the error propagation in Stokes' formula

$$N = \frac{R}{4\pi G}\int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\pi} \Delta g(\psi, \alpha)S(\psi)\sin\psi\, d\psi\, d\alpha.$$

This is done in quite the same way as in the preceding section. The individual error of N is given by

$$\eta = \frac{R}{4\pi G}\int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\pi} \epsilon(\psi, \alpha)S(\psi)\sin\psi\, d\psi\, d\alpha;$$

and its square becomes

$$\eta^2 = \left(\frac{R}{4\pi G}\right)^2 \int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\pi} \epsilon(\psi, \alpha)S(\psi)\sin\psi\, d\psi\, d\alpha \cdot \int_{\alpha'=0}^{2\pi}\int_{\psi'=0}^{\pi} \epsilon(\psi', \alpha')S(\psi')\sin\psi'\, d\psi'\, d\alpha'$$

or

$$\eta^2 = \left(\frac{R}{4\pi G}\right)^2 \int_{\alpha=0}^{2\pi}\int_{\psi=0}^{\pi}\int_{\alpha'=0}^{2\pi}\int_{\psi'=0}^{\pi} \epsilon(\psi, \alpha)\,\epsilon(\psi', \alpha')S(\psi)S(\psi')\sin\psi \sin\psi'\, d\psi\, d\alpha\, d\psi'\, d\alpha'.$$

Forming the average M of both sides of this equation we find

$$m^2 = \left(\frac{R}{4\pi G}\right)^2 \int_{\alpha=0}^{2\pi} \int_{\psi=0}^{\pi} \int_{\alpha'=0}^{2\pi} \int_{\psi'=0}^{\pi} \sigma(\psi, \alpha, \psi', \alpha') S(\psi) S(\psi') \sin\psi \sin\psi' \, d\psi \, d\alpha \, d\psi' \, d\alpha'. \quad (7\text{-}74)$$

Here m is the standard error of N, and $\sigma(\theta, \lambda, \theta', \lambda')$ is the error function of the gravity anomalies. This is the general formula for error propagation in Stokes' formula. It is valid for an arbitrary form of the error function.

This equation may again be simplified drastically if we make the two assumptions—no correlation of errors beyond a certain small distance and uniform accuracy—which we already used in the preceding section. We apply the same trick as with equation (7-70). We set $S(\psi') \doteq S(\psi)$ and then perform the integration over ψ'.

Using the error constant S according to (7-71) we obtain

$$m^2 = \frac{S}{16\pi^2 G^2} \int_{\alpha=0}^{2\pi} \int_{\psi=0}^{\pi} [S(\psi)]^2 \sin\psi \, d\psi \, d\alpha.$$

The integration with respect to α may now be performed at once; we finally get

$$m^2 = \frac{S}{8\pi G^2} \int_{\psi=0}^{\pi} [S(\psi)]^2 \sin\psi \, d\psi.$$

This formula is very simple, but unfortunately it does not hold in this form; in fact, it yields the value ∞. The reason is that if we approximately replace $S(\psi')$ by $S(\psi)$ we must assume that for $\psi' \doteq \psi$ we also have $S(\psi') \doteq S(\psi)$. This is not so at the neighborhood of the origin $\psi = 0$, because $S(\psi)$ increases rapidly there and is, in fact, discontinuous at the origin: $S(\psi) \longrightarrow \infty$ if $\psi \longrightarrow 0$.

We must therefore exclude the origin by beginning the integration with $\psi = \psi_0$ (ψ_0 small) instead of with $\psi = 0$:

$$m^2 = \frac{S}{8\pi G^2} \int_{\psi=\psi_0}^{\pi} [S(\psi)]^2 \sin\psi \, d\psi. \quad (7\text{-}75)$$

The small neighborhood $\psi < \psi_0$ must then be taken into account in some other way, for which the reader is referred to Groten and Moritz (1964).

The integral in (7-75) may be evaluated in several ways. One possibility is to take the functions $S(\psi)$ and $S(\psi) \sin\psi$ from the tables of Lambert and Darling (1936), referred to in Chapter 2, and compute the integral by numerical integration. For certain values of ψ_0 the integral

$$\int_{\psi_0}^{\pi} [S(\psi)]^2 \sin\psi \, d\psi,$$

computed in this way, is tabulated in the paper by Groten and Moritz referred to above. There is also a closed formula for the integral, given in Molodenskii et al. (1962, p. 157), but it is rather complicated.

The numerical values of Table 7-3 were computed on the basis of the results, particularly for the error constant S, of the following section. They also in-

Table 7-3
Standard Error of the Geoidal Undulation for Idealized Uniform Gravity Distributions
Unit 1 meter

Block	$1^\circ \times 1^\circ$	$2^\circ \times 2^\circ$	$5^\circ \times 5^\circ$	$10^\circ \times 10^\circ$
Point	± 1.5	± 5	± 13	± 25
Profile	± 1.2	± 3	± 7	± 9

corporate the central zone $\psi < \psi_0$, which is excluded in (7-75), and correspond to the cases in which there is one arbitrarily situated point, or one central east-west profile gravity measurement in each $1^\circ \times 1^\circ$, $2^\circ \times 2^\circ$, $5^\circ \times 5^\circ$, or $10^\circ \times 10^\circ$ block, uniformly over the entire surface of the earth.

A final remark seems to be in order about the error constant S, which should not be confused with Stokes' function $S(\psi)$, in general problems of error propagation. Assume that the error covariance function $\sigma(\theta, \lambda, \theta', \lambda')$ satisfies the assumptions 1 and 2 of the preceding section and that we can apply the trick of replacing θ', λ' by θ, λ in part of the integrand. This is possible if the particular part of the integrand changes slowly and continuously with θ and λ, instances being found in the last two sections. Then the error covariance function enters into the error propagation formula only through the error constant S, which can be computed once for all and is independent of the particular problem of error propagation. Thus the central role of S becomes evident.

7-9. Accuracy of Mean Anomalies

The mean gravity anomaly $\overline{\Delta g}$ of a rectangular block $ABCD$ of sides a and b is expressed by

$$\overline{\Delta g} = \frac{1}{ab} \int_{x=0}^{a} \int_{y=0}^{b} \Delta g(x, y)\, dx\, dy \tag{7-76}$$

(Fig. 7-5). This rigorous formula presupposes that the gravity anomaly Δg be given at every point (x, y) inside the rectangle $ABCD$.

In practice we have measured Δg at only a few points within the rectangle;

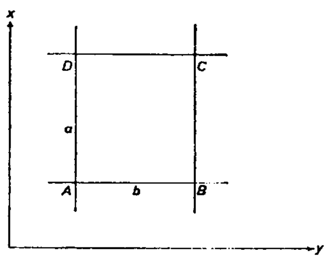

FIGURE 7-5
The mean anomaly of a rectangle.

the problem is to estimate the mean anomaly $\overline{\Delta g}$ from these measurements. One way is to interpolate or predict Δg at all other points of the block by the methods of Sec. 7-5 and to compute $\overline{\Delta g}$ from these estimated point anomalies Δg by formula (7-76).

We may also use a more direct way. In analogy to (7-47), we may approximate $\overline{\Delta g}$ by a linear combination of the measured values $\Delta g_1, \Delta g_2, \ldots, \Delta g_n$:

$$\widetilde{\overline{\Delta g}} = \alpha_1 \Delta g_1 + \alpha_2 \Delta g_2 + \cdots + \alpha_n \Delta g_n = \sum_{i=1}^{n} \alpha_i \Delta g_i. \tag{7-77}$$

The error of the predicted value $\widetilde{\overline{\Delta g}}$ is clearly the difference

$$\epsilon = \overline{\Delta g} - \widetilde{\overline{\Delta g}} = \overline{\Delta g} - \sum_{i=1}^{n} \alpha_i \Delta g_i. \tag{7-78}$$

By squaring we get

$$\epsilon^2 = \overline{\Delta g}^2 - 2 \sum_{i=1}^{n} \alpha_i \Delta g_i \overline{\Delta g} + \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_i \alpha_k \Delta g_i \Delta g_k.$$

To find the standard error m of the estimated mean anomaly, we form the average M, obtaining

$$m^2 = \bar{\bar{C}} - 2 \sum_{i=1}^{n} \alpha_i \bar{C}_i + \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_i \alpha_k C_{ik}. \tag{7-79}$$

The quantity C_{ik} is defined by (7-54); the quantity

$$\bar{\bar{C}} \equiv M\{\overline{\Delta g}^2\} \tag{7-80}$$

is the mean square of the mean block anomaly $\overline{\Delta g}$, or its variance; and

$$\bar{C}_i \equiv M\{\Delta g_i \overline{\Delta g}\} \tag{7-81}$$

is the covariance between the point Δg_i and the mean anomaly $\overline{\Delta g}$.

These quantities can be expressed in terms of the covariance function $C(s)$. On inserting (7-76) into (7-80) and taking the definition (7-5) of the covariance function into account, we readily find

$$\bar{\bar{C}} = \frac{1}{a^2 b^2} \int_{x=0}^{a} \int_{y=0}^{b} \int_{x'=0}^{a} \int_{y'=0}^{b} C(\sqrt{(x - x')^2 + (y - y')^2})\, dx\, dy\, dx'\, dy'. \tag{7-82}$$

Similarly

$$\bar{C}_i = \frac{1}{ab} \int_{x=0}^{a} \int_{y=0}^{b} C(\sqrt{(x - x_i)^2 + (y - y_i)^2})\, dx\, dy, \tag{7-83}$$

where (x_i, y_i) are the coordinates of the point at which Δg_i is measured.

If there is only one measured gravity anomaly Δg_1 in the block, the prediction formula (7-77) becomes

$$\widetilde{\overline{\Delta g}} = \alpha \Delta g_1, \tag{7-84}$$

and equation (7-79) is simplified to

$$m^2 = \bar{\bar{C}} - 2\alpha\bar{C}_1 + \alpha^2 C_0, \tag{7-85}$$

where we have set $\alpha_1 = \alpha$ and $C_0 = C(0)$.

The α_i in (7-77) and α in (7-84) may be chosen in different ways. Particularly simple is the case of *direct representation*, $\alpha = 1$. The mean anomaly $\overline{\Delta g}$ is directly approximated, or represented, by the measured anomaly Δg_1. Equation (7-84) then becomes

$$\overline{\overline{\Delta g}} = \Delta g_1, \tag{7-86}$$

and (7-85) reduces to

$$m^2 = \bar{\bar{C}} - 2\bar{C}_1 + C_0. \tag{7-87}$$

Equation (7-87) depends on the location (x_1, y_1) of the gravity station through $\bar{C}_1$, equation (7-83). It is also useful to consider the average error variance $\overline{m}^2$ for an arbitrary situation of the gravity observation within the square:

$$\overline{m}^2 = \frac{1}{ab}\int_{x_1=0}^{a}\int_{y_1=0}^{b} m^2(x_1, y_1)\,dx_1\,dy_1. \tag{7-88}$$

Averaging (7-87) we must keep in mind that $\bar{\bar{C}}$ and C_0, being constants, remain unchanged, whereas the average of C_1 becomes

$$\frac{1}{ab}\int_{x_1=0}^{a}\int_{y_1=0}^{b} \bar{C}_1\,dx_1\,dy_1 = \bar{\bar{C}}.$$

This is immediately seen on comparing (7-82) and (7-83).

Hence we get simply

$$\overline{m}^2 = C_0 - \bar{\bar{C}}. \tag{7-89}$$

Hirvonen (1956), to whom this formula is due, wrote it in a particularly elegant and instructive form:

$$E_s^2 = G_0^2 - G_s^2. \tag{7-89'}$$

He called E_s the (standard) error of representation. The symbol G_s is the r.m.s. mean gravity anomaly of a block of side s[1] (he considered square blocks with $a = b = s$); this follows from the definition (7-80), $\bar{\bar{C}} = G_s^2$. Accordingly, G_0 is the r.m.s. point anomaly, which may be considered as a mean anomaly of a block of side $s = 0$; in our notation $G_0^2 = C_0$.

The preceding formulas may be investigated for other prediction methods by assuming different values of α_i. The values of α_i that minimize m^2, equation (7-79), are readily found (least squares prediction). All this is done along lines similar to those in Secs. 7-5 and 7-6.

Generalizations and extensions are obvious. Besides error variances m^2 we may also consider error covariances of different blocks. These can be used to compute the error constant S mentioned in the preceding sections. Another extension is to profile observations, where gravity is measured along profiles

[1] In the present chapter we are using s in a different meaning!

rather than at point stations. All this, however, is beyond the scope of the present book; the reader is referred to Moritz (1964b).

Numerical results. We shall merely give some numerical values from Moritz (1964b), with explanation but without detailed formulas. Basically, the error variances m^2 of (7-85), and the corresponding error covariances, were computed for different α. This is the case in which there is one gravity station in each block. There is a similar set of formulas for the error variances and covariances for one measured gravity profile in each block; these formulas were also evaluated. The integrations were performed on the basis of the estimated covariances $C(\psi)$ of Table 7-1, using an electronic computer. The author considered $1° \times 1°$, $2° \times 2°$, $5° \times 5°$, and $10° \times 10°$ blocks in 45° latitude, so that a $10° \times 10°$ block is a rectangle, 1112 km $\times$ 788 km.

Table 7-4 shows the error variances and error covariances for point gravity

Table 7-4

Error Variances and Covariances (mgals²). Point Observations

	$1° \times 1°$			$2° \times 2°$			$5° \times 5°$			$10° \times 10°$		
Zero Anomaly	844	584	379	591	388	298	375	270	155	266	144	55
($\alpha = 0$)	508	430	348	356	318	282	230	190	114	110	76	42
	334	327	313	269	262	243	95	85	68	34	30	22
Representation	153	4	−1	434	−1	0	763	1	2	852	4	1
($\alpha = 1$)	−2	−4	−1	6	4	−1	8	5	−4	−13	1	0
	1	1	0	0	0	0	0	−1	1	0	0	0
	($\alpha = 0.788$)			($\alpha = 0.565$)			($\alpha = 0.338$)			($\alpha = 0.256$)		
Least Standard Error	99	26	18	208	77	56	238	116	67	188	77	31
	27	20	16	77	63	53	100	82	51	63	44	23
	17	16	14	50	49	45	44	39	30	19	17	12
	($\alpha = 0.991$)			($\alpha = 0.959$)			($\alpha = 0.862$)			($\alpha = 0.659$)		
Least Error Constant	148	4	0	394	0	0	567	5	4	383	15	7
	−2	−3	−1	8	4	0	9	6	0	10	12	5
	1	1	0	0	1	0	3	2	2	4	4	2
Average	356	22	−6	609	−15	2	825	12	7	934	18	1
Representation	−34	−21	−7	−26	−5	1	13	12	−9	−26	−9	0
($\alpha = 1$)	−5	−2	0	2	3	3	−10	−8	−1	−2	−2	−1

observations. The first value in the top line of each section (for zero anomaly, representation, etc.) is the error variance; the second value in the top line of each section is the error covariance between a block and its neighbor to the east (or west); the third value in each top line is the error covariance between two blocks that have the same latitude and are separated by one other block, etc.

The relative position of any two blocks under consideration is thus directly represented by the place occupied by their covariance in the table.

The meaning of zero anomaly ($\alpha = 0$) and representation ($\alpha = 1$) is clear. "Least standard error" corresponds to that value of α which minimizes m^2 (7-85):

$$\alpha = \frac{\bar{C}_1}{C_0}; \tag{7-90}$$

"least error constant" refers to that α which minimizes the error constant S.

For these first four items the gravity station was assumed to lie at the center of each block. The last item, "average representation," refers to a random position of the gravity station within the block. The corresponding error variances are expressed by (7-89), whereas the error variance for "representation" was given by (7-87).

Table 7-5 shows the analogous results for the accuracy of profile gravity

Table 7-5

Error Variances and Covariances (mgals²). Profile Observations

	1° × 1°			2° × 2°			5° × 5°			10° × 10°		
Zero Anomaly	844	584	379	591	388	298	375	270	155	266	144	55
($\alpha = 0$)	508	430	348	356	318	282	230	190	114	110	76	42
	334	327	313	269	262	243	95	85	68	34	30	22
Representation	80	23	2	165	26	0	180	13	0	125	10	0
($\alpha = 1$)	1	−3	−1	10	2	0	7	3	−4	−9	−1	0
	2	1	0	0	1	0	0	0	0	0	0	0
	($\alpha = 0.868$)			($\alpha = 0.752$)			($\alpha = 0.666$)			($\alpha = 0.660$)		
Least Standard Error	62	22	7	113	34	18	115	32	15	74	15	6
	15	8	6	35	22	17	29	22	11	14	10	5
	8	7	5	17	16	15	13	11	8	4	4	2
	($\alpha = 0.992$)			($\alpha = 0.977$)			($\alpha = 0.956$)			($\alpha = 0.929$)		
Least Error Constant	78	22	1	156	25	0	164	11	0	106	7	0
	1	−2	−1	11	2	0	6	2	−4	−6	0	0
	2	1	0	0	1	0	1	0	0	0	0	0
Average	193	70	10	254	51	4	211	28	10	176	34	3
Representation	−46	−20	−2	−29	−6	1	6	6	−5	−30	−10	0
($\alpha = 1$)	−5	−2	0	1	2	2	−12	−8	−1	−3	−2	0

measurements. "Representation," "least standard error," and "least error constant" refer to equally spaced east-west profiles through the center of each block, whereas "average representation" corresponds to a random position of the east-west profile within the block.

The variances are, of course, smallest for the assumption of a centrally located gravity station or profile. They will be larger for other situations of the observations. This may be recognized by comparing "representation," which refers to the central case, to "average representation," where we have averaged over observations distributed over the entire block. Note that for larger blocks the location of the observations has less influence.

Table 7-6
Error Constants
S/R^2 (mgals^2)

	Point				Profile			
	1° × 1°	2° × 2°	5° × 5°	10° × 10°	1° × 1°	2° × 2°	5° × 5°	10° × 10°
Representation	0.029	0.40	4.2	18.1	0.027	0.21	1.2	2.5
Least Standard Error	0.184	2.14	12.3	20.6	0.084	0.77	3.5	5.1
Least Error Constant	0.029	0.38	3.6	11.8	0.026	0.21	1.1	2.4
Average Representation	0.040	0.52	4.5	18.7	0.032	0.30	1.3	2.9

Finally, Table 7-6 shows the corresponding error constants, or rather the quantities S/R^2, where $R = 6371$ km.

These tables show that the different methods of estimation differ widely in accuracy and error correlation. "Least standard error" has a rather large error correlation, so that it is not best with respect to error propagation. This is particularly clear from the error constants of Table 7-6: we have seen in the preceding section that the error constant S, rather than the standard error m, is relevant for error propagation. Thus the error constant should in general be minimized rather than the standard error, but the results of direct representation ($\alpha = 1$) are almost as good. "Least standard error" is distinctly inferior with respect to error propagation. It gives too small an α; if we interpret $\alpha < 1$ as a weighted average of observed anomaly and zero anomaly, then too much weight is given to the zero anomaly, which has a large correlation.

7-10. Correlation with Elevation

So far we have taken into account only the mutual correlation of the gravity anomalies, their autocorrelation, disregarding the correlation with elevation, which is important in many cases. Therefore our formulas were valid only for gravity anomalies uncorrelated with elevation, such as isostatic or, to a certain extent, Bouguer anomalies; or for free-air anomalies in moderately flat areas. Free-air anomalies in mountains must be treated differently.

Fig. 7-6, following Uotila (1960), shows the correlation of free-air anomalies with elevation. The gravity anomalies Δg are plotted against the elevation h. If there were an exact functional dependence between Δg and h, then all points would lie on a straight line (or, more generally, on a curve). In reality, there is only an approximate functional relation, a general trend or tendency of the free-air anomalies to increase linearly with elevation; exceptions, even large ones, are possible. This shows very well the meaning of correlation.

We have characterized the mutual correlation of the gravity anomalies by the autocovariance function (7-5),

$$C(s) = M\{\Delta g\, \Delta g'\},$$

where $s = PP'$. In a similar way we may form the functions

$$B(s) = M\{\Delta g\, \Delta h'\} = M\{\Delta g'\, \Delta h\}, \tag{7-91}$$

expressing the correlation between gravity and elevation, and

$$A(s) = M\{\Delta h\, \Delta h'\}, \tag{7-92}$$

which is the autocovariance function of the elevation differences

$$\Delta h = h - M\{h\}; \tag{7-93}$$

the symbol $M\{h\}$ denotes the mean elevation of the whole area considered.

If Δg and Δh are not correlated, then the function $B(s)$ is identically zero. If this is not the case, then we should also take the elevation into account in our interpolation.

It is easy to extend the prediction formula (7-47) for this purpose. Restricting ourselves to predictions linear in Δh as well as in Δg, we may write

$$\widetilde{\Delta g}_P = \sum \alpha_{Pi}\, \Delta g_i + \sum_i \beta_{Pi}\, \Delta h_i - \beta \Delta h_P, \tag{7-94}$$

where the coefficients α_{Pi}, β_{Pi}, and β do not depend on Δg or Δh.

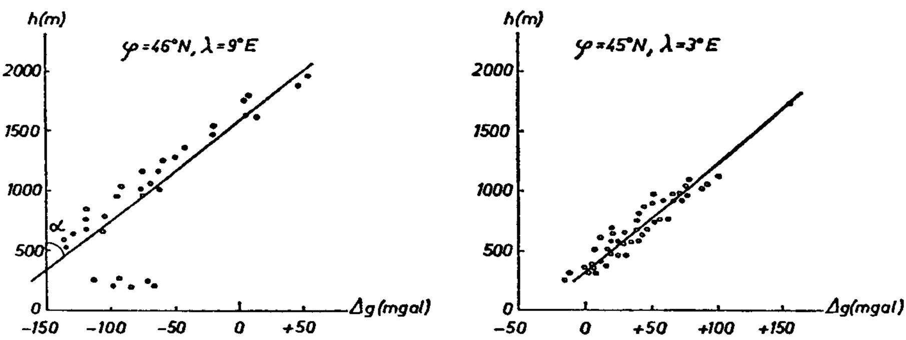

FIGURE 7-6

Correlation of the free-air anomalies with elevation.

In statistical terminology, this is the removal of trend (with respect to elevation) by *linear regression*. Similarly, (7-47) is an *autoregressive* formula.

The error of prediction is

$$\epsilon_P = \Delta g_P - \widetilde{\Delta g}_P = \Delta g_P + \beta \Delta h_P - \sum_i \alpha_{Pi}\,\Delta g_i - \sum_i \beta_{Pi}\,\Delta h_i.$$

Squaring and averaging in the usual way yields

$$\begin{aligned} m_P^2 = C_0 &+ 2\beta B_0 + \beta^2 A_0 - 2\sum_i \alpha_{Pi} C_{Pi} - 2\left(\sum_i \beta_{Pi} + \beta \sum_i \alpha_{Pi}\right) B_{Pi} \\ &- 2\beta \sum_i \beta_{Pi} A_{Pi} + \sum_i \sum_k \alpha_{Pi}\alpha_{Pk} C_{ik} + 2\sum_i \sum_k \alpha_{Pi}\beta_{Pk} B_{ik} \\ &+ \sum_i \sum_k \beta_{Pi}\beta_{Pk} A_{ik}, \end{aligned} \tag{7-95}$$

where

$$\begin{aligned} A_0 &= A(0), & B_0 &= B(0), & C_0 &= C(0), \\ A_{Pi} &= A(Pi), & B_{Pi} &= B(Pi), & C_{Pi} &= C(Pi), \\ A_{ik} &= A(ik), & B_{ik} &= B(ik), & C_{ik} &= C(ik); \end{aligned}$$

P being the point at which Δg is to be predicted, and i or k denoting the given gravity stations.

This formula, which is clearly an extension of (7-56), gives the standard error of prediction if correlation with elevation is taken into account. It is easy to find a formula for the error covariance function, generalizing (7-60), and matrix formulas corresponding to (7-63) through (7-65) for a least squares prediction that minimizes (7-95); see Moritz (1963). It should be noted that the functions A, B, and C, *but no other statistical quantities*, enter into these formulas.

Application to Bouguer anomalies. Of great practical importance is the question whether it is possible to render the free-air anomalies independent of elevation by adding a term that is proportional to the elevation. In other words, when is the quantity

$$z = \Delta g - b\,\Delta h, \tag{7-96}$$

with a certain coefficient b, uncorrelated with elevation?

The form of z is that of a Bouguer anomaly; for a real Bouguer anomaly we have, according to Sec. 3-3,

$$b = 2\pi k\rho; \tag{7-97}$$

if the density $\rho = 2.67$ g/cm^3, then

$$b = +0.112 \text{ mgal/meter}. \tag{7-97'}$$

Let us form the covariance function $Z(s)$ of the "Bouguer anomaly" (7-96) with elevation:

$$Z(s) \equiv M\{z\,\Delta h'\} = M\{\Delta g\,\Delta h' - b\,\Delta h\,\Delta h'\} = B(s) - bA(s).$$

If z is to be independent of h, then $Z(s)$ must be identically zero. The condition is

$$B(s) - bA(s) \equiv 0, \tag{7-98}$$

which must be satisfied for all s and a certain constant b.

We see that the "Bouguer anomaly" z is uncorrelated with elevation if the functions $A(s)$ and $B(s)$ are proportional for the area considered; the constant b is then represented by

$$b = \frac{B(s)}{A(s)}. \tag{7-99}$$

It may be shown that this is equivalent to the condition that the points of Fig. 7-6 lie approximately on a straight line, and not on some other curve. The coefficient b is then given by

$$b = \tan \alpha \tag{7-100}$$

as the inclination of the line towards the h-axis.

In practice these conditions are very often fulfilled to a good approximation; and furthermore, by computing b from equation (7-99) or determining it graphically by means of (7-100), we often get a value that is close to the normal Bouguer gradient (7-97′).

If we assume that b depends only on the rock density ρ, then we obtain a means for determining the average density, which is often difficult to measure directly. This is the "Nettleton method," used in geophysical prospecting: the coefficient b is found statistically by means of equations (7-99) or (7-100), and the rock density ρ is then computed from (7-97). Figure 7-7 illustrates the principle of this method; see also Jung (1956, p. 600).

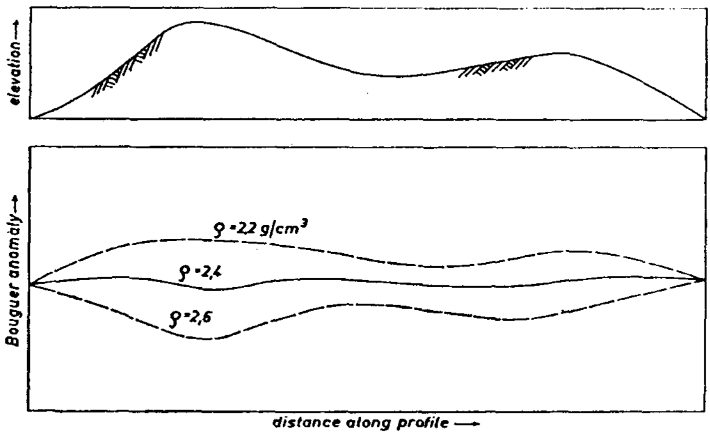

FIGURE 7-7

Bouguer anomalies corresponding to different densities ρ. The best density is $\rho = 2.4$ g/cm³ (no correlation); for other densities the Bouguer anomalies are correlated with elevation (positive correlation for $\rho = 2.2$, negative correlation for $\rho = 2.6$).

If condition (7-98) is fulfilled, then we may consider the "Bouguer anomaly" z as a gravity anomaly that is completely uncorrelated with elevation; we can directly apply to it the whole theory of the preceding sections. But even when this condition is not quite satisfied, Bouguer anomalies will in general be far less correlated with elevation than free-air anomalies. The fact that in (7-96) gravity is reduced to a mean elevation and not to sea level, is quite irrelevant in this connection, because this is only a question of an additive constant. From this crude statistical point of view, such refinements as terrain correction, etc., may also be disregarded.

It is thus possible to consider the Bouguer reduction as a means of obtaining gravity anomalies that are less dependent on elevation and hence more representative than free-air anomalies. More precisely, the Bouguer anomalies take care of the dependence on the *local* irregularities of elevation. The isostatic anomalies are, in addition, also largely independent of the *regional* features of topography. See also Chapter 3.

References

de Graaff-Hunter, J. (1935). The figure of the earth from gravity observations and the precision obtainable. Phil. Trans. Roy. Soc. London, ser. A., v. 234, pp. 377–431.

Groten, E., and H. Moritz (1964). On the accuracy of geoid heights and deflections of the vertical. Report No. 38, Inst. Geod. Phot. Cart., Ohio State Univ.

Hirvonen, R. A. (1956). On the precision of the gravimetric determination of the geoid. Trans. Am. Geophys. Union, v. 37, pp. 1–8.

Hirvonen, R. A. (1962). On the statistical analysis of gravity anomalies. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 37.

Jeffreys, H. (1962). The earth. 4th ed., Cambridge University Press.

Jung, K. (1956). Figur der Erde. Handbuch der Physik, v. 47, pp. 534–639, Berlin, Springer.

Kaula, W. M. (1957). Accuracy of gravimetrically computed deflections of the vertical. Trans. Am. Geophys. Union, v. 38, pp. 297–305.

Kaula, W. M. (1959). Statistical and harmonic analysis of gravity. J. Geophys. Research, v. 64, pp. 2401–2421.

Kaula, W. M. (1963). Determination of the earth's gravitational field. Reviews of Geophysics, v. 1, pp. 507–551.

Kaula, W. M. (1966). Global harmonic and statistical analysis of gravimetry. Proceedings of the Symposium "Extension of Gravity Anomalies to Unsurveyed Areas" (H. Orlin, ed.), Am. Geophys. Union, Geophys. Monograph No. 9, pp. 58–67.

Miller, K. S. (1956). Engineering mathematics. New York, Holt, Rinehart and Winston (New York, Dover Publications, 1963).

Molodenskii, M. S., V. F. Eremeev, and M. I. Yurkina (1962). Methods for study of the external gravitational field and figure of the earth. Transl. from Russian (1960). Jerusalem, Israel Program for Scientific Translations.

Moritz, H. (1961). Fehlertheorie der graphisch-mechanischen Integration; Grundzüge einer allgemeinen Fehlertheorie im Funktionenraum. Österreichische Zeitschrift für Vermessungswesen, Sonderheft 22.

Moritz, H. (1963). Interpolation and prediction of point gravity anomalies. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 40.

Moritz, H. (1964a). An extension of error theory with application to spherical harmonics. *Ibid.*, No. 44.

Moritz, H. (1964b). Accuracy of mean gravity anomalies obtained from point and profile measurements. *Ibid.*, No. 45.

Moritz, H. (1965). Schwerevorhersage und Ausgleichungsrechnung. Zeitschrift für Vermessungswesen, v. 90, pp. 181–184.

Rapp, R. (1964). The prediction of point and mean gravity anomalies through the use of a digital computer. Report No. 43, Inst. Geod. Phot. Cart., Ohio State Univ.

Uotila, U. A. (1960). Investigations on the gravity field and shape of the earth. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 33.

8

Modern Methods for Determining the Figure of the Earth

8-1. Introduction

In the preceding chapters we have usually followed what might be called the conservative approach to the problems of physical geodesy. The geodetic measurements—astronomical coordinates and azimuths, horizontal angles, gravity observations, etc.—are reduced to the geoid, and the "geodetic boundary-value problem" is solved for the geoid by means of Stokes' integral and similar formulas. The geoid then serves as a basis for establishing the position of points of the earth's surface.

The advantage of this approach is that the geoid is a level surface, capable of simple definition in terms of the physically meaningful and geodetically important potential W. The geoid represents the most obvious mathematical formulation of a horizontal surface at sea level. This is why the use of the geoid simplifies geodetic problems and makes them accessible to geometrical intuition.

The disadvantage is that the potential W inside the earth, and hence the geoid $W =$ const., depends on the density ρ because of Poisson's equation (2-6)

$$\Delta W = -4\pi k\rho + 2\omega^2.$$

Therefore, in order to determine or to use the geoid, the density of the masses at every point between the geoid and the ground must be known, at least theoretically. This is clearly impossible, and therefore some assumptions concerning the density must be made, which is unsatisfactory theoretically, even though the practical influence of these assumptions is usually very small.

For this reason it is of basic importance that Molodensky in 1945 was able to show that the physical surface of the earth can be determined from geodetic measurements alone, without using the density of the earth's crust. This requires that the concept of the geoid be abandoned. The mathematical formulation becomes more abstract and more difficult. Both the gravimetric method and the astrogeodetic method can be modified for this purpose. The gravity anomalies and the deflections of the vertical now refer to the ground, and no longer to sea level; the "height anomalies" at ground level take the place of the geoidal undulations.

These recent developments have considerably broadened our insight into the principles of physical geodesy and have also introduced powerful new methods for attacking classical problems. Hence their basic theoretical significance is hardly lessened by the fact that many scientists prefer to retain the geoid because of its conceptual and practical advantages.

In this chapter we shall first give a concise survey of the conventional determination of the geoid by means of gravity reductions, in order to understand better the modern ideas. After an exposition of Molodensky's theory we shall show how the new methods may be applied to classical problems such as gravity reduction or the determination of the geoid.

It should be mentioned that the terms "modern" and "conventional" merely serve as convenient labels; they do not imply any connotation of value or preferability.

8-2. Gravity Reductions and the Geoid

The integrals of Stokes and of Vening Meinesz and similar formulas presuppose that the disturbing potential T is harmonic on the geoid, which implies that there are no masses outside the geoid. This assumption—no masses outside the bounding surface—is necessary if we wish to treat any problem of physical geodesy as a boundary-value problem in the sense of potential theory. The reason is that the boundary-value problems of potential theory always involve harmonic functions, that is, solutions of Laplace's equation

$$\Delta T = 0.$$

We know, for instance, that the determination of T or N from the gravity anomalies Δg may be considered as a third boundary-value problem; see Sec. 2-13.

Since there are masses outside the geoid, they must be moved inside the geoid or completely removed, before we can apply Stokes' integral or related formulas. *This is the purpose of the various gravity reductions.* They were considered extensively in Chapter 3; we therefore can limit ourselves to pointing out those theoretical features that are relevant to our present problem.

If the external masses, the masses outside the geoid, are removed or moved inside the geoid, gravity will be changed. Furthermore, gravity is observed at ground level, but is needed at sea level. The reduction of gravity thus involves the consideration of these two effects, in order to obtain *boundary values* on the geoid.

This so-called *regularization* of the geoid by removing the external masses unfortunately also changes the level surfaces and hence, in general, the geoid. This is the *indirect effect;* the changed geoid is called the *cogeoid* or the regularized geoid.

The principle of this method may be described as follows (Jung, 1956, p. 578); see Fig. 8-1.

1. The masses outside the geoid are, by computation, either removed entirely or else moved inside the geoid. The effect of this procedure on the value of gravity g at the station P is considered.
2. The gravity station is moved from P down to the geoid, to the point P_0. Again, the corresponding effect on the gravity is considered.
3. The indirect effect, the distance $\delta N = P_0P^c$, is obtained by dividing the change in potential at the geoid, δW, by normal gravity (Bruns' theorem):

$$\delta N = \frac{\delta W}{\gamma}. \tag{8-1}$$

4. The gravity station is now moved from the geoidal point P_0 to the cogeoid, to the point P^c. This gives the boundary value of gravity at the cogeoid, g^c.
5. The shape of the cogeoid is computed from the reduced gravity anomalies

$$\Delta g^c = g^c - \gamma \tag{8-2}$$

by Stokes' formula, which gives $N^c = QP^c$.

6. Finally, the geoid is determined by considering the indirect effect. The geoidal undulation N is thus obtained as

$$N = N^c + \delta N. \tag{8-3}$$

At first sight it may seem that the masses between the geoid and the cogeoid should be removed if the cogeoid happens to be below the geoid, because Stokes' formula is applied to the cogeoid. However, this is not necessary, and therefore we need not be concerned with a "secondary indirect effect." The argument is a little too technical to be presented here; see Moritz (1965, p. 26).

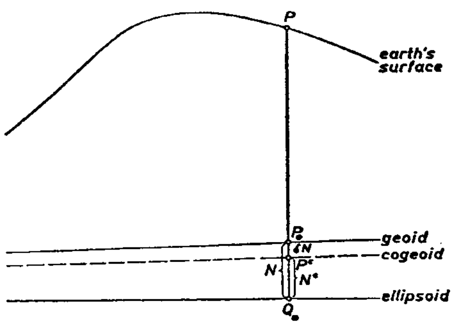

FIGURE 8-1

Geoid and cogeoid.

In principle every gravity reduction that gives boundary values at the geoid is equally suited for the determination of the geoid, provided the indirect effect is properly taken into account.[1] Thus the selection of a good reduction method should be made from other points of view, such as the geophysical meaning of the reduced gravity anomalies, the simplicity of computation, the feasibility of interpolation between the gravity stations, the smallness or even absence of the indirect effect, etc.; see Sec. 3-9.

The *Bouguer reduction* corresponds to a complete removal of the external masses. In the *isostatic reduction* these masses are shifted vertically downward according to some theory of isostasy. In Helmert's *condensation reduction* the external masses are compressed to form a surface layer on the geoid. In the *Rudzki reduction* they are moved inside the geoid in such a way that the potential on the geoid, and hence the geoid itself, remains unchanged (the external potential and the external level surfaces, however, are changed); hence there is no indirect effect in this case.

Quite different are the Prey reduction and the free-air reduction. The *Poincaré-Prey reduction* (Sec. 4-3) gives the actual gravity inside the earth; it does not give boundary values. The *free-air reduction*, in the present context, requires that the masses outside the geoid have been removed beforehand; it is here part of every gravity reduction to the geoid rather than an independent reduction. In Sec. 8-10 we shall deal with another aspect of this problem.

In all reduction methods it is necessary to know the density of the masses above the geoid. In practice this involves some kind of an assumption—for instance, putting $\rho = 2.67 \text{ g/cm}^3$. A second assumption is usually made in the free-air reduction, which is part of the reduction of gravity to the geoid: the actual free-air gravity gradient is assumed to be equal to the normal gradient

$$\frac{\partial \gamma}{\partial h} \doteq -0.3086 \text{ mgal/meter}.$$

These two assumptions falsify our results, at least theoretically (Moritz, 1962).

The second assumption can be avoided by using the actual free-air gradient as computed by the methods of Sec. 2-23. The anomalies Δg to be used in formula (2-217) must be the reduced gravity anomalies at the geoid: gravity g after steps 1 and 2 of the above description, minus theoretical gravity γ on the ellipsoid. This presupposes that in step 2 a preliminary free-air reduction using the normal gradient has been applied first.

Deflections of the vertical. The indirect effect affects the deflection of the vertical as well as the geoidal height. We have found

$$N = N^c + \delta N,$$

where N^c is the undulation of the cogeoid, the immediate result of Stokes'

[1] A formal proof, based on a transformation of a certain integral equation, may be found in Moritz (1965, Sec. 4).

formula, and δN is the indirect effect. By differentiating N in a horizontal direction we get the deflection component along this direction:

$$\epsilon = -\frac{\partial N}{\partial s} = -\frac{\partial N^c}{\partial s} - \frac{\partial(\delta N)}{\partial s}. \tag{8-4}$$

This means that to the immediate result of Vening Meinesz' formula, $-\partial N^c/\partial s$, we must add a term representing the horizontal derivative of δN; see also Sec. 3-6.

In the case of the Rudzki reduction, where the indirect effect is zero, Vening Meinesz' formula will give deflections of the vertical that refer directly to the geoid.

8-3. Molodensky's Problem

We have just seen that the reduction of gravity to sea level necessarily involves assumptions concerning the density of the masses above the geoid. This is equally true of other geodetic computations when performed in the conventional way.

To see this, consider the problem of computing the geodetic coordinates ϕ, λ, h, from the natural coordinates Φ, Λ, H, as described in Chapter 5. The geometric height h above the ellipsoid is obtained from the orthometric height H above the geoid and the geoidal undulation N by

$$h = H + N.$$

The determination of N was considered in the preceding section. To compute H from the results of leveling, we need the mean gravity $\bar{g}$ along the plumb line between the geoid and the ground (Sec. 4-4). Since gravity g cannot be measured inside the earth, we compute it by Prey's reduction, for which we must know the density of the masses above the geoid.

The geodetic coordinates ϕ and λ are obtained from the astronomical coordinates Φ and Λ and the deflection components ξ and η by

$$\phi = \Phi - \xi, \quad \lambda = \Lambda - \eta \sec \phi.$$

The coordinates Φ and Λ are measured on the ground; ξ and η can be computed for the geoid by Vening Meinesz' formula, the indirect effect being taken into account according to the preceding section. To apply the above formulas, either Φ and Λ must be reduced down to the geoid or ξ and η must be reduced up to the ground. In both cases this involves the reduction for the curvature of the plumb line (Sec. 5-6), which also depends on the mean value $\bar{g}$ through its horizontal derivatives. Hence Prey's reduction enters here too.

Thus we see that in the conventional approach to the problems of physical geodesy we must know the density of the outer masses or make assumptions concerning it. To avoid this, Molodensky in 1945 proposed a different approach.

Figure 8-2 shows the geometrical principles of this method. The ground point

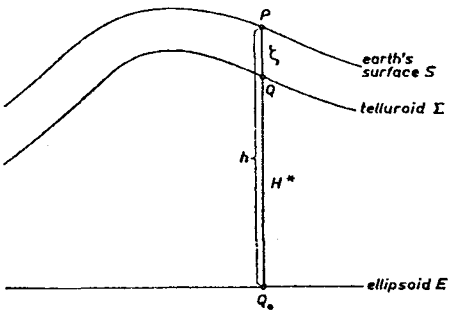

FIGURE 8-2
The telluroid. The normal height H^* and the height anomaly ζ.

P is again projected onto the ellipsoid according to Helmert. However, the geometric height h is now determined by

$$h = H^* + \zeta, \tag{8-5}$$

the *normal height* H^* replacing the orthometric height H, and the *height anomaly* ζ replacing the geoidal undulation N.

This will be clear if one considers the surface whose normal potential U at every point Q is equal to the actual potential W at the corresponding point P, so that $U_Q = W_P$, corresponding points P and Q being situated on the same ellipsoidal normal. This surface is called the *telluroid* (Hirvonen 1960, 1961). The vertical distance from the ellipsoid to the telluroid is the normal height H^* (Sec. 4-5), whereas the geometric height h is the vertical distance from the ellipsoid to the ground. The difference between these two heights is thus the height anomaly

$$\zeta = h - H^*, \tag{8-6}$$

closely corresponding to the geoidal undulation $N = h - H$, which is the difference between the geometric and the orthometric height.

The normal height H^*, and hence the telluroid Σ, can be determined by leveling combined with gravity measurements, according to Sec. 4-5. First the geopotential number of P, $C = W_0 - W_P$, is computed by

$$C = \int_0^P g\, dn,$$

where g is the measured gravity and dn is the leveling increment. The normal height H^* is then related to C by an *analytical* expression such as (4-44),

$$H^* = \frac{C}{\gamma_0}\left[1 + (1 + f + m - 2f\sin^2\phi)\frac{C}{a\gamma_0} + \left(\frac{C}{a\gamma_0}\right)^2\right],$$

where γ_0 is the normal gravity at the ellipsoidal point Q_0. Obviously H^* is independent of the density.

The normal height H^* of a ground point P is identical with the height above the ellipsoid, h, of the corresponding telluroid point Q. If the geopotential function W were equal to the normal potential function U at every point, then

Q would coincide with P, the telluroid would coincide with the physical surface of the earth, and the normal height of every point would be equal to its geometric height. Actually, however, $W_P \neq U_P$; hence the difference

$$\zeta_P = h_P - H_P^* = h_P - h_Q$$

is not zero. This explains the term "height anomaly" for ζ.

The gravity anomaly is now defined as

$$\Delta g = g_P - \gamma_Q; \tag{8-7}$$

it is the difference between the actual gravity as measured on the ground and the normal gravity on the telluroid. The normal gravity on the telluroid, which we shall briefly denote by γ, is computed from the normal gravity at the ellipsoid, γ_0, by the normal free-air reduction, but now applied *upward:*

$$\gamma \equiv \gamma_Q = \gamma_0 + \frac{\partial \gamma}{\partial h} H^* + \frac{1}{2!}\frac{\partial^2 \gamma}{\partial h^2} H^{*2} + \cdots. \tag{8-8}$$

For this reason the new gravity anomalies (8-7) are called *free-air anomalies.* They are *referred to ground level,* whereas the conventional gravity anomalies have been referred to sea level. Therefore the new free-air anomalies have nothing in common with a free-air reduction of actual gravity to sea level, except the name. This distinction should be carefully kept in mind.

A direct formula for computing γ at Q is (2-123),

$$\gamma = \gamma_0 \left[1 - 2(1 + f + m - 2f \sin^2 \phi)\frac{H^*}{a} + 3\left(\frac{H^*}{a}\right)^2\right], \tag{8-9}$$

where γ_0 is the corresponding value on the ellipsoid.

The height anomaly ζ may be considered as the distance between the geopotential surface $W = W_P =$ const. and the corresponding spheropotential surface $U = W_P =$ const. at the point P. In Sec. 2-16 we have denoted this distance by N_P, and have found that Bruns' formula (2-144) also applies to this quantity. Hence we have for $\zeta = N_P$

$$\zeta = \frac{T}{\gamma}, \tag{8-10}$$

$T = W_P - U_P$ being the disturbing potential at ground level, and γ being the normal gravity at the telluroid.

It may be expected that ζ is connected with the ground-level anomalies Δg by an expression analogous to Stokes' formula for the geoidal height N. This is indeed true. However, the telluroid is not a level surface, and to every point P on the earth's surface corresponds in general a different geopotential surface $W = W_P$. Therefore the relation between Δg and ζ in the new theory is considerably more complicated than for the geoid. The problem involves an integral equation, which may be solved by an iteration, the first term of which is given by Stokes' formula.

Finally we remark that we may also plot the height anomalies ζ above the ellipsoid. In this way we get a surface that is identical with the geoid over the oceans, because there $\zeta = N$, and is very close to the geoid anywhere else. This surface has been called the *quasigeoid* by Molodensky. However, the quasigeoid is not a level surface and has no physical meaning whatever. It must be considered as a concession to conventional conceptions that call for a geoidlike surface. From this point of view the normal height of a point is its elevation above the quasigeoid, just as the orthometric height is its elevation above the geoid.

8-4. Linear Integral Equations

In the following sections we shall make use of linear integral equations. For the reader who is not familiar with this subject we shall now give a brief intuitive introduction to linear integral equations, which is sufficient for our purpose. More details may be found in such standard treatises as Courant and Hilbert (1953).

The functions defined on the surface of the earth are functions of two variables (for instance, latitude and longitude). For simplicity, however, we shall be concerned here with functions of *one* variable only; this is sufficient for a general understanding.

Consider the equation

$$\int_a^b K(s, t)u(t)\, dt = f(s). \tag{8-11}$$

It is called a *linear integral equation of the first kind.* The functions $f(s)$ and $K(s, t)$ (called the *kernel* of the integral equation) are given; the problem is to determine the unknown function $u(t)$ from this equation.

The analogy of this integral equation with the system of linear equations

$$\sum_{j=1}^{n} K_{ij}u_j = f_i \qquad (i = 1, 2, \ldots, n), \tag{8-12}$$

which may be fully written as

$$\begin{aligned} K_{11}u_1 + K_{12}u_2 + \cdots + K_{1n}u_n &= f_1, \\ K_{21}u_1 + K_{22}u_2 + \cdots + K_{2n}u_n &= f_2, \\ \cdot \qquad\qquad\qquad\qquad & \quad \cdot \\ \cdot \qquad\qquad\qquad\qquad & \quad \cdot \\ \cdot \qquad\qquad\qquad\qquad & \quad \cdot \\ K_{n1}u_1 + K_{n2}u_2 + \cdots + K_{nn}u_n &= f_n, \end{aligned} \tag{8-12$'$}$$

is evident. There correspond:

integral $\int_{t=a}^{b}$ to sum $\sum_{j=1}^{n}$;

variables s, t to indices i, j.

This shows that a linear integral equation may be considered as an *analogue of a system of linear equations*.

It is also very easy to approximate the integral equation (8-11) by a system of linear equations. We may divide the interval of integration (a, b) into n equal parts and set

$$h = \frac{b-a}{n};$$

$$t_1 = a + \frac{h}{2}, \quad t_2 = a + \frac{3h}{2}, \quad t_3 = a + \frac{5h}{2}, \quad \cdots, \quad t_n = a + \frac{(2n-1)h}{2}.$$

Then Fig. 8-3 shows that the integral may be approximated by

$$\int_a^b K(s, t)u(t)\, dt \doteq K(s, t_1)u(t_1)\cdot h + K(s, t_2)u(t_2)\cdot h + \cdots + K(s, t_n)u(t_n)\cdot h.$$

This is nothing but the usual approximation of an area by a sum of rectangles; s is considered here as a fixed parameter. Thus the integral equation (8-11) becomes approximately

$$h[K(s, t_1)u(t_1) + K(s, t_2)u(t_2) + \cdots + K(s, t_n)u(t_n)] = f(s).$$

Letting s successively be equal to $t_1, t_2, \ldots, t_n$, we obtain

$$\begin{aligned} h[K(t_1, t_1)u(t_1) + K(t_1, t_2)u(t_2) + \cdots + K(t_1, t_n)u(t_n)] &= f(t_1), \\ h[K(t_2, t_1)u(t_1) + K(t_2, t_2)u(t_2) + \cdots + K(t_2, t_n)u(t_n)] &= f(t_2), \\ \cdot \qquad\qquad\qquad\qquad\qquad\qquad\qquad &\quad \cdot \\ \cdot \qquad\qquad\qquad\qquad\qquad\qquad\qquad &\quad \cdot \\ \cdot \qquad\qquad\qquad\qquad\qquad\qquad\qquad &\quad \cdot \\ h[K(t_n, t_1)u(t_1) + K(t_n, t_2)u(t_2) + \cdots + K(t_n, t_n)u(t_n)] &= f(t_n). \end{aligned} \tag{8-13}$$

On substituting

$$hK(t_i, t_j) = K_{ij}, \quad u(t_j) = u_j, \quad f(t_i) = f_i,$$

the system (8-13) becomes the system (8-12) or (8-12′).

As $n \to \infty$, the approximate system of linear equations (8-13) changes rigorously into the integral equation (8-11). We may thus regard a linear integral equation as the *limit* $(n \to \infty)$ *of a system of linear equations*.

The system (8-13) may also be used for an approximate solution of the integral equation (8-11): we may compute the values of $u(t)$ for $t = t_1, t_2, \ldots, t_n$ by

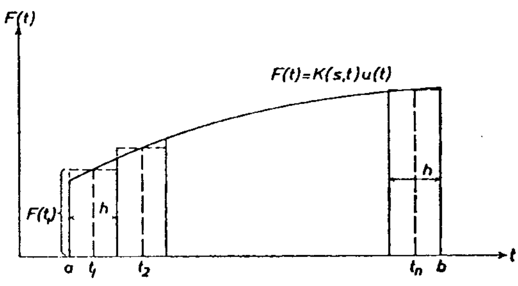

FIGURE 8-3

Approximation of an area by a sum of rectangles.

solving (8-13) and interpolate for other arguments t (in the same way as in a table of functions).

Considerably more important theoretically and in application are the *linear integral equations of the second kind.* They have the form

$$u(s) + \int_a^b K(s, t)u(t)\, dt = f(s). \tag{8-14}$$

As a matter of fact, such an integral equation is equivalent to a system of linear equations of the form

$$u_i + \sum_{j=1}^{n} K_{ij}u_j = f_i \tag{8-15}$$

and may be approximated by it, in the way shown above.

Since it is easy to think of a linear integral equation as a limit of a system of linear equations, we have chosen this approach here, although later we shall solve our integral equations differently, by an iterative process rather than by approximating them by systems of linear equations.

8-5. Application of Green's Identities

By applying Green's third identity to the gravity potential W we derived in Sec. 1-6 the formula (1-34),

$$-2\pi W + \iint_S \left[W \frac{\partial}{\partial n}\left(\frac{1}{l}\right) - \frac{1}{l}\frac{\partial W}{\partial n}\right] dS + 2\pi\omega^2(x^2 + y^2) + 2\omega^2 \iiint_{\text{earth}} \frac{dv}{l'} = 0. \tag{8-16}$$

Here S is the physical surface of the earth; l is the distance between a fixed point P, to which the first and third terms refer, and the variable surface element dS; n is the outward directed normal to the physical surface at dS; $\partial W/\partial n$ is the component of the gravity vector normal to S; the z-axis is the earth's axis of rotation; ω is the angular velocity; and l' is the distance between P and the volume element dv. Slight changes of notation are obvious.

This equation, which was also obtained by de Graaff-Hunter (1960), relates the earth's surface S to the potential W and its normal derivative $\partial W/\partial n$. It represents the most direct mathematical formulation of the problem of the gravimetric determination of the figure of the earth S, or in other words, of the boundary-value problem of physical geodesy according to Molodensky. It is therefore worthwhile to examine its meaning in some detail.

The gravity potential W at any point P is obtained, apart from an additive constant W_0, by leveling combined with gravity measurements according to

$$W = W_0 - \int_0^P g\, dn. \tag{8-17}$$

The normal component $\partial W/\partial n$ of the gravity vector **g** is determined by the measurement of g, which is the magnitude of **g**, and of the astronomical latitude and longitude, which establish the direction of **g**.

Hence the only unknown in (8-16) is the surface S itself, because l, l', x, y are determined by S and by the astronomical coordinates of the points involved. It is therefore plausible to assume that this equation can be solved for S in some way. Thus we see how it is possible to determine a purely geometrical quantity—namely, the shape of the surface S—solely from physical quantities connected with the earth's gravitational field (gravity potential and gravity vector).

So far we have considered the constant W_0, which can be taken as the potential at sea level, as known. As we have seen in Sec. 2-20, it is related to the *linear scale* of the earth; see also Molodenskii et al. (1962a, p. 113). If W_0 is only approximately known, as it is at present, then the shape of the earth is determined only up to a scale factor. The measurement of a single distance (preferably a long arc) is sufficient to establish the scale. No other measurements of distances or angles, neither triangulation nor trilateration, are necessary in principle.

Thus the geodetic measurements that are necessary and sufficient for the gravimetric determination of the physical surface of the earth may be summarized as:

1. gravity measurements;
2. astronomical determination of latitude and longitude;
3. leveling; and
4. measurement of *one* distance.

This, of course, is the theoretical minimum; in practice triangulation and trilateration are very useful because of their high relative accuracy.

Linearization. The basic equation (8-16) has the symbolic form

$$F\left(S,\, W,\, \frac{\partial W}{\partial n}\right) = 0;$$

the problem is to solve it for S. Unfortunately it is a nonlinear integral equation that cannot be solved directly. However, we can apply to it the usual treatment of any complicated nonlinear equation: we linearize it by introducing suitable approximate values, so that we finally get a linear equation for the deviation of the actual from the approximate solution. The actual potential W is thus approximated by a normal potential U; the approximate solution for S is the telluroid Σ. The deviation of W from U is the disturbing potential $T = W - U$, and the deviation of S from Σ is the height anomaly ζ.

We shall now proceed to the linearization of (8-16). Since W in this equation is quite an arbitrary function, we may apply (8-16) also to the normal potential U, getting

$$-2\pi U + \iint_S \left[U \frac{\partial}{\partial n}\left(\frac{1}{l}\right) - \frac{1}{l}\frac{\partial U}{\partial n} \right] dS + 2\pi\omega^2(x^2 + y^2) + 2\omega^2 \iiint_{\text{earth}} \frac{dv}{l'} = 0.$$

Subtracting this from the original equation (8-16) we obtain

$$-2\pi T + \iint_S \left[T \frac{\partial}{\partial n}\left(\frac{1}{l}\right) - \frac{1}{l}\frac{\partial T}{\partial n} \right] dS = 0.$$

This equation is already much simpler than (8-16). The essential point, however, is that in this equation we can replace the integration over the unknown surface S by an integration over the known telluroid Σ, getting

$$-2\pi T + \iint_\Sigma \left[T \frac{\partial}{\partial n}\left(\frac{1}{l}\right) - \frac{1}{l}\frac{\partial T}{\partial n} \right] d\Sigma = 0. \tag{8-18}$$

The reason why this is possible is that dS differs from $d\Sigma$ only by quantities of the order of the height anomaly ζ:

$$dS = d\Sigma(1 + A\zeta + B\zeta^2 + \cdots).$$

Thus we have

$$T\,dS = T\,d\Sigma + \frac{A}{\gamma} T^2\,d\Sigma + \cdots.$$

If we limit ourselves to terms linear in T or $\zeta = T/\gamma$, then the terms containing T^2 and higher powers of T will be neglected, and there remains

$$T\,dS = T\,d\Sigma$$

or, as long as it is multiplied by terms of the order of T,

$$dS = d\Sigma. \tag{8-19}$$

Note that this replacement of S by Σ is not permissible in the original equation (8-16), because W is of a greater order of magnitude than T.

In (8-18) the normal n is the normal to the physical surface of the earth or, to the same accuracy, the normal to the telluroid. It is thus in general not vertical. For this reason $\partial T/\partial n$ is not equal to

$$\frac{\partial T}{\partial h} = -\Delta g + \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} T \tag{8-20}$$

[this is equation (2-147c) applied to ground level], but it contains, besides Δg, the components ξ and η of the deflection of the vertical.

The actual evaluation of $\partial T/\partial n$ in terms of Δg, ξ, η is somewhat laborious (Molodenskii et al., 1962a, Chapter V; Moritz, 1965, p. 18). Therefore, and because we shall find a more convenient solution to Molodensky's problem in the next section, we omit the derivation here. The result is

$$\frac{\partial T}{\partial n} = \left[-\Delta g + \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} T + \gamma(\xi \tan \beta_1 + \eta \tan \beta_2) \right] \cos \beta, \tag{8-21}$$

where β_1 is the angle of inclination of a north-south terrain profile with respect to the horizontal; similarly, β_2 is the inclination of an east-west profile; β is the angle of maximum inclination of the terrain.

The insertion of (8-21) into (8-18) yields

$$T - \frac{1}{2\pi}\iint_{\Sigma}\left[\frac{\partial}{\partial n}\left(\frac{1}{l}\right) - \frac{1}{\gamma}\frac{\partial \gamma}{\partial h}\frac{\cos\beta}{l}\right]T\,d\Sigma$$
$$= \frac{1}{2\pi}\iint_{\Sigma}\frac{1}{l}\left[\Delta g - \gamma(\xi\tan\beta_1 + \eta\tan\beta_2)\right]\cos\beta\,d\Sigma. \quad (8\text{-}22)$$

This is a linear integral equation of the second kind for the disturbing potential T or for the height anomaly $\zeta = T/\gamma$. If we compare it with (8-14), we see that the unknown function u is now represented by T, the known function f is given by the right-hand side of (8-22), and the kernel K is $-1/2\pi$ times the expression in brackets in the integral on the left-hand side of (8-22). This integral equation is also considered by Levallois (1958).

If we want to solve this equation for T we must know, besides Δg, the deflection components ξ and η. Since the inclinations β_1 and β_2 are usually small, and since ξ and η are multiplied by $\tan\beta_1$ and $\tan\beta_2$, approximate values for the deflection components are usually sufficient. Molodensky has even succeeded in eliminating ξ and η from (8-22) in a highly ingenious way.

As we have indicated, we shall deal with a considerably simpler method in the following section. We shall therefore pursue the present approach no further but mention only that the integral equation (8-22) can be solved by an iteration analogous to that to be described in Sec. 8-7.

Application to the geoid. The integral equation (8-22) may also be applied to the geoid, provided the geoid has been "regularized" by removing the masses outside it. Instead of the telluroid Σ we then have the reference ellipsoid E; furthermore, $\beta_1 = \beta_2 = \beta = 0$, and $\partial/\partial n = \partial/\partial h$. Hence we obtain

$$T - \frac{1}{2\pi}\iint_{E}\left[\frac{\partial}{\partial h}\left(\frac{1}{l}\right) - \frac{1}{\gamma}\frac{\partial \gamma}{\partial h}\frac{1}{l}\right]T\,dE = \frac{1}{2\pi}\iint_{E}\frac{\Delta g}{l}\,dE. \quad (8\text{-}23)$$

This equation is much simpler than (8-22), because it does not contain the deflection components ξ and η.

If the reference ellipsoid is approximated by a sphere, or in other words, if we make a spherical approximation, the solution of (8-23) is given simply by Stokes' formula. This is immediately obvious because Stokes' formula expresses T in terms of Δg as a spherical approximation.

By expanding the ellipsoidal quantities of (8-23) in terms of e'^2 or a similar parameter of the order of the flattening, this integral equation may be solved iteratively, using Stokes' formula as a first approximation. In this way it is possible to find a relatively simple solution of "Zagrebin's problem," the determination of the regularized geoid by means of a reference ellipsoid to the order of e'^2 (Molodenskii et al., 1962a, p. 53).

The integral equation method thus makes possible the numerical solution of boundary-value problems of physical geodesy—problems whose solution by some other method may be much more complicated or even impossible. Besides this advantage in *solving* problems we have also an advantage in *formulating* problems. By the integral equations (8-22) and (8-23) the respective problems are completely described. The corresponding conventional formulation would be to determine a function T that outside a certain surface (the earth's surface or the regularized geoid) satisfies Laplace's differential equation

$$\Delta T = 0 \tag{8-24a}$$

and is subject to the boundary condition (8-20),

$$\frac{\partial T}{\partial h} - \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} T = -\Delta g \tag{8-24b}$$

on this surface. It is evident that the formulation by means of one integral equation (instead of a differential equation and a boundary condition) is shorter; it also brings the problem closer to a solution in many cases.

Even if T is not harmonic, in the case of the actual geoid, a direct integral equation approach is still possible, whereas the conventional approach can no longer be used directly. The application of an integral equation, corresponding to (8-23), to the actual geoid provides a direct mathematical treatment of the problem of gravity reductions for the determination of the geoid (Moritz, 1965, Sec. 4). Hence the integral equation method is also an efficient tool for problems of classical geodesy.

8-6. Integral Equation for the Surface Layer

The integral equation (8-22) has the disadvantage that it contains, besides the gravity anomaly Δg, the deflection components ξ and η. As we have mentioned, it is possible to transform (8-22) so that it contains Δg only, but then it becomes rather complicated.

A simpler and more practical integral equation may be obtained in the following way. We may express the anomalous potential T as the potential of a surface layer (Sec. 1-3) on the earth's surface or, to the same accuracy, on the telluroid:

$$T = \iint_{\Sigma} \frac{\phi}{l}\, d\Sigma. \tag{8-25}$$

The symbol ϕ is the surface density κ multiplied by the gravitational constant.

We insert this expression into the boundary condition (8-20),

$$-\frac{\partial T}{\partial h} + \frac{1}{\gamma}\frac{\partial \gamma}{\partial h} T = \Delta g.$$

If we wish to differentiate equation (8-25) with respect to h, we must remember from Sec. 1-3 that the derivatives of a surface layer potential are discontinuous at the surface. For the boundary condition we obviously need the *outer* derivative, which is given by equation (1-19a):

$$\left(\frac{\partial T}{\partial h_P}\right)_e = -2\pi\phi\cos\beta + \iint_\Sigma \phi \frac{\partial}{\partial h_P}\left(\frac{1}{l}\right) d\Sigma, \tag{8-26}$$

where the direction of m is now the vertical of the point P to which both T in (8-25) and the boundary condition (8-20) refer; we have therefore written $\partial/\partial m = \partial/\partial h_P$. The angle (m, n) is now the angle between this vertical and the surface normal, which is the inclination angle β.

On inserting this expression into the boundary condition we obtain

$$2\pi\phi\cos\beta - \iint_\Sigma \left[\frac{\partial}{\partial h_P}\left(\frac{1}{l}\right) - \frac{1}{\gamma_P}\frac{\partial\gamma}{\partial h_P}\frac{1}{l}\right]\phi\, d\Sigma = \Delta g. \tag{8-27}$$

The quantities outside the integral are always taken at the point P. If quantities inside the integral are to refer to this point, they are specially marked by the subscript P, otherwise they are taken at the surface element $d\Sigma$.

It is instructive to compare this equation with (8-22). Both are linear integral equations of the second kind. The coefficient of T inside the integral in (8-22) is apparently quite similar to the corresponding coefficient of ϕ in (8-27). However, γ and the partial derivatives $\partial/\partial n$ and $\partial/\partial h$ in (8-22) are taken at $d\Sigma$, whereas γ and $\partial/\partial h$ in (8-27) refer to P.

The advantage of the new integral equation (8-27) is that it depends only on Δg.

Spherical approximation. We now write the integral equation (8-27) as a spherical approximation. Note that this means that the reference ellipsoid, but not the telluroid, is approximated by a sphere.

Then the geocentric radius vectors of P and of $d\Sigma$ are approximated by (see Fig. 8-4)

$$\begin{aligned} r_P &= R + h_P, \\ r &= R + h, \end{aligned} \tag{8-28}$$

where R is a mean radius of the earth and h is the height above the ellipsoid or, to the same approximation, the orthometric or also the normal height.

We have

$$\begin{aligned} l &= \sqrt{r_P^2 + r^2 - 2r_P r\cos\psi}, \\ \frac{\partial}{\partial h_P}\left(\frac{1}{l}\right) &= \frac{\partial}{\partial r_P}\left(\frac{1}{l}\right) = -\frac{r_P - r\cos\psi}{l^3}, \\ -\frac{1}{\gamma_P}\frac{\partial\gamma}{\partial h_P} &= \frac{2}{r_P}, \end{aligned}$$

so that we find, after a simple calculation,

$$\frac{\partial}{\partial h_P}\left(\frac{1}{l}\right) - \frac{1}{\gamma_P}\frac{\partial\gamma}{\partial h_P}\frac{1}{l} = \frac{3}{2r_P l} + \frac{r^2 - r_P^2}{2r_P l^3}. \tag{8-29}$$

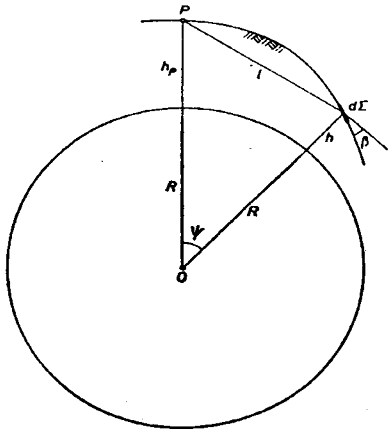

FIGURE 8-4
Spherical approximation.

Thus (8-27) becomes

$$2\pi\phi\cos\beta - \iint_{\Sigma}\left(\frac{3}{2r_P l} + \frac{r^2 - r_P^2}{2r_P l^3}\right)\phi\, d\Sigma = \Delta g.$$

The surface element $d\Sigma$ may be eliminated by noting that the projection of $d\Sigma$ onto the local horizon is given by

$$d\Sigma\cos\beta.$$

This is also equal to

$$r^2\, d\sigma,$$

where $d\sigma$ is the element of solid angle, because r is the radius vector of $d\Sigma$. Hence we have

$$d\Sigma = r^2\sec\beta\, d\sigma.$$

Thus our integral equation finally becomes

$$2\pi\phi\cos\beta - \iint_{\sigma}\left(\frac{3}{2l} + \frac{r^2 - r_P^2}{2l^3}\right)\frac{r^2}{r_P}\sec\beta\cdot\phi\, d\sigma = \Delta g. \tag{8-30}$$

This equation will be simplified and solved in the next section.

If ϕ is known, then T and ζ are determined by (8-25), which may be written

$$T = \gamma\zeta = \iint_{\sigma}\frac{\phi}{l}\, r^2\sec\beta\, d\sigma. \tag{8-31}$$

Application to the geoid. The integral equation (8-30) may also be applied to the regularized geoid. Then we have

$$h = h_P = \beta = 0, \quad r = r_P = R,$$

and (8-30) becomes

$$2\pi\phi - \frac{3R}{2}\iint_\sigma \frac{\phi}{l_0}\,d\sigma = \Delta g, \tag{8-32}$$

where

$$l_0 = 2R \sin\frac{\psi}{2}; \tag{8-33}$$

see Fig. 1-13.

T and N are expressed in terms of ϕ by (8-31), which now becomes

$$T = GN = R^2 \iint_\sigma \frac{\phi}{l_0}\,d\sigma, \tag{8-34}$$

where G is an average value for the gravity.

On inserting (8-34) into (8-32) we find

$$\phi = \frac{1}{2\pi}\left(\Delta g + \frac{3}{2R}T\right) = \frac{1}{2\pi}\left(\Delta g + \frac{3G}{2R}N\right). \tag{8-35}$$

This expression of ϕ in terms of Δg and N is equivalent to (6-57), since $\mu = 2\pi\phi$. The geoidal height N is given as a spherical approximation by Stokes' formula

$$N = \frac{R}{4\pi G}\iint_\sigma \Delta g\, S(\psi)\,d\sigma. \tag{8-36}$$

This is inserted into (8-35), which gives

$$2\pi\phi = \Delta g + \frac{3}{8\pi}\iint_\sigma \Delta g\, S(\psi)\,d\sigma. \tag{8-37}$$

This formula expresses ϕ in terms of Δg and is thus a solution of the integral equation (8-32).

Solving (8-35) we find

$$T = \frac{2R}{3}(2\pi\phi - \Delta g). \tag{8-38}$$

These simple formulas are valid for the regularized geoid to a spherical approximation.

8-7. Solution of the Integral Equation

Before solving the integral equation (8-30) we simplify it by noting that

$$r = R + h = R\left(1 + \frac{h}{R}\right)$$

differs from R by less than 10^{-3}, which is smaller than the error of the spherical approximation. Thus we may safely put

$$\frac{r^2}{r_P} = R,$$

$$r^2 - r_P^2 = (h - h_P)(r + r_P) = 2R(h - h_P),$$

obtaining

$$2\pi\phi \cos\beta - \iint_\sigma \left(\frac{3R}{2l} + \frac{R^2(h - h_P)}{l^3}\right) \sec\beta \cdot \phi \, d\sigma = \Delta g. \tag{8-39}$$

This equation is simpler than (8-30), but quite as accurate.

We can also simplify the expression for the distance l. We find

$$\begin{aligned} l^2 &= r_P^2 + r^2 - 2r_P r \cos\psi \\ &= (R + h_P)^2 + (R + h)^2 - 2(R + h_P)(R + h)\cos\psi \\ &= 2R^2(1 - \cos\psi) + 2R(h + h_P)(1 - \cos\psi) + h_P^2 + h^2 - 2h_P h \cos\psi \\ &= 4R^2 \sin^2\frac{\psi}{2}\left(1 + \frac{h + h_P}{R} + \frac{h_P h}{R^2}\right) + (h - h_P)^2. \end{aligned}$$

For the same reason as above we may neglect $(h + h_P)/R$ and $h_P h/R^2$, obtaining

$$l^2 = l_0^2 + (h - h_P)^2,$$

$$l = l_0\sqrt{1 + \left(\frac{h - h_P}{l_0}\right)^2}. \tag{8-40}$$

Here l_0 denotes the spherical distance (8-33).

With these preliminary steps completed we are now able to solve the integral equation (8-39). The basic principle is to use an expansion in powers of the quantities

$$\frac{h - h_P}{l_0} \quad \text{and} \quad \tan\beta, \tag{8-41}$$

These quantities are of the same order of magnitude because, as $l_0 \longrightarrow 0$, then $(h - h_P)/l_0$ obviously approaches $\tan\beta'$, where β' is the angle of inclination in the direction of l_0.

Note that the quantities (8-41) are of a greater order of magnitude than h/R. Hence they must be taken into account, although h/R was neglected in (8-39). To give a numerical example, we assume a moderate mountain slope of inclination $\beta = 15°$ at an elevation h of 1000 meters. Then

$$\frac{h}{R} = 0.00016, \quad \text{but} \quad \tan\beta = 0.27.$$

Solution. The solution of (8-39) is obtained by successive approximations.

a) As a first step, the quantities (8-41) are neglected. Then (8-39) becomes

$$2\pi\phi_0 - \frac{3R}{2}\iint_\sigma \frac{\phi_0}{l_0} \, d\sigma = G_0, \tag{8-42}$$

where we have put

$$G_0 = \Delta g \tag{8-43}$$

and denoted the "zero-order approximation" of ϕ by ϕ_0.

Since (8-42) is of the form (8-32), its solution is given by (8-37), which in the present notation reads

$$2\pi\phi_0 = G_0 + \frac{3}{8\pi}\iint_\sigma G_0\, S(\psi)\, d\sigma. \tag{8-44}$$

b) After this, the quantities (8-41) are taken into account, but only to the first power; second and higher powers are neglected. Then ϕ will receive a small correction ϕ_1, so that as a "first-order approximation"

$$\phi \doteq \phi_0 + \phi_1. \tag{8-45}$$

To this approximation we still have

$$l \doteq l_0,$$
$$\cos\beta \doteq \sec\beta \doteq 1,$$

because the quadratic terms in the series

$$l = l_0\sqrt{1 + \left(\frac{h - h_P}{l_0}\right)^2} = l_0\left[1 + \frac{1}{2}\left(\frac{h - h_P}{l_0}\right)^2 + \cdots\right],$$

$$\cos\beta = \frac{1}{\sqrt{1 + \tan^2\beta}} = 1 - \frac{1}{2}\tan^2\beta + \cdots$$

are neglected.

Hence (8-39) becomes

$$2\pi(\phi_0 + \phi_1) - \frac{3R}{2}\iint_\sigma \frac{\phi_0 + \phi_1}{l_0}\, d\sigma - R^2\iint_\sigma \frac{h - h_P}{l_0^3}(\phi_0 + \phi_1)\, d\sigma = \Delta g.$$

Since both $(h - h_P)/l_0$ and ϕ_1 are first-order quantities, their product is to be neglected in the second integral, and we obtain

$$2\pi\phi_0 - \frac{3R}{2}\iint_\sigma \frac{\phi_0}{l_0}\, d\sigma + 2\pi\phi_1 - \frac{3R}{2}\iint_\sigma \frac{\phi_1}{l_0}\, d\sigma - R^2\iint_\sigma \frac{h - h_P}{l_0^3}\phi_0\, d\sigma = G_0.$$

The first two terms on the left-hand side are equal to the right-hand side, according to (8-42). There remains

$$2\pi\phi_1 - \frac{3R}{2}\iint_\sigma \frac{\phi_1}{l_0}\, d\sigma - R^2\iint_\sigma \frac{h - h_P}{l_0^3}\phi_0\, d\sigma = 0$$

or

$$2\pi\phi_1 - \frac{3R}{2}\iint_\sigma \frac{\phi_1}{l_0}\, d\sigma = G_1, \tag{8-46}$$

where

$$G_1 = R^2\iint_\sigma \frac{h - h_P}{l_0^3}\phi_0\, d\sigma. \tag{8-47}$$

Equation (8-46) is the same as (8-42), with ϕ_1 and G_1 instead of ϕ_0 and G_0. Its solution is therefore given by (8-44),

$$2\pi\phi_1 = G_1 + \frac{3}{8\pi}\iint_\sigma G_1\, S(\psi)\, d\sigma. \tag{8-48}$$

c) As a next step we could take the squares of the quantities (8-41) into account, neglecting third and higher powers. The procedure is essentially the same as in (b). In this way we may proceed to higher and higher approximations.

Molodensky (Molodenskii et al., 1962a, p. 118) has devised an elegant method for this purpose and has also computed the second- and third-order approximations. However, practical tests have indicated that the first-order approximation is accurate enough in most cases. Accordingly, we shall here limit ourselves to this approximation.

To obtain T and ζ from ϕ, we use (8-31), in which we again set $r \doteq R$:

$$T = R^2 \iint_\sigma \frac{\phi}{l}\sec\beta\, d\sigma = R^2 \iint_\sigma \frac{\phi_0}{l_0}\, d\sigma + R^2 \iint_\sigma \frac{\phi_1}{l_0}\, d\sigma + \cdots = T_0 + T_1 + \cdots.$$

Since both ϕ_0 and ϕ_1 satisfy equations of the form (8-32) and are related to T_0 and T_1 by equations of the form (8-34), we may apply (8-38), obtaining

$$T_0 = \frac{2R}{3}(2\pi\phi_0 - G_0),$$

$$T_1 = \frac{2R}{3}(2\pi\phi_1 - G_1).$$

On inserting (8-44) and (8-48) we find

$$\begin{aligned} T_0 &= \frac{R}{4\pi}\iint_\sigma G_0\, S(\psi)\, d\sigma, \\ T_1 &= \frac{R}{4\pi}\iint_\sigma G_1\, S(\psi)\, d\sigma. \end{aligned} \tag{8-49}$$

Bruns' formula, $\zeta = T/\gamma$, thus finally gives

$$\zeta = \zeta_0 + \zeta_1 = \frac{R}{4\pi\gamma}\iint_\sigma \Delta g\, S(\psi)\, d\sigma + \frac{R}{4\pi\gamma}\iint_\sigma G_1\, S(\psi)\, d\sigma, \tag{8-50}$$

where, according to (8-47) and (8-35)

$$G_1 = \frac{R^2}{2\pi}\iint_\sigma \frac{h - h_P}{l_0^3}\left(\Delta g + \frac{3G}{2R}\zeta_0\right) d\sigma. \tag{8-51}$$

Thus ζ is again approximately given by Stokes' formula; this is the term ζ_0. In addition we have a small correction ζ_1. The steps of computation are as follows: first, compute ζ_0 by Stokes' formula; then, evaluate G_1 by means of (8-51); and finally, use G_1 to calculate the correction term ζ_1 in (8-50).

In the following section we shall see that the term containing ζ_0 in (8-51) may even be neglected without impairing the accuracy.

The integral formula (8-51) can be evaluated by the usual methods, as outlined in Sec. 2-24; see also Burša (1965).

The method of using the potential of a fictitious surface layer to obtain a suitable integral equation, described in the preceding section, may be generalized to construct other integral equations for Molodensky's problem. They can be solved by the iterative method used in the present section (Brovar, 1964).

8-8. Geometrical Interpretation

We shall now describe a geometrical interpretation of Molodensky's approximate solution (8-50),

$$\zeta = \frac{R}{4\pi\gamma} \iint_\sigma (\Delta g + G_1)\, S(\psi)\, d\sigma. \tag{8-52}$$

Using the notation of Sec. 6-5, we put

$$\mu = \Delta g + \frac{3G}{2R}\zeta_0, \tag{8-53}$$

so that (8-51) takes the form

$$G_1 = \frac{R^2}{2\pi} \iint_\sigma \frac{h - h_P}{l_0^3}\, \mu\, d\sigma. \tag{8-54}$$

We shall now apply a transformation whose principle was given by Molodensky et al. (1962b). We write

$$\begin{aligned}(h - h_P)\mu &= (h - h_P)\mu + h_P\mu_P - h_P\mu_P \\ &= -h_P(\mu - \mu_P) + (h\mu - h_P\mu_P).\end{aligned}$$

Then (8-54) becomes

$$G_1 = -h\frac{R^2}{2\pi} \iint_\sigma \frac{\mu - \mu_P}{l_0^3}\, d\sigma + \frac{R^2}{2\pi} \iint_\sigma \frac{h\mu - (h\mu)_P}{l_0^3}\, d\sigma. \tag{8-55}$$

Note that if h_P is taken out of the integral, it may be denoted simply by h because, apart from the quantities within the integral sign, everything refers to the point P.

By means of equations (1-101) and (1-102) we may express (8-55) in terms of spherical harmonics. Let the spherical-harmonic expansions of the functions μ and $h\mu$ be

$$\mu = \sum_{n=0}^{\infty} \mu_n, \quad h\mu = \sum_{n=0}^{\infty} (h\mu)_n. \tag{8-56}$$

Then (8-55) becomes

$$G_1 = \frac{h}{R} \sum_0^\infty n\mu_n - \frac{1}{R} \sum_0^\infty n(h\mu)_n.$$

Subtracting and adding $1/R$ times

$$h \sum_0^\infty \mu_n = h\mu = \sum_0^\infty (h\mu)_n$$

we obtain

$$G_1 = \frac{h}{R} \sum_0^\infty (n-1)\mu_n - \frac{1}{R} \sum_0^\infty (n-1)(h\mu)_n. \tag{8-57}$$

We may thus split G_1 into two parts:

$$G_1 = G_{11} + G_{12}, \tag{8-58}$$

where

$$G_{11} = \frac{h}{R} \sum_0^\infty (n-1)\mu_n = -h \frac{R^2}{2\pi} \iint_\sigma \frac{\mu - \mu_P}{l_0^3} \, d\sigma - \frac{h}{R} \mu, \tag{8-59a}$$

$$G_{12} = -\frac{1}{R} \sum_0^\infty (n-1)(h\mu)_n = \frac{R^2}{2\pi} \iint_\sigma \frac{h\mu - (h\mu)_P}{l_0^3} \, d\sigma + \frac{h}{R} \mu. \tag{8-59b}$$

Let us first consider the term G_{11}. Writing $\Delta g = \sum \Delta g_n$ and $T_0 = \sum T_n$ (note that T_0 here means the zero-order approximation of the function T and not the harmonic of degree zero!), we have

$$\mu_n = \Delta g_n + \frac{3}{2R} T_n.$$

Hence (8-59a) becomes

$$G_{11} = \frac{h}{R} \sum_0^\infty (n-1)\, \Delta g_n + \frac{3h}{2R^2} \sum_0^\infty (n-1) T_n$$

$$= \frac{h}{R} \sum_0^\infty (n+2)\, \Delta g_n - 3\frac{h}{R} \Delta g + \frac{3h}{2R^2} \sum_0^\infty (n-1) T_n.$$

According to equations (2-216) and (2-155′) we have

$$\frac{1}{R} \sum_0^\infty (n+2)\, \Delta g_n = -\frac{\partial \Delta g}{\partial h}, \quad \frac{1}{R} \sum_0^\infty (n-1) T_n = \Delta g,$$

so that

$$G_{11} = -h \frac{\partial \Delta g}{\partial h} - \frac{3h}{2R} \Delta g. \tag{8-60}$$

Since G_{11} will be added to Δg, according to (8-52) and (8-58), the quantity $(h/R)\,\Delta g$, which is at most of the order of $10^{-3}\,\Delta g$, may be safely neglected as compared to Δg, and there remains

$$G_{11} = -h \frac{\partial \Delta g}{\partial h}. \tag{8-60′}$$

Thus we see that the term G_{11} *corresponds to the reduction of the free-air gravity anomaly from the ground to sea level,* by the topographic elevation h. Neglecting again a relative error of h/R we have by (2-217)

$$G_{11} = -h\frac{R^2}{2\pi}\iint_\sigma \frac{\Delta g - \Delta g_P}{l_0^3}\, d\sigma. \tag{8-61}$$

Before we consider G_{12} we note that the correction term ζ_1, which represents the effect of G_1, may be split in the same way as G_1.

$$\zeta_1 = \zeta_{11} + \zeta_{12}. \tag{8-62}$$

Then

$$\zeta_{11} = \frac{R}{4\pi\gamma}\iint_\sigma G_{11}\, S(\psi)\, d\sigma = -\frac{R}{4\pi\gamma}\iint_\sigma h\frac{\partial \Delta g}{\partial h} S(\psi)\, d\sigma. \tag{8-63a}$$

The second component

$$\zeta_{12} = \frac{R}{4\pi\gamma}\iint_\sigma G_{12}\, S(\psi)\, d\sigma \tag{8-63b}$$

can be directly evaluated. Remember that the equivalent of Stokes' formula

$$\zeta = \frac{R}{4\pi\gamma}\iint_\sigma \Delta g\, S(\psi)\, d\sigma$$

in terms of spherical harmonics is

$$\zeta_n = \frac{R}{(n-1)\gamma}\Delta g_n.$$

If we replace ζ by ζ_{12}, Δg by G_{12}, and Δg_n by $-(n-1)(h\mu)_n/R$, in agreement with (8-59b), then the conversion of (8-63b) to an expression in spherical harmonics becomes

$$(\zeta_{12})_n = \frac{R}{(n-1)\gamma}\left(-\frac{1}{R}\right)(n-1)(h\mu)_n = -\frac{1}{\gamma}(h\mu)_n.$$

The summation from $n = 0$ to ∞ yields the simple formula

$$\zeta_{12} = -\frac{h\mu}{\gamma}. \tag{8-64}$$

On inserting (8-53) with $G \doteq \gamma$ this becomes

$$\zeta_{12} = -\frac{h\Delta g}{\gamma} - \frac{3h}{2R}\zeta_0. \tag{8-65}$$

Since ζ_{12} is added to ζ_0, we again introduce a relative error of the order of only h/R if we neglect the second term on the right-hand side of this equation. Thus we finally obtain

$$\zeta_{12} = -\frac{\Delta g}{\gamma}h. \tag{8-65'}$$

This term is as simple as (8-60′) and admits of a corresponding geometrical interpretation. Consider the vertical derivative of the height anomaly ζ. We find

$$\frac{\partial \zeta}{\partial h} = \frac{\partial}{\partial h}\left(\frac{T}{\gamma}\right) = \frac{1}{\gamma}\frac{\partial T}{\partial h} - \frac{1}{\gamma^2}\frac{\partial \gamma}{\partial h}T = -\frac{1}{\gamma}\left(-\frac{\partial T}{\partial h} + \frac{1}{\gamma}\frac{\partial \gamma}{\partial h}T\right);$$

according to equation (2-147c) this is equal to

$$\frac{\partial \zeta}{\partial h} = -\frac{\Delta g}{\gamma}. \tag{8-66}$$

Hence (8-65′) is equivalent to

$$\zeta_{12} = \frac{\partial \zeta}{\partial h} h. \tag{8-67}$$

Thus we see that the term ζ_{12} *corresponds to the reduction of the height anomaly from sea level to the ground*, the sign of this reduction being opposite to that of (8-60′).

Using (8-63a) and (8-67) we may write the solution (8-52) in the alternative form

$$\zeta = \frac{R}{4\pi\gamma}\iint_\sigma \left(\Delta g - \frac{\partial \Delta g}{\partial h} h\right) S(\psi)\, d\sigma + \frac{\partial \zeta}{\partial h} h. \tag{8-68}$$

The geometrical interpretation of this equation is evident from what has been said: the free-air anomalies Δg at ground level are reduced to sea level to become

$$\Delta g^* = \Delta g - \frac{\partial \Delta g}{\partial h} h; \tag{8-69}$$

then Stokes' integral gives height anomalies at sea level, which are reduced upward to ground level by adding the term (8-67).

As a by-product we have obtained the important result that μ (8-53) can be replaced by Δg without introducing a relative error of an order greater than h/R; this error has been neglected already in equation (8-39). This is seen by comparing (8-59a) with (8-61), and (8-64) with (8-65′). Thus equation (8-51) for G_1 may be essentially simplified to become

$$G_1 = \frac{R^2}{2\pi}\iint_\sigma \frac{h - h_P}{l_0^3} \Delta g\, d\sigma. \tag{8-70}$$

Equation (8-52), together with (8-70), and equation (8-68), together with (8-61) and (8-66), are two equally simple and practical formulas for the height anomaly ζ. They are both valid to a *linear approximation*, an approximation linear in the quantities (8-41).

Hence we see why gravity anomalies Δg *at ground level* may be used in (8-61), whereas the equivalent expression (2-217) was originally derived for gravity anomalies *at sea level*, which in the present case are given by (8-69). Since Δg^* and Δg differ only by terms of the order of h, the difference between using Δg or Δg^* in (8-61) is of the order of h^2, which is negligible in the linear approximation.

Reduction to point level. The elevation h in (8-68) is taken above sea level (Fig. 8-5a). If we examine the arguments leading to this equation, we shall find that the sea level is not distinguished from any other level. If we reckon the elevation above some other reference level, which has the elevation h_0

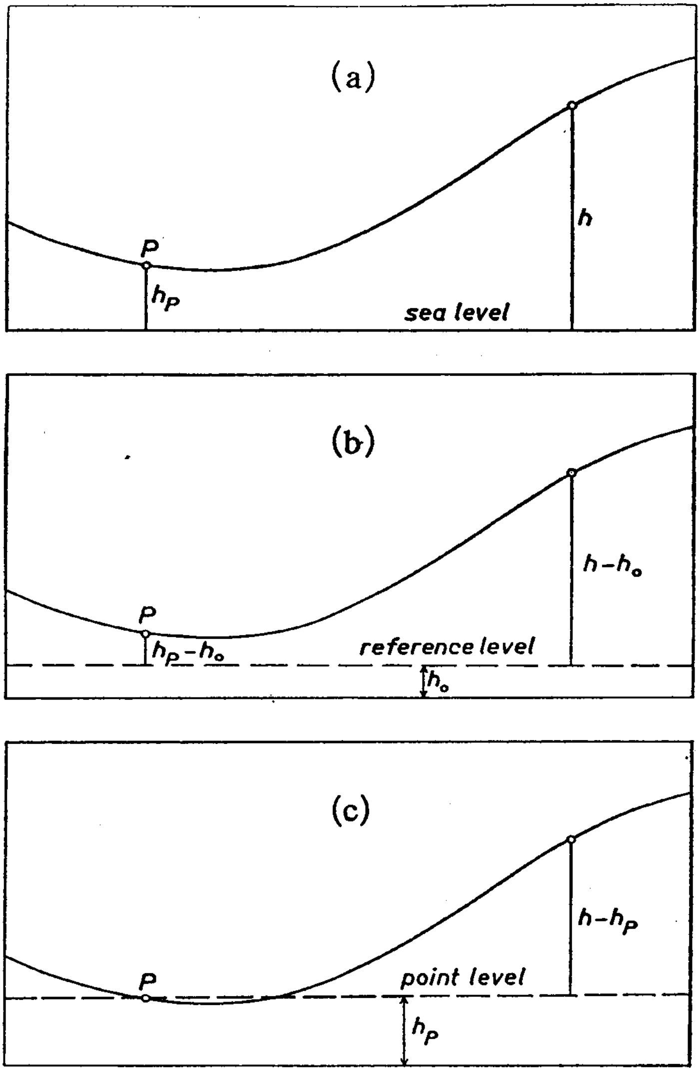

FIGURE 8-5
Reduction (a) to sea level, (b) to an arbitrary level, and (c) to the level of point P

above sea level, we must replace h by $h - h_0$ (Fig. 8-5b). Thus (8-68) is equivalent to

$$\zeta = \frac{R}{4\pi\gamma} \iint_\sigma \left[\Delta g - \frac{\partial \Delta g}{\partial h}(h - h_0)\right] S(\psi)\, d\sigma + \frac{\partial \zeta}{\partial h}(h - h_0).$$

In particular we may take as reference level the level of the point P itself, so that

$$h_0 = h_P,$$

P being the point at which the height anomaly ζ is computed. If this choice is made, the last term in the above expression will be zero, because outside the integral h always means h_P, so that $h - h_0 = h_P - h_P = 0$. Thus we have

$$\zeta = \frac{R}{4\pi\gamma} \iint_\sigma \left[\Delta g - \frac{\partial \Delta g}{\partial h}(h - h_P)\right] S(\psi)\, d\sigma. \tag{8-71}$$

This formula is particularly simple; geometrically it means that the free-air anomalies are reduced from the ground to the level of the computation point P (Fig. 8-5c). The reference level is thus different for different computation points.

8-9. Deflections of the Vertical

In Sec. 2-22 we have found that the components of the deflection of the vertical at the geoid are given by

$$\begin{aligned} \xi &= -\frac{\partial N}{\partial s_\phi} = -\frac{1}{R}\frac{\partial N}{\partial \phi}, \\ \eta &= -\frac{\partial N}{\partial s_\lambda} = -\frac{1}{R\cos\phi}\frac{\partial N}{\partial \lambda}. \end{aligned} \tag{8-72}$$

In the same way, the deflection components at the earth's surface are expressed by

$$\xi = -\frac{1}{R}\frac{\partial \zeta}{\partial \phi}, \quad \eta = -\frac{1}{R\cos\phi}\frac{\partial \zeta}{\partial \lambda}. \tag{8-73}$$

However, we must be careful in forming the partial derivatives of ζ. In (8-73) the derivatives are taken *along the level surface* $W = W_P$ containing the point P, just as the derivatives in (8-72) are taken along the geoid $W = W_0$. On the other hand, by solving Molodensky's problem we obtain the height anomalies *along the physical surface of the earth* S.

Let S be given by an expression of the form

$$W = W(\phi, \lambda);$$

that is, the potential at the physical surface of the earth is expressed as a function of the geodetic coordinates ϕ and λ. The height anomalies along S may be symbolized by

$$\zeta = \zeta(\phi, \lambda, W) = \zeta[\phi, \lambda, W(\phi, \lambda)].$$

Differentiating this with respect to ϕ along S yields

$$\frac{\partial\zeta[\phi, \lambda, W(\phi, \lambda)]}{\partial\phi} = \left[\frac{\partial\zeta(\phi, \lambda, W)}{\partial\phi}\right]_{W=\text{const.}} + \frac{\partial\zeta(\phi, \lambda, W)}{\partial W}\frac{\partial W(\phi, \lambda)}{\partial\phi},$$

according to well-known rules of the calculus. We abbreviate this as

$$\left(\frac{\partial\zeta}{\partial\phi}\right)_S = \left(\frac{\partial\zeta}{\partial\phi}\right)_{W=W_P} + \frac{\partial\zeta}{\partial W}\frac{\partial W}{\partial\phi}.$$

Since

$$dW = -g\,dh,$$

this is equivalent to

$$\left(\frac{\partial\zeta}{\partial\phi}\right)_S = \left(\frac{\partial\zeta}{\partial\phi}\right)_{W=W_P} + \frac{\partial\zeta}{\partial h}\frac{\partial h}{\partial\phi}$$

or

$$\left(\frac{\partial\zeta}{\partial\phi}\right)_{W=W_P} = \left(\frac{\partial\zeta}{\partial\phi}\right)_S - \frac{\partial\zeta}{\partial h}\frac{\partial h}{\partial\phi}. \tag{8-74}$$

The partial derivative $\partial\zeta/\partial h$ is given by (8-66); furthermore, we have

$$\frac{1}{R}\frac{\partial h}{\partial\phi} = \frac{\partial h}{\partial s_\phi} = \tan\beta_1,$$

where β_1, in the terminology of Secs. 5-6 and 8-5, is the angle of inclination of a north-south terrain profile with respect to the horizontal. Hence we find

$$\left(\frac{\partial\zeta}{\partial\phi}\right)_{W=W_P} = \left(\frac{\partial\zeta}{\partial\phi}\right)_S + R\frac{\Delta g}{\gamma}\tan\beta_1.$$

We recall that $\partial\zeta/\partial\phi$ along the level surface $W = W_P$ is needed in (8-73), and that $\partial\zeta/\partial\phi$ along S is obtained by differentiating (8-52) or a similar solution for ζ. Hence we get from (8-73)

$$\begin{aligned}\xi &= -\left(\frac{1}{R}\frac{\partial\zeta}{\partial\phi}\right)_S - \frac{\Delta g}{\gamma}\tan\beta_1,\\ \eta &= -\left(\frac{1}{R\cos\phi}\frac{\partial\zeta}{\partial\lambda}\right)_S - \frac{\Delta g}{\gamma}\tan\beta_2,\end{aligned} \tag{8-75}$$

where β_2 correspondingly is the angle of inclination of an east-west profile.

Finally we must differentiate (8-52). Since (8-52) is Stokes' formula, applied to $\Delta g + G_1$ instead of Δg, the result is Vening Meinesz' formulas (Sec. 2-22). Replacing Δg in equations (2-210′) by $\Delta g + G_1$ gives the first term on the right-hand sides of (8-75). We thus obtain

$$\begin{aligned}\xi &= \frac{1}{4\pi\gamma}\iint_\sigma (\Delta g + G_1)\frac{dS}{d\psi}\cos\alpha\,d\sigma - \frac{\Delta g}{\gamma}\tan\beta_1,\\ \eta &= \frac{1}{4\pi\gamma}\iint_\sigma (\Delta g + G_1)\frac{dS}{d\psi}\sin\alpha\,d\sigma - \frac{\Delta g}{\gamma}\tan\beta_2.\end{aligned} \tag{8-76}$$

These are Molodensky's formulas for the deflection of the vertical at the earth's surface to a linear approximation. If we neglect the first-order terms—the correction G_1 and the inclinations β_1 and β_2—we are left with the original Vening Meinesz formulas (2-210′).

An alternative formula. Numerical calculations (Molodenskii et al., 1962a, p. 223) indicate that the two correction terms to Vening Meinesz' formula,

$$\frac{1}{4\pi\gamma}\iint_\sigma G_1 \frac{dS}{d\psi}\begin{Bmatrix}\cos\alpha\\ \sin\alpha\end{Bmatrix} d\sigma \quad \text{and} \quad -\frac{\Delta g}{\gamma}\begin{Bmatrix}\tan\beta_1\\ \tan\beta_2\end{Bmatrix}, \tag{8-77}$$

can each attain quite large values in steep mountains, but that they tend to cancel each other in equations (8-76). The reason for this becomes evident on splitting G_1 according to formula (8-58) of the preceding section. Then the first term in (8-77) for the component ξ becomes

$$\begin{aligned}\frac{1}{4\pi\gamma}\iint_\sigma G_1 \frac{dS}{d\psi}\cos\alpha\, d\sigma &= \frac{1}{4\pi\gamma}\iint_\sigma G_{11}\frac{dS}{d\psi}\cos\alpha\, d\sigma + \frac{1}{4\pi\gamma}\iint_\sigma G_{12}\frac{dS}{d\psi}\cos\alpha\, d\sigma\\ &= -\frac{1}{4\pi\gamma}\iint_\sigma \left(\frac{\partial\Delta g}{\partial h}h\right)\frac{dS}{d\psi}\cos\alpha\, d\sigma - \frac{\partial\zeta_{12}}{R\,\partial\phi}\\ &= -\frac{1}{4\pi\gamma}\iint_\sigma \left(\frac{\partial\Delta g}{\partial h}h\right)\frac{dS}{d\psi}\cos\alpha\, d\sigma + \frac{1}{R\gamma}\frac{\partial(h\Delta g)}{\partial\phi},\end{aligned}$$

and finally,

$$\frac{1}{4\pi\gamma}\iint_\sigma G_1\frac{dS}{d\psi}\cos\alpha\, d\sigma = -\frac{1}{4\pi\gamma}\iint_\sigma \left(\frac{\partial\Delta g}{\partial h}h\right)\frac{dS}{d\psi}\cos\alpha\, d\sigma + \frac{\Delta g}{\gamma}\tan\beta_1 + \frac{h}{\gamma}\frac{\partial\Delta g}{R\,\partial\phi}. \tag{8-78}$$

Here we have used equations (8-63) and (8-65′); because the horizontal derivatives of γ are so small, they have been neglected in these correction terms.

As we did at the end of the preceding section, we may refer the elevation h to the level of point P instead of to sea level by replacing h by $h - h_P$. Then the last term of (8-78) becomes zero, because h denotes h_P there and is replaced by $h_P - h_P = 0$. We thus obtain

$$\frac{1}{4\pi\gamma}\iint_\sigma G_1\frac{dS}{d\psi}\cos\alpha\, d\sigma = -\frac{1}{4\pi\gamma}\iint_\sigma \frac{\partial\Delta g}{\partial h}(h - h_P)\frac{dS}{d\psi}\cos\alpha\, d\sigma + \frac{\Delta g}{\gamma}\tan\beta_1. \tag{8-78′}$$

On adding the two terms (8-77), the first being expressed by (8-78′), we see that

$$\frac{\Delta g}{\gamma}\begin{Bmatrix}\tan\beta_1\\ \tan\beta_2\end{Bmatrix}$$

will cancel out, leaving

$$-\frac{1}{4\pi\gamma}\iint_{\sigma} \frac{\partial \Delta g}{\partial h}(h - h_P)\frac{dS}{d\psi}\begin{Bmatrix}\cos\alpha\\ \sin\alpha\end{Bmatrix} d\sigma \tag{8-79}$$

as the correction to Vening Meinesz' formula. This correction will in general be much smaller than either of the terms (8-77), which explains their above-mentioned tendency to cancel out.

Hence the generalized Vening Meinesz equations (8-76) become

$$\begin{Bmatrix}\xi\\ \eta\end{Bmatrix} = \frac{1}{4\pi\gamma}\iint_{\sigma}\left[\Delta g - \frac{\partial \Delta g}{\partial h}(h - h_P)\right]\frac{dS}{d\psi}\begin{Bmatrix}\cos\alpha\\ \sin\alpha\end{Bmatrix} d\sigma. \tag{8-80}$$

This formula is as accurate as (8-76), but simpler. Its geometrical interpretation is analogous to that of (8-71). The gravity anomalies Δg are reduced to the level of point P, so that we obtain

$$\Delta g^* = \Delta g - \frac{\partial \Delta g}{\partial h}(h - h_P).$$

Since these anomalies *refer to a level surface*, Vening Meinesz' formula can now be applied directly and gives (8-80).

This shows the relation between the equivalent formulas (8-76) and (8-80), and also indicates why (8-80) seems to be preferable in practice. In (8-76) the two cancelling terms are computed in completely different ways, whereby additional errors may be introduced. In (8-80) they are omitted from the beginning.

Equation (8-80), supplemented by nonlinear terms, is also used by Arnold (1962a).

Relation with the geographical coordinates. The deflection components ξ and η as given by the above expressions represent the deviation of the actual plumb line from the normal plumb line at the ground point P. They are therefore defined by

$$\begin{aligned}\xi &= \Phi - \phi^*,\\ \eta &= (\Lambda - \lambda^*)\cos\phi.\end{aligned} \tag{8-81}$$

The symbols Φ and Λ represent the *actual geographical coordinates* of P; they are thus astronomical coordinates referred to the ground. The symbols ϕ^* and λ^* represent the *normal geographical coordinates* of P, defining the direction of the normal plumb line at P; they are not identical with the geodetic coordinates of P, ϕ and λ, which are the normal geographical coordinates of the foot point Q_0 of the straight perpendicular to the ellipsoid (Fig. 8-6).

The normal coordinates of P, ϕ^* and λ^*, differ from the normal coordinates of Q_{00}, ϕ' and λ', by the correction for the normal curvature of the plumb line; see Sec. 5-6. Formula (5-34) gives

$$\phi^* = \phi' + f^*\frac{h}{R}\sin 2\phi,$$

$$\lambda^* = \lambda'.$$

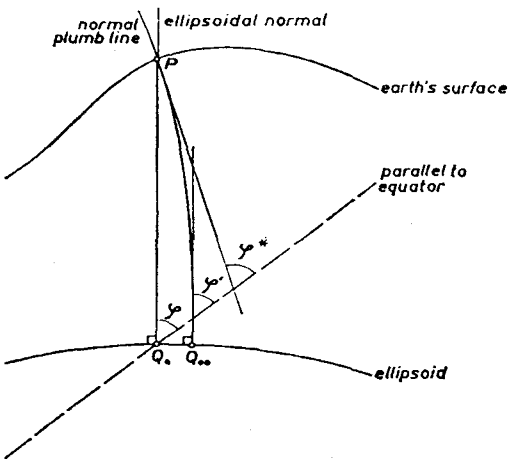

FIGURE 8-5
Normal geographical latitude ϕ^* and geodetic latitude ϕ.

Because of the rotational symmetry we have rigorously

$$\lambda' = \lambda,$$

because Q_0 and Q_{00} lie on the same ellipsoidal meridian. Furthermore, even in extreme cases the distance between Q_0 and Q_{00} can never exceed a few centimeters. For this reason we may also set

$$\phi' = \phi,$$

without introducing a perceptible error. Hence we can identify ϕ' and λ' with ϕ and λ, which are the *geodetic coordinates* of P according to Helmert's projection (Sec. 5-3). Therefore, we may replace the above equations for ϕ^* and λ^* by

$$\phi^* = \phi + f^* \frac{h}{R} \sin 2\phi, \qquad \lambda^* = \lambda. \tag{8-82}$$

If we introduce the deflection components according to Helmert's projection, defined as

$$\xi_{\text{Helmert}} = \Phi - \phi, \qquad \eta_{\text{Helmert}} = (\Lambda - \lambda) \cos \phi, \tag{8-83}$$

we see that they are related to ξ and η by the equations

$$\xi_{\text{Helmert}} = \xi + f^* \frac{h}{R} \sin 2\phi, \qquad \eta_{\text{Helmert}} = \eta. \tag{8-84}$$

Thus ξ and ξ_{Helmert} differ by the normal reduction for the curvature of the plumb line,

$$-\delta\phi_{\text{normal}} = f^* \frac{h}{R} \sin 2\phi.$$

The deflection components ξ_{Helmert} and η_{Helmert} are used in astrogeodetic computations; ξ and η are those obtained gravimetrically from formulas (8-76) or (8-80).

These relations are quite analogous to the corresponding equations (5-33) for the conventional method using the geoid. The corrections $\delta\phi$ and $\delta\lambda$ for the curvature of the plumb line are now replaced by their normal values $\delta\phi_{\text{normal}}$ and $\delta\lambda_{\text{normal}} = 0$. This is comparable to the replacement of the orthometric height in the conventional theory by the normal height in the new theory.

8-10. Downward Continuation to Sea Level

The basic integral equation (8-27) is extended over the physical surface of the earth. We may, however, interpret its solution as a free-air reduction to sea level by means of the anomalous gradient $\partial\Delta g/\partial h$, as we have seen in Sec. 8-8. This forms a link between the modern and the conventional point of view.

The interpretation of Sec. 8-8 was limited to the linear approximation, terms of second and higher degree in the quantities (8-41) having been neglected. However, we may also extend the principle of free-air reduction to sea level to higher approximations.

Assume a fictitious field of gravity anomalies Δg^* on the ellipsoid E (at sea level), which generate on the earth's surface S (at ground level) the measured free-air anomalies Δg and the height anomalies ζ (Fig. 8-7). This is to be understood in the sense that Δg and Δg^* are related to each other by upward continuation according to Sec. 6-8.

An integral equation. According to equation (6-75) we have

$$\Delta g = \frac{t^2(1 - t^2)}{4\pi} \iint_\sigma \frac{\Delta g^*}{D^3} \, d\sigma, \tag{8-85}$$

where

$$t = \frac{R}{r} \quad \text{and} \quad r = R + h_P.$$

In this equation the terms that depend on the flattening of the reference ellipsoid are neglected (spherical approximation); R is a mean radius of the earth, and h is the topographic height.

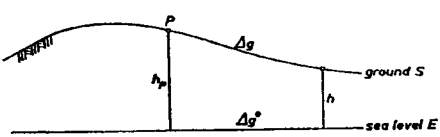

FIGURE 8-7

Free-air anomalies at ground level, Δg, and at sea level, Δg^*.

Given Δg^* we may compute Δg from (8-85). But since Δg at ground level is obtained by measurement, we shall obtain the fictitious anomalies Δg^* at sea level by solving (8-85) for Δg^*. In this sense, equation (8-85) is a linear integral equation of the first kind for the function Δg^*; see Sec. 8-4. It may again be solved by an expansion in powers of the quantities (8-41), according to the method of Sec. 8-7. The linear approximation to the solution is then obviously given by (8-69).

Another way of solving (8-85) is to use an iterative method. Let us transform (8-85) in the following way. The identity

$$t^2 = \frac{t^2(1-t^2)}{4\pi} \iint_\sigma \frac{d\sigma}{D^3} \tag{8-86}$$

may be verified by a straightforward evaluation of the integral, taking the definition (6-44) of D into account. Since Δg outside the integral refers to the point P, we may write (8-85) more precisely as

$$\Delta g_P = \frac{t^2(1-t^2)}{4\pi} \iint_\sigma \frac{\Delta g^*}{D^3}\, d\sigma. \tag{8-85$'$}$$

Next we multiply (8-86) by Δg_P^*, which is a constant with respect to the integration, and subtract the product from (8-85′). The result is

$$\Delta g_P - t^2\, \Delta g_P^* = \frac{t^2(1-t^2)}{4\pi} \iint_\sigma \frac{\Delta g^* - \Delta g_P^*}{D^3}\, d\sigma.$$

The factor t^2 of Δg_P^* may be set equal to unity without appreciable loss of accuracy. Thus we obtain, as an alternative expression for (8-85), the integral equation

$$\Delta g_P^* = \Delta g_P - \frac{t^2(1-t^2)}{4\pi} \iint_\sigma \frac{\Delta g^* - \Delta g_P^*}{D^3}\, d\sigma, \tag{8-87}$$

which lends itself to iterative solution by automatic computer.

As a first approximation we set

$$\overset{(1)}{\Delta g^*} = \Delta g,$$

then compute a second approximation by

$$\overset{(2)}{\Delta g_P^*} = \Delta g_P - \frac{t^2(1-t^2)}{4\pi} \iint_\sigma \frac{\overset{(1)}{\Delta g^*} - \overset{(1)}{\Delta g_P^*}}{D^3}\, d\sigma,$$

a third approximation by

$$\overset{(3)}{\Delta g_P^*} = \Delta g_P - \frac{t^2(1-t^2)}{4\pi} \iint_\sigma \frac{\overset{(2)}{\Delta g^*} - \overset{(2)}{\Delta g_P^*}}{D^3}\, d\sigma,$$

and so forth. This process will usually converge well.

In this way the "downward continuation" of the gravity anomalies from ground level (Δg) to sea level (Δg^*) may be carried out. This iterative process may also be performed as a plane approximation, using equation (6-76).

Physical interpretation. The sea-level anomalies Δg^* thus obtained generate on the physical surface of the earth a field of gravity anomalies Δg that is identical with the actual gravity anomalies on the earth's surface as obtained from observation. Therefore, the gravity anomalies that they generate outside the earth must also be identical with the actual gravity anomalies outside the earth, according to an extension of Stokes' uniqueness theorem for harmonic functions (Sec. 1-7); remember that the function $r\,\Delta g$ is harmonic according to Sec. 2-15.

It follows that the harmonic function T that is produced by Δg^* according to Pizzetti's generalization (2-161) of Stokes' formula

$$T(r, \theta, \lambda) = \frac{R}{4\pi} \iint_\sigma \Delta g^* \, S(r, \psi) \, d\sigma, \tag{8-88}$$

is identical with the actual disturbing potential of the earth *outside and on its surface.*

Inside the earth, however, the harmonic function defined by (8-88) cannot coincide with the actual disturbing potential, which is not harmonic. We say that (8-88) inside the earth, but above the reference ellipsoid, defines the *analytical continuation* of the external potential. Hence Δg^* corresponds to this analytical continuation of the external field, and not to the actual internal field. From this we see that Δg^* and the mathematical model based on it have no direct correspondence to physical reality. Nevertheless, a physical interpretation is possible. Imagine that the masses outside the ellipsoid are shifted into its interior in such a way that the potential on and outside the earth's surface S remains unchanged. The new disturbing potential is now harmonic outside the ellipsoid, since the exterior masses have been removed, and coincides with the original potential on and outside S. It is therefore given by (8-88) everywhere outside the ellipsoid.

This interpretation furnishes a relation with the ideas of Sec. 8-2, because this free-air reduction to sea level can indeed be considered a mass-transporting gravity reduction. What characterizes our present reduction is the fact that this mass-shift is not actually carried out, the anomalies Δg^* being obtained by solving the integral equation (8-85), *which does not explicitly contain the masses outside the ellipsoid*—an important distinction from the methods discussed in Sec. 8-2.

Applications. This method is very attractive theoretically and practically. It combines the advantages of the modern and the conventional approaches, while avoiding the drawbacks of both. In agreement with Molodensky's ideas, the method does not require that the rock density be known, but avoids Molodensky's integral equations extended over the surface of the earth and replaces them by the very simple equation (8-85), which is extended over the reference ellipsoid as approximated by a sphere or even by a plane. As in the conventional method of using gravity reductions, the simple formulas referring to equipotential surfaces that were considered in Chapter 2, such as Stokes' formula, can now be applied *rigorously*, but unlike the conventional reductions the downward continuation does not distort the external gravity field, which is of particular interest at present.

After reduction, or rather downward continuation, of Δg^* to sea level by solving (8-87) for Δg^* we can compute the external gravity field, the spherical harmonics, etc., rigorously by means of the conventional formulas of Chapters 2 and 6, provided we use Δg^* rather than Δg in these formulas. For instance, the coefficients of the spherical harmonics of the gravitational potential may be obtained by expanding the function Δg^* according to equations (2-195b) and (2-196); see also Sec. 6-7.

If we wish to compute the height anomaly ζ at a point P at ground level we must remember that P lies *above* the ellipsoid, so that the formulas for the external gravity field are to be applied. On dividing (8-88) by γ we obtain

$$\zeta = \frac{R}{4\pi\gamma} \iint_\sigma \Delta g^* \, S(r, \psi) \, d\sigma, \tag{8-89}$$

where $r = R + h$, h being the topographic height of P. The function $S(r, \psi)$ is expressed by (6-31) or (6-45). Similarly ξ and η, being deflections of the vertical *above* sea level, must be computed by (6-49) and (6-39b). This gives

$$\begin{Bmatrix} \xi \\ \eta \end{Bmatrix} = \frac{t}{4\pi\gamma} \iint_\sigma \Delta g^* \frac{\partial S(r, \psi)}{\partial \psi} \begin{Bmatrix} \cos \alpha \\ \sin \alpha \end{Bmatrix} d\sigma, \tag{8-90}$$

where $\partial S(r, \psi)/\partial \psi$ is expressed by (6-42) or (6-46b). The linear approximation of (8-89) is evidently represented by (8-68).

This solution of Molodensky's problem by downward continuation of Δg to sea level is probably the most comprehensive and versatile method. The solution is carried out in two steps:

1. Downward continuation of Δg from ground level to sea level, using (8-69) or an iterative solution of (8-87), which is more accurate.
2. Computation of height anomalies, deflections of the vertical at ground level, the external gravity field, etc., by means of the conventional spherical formulas.

This indirect procedure, downward continuation to sea level and again upward continuation to ground level or above, has the advantage that only the conventional spherical formulas are needed; yet at the same time the irregularities of the earth's topography are fully taken into account. Step (1) need be performed only once; the resulting anomalies Δg^* may be stored and used for all further computations.

This solution has been advocated by Bjerhammar (1964), who solved the integral equation (8-85) in a different way.

Validity of this method. So far we have assumed that the integral equation (8-85) can actually be solved for Δg^*. Obviously this is the case if, and only if, it is possible to shift the masses outside the ellipsoid into its interior in such a way that the potential outside the earth remains unchanged or, in other words, if the analytical continuation of the disturbing potential T is a regular function everywhere between the earth's surface and the ellipsoid. Thus the question arises whether the external potential can be analytically continued down to sea level.

Rigorously, the answer must be in the negative, in view of the irregularities of topography (Molodenskii et al., 1962a, p. 120; Moritz, 1965, Sec. 6.4). This fact is also related to the divergence at the earth's surface of the spherical-harmonic expansion for the external potential (Sec. 2-5).

However, the analytical continuation of the external potential down to sea level is probably possible *with sufficient accuracy for all practical purposes.* Bjerhammar has pointed out that the assumption of a complete continuous gravity coverage at every point of the earth's surface, from which the above negative answer follows, is unrealistic because we can measure gravity only at discrete points. If the purpose of physical geodesy is understood as the *determination of a gravity field that is compatible with the given discrete observations*, then it is always possible to find a potential that can be analytically continued down to the ellipsoid.

The practical applicability of downward continuation to sea level is also borne out by the fact that its linear approximation may be obtained by a formal transformation of (8-52), as discussed in Sec. 8-8, in a way which is not affected by the difficulties of analytical continuation.

8-11. Gravity Reduction in the Modern Theory

In Sec. 8-2 we have considered gravity reductions from the point of view of the determination of the geoid. It is quite remarkable that these reductions, such as the Bouguer or the isostatic reduction, can also be incorporated into the new method of direct determination of the earth's physical surface, although with essentially changed meaning (Arnold, 1962b; Levallois, 1962; Moritz, 1965, Sec. 5.2; Pellinen, 1962).

Let the masses outside the geoid be removed or moved inside the geoid, as described in Sec. 8-2, and consider the effect of this procedure on quantities *referred to the ground.*

We denote the changes in potential and in gravity by δW and δg; then the new values at the ground will be

$$\begin{aligned} W^c &= W - \delta W, \\ g^c &= g - \delta g. \end{aligned} \tag{8-91}$$

The disturbing potential $T = W - U$ becomes

$$T^c = T - \delta W.$$

The physical surface S as such will remain unchanged, but the telluroid Σ will change, because its points Q are defined by $U_Q = W_P$, and the potential W at any surface point P will be affected by the mass displacements according to (8-91). The distance QQ^c between the original telluroid Σ and the changed telluroid Σ^c (Fig. 8-8) is given by

$$QQ^c = \frac{\delta W}{\gamma}$$

according to Bruns' theorem. This is identical with the variation of the height anomaly ζ, so that

$$\delta\zeta = \zeta - \zeta^c = \frac{\delta W}{\gamma}. \tag{8-92}$$

Normal gravity γ on the telluroid Σ becomes on the changed telluroid Σ^c:

$$\gamma^c = \gamma + \frac{\partial\gamma}{\partial h}\delta\zeta = \gamma + \frac{1}{\gamma}\frac{\partial\gamma}{\partial h}\delta W, \tag{8-93}$$

so that the new gravity anomaly will be

$$\Delta g^c = g^c - \gamma^c = (g - \delta g) - \left(\gamma + \frac{1}{\gamma}\frac{\partial\gamma}{\partial h}\delta W\right)$$

or

$$\Delta g^c = \Delta g - \delta g - \frac{1}{\gamma}\frac{\partial\gamma}{\partial h}\delta W. \tag{8-94}$$

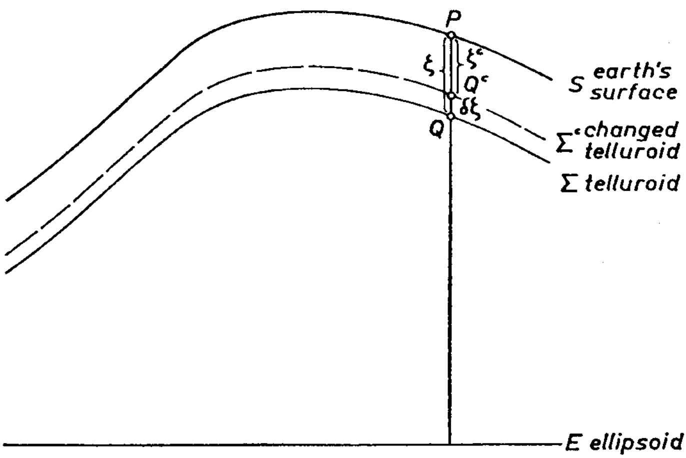

FIGURE 8-8

The telluroid before and after gravity reduction, Σ and Σ^c.

The reduced gravity anomaly Δg^c consists of the free-air anomaly Δg and two reductions:

1. the direct effect, $-\delta g$, of the shift of the outer masses on g; and
2. the "indirect effect,"

$$-\frac{1}{\gamma}\frac{\partial \gamma}{\partial h}\,\delta W,$$

   of this shift on γ, because of the change of the telluroid to which γ refers.

If the masses outside the geoid are completely removed, then Δg^c is a Bouguer anomaly; if the outer masses are shifted vertically downward according to some isostatic hypothesis, then Δg^c is an isostatic anomaly, etc. In this way we may get a "ground equivalent" for each conventional gravity reduction.

Now we may describe the determination of the height anomalies ζ in a way that is similar to the corresponding procedure for the geoidal undulations N of Sec. 8-2:

1. The masses outside the geoid are, by computation, removed entirely or else moved inside the geoid; W and g change to W^c and g^c according to (8-91).
2. The point at which normal gravity is computed is moved from the ellipsoid upward to the telluroid point Q.
3. The indirect effect, the distance $QQ^c = \delta\zeta$, is computed by (8-92).
4. The point to which normal gravity refers is now moved from the point Q of the telluroid Σ to the point Q^c of the changed telluroid Σ^c, according to (8-93).
5. The changed height anomalies ζ^c are computed from the "reduced" gravity anomalies Δg^c (8-94) by any solution of Molodensky's problem, such as equation (8-50).
6. Finally, the original height anomalies ζ are obtained by considering the indirect effect according to

$$\zeta = \zeta^c + \delta\zeta. \tag{8-95}$$

The purpose of this somewhat complicated procedure is to make use of the well-known advantages of Bouguer and isostatic anomalies. The Bouguer anomalies, and even more so the isostatic anomalies, are *smoother and more representative* than the free-air anomalies and can therefore be interpolated more easily and more accurately.

The isostatic gravity anomalies Δg^c in the new sense are thus quite analogous to the conventional isostatic anomalies. (The same holds, of course, for any other type of gravity reduction.) The difference is that now the isostatic anomalies, etc., *refer to the physical surface of the earth* as well as the free-air anomalies. If the isostatic anomalies in this new sense are analytically continued from the earth's surface down to the geoid, then the isostatic anomalies in the conventional sense are obtained.

Hence the isostatic anomalies according to the conventional definition and those according to the new definition are related through analytical continuation. This fact leads to two conclusions. First, the difference between the isostatic anomalies according to these two definitions will be small, because the distance along which this analytical continuation is made is only the height above sea level and because the isostatic reduction achieves a strong smoothing of the anomalous gravity field. This difference is considerably smaller than the corresponding difference between free-air anomalies at ground level and at sea level (Groten, 1964). This fact clearly provides a computational advantage if isostatic anomalies are used in a formula such as (8-87).

Second, we obtain a relation between the conventional and the modern use of gravity reduction if the method of downward continuation, as discussed in the preceding section, is applied for obtaining the height anomalies. As we have just seen, the gravity anomalies Δg^{c*} at sea level, obtained by downward continuation of the isostatic ground-level anomalies Δg^c, are identical with the isostatic anomalies in the conventional sense. Hence we obtain on the one hand the height anomalies by

$$\zeta = \frac{R}{4\pi\gamma} \iint_\sigma \Delta g^{c*} S(R + h, \psi)\, d\sigma + \left(\frac{\delta W}{\gamma}\right)_{\text{ground}}, \tag{8-96}$$

according to (8-89) and (8-92), and on the other hand the geoidal undulations by

$$N = \frac{R}{4\pi\gamma_0} \iint_\sigma \Delta g^{c*} S(\psi)\, d\sigma + \left(\frac{\delta W}{\gamma}\right)_{\text{geoid}}, \tag{8-97}$$

according to (8-3). Since the height anomalies refer to the elevation h, the function $S(R + h, \psi)$ replaces in (8-96) the original function of Stokes $S(\psi) \equiv S(R, \psi)$, which occurs in (8-97) because the geoidal undulation refers to zero elevation. Normal gravity at telluroid and ellipsoid are denoted by γ and γ_0.

Summarizing, we have the following steps:

1. Computation of the free-air anomaly at ground level, Δg, according to (8-7).
2. Computation of the isostatic anomaly at ground level, Δg^c, according to (8-94).
3. Downward continuation of Δg^c by an iterative solution of (8-87) where Δg and Δg^* are replaced by Δg^c and Δg^{c*}. The resulting isostatic anomalies at sea level, Δg^{c*}, may now be used for two purposes: either for

4a. the determination of the physical surface of the earth according to (8-96), or for

4b. the determination of the geoid according to (8-97).

An error in the assumed density of the masses below the earth's surface affects the geoidal undulations as determined from (8-97) but does not influence the height anomalies resulting from (8-96). This is clear because a wrong guess of the density means only that the masses above sea level are not completely removed, which is no worse than not removing them at all when using free-air anomalies.

It is of historic interest to note that as early as 1912 Hayford referred his isostatic gravity anomalies to the earth's surface by reducing γ upward to ground level instead of reducing g to sea level; of course he was not aware of the theoretical refinements that go with this procedure.

8-12. Determination of the Geoid from Ground-level Anomalies

We have seen that it is possible to determine the physical surface of the earth, by means of the height anomalies ζ, and the direction of the plumb line on it, by means of the deflection components ξ and η, from free-air anomalies referred to the ground. If we plot the orthometric height H downward along the plumb line, starting from the physical surface, then the locus of the points so obtained will be the geoid (Fig. 8-9).

This geometrical idea may be formulated analytically in the following way. Conventionally, the height h above the ellipsoid is given by

$$h = H + N;$$

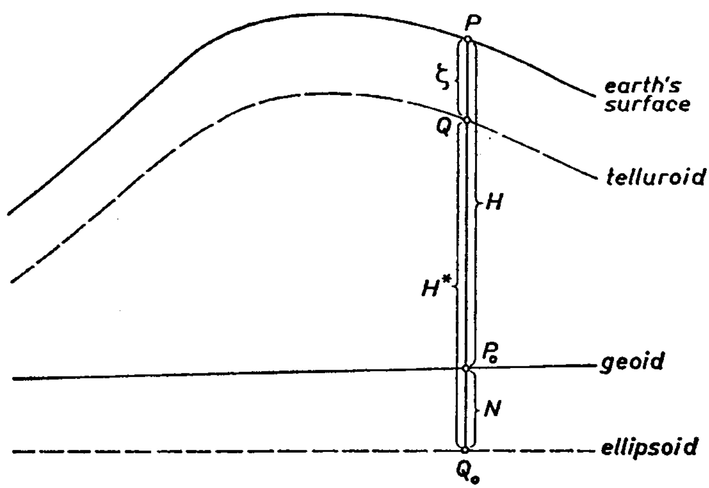

FIGURE 8-9

The geoid lies at a depth H below the earth's surface.

according to the modern theory, by

$$h = H^* + \zeta.$$

From these two equations we get

$$N = \zeta + H^* - H. \tag{8-98}$$

This means that the difference between the geoidal undulation N and the height anomaly ζ is equal to the difference between the normal height H^* and the orthometric height H. Since ζ is also the undulation of the quasigeoid, this difference is also the distance between geoid and quasigeoid.

According to Sec. 4-6 the two heights are defined by

$$H = \frac{C}{\bar{g}}, \quad H^* = \frac{C}{\bar{\gamma}},$$

where C is the geopotential number, $\bar{g}$ is the mean gravity along the plumb line between geoid and ground, and $\bar{\gamma}$ is the mean normal gravity along the normal plumb line between ellipsoid and telluroid. By eliminating C between these two equations we readily find

$$H^* - H = \frac{\bar{g} - \bar{\gamma}}{\bar{\gamma}} H, \tag{8-99}$$

which is also the distance between the geoid and the quasigeoid; hence

$$N = \zeta + \frac{\bar{g} - \bar{\gamma}}{\bar{\gamma}} H. \tag{8-100}$$

The height anomaly ζ may be expressed, for instance, by Molodensky's formula (8-50). Then we obtain

$$N = \frac{R}{4\pi\gamma} \iint_\sigma \Delta g\, S(\psi)\, d\sigma + \frac{R}{4\pi\gamma} \iint_\sigma G_1\, S(\psi)\, d\sigma + \frac{\bar{g} - \bar{\gamma}}{\bar{\gamma}} H, \tag{8-101}$$

where G_1 is the term (8-70). Thus N is given by Stokes' integral, *applied to free-air anomalies at ground level,* and two small corrections, where

1. the term containing G_1 represents the effect of topography; and
2. the term containing $\bar{g} - \bar{\gamma}$ represents the distance between the geoid and the quasigeoid.

If we neglect these two corrections, then the geoidal undulations N are given by Stokes' integral using free-air anomalies. This was first noted by Stokes in 1849. A new approach by Jeffreys (1931) by means of Green's identities started the recent developments, which culminated in the work of Molodensky and others. The ideas of the present section are essentially those of Arnold (1962a), although his treatment is different. His final formulas are equivalent to the expression of ζ in terms of free-air gravity anomalies reduced to point level, equation (8-71), but he also considers nonlinear terms.

The advantage of this method for the determination of N is that the density of the masses above sea level enters only indirectly, as an effect on the orthometric height H through the mean gravity $\bar{g}$, which must be computed by a Prey reduction (Sec. 4-4). Hence, as far as errors in the density are concerned, the geoidal undulation N as obtained by this method is as accurate as the orthometric height.

As a matter of fact, the gravity anomaly Δg in this method refers to ground level; it is the difference between gravity at ground and normal gravity at the telluroid. Instead of using directly this free-air anomaly we may also use other gravity anomalies—for instance, the isostatic anomaly in the sense of Sec. 8-11.

8-13. Review

The new methods described in this chapter are primarily intended for the determination of the physical surface of the earth, but they are also well suited for the determination of the geoid (Sec. 8-12). Their essential feature is that the gravity anomalies now refer to the ground, whether we deal with free-air anomalies or with isostatic, or other similarly reduced, gravity anomalies (Sec. 8-11).

The immediate result is the height anomaly ζ, the separation between the geopotential and the corresponding spheropotential surface at ground level. By plotting the height anomalies above the ellipsoid we get the quasigeoid. This geoid-like surface has no physical significance, but it furnishes a convenient visualization of the height anomalies. By plotting the orthometric height from the earth's surface vertically downward, we obtain the geoid.

It is instructive to compare the geoid and the quasigeoid. The geoidal undulation N and ζ, the undulation of the quasigeoid, are related by (8-100), or

$$N - \zeta = \frac{\bar{g} - \bar{\gamma}}{\bar{\gamma}} H = H^* - H. \tag{8-102}$$

The term $\bar{g} - \bar{\gamma}$ is approximately equal to the Bouguer anomaly; this may be immediately seen by using (4-25) together with

$$\bar{\gamma} \doteq \gamma - \frac{1}{2}\frac{\partial \gamma}{\partial h} H.$$

The quantity $\bar{\gamma}$ in the denominator may be replaced by a constant average value, say 981 gals. Since the Bouguer anomaly is rather insensitive to local topographic irregularities, the coefficient is locally constant, so that there is approximately a linear relation between ζ and the local irregularities of the height H. In other words, the quasigeoid mirrors the topography (Fig. 8-10).

To get a quantitative estimate of the difference $N - \zeta$, we again use the fact that

$$\frac{\bar{g} - \bar{\gamma}}{\bar{\gamma}} \doteq \frac{\Delta g_B}{981 \text{ gals}} \doteq 10^{-3}\, \Delta g_B,$$

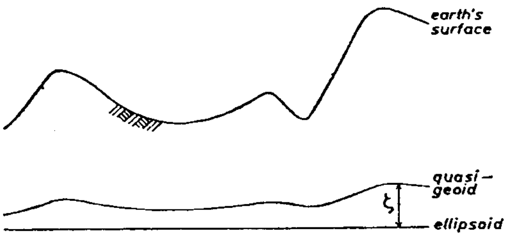

FIGURE 8-10
The quasigeoid mirrors the topography.

where Δg_B is the Bouguer anomaly in gals, so that

$$(\zeta - N)_{\text{in meters}} \doteq -\Delta g_{B \text{ in gals}} \cdot H_{\text{in km}}. \tag{8-103}$$

Since Δg_B is usually negative on the continents, the differences $\zeta - N$ are usually positive there. In other words, the height anomaly ζ is in general greater than the corresponding geoidal undulation N on land. On the oceans we of course have $\zeta = N$. If $\Delta g_B = -100$ mgals $= -0.1$ gal and $H = 1$ km, then

$$\zeta - N = 0.1 \text{ meter.}$$

Furthermore, the Bouguer anomaly depends on the *mean* elevation of the terrain, decreasing approximately by 0.1 gal per 1 km average elevation. Assuming as a rough estimate, which may be verified by inspecting maps of Bouguer anomalies,

$$\Delta g_{B \text{ in gals}} = -0.1 H^{\text{av}}{}_{\text{in km}},$$

we obtain

$$(\zeta - N)_{\text{in meters}} \doteq +0.1 H^{\text{av}}{}_{\text{in km}} H_{\text{in km}}, \tag{8-104}$$

where H is the height of the station and H^{av} is an average height of the area considered. We see that the difference $\zeta - N$ increases faster than the elevation, almost as the square of the elevation.

As a matter of fact, this formula is suited only to giving an idea of the order of magnitude. As an extreme example, take Mt. Blanc in the Alps, $H = 4807$ meters. If the average elevation is taken as 3 km, the above formula gives $\zeta - N = 1.4$ meters, whereas, more accurately, according to Arnold (1960, p. 66),

$$\bar{g} - \bar{\gamma} = -360 \text{ mgals}, \quad \zeta - N = 1.8 \text{ meters.}$$

Since $\zeta - N = H - H^*$, the approximate formulas given above may also be used to estimate the differences between the orthometric height H and the normal height H^*.

A theoretically important point is that the quasigeoid can be determined without hypothetical assumptions concerning the density, but not so the geoid. The avoidance of such assumptions has been the guiding idea of Molodensky's research. However, orthometric heights are but little affected by errors in density. The error in H due to the imperfect knowledge of the density hardly ever exceeds 1–2 decimeters even in extreme cases (Sec. 4-4). It is presumably smaller than the inaccuracy of the corresponding ζ even with very good gravity coverage, because of inevitable errors of interpolation, etc. If, therefore, the method of Sec. 8-12 is used, the geoid can be determined with practically the same accuracy as the quasigeoid.[1] Thus we may well retain the geoid with its physical significance and its other advantages.

How much do Molodensky's formulas differ from the corresponding equations of Stokes and Vening Meinesz? The deviation of ζ from the result of the original Stokes formula is given by the equivalent expressions

$$\zeta_1 = \frac{R}{4\pi\gamma} \iint_\sigma G_1 S(\psi)\, d\sigma \quad \text{or} \quad \zeta_1 = -\frac{R}{4\pi\gamma} \iint_\sigma \frac{\partial \Delta g}{\partial h} (h - h_P) S(\psi)\, d\sigma,$$

according to equations (8-50) and (8-71). This correction is even considerably smaller than the difference $\zeta - N$. For Mt. Blanc, Arnold (1960, p. 57) found ζ_1 to be only -0.2 meter, as compared to $\zeta - N = 1.8$ meters.

The deflection of the vertical is relatively more affected by the Molodensky correction than is the height anomaly. In extreme cases this correction may attain values of a few seconds, as test computations by Arnold (1960) and studies of models by Molodensky (Molodenskii et al., 1962a, pp. 217–225) indicate. This is considerable, since $1''$ in the deflection corresponds to 30 meters in position.

We may summarize the result of applying Stokes' and Vening Meinesz' formulas to free-air anomalies directly, without any corrections. Stokes' formula yields height anomalies ζ with high accuracy; for many practical purposes we may, in addition, identify these height anomalies with the corresponding geoidal undulations N. Vening Meinesz' formula gives deflections of the vertical at ground level that are relatively less accurate, but often acceptable.

An advantage of the modern theory is its direct relation to the external gravity field of the earth, which is particularly important nowadays for the computation of the effect of gravitational disturbances on space trajectories and satellite orbits. It is immediately clear that ground-level quantities, such as free-air gravity anomalies, are better suited for this purpose than the corresponding quantities referred to the geoid, which is separated from the external field by the outer masses. For the computation of the external field and of spherical harmonics the method described in Sec. 8-10 is particularly appropriate; see also Sec. 6-7.

Practically it is usually adequate to consider the linear approximation only by using (8-69). In many cases it is even possible to neglect the correction $-(\partial \Delta g/\partial h)h$, identifying the sea-level free-air anomalies Δg^* with the corresponding ground-level anomalies Δg. In agreement with Sec. 3-7, these free-air anomalies $\Delta g^* \doteq \Delta g$ may also be considered approximations to condensation anomalies in the sense of Helmert. This approximation is particularly sufficient for the external gravity field, spherical harmonics, and geoidal undulations or

[1] Theoretically it is even possible to eliminate completely the errors arising from the use of the geoid (Moritz, 1962, 1964).

height anomalies. For deflections of the vertical it is often necessary to use a more careful approach, such as the consideration of the indirect effect with mass-transporting gravity reductions (Sec. 8-2) or the modern methods of Sec. 8-9.

In high and steep mountains the approach of Molodensky and others through free-air anomalies meets with practical difficulties, such as unreliability of interpolation, large corrections, and other computational problems. To avoid this, even the "orthodox" Russians use a kind of Bouguer reduction in the sense of Sec. 8-11 and suggest a smoothing of the physical surface somewhat similar to de Graaff-Hunter's (1958) model earth (Pellinen, 1962); Arnold (1962b) advocates the isostatic reduction. Thus the clash between "conventional" and "modern" ideas gives way to an important reconciliation.

For further study the reader is referred to the textbooks by Molodenskii et al. (1962a) and Magnitzki et al. (1964).

References

Arnold, K. (1960). Numerische Beispiele zur strengen Theorie der Figur der Erde. Veröff. Geod. Inst. Potsdam, No. 16.

Arnold, K. (1962a). Recent achievements in the theory of the level-surfaces of the earth. Bull. Géod., No. 63, pp. 39–42.

Arnold, K. (1962b). The deflections of the vertical expressed in free-air anomalies and their relationship to the gravimetric-isostatic deflections. *Ibid.*, No. 65, pp. 259–264.

Bjerhammar, A. (1964). A new theory of geodetic gravity. Stockholm, Trans. Roy. Inst. Technol., No. 243.

Brovar, V. V. (1964). On the solutions of Molodensky's boundary value problem. Bull. Géod., No. 72, pp. 167–173.

Burša, M. (1965). On practical application of the solution of Molodensky's integral equation in the first approximation. Studia geophysica et geodaetica, v. 9, pp. 144–149.

Courant, R., and D. Hilbert (1953). Methods of mathematical physics, vol. 1. New York, Interscience Publishers.

de Graaff-Hunter, J. (1958). Report of Study-group No. 8. Reduction of observed gravity. Bull. Géod., No. 50, pp. 1–16.

de Graaff-Hunter, J. (1960). The shape of the earth's surface expressed in terms of gravity at ground level. *Ibid.*, No. 56, pp. 191–200.

Groten, E. (1964). On gravity reductions. Report No. 44, Inst. Geod. Phot. Cart., Ohio State Univ.

Hirvonen, R. A. (1960). New theory of the gravimetric geodesy. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 32.

Hirvonen, R. A. (1961). The reformation of geodesy. J. Geophys. Research, v. 66, pp. 1471–1478.

Jeffreys, H. (1931). An application of the free-air reduction of gravity. Gerlands Beitr. Geophys., v. 31, pp. 378–386.

Jung, K. (1956). Figur der Erde. Handbuch der Physik, v. 47, pp. 534–639, Berlin, Springer.

Levallois, J. J. (1958). Sur une équation intégrale très générale de la gravimétrie. Bull. Géod., No. 50, pp. 36–49.

Levallois, J. J. (1962). Considérations générales sur les réductions de la pesanteur. *Ibid.*, No. 63, pp. 79–93.

Magnitzki, W. A., W. W. Browar, and B. P. Schimbirew (1964). Theorie der Figur der Erde. Berlin, VEB Verlag für Bauwesen.

Molodenskii, M. S., V. F. Eremeev, and M. I. Yurkina (1962a). Methods for study of the external gravitational field and figure of the earth. Transl. from Russian (1960). Jerusalem, Israel Program for Scientific Translations.

Molodensky, M. S., V. F. Yeremeyev, and M. I. Yourkina (1962b). An evaluation of accuracy of Stokes' series and of some attempts to improve his theory. Bull. Géod., No. 63, pp. 19–38.

Moritz, H. (1962). Über die Hypothesen bei der Geoidbestimmung. Zeitschrift für Vermessungswesen, v. 87, pp. 138–145.

Moritz, H. (1964). Zur Bestimmung des Geoides und seiner Verwendung als Reduktionsfläche. Zeitschrift für Vermessungswesen, v. 89, pp. 200–202.

Moritz, H. (1965). The boundary value problem of physical geodesy. Helsinki, Publ. Isostat. Inst. Int. Assoc. Geod., No. 50.

Pellinen, L. P. (1962). Accounting for topography in the calculation of quasigeoidal heights and plumb-line deflections from gravity anomalies. Bull. Géod., No. 63, pp. 57–65.

9
Celestial Methods

9-1. Introduction. Observational Methods

The subject of this chapter is the use of observations of the motion of the moon and of artificial satellites, and also of the precession of the earth's axis, for determining features of the gravity field and figure of the earth. Only the barest essentials can be presented within the scope of a chapter. The reader will find more information in special textbooks such as Berroth and Hofmann (1960) and Mueller (1964), in review articles such as Kaula (1962) and Cook (1963), and in the published proceedings of the two international symposia on the use of artificial satellites for geodesy (Veis, 1963, 1966). These publications also give extensive reference to earlier papers.

The observational methods are intended to determine the spatial direction and the distance to the satellite. *Directions* may be measured visually by means of theodolites, photographically by photographing the satellite against the background of stars, or by means of radio waves transmitted from the satellite, using the principle of interference. Photography can achieve an accuracy of about one second of arc, whereas the accuracy of the other methods is less by at least one order of magnitude. The principle of the photographic method is this. On the photographic plate the image of the satellite is surrounded by images of stars. The directions to the surrounding stars are defined by their right ascensions and declinations (see Sec. 9-8), which are known from spherical astronomy. By interpolation we find the right ascension of the direction to the artificial satellite, or to the center of the moon. A modification is the method of occulation: at the moment at which a point on the moon's limb covers a star, the direction to this point is equal to the known direction to the star.

Distances are measured by radar or by means of modulated light ("optical radar"). The time rate of change of the distance is found by observing the Doppler effect with radio waves transmitted from an artificial satellite.

Radio methods (interferometer, radar, doppler) are very convenient for tracking artificial satellites, since they are independent of weather and time of day, on which the photographic methods strongly depend. Radar and doppler are capable of geodetic accuracy, whereas precise determination of directions requires photographic techniques.

The precession of the earth's axis is determined by the classical methods of spherical astronomy, by observing the change of right ascension and declination of the stars with time.

9-2. Determination of the Size of the Earth from Observations of the Moon

Although the geodetic application of the artificial satellites is more important, we shall start with the use of our natural satellite, the moon. The reasons are partly historical (much interesting effort was spent on this problem long before the artificial satellites were launched) and partly didactic. Since the moon is very far away from the earth, the effect of the irregularities of the earth, apart from the flattening, can be neglected. Therefore, because of their simplicity, the methods using the moon provide an excellent introduction to the more complex techniques using artificial satellites.

Geometric use of the moon. Figure 9-1 shows the principle. The astronomical determination of the position of the moon with respect to the stars at two consecutive times t_1 and t_2 furnishes two spatial directions PM_1 and PM_2. Since the orbit of the moon is very accurately known as a function of time, having been tabulated in Lunar Ephemerides, we know the positions M_1 and M_2. Then the spatial position of P is found by a resection from M_1 and M_2. One may use either photography of the moon against the background of stars (method of W. Markowitz) or occultation of stars by the moon (method of J. A. O'Keefe). Difficulties arise from the finite size and slightly irregular shape of the moon.

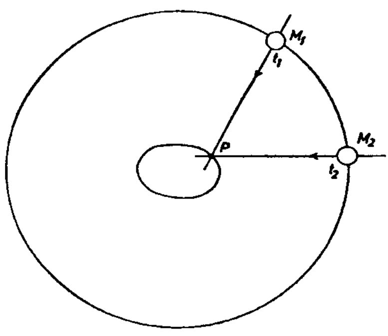

FIGURE 9-1

Determination of a terrestrial point P by resection from two positions of the moon, M_1 and M_2.

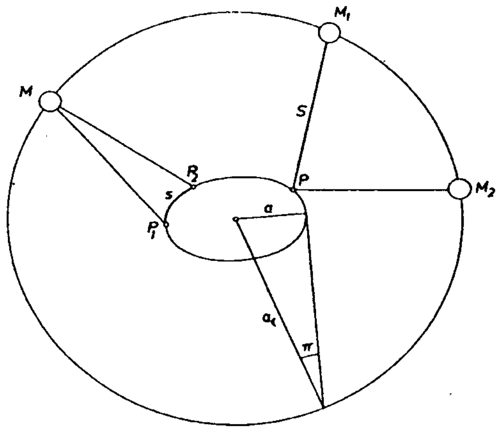

FIGURE 9-2

Determination of the scale of the earth-moon system.

Geometric determination of the equatorial radius. The moon's orbit is known with extreme precision, but its scale is less well determined. The radius of the moon's orbit (which may be considered circular in this context) is not too accurately known either in relation to the size of the earth or in metrical units. The problem is thus first to establish the relation between the radius $a_☾$ of the moon's orbit and the earth's equatorial radius a. By measuring one distance we shall then know both a and $a_☾$.

Let us for the moment consider the earth as an exact ellipsoid of revolution of known flattening f. According to Fig. 9-2, the relation between a and $a_☾$ may be established either by intersecting an ellipsoidal point P of known latitude and longitude from two spatial positions M_1 and M_2 of the moon, or by intersecting one position M of the moon from two ellipsoidal points P_1 and P_2 whose latitude and longitude are known. Both methods were applied, the first by J. A. O'Keefe around 1952, and the second by A. C. D. Crommelin around 1910.

So far no distance measurement has been considered. Figure 9-2 is therefore determined only apart from the scale. What is known is the ratio $a{:}a_☾$. This ratio may be interpreted geometrically by the angle π, the *equatorial horizontal parallax*. From Fig. 9-2 we read

$$\sin \pi = \frac{a}{a_☾}. \tag{9-1}$$

In order to find both a and $a_☾$, it is sufficient to determine the scale of the figure by measuring *one* distance—for instance, the arc $s = P_1P_2$ or the straight distance $S = PM_1$. The terrestrial distance s may be obtained by triangulation, etc.; S has been measured recently by B. S. Yaplee, using radar.

This is the principle of the method. It is easy to generalize it to the actual case in which the earth is not an exact ellipsoid. This is done by applying small corrections. The geodetic coordinates ϕ, λ, with respect to a certain reference ellipsoid, of a point P on the earth's surface are related to its astronomical coordinates Φ, Λ by (5-17):

$$\phi = \Phi - \xi,$$
$$\lambda = \Lambda - \eta \sec \phi.$$

By plotting the height above the ellipsoid,

$$h = H + N,$$

downward along the ellipsoidal normal, starting from P, we get a point P_0 that really lies on the ellipsoid and whose ellipsoidal latitude ϕ and longitude λ are known (Fig. 9-3). Thus we have reduced our actual problem to the above simplified one.

Dynamic determination of mass and equatorial radius. The basic relation is Kepler's third law for planetary motion, applied to the moon:

$$n^2 a_{☾}^3 = k(M + M_{☾}), \tag{9-2}$$

where n is the mean angular velocity of the moon's revolution, k is the gravitational constant, and M and $M_{☾}$ are the masses of earth and moon. Introducing by

$$n = \frac{2\pi}{P},$$

the period P of the moon's revolution (the time it takes to complete one revolution around the earth), we obtain the more familiar form

$$\frac{P^2}{a_{☾}^3} = \frac{4\pi^2}{k(M + M_{☾})}, \tag{9-2'}$$

relating the square of the period and the cube of the semimajor axis of the orbit.

Introducing the ratio

$$\mu = \frac{M_{☾}}{M}, \tag{9-3}$$

which is about $1/80$, we may write (9-2) as

$$kM = \frac{n^2}{1 + \mu} a_{☾}^3. \tag{9-4}$$

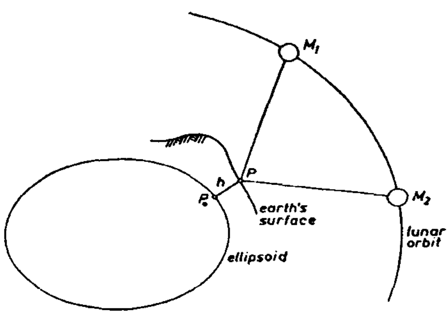

FIGURE 9-3
Reduction to the ellipsoid.

By measuring radar distances to the moon, and by means of an approximate value for a, we may compute the orbital radius $a_☾$. Then we obtain kM by the above equation.

We can also express kM in terms of equatorial gravity γ_a by using (2-111), which may be written

$$kM = a^2\gamma_a(1 - f + \tfrac{3}{2}m \cdots). \tag{9-5}$$

Here we have introduced $b = a(1 - f)$ and neglected terms of the second and higher order.

The basic relations (9-1), (9-4), and (9-5) may be combined and applied in various ways. Solving (9-5) for a we obtain

$$a = \sqrt{\frac{kM}{\gamma_a}}\,(1 + \tfrac{1}{2}f - \tfrac{3}{4}m \cdots). \tag{9-6}$$

This formula gives the equatorial radius in terms of the celestially obtained constant kM and of equatorial gravity. Thus a is obtained *without terrestrial distance measurements.*

By expressing $a_☾$ in terms of a and π by (9-1) we may write (9-4) as

$$kM = \frac{n^2}{1 + \mu}\left(\frac{a}{\sin \pi}\right)^3.$$

Equating the right-hand sides of this equation and of (9-5) and solving for a we find

$$a = \frac{1 + \mu}{n^2}\sin^3 \pi\, \gamma_a(1 - f + \tfrac{3}{2}m \cdots). \tag{9-7}$$

This formula gives a in terms of the parallax π and of equatorial gravity. Here distance measurement does not enter at all, not even implicitly as in the application of (9-6), where the celestial determination of kM by means of (9-4) involves measurements of extraterrestrial distances.

It is also possible to solve (9-7) for $\sin \pi$ with the result

$$\sin \pi = \sqrt[3]{\frac{n^2 a}{\gamma_a(1 + \mu)}}\,(1 + \tfrac{1}{3}f - \tfrac{1}{2}m \cdots), \tag{9-8}$$

expressing the angle π in terms of a and γ_a. When calculated in this way, π is called *dynamic parallax.*

We thus have a connection between the parallax π, the equatorial radius a, and equatorial gravity γ_a, which must be taken into account if we wish to obtain a consistent system of geodetic and astronomical constants. A study by Fischer (1962) shows that there are discrepancies in this respect. This is to be expected in view of the scarcity of our observational data.

9-3. Dynamic Effects of the Earth's Flattening

If the earth were a homogeneous sphere, then the orbit of a satellite such as the moon would be an exact Keplerian ellipse. The oblateness of the earth, however, gives rise to a slow *variation of the orbital plane* with time, which is accompanied by a corresponding variation of the direction of the earth's axis in space, the *precession*. Both effects will now be derived jointly.

Figure 9-4 shows the intersection of the equatorial plane and the orbital plane with a unit sphere centered at the earth's center of mass. The vector **L** of angular momentum of the earth's rotation coincides with the earth's axis of rotation (to within 0.3″), thus being normal to the equatorial plane. Its magnitude is

$$L = C\omega,$$

where C is the moment of inertia of the earth about its axis of rotation, and ω is its angular velocity. The vector $\mathbf{L}_☾$ of angular momentum of the moon's revolution in its orbit is normal to the orbital plane. Its magnitude is

$$L_☾ = M_☾ a_☾^2 \cdot n,$$

because $M_☾ a_☾^2$ is the moment of inertia of the moon's orbital motion, and n is the mean angular velocity of this motion. The small ellipticity of the moon's orbit is again neglected.

Because of the oblateness of the earth there are forces that tend to make the orbital and the equatorial planes coincide. It is a well-known peculiarity of rotational motion that the rotating systems cannot follow this tendency, so that the effect is quite different. For reasons of symmetry, the average resultant of these forces lies in the plane defined by the vectors **L** and $\mathbf{L}_☾$. The moment **N** of the resulting force is therefore normal to this plane. Its magnitude is given by

$$N = \frac{3kM_☾}{2a_☾^3}(C - \bar{A}) \sin i \cos i, \tag{9-9}$$

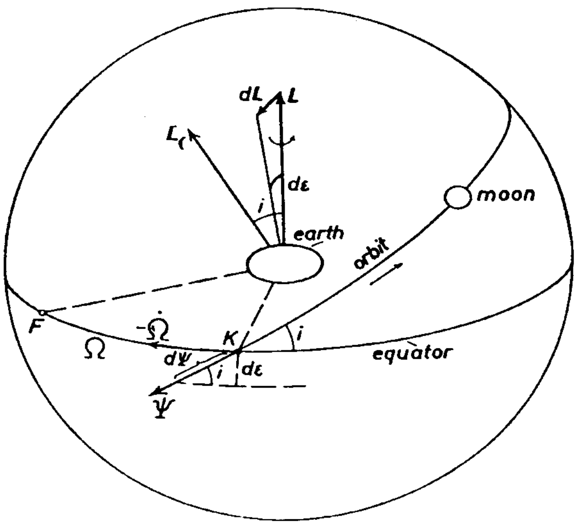

FIGURE 9-4

Variation of the earth's axis and the orbit because of the oblateness of the earth.

where $\bar{A}$ is the mean equatorial moment of inertia, defined by (2-199), and i is the inclination of the orbit (see Fig. 9-4). The derivation of this formula may be found in (Thomson, 1961, pp. 94–97).[1]

According to a well-known law of mechanics, the moment **N** is the derivative of the angular momentum **L** with respect to time,

$$\mathbf{N} = \frac{d\mathbf{L}}{dt}.$$

Hence we have

$$dL = N\,dt.$$

The vector $d\mathbf{L}$, being parallel to **N**, is normal to the vector **L**, and we read from Fig. 9-4

$$dL = L\,d\epsilon.$$

Comparing the last two equations we find

$$\frac{d\epsilon}{dt} = \frac{N}{L}. \tag{9-10}$$

Rotating the earth's axis by $d\epsilon$ is equivalent to tilting the equatorial plane by the same angle. This gives rise to a precession of the node K along the orbit by the amount (see Fig. 9-4)

$$d\Psi = \frac{d\epsilon}{\sin i}.$$

Hence we get, using (9-10),

$$\dot{\Psi} = \frac{N}{L \sin i}, \tag{9-11a}$$

where the dot is the usual symbol for the temporal derivative, $\dot{\Psi} = d\Psi/dt$. This formula gives the *precession of the earth's axis*, measured along the moon's orbit.

It follows by analogy that a corresponding formula holds for the *regression of the orbital plane*, as expressed by the temporal variation of the angle $\Omega = FK$ measured from a fixed point F on the equator. Replacing L by $L_☾$ in (9-11a) we thus obtain

$$-\dot{\Omega} = \frac{N}{L_☾ \sin i}, \tag{9-11b}$$

the minus sign indicating the retrogressive motion of the node K.

Inserting the above expressions for L, $L_☾$, and N into (9-11a, b) and taking (9-4) into account we find

$$\dot{\Psi} = \frac{3n^2}{2\omega}\frac{\mu}{1+\mu}\frac{C-\bar{A}}{C}\cos i, \tag{9-12a}$$

$$-\dot{\Omega} = \frac{3n}{2}\frac{1}{1+\mu}\left(\frac{a}{a_☾}\right)^2\frac{C-\bar{A}}{Ma^2}\cos i. \tag{9-12b}$$

[1] The relevant formula there is (4.18-7); since we are concerned with the average moment N, we have replaced $\sin^2\phi$ by its average value 1/2.

Although the original expressions (9-11a) and (9-11b) are closely similar, the effect of the earth's oblateness as expressed by the inequality of the moments of inertia C and $\bar{A}$ is quite different on $\dot{\Psi}$ and $\dot{\Omega}$. The precession of the earth's axis, $\dot{\Psi}$, is proportional to the ratio

$$H = \frac{C - \bar{A}}{C}, \tag{9-13}$$

which is called *mechanical ellipticity*. The regression of the orbital plane, $-\dot{\Omega}$, is proportional to the quantity

$$J_2 = \frac{C - \bar{A}}{Ma^2}, \tag{9-14}$$

which is nothing but the well-known spherical-harmonic coefficient (2-198).

The development of the external gravitational potential in spherical harmonics does not explicitly depend on the internal structure of the earth, since many different configurations of mass may give rise to the same external potential (Sec. 1-7). The mechanical ellipticity H, however, depends essentially on the distribution of mass inside the earth. Hence it is evident that the geometric flattening f of the mean earth ellipsoid is directly related to the coefficient J_2 (see Sec. 2-21), whereas the connection between f and H is more complicated and indirect, involving the mass distribution.

Because of these relations it is possible to determine f from both J_2 and H, but the determination from J_2 is more direct and therefore preferable. Unfortunately the accuracy of J_2 as obtained from the regression of the moon's node is rather poor (see also Sec. 9-6). A precise determination of J_2 was possible only through the use of artificial satellites. The reason is seen on inspecting equation (9-12b). The regression of the orbital plane, $-\dot{\Omega}$, contains the factor $(a/a_☾)^2$. It is therefore largest for close artificial satellites, where

$$\frac{a}{a_☾} \doteq 1$$

($a_☾$ being here the radius of the satellite's orbit), and very small for the moon, where

$$\frac{a}{a_☾} \doteq \frac{1}{60}.$$

Matters are just the opposite with the precession, equation (9-12a). It is evident that the motion of an artificial satellite cannot noticeably affect the direction of the earth's axis. This is expressed in equation (9-12a) by the fact that the ratio μ, the mass of the satellite to the mass of the earth, is negligibly small. For the moon, however, this mass ratio is appreciable, $\mu \doteq \frac{1}{80}$. The precessional effect of the moon is combined with the effect of the sun, which also depends on H through a formula analogous to (9-12a). The influence of the sun is about half of that of the moon. This "luni-solar precession," and hence H, are known very accurately. The determination of f from H will be the subject of the next section.

In equations (9-12a, b) only the mutual influence of earth and moon was taken into account. This is a rather unrealistic simplification, because the sun, whose dominant attraction affects the moon's orbit even more than the earth's oblateness, was left out of consideration. The interactions between earth, sun, and moon make an accurate analysis exceedingly difficult, but our simplified approach illustrates the principles to a sufficient extent.

9-4. Determination of the Flattening from Precession. Hydrostatic Equilibrium

We have seen in the preceding section that observing the rate of precession of the earth's axis yields the mechanical ellipticity

$$H = \frac{C - A}{C}.$$

The geometrical flattening f of the mean earth ellipsoid can be determined from H as well as from J_2. There is, however, an essential difference. The determination of f from J_2 is straightforward, and no assumptions about the earth's internal structure are needed; see Sec. 2-21. If we wish to determine f from H, however, we must assume that the earth is approximately in hydrostatic or fluid equilibrium, which is only superficially disturbed by the solid crust.

The theory of spheroidal figures of fluid equilibrium is governed by the famous *differential equation of Clairaut*, published in 1743:

$$\frac{d^2f}{dr^2} + \frac{2\rho r^2}{\int_0^r \rho r^2\, dr}\frac{df}{dr} + \left(\frac{2\rho r}{\int_0^r \rho r^2\, dr} - \frac{6}{r^2}\right)f = 0. \tag{9-15}$$

This second-order linear differential equation connects the flattening f of any level surface of the earth with its mean radius r and with the density ρ.

The solution of Clairaut's equation yields a relation between J_2 and H:

$$J_2 = \frac{2}{3}\left(1 - \frac{2}{5}\sqrt{\frac{5m}{2f} - 1}\right)H, \tag{9-16}$$

where m is defined by equation (2-100). Equation (2-118) becomes, on neglecting second-order terms,

$$J_2 = \tfrac{2}{3}(f - \tfrac{1}{2}m). \tag{9-17}$$

By equating the right-hand sides of these two equations we obtain

$$H = \frac{f - \frac{1}{2}m}{1 - \frac{2}{5}\sqrt{\frac{5m}{2f} - 1}}, \tag{9-18}$$

or in a slightly more practical form,

$$\frac{1}{H} = \frac{1}{f}\,\frac{1 - \frac{2}{5}\sqrt{\frac{5m}{2f} - 1}}{1 - \frac{m}{2f}}. \tag{9-18'}$$

Since H and m are given, we must solve (9-18) or (9-18′) for f. The simplest way is to compute by (9-18′) the values of $1/H$ corresponding to a set of assumed values of f, and to interpolate f for the given value of $1/H$.

Derivations of (9-15) and (9-16) may be found in Jeffreys (1962, Sec. 4.03).

Discussion and numerical results. Equation (9-16) does not explicitly contain the density ρ inside the earth; it connects surface quantities only. Like (9-17) it is valid to a first approximation: terms of the second degree and higher in f are neglected. In the higher approximations of (9-16) we must consider zonal harmonics of higher degrees and their squares and cross-products; furthermore, the density distribution enters more directly; see Wavre (1932) or Kopal (1960).

Thus for computing f from H we need only presuppose the earth to be in approximate hydrostatic equilibrium; the result is almost independent of what density distribution we assume inside the earth. This is what makes the geodetic application of the theory of equilibrium figures possible, since we do not know the internal densities very reliably.

In 1948 E. C. Bullard made a careful determination of the flattening by this method. His starting value was

$$\frac{1}{H} = 305.59.$$

Using the best density distribution known at this date and taking second-order corrections into account he found

$$\frac{1}{f} = 297.34. \tag{9-19}$$

Observations of artificial satellites, however, have shown quite conclusively that f^{-1} lies between 298.2 and 298.3. This result is derived from determinations of J_2 by orbital analysis (Sec. 9-6) and is independent of any assumption about hydrostatic equilibrium.

The question whether discrepancies between hydrostatically determined values of f and those obtained by other methods point to a significant deviation of the earth's interior from hydrostatic equilibrium is an old one.[1] Recently the discussion of this point was renewed because at present the accuracies of the precessional determination of H and the satellite determination of J_2 are equally high, so that an agreement of f^{-1} as obtained from H and J_2 to about 0.1 could be expected (Ledersteger, 1966).

9-5. Orbits of Artificial Satellites

In Sec. 9-3 we have analyzed the effect of the earth's flattening on a satellite such as the moon. We have found that it produces a slow rotation of the orbital

[1] As early as about 1885 Poincaré recognized that the results of precession under the assumption of hydrostatic equilibrium were incompatible with the value $f = 1/293$, which was current at that time.

plane around the earth's axis, the inclination remaining constant on the average.

The earth's flattening causes the largest but not the only deviation of the gravitational field of the earth from that of a homogeneous sphere. Generally the gravitational potential can be expanded into a series of spherical harmonics according to Sec. 2-5, equation (2-39):

$$V = \frac{kM}{r}\left\{1 - \sum_{n=2}^{\infty}\left(\frac{a}{r}\right)^n J_n P_n(\cos\theta) - \sum_{n=2}^{\infty}\sum_{m=1}^{n}\left(\frac{a}{r}\right)^n (J_{nm}\cos m\lambda + K_{nm}\sin m\lambda)P_{nm}(\cos\theta)\right\}. \quad (9\text{-}20)$$

Here the terms containing J_n are the zonal harmonics, and those containing J_{nm} and K_{nm} the tesseral harmonics.

Since the moon is far away, the only term of appreciable influence is J_2, which represents the flattening. Artificial satellites are much closer to the earth, a typical height above ground of a geodetically used satellite being 1000 km. Hence they are also influenced by harmonics other than J_2 and can therefore be used to determine harmonics of low degree. For this purpose we must study the effect of gravitational disturbances on the orbits of close satellites.

Before we can do this we must briefly review the theory of an undisturbed orbit. By this we mean that the gravitational potential has the form

$$V = \frac{kM}{r},$$

all J's and K's being zero. This represents the gravitational field of a point mass or a homogeneous sphere. Then the motion of a satellite is described by Kepler's three laws for planetary motion.[1]

According to *Kepler's first law*, the orbit is an ellipse of which the center of the earth occupies one focus. The position of the orbit in space is defined by the six *orbital elements:*

a	semimajor axis,
e	eccentricity,
i	inclination,
Ω	right ascension of the node,
ω	argument of perigee,
T	time of perigee passage.

If a and b are the semiaxes of the orbital ellipse (not to be confused with those of the terrestrial ellipsoid), then the eccentricity is defined by

$$e = \frac{\sqrt{a^2 - b^2}}{a}. \quad (9\text{-}21)$$

Figure 9-5 shows the projection of the orbit onto a geocentric unit sphere. The

[1] Satellites with parabolic or hyperbolic orbits are of no interest in this context.

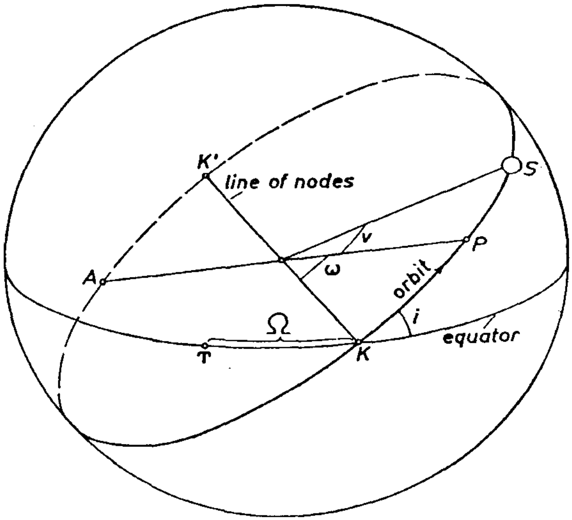

FIGURE 9-5

The satellite orbit as projected onto a unit sphere. P = perigee, A = apogee, K = ascending node, K' = descending node, S = instantaneous position of satellite.

line of nodes is the intersection of the orbital plane with the plane of the equator; it connects the ascending and the descending node. The right ascension of the node, Ω, is the angle between the line of nodes and the direction to the vernal equinox ♈.[1] The major axis of the orbit intersects the orbital ellipse at the perigee, the position where the satellite is closest to the earth, and at the apogee, where the satellite is farthest away. The angle ω between the nodes and the major axis is the argument of perigee.

The angular distance of the satellite S from perigee is called *true anomaly* and denoted by v; it is a function of time. The equation of the orbital ellipse may be written

$$r = \frac{p}{1 + e \cos v}, \tag{9-22}$$

where r is the distance of the satellite from the earth's center of mass and

$$p = \frac{b^2}{a} = a(1 - e^2) \tag{9-23}$$

is the length of the radius vector r for $v = 90°$. The radius vector r and the true anomaly v form a pair of polar coordinates, and (9-22) is the well-known polar equation of an ellipse. See Fig. 9-6 for an illustration of these quantities.

According to *Kepler's second law*, the area of the elliptical sector swept by the radius vector r between any two positions of the satellite is proportional to

[1] The symbol Ω is also called longitude of the node, but in conformity with astronomical terminology it is the right ascension of the (ascending) node. For the definition of the vernal equinox and of right ascension see Secs. 9-7 and 9-8.

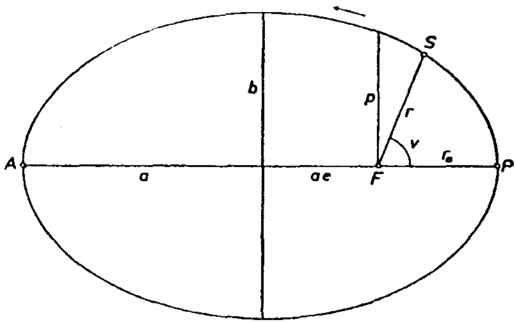

FIGURE 9-6

The orbital ellipse. P = perigee, A = apogee, F = earth's center of mass, S = instantaneous position of satellite.

the time it takes the satellite to pass from one position to the other. In other words, the time rate of change of the area swept by the radius vector is constant. Since the element of area of a sector in polar coordinates r and v is $\frac{1}{2}r^2\,dv$, this law may be formulated mathematically as

$$r^2 \frac{dv}{dt} = \sqrt{kMa(1 - e^2)}, \tag{9-24}$$

where the constant has already been given its proper value.

Kepler's third law has been anticipated by (9-2). Since the mass of the satellite is negligibly small, we now have

$$n^2a^3 = kM, \tag{9-25}$$

where

$$n = \frac{2\pi}{P}$$

is the "mean motion" (mean angular velocity) of the satellite, P being its period.

So far we have assumed that all J_n, J_{nm}, and K_{nm} in (9-20) are zero. Because of the irregularities of the earth's gravitational field this is not true, even though these coefficients are small. Therefore, the satellite is subject to small disturbing forces. We may still consider the satellite orbit as an ellipse, but then the parameters of this ellipse, the orbital elements, will no longer be constant but will change slowly. At each instant this so-called *osculating ellipse* will be slightly different. It is defined as follows. Imagine that at the instant under consideration all disturbing forces suddenly vanish. Then the satellite will continue its motion along an exact ellipse; this is the osculating ellipse.

If we resolve the total disturbing force into rectangular components S, T, and W—S being directed along the radius vector, W being normal to the orbital plane, and T being normal to S and W[1]—then the time rate of change of the orbital parameters can be expressed in terms of these components:

[1] This notation follows astronomical usage; there is no relation to the geodetic use of T and W for potentials.

$$\dot{a} = \frac{2a^2}{b}\sqrt{\frac{a}{kM}}\left(eS\sin v + \frac{p}{r}T\right),$$

$$\dot{e} = \frac{b}{a}\sqrt{\frac{a}{kM}}\left[S\sin v + \left(\frac{r+p}{p}\cos v + \frac{er}{p}\right)T\right],$$

$$\frac{di}{dt} = \frac{r}{b}\sqrt{\frac{a}{kM}}\,W\cos(\omega + v), \tag{9-26}$$

$$\dot{\Omega} = \frac{r}{b}\sqrt{\frac{a}{kM}}\,W\frac{\sin(\omega + v)}{\sin i},$$

$$\dot{\omega} = \frac{b}{a}\sqrt{\frac{a}{kM}}\left[-\frac{1}{e}S\cos v + \frac{r+p}{ep}T\sin v - \frac{r}{p}W\sin(\omega + v)\cot i\right].$$

As usual, $\dot{a}$ denotes da/dt, etc. The derivation of these equations may be found in any textbook on celestial mechanics—for instance, Brouwer and Clemence (1961, p. 301) or Plummer (1918, p. 151).

9-6. Determination of Zonal Harmonics

The effect of the zonal harmonics on satellite orbits is much greater than that of the tesseral harmonics. Only zonal harmonics ($J_2, J_3, J_4, \ldots$) will give observable variations of the orbital elements themselves. The tesseral harmonics cause oscillatory disturbances that rapidly change their sign, whereas the effect of the zonal harmonics is cumulative. For this reason we shall consider first the effect of zonal harmonics, that is, the effect of those independent of longitude λ.

Hence we set

$$V = \frac{kM}{r} + R, \tag{9-27}$$

where the *perturbing potential*

$$R = -\frac{kM}{a_e}\sum_{n=2}^{\infty}\left(\frac{a_e}{r}\right)^{n+1} J_n P_n(\cos\theta) \tag{9-28}$$

is a function of r and θ only.[1,2]

Note that the equatorial radius of the earth (the semimajor axis of the terrestrial ellipsoid) has been denoted by a_e, in order to distinguish it from a which now denotes the semimajor axis of the orbital ellipse. This notation will be used in what follows.

[1] The main difference between the perturbing potential R of celestial mechanics and the disturbing potential T of physical geodesy is that R, but not T, also incorporates the effect of the flattening through J_2.

[2] There are also other perturbing forces acting on a satellite, such as the resistance of the atmosphere (atmospheric drag), radiation pressure exerted by the sunlight, etc. These nongravitational perturbances must be taken into account separately and will not be considered here.

Since S is the component of the perturbing force along the radius vector, we have

$$S = \frac{\partial R}{\partial r}. \tag{9-29a}$$

The components of the perturbing force along the meridian and the prime vertical are

$$-\frac{1}{r}\frac{\partial R}{\partial \theta} \quad \text{and} \quad \frac{1}{r \sin\theta}\frac{\partial R}{\partial \lambda}.$$

The components T and W are obtained from these by a plane rotation (Fig. 9-7):

$$T = -\frac{1}{r}\frac{\partial R}{\partial \theta}\cos\alpha + \frac{1}{r\sin\theta}\frac{\partial R}{\partial \lambda}\sin\alpha,$$

$$W = -\frac{1}{r}\frac{\partial R}{\partial \theta}\sin\alpha - \frac{1}{r\sin\theta}\frac{\partial R}{\partial \lambda}\cos\alpha.$$

From the rectangular spherical triangle ABC it follows that

$$\cos\alpha = \frac{\cos(\omega + v)\sin i}{\sin\theta}, \quad \sin\alpha = \frac{\cos i}{\sin\theta},$$

so that finally

$$T = -\frac{\cos(\omega + v)\sin i}{r\sin\theta}\frac{\partial R}{\partial \theta} + \frac{\cos i}{r\sin^2\theta}\frac{\partial R}{\partial \lambda}, \tag{9-29b}$$

$$W = -\frac{\cos i}{r\sin\theta}\frac{\partial R}{\partial \theta} - \frac{\cos(\omega + v)\sin i}{r\sin^2\theta}\frac{\partial R}{\partial \lambda}. \tag{9-29c}$$

We have included $\partial R/\partial\lambda$ because of the presence of longitude-dependent

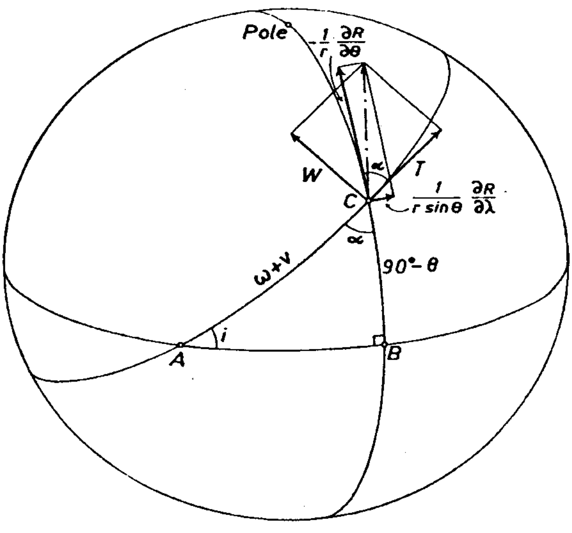

FIGURE 9-7

Components of the perturbing force. A = node, C = satellite.

tesseral harmonics in the general case; see Sec. 9–8. In our present case, where R is given by (9-28), $\partial R/\partial\lambda$ is zero.

Now we must differentiate (9-28) with respect to r and θ, compute the components S, T, W from equations (9-29), and insert them into the system (9-26). In this way we can express the rates of change $\dot{a}$, $\dot{e}$, . . . of the orbital elements in terms of the coefficients J_2, J_3, J_4, ...

We cannot, however, observe these rates of change directly. Rather, we observe the changes of the orbital elements after several revolutions. The changes after *one* revolution, with period P, are

$$\Delta a = \int_{t_0}^{t_0+P} \dot{a}\,dt, \quad \Delta e = \int_{t_0}^{t_0+P} \dot{e}\,dt, \quad \Delta i = \int_{t_0}^{t_0+P} \frac{di}{dt}\,dt, \quad \text{etc.}$$

The t_0 is an arbitrary "epoch" (instant of time). In order to perform these integrations, we must express $\dot{a}$, $\dot{e}$, . . . in terms of one independent variable. For this independent variable we may take the time t or the true anomaly v. The second possibility will be adopted here.

The polar distance θ is expressed as a function of v through the relation

$$\cos\theta = \sin(\omega + v)\sin i, \tag{9-30}$$

which follows from the spherical triangle ABC of Fig. 9-7. The radius vector r is also a function of v according to (9-22). Finally, Kepler's second law (9-24) furnishes the relation between v and the time t:

$$\frac{dt}{dv} = \frac{r^2}{\sqrt{kMa(1-e^2)}}.$$

Hence we may change the integration variable from t to v, obtaining, for instance,

$$\Delta a = \int_{t=t_0}^{t_0+P} \dot{a}\,dt = \int_{v=0}^{2\pi} \frac{da}{dv}\,dv,$$

where

$$\frac{da}{dv} = \frac{da}{dt}\frac{dt}{dv} = \frac{r^2}{\sqrt{kMa(1-e^2)}}\,\dot{a}.$$

Analogous formulas hold for the other orbital elements.

After performing all these operations, which are lengthy but not too difficult, we find

$$\Delta a = 0;$$

$$\Delta e = -\frac{1-e^2}{e}\tan i\cdot\Delta i;$$

$$\begin{aligned}\Delta i = {} & 3\pi e\left(\frac{a_e}{p}\right)^3\left(1 - \frac{5}{4}\sin^2 i\right)\cos i\cos\omega\cdot J_3 \\ & + \frac{45}{16}\pi e\left(\frac{a_e}{p}\right)^4\left(1 - \frac{7}{6}\sin^2 i\right)\sin 2i\sin 2\omega\cdot eJ_4\cdots;\end{aligned}$$

$$\Delta\Omega = -3\pi\left(\frac{a_e}{p}\right)^2 \cos i \cdot J_2 + 3\pi\left(\frac{a_e}{p}\right)^3\left(1 - \frac{15}{4}\sin^2 i\right)\cot i \sin\omega \cdot eJ_3 + \frac{15}{2}\pi\left(\frac{a_e}{p}\right)^4\left(1 - \frac{7}{4}\sin^2 i\right)\cos i \cdot J_4 \cdots;$$

$$\Delta\omega = 6\pi\left(\frac{a_e}{p}\right)^2\left(1 - \frac{5}{4}\sin^2 i\right)\cdot J_2 + 3\pi\left(\frac{a_e}{p}\right)^3\left(1 - \frac{5}{4}\sin^2 i\right)\sin i \sin\omega \cdot eJ_3 - 15\pi\left(\frac{a_e}{p}\right)^4\left[\left(1 - \frac{31}{8}\sin^2 i + \frac{49}{16}\sin^4 i\right) + \left(\frac{3}{8} - \frac{7}{16}\sin^2 i\right)\sin^2 i \cos 2\omega\right]\cdot J_4 \cdots. \tag{9-31}$$

Terms of the order of e^2J_3 and e^2J_4, which are very small, have been neglected in these equations. The proportionality of Δe and Δi is more or less accidental: it holds only with respect to long-periodic disturbances; $\dot{e}$ and di/dt themselves are not proportional. The quantity p is defined by (9-23); it is hardly necessary to repeat that a, p, e, etc. refer to the orbital ellipse and not to the terrestrial ellipsoid, of which a_e is the equatorial radius.

By integrating over one revolution we have removed the *short-periodic* terms of periods P, $2P$, $3P$, . . . , such as $\cos v$, $\cos 2v$, etc. What remains are *secular* terms, which are constant for one revolution and increase steadily with the number of revolutions, and the *long-periodic* terms, which change very slowly with time in a periodic manner. The argument of perigee ω increases slowly but steadily, so that the perigee of a satellite orbit also rotates around the earth, but much slower than the satellite itself; a typical period of ω is two months. Therefore, terms containing $\cos\omega$, $\sin\omega$, or $\sin 2\omega$ are called long-periodic.

The first equation of (9-31) shows that the semimajor axis of the orbit does not change secularly or long-periodically. The eccentricity and the inclination undergo long-period, but not secular, variations, whereas Ω and ω change both secularly and long-periodically.

The equations (9-31) are linear in $J_2, J_3, J_4, \ldots$. For practical applications nonlinear terms containing J_2^2, J_2J_3, J_2J_4, etc. must also be taken into account, since J_2^2 is of the order of J_4. The derivation of these nonlinear terms is much more difficult, and their expressions are different in the various orbital theories that have been proposed. For these reasons such expressions will not be given here,

Equations (9-31), supplemented by certain nonlinear terms, can be used to determine coefficients J_2, J_3, J_4, etc. The secular or long-periodic variations $\Delta\Omega$, $\Delta\omega$, Δe, Δi being known from observation for a sufficient number of satellites, we obtain equations of the form

$$\begin{aligned} a_2J_2 + a_3J_3 + a_4J_4 + \cdots + a_{22}J_2^2 + a_{23}J_2J_3 + \cdots &= A, \\ b_2J_2 + b_3J_3 + b_4J_4 + \cdots + b_{22}J_2^2 + b_{23}J_2J_3 + \cdots &= B, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned} \tag{9-32}$$

which can be solved for J_2, J_3, J_4, ... Since there can be only a finite number of these equations, we must neglect all J_n with n greater than a certain number n_0 which depends on the number of equations available, on their degree of mutual independence, etc. This, of course, is a difficulty with this method.

From (9-31) it is seen that the coefficients of the J_n depend essentially on the inclination i. It is therefore important to use satellites with a wide variety of inclinations, in order to get equations with a high mutual independence.

Now the question arises as to which orbital elements are to be used for determining the coefficients J_n. The semimajor axis a clearly cannot be used at all. As for the other elements, we must distinguish between coefficients of even and of odd degree n. The even coefficients J_2, J_4, . . . can be well determined from the regression of the node, $\Delta\Omega$, and the rotation of perigee, $\Delta\omega$. This is seen on inspecting (9-31). The even harmonics cause secular disturbances of Ω and ω, which are much larger than the long-periodic effects of the odd coefficients, since J_3, J_5, . . . are multiplied by the small eccentricity e.

On the other hand, in Δe and Δi the odd coefficients J_3, J_5, . . . have a much larger effect than the even coefficients, which here appear with the small factor e. Therefore, the odd coefficients are determined from Δe or Δi, or from the change of perigee distance $r_0 = FP$ (Fig. 9-6). Since r_0 is the radius vector for $v = 0$, we have from (9-22) and (9-23)

$$r_0 = \frac{p}{1 + e} = a(1 - e),$$

so that

$$\Delta r_0 = -a\Delta e$$

because $\Delta a = 0$. Thus the variation of perigee distance is proportional to the variation of eccentricity and may be used instead of Δe.

Numerical values. Helmert (1884, p. 472) used the regression of the node of the moon's orbit to determine J_2, which is the only coefficient to have an appreciable effect on it.[1] He found

$$J_2 = 1086.5 \times 10^{-6}$$

by averaging two widely different values. This corresponds to a flattening of

$$\frac{1}{f} = 297.8 \pm 2.2.$$

[1] Note that for $e \doteq 0$ and $p \doteq a \gg a_e$, the equation for $\Delta\Omega$ in (9-31) becomes

$$\Delta\Omega = -3\pi \left(\frac{a_e}{a}\right)^2 J_2 \cos i,$$

which, apart from a slightly different notation and other minor differences, is the same as (9-12b), derived in a completely different way.

This value is quite close to the recent results, but has a much larger uncertainty.

Reliable values by this method can only be obtained from close artificial satellites. Currently accepted values are, for example,

$$\begin{aligned} J_2 &= 1082.6 \times 10^{-6}, \\ J_3 &= -2.5 \times 10^{-6}, \\ J_4 &= -1.6 \times 10^{-6}, \end{aligned} \tag{9-33}$$

whose standard errors are assumed to be around 0.2×10^{-6} or better; see the review paper by Kozai in Veis (1966).

The most significant geodetic result is the reliable determination of J_2, and hence of the flattening f. At present $1/f$ is believed to have a value between 298.2 and 298.3 with a standard error of better than 0.1; in 1964 the International Astronomical Union adopted the value 298.25 corresponding to $J_2 = 1082.7 \times 10^{-6}$ (see Sec. 2-11).

9-7. Rectangular Coordinates of the Satellite and Their Perturbations

We shall now describe how the rectangular coordinates of the satellite are computed from the orbital elements. Then we shall outline how they are affected by the irregularities of the gravity field. These considerations are necessary for the determination of tesseral harmonics from satellite observations.

We introduce a coordinate system XYZ that is at rest with respect to the stars. The origin is at the earth's center of mass. The Z-axis coincides with its axis of rotation; the XY-plane is the equatorial plane. The X-axis is the line of intersection of the equatorial plane and the ecliptic (the plane of the earth's orbit around the sun); according to astronomical terminology, it points to the *vernal equinox* ♈. This coordinate system XYZ is fundamental in spherical astronomy.[1]

The relation between the rectangular coordinates of a satellite and the elements of its osculating ellipse (Sec. 9-5) at a certain time is found as follows. Let ξ, η, ζ be the angles between the radius vector r and the X-, Y-, Z-axis, respectively; then

$$X = r\cos\xi, \quad Y = r\cos\eta, \quad Z = r\cos\zeta.$$

The relevant relations between directions are represented by means of the unit sphere of Fig. 9-8. Its center is at the origin O; and the points X, Y, Z denote the intersection of the respective coordinate axes with the unit sphere. The point S represents the direction of the radius vector to the satellite; K is the node. Then ξ is the arc XS, etc.

[1] The directions of the coordinate axes so defined are not completely constant in time. This fact requires certain refinements, for which the reader is referred to the article by G. Veis, "Precise Aspects of Terrestrial and Celestial Reference Frames" in Veis (1963). In the present context we shall consider the XYZ-system as constant in time.

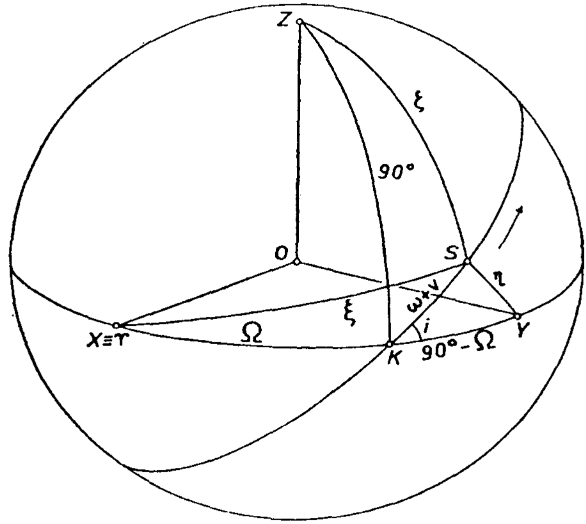

FIGURE 9-8

Unit sphere illustrating the relation between orbital parameters and rectangular coordinates.

The application of the law of cosines to the side ξ of the spherical triangle XKS gives

$$\cos \xi = \cos \Omega \cos (\omega + v) - \sin \Omega \sin (\omega + v) \cos i.$$

From the triangle KYS we find in the same way

$$\cos \eta = \sin \Omega \cos (\omega + v) + \cos \Omega \sin (\omega + v) \cos i.$$

Finally the triangle KSZ yields

$$\cos \zeta = \sin (\omega + v) \sin i.$$

Hence we have

$$\begin{aligned} X &= r[\cos \Omega \cos (\omega + v) - \sin \Omega \sin (\omega + v) \cos i], \\ Y &= r[\sin \Omega \cos (\omega + v) + \cos \Omega \sin (\omega + v) \cos i], \\ Z &= r \sin (\omega + v) \sin i, \end{aligned} \tag{9-34}$$

where, according to (9-22),

$$r = \frac{a(1 - e^2)}{1 + e \cos v}.$$

This expresses the rectangular coordinates of the satellite in terms of the elements of its osculating orbit, the true anomaly v fixing its position as a function of time.

Since the osculating ellipse does not remain constant, it is convenient to use a fixed *reference orbit*—for instance, the osculating ellipse E_0 at a certain instant t_0, having the elements a_0, e_0, i_0, Ω_0, ω_0, T_0. At a later instant t the orbital elements will have changed to $a_0 + \Delta_t a$, $e_0 + \Delta_t e$, $i_0 + \Delta_t i$, $\Omega_0 + \Delta_t \Omega$, $\omega_0 + \Delta_t \omega$, $T_0 + \Delta_t T$, corresponding to an osculating ellipse E_t.

The orbital elements in (9-34) refer to this instantaneous osculating ellipse, so that $a = a_0 + \Delta_t a$, etc. Therefore, the coordinates X, Y, Z depend on the time in two ways: *explicitly*, through the true anomaly v, and *implicitly*, through the variable elements of the osculating orbit. We shall eliminate the implicit dependence in the following way. We evaluate (9-34) using the elements a_0, etc., of the fixed reference ellipse. Then the coordinates X_0, Y_0, Z_0 so obtained depend on the time only explicitly, and correspond to a Keplerian motion in space along a fixed ellipse. To convert them into true coordinates X, Y, Z, these values X_0, Y_0, Z_0 must be corrected by $\Delta_t X$, $\Delta_t Y$, $\Delta_t Z$, for which the linear terms of a Taylor expansion of (9-34) give

$$\begin{aligned}
\Delta_t X &= \frac{\partial X}{\partial a}\Delta_t a + \frac{\partial X}{\partial e}\Delta_t e + \frac{\partial X}{\partial i}\Delta_t i + \frac{\partial X}{\partial \Omega}\Delta_t \Omega + \frac{\partial X}{\partial \omega}\Delta_t \omega + \frac{\partial X}{\partial v}\Delta_t v, \\
\Delta_t Y &= \frac{\partial Y}{\partial a}\Delta_t a + \frac{\partial Y}{\partial e}\Delta_t e + \frac{\partial Y}{\partial i}\Delta_t i + \frac{\partial Y}{\partial \Omega}\Delta_t \Omega + \frac{\partial Y}{\partial \omega}\Delta_t \omega + \frac{\partial Y}{\partial v}\Delta_t v, \qquad (9\text{-}35) \\
\Delta_t Z &= \frac{\partial Z}{\partial a}\Delta_t a + \frac{\partial Z}{\partial e}\Delta_t e + \frac{\partial Z}{\partial i}\Delta_t i + \frac{\partial Z}{\partial \Omega}\Delta_t \Omega + \frac{\partial Z}{\partial \omega}\Delta_t \omega + \frac{\partial Z}{\partial v}\Delta_t v.
\end{aligned}$$

The partial derivatives are readily obtained by differentiating (9-34); note that r is a function of a, e, and v.

In these equations we have used the perturbation of the true anomaly, $\Delta_t v$, instead of the perturbation of perigee epoch, $\Delta_t T$.

Perturbations expressed in terms of J_{nm} and K_{nm}. The perturbations of the orbital elements are found by integrating (9-26):

$$\Delta_t a = \int_{t_0}^{t} \dot{a}\,dt, \quad \Delta_t e = \int_{t_0}^{t} \dot{e}\,dt, \quad \ldots . \qquad (9\text{-}36)$$

A similar expression can be written for $\Delta_t v$. The components S, T, W of the perturbing force are expressed in terms of J_n, J_{nm}, and K_{nm} by equations (9-27) and (9-29), where the perturbing potential

$$R = -\frac{kM}{a_e}\sum_{n=2}^{\infty}\left(\frac{a_e}{r}\right)^{n+1}\left[J_n P_n(\cos\theta) + \sum_{m=1}^{n}(J_{nm}\cos m\lambda + K_{nm}\sin m\lambda)P_{nm}(\cos\theta)\right]$$

now also contains the tesseral harmonics.

By performing the integrations in (9-36) we obtain equations of the form

$$\begin{aligned}
\Delta_t a &= \sum_{n,m}(A_{nm}J_{nm} + \bar{A}_{nm}K_{nm}), \\
\Delta_t e &= \sum_{n,m}(B_{nm}J_{nm} + \bar{B}_{nm}K_{nm}), \qquad (9\text{-}37) \\
&\cdots\cdots\cdots\cdots
\end{aligned}$$

the coefficients A_{nm}, etc., being functions of the time t, which as a rule are periodic. Zonal and tesseral harmonics have been combined in (9-37) by setting

$J_n = J_{n0}$ and admitting the value $m = 0$; this practice will be continued in what follows.

The insertion of (9-37) into (9-35) gives the perturbation of the rectangular coordinates X, Y, Z as functions of the harmonic coefficients J_{nm} and K_{nm} in the form

$$\begin{aligned} \Delta_t X &= \sum_{n,m} (L_{nm} J_{nm} + \bar{L}_{nm} K_{nm}), \\ \Delta_t Y &= \sum_{n,m} (M_{nm} J_{nm} + \bar{M}_{nm} K_{nm}), \\ \Delta_t Z &= \sum_{n,m} (N_{nm} J_{nm} + \bar{N}_{nm} K_{nm}), \end{aligned} \tag{9-38}$$

where again L_{nm}, $\bar{L}_{nm}$, M_{nm}, etc. are functions of the time t.

These perturbations are added to the coordinates X_0, Y_0, Z_0 computed from (9-34) using the orbital elements of the reference ellipse E_0. In this way we obtain the rectangular coordinates of the satellite in the form

$$\begin{aligned} X &= X(t;\, a_0,\, e_0,\, i_0,\, \Omega_0,\, \omega_0,\, T_0;\, J_{nm},\, K_{nm}), \\ Y &= Y(t;\, a_0,\, e_0,\, i_0,\, \Omega_0,\, \omega_0,\, T_0;\, J_{nm},\, K_{nm}), \\ Z &= Z(t;\, a_0,\, e_0,\, i_0,\, \Omega_0,\, \omega_0,\, T_0;\, J_{nm},\, K_{nm}), \end{aligned} \tag{9-39}$$

as *explicit* functions of the time t, containing as *constant* parameters the orbital element of the reference ellipse E_0 and the gravitational coefficients J_{nm} and K_{nm}. This is the advantage of (9-39) over the system (9-34), which formally is much simpler, but depends on the *variable* orbital parameters of the osculating ellipse.

The actual expressions for (9-39) are very complicated. Therefore we have been satisfied with outlining the procedure, referring the reader for details to Kaula (1962) and to the literature given there.

9-8. Determination of Tesseral Harmonics and Station Positions

Zonal harmonics give rise to secular and long-periodic perturbations of the orbital elements a, e, etc. Therefore, their influence can be detected in changes of orbital parameters obtained by integrating over many revolutions of the satellite.

The perturbations due to tesseral harmonics have a much shorter period. The longest period of a harmonic of the order $m = 1$ is one day, for $m = 2$ it is only half a day, etc. Therefore, we must look for another method, which is sensitive enough to detect even short-periodic effects and extracts as much information as possible from the observations.

The observed elements are essentially spatial polar coordinates of the satellite with respect to the observing station: the distance s and the direction as determined by two angles. Corresponding to our coordinate system X, Y, Z introduced in the preceding section, these two angles are the *right ascension* α and

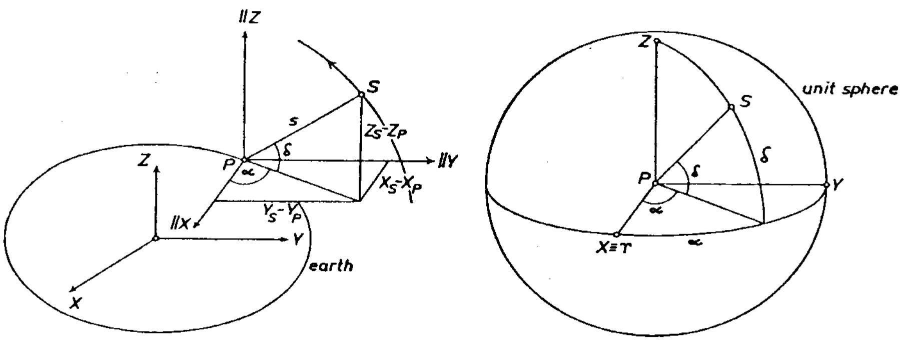

FIGURE 9-9

The direction to the satellite as defined by right ascension α and declination δ.

the *declination* δ, whose definition may be seen in Fig. 9-9. The angles α and δ are obtained by photographing the satellite against the background of stars, as outlined in Sec. 9-1. We also recall that s can be measured by radar, either ordinary or optical.[1]

Denoting the rectangular coordinates of the terrestrial station P by X_P, Y_P, Z_P, and of the satellite S by X_S, Y_S, Z_S, we find by inspecting Fig. 9-9

$$\begin{aligned} X_S - X_P &= s\cos\delta\cos\alpha, \\ Y_S - Y_P &= s\cos\delta\sin\alpha, \\ Z_S - Z_P &= s\sin\delta, \end{aligned} \tag{9-40}$$

so that

$$\begin{aligned} \alpha &= \tan^{-1}\frac{Y_S - Y_P}{X_S - X_P}, \\ \delta &= \tan^{-1}\frac{Z_S - Z_P}{\sqrt{(X_S - X_P)^2 + (Y_S - Y_P)^2}}, \\ s &= \sqrt{(X_S - X_P)^2 + (Y_S - Y_P)^2 + (Z_S - Z_P)^2}. \end{aligned} \tag{9-41}$$

We shall now compute the rectangular coordinates X_P, Y_P, Z_P of the observing station P. The system XYZ, being fixed with respect to the stars, rotates with respect to the earth. The coordinates of P in this system will therefore be functions of time. Let x_P, y_P, z_P be the coordinates of P in the usual geocentric coordinate system fixed with respect to the earth. In this system, the z-axis, coinciding with the Z-axis, is the earth's axis of rotation; the x-axis lies in the mean meridian plane of Greenwich, corresponding to the longitude $\lambda = 0°$; and the y-axis points to $\lambda = 90°$ east. Figure 9-10 shows that

[1] The measurement of the radial velocity ds/dt of the satellite by means of the Doppler effect is also very important for the determination of tesseral harmonics and station positions. Doppler measurements are treated in essentially the same way as measurements of directions and distances.

$$\begin{aligned} X_P &= x_P \cos \tau - y_P \sin \tau, \\ Y_P &= x_P \sin \tau + y_P \cos \tau, \\ Z_P &= z_P. \end{aligned} \tag{9-42}$$

The angle τ is called *Greenwich sidereal time*; its value is

$$\tau = \omega t,$$

where ω is the angular velocity of the earth's rotation. It is proportional to the time t and, in appropriate units, measures it. Thus absolute Greenwich time is needed to convert the constant (in time) coordinates x_P, y_P, z_P to the periodically varying coordinates X_P, Y_P, Z_P that are required in (9-40) and (9-41).

As a final step we insert the station coordinates, as given by (9-42), and the satellite coordinates, as symbolized by (9-39), into (9-41), obtaining expressions of the form

$$\begin{aligned} \alpha &= \alpha(x_P, y_P, z_P; \; t; \; a_0, e_0, i_0, \Omega_0, \omega_0, T_0; \; J_{nm}, K_{nm}), \\ \delta &= \delta(x_P, y_P, z_P; \; t; \; a_0, e_0, i_0, \Omega_0, \omega_0, T_0; \; J_{nm}, K_{nm}), \\ s &= s(x_P, y_P, z_P; \; t; \; a_0, e_0, i_0, \Omega_0, \omega_0, T_0; \; J_{nm}, K_{nm}). \end{aligned} \tag{9-43}$$

Besides depending on the station coordinates and the time, they also contain the orbital and gravitational parameters.

Every observed α, δ, or s furnishes an equation of type (9-43). Provided we have a sufficient number of such observation equations, we can solve them for the station coordinates x_P, y_P, z_P, for the orbital parameters a_0, e_0, etc., of the reference ellipse, and for a certain number of gravitational parameters J_{nm} and K_{nm}. This is the principle of the *orbital method*. In practice, differential formulas will be applied to determine corrections to assume approximate values by means of a least squares adjustment.

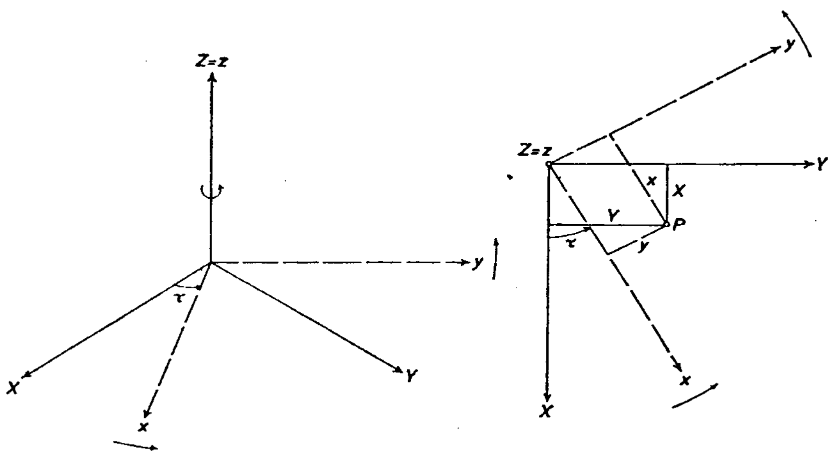

FIGURE 9-10

Geocentric coordinate systems XYZ (celestial) and xyz (terrestrial).

Therefore, the actual analytical developments are from the outset directed toward obtaining differential formulas corresponding to (9-43). The substitutions indicated above are thus consistently performed in terms of the corresponding differential expressions. In this way one is able to operate with linear equations and to employ that efficient tool of linear analysis, matrix calculus.

Simple though the principle of this procedure is, the details when written out are nevertheless so complicated that the reader must again be referred to the literature (e.g., Kaula, 1962).

Besides these analytical problems, which have been satisfactorily solved, the geodetic application of (9-43) raises difficulties similar in principle to those involved in the determination of zonal harmonics by means of (9-32), but even more serious in practice. Strictly speaking, an infinite number of unknowns, J_{nm}, K_{nm}, etc., are to be determined from a finite number of observations. In order to get a definite solution it must be assumed that the effect of higher degree terms is negligibly small. But even then there are very many unknowns: coordinates of the observing stations, parameters of the reference orbit, and gravitational parameters; in addition, other unknowns must be included to take into account nongravitational forces acting on the satellite, such as air drag.

To get a strong solution, observations should be evenly distributed both in space (with respect to the inclination of the satellites used) and in time. This is difficult to achieve. From this point of view radio methods such as the doppler method are superior to photographic observations.

Present results. At present (1966) several determinations of tesseral harmonics up to the seventh degree are available; see the review paper by Y. Kozai in Veis (1966). These coefficients represent the large-scale features of the disturbing potential T and hence of the geoid, since the geoidal height is given by $N = T/\gamma$. There is a general agreement between the broad qualitative aspects of these determinations as expressed in geoidal maps, although the details of these maps, and even more so the individual coefficients, are rather different.

As an example we take the first nonzonal coefficients, J_{22} and K_{22}, which according to Sec. 2-6, equation (2-49) express the inequality of the earth's principal equatorial moments of inertia or, somewhat loosely speaking, its triaxiality. By analyzing optical observations in 1964, I. G. Izsak found the values

$$J_{22} = -0.83 \times 10^{-6}, \quad K_{22} = 0.56 \times 10^{-6},$$

whereas R. J. Anderle in 1965 obtained

$$J_{22} = -1.58 \times 10^{-6}, \quad K_{22} = 0.98 \times 10^{-6}$$

from doppler observations.

As far as station positions are concerned, it is hoped that an accuracy of absolute geocentric position of ± 10 meters will eventually be achieved. The present accuracy is perhaps around ± 30 meters.

Stellar triangulation. It is also possible to perform a three-dimensional triangulation by applying equations (9-40) through (9-42) in a purely geometric way. In this method the orbit itself is not used, the satellite being considered merely as an elevated target, in place of which a flashing rocket or an illuminated balloon might also be employed. Then the satellite must be observed *simultaneously* from several ground stations. The position S of the satellite at that instant forms a corner of a three-dimensional triangulation net, in much the same way as the ground stations do; and the coordinates of S are directly determined as unknowns just as the coordinates of any ground station P are determined.

This procedure is called *stellar triangulation*; its principles were given by Y. Väisälä in 1946. This astronomical method of triangulation is basically the same as terrestrial three-dimensional triangulation described in Sec. 5-12. Azimuth and zenith distance are replaced by right ascension and declination, and equations (9-41), together with (9-42), take the place of (5-82). In the same way as in Sec. 5-12 we may introduce approximate positions, which are corrected by differential formulas analogous to (5-83) or (5-85).

The main advantage of stellar triangulation as compared to terrestrial triangulation rests on the fact that atmospheric refraction is made harmless by referring the satellite direction to the star background, since the main part of the refraction affects the images of the satellite and the stars in the same way and hence cancels out. Other advantages of stellar triangulation are the possibility of measuring larger triangles, because intervisibility between the ground stations is not required, and its independence of the direction of the plumb line, since equations (9-41) do not contain Φ and Λ.

A practical difficulty is caused by the need of observing simultaneously from several stations. This requires a synchronization accurate to one millisecond or better, because of the rapid motion of the satellite.[1]

As compared to the joint determination of station positions and tesseral harmonics by means of the orbital method expressed by (9-43), stellar triangulation has the advantage that it does not depend on an infinite number of gravitational parameters. On the other hand, it cannot give absolute geocentric positions. It is a common feature of all geometric methods—the conventional astrogeodetic method, the spatial triangulation of Bruns-Hotine, and stellar triangulation—that they can only furnish *relative* positions. In the case of stellar triangulation this is expressed by the fact that the observables α, δ, and s depend only on the coordinate *differences* according to (9-41).

In order to obtain geocentric coordinates we must have recourse to physical principles—to potential theory in the gravimetric method and to celestial mechanics in the orbital method.

[1] For the same reason, accurate timing is also necessary in the orbital method using nonsimultaneous observations. As a rough estimate, one millisecond of time and one second of arc in direction correspond to the desired positional accuracy of ± 10 meters.

References

Berroth, A., and W. Hofmann (1960). Kosmische Geodäsie. Karlsruhe, G. Braun.

Brouwer, D., and G. M. Clemence (1961). Methods of celestial mechanics. New York and London, Academic Press.

Cook, A. H. (1963). The contribution of observations of satellites to the determination of the earth's gravitational potential. Space Science Reviews, v. 2, pp. 355–437.

Fischer, I. (1962). The parallax of the moon in terms of a world geodetic system. Astron. J., v. 67, pp. 373–378.

Helmert, F. R. (1884). Die mathematischen und physikalischen Theorien der höheren Geodäsie, vol. 2. Leipzig, B. G. Teubner (reprinted 1962).

Jeffreys, H. (1962). The earth. 4th ed., Cambridge University Press.

Kaula, W. M. (1962). Celestial geodesy. Advances in Geophysics, v. 9, pp. 191–293.

Kopal, Z. (1960). Figures of equilibrium of celestial bodies. Madison, The University of Wisconsin Press.

Ledersteger, K. (1966). Multi-parametric theory of spheroidal equilibrium figures and the normal-spheroids of earth and moon (Final Technical Report). Vienna, Technical University, Institute of Geodesy.

Mueller, I. I. (1964). Introduction to satellite geodesy. New York, Ungar.

Plummer, H. C. (1918). An introductory treatise on dynamical astronomy. Cambridge University Press (New York, Dover Publications, 1960).

Thomson, W. T. (1961). Introduction to space dynamics. New York and London, Wiley.

Veis, G., ed. (1963). The use of artificial satellites for geodesy (Proceedings of the First International Symposium on the Use of Artificial Satellites for Geodesy, Washington, D.C., April 1962). Amsterdam, North-Holland Publishing Co.

Veis, G., ed. (1966). Proceedings of the Second International Symposium on the Use of Artificial Satellites for Geodesy, Athens, April 1965. In press.

Wavre, R. (1932). Figures planétaires et Géodésie. Paris, Gauthier-Villars.

Index