GEOMETRIC GEODESY
PART II

by

Richard H. Rapp

The Ohio State University
Department of Geodetic Science and Surveying
1958 Neil Avenue
Columbus, Ohio 43210

March 1993

Foreword

Geometric Geodesy, Volume II, is a continuation of Volume I. While the first volume emphasizes the geometry of the ellipsoid, the second volume emphasizes problems related to geometric geodesy in several diverse ways. The four main topic areas covered in Volume II are the following: the solution of the direct and inverse problem for arbitrary length lines; the transformation of geodetic data from one reference frame to another; the definition and determination of geodetic datums (including ellipsoid parameters) with terrestrial and space derived data; the theory and methods of geometric three-dimensional geodesy.

These notes represent an evolution of discussions on the relevant topics. Chapter 1 (long lines) was revised in 1987 and retyped for the present version. Chapter 2 (datum transformation) and Chapter 3 (datum determination) have been completely revised from past versions. Chapter 4 (three-dimensional geodesy) remains basically unchanged from previous versions.

The original version of the revised notes was printed in September 1990. Slight revisions were made in the 1990 version in January 1992. For this printing, several corrections were made in Table 1.4 (line E and F). The need for such corrections, and several others, was noted by B.K. Meade whose comments are appreciated.

Richard H. Rapp
March 25, 1993

Table of Contents

1. Long Geodesics on the Ellipsoid

1.1 Introduction

The purpose of this section is to discuss methods for the solution of the direct and inverse problem without limitation on distance. There are several solutions that have been derived for lines whose length does not exceed 500 or 1000 km with a number of solutions for considerably shorter distances. The most familiar shorter solution is the Puissant's equations, where the result is interpreted for the normal section or the geodesic as the line is too short for distinction. A desription of several of the methods and the resultant equations may be found in Bomford (1980, Sec. 2.14). A discussion of methods for lines up to 200 km in length may be found in Rapp (1984).

1.2 An Iterative Solution for Long Geodesics

The discussion given here has evolved from the English translation of Helmert's "Higher Geodesy" written in 1898 and from the Army Map Service translation of Jordan's "Handbook of Geodesy" - Volume III, second half, dated 1962 (original 1941), sections 23 and 24.

The problem of computing "long" geodesics is attacked by considering the relationship between the ellipsoid and sphere, in terms of distance and longitude. The main concept of the derivation is to use the sphere as an auxiliary surface and relate it to the ellipsoid. We do not approximate the ellipsoid by a sphere. The radius of the sphere is immaterial and in fact, the sphere may be considered to have a unit radius.

First let us establish some differential relationships between the ellipsoid and the sphere. We define first:

ϕ, L geodetic latitude and longitude on the ellipsoid

β reduced latitude, (latitude on auxiliary sphere)

λ longitude on the sphere

α geodesic azimuth

σ spherical arc on the sphere

A fundamental property of the geodesic on the ellipsoid follows from Clairaut's equation such that:

$$\cos \beta_1 \sin \alpha_1 = \cos \beta_2 \sin \alpha_2 = \cos \beta_i \sin \alpha_i = \cos \beta_0 \qquad (1.1)$$

where β and α are the reduced latitude and geodesic azimuth at any point on the geodesic, and β_0 is the highest reduced latitude that this geodesic has reached. Equation (1.1) represents a property of all geodesics, whether on the ellipsoid or sphere. We now construct an auxiliary sphere. A geodesic is mapped from the ellipsoid to a great circle on the auxiliary sphere by specifying that the highest reduced latitude of the geodesic (extended if necessary) will be the same as the highest latitude of the corresponding great circle on the sphere. See Figure 1.1.
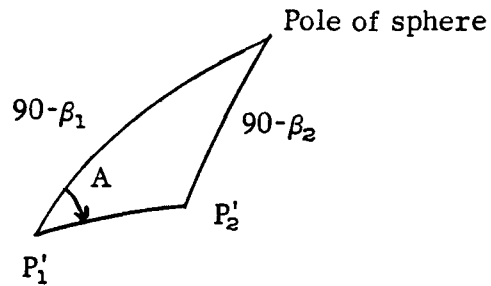
Figure 1.1
Polar Triangle on the Auxiliarly Sphere

A is the azimuth of the geodesic (great circle on the sphere) from $P_1'P_2'$. Using the property expressed in equation (1.1) we have:

$$\cos \beta_1 \sin A_1 = \cos \beta_2 \sin A_2 = \cos \beta_i \sin A_i = \cos \beta_0 \qquad (1.2)$$

By definition, the β_0 in (1.1) must be equal to the β_0 in (1.2). We must then have the azimuths on the ellipsoid and on the sphere the same, i.e. $A_i = \alpha_i$.

Next we consider a differential figure on the ellipsoid and sphere as shown in Figure 1.2.
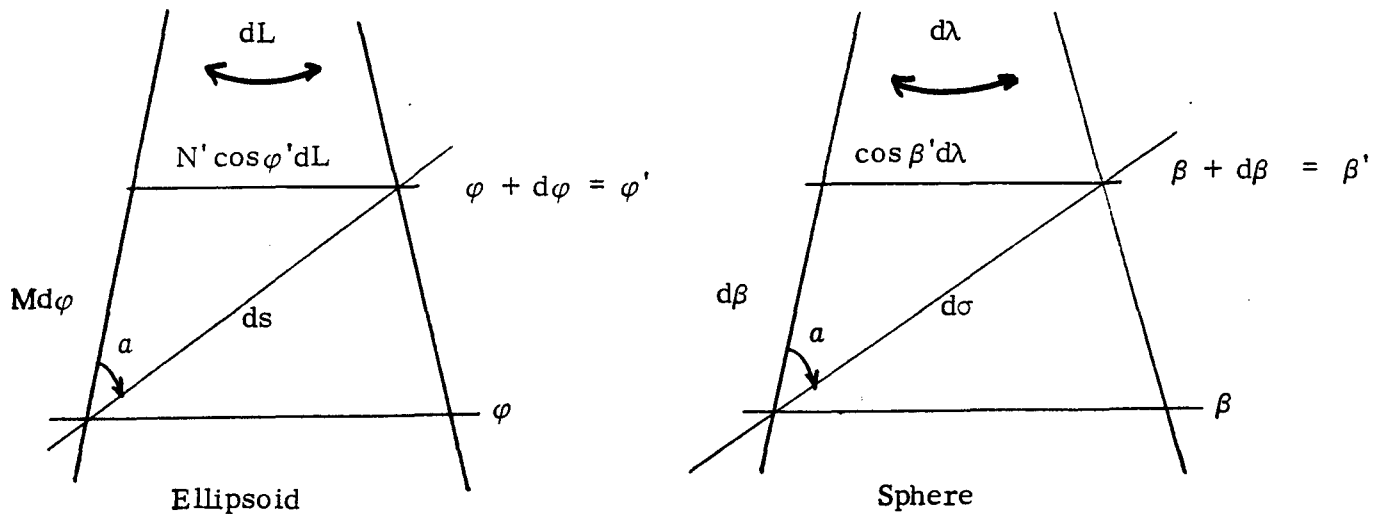
Figure 1.2
Differential Figures on the Ellipsoid and the Sphere

Then we have for the ellipsoid:

$$ds \cos\alpha = M d\phi$$
$$ds \sin\alpha = N' \cos\phi' \, dL \tag{1.3}$$

and for the sphere:

$$d\sigma \cos\alpha = d\beta$$
$$d\sigma \sin\alpha = \cos\beta' \, d\lambda \tag{1.4}$$

Dividing the first equation in (1.3) by the first equation in (1.4), and repeating for the second equations we have:

$$\frac{ds}{d\sigma} = M \frac{d\phi}{d\beta} = a \frac{dL}{d\lambda} \tag{1.5}$$

But $N' \cos\phi' = a \cos\beta'$ so that:

$$\frac{ds}{d\sigma} = M \frac{d\phi}{d\beta} = a \frac{dL}{d\lambda} \tag{1.6}$$

Equation (1.6) may be written in several forms. For example:

$$\frac{dL}{d\lambda} = \frac{1}{a} \frac{ds}{d\sigma} \tag{1.7}$$

or

$$\frac{ds}{d\sigma} = M \frac{d\phi}{d\beta} \tag{1.8}$$

We consider equation (1.8) by recalling:

$$\tan\beta = \left(1 - e^2\right)^{1/2} \tan\phi \tag{1.9}$$

or upon differentiation:

$$\frac{d\beta}{\cos^2\beta} = \left(1 - e^2\right)^{1/2} \frac{d\phi}{\cos^2\phi} \qquad \text{so that}$$

$$\frac{d\phi}{d\beta} = \frac{1}{\left(1 - e^2\right)^{1/2}} \frac{\cos^2\phi}{\cos^2\beta} \tag{1.10}$$

Equation (1.7) then becomes (with 1.8):

$$\frac{dL}{d\lambda} = \frac{M}{a\left(1 - e^2\right)^{1/2}} \frac{\cos^2 \phi}{\cos^2 \beta}$$

(1.11)

We also recall at this point the expression for the x coordinate of a point located on a meridian ellipse (Rapp, 1984, Sec 3.3):

$$x = a\cos\beta = \frac{c}{V} \cos\phi$$

(1.12)

where $c = a^2/b$ and $V^2 = 1 + e'^2 \cos^2 \phi$. From (1.12):

$$\frac{\cos^2 \phi}{\cos^2 \beta} = \frac{V^2 a^2}{c^2}$$

(1.13)

Using the relation $M = c/V^3$ we many write equation (1.12) as:

$$\frac{dL}{d\lambda} = \frac{c}{V^3} \frac{1}{a\left(1 - e^2\right)^{1/2}} \frac{V^2 a^2}{c^2} \quad \text{or}$$

(1.14)

$$\frac{dL}{d\lambda} = \frac{a}{Vc} \frac{1}{\left(1 - e^2\right)^{1/2}}$$

(1.15)

Noting that $c = a^2/b$ and $b = a (1 - e^2)^{1/2}$ equations (1.14) and (1.7) become:

$$\frac{dL}{d\lambda} = \frac{1}{V} = \frac{1}{a} \frac{ds}{d\sigma}$$

(1.16)

From Rapp (ibid, eq 3.41):

$$V = \left(1 - e^2 \cos^2 \beta\right)^{-1/2}$$

(1.17)

so that equation (1.16) can be written as:

$$\frac{ds}{d\sigma} = a \sqrt{1 - e^2 \cos^2 \beta}$$

(1.18)

and

$$\frac{dL}{d\lambda} = \sqrt{1 - e^2 \cos^2 \beta}$$

<div align="right">(1.19)</div>

We now must consider the integration of equations (1.18) and (1.19). Consider two points P_1 and P_2 on the sphere shown in Figure 1.3.
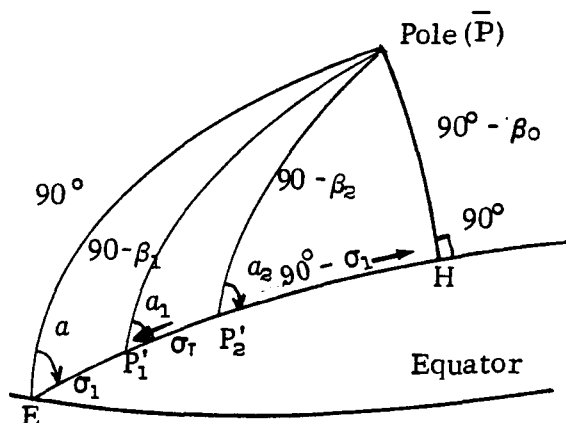
Figure 1.3
Geometry of Auxiliarly Spherical Triangle

We let σ be an arc length on the great circle and define the following:

σ = arc from P_1' to an arbitrary P'
σ_1 = arc from E to P_1
σ_2 = arc from E to P_2
σ_T = arc from P_1 to P_2

We also note that the arc from P_1' to H is $90° - \sigma_1$ and $\sigma_T = \sigma_2 - \sigma_1$. We let:

α = azimuth of specific geodesic at the equator
α_1 = azimuth of specific geodesic at P_1
α_2 = azimuth of specific geodesic at P_2
β_0 = highest reduced latitude geodesic reaches.

From the spherical triangle $\bar{P} \, P_1' \, H$ or using (1.1) we have:

$$\frac{\sin \alpha_1}{\cos \beta_0} = \frac{1}{\cos \beta_1} \quad \text{so that:}$$

$$\cos \beta_0 = \sin \alpha_1 \cos \beta_1 \tag{1.20}$$

Applying Napier's Rules to triangle $P_1' \overline{P}H$ we have:

$$\cos \alpha_1 = \tan(90 - \sigma_1) \cot\left(90 - \beta_1\right) = \cot \sigma_1 \tan \beta_1 \qquad \text{so that:}$$

$$\tan \sigma_1 = \frac{\tan \beta_1}{\cos \alpha_1} \tag{1.21}$$

From the spherical triangle $\overline{P}P_2'H$ we have:

$$\sin\beta_2 = \sin\left(\sigma_1 + \sigma_T\right) \sin\beta_0 \tag{1.22}$$

If we apply (1.22) at some arbitrary point P' (where β_2 becomes an arbitrary β, and σ_T is associated with σ) we write:

$$\sin\beta = \sin\left(\sigma_1 + \sigma\right) \sin\beta_0 \tag{1.23}$$

From equation (1.18) or (1.19) we need to find an expression for $\cos^2\beta$. Thus using $\cos^2\beta = 1 - \sin^2\beta$ with $\sin\beta$ from equation (1.23) we have:

$$\cos^2\beta = 1 - \sin^2\left(\sigma_1 + \sigma\right) \sin^2\beta_0 \tag{1.24}$$

If we let $x = \sigma_1 + \sigma$ so that

$$\cos^2\beta = 1 - \sin^2 x \sin^2\beta_0 \tag{1.25}$$

and noting $dx = d\sigma$ since σ_1 is a constant, we write (1.18) as:

$$ds = a \sqrt{1 - e^2 + e^2 \sin^2\beta_0 \sin^2 x} \; dx$$

Now

$$e^2 = \frac{e'^2}{1 + e'^2}, \qquad 1 - e^2 = \frac{1}{1 + e'^2} \qquad \text{so:}$$

$$ds = a \left[\frac{1}{1 + e^{'2}} + \frac{e^{'2}}{1 + e^{'2}} \sin^2\beta_0 \sin^2 x \right]^{1/2} dx$$

or:

$$ds = \frac{a}{\left(1 + e^{'2}\right)^{1/2}} \sqrt{1 + e^{'2} \sin^2\beta_0 \sin^2 x} \; dx$$

We note however: $\dfrac{1}{\left(1 + e^{'2}\right)^{1/2}} = \dfrac{b}{a}$

We define: $k^2 = e^{'2} \sin^2\beta_0$ (1.26)

so that we now have:

$$ds = b \sqrt{1 + k^2 \sin^2 x} \; dx \tag{1.27}$$

Before we integrate this expression we must establish the limits on x. Recall $x = \sigma_1 + \sigma$. At the start of a line $\sigma = 0$ yielding the lower limit on x: $x = \sigma_1$. At the end of the line $\sigma = \sigma_T$. Thus, in integral form, equation (1.27) is written:

$$s = b \int_{\sigma_1}^{\sigma_1 + \sigma_T} \sqrt{1 + k^2 \sin^2 x} \; dx \tag{1.28}$$

This integral is similar to what are called elliptic integrals (Bulirsch and Gerstl, 1983). The evaluation of these integrals could take place in two ways: by numerical integration or by analytic integration. The first form is possible using various numerical integration methods. The second procedure, although more complicated than the first, allows a better accuracy control on the solution and permits a unique set of equations to be established.

We thus look at the integration of (1.28) by analytic procedures. We first expand the kernel of (1.28):

$$\left(1 + k^2 \sin^2 x\right)^{1/2} = 1 + \frac{1}{2} k^2 \sin^2 x - \frac{1}{8} k^4 \sin^4 x + \frac{1}{16} k^6 \sin^6 x + \ldots \tag{1.29}$$

Next we convert from powers of angles to multiple angles. We use the relationships given in Rapp (ibid, Section 2.5).

Then equation (1.29) becomes, after combining terms:

$$\left(1 + k^2 \sin^2 x\right)^{1/2} = \left(1 + \frac{k^2}{4} - \frac{3}{64} k^4 + \frac{5}{256} k^6 + --\right) + \left(-\frac{1}{4} k^2 + \frac{1}{16} k^4 - \frac{15}{512} k^6 + --\right) \cos 2x$$

$$+ \left(-\frac{k^4}{64} + \frac{3}{256} k^6 + --\right) \cos 4x + \left(-\frac{1}{512} k^6 + --\right) \cos 6x + -- \tag{1.30}$$

We now define the coefficients of cos (nx) as A, B, C, D, -- respectively. That is:

$$A = 1 + \frac{k^2}{4} - \frac{3}{64} k^4 + \frac{5}{256} k^6 - \frac{175}{16384} k^8 + --$$

$$B = -\frac{1}{4} k^2 + \frac{1}{16} k^4 - \frac{15}{512} k^6 + \frac{35}{2048} k^8 + --$$

$$C = -\frac{k^4}{64} + \frac{3}{256} k^6 - \frac{35}{4096} k^8 + --$$

$$D = -\frac{1}{512} k^6 + \frac{5}{2048} k^8 + -- \tag{1.31}$$

etc.

Then:

$$\left(1 + k^2 \sin^2 x\right)^{1/2} = A + B\cos 2x + C\cos 4x + D\cos 6x + -- \tag{1.32}$$

which we now insert for integration into (1.28) yielding

$$\frac{s}{b} = A \int_{\sigma_1}^{\sigma_1 + \sigma_T} dx + B \int_{\sigma_1}^{\sigma_1 + \sigma_T} \cos 2x \, dx + C \int_{\sigma_1}^{\sigma_1 + \sigma_T} \cos 4x \, dx + D \int_{\sigma_1}^{\sigma_1 + \sigma_T} \cos 6x \, dx + -- \tag{1.33}$$

First consider the general integral:

$$\int_{\sigma_1}^{\sigma_1 + \sigma_T} \cos nx \, dx = \frac{1}{n} \sin nx \Bigg|_{\sigma_1}^{\sigma_1 + \sigma_T} = \frac{1}{n} \left[\sin n \left(\sigma_1 + \sigma_T\right) - \sin n\sigma_1\right] \tag{1.34}$$

We may abbreviate this by recalling the trigonometric identity:

$$\sin nX - \sin nY = 2\cos \frac{n}{2} (X + Y) \sin \frac{n}{2} (X - Y) \tag{1.35}$$

In our case:

$$X = \sigma_1 + \sigma_T$$

$$Y = \sigma_1$$

$$X + Y = 2\sigma_1 + \sigma_T$$

$$X - Y = \sigma_T$$

Now (1.35) becomes:

$$\sin n\left(\sigma_1 + \sigma_T\right) - \sin n\sigma_1 = 2\cos \frac{n}{2}\left(2\sigma_1 + \sigma_T\right) \sin \frac{n}{2} \sigma_T \tag{1.36}$$

Recalling that $\sigma_T = \sigma_2 - \sigma_1$, we have: $2\sigma_1 + \sigma_T = 2\sigma_1 + \sigma_2 - \sigma_1 = \sigma_1 + \sigma_2$. If we then define

$$\sigma_m = \frac{\sigma_1 + \sigma_2}{2} \tag{1.37}$$

equation (1.36) becomes:

$$\sin n\left(\sigma_1 + \sigma_T\right) - \sin n\sigma_1 = 2\cos n\sigma_m \sin \frac{n}{2} \sigma_T \tag{1.38}$$

Thus (1.34) now can be written as:

$$\int_{\sigma_1}^{\sigma_1 + \sigma_T} \cos nx \, dx = \frac{2}{n} \cos n\sigma_m \sin \frac{n}{2} \sigma_T \tag{1.38a}$$

Now we go back to (1.33), using (1.34) with (1.38) to write:

$$\int_{\sigma_1}^{\sigma_1 + \sigma_T} dx = \sigma_T$$

$$\int_{\sigma_1}^{\sigma_1 + \sigma_T} \cos 2x \, dx = \cos 2\sigma_m \sin \sigma_T$$

$$\int_{\sigma_1}^{\sigma_1 + \sigma_T} \cos 4x \, dx = \frac{1}{2} \cos 4\sigma_m \sin 2\sigma_T$$

$$\int_{\sigma_1}^{\sigma_1 + \sigma_T} \cos 6x \, dx = \frac{1}{3} \cos 6\sigma_m \sin \sigma_T$$

Then equation (1.33) becomes:

$$s = b\left(A\sigma_T + B\cos2\sigma_m \sin\sigma_T + \frac{C}{2}\cos4\sigma_m \sin2\sigma_T + \frac{D}{3}\cos6\sigma_m \sin3\sigma_T + --\right)$$

(1.39)

This equation is an important part of the iterative solution of the direct solution. Before we go on we define a new set of constants to be consistent with that in a paper of Rainsford (1955): In addition, we add additional terms as given by Rainsford. We define:

$B_0 = A$
$B_2 = B$
$B_4 = C/2$
$B_6 = D/3$
$B_8 = E/4$ etc.

We also let $u^2 = k^2 = e'^2 \sin^2\beta_0 = e'^2 \cos^2\alpha$, recalling that α is the azimuth of the geodesic at the equator. Then (1.39) becomes (dropping the subscript, T, on the σ):

$$s = b\left(B_0\sigma + B_2 \sin\sigma \cos2\sigma_m + B_4 \sin2\sigma \cos4\sigma_m + B_6 \sin3\sigma \cos6\sigma_m\right.$$

$$\left. + B_8 \sin4\sigma \cos8\sigma_m + --\right)$$

(1.40)

In equation (1.40), we have the following coefficients:

$$B_0 = 1 + \frac{1}{4}u^2 - \frac{3}{64}u^4 + \frac{5}{256}u^6 - \frac{175}{16384}u^8 + --$$

$$B_2 = -\frac{1}{4}u^2 + \frac{1}{16}u^4 - \frac{15}{512}u^6 + \frac{35}{2048}u^8 + --$$

$$B_4 = -\frac{1}{128}u^4 + \frac{3}{512}u^6 - \frac{35}{8192}u^8 + --$$

(1.41)

$$B_6 = -\frac{1}{1536}u^6 + \frac{5}{6144}u^8 + --$$

$$B_8 = -\frac{5}{65536}u^8 + --$$

Equation (1.40) may be used in two ways which will be discussed in detail later. Briefly, however, we may use it to solve iteratively for σ (given s) by first computing a zeroth approximation as $\sigma_0 = s/Ab$, using this on the right side of the equation and solving iteratively to convergence. The value of σ_m may be found from spherical trigonometry formulas as will be shown later. A second application of (1.40) is in the computation of s once σ is determined.

At this point we have derived a connection between a distance on a sphere and the distance on the ellipsoid. However, we do not have a relation (other than in differential form) between the longitude on the ellipsoid and the longitude on the sphere. We now do this by integrating equation (1.19). We consider Figure 1.4.
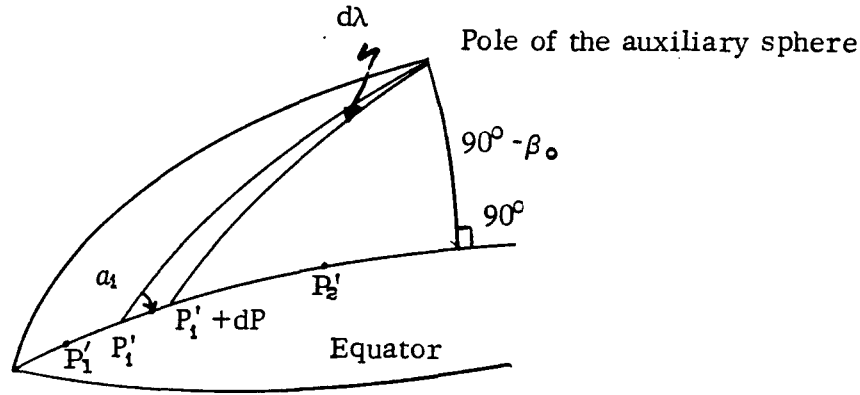
Figure 1.4
Auxiliary Spherical Triangle for Longitude Determination

In this figure P_i' is an arbitrary point on the great circle between P_1' and P_2'. This differential triangle may be enlarged to look as follows:
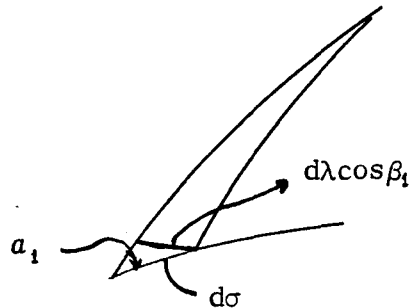
Figure 1.5
Differential Triangle for Longitude Determination

We have:

$$d\lambda \cos\beta_i = d\sigma \sin\alpha_i$$

or

$$d\lambda = \frac{\sin\alpha_i}{\cos\beta_i} d\sigma \qquad (1.42)$$

From equation (1.20) we write:

$$\sin\alpha_i = \frac{\cos\beta_0}{\cos\beta_i} \qquad (1.43)$$

Substituting this into (1.42) we have:

$$d\lambda = \frac{\cos\beta_0}{\cos^2\beta_i} d\sigma \qquad (1.44)$$

Using (1.44) in equation (1.19) we may write (with β_i now simply β):

$$dL = \cos\beta_0 \frac{\sqrt{1 - e^2 \cos^2\beta}}{\cos^2\beta} d\sigma$$

(1.45)

Now subtract equation (1.44) from equation (1.45):

$$dL - d\lambda = \cos\beta_0 \left[\frac{\sqrt{1 - e^2 \cos^2\beta}}{\cos^2\beta} - \frac{1}{\cos^2\beta} \right] d\sigma$$

(1.46)

In order to simplify the bracketed expression we expand the radical term:

$$\left(1 - e^2 \cos^2\beta\right)^{1/2} = 1 - \frac{e^2}{2}\cos^2\beta - \frac{e^4}{8}\cos^4\beta - \frac{e^6}{16}\cos^6\beta ---$$

so that:

$$\frac{\left(1 - e^2 \cos^2\beta\right)^{1/2}}{\cos^2\beta} = \frac{1}{\cos^2\beta} - \frac{e^2}{2} - \frac{e^4}{8}\cos^2\beta - \frac{e^6}{16}\cos^4\beta --$$

Subtracting $1/\cos^2\beta$ from this expression, equation (1.46) may now be written as:

$$dL = d\lambda - \cos\beta_0 \left(\frac{e^2}{2} + \frac{e^4}{8}\cos^2\beta + \frac{e^6}{16}\cos^4\beta + -- \right) d\sigma$$

which may be re-written:

$$d\lambda - dL = \frac{e^2}{2}\cos\beta_0 \left(1 + \frac{e^2}{4}\cos^2\beta + \frac{e^4}{8}\cos^4\beta + -- \right) d\sigma$$

(1.47)

We now have to put (1.47) into an integrable form. From equation (1.23) we had:

$$\sin\beta = \sin\left(\sigma_1 + \sigma\right)\sin\beta_0$$

With $x = \sigma_1 + \sigma$ this becomes:

$$\sin\beta = \sin x \sin\beta_0$$

Then:

$$\cos^2 \beta = 1 - \sin^2 \beta_0 \sin^2 x \tag{1.48}$$

$$\cos^4 \beta = 1 - 2\sin^2 \beta_0 \sin^2 x + \sin^4 \beta_0 \sin^4 x$$

Now insert these into (1.47):

$$d\lambda - dL = \frac{e^2}{2} \cos\beta_0 \left[1 + \frac{e^2}{4}\left(1 - \sin^2 \beta_0 \sin^2 x\right) + \frac{e^4}{8}\left(1 - 2\sin^2 \beta_0 \sin^2 x\right.\right.$$

$$\left.\left. + \sin^4 \beta_0 \sin^4 x\right) + -- \right] dx \tag{1.50}$$

Now substitute the multiple angle expressions for $\sin^2 x$, $\sin^4 x$, etc.:

$$d\lambda - dL = \frac{e^2}{2} \cos\beta_0 \left[1 + \frac{e^2}{4}\left(1 - \sin^2 \beta_0 \left[\frac{1}{2} - \frac{1}{2}\cos 2x\right]\right) \right.$$

$$+ \frac{e^2}{8}\left(1 - 2\sin^2 \beta_0 \left[\frac{1}{2} - \frac{1}{2}\cos 2x\right] + \sin^4 \beta_0 \left(\frac{3}{8} - \frac{1}{2}\cos 2x\right.\right.$$

$$\left.\left.\left. + \frac{1}{8}\cos 4x\right)\right) + -- \right] dx \tag{1.51}$$

Collecting terms:

$$d\lambda - dL = \frac{e^2}{2} \cos\beta_0 \left[1 + \frac{e^2}{4} + \frac{e^4}{8} - \frac{e^2}{8}\sin^2 \beta_0 - \frac{e^4}{8}\sin^2 \beta_0 + \frac{3}{64} e^4 \sin^4 \beta_0 \right.$$

$$\left. + \left(\frac{e^2}{8}\sin^2 \beta_0 + \frac{e^4}{8}\sin^2 \beta_0 + ---\right)\cos 2x + \frac{e^4}{64}\sin^4 \beta_0 \cos 4x + -- \right] dx \tag{1.52}$$

We may substitute:

$$A' = 1 + \frac{e^2}{4} + \frac{e^4}{8} + \frac{5}{64} e^6 - \left(\frac{e^2}{8} + \frac{e^4}{8} + \frac{15}{128} e^6\right)\sin^2 \beta_0 +$$

$$\left(\frac{3}{64} e^4 + \frac{45}{512} e^6\right)\sin^4 \beta_0 - \frac{25}{1024} e^6 \sin^6 \beta_0$$

$$B' = \left(\frac{e^2}{8} + \frac{e^4}{8} + \frac{15}{128}e^6\right)\sin^2\beta_0 - \left(\frac{e^4}{16} + \frac{15}{128}e^6\right)\sin^4\beta_0 + \frac{75}{2048}e^6\sin^6\beta_0$$

$$C' = \left(\frac{e^4}{64} + \frac{15}{512}e^6\right)\sin^4\beta_0 + \text{---}$$

$$D' = \frac{5}{2024}e^6\sin^6\beta_0 + \text{---}$$

so that (1.52) becomes:

$$d\lambda - dL = \frac{e^2}{2}\cos\beta_0\left[A' + B'\cos2x + C'\cos4x + \text{--}\right]dx \tag{1.53}$$

Integrating, we note:

$$\int_0^\lambda d\lambda = \lambda, \quad \int_0^L dL = L \quad \text{so that:}$$

$$(\lambda - L) = \frac{e^2}{2}\cos\beta_0\int_{\sigma_1}^{\sigma_1+\sigma_T}\left(A' + B'\cos2x + C'\cos4x + \text{--}\right)dx \tag{1.54}$$

The integration required in (1.54) is identical to that in equation (1.33). By inspection we may write the result:

$$(\lambda - L) = \frac{e^2}{2}\cos\beta_0\left[A'\sigma_T + B'\sin\sigma_T\cos2\sigma_m + \frac{C'}{2}\sin2\sigma_T\cos4\sigma_m + \text{--}\right] \tag{1.55}$$

Rainsford (1955) expressed this equation in terms of the flattening, f. Letting $\cos\beta_0 = \sin\alpha$, as before, we re-write equation (1.55), with $\sigma = \sigma_T$, as:

$$(\lambda - L) = f\sin\alpha\left(A_0\sigma + A_2\sin\sigma\cos2\sigma_m + A_4\sin2\sigma\cos4\sigma_m\right.$$

$$\left. + A_6\sin3\sigma\cos6\sigma_m + \text{--}\right) \tag{1.56}$$

where:

$$A_0 = 1 - \frac{1}{4}f\left(1 + f + f^2\right)\cos^2\alpha + \frac{3}{16}f^2\left(1 + \frac{9}{4}f\right)\cos^4\alpha - \frac{25}{128}f^3\cos^6\alpha + \text{--}$$

$$A_2 = \frac{1}{4}f\left(1 + f + f^2\right)\cos^2\alpha - \frac{1}{4}f^2\left(1 + \frac{9}{4}f\right)\cos^4\alpha + \frac{75}{256}f^3\cos^6\alpha + \text{--}$$

$$A_4 = \frac{1}{32} f^2 \left(1 + \frac{9}{4} f\right) \cos^4 \alpha - \frac{15}{256} f^3 \cos^6 \alpha + --$$

$$A_6 = \frac{5}{768} f^3 \cos^6 \alpha + --$$

Certain terms may be dropped from the above coefficients if maximum accuracy is not required.

At this point equation (1.40) for s and equation (1.56) for (λ-L) are the important equations required. We next show how these equations are specifically applied to the inverse and direct problem.

1.21 The Iterative Inverse Problem

We assume we are given the latitude and longitude of the points for which the distance and azimuth are to be obtained. We then have given (ϕ_1, L_1), (ϕ_2, L_2) where all longitudes are positive east. We can now compute the reduced latitude for each of these points using equation (1.9). Next consider Figure 1.6 showing the auxiliary sphere:
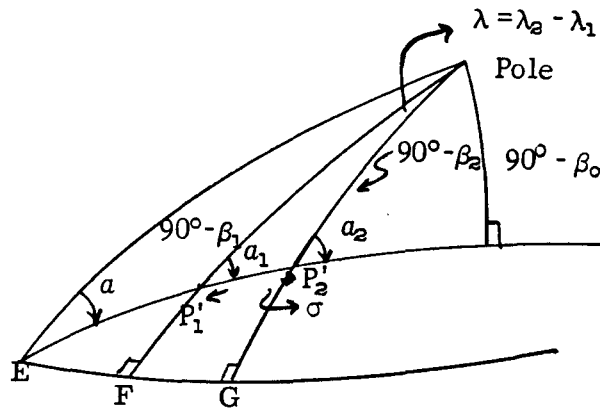
Figure 1.6.
The Auxiliarly Sphere as Used for the Inverse Problem

From the triangle P_1' Pole P_2' we can apply the spherical law of cosines to yield:

$$\cos\sigma = \sin\beta_1 \sin\beta_2 + \cos\beta_1 \cos\beta_2 \cos\lambda \qquad (1.57)$$

This formula weakly determines σ when σ is very small, so that the following equation is recommended (Sodano, 1963) when $\cos\sigma$ is close to one or when both $\sin\sigma$ and $\cos\sigma$ are to be used in subsequent computations.

$$\sin\sigma = \left[\left(\sin\lambda\,\cos\beta_2\right)^2 + \left(\sin\beta_2\cos\beta_1 - \sin\beta_1\cos\beta_2\cos\lambda\right)^2\right]^{1/2}$$

(1.58)

σ can then be determined (with a proper quadrant) using arc tangent subroutines where both $\sin\sigma$ and $\cos\sigma$ are input. If σ is regarded as $\le 180°$, quadrant determination is provided only by (1.57). Starting with the data of the inverse problem we could not evaluate (1.57) or (1.58) since we do not know λ. However, as a first approximation we may let $\lambda = L$ so that an approximate value of σ may be found. Iteration procedures will be described shortly to assure a precise determination of λ and consequently, σ.

In seeking to apply equation (1.56) we need to find α and functions of $2\sigma_m$, as well as have σ. We note from Figure 1.6:

$$\frac{\sin\alpha_1}{\cos\beta_2} = \frac{\sin\lambda}{\sin\sigma}$$

so that:

$$\sin\alpha_1 = \frac{\sin\lambda\,\cos\beta_2}{\sin\sigma}$$

(1.59)

Applying equation (1.1) to the problem of the geodesic (great circle) passing through P_2', P_1' and the point on the equator we have:

$$\sin\alpha_2\cos\beta_2 = \sin\alpha_1\cos\beta_1 = \sin\alpha\,\cos 0°$$

(1.60)

Using (1.59) we may write from (1.60):

$$\sin\alpha = \sin\alpha_1\cos\beta_1 = \frac{\sin\lambda\,\cos\beta_1\,\cos\beta_2}{\sin\sigma}$$

(1.61)

from which we could find $\sin\alpha$ and thus cosine α. In order to find $2\sigma_m$ we first write:

$$\sigma_m = \frac{1}{2}\left(\sigma_1 + \sigma_2\right) = \frac{1}{2}\left(2\sigma_1 + \sigma\right)$$

(1.62)

$$\cos 2\sigma_m = \cos\left(2\sigma_1 + \sigma\right) = \cos 2\sigma_1\cos\sigma - \sin 2\sigma_1\sin\sigma$$

$$= \cos\sigma\left(1 - 2\sin^2\sigma_1\right) - 2\sin\sigma_1\cos\sigma_1\sin\sigma$$

$$= \cos\sigma - 2\sin\sigma_1\left(\sin\sigma_1\cos\sigma + \cos\sigma_1\sin\sigma\right)$$

$$= \cos\sigma - 2\sin\sigma_1\sin\left(\sigma_1 + \sigma\right)$$

(1.63)

Now we can show:

$$\sin\sigma_1 = \sin\beta_1/\cos\alpha \quad \text{(using the law of sines in triangle } P_1'EF) \tag{1.64}$$

and

$$\sin(\sigma_1 + \sigma) = \sin\beta_2/\cos\alpha \quad \text{(using the law of sines in triangle } P_2'EG) \tag{1.65}$$

Then equation (1.63) becomes:

$$\cos2\sigma_m = \cos\sigma - \frac{2\sin\beta_1 \sin\beta_2}{\cos^2\alpha} \tag{1.66}$$

from which we can find $2\sigma_m$, $4\sigma_m$, $6\sigma_m$, etc. using half angle formulas.

With these values we may compute $(\lambda - L)$ from equation (1.56). Recall, however, at this time, the value of $(\lambda - L)$ is not exact as we needed to assume $\lambda = L$ in the initial evaluation of equation (1.57) or (1.58). However, with this new computation we can compute a new, better value of λ by using:

$$\lambda = L + (\lambda - L) \tag{1.67}$$

Using this value we return to (1.57) and (1.58), compute a new σ, find a new α from (1.61), and $2\sigma_m$ from (1.62), and finally a new $(\lambda - L)$ from (1.56). The iteration process is considered complete when the value of $(\lambda - L)$ does not differ by a certain amount from the preceding computed value. The amount may be on the order of 0."0001 to 0."00001 for most applications. The number of iterations to be expected is about 4 although certain special cases to be discussed later will not converge.

At the conclusion of the iteration, we can evaluate equation (1.40) for the distances. In order to determine the azimuths we may use equation (1.60) to write:

$$\sin\alpha_1 = \frac{\sin\alpha}{\cos\beta_1} \tag{1.68}$$

$$\sin\alpha_2 = \frac{\sin\alpha}{\cos\beta_2} \tag{1.69}$$

where α would be that value found from (1.61) at the last iteration for $(\lambda\text{-L})$.

Somewhat more stable equations are recommended by Sodano (1963) for azimuth determinations:

$$\tan\alpha_{12} = \frac{\sin\lambda \cos\beta_2}{\sin\beta_2 \cos\beta_1 - \cos\lambda \sin\beta_1 \cos\beta_2} \tag{1.70}$$

$$\tan\alpha_{21} = \frac{\sin\lambda \cos\beta_1}{\sin\beta_2 \cos\beta_1 \cos\lambda - \sin\beta_1 \cos\beta_2}$$

(1.71)

Proper quadrant determinations for the azimuths can be made by using arc tangent subroutines where the input parameters are sin, and sin/tan. Sodano (1963) points out that for short lines the denominators of (1.70) and (1.71) may be close to zero and therefore he suggests the following alternate forms:

$$\tan\alpha_{12} = \frac{\sin\lambda \cos\beta_2}{\sin\left(\beta_2 - \beta_1\right) + 2\sin\beta_1 \cos\beta_2 \sin^2\frac{\lambda}{2}}$$

(1.72)

$$\tan\alpha_{21} = \frac{\sin\lambda \cos\beta_1}{\sin\left(\beta_2 - \beta_1\right) - 2\cos\beta_1 \sin\beta_2 \sin^2\frac{\lambda}{2}}$$

(1.73)

This completes the discussion of the iterative inverse problem. Maintaining the coefficients given in the (λ-L) and s expressions, the accuracies are on the order of 0."00001 in azimuths and a millimeter in distance for any length lines. This, of course, would assume that all calculations carried the proper number of significant digits. The actual accuracy will depend on the number of series terms carried and the geometry of the line.

1.22 The Iterative Direct Problem

Using the equations previously derived, it is possible to formulate an iterative solution to the direct problem. We assume we are given the following quantities.

ϕ_1, L_1

α_{12}, s geodesic referenced quantities.

Knowing ϕ_1, we may compute the reduced latitude, β_1, of the first point using equation (1.9). In addition we may determine the azimuth (α) at the equator of the geodesic using equation (1.61). The next step requires the computation of σ by an iteration process using the inversion of equation (1.40). We note that we may write from (1.40):

$$\sigma = \frac{s}{bB_0} - \frac{B_2}{B_0} \sin\sigma \cos2\sigma_m - \frac{B_4}{B_0} \sin2\sigma \cos4\sigma_m - \frac{B_6}{B_0} \sin3\sigma \cos6\sigma_m --$$

(1.74)

Noting that the coefficients B_2, B_4, B_6 (which may be computed from the given information) are small, we may write a first approximation to $\sigma = \sigma^0$ as:

$$\sigma^0 = \frac{s}{bB_0} \tag{1.75}$$

In order to iterate for σ in equation (1.74) we must determine $2\sigma_m$. Recalling from (1.22) and immediately preceding it that $2\sigma_m = 2\sigma_1 + \sigma$ we need to know at this point σ_1, as an approximation to σ has been obtained through (1.75). This may be done by using equation (1.21) for $\tan\sigma_1$. We can also find σ_1 using (1.64). We thus have all the information required to iterate equation (1.74) to convergence.

Assuming we now know σ we can apply equation (1.22) to find β_2. We may note here that $\sin\alpha = \cos\beta_0$, so (1.22) may be written:

$$\sin\beta_2 = \sin\left(\sigma_1 + \sigma\right)\cos\alpha$$

Knowing β_2 we can then find ϕ_2. With β_2 found we can find λ by applying equation (1.59) to yield:

$$\sin\lambda = \frac{\sin\sigma\,\sin\alpha_1}{\cos\beta_2} \tag{1.76}$$

We can also use equation (1.57) to determine $\cos\lambda$, which then, in conjunction with (1.76), allows the proper quadrant determination for λ. We then evaluate $(\lambda-L)$ using equation (1.56) and find L by computing:

$$L = \lambda - (\lambda - L) \tag{1.77}$$

Finally the back azimuth may be computed by applying equation (1.73).

We thus see that this form of the direct problem required iteration. This iteration is required in only one equation. Vincenty (1975) has given step by step procedures and compact equations to invoke the procedures described in sections 1.21 and 1.22. These are as follows:

Direct Problem - Given ϕ_1, L_1, α_1, s - Vincenty Formulation

$$\tan\beta = (1 - f)\,\tan\phi$$

$$\tan\sigma_1 = \tan\beta_1/\cos\alpha_1$$

$$\sin\alpha = \cos\beta_1\,\sin\alpha_1$$

$$u^2 = e'^{\,2}\cos^2\alpha$$

$$A = 1 + \frac{u^2}{16384}\left\{4096 + u^2\left[-768 + u^2\left(320 - 175u^2\right)\right]\right\} \tag{1.78}$$

$$B = \frac{u^2}{1024}\left\{256 + u^2\left[-128 + u^2\left(74 - 47\,u^2\right)\right]\right\}$$

$$2\sigma_m = 2\sigma_1 + \sigma$$

$$\Delta\sigma = B\sin\sigma\left\{\cos 2\sigma_m + \frac{1}{4}\,B\left[\cos\sigma\left(-1 + 2\cos^2 2\sigma_m\right)\right.\right.$$

$$\left.\left. - \frac{1}{6}\,B\cos 2\sigma_m\left(-3 + 4\sin^2\sigma\right)\left(-3 + 4\cos^2 2\sigma_m\right)\right]\right\}$$

(1.79)

$$\sigma = \frac{s}{bA} + \Delta\sigma$$

(1.80)

Equation (1.78), (1.79) and (1.80) are iterated until there is a negligible change in σ. The first approximation for σ, needed in (1.79) is taken as the first term in (1.80). The following equations are then evaluated:

$$\tan\phi_2 = \frac{\sin\beta_1\cos\sigma + \cos\beta_1\sin\sigma\cos\alpha_1}{(1 - f)\left[\sin^2\alpha + \left(\sin\beta_1\sin\sigma - \cos\beta_1\cos\sigma\cos\alpha_1\right)^2\right]^{1/2}}$$

(1.81)

$$\tan\lambda = \frac{\sin\sigma\,\sin\alpha_1}{\cos\beta_1\cos\sigma - \sin\beta_1\sin\sigma\cos\alpha_1}$$

(1.82)

$$C = \frac{f}{16}\cos^2\alpha\left[4 + f\left(4 - 3\cos^2\alpha\right)\right]$$

(1.83)

$$L = \lambda - (1 - C)\,f\sin\alpha\left\{\sigma + C\sin\sigma\left[\cos 2\sigma_m + C\cos\sigma\left(-1 + 2\cos^2 2\sigma_m\right)\right]\right\}$$

(1.84)

$$\tan\alpha_2 = \frac{\sin\alpha}{-\sin\beta_1\sin\sigma + \cos\beta_1\cos\sigma\cos\alpha_1}$$

(1.85)

Inverse Problem - Given ϕ_1, L_1, ϕ_2, L_2 - Vincenty formulation.

$$\lambda = L \quad \text{(first approximation)}$$

(1.86)

$$\sin^2\sigma = \left(\cos\beta_2 \sin\lambda\right)^2 + \left(\cos\beta_1 \sin\beta_2 - \sin\beta_1 \cos\beta_2 \cos\lambda\right)^2 \tag{1.87}$$

$$\cos\sigma = \sin\beta_1 \sin\beta_2 + \cos\beta_1 \cos\beta_2 \cos\lambda \tag{1.88}$$

$$\tan\sigma = \frac{\sin\sigma}{\cos\sigma} \tag{1.89}$$

$$\sin\alpha = \frac{\cos\beta_1 \cos\beta_2 \sin\lambda}{\sin\sigma} \tag{1.90}$$

$$\cos2\sigma_m = \cos\sigma - \frac{2\sin\beta_1 \sin\beta_2}{\cos^2\alpha} \tag{1.91}$$

λ is obtained by equation (1.82) or (1.84). This procedure is iterated starting with equation (1.87) until the change in λ is less than some specified value. Then:

$$s = bA\left(\sigma - \Delta\sigma\right) \tag{1.92}$$

where $\Delta\sigma$ is obtained from (1.76), (1.77) and (1.79). Finally:

$$\tan\alpha_1 = \frac{\cos\beta_2 \sin\lambda}{\cos\beta_1 \sin\beta_2 - \sin\beta_1 \cos\beta_2 \cos\lambda} \tag{1.93}$$

$$\tan\alpha_2 = \frac{\cos\beta_1 \sin\lambda}{-\sin\beta_1 \cos\beta_2 + \cos\beta_1 \sin\beta_2 \cos\lambda} \tag{1.94}$$

1.23 Improved Iteration Procedures for the Inverse Problem

Bowring (1983) has discussed several ways in which the iterative inverse problem can be improved by the implementaion of various iteration procedures. Bowring first expresses our equation (1.56) in the following form:

$$\lambda - L = E = (1 - D) f\gamma \left\{\sigma + D\sin\sigma\left[\zeta + D\cos\sigma\left(2\zeta^2 - 1\right)\right]\right\} \tag{1.95}$$

where

$$x = \sin\beta_1 \sin\beta_2$$

$$y = \cos\beta_1 \cos\beta_2$$

$$\bar{x} = \cos\beta_1 \sin\beta_2$$

$$\bar{y} = \sin\beta_1 \cos\beta_2$$

$$\gamma = \frac{y\sin\lambda}{\sin\sigma} = \sin\alpha \text{ (see 1.90)}$$

(1.96)

$$\Gamma = 1 - \gamma^2$$

$$\zeta = \cos\sigma - \frac{2x}{\Gamma}$$

$$\bar{\zeta} = \Gamma\cos\sigma - x$$

$$D = \frac{1}{16}\, f\Gamma\left(4 + 4f - 3f\Gamma\right)$$

With this notation the simple interation procedure previously discussed could be written as:

$$\lambda_{n+1} = L + E\left(\lambda_n\right)$$

(1.97)

where $\lambda_0 = L$.

The Newton-Raphson method can be first implemented by writing the ideal function:

$$F\left(\lambda\right) = \lambda - L - E\left(\lambda\right) = 0$$

(1.98)

We differentiate (1.98) with respect to λ:

$$F'\left(\lambda\right) = 1 - E'\left(\lambda\right)$$

(1.99)

Then the Newton-Raphson procedure yields the following iterative procedure:

$$\lambda_{n+1} = \lambda_n - \frac{F\left(\lambda_n\right)}{F'\left(\lambda_n\right)}$$

(1.100)

This can be written as:

$$\lambda_{n+1} = \lambda_n - \frac{\left[\lambda_n - L - E\left(\lambda_n\right)\right]}{\left[1 - E'\left(\lambda_n\right)\right]}$$

(1.101)

where

$$E'(\lambda) = f\left(\gamma^2 + \frac{\overline{\zeta}\sigma}{\sin\sigma}\right)$$

(1.102)

Bowring also discusses an extended Newton-Raphson procedure and Lagrange's method. The extended Newton-Raphson procedure uses first and second derivatives of $F(\lambda)$. It should be more accurate than the simple Newton-Raphson procedure. The Lagrangian method creates a series expansion for λ of the following form:

$$\lambda = L + E(L) + E(L) E'(L) + \frac{1}{2} E^2(L) E''(L) + E(L)\left(E'(L)\right)^2$$

(1.103)

Note that the right hand side of this equation is a function of L alone and this in reality is a non-iterative procedure.

These improved procedures have been tested for a series of lines described in section 1.7. Results show that the number of iterations required in the Newton-Raphson procedure is about half that of simple iteration. This is done with a reduction of computer time needs by about 20%. The extended Newton-Raphson procedure shows a small improvement over the Newton-Raphson procedure.

The Lagrange method gave no iterations but yielded results that were not as accurate as the other methods. It seems clear that the software for the iterative inverse problem should include either the simple or the extended Newton-Raphson procedure.

The methods described in this section have not been applied to the iterative direct problem. This may not be necessary because of the existence of accurate, non-iterative inverse problem procedures to be discussed later.

1.24 The Non-Iterative Direct Problem

There are several solutions to the direct problem that are quite accurate and require no iteration. Papers of interest include those of McCaw (1930), referenced in Rainsford (1955), a report by Sodano and Robinson (1963) that expands a report of Sodano (1963), and a thesis by Singh (1980) that discusses a non-iterative procedure based on some McCaw procedures. For the purposes of this text we examine first the principles involved with the McCaw solution with more detailed discussion being found in Ganshin (1969, p.86) or Singh (1980).

McCaw's solution also uses an auxiliary sphere for computational purposes. But this sphere is used such that a point on the ellipsoid with latitude ϕ, has a corresponding point on the sphere with the same geodetic latitude. With this correspondence the longitude difference on the sphere must be different than on the ellipsoid, and the azimuths on the sphere will differ from the corresponding azimuths on the ellipsoid. To show the relationship between the azimuths we first write equation (1.1):

$$\cos\beta_1 \sin\alpha_1 = \cos\beta_2 \sin\alpha_2 = \cos\beta_0 \sin\alpha$$

(1.1)

On the McCaw sphere, the corresponding equation will be:

$$\cos\phi_1 \sin\alpha_1^* = \cos\phi_2 \sin\alpha_2^* = \cos\phi_0 = \sin\alpha^* \tag{1.104}$$

where the α^* are called the reduced azimuths of the geodesic line. Note that an α^* corresponds to an (α) used in the Rainsford (1955) paper. Now we know that:

$$\cos\phi = \frac{\left(1 - e^2\right)^{1/2} \cos\beta}{\left(1 - e^2 \cos^2\beta\right)^{1/2}} \tag{1.105}$$

At ϕ_0, (1.105) becomes:

$$\cos\phi_0 = \frac{\left(1 - e^2\right)^{1/2} \cos\beta_0}{\left(1 - e^2 \cos^2\beta_0\right)^{1/2}} = \frac{\left(1 - e^2\right)^{1/2} \cos\beta_0}{\left(1 - e^2 \sin^2\alpha\right)^{1/2}} \tag{1.106}$$

Then from (1.1), (1.104), and (1.105) we have:

$$\frac{\sin\alpha}{\sin\alpha^*} = \frac{\cos\beta_0}{\cos\phi_0} = \frac{\left(1 - e^2 \sin^2\alpha\right)^{1/2}}{\left(1 - e^2\right)^{1/2}} \tag{1.107}$$

Now we solve (1.107) for $\sin\alpha^*$ and substitute it into (1.104) to find:

$$\cos\phi_1 \sin\alpha_1^* = \frac{\sin\alpha\left(1 - e^2\right)^{1/2}}{\left(1 - e^2 \sin^2\alpha\right)^{1/2}} \tag{1.108}$$

Substituting on the left side of (1.108) for $\cos\phi_1$ from (1.105), dividing the left side by $\sin\alpha_1$ $\cos\beta_1$ and the right side by $\sin\alpha$ we find:

$$\frac{\sin\alpha_1}{\sin\alpha_1^*} = \frac{\left(1 - e^2 \sin^2\alpha\right)^{1/2}}{\left(1 - e^2 \cos^2\beta_1\right)^{1/2}} \tag{1.109}$$

Squaring (1.109), substituting for $\sin^2\alpha$ and $\sin^2\alpha_1^*$ by $1 - \cos^2\alpha_1^*$ and substituting for e^2 in terms of e'^2 we find:

$$\cos\alpha_1 = k \cos\alpha_1^*$$

<div align="right">(1.110)</div>

where

$$k = \frac{\left(1 + e'^2 \cos^2\alpha\right)^{1/2}}{\left(1 + e'^2\right)^{1/2}}$$

Equation (1.110) is valid for a point on the geodesic under consideration. Now, since $\sin\alpha = \cos\beta_0$ we can write k in the form:

$$k^2 = 1 - e^2 \cos^2\beta_0 \equiv \frac{1}{V_0^2}$$

so that

$$\cos\alpha_1^* = V_0 \cos\alpha_1$$

<div align="right">(1.111)</div>

We can also show that:

$$\sin\alpha_1^* = \frac{V_0}{V} \sin\alpha_1$$

<div align="right">(1.112)</div>

The method of solution of the problem from this point may be found in McCaw (1930) or in Ganshin. Singh discusses the general philosophy of the mapping from the ellipsoid to a sphere and the development of equations similar to the above for different mappings. The general mapping is represented by:

$$\tan\eta = J\tan\phi$$

<div align="right">(1.113)</div>

where η is the auxiliary latitude on the sphere and $\bar{\alpha}$ is the corresponding auxiliary azimuth. Singh develops the differential relationships between s (the distance on the geodesic) and σ, and L and λ. We have:

$$\frac{ds}{d\sigma} = \frac{a\left(1 - e^2\right)}{J} \frac{\left(1 + G_1 \sin^2\eta\right)^{1/2}}{\left(1 + G_2 \sin^2\alpha\right)^{1/2}} \frac{1}{\left(1 + G_2 \sin^2\eta\right)^{3/2}}$$

<div align="right">(1.114)</div>

where

$$G_1 = \frac{\left(1 - J^2\right)}{J^2}$$

$$G_2 = \frac{\left(1 - e^2 - J^2\right)}{J^2}$$

$$\sin\eta = \sin\sigma \cos\overline{\alpha} \tag{1.115}$$

The longitude relationship is:

$$dL - d\lambda = \frac{\left[1 + \left(J^2\left(1 + e'^2\right) - 1\right)\cos^2\eta\right]^{1/2}\left[\left(1 + \left(J^2-1\right)\cos^2\eta\right)^{-1/2}\right]\sin\overline{\alpha}d\sigma}{\cos^2\eta} \tag{1.116}$$

The mapping $J = 1$ corresponds to the McCaw case; $J = (1 - e^2)^{1/2}$ corresponds to the classical (Bessel, Helmert) case; and $J = (1 - e^2)$ corresponds to the case of the auxiliary latitude being the geocentric latitude.

The equations of the original McCaw solution were re-cast by Rainsford (1955) and put into the following computational form given ϕ_1, L_1, α_{12}, and s:

$$\tan\beta = \left(1 - e^2\right)^{1/2}\tan\phi \tag{1.9}$$

$$\sin\alpha = \sin\alpha_1 \cos\beta_1 \tag{1.60}$$

$$u^2 = e'^2 \cos^2\alpha \tag{1.117}$$

$$k^2 = \frac{\left(1 + u^2\right)}{\left(1 + e'^2\right)} \tag{1.118}$$

$$\tan G_1 = \frac{k\tan\phi_1}{\cos\alpha_1} \tag{1.119}$$

$$K = \frac{\sqrt{\left(1 + u^2\right)}}{b} \tag{1.120}$$

$$\gamma = KC_0 s \tag{1.121}$$

$$\gamma_1 = G_1 - C_2 \sin 2G_1 + C_4 \sin 4G_1 - C_6 \sin 6G_1 \tag{1.122}$$

$$\gamma_2 = \gamma_1 + \gamma \tag{1.123}$$

$$2\gamma_m = \gamma_1 + \gamma_2 \tag{1.124}$$

$$G = \gamma + D_2 \sin\gamma \cos2\gamma_m + D_4 \sin2\gamma \cos4\gamma_m + D_6 \sin3\gamma \cos6\gamma_m \tag{1.125}$$

$$G_2 = G_1 + G, \quad 2G_m = G_1 + G_2 \tag{1.126}$$

$$C_0 = 1 - \frac{3}{4}u^2 + \frac{39}{64}u^4 - \frac{133}{256}u^6 + \frac{7491}{16384}u^8 \tag{1.127}$$

$$C_2 = \frac{3}{8}u^2 - \frac{3}{16}u^4 + \frac{111}{1024}u^6 - \frac{141}{2048}u^8$$

$$C_4 = \frac{15}{256}u^4 - \frac{15}{256}u^6 + \frac{405}{8192}u^8$$

$$C_6 = \frac{35}{256}u^6 - \frac{105}{6144}u^8$$

$$C_8 = \frac{315}{131072}u^8$$

$$D_2 = \frac{3}{4}u^2 - \frac{3}{8}u^4 + \frac{213}{1024}u^6 - \frac{255}{2048}u^8 \tag{1.128}$$

$$D_4 = \frac{21}{128}u^4 - \frac{21}{128}u^6 + \frac{1599}{12288}u^8$$

$$D_6 = \frac{151}{3072}u^6 - \frac{453}{6144}u^8$$

$$D_8 = \frac{1097}{65536}u^8$$

$$\sin\phi_2 = \frac{\sin G_2 \cos\alpha}{k} \tag{1.129}$$

$$\cos\alpha_2 = k \cot G_2 \tan\phi_2 \tag{1.130}$$

$$\cos\lambda = \frac{\left(\cos G - \sin\phi_1 \sin\phi_2\right)}{\cos\phi_1 \cos\phi_2}; \quad \sin\lambda = \frac{\sin G \sqrt{\left(k^2 - \cos^2\alpha_1\right)}}{k\cos\phi_2} \tag{1.131}$$

$$(\lambda\text{-}L) = f\sin\alpha \left(E_0 G - E_2 \sin G \cos 2G_m + E_4 \sin 2G \cos 4G_m\right.$$

$$\left. - E_6 \sin 3G \cos 6G_m\right) \tag{1.132}$$

$$E_0 = 1 - \frac{1}{4}f\left(1 + f + f^2\right)\cos^2\alpha + \frac{3}{16}f^2\left(1 + \frac{9}{4}f\right)\cos^4\alpha - \frac{25}{128}f^3\cos^6\alpha \tag{1.133}$$

$$E_2 = \frac{1}{4}f\left(3 + 5f + 7f^2\right)\cos^2\alpha - f^2\left(1 + \frac{49}{16}f\right)\cos^4\alpha + \frac{365}{256}f^3\cos^6\alpha$$

$$E_4 = \frac{5}{32}f^2\left(1 + \frac{13}{4}f\right)\cos^4\alpha - \frac{95}{256}f^3\cos^6\alpha$$

$$E_6 = \frac{35}{768}f^3\cos^6\alpha$$

We thus have found ϕ_2 from equation (1.129), the azimuth at the second point from equation (1.130), and the longitude of the second point by using:

$$L_2 = L_1 + \lambda - (\lambda - L) \tag{1.134}$$

The accuracy of these equations is fully compatible with the set used in the iterative inverse problem.

Another version of the non-iterative direct problem has been described by Singh (ibid) based on some procedures developed by McCaw applied to the original iterative solution discussed in section 1.21. To consider the new procedure we start with equation (1.33) written in the following indefinite form:

$$\frac{s}{b} = \int \left(A + B\cos 2x + C\cos 4x + \ldots\right)dx \tag{1.135}$$

Integrating this we have:

$$\frac{s}{b} = \left(Ax + \frac{B}{2}\sin 2x + \frac{C}{4}\sin 4x + \ldots\right) \tag{1.136}$$

Divide each side by A to write:

$$\frac{s}{b}C_0 = x + C_2\sin 2x + C_4\sin 4x, \quad \text{where} \tag{1.137}$$

$$C_0 = \frac{1}{A}, \quad C_2 = \frac{B}{2A}; \quad C_4 = \frac{C}{4A}, \text{ etc.}$$

Note that the C values appearing in (1.137) are not the C values defined in (1.127). Similar notice should be taken for the D values to be defined in (1.142) which are not the same as the D values in (1.128).

Now evaluate (1.137) between $\sigma = 0$ and $\sigma = \sigma_1$ where $s = s_1$. Then we define γ_1 which becomes

$$\gamma_1 \equiv \frac{s_1 C_0}{b} = \sigma_1 + C_2 \sin2\sigma_1 + C_4 \sin4\sigma_1 + \ldots \tag{1.138}$$

Now evaluate (1.137) for the distance 0 to $s + s_1 = s_2$ where s is the length of the line between the two points of interest. We have:

$$\frac{C_0 (s+s_1)}{b} = \gamma + \gamma_1 = \sigma_2 + C_2 \sin2\sigma_2 + C_4 \sin4\sigma_2 + \ldots \tag{1.139}$$

where σ_2 is the arc corresponding to $s + s_1$. Here $\gamma = C_0 s/b$ and would be a known quantity in the direct problem. Now perform a series inversion (Rapp, 1984) of (1.138) and (1.139) to find:

$$\sigma_1 = \gamma_1 + \overline{C}_2 \sin2\gamma_1 + \overline{C}_4 \sin4\gamma_1 + \overline{C}_6 \sin6\gamma_1 + \ldots \tag{1.140}$$

$$\sigma_2 = \left(\gamma + \gamma_1\right) + \overline{C}_2 \sin2\left(\gamma + \gamma_1\right) + \overline{C}_4 \sin4\left(\gamma + \gamma_1\right) + \ldots \tag{1.141}$$

The arc between the two points is $\sigma = \sigma_2 - \sigma_1$ which can be found by differencing (1.140) and (1.141) and using (1.35). We have:

$$\sigma = \gamma + D_2 \sin\gamma \cos2\gamma_m + D_4 \sin2\gamma \cos4\gamma_m + D_6 \sin3\gamma \cos6\gamma_m + \ldots \tag{1.142}$$

where:

$$2\gamma_m = \gamma + 2\gamma_1 \tag{1.143}$$

$$D_2 = \frac{1}{4} u^2 - \frac{1}{8} u^4 + \frac{71}{1024} u^6 - \frac{85}{2048} u^8 +$$

$$D_4 = \frac{5}{128} u^4 - \frac{5}{128} u^6 + \frac{383}{12288} u^8 + \ldots$$

$$D_6 = \frac{29}{3072} u^6 - \frac{29}{2048} u^8 + \ldots \tag{1.144}$$

$$D_8 = \frac{539}{196608} u^8 + \ldots$$

$$u^2 = e'^2 \sin^2\beta_0$$

In the actual computations for the direct problem using the Singh procedure the value of σ_1 is found using equation (1.21). Knowing u^2 the C and D coefficients can be computed. Then find

$\gamma(=C_0s/b)$ and γ_1 from (1.138). Using (1.143) we can then find σ from (1.142) after which the usual equations developed for the iterative direct solution can be used.

Numerical tests conducted by Singh indicated the procedures give accuracies equivalent to the iterative procedure. Due to the way in which the inverse problem is developed the procedure of Singh does not appear to be applicable. However other techniques are available as discussed in the next section.

1.3 The Non-Iterative Inverse Problem

The computation of an iterative solution to a high accuracy can be time consuming. Requirements for a non-iterative approach led Sodano (1958) to the development of such a system. In the following paragraphs we outline the method of derivation and present working formulas.

If we consider (1.56) we see that it can be written in the form:

$$\lambda = L + x \tag{1.145}$$

where x is a small quantity equal to the right-hand side of (1.56). We may use (1.145) wherever the value of λ is required. For example, we need $\cos\lambda$ in equation (1.57). We may write:

$$\cos\lambda = \cos(L + x) = \cos L \cos x - \sin L \sin x$$

$$\cos\lambda = \cos L \left(1 - \frac{x^2}{2} + \cdots\right) - (\sin L)\left(x - \frac{x^3}{3!} + \cdots\right)$$

and finally:

$$\cos\lambda = \cos L - (\sin L) x - \left(\frac{1}{2}\cos L\right)x^2 + \cdots \tag{1.146}$$

The process of developing the non-iterative procedure consists of substituting series such as (1.146) and all subsequent series into the usual iterative procedures. For example, equation (1.57) could be written:

$$\cos\sigma = \sin\beta_1 \sin\beta_2 + \cos\beta_1 \cos\beta_2 \left(\cos L - \sin L x - \frac{1}{2}\cos L x^2 \cdots\right) \tag{1.147}$$

$$= \sin\beta_1 \sin\beta_2 + \cos\beta_1 \cos\beta_2 \cos L$$

$$- \cos\beta_1 \cos\beta_2 \sin L \ x$$

$$- \frac{1}{2}\cos\beta_1 \cos\beta_2 \cos L \ x^2$$

If we let

$$\cos\sigma = \cos\sigma_0 - \cos\beta_1 \cos\beta_2 \sin L \ x - \frac{1}{2}\cos\beta_1 \cos\beta_2 \cos L \ x^2 + --$$

(1.148)

which may be written:

$$\cos\sigma = \cos\sigma_0 + k_1 x + k_2 x^2 + ---$$

where k_1, k_2 are appropriate constants that may be read from equation (1.148). We could continue writing:

$$\sigma = \sigma_0 + k_3 x + k_4 x^2 + --$$

(1.149)

$$\sin\lambda = \sin L + k_5 x + k_6 x^2 = --$$

(1.150)

Continuing through the equations we find equation (1.56) may be written in the form:

$$(\lambda\text{-}L) = k_7 + k_8 x + k_9 x^2 + --$$

(1.151)

where k_7, k_8, and k_9 are complicated expressions. Now we note that from (1.145) $\lambda - L = x$ or using (1.151):

$$(\lambda\text{-}L) = x = k_7 + k_8 x + k_9 x^2$$

(1.152)

Equation (1.152) may then be solved for x to yield:

$$x = k_7\left(1 + k_8 + k_8^2 + k_7 k_9\right)$$

(1.153)

Since we know expressions for k_7, and k_8, and k_9 it is possible to develop an algebraic expression for x or (λ-L) without recourse to iteration. Before we give this expression we may note that it is also possible to modify the distance expression, equation (1.40), by using the series expressions for σ, or $\sin\sigma$ and its multiples, that are a function of the parameter x. This expression will be a function of the ellipsoidal longitude difference as opposed to equation (1.40), which is basically a function of the longitude difference on the auxiliary sphere. In this case we could write:

$$s = b\left(k_{10} + k_{11}x + k_{12}x^2 + --\right)$$

(1.154)

It is also possible to develop expressions for the azimuths that will be a function of the ellipsoidal longitude difference and the parameter x. Although these expressions have been developed by Sodano, they are not specifically required as previously derived expressions may be used since we will have the value of λ using the x value found from equation (1.153).

Once the value of (λ-L) has been established through equation (1.153) we may go back and find σ from equation (1.57) or (1.58), proceed to find the azimuths as in (1.59) and (1.60), and finally the distance from equation (1.40). In the latter case, however, an alternative is to use equation (1.154) for s. This is accomplished by algebraically substituting the expression for x

found from equation (1.153) into equation (1.154). Although algebraically complex, the result is a fairly patterned equation.

Sodano (1965) published the following recommended working equations for his non-iterative solution. These equations, given to the order of f^3, are as follows for the inverse solution:

$$a = \sin\beta_1 \sin\beta_2$$
$$b = \cos\beta_1 \cos\beta_2 \tag{1.155}$$

$$\cos\Phi = a + b \cos L$$

$$\sin\Phi = \left[\left(\sin L \cos\beta_2\right)^2 + \left(\sin\beta_2 \cos\beta_1 - \sin\beta_1 \cos\beta_2 \cos L\right)^2\right]^{1/2} \tag{1.156}$$

These equations should be compared with (1.57) and (1.58) where the only difference is seen to be the replacement of λ by L to obtain (1.156).

Next define:

$$c = \frac{b \sin L}{\sin\Phi}.$$

$$m = 1 - c^2 \tag{1.157}$$

Then the following equations for s as taken from Sodano and Robinson (1963) are

$$\frac{s}{b_0} = \left(1 + f + f^2 + f^3\right)\Phi$$

$$+ a\left[\left(f + f^2 + f^3\right)\sin\Phi + \left(-\frac{1}{2}f^2 - f^3\right)\Phi^2 \csc\Phi + \frac{1}{2}f^3\Phi^3 \csc\Phi \cot\Phi\right]$$

$$+ m\left[\left(-\frac{1}{2}f - \frac{1}{2}f^2 - \frac{1}{2}f^3\right)\Phi + \left(-\frac{1}{2}f - \frac{1}{2}f^2 - \frac{1}{2}f^3\right)\sin\Phi \cos\Phi + \right.$$

$$+ \left(\frac{1}{2}f^2 + f^3\right)\Phi^2 \cot\Phi - \frac{1}{6}f^3\Phi^3 - \frac{1}{2}f^3\Phi^3\cot^2\Phi\Bigg]$$

$$+ a^2\left[\left(-\frac{1}{2}f^2 - f^3\right)\sin\Phi\cos\Phi + \frac{1}{2}f^3\Phi^3\csc^2\Phi + \frac{1}{2}f^3\Phi\right]$$

$$+ m^2\left[\left(\frac{1}{16}f^2 + \frac{1}{8}f^3\right)\Phi + \left(\frac{1}{16}f^2 + \frac{1}{8}f^3\right)\sin\Phi\cos\Phi\right.$$

$$+ \left(-\frac{1}{2}f^2 - \frac{7}{4}f^3\right)\Phi^2\cot\Phi + \left(-\frac{1}{8}f^2 - \frac{1}{4}f^3\right)\sin\Phi\cos^3\Phi$$

$$\left. + \frac{1}{4}f^3\Phi\cos^2\Phi + \frac{1}{3}f^3\Phi^3 + \frac{3}{2}f^3\Phi^3\cot^2\Phi\right] \qquad\qquad (1.158)$$

$$+ am\left[\left(\frac{1}{2}f^2 + \frac{7}{4}f^3\right)\Phi^2\csc\Phi + \left(\frac{1}{2}f^2 + f^3\right)\sin\Phi\cos^2\Phi\right.$$

$$\left. - \frac{3}{4}f^3\Phi\cos\Phi - 2f^3\Phi^3\csc\Phi\cot\Phi\right]$$

$$+ a^2m\left[-\frac{1}{2}f^3\Phi - \frac{1}{2}f^3\sin\Phi\cos\Phi - \frac{1}{2}f^3\Phi^3\csc^2\Phi + f^3\sin^3\Phi\cos\Phi\right]$$

$$+ am^2\left[-\frac{3}{4}f^3\Phi^2\csc\Phi + \frac{1}{2}f^3\sin\Phi\cos^2\Phi + \frac{3}{4}f^3\Phi\cos\Phi\right.$$

$$\left. + \frac{3}{2}f^3\Phi^2\csc\Phi\cot\Phi - \frac{1}{2}f^3\sin\Phi + \frac{1}{2}f^3\sin^5\Phi\right]$$

$$+ m^3\left[-\frac{1}{32}f^3\Phi + \frac{3}{4}f^3\Phi^2\cot\Phi - \frac{1}{32}f^3\sin\Phi\cos\Phi\right]$$

$$+ \frac{1}{16}f^3\sin\Phi\cos^3\Phi - \frac{1}{4}f^3\Phi\cos^2\Phi - \frac{1}{6}f^3\Phi^3 - f^3\Phi^3\cot^2\Phi$$

$$+ \frac{1}{12}f^3\sin^3\Phi\cos^3\Phi$$

$$+ a^3\left[\frac{1}{2}f^3\sin\Phi - \frac{2}{3}f^3\sin^3\Phi\right]$$

In addition:

$$\frac{\lambda \text{-L}}{c} = \left[\left(f + f^2 + f^3\right)\Phi\right] + a\left[\left(-\frac{1}{2}f^2 - f^3\right)\sin\Phi + \left(-f^2 - 4f^3\right)\Phi^2\csc\Phi\right.$$

(1.159)

$$\left. + \frac{3}{2}f^3\Phi^3\csc\Phi\cot\Phi\right] + m\left[\left(-\frac{5}{4}f^2 - 3f^3\right)\Phi + \left(\frac{1}{4}f^2 + \frac{1}{2}f^3\right)\sin\Phi\cos\Phi\right.$$

$$\left. + \left(f^2 + 4f^3\right)\Phi^2\cot\Phi - \frac{1}{2}f^3\Phi^3 - \frac{3}{2}f^3\Phi^3\cot^2\Phi\right]$$

$$+ m^2\left[\frac{31}{16}f^3\Phi - \frac{7}{16}f^3\sin\Phi\cos\Phi + \frac{1}{2}f^3\Phi^3 - \frac{1}{8}f^3\sin^3\Phi\cos\Phi\right.$$

$$\left. - \frac{9}{2}f^3\Phi^2\cot\Phi + \frac{1}{2}f^3\Phi\cos^2\Phi + \frac{5}{2}f^3\Phi^3\cot^2\Phi\right]$$

$$+ am\left[\frac{9}{2}f^3\Phi^2\csc\Phi - \frac{3}{2}f^3\Phi\cos\Phi - \frac{7}{2}f^3\Phi^3\csc\Phi\cot\Phi - \frac{f^3}{2}\sin\Phi\cos^2\Phi + f^3\sin\Phi\right]$$

$$+ a^2\left[f^3\Phi + \frac{1}{2}f^3\sin\Phi\cos\Phi + f^3\Phi^3\csc^2\Phi\right]$$

Finding the value of λ from (1.159) we may use equation (1.70) and (1.71) to find the required azimuths.

In the development of the Sodano non-iterative equations a problem arose in numerically checking the iterative inverse problems with lines whose σ value was nearly 180°. Such lines are called anti-podal lines, or near anti-podal lines. The discrepancies that arose were caused by the increase of some terms in equations similar to (1.158) and (1.159). This may be seen from these equations in the terms involving $\csc\Phi$ and $\cot\Phi$. As Φ approaches 180° these terms become quite large, and in the limit go to infinity. Examination of equations (1.151) and (1.154) would show that the rapid increase in certain terms does not occur in the constant coefficients (e.g. k_7 or k_{10}), but in the coefficients of x. Thus it was reasoned that if x could be made sufficiently small the increase previously noted would be balanced out. These problems do not occur when Φ approaches zero because the terms Φ and $\sin\Phi$ will approach zero.

To this end, equation (1.145) may be reformulated to read:

$$\lambda = L_n + z$$

(1.160)

where L_n is a value closer to λ than L, and z is a value smaller than x. We could take the value of L_n to be that value of λ calculated using equation (1.145). Thus, we could write:

$$\lambda = (L + x) + z \qquad (1.161)$$

Using equation (1.161) the complete procedure deriving the non-iterative procedure may be repeated, this time with the equations being a function of L_n instead of L. Thus considering equation (1.56):

$$(\lambda - L) = f\sin\alpha \left(A_0\sigma + \text{---}\right) = L_n + z - L \qquad (1.162)$$

so that

$$(L - L_n) + f\sin\alpha \left(A_0\sigma + \text{---}\right) = z \qquad (1.163)$$

Expressing the series terms in the manner of developing equation (1.151), we will have $(\lambda\text{-}L)$ as a function of L_n and z. In fact the series expressions will be the same as previously except that the coefficients will be a function of L_n instead of L, and of z instead of x. We have:

$$f\sin\alpha \left(A_0\sigma + \text{---}\right) = F\left(L_n, z\right) \qquad (1.164)$$

A solution directly for z may be obtained in a manner similar to that expressed in (1.153). Carrying these computations out, Sodano found the following:

$$z = \frac{16(L\text{-}L_n) + \left(16e^2 Nc\Phi - e^2hc\Phi - e^2hc\sin\Phi\cos\Phi + 2e^2e'^2cP\sin^2\Phi\right)_n}{16\left(1 - e^2Nc^2 - e^2NP\Phi\right)_n} \qquad (1.165)$$

where

$$N = \frac{e'}{\left(e' + e\right)}$$

$$h = e'^2 m$$

$$P = \left(1 - c^2\right)\cot\Phi - a\csc\Phi \qquad (1.166)$$

where the subscript n signifies that all evaluations must be made with the value of L_n instead of L.

The procedure in applying the non-iterative inverse problem of Sodano for near anti-podal lines is to first find the value of x using the right side of (1.157). This will give a value of $L_n=\lambda$ which is then used in (1.165) to find z and consequently the better value of λ through equation (1.161).

The value of z primarily depends on the length of the line. For lines under 170° in arc length z can be on the order of 0."001. However for lines whose length is a degree or so less than 180°, z

can reach 4 or 5" depending on azimuth, starting latitude, etc. The Sodano procedure fails for lines in the antipodal region that is described later on.

Sodano has also applied his reduction process to the direct problem. Equations for this computation given to $O(f^2)$ are described in Sodano (1965). Extension of the equations to terms of f^3 may be found in Sodano (1963). The application of these equations could be compared with those of McCaw and Singh. The critical development of the Sodano was in the area of the non-iterative inverse problem.

1.4 A Numerical Integration Approach to the Solution of the Direct and Inverse Problem.

In previous sections we were concerned with solutions obtained by the integration, through series expansions of equations (1.28) and (1.53). An alternate procedure has been described by Saito (1970) where the needed integrations are carried out numerically. To develop this procedure we re-write equation (1.46) in the following form:

$$dL - d\lambda = \cos\beta_0 \left[\frac{\left(1 e^2 \cos^2\beta\right)^{1/2} - 1}{\cos^2\beta} \right] d\sigma$$

$$(1.167)$$

We next multiply the numerator and denominator on the right hand side of (1.167) by $(1- e^2 \cos^2\beta)^{1/2} + 1$ to obtain:

$$d\lambda - dL = \frac{e^2 \cos\beta_0 \, d\sigma}{\left(1 - e^2 \cos^2\beta\right)^{1/2} + 1}$$

$$(1.168)$$

We now let $x = \sigma_1 + \sigma$ so that $dx = d\sigma$. Using (1.26), (1.168) becomes:

$$d\lambda - dL = \frac{e^2 \cos\beta_0 \, dx}{1 + \frac{b}{a}\left(1 + k^2 \sin^2 x\right)^{1/2} + 1}$$

$$(1.169)$$

where k^2 is defined in equation (1.26). Integrating (1.169) we have:

$$\lambda - L = e^2 \cos\beta_0 \int_{\sigma_1}^{\sigma_1 + \sigma_T} \frac{dx}{1 + \frac{b}{a}\left(1 + k^2 \sin^2 x\right)^{1/2}}$$

$$(1.170)$$

In order to normalize the interval of integration we define a quantity z so that:

$$x = \sigma_1 + \sigma_T z \text{ with } dx = \sigma_T dz \tag{1.171}$$

with $0 \le z \le 1$. Then we can write equations (1.28) and (1.170):

$$s = b\sigma_T \int_0^1 \sqrt{1 + k^2 \sin^2(\sigma_1 + \sigma_T z)} \, dz \tag{1.172}$$

$$\lambda - L = e^2 \cos\beta_0 \, \sigma_T \int_0^1 \frac{dz}{1 + \frac{b}{a}\sqrt{1 + k^2 \sin^2(\sigma_1 + \sigma_T z)}} \tag{1.173}$$

The integrals in (1.172) and (1.173) are in a form that can be numerically integrated using any appropriate numerical integration method of sufficient accuracy.

Saito (1979) has also discussed a specific numerical integration procedure to use for evaluation of equations such as (1.170), (1.172) or (1.173) considering Gaussian quadrature formulas. To do this we introduce a new quantity z' so that:

$$x = \sigma_m + \frac{\sigma}{2} z' \tag{1.174}$$

where σ_m is given in (1.37). Equations (1.28) and (1.170) can then be written as:

$$s = \frac{b}{2}\sigma_T \int_{-1}^1 \sqrt{1 + k^2 \sin^2\left(\sigma_m + \frac{\sigma_T}{2}z'\right)} \, dz' \tag{1.175}$$

$$\lambda - L = \frac{e^2 \cos\beta_0 \, \sigma_T}{2} \int_{-1}^1 \frac{dz'}{1 + \frac{b}{a}\sqrt{1 + k^2 \sin^2\left(\sigma_m + \frac{\sigma_T}{2}z'\right)}} \tag{1.176}$$

Now the Gaussian quadrature procedure applied to $f(x)$ can be written:

$$\int_{-1}^1 f(x) \, dx \approx \sum_{k=1}^n w_k \, f(x_k) \tag{1.177}$$

where w_k are the weights and the x_k are the corresponding nodes. The accuracy of the evaluation will depend on n. With this formulation equation (1.175) and (1.176) can be written as:

$$s = \frac{b}{2} \sigma_T G$$

(1.178)

with

$$G = \sum_{i=1}^{n} w_i \sqrt{1+k^2 \sin^2\left(\sigma_m + \frac{\sigma}{2} z_i'\right)}$$

$$\lambda - L = \frac{F}{2} e^2 \sigma_T \cos\beta_0$$

(1.179)

$$F = \sum_{i=1}^{n} \frac{w_i}{1 + \frac{b}{a}\sqrt{1+k^2 \sin^2\left(\sigma_m + \frac{\sigma}{2} z_i'\right)}}$$

(1.180)

For the case of n=8 we have the following weights and nodes (Saito, 1979):

i	z_i'				w_i			
1	.96028	98564	97536	23168	.10122	85362	90376	25915
2	.79666	64774	13626	73959	.22238	10344	53374	47054
3	.52553	24099	16328	98582	.31370	66458	77887	28734
4	.18343	46424	95649	80494	.36268	37833	78361	98297
5	$-z_4'$				w_4			
6	$-z_3'$				w_3			
7	$-z_2'$				w_2			
8	$-z_1'$				w_1			

We conclude this section by giving a step by step procedure for the solution of the direct and inverse problem using this integration procedure.

For the inverse problem: Given: ϕ_1, ϕ_2, L_1, L_2

Steps

1. $L = L_2 - L_1$
2. Compute β_1, β_2 from (1.9)
3. Assume $\lambda = L$
4. Compute σ_T using (1.57) or (1.58)
5. Compute $(\cos\beta_0 = \sin\alpha)$ from (1.61)
6. Compute k^2 from (1.26)
7. Compute σ_1 using (1.23) with $\sigma = \sigma_T$ and $\beta = \beta_2$
8. Evaluate (1.173) or (1.179) to find $(\lambda - L)$
9. Using the result in 8, update λ to $\lambda=(\lambda-L)+L$. Repeat solution from step 4 until convergence.
10. After convergence evaluate (1.172) or (1.178).
11. Compute azimuths using (1.70) and (1.71).

<u>For the direct problem:</u> Given ϕ_1, s, α_{12}, L_1

<u>Steps</u>

1. Compute β_1 from (1.9).
2. Compute β_0 from $\cos\beta_0 = \sin\alpha_1 \cos\beta_1$.
3. Compute k^2 from (1.26).
4. Compute σ_1 from (1.21).
5. Compute the first approximation to σ_T as s/b.
6. Evaluate (1.172) or (1.178) for σ_T using the value of σ_T from step 5 as the values needed in the integral.
7. Repeat 6 until convergence.
8. Compute β_2 (and then ϕ_2) from (1.23).
9. Compute λ from (1.59).
10. Evaluate (1.173) or (1.179) to find (λ-L). Then $L_2 = L - (\lambda - L)$.
11. Compute azimuths using (1.70) and (1.71).

Tests described by Saito (ibid) show that this procedure gives results equivalent to the usual series solution of the problem.

A completely different numerical integration approach was given by Kivioja (1971). He uses as a starting premise the following equations taken from (1.3) and Clairaut's equation:

$$ds \cos\alpha_1 = M_1 d\phi \tag{1.181}$$

$$ds \sin\alpha_1 = N_1 \cos\phi_1 dL \tag{1.182}$$

$$N_1 \cos\phi_1 \sin\alpha_1 = constant = c \tag{1.183}$$

For the direct problem a suitable ds increment is chosen, the initial azimuth is used as the starting azimuth and increments of ϕ and L computed using (1.181) and (1.182) with (1.183) being used to compute a new azimuth. Analogous procedures are used for the inverse problem. Jank and Kivioja (1980) have discussed additional application of this procedure but the significant amount of computer time needed for the technique and other concerns may limit the application of this method. Meade (1981) discusses some of these limitations.

1.5 Geodesic Behavior for Near Anti-Podal Lines

Two points on the ellipsoid are defined to be anti-podal when L=180° and $\phi_2 = - \phi_1$. Near anti-podal points will have these conditions approximately met in a sense to be clarified later. When the inverse solutions previously discussed are applied to anti-podal lines they fail to converge. It is thus important to understand the general behavior of these anti-podal lines. The general case of two points located at an arbitrary ϕ is discussed by Fichot and Gerson (1937). A special case of the general problem occurs when the two points lie on the equator. This situation is discussed in the next section which is followed by a discussion of the general case.

1.51 Anti-Podal Behavior for Two Points on the Equator

Consider two points located on the equator not too far apart. The geodesic will be the equator itself with the forward azimuth at the first point 90° and the distance between the two points on the equator is simply the arc of the equatorial circle given by:

$$S = aL \tag{1.184}$$

Now consider the two points exactly 180° apart. The aximuth from the first point will be 0°, and the geodesic distance will be twice the quadrant arc. Note that the geodesic is not along the equator as it is not the shortest distance between the two points.

There, thus, must be a region in which the azimuth changes from 90° to 0° and a formulation of the distance problem to regard the above two situations. Helmert (1896) discussed some of these problems as well as Lambert (1942), and Lewis (1963). Thien (1967) has shown the formulation of this problem to a high degree of accuracy.

We first consider the form that the Rainsford formulations take when the two points are on the equator. Since $\phi_1 = \phi_2 = 0$, we have $\beta_1 = \beta_2 = 0$ and equation (1.57) reduces to:

$$\cos\sigma = \cos\lambda \qquad \text{or} \qquad \sigma = \lambda \tag{1.185}$$

From equation (1.59):

$$\sin\alpha_1 = \frac{\sin\lambda}{\sin\sigma} \tag{1.186}$$

which gives $\sin\alpha_1 = 1$ or $\alpha_1 = 90°$, provided λ (or σ) is not 0° or 180° at which time α_1 is indeterminate. In order to determine $2\sigma_m$, needed both in equation (1.39) and (1.56), we recall from (1.37) and preceeding, that:

$$2\sigma_m = 2\sigma_1 + \sigma_T \tag{1.187}$$

However, $\sigma_1 = 0$ and we have let $\sigma_T = \sigma$, so we conclude $2\sigma_m = \sigma$. Since $\cos\alpha = 0$, $\sin\alpha = 1$ we may write (1.56) as:

$$(\lambda - L) = f\sigma = f\lambda \tag{1.188}$$

$$\lambda = \frac{L}{1-f}$$

(1.189)

$$L = \lambda \, (1 - f)$$

(1.190)

If we let L=180° in (1.189) we obtain a value of λ greater than 180°. This cannot be correct as there would then be a geodesic of length smaller than 180° on the sphere. This would imply an inconsistency in the method. This inconsistency is resolved by noting that the maximum value of λ is 180°. When λ reaches this value L may be computed from:

$$L = 180° \, (1 - f)$$

(1.191)

Although we know L can be greater than 180° (1-f), we cannot formulate the behavior of the geodesic after L = 180° (1-f) as we meet with the inconsistencies in the value of λ previously mentioned. The longitude given by (1.191) is the maximum longitude that can be reached with the assumption that $\alpha_1=90°$ or that equation (1.186) is determinate. At the point given by (1.191) this assumption is no longer valid and other steps must be taken for the solution.

To do this we go back to equation (1.56) and consider it for the case $\lambda=180°$. Now we cannot consider $\alpha=90°$. However, we still have $2\sigma_m = \sigma =180°$ so that we now write (1.56) as:

$$\left(180° - L\right) = f \sin\alpha \, A_0 \, 180°$$

or

$$L = 180° \left(1 - f \sin\alpha \, A_0\right)$$

(1.192)

When $\alpha=90°$ equation (1.192) reduces to (1.191). Thus, equation (1.192) shows how L will be a function of the azimuth of the geodesic, after the longitude indicated by (1.191) is reached.

At this point we may summarize the behavior of the geodesic for our special case. For longitudes on the ellipsoid less than a certain amount, the azimuth from the first point is 90° (assuming the second point is east of the first) and the relationship between L and λ is given by equation (1.190). In this region the path is equatorial. At the longitude given by (1.191) (i.e. $\lambda=180°$) the critical point of "lift off" is reached. That is, beyond this point the path is no longer equatorial, but rises from the equator with the azimuth of the geodesic such that equation (1.192) is maintained.

We may also be interested in the difference in length between the geodesic and the equatorial arc for the special case. Up to the lift off point, the geodesic coincides with the equatorial arc and thus there is no distance difference. We now consider what happens beyond the lift off point. Using the fact that for this special case beyond lift off $\lambda = \sigma = 180°$, we write equation (1.40) as:

$$s = \pi b B_0 = \pi a \, (1 - f) \, B_0$$

(1.193)

where B_0 is given in equation (1.41). By substituting equation (1.192) into (1.184) we may write:

$$S = \pi a \left(1 - f \sin\alpha A_0\right) \tag{1.194}$$

Subtracting (1.193) from (1.194) we have:

$$S - s = \pi a \left(1 - f \sin\alpha A_0 - B_0 (1 - f)\right) \tag{1.195}$$

Neglecting higher order terms we can write:

$$A_0 \approx 1 - \frac{f}{4} \cos^2\alpha$$

$$B_0 \approx 1 + \frac{1}{4} e'^2 \cos^2\alpha \approx 1 + \frac{f}{2} \cos^2\alpha \tag{1.196}$$

Substituting (1.196) into (1.195) we find approximately:

$$S - s = \frac{1}{2} \pi a f \left(1 - \sin\alpha\right)^2 \tag{1.197}$$

When $\alpha = 90°$, S-s=0, and when $\alpha = 0°$ we have the maximum difference, $\pi a f / 2$, approximately.

It is of interest to apply some of the equations previously derived in this section. For this purpose we take the parameters a = 6378388 m, f = 1/297. From equation (1.191) the lift off longitude is 179° 23′ 38″.18182. Beyond this point the geodesic rises off the equator. If we desire to compute the azimuth of the geodesic beyond this critical point by specifying L, we use equation (1.192). This equation can be solved by iteration as follows: Noting that A_0 is approximatelly one, we may write from (1.192):

$$\sin\alpha^{(0)} = \frac{180° - L}{180° f} \tag{1.198}$$

where $\alpha^{(0)}$ is the initial approximation to the desired azimuth. After this is obtained, the more precise value may be obtained by iteration of (1.192), written in the form:

$$\sin\alpha = \frac{180° - L}{180° f A_0} \tag{1.199}$$

Alternately we may specify the value of α and compute from (1.192) the value of L at which the geodesic will intersect the equator.

An alternate solution to (1.192) has been carried out by Vincenty (1975) and Bowring (1983) who gives the following direct solution for $\sin\alpha$:

$$\sin\alpha = b_1 Q + b_3 Q^3 + b_5 Q^5 + b_7 Q^7 \tag{1.200}$$

$$\sin\alpha = b_1 Q + b_3 Q^3 + b_5 Q^5 + b_7 Q^7 \tag{1.200}$$

where

$$Q = \frac{(\pi - L)}{\pi f}$$

$$b_1 = 1 + \frac{f}{4} + \frac{f^2}{8} + \frac{9}{128} f^3$$

$$b_3 = -\frac{f}{4} - \frac{f^2}{8} - \frac{11}{128} f^3$$

$$b_5 = \frac{3}{128} f^3$$

$$b_7 = -\frac{1}{128} f^3 \tag{1.201}$$

With the ellipsoid parameters above some compatible values of α and L are the following:

α	L	
90°	179°23′	38″.18182
70°	179°25′	49″.96405
50°	179°32′	09″.21295
30°	179°41′	49″.78063
10°	179°53′	41″.44083

A sketch of this variation is shown in Figure 1.6:
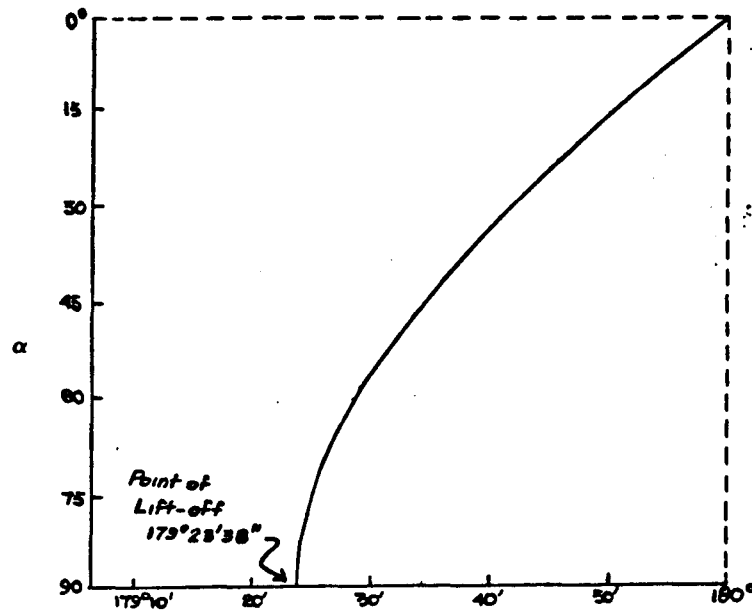
Figure 1.6
Azimuth vs Longitude Difference For Two Points on the Equator.

We note from this figure that the fastest change of α with L takes place at the lift off point. This could be verified by differentiating (1.198) or (1.199).

For the same case the values of S-s have been computed from (1.195) with the values plotted in Figure 1.7.
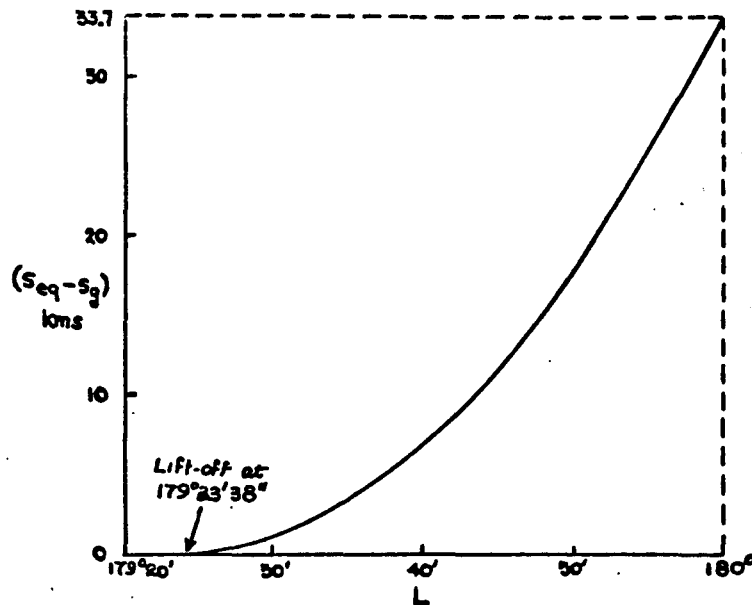
Figure 1.7
Difference in Length Between a Geodesic and an Equatorial Arc for Two Points on the Equator.

One final property of the geodesic which is of some theoretical interest lies in the fact that as a geodesic is extended around the ellipsoid it, in general, will not close back on itself. To demonstrate this, consider the Figure 1.8 (taken from Lewis (1963)) which is a view from above the pole (designated N) of the ellipsoid where a geodesic crosses the equator at point P_1 and continues until it intersects the equator again at P_2 which does not coincide with the point A which is 180° apart from P_1. The geodesic then continues around the back of the ellipsoid until it reaches the equator again at point B which does not coincide with the starting point P_1. The shift between P_1 and B may be computed using the equations previously discussed.
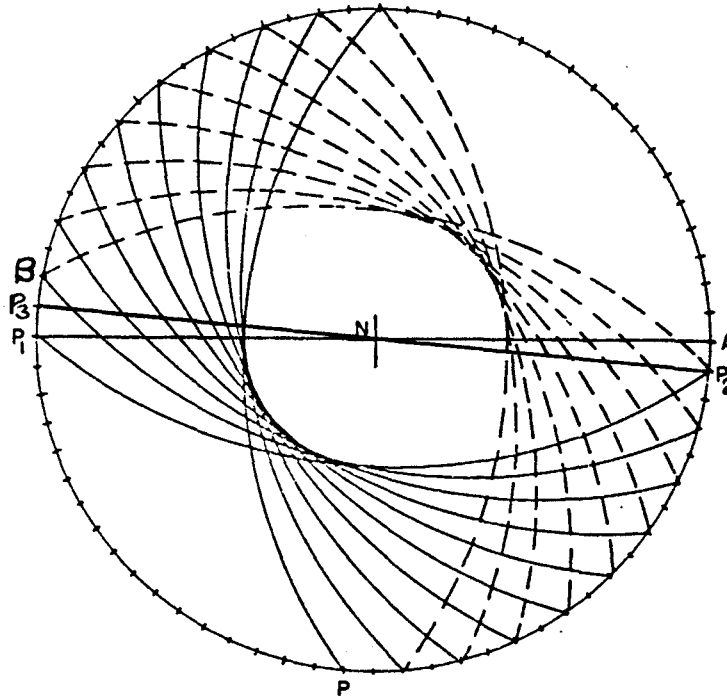
Figure 1.8
A Geodesic Extended as a Continuous Curve.

To compute P_1B we first use (1.192) to express the longitude difference between A and P_2. We have

$$AP_2 = L_A - L_{P_2} = 180° - 180° \left(1 - f\sin\alpha A_0\right)$$

$$AP_2 = 180° \, f\sin\alpha A_0 \tag{1.202}$$

Now, the angular distance BP_3, by the same procedures will be

$$BP_3 = 180° \, f\sin\left(180° - \alpha\right) A_0 \tag{1.203}$$

where 180°- α is the azimuth of the geodesic at P_2. Adding (1.202) and (1.203) yields the distance BP_1 by which the geodesic does not close back upon itself. Thus:

$$BP_1 = 360° \ f \sin\alpha A_0$$

(1.204)

In terms of distance:

$$BP_1 = 2\pi a \ f \sin\alpha A_0$$

(1.205)

We should finally note that unless the distance $BP_1/2\pi$ is a rational number the geodesic will never close on itself but will continue to creep around the ellipsoid. It should be clear that the curve we are discussing here is not the shortest distance geodesic. Instead it is a curve that has all the properties of a geodesic except for the shortest distance property.

1.52 Geodesic Behavior for Near Anti-Podal Points - General Case

The discussion in the previous section has addressed a special case of a geodesic when the two points involved are on the equator. A similar problem must occur when two points on the ellipsoid are approximately opposite each other. By exactly opposite we mean $\phi_1 = - \phi_2$ and L is equal to 180°.

If two points are nearly opposite (antipodal) the standard iterative inverse procedure described in sections 1.21 and 1.23 will fail to converge. Such a case can be detected when $|\lambda|$ is greater than π as computed from equation (1.67) with (1.56) or from (1.101) at the first iteration. This is because the maximum allowable value of λ is π. If we consider P_1 fixed and P_2 exactly antipodal, then there will be a region on the ellipsoid about P_2 in which the iterative solution will fail to converge for the line between P_1 and an arbitrary point in the region. This region will depend on the equations (e.g., 1.67 or 1.101) being used to calculate λ.

Since the standard procedure fails to work an alternate procedure must be used (Vincenty, 1975). To do this we first assume as first appromxation that λ is 180°. At this point equation (1.57) can be written as:

$$\cos\sigma = - \left(\cos\beta_1 \cos\beta_2 - \sin\beta_1 \sin\beta_2 \right)$$

(1.206)

This formula is consistent with:

$$\sigma = 180° - \left| \beta_1 + \beta_2 \right|$$

(1.207)

Note that in the near antipodal case σ is approximately 180°. At this point we need to find the azimuth and length of the line between the two near antipodal points. This procedure will be an iterative procedure.

We first rewrite equation (1.56):

$$\sin\alpha = \frac{\lambda - L}{f \left(A_0 \ \sigma + A_2 \sin\sigma \cos2\sigma_m + --- \right)}$$

(1.208)

where the coefficients are function of f and $\cos^2\alpha$; $\cos2\sigma_m$ can be computed from (1.62). Knowing L, and taking $\lambda = (\text{sign L})\pi$, an approximate value of α can be found from (1.208). We next solve (1.61) to find an updated value of λ:

$$\sin\lambda = \frac{\sin\alpha \, \sin\sigma}{\cos\beta_1 \cos\beta_2} \qquad (1.209)$$

This λ value can then be used to determine an improved value of σ from equation (1.58) for example. The process is then iterated back through (1.208) until the change in $\sin\alpha$ from the previous value is less than a specified amount.

At the completion of the iteration the following values would be known: σ, λ, β_1, β_2 and α. We then can use (1.61) to determine α_1:

$$\sin\alpha_1 = \frac{\sin\alpha}{\cos\beta_1} \qquad (1.210)$$

Then:

$$\cos\alpha_1 = \pm \left(1 - \sin^2\alpha_1\right)^{1/2}$$

when the minus sign is chosen if:

$$\cos\beta_1 \sin\beta_2 - \sin\beta_1 \cos\beta_2 \cos\lambda < 0$$

Equation (1.73) in conjunction with (1.69) can be used to determine α_2 while the distance is determined using equation (1.40).

A special case of these equations occurs when $\phi_1 = -\phi_2$ and we are interested in the behavior of α_1 and L in the antipodal region. In this region $\lambda = \pi$ so that for this case (1.207) yields $\sigma=180°$. Then (1.56) becomes the same as (1.192). Thus (1.192) holds not only for two points on the equator but for two points of opposite latitude provided the points are within the antipodal region. The geodesic distance is then found from equation (1.193). Vincenty notes that in this special antipodal case the value of α and s do not depend on latitude but only on the longitude difference of the two points which must be _within_ the antipodal region.

As an example, consider two cases with L=179°54′ for both cases. Case one has $\phi_1=80° =-\phi_2$ and case two has $\phi_1= 1° = -\phi_2$. If one solves (1.192 or 1.200) and (1.193) we find (for the International Ellipsoid):

$\alpha = 9° \, 30′ \, 18″34717$
$s = 20003657.4122$ m

Note that only the azimuth _at the equator_ is the same in this example.

Let's now return to the discussion of the more general antipodal problem. Again consider P_1 fixed and P_2 exactly antipodal. About P_2 there is a locus of points inside of which the standard iterative solution will fail. If (1.67) with (1.56) is used, the points form an approximate circle

about P_2. If the Newton-Raphson procedure (e.g. eq. (1.101)) is used the region corresponds to the geodesic envelope described by Fichot and Gerson (1937).
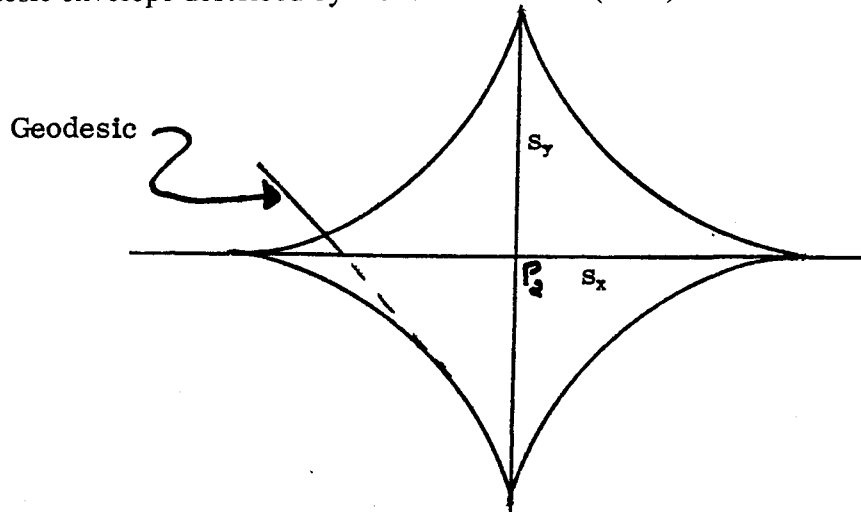
Figure 1.9

The Geodesic Envelope about P_2.

To define this latter region we can construct the envelope of the tangents to the geodesic that passes throught the same parallel on which P_2 lies. Such an envelope is called the geodesic evolute by Thomas (1970). Let s_y and s_x be the axis lengths shown in Figure 1.9. Then Bowring (1976) shows that:

$$s_y = \pi a f \left(1 - \frac{1}{4} f \right) \cos^2 \beta_1$$

(1.211)

$$s_x = \pi a f \left(1 - \frac{1}{4} f \sin^2 \beta_1 \right) \cos^2 \beta_1$$

(1.212)

The ratio of these two lengths is:

$$\frac{s_y}{s_x} = 1 - \frac{1}{4} f \cos^2 \beta_1$$

(1.213)

which shows the envelope is not exactly symmetric. If $\beta_1 = 30°$ we find that $s_y = 50559$ m and $s_x = 50591$ m. As the latitude increases the radius of this region decreases.

The equation of the envelope would be (Bowring, 1976, p. 100, Fichot and Gerson, 1937, p.66):

$$\left[\frac{x}{1 - \frac{1}{4} f \sin^2 \beta_1} \right]^{2/3} + \left[\frac{y}{\left(1 - \frac{1}{4} f \right)} \right]^{2/3} = \left(\pi a f \cos^2 \beta_1 \right)^{2/3}$$

(1.214)

Here x and y are local plane coordinates whose origin is at the antipodal point. The x and y coordinates are (Fichot and Gerson, ibid, p.65):

$$x = -af\pi\cos^2\beta_1\left(1 - \frac{f}{4}\sin^2\beta_1\right)\left(1 - \frac{3f}{4}\cos^2\beta_1\cos^2\alpha_1\right)\sin^3\alpha_1 \tag{1.215}$$

$$y = -af\pi\cos^2\beta_1\left(1 - \frac{f}{4}\right)\left(1 + \frac{3f}{4}\cos^2\beta_1\sin^2\alpha_1\right)\cos^3\alpha_1 - \frac{a\pi^2 f^2}{2}\sin\beta_1\cos^3\beta_1 \cdot$$

$$\left(1 - 3\cos^2\alpha_1\right) \tag{1.216}$$

Given the starting latitude and azimuth the coordinates of the envelope may be computed. Figure 1.10 shows a part of the anti-podal envelope for the case of $\phi_1=30°$, and $L_1=0°$.

1.521 A Convergence Problem

One problem noted by Vincenty (1975, private communication) was the slow convergence of the standard iterative procedure when P_2 is just outside the antipodal region. A similar problem was encountered with the antipodal solution when P_2 was just inside the antipodal circle. In each case the problem was caused by the oscillation of λ or α respectively during the iteration. Vincenty has suggested that in carrying out a calculation that a test for oscillation be made during the iteration. If such is the case, faster convergence can be obtained by computing a weighted mean values of λ or α as follows:

$$A_{i+1} = \frac{\left(2A_i + 3A_{i-1} + A_{i-2}\right)}{6} \tag{1.217}$$

1.6 The Behavior of "Backside lines"

To this point we have been investigating the single shortest line between two points. We could clearly imagine another path from P1 to P2 that goes around the "backside" of the ellipsoid. This line is not a geodesic since it is not the shortest distance, but it has all the other properties of the geodesic.

In order to compute a backside line the usual inverse procedure can be used with some minor changes. Vincenty suggests the following: Compute the usual σ from (1.57) and/or (1.58). Then compute a backside σ:

$$\sigma_{BS} = 2\pi + \tan^{-1}\left(-\sin\sigma, \cos\sigma\right) \tag{1.218}$$

Now change the two azimuths by $\pm\pi$ which is easily done by changing the sign on the numerator and denominator in equations (1.70) and (1.71).

Several special cases arise in the backside solution. If the second point lies within the antipodal envelope there are four distinct "geodetic connections" between the two points (Fichot and Gerson, 1937, p. 47). And if the two points are very close, there can also be three backside lines. One can always check a backside inverse by performing a direct solution with the computed distance and azimuth.

(Based on Vincenty, 1975)
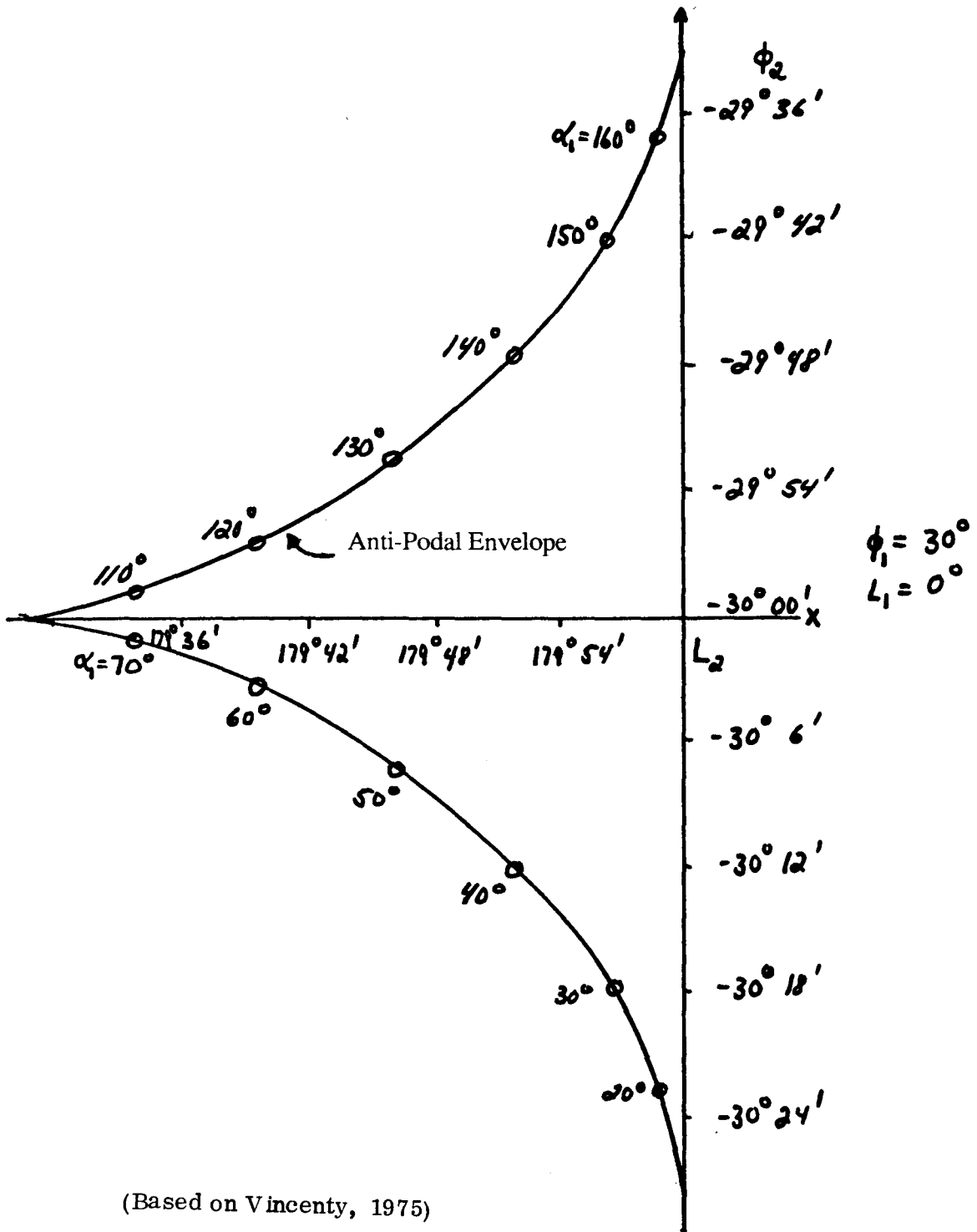
Figure 1.10
The Anti-Podal Envelope when $\phi_1 = 30°$.

1.7 Test Lines

When programs have been written to solve the so called long line problems it is convenient to have test lines for which previously computed results are available. This section gives such results for three types of cases on the International Ellipsoid (i.e. a=6378388 m, f=1/297).

1.71 Standard Test Lines

The first set of lines are rather standard not involving anti-podal or backside cases. The first four lines have been previously used by Rainsford (1955). The fourth line was designed to be a short line while the fifth line was one where σ was forced to be close to 90°. The seventh line is one where α_{12} was chosen greater than 180°. Table 1.1 gives the latitudes of the end points and the longitude difference. Table 1.2 gives the geodesic distance and azimuths.

Table 1.1
Standard Test Lines - Position Definition

	ϕ_1			ϕ_2			L		
Line 1	37°	19′	54.95367″	26°	07′	42.83946″	41°	28′	35.50729″
2	35°	16′	11.24862″	67°	22′	14.77638″	137°	47′	28.31435″
3	1°	00′	00.00000″	-0°	59′	53.83076″	179°	17′	48.02997″
4	1°	00′	00.00000″	1°	01′	15.18952″	179°	46′	17.84244″
5	41°	41′	45.88000″	41°	41′	46.20000″	0°	0′	0.56000″
6	30°	00′	00.00000″	37°	53′	32.46584″	116°	19′	16.68843″
7	37°	00′	00.00000″	28°	15′	36.69535″	-2°	37′	39.52918″

Table 1.2
Standard Test Lines - Distance and Azimuths

	S	α_{12}			α_2		
Line 1	4085966.7026	95°	27′	59.630888″	118°	05′	58.961608″
2	8084823.8383	15°	44′	23.748498″	144°	55′	39.921473″
3	19959999.9998	88°	59′	59.998970″	91°	00′	06.118357″
4	19780006.5588	4°	59′	59.999953″	174°	59′	59.884804″
5	16.2839751	52°	40′	39.390667″	52°	40′	39.763168″
6	10002499.9999	45°	00′	00.000004″	129°	8′	12.326010″
7	1000000.0000	195°	00′	00.000000″	193°	34′	43.74060″

In this computation the iteration on λ-L was stopped when this value changes less than 0.5×10^{-14} radians or 0.000000001″. The number of iterations needed averaged 5 but line 4 needs 23 iterations. A fluctuation of a single digit in the last place of the results could be expected.

Checks of a direct problem program may be made by using the results of the inverse problem and comparing them with the original starting values.

1.72 Anti-Podal Lines

We now consider six test lines for which the second point is near the antipodal region. Table 1.3 gives the information on the point coordinates while Table 1.4 gives the azimuths and distances

between the points. Also given is the longitude difference λ and the number of iterations to converge the solution when the Newton-Raphson procedure is used for iterations on λ (non-antipodal) and simple iteration when the sin α iteration is used.

Table 1.3
Anti-Podal Lines and Near Anti-Podal (*) Lines - Position Definition and σ

Line	ϕ_1	ϕ_2	L	σ
A	41° 41′ 45.88	-41° 41′ 46.20	179° 59′ 59.44	179° 59′ 59.68013
B	0°	0°	179° 41′ 49.78063	180°
C	30°	-30°	179° 40′	180°
D	60°	-59° 59′	179° 50′	179° 58′ 51.15676
E	30°	-29° 50′	179° 48′	179° 49′ 36.79418
F	30°	-29° 55′	179° 48′	179° 54′ 43.94956

Table 1.4
Anti-Podal Lines and Near Anti-Podal (*) Lines - α_1, α_2, λ, and s

Line	α_1	α_2	λ	s (m)	iteration
A	179° 58′ 49.1625	0° 01′ 10.8376	179° 59′ 59.99985	20004566.7228	3
B	29° 59′ 59.9999	150°	180°	19996147.4168	21
C	39° 24′ 51.8058	140° 35′ 08.1942	180°	19994364.6069	21
D	29° 11′ 51.0700	150° 49′ 06.8680	179° 58′ 53.03674	20000433.9629	14
E*	16° 2′ 28.3389	163° 59′ 10.3369	179° 56′ 41.64754	19983420.1536	6
F	18° 38′ 12.5568	161° 22′ 45.4373	179° 58′ 3.57082	19992241.7634	22

Lines A, B, C, and D are the anti-podal lines described by Vincenty (1975, Table 1). Line E is a point just outside the anti-podal envelope (see Figure 1.10) and line F is a point just inside the envelope.

1.73 Backside Lines

Backside lines have been discussed in section 1.6. The four examples given in Table 1.5 are taken from Vincenty (ibid):

Table 1.5
Backside Geodesic Lines

Line	A	B	C	D
ϕ_1	41° 41′ 45.88000	00° 00′ 00.00000	30° 00′ 00.00000	60° 00′ 00.00000
ϕ_2	41° 41′ 46.20000	0° 00′ 00.00000	30° 00′ 00.00000	59° 59′ 00.00000
L	0° 00′ 56.000	0° 18′ 10.21937	0° 20′ 00.00000	0° 10′ 00.00000
α_1	180° 00′ 35.423	194° 28′ 47.448	198° 30′ 47.488	344° 56′ 31.727
α_2	180° 00′ 35.423	194° 28′ 47.448	198° 30′ 47.488	344° 56′ 59.622
s	40 009 143.3208	40 004 938.2722	40 004 046.7114	40 006 087.0024

As noted in section 1.6 it is possible to have four geodetic connections between **two points** provided the second point lies within the antipodal envelope. Vincenty has constructed **a test case** for this situation. The first point has a latitude of 40° and a longitude of 0°. The second **point has a** latitude of -40° 1′ 5″.75932 and a longitude of 179° 55′ 15″.59578. The azimuths and **distances** between the two points are shown in Table 1.6.

Table 1.6
Four Geodetic Connections

Method	α_{12}			α_{21}			s (m)
1	170°	41′	42″.49809	189°	18′	26″.51095	20002002.7295
2	272°	40′	42″.01097	87°	40′	15″.32624	20031200.7134
3	86°	20′	38″.15306	273°	24′	31″.22084	20017727.6841
4	10°	20′	3″.7186	349°	39′	46″.25456	20005999.9995

Method 1 is the usual anti-podal solution. Method 2 is the backside solution described **in** section 1.6. Method 3 uses the initial value of λ as -L and reverses the sign of the tan on the σ backside determination. Method 4 carries out a backside solution but iterating for $\sin\sigma$ and specifying $\sin\sigma < 0$. All values have been computed on the International Ellipsoid.

3/18/93

References

Bowring, B.R., Analytical Geodesy, unpublished manuscript, (26 Queens Drive, Thames Ditton, Survey, England), 1975.

Bowring, B.R., The Geodesic Inverse Problem, Bulletin Geodesique, 57, 109-120, 1983.

Bulirsch, R., and M. Gerstl, Numerical Evaluation of Elliptic Integrals for Geodetic Applications, in Proc., Eighth Symposium of Mathematical Geodesy, held in Como, Italy 1981, Florence 1983.

Fichot, M., and M. Gerson, La Zone Geodesique Antipode, in Annales Hydrographiques, 3° serie, Tome Quinzieme, Service Hydrographique De Le Marine, Paris, 1937.

Ganshin, V.N., Geometry of the Earth Ellipsoid, translated from the Russian original, dated 1967, by the Aeronautical Chart and Information Center, St. Louis, 63125, 1969, AD 689507.

Helmert, F.R., Die Mathematischen Und Physikalischen Theorium Der Höheren Geodasie, B.G. Teubner, Leipzig, 1880 (reprinted 1962), (translation by Aeronautical Chart and Information Center, St. Louis, 1964).

Jank, W., and L.A. Kivioja, Solution of the Direct and Inverse Problems on Reference Ellipsoids by Point-by-Point Integration Using Programmable Pocket Calculators, Surveying and Mapping, Vol. XL, No. 3, 325-337, September, 1980.

Jordan-Eggert, Jordan's Handbook of Geodesy, translation of 8th edition, 1941 by Martha W. Carta, (U.S. Army Map Service, Washington, D.C., 1962).

Kivioja, L.A., Computation of Geodetic Direct and Indirect Problems by Computers Accumulating Increments from Geodetic Line Elements, Bulletin Geodesique, No. 99, March 1971.

Lambert, W.D., The distance between two widely separated points on the surface of the earth, Journal of the Washington Academy of Sciences, Vol. 32, No. 5 (May 15, 1942), p. 125-30.

Lewis, E.A., Parametric Formulas for Geodesic Curves and Distances on a Slightly Oblate Earth, Air Force Cambridge Research Laboratories, Note No. 63-485, April, 1963, AD412501.

Meade, B.K., Comments on Formulas for the Solution of Direct and Inverse Problems on Reference Ellipsoids Using Pocket Calculators, Surveying and Mapping, Vol. 41, No. 1, pp. 35-41, 1981.

McCaw, G.T., Long Lines on the Earth, Empire Survey Review, Vol. II, p. 156-63, 346-52, 505-8, 1932.

Rainsford, H.F., Long Geodesics on the Ellipsoid, Bulletin Geodesique, No. 37, September, 1955, pp. 12-21.

Rapp, R.H., Geometric Geodesy, Volume I (Basic Principles), Dept. of Geodetic Science, The Ohio State University, 1984.

Saito, T., The Computation of Long Geodesics on the Ellipsoid by Non-Series Expanding Procedure, Bulletin Geodesique, No. 98, December, 1970, p. 341-374.

Saito, T., The Computation of Long Geodesies on the Ellipsoid Through Gaussian Quadrature, Bulletin Geodesique, Vol. 53, 165-177, 1979.

Singh, A., An Analysis of McCaw's Solution of the Direct Problem of Long Lines on the Ellipsoid, M. Sc. Thesis, The Ohio State University, Columbus, Ohio 1980.

Sodano, E.M., A Rigorous Non-Iterative Procedure for Rapid Inverse Solutions of Very Long Geodesics, Bulletin Geodesique, No. 47/48, 1958.

Sodano, E.M., General Non-Iterative Solution of the Inverse and Direct Geodetic Problems, Bulletin Geodesique, March, No. 75, p. 69-89, 1965.

Sodano, E.M., and T. Robinson, Direct and Inverse Solutions of Geodesics, Army Map Service Technical Report No. 7 (Rev.), July, 1963, AD 657591.

Thien, G., A Solution to the Inverse Problem for Nearly-Antipodal Points on the Equator of the Ellipsoid of Revolution, M.S. thesis, The Ohio State University, 1967.

Thomas, P.D., Spheroidal Geodesics, Reference Systems and Local Geometry, U.S. Naval Oceanographic Office, SP-138, Washington, D.D., 20390, 1970.

Vincenty, T., Geodetic Inverse Solution Between Antipodal Points, unpublished paper dated Aug. 28, 1975.

Vincenty, T., Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of Nested Equations, Survey Review, XXII, 176, April 1975.

2. Transformation of Geodetic Data Between Reference Datums

2.1 Introduction

The history of Geodesy must include a discussion of positioning for one of the fundamental goals of geodesy is to precisely define positions of points on the surface of the Earth. In order to do this it was necessary to define some starting point and reference ellipsoid. With this information and with the measured angles and distances, the usual computation of the geodetic positions took place. In Europe, individual countries started in the 18th century the development of national triangulation networks. These national networks were subsequently extended over Europe in the 19th century with connections between various countries. After World War II, extensive efforts were made to combine the national networks into a consistent system which became known as the European Datum (1950). The development of an improved, consistent network, incorporating precise distance and angle observations, as well as VLBI, Doppler and SLR derived positions, continues. New networks such as ED79 and ED87 have been developed.

In the United States the development of the geodetic network started in 1815 when F. Hassler started geodetic measurements near New York City (Dracup, 1976). During the remaining part of the 19th century, a number of major areas were developed including the Eastern Oblique Arc from Calais, Maine to New Orleans and the first Transcontinental Arc along the 39th Parallel. In 1879 the New England Datum was adopted for triangulation in the northeast and eastern United States. The origin was chosen at station Principio in Maryland. In 1901 the New England Datum was adopted as the United States Standard Datum with the origin point moved, by definition, to Meades Ranch, Kansas. In 1913 the Standard Datum was adopted for use by Mexico and Canada, and its name changed to the North American Datum. In 1927 a readjustment took place fixing the coordinates of Meades Ranch. This led to the North American Datum 1927 which served for almost sixty years as a reference system for the United States Improvement in measuring techniques, and errors in the NAD27 led to the development of NAD83 which was completed in 1987 (Bossler, 1987). Additional discussion on this system will be found in Section 3.

The two examples described above will be typical of various countries and areas. Clearly, each system will have its own coordinate system and reference ellipsoid. One easy task to visualize is the conversion of coordinates from one geodetic system to another. However we now have a number of fundamental reference systems or, in practice, a conventional reference system. This system can be associated with a particular satellite (Doppler or laser, for example) system. Consequently, we will be interested in the transformation between geodetic systems and some externally defined system.

However, we must recognize that most geodetic systems are essentially horizontal in nature. We have been speaking of horizontal datums where latitude and longitude are determined. Vertical datums have historically been treated separately. The conversion of a horizontal system and a vertical system into a consistent three dimensional system is difficult because of the role of the geoid or the height reference surface. The development of horizontal networks was hindered because of the lack of knowledge of the separation between the ellipsoid and the geoid. This lack of knowledge made it impossible to reduce angles and, most importantly, distances down to the ellipsoid which was the actual computational surface. Instead, the measurements were reduced to the geoid with computations taking place as if they were on the ellipsoid. This method of reduction to the geoid was called the development method where the observations are "developed" on the geoid. Because the geoid undulations can vary substantially in a large country, the neglect of geoid undulations can cause systematic errors in the computed positions.

An alternate method of triangulation and triliteration computation is known as the projective method. In this procedure the observations are rigorously reduced to the ellipsoid taking into account deflections of the vertical and the separation between the geoid and ellipsoid. This projective method has not been widely used because of the lack of knowledge of geoid undulations as historical networks were determined. Today the situation is much easier, but this does not help the problems of the past.

We should also note that there are several methods in which the projective method can be implemented. In the Pizzetti method, a point is reduce from the surface along a curved vertical to the geoid and from there to the ellipsoid on a perpendicular to the ellipsoid. The Helmert procedure projects the surface point to the ellipsoid along the ellipsoidal normal. The two projection methods are shown in Figure 2.1 which represents a section in an arbitrary direction.
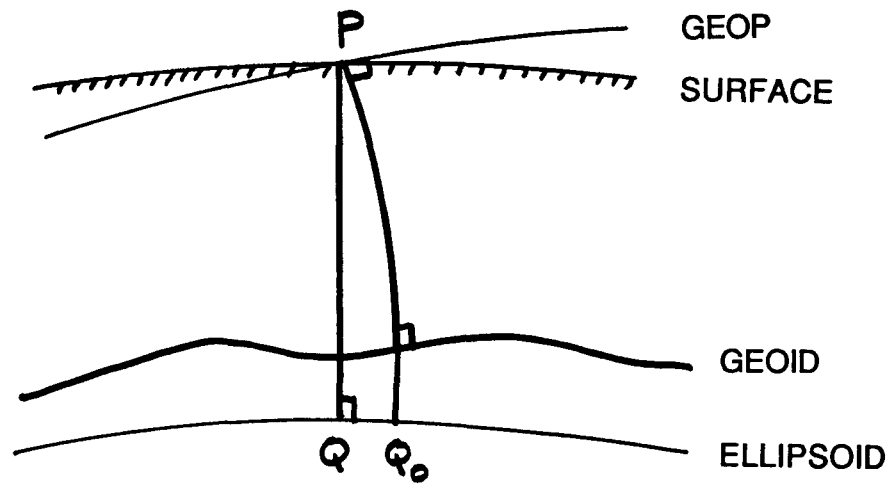
Figure 2.1   Two Projective Techniques

A discussion of the projection of point P to the ellipsoid may be found in Heiskanen and Moritz (1967, p. 180). A discussion of the projection method and the development method may be found in Wilcox (1963).

With this section as a background we now turn to transformation procedures. In principle we should define whether we are working with a development or projective geodetic network. We should also distinguish between horizontal or vertical network transformations. In practice this is rarely done and we simply form three dimensional systems although such systems may have never been computed in three dimensions originally.

2.2   Similarity Transformations

We are given a set of rectangular coordinates, (x,y,z), in an "old" system and we want to transform these coordinates into the "new" system to obtain (X,Y,Z) ($\underline{X}$). We can first postulate a general linear (affine) transformation of the form (Leick and van Gelder, 1975):

$$\underline{X} = A\underline{x} + \underline{A}_0 \qquad (2.1)$$

where A is a 3x3 matrix while A_0 is a 3x1 vector. There are a total of 12 parameters describing this linear transformation as can be seen from the component form of (2.1):

58

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} a_{10} \\ a_{20} \\ a_{30} \end{bmatrix} \qquad (2.2)$$

The 12 parameters can be interpreted as follows (ibid): six for the orthogonality transformation (three parameters for translation and three parameters for rotation) and 6 parameters describing the scaling transformation (three scale parameters along three perpendicular axes whose orientation is defined by the remaining three parameters).

A special case of the general affine transformation is the <u>orthogonal</u> transformation. Such a transformation preserves lengths and an orthogonal system of axes. The coefficients in A must meet the following conditions (Leick and van Gelder, 1975, p. 15).

$$a_{11}a_{12} + a_{21}a_{22} + a_{31}a_{32} = 0$$
$$a_{11}a_{13} + a_{21}a_{23} + a_{31}a_{33} = 0$$
$$a_{12}a_{13} + a_{22}a_{23} + a_{32}a_{33} = 0$$
$$a_{11}^2 + a_{21}^2 + a_{31}^2 = 1$$
$$a_{12}^2 + a_{22}^2 + a_{32}^2 = 1$$
$$a_{13}^2 + a_{23}^2 + a_{33}^2 = 1$$

$$(2.3)$$

Under these six conditions, the number of parameters of the general transformation is reduced to 6: three in A_0 and three in A. The latter three are rotations about each of the "old" axes. We will designate these rotations as ω_x, ω_y, and ω_z so that this orthogonal transformation can be written in the form:

$$\underline{X} = R(\omega_x, \omega_y, \omega_z)\underline{x} + \underline{A}_0 \qquad (2.4)$$

where R is a 3x3 orthogonal matrix that will be derived shortly. An alternate form of (2.4) can be written if the rotations are applied to the translated axes. We then would have:

$$\underline{X} = R(\underline{x} + \underline{A}_0) \qquad (2.4A)$$

We may now introduce a <u>single</u> scaling parameter, s, into the process, which yields a seven parameter similarity transformation. Leick and van Gelder point out that two versions of this type of transformation given by (2.4) can be written:

$$\underline{X} = sR\underline{x} + \underline{A}_0 \qquad (2.5)$$

$$\underline{X} = s(R\underline{x} + \underline{A}_0) \qquad (2.6)$$

Similarly two versions of the (2.4A) form can be written:

$$\underline{X} = sR(\underline{x} + \underline{A}_0) \qquad (2.7)$$

$$\underline{X} = R(s\underline{x} + \underline{A}_0) \qquad (2.8)$$

The easiest form to interpret is (2.5) where $\underline{A}_0$ represents the three translations between the origins of the two systems; R represents the rotation from the old to the new

system and s is the scale between the two systems. If there is no scale difference, s = 1. If there are no rotations between the systems, R is an identity matrix, and if there are no translations, $\underline{A}_0$ is zero.

In the next sections we will examine in detail a number of similarity transformations.

2.21 The Bursa - Wolf Transformation Model

We now consider the seven parameter similarity transformation discussed by Bursa (1962) and by Wolf (1963). The general geometry of the transformation is shown in Figure 2.2.
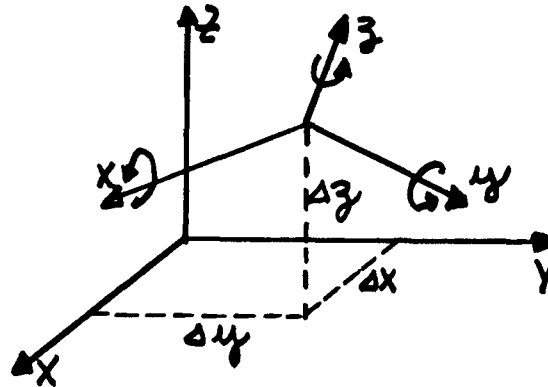
Figure 2.2. A Translated and Rotated Coordinate System

In Figure 2.2, we have indicated the translation parameters Δx, Δy, Δz which will be designated $\underline{T}$ in vector form. We have also shown the three rotation angles ω_x, ω_y, and ω_z. A positive rotation is a counterclockwise rotation about an axis when viewed from the end of the positive axis in right-handed coordinate systems. Equation (2.5) can now be written as:

$$\underline{X} = sR_z(\omega_z)R_y(\omega_y)R_x(\omega_x)\underline{x} + \underline{T} \tag{2.9}$$

where R_x, R_y, R_z are the following rotation angles (also see Rapp, 1984, p. 69).

$$R_x(\omega) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \sin\omega \\ 0 & -\sin\omega & \cos\omega \end{bmatrix} \tag{2.10}$$

$$R_y(\omega) = \begin{bmatrix} \cos\omega & 0 & -\sin\omega \\ 0 & 1 & 0 \\ \sin\omega & 0 & \cos\omega \end{bmatrix} \tag{2.11}$$

$$R_z(\omega) = \begin{bmatrix} \cos\omega & \sin\omega & 0 \\ -\sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.12}$$

The product of the three rotation matrices yields the following:

$$R_zR_yR_x =$$

$$\begin{bmatrix} \cos\omega_y \cos\omega_z & \cos\omega_x \sin\omega_z + \sin\omega_x \sin\omega_y \cos\omega_z & \sin\omega_x \sin\omega_z - \cos\omega_x \sin\omega_y \cos\omega_z \\ -\cos\omega_y \sin\omega_z & \cos\omega_x \cos\omega_z - \sin\omega_x \sin\omega_y \sin\omega_z & \sin\omega_x \cos\omega_z + \cos\omega_x \sin\omega_y \sin\omega_z \\ \sin\omega_y & -\sin\omega_x \cos\omega_y & \cos\omega_x \cos\omega_y \end{bmatrix}$$

$$(2.13)$$

Equation (2.13) can be evaluated assuming the rotation angles are small (a few seconds of arc) as they are in the cases we are concerned with. Under these circumstances (2.13) becomes:

$$R_zR_yR_x = \begin{pmatrix} 1 & \omega_z & -\omega_y \\ -\omega_z & 1 & \omega_x \\ \omega_y & -\omega_x & 1 \end{pmatrix}$$

$$(2.14)$$

Malys (1988) has studied the numerical impact of the small angle approximation in obtaining (2.14). He found that the disagreement between an element of (2.13) and (2.14) was at the level of 0.5×10^{-11} when the rotation angles were on the order of 1"; on the order of 0.5×10^{-10} when the angles were on the order of 3"; and on the order of 0.5×10^{-9} when the angles were on the order of 9". An error of 0.5×10^{-9} propagates into a coordinate error on the order of 3mm. We should note that the order of rotation is not important when the angles are small, as (2.14) is independent of the order of rotation.

We can now write (2.9), with (2.14), as

$$\underline{X} = s \begin{pmatrix} 1 & \omega_z & -\omega_y \\ -\omega_z & 1 & \omega_x \\ \omega_y & -\omega_x & 1 \end{pmatrix} \underline{x} + \underline{T}$$

$$(2.15)$$

We now introduce a scale difference quantity, Δs, defined such that:

$$s = (1 + \Delta s)$$

$$(2.16)$$

We can introduce this into (2.15) to write:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} + (1 + \Delta s) \begin{bmatrix} 1 & \omega_z & -\omega_y \\ -\omega_z & 1 & \omega_x \\ \omega_y & -\omega_x & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$(2.17)$$

Multiplying out and neglecting higher order terms such as $\omega \Delta s$ we have

$$X = x + \Delta x + x\Delta s + \omega_z y - \omega_y z$$

$$Y = y + \Delta y + y\Delta s - \omega_z x + \omega_x z \tag{2.18}$$

$$Z = z + \Delta z + z\Delta s + \omega_y x - \omega_x y$$

Equation (2.17) may be written in an alternate form which is convenient for some modeling problems:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} + U \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} + (1 + \Delta s) \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{2.19}$$

where

$$U = \begin{bmatrix} 0 & -z(1 + \Delta s) & y(1 + \Delta s) \\ z(1 + \Delta s) & 0 & -x(1 + \Delta s) \\ -y(1 + \Delta s) & x(1 + \Delta s) & 0 \end{bmatrix} \tag{2.20}$$

The above form (i.e., 2.19) has been used by Vincenty (1982) who neglected the Δs terms in (2.20) which is a reasonable assumption.

From (2.18) we can identify specific changes in the rectangular coordinates due to scale and due to rotation effects. We define the following quantities:

$$\begin{aligned} \Delta x_s &= x\Delta s \\ \Delta y_s &= y\Delta s \\ \Delta z_s &= z\Delta s \end{aligned} \tag{2.21}$$

$$\begin{aligned} \Delta x_r &= \omega_z y - \omega_y z \\ \Delta y_r &= -\omega_z x - \omega_x z \\ \Delta z_r &= -\omega_y x - \omega_x y \end{aligned} \tag{2.22}$$

With these symbols, our seven parameter similarity transformation can be written in the form:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} + \begin{bmatrix} \Delta x_s \\ \Delta y_s \\ \Delta z_s \end{bmatrix} + \begin{bmatrix} \Delta x_r \\ \Delta y_r \\ \Delta z_r \end{bmatrix} \tag{2.23}$$

We see that each effect takes on the form of a translation that will depend on scale change or rotation effects.

Some physical significance can be given to the rotation parameters if we recognize that the diagonal elements of the rotation matrix, R, (in 2.4) represent the direction cosines between the new and old like axes. We can write, for example, from (2.13):

$$\cos (x, X) = \cos \omega_y \cos \omega_z \qquad (2.24)$$

$$\cos (y, Y) = \cos \omega_x \cos \omega_z - \sin \omega_x \sin \omega_y \sin \omega_z \qquad (2.25)$$

$$\cos (z, Z) = \cos \omega_x \cos \omega_y \qquad (2.26)$$

These angles can be expressed in the following form:

$$\cos (x, X) \equiv \cos \delta_{\lambda'} \qquad (2.27)$$

$$\cos (y, Y) \equiv \cos \delta_\lambda \qquad (2.28)$$

$$\cos (z, Z) \equiv \cos \delta_z \qquad (2.29)$$

We then have:

$$\delta_{\lambda'} \cong \left(\omega_y^2 + \omega_z^2 \right)^{1/2} \qquad (2.30)$$

$$\delta_\lambda \cong \left(\omega_x^2 + \omega_z^2 \right)^{1/2} \qquad (2.31)$$

$$\delta_z \cong \left(\omega_x^2 + \omega_y^2 \right)^{1/2} \qquad (2.32)$$

The δ_z angle is the angle between the directions fo the z and Z axes. The $\delta_{\lambda'}$ and δ_λ angles will represent the angle between the initial meridian of the two systems only when ω_x and ω_y are zero. Figure 2.3 shows a geometric interpretation of the three rotations.
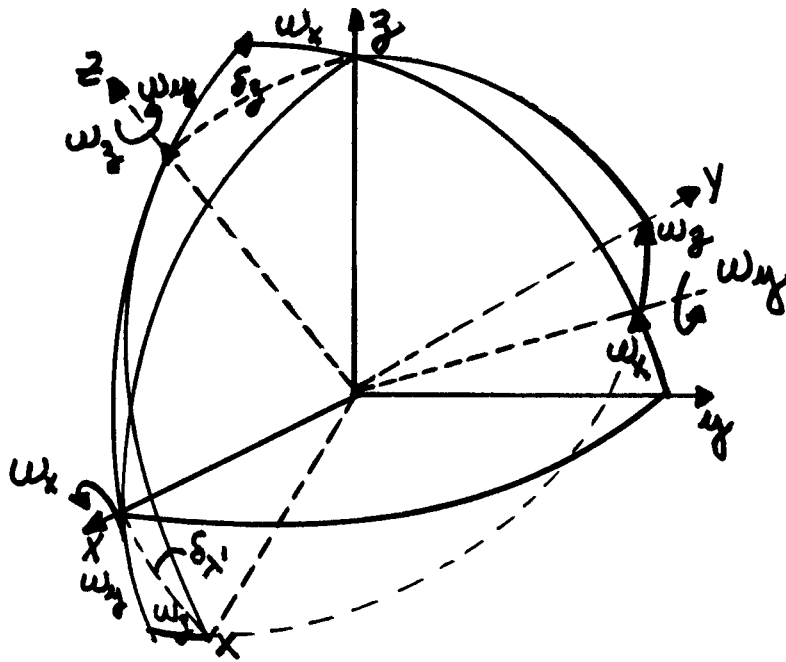
Figure 2.3 Angular Rotations in Going from x,y,z to X,Y,Z

A major problem in transformation work is the estimate of the seven (or less) parameters given estimates of the coordinates in the new and old system along with, in principle, the error variance matrix of these coordinates. This problem has been studied in several reports including those by Kumar (1972), Leick and van Gelder (1975), Adam (1982) and Malys (1988). To formulate our observation equation, we consider from (2.19) the observables as X,Y,Z,x,y,z while the parameters to be determined are: $\Delta x, \Delta y, \Delta z, \omega_x, \omega_y, \omega_z$, and Δs. We formulate the mathematical model for adjustment purposes as:

$$F \equiv \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} + U \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} + (1 + \Delta s) \begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

(2.33)

The linearized observation equation is:

$$BV + Ax^* + \omega = 0 \tag{2.34}$$

where V is the observation residuals and x* are the parameters, which may be corrections to assumed values.

We have:

$$V_i = \begin{bmatrix} v_x \\ v_y \\ v_z \\ v_X \\ v_Y \\ v_Z \end{bmatrix} ; \ x^* = \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \omega_x \\ \omega_y \\ \omega_z \\ \Delta s \end{bmatrix}$$

(2.35)

64

where i is the ith station. The elements of the B matrix are (for a given station):

$$
B_i = \begin{bmatrix} 1 & 0 & 0 \vdots -1 & 0 & 0 \\ 0 & 1 & 0 \vdots 0 & -1 & 0 \\ 0 & 0 & 1 \vdots 0 & 0 & -1 \end{bmatrix}
$$

(2.36)

The elements of A would be (again for a given station and neglecting the Δs term in (2.20):

$$
\begin{bmatrix} 1 & 0 & 0 \vdots 0 & -z & y \vdots x \\ 0 & 1 & 0 \vdots z & 0 & -x \vdots y \\ 0 & 0 & 1 \vdots -y & x & 0 \vdots z \end{bmatrix}
$$

(2.37)

A normal adjustment can be carried out to estimate the parameters under the least squares principle. A complication arrises when dealing with geodetic systems as the x,y,z coordinates are not generally derived. Usually given is information on the ϕ,λ,h triplet where h represents the height above the ellipsoid of the given datum. The ellipsoidal height is the sum of the orthometeric elevation and the astrogeodetic undulation (Rapp, 1984, Chapter 7). Since the astro geodetic undulations are determined from information including the geodetic coordinates, the h value is intrinsically correlated with both ϕ and λ. Therefore the error correlation matrix of ϕ,λ,h is a 3x3 full matrix which could be represented as:

$$
\Sigma_{\phi, \lambda, h} = \begin{bmatrix} \sigma_{\phi\phi} & \sigma_{\phi\lambda} & \sigma_{\phi h} \\ \sigma_{\lambda\phi} & \sigma_{\lambda\lambda} & \sigma_{\lambda h} \\ \sigma_{h\phi} & \sigma_{h\lambda} & \sigma_{hh} \end{bmatrix}
$$

(2.38)

This matrix can be propagated into the error correlation matrix for x,y,z, (or X,Y,Z). We can write:

$$
\Sigma_{x, y, z} = G \, \Sigma_{\phi, \lambda, h} \, G'
$$

(2.39)

where G is a matrix representing the partial derivatives of the transformation from ϕ, λ, h to x, y, z.

Specifically we have:

$$
G = \begin{bmatrix} -(M+h)\sin\phi\cos\lambda & -(N+h)\cos\phi\sin\lambda & \cos\phi\cos\lambda \\ -(M+h)\sin\phi\sin\lambda & (N+h)\cos\phi\cos\lambda & \cos\phi\sin\lambda \\ (M+h)\cos\phi & 0 & \sin\phi \end{bmatrix}
$$

(2.40)

We can see that, even if $\Sigma_{\phi,\lambda,h}$ is a diagonal matrix $\Sigma_{x,y,z}$ will not be.

In most geodetic networks the $\Sigma_{\phi,\lambda,h}$ is not rigorously available. Instead, various rules have been suggested that represent the proportional accuracy of a given network. One such rule was developed by Simmons (1950) based on the analysis of triangulation loop closures in NAD 1927. This rule states that the 2σ (standard deviation) proportional accuracy between two points in NAD27 can be given by:

$$2\sigma = 1 \text{ part in } 20,000\sqrt[3]{M} \tag{2.41}$$

where M is in miles. An equivalent statement is that the standard error in meters between two points separated by a distance of K (km):

$$E = 0.029 \, K^{2/3} \, (m) \tag{2.42}$$

Other accuracy estimates determine the accuracy of the distance (r) from the initial (origin) point to an arbitrary point in the network. Wells and Vaniček (1975) have used the following form:

$$\sigma_\phi = \sigma_\lambda = r_i^{2/3} k \quad \text{(meters)} \tag{2.43}$$

where they suggest $k = 0.0004 \text{ m}^{1/3}$ for the NAD27 and Australian datum, and $k = 0.0008$ $\text{m}^{1/3}$ for ED50 and the South American datum.

Other procedures for estimating triangulation accuracy have been discussed by Bomford (1980, p.172). He expressed the standard error of position as of function of the length of the chain, scale errors, and angular errors in the networks.

We finally turn to height accuracy. Estimates on the accuracy of the orthometric height can be derived from the levelling process. The accuracy of astro-geodetic undulations can also be estimated from rules (Rapp, ibid, Chapter 7). The magnitude of error correlations between the ϕ, λ quantities and h would be small.

The above guidelines are only approximations that enable some estimate of $\Sigma_{x,y,z}$ to be made. For proper weighting in the adjustment leading to proper statistical results, it is important that reliable statistical information on the accuracy of the geodetic networks be part of the solution process.

We should note here that the accuracy of the parameters being determined is sensitive to the geometry of the given points. Ideally, a global distribution of points is needed for good (i.e., low correlations between parameters), parameter determinations. If stations in a small area are analyzed, it may not be possible to effectively find all seven parameters since some will be highly correlated. For example, in some areas there will be insufficient information to determine ω_x and ω_y.

Malys (1988) has studied various station configurations to learn what parameters are best estimated with different station geometry. He did this by carrying out an adjustment with the simulated station positions and examining the error covariance matrix of the estimated parameters as a function, not only of station geometry, but of the error covariance matrix (specifically cross covariance terms) of the observed coordinates. One test carried out postulated 28 stations in the United States area ($20°\leq \phi \leq 50°$; $240°\leq \lambda \leq 300°$) at $10°$ increments in latitude and longitude. Malys (ibid) then examined the correlations between various parameters. He found that the scale parameter was never significantly correlated

66

with a rotation angle and that the rotation angles are only slightly correlated with each other. He found that the dominant correlations are between the translation parameters of one axis and the rotation parameters of another axis. For example, a correlation of 0.8 was found between Δx and ω_z, and -0.9 between Δz and ω_x. Malys pointed out on the basis of these results that x translation can be disguised as a rotation about a distant axis.

The previous discussion has outlined a method where the seven parameters can be simultaneously estimated. Alternate procedures have been developed that can determine Δs and the rotation angles independently of the other parameters. The scale difference can be estimated by comparing the chord distance between two points in the new and old system. For a single line, we can write:

$$\Delta s = \frac{C-c}{c}$$

(2.44)

where C is the chord distance between two points in the new system and c is the corresponding distance in the old. Note that this determination is independent of translation and rotation effects. A best estimate of Δs could be obtained by combining individual estimates of Δs from independent lines. Special care must be taken to recognize the various error correlations between station coordinates.

A procedure suggested by Bursa (1966, sec.5.28) enables the determination of the rotation angles independently of the scale and translation parameters. One version of this procedure derives the direction cosines of a line between two points in the old and the new system. In the old system, we can write for the line between stations i and k:

$$\cos(x,l) \equiv a = \frac{x_k - x_i}{c_{ik}}$$

$$\cos(y,l) \equiv b = \frac{y_k - y_i}{c_{ik}}$$

$$\cos(z,l) \equiv c = \frac{z_k - z_i}{c_{ik}}$$

(2.45)

where l indicates the direction of the line between i and k. The direction cosines in the new system would be designated A,B,C. We now can substitute the relationships given in (2.18) into the expressions for A,B,C to find:

$$A = a + \omega_z b - \omega_y c$$

$$B = b - \omega_z a + \omega_x c$$

$$C = c + \omega_y a - \omega_x b$$

(2.46)

Given the station information, the values of a,b,c,A,B,C along with their rigorously determined error covariance matrix are to be computed. A least squares adjustment is then carried out to determine the three rotation angles. However this adjustment does not recognize the implicit condition between the direction cosines (i.e., $A^2 + B^2 + C^2 = 1$, $a^2 + b^2 + c^2 = 1$).

An alternate procedure that could reduce the effect of neglected error correlations is to calculate two quantities (α and δ) from the direction cosines and to formulate two observation equations in these variables. We first define the following quantities in the old system.

$$T_{old} = -\tan^{-1} \frac{b}{a}$$

$$\delta_{old} = \tan^{-1} \frac{c}{(a^2 + b^2)^{1/2}}$$

(2.47)

with similar expressions for the _new_ system. The three direction cosine equations now become two equations in the two new variables:

$$\omega_z - \omega_x \cos T \tan\delta + \omega_y \sin T \tan\delta + (T_{old} - T_{new}) = v(t)$$

$$\omega_x \sin T + \omega_y \cos T + (\delta_{old} - \delta_{new}) = v(\delta)$$

(2.48)

Again a rigorous least squares adjustment can take place to estimate the rotation parameters independently of the other parameters.

These methods involving Δs, ω_x, ω_y, ω_z are attempts to solve the transformation problem using quantities that are invariant with respect to one or more other quantities. For example scale is invariant to translations and rotations; rotations are invariant with respect to scale and translations. At issue is the value of splitting up the adjustment into the various components. Leick and van Gelder (1975) have carried out tests with the same given data. They show that the results from either approach must be identical provided all assumptions made are the same. They recommend that the simultaneous adjustment process should be the preferred procedures since all seven transformation parameters and the corresponding error covariances are obtained at the same time.

The discussion concerning the seven parameter adjustment has used the usual least squares technique. Alternate adjustment procedures are possible. For example, Somogyi (1988) suggests the application of the robust estimation method for the parameter determination. In this method the weights for the observations are made dependent on the magnitude of the residuals in various ways that are specified.

2.22 The Veis Transformation Model

This similarity transformation model was proposed by Veis (1960). This form was an attempt to introduce rotations that could be associated with some process that took place at the datum origin point when the geodetic datum was originally defined. We first define a local right handed coordinate system at the datum origin point which is defined by ϕ_o, λ_o in the old (datum) system. The local axes are u (tangent to the geodetic meridian, positive south); v (perpendicular to the geodetic meridian passing through the datum origin, positive east); and w (in the direction of the (old) ellipsoid normal at the datum origin, positive up). This system, along with the new (X,Y,Z) and old (x,y,z) rectangular coordinate system are shown in Figure 2.4.
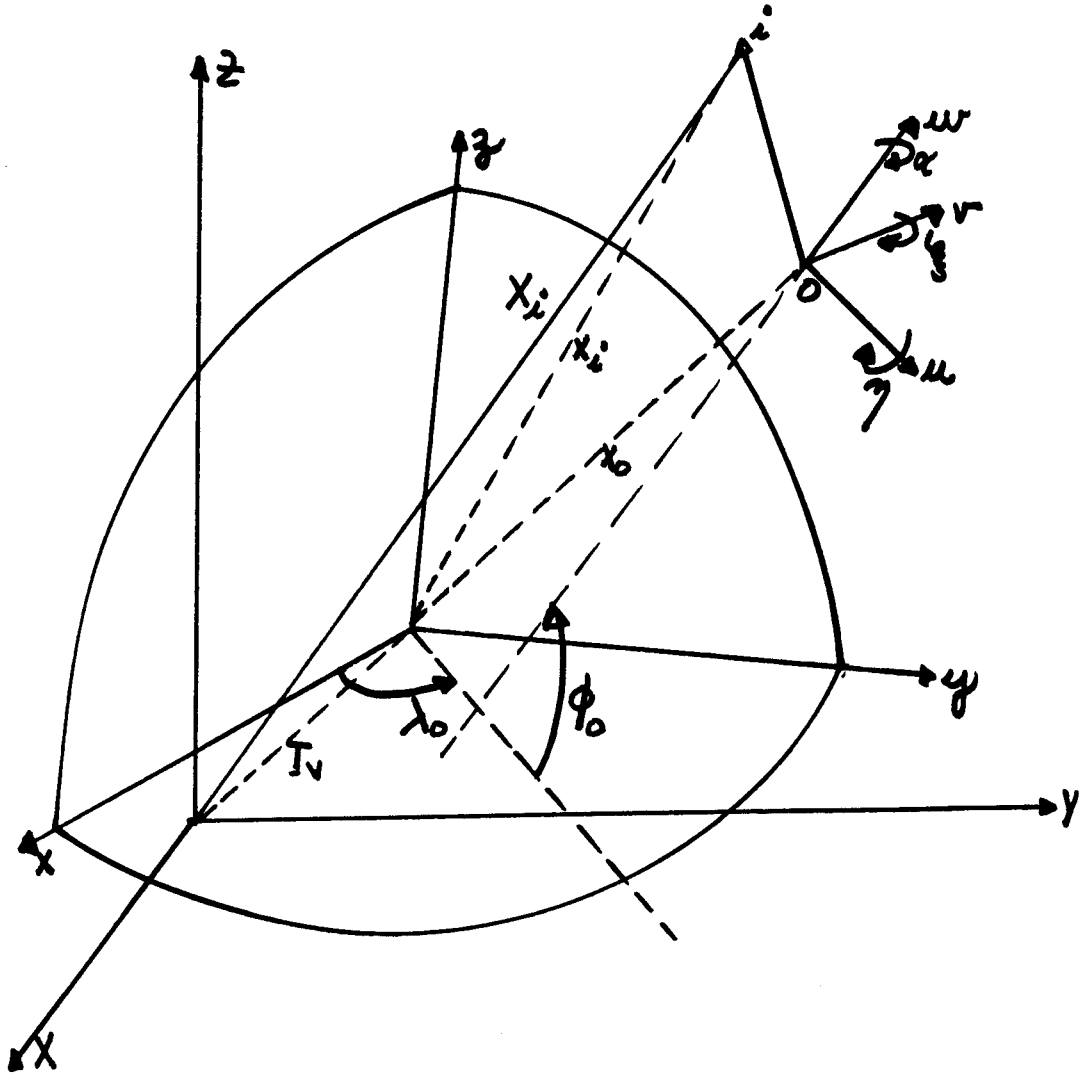
Figure 2.4
The Veis Transformation Method

The vector from the datum origin to an arbitrary point (i) in the datum would be $\underline{X}_i - \underline{X}_0$. Now the original alignment of the datum can be changed by considering small rotations about the local origin axes; α about w, ξ about v, η about u. The α rotation, for example, could be due to an azimuth error in the original azimuth definition. We want to apply these rotations to the vector from the origin to the ith point. To do this we rotate $\underline{X}_i - \underline{X}_0$ into the local system, apply the rotations, then rotate back into the rectangular system. This rotation can be accomplished using the following (see Rapp (1984, section 4.19)):

$$M \equiv R_z^T (\lambda_o)\, R_y^T (90^\circ - \phi_o)\, R_x(\eta)\, R_y(\xi)\, R_z(\alpha)\, R_y (90^\circ - \phi_o)\, R_z (\lambda_o)$$

(2.49)

The rotated vector is then scaled by a factor $(1 + \Delta s_v)$ where Δs_v is the scale change parameter associated with the Veis transformation. The complete transformation follows, somewhat, from the inspection of Figure 2.4. Actually we define the Veis transformation as follows:

69

$$X_v = \vec{T}_v + x_o + \left(1 + \Delta s_v\right) M\left(\phi_o, \lambda_o, \alpha, \xi, \eta\right)\left(x - x_o\right)$$

(2.50)

where T_v is the translation vector measured in the <u>new x</u> frame. Multiplying out (2.49) we have:

$$M = \begin{bmatrix} 1 & \begin{matrix} \alpha\sin\phi_o \\ -\eta\cos\phi_o \end{matrix} & \begin{matrix} -\alpha\cos\phi_o \sin\lambda_o \\ -\xi\cos\lambda_o \\ -\eta\sin\phi_o \sin\lambda_o \end{matrix} \\ \begin{matrix} -\alpha\sin\phi_o \\ +\eta\cos\phi_o \end{matrix} & 1 & \begin{matrix} \alpha\cos\phi_o\cos\lambda_o \\ -\xi\sin\lambda_o \\ +\eta\sin\phi_o\cos\lambda_o \end{matrix} \\ \begin{matrix} \alpha\cos\phi_o\sin\lambda_o \\ +\xi\cos\lambda_o \\ +\eta\sin\phi_o\sin\lambda_o \end{matrix} & \begin{matrix} -\alpha\cos\phi_o\cos\lambda_o \\ +\xi\sin\lambda_o \\ -\eta\sin\phi_o\cos\lambda_o \end{matrix} & 1 \end{bmatrix}$$

(2.51)

Equation (2.50) can be modified such that a linear adjustment model can be established to estimate the seven $(\Delta x, \Delta y, \Delta z, \Delta s_v, \alpha, \xi, \eta)$ parameters of this model.

We can compare new coordinate difference computed with the Bursa-Wolf system with those computed by the Veis system. From the origin to the ith point, in the new system, we have, from (2.18):

$$\Delta X_{io(B)} = (1 + \Delta s_B) R_B (\omega_x, \omega_y, \omega_z)\left(x_i - x_0\right)$$

(2.52)

This difference in the Veis model would be, from (2.50):

$$\Delta X_{io(V)} = \left(1 + \Delta s_v\right) M\left(\phi_0, \lambda_0, \alpha, \xi, \eta\right)\left(x_i - x_0\right)$$

(2.53)

Since ΔX must be the same from both equations, and $(x_i - x_0)$ is the same, the equality of (2.52) and (2.53) implies the following:

$$\Delta s_B = \Delta s_v$$

(2.54)

$$R_B (\omega_x, \omega_y, \omega_z) = M (\phi_0, \lambda_0, \alpha, \xi, \eta)$$

(2.55)

The scale factor in the Veis and Bursa-Wolf system are the same. The implications of (2.55) is found by equating (2.14) to (2.51). We have:

70

$$\omega_x = \alpha \cos\phi_0 \cos\lambda_0 - \xi \sin\lambda_0 + \eta \sin\phi_0 \cos\lambda_0$$

$$\omega_y = \alpha \cos\phi_0 \sin\lambda_0 + \xi \cos\lambda_0 + \eta \sin\phi_0 \sin\lambda_0 \qquad (2.56)$$

$$\omega_z = \alpha \sin\phi_0 - y \cos\phi_0$$

We can invert (2.56) to write:

$$\begin{pmatrix} \alpha \\ \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \sin\phi_0 & \cos\phi_0 \sin\lambda_0 & \cos\phi_0 \cos\lambda_0 \\ 0 & \cos\lambda_0 & -\sin\lambda_0 \\ -\cos\phi_0 & \sin\phi_0 \sin\lambda_0 & \sin\phi_0 \cos\lambda_0 \end{pmatrix} \begin{pmatrix} \omega_z \\ \omega_y \\ \omega_x \end{pmatrix} \qquad (2.57)$$

It is clear that the rotations of one system have a complete analogy with the rotations in the other system through (2.57).

A special case of (2.51) and (2.55) occur if we assume ξ and η are zero. That such quantities should be zero has been discussed by Wells and Vanicek (1975) and Vanicek and Carrera (1985). In essence the authors argue that if the parallel conditions involving astronomic coordinates, geodetic coordinates and deflections of the vertical are maintained (Rapp, 1984, Chapter 7) at the datum origin point, ξ and η (i.e., the _corrections_ to the assumed deflections) should be zero. This would leave only a rotation, α, about the geodetic normal, as the remaining rotation. Under these circumstances equation (2.51) becomes.

$$M_\alpha = \begin{bmatrix} 1 & \alpha \sin\phi_0 & -\alpha \cos\phi_0 \sin\lambda_0 \\ -\alpha \sin\phi_0 & 1 & \alpha \cos\phi_0 \cos\lambda_0 \\ \alpha \cos\phi_0 \sin\lambda_0 & -\alpha \cos\phi_0 \cos\lambda_0 & 1 \end{bmatrix} \qquad (2.58)$$

In addition the rotations about the x, y, z axes would be given by (2.56) with ξ and η zero:

$$\omega_{x\alpha} = \alpha \cos\phi_0 \cos\lambda_0$$

$$\omega_{y\alpha} = \alpha \cos\phi_0 \sin\lambda_0 \qquad (2.59)$$

$$\omega_{z\alpha} = \alpha \sin\phi_0$$

Equation (2.59) is also equation (3.16) in Vincenty (1985) and equation (1) in Vanicek and Carrera (1985).

We next compare equation (2.9) (Bursa/Wolf) and (2.50) (Veis) for the transformed coordinates recognizing the equality given in (2.54) and (2.55). We have:

$$\underline{T}_B = \underline{x}_0 - (1 + \Delta s) \; R\left(\omega_x, \omega_y, \omega_z\right) \underline{x}_0 + \underline{T}_v \tag{2.60}$$

It is clear from (2.60) that the translation vectors of the two models are generally different. This occurs because of the manner in which the Veis transformation is defined by equation (2.50). Such a definition leads to a translation vector without geometric meaning.

2.23 The Molodensky Transformation Model

A set of differential formulas for transformation to a new coordinate system is described in Molodensky et al., (1962, Section 3). The discussion in this book relates to the calculation of latitude, longitude, and height changes considering eight parameter changes. These are three rotations, three translations, and two ellipsoid parameter changes. Our previous discussion has excluded ellipsoid parameter changes and we will modify the Molodensky discussion, for now, to continue this exclusion. We also note that Molodensky allowed a change to the geodetic coordinates in the old system. Our prior discussion has not introduced such a change and therefore, in this discussion, we will set such changes (dB, dL, dH in Molodensky's notation) to zero. Various interpretations and/or application of the Molodensky transformation have been given in Badekas (1969), Leick and van Gelder (1975) Soler (1975) and others. Our discussion will follow that of Soler (ibid).

The Molodensky transformation is designed to consider a translation and the rotation of a vector from the datum origin point to a arbitrary point in the system. The rotations are about the old (x, y, z) axes. We have from Molodensky (ibid, eq. (I-3.2)) and Soler (1976, equations (4.3-2), (4.4-5) or (A.1-5)):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \delta x_0 \\ \delta y_0 \\ \delta z_0 \end{bmatrix} + \begin{bmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{bmatrix} \begin{bmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{bmatrix} \tag{2.61}$$

where x_0, y_0, z_0 are the coordinates of the datum origin point. The δx_0, δy_0, δz_0 are quantities that require careful interpretation.

Let the rectangular shifts between the datum point, in the new system, and the old system be dX, dY, dZ. We have:

$$\begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix}_X = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix}_X - \begin{bmatrix} \overline{X}_0 \\ \overline{Y}_0 \\ \overline{Z}_0 \end{bmatrix}_X \tag{2.62}$$

where X_0, Y_0, Z_0 are the new coordinates based on the new origin, and $\overline{X}_0$, $\overline{Y}_0$, $\overline{Z}_0$ are the coordinates of the datum origin based on the old coordinate origin but with the new axis alignment as seen in Figure 2.5. Note that dX, dY, dZ correspond to $d\overline{x}$, $d\overline{y}$, $d\overline{z}$ given in Molodensky (ibid, eq. (I.3.4)).
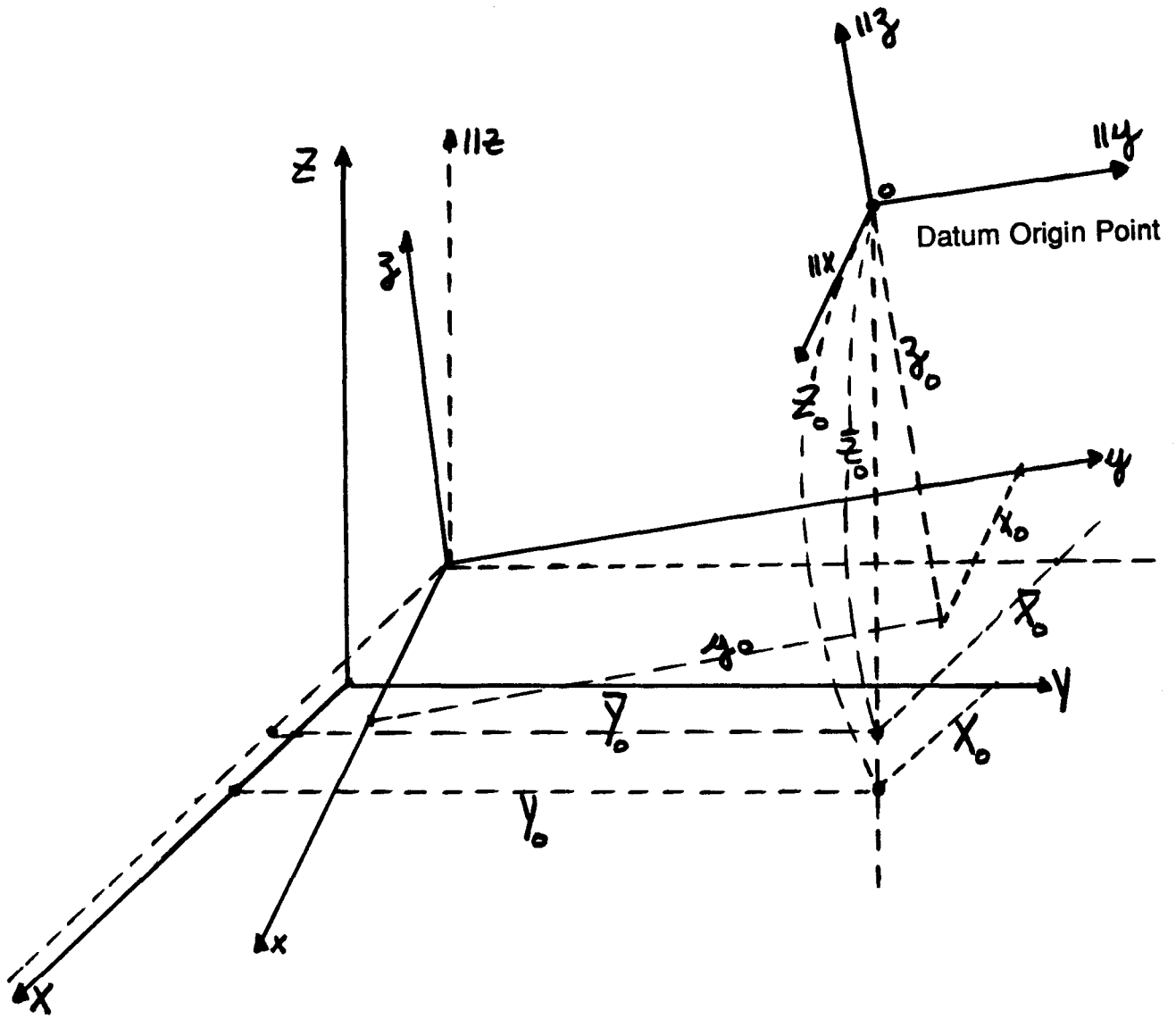
Figure 2.5
Geometry of the Molodensky Transformation Model

Now the $(\overline{X}, \overline{Y}, \overline{Z})_0$ vector can be obtained by rotating the $(x, y, z)_0$ vector from the old system to the new system. From eq. (2.17) we can write (letting $\Delta x=\Delta y=\Delta z=\Delta s=0$):

$$\begin{bmatrix} \overline{X}_0 \\ \overline{Y}_0 \\ \overline{Z}_0 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

(2.63)

Then eq. (2.62) becomes:

$$\begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix}_x = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix}_x - \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} - \begin{bmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

(2.64)

73

We now write equation (I.3.4) in Molodensky in our notation:

$$
\begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix}_M = \begin{bmatrix} \delta x_0 \\ \delta y_0 \\ \delta z_0 \end{bmatrix} - \begin{bmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}
\tag{2.65}
$$

Comparing (2.64) and (2.65) we can see that

$$
\begin{bmatrix} \delta x_0 \\ \delta y_0 \\ \delta z_0 \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} - \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}
\tag{2.66}
$$

Note that (2.65) does not strictly give a translation vector as the coordinates used are defined in <u>different</u> coordinate systems. Now solve (2.65) for $(\delta x_0, \delta y_0, \delta z_0)$ and substitute into (2.61):

$$
\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix} + \begin{bmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}
\tag{2.67}
$$

We can compare this equation with the Bursa/Wolf model given by equation (2.17) (setting $\Delta s = 0$). We see that the equations are the same so that the Molodensky transformation model is, in reality, the same as the Bursa/Wolf similarity transformation.

2.24 The Vaniček-Wells Transformation Models

A transformation described by Wells and Vaniček (1975) introduces several coordinate systems in dealing with a network (datum) coordinate system, a coordinate frame defined by a particular satellite observation system, and an ideal system such as the Conventional Terrestrial System (CTS) (or its successor, the IERS Reference Frame). Vaniček and Wells postulate station positions given in the network system is assumed properly aligned with the exception of a single azimuth rotation about the ellipsoidal normal at the datum origin. The space system axes (X', Y', Z') are considered rotated by amounts $(\omega_x, \omega_y, \omega_z)$ with respect to the CTS (X, Y, Z). This information is portrayed in Figure 2.6 where the datum origin is at O, and an arbitrary point in the system is i.

74
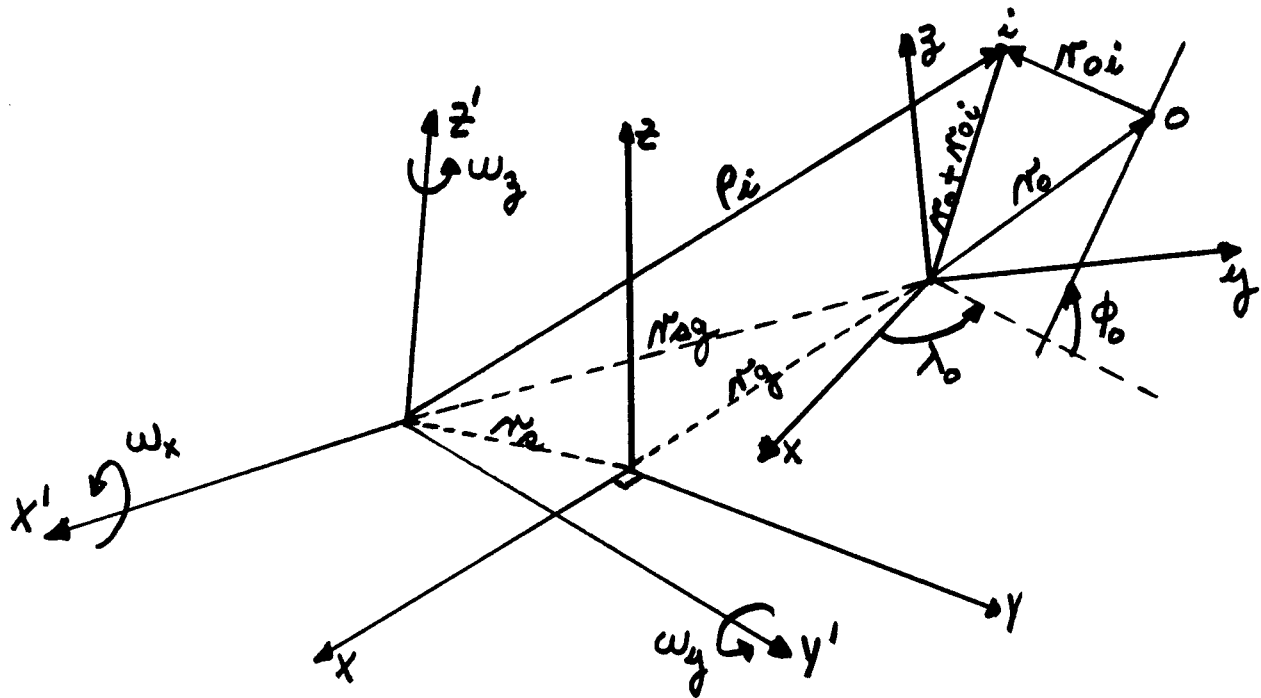
Figure 2.6
The Wells-Vaniček Transformation Method

The following quantities are three component vectors: ρ_i, r_s, r_g, r_{sg}, r_o, r_i. From the figure we can see that there are two ways in which the vector to the arbitrary point can be represented in the CTS. We have:

$$\underline{X}_i = \underline{r}_s + R(\omega_x, \omega_y, \omega_z)\underline{\rho}_i \tag{2.68}$$

$$\underline{X}_i = \underline{r}_g + (1 + \Delta s) R(\alpha) (\underline{r}_o + \underline{r}_{oi}) \tag{2.69}$$

Equation (2.68) assumes the scale of the satellite system and CTS are the same. In (2.69) the scale difference between the CTS and network is Δs as dealt with earlier. In (2.69) the Δs is applied to the vector from the center of the datum to the point after this vector is rotated about the ellipsoid normal at the datum origin.

The elements of $R(\omega_x, \omega_y, \omega_z)$ are given by (2.14) while the elements of $R(\alpha)$ are given by (2.57). If we let $R(\omega_x, \omega_y, \omega_z) = I + Q$, and $R(\alpha) = I + A(\alpha)$ we can equate (2.68) and (2.69):

$$\underline{r}_s + \underline{\rho}_i + Q(\omega_x, \omega_y, \omega_z)\underline{\rho}_i = \underline{r}_g + \underline{r}_o + \underline{r}_{oi} + A(\alpha)(\underline{r}_o + \underline{r}_{oi}) + \Delta s(\underline{r}_o + \underline{r}_{oi}) \tag{2.70}$$

With sufficient accuracy we can let $\underline{r}_o + \underline{r}_{oi} = \rho_i$ in the last two terms in (2.70). We then can write:

$$[Q(\omega_x, \omega_y, \omega_z) - \underline{A}(\alpha) - I\Delta s] \underline{\rho}_i - (\underline{r}_g - \underline{r}_s) = \underline{r}_o + \underline{r}_i - \underline{\rho}_i \tag{2.71}$$

75

In this equation we have the right hand side known while there are a set of parameters to be determined. These values are ω_x, ω_y, ω_z, α, Δs, and three translation difference terms of $(\underline{r}_g - \underline{r}_s)$ for a total of 8 parameters if only one datum is being considered. If we consider data from several different systems we will have an additional five parameters per datum. Since $Q(\omega_x, \omega_y, \omega_z)$ and $A(\alpha)$ enter in the same way on ρ_i, it will not be possible to separate α from ω_x, ω_y, ω_z if only one datum is being considered. We also conclude that we must have a minimum of two stations per datum to achieve a solution.

Wells and Vaniček (ibid) applied this transformation model to data given on several geodetic datums. Since the data available at that time was sparse, their results would be regarded as encouraging rather than definitive. Additional computations are now warranted with the improved satellite derived station coordinate that are available.

2.3 Geodetic Coordinate Transformation

The discussion in the previous section has been directed to the conversion of rectangular coordinates in an "old" system to coordinates in a "new" system. We recognized that the "old" coordinates would be determined by combining horizontal and vertical datum information together. Now that such transformations have been developed it is time to consider going back to a latitude, longitude, and ellipsoid height. Assume that we have the transformed rectangular coordinates X, Y, Z. We want to obtain the ϕ, λ, h with respect to some ellipsoid whose parameters (a,f) are defined. The procedure for doing this has been discussed via several techniques in Rapp (1984, Section 6.8) and presents no special problems.

An alternate method is to develop a differential procedure. We can write (2.23) in the following form:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}_T \tag{2.72}$$

where $[\Delta x, \Delta y, \Delta z]_T$ represent the sum of the translation, scale, and rotation effects shown on the right hand side of (2.23). An analogous equation in geodetic coordinates would be:

$$\begin{bmatrix} \bar{\phi} \\ \bar{\lambda} \\ \bar{h} \end{bmatrix} = \begin{bmatrix} \phi \\ \lambda \\ h \end{bmatrix} + \begin{bmatrix} \Delta\phi \\ \Delta\lambda \\ \Delta h \end{bmatrix} \tag{2.73}$$

In both (2.72) and (2.73) we regard the quantities as differential in nature. Our next task is to calculate $\Delta\phi$, $\Delta\lambda$, Δh as a function of $[\Delta x, \Delta y, \Delta z]_T$ and ellipsoid change (old to new) parameters.

2.31 A Differential Projective Transformation Procedure

We first repeat the standard equations relating rectangular and geodetic coordinates:

$$\begin{aligned} x &= (N + h) \cos\phi \cos\lambda \\ y &= (N + h) \cos\phi \sin\lambda \\ z &= (N(1-e^2) + h) \sin\phi \end{aligned} \tag{2.74}$$

76

We differentiate each of these equations with respect to five variables: ϕ, λ, h, a, f. For example:

$$dx = \frac{\partial x}{\partial \phi} d\phi + \frac{\partial x}{\partial \lambda} d\lambda + \frac{\partial x}{\partial h} dh + \frac{\partial x}{\partial a} da + \frac{\partial x}{\partial f} df$$

(2.75)

with similar equations for dy and dz. The dx, dy, dz quantities can be associated with the total changes $[\Delta x, \Delta y, \Delta z]_T$ or with any of the specific changes associated with translation, scale, or rotation.

The derivatives needed for (2.75) and the other expressions are as follows:

$$\frac{\partial x}{\partial \phi} = -(M+h)\sin\varphi \cos\lambda, \qquad \frac{\partial x}{\partial \lambda} = -(N+h)\cos\varphi \sin\lambda,$$

$$\frac{\partial y}{\partial \phi} = -(M+h)\sin\varphi \sin\lambda, \qquad \frac{\partial y}{\partial \lambda} = (N+h)\cos\varphi \cos\lambda,$$

$$\frac{\partial z}{\partial \phi} = (M+h)\cos\varphi, \qquad \frac{\partial z}{\partial \lambda} = 0,$$

$$\frac{\partial x}{\partial a} = \frac{\cos\varphi \cos\lambda}{W}, \qquad \frac{\partial x}{\partial e^2} = \frac{a \sin^2\varphi \cos\varphi \cos\lambda}{2W^3} \qquad \frac{\partial x}{\partial h} = \cos\phi \cos\lambda$$

$$\frac{\partial y}{\partial a} = \frac{\cos\varphi \sin\lambda}{W}, \qquad \frac{\partial y}{\partial e^2} = \frac{a \sin^2\varphi \cos\varphi \sin\lambda}{2W^3} \qquad \frac{\partial y}{\partial h} = \cos\phi \sin\lambda$$

$$\frac{\partial z}{\partial a} = \frac{(1-e^2)\sin\varphi}{W}, \qquad \frac{\partial z}{\partial e^2} = \frac{1}{2}(M\sin^2\varphi - 2N)\sin\varphi \qquad \frac{\partial z}{\partial h} = \sin\phi$$

(2.76)

where:

$$W^2 = 1 - e^2\sin^2\varphi$$

$$M = \frac{a(1-e^2)}{W^3}, \quad N = \frac{a}{W}$$

(2.77)

To find changes with respect to the flattening, we note that:

$$\frac{\partial}{\partial f} = \frac{\partial}{\partial e^2} \cdot \frac{\partial e^2}{\partial f}$$

(2.78)

Since $e^2 = 2f - f^2$ we have:

$$\frac{\partial e^2}{\partial f} = 2(1-f)$$

(2.79)

Using the derivatives in equations (2.76) and (2.79), we can write the three equations implied in equation (2.75) as follows:

$$dx = -(M+h)\sin\phi\cos\lambda d\phi - (N+h)\cos\phi\sin\lambda d\lambda + \cos\phi\cos\lambda dh$$
$$+ \frac{\cos\phi\cos\lambda}{W}da + \frac{a(1-f)\sin^2\phi\cos\phi\cos\lambda}{W^3}df$$

(2.80)

$$dy = -(M+h)\sin\phi\sin\lambda d\phi + (N+h)\cos\phi\cos\lambda d\lambda + \cos\phi\sin\lambda dh$$
$$+ \frac{\cos\phi\sin\lambda}{W}da + \frac{a(1-f)\sin^2\phi\cos\phi\sin\lambda}{W^3}df$$

(2.81)

$$dz = (M+h)\cos\phi d\phi \qquad\qquad + \sin\phi dh$$
$$+ \frac{(1-e^2)\sin\phi da}{W} + (M\sin^2\phi - 2N)(1-f)\sin\phi df$$

(2.82)

Various approximations may be made to equations (2.80), (2.81), and (2.82) to simplify them . For some computations this may be desirable, but when calculations are done on a computer no reduction appears called for. As an example of a simplification, we write the equations assuming the coefficients of the differential changes refer to a spherical earth of radius a, and that $h = 0$. We find:

$$dx = -a\sin\varphi\cos\lambda d\varphi - a\cos\phi\sin\lambda d\lambda + \cos\varphi\cos\lambda(dh + da + a\sin^2\varphi df)$$

(2.83)

$$dy = -a\sin\varphi\sin\lambda d\varphi + a\cos\phi\cos\lambda d\lambda + \cos\varphi\sin\lambda(dh + da + a\sin^2\varphi df)$$

(2.84)

$$dz = a\cos\varphi d\varphi + \sin\varphi(da + dh) + a\sin\varphi(\sin^2\varphi - 2)df$$

(2.85)

These equations may also be found in Heiskanen and Moritz (1967, p. 206, eq. 5-54). A better approximation to equations (2.80), (2.81) and (2.82) may be found in Vincenty (1966, eq. 5). It should be noted, however, that the equations (2.80), (2.81) and (2.82) are the exact <u>differential</u> equations. The accuracy of these equations will depend on the magnitude of the changes since, implicitly, the terms are first terms in a Taylor's series, with terms in $\Delta\varphi^2$, $\Delta\lambda^2$, Δh^2, Δa^2, Δf^2 and higher powers neglected.

Given dx, dy, and dz as well as da and df, we must now develop equations to give us $d\varphi$, $d\lambda$, and dh. We may note that solution 2 obtained by regarding (2.80), (2.81) and (2.82) as three equations in three unknowns. If we were to rewrite these equations, we could put them in the matrix form shown symbolically as follows:

78

$$\begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix} \begin{bmatrix} d\varphi \\ d\lambda \\ dk \end{bmatrix} = \begin{bmatrix} D_1 \\ D_a \\ D_3 \end{bmatrix} \tag{2.86}$$

where the terms A, B, C, and D are known quantities. Consequently, $d\varphi$, $d\lambda$, and dh may be found by inverting the coefficient matrix and multiplying by the D vector. This procedure is inconvenient if we need to consider only a single change, and consequently, we proceed to find separate expressions for $d\varphi$, $d\lambda$ and dh.

To find $d\varphi$, multiply (2.80) by $-\sin\varphi \cos\lambda$; (2.81) by $-\sin\varphi \sin\lambda$; and (2.82) by $\cos\varphi$. Add the resulting equations to find $d\phi$ separately. To find $d\lambda$, multiply (2.80) by $-\sin\lambda$; (2.81) by $\cos\lambda$; and (2.82) by zero, and then add as before. To find dh, multiply (2.80) by $\cos\varphi \cos\lambda$; (2.81) by $\cos\varphi \sin\lambda$; and (2.82) by $\sin\varphi$. We find:

$$(M + h)d\varphi = -\sin\varphi \cos\lambda \, dx - \sin\varphi \sin\lambda \, dy + \cos\varphi \, dz + \frac{e^2 \sin\varphi \cos\varphi}{W} da$$
$$+ \sin\varphi \cos\phi \left(2N + e'^2 M \sin^2 \varphi\right)(1 - f) df \tag{2.87}$$

$$(N + h)\cos\varphi d\lambda = -\sin\lambda dx + \cos\lambda dy \tag{2.88}$$

$$dh = \cos\varphi \cos\lambda dx + \cos\varphi \sin\lambda dy + \sin\varphi dz - W da + \frac{a(1 - f)}{W} \sin^2\varphi df \tag{2.89}$$

Equations (2.87), (2.88), and (2.89) represent working formulas for converting geodetic coordinates referred to an old system to a new system. We must specify the shifts by Δx_T, Δy_T, Δz_T (which are dx, dy, and dz) and the parameters of the new ellipsoid. Note that dx, dy, dz will only be constants if there is no orientation and scale effects. If this is not the case, dx, dy, dz will be position dependent. Spherical approximations to (2.87), (2.88), (2.89) are given in Heiskanen and Moritz (1967, p. 207, equation (5-55) with a more accurate approximation being given by Vincenty (1966, equation (10)).

2.31.1 The Molodensky Geodetic Coordinate Transformation

Section 2.23 discussed the rectangular coordinate transformation using a method described in Molodensky, et al., (1962). Equations (I.3.5, I.3.6., and I.3.7) in Molodensky, et al., can be used to calculate changes in latitude, longitude and height as (2.87), (2.88), and (2.89) do. As the Molodensky formulas are used by a number of different groups (e.g., see DMA WGS84 report, 1987), they are repeated here in a form similiar to our previously derived values. We have:

$$(M + h)d\phi = -\sin\phi \cos\lambda dx - \sin\phi \sin\lambda dy + \cos\phi dz + \frac{e^2 \sin\phi \cos\phi}{W} da$$
$$+ \sin\phi \cos\phi \left(M\frac{a}{b} + N\frac{b}{a}\right) df \tag{2.90}$$

$$(N + h)\cos\phi d\lambda = -\sin\lambda dx + \cos\lambda dy \tag{2.91}$$

$$dh = \cos\phi \cos\lambda dx + \cos\phi \sin\lambda dy + \sin\phi dz - W da + \frac{a(1 - f)}{W} \sin^2\phi df \tag{2.92}$$

We see that (2.91) is identical to (2.88); (2.92) is identical to (2.89) and (2.90) differs from (2.87) only in the coefficient of df.

A set of "Abridged Molodensky Formulas" can be obtained by setting h equal to zero and simplifying the coefficients of da and df. These formulas are:

$$Md\phi = -\sin\phi \cos\lambda \, dx - \sin\phi \sin\lambda \, dy + \cos\phi \, dz + (adf + fda) \sin2\phi \qquad (2.93)$$

$$Nd\lambda = -\sin\lambda \, dx + \cos\lambda \, dy \qquad (2.94)$$

$$dh = \cos\phi \cos\lambda \, dx + \cos\phi \sin\lambda \, dy + \sin\phi \, dz + (adf + fda) \sin^2\phi - da \qquad (2.95)$$

2.31.2 Geodetic Coordinate Changes Caused by Changes at the Datum Origin Point Due to Shift and Ellipsoid Changes.

Equations (2.87), (2.88) and (2.89) are convenient if the values of dx, dy, and dz are given. In some cases we desire to know the changes in coordinates at any point in our system if the coordinate changes at the origin, and ellipsoid changes are given. This, of course, may be done as a two-step problem, first computing dx, dy, dz from (2.80), (2.81) and (2.82) and then applying these values in (2.87), (2.88) and (2.89). However, we seek a set of equations that eliminates the two-step procedure. First we assume in the following discussion that the changes being considered are due solely to the origin shifts (dx, dy, dz) and ellipsoid parameter changes. The effects due to the other quantities will be considered later.

Now, evaluate (2.80), (2.81), (2.82) at the origin point designated by subscript 0. We then have:

$$dx = -(M + h)_0 \sin\varphi_0 \cos\lambda_0 d\varphi_0 - (N + h)_0 \cos\varphi_0 \sin\lambda_0 d\lambda_0 + \cos\varphi_0 \cos\lambda_0 dh_0$$

$$+ \frac{\cos\varphi_0 \sin\lambda_0}{W_0} da + \frac{a(1-f)\sin^2\varphi_0 \cos\varphi_0 \cos\lambda_0}{W_0^3} df \qquad (2.96)$$

$$dy = -(M + h)_0 \sin\varphi_0 \sin\lambda_0 d\varphi_0 + (N + h)_0 \cos\varphi_0 \cos\lambda_0 d\lambda_0 + \cos\varphi_0 \sin\lambda_0 dh_0$$

$$+ \frac{\cos\varphi_0 \sin\lambda_0}{W_0} da + \frac{a(1-f)\sin^2\varphi_0 \cos\varphi_0 \sin\lambda_0}{W_0^3} df \qquad (2.97)$$

80

$$dz = (M + h)_0 \cos \varphi_0 \, d\varphi_0 + \sin \lambda_0 dh_0 + \frac{(1 - e^2) \sin\varphi_0}{W_0} da$$

$$+ (M \sin^2 \varphi - 2N \sin \varphi)_0 (1 - f) \, df \qquad (2.98)$$

Now substitute these equations in (2.87), (2.88), (2.89) to find:

$$(M + h) \, d\varphi = (M + h)_0 (\cos\varphi \cos\varphi_0 + \sin \varphi_0 \sin \varphi \cos \Delta\lambda) \, d\varphi_0$$

$$-(N + h)_0 \sin \varphi \cos \varphi_0 \sin \Delta\lambda d\lambda_0$$

$$+ (\sin \varphi_0 \cos \varphi - \cos \varphi_0 \sin \varphi \cos \Delta\lambda) \, dh_0$$

$$+ [\frac{1}{W_0} (\sin \varphi_0 \cos \varphi \, (1 - e^2) - \cos \varphi_0 \sin \varphi \cos \Delta\lambda) + \frac{e^2}{W} \sin \varphi \cos \varphi] \, da$$

$$+ (\sin \varphi_0 \cos \varphi \, (M\sin^2 \phi - 2N)_0 (1 - f) - K_0 \sin \varphi \cos \Delta\lambda$$

$$+ \sin \varphi \cos \varphi \, (1 - f) \, (2N + e'^2 M\sin\varphi)) \, df \qquad (2.99)$$

$$(N + h) \cos \varphi d\lambda = (M + h)_0 \sin\varphi_0 \sin \Delta\lambda d\varphi_0$$

$$+ (N + h)_0 \cos \varphi_0 \cos \Delta\lambda \, d\lambda_0$$

$$- \cos\varphi_0 \sin \Delta\lambda dh_0$$

$$- \frac{\cos \varphi_0}{W_0} \sin \Delta\lambda da$$

$$- K_0 \sin \Delta\lambda df \qquad (2.100)$$

$$dh = (M + h)_0 (\cos\varphi_0 \sin\varphi - \sin\varphi_0 \cos\varphi \cos\Delta\lambda) \, d\varphi_0$$

$$+ (N + h)_0 \cos \varphi \cos \varphi_0 \sin \Delta\lambda \, d\lambda_0$$

$$+ (\sin \varphi_0 \sin \varphi + \cos \varphi_0 \cos \varphi \cos \Delta\lambda) \, dh_0$$

$$+ \frac{1}{W_0} (\sin \varphi_0 \sin \varphi \, (1 - e^2) + \cos \varphi_0 \cos \varphi \cos \Delta\lambda - WW_0) \, da$$

$$+ (\cos \varphi \cos \Delta\lambda K_0 + \sin \varphi \, (1 - f) \, ((M\sin^3 \varphi - 2N\sin\varphi)_0 + \frac{a}{W} \sin \varphi)) \, df \qquad (2.101)$$

where:

$$\Delta\lambda = \lambda - \lambda_0$$

$$K_0 = \frac{a \, (1 - f) \sin^2 \varphi_0 \cos\varphi_0}{W_0^3}$$

The spherical form of equations (2.99), (2.100), and (2.101) may be found in Heiskanen and Moritz (1967, p. 207, equation (5-57)).

Vening-Meinesz (1950) derived equations similar to (2.99), (2.100), and (2.101). Although in terms of the deflection of the vertical they may easily be converted to the form of the above equations. His derivation makes use of some series expansions that generally retain terms including f^2.

Examination of equations (2.99), (2.100) and (2.101) shows that they may be written in the general form as follows:

$$d\varphi = E_1 d\varphi_0 + E_2 d\lambda_0 + E_3 dh_0 + E_4 da + E_5 df$$

$$d\lambda = F_1 d\varphi_0 + F_2 d\lambda_0 + F_3 dh_0 + F_4 da + F_5 df$$

$$dh = G_1 d\varphi_0 + G_2 d\lambda_0 + G_3 dh_0 + G_4 da + G_5 df \qquad (2.102)$$

where E_1, F_1, and G_1 are coefficients determined by comparison of (2.102) with (2.99), (2.100) and (2.101).

Again we should note that the $d\varphi$, $d\lambda$ and dh terms in (2.102) are due only to origin shifts and ellipsoid changes. We have implicitly assumed that the axes of the two systems are parallel and the scale difference is zero.

2.31.3 Differential Change Formulas in Terms of Deflections of the Vertical and Geoid Undulations (or Height anomalies)

From previous discussion we know that we can write, with sufficient accuracy for this differential purpose:

$$\xi = \Phi - \varphi$$
$$\eta = (\Lambda - \lambda)\cos \varphi \qquad (2.103)$$

If we let H be the orthometric height of a point P and N the geoid undulation at the point we have:

$$h = H + N \qquad (2.104)$$

To find the change in these quantities we differentiate them, noting the Φ, Λ and H are independent of the geodetic datum coordinate system. Thus we have:

$$d\xi = -d\varphi$$
$$d\eta = -d\lambda \cos\varphi$$
$$dh = dN \qquad (2.105)$$

At the datum origin we write from (2.105):

$$d\varphi_0 = -d\xi_0$$

$$d\lambda_0 = \frac{-d\eta_0}{\cos\varphi_0}$$

$$dh_0 = dN_0 \tag{2.106}$$

with similar expressions holding for the arbitrary point in the system. Then we may write (2.102) in the form:

$$d\xi = E_1{'}d\xi_0 + E_2{'}d\eta_0 + E_3{'}dN_0 + E_4{'}da + E_5{'}df$$

$$d\eta = F_1{'}d\xi_0 + F_2{'}d\eta_0 + F_3{'}dN_0 + F_4{'}da + F_5{'}df$$

$$dN = G_1{'}d\xi_0 + G_2{'}d\eta_0 + G_3{'}dN_0 + G_4{'}da + G_5{'}df \tag{2.107}$$

Thus we interpret $d\xi_0$, and $d\eta_0$ as changes of the deflections of the vertical at the origin. $d\eta_0$ may be considered as the change from an adopted geoid height to a better or absolute value.

We can also express (2.87), (2.88) and (2.89) in a form where the changes computed are of deflections and undulations. We have:

$$d\xi = E_1{''}dx + E_2{''}dy + E_3{''}dz + E_4{''}da + E_5{''}df$$

$$d\eta = F_1{''}dx + F_2{''}dy + F_3{''}dz + F_4{''}da + F_5{''}df$$

$$dN = G_1{''}dx + G_2{''}dy + G_3{''}dz + G_4{''}da + G_5{''}df \tag{2.108}$$

In equations (2.103) and (2.104) the E, F, and G coefficients (single and double prime) can be found by simple substitution and comparison with the original equations.

2.31.4 Special Cases of Transformation Involving Origin Shifts and Ellipsoid Parameter Changes.

There are certain cases where the general cases derived here reduce to a simpler form. Suppose we consider the case when the rectangular coordinates change due to changes of φ, λ, and h with no change of the reference ellipsoid parameters. Thus, da = df = 0 so that equations (2.80), (2.81) and (2.82) become:

$$dx = -(M+h)\sin\varphi \cos\lambda d\varphi -(N+h)\cos\varphi \sin\lambda d\lambda +\cos\varphi \cos\lambda dh \tag{2.109}$$

$$dy = -(M+h)\sin\varphi \sin\lambda d\varphi +(N+h)\cos\varphi \cos\lambda d\lambda +\cos\varphi \sin\lambda dh \tag{2.110}$$

$$dz = (M+h)\cos\varphi \, d\varphi + \sin \varphi dh \tag{2.111}$$

Under the specification that da = df = 0, equations (2.87), (2.88) and (2.89) become:

$$(M+h)d\varphi = -\sin\varphi \cos\lambda dx - \sin\varphi \sin\lambda dy + \cos\varphi dz \tag{2.112}$$

$$(N+H)\cos\varphi d\lambda = -\sin\lambda dx +\cos\lambda dy \tag{2.113}$$

$$dh = \cos\varphi \cos\lambda dx +\cos\varphi \sin\lambda dy + \sin\varphi dz \tag{2.114}$$

83

Both cases could be represented in matrix form as:

$$
\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} -(M+h)\sin\varphi\cos\lambda & -(N+h)\cos\varphi\sin\lambda & \cos\varphi\cos\lambda \\ -(M+h)\sin\varphi\sin\lambda & (N+h)\cos\varphi\cos\lambda & \cos\varphi\sin\lambda \\ (M+h)\cos\varphi & 0 & \sin\varphi \end{pmatrix} \begin{pmatrix} d\varphi \\ d\lambda \\ dh \end{pmatrix}
$$

(2.115)

and

$$
\begin{pmatrix} (M+h)d\varphi \\ (N+h)\cos\varphi d\lambda \\ dh \end{pmatrix} = \begin{pmatrix} -\sin\varphi\cos\lambda & -\sin\varphi\sin\lambda & \cos\varphi \\ -\sin\lambda & \cos\lambda & 0 \\ \cos\varphi\cos\lambda & \cos\varphi\sin\lambda & \sin\varphi \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}
$$

(2.116)

Equations (2.115) and (2.116) are only valid when no ellipsoid parameters are changed.

A similar procedure could be adopted if no change is made in the coordinates at one point in the geodetic network, but the ellipsoid parameters are changed. From (2.80), (2.81) and (2.82) we find:

$$
dx = \frac{\cos\varphi\cos\lambda}{W} da + \frac{a(1-f)\sin^2\varphi\cos\varphi\cos\lambda}{W^3} df
$$

(2.117)

$$
dy = \frac{\cos\varphi\sin\lambda}{W} da + \frac{a(1-f)\sin^2\varphi\cos\varphi\sin\lambda}{W^3} df
$$

(2.118)

$$
dz = \frac{(1-e)^2\sin\varphi}{W} da + (M\sin^2\varphi - 2N)(1-f)\sin\varphi\, df
$$

(2.119)

The change produced at any other point in the system would be given by substituting these equations into (2.87), (2.88) and (2.89). A similar procedure may be applied directly through equations (2.99), (2.100) and (2.101) where in this special case $d\varphi_0$, $d\lambda_0$ and dh_0 are zero.

Another special case occurs when we define the centers of the two ellipsoids to be coincident. We then set $dx = dy = dz = 0$ in equations (2.87), (2.88) and (2.89) to obtain:

$$
(M+h)d\varphi = \frac{e^2\sin\varphi\cos\varphi}{W} da + \sin\varphi\cos\varphi(2N + e'^2M\sin^2\varphi)(1-f)\, df
$$

(2.120)

$$
(N+h)\cos\varphi d\lambda = 0
$$

(2.121)

$$
dh = -Wda + \frac{a(1-f)}{W}\sin^2\varphi df
$$

(2.122)

Note that the change in latitude due to changes in a are small (they depend on e^2); there is no change in λ (due to symmetry reasons), and the change in height is essentially the negative change in the equatorial radius.

2.31.5 Geodetic Coordinate Changes Due to the Scale Change

We are now interested in the changes of ϕ, λ, and h due to the scale change Δs. We take the rectangular coordinate changes from (2.21) and substitute them into (2.116) where x,y and z are given by (2.74) we find:

$$d\phi_{\Delta s} = -\frac{Ne^2 \sin 2\phi}{2(M+h)} \Delta s \tag{2.123}$$

$$d\lambda_{\Delta s} = 0 \tag{2.124}$$

$$dh_{\Delta s} = (aW+h)\, \Delta s \tag{2.125}$$

We see that the latitude change is zero, in a spherical approximation, indicating the insensitivity of the latitude to scale. The longitude change is zero due to symmetry reasons. The dominant effect of the scale change is on height. If $\Delta s = 10^{-6}$, dh is approximately 6.4 m.

2.31.6 Geodetic Coordinate Changes Due to the Three Rotation Angles

The rectangular coordinate changes introduced by the ω_x, ω_y, ω_z rotations in the Bursa/Wolf model are given by equations (2.22). We can substitute these equations into (2.112) using (2.74) for x,y and z. The results are Soler (1976, p.70):

$$d\phi_R = -\omega_x \left[\frac{aW+h}{M+h}\right] \sin\lambda + \omega_y \left[\frac{aW+h}{M+h}\right] \cos\lambda \tag{2.126}$$

$$d\lambda_R = -\omega_z + \omega_x \left[1 - \frac{Ne^2}{N+h}\right] \tan\phi\cos\lambda + \omega_y \left[1 - \frac{Ne^2}{N+h}\right] \tan\phi\sin\lambda \tag{2.127}$$

$$dh_R = -\omega_x Ne^2 \sin\phi\cos\phi\sin\lambda + \omega_y Ne^2 \sin\phi\cos\phi\cos\lambda \tag{2.128}$$

An alternate form for (2.126) and (2.127) has been given in Bursa (1965, eq. (18)):

$$d\phi_R = -\omega_x(1 + e^2 \cos 2\phi)\sin\lambda + \omega_y(1 + e^2 \cos 2\phi)\cos\lambda \tag{2.129}$$

$$d\lambda_R = -\omega_z + \omega_x(1-e^2)\tan\phi\cos\lambda + \omega_y(1-e^2)\tan\phi\sin\lambda \tag{2.130}$$

We see that the latitude change is primarily a function of ω_x, ω_y and the longitude. The longitude change is a function of the three rotation angles, latitude and longitude. Note that at low latitudes ($\tan\phi$ is small) the longitude change will be primarily $-\omega_z$. The change in height does not depend on ω_z. For a 1″ rotation the maximum effect on height is 21 cm.

2.31.7 The Total Change in Geodetic Coordinates From the Sum of the Individual Components

In our previous discussion we identified 9 change parameters (3 rotations, 3 translations, one scale, 2 ellipsoid). We have now isolated these changes with the total change being the sum of the individual changes. In brief summary we have:

Translations	Equation (2.116)
Ellipsoid Parameters	Equations (2.120, 2.121, 2.122)
Scale	Equations (2.123, 2.124, 2.125)
Rotations	Equations (2.126, 2.127, 2.128)

If one is not interested in the individual component changes the direct approach as discussed in Section (2.3, eq. (2.72)) can be used. A similar form of these equations may be found in Vincenty (1985, p. 191).

2.31.8 Azimuth Changes Due to Rotation Parameters

As the coordinates change in going form an old to a new system, so must the geodetic and astronomic change. We first examine the geodetic azimuth change by expressing the normal section azimuth between two points in the following form Rapp (1984, eq. (4.71)); Vincenty (1985, eq. (4.2)):

$$\tan\alpha = \frac{-\Delta x\sin\lambda + \Delta y\cos\lambda}{-\sin\phi(\Delta x\cos\lambda + \Delta y\sin\lambda) + \Delta z\cos\phi} \tag{2.131}$$

Also of interest here is the Laplace equation (Rapp, 1984, eq. (7.29)). We write:

$$\alpha = A - (\sin\phi - \cos\phi\cos\alpha\tan V)(\Lambda - \lambda) - \sin\alpha\,\tan V(\Phi - \phi) \tag{2.132}$$

where the Φ and Λ designate astronomic quantities and V is the vertical angle from the observing point to the observed point. In evaluating (2.132) we first consider corrections to the astronomic coordinates associated with changes in the astronomic system reflected in ω_x, ω_y and ω_z. In analogy with (2.129) and (2.130) (and with Rapp (1984, eq. (7.1) and (7.2)) we have:

$$d\Phi = -\sin\lambda\omega_x + \cos\lambda\omega_y \tag{2.133}$$

$$d\Lambda = \tan\phi\,(\cos\lambda\,\omega_x + \sin\lambda\omega_y) - \omega_z \tag{2.134}$$

The change in the astronomic azimuth follows from Rapp (ibid, eq. (7.3) or Vincenty (1982, eq. (4.9)):

$$dA = \left(\cos\lambda\omega_x + \sin\lambda\omega_y\right)/\cos\phi \tag{2.135}$$

These changes implicitly reflect a rotation about a pivot point at the center of mass of the system.

The changes in the geodetic Laplace azimuth caused by change in the astronomic system would be (from (2.132)):

$$d\alpha_a = -\sin\alpha\tan V d\Phi - (\sin\phi - \cos\phi\cos\alpha\tan V)\,d\Lambda + dA \tag{2.136}$$

Using (2.133), (2.134), and (2.135) this becomes (Vincenty, 1885 eq. (4.8)):

86

$$d\alpha_a = (\cos\phi\cos\lambda + \tan V(\sin\phi\cos\lambda \, \cos\alpha + \sin\lambda\sin\alpha)) \, \omega_x +$$
$$(\cos\phi\sin\lambda + \tan V(\sin\phi\sin\lambda\cos\alpha - \cos\lambda\sin\alpha))\omega_y + (\sin\phi - \cos\phi \, \cos\alpha\tan V) \, \omega_z \qquad (2.137)$$

We next consider the change in the geodetic azimuth due to change in the geodetic coordinates. Differentiating (2.132) we have:

$$d\alpha_g = \sin\alpha\tan\phi \, d\phi + (\sin\phi - \cos\phi\cos\alpha\tan V) \, d\lambda \qquad (2.138)$$

We now must consider the appropriate procedure for the calculation of $d\phi$ and $d\lambda$. Vincenty (ibid) used the Molodensky approach (see section 2.23) where the rotations take place about the datum origin point. The rectangular coordinates changes would be given by the third term on the right hand side of (2.61). Following Vincenty we write:

$$\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}_r - \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}_t \qquad (2.139)$$

where:

$$\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}_r = U(x,y,z) \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \qquad (2.140)$$

and:

$$\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}_t = U(x_0,y_0,z_0) \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \qquad (2.141)$$

where U is given by equation (2.20) where Δs can be neglected. The "r" subscript in (2.140) indicates the rotation effect while the "t" indicates the translation of the center of the coordinate system.

The value of $d\phi$ and $d\lambda$ in (2.138) is now considered to be made up of 2 components: one due to the rotation and one due to the translation caused by the rotation about the datum origin point. We write in analogy to (2.139):

$$d\phi = d\phi_r - d\phi_t$$

$$d\lambda = d\lambda_r - d\lambda_t \qquad (2.142)$$

Using (2.129) and (2.130) as $d\phi_r$ and $d\lambda_r$, and substituting (2.142) into (2.138) we have (Vincenty, 1985, eq. (4.01)):

$$d\alpha_g = (\sin\phi\tan\phi\cos\lambda - \tan V(\sin\phi\cos\lambda\cos\alpha + \sin\lambda\sin\alpha)\omega_x +$$
$$(\sin\phi\tan\phi\sin\lambda - \tan V(\sin\phi\sin\lambda\cos\alpha - \cos\lambda\sin\alpha))\omega_y + (-\sin\phi +$$
$$\cos\phi\cos\alpha\tan V)\omega_z - \sin\alpha\tan V d\phi_t - (\sin\phi - \cos\phi\cos\alpha\tan V)d\lambda_t \qquad (2.143)$$

87

We now add (2.137) (for $d\alpha_a$) and (2.143) (for $d\alpha_g$) together to obtain the total change in a geodetic azimuth computed through the Laplace azimuth (Vincenty, ibid, eq. (4.11):

$$d\alpha = (\cos\lambda\omega_x + \sin\lambda\omega_y)/\cos\phi - \sin\alpha\tan V d\phi_t - (\sin\phi - \cos\phi\cos\alpha\tan V)d\lambda_t \quad (2.144)$$

The $d\phi_t$ and $d\lambda_t$ terms are found by substituting (2.141) into (2.116) (See Vincenty, ibid, 2.16, 2.17 for $d\phi_t$, $d\lambda_t$).

Special cases of these transformations are discussed in Vincenty (ibid). The general equations can be used in the adjustment of terrestrial networks with space defined positions as will be discussed in a later section.

2.32 A Differential Development Transformation Procedure

The equations of the previous section have been used assuming the projective method has been used in the calculation of our geodetic network. If the development method has been used, there is an argument that an alternate procedure—a development based procedure—should be used. In establishing this method we consider the following change possible:

$d\phi_0$, $d\lambda_0$, coordinate changes at the datum origin point;

$d\alpha_0$, a change in the azimuth of an initial line although it is an idealism to believe an actual network is oriented by a single azimuth;

ds, the effect on ϕ and λ of the lack of a reduction of distances from the geoid to the ellipsoid;

da, df, the usual ellipsoid parameter changes.

We may represent the above changes in the following form:

$$d\phi = F_1(d\phi_0, d\lambda_0, d\alpha_0, ds, da, df)$$

$$d\lambda = F_2(d\phi_0, d\lambda_0, d\alpha_0, ds, da, df) \quad (2.145)$$

If we determine the geodetic azimuth at the origin point such that the Laplace equation is fulfilled we have:

$$\alpha_0 = A_0 - (\Lambda_0 - \lambda_0)\sin\phi_0 \quad (2.146)$$

If we consider that both the astronomic azimuth and geodetic longitude are subject to change at the origin, we have

$$d\alpha_0 = dA + d\lambda_0\sin\phi_0 \quad (2.147)$$

On the other hand, assuming an astronomic azimuth is fixed, the value $d\alpha_0$ is simply $d\lambda_0\sin\phi_0$ so that such change may be combined with the $d\lambda_0$ indicated in equation (2.141).

Equation (2.141) expresses changes previously discussed as differential formulas of the first and second kind (Zakatov, 1962, Chapter II, or Rapp (1984)). Formulas of the first kind define the effect at an arbitrary point of $d\phi_0$, $d\lambda_0$, $d\alpha_0$, and a ds change while

formulas of the second kind consider ellipsoid parameter changes (i.e., da and df). Of the various formulas available, the earliest and most comprehensive are probably due to Helmert (1880, Chapter 12, Section 15). We now give these equations as taken from Bursa (1957) for an arbitrary point in the development computed geodetic network:

$$d\phi_i = p_1 d\phi_0 + p_2 \cos\phi_0 d\lambda_0 + p_3 ds_i + p_4 d\alpha_0 + p_5 \frac{da}{a} + p_6 df \qquad (2.148)$$

$$\cos\phi_i d\lambda_i = q_1 d\phi_0 + q_2 \cos\phi_0 d\lambda_0 + q_3 ds_i + q_4 d\alpha_0 + q_5 \frac{da}{a} + q_6 df \qquad (2.149)$$

The p and q coefficents are given below:

$$p_1 = \frac{M_0}{M_i} \cos\Delta\lambda \qquad\qquad q_1 = \frac{M_0}{N_i} \sin\phi_0 \sin\Delta\lambda$$

$$p_2 = 0 \qquad\qquad q_2 = \sec\phi_0 \cos\phi_i$$

$$p_3 = \frac{-\cos\alpha_{i0}}{M_i} \qquad\qquad q_3 = \frac{-\sin\alpha_{i0}}{N_i}$$

$$p_4 = \frac{R_0}{M_i} \sin\frac{s_i}{R_0} \sin\alpha_{i0} \qquad\qquad q_4 = \frac{-R_0}{N_i} \sin\frac{s_i}{R_0} \cos\alpha_{i0}$$

$$p_5 = \frac{s_i \cos\alpha_{i0}}{M_i} \qquad\qquad q_5 = \frac{s_i \sin\alpha_{i0}}{N_i}$$

$$p_6 = [\Delta\phi(2 - \frac{3}{W_m^2} \sin^2\phi_m) + \qquad\qquad q = -\Delta\lambda \sin^2\phi_0 \cos\phi_0 \frac{\sqrt{1-e^2}}{W_0}$$

$$+ \frac{\Delta\lambda^2}{2} \sin^3\phi_m \cos\phi_m] \frac{1}{\sqrt{1-e^2}}$$

$$\phi_m = 1/2\,(\phi_0 + \phi_i)\,,\ \Delta\lambda = \lambda_i - \lambda_0,\ \text{(positive east)},\ R_0 = \sqrt{M_0 N_0} \qquad (2.150)$$

In these equations i indicates an arbitrary point in the system while the subscript zero refers to the origin at which the changes are originated. α_{i0} indicates the azimuth from point i to the origin.

ds_i represents the desired change in the length of the line between the origin and i caused by the reduction from the geoid to the ellipsoid. Recalling the formula for base line reductions we may estimate ds_i as follows:

$$ds_i = -\frac{s_i}{R} \bar{N} \qquad (2.151)$$

where $\bar{N}$ is the average astro geodetic geoid undulation from the origin to point i with respect to the old ellipsoid, and R is a mean radius of curvature along the line.

It should be clear now that we do not consider in (2.144) and (2.145) quantities considered in the projective system such as orientation changes, scale changes, etc. Such changes do not play a direct role in the development method transformation formulas. In addition (2.148) and (2.149) are generally considered to have a working radius of 600-800km (Zakatov, 1962, p.113).

Equations (2.144) and (2.145) may also be written in terms of deflections of the vertical using equation (2.105) applied at the origin and at the arbitrary point in the system. Then we write:

$$d\xi_i = p'_1 d\xi_0 + p'_2 d\eta_0 + p'_3 ds_i + p'_4 d\alpha_0 + p'_5 \frac{da}{a} + p'_6 df$$

(2.152)

$$d\eta_i = q'_1 d\xi_0 + q'_2 d\eta_0 + q'_3 ds_i + q'_4 d\alpha_0 + q'_5 \frac{da}{a} + q'_6 df$$

(2.153)

For consistency purpose all coefficients, p and q in (2.152) and (2.153) have been primed, even if they do not change in going from (2.148) and (2.149) to (2.152) and (2.153).

If we assume that dA of equation (2.147) is zero, we note that:

$$d\alpha_0 = d\lambda_0 \sin\varphi_0 = -d\eta_0 \tan\varphi_0$$

(2.154)

which may be substituted into (2.152) and (2.153) to yield:

$$d\xi_i = p_1' d\xi_0 + \left(p_2' - p_4' \tan\varphi_0\right) d\eta_0 + p_3' ds_i + p_5' \frac{da}{a} + p_6' df$$

(2.155)

$$d\eta_i = q_1' d\xi_0 + \left(q_2' - q_4' \tan\varphi_0\right) d\eta_0 + q_3' ds_i + q_5' \frac{da}{a} + q_6' df$$

(2.156)

Equations (2.148) and (2.149) or (2.152) and (2.153) may be used to implement system changes in development computed triangulation. Notice that they are not written in terms of dx, dy and dz as the corresponding projective method equations. In addition the development transformation considers an azimuth change and a distance change which is not found, or required, in the projective system transformations, except when network scale and orientation is being considered. The derivation of the development equations is not as concise as the projective transformation. It obviously involves some assumptions not required in the projective system.

2.32.1 Comparison of Certain Projective and Development Change Formulas

We are now in a position to compare changes in φ and λ due to changes at the origin form either the projective method as expressed through equation (2.99), and (2.100) or by the development method as expressed through (2.148) and (2.149). Comparisons can be made analytically and/or numerically to determine the differences between the two methods of computing the differential change. Such a study has been carried out for all change expected by Rais (1969). His results show that for small arc distances away form the origin the results form the projective and development methods (with ds = 0) are very close. However, as the arc distance increases so does the difference between the methods. For example, out to 20° from the origin the differences are on the order of 0."05. This is to be expected as (2.148) and (2.149) have a limited distance over which they are to be considered highly accurate. In addition the inclusion of a ds term not equal to zero in the development formulas causes a greater difference with the projective method results that when ds was set to zero.

2.33 Non-Conventional Transformation of Geodetic Coordinates

The methods discussed in the previous sections assume that there is some relationship that can be simply established between a new and old coordinate system. This relationship is modeled by a selected number of parameters which is usually nine. In reality the actual parameterization is not as simple as implied by our models. The coordinates in a typical geodetic datum, that has been built up over a period of time, do not have a uniform accuracy. Distortions can exist as new (and more precise) data are fitted to an older geodetic frame. That such distortions exist was used as one argument for the development of the North American Datum 83 to replace the North American Datum 1927. Because of the complex nature of these distortions, it is not possible to use the simple models described so far in this report.

An alternate method has been used in converting NAD27 coordinates to WGS84 (DMA, 1987). In this method the differences between the geodetic coordinates of both systems are modelled by a polynomial of sufficient terms to represent the differences over the network, to a given degree of accuracy. In the specific case of the NAD27 to WGS84 the transformation equations took the following form:

$$\Delta\phi'' = A_0 + A_1U + A_2V + A_3U^2 + A_4U^3 + A_5U^2V + A_6U^2V + A_7V^3 + A_8U^3V + A_9UV^3$$

$$+ A_{10}V^4 + A_{11}U^5 + A_{12}U^4V + A_{13}U^2V^3 + A_{14}V^5 + A_{15}V^6 + A_{16}U^7 + A_{17}V^7 + A_{18}U^8 +$$

$$A_{19}V^8 + A_{20}U^9 + A_{21}U^6V^3 + A_{22}U^3V^9 + A_{23}U^4V^9 \tag{2.157}$$

where:

$$
\begin{aligned}
U &= K(\phi - 37) \\
V &= K(\lambda - 265) \\
K &= 0.05235988 \\
\phi &= \text{latitude in degrees} \\
\lambda &= \text{longitude (positive east) in degrees}
\end{aligned}
\tag{2.158}
$$

Similar, but not identical equations were used for $\Delta\lambda$ and ΔH. The number of terms to be retained can be determined by usual significance tests.

An empirical transformation between NAD 1927 and NAD 1983 coordinates (the datums will be discussed in the next chapter) has been developed by Dewhurst (1990). This procedure interpolates datum position differences at known points using a procedure that minimizes "the total curvature associated with surfaces defining the differences between the datums" (Dewhurst, ibid.).

91

REFERENCES

Adam, J., On the Determination of Similarity Coordinate Transformation Parameters Bolletino de Geodesia e Science Affini, XLI, 3, 283-290, 1982.

Badekas, J., Investigations Related to the Establishment of a World Geodetic System, Dept. of Geodetic Science Report No. 124, Ohio State University, Columbus, 1969.

Baldini, A.A., New formulas useful when changing ellipsoidal parameters or orientation, GIMRADA Res. Note 2, 1962.

Bossler, J., Geodesy Solves 900,000 Equations Simultaneously, EOS, (Trans. AGU), 68, p. 569, June 9, 1987.

Bowring, B., Transformation from Spatial to Geographical Coordinates, Survey Review, XXIII, 191, 323-327, 1976.

Bursa, M., Ausdrücke Für Differentialquotienten Der Gradmessungsgleichungen Als Funktionen Ausschliesslich Von Ellipsoidischen Koordinaten, Studia geoph. et geod., 1, 76-87, 1957 (in Russian).

Bursa, M., Fundamentals of the Theory of Geometric Satellite Geodesy, Travaux De L'Institut Geophysique De L'Academic Tchecoslovaque Des Sciences, No. 241, 1966.

Bursa, M., On the Determination of the Direction of the Minor Axis of the Reference Ellipsoid and the Plane of the Initial Geodetic Meridian from Observations of Artificial Earth Satellites, Studia geoph. et geod, Vol. 9, no. 1, p. 14-22, 1965 (in Russian).

Bursa, M., The Theory of the Determination of the Nonparallelism of the Minor Axis of the Earth, and Initial Astronomical and Geodetic Meridians from Observations of Artificial Earth Satellites, translation from Studia Geophysica et Geodaetica, Vol. 6, No. 2, 1962.

Bursa, Mo, On the Possibility of Determining the Rotating Elements of Geodetic Reference Systems on the Basis of Satellite Observation, Studia. geoph. et geod., 11 (1967).

Dewhurst, W., The Application of Minimum—Curvature—Derived Surfaces in the Transformation of Positional Data from the North American Datum of 1983, NOAA Technical Memorandum NOS NGS-50, National Geodetic Information Center, Rockville, MD 20852, 1990.

DMA, Department of Defense World Geodetic System 1984, It's Definition and Relationship with Local Geodetic Systems, DMA TR 8350.2, Sept. 1987.

Dracup, J., The United States Horizontal Control Network 1816-1976, in Proc. ACSM 36th Annual Meeting, Washington, Feb. 1976.

Harvey, B., Transformation of 3D Coordinates, The Australian Surveyor, Vol. 33, No. 2, p. 105-125, 1986.

Heiskanen, W., and H. Moritz, Physical Geodesy, W.A. Freeman, San Francisco, 1967.

Helmert, F.R., Mathematischen Und Physikalischen Theorien Der Hoheren Geodasie, 1880, (Minerva G.M.B.H. Frankfurt/Main, Stuttgart, 1962).

Herring, J., Changes in the Geodetic Coordinates Due to a Change in the Reference Ellipsoid, M.S. thesis, The Ohio State University, Columbus, 1965.

Hotine, M., Mathematical Geodesy, ESSA Monograph 2, 1969, COM-74-11127.

Kumar, M., Coordinate Transformation by Minimizing Correlations between Parameters, Dept. of Geodetic Science Report No. 184, Ohio State University, July, 1972.

Leick and van Gelder, On Similarity Transformations and Geodetic Network Distortions Based on Doppler Satellite Observations, Dept. of Geodetic Science Report No. 235, The Ohio State University, Columbus, 1975.

Malys, S., Dispersion and Correlations among Transformation Parameters, Relating Two Satellite Reference Frames, Report No. 329, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, 95 p., 1988.

Molodensky, M.S., et al., "Methods for Study of the External Gravitational Field and Figure of the Earth", Translation for National Science Foundation and Department of Commerce by Israel Program for Scientific Translations, 1962, (OTS-61-31207).

Rais, J., A Comparison of the Projective and the Development Methods for Computing the Best Fitting Reference Ellipsoid, M.Sc. thesis , The Ohio State University, Columbus, 1969.

Rapp, R.H., Geometric Geodesy, Volume I (Basic Principles), Dept. of Geodetic Science, The Ohio State University, Columbus, 1984.

Simmons, L., How Accurate is First-Order Triangulation, The Journal, (U.S.) Coast and a Geodetic Survey, No. 3, p. 52-56, April 1980.

Soler, T., On Differential Transformation Between Cartesian and Curvilinear (Geodetic) Coordinates, Dept. of Geodetic Science Report No. 236, The Ohio State University, Columbus, 1976.

Somogyi, J., Robust Estimation of the Parameters for a Three Dimensional Transformation, in Festschrift Rudolf Sigh, Deutsche Geodatische Kommission, Reihe B, No. 287, Munchen, 1988.

Thomson, D., and E. Krakiwsky, Concepts of the Combination of Geodetic Networks, in Proc Int. Geodetic Symp., "Satellite Doppler Positioning", Physical Sciences Lab., New Mexico State University, 1976.

Vaníček, P. and G. Carrera, Reference Ellipsoid Misalignment, Deflection Components, and Geodetic Azimuth, The Canadian Surveyor, Vol. 39, No. 2, 1985.

Veis, G., Geodetic Uses of Artificial Satellites, Smithsonian Contributions to Astrophysics, Vol. 3, No. 2, Smithsonian Institution, Washington, D.C., 1960.

Vening Meinesz, F.A. (a) New Formulas for systems of deflections of the plumb-line and Laplace's theorem; (b) Changes of deflections of the plumb-line brought about by a change of reference ellipsoid, Bulletin Geodesique, no. 15, 1950.

Vincenty, T., Methods of adjusting space sytems data and terrestrial measurements, Bulletin Geodesique, 56, 231-241, 1982.

Vincenty, T., On the Meaning of Geodetic Orientation, Bulletin Geodesique, 59, 2, 189-199, 1985.

Vincenty, T., Transformation of Geodetic Data between Reference Ellipsoids, Journal of Geophysical Research, Vol. 71, No. 10, May 15, 1966.

Vincenty, T., Zur räumlich-ellipsoidischen Koordinaten-Transformation, Zeitschrift für Vermessugswesen, Heft 11, 519-521, November 1980.

Wells, D., and P. Vaniček, Alignment of Geodetic And Satellite Coordinate Systems to the Average Terrestrial System, Bulletin Geodesique, No. 117, 241-257, 1975.

Wilcox, L., The Development and the Projective Methods of Geodesy and Their Effect on Large Triangulation and Trilateration Networks, Bulletin Geodesique, No. 68, June 1963, p. 127-138.

Wolf, H., Geometric Connection and Re-Orientation of Three-Dimensional Triangulation Nets, Bulletin Geodesique, no. 68, June 1963, p. 165-169.

Zakatov, P.S., A Course in Higher Geodesy, translated from Russian by Israel Program for Scientific Translation, for the National Science Foundation, OTS 61-31212, 1962.

3. The Determination of Geodetic Datums and Ellipsoid Parameters

3.1 Introduction

The discussion in Section 2 has assumed that we have been given geodetic information on a defined geodetic datum. The horizontal coordinates (usually ϕ and λ) were combined with vertical coordinate information (orthometric height and astro-geodetic undulation) to defined a three dimensional position with respect to the ellipsoid associated with the datum. Given the ellipsoid parameters we can then calculate the three rectangular coordinates.

In Rapp (1984, Chapter 9) we have discussed the adjustment of a triangulation/trilateration network on the ellipsoid through the development of observation equations for direction measurements, distances, astronomic azimuths, and Laplace azimuths. In these discussions we assumed that a geodetic datum was defined so that there would be no rank defect in the normal equation matrix of the adjustment process.

In this section we examine various definitions of geodetic datums based on our experience with the transformation process discussed in Chapter 2. In addition we will examine various ways in which the parameters of the reference ellipsoid can be estimated by classical (triangulation/trilateration) and non-classical (e.g., sea surface heights from satellite altimeter data) data types.

3.2 Horizontal Geodetic Datums—Theory

Discussions of the manner in which horizontal datums are defined have often been carried out in the literature (e.g., Hotine (1969), Jones (1973), Vaníček and Wells (1974), Mueller (1974), Moritz (1978), Bomford (1980, Section 2), Vaníček and Carrera (1985), Vincenty (1985) etc.). In the discussions one needs to distinguish between the ideal situation and a situation that may have existed a number of years ago when the horizontal geodetic datum was being established.

We might start from an ideal definition of the coordinate frame, its center, and an ellipsoid to be associated with this system. We will argue here (but with counter arguments to come later) that the ideal system should be one whose center is at the center of mass of the earth. The alignment of the axes of this system should coincide with an internationally adopted Conventional Terrestrial System (CTS). In practice there may be several candidate CTS. International agreement does exist on the establishment of the ideal CTS from 1988 onwards (Mueller, 1985, 1988). Before the establishment of the new system various estimates of such a system have been made. A widely used one is the Bureau International De L'Heure, (BIH) Terrestrial System (BTS). Various BTS systems have been defined. For example, the definition and estimation of BTS (1987) is described by Boucher and Altamimi (1988). In the future the ideal frame will be defined by the International Earth Rotation Service (Mueller, ibid).

The ideal reference system has only become a near reality due to the rapid progress made in the development of space related observation systems. In the development of geodetic datums in the early 20th century access to the ideal system was not available. In practice astronomic observations were used to obtain access to some reference system and an ellipsoid, derived using existing geodetic data, was used as the reference surface. Before we consider some specific details we need to consider a very simple definition of a horizontal datum.

A simple definition of a horizontal datum involves the definition of the latitude (ϕ_0) and longitude (λ_0) of the datum origin point; the azimuth from the origin point to an

arbitrary point in the datum; the equatorial radius (a) and the flattening (f). These five parameters (ϕ_0, λ_0, α_0, a, f) constitute a minimal definition of a horizontal datum. It is a minimal definition as nothing is said about the alignment of the axes of the geodetic system or about the location of the ellipsoid with respect to the origin of the datum (or its location with respect to the center of mass of the earth.)

This simple definition does not take into account the realities of the observational procedures used in the development of geodetic networks in the first half of this century. For example, we know that in practice Laplace (geodetic) azimuths are derived for various lines in a network. Such azimuths provide the orientation to the network and thus the azimuth at the datum origin is not, in reality, needed.

In order to be more complete we must now extend our simple datum definition so that it's realization in terms of a reference system can be obtained. We start by specifying that the minor axis of the reference ellipsoid should be parallel to the Z axis of a specified reference system (such as the CTS). We also wish to have the initial meridian of the datum system to be parallel to the XZ plane defined by some recognized reference system (again such as the CTS). In order to implement such requirements we must consider the measurements that are possible in order that we can gain access to our ideal coordinate axes. We can measure astronomic latitude, Φ_0, astronomic longitude Λ_0, and astronomic azimuth A_0. In addition, for the most general case we may observe a zenith distance z_0' from the origin point to another point in the system. We can also have access to coordinate systems implied by satellite positioning and VLBI measurements.

Now assume that we have the deflections of the vertical ξ_0 and η_0 in the meridian and prime vertical respectively. (ξ_0 and η_0 may be initially set to zero, or computed gravimetrically, or estimated from adjustment techniques to be discussed later). We can then connect the astronomic and geodetic latitudes and longitudes using the following equations (Heiskanen and Moritz, 1967, equation (5-17), Rapp (1984, Section 7.2)

$$\varphi_0 = \Phi_0 - \xi_0 ; \qquad\qquad \xi_0 = \Phi_0 - \varphi_0 \qquad\qquad (3.1)$$

$$\lambda_0 = \Lambda_0 - \eta_0 \sec \varphi_0 \qquad\qquad \eta_0 = \left(\Lambda_0 - \lambda_0 \right) c\ddot{}s\, \varphi_0$$

These equations are valid if the axes of the astronomic and geodetic coordinate systems are parallel and higher order terms are negligible. Higher order terms may be found in Pick et al. (1973, Chapter XV, Section 4). For the case when the axes of the two systems are rotated, equations corresponding to (1) are given in (ibid, Chapter XV, Section 6), in Grafarend and Richter (1977) and in Vincenty (1985). Next we relate the astronomic and geodetic azimuths through the extended Laplace equation (Heiskanen and Moritz, 1967, equation (5-13), Rapp (ibid, eq. 7.25)):

$$\alpha = A - \eta \tan \varphi - \left(\xi \sin \alpha - \eta \cos \alpha \right) \cot z \qquad\qquad (3.2)$$

where z is the zenith distance. Substitution from (3.1) we have:

$$\alpha = A - \left(\Lambda - \lambda \right) \sin \varphi - \left[\left(\Phi - \varphi \right) \sin \alpha - \left(\Lambda - \lambda \right) \cos \varphi \cos \alpha \right] \cot z \qquad\qquad (3.3)$$

An extended form of Laplace's equations when the axes are not parallel is given in Pick et al. (1973, Chapter XV, Section 6).

An equation relating the astronomic and geodetic zenith distance is given in Hotine (1969, equation 19.29) or Rapp (ibid, eq. 7.32):

$$z = z' + \cos \varphi \sin \alpha \left(\Lambda - \lambda \right) + \cos \alpha \left(\Phi - \varphi \right) \tag{3.4}$$

or

$$z = z' + \eta \sin \alpha + \xi \cos \alpha \tag{3.5}$$

where z' is the astronomic zenith distance. Although equations (3.2), (3.3) and (3.4) hold at any point in our network we are specifically interested in them at the origin point.

Using equation (3.1) at the datum origin and other points in the geodetic system will impose two orientation conditions because of the assumptions made in their derivation. The third condition (now involving a rotation about the ellipsoid normal) is introduced by using the Laplace azimuth equation, (3.2) or (3.3). Hotine (1969) argues that the vertical angle equation (3.5) must be fulfilled through separate observations but Vaniček and Wells (1974) point out that (3.5) will be fulfilled if ξ and η are computed through (3.1). Since the conditions are imposed through astronomic observations a parallelism attempt is not exact being subject to observational errors. Using the needed equations at many points in a geodetic network, and not just at a datum origin point, will reduce the effect of observational errors in our alignment attempt.

At this point we have seen how we can relate the axes of our ellipsoid to a measurable system. Specifically the rotation axis of the ellipsoid will be parallel to the "z" axis of the astronomic system. The initial geodetic meridian will be parallel to the initial meridian of the "astronomic" system. We now have to locate the center of the ellipsoid with respect to a point located (typically) on the surface of the earth. If we consider the origin point as a monument in the field we require the distance between the ellipsoid of the datum and the monument measured along the normal to the ellipsoid through the origin point. This could be specified as $h_0 = H_0 + N_0$ where H_0 is the orthometric height of the origin and N_0 is the separation between the ellipsoid and geoid at the origin. If we consider the origin to be defined as a point on the geoid then we need only specify N_0 to determine the geoid ellipsoid separation at the origin. In the case of the origin point on the geoid it is of course necessary to reduce all astronomic observations from the height at which they were made down to corresponding values on the geoid.

Now let us review the information in the past few paragraphs. We first list the quantities needed at an origin point to determine a datum in the classical sense. These are: Φ, Λ, A, H, ξ, η, N, a, f and two equations (3.1 and 3.3) relating the astronomic measurements to the geodetic coordinates. In choosing these parameters various approximations can be made. For example, by specifying that the ellipsoid and the geoid coincide at the origin point (on the geoid) we would have $N_0 = 0$. We could also make the deflections of the vertical zero so that the ellipsoid would be tangent (if $N_0 = 0$) to the geoid at the origin. Clearly if this is done the center of the ellipsoid and the center of mass of the earth could be far apart, and the separation between the ellipsoid and the geoid could rapidly increase as we get away from the datum origin. This is demonstrated in Figure 3.1.
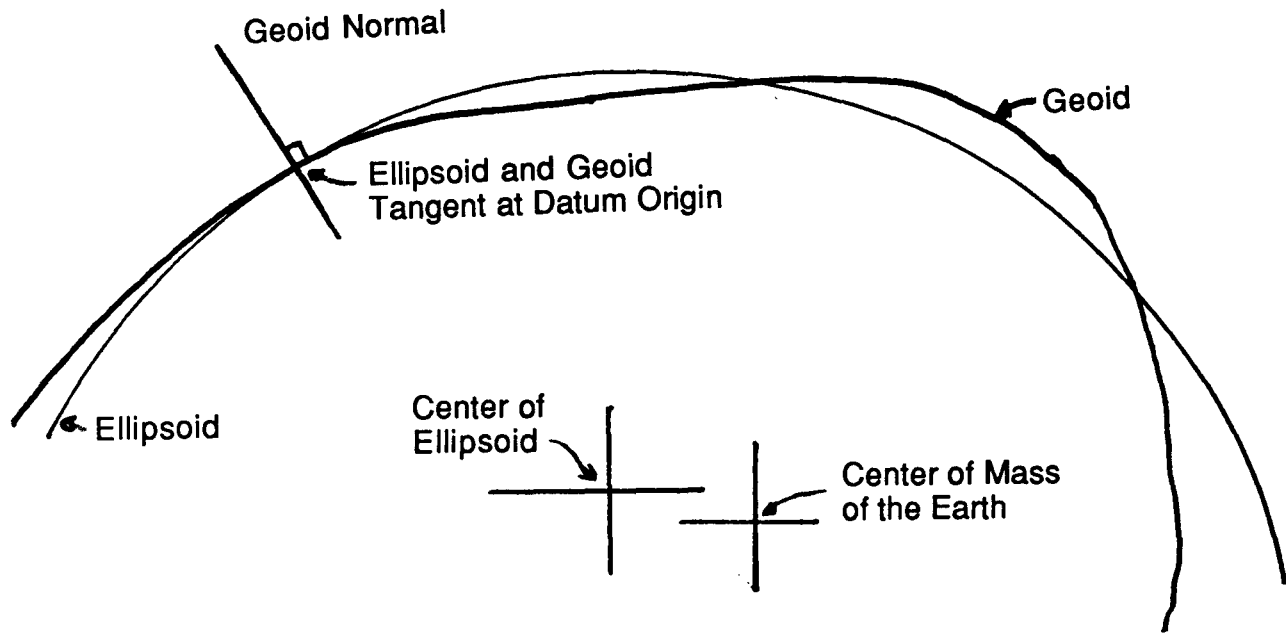
Figure 3.1

A Datum with the Ellipsoid Tangent at the Datum Origin Point

Now a somewhat different view of a horizontal geodetic datum can be taken if we regard the datum as a system defined by an origin near the center of mass of the earth with rectangular coordinates axes aligned parallel to the Conventional Terrestrial System (or other suitable reference system). In this case we might specify an ideal datum as one such that the shifts (Δx, Δy, Δz) between the center of mass of the earth and the datum (ellipsoid) center be zero, and that the three rotation angles ω_x, ω_y and ω_z also be zero. If these quantities are not zero then we would want to define our datum through specified values of Δx, Δy, Δz and ω_x, ω_y and ω_z, which constitute 6 parameters needed to locate the center of the ellipsoid of our datum and to orient the axes of the ellipsoid (Pick et al. 1973).

In practice it clearly is unrealistic to base the determination of a continental network on measurements made at a single point (the origin). Consequently the procedure used for the determination of a horizontal datum is one of adopting preliminary origin coordinates and pertinent parameters sufficient to compute a geodetic network. This preliminary network is then examined to determine better origin parameters and ellipsoid parameters determined such that certain quantities may be minimized in a least squares adjustment. Such procedures will be discussed in a later section. Clearly if the initial data is not the 'best', we would expect to see errors in our geodetic network as it is expanded from the origin.

3.3 Datum Definition and Horizontal Networks Through the Use of Positions Derived from Space Observations

3.3.1 Introduction

Space techniques enable us to determine rectangular coordinates (x, y, z) or coordinate differences (Δx, Δy, Δz) in some defined reference system. Given the x, y, z coordinates and a set of ellipsoid parameters we can determine the latitude, longitude, and height of points that are connected to our usual horizontal network. These coordinates refer to a datum that is implied by the space system. Specifically we have axes orientation, scale, and the origin (center) implied by the specific system we are clearly with. These coordinates could, in a simple sense, be used as fixed points (or more correctly as information with an error variance - covariance matrix) that can be incorporated into a horizontal network. We can thus let the space system provided the ultimate datum origin and no specific datum origin point, in the classical sense is involved.

The actual procedures to be used are not so simple. In practice we have a number of different procedures that can be used for incorporating space positions in our horizontal networks. In one procedure the space positions are first transformed into the datum system using some or all of the transformation parameters treated in Chapter 2. The transformed coordinates are then used in a two dimensional adjustment to combine the space and terrestrial data. Such procedures have been used in the U.S. (Dracup, 1975); in Great Britain (Ashkenazi, Crane, and Williams, 1981, Ashkenazi and Crane, 1985), in Australia (Allman, 1981 Allman and Veenstra, 1984); in some aspects of the readjustment of the European Datum, and most probably in other areas. Various assumptions are made with these procedures depending on how the transformation is performed and what reference system (scale and orientation) is implied by the terrestrial observations.

A somewhat different point of view can be taken that eliminates any reference to the original geodetic datum. In this case various space positioning systems are used to define the orientation, scale, and origin of the final system. This data is merged with the terrestrial observations with due regard to the possible inconsistencies of the reference system (orientation and scale) of the terrestrial observations. This general procedure has been used in the definition of NAD83 (Bossler, 1987). Vincenty (1982) and Steeves (1984) describes the various forms of observation equations that may be used on this type of data merging.

It is important to note that the merger of space and terrestrial data is a merger of data that yields different information. With space observations we deal almost exclusively with three-dimensional observations. In our horizontal networks, we are dealing with two dimensions. Various techniques have been described (e.g., Wolf, 1980, 1982a) to carry these procedures out.

In the following two sections we will examine one specific merger procedure for each type of combination procedure.

3.3.2 Space Positions to Horizontal Datum System

The method to be discussed here was proposed by Wolf (1981, 1982b) and has been used in the new adjustments of the European triangulation (Ehrnsperger, 1985, Kelm, 1987).

As a first step assume that an adjustment has been made of the classical network type where the usual geodetic datum has been defined. This adjustment is done with the

99

projective method where the terrestrial observations are reduced to a defined ellipsoid. This adjustment can be carried out with different scale factors for different distant measuring equipment, or for distance measurements from different geographic areas or countries. Orientation unknows for the azimuth (e.g., one per region or country) can also be introduced to reflect different observational procedures. One such adjustment was ED79 (Hornik and Reinhart, 1980). Another adjustment holding the ED50 coordinates of station D 7835 München fixed was carried out in 1985 (Ehrnsperger, 1985). The reference ellipsoid was retained as the International Ellipsoid.

Now consider a set of stations whose rectangular coordinates are defined in a space system (e.g., Doppler or laser positioning). Let the geodetic positions, in the local datum, be ϕ_g, λ_g, h_g where hg is the sum of the orthometric height plus the astrogeodetic geoid undulation. From this data calculate the rectangular coordinates in the local datum. These coordinates are compared to the coordinates of the points in the satellite system to estimate one scale and three translation parameters. An adjustment procedure has been described starting with equation (2.33). However we now set the rotation angles to zero and establish the transformation parameters going from the space system to the local system. In doing this one must decide if the space systems needs any scale or orientation corrections of its own. (Such corrections were needed for the Doppler coordinate systems used prior to WGS84). Using the notation of Wolf (1982a) we write:

$$\underline{d} = \underline{B}\,\delta\underline{r}_0^0 + \underline{r}\Delta s + \left(\underline{r}_g - \underline{r}_s\right)$$

(3.6)

where:

r_g is the vector of x, y, z coordinates in the local datum;
r_s is the vector of x, y, z coordinates in the satellite system;

Δs is the scale parameter;

δr_0 is the shift vector $(\Delta x_0^0, \Delta y_0^0, \Delta z_0^0)^T$
d is the vector of the residuals;
B^T is [I, I, ... I], I = the identity matrix;

One forms the normal equations and solves for δr_0^0 and Δs.

We next turn to the "fusion" of the terrestrial data (actually reduced normal equations from the local datum adjustment) with the satellite system normal equations. We start by the comparison of the datum positions derived from the terrestrial network with the corresponding positions from the satellite system. In doing this comparison we consider as known Δs and δr_0 determined previously, and we introduce three rotation angles that represent bias between the satellite system and the terrestrial system. At a given station we let the corrections (in a rectangular coordinate system) to the terrestrial network values be $\delta\underline{r}_g$. The final coordinates of the station must be the same in the terrestrial side and from the satellite system. The observation equation takes the form (Wolf, 1982e, eq. 9 and 10):

$$\underline{e}_s = \underline{r}_g^0 + \delta\underline{r}_g - \left(\underline{r}_s - \underline{A}\delta\underline{\varepsilon} - B\delta\underline{r}_0^0 - \underline{r}\Delta s\right)$$

(3.7)

$$= \delta\underline{r}_g + \underline{r}\Delta s + \underline{A}\delta\underline{\varepsilon} + \underline{B}\delta\underline{r}_0^0 + \left(\underline{r}_g^0 - \underline{r}_s\right)$$

(3.8)

Note that $\underline{r}_s$ is the position vector in the satellite system and Δs and δr_0^0 are known from the previous adjustment. We then introduce the quasi-parameter vector $\delta\underline{r}_s$ as follows:

$$\underline{\delta r}_s = \delta\underline{r}_g + \underline{r}\Delta s + \underline{A}\delta\underline{\varepsilon} + B\delta\underline{r}_0^0$$

(3.9)

where $\delta\varepsilon$ are the ω_x, ω_y, ω_z values as used in Chapter 2 and as given by (2.20). If we introduce the rotations about the rectangular axes at an average location (P_0) in the system the coordinate values in the U matrix of (2.20) are replaced by coordinate differences: $\underline{X_i}$ - $\underline{X_0}$. The normal equations from the satellite and common terrestrial stations stations are now written as:

$$\underline{N}_s \delta \underline{r}_s + \underline{C}_s = 0$$

(3.10)

where $\underline{C}_s$ are the constant terms and $\underline{N}_s$ represents the reduced normals after any nuissance parameters are eliminated.

We now introduce corrections to the geodetic coordinates of the terrestrial points: ($\delta\phi$, $\delta\lambda$, δh). The corresponding linear corrections will be designated δt_g where for a given point:

$$\delta \underline{t}_{gi} = \begin{bmatrix} \delta\phi M \\ \delta\lambda N & \cos\phi \\ \delta h \end{bmatrix}_i$$

(3.11)

These values can be related to δr_g through (2.115). Formally we write:

$$\delta \underline{r}_g = \underline{C}\,\underline{K}^{-1} \delta \underline{t}_g$$

(3.12)

where the elements of C and K are clear from (2.115).

We now introduce rotations in a local coordinate system about axes passing through P_0 (ϕ_0, λ_0). Such rotations were used in the Veis transformation method. Wolf (1982b, eq. II.5) represents this form as follows:

$$\delta\underline{\beta}_0 = \left[\delta\overline{x}_0 / R_0, \delta\overline{y}_0 / R_0, \delta\overline{A}_0 \right]$$

(3.13)

where R_0 is a mean earth radius, and the three values in (3.13) represent rotations analogous to ξ, η, α used in the Veis procedure. Wolf designates $\delta\overline{x}_0$ and $\delta\overline{y}_0$ as the horizontal shift components at P_0 and the azimuthal rotation angle at P_0. This was done to reduce the correlations of the estimated rotation parameters. Values of $\delta\underline{\beta}_0$ are related to $\delta\varepsilon$ using (2.56). Specifically we have:

$$\delta\underline{\varepsilon} = \underline{S}\,\delta\underline{\beta}_0$$

(3.14)

where:

$$\underline{S} = \begin{bmatrix} -\sin\lambda_0 & \sin\phi_0\cos\lambda_0 & \cos\phi_0\cos\lambda_0 \\ \cos\lambda_0 & \sin\phi_0\sin\lambda_0 & \cos\phi_0\sin\lambda_0 \\ 0 & -\cos\phi_0 & \sin\phi_0 \end{bmatrix}$$

(3.15)

We now substitue (3.9), (3.12), and (3.14) into (3.10) to obtain:

$$\underline{N}_s \left(\underline{C}\,\underline{K}^{-1} \delta \underline{t}_g + A\underline{\overline{S}}\,\delta\underline{\beta}_0 \right) + \underline{\overline{C}}_s = 0$$

(3.16)

where:

$$\bar{\bar{C}}_s = \bar{C}_s + \bar{N}_s \left(\mathbf{I} \Delta s + B \delta r_0^0 \right)$$

(3.17)

This system contains three unknowns per point plus the 3 rotation unknowns per satellite system. Since we are dealing with a two dimensional terrestrial network we eliminate all height parameters from δt_g. The resultant normal equations take the form:

$$\begin{bmatrix} M_{tt} & M_{t\beta} \\ M_{\beta t} & M_{\beta\beta} \end{bmatrix} \begin{bmatrix} \delta t'_g \\ \delta\beta_0 \end{bmatrix} + \begin{bmatrix} w_t^s \\ w_\beta^s \end{bmatrix} = 0$$

(3.18)

where the $\delta t'$ contains only latitude and longitude corrections.

We next consider the normal equations of the terrestrial data where a reduced set of normal equations have been formed containing only the corrections to the positions at the satellite stations. We write these equations as follows:

$$\underline{N}'_g \delta \underline{t}'_g + \underline{C}'_g = 0$$

(3.19)

We now add the two sets (3.18 and 3.19) of normal equations to find (Wolf, ibid, II, 10):

$$\begin{bmatrix} \underline{M}_{tt} + \underline{N}'_g & \underline{M}_{t\beta} \\ \underline{M}_{\beta t} & \underline{M}_{\beta\beta} \end{bmatrix} \begin{bmatrix} \delta \underline{t}'_g \\ \delta \underline{\beta}_0 \end{bmatrix} + \begin{bmatrix} \underline{w}_t^s + \underline{C}'_g \\ \underline{w}_\beta^s \end{bmatrix} = 0$$

(3.20)

This system is solved for $\delta \underline{t}'_g$ and $\delta\beta_0$. These values are then used in the full set of normal equations for the terrestrial system. The solution then yields the adjusted horizontal coordinates for the stations at which no satellite positions are available.

This method of adjustment effectively uses a seven parameter transformation model between the terrestrial and satellite system. However it does it in two steps. The first step calculates four parameters while the second calculates three. We must realize that the results are dependent on height information in the terrestrial system. But the sensitivity to the heights (or actually geoid undulations) should be low. Preliminary results using this combination procedure are given in Ehrnsperger (1985).

3.3.3 Horizontal Positions to Space Positions

We next consider the case in which terrestrial observations (directions, distances, astronomic azimuths, etc.) are to be placed in a frame to be defined by a particular space system, or a combination of several space systems. Such a procedure would be followed if we wanted to define a geodetic system to have the attributes (e.g., a center of mass origin) of the space system(s). We follow Vincenty (1982) in this section.

Let x, y, z be the coordinates in the ideal system and X, Y, Z be the "observed" coordinates defined in a space system. The connection between these two systems is represented by eq. (2.33). We now postulate residuals (v_x, v_y, v_z) on the "observed" coordinates and corrections dx, dy, dz to the assumed approximate coordinates (x_0, y_0, z_0). We then write:

$$\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} + U \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} + \begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix} + \Delta s \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

(3.21)

Re-arranging this equation we have

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} - \begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix} - U \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} - \Delta s \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} x_0\text{-}X \\ y_0\text{-}Y \\ z_0\text{-}Z \end{bmatrix}$$

(3.22)

This equation corresponds to (2.3) in Vincenty (ibid).

We now introduce a local coordinate system (u, v, h) at the point. The change in u and v will depend on $d\phi$ and $d\lambda$:

$$du = (M + h) \, d\phi$$

(3.23)

$$dv = (N + h) \cos \phi \, d\lambda$$

We can write, using (2.115):

$$\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} -\sin\phi \cos\lambda & -\sin\lambda & \cos\phi \cos\lambda \\ -\sin\phi \sin\lambda & \cos\lambda & \cos\phi \sin\lambda \\ \cos\phi & 0 & \sin\phi \end{bmatrix} \begin{bmatrix} du \\ dv \\ dh \end{bmatrix}$$

(3.24)

We introduce the R matrix in the following form:

$$\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = R^T \begin{bmatrix} du \\ dv \\ dh \end{bmatrix}$$

(3.25)

Where R^T follows directly from (3.24). We now substitute (3.25) into (3.22):

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = R^T \begin{bmatrix} du \\ dv \\ dh \end{bmatrix} - \begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix} - U \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} - \Delta s \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} x_0\text{-}X \\ y_0\text{-}Y \\ z_0\text{-}Z \end{bmatrix}$$

(3.26)

Multiplying from the left by R we have:

$$R \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \begin{bmatrix} v_u \\ v_v \\ v_h \end{bmatrix} = \begin{bmatrix} du \\ dv \\ dh \end{bmatrix} - R \begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix} - RU \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} - \Delta s R \begin{bmatrix} x \\ y \\ z \end{bmatrix} + R \begin{bmatrix} x_0\text{-}X \\ y_0\text{-}Y \\ z_0\text{-}Z \end{bmatrix}$$

(3.27)

where the discrepency vector is expressed in the local geodetic horizon system.

The product of R $[x, y, z]^T$ is effectively given by (2.123), (2.124), and (2.125). Vincenty (ibid, eq. 2.14) writes this product in the following form:

$$R \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -e^2 r \sin \phi \cos \phi \\ 0 \\ r \end{bmatrix} \tag{3.28}$$

where r is the geocentric radius to the point. The RU product is given by Vincenty (ibid, eq. 2.13):

$$RU = \begin{bmatrix} -yr/p & xr/p & 0 \\ xz/p & yz/p & -p \\ -e^2 yz/r & e^2 xz/r & 0 \end{bmatrix} \tag{3.29}$$

where

$$p = (x^2 + y^2)^{1/2} \tag{3.30}$$

Equation (3.29) shows, (as did 2.126 and 2.128) that ω_z does not affect lattitude or heights. From examination of eq. (3.28) we see (as also seen from (2.123, 2.124, 2.125)) that the scale change (Δs) does not affect longitude and has only a very minor affect on latitude. These equations point out that in a usual 2 dimensional terrestrial network it is not possible to determine Δs on the basis of terrestrial distances.

The du, dv corrections can be expressed in $d\phi$ and $d\lambda$ by using (3.23). We define a diagonal matrix D:

$$D = \begin{bmatrix} (M+h) & & \\ & (N+h)\cos\phi & \\ & & 1 \end{bmatrix} \tag{3.31}$$

Then (3.27) can be written as (with $T=D^{-1}R$):

$$\begin{bmatrix} v_\phi \\ v_\lambda \\ v_h \end{bmatrix} = \begin{bmatrix} d\phi \\ d\lambda \\ dh \end{bmatrix} - T \begin{bmatrix} dX \\ dY \\ dZ \end{bmatrix} - TU \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} - \Delta sT \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} \phi_0 - \bar{\phi} \\ \lambda_0 - \bar{\lambda} \\ h_0 - \bar{h} \end{bmatrix} \tag{3.32}$$

where ϕ_0, λ_0, h_0 are the approximate coordinates and $\bar{\phi}, \bar{\lambda}, \bar{h}$ are the "observed" coordinates.

Equation (3.32) is the main form for introducing space positions into a geodetic network. Several different types of systems can be used with various aspects of each system selected to define the final reference system. For example, a laser system could be used to define the origin of the coordinate system, while VLBI results may define the orientation. In each case one or more of the seven transformation parameters are set to zero

104

for that part chosen to define the system. For example, if a laser system is used to define the center of the system, the values of dX, dY, dZ would be set to zero. Using the data from the space positions, the observation equations (3.32) can be used to form normal equations where unknowns are the values of the transformation selected for estimation for each system.

We now consider the terrestrial network, the observation equations, and finally the normal equations. The observation equations can take the forms described in Rapp (1984, Section 9) or using the form where the ellipsoidal heights of points are considered known (Vincenty, 1980) and Section 4 of these notes.

The astronomic azimuth will depend on the coordinate system to be used. If we let the space system define the alignment of the system we must introduce terms into the azimuth observation equation that takes into account that the astronomic system and the chosen space system may not have (in general) parallel axes. Let these rotation angles be defined by $\omega'_x, \omega'_y, \omega'_z$ where the prime is used to specify the ω rotations that relate a space system to the ideal system. We can write the form of the observation equation as:

$$A + dA + v = A_0 + \Delta A \tag{3.33}$$

where A is the observed azimuth, dA is the correction for the changed orientation of the axes, A_0 is the approximate azimuth based on the approximate station coordinates, and ΔA is the differential effect of the coordinate correction. The value of ΔA must be computed recognizing A is an astronomic azimuth. Techniques for doing this are discussed in Chapter 4. Equation (2.137) can be used for dA. Then (3.33) can be written as (Vincenty, 1982, eq. (4.4)):

$$v = a_1 d\phi_1 + a_2 d\lambda_1 + a_3 d\phi_2 + a_4 d\lambda_2$$

$$- \left(\cos\phi \, \cos\lambda + \tan V \left(\sin\phi \, \cos\lambda \, \cos\alpha + \sin\lambda \, \sin\alpha \right) \right) \omega'_x$$

$$- \left(\cos\phi \, \sin\lambda + \tan V \left(\sin\phi \, \sin\lambda \, \cos\alpha - \cos\lambda \, \sin\alpha \right) \right) \omega'_y$$

$$- \left(\sin\phi - \cos\phi \, \cos\alpha \tan V \right) \omega'_z + \left(A_0 - A \right) \tag{3.34}$$

Note that in the use of (3.34) the value of A_0 must be computed using equations that will require observed or estimated (through deflections of the vertical) astronomic positions.

The terrestrial normal equations are created for the whole network. Reduced normal equations are formed eliminating all unknowns except for those related to the stations common in the terrestrial system and in the space systems. These equations are merged with the normal equations from the space systems. The solution of this system yields the adjusted positions at the common stations, the orientation parameters, and the other parameters relating the space systems. Back substitutions yield the remaining parameters (including station positions) of the adjustment.

A modification of this procedure does not require the estimation of the space system relationships. This is done outside of the merger of the space and terrestrial systems. The parameters of the transformations are then used to convert the space system coordinates into the ideal system coordinates. These coordinates are then used, with an appropriate variance-covariance matrix, with the terrestrial data. This procedure is desirable when it is

105

doubtful the terrestrial data can yield information on the transformation parameters. For example, Vincenty (1982) indicates that the three rotation elements in the azimuth equation cannot be adequately separated, even in areas of continental extent.

3.4 Local Geodetic Datums of the World

Many different geodetic datums have been developed in the past geodetic history. These datums have been used for continental areas or for special purpose applications in a small region. Some datums have been superceded by later and move accurate datums. Some datums have been in use for more than 60 years. A listing of 55 datums (with the origin point and ellipsoid parameters) is given in Table 3.1. A listing of the names of 83 datums is given in Table 7.3 in the WGS84 report (DMA, 1987). Table 7.5 of the same report lists datum transformations for 97 local areas to the WGS84 system (See Section 3.8). However, not all 97 translation sets represent different datums. The ellipsoid parameters that are most often used with the different datums are given in Table 3.2.

Each datum has its own background. This background is given for some datums in Appendix A. In the following sections we discuss a few selected datums.

Table 3.1

Selected Geodetic Datums[+]

DATUM	SPHEROID	ORIGIN	LATITUDE	LONGITUDE (E)
Adindàn	Clarke 1880	STATION Z_x	22°10'07".110	31°29'21".608
American Samoa 1962	Clarke 1866	BETTY 13 ECC	-14 20 08.341	189 17 07.750
Arc-Cape (South Africa)	Clarke 1880	Buffelsfontein	-33 59 32.000	25 30 44.622
Argentine	International	Campo Inchauspe	-35 58 17	297 49 48
Ascension Island 1958	International	Mean of three stations	-07 57	345 37
Australian Geodetic 1966	Australian National	Johnston Geodetic Station	-25 56 54.55	133 12 30.08
Bermuda 1957	Clarke 1866	FT. GEORGE B 1937	32 22 44.360	295 19 01.890
Berne 1898	Bessel	Berne Observatory	46 57 08.660	07 26 22.335
Betio Island, 1966	International	1966 SECOR ASTRO	01 21 42.03	172 55 47.90 ·
Camp Area Astro 1961-62 USGS	International	CAMP AREA ASTRO	-77 50 52.521	166 40 13.753
Canton Astro 1966	International	1966 CANTON SECOR ASTRO	-02 46 28.99	188 16 43.47
Cape Canaveral*	Clarke 1866	CENTRAL	28 29 32.364	279 25 21.230
Christmas Island Astro 1967	International	SAT.TRI.STA. 059 RM3	02 00 35.91	202 35 21.82
Chua Astro (Brazil-Geodetic)	International	CHUA	-19 45 41.16	311 53 52.44
Corrego Alegre (Brazil-Mapping)	International	CORREGO ALEGRE	-19 50 15.140	311 02 17.250
Easter Island 1967 Astro	International	SATRIG RM No. 1	-27 10 39.95	250 34 16.81
Efate (New Hebrides)	International	BELLE VUE IGN	-17 44 17.400	168 20 33.250
European (Europe 50)	International	Helmertturm	52 22 51.446	13 03 58.928
Graciosa Island (Azores)	International	SW BASE	39 03 54.934	331 57 36.118
Gizo, Provisional DOS	International	GUX 1	-09 27 05.272	159 58 31.752
Guam 1963	Clarke 1866	TOGCHA LEE NO. 7	13 22 38.49	144 45 51.56
Heard Astro 1969	International	INTSATRIG 0044 ASTRO	-53 01 11.68	73 23 22.64
Iben Astro, Navy 1947 (Truk)	Clarke 1866	IBEN ASTRO	07 29 13.05	151 49 44.42
Indian	Everest	Kalianpur	24 07 11.26	77 39 17.57
Isla Socorro Astro	Clarke 1866	Station 038	18 43 44.93	249 02 39.28
Johnston Island 1961	International	JOHNSTON ISLAND 1961	16 44 49.729	190 29 04.781
Kourou (French Guiana)	International	POINT FONDAMENTAL	05 15 53.699	-52 48 09.149
Kusaie, Astro 1962, 1965	International	ALLEN SODANO LIGHT	05 21 48.80	162 58 03.28
Luzon 1911 (Philippines)	Clarke 1866	BALANCAN	13 33 41.000	121 52 03.000
Midway Astro 1961	International	MIDWAY ASTRO 1961	28 11 34.50	182 36 24.28
New Zealand 1949	International	PAPATAHI	-41 19 08.900	175 02 51.000
North American 1927	Clarke 1866	MEADES RANCH	39 13 26.686	261 27 29.494
Old Bavarian	Bessel	Munich	48 08 20.000	11 34 26.483
Old Hawaiian	Clarke 1866	OAHU WEST BASE	21 18 13.89	202 09 04.21
Ordnance Survey G.B. 1936	Airy,	Herstmonceux	50 51 55.271	00 20 45.882
OSGB 1970 (SN)	Airy	Herstmonceux	50 51 55.271	00 20 45.882
Palmer Astro 1969 (Antarctica)	International	ISTS 050	-64 46 35.71	295 56 39.53
Pico de las Nieves (Canaries)	International	PICO DE LAS NIEVES	27 57 41.273	344 25 49.476
Pitcairn Island Astro	International	PITCAIRN ASTRO 1967	-25 04 06.97	229 53 12.17
Potsdam	Bessel	Helmertturm	52 22 53.954	13 04 01.153
Provisional S. American 1956	International	LA CANOA	08 34 17.17	296 08 25.12
Provisional S. Chile 1963	International	HITO XVIII	-53 57 07.76	291 23 28.76
Pulkovo 1942	Krassovski	Pulkovo Observatory	59 46 18.55	30 19 42.09
Qornoq (Greenland)	International	No. 7008		
South American 1969	South American 1969	CHUA	-19 45 41.653	311 53 55.936
Southeast Island (Mahe)	Clarke 1880		-04 40 39.460	55 32 00.166
South Georgia Astro	International	ISTS 061 ASTRO POINT 1968	-54 16 38.93	323 30 43.97
Swallow Islands (Solomons)	International	1966 SECOR ASTRO	-10 18 21.42	166 17 56.79
Tananarive	International	Tananarive Observatory	-18 55 02.10	47 33 06.75
Tokyo	Bessel	Tokyo Observatory (AZABU)	35 39 17.5148	139 44 40.90
Tristan Astro 1968	International	INTSATRIG 069 RM No. 2	-37 03 26.79	347 40 53.21
USAFETR*	Clarke 1866	PAD 3	28 27 57.7564	279 27 43.1180
Viti Levu 1916 (Fiji)	Clarke 1880	MÒNAVATU (latitude only) SUVA (longitude only)	-17 53 28.285	178 25 35.835 ·
Wake Island, Astronomic 1952	International	ASTRO 1952	19 17 19.991	166 38 46.294
Wake-Eniwetok 1960	Hough	WAKE	19 16 19.606	166 39 21.798
WCT Vandenberg Adjustment*	Clarke 1866	ARGUELLO 2, 1959	34 34 58.021	239 26 22.361
White Sands*	Clarke 1866	KENT 1909	32 30 27.079	253 31 01.306
Yof Astro 1967 (Dakar)	Clarke 1880	YOF ASTRO 1967	14 44 41.62	342 30 52.98

* Local datums of special purpose, based on NAD 1927 values for the origin stations.

[+] from NASA Directory of Station Locations, 5th edition, Computer Sciences Corp., Silver Springs, MD, 1978

Table 3.2

Reference Ellipsoid Parameters

Ellipsoid Name (Year computed)	Semi-Major Axis (a) (m)	Inverse Flattening 1/f
Airy (1830)	6378563.396	299.3249646
Modified Airy	6377340.189	299.3249646
Bessel (1841)	6377397.155	299.1528128
Clarke 1866	6378206.4	294.9786982
Clarke 1880 (modified)	6378249.145	293.4663
Clarke 1880	6378249.145	293.465
Everest (1830)*	6377276.345	300.8017
Modified Everest	6377304.063	300.8017
International (1909)	6378388	297
Krassovski (1940)	6378245	298.3
Mercury 1960	6378166	298.3
Modified Mercury 1968	6378150	298.3
Australian National	6378160	298.25
South American 1969	6378160	298.25
Geodetic Reference System 1967	6378160	298.2471674273
WGS 60	6378165	298.3
WGS 66	6378145	298.25
WGS 72	6378135	298.26
WGS 84	6378137	298.257223563
Rapp (1987)	6378136.2	298.257222101
IAG Recommendation (1987)	6378136±1m	1/298.257
Geodetic Reference System 1980	6378137	298.257222101

3.4.1 The European Datum

A number of different country organized datums were the rule in Europe prior to 1950. After World War II a substantial effort was made to connect the separate country triangulation networks. The original adjustment was for a Central European Network. A total of 52 base lines and 106 Laplace azimuths scaled and oriented this system with no particular point considered as a datum point. For convenience, for comparison purposes, Helmert Tower, near Potsdam, is considered the origin. Two large networks, the Southwestern Block and the Northern Block were added to the Central European Network. The merged networks were referred to as the European Datum 1950. The reference ellipsoid was the International Ellipsoid, although base lines were apparently not reduced to this ellipsoid.

At the 1954 General Assembly of the International Association of Geodesy in Rome, an agreement was reached to carry out a new adjustment of the European triangulation data that would be "more complete and rigerous than ED50", (Kobold, 1980). The structure under which this work was to be done was a commission, of IAG, on the "Readjustment of the European Triangulation Network (RETrig)". The RETrig Commission became a Subcommission of Commission X "Continental Networks" that was established by IAG in 1975 at Grenoble, France. The RETrig work had been carefully planned to be carried out in phases. Phase I was completed in 1975. Phase II, which was to incorporate Laplace - azimuths and distances, had first results presented in 1977. A second part of Phase II was completed in 1979 with the resulting adjusted system called ED79 (Hornik and Reinhart, 1980). To provide continuity with ED50 one station, D 7835 München, was held fixed in ED79, at its ED50 coordinates and deflections of the vertical, and the International Ellipsoid was retained as the reference surface. A classical two-dimensional adjustment was carried out.

The basic method used was one of Helmert blocking where various countries (or areas) carried out the normal equation computations for their respective areas. Reduced normal equations were formed leaving only common (or border) stations in the reduced system. The ten RETrig blocks were then combined to obtain the adjusted coordinates for the junction points. Back substitution for the adjusted coordinates in each block was accomplished by the separate responsible groups. The total number of observations in ED79 were 25111 and the total number of unknowns was 11168. Figure 3.2 (taken from Ehrnsperger et al. (1982)) shows the triangulation networks that formed ED79.

Figure 3.2

European Triangulation Networks used in ED79
(Ehrnsperger et al. (1982))
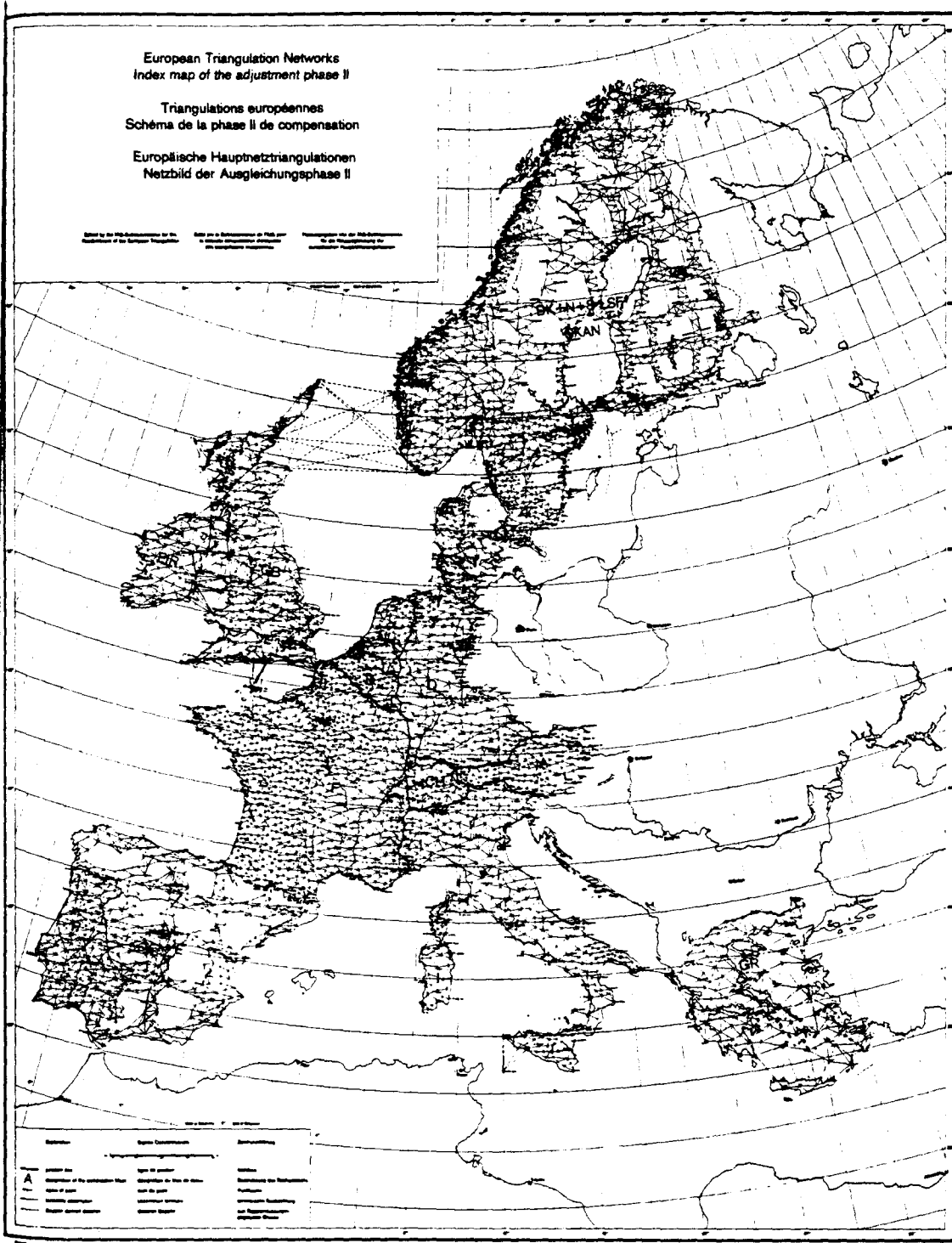
In 1985 a revised terrestrial solution was carried out (Ehrnsperger, 1985). This solution incorporated additional data (from the ED79) solution and introduced additional scale and orientation unknowns into the different blocks. The scale unknowns could be applied for the usual baselines, and for Geodimeter and Tellurometer measurements. It was found in the terrestrial solution that scale parameters varied by instrument and by block.

A combined solution with Doppler stations was carried out with the terrestrial network. The method used for the combination was that of Wolf (1982) as described earlier in these notes. The resultant solution was described as the "Rough Solution" of RETrig Phase III.

In 1987 the work of the RETrig Subcommission was completed with the calculation of the European Datum 1987 (ED87). This solution combined a terrestrial network data with positions derived from Doppler satellite tracking, laser tracking of Lageos, and with rectangular coordinate differences derived from VLBI measurements. The details of ED87 are described by Kelm (1987) and Ehrnsperger (1987). The principles for the adjustment of ED87 are as follows (Kelm, 1987):

A: The terrestrial observations are projected onto a rotational ellipsoid (Hayford).

B: The adjusted coordinates are defined as geographical latitudes and longitudes on terrestrial points and additionally ellipsoidal heights on satellite stations.

C: Bias parameters for the scale of terrestrial distance and the rotation of azimuth measurements are estimated.

D: The free net adjustment is relative to a local datum.

E: Three-dimensional Cartesian coordinate sets (satellite- or VLBI-derived) are transformed to the local datum with respect to given ellipsoid coordinates before entering the 3D adjustment on the local datum.

F: The adjusted coordinates on the local datum (ED87) are related to a satellite reference system or a conventional terrestrial system by an appropriate similarity transformation (3 translation, 3 rotation and 1 scale).

G: An approximated variance factor estimation is applied by standardizing the normal equations of the national blocks by the variance of unit weight obtained by the internal free adjustment of each block.

H: The linearized adjustment model will be iterated up to a significant convergence level.

The final phase, III, of the development of the new European datum (ED87) is described by Ehrnsperger (1988). The final solution was carried out by combining the terrestrial network normal equations with space positions through the Wolf procedure. In doing this the terrestrial data from 16 European countries was merged with space data from approximately ten systems. The scale of the system was taken from the space networks and the terrestrial data. The overall orientation is defined by the space networks.

3.4.2 The Australian Geodetic Datum

The Australian Geodetic Datum of 1966 (AGD66) was based on an adjustment of terrestrial geodetic data fixing the coordinates of Johnston Geodetic Station as:

111

latitude	- 25°	56′	54″5515
longitude	133°	12′	30″0771
Ellipsoidal Height	571.2 m		

This station is located near the center of the country and the coordinates were based on a previous solution. In the adjustment, leading to AGD66 there were 2506 stations of which 533 were Laplace stations. The distances were reduced to the geoid because of a lack of knowledge of the geoid undulations at that time. The ellipsoid used for AGD 66 was: a = 6378160 m, f = 1/298.25.

In 1981 a proposal was made for the readjustment of the Australian Primary Network. The new adjustment would take into account the additional geodetic data that had been acquired since AGD66 had been calculated, and also incorporate space derived position information. This latter information would include point-position Doppler values using the precise satellite ephemeres; relative Doppler positions; satellite laser ranging distances; and VLBI chord distances. The terrestrial and space data available as of, approximately, 1983 was used in the development of the new model. In order to have continuity with previous models it was decided to retain the Johnston Geodetic Station at the coordinates specified for AGD66 and to use the same ellipsoid as in AGD66.

A long wavelength geoid model based on a spherical harmonic expansion to degree 20 was used for distance reductions. However, before use, the geocentric undulations were transformed to the AGD datum through a three parameter translation, and bias term, transformation. Deflections of the vertical were also introduced for direction reductions. This terrestrial data set was divided into 35 blocks containing 5,498 stations, 30063 directions, 12506 distances, and 1,292 Laplace azimuths. A preliminary adjustment of this data was carried out.

The satellite laser ranging and very long baseline interferometry (VLBI) data were used to compute eleven chord distances between a number of points in the network. Doppler positions, in the coordinate system, were determined at 156 stations. A seven parameter transformation was developed, in two stages, for converting from the NSWC 9Z-2 system to the AGD system. This was done by first using the coordinates of the terrestrial data adjustment. After a preliminary combination of the terrestrial and satellite data, a final set of transformation parameters was developed. These transformation parameters are given in Table 3.3 (Allman and Veenstra, 1984):

Table 3.3

Transformation Parameters Going From
NSWC 9Z-2 to AGD84 (GMA82)

Parameter	Initial Value	Final Value	Std. Dev.
ΔX	116.47 m	116.00 m	±1.2 m
ΔY	50.25	50.47	1.2
ΔZ	-138.87	-137.19	1.5
ω_x	0″21	0″23	0″04
ω_y	0.36	0.39	0.04
ω_z	-0.47	-0.47	0.04
Δs	-0.75 ppm	-0.699 ppm	0.07 ppm

These transformation parameters were used to convert the NSWC 9Z-2 positions into the AGD systems for a subsequent combined adjustment.

A classical adjustment procedure was used where there were two unknowns for each free station plus one orientation parameter for each set of observations. The final merged adjustment took place in March 1984. The model, called the Geodetic Model of Australia 1982 was adopted by the National Mapping Council in October 1984. This system was also to be referred to as the Australian Geodetic Datum 1984 (AGD84).

3.4.3 The North American Datum 1983 (NAD83)

The NAD27 was developed in the 1920's on the basis of triangulation data available at that time, fixing the coordinates of a single point (Meades Rarch) and retaining the Clarke 1866 ellipsoid that had been used in prior geodetic work in the United States. Over time this network became inadequate due to improving accuracy of measuring devices, and due to inherent weakness in the network saved by lack of data and incomplete data reduction.

In the early 1970's a decision was made to readjust the whole network using as much data as possible, with as many stations as feasible, and with the introduction of space positions so that the new system (to be called NAD83) would be essentially geocentric with coordinates referred to a modern ellipsoid, that of the Geodetic Reference System 1980.

The task of data management and subsequent adjustment was immense considering that approximately 1.8 million observation equations with 900,000 unknowns, involving 270,000 stations, were to be analyzed. The data involved came from the United States, Canada, Mexico, Central America, some Caribbean Islands, and Denmark (for Greenland) (Bossler, 1987).

Space information from Doppler positioning, laser ranging, and VLBI was incorporated into this system. Details may be found in Schwarz and Wade (1990) and Schwarz (1989).

3.5 Space Based Reference Systems

The determination of station positions through the observations of satellites is a complex process involving the estimation of many parameters (Malys, 1988). Such parameters include force model components (potential coefficients, atmospheric drag, radiation pressure, etc.), environmental factors (e.g., tropospheric and ionospheric refraction), earth orientation changes (e.g., earth rotation and polar motion), and station positions. Some of these quantities, as well as other quantities (such as precession and nutation) can be held fixed in a solution, depending on what type of solution is being sought. In some cases a simultaneous solution is made for all estimable quantities, while in other cases only selected quantities are estimated. Basic to this process is the definition of a coordinate system. In the most general case, the coordinate system is related to the complete "description of the physical environment as well as theories under the definition of the coordinates" (Malys, 1988). When some quantities, specifically station positions, are considered fixed, the resultant system is said to define a reference frame. Positions of points, other than the fundamental positions, are defined in the frame defined by the fixed stations.

Laser observations of Lageos have formed a basis for the determination of station positions to an accuracy of ±10 cm in a specific coordinate system defined by the group that is carrying out the analysis of the data. The laser systems provide station positions that conceptually refer to the center of mass of the Earth. A discussion of a number of these systems may be found in BIH (1988). Very long line interferometry (VLBI) provides precise coordinate differences in a well defined coordinate systems. By tieing a VLBI station into a laser system the VLBI stations can be placed in a geocentric system. A discussion of several VLBI analysis may be found in BIH (1988).

In the past a widely used positioning system has been the Navy Navigation Satellite System (NAVSAT). This system has provided positions through the analysis of Doppler data obtained from Doppler receivers. Since the start of the Navy program a number of different reference frames have been used. Between June 1977 and December 31, 1986 the reference frame was designated NSWC 9Z-2. The NAVSAT precise ephemerides were calculated in this frame with the NWL/10E-1 gravity model (Malys, 1988). The position accuracy for stations estimated in the NSWC 9Z-2 frame is approximately ±1m. The coordinates determined in this frame are consistently determined but they are not determined with respect to the center of mass of the Earth, nor the ideal orientation, as we shall subsequently see.

Starting January 1, 1987 the mean reference frame for the NAVSAT system became the World Geodetic System 1984 WGS84 (DMA; 1987) (See Section 3.8). WGS84 is also the coordinate system for the Global Positioning System (GPS).

3.6 The BIH Terrestrial System (BTS)

The information from the various systems for positioning and the determination of earth-rotation parameters can be merged to form a single consistent reference frame to which the individual system can be related. This merger can be done for sites for which station positions are available in at least two networks (Boucher and Altamimi, 1985). One example of the determination of such relationships is that of Hothem, Vincenty, and Moore (1982). In this work the authors related certain space positions to a three-dimensional geodetic network (the Trancontinental Traverse (TCT)).

The development of the Bureau International de l'Heure (BIH) Terrestrial Systems (BTS) is described in a sequence of papers by Feissel (1985), for BTS84 and by Boucher and Altamimi (1985, 1987, 1988). The basic method is to adopt a seven parameter similarity transformation between the various systems. Based on co-located, or connected stations the estimates of the transformation parameters are determined after agreeing on what quantities are to be considered fixed by one or more systems. The calculations use data over a period of time that is of sufficient accuracy that plate motion models are required. Because of this motion the BTS system is defined at an epoch 1984.0. In the 1987 solution information from 13 systems involving 64 co-located stations was used. The data was from lunar and satellite laser ranging; VLBI; and Doppler measurements. We give in Table 3.4 the transformation parameters (BIH, Table 3, 1988) from the individual systems into the BTS 87 system. A value given as zero means the system is defined by that quantity. The orientation was fixed to the BTS 86 system.

114

Table 3.4

Transformation from the individual system to BTS 87

Network*	SSC+	−Δx cm	−Δy cm	−Δz cm	−Δs 10^{-8}	ω_x ".001	ω_y ".001	ω_z ".001
NGS (VLBI)	88 R 01	-8.9	14.3	-1.6	0.9	-4.3	9.3	-3.3
GSFC (VLBI)	88 R 01	167.1	-105.4	40.0	-2.3	-7.2	5.0	-3.6
JPL (VLBI)	83 R 05	24.2	-12.1	-18.3	-3.9	-1.4	-4.8	-5.7
JPL (Lunar)	88 M 01	0.0	0.0	0.0	-2.4	-5.6	-0.8	-3.8
CSR (Lageos)	88 L 01	0.0	0.0	0.0	0.0	-4.7	2.3	-10.5
GSFC (Lageos)	87 L 14	0.0	0.0	0.0	0.0	-1.8	6.2	-7.5
DMA (Doppler)†	77 D 01	7.1	-50.9	-466.6	58.3	17.9	-0.5	-807.3
WGS 84**		7.1	-50.9	-16.6	-1.7	17.9	-0.5	6.7

*NGS = National Geodetic Survey
GSFC = NASA Goddard Space Flight Center
JPL = Jet Propulsion Laboratory
CSR = Center for Space Research (University of Texas at Austin)
DMA = Defense Mapping Agency (NSWC 9Z-2)
** see discussion in Section 3.8
+ Station coordinate identification
† NSWC 9Z-2 system

The large translation values for GSFC (VLBI) occur because the stations had not been moved to a geocentric system. The Δz value for the Doppler system shows the well known z bias while the large ω_z represents the 0".8 longitude bias of the NSWC 9Z-2 system. (The origin of the Doppler system is 0".8 to the east of the BTS system.)

3.7 The IERS Terrestrial System

In 1988 the International Earth Rotation Service (Mueller, 1988, IERS, 1988) became operational. This service replaces the International Polar Motion Service (IPMS) and the earth-rotation section of the Bureau International de l'Heure (BIH). The activities of IERS have led to the definition and maintenance of a conventional terrestrial reference system (CRTS). This system will be analogous to the BIH CTS (BTS). The new system, in its first reference frame implementation, is described by Boucher and Altamimi (1989). The initial IERS Terrestrial Reference Frame is designated ITRF-0 and has a reference epoch of 1988.0. Transformation parameters between the individual systems and ITRF-0 are given in Table 3 of Boucher and Altamimi (ibid). Analysis on a yearly basis is carried out in defining the ITRF. The ITRF89 is described by Boucher and Altamimi (1991).

3.8 The World Geodetic System 1984 (WGS84)

WGS84 (DMA, 1987) is a successor to three (WGS60, 66, 72) prior global systems defined by the Defense Mapping Agency. The WGS system includes a coordinate frame definition, a gravity field model defined by a set of potential coefficients, a reference ellipsoid, and related quantities. The WGS84 coordinate system was defined by applying selected transformation parameters to the NSWC 9Z-2 coordinate system. The parameters used were close to those calculated by Boucher and Altamini (1985). The non-zero parameters adopted for the transformation were Δz = 4.5 m; ω_z = -0".814; Δs = -0.6X10^{-6}. Because of this procedure the conversion from WGS84 to a BTS system can be derived

using the DMA (Doppler) transformation parameters minus the defining parameters of WGS84. Other defining parameters of WGS84 are shown in Table 3.4.

Table 3.5

WGS 84 Parameters

Quantity	Value
a	6378137
GM	3986005×10^8 m^3 s^{-2}
$\bar{C}_{2,0}$	$-484.16685 \times 10^{-6}$
ω	7292115×10^{-11} radians s^{-1}
c	299792458 m s^{-1}
f*	1/298.257223563

* a derived quantity

The WGS84 values are the same (except for f) as GRS80 (Moritz, 1980). ($\bar{C}_{2,0}$ was derived from the J_2 defined as part of GRS80.)

Malys (1988) has studied the relationship between WGS84 and NSWC 9Z-2 by using station coordinates that had been determined in NSWC 9Z-2, and in WGS84 through the precise orbits that were computed in the WGS84 system. The differences found were quite similar to the values shown in Table 3.4.

The DMA WGS84 report provides the three translation parameters that can be used to convert local geodetic systems to WGS84. Because only three parameters are given for some datums, one can expect that, in large networks, there will be residual distortions that can be related to the neglected transformation parameters (three rotations and one scale), as well as distortions in the geodetic network. This is demonstrated by considering the variations of Δx for the NAD27 to WGS84 conversion shown in Figure 3.3. The table of transformation parameters as taken directly from DMA (1987) is given in Appendix B. The transformation from the local datum to WGS84 using multiple regression formulas (see Section 2.33) has also been developed for several datums. An example of such a transformation is given in Appendix B for NAD27 to WGS84.

We finally note that NAD83 and WGS84 are in the same coordinate system as both NGS and DMA adopted the same transformation parameters in going from NSWC 9Z-2 to the "ideal" system.
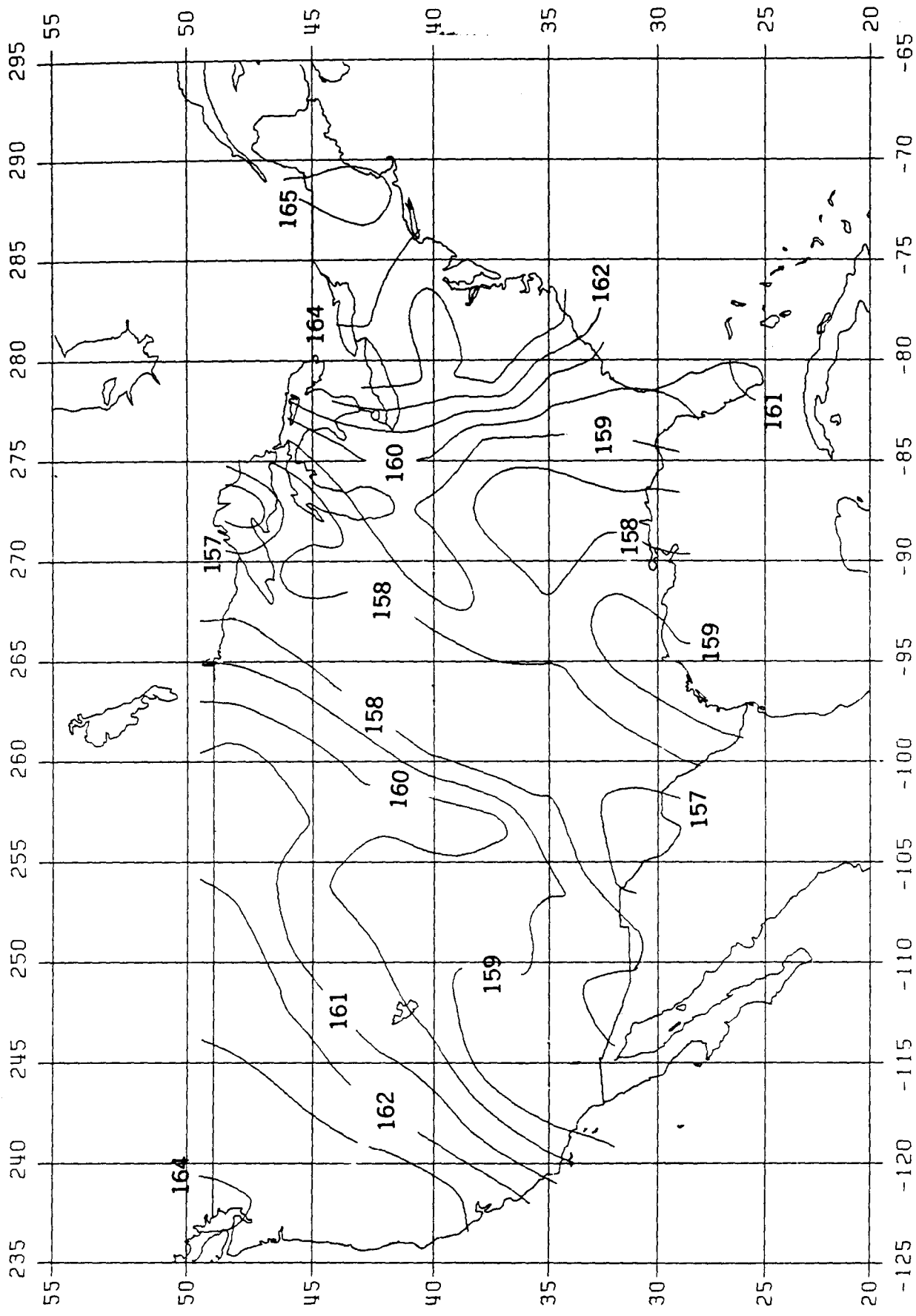
Figure 3.3 ΔY Datum Shift, NAD-27 to WGS84 (DMA, 87) Contour Interval is 1 m

117

3.9 The Estimation of the Datum Origin Coordinates and Ellipsoid Parameters

In the previous sections we have considered global coordinate systems and their relationship to local geodetic systems. In the following sections we examine how most geodetic datums were established in the past. We also consider how ellipsoid parameters are estimated, simultaneously with datum positions, and separately using space measurements. Allan and Audson (1987) describe a number of techniques that have been used historically for the determination of the size and shape of the Earth.

3.9.1 The Determination of a Best Fitting Datum and Reference Ellipsoid

We now consider the case where we wish to define a datum where the ellipsoid is matched, in some specified way, to the true geoid, in a given region (σ). The condition that can be specified, in analogy, to a least squares adjustment condition is:

$$\int_\sigma \int N^2 \, d\sigma = \text{a minimum}$$

(3.35)

In practice the integral is replaced by a summation:

$$\sum_\sigma N^2 = \text{a minimum}$$

(3.36)

The precise interpretation of N will be discussed later although we can start by interpreting N as an astro-geodetic undulation with respect to the ideal datum and ellipsoid.

An alternative specification to (3.37), and a more convenient form to apply, is the following:

$$\sum_\sigma \left(\bar{\xi}_{AG}^2 + \bar{\eta}_{AG}^2 \right) = \text{a minimum}$$

(3.37)

where $\bar{\xi}_{AG}$ and $\bar{\eta}_{AG}$ are astro-geodetic deflections of the vertical with respect to our "best" datum. Equation (3.37) indicates that we try to minimize the square sum of the deflections over the area for which the datum is to be determined. Equation (3.37) does not determine (3.36), but (3.36) does imply (3.37) on a global basis (Heiskanen and Moritz, 1967, sect. 5-11).

Consider an existing triangulation network where we have values of ξ_{AG}, and η_{AG} throughout the system. They will depend, in part, on the geodetic coordinates of these stations. In turn these geodetic coordinates will depend on the adopted origin parameters and the parameters of the original reference ellipsoid. If we wish to use a condition such as (3.37), we must relate it to these origins or ellipsoid changes. We thus may interpret our condition such that:

$$\bar{\xi}_{AG} = \xi_{AG} + d\xi$$

$$\bar{\eta}_{AG} = \eta_{AG} + d\eta$$

(3.38)

Then using equation (2.107) and (3.38) we have for an observation equation:

$$\overline{\xi}_{AG} = \xi_{AG} + E_1^{'} d\xi_0 + E_2^{'} d\eta_0 + E_3^{'} dN_0 + E_4^{'} da + E_5^{'} df$$

$$\overline{\eta}_{AG} = \eta_{AG} + F_1^{'} d\xi_0 + F_2^{'} d\eta_0 + F_3^{'} dN_0 + F_4^{'} da + F_5^{'} df \tag{3.39}$$

In principle we use (3.39) as observation equations given astrogeodetic data in a network and determine the parameters subject to the minimization condition expressed by equation (3.37). We may also express $d\xi$ and $d\eta$ needed in (3.38) using equation (2.108). We have:

$$\overline{\xi}_{AG} = \xi_{AG} + E_1^{''} dx + E_2^{''} dy + E_3^{''} dz + E_4^{''} da + E_5^{''} df$$

$$\overline{\eta}_{AG} = \eta_{AG} + F_1^{''} dx + F_2^{''} dy + F_3^{''} dz + F_4^{''} da + F_5^{''} df \tag{3.40}$$

In applying these projective transformation formulas, however, we cannot determine (accurately) the da correction. To show this we first write the $E_3^{'}$, $E_4^{'}$ and $F_3^{'}$ and $F_4^{'}$ coefficients. We have:

$$E_3^{'} = - \frac{1}{(M+h)} \left[\sin\varphi_0 \cos\varphi - \cos\varphi_0 \sin\varphi \cos \Delta\lambda \right]$$

$$E_4^{'} = - \frac{1}{(M+h)} \left[\frac{1}{W_0} \left(\sin\varphi_0 \cos\varphi \, (1-e^2) - \cos\varphi_0 \sin\varphi \cos \Delta\lambda \right) + \frac{e^2}{W} \sin\varphi \cos\varphi \right] \tag{3.41}$$

$$F_3^{'} = \frac{1}{(N+h)} \left[\cos\varphi_0 \sin \Delta\lambda \right]$$

$$F_4^{'} = \frac{1}{(N+h)} \left[\frac{\cos\varphi_0 \sin \Delta\lambda}{W_0} \right] \tag{3.42}$$

Comparing $E_3^{'}$ to, $E_4^{'}$ and $F_3^{'}$ to $F_4^{'}$ we find that the coefficients are very close to being the same, and in a spherical approximation they are equal. This means that we cannot determine dN_0 and da separately. Rather we must solve for dN_0 or da regarding one as known or solve for the sum $dN_0 + da$.

If we consider equation (3.40) we can express $E_3^{''}$ and $F_4^{''}$ as:

$$E_4^{''} = - \frac{1}{(M+h)} \frac{e^2 \sin\varphi \cos\varphi}{W}$$

$$F_4^{''} = 0 \tag{3.43}$$

In this case the coefficient of da is very small because of the e^2 in the top equation of (3.43), and exactly zero in the second equation. This then indicates that we cannot effectively determine an equatorial radius using the projective method equations with deflection of the vertical data.

If our geodetic network has been computed using the development method, then the $d\xi$ and $d\eta$ values in (3.38) must be taken from equations (2.152) and (2.153). We have in this case:

$$\overline{\xi}_{AG} = \xi_{AG} + p_1' d\xi_0 + p_2' d\eta_0 + p_3' ds_i + p_4' d\alpha_0 + p_5' \frac{da}{a} + p_6' df$$

$$\overline{\eta}_{AG} = \eta_{AG} + q_1' d\xi_0 + q_2' d\eta_0 + q_3' ds_i + q_4' d\alpha_0 + q_5' \frac{da}{a} + q_6' df \tag{3.44}$$

These equations are then used in a best fitting ellipsoid determination using (3.37). In such a procedure ds_i may be set to zero or it may be determined through (2.151). If we assume that the astronomic azimuth is correct at the origin, then (3.44) should be rewritten using (2.154). Examination of the coefficents p_5' and q_5' indicates that there is no problem in determining an equatorial radius from these equations. Based on our previous discussion in Section 2.32.1 we could use our projective equations for an equatorial radius determination, in analogy to that obtained from the application of (3.44) if we assume that $dN_0 = 0$, and restrict that analysis to data not too distant from the datum origin.

To this point we have considered deflections of the vertical. If we have astro-geodetic undulations in the projective system we can write that the new undulation, $\overline{N}_{AG}$, should be equal to the original values, N_{AG}, plus a correction due to the datum changes. We have:

$$\overline{N}_{AG} = N_{AG} + dN \tag{3.45}$$

Using (2.107) or (2.108) we have:

$$\overline{N}_{AG} = N_{AG} + G_1' d\xi_0 + G_2' d\eta_0 + G_3' dN_0 + G_4' da + G_5' df \tag{3.46}$$

or

$$\overline{N}_{AG} = N_{AG} + G_1'' dx + G_2'' dy + G_3'' dz + G_4'' da + G_5'' df \tag{3.47}$$

where the best fitting ellipsoid condition is given by (3.36). Examination of the coefficients G_3' and G_4' indicate that they are sufficiently different to allow a determination of both dN_0 and da. In addition G_4'' is not close to zero, also indicating that an equatorial radius can be determined using the astro-geodetic undulation information.

Once the correction terms are found using (3.39), (3.44), (3.46) and (3.47) the new datum origin coordinates and the new ellipsoid parameters are found by adding the corrections to the starting values.

We note that no specification has been made that would make the center of the best fitting ellipsoid coincide with the center of mass of the earth. In general it will not. Consequently the interpretation of a best fitting ellipsoid does not lead to one whose center is at the center of mass of the earth.

3.9.1.1 A Modified Best Fitting Ellipsoid

The condition expressed by (3.36) must be reconsidered noting that the forcing of sums of the astrogeodetic deflections squared to be a minimum is somewhat unwarranted because, at each point in our network, there is a deflection of the vertical. This deflection may be estimated through the computation of topographic-isostatic deflections (Heiskanen and Vening-Meinesz, 1958, pp. 252-255) or through the computation of gravimetric deflections (Heiskanen and Moritz, 1967). In this section we consider the topographic-isostatic deflections ξ_{TI} and η_{TI}. These deflections are those implied by the topography and its isostatic compensation.

We now express our modified best fitting ellipsoid as one where:

$$\bar{\xi}_{AG} = \xi_{TI} + v_\xi = \xi_{AG} + d\xi$$

$$\bar{\eta}_{AG} = \eta_{TI} + v_\eta = \eta_{AG} + d\eta \tag{3.48}$$

with the specific condition that:

$$\sum_\sigma \left(v_\xi^2 + v_\eta^2 \right) = a \text{ minimum} \tag{3.49}$$

If we use the development method equations for $d\xi$ and $d\eta$ we can write from (3.48):

$$v_\xi = \xi_{AG} - \xi_{TI} + p_1' \, d\xi_0 + p_2' \, d\eta_0 + p_3' \, ds_i + p_4' \, d\alpha_0 + p_5' \frac{da}{a} + p_6' \, df$$

$$v_\eta = \eta_{AG} - \eta_{TI} + q_1' \, d\xi_0 + q_2' \, d\eta_0 + q_3' \, ds_i + q_4' \, d\alpha_0 + q_5' \frac{da}{a} + q_6' \, df \tag{3.50}$$

The limitations discussed in the previous sections all apply here. The advantage of this method over that of the previous section is simply that we are now trying to add some additional data (ξ_{TI} and η_{TI}) to our datum determination problem. This method of using topographic isostatic deflections was carried out in the United States by Hayford (1909, 1910). Note that we do not consider the case of topographic-isostatic undulations as they are not generally computed.

3.9.2 A General Terrestrial Ellipsoid Based on Astro-Geodetic Data

The ellipsoid determined using the methods considered in the previous sections will not have their centers at the center of mass of the earth since no information concerning the earth's gravitational field has entered the discussion. Now a general terrestrial ellipsoid or mean earth ellipsoid (Heiskanen and Moritz, 1967, Section 2-21) is one whose center is at the center of mass of the Earth and whose parameters meet certain conditions related to the minimization of astro-geodetic deflections or astro-geodetic undulations.

In order to obtain a general terrestrial ellipsoid using the methods being discussed in this chapter we compute the gravimetric deflections of the vertical (ξ_g, η_g) and the gravimetric undulations (N_g) using the Stokes' and Vening-Meinesz equations (Heiskanen and Moritz, 1967). These quantities will refer to an ellipsoid whose center is at the center of mass of the earth. In carrying out such computations we will need to adopt a flattening of the reference ellipsoid for use in determining the gravity formula used in computing the

anomalies that were used to determine ξ_g, η_g, N_g. The simplest case arises if we assume that this flattening is the value we wish to consider fixed for our new system. If this is not the case we will have to consider the change in ξ_g, η_g, N_g caused by a flattening change using equations given, for example, by Pick et al. (1973, equation 862). In our subsequent discussion we will assume that the flattening of the general terrestrial ellipsoid is adequately known from satellite data, for example, so that the gravimetric terms are considered without any dependence on the flattening:

We now can write:

$$\overline{\xi}_{AG} = \xi_g + v_\xi = \xi_{AG} + d\xi$$

$$\overline{\eta}_{AG} = \eta_g + v_\eta = \eta_{AG} + d\eta$$

$$\overline{N}_{AG} = N_g + v_N = N_{AG} + dN \tag{3.51}$$

Using the projective transformation formulas, (3.51) can be expressed as:

$$v_\xi = \xi_{AG} - \xi_g + E_1' d\xi_0 + E_2' d\eta_0 + E_3' dN_0 + E_4' da + \left[E_5' df \right]$$

$$v_\eta = \eta_{AG} - \eta_g + F_1' d\xi_0 + F_2' d\eta_0 + F_3' dN_0 + F_4' da + \left[F_5' df \right]$$

$$v_N = N_{AG} - N_g + G_1' d\xi_0 + G_2' d\eta_0 + G_3' dN_0 + G_4' da + \left[G_5' df \right] \tag{3.52}$$

where the term in brackets is considered known. Equation (3.52) can also be expressed in terms of dx, dy and dz as unknowns as follows:

$$v_\xi = \xi_{AG} - \xi_g + E_1'' dx + E_2'' dy + E_3'' dz + E_4'' da + \left[E_5'' df \right]$$

$$v_\eta = \eta_{AG} - \eta_g + F_1'' dx + F_2'' dy + F_3'' dz + F_4'' da + \left[F_5'' df \right]$$

$$v_N = N_{AG} - N_g + G_1'' dx + G_2'' dy + G_3'' dz + G_4'' da + \left[G_5'' df \right] \tag{3.53}$$

We should point out that we could have also written the deflection of the vertical equations (but not the undulation equations) using the development transformation equation.

We should finally note that in imposing our conditions on the adjustment we can use:

$$\sum_\sigma \left(v_\xi^2 + v_\eta^2 \right) = \text{a minimum}$$

122

or

$$\sum_\sigma \left(v_N^2\right) = a \text{ minimum}$$

but we cannot impose both conditions as the astrogeodetic undulation data is not independent of the astrogeodetic deflection data.

3.9.3 Remarks on the Area Method

The discussion in the previous sections has revolved around geodetic data given in a geodetic network. Since this data is considered to be distributed in an area in which the datum is to be established the method is often called the area method of best fitting ellipsoid determination. This will be contrasted in a subsequent section to the arc method where data is given along a meridian or a parallel.

Clearly, if we are determining a datum for an area, whether this area is a local region or a continent, we should use data as widely distributed in this area as possible. If we do not do this, the extension of the datum into areas where no data was used in the datum determination, may lead to large astrogeodetic deflections and undulations in such areas. We also have to note that these astrogeodetic techniques are generally only applied to land areas because of the difficulty in determining astrogeodetic information in oceanic areas.

We need also to discuss the advantages and disadvantages of a best fitting ellipsoid as opposed to a general terrestrial ellipsoid. In the former case we determine a datum in an area such that the ellipsoid comes as close as possible to the geoid in the area. In the case of the general terrestrial ellipsoid our primary concern is to obtain a geocentric ellipsoid after which a best fitting principle is applied. For a given area or country arguments can be presented for a best fitting ellipsoid or for a general terrestrial ellipsoid. A review of some of them may be found in Mueller (1974).

Basically the argument for a best fitting ellipsoid is that the astro-geodetic deflections and undulations are smaller than in the case of a general terrestrial ellipsoid. Thus any errors caused by the neglect of deflections of the vertical or geoid undulations in the reduction of geodetic data will be smaller with a best fitting ellipsoid. On the other hand, if we determine best fitting ellipsoids for each area of interest we will have a large number of datums in the world which will make the comparison of geodetic coordinates in different parts of the world difficult at best. In the general terrestrial ellipsoid system a system of coordinates that is unique and that may be used on a global basis can be established, thus determining a unique set of coordinates consistent throughout the world.

In order to demonstrate the geoid undulation behavior with respect to a global system and a local system two maps have been prepared for the Australia area. Figure 3.4 shows the geoid undulations computed from the OSU86F (Rapp and Cruz, 1986) potential coefficient model to degree 180. These undulations refer to a geocentric ellipsoid whose dimensions are those of the ideal global ellipsoid. Figure 3.5 shows the geoid undulations referred to the local Australian datum (GMA82) described in Section 3.4.2. These values were calculated from the undulations shown in Figure 3.4 plus the differential height change given by equation (2.89) using the following translations (Allman and Veenstra, 1984): $\Delta x = 116.5$ m, $\Delta y = 50.3$ m, $\Delta z = -138.9$m. Assuming the ideal ellipsoid radius

123

to be a = 6378136 m and knowing the a (6378160 m) of GMA82 we have $\Delta a = 24$ m. The Δf value was taken as zero.

Figure 3.4 shows the strong geoid slope across Australia when in the geocentric system. Figure 3.5 shows the much different character of the undulations when the local system is the reference. The undulations are now much smaller without the massive slope apparent in Figure 3.4. The undulations in Figure 3.5 are similar to those calculated from astro geodetic deflections of the vertical by Fischer and Slutsky (1967).

One final note relates to the role of datum definition and extension through satellite positioning techniques. Positions determined by such techniques are related to the satellite reference frame. These positions can be used in the datum defined by the satellite frame, or converted to a previously existing geodetic datum through datum transformation procedures. In some applications such as with the Global Positioning System, relative positions are determined. The relative positions may be in the local geodetic datum system, or in the satellite frame. Although the satellite datum definition gives a more consistent coordinate determination the advantages and disadvantages of geocentric and local datums still need to be addressed.
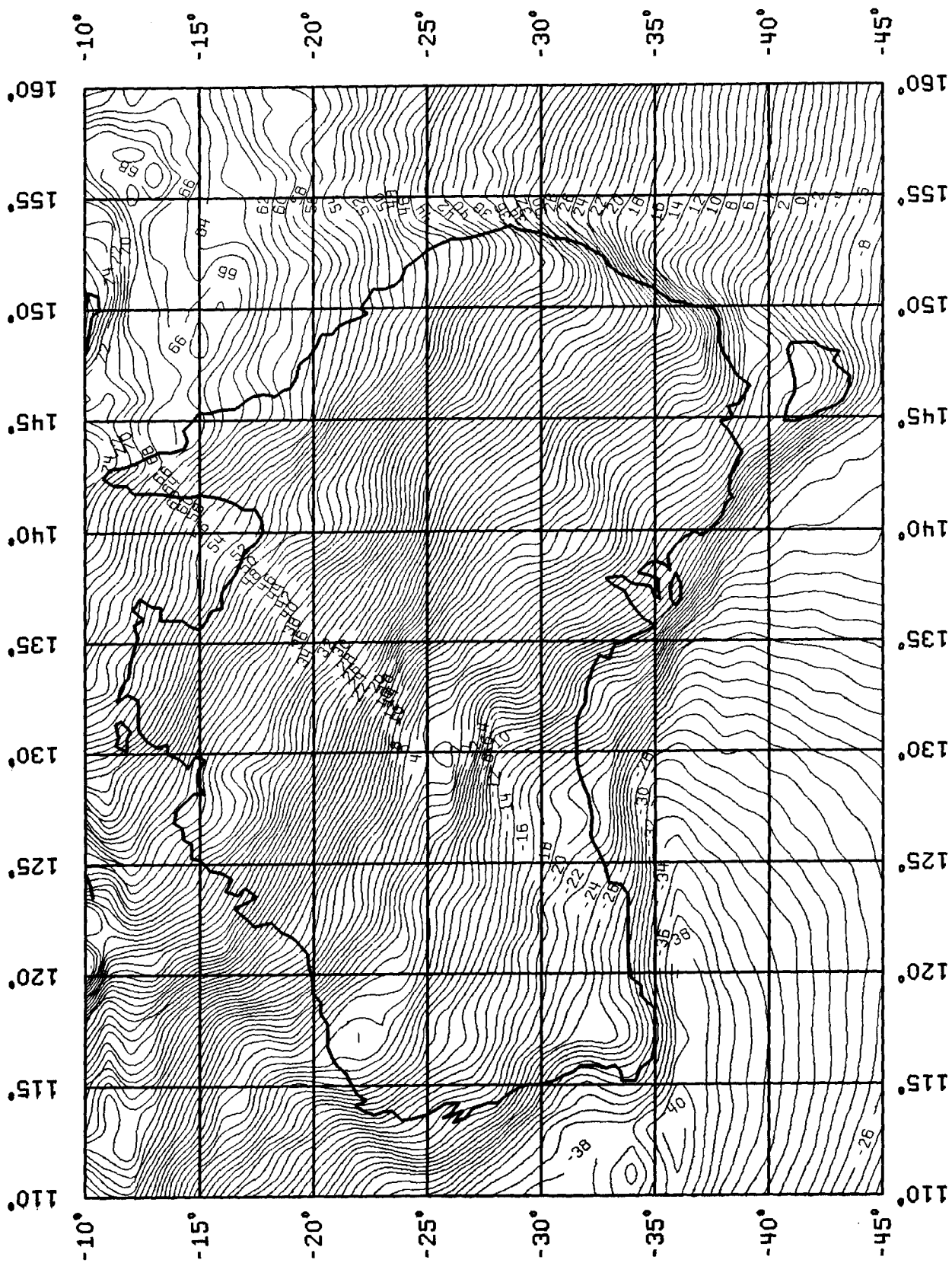
Figure 3.4. Geoid Undulations in a Geocentric System in the Australian Region Based on the OSU86F Potential Coefficient Model to Degree 180
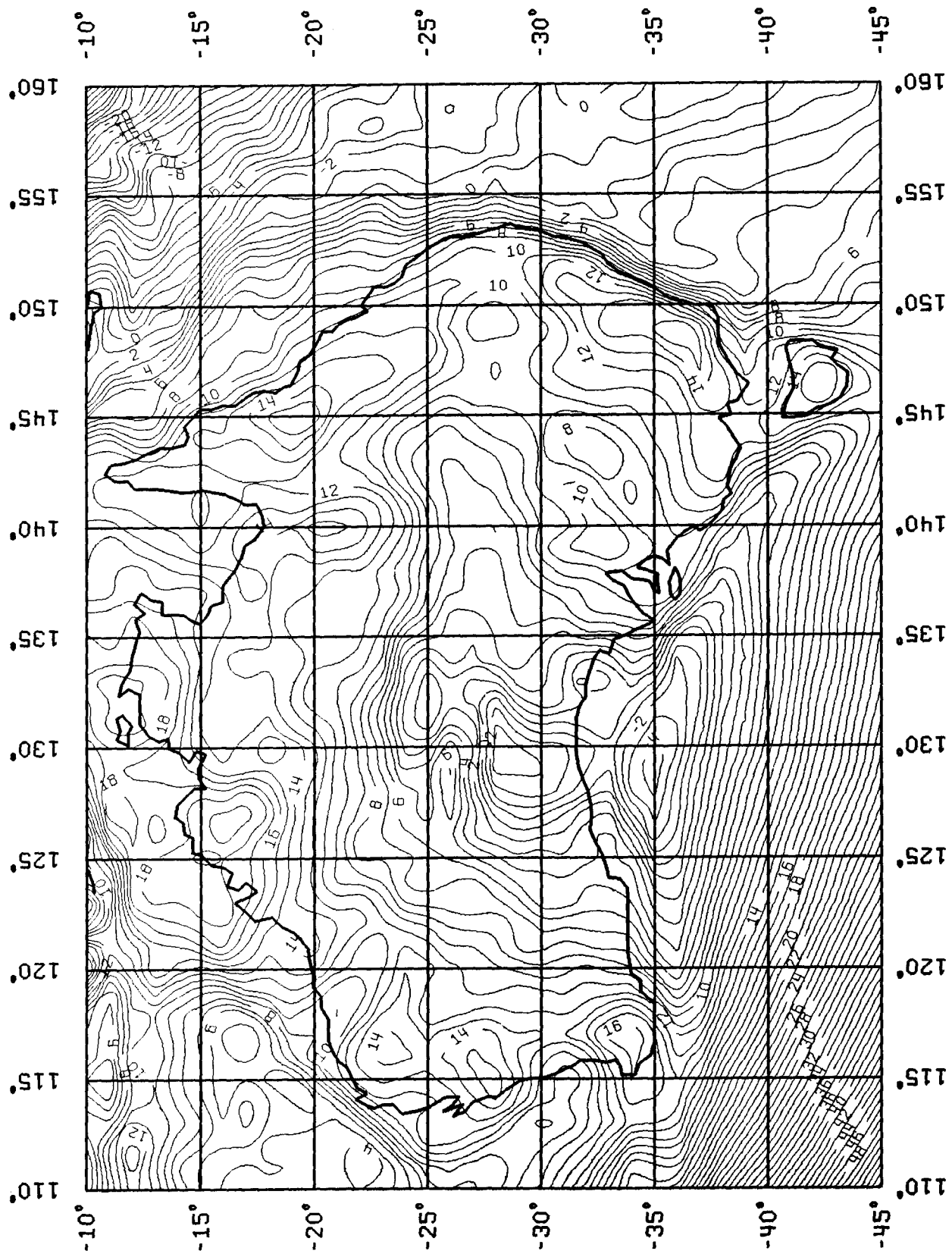
Figure 3.5. Geoid Undulations in a Local System (GMA82) in the Australian Region Based on the OSU86F Potential Coefficient Model to Degree 180

126

3.9.4. The Molodensky Correction to Development Computed Astrogeodetic Data

We now consider the conversion of astrogeodetic data computed on the basis of development procedures in the geodetic network to corresponding values that would apply if the projective reduction system had been used. Such a conversion would be used if we, for example, wished to process development computed astro geodetic undulations with the projective formulas for a best fitting ellipsoid determination, or if we were constructing an astrogeodetic undulation map that was to refer to a projective system using development deflections.

The procedures to be used were developed to Molodensky in 1944 with details given in Molodensky et al. (1962, p. 29). Practical implementation of the Molodensky correction procedures has been discussed by Fischer (1966) and Fischer et al. (1967). Obenson (1971) reviews the derivation and discussed the application.

3.9.5 The Arc Methods for Datum and Ellipsoid Parameter Determination

The previous discussion has considered geodetic data acquired in an areal sense. Prior to having such data it was of geodetic interest to analyze data acquired along an arc of a meridian (primarily) or along an arc of a parallel (sometimes). This information was primarily used to derive parameters of an ellipsoid that was a "best fit" to the geoid along the given arc. Although such procedures now are of historic interest, it is instructive to examine how such determinations were made.

To a certain approximation, the meridian arc, s, between two latitudes, φ_1, and φ_2 is:

$$s = \bar{a}\left(\bar{\varphi}_2 - \bar{\varphi}_1\right)\left(1 - \left(\frac{1}{4} + \frac{3}{4}\cos 2\,\bar{\varphi}_m\right)\bar{e}^2\right)$$

(3.54)

If we consider a similar distance, s_0, based on approximate parameters a and e^2 of the reference ellipsoid we may write:

$$s = s_0 + \frac{\partial s}{\partial a}\,da + \frac{\partial s}{\partial e^2}\,de^2$$

(3.55)

$$\frac{\partial s}{\partial a} = \left(\bar{\varphi}_2 - \bar{\varphi}_1\right)\left(1 - \left(\frac{1}{4} + \frac{3}{4}\cos 2\,\varphi_m\right)e^2\right)$$

$$\frac{\partial s}{\partial e^2} = -a\left(\bar{\varphi}_2 - \bar{\varphi}_1\right)\left(\frac{1}{4} + \frac{3}{4}\cos 2\,\varphi_m\right)$$

(3.56)

Now (3.55) can be written:

$$s = a\left(\bar{\varphi}_2 - \bar{\varphi}_1\right)\left(1 - ke^2\right) + da\left(\bar{\varphi}_2 - \bar{\varphi}_1\right)\left(1 - ke^2\right) - a\left(\bar{\varphi}_2 - \bar{\varphi}_1\right)kde^2$$

(3.57)

where

$$k = \frac{1}{4} + \frac{3}{4} \cos 2\, \varphi_m$$

We may also recall the radius of curvature in the meridian at the mean latitude of the arc. Thus computing this radius based on the original ellipsoid values we write:

$$M_m = \frac{a\left(1 - e^2\right)}{\left(1 - e^2 \sin^2 \varphi_m\right)^3} = a\left(1 - e^2\right)\left(1 + \frac{3}{2} e^2 \sin^2 \varphi_m\right)$$

(3.58)

Using:

$$\sin^2 \varphi_m = \frac{1}{2} - \frac{1}{2} \cos 2\, \varphi_m$$

we have:

$$M_m = a\left(1 - \left(\frac{1}{4} + \frac{3}{4} \cos 2\, \varphi_m\right) e^2\right) = a\left(1 - ke^2\right)$$

(3.59)

If we divide each side of (3.57) by (3.59) we have:

$$\frac{s}{M_m} = \left(\overline{\varphi}_2 - \overline{\varphi}_1\right) + \left(\overline{\varphi}_2 - \overline{\varphi}_1\right)\frac{da}{a} - \left(\overline{\varphi}_2 - \overline{\varphi}_1\right) kde^2$$

(3.60)

where the last term is approximate but sufficiently accurate for this derivation.

We now make two substitutions. First define: $(\varphi_2 - \varphi_1) = s/M_m$. In addition using the deflections at the two points we have:

$$\overline{\varphi}_1 = \Phi_1 - \xi_1$$

$$\overline{\varphi}_2 = \Phi_2 - \xi_2$$

(3.61)

so that (3.60) may be written:

$$\left(\varphi_2 - \varphi_1\right) = \left[\left(\Phi_2 - \xi_2\right) - \left(\Phi_1 - \xi_1\right) + \left(\overline{\varphi}_2 - \overline{\varphi}_1\right)\frac{da}{a} - \left(\overline{\varphi}_2 - \overline{\varphi}_1\right) kde^2\right]$$

(3.62)

or simplifying and noting to sufficient approximation in the coefficients of da, and de^2 that: $(\overline{\varphi}_2 - \overline{\varphi}_1) \approx (\varphi_2 - \varphi_1)$ we write:

$$\xi_2 = \xi_1 + \left[\left(\Phi_2 - \Phi_1\right) - \left(\varphi_2 - \varphi_1\right)\right] + \left(\varphi_2 - \varphi_1\right)\frac{da}{a} - \left(\varphi_2 - \varphi_1\right) k\, de^2$$

(3.63)

or rewriting:

$$\xi_2 = \xi_1 + \left(\varphi_2 - \varphi_1\right)\frac{da}{a} - \left(\varphi_2 - \varphi_1\right)k\,de^2 + \left(\Phi_2 - \varphi_2\right) - \left(\Phi_1 - \varphi_1\right) \tag{3.64}$$

Recall that φ_1 and φ_2 would be considered as the geodetic coordinates computed from the observed triangulation using the existing ellipsoid parameters. If we let:

$$L = \left(\Phi_2 - \varphi_2\right) - \left(\Phi_1 - \varphi_1\right) \tag{3.65}$$

$$p = \left(\varphi_2 - \varphi_1\right)$$

$$q = -\left(\varphi_2 - \varphi_1\right)k$$

we have

$$\xi_2 = \xi_1 + p\frac{da}{a} + q\,de^2 + \tag{3.66}$$

Now suppose we consider an arc that consists of n points for which we can write an equation of the type (3.66). At the first point, however, we specify the deflection as ξ_1. Then:

$$\xi_1 = \xi_1$$

$$\xi_2 = \xi_1 + p_1\frac{da}{a} + q_1\,de^2 + L_1$$

$$\xi_3 = \xi_2 + p_2\frac{da}{a} + q_2\,de^2 + L_2$$

$$\cdot$$
$$\cdot$$
$$\cdot$$

$$\xi_n = \xi_{n-1} + p_{n-1}\frac{da}{a} + q_{n-1}\,de^2 + L_{n-1} \tag{3.67}$$

Next we eliminate the reference to a preceding deflection in all equations of (3.67). For example, we may write:

$$\xi_3 = \xi_1 + \left(p_1 + p_2\right)\frac{da}{a} + \left(q_1 + q_2\right)de^2 + L_1 + L_2$$

Then:

$$\xi_4 = \xi_1 + \left(p_1 + p_2 + p_3\right)\frac{da}{a} + \left(q_1 + q_2 + q_3\right)de^2 + L_1 + L_2 + L_3$$

Or for a general term:

$$\xi_n = \xi_1 + \left(p_1 + p_2 + \dots p_{n-1}\right)\frac{da}{a} + \left(q_1 + q_2 + \dots q_{n-1}\right)de^2 + L_1 + L_2 + \dots L_{n-1} \tag{3.68}$$

If we insert into (3.68) the values of p, q, and L we have for the ith deflection:

$$\xi_i = \xi_1 + \left[\left(\varphi_2 - \varphi_1\right) + \left(\varphi_3 - \varphi_2\right) + \dots \left(\varphi_i - \varphi_{i-1}\right)\right]\frac{da}{a}$$

$$+ \left[-\frac{1}{4}\left(\varphi_2 - \varphi_1\right) - \frac{1}{4}\left(\varphi_3 - \varphi_2\right) + \dots \frac{1}{4}\left(\varphi_i - \varphi_{i-1}\right)\right.$$

$$\left. - \frac{3}{4}\sum_{k=2}^{i}\left(\varphi_k - \varphi_{k-1}\right)\cos 2\varphi_{mk}\right]de^2 + \left(\Phi_2 - \varphi_2\right) - \left(\Phi_1 - \varphi_1\right) + \left(\Phi_3 - \varphi_3\right)$$

$$- \left(\Phi_2 - \varphi_2\right)\dots + \left(\Phi_i - \varphi_1\right) - \left(\Phi_{i-1} - \varphi_{i-1}\right) \tag{3.69}$$

Noticing the large cancellation of terms we now have:

$$\xi_i = \xi_1 + \left(\varphi_i - \varphi_1\right)\frac{da}{a} + \left(-\frac{1}{4}\left(\varphi_i - \varphi_1\right) - \sum_{k=2}^{i}\left(\varphi_k - \varphi_{k-1}\right)\cdot\right.$$

$$\left. \cdot \frac{3}{4}\cos 2\varphi_{mk}\right)de^2 + \left(\Phi_1 - \varphi_i\right) - \left(\Phi_1 - \varphi_1\right) \tag{3.70}$$

Noting that there are three unknowns, ξ_1, da, and de^2 a least squares solution may be found by specifying (in analogy to (3.37)):

$$\Sigma\,\xi_i\,\xi_i = \text{a minimum} \tag{3.71}$$

A more exact equation than 3.70 has been given by Chovitz and Fischer (1956):

$$\xi_i = \xi_1 + \left(\varphi_i - \varphi_1\right)\frac{da}{a}\left[I\left(\varphi_i - \varphi_1\right) - \sum_{k=2}^{i}\left(\varphi_k - \varphi_{k-1}\right)\right]\cdot$$

$$\cdot \left[II\cos 2\varphi_{mk} - III\cos 4\varphi_{mk} + IV\cos 6\varphi_{mk}\right]de^2 + \left(\Phi_i - \varphi_i\right) - \left(\Phi_1 - \varphi_1\right) \tag{3.72}$$

where:

130

$$I = -\frac{1}{4} - \frac{7}{16}e^2 - \frac{17}{32}e^4$$

$$II = \frac{3}{4} + \frac{3}{4}e^2 - \frac{45}{64}e^4$$

$$III = \frac{3}{16}e^2 + \frac{9}{32}e^4$$

$$IV = \frac{3}{64}e^4$$

Equation (3.70) is valid for arc segments up to 3° in length.

If prime vertical deflection data is available along a parallel the following observation equation can be used (ibid, 1956) or see Zakatov (1962):

$$\eta_i = \eta_1 + \frac{da}{a}\left(\lambda_i - \lambda_1\right)\cos\varphi + de^2\left(\lambda_i - \lambda_1\right).$$

$$\cdot \frac{\sin^2\varphi\cos\varphi}{2}\left(1 + e^2\sin^2\varphi + e^4\sin^4\varphi\right)$$

$$+ \left(\Lambda_i - \lambda_1\right)\cos\varphi - \left(\Lambda_1 - \lambda_1\right)\cos\varphi$$

$$(3.74)$$

Chovitz and Fischer (1956) provide results from the analysis of meridian and parallel arcs from several different geographic areas. Also incorporated in their studies is the use of topographic-isostatic deflections of the vertical in a manner similar to that used in Section 3.9.1.1.

3.10 The Determination of the Parameters of a General Terrestrial Ellipsoid

An equipotential ellipsoid of revolution is a standard reference figure for problems of geometric and gravimetric geodesy. This ellipsoid and its gravity and gravity potential can be defined by four quantities. Although a number of different combinations can be used the most common are the following:

GM - the geocentric gravitational constant;
J_2 - the second degree zonal harmonic;
a - the equatorial radius;
ω - the angular velocity of the earth.

Knowing these quantities, other quantities such as the ellipsoid surface potential, U_0, normal gravity, γ, or the ellipsoid flattening, may be derived.

The parameters are said to define the general terrestrial ellipsoid when the values have some defined relationship with the Earth. For example, GM should be that of the actual Earth, including the mass of the atmosphere (in most cases). The J_2 harmonic should be that of the Earth as derived, for example, from the analysis of satellite orbital perturbations. And ideally the potential (U_0) on the surface of the ellipsoid should be equal to the potential

131

of the geoid (W_0). In addition the angular velocity (ω) should be equal to some average velocity of the Earth, defined over some time period such as one year.

In some determination of these parameters the values are solved separately while in other cases simultaneous solutions of some of the parameters can be made. We now consider a number of these methods using different types of space data. We first remark about the estimation of GM and J_2.

3.10.1 GM and J_2 Determinations

The value of GM can be determined from the analysis of laser ranging to the satellite or lunar laser ranging. Determination of GM are usually made simultaneously with other parameter (e.g., station positions) estimations. Estimates of GM from satellite solutions have been given by the Goddard Space Flight Center and The University of Texas as $3986004.40 \pm .02 \times 10^8$ m^3 s^{-2} (see Chovitz, 1987 and Bursa (1991)). Note that the values of GM include the mass of the atmosphere (3.5×10^8 m^3 s^{-2}) since the satellite orbits the Earth outside the atmosphere. We also note that the value of GM depends on the velocity of light, c, adopted as a fundamental constant. The standard value of c is: 299792458 ms^{-1}. As the value of c changes so will GM but in a non-linear way.

In some satellite solutions the value of GM is held fixed. This GM then scales the orbit and any station positions derived from the orbital analysis. If the GM is changed so will the station coordinates.

The value of J_2 is usually determined in a general geopotential modeling effort. Knowing J_2 the flattening of the reference ellipsoid can be derived having approximate values for the other three quantities (Heiskanen and Moritz, 1967, p. 73). The precise interpretation of J_2 must be considered in light of a special tidal consideration and considering the time variations of J_2 (i.e., $\dot{J}_2$).

The tidal concern arrises from the consideration of the permanent deformation of the Earth's surface by the gravitational attraction of the sun and moon. In some orbital analysis it is customary to remove the indirect deformation (see Moritz, (1979)) while in other solutions this deformation is included. Moritz (1979) shows that the relationship between the second degree zonal harmonics of the two cases is as follows:

$$J_2 \text{ (with indirect deformation)} = J_2 \text{ (without indirect deformation)} + 9 \times 10^{-9} \qquad (3.75)$$

In the definition of the Geodetic Reference Systems of the International Association of Geodesy it has been customary to calculate the flattening with a J_2 defined without the induced deformation. As pointed out by Engelis (1985) and others, for oceanographic applications, a flattening derived from the corrected J_2 is appropriate. This effect is small. The flattening for GRS80 is 1/298.2572 (derived from $J_2 = 1082\ 63 \times 10^{-8}$) while the flattening corresponding to the J_2 with the indirect deformation 1/298.2566. Chovitz (1988) gives a representative value of J_2 as $(1082\ 626 \pm 2) \times 10^{-9}$ which excludes the indirect deformation. More details on the role of permanent tidal deformation in geodetic parameter definition are found in Rapp et al. (1991a).

3.10.2 The Angular Velocity ω

The Earth's angular velocity fluctuates in a regular and irregular way. Fortunately the fluctuation is sufficiently small so that a value given to seven digits is suitable for the definition of the complete general terrestrial ellipsoid. Annual values of an average ω are

given in the IERS annual reports from which the values given in Table 3.5 have been taken.

Table 3.6

Year Average Angular Velocity of the Earth by Year

Year	ω
1978	$7292114.903 \times 10^{-11}$ rads^{-1}
1979	4.925
1980	4.952
1981	4.964
1982	4.964
1983	4.954
1984	5.019
1985	5.025
1986	5.043
1987	5.032
1988	5.035
1989	5.018
1990	4.983

The value adopted for GRS 80 is $7292\ 115 \times 10^{-11}$ radians s^{-1}.

3.10.3 The Equatorial Radius

The Earth's equatorial radius can be derived through a number of different techniques involving space related data. In most of these determinations the equatorial radius is not the only quantity being estimated. We now consider two such methods.

3.10.3.1 The Determination of the Equatorial Radius from Space Derived Station Positions

Assume we are given a set of rectangular coordinates defined in some space related reference frame. At each station we have the orthometric height (H) which we assume is refered to the same datum for all stations. (Because of vertical datum inconsistencies the assumption is not precisely correct.) In addition we define the equatorial radius and flattening of a reference ellipsoid so that the geometric height (h) of each station can be determined. We then calculate the geometric geoid undulation:

$$N_{GE} = h - H \tag{3.76}$$

Note that the value of N_{GE} will depend on the scale of the satellite system, as well as the dimensions of the adopted reference ellipsoid.

The value of N_{GE} can now be compared to the undulation computed from a set of potential coeffcients or from a combination of potential coefficient information and terrestrial gravity data as described by Despotakis (1987), for example. We deal here only with the potential coefficient determination of N through the following:

$$N = \frac{GM}{r\gamma} \sum_{n=2}^{\infty} \left(\frac{a}{r}\right)^n \sum_{m=0}^{n} \left(\overline{C}_{nm} \cos m\lambda + \overline{S}_{nm} \sin m\lambda\right) \overline{P}_{nm}(\cos \theta) \tag{3.77}$$

133

where:

r is the geocentric radius to the point on the geoid.
$\bar{C}$, $\bar{S}$ are fully normalized potential coefficients;
a is the scaling parameter associated with the potential coefficients;
θ, λ are the co-latitude and longitude.

The summation to infinity in (3.77) is replaced by a summation to a finite degree varying from 36 to 360. Note that N from (3.77) will refer to a geocentric system and an ellipsoid whose parameters are those of the general terrestrial ellipsoid. The values of the $C_{2n,0}$ will be referred to a specified flattening.

We now formulate a mathematical model under the assumption that the corrected geometric undulation and the gravimetric undulation should be the same. We use (2.89) to represent the charge in the undulation calculated via h from (3.76). We write (introducing the residual v):

$$N + v = N_{GE} + \cos \phi \, \cos \lambda \, \Delta x + \cos \phi \, \sin \lambda \, \Delta y$$

$$+ \sin \phi \, \Delta z - W\Delta a + \frac{a \, (1\text{-}f) \, \sin^2 \phi}{W} \, \Delta f \tag{3.78}$$

We have retained a df term on the right hand side of (3.78) for generality. Since N will also depend slightly on the flattening an iterative solution may be necessary.

Given a set of stations equation (3.78) constitutes an observation equation that is used in the formation of normal equations such that a least squares solution for the parameters in (3.78) are determined. Grappo (1980) describes solutions for the parameters of interest using Doppler derived positions for various stations configurations and various quantities constrained.

Rapp and Cruz (1986) used a technique where the translation parameters for the Doppler stations were defined through the BTS85 reference system. In addition the scale correction of 0.6 ppm was applied to the Doppler coordinates (See Section 3.6). The resultant equatorial radius using a global, but not uniformly distributed station set, was 6378136.2 m. An accuracy assesment is difficult because of the many factors affecting its estimates. A nominal value of ± 1 m is not unreasonable. Using the same Doppler station file Rapp et al (1991b), using the OSU91A potential coefficient model to degree 360, obtain an equatorial radius of 6378136.35, using the ITRF-0 transformation parameters.

3.10.3.2 The Determination of the Equatorial Radius from Satellite Altimeter Data

A satellite altimeter determines the distance from the satellite to the instantaneous ocean surface. Given the position of the satellite in a specified reference frame the sea surface height (ζ) with respect to a defined reference ellipsoid can be determined. Now define a sea surface height, ζ, with respect to an ideal, geocentric ellipsoid. Since the original sea surface heights (ζ_0) may not be in a geocentric system we introduce coordinate translations Δx, Δy, Δz. Since the original ellipsoid may not be the optimal one we define changes Δa and Δf to the approximate ellipsoid parameters. Using (2.89) we may transform ζ_0 to ζ as follows:

$$\zeta = \zeta_0 + \cos \phi \cos \lambda \, \Delta x + \cos \phi \sin \lambda \, \Delta y + \sin \phi \, \Delta z - W\Delta a + \frac{a \, (1\text{-}f)}{W} \sin^2 \phi \, \Delta f \tag{3.79}$$

We now define sea surface topography ζ_T, as the difference between the sea surface height, ζ_T and the ideal (geocentric) geoid undulation (also see Section 3.13.2):

134

$$\zeta_T = \zeta - N = \zeta_0 - N + \cos \phi \cos \lambda \, \Delta x + \cos \phi \sin \lambda \, \Delta y$$
$$+ \sin \phi \, \Delta z - W \Delta a + \frac{a(1-f)}{W} \sin^2 \phi \, \Delta f \tag{3.80}$$

We now define our adjustment model as one where the sum of the squares of ζ_T becomes a minimum. With this interpretation, (3.80) becomes an observation equation which, given estimates of N from (for example) potential coefficient models, can be used to form normal equations and then the estimation of the parameters Δx, Δy, Δz, Δa and Δf. With these values new ellipsoid parameters may be determined. This new ellipsoid will have an overall least squares (via ζ_T) fit to the sea surface. Although N may not be precisely known on a point by point basis the analysis taken over the whole oceans will average out missing high frequency effects in N. The analysis suggested in this section has been discussed by West (1982), by Engelis (1985) and others.

Rizos (1980) and Engelis (1985) have discussed other ways in which the ellipsoid may be defined. One procedure is to define the ellipsoid such that the mean sea surface topography is zero over the oceans. Specifically we introduce the condition that M (ζ_T) = 0 where M is an averaging operator over the oceans. Taking Δx, Δy, Δz, and Δf to be zero, and W to be one in (3.80) our condition implies

$$\Delta a = M (\zeta_0 - N) \tag{3.81}$$

where high frequency effects are assumed to average to zero.

Engelis (1985) described calculations using the (3.80) observation equation and also (3.81) using different models for N. Using Seasat altimeter data as adjusted by Rowlands (1981) using the original JPL geophysical data records, Engelis found (from both (3.80) and (3.81) an equatorial radius of 6378136.0 m. This value is dependent on the scale of the Seasat orbits which depends on the GM used in the orbit calculation.

Global modeling efforts incorporating Seasat data have been described by Marsh et al. (1989). The equatorial radius computed from the Marsh et al. analysis was 6378136.14 m which is quite close to results from the Doppler station analysis described in Section 3.10.3.1. Denker and Rapp (1990) found an equatorial radius of 6378136.4 m based on Geosat altimeter analysis while Rapp et al (ibid, p. 27) report on equatorial radius of 6378136.38 m.

3.11 Other Considerations on Ellipsoid Determination

Bursa and Sima (1985) have discussed ways in which the dimensions of celestial bodies can be determined given the harmonic coefficients that represent the gravitational potential of the body. The authors define an equipotential boundary surface S with a geocentric radius-vector ρ_s. Let ρ_Σ be the radius vector to the actual topographic surface. The surface S should be chosen such that

$$\iint_S \left(\rho_s - \rho_\Sigma \right)^2 dS = a \text{ minimum} \tag{3.82}$$

The ρ_s is expressed in terms of a parameter R_0 (= GM/W_0) and dimensionless coefficients that are assumed known.

135

3.12 Future Determinations

3.13 Vertical Datums

The height above the ellipsoid, the geodetic height, although rigorously defined, is not the height conventionally used for mapping. Instead, it is common practice to introduce, as a vertical reference datum a surface that is associated in some average way with mean sea level or the mean ocean surface. Heights, now called orthometric heights, are measured with respect to this mean sea level surface. There are different ways in which mean sea level may be defined and determined, and various ways in which heights, given with respect to this surface can be defined. The following gives more specific details on the definition nad realization of vertical reference systems.

3.13.1 The Geoid

A fundamental surface of gravimetric geodesy, and of high importance to vertical reference systems, is the geoid. The geoid is a specific equipotential surface of the earth's gravity field. In this discussion we will adopt a geoid definition that excludes the direct effects of the Sun and Moon although for some applications (e.g., in oceanography) it is appropriate to consider such effects. The gravity potential on the surface of the geoid is defined to be W_0. By definition, there should be only one geoid, although the estimation of the location of the geoid yields many values. The geoid can be located with respect to a reference ellipsoid through geoid undulations, N. These undulations can be determined from knowledge of the gravity field of the earth. Calculation of N with such data can only be done up to some constant value which is known within one meter. Since the geoid is determined by variations in the gravity field, the geoid is an irregular surface with a maximum positive geoid undulation of 78 meters and a maximum negative undulation of -108 m.

The gravity potential on the geoid is not directly observable. However, it can be computed given knowledge of the Earth's gravity field, and the position of a point on the geoid. One could write:

$$W_0 = f\left(r, \psi, \lambda, C_{nm}, S_{nm}, GM, \omega\right)$$

(3.83)

Here, r, ψ, λ are the geocentric radius, geocentric latitude, and longitude of a point on the geoid. C_{nm} and S_{nm} are potential coefficients of degree n and order m in a spherical harmonic expansion of the Earth's gravitational field (see eq. (3.77)); GM is the gravitational constant times the mass of the Earth, and ω is the rotational velocity of the Earth. The calculation of W_0 is hindered by the lack of knowledge of the physical location of the geoid (so that r, ψ, λ cannot be accurately determined) and lack of knowledge of the potential coefficients. If one could identify points on the geoid, on a global basis, the averaging of many W_0 values could lead to an estimate of the potential of the geoid where random errors have been reduced but systematic effects remain.

3.13.2 The Mean Sea Level

Mean sea level is a surface defined by averaging sea level over time and in some cases spatially. Tide gauge stations are the principal source of information on sea level. Such stations continually monitor the rise and fall of sea level. The largest signature will be that of tides. Averaging, at a point, over appropriate time intervals (one year to 18 years) yields

an average location of local mean sea level. Mean sea level is not constant as it can be affected by ice cap melting, wind variations and changing ocean current (e.g., El Ninó) patterns. The determination of mean sea level in coastal regions may be sensitive to the location of the site. For example, the location near a river discharge to the ocean could give unreliable readings in time of drought or excessive rainfall.

Mean sea level is not an equipotential surface. This is due to the fact that currents exist in the ocean where water will flow from one equipotential level to another. The geoid can now be defined as the equipotential surface that has the same physical location as a global mean sea surface when tidal, atmospheric and current effects are removed. The separation between the mean sea level and the geoid is called sea surface topography (SST). Sea surface topography can be estimated from oceanographic information (such as water density, salinity, pressure, current flow, etc.) in conjunction with assumptions on a level of no motion in the oceans. However, its determination on a global basis is complex due to the need for substantial information that is difficult to collect on a large, ocean-wide scale. The estimation of long-term sea surfce topography has been discussed by Lisitzen (1974), Levitus (1982), and others. The estimates of SST made by these authors indicate the separation between the geoid and mean sea level to be on the order of ±1m. At this time there is no uniform agreement on the deep reference levels in the oceans to be used in SST computations. In addition, SST is especially difficult to compute in the coastal waters where tide gauge measurements are made. Future prospects for SST determination would be enhanced with improved gravity field information from special satellite missions (e.g., using a gradiometer) and through the direct measurements to the ocean surface from satellite altimeters.

To summarize this discussion, consider Fig. 3.6 which shows a meridian section of the ellipsoid and various surfaces of interest. The ellipsoidal (h) and orthometric (H) heights are shown. The orthometric height is formally measured along the curved vertical between the point P and the reference equipotential surface, the geoid. Fig. 3.7 portrays the information at a tide gauge station and its connection to a reference benchmark.
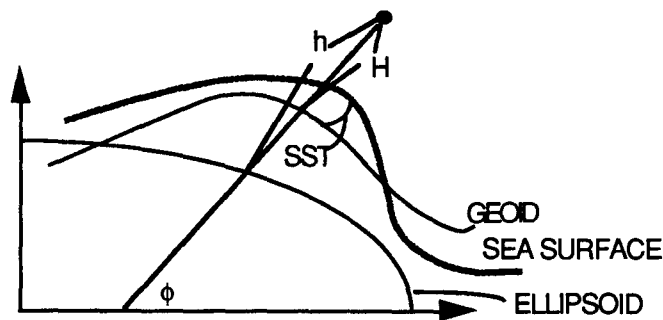

Figure 3.6. Location of ellipsoid height (h), orthometric height (H), geoid undulation (N), and sea surface topography (SST).
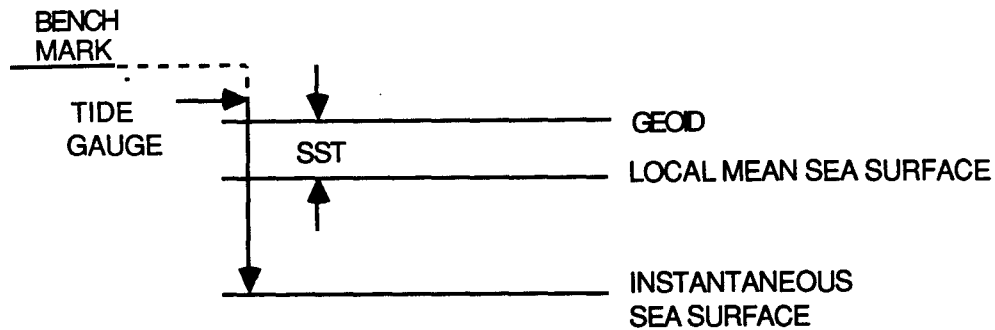
Figure 3.7. Measurements at a tide gauge site.

3.13.3 Determination of Orthometric Heights

In order to determine orthometric heights, we must determine a reference surface from which these heights are measured. Ideally, this surface should be the same for the whole world; and, therefore, conceptually the geoid is the appropriate surface. Since it is essentially impossible to physically determine the geoid, mean sea level is used. There are several ways in which MSL is introduced into a vertical network. The simplest procedure is one in which mean sea level, at one site, is transferred to a nearby fundamental benchmark. The elevation of this benchmark is found by measuring the small elevation difference between the MSL determination at the tide gauge and the benchmark. This benchmark then becomes the fundamental point of the vertical network. That is, the reference equipotential surface is that surface passing through the benchmark. This surface is traceable to the MSL at the nearby tide gauge site.

Starting from this point, vertical control measurements, consisting of leveled height differences and gravity measurements, are made. With this information, the potential difference or orthometric height, with respect to the reference surface can be determined.

Consider the determination of the elevation of MSL at a site some distance form the fundamental tide gauge. The elevation at this new point would not be expected to be zero because we have previously noted that mean sea level at various locations does not define an equipotential surface.

In contrast to using a single tide gauge station to define the fundamental reference surface, an alternate technique incorporates multiple tide gauge determinations of mean sea level. In this case, a vertical network is adjusted to maintain consistency betweeen the various loops of the network. In addition, constraints are imposed on the adjustment to force the equivalent of a zero elevation at each of the local mean sea levels. This procedure has the advantage that elevations near coast line will be close to zero. However, it has the disadvantage that the datum surface is no longer associated with a single station. In fact, the reference surface is no longer an equipotential surface due to the warping necessary to uphold the constraints of the adjustment.

Another procedure for beginning a vertical datum is to carry out a preliminary adjustment with one station held fixed. At the completion of the adjustment, the heights of the local mean sea levels throughout the network are examined. A mean discrepancy is computed and applied to the station originally held fixed. This procedure forces the

138

average elevation of all local MSL determinations to be zero. It does leave the reference surface unattached to any specific station.

We should finally emphasize that this discussion has ignored the time variations of mean sea level determinations. As noted earlier, mean sea level can change with time so that it is appropriate to associate a vertical datum with a mean sea level at a specified epoch. An alternative is to refer the datum to a defined elevation at a specified datum benchmark. Another complication relates to the motion of the crust, which for this discussion is assumed fixed.

With this discussion in mind, it is clear that there will be many vertical datums in the world. Each datum may be traced to some local mean sea level determination, or to some fixed reference point, or to some implicit surface defined by an adjustment procedure. Each country (or region) may have its own datum.

3.13.4 Specific Vertical Datums

The vertical datum now used in the United States is called the National Geodetic Vertical Datum of 1929 (NGVD 29). The date reflects the time at which the leveling data of the United States and Canada were adjusted. At that time there was 75,159 km of leveling in the U.S. and 31,565 km in Canada. The adjustment was carried out by holding local mean sea level at zero elevation at 21 tide gauge stations in the U.S. and five in Canada. This procedure led to a datum that is warped to local mean sea level. The development of a new vertical network for North America will take into account the fact that local mean sea levels do not fall on the same equipotential surface.

The height system in Australia is called the Australian Height Datum (AHD). It was developed in 1971 through the adjustment of 97,320 km of leveling, holding mean sea level, for the 1966-68 epoch, fixed at zero at 30 tide gauge sites around the coast of Australia. This procedure was analogous to that used in the development of NGVD 29. The reference surface for the AHD is not an equipotential surface, but a surface warped to mean sea level around the continent.

In Europe one finds two types of vertical datums. The first type is that associated with a particular country or region. This type has grown from the historical need for height information. The second type of datum is that associated with the development of the United European Levelling Net (UELN).

In 1973, a new subcommission of the International Association of Geodesy was formed with the task of continuing the work of prior groups involved with the United European Levelling Net (Ehrnsperger et al. 1982). The purpose of the Net was to combine all leveling data from the European countries into one consistent system. For datum definition purposes, a single point was held fixed at a specified elevation (or geopotential number). This station is No. 4019, Normal Amsterdam Piel (NAP). The datum for the UELN-73 network is the equipotential surface which is a defined potential below the surface that passes through NAP. Since no other constraints have been imposed at tide gauge stations, the UELN-73 datum is free of internal distortions caused by the departure of local mean sea levels from the same equipotential surface.

Arur and Baveja (1984) have discussed the vertical datum for India. The First Level Net of India was adjusted in 1909 holding the elevation zero at mean sea level at nine tide gauge stations. Later preliminary adjustments between local mean sea levels between the east and west coast of the country. It was then decided to define the vertical datum origin at

a single tide gauge station in Bombay based on a local mean sea level determined from 38 years of observations.

It should be clear from these discussions that most countries have adopted varying procedures for the definition of their vertical datums. Such procedures make it impossible to have a vertical datum that is truly global in nature at this time. Fortunately, since vertical datums are tied to local mean level, the inconsistency of the reference levels should not exceed 2 meters, which is the range of sea surface topography described by Lisitzen or Levitus.

3.14 Future Vertical Datums

The above discussion indicates the variety of vertical datums that exist in the world. This leads one to ask two questions:
1) Is it possible to determine the height (or potential) difference between two or more vertical datums?
2) Is it possible to construct a world vertical datum? A general discussion of possible solutions to these questions is found in (Rapp, 1983).

The calculation of a potential difference between two datums has been discussed by Colombo (1985), Hajela (1985), and others. In these discussions, several different types of information are brought together. This information includes the geocentric position of fundamental benchmarks as derived from laser tracking of satellites; global gravity field models; detailed gravity surveys within several hundred kilometers of the benchmarks whose geocentric coordinates are known; and potential difference determinations between the geocentrially positioned benchmarks. The simulation studies of Hajela indicated that it would be possible to determine the height difference between Europe and the United States to an accuracy of about ± 0.5 m. Since this is about the accuracy we could obtain with existing oceanographic data, it appears that we need to wait for more accurate gravity field models to determine the height difference more accurately.

Of great future interest is the need for a common surface that is ultimately accessible to all countries for vertical reference purposes. Cartwright (1985) has suggested that such a surface may be a surface of no motion in the oceans. Such a surface may exist at locations where the pressure reaches some defined value. One such surface might be the 4000 decibar surface. Using oceanographic measurements, it is possible to calculate the sea surface topography in the open oceans with respect to this reference surface. This information is then brought into the tide gauge stations through stellite altimeter measurements and geostrophic leveling using current measuring devices.

This proposal would enable the local mean sea level heights to be converted to refer to the deep pressure surface. This method could be an important step in defining a world vertical datum. An error analysis of the procedure needs to be done to verify that the accuracy would be substantially better than that which could be accomplished using ellipsoidal heights and geoid undulations.

REFERENCES

Allan, A.L., and C. Hudson, The Figure of the Earth, Survey Reviews, 29,225, p. 120 1987.

Allman, J., Readjustment of the Australian Primary Geodetic Network Using a Combination of Terrestrial and Satellite Doppler Observations, Annales de Geophysique, t. 37, fasc. 1, 55-58, 1981.

Allman, J.S., and C. Veenstra, Geodetic Model of Australia 1982, Technical Report 33, Division of National Mapping, Dept. of Resources and Mapping, Canberra, 1984.

Anderle, R., Accuracy of Mean Earth Ellipsoid Based on Doppler, Laser, and Altimeter Observations, NSWC TR80-4, Naval Surface Weapons Center, Dahlgren, Virginia 22448, 1980.

Arur, M.G., and S.D. Baveja, Status of the Establishment of Vertical Datum and the Level-net in India, umpublished manuscript, Survey of India, Dehra Dun, 1984.

Ashkenazi, V., S.A. Crane, and J. Williams, The OSGB Scientific Network: Combination of Satellite and Terrestrial Data; Annales de Geophysique, t. 37, fasc. 1, 49-54, 1981.

BIH, Annual Report for 1987, Bureau International De L'Heure, Annual Report for 1987, Paris, July, 1988.

Bomford, G., Geodesy, Clarendon Press, Oxford, 1980.

Bossler, J., Geodesy Solves 900,000 Equations Simultaneously, EOS, (Trans. AGU), 68, p. 569 June 9, 1987.

Boucher, C. and Z. Altamimi, Towards an Improved Realization of the BIH Terrestrial Frame, in Proc. of the Int. Conference on Earth Rotation and the Terrestrial Reference Frame, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, OH 43210-1247, 1985.

Boucher, C and Z. Altamimi, Complement of Analysis of BIH Terrestrial System for 1986, in Annual Report of the BIH, Paris, 1987

Boucher, C. and Z. Altamimi, Complement of Analysis of BIH Terrestrial System for 1987, in Annual Report for 1987, BIH, Observatoire de Paris, Paris, 1988.

Boucher, C. and Z. Altamimi, The Initial IERS Terrestrial Reference Frame, IERS Technical Note 1, Observatoire de Paris, June 1989.

Boucher, C. and Z. Altamimi, ITRF89 and other Realizations of the IERS Terrestrial Reference System for 1989, IERS Technical Note 6, Observatoire de Paris, April 1991.

Bursa, M., and Z. Sima, Dimension Parameters of the Moon, Mars, and Venus, manuscript, Oct. 1988.

Bursa, M., Parameters of Common Relevance of Astronomy, Geodesy, and Geodynamics, Report of IAG Special Study Group 5-100, IAG/IUGG Meeting, Vienna, 1991.

Cartwright, D., A Practical Oceanographic Approach to the Problem of Defining a World Vertical Datum in Proc. North American Vertial Datum 85, 155-160, Nat. Geod. Info. Ctr., NOAA, Rockville, MD 1985.

Chovitz, B., and I. Fischer, A New Determination of the Figure of the Earth from Arcs, Trans. Amer. Geophy. Union, 37, 534-545, 1956.

Chovitz, B., Parameters of Common Relevance of Astronomy, Geodesy, and Geodynamics, in Bulletin Geodesique, (The Geodesist's Handbook), 62, 359-367, 1988.

Colombo, O., Levelling with the Help of Space Techniques, in Proc., North American Vertical Datum 85, 137-144, Nat. Geod. Info. Ctr., NOAA, Rockville, MD, 1985.

Denker, H., and R. Rapp, Geodetic and Oceanographic Results from the Analysis of One Year of Geosat Data, J. Geophysical Research (c), August 1990.

Despotakis, V., Geoid Undulation Computations at Laser Tracking Stations, Report No. 383, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, p. 137p, 1987.

DMA, Department of Defense World Geodetic System 1984, Its Definition and Relationship with Local Geodetic Systems, DMA TR 8350.2, DMA, Washington, DC 20305-3000, p. 121, September, 1987.

Dracup, J.F., Use of Doppler Positions to Control Classical Geodetic Networks, National Ocean Survey Reprints 1975, National Geodetic Survey, Rockville, MD 20852, 1975.

Engelis, T., Global Circulation from SEASAT Altimeter Data, Marine Geodesy, 9, 45-69, 1985.

Ehrnsperger, W., Report on the Computations of the "Rough Solution" of Retrig, Phas III, in publication No. 15, IAG Subcommission for the New Adjustment of the European Triangulation, Deutshi Geodatische Kommission, Number 22, 1985

Ehrnsperger, W., et al. Activities of the RETrig Computing Center Munich since the last Symposium in May 1979, in Publication No. 14, IAG Subcommission for the New Adjustment of the European Triangulation, Deutsches Geodatisches Forscheungsinstitut, München 22, 1982.

Ehrnsperger, W., J.J. Rok, and Ivan Mierlo, Status and Provisional Results of the 1981 Adjustment of the United European Levelling Network - UELN73, in Proc. Int. Symposium on Geodetic Networks and Computations, Deutsche Geodatische Kommission, Reihe B, No. 258/II, 1982.

Ehrnsperger, W., Computations for the Final Solution of RETrig, paper submitted to the XIX General Assembly of IUGG, Vancouver, 1987.

Ehrnsperger, W., Computations of the Final Solution of RETrig (ED87), presented at the meeting of the RETrig Subcommission (IAG Section I), Lisboa, Portugal, 1988.

Feissel, M., Realization of the BIH System in 1984, in Annual Report of the BIH, Paris, 1985.

Fischer, I., The Hough Ellipsoid (or the Figure of the Earth from Geoidal Heights), Bulletin Geodesique, 54, 45-52, 1959.

Fischer, I., A Revision of the Geoid Map of North America, J. Geophys. Res., 71, 4905-4908, 1966.

Fischer, I., et al., Geoid Chart of North and Central America, Army Map Service Technical Report No. 62, Washington, D.C., 1967.

Fischer, I., and M. Slutsky, A Preliminary Geoid Chart of Australia, The Australian Surveyor, December, 1967.

Grafarend, E., and B. Richter, The Generalized Laplace Condition, Bulletin Geodesique, 51, 287-293, 1977.

Grappo, G., Determination of the Earth's Mean Equatorial Radius and Center of Mass from Doppler Derived and Gravimetric Geoid Heights, manuscripta geodaetica, 5, 201-216, 1980.

Grushinskiy, N.P., The Theory of the Figure of the Earth, 1963, translated from Russian 1969, AD694748.

Hajela, D.P., Accuracy Estimates of Intercontinental Vertical Datum Connections, in Proc., North American Vertical Datum 85, 145-154, Nat. Geod. Info. Ctr., NOAA, Rockville, MD, 1985.

Hayford, J., The Figure of the Earth and Isostasy from Measurements in the United States, U.S. Dept. of Commerce (Coast and Geodetic Survey), 1909, Washington, D.C.

Hayford, J., Supplementary investigation of the figure of the earth and isostasy, U.S. Dept of Commerce (Coast and Geodetic Survey), 1910, Washington, D.C.

Heiskanen, W., and H. Moritz, Physical Geodesy, W.A. Freeman, San Francisco, 1967.

Heiskanen, W.A. and F.A. Vening Meinesz, The Earth and Its Gravity Field, New York, McGraw Hill, 1958.

Hornik, H. and E. Reinhart, Report on the Readjustment of the European Triangulation Network, Retrig-Phase II-1979;in Publication No. 13, IAG Subcommission for the New Adjustment of the European Triangulation (RETrig), "The European Datum 1979 (ED79), Deutsches Geodatisches Forschungsinstitut, München 22, 1980.

Hosmer, G., Geodesy, John Wiley & Sons, New York, second edition, 1930.

Hothem, L., T. Vincenty, R. Moose, Relationship Between Doppler and Other Advanced Geodetic Space Measurements Based on Global Data, Proc. 3rd Geodetic Symposium on Satellite Doppler Positioning, Vol. 1, 109-128, New Mexico State University, 1982.

Hotine, M., Mathematical Geodesy, ESSA Monograph 2, 1969, COM-74-11127.

IERS, A brochure assembled for the IERS by the National Geodetic Survey and the U.S. Naval Observatory, 1988.

Jones, H.E., Geodetic Datums in Canada, Canadian Surveyor, Vol. 27, No. 3, pp. 195-207, Sept., 1973.

Kelm, R., European Datum 1987: Aspects to the Final Report of the IAG-Subcommission "Readjustment of the European Triangulation Network", paper presented at the XIX IUGG General Assembly, IAG August, 1987, Vancouver, Canada.

Kohold, F., Introduction by the President of the RETRIG Subcommission, in Publication No. 13, IAG Subcommission for the New Adjustment of the European Triangulation (RETrig), "The European Datum 1979 (ED79), Deutsches Geodatisches Forschungsinstitut, München 22, 1980.

Levitus, S., Climatological Atlas of the World Ocean, NOAA, Geophysical Fluid Dynamics Laboratory, Professional Paper 13, Rockville, MD 1982.

Lisitzin, E., Sea Level Changes, Elsevier Science Publ. Co., Amsterdam and N.Y., 1985.

Malys, S., Dispersion and Correlation Among Transformation Parameters Relating Two Satellite Reference Frames, Report No. 392, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, p. 95, 1988.

Marsh, J., et al. Dynamic Sea Surface Topography Gravity, and Improved Orbit Accuracies from the Direct Evaluation of SEASAT Altimeter Data, NASA Technical Memorandum 100735, Goddard Space Flight Center, Greenbelt, MD, 1989.

Mather, R., and C. Rizos, The Shape of Global Mean Sea Level from GEOS-3 Altimetry, Australian Journal of Geodesy, Photogrammetry and Surveying, NO. 31, Dec., 1979.

Molodensky, M.S., V.F. Eremeev, and M.I. Yurkina, Methods for study of the external gravitational field and figure of the earth, Transl. from Russian (1960), by Israel Program for Scientific Translations, National Science Foundation, OTS 61-31207, 1962.

Moritz, H., Report of Special Study Group No. 5.39 of IAG: Fundamental Geodetic Constants, Bulletin Geodesique, 118, 398-408, 1975.

Moritz, H., The Definition of a Geodetic Datum, in proced., Second Int. Symp. on Problems Related to the Redefinition of North American Geodetic Networks, Superintendent of Documents, U.S. Govt. Printing Office, Washington, D.C. 20402, (No. 003-017-0426-1), 1978.

Moritz, H., Report of Special Study Group No. 5.39 of IAG: Fundamental Geodetic Constants; presented at IAG, IUGG meeting, Canberra, 1979.

Mueller, I.I., Proposed Optical Network for the National Geodetic Satellite Program, Dept. of Geodetic Science Report No. 71, The Ohio State University, Columbus, 1966.

Mueller, I., Review of Problems Associated with Conventional Geodetic Datums, presented at the International Symposium on Problems Related to the Redefinition of North American Geodetic Network, Fredericton, N.B., Canada, 1974.

Mueller, I., Reference Coordinate Systems and Frames: Concepts and Realization, Bulletin Geodesique, 59, 181-188, 1985.

Mueller, I., Reference Coordinate Systems: An Update, Report No. 394, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, 58 pp, 1988.

NASA, NASA Directory of Observation Station Locations, Goddard Space Flight Center, Greenbelt, Maryland, Vol. 1, Third Edition, November, 1973 (N74-22890/9).

Obenson, G., Reduction of Geodetic Coordinates of a Triangulation Scheme Computed by the Development Method to their Equivalent Projection Method Values, Survey Review, No. 162, p. 177-182, October, 1971.

Pick, M., J. Picha, and V. Vyskočil, Theory of the Earth's Gravity Field, Elsevier Scientific Publishing Co., New York, 1973.

Rais, J., A Comparison of the Projective and the Development Methods for Computing the Best Fitting Reference Ellipsoid, M.Sc. thesis, The Ohio State University, Columbus, 1969.

Rapp, R.H., Methods for the computation of geoid undulations from potential coefficients, Bulletin Geodesique, No. 101, September, 1971a.

Rapp, R.H., The Need and Prospects for a World Vertical Datum in Proc., IAG Symposia, IAG, XVIII General Assembly of IUGG, Vol. II, 432-446, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, 1983.

Rapp, R.H., and J. Cruz, The Representation of the Earth's Gravitational Potential in a Spherical Harmonic Expansion to Degree 250, Report No. 372, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, 1986.

Rapp, R.H., et al., Consideration of Permanent Tidal Deformaton in the Orbit Determination and Data Analysis for the Topex/Poseidon Mission, NASA TM 100775, Goddard Space Flight Center, Greenbelt, MD, 1991a.

Rapp, R.H., Y. M. Wang, N.K. Pavlis, The Ohio State 1991 Geopotential and Sea Surface Topography Harmonic Coefficient Models, Report 410, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, OH 43210, 1991b.

Rizos, C., The Contribution of Sea Surface Topography Studies to the Understanding of Global Ocean Surface Circulation, Aust. J. Geod., Photo., and Surveying, No. 32, 59-92, June 1980.

Rowlands, D., The Adjustment of SEASAT Altimeter Data on a Global Basis for Geoid and Sea Surface Height Determination, Report No. 325, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, 1981.

Schwarz, C. and E. Wade, The North American Datum of 1983: Project Methodology and Execution; Bulletin Geodesique, 64, 28-62, 1990.

Schwarz, C. (editor), North American Datum of 1983, NOAA Profesional Paper NOS2, National Geodetic Survey Rockvill, MD, 260 pp., December 1989.

Seppelin, T., The Department of Defense World Geodetic System 1972, presented at the International Symposium on Problems Related to the Redefinition of North American Geodetic Networks, the University of New Brunswick, Fredericton, N.B., Canada, 1974.

Seppelin, T., The Department of Defense World Geodetic System 1972, The Canadian Surveyor, Vol. 28, No. 5, 496-506, 1974b.

Steeves, R., Mathematical models for use in the readjustment of the North American geodetic networks, Technical Report No. 1, Geodetic Survey of Canada, Ottawa, 1984.

Vaniček, P., and D.E. Wells, Positioning of Geodetic Datums, University of New Brunswick, Fredericton, N.B., Canada, 1974. (Paper presented at the International Symposium on Problems Related to the Definition of North American Geodetic Networks).

Vincenty, T., Rotation Unknown in Adjustments of Large Geodetic Networks, manuscript, National Geodetic Survey, Rockville, MD, 12.27, 1980.

Vincenty, T., Methods of Adjusting Space Systems Data and Terrestrial Measurements, Bulletin Geodesique, 56, 231-241, 1982.

West, G., Mean Earth Ellipsoid Determined from SEASAT Altimeter Observations, J. Geophysical Research, 87, 5538-5540, 1982.

Wolf, H., Scale and Orientation in Combined Doppler and Triangulation Nets, Bulletin Geodesique, 54, 45-53, 1980.

Wolf, H., Stochastic Aspects in Combined Doppler and Trangulation Networks, Bulletin Geodesique, 56, 63-69, 1982a.

Wolf, H., Minutes on the Combining Procedure of Doppler Observations with the RETRIG's Phase III, Publication No. 14, 212-221, Subcommission for the New Adjustment of the European Triangulation (RETrig), L. Asplund, ed., Munchen, 1982b.

Wolf, H., Alternate Procedure of how to Combine Doppler Measurements with the RETrig Adjustment Phase III, Publication No. 14, 115-128, Subcommission for the New Adjustment of the European Triangulation (RETrig), L. Asplund, ed., Munchen, 1982c.

Xu, P., and R. Rummel, A Quality Investigation of Global Vertical Datum Connection, Number 34, Publications on Geodesy, Netherlands Geodetic Commission, Delft, 1991.

4. Fundamentals of Three Dimensional Geodesy

4.1 Introduction

The classical techniques of geodetic control are divided into the solution of two separate networks: the horizontal control network and the vertical control network. The object of the horizontal control network was to establish the geodetic coordinates of points on a reference ellipsoid that have a correspondence to points on the surface of the earth. In order to do this, measurements were made on the surface of the earth and reduced, in theory, to corresponding values with respect to the ellipsoid. This required corrections for skew normals, normal section to geodesic, and most important, deflections of the vertical, for the angular or direction measurements. In addition it was required that all measured distances be reduced to the ellipsoid. This latter requirement yielded a need for the orthometric heights of the baselines, and the geoid-ellipsoid separations, or astro-geodetic undulations. In practice the orthometric heights can be determined accurately, but the astro-geodetic undulations are not known accurately until an accurate horizontal control network is available. We thus can see that the processing data observed on the surface of the earth to obtain geodetic control is a difficult one requiring at many stages certain approximations and assumptions.

Since there is no direct requirement for determination of heights of surface points in the usual horizontal control network, it was reasonable to separate the vertical control and carry out an adjustment independent of horizontal control data using primarily the results of the usual geodetic leveling.

If we now re-consider the problem in the light of current observational data, it is apparent that an adjustment of all such data would be a goal recognizing that the observational material is generated from the surface of the earth. Thus we wish to use some or all of the following measurements: horizontal direction or angular measurements; zenith distance or vertical angle observations; chord distances such as determined from electromagnetic distance measuring equipment, astronomic latitude, longitude and azimuth measurements; leveled height differences as determined through usual geodetic practices. x, y, z positions from satellite or lunar data, chord distances from very long line baseline interferometry (VLBI) etc. This data could be processed without a reference surface such as a reference ellipsoid, or we could use such a surface if the results are needed with respect to such a system. The method that incorporates some or all of the data types into a single adjustment of geodetic data to determine the position of points is known as three dimensional or spatial geodesy.

The original suggestion of computing a triangulation network in space is attributed to Bruns (1878). In 1957 Hotine developed equations that could be used for the adjustment of a three dimensional network and Brazier and Windsor (1957) described a test network in which an adjustment of a simulated three dimensional network was carried out and compared with the adjustment of a conventional horizontal control network. Subsequent papers by various authors have refined

147

the mathematics involved in three-dimensional geodesy and have carried out adjustments using real data. Especially important in this respect is the paper of Wolf (1963) where the differential equations needed for the adjustment process are derived in a manner different from Hotine, and the discussions on three dimensional geodesy and vertical refraction by Hotine (1969). Additional discussion on the background and principles of three-dimensional geodesy may be found in Ramsayer (1972).

In the following sections we will derive the more important concepts in establishing the mathematical model of three dimensional geodesy, and we will discuss some of the practical implementation techniques and results.

4.2 Coordinate Systems and Coordinate Relationship

We first introduce the rectangular coordinates, x, y, z for some point of interest. This rectangular coordinate system is defined such that the z axis is parallel to the mean rotation axis as defined by the CIO pole, and the x axis passes through the mean Greenwich astronomic meridian as defined by the BIH. x is perpendicular to z and y is perpendicular to x and z. We may also define the origin of the rectangular coordinate system as being at the center of mass of the earth.

If we were to introduce an ellipsoidal surface for reference purposes we would have as usual:

$$x = (N+h) \cos \varphi \cos \lambda$$

$$y = (N+h) \cos \varphi \sin \lambda \qquad (1)$$

$$z = (N(1-e^2)+h) \sin \varphi$$

where φ and λ are the geodetic latitude and longitude of a point on the ellipsoid through which a normal to the ellipsoid is passed through the point in question. h is the geometric height of the point above the ellipsoid measured along the normal.

We next define a local coordinate system whose axes are u, v and w. The origin of this system is at a point from which observations might be made. The w axis coincides with the local vertical and is positive up. The u axis point along the astronomic meridian (positive north) and v points east being perpendicular to u and w. The actual directions involved in u, v, and w must be related to the same mean astronomic system used in defining the x, y, z coordinate system. These local coordinates are shown in the rectangular coordinate system in Figure 4-1 for a point whose astronomic latitude and longitude are φ' and λ'.

148
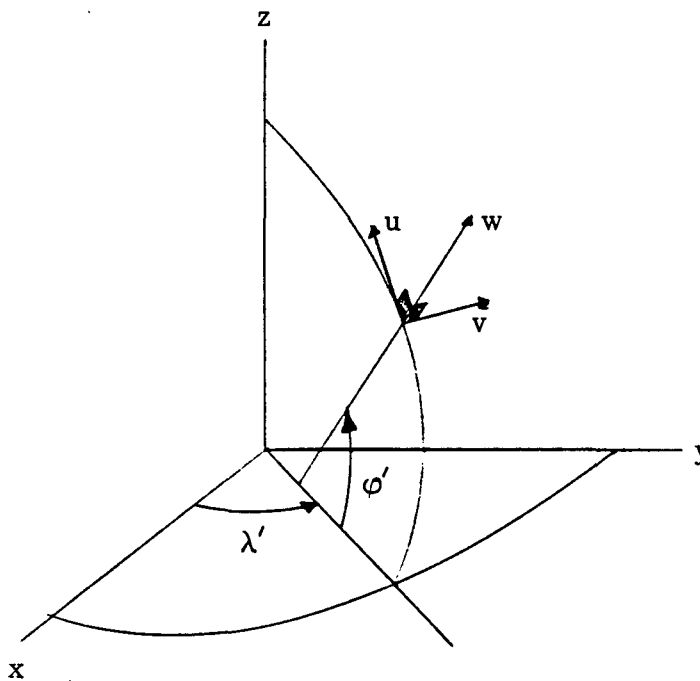
Figure 4-1

The Local and Rectangular Coordinate System

Now let the observations from point P to another point be designated α' for the astronomic azimuth, V' for astronomic vertical angle and s for the chord distance between the points. The observed quantities are illustrated in Figure 4-2 in the u, v, w coordinate system.
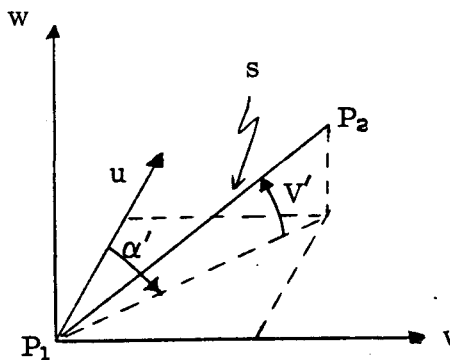
Figure 4-2

α', V' and s in the u, v, w System

149

In terms of the measured quantities at P_1, the u, v, w coordinates of P_2 can be seen from Figure 4-2 to be as follows:

$$u = s \cos V' \cos \alpha'$$

$$v = s \cos V' \sin \alpha' \tag{2}$$

$$w = s \sin V'$$

We can solve equations (2) to express the measured quantities in terms of the u, v and w coordinates. We have:

$$\alpha' = \tan^{-1} \frac{v}{u}$$

$$V' = \tan^{-1} \frac{w}{\sqrt{u^2 + v^2}} = \sin^{-1} \frac{w}{s} \tag{3}$$

$$s = \sqrt{u^2 + v^2 + w^2}$$

The rectangular coordinate difference between points 1 and 2 can now be written as:

$$\Delta x = x_2 - x_1$$

$$\Delta y = y_2 - y_1 \tag{4}$$

$$\Delta z = z_2 - z_1$$

We now need to convert these spatial coordinate differences into the local astronomic (u, v, w) system.

To do this we consider Figure 4.3 drawn with an origin at the point in question. The translated x, y, z axes correspond to the coordinate differences Δx, Δy, Δz of the two points being considered. For convience we introduce a $\bar{v}$ axis so that $\bar{v} = -v$. There are two rotations involved. The first is a R_3 rotation about the z axis of $-(180° - \lambda')$ and the second is a R_2 rotation of $-(90° - \varphi')$. We then have:

$$(5) \begin{pmatrix} u \\ -v \\ w \end{pmatrix} = R_2 (\varphi' - 90°) \; R_3 (\lambda' - 180°) \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}$$
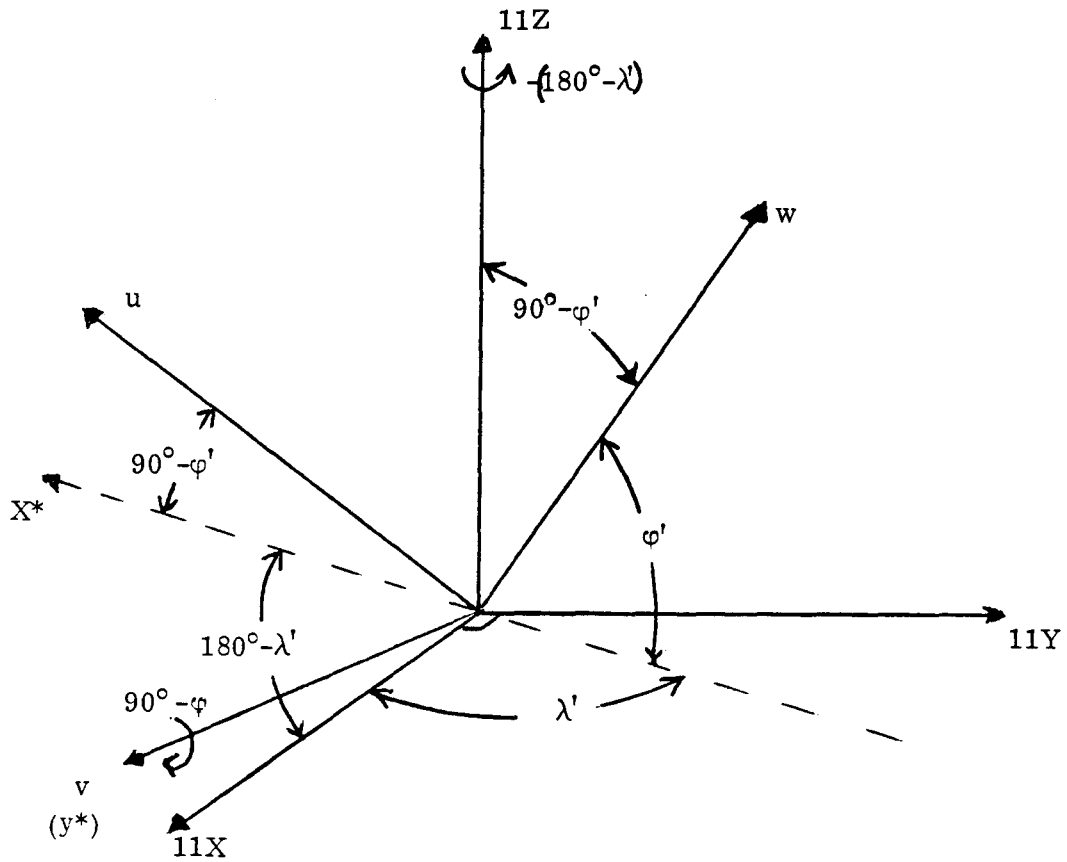
Note that there are only two rotations involved because the u and w axes form a plane that, by definition is parallel to the z axis. Multiplying out the rotation matrices and considering the sign on v, equation (5) can be written as:

151

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -\sin\varphi'\cos\lambda' & -\sin\varphi'\sin\lambda' & \cos\varphi' \\ -\sin\lambda' & \cos\lambda' & 0 \\ \cos\varphi'\cos\lambda' & \cos\varphi'\sin\lambda' & \sin\varphi' \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \qquad (6)$$

We may use equation (6) to determine unit vectors along the u, v, w axes in terms of the unit vectors along the x, y, z axes. We can write:

$$\vec{u} = -\sin\varphi'\cos\lambda'\,\vec{i} - \sin\varphi'\sin\lambda'\,\vec{j} + \cos\varphi'\,\vec{k}$$

$$\vec{v} = -\sin\lambda'\,\vec{i} + \cos\lambda'\,\vec{j} \qquad (7)$$

$$\vec{w} = \cos\varphi'\cos\lambda'\,\vec{i} + \cos\varphi'\sin\lambda'\,\vec{j} + \sin\varphi'\,\vec{k}$$

where $\vec{u}$, $\vec{v}$, $\vec{w}$ are unit vectors along the u, v, w axes and $\vec{i}$, $\vec{j}$, $\vec{k}$ are unit vectors along the x, y, z axes. If we wished to determine unit vectors along a local geodetic system we would simply replace the astronomic coordinates in (7) by the corresponding geodetic coordinates.

Next we express (6) in the following form:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = R(\varphi', \lambda') \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \qquad (8)$$

Where R is the coefficient matrix in (6). We then solve (8) for Δx, Δy, Δz:

$$\begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = R^{-1}(\varphi', \lambda') \begin{pmatrix} u \\ v \\ w \end{pmatrix} \qquad (9)$$

Carrying out the inversion or noting that R^{-1} equals R^T we can write (9) as follows:

$$\begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} -\sin\varphi'\cos\lambda' & -\sin\lambda' & \cos\varphi'\cos\lambda' \\ -\sin\varphi'\sin\lambda' & \cos\lambda' & \cos\varphi'\sin\lambda' \\ \cos\varphi' & 0 & \sin\varphi' \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} \qquad (10)$$

We can express the rectangular coordinates differences in terms of the measured quantities by substituting (2) into (10) to find:

$$\Delta x = s\,[\cos\lambda'\,(\cos\varphi'\sin V' - \sin\varphi'\cos\alpha'\cos V') - \sin\lambda'\sin\alpha'\cos V'] \qquad (11)$$

152

$$\Delta y = s \left[\sin \lambda' (\cos \omega' \sin V' - \sin \omega' \cos \alpha' \cos V') + \cos \lambda' \sin \alpha' \cos V' \right] \qquad (12)$$

$$\Delta z = s \left[\cos \omega' \cos V' \cos \alpha' + \sin \omega' \sin V' \right] \qquad (13)$$

We can also re-write (3) by substituting equation (6). We find:

$$\alpha' = \tan^{-1} \left[\frac{-\sin \lambda' \Delta x + \cos \lambda' \Delta y}{-\sin \omega' \cos \lambda' \Delta x - \sin \omega' \sin \lambda' \Delta y + \cos \omega' \Delta z} \right] \qquad (14)$$

$$V' = \sin^{-1} \frac{1}{s} (\cos \omega' \cos \lambda' \Delta x + \cos \omega' \sin \lambda' \Delta y + \sin \omega' \Delta z) \qquad (15)$$

$$s = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} \qquad (16)$$

Equations (14) and (15) may be used to compute the geodetic normal section azimuth, and geodetic vertical angle, simply by replacing the astronomic coordinates by the geodetic coordinates of the point at which the azimuth and vertical angle are being computed.

In some-cases it is convenient to express (14) and (15) in terms of geodetic coordinates by replacing Δx, Δy, Δz with corresponding values computed from equation (1). In this case we can write (Mitchell, 1963):

$$\alpha' = \cot^{-1} \left[\frac{\begin{aligned} &[(N_2 + h_2)(\cos \varphi_2 \sin \varphi_1' \cos(\lambda_1' - \lambda_2) - \sin \varphi_2 \cos \varphi_1') \\ &- (N_1 + h_1)(\cos \varphi_1 \sin \varphi_1' \cos(\lambda_1' - \lambda_1) - \sin \varphi_1 \cos \varphi_1') \\ &+ e^2 \cos \varphi_1' (N_2 \sin \varphi_2 - N_1 \sin \varphi_1)] \end{aligned}}{(N_2 + h_2) \cos \varphi_2 \sin(\lambda_1' - \lambda_2) - (N_1 + h_1) \cos \varphi_1 \sin(\lambda_1' - \lambda_1)} \right] \qquad (17)$$

$$\begin{aligned} V' = \sin^{-1} \frac{1}{s} \Big(&(N_2 + h_2)(\cos \varphi_2 \cos \varphi_1' \cos(\lambda_2 - \lambda_1') + \sin \varphi_2 \sin \varphi_1') \\ &- (N_1 + h_1)(\cos \varphi_1 \cos \varphi_1' \cos(\lambda_1 - \lambda_1') + \sin \varphi_1 \sin \varphi_1') \\ &- e^2 \sin \varphi_1' (N_2 \sin \varphi_2 - N_1 \sin \varphi_1) \Big) \end{aligned} \qquad (18)$$

At this point we can recognize the general functional relationships between the quantities discussed so far:

$$\alpha' = f_1(x_1, y_1, z_1, x_2, y_2, z_2, \varphi_1', \lambda_1')$$

$$V' = f_2(x_1, y_1, z_1, x_2, y_2, z_2, \varphi_1', \lambda_1') \tag{19}$$

$$s = f_3(x_1, y_1, z_1, x_2, y_2, z_2)$$

or

$$\alpha' = f_4(\varphi_1, \lambda_1, h_1, \varphi_2, \lambda_2, h_2, \varphi_1', \lambda_1')$$

$$V' = f_5(\varphi_1, \lambda_1, h_1, \varphi_2, \lambda_2, h_2, \varphi_1', \lambda_1') \tag{20}$$

$$s = f_6(\varphi_1, \lambda_1, h_1, \varphi_2, \lambda_2, h_2)$$

The equations derived in this section may be used for several different purposes as well as providing the basic equations to be used for deriving the observation equation needed for a three-dimensional geodesy adjustment. For example, if we know the geodetic coordinates (φ, λ, h) of one point, the geodetic coordinates of a second point can be computed if we observed φ', λ', α', V' and s, compute the rectangular coordinate differences from (11), (12), and (13), find the rectangular coordinates of the second point by adding the coordinate differences to the rectangular coordinates of the first point and then solving for φ, λ, h of the second point from the x, y, z value. In addition, equations such as (14), (15) and (16) may be used to compute the approximate values of the observed quantities based on the approximate coordinates of the stations involved when we are considering an adjustment procedure.

4.3 Differential Relationships

At this point we need to develop the differential relationships that relate the changes in the quantities given in the functions of (19) and (20) to the corresponding changes in α', V' and s. To do this we first consider the variation of α', and V' through equation (3). For $d\alpha'$ we have:

$$d\alpha' = \frac{u^2}{v^2 + u^2} \cdot \frac{u\,dv - v\,du}{u^2} \tag{21}$$

Using equation (2), (21) reduces to:

$$d\alpha' = \frac{1}{s \cos V'} (\cos \alpha' \, dv - \sin \alpha' \, du) \tag{22}$$

154

In a similar fashion the arcsine expression for V' in (3) yields:

$$dV' = -\frac{s\ dw - w\ ds}{s^2 \cos V'} \tag{23}$$

We next find the variation of the local coordinates by differentiating (8). We have:

$$\begin{pmatrix} du \\ dv \\ dw \end{pmatrix} = R \begin{pmatrix} d\Delta x \\ d\Delta y \\ d\Delta z \end{pmatrix} + \frac{\partial R}{\partial \varphi'} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} d\varphi' + \frac{\partial R}{\partial \lambda'} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} d\lambda' \tag{24}$$

We find:

$$\frac{\partial R}{\partial \varphi'} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} -w \\ o \\ u \end{pmatrix} \tag{25}$$

and:

$$\frac{\partial R}{\partial \lambda'} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} -v \sin \varphi' \\ -\cos \lambda' \Delta x - \sin \lambda' \Delta y \\ v \cos \varphi' \end{pmatrix} \tag{26}$$

We now can write (24) as follows:

$$\begin{pmatrix} du \\ dv \\ dw \end{pmatrix} = R \begin{pmatrix} d\Delta x \\ d\Delta y \\ d\Delta z \end{pmatrix} + \begin{pmatrix} -w \\ o \\ u \end{pmatrix} d\varphi' + \begin{pmatrix} -v \sin \varphi' \\ -\cos \lambda' \Delta x - \sin \lambda' \Delta y \\ v \cos \varphi' \end{pmatrix} d\lambda' \tag{27}$$

We can also compute ds from (16):

$$ds = \frac{1}{s} (\Delta x\ d\Delta x + \Delta y\ d\Delta y + \Delta z\ d\Delta z) \tag{28}$$

155

We now substitute equations (27) into (21), and (27) and (28) into (23), and using (2) we determine the following equation:

$$d\alpha' = a_1 (dx_2 - dx_1) + a_2 (dy_2 - dy_1) + a_3 (dz_2 - dz_1)$$
$$+ a_4 d\varphi_1' + a_5 d\lambda_1' \tag{29}$$

$$dV' = b_1 (dx_2 - dx_1) + b_2 (dy_2 - dy_1) + b_3 (dz_3 - dz_1)$$
$$+ b_4 d\varphi_1' + b_5 d\lambda_1' \tag{30}$$

$$ds = c_1 (dx_2 - dx_1) + c_2 (dy_2 - dy_1) + c_3 (dz_2 - dz_1) \tag{31}$$

The coefficients in (29) through (31) are as follows (Wolf, 1963):

$$a_1 = \frac{\sin\varphi' \cos\lambda' \sin\alpha' - \sin\lambda' \cos\alpha'}{s \cos V'}$$

$$a_2 = \frac{\sin\varphi' \sin\lambda' \sin\alpha' + \cos\lambda' \cos\alpha'}{s \cos V'}$$

$$a_3 = \frac{-\cos\varphi' \sin\alpha'}{s \cos V'} \tag{32}$$

$$a_4 = \sin\alpha' \tan V'$$

$$a_5 = (\sin\varphi' - \cos\varphi' \cos\alpha' \tan V')$$

$$b_1 = \frac{s \cos\varphi' \cos\lambda' - \sin V' \Delta x}{s^2 \cos V'}$$

$$b_2 = \frac{s \cos\varphi' \sin\lambda' - \sin V' \Delta y}{s^2 \cos V'}$$

$$b_3 = \frac{s \sin\varphi' - \sin V' \Delta z}{s^2 \cos V'} \tag{33}$$

$$b_4 = \cos\alpha'$$

$$b_5 = \sin\alpha' \cos\varphi'$$

156

$$c_1 = \frac{x_2 - x_1}{s}$$

$$c_2 = \frac{y_2 - y_1}{s}$$

(34)

$$c_3 = \frac{z_2 - z_1}{s}$$

In these coefficients the astronomic coordinates refer to the point at which the observations are made, α' is the astronomic azimuth from the first point to the second point, and V' is the vertical angle from point 1 to point 2. In practice, if the astronomic observations are not available, the corresponding geodetic values may be used in the above coefficients.

Equation (29) through (31) may also be re-written with the variable quantities on the right hand side in terms of $d\varphi$, $d\lambda$, and dh by using equation (2.109, 110, 111) to replace dx, dy, and dz. After considerable algebraic manipulation we find (Mitchell, 1963, Wolf, 1963):

$$d\alpha' = d_1\, d\varphi_1 + d_2\, d\lambda_1 + d_3\, dh_1 + d_4\, d\varphi_2 + d_5\, d\lambda_2 + d_6\, dh_2$$
$$+ d_7\, d\varphi_1' + d_8\, d\lambda_1'$$

(35)

$$dV' = e_1\, d\varphi_1 + e_2\, d\lambda_1 + e_3\, dh_1 + e_4\, d\varphi_2 + e_5\, d\lambda_2 + e_6\, dh_2$$
$$+ e_7\, d\varphi_1' + e_8\, d\lambda_1'$$

(36)

$$ds = f_1\, d\varphi_1 + f_2\, d\lambda_1 + f_3\, dh_1 + f_4\, d\varphi_2 + f_5\, d\lambda_2 + f_6\, dh_2$$

(37)

The coefficients in (35) through (37) are as follows:

$$d_1 = \frac{(M_1 + h_1)\sin\alpha_1}{s\cos V_1}$$

$$d_2 = \frac{-(N_1 + h_1)\cos\varphi_1 \cos\alpha_1}{s\cos V_1}$$

$$d_3 = 0$$

157

$$d_4 = \frac{-(M_2 + h_2)}{s \cos V_1} \sin \alpha_1 (\sin \varphi_1 \sin \varphi_2 \cos \Delta\lambda + \cot \alpha_1 \sin \varphi_2 \sin \Delta\lambda +$$

$$+ \cos \varphi_1 \cos \varphi_2)$$

(38)

$$d_5 = \frac{(N_2 + h_2) \cos \varphi_2 \cos \alpha_1 (\cos \Delta\lambda - \sin \varphi_1 \sin \Delta\lambda \tan \alpha_1)}{s \cos V_1}$$

$$d_6 = \frac{\cos \varphi_2 \cos \alpha_1 [\tan \alpha_1 (\sin \varphi_1 \cos \Delta\lambda - \tan \varphi_2 \cos \varphi_1) + \sin \Delta\lambda]}{s \cos V_1}$$

$$d_7 = \tan V_1 \sin \alpha_1$$

$$d_8 = \sin \varphi_1 - \cos \varphi_1 \tan V_1 \cos \alpha_1$$

$$e_1 = \frac{(M_1 + h_1) \sin V_1 \cos \alpha_1}{s}$$

$$e_2 = \frac{(N_1 + h_1) \sin V_1 \sin \alpha_1 \cos \varphi_1}{s}$$

$$e_3 = \frac{-\cos V_1}{s}$$

$$e_4 = \frac{-(M_2 + h_2)}{s \cos V_1} (\sin \varphi_2 \cos \varphi_1 \cos \Delta\lambda - \cos \varphi_2 \sin \varphi_1 - \sin V_1 \cos V_2 \cos \alpha_2)$$

(39)

$$e_5 = \frac{-(N_2 + h_2) \cos \varphi_2}{s \cos V_1} (\cos \varphi_1 \sin \Delta\lambda - \sin V_1 \cos V_2 \sin \alpha_2)$$

$$e_6 = \frac{1}{s \cos V_1} (\cos \varphi_1 \cos \varphi_2 \cos \Delta\lambda + \sin \varphi_1 \sin \varphi_2 + \sin V_1 \sin V_2)$$

$$e_7 = \cos \alpha_1$$

$$e_8 = \cos \varphi_1 \sin \alpha_1$$

158

$$f_1 = -(M_1 + h_1) \cos V_1 \cos \alpha_1$$

$$f_2 = -(N_1 + h_1) \cos V_1 \sin \alpha_1 \cos \varphi_1$$

$$f_3 = -\sin V_1$$

$$f_4 = -(M_2 + h_2) \cos V_2 \cos \alpha_2 \tag{40}$$

$$f_5 = -(N_2 + h_2) \cos V_2 \sin \alpha_2 \cos \varphi_2$$

$$f_6 = -\sin V_2$$

In these expressions $\Delta\lambda = \lambda_2 - \lambda_1$, α_2 is the azimuth from point 2 to point 1 and V_2 is the vertical angle from point 2 to point 1. In addition, no distinction has been made in these coefficients between astronomic and geodetic quantities.

In some cases simplification of certain of the above coefficients is possible under certain assumptions (for example, short lines). We have (Ramsayer, 1971) for lines of less than 20 km:

$$d_4 = -d_1$$

$$d_5 = -d_2$$

$$e_4 = -e_1$$

$$e_5 = -e_2 \tag{41}$$

$$f_4 = -f_1$$

$$f_5 = -f_2$$

In addition d_6 is very close to zero and is in fact taken as zero by Vincenty (1974). Some alternate approximations to the rigorous coefficients may also be found in Vincenty.

159

4.4   General Adjustment Procedures

Now having available the differential relationships derived in section 4.3, we are in a position to develop the observation equations needed to perform a least squares adjustment. To start, we write a general observation in the form:

$$F(X_0) + \frac{\partial F}{\partial X}\, dX = L_{0BS} + v \tag{42}$$

where F is the function relating the observations, L_{0BS}, and the parameters, X of the problem. dX are the corrections to the approximate values X_0 of the parameters and v is the observation residual. From (42) we can write:

$$v = F(X_0) - L_{0BS} + \frac{\partial F}{\partial X}\, dX \tag{43}$$

In section 4.3 we have developed the expression for $\frac{\partial F}{\partial X}\, dX$ and we now proceed to apply them for each possible observation. Before an adjustment is started the observations may be reduced to corresponding observations between reference marks on the ground. If this is not done the final adjusted positions of the mark must be found by reducing the adjusted positions of the observation instrument. There are a number of arguments for not first reducing the data to the marks. For example, Vincenty (1979a) points out that any reduction assumes heights which may not accurately be known, and such re-computation after each iteration of adjustment may slow down the convergence of the process.

4.41 Astronomic Azimuth Observations.

In this case we assume we have measured α' and that we compute a corresponding approximate value α' based on the approximate values adopted for the unknown quantities which are the quantities listed on the right hand side of the first equations in either (19) or (20). Using (35) (for example) and (43) we can write:

$$v_{\alpha'} = \alpha'_0 - \alpha' + d_1\, d\omega_1 + d_2\, d\lambda_1 + d_3\, dh_1 + d_4\, d\omega_2 + d_5\, d\lambda_2 + d_6\, dh_2 + d_7\, d\omega'_1 + d_8\, d\lambda'_1 \tag{44}$$

4.42 Horizontal Direction Measurements

Let a set of directions referred to an initial direction D_1 be designated by $D_1, \ldots D_1$. The astronomic azimuth of the initial line is α'_1 which may be only approximately known (α'_{10}) so that we write:

$$\alpha'_1 = \alpha'_{10} + Z \tag{45}$$

where Z is known as the orientation or station correction. α'_{10} can be computed given the approximate coordinates of the two points involved with the initial line using (14). The "observed" astronomic azimuth for some direction would then be:

$$\alpha_1' = \alpha_1' + D_i - D_1 = \alpha_{10}' + Z + D_i - D_1 \tag{46}$$

Using α_1' as the observed quantity in (43) we can write:

$$v_0 = \alpha_0' - (\alpha_{10}' + D_i - D_1) - Z + d\alpha' \tag{47}$$

where $d\alpha'$ is given by (35) or (29). Note that no corrections for deflection of the vertical are to be made to the observed direction. The discussion here can be compared to that given for classical direction observations equations as described in section 7 of Rapp (1984).

4.43 Vertical Angle Measurements

In deriving equations in this chapter we have assumed that no refraction of the light rays has taken place. In fact it does, with the greatest effect occurring on vertical angles. Consequently before we can finalize the vertical angle observation equation it is necessary to consider the effect of vertical refraction.

4.431 Vertical Refraction Modeling

We start by designating the measured vertical angle by $\overline{V}'$ recalling that the corresponding value unaffected by refraction has been called V'. We write:

$$\overline{V}' = V' + \Delta V \tag{48}$$

as may also be seen in Figure 4-3
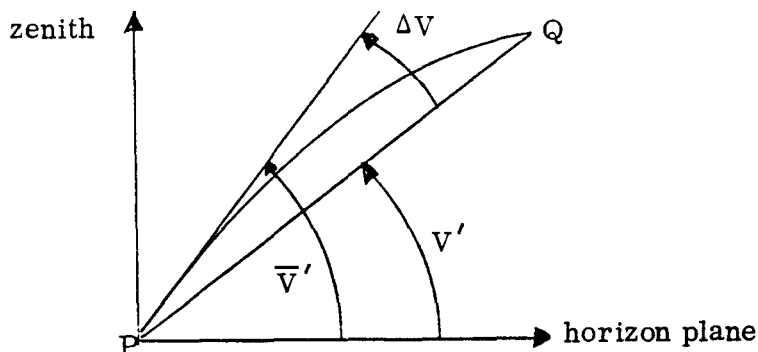
Figure 4-3

Vertical Refraction Effect

A general discussion of vertical refraction may be found in Hotine (1969, Chapter 24) and Bomford (1971, Sections 3.19 and 3.20).

To start we first express the angle of refraction at P in terms of the curvature of the light ray and the length of the line. We have (Bomford, 1980, e.g. 3.49):

$$\Delta V = \frac{1}{s} \int_{o}^{s} \frac{s - \ell}{\sigma} \, d\ell \tag{49}$$

where s is the length of the line from P to Q, and σ is the radius of curvature of the light path. The value of σ can be expressed as (Bomford, eq. 3.44, Hotine, eq. 24.59):

$$\frac{1}{\sigma} = - \frac{1}{n} \frac{dn}{dh} \cos V \tag{50}$$

where n is the refractive index and h is the geodetic height. The refractive index for optical wavelengths can be expressed as follows (Bomford, eq. 1.70; Hotine eq. 24.54):

$$(n-1) = \frac{(n_{0} - 1)}{\alpha T} \cdot \frac{P}{1013.25} - \frac{0.000000042 \, e}{\alpha T} \tag{51}$$

where:

> n_{0} is the refractive index for a specific wavelength at standard atmospheric conditions;
> α is the coefficient of expansion of air;
> T is the temperature in $^{\circ}$K;
> P is the total air pressure in millibars;
> e is the partial pressure of water vapor in millibars.

Differentiating (51), neglecting the small effect of e, and substituting nominal numerical values we have from (50) (Bomford, eq. 3.48):

$$\frac{1}{\sigma} = 16.3 \frac{P}{T^{2}} (.0342 + \frac{dT}{dh}) \cos V \text{ seconds/meter} \tag{52}$$

If P, T, and $\frac{dT}{dh}$ were known along the light path we could substitute (52) into (49) to determine ΔV. Alternately we can postulate some average conditions in the evaluation of (49). Thus, if we assume σ is constant over the ray path we can evaluate (49) to give:

162

$$\Delta V = \frac{s}{2\sigma} \tag{53}$$

or using (52) we have:

$$\Delta V = s \left[\frac{16.3}{2} \frac{P}{T^2} \left(.0342 + \frac{dT}{dh}\right) \right] \cos V \tag{54}$$

We could insert nominal values at sea level of P (1000 millibars); of T (300°K); and of dT/dh = -0.0055°C/m to find from (54):

$$\Delta V = 0\overset{.}{.}0026 \, s \cos V \quad \text{(s in meters)} \tag{55}$$

At this point we have developed a model that can be used to compute the effect of refraction on the measured vertical angle. Thus from (54) or (55) we could write:

$$\Delta V = qs \cos V \tag{56}$$

where q is a constant that is to be determined in an adjustment starting from a nominal value for q such as given in (55) as $0\overset{.}{.}0026$/meter. Additional refinement could be made, for example, by assuming that q is composed of two parts: one a constant part and another part dependent on h in an inverse manner. Thus we might write:

$$q = q_1 + \frac{q_2}{h} \tag{57}$$

where q_1 and q_2 will be quantities to be determined.

Another analysis of vertical angle refraction has been carried out by Pfeifer (1973) where he obtained the following expression for ΔV:

$$\Delta V = \tfrac{1}{2} \, Gs \sqrt{2(U+V) - (U-V)^2 - 1} \left[\frac{K-H}{s} + \tfrac{1}{2}(1 + U-V)\ell n \left\{ \frac{H+K+s}{H+K-s} \right\} \right] \tag{58}$$

where: $\quad$ H = h_1 + R
$\qquad\qquad$ K = h_2 + R
$\qquad\qquad$ U = $(H/s)^2$
$\qquad\qquad$ V = $(K/s)^2$
$\qquad\qquad$ R = mean earth radius,

and G is the vertical gradient of the index of refraction which is regarded as constant throughout the local airmass of the station from which the observations are made. Equation (58) has been transformed into its present form from a form in which there was a dependence on cos V such as we have seen in equation (56). Equation (58) is then used as the refraction model where G is the parameter to be determined.

Saito (1974) also discusses models for vertical refraction assuming a two parameter model for the curvature in equation (53).

In some literature the concept of the coefficient of refraction, **k**, is introduced in this subject through the defining equation:

$$\Delta V = k \, \theta \tag{59}$$

where θ is the angle subtended by the line PQ at the center of the earth. With θ given by s/R and ΔV by (54) we can write (Bomford, eq. 3.52):

$$k = 252 \, \frac{P}{T^2} \, (.0342 + \frac{dT}{dh}) \, \cos V \tag{60}$$

Inserting the normal values used in obtaining (55), into (60) yields a normal k value equal to 0.080. We can write (60) in the general form:

$$k = k_1 + k_2 \cdot \frac{dT}{dh} \tag{61}$$

If dT/dh is constant then the second term can be combined with the first. If we assume dT/dh is an inverse function of height we can write:

$$k = k_1 + \frac{k_2}{h} \tag{62}$$

where k_1 and k_2 become parameters of the vertical refraction model through equation (59). Such a model has been used by Hradilek (1968).

For the purposes of this section we will regard our vertical refraction model to be given by equation (56) leaving the use of more sophisticated models to the reader.

4.432 Vertical Angle Observation Equations

From (48) and (56) we write:

$$V' = \overline{V}' - \Delta V = \overline{V}' - qs \cos V \qquad (63)$$

Regarding V' as the observed quantity consistent with our previous derivations we write using (43)

$$v_{V'} = V_0' - (\overline{V}' - q \ s \cos V) + dV \qquad (64)$$

where V_0' is the vertical angle (without refraction effects) based on the assumed data and computed through equation (15). Letting q_0 be an approximate value of q and dq the correction to this approximate value we can write:

$$v_{V'} = V_0' - (\overline{V}' - q_0 \ s \cos V) + dq \ s \cos V + dV \qquad (65)$$

where dV would be given by equation (30) or (36). In some computations over non-steep lines it may be permissable to set $\cos V$ equal to 1, but in general the $\cos V$ should be retained. Hradilek (1973) reports however no correlation between the vertical angle and the magnitude of residuals when adjustments were made with $\cos V$ equal to one.

Clearly we can not consider a q value as an unknown quantity for each vertical angle since we would introduce too many unknowns into the problem. We can make certain reasonable assumptions, however, the most widely accepted being one where q is regarded as unique for all vertical angle measurements made from a given station during a specific observation session.

Another procedure for handling refraction effects on vertical angles is to make reciprocal vertical angle measurements between two stations, either simultaneously or at the same time of day such that refraction effects are the same. We then assume that the refraction unknown (q) is the same from both ends of the line. Then we form the vertical angle observation equation (equation 65) for the two reciprocal measurements, and subtract these two equations forming a single observation equation in which the refraction unknown has been eliminated through cancellation. This new observation equation is then used with a weight based on the weights of the initial reciprocal observations.

Other procedures such as adopting a known q value or a unique q value adjusted for a whole network are described by Hradilek (1972). A discussion of the various effects of different refraction models used in three dimensional geodesy adjustments may be found in Körner (1968).

4.433 Measurement of the Vertical Refraction Angle

Clearly if it were possible to measure ΔV we would have no need to try to model it, or to carry out special observation techniques to cancel out model parameters. Tengström (1967) has described a method of measuring ΔV using measurements made at two different known wavelengths. This technique is also discussed theoretically by Hotine (1969, p.226).

The principles of this technique start with the fact that ΔV can be expressed in the following form:

$$\Delta V = (n_0 - 1) R + Q \tag{66}$$

where R and Q 'meteorological integrals' since they are quantities dependent on the integration of functions which depend on the atmospheric conditions over the line. We next consider ΔV values dependent on two different wavelengths of light so we have:

$$\Delta V(\lambda_1) = (n_0(\lambda_1) - 1) R + Q \tag{67}$$

$$\Delta V(\lambda_2) = (n_0(\lambda_2) - 1) R + Q \tag{68}$$

We difference (67) and (68) to write:

$$\delta = [n_0(\lambda_1) - n_0(\lambda_2)] R \tag{69}$$

Now assume that we can measure δ so that we can determine R since $n(\lambda)$ is known. Using this R we can evaluate (66), (67), or (68) to determine ΔV provided Q is known. Tengström shows that Q can be estimated from:

$$Q = -5.5 \times 10^{-8} \left(\frac{e}{p} \right)_m \cdot 760 R \tag{70}$$

where e is the absolute humidity pressure in mm H_g and p is the total air pressure in mm H_g. Fortunately Q is a very small term and in many cases can be neglected if no information is available for its approximate computation through (70).

166

This technique requires a highly accurate measurement of the dispersion angle δ for the satisfactory determination of ΔV. For example Tengström indicates that δ needs to be measured to 0".003 which places a stringent criteria on the instrumental techniques needed for this method of determining ΔV. Prilepin (1973) indicates that the standard deviation of ΔV determined from the two wavelength method is 6000 times the standard deviation of the measurement of the dispersion angle, again pointing out the need for a very accurate measurement of δ.

A technique somewhat analogous to the method described here can be used to correct distances measured with optical distance measuring equipment. Recent results indicating satisfactory implementation of a two wavelength equipment is described by Bouricius and Earnshaw (1974).

4.434 Weighting of Vertical Angle Measurements

Because of the problems involved with vertical angle measurements and refraction modelling there is a need to carefully consider the proper weighting of the vertical angle observation equations. Thus the weighting itself may not depend solely on the actual observational accuracy of the vertical angle measurements. For example, in some cases it has been suggested to down weight the vertical angle observation equations because of the uncertainty due to refraction. Hradilek (1973) suggests that the a priori variance $m^2(\overline{V}')$ be assigned by the following equation:

$$m^2(\overline{V}') = m^2(a) + [C \tfrac{1}{2} \theta \, m \, (k)]^2 \tag{71}$$

where m(a) represents the accidental observation errors that are associated with the instrument used to measure the vertical angle, m(k) is the mean square error of the coefficient of refraction; and C is a quantity between 0.5 and 1.5 that depends on the number of observations made and the variation of their changes with time. Typical values of C are:number of observations: >3, C =0.5 to 0.8; 3, C= 1.3; 1, C = 1.5. Values of m(k) are given by Hradilek for various regions based on past adjustments and vary from ±0.011 to ±0.028.

Equation (71) is directly applicable when a refraction model is assumed known in the adjustment. When a refraction model is being determined in the adjustment, the standard deviation of the vertical angle observation equation should be based only on the first term on the right hand side of (71).

4.44 Distance Measurements

We let s_b be the observed chord distance between two stations after the measurement has been corrected for refraction and instrumental correction terms.

167

Then we write the distance observation equation as:

$$v_s = s_o - s_b + ds \qquad (72)$$

where ds is given by equation (31) or (37).

4.45 Astronomic Latitude and Longitude

Let φ_b' and λ_b' be the observed astronomic latitude and longitude of the actual station on the surface of the earth referred to the mean astronomic system used in defining the basic coordinate system. Now let φ_o' and λ_o' be corresponding approximate values so that equation (43) can be written:

$$v_{\varphi'} = \varphi_o' - \varphi_b' + d\varphi' \qquad (73)$$

$$v_{\lambda'} = \lambda_o' - \lambda_b' + d\lambda' \qquad (74)$$

These equations are then simply used as observation equations with appropriate weighting. φ_o' and λ_o' may be chosen the same as the observed quantities or they may be chosen as the approximate geodetic coordinates of the station.

An alternate procedure for incorporating this astronomic information is to regard it as "a priori" parameter data such that the usual normal equations are modified by adding the weight matrix of the 'observed' parameters to the elements of the normal equations in a procedure such as described by Mikhail (1970). This procedure allows the very simple fixing of one or any number astronomic coordinates.

4.46   Height Differences from Spirit Levelling

We let H be the orthometric height of a point, and N the geoid undulation so that the geometric height of the point above the ellipsoid is essentially given by:

$$h = H + N \qquad (75)$$

If we write this equation for two stations and take the difference we have:

168

$$h_2 - h_1 = H_2 - H_1 + N_2 - N_1 \tag{76}$$

Now $H_2 - H_1$ is a quantity that is accurately determined by standard leveling procedures. If we were working with lines sufficiently short such that $N_2 - N_1$ is zero, we could form a simple observation equation from (75) by writing:

$$\Delta H = h_2 - h_1 \tag{77}$$

$$v_{\Delta H} = (h_2 - h_1)_0 - \Delta H + dh_2 - dh_1 \tag{78}$$

Over longer lines where $N_2 - N_1$ can not be assumed zero we must develop a more complex observation equation. Hotine (1969, p. 245) indicates, and Chovitz (1974) proves that under the assumption that the vertical angles involved are small, the following equation can be used to connect the measured orthometric height differences and vertical angles:

$$\Delta H = s \, \cos \frac{V_1 + V_2}{2} \, \sin \tfrac{1}{2}(V_1 - V_2) \approx \tfrac{1}{2} s \, (V_1 - V_2) \tag{79}$$

where V_2 is the vertical angle of the line from point 2 to point 1. We use (79) to form the following observation equation:

$$v_{\Delta H} = (\Delta H_0 - \Delta H_b) + \tfrac{1}{2} s \, dV_1 - \tfrac{1}{2} s \, dV_2 \tag{80}$$

where ΔH_0 is the approximate orthometric height difference computed from (79) on the basis of the assumed approximate coordinates. dV_1 and dV_2 are taken from (30) or (36) being evaluated for the two points involved.

A somewhat similar procedure has been suggested by Vincenty (1974, 1979a) where the following observations equation has been given assuming that the astrogeodetic deflections vary uniformly between the two stations and that the preliminary astronomic coordinates are set to the corresponding most recent geodetic values.

$$v_{\Delta H} = -dh_1 + dh_2 + (s \cos \alpha_{12} \cos V_1 /2) d\omega_1' + (s \cos \omega_1 \sin \alpha_{12} \cos V_1 /2) d\lambda_1'$$

$$-(s \cos \alpha_{21} \cos V_2 /2) d\omega_2' - (s \cos \omega_2 \sin \alpha_{21} \cos V_2 /2) d\lambda_2' \tag{81}$$

$$+ (h_2 - h_1)_0 - \Delta H$$

The weights for the various observation equations discussed here would be computed on the basis of the standard deviation of the measured elevation difference.

4.5 The Use of Three-Dimensional Adjustment Procedures in Horizontal Networks

The previous discussion has been related to the use of many different types of observations in a three dimensional sense so that no distinction between a horizontal and vertical network was made. It is possible however, to apply our three dimensional equations to just data acquired in a horizontal network. Such a procedure has been discussed by Vincenty and Bowring (1978) with a computer program discussed by Vincenty (1979b, 80a) and additional information in Bowring (1980).

We start by assuming that both heights and astronomic coordinates are known and are to be held fixed in the adjustment. The observations will be the usual horizontal directions, astronomic azimuths, chord distances etc. These observations are not reduced to the ellipsoid as is done in the classic horizontal network adjustment.

The adjustment can be carried out in geographic coordinates or rectangular coordinates. Consider first the geographic coordinate adjustment. For the direction observation equation we can write from (47) and (35):

$$(82) \quad v_D = \alpha'_0 - (\alpha'_0 + D_i - D_i) - Z + d_1 d\varphi_1 + d_2 d\lambda_1 + d_4 d\varphi_2 + d_5 d\lambda_2$$

The chord distance observation equation would be (from 72 and 37):

$$(83) \quad v_s = s_0 - s_b + f_1 d\varphi_1 + f_2 d\lambda_1 + f_4 d\varphi_2 + f_5 d\lambda_2$$

Note that in the computation of the approximate values of the observations the values of the terms to be held fixed are used.

A more complicated procedure takes place when an adjustment in rectangular coordinates is to be carried out. Here the general form of the observation equation is (see 29, 30, 31 and 43):

$$(84) \quad v = F(X_0) - L_{obs} + a(dx_2 - dx_1) + b(dy_2 - dy_1) + c(dz_2 - dz_1)$$

where the a, b, c coefficients are readily identified with the coefficients given in equation (32), (33) and (34). Note that no $d\varphi'$ and $d\lambda'$ appear in these expressions as these quantities are to be held fixed. We now need to consider the constraint imposed by saying the height is to be held fixed in the adjustment. Several ways have been discussed in the literature.

170

Vincenty and Bowring (1978) and Vincenty (1980b) discuss a procedure to eliminate one coordinate unknown per station from (84). To do this an auxiliary ellipsoid is introduced so that the normal at the point in question is the same as the normal to the usual reference ellipsoid. This implies that the rectangular coordinates of the point with respect to the auxiliary ellipsoid are:

$$\begin{aligned} x &= N_0 \cos\varphi \cos\lambda \\ (85) \quad y &= N_0 \cos\varphi \sin\lambda \\ z &= N_0 (1 - e_0^2) \sin\varphi \end{aligned}$$

where $N_0 = N + h$ and e_0 is the eccentricity of the auxiliary ellipsoid. Equating (1) and (85) Vincenty (1980b) finds:

$$(86) \quad e_0^2 = \frac{e^2}{1 + \dfrac{h}{N}} \approx \frac{e^2}{1 + \dfrac{h}{a}}$$

The equation of the auxiliary ellipsoid is

$$(87) \quad x^2 + y^2 + \frac{z^2}{(1 - e_0^2)} = a_0^2$$

which in differential form is:

$$(88) \quad x \, dx + y \, dy + \frac{z \, dz}{(1 - e_0^2)} = 0$$

This equation can them be used to eliminate one of the unknowns at a given station that will assume the fixing of the height. As an example consider the elimination of dz_2 and dz_1 appearing in (84). Our observation equation will now become:

$$(89) \quad v = F(X_0) - L_{obs} + \bar{b}_1 \, dx_1 + \bar{b}_2 \, d \, y_1 + \bar{b}_3 \, dx_2 + \bar{b}_4 \, dy_2$$

where (Vincenty, 1980b)

$$(90) \quad \begin{aligned} \bar{b}_1 &= -a + \frac{c \, k_1 \, x_1}{z_1} & \bar{b}_3 &= a - \frac{c \, k_2 \, x_2}{z_2} \\ \bar{b}_2 &= -b + \frac{c \, k_1 \, y_1}{z_1} & \bar{b}_4 &= b - \frac{c \, k_2 \, y_2}{z_2} \end{aligned}$$

with $k = 1 - e_0^2$

The coefficients where other unknowns are eliminated are given in Vincenty and Bowring (1978) and Vincenty (1980b). Once the adjustment is completed the eliminated unknowns can be found from equation (88).

Another approach to solving the rectangular coordinate adjustment is described by Bowring (1980) and Vincenty (1980b). Here equation (10) of Section 4 is differentiated where the height (w) is held fixed. Then (10) can be written:

171

$$(91) \quad \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = Q \begin{pmatrix} du \\ dv \\ 0 \end{pmatrix}$$

where Q is the obvious coefficient matrix in (10) but with geodetic ϕ and λ.

If we substitute (91) into (84) we can write:

$$(92) \quad v = F(X_0) - L_{obs} + F\, du_1 + G\, dv_1 - \overline{F}\, du_2 - \overline{G}\, dv_2$$

where
$$(93) \quad \begin{bmatrix} F \\ G \\ H \end{bmatrix} = -Q^\tau \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

with a similar expression for $\overline{F}$ and $\overline{G}$. The corrections du and dv appearing in (92) are related to $d\phi$ and $d\lambda$ by:

$$(94) \quad du = (M + h)\, d\varphi$$
$$dv = (N + h)\, \cos\varphi\, d\lambda$$

Thus the use of (92) in the adjustment effectively inforces the height. Simplified coefficients appearing in (92) for both azimuth and chord distance are given in Bowring (1980):

For azimuth:

$$(95) \quad \begin{bmatrix} F \\ G \end{bmatrix} = \frac{-1}{R^2}\, Q^\tau\, P \begin{bmatrix} -D \\ C \end{bmatrix}$$

$$\begin{bmatrix} \overline{F} \\ \overline{G} \end{bmatrix} = \frac{-1}{R^2}\, \overline{Q}^{\,\tau}\, P \begin{bmatrix} -D \\ C \end{bmatrix}$$

For chord distance:

$$(96) \quad \begin{bmatrix} F \\ G \end{bmatrix} = \frac{-1}{S}\, Q^\tau \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}$$

$$\begin{bmatrix} \overline{F} \\ \overline{G} \end{bmatrix} = \frac{-1}{S}\, \overline{Q}^{-\tau} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} \quad ; \quad R^2 = C^2 + D^2$$

The P matrix is Q evaluated with astronomic coordinates; S is the chord distance and:

$$(97) \quad \begin{bmatrix} C \\ D \end{bmatrix} = P^\tau \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}$$

The implications for horizontal network adjustments are several. First and foremost we are not required to carry out reductions to the ellipsoid of our data. Second it is possible to formulate a computer program (Vincenty, 1979b, 1980a) that can be significantly faster than the classical network adjustment program because of a considerable reduction in computational effort.

4.6   Summary and Conclusions

The techniques developed in this chapter allow the incorporation of a variety of different measurements into a coherent system relating position information to this data. Thus we have achieved the goal of determining simultaneously the vertical and horizontal position of a point in a geodetic network.

The observation equations were developed with either rectangular (x, y, z) coordinates, or geodetic (φ, λ, h) coordinates as the unknowns, along with the astronomic values of φ', λ'. The choice of which type of unknown to use should be arbitrary because either choice should yield the same final coordinates. In actual application of these equations it may be convenient to alter the units so that angular unknowns are in seconds of arc, for example. Careful attention needs to be paid to magnitudes of various numbers so that significant digits are not lost, which may cause errors on the observation equations or instability in the normal equation matrix to be inverted. Since there are 5 to 7 unknowns or more per station, the matrix to be inverted in a three dimensional geodesy adjustment is considerably larger than that found for a corresponding horizontal network in a classical adjustment. In some respects this represents a disadvantage of this type of model, but it is a penalty we have to pay for the consistency that we obtain.

One interesting point related to our adjustment model is that we can obtain the astronomic latitude and longitude of a point without having to make actual astronomic observations at the point. Unfortunately the determination of φ' and λ' in this manner is not accurate because such determinations depend primarily on the vertical angle measurements. Using realistic vertical angle standard deviations, the simulation studies of Fubura (1972) showed average standard deviations of φ' and λ' varying from 4" to 20" at points where no astronomic determination of φ'' and λ' were made.

We note that the method of three dimensional geodesy can easily be incorporated with satellite determinations of the coordinates of points on the surface of the earth. If such coordinates are given as x, y, z values the adjustment of the three dimensional network can proceed with the rectangular coordinates being considered as data with an a priori known variance-covariance matrix. A similar procedure could be used with φ, λ and h simply by adopting a reference ellipsoid and finding the φ, λ, h values corresponding to the satellite derived x, y, z coordinates. Alternate to using the a priori weighting of the parameters (φ, λ, h or x, y, z) separate observation equations for each coordinate might be used. This may cause some problems because of the correlation between the determinations of φ, λ and h using satellite techniques.

173

Several computer programs for the evaluation of three dimensional networks have been published or described (Vincenty, 1979a, Sikonia, 1977).

Large continental networks will probably never be adjusted by three dimensional geodesy techniques. However its applicaton to special net adjustments seem to be a much more feasable approach. Such applications are described by Torge and Wenzel (1978), Carter and Pettey (1978) and in Sikonia (1977).

We should note that the discussion in these pages has generally ignored information on the gravity field and equipotential surfaces to some extent. There are more general theories available that may be useful in certain cases. These cases are discussed by several authors including Moritz (1978), Grafarend (1980) Reilly (1980).

References

Bomford, G., Geodesy, Oxford at the Clarendon Press, 3rd edition, 1971.

Bouricius, G.M.S., and K.B. Earnshaw, Results of Field Testing a Two-Wavelength Optical Distance-Measuring Instrument, J. Geophys. Res., 79, 3015-3018, 1974.

Bowring, B., Notes on Space Adjustment Coefficients, Bulletin Geodesique 54, 191-199, 1980.

Brazier, H., and L. Windsor, A Test Triangulation, paper presented to Study Group No. 1, Toronto Assembly of the International Association of Geodesy, 1957.

Bruns, H., Die Figur der Erde, Publ. of the Preuss. Geod. Inst., 1878.

Carter, W., and J. Pettey, Goldstone Validation Survey - Phase I, Noaa Technical Memorandum NOS NGS-15, National Geodetic Survey, Rockville, Md. 20851, 1978.

Chovitz, B., Three-Dimensional Model Based on Hotine's "Mathematical Geodesy", presented at the International Symposium on Problems Related to the Redefinition of North American Geodetic Network, Fredericton, N.B., Canada, 1974.

Fubura, D., Three-Dimensional Geodesy for Terrestrial Network Adjustment, J. Geophys. Res., 77, 796-807, 1972.

Grafarend, E., The Bruns Transformation and a Dual Setup of Geodetic Observational Equations, NOAA Technical Report NOS NGS 16, Rockville, MD April, 1980.

Hotine, M., A Primer of Non-Classical Geodesy, paper presented to Study Group No. 1, Toronto Assembly of the International Association of Geodesy, 1957.

Hotine, M., Mathematical Geodesy, ESSA Monograph 2, 1969, COM-74-11127.

Hradilek, L., Trigonometric Levelling and Spatial Triangulation in Mountain Regions, Bulletin Geodesique, 87, 33-52, 1968.

Hradilek, L., Refraction in Trigonometric and Three-Dimensional Networks, Canadian Surveyor, 26, 59-70, 1972.

Hradilek, L., Weighting of Vertical Angles, Studia geoph. et geod., 17, 169-172, 1973.

Körner, H., Die Berechnung einer Raumtriangulation im System der Dreidimensionalen Geodäsie, Inaugural-Dissertation, University of Bonn, October, 1968.

Mikhail, E.M., Parameter Constraints on Least Squares, Photogrammetric Engineering, 36 (12), 1277-91, 1970.

Mitchell, M., Fundamentals of Spatial Coordinates in Terrestrial Geodesy, M.S. Thesis, The Ohio State University, Columbus, 1963.

Moritz, H., The Operational Approach to Physical Geodesy, Dept. of Geodetic Science Report No. 277, Oct. 1978.

Pfeifer, L., A Treatment of Vertical Refraction in the Adjustment of Three-Dimensional Geodetic Networks, unpublished manuscript, NOAA, Rockville, Maryland, 1973.

Prilepin, J.T., Some Problems in the Theory of Determining Geodetic Refraction by Dispersion Method, Bulletin Geodesique, 108, 115-123, 1973.

Ramsayer, K., Strenge und genäherte Ausgleichung von Raumnetzen in einem Cokalin kartisischen Koordinatensystem und Erprobung in den Raumnetzen Heerbrugg und Stuttgart, Deutsche Geod. Komm., Reihe A, Höhere Geodäsie-Heft Nr. 71, München, 1971.

Ramsayer, K., Von der zwei-zur dreidimensionalen Geodäsie, Allgemeine Vermessunge-Nachrechten, 79, 334-349, 1972.

Rapp, R.H., Geometric Geodesy, Volume I (Basic Principles), Dept. of Geodetic Science, Ohio State University, Columbus, p. 134, 1974.

Reilly, E. I., Three-Dimensional Adjustment of Geodetic Networks with the Incorporation of Gravity Field Data, Report No. 160, Geophysics Division, Dept. of Scientific and Industrial Research, Wellington, New Zealand, June, 1980.

Saito, T., The Determination of the Curvature of Ray Path for the Adjustment of a Geodetic Traverse, J. Geodetic Society of Japan, 20, 1-17, 1974.

Sikonia, W., Three Dimensional Geodetic Survey Adjustment USGS, National Technical Infor. Service, Springfield, VA 22152, PB278660, 1977.

Tengström, E., Elimination of Refraction at Vertical Angle Measurements, Using Casers of Different Wavelengths, Proceedings of the International Symposium Figure of the Earth and Refraction, Vienna, Austria, 292-303, March 14-17, 1967.

Torge, W. and H-G Wenzel, Dreidemensionale Ausgleichung Des Testnetzes Westharz, Deut. Geod. Kormmin, Reihe B, Angewandte Geodäsie, Hefta Nr. 234, München, West Germany, 1978.

176

Vincenty, T., Spatial Adjustments of Geodetic Networks in Practice, unpublished manuscript dated August 27, 1974.

Vincenty, T. and B. Bowring, Application of Three-Dimensional Geodesy to Adjustments of Horizontal Networks, in proceedings Second Int. Symp. on Problems Related to the Redefinition of North American Geodetic Networks, Superintendant of Documents, U.S. Govt. Print. Office, Washington, D.C. (#003-017-0426-1), 1978; (also NOAA Technical Memorandum NOS NGS-13, National Geodetic Survey, Rockville, Md. 20842 1978)

Vincenty, T., The HAVAGO Three Dimensional Adjustment Program, NOAA Technical Memorandum NOS NGS 17, Rockville, Md. 20842, 1979a.

Vincenty, T., A Program for Adjusting Horizontal Networks in Three Dimensions, NOAA Technical Memorandum NOS NGS 19, Rockville, MD. 208521, 1979b.

Vincenty, T., Revisions of the HOACOS Height-Controlled Network Adjustment Program, NOAA Technical Memorandum NOG NGS25, Rockville, Md. 208521, 1980a.

Vincenty, T., Height-Controlled Three-Dimensional Adjustment of Horizontal Networks, Bulletin Geodesique, 54, 37-43, 1980b.

Wolf, H., Die Grundgleichungen der Dreidimensionalen Geodäsie in elementarer Darstellung, Zeitschrift für Vermessungswesen, Volume 88, No. 6, 225-233, 1963.

Appendix A

Historical Information on the Development of Major Horizontal Geodetic Datums

The following is taken from the "NASA Directory of Station Locations"
prepared by T. Gunther of the Computer Sciences Corporation, February 1978.

SECTION 3 - DEVELOPMENT OF THE MAJOR GEODETIC DATUMS

3.1 INTRODUCTION

Much of the inhabited area of the world is covered with geodetic networks con-
sisting mostly of triangulation, although some are in the form of traverse sur-
veys such as those established by Australia in the 1960s or Shoran trilateration
as established by Canada in the 1950s. The most notable voids of great extent
are the interior of Brazil; portions of west, central, and northern Africa; much
of China; and northern Siberia.

These geodetic operations date back to the last part of the 18th century, and it
was common practice from that time to the early 20th century to employ sep-
arate origins or datums in each country, and even more than one origin in some
countries, e.g., the United States. Even in the early days astronomically
determined latitudes were rather easily established as one coordinate of the
origin. But longitudes were another matter for two reasons: (1) there is no
natural common plane of reference like the equator for latitude, and (2) even
if a common plane, such as that of the Greenwich meridian, were agreed upon,
there was no accurate method of observing longitude before the electric tele-
graph and the associated lines of transmission, including submarine cables,
were developed.

The longitude problem taxed the ingenuity of the astronomers in the first half
of the 18th century. Lunar culminations, occultations, and distances were
observed along with solar eclipses in an attempt to determine differences of
longitude of widely separated points. These methods depended on "fixing" the
Moon as it moves among the stars, but because of the relatively slow movement
of the Moon among the stars and the irregularity of the Moon's limb, this ap-
proach was inherently inaccurate. It gave way to the transportation of chron-
ometers for timing observations of the stars. This method, which reached its
peak about the middle of the 19th century, was replaced by telegraph and, later,

radio time signals. With the recent development of portable crystal and atomic clocks, transportation of time is again in use for correlations of the highest precision.

In the early days, longitudes of a geodetic system were often based on the position of an astronomic observatory situated in or near the capital city of a country. A reference ellipsoid was chosen for the datum, and the latitudes and longitudes of all other geodetic points were derived by computation through the triangulation. This meant that the many datums, computed on different ellipsoids and based on astronomic observations at separate origins, were not accurately related to each other in a geodetic sense, although the astronomic latitudes were of high caliber.

There was a slow trend toward accepting the Greenwich meridian as the basis for longitude, and by 1940 practically all important geodetic networks were based on it. But there still remained the separate geodetic datums employing a variety of ellipsoids and methods for determining the coordinates of the origins. The only computations of extensive geodetic work of an international nature, based on a single datum, were those for long arcs done in an effort to improve the knowledge of the size and shape of the Earth.

Since World War II, much has been accomplished in combining separate datums on the continents and in relating datums between the continents. The advent of artificial satellites has made possible the tremendous task of correlating all datums and, ultimately, of placing all geodetic points on a single, worldwide geodetic system. The first step in this process, taken after World War II, was the selection of several so-called "preferred datums," into which many local geodetic systems were reduced. The more important datums appear on the accompanying map.
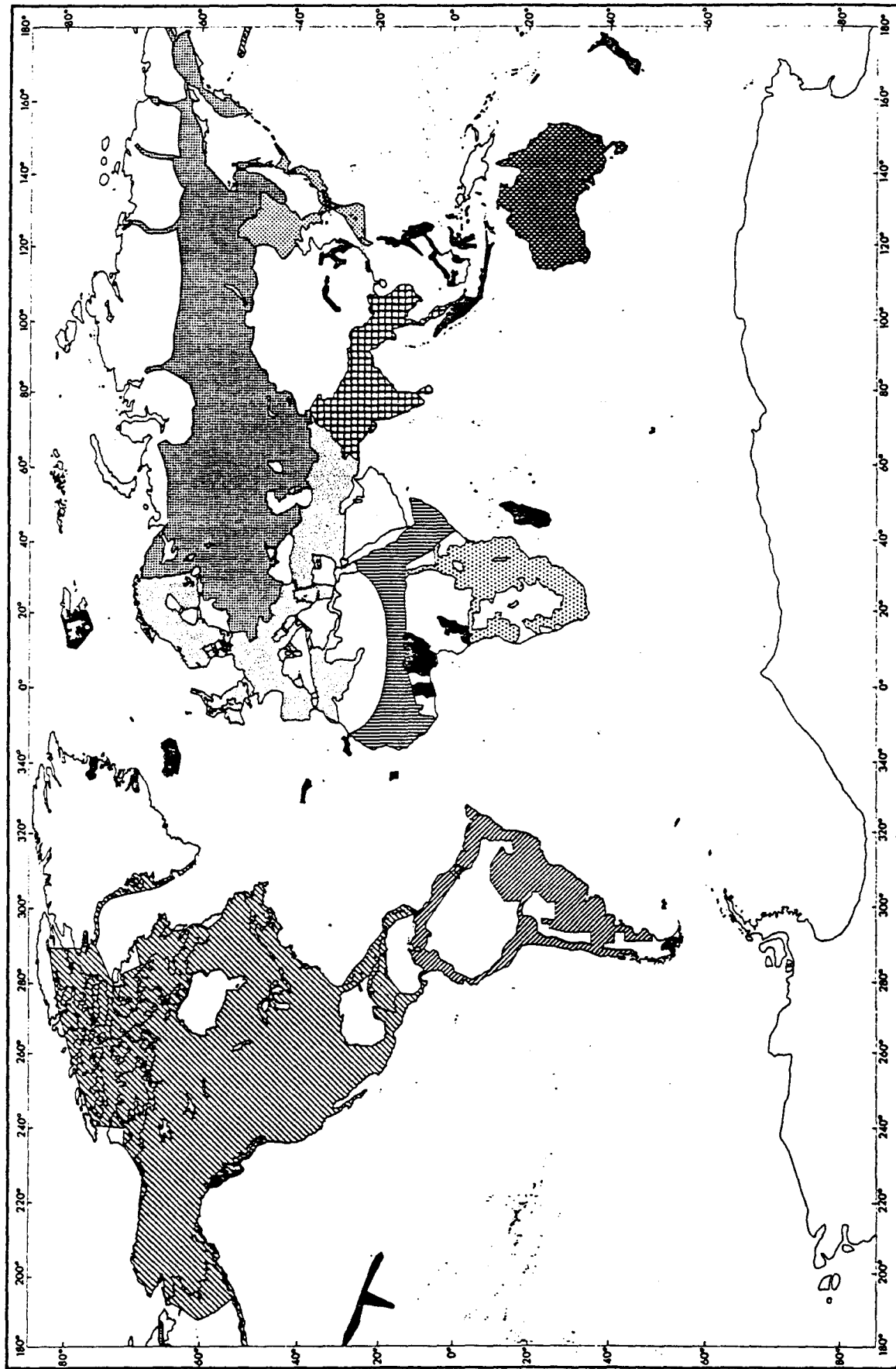
DATUMS

North American	Tokyo	Adindãn	Australian National	Pulkovo 42
European	Cape (Arc)	Indian	South American	Independent

MAJOR GEODETIC DATUM BLOCKS

3.2 THE NORTH AMERICAN DATUM OF 1927

Most extensive of the preferred datums, the North American Datum of 1927 is the basis of all geodetic surveys on the North American Continent. This datum is based ultimately on the New England Datum, adopted in 1879 for triangulation in the northeastern and eastern areas of the United States. The position of the origin of this datum, station Principio in Maryland, was based on 58 astronomic latitude and 7 astronomic longitude stations between Maine and Georgia.

At the turn of the century, when the computations for the transcontinental triangulation were complete, it was feasible to adopt a single datum for the entire country. Preliminary investigation indicated that the New England Datum might well serve as a continental datum. In 1901 the New England Datum was officially adopted and became known as the United States Standard Datum. A subsequent examination of the astrogeodetic deflections available at that time at 204 latitude, 68 longitude, and 126 azimuth stations scattered across the entire country indicated that the adopted datum approached closely the ideal under which the algebraic sum of the deflection components is zero.

A later test was applied to the U.S. Standard Datum. Using Hayford's observation equations based on astronomic observations for 381 latitude, 131 longitude, and 253 azimuth stations available in 1909, a solution was made for the shift at Meades Ranch, the chosen datum point, to best satisfy the observed data. Observed deflections uncorrected for topography were used, and the elements of the Clarke Spheroid of 1866 were held fixed. The computed corrections to the latitude and longitude were, respectively, only 0!'41 and 0!'11. In 1913, after Canada and Mexico had adopted the U.S. Standard Datum as the basis for their triangulation, the designation was changed to "North American Datum" with no difference in definition.

Beginning in 1927, a readjustment was made of the triangulation in the United States, and the resulting positions were listed on the North American Datum of 1927. In this readjustment, the position of only Meades Ranch was held

fixed. As a matter of fact, this is really all that sets Meades Ranch apart from all other triangulation stations. Its choice as the datum origin was purely arbitrary and was made because it was near the center of the United States and at the intersection of the Transcontinental and 98th Meridian Arcs of the triangulation. The deflection at Meades Ranch is not zero, as is sometimes assumed; in fact, it was not determined until the late 1940s. Its deflection components in the meridian and prime vertical are, respectively, approximately -1!'3 and +1!'9, in the sense astronomic minus geodetic, with latitude and longitude measured positively north and east.

Loop closures and corrections to sections in the 1927 readjustment of the triangulation in the United States indicate that distances between points separated by at least 2000 kilometers are determined to an accuracy of 5 parts per million, and transcontinental distances are known to 4 parts per million. Gravimetric and other studies suggest that the position of the datum origin is within 1 arc-second in an absolute sense, and recent satellite triangulation indicates an accuracy of better than 1 arc-second in the overall orientation of the 1927 adjustment. (These statements do not necessarily apply to the extension of the North American Datum of 1927 into Mexico, Canada, and Alaska.)

In summary, the North American Datum of 1927 (NAD 27) is defined by the following position and azimuth at Meades Ranch: latitude 39° 13' 26!'686 N, longitude 98° 32' 30!'506 W, azimuth to Waldo (from south) 75° 28' 09!'64.

Although a geodetic azimuth is included in the fundamental data of Meades Ranch, this is of only minor importance, since the orientation of the triangulation is controlled by many Laplace azimuths scattered throughout the network. The latitude is based on 58 astronomical latitude stations, the longitude is based on 7 astronomical longitude stations, and the azimuth is based on nearby Laplace azimuth control. The basis for computations is the Clarke Spheroid of 1866. All measured lengths are reduced to the geoid (mean sea level), not to the spheroid.

Revision of NAD 1927 is long overdue. Local distortions of 10 arc–seconds in azimuth are known to exist, and closures within limited areas may be as poor as 1/20,000. An entirely new adjustment, which will include geodimeter and satellite observations, is underway. When completed in 1983, it is expected to have an overall accuracy of $1/10^6$, with errors between adjacent stations no greater than $1/10^5$, an improvement in accuracy by a factor of 3 or 4.

3.3 EUROPEAN DATUM (1950)

Until 1947 each country in Europe had established its own triangulation, computed on its own datum, which usually consisted of a single astronomic latitude and longitude of a selected origin. Moreover, at least three different spheroids were used. This situation, coupled with the inevitable accumulation of errors in the networks, led to differences at international boundaries of nearly 500 meters in extreme cases.

Although considerable thought was given to unification of the European triangulation, no results became available until after World War II. For several years before the war, extensive surveys were conducted to connect many separate national triangulations; thus, the groundwork was laid for a general adjustment of the major European networks. Under the supervision of the U.S. Army Map Service and with the assistance of the U.S. Coast and Geodetic Survey, the Land Survey Office at Bamberg, Germany, commenced the adjustment of the Central European Network in June 1945 and completed it 2 years later. This triangulation network roughly covers the region that lies between 47° to 56° North latitude and between 6° and 27° East longitude, and is generally in the form of area, rather than arc, coverage. The basis for the computation is the International Ellipsoid.

To expedite the work, triangles were selected to form a few strong arcs of the parallel and meridian to build a network susceptible of the Bowie junction method of adjustment. A scheme was selected which included 23 junction figures, each of which contained at least one base line and one Laplace azimuth.

A total of 52 base lines and 106 Laplace azimuths scaled and oriented the Central European Network.

The datum of this network depends on the study of 173 astronomic latitudes, 126 astronomic longitudes, and 152 azimuths, of which 106 are the Laplace type. No one station can be logically designated as the datum point. The Central European Datum has been referred to as a "condition of the whole," not to any single point. However, as a matter of convenience, Helmert Tower near Potsdam is often referred to as the origin for comparison of the Central European Datum with other datums.

The Central European Network was extended by the addition of two separate adjustments of large networks of triangulation known as the Southwestern Block and the Northern Block. The Central Network was substantially held fixed and, with the addition of the two blocks, forms the European Triangulation based on what is now designated as the European Datum 1950.

The Southwestern Block comprises 1230 triangulation stations in Belgium, France, Spain, Portugal, Switzerland, Austria, Italy, and North Africa; the Northern Block includes 822 stations in Finland, Estonia, Latvia, Denmark, Norway, and Sweden. As in the Central European Adjustment, arcs were selected and adjusted in loops, not by the Bowie junction method but by a modified simultaneous approach. Triangle and loop closures indicate that, on the average, the accuracy of the Central Network and the Northern Block of triangulation is somewhat greater than that in NAD 1927, possibly 3 parts per million for determination of distances of several hundred kilometers. On the average the accuracy of the Southwestern Block is not as high, probably nearer 5 or 6 parts per million. These are average estimates: the accuracies vary considerably within the blocks. There is no evidence that any of the base lines were reduced to a common spheroid, certainly not to the International Ellipsoid.

Since the completion of the original adjustment of the European triangulation networks, the European Datum has been connected to work in Africa and, upon

completion of the 30th Meridian Arc, as far as South Africa, as well as to the Indian Datum through ties made in the Middle East. It is also possible by computation to carry the European Datum to the North American Datum of 1927 by way of the North Atlantic Hiran connection.

It has long been apparent that the European 50 adjustment falls short of meeting current needs. In 1954 the International Association of Geodesy initiated a more rigorous combination of the triangulations of Europe. Called RETrig, it is being undertaken in three phases. In Phase I, completed in 1975, the national nets were independently adjusted after being strengthened with newly observed distances and directions. RETrig II will be a quick computation of the adjusted junction points between the national nets, with the addition of long base lines and Laplace azimuths. Phase III will follow the procedure of Phase II, but using all scientific and mathematical sophistication available. The results will be compared with satellite solutions and may be blended with them.

All stations will be reduced to the International Spheroid (using the 1975 geoid of Levallois and Monge), from which transition to a world datum can be made. An inverse solution from the adjusted junction points will position the stations within the national blocks. Completion of the adjustment, perhaps in the 1980s, should fix the area of Western Europe with precision and stability.

3.4 INDIAN DATUM

A brief history of the Great Trigonometric Survey of India and of the Indian Datum is of interest, if for no other reason than that geodetic operations were commenced at such an early date in an area so remote from any similar activity and from the country responsible for conducting them. Operations were begun about 1802, and the Madras Observatory was first selected as the origin of the trigonometric coordinates because it was the only institution equipped with precision instruments.

It was, however, many years before any real progress was made on what is now known as the primary triangulation. Colonel George Everest, who was appointed Surveyor General of India in 1830, decided in 1840 to adopt as the origin the triangulation station at Kalianpur H.S. This station was selected because it is on a broad plateau at what was thought to be a safe distance from the Himalayan mass and its adverse effect on the plumb line.

In 1847 a value of 77° 41' 44''75 E was accepted as the astronomic and geodetic longitude at Kalianpur. It was based on a preliminary value of the position of Madras Observatory. But in 1894-1895, a reliable determination of the longitude of Karachi was made possible by telegraphic observations, and it was learned that the Indian longitudes should be corrected by -2' 27''18. Thus, the corrected longitude at the origin is 77° 39' 17''57 E. But because this was considered as the astronomic longitude and a deflection of +2''89 in the prime vertical had been adopted, a further correction to the geodetic longitude was needed to maintain this deflection. These modern longitudes were introduced in India in 1905; prior to this, the mapping longitudes of India were off by about 4 kilometers.

The first comprehensive adjustment of the Indian triangulation was undertaken about 1880. There were no Laplace stations in the strict sense of the word at this time, but expedients were adopted to approximate the Laplace correction from telegraphic longitudes available at certain cities. There appear to have been only about 11 base lines at the time.

After the recommendation of the International Spheroid by the I.U.G.G. in 1927, it was decided to use this spheroid in India for scientific purposes. The Everest Spheroid which was used had long been known to be unsuitable. A least squares solution was accomplished to best fit the geoid in India to the International Spheroid. In this adjustment the deflections at Kalianpur were +2''42 and +3''17 in the meridian and prime vertical, respectively, and the geoid height was 31 feet. In 1938 a detailed adjustment of the Indian triangulation was made on

the Everest Spheroid, but it lacked the rigor of least squares; it employed detailed diagrams of misclosures in scale, azimuth and circuit closures, and personal judgment in the distribution of these errors of closure.

The Indian work comprises about 9400 miles of primary arcs of triangulation and nearly as many more miles of secondary arcs. In the primary work, the mean square error of an observed angle ranges among the various sections from 0''15 to 1''00, and averages about 0''5. Thus the angle observations are of very high caliber, but the number of base lines and Laplace azimuths is deficient. There are now about 127 Laplace stations available in India, which will greatly strengthen any future readjustment of the work. Before this is done, however, the plan is to raise the accuracy of the secondary work to primary standards by reobservation and to provide additional work in many of the existing gaps.

It has been the custom in India to give the deflections rather than the position coordinates at the origin. For Kalianpur, in the 1938 adjustment, these were −0''29 in the meridian and +2''29 in the prime vertical (a plus sign indicates that the plumb line is south or west of the spheroid normal). The geoid height is zero at the origin by definition. The spheroid is the Everest: f = 1/300.8017, a = 6 377 301 meters. This value for a, used in India and Pakistan, is based on the ratio of 0.304 7996 meters to the Indian foot, rather than Benoit's ratio (0.304 798 41), for which a = 6 377 276 meters. The Benoit ratio continues to be used in U.S. and U.K. tables for historical convenience, with a scale factor introduced when appropriate.

3.5 TOKYO DATUM

The origin of the Tokyo Datum is the astronomic position of the meridian circle of the old Tokyo Observatory. The adopted coordinates were: latitude 35° 19' 17''5148 N, longitude 139° 44' 40''9000 E; reference surface: Bessel Spheroid, 1841. The latitude was determined from observations by the Tokyo Observatory, and the longitude by the Hydrographic Department of the Imperial Navy by

telegraphic submarine cable between Tokyo and the United States longitude station at Guam. This datum is known to be in considerable error as related to an ideal world datum because of large deflections of the plumb line in the region of Tokyo.

The primary triangulation of Japan proper consists of 426 stations and 15 base lines established between 1883 and 1916. The mean error of an observed angle is 0''66, which is roughly equivalent to a probable error of 0''3 as applied to an observed direction. This puts the accuracy of the work about on a par with that of the United States in this respect.

After completion of the primary work in Japan proper, the Tokyo Datum was extended in the mid-1920s into the Karahuto portion of Sakhalin. The Manchurian triangulation, established by the Japanese Army after 1935, has been connected through Korea to the Tokyo Datum. The quality of the primary triangulation in Korea and Manchuria is believed to be about, though not quite, equal to that of Japan proper.

3.6 AUSTRALIAN GEODETIC DATUM (1966)

Until 1961 the spheroid generally used in Australia was the Clarke of 1858. Because the triangulation in Australia was initiated in several separate areas, there were several distinct origins rather than a single national datum. The most important were Sydney Observatory, Perth Observatory - 1899, and Darwin Origin Pillar.

During the early 1960s an ambitious geodetic survey was started to establish complete coverage of the continent and connect all important existing geodetic surveys. For a short period of 1962 computations were performed on the so-called "NASA" Spheroid (a = 6 378 148 meters; f = 1/298.3) with the origin at Maurice, but these have been completely superseded. The first comprehensive computation of the new geodetic survey was made on the "165" Spheroid (a = 6 378 165, f = 1/298.3). This was based on the "Central Origin," in use since

1963, and depended on 155 astrogeodetic stations distributed over most of Australia except Cape York and Tasmania.

It appeared at this time that there might be international agreement on one spheroid, which Australia might adopt officially. Many modern determinations had been made for which the ranges in a and f were so narrow as to have no practical significance. On the strength of the acceptance of a spheroid by the International Astronomical Union, it was adopted in April 1965 as the Australian National Spheroid, with the only difference that the flattening of the spheroid used for astronomy was rounded to 1/298.25 exactly. The semimajor axis is 6 378 160 meters.

Holding the Central Origin, which was defined by the coordinates of station Grundy, a complete readjustment of the geodetic network was made in 1966, using the Australian National Spheroid. The mean deflection, uncorrected for topography, at 275 well-distributed stations was +0!'12 in meridian and -0!'33 in prime vertical. Although the Central Origin has in effect been retained, instead of being defined as originally in terms of station Grundy it is now defined by equivalent coordinates for the Johnston Geodetic Station. These are: latitude 25° 56' 54!'5515 S, longitude 133° 12' 30!'0771 E. The geoid separation at this point is -6 meters, as of 1 November 1971.

A study of the observations of satellite orbits indicates there is a rather uniform and relatively heavy tilt of the geoidal surface over Australia, which would introduce a bias to the astrogeodetic deflections determined on the Australian Geodetic Datum (AGD) of 4!'7 and 4!'4 in the meridian and prime vertical, respectively. This tilt is in such a direction that the astronomic zenith is pulled approximately 6!'5, on the average, southwest of where an ideal or absolute geodetic zenith would be.

The survey net for AGD 1966 consists of 161 sections which connect 101 junction points and form 58 loops. Virtually all the surveys were of the traverse type

in which distances were determined by Tellurometer. There are 2506 stations, of which 533 are Laplace points, and the total length of the traverse is 53,300 kilometers.

Measured lengths were reduced to the geoid, not the spheroid, because of lack of knowledge of the separation of these surfaces at the time of the general adjustment. Development of the geoid for the continent by 1971 showed its effect on the adjustment to be insignificant. The method of adjustment may briefly be described as follows: each section was given a free adjustment by which the length and azimuth between the end points were determined; these lengths and azimuths were then put into a single adjustment to determine the final coordinates of the junction points connected by the sections; each section was then adjusted to the final coordinates of the pertinent junctions. The average loop length is about 1500 kilometers; the average closure is 2.2 parts per million, with a maximum closure of 4.3 parts per million.

Tasmania has been connected by two new sections across Bass Strait via King and Flinders Islands. A connection to New Guinea and the Bismarck Archipelago has been effected by a Tellurometer traverse up Cape York and the USAF Hiran network of 1965. A large section in eastern New South Wales and the Australian Capital Territory has been strengthened and adjusted into AGD 1966. Similar substitution of new work into the AGD is planned in Victoria, around Adelaide, and around Perth.

While AGD 1966 remains the basis for normal surveying and mapping at this time, in 1973 a new adjustment called the Australian Geodetic Model 1973 was made. It incorporated many new observations, including more accurate heights, accurate geoid-spheroid separation, and more recent high-precision traverses, for which the AGA 8 Geodimeter is (with the Wild T3) the principal instrument. Comparison with AGD 1966 showed no shift at any station of as much as 5 meters. Annual mathematical readjustments using all suitable data, including satellite observations, continue to be made.

A-14

3.7 SOUTH AMERICAN DATUM

By 1953 the Inter-American Geodetic Survey of the U.S. Corps of Engineers had completed the triangulation from Mexico through Central America and down the west coast of South America to southern Chile. This was done in cooperation with the various countries through which the work extended, and marked the completion of the longest north-south arc of triangulation ever accomplished. It had an amplitude of over 100 arc degrees through North and South America.

In 1956 the Provisional South American Datum was adopted as an interim reference datum for the adjustment of the triangulation in Venezuela, Colombia, and the meridional arc along the West Coast. Instead of depending on one astronomic station as the origin and assuming its deflection components to be zero, or attempting to average out the deflections at many astronomic stations by the astrogeodetic method, one astronomic station was chosen as the datum origin, but its deflection components were determined gravimetrically. The gravity survey covered an area about 75 kilometers in radius centered on the origin, station La Canoa in Venezuela. The reference figure was the International Ellipsoid, and the geoid height at La Canoa was zero by definition. A major portion of the South American work was adjusted on the Provisional South American Datum, including the extensive Hiran trilateration along the northeast coast of the continent. The principal exceptions were the networks in Argentina, Uruguay, and Paraguay.

Considering the geographic location of La Canoa, with all of the continent on one side and the Puerto Rican ocean trench on the other, the gravity coverage was insufficient to produce a deflection for a continentally well-fitting datum. From the astrogeodetic deflections based on this datum it can be inferred that the geoid drops about 280 meters below the spheroid in Chile at latitude 41° South. This drop is more or less uniform in a southerly direction for a distance of roughly 5500 kilometers. In 5500 kilometers, 280 meters is very

A-15

nearly 10 seconds of arc; such a correction to the meridian deflection compo-
nent at La Canoa would produce a better fit of the International Ellipsoid to the
area of the South American adjustment. But the La Canoa Datum has not been
corrected for this large and increasing geoidal separation, and thus contains
large distortions. For example, cross-continental distances may be several
tens of meters too short. In addition, the Hiran net has also been shown to be
tens of meters too short.

An investigation of the astrogeodetic data from the long meridional arc in the
Americas and the 30th Meridian Arc from Finland to South Africa led to the
conclusion that the equatorial radius of the International Ellipsoid should be
reduced by at least 100 meters (a subsequent change in the flattening inferred
from satellite observations suggested another 100-meter reduction), and that
the North American and European Datums were not at all well suited for the
continents to the south. Thus it became apparent that consideration must be
given to the selection of another datum for South America.

A Working Group for the Study of the South American Datum was asked in 1965
by the Committee for Geodesy of the Cartographic Commission of the Pan
American Institute of Geography and History to select a suitable geodetic datum
for South America and to establish a coherent geodetic system for the entire
continent. This was achieved, and the South American Datum 1969 (SAD 1969)
was accepted by the Cartographic Commission in June 1969 at the IX General
Assembly of PAIGH in Washington, D.C. This new datum is computed on the
Reference Ellipsoid 1967, accepted by the International Union of Geodesy and
Geophysics in Lucerne in 1967, with the minor difference that the flattening is
rounded (a = 6 378 160 meters, f = 1/298.25 exactly). Both Chua and Campo
Inchauspe, the National datum points of Brazil and Argentina, respectively,
were assigned minimal geoid heights (0 and 2 meters). Chua is taken to be
the nominal origin. A vast amount of recent triangulation, Hiran, astronomic,
and satellite data were incorporated in the solution, and SAD 1969 now provides
the basis for a homogeneous geodetic control system for the continent.

A-16

3.8 ARC DATUM (CAPE)

The origin of the old South African, or Cape, Datum is at Buffelsfontein. The latitude at this origin was adopted after a preliminary comparison of the astronomic and geodetic results, rejecting those stations at which the astronomic observations were probably affected by abnormal deflections of the plumb line. The longitude of this origin depends upon the telegraphic determination of longitude of the Cape Transit Circle, to which was added the difference of geodetic longitude computed through the triangulation. Computations were based on the modified Clarke Spheroid of 1880. The geodetic coordinates of Buffelsfontein are latitude 33° 59' 32".000 S, longitude 25° 30' 44".622 E.

Over the years this datum has been extended over much of south, east, and central Africa. Through the 30th Meridian Arc, completed in the 1950s, it has been connected to the European Datum. Because the 30th Meridian Arc is the backbone of this work, which also includes triangulation in Zaire and former Portuguese Africa, the published geodetic coordinates are now referred to the Arc Datum. The whole comprises a uniform system from the Cape to the equator.

The accuracy of the South African work and of the 30th Meridian Arc compares favorably with that of the other major systems of the world, but some of the related triangulation requires additional length control and Laplace azimuths.

3.9 PULKOVO DATUM 1942

The development of the triangulation network in the U.S.S.R. parallels to some extent the development of the network in the United States. The Russian work began in 1816 in the Baltic states, and was gradually extended by the Corps of Military Topographers (KTV) as well as by provincial organizations. An important early accomplishment was the establishment of the Struve-Tenner arc of the meridian from Finland to the mouth of the Danube, the results of which were used for figure-of-the-Earth studies.

These early surveys were established independently and were based on different ellipsoids and datum points. By the turn of the century, over 20 independent sets of coordinates were in use. About this time the first effort was made to unify the many systems and place them on the Bessel Ellipsoid, with the Tartu Observatory as the initial point. Not much was done until a new plan was formulated by the KTV in which arcs of triangulation were to be observed along parallels and meridians, spaced from 300 to 500 kilometers, with Laplace azimuths and base lines at their intersections. The Bessel Ellipsoid was chosen again, but the initial point was changed to the Pulkovo Observatory. The coordinates assigned to Pulkovo are now referred to as the Old Pulkovo Datum.

This plan was implemented in 1910 and, after interruption by World War I and the Revolution, was pursued vigorously until 1944, at which time 75,000 kilometers of arc and associated astronomic observations and base lines were completed. In 1928 Professor Krassovski was commissioned to augment the original plan. He called for closer spacing of arcs, Laplace stations, and base lines, and a breakdown between primary arcs by lower order work. The standards of accuracy were comparable to those in North America.

During this period triangulation had begun in the Far East, and by 1932 two basic datums were in use, both on the Bessel Ellipsoid but with different initial points--Pulkovo and an astronomic position in the Amur Valley of Siberia. The coordinates of Pulkovo were changed slightly (less than 1 second) from those of the Old Pulkovo Datum. When the two systems were finally joined, a discrepancy of about 900 meters in coordinates of the common points naturally developed. This was due principally to the use of the Bessel Ellipsoid, now known to be seriously in error.

In 1946 a new unified datum was established, designated the "1942 Pulkovo System of Survey Coordinates." This datum employs the ellipsoid determined by Krassovski and Izotov and new values for the coordinates of Pulkovo. The ellipsoid is defined by an equatorial radius of 6 378 245 meters and a flattening of

1/298.3. The coordinates of Pulkovo are latitude 59° 46' 18".55 North, longitude 30° 19' 42".09 East of Greenwich. Deflections at the origin are +0".16 and -1".78 in the meridian and prime vertical, respectively.

3.10 BRITISH DATUM

The original primary network of Great Britain was the result of a selection of observations from a large amount of accumulated triangulation done in a piece-meal fashion. The selected network covered the whole of the British Isles, was scaled by two base lines, and was positioned and oriented by observation at the Royal Observatory, Greenwich. The adjustment was accomplished in 21 blocks, computed on the Airy Spheroid.

In the Retriangulation of 1936, only the original work in England, Scotland, and Wales was included. Original stations were used when practicable, and many stations were added, including secondary and tertiary points. The adjustment was carried out in seven main blocks. The scale, orientation, and position were an average derived from comparison with 11 stations in Block 2 (central England) common to the two triangulations. Other blocks were adjusted sequentially, holding fixed previously adjusted blocks. The result, known as OSGB 1936 Datum, has not proved to be entirely satisfactory. No new base lines were included, and subsequent checks with Geodimeter and Tellurometer indicated that the scale of the Retriangulation was not only too large, but varied alarmingly.

To correct this situation a new adjustment has been made, described as the Ordnance Survey of Great Britain Scientific Network 1970 (OSGB 1970 (SN)). This is a variable quantity and consists, at any moment, of the best selection of observations available. It consists now of 292 primary stations connected by 1900 observed directions, 180 measured distances, and 15 Laplace azimuths. Published positions of all orders on the OSGB 1936 Datum (given as rectangular coordinates on the National Grid) are not altered, nor is the grid on Ordnance Survey maps to be changed, under present policy. Initially only the values of

the first-order stations will be available on OSGB 1970 (SN). More accurate conversions to the European Datum became available when Block 6 of the European readjustment was completed.

The Airy Spheroid was used for all three British datums. The origin is the Royal Observatory at Herstmonceux.

3.11 ADINDÂN DATUM

Between 1967 and 1970, a precise traverse was run across Africa roughly following the Twelfth Parallel North. Starting at the Chad-Sudan border, it extended 4654 kilometers of traverse length to Dakar, Senegal, passing through Nigeria, Niger, Upper Volta, and Mali. The portion in Nigeria was done by the U.S. Defense Mapping Agency Topographic Center (USDMATC) in cooperation with the Nigerian Survey Department; the remainder was done by the French Institut Geographique National (IGN) under contract to DMATC, with the cooperation of the countries through which it passed.

All distances were measured with a Geodimeter and checked with a Tellurometer. First-order angles were used. Trigonometric elevations carried between stations were referred frequently to first-order bench marks. Because first-order astronomic observations with a Wild T-4 were made at every other station (about 40-kilometer spacing), a geoid profile across the continent made it possible to adjust the traverse to the spheroid. The final adjustment by DMATC of April 1971 indicates an accuracy of better than one part in 10^6, or nearly that of the U.S. precise transcontinental traverse.

All triangulation, trilateration, and traverse work in Sudan and Ethiopia has subsequently been computed in this datum. The Adindân base terminal Z_Y was chosen as the origin: latitude 22° 10' 07!'1098 N, longitude 31° 29' 21!'6079 E, with azimuth (from North) to Y_Y 58° 14' 28!'45. The Clarke 1880 Spheroid is used (a = 6 378 249.145 meters, f = 1/293.465). Z_Y is now about 10 meters below the surface of Lake Nasser.

3.12 WORLD GEODETIC SYSTEMS

A world geodetic system may be defined as that in which all points of the system are located with respect to the Earth's center of mass. A practical addendum to this definition is usually the figure of an ellipsoid which best fits the geoid as a whole. In such a system the locations of datum origins with respect to the center of mass are expressed by rectangular space coordinates, X, Y, and Z. This implies three more designations to specify the directions of the axes unambiguously. Conventionally, in reference to the Earth-centered ellipsoid, X and Y are in the equatorial plane, with X positive toward zero longitude, Y positive toward 90° East, and Z positive toward North. The relationship between the X, Y, and Z coordinates and the ellipsoidal coordinates of latitude, longitude, and height is expressed by simple transformations.

The preferred datums provide satisfactory solutions to large areas, even continental in extent. The points within each datum are interrelated with a high order of accuracy. Some connections have been made between these datums by terrestrial surveys, but these are often tenuous. Part of the difficulty in extending datum connections is that the chosen spheroid is usually not suitable for areas remote from the datum proper, which results in excessive deflections and geoid separations. These can seriously distort the triangulation if the geoid heights are not taken into account in base line reduction, and even when the geoid heights are taken into account, the result is not satisfactory.

Realizing that a world geodetic system is desirable for scientific purposes, some of which are of a practical nature, geodesists attacked the problem. Observation of satellite orbits from points around the world require better approximations of the coordinates of the observing stations on a world basis; worldwide oceanographic programs demand accurate positioning at sea; and such approaches as Loran C and Doppler satellite navigation need a coherent worldwide geodetic framework.

A brief assessment of the uncertainties in positioning geodetic datums by classical methods may be made by considering the North American Datum of 1927, the European Datum, and the Tokyo Datum. The uncertainties are given in the two-sigma sense, or twice the standard error. Such a figure approaches the outside error, and might be considered a practical limit of uncertainty. The relative positions of the datum points of North America and Europe were probably known within 300 meters, whereas the figure for North America and Tokyo was 600 or 700 meters. The positions of an island determined astronomically at a single point may be in error, in an absolute geodetic sense, by 1 or 2 kilometers.

In recent years the satellite development of world geodetic systems has greatly reduced the uncertainties of the relative positions of the major datums. The goal of the National Geodetic Satellite Program in positioning primary geodetic points with 10-meter accuracy (standard deviation) in an absolute sense was in general achieved, and in 1978 a 1- to 3-meter accuracy is probably possible.

3.12.1 Vanguard

The first operational world geodetic system was the Vanguard Datum, developed by I. Fischer of the U.S. Army Map Service in 1956. It combined the results from the two long individual arcs of 30° East and down the west coast of the Americas with shorter arcs (35°W, 24°E, and Struve's 52°N among them), corrected by geoid heights instead of by deflections. Vanguard was used to position early satellite tracking stations. The Hough spheroid was derived from the study and used for the system (a = 6 378 270 meters, f = 1/297).

3.12.2 Mercury Datum

With an early determination of the Earth's ellipticity (1/298.3) from observations on the Sputnik I and Vanguard satellites, Army Map Service geodesists proposed in 1960 that by minimizing the differences between astrogeodetically and gravimetrically derived geoid heights, the major datums could be placed in

proper relative position. Through terrestrial ties, other datums--South American, Cape, and Indian--could be connected to the system. This datum was selected by NASA to position the original Project Mercury tracking stations, and from this took its name. The semimajor axis for its spheroid is 6 378 166 meters.

In 1968 it was modified to reflect the accumulation of new data, particularly dynamic satellite results, which provide a superior method for determining the relationships of isolated datum blocks to the Earth's center of mass. The Modified Mercury Datum retained the 1/298.3 flattening, but had a shorter semimajor axis (6 378 150 meters). Translation components for 24 datums were published.

3.12.3 SAO Standard Earth III

The Smithsonian Astrophysical Observatory has long been engaged in satellite observations. Its original world network of 12 (now 8) Baker-Nunn cameras is supported with lasers, and the several solutions published since 1966 have been based on increasing amounts and types of data. Orbital elements derived from single photographic observations were strengthened with paired observations for geometric support. Laser data from GSFC and French stations as well as their own contributed to the results. Data from the BC-4 camera network, from individual observatories, and from the Jet Propulsion Laboratory deep space observations have been incorporated in the later solutions. Surface gravity data were utilized for determination of the geopotential.

Solutions C5, C6, C7, and Standard Earth II were followed in 1973 with SAO Standard Earth III. The analysis of satellite data combined with surface measurements has resulted in a reference gravity field complete to 18th degree and order and the coordinates of 90 satellite tracking sites. The values adopted for the reference ellipsoid are: a = 6 378 155 meters, f = 1/298.257.

3.12.4 NWL-9D, -10F

The U.S. Naval Surface Weapons Center (formerly the Naval Weapons Laboratory) has conducted research in satellite geodesy since 1959 in the development of the Navy Navigation Satellite System. Objectives have included connecting the major datums and isolated sites into a unified world system, relating this system to a best fitting Earth-centered ellipsoid, refining the gravity field, and determining the motion of the pole. The system is now used routinely by other domestic and foreign agencies.

Several solutions have been published. The latest, NWL-9D, includes the positions of 40 stations with worldwide distribution and the shifts of 26 datums to the system. Because the longitude origin of the Doppler system is arbitrary, a rotation may be applied to NWL-9D so that it agrees with the gravimetric deflection in longitude of NAD 1927, and a correction may be applied for a discrepancy in scale with respect to independent determinations. The resulting system is termed NWL-10F and is consistent with datum transformations of the DOD WGS-72 system. For NWL-9D, $a = 6\ 378\ 145$ meters, $f = 1/298.25$; for 10F, $a = 6\ 378\ 135$ meters, $f = 1/298.26$.

3.12.5 World Geodetic System 1972

WGS 72 was developed to meet the mapping, charting, and geodetic needs of the Department of Defense. It represents 5 years of data collection; its development involved primarily the U.S. Air Force (USAF), the Defense Mapping Agency, the Naval Weapons Laboratory, and the Naval Oceanographic Office. It is characterized by the formation of a large-scale matrix by combining normal matrices from each of the major data input sets. It is referenced to the WGS 72 Ellipsoid ($a = 6\ 378\ 135$ meters, $f = 1/298.26$).

3.12.6 Spaceflight Tracking and Data Network System

STDN is a worldwide geodetic system with transformations available to most major or local geodetic datums. It is an outgrowth of the Mercury 1960 Datum

and is referenced to its spheroid (a = 6 378 166 meters, f = 1/298.3). Results from Apollo, Mariner-Mars, Landsat (ERTS), GEOS, and other missions have contributed to the definition of the geodetic locations within the system. Continuing analysis of tracking and geodetic data causes revisions to be made to this system as new tracking and geodetic data are obtained and additional geodetic refinements are made. STDN positions are those currently used by NASA for space flight operations and are tabulated in this directory.

3.12.7 Other Systems

In addition to the systems mentioned above, primarily for historical reasons, other solutions have been developed in recent years based on different instrumentation, different satellites, and different mathematical techniques. Among them are the National Geodetic Survey's BC-4 solution, Ohio State University's WN-14, the Goddard Space Flight Center 1973 system, and the Goddard Earth Model (GEM) series up to GEM 9/10. (Reports on these programs were published in the Journal of Geophysical Research of 10 December 1974 and 10 February 1976.) Although differences between the results of the different solutions have narrowed, agreement on a single system is not yet at hand.

Appendix B

WGS-84 Datum Transformation Information

The following information is contained in this Appendix:

1. Transformation Parameters - Local Geodetic Systems to WGS84, Table 7.5 of WGS84 Report.

2. Local Geodetic System to WGS84, Datum Transformation Multiple Regression Equations, NAD27 to WGS84, Table 7.6 of WGS84 Report.

Note: The information in this appendix has been taken from DMA Technical Report TR 8350.2, 1987. The original page numbers have been left on for convenience.

Table 7.5 (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	$\Delta a(m)$	$\Delta f \times 10^4$		$\Delta X(m)$	$\Delta Y(m)$	$\Delta Z(m)$
CANTON ISLAND 1966 Phoenix Islands	International	-251	-0.14192702	4	298	-304	-375
CAPE South Africa	Clarke 1880	-112.145	-0.54750714	5	-136	-108	-292
CAPE CANAVERAL Mean Value (Florida and Bahama Islands)	Clarke 1866	-69.4	-0.37264639	16	-2	150	181
CARTHAGE Tunisia	Clarke 1880	-112.145	-0.54750714	5	-263	6	431
CHATHAM 1971 Chatham Island (New Zealand)	International	-251	-0.14192702	4	175	-38	113
CHUA ASTRO Paraguay	International	-251	-0.14192702	6	-134	229	-29
CORREGO ALEGRE Brazil	International	-251	-0.14192702	17	-206	172	-6

* Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.
** WGS 84 minus local geodetic system

Table 7.5   (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	Δf x 10⁴		ΔX(m)	ΔY(m)	ΔZ(m)
DJAKARTA (BATAVIA) <u>Sumatra Island</u> (Indonesia)	Bessel 1841	739.845	0.10037483	5	-377	681	-50
DOS 1968 <u>Gizo Island (New Georgia Islands)</u>	International	-251	-0.14192702	1	230	-199	-752
EASTER ISLAND 1967 <u>Easter Island</u>	International	-251	-0.14192702	1	211	147	111
EUROPEAN 1950 <u>Mean Value</u> [Austria, Denmark, Finland, France, FRG (Federal Republic of Germany), Gibraltar, Greece, Italy, Netherlands, Norway, Portugal, Spain, and Switzerland]	International	-251	-0.14192702	85	-87	-98	-121

*   Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.
** WGS 84 minus local geodetic system

7-17

Table 7.5 (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	Δf x 10^4		ΔX(m)	ΔY(m)	ΔZ(m)
EUROPEAN 1979 Mean Value (Austria, Finland, Netherlands, Norway, Spain, Sweden, and Switzerland)	International	-251	-0.14192702	22	-86	-98	-119
GANDAJIKA BASE Republic of Maldives	International	-251	-0.14192702	1	-133	-321	50
GEODETIC DATUM 1949 New Zealand	International	-251	-0.14192702	14	84	-22	209
GUAM 1963 Guam Island	Clarke 1866	-69.4	-0.37264639	5	-100	-248	259
GUX 1 ASTRO Guadalcanal Island	International	-251	-0.14192702	1	252	-209	-751
HJORSEY 1955 Iceland	International	-251	-0.14192702	6	-73	46	-86

*   Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.

** WGS 84 minus local geodetic system

Table 7.5   (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	Δf x 10^4		ΔX(m)	ΔY(m)	ΔZ(m)
HONG KONG 1963 / Hong Kong	International	-251	-0.14192702	2	-156	-271	-189
INDIAN / Thailand and Vietnam; Bangladesh, India, and Nepal	Everest	860.655	0.28361368	14 / 13	214 / 289	836 / 734	303 / 257
IRELAND 1965 / Ireland	Modified Airy	796.811	0.11960023	7	506	-122	611
ISTS 073 ASTRO 1969 / Diego Garcia	International	-251	-0.14192702	2	208	-435	-229
JOHNSTON ISLAND 1961 / Johnston Island	International	-251	-0.14192702	1	191	-77	-204
KANDAWALA / Sri Lanka	Everest	860.655	0.28361368	3	-97	787	86
KERGUELEN ISLAND / Kerguelen Island	International	-251	-0.14192702	1	145	-187	103

*   Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.
** WGS 84 minus local geodetic system

Table 7.5 (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	$\Delta f \times 10^4$		ΔX(m)	ΔY(m)	ΔZ(m)
KERTAU 1948 West Malaysia and Singapore	Modified Everest	832.937	0.28361368	6	-11	851	5
LA REUNION Mascarene Island	International	-251	-0.14192702	1	94	-948	-1262
L.C. 5 ASTRO Cayman Brac Island	Clarke 1866	-69.4	-0.37264639	1	42	124	147
LIBERIA 1964 Liberia	Clarke 1880	-112.145	-0.54750714	4	-90	40	88
LUZON Philippines (Excluding Mindanao Island)	Clarke 1866	-69.4	-0.37264639	6	-133	-77	-51
Mindanao Island				1	-133	-79	-72
MAHE 1971 Mahe Island	Clarke 1880	-112.145	-0.54750714	1	41	-220	-134

* Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.

** WGS 84 minus local geodetic system

Table 7.5   (Cont'd)

Transformation Parameters

- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	Δf × 10^4		ΔX(m)	ΔY(m)	ΔZ(m)
MARCO ASTRO Salvage Islands	International	-251	-0.14192702	1	-289	-124	60
MASSAWA Eritrea (Ethiopia)	Bessel 1841	739.845	0.10037483	1	639	405	60
MERCHICH Morocco	Clarke 1880	-112.145	-0.54750714	9	31	146	47
MIDWAY ASTRO 1961 Midway Island	International	-251	-0.14192702	1	912	-58	1227
MINNA Nigeria	Clarke 1880	-112.145	-0.54750714	6	-92	-93	122
NAHRWAN Masīrah Island (Oman)	Clarke 1880	-112.145	-0.54750714	2	-247	-148	369
United Arab Emirates				2	-249	-156	381
Saudi Arabia				1	-231	-196	482

*   Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180),
then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.

**   WGS 84 minus local geodetic system

Table 7.5 (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	Δf × 10⁴		ΔX(m)	ΔY(m)	ΔZ(m)
NAMIBIA Namibia	Bessel 1841	653.135***	0.10037483	3	616	97	-251
NAPARIMA, BWI Trinidad and Tobago	International	-251	-0.14192702	1	-2	374	172
NORTH AMERICAN 1927 Mean Value (CONUS)	Clarke 1866	-69.4	-0.37264639	405	-8	160	176
Alaska				47	-5	135	172
Bahamas (Excluding San Salvador Island)				11	-4	154	178
San Salvador Island				1	1	140	165
Canada (Including Newfoundland Island)				112	-10	158	187
Canal Zone				3	0	125	201

* Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.

** WGS 84 minus local geodetic system

*** This value reflects a difference in the definition of legal and international meters in Namibia. (The semimajor axis for the Bessel 1841 Ellipsoid in Namibia is a = 6377483.865 legal meters.)

Table 7.5   (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	Δf × 10⁴		ΔX(m)	ΔY(m)	ΔZ(m)
NORTH AMERICAN 1927 (Cont'd)	Clarke 1866	-69.4	-0.37264639				
Caribbean (Barbados, Caicos Islands, Cuba, Dominican Republic, Grand Cayman, Jamaica, Leeward Islands, and Turks Islands)				14	- 7	152	178
Central America (Belize, Costa Rica, El Salvador, Guatemala, Honduras, and Nicaragua)				19	0	125	194
Cuba				1	-9	152	178
Greenland (Hayes Peninsula)				2	11	114	195
Mexico				22	-12	130	190

*   Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.

** WGS 84 minus local geodetic system

Table 7.5 (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	Δf x 10⁴		ΔX(m)	ΔY(m)	ΔZ(m)
NORTH AMERICAN 1983 Alaska, Canada, Central America, CONUS, Mexico	GRS 80	0	-0.00000016	379	0	0	0
OBSERVATORIO 1966 Corro, Santa Cruz, and Flores Islands (Azores)	International	-251	-0.14192702	3	-425	-169	81
OLD EGYPTIAN 1930 Egypt	Helmert 1906	-63	0.00480795	14	-130	110	-13
OLD HAWAIIAN Mean Value	Clarke 1866	-69.4	-0.37264639	13	61	-285	-181
OMAN Oman	Clarke 1880	-112.145	-0.54750714	7	-346	-1	224

* Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.

** WGS 84 minus local geodetic system

Table 7.5 (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	Δf x 10⁴		ΔX(m)	ΔY(m)	ΔZ(m)
ORDNANCE SURVEY OF GREAT BRITAIN 1936 Mean Value (England, Isle of Man, Scotland, Shetland Islands, and Wales)	Airy	573.604	0.11960023	38	375	-111	431
PICO DE LAS NIEVES Canary Islands	International	-251	-0.14192702	1	-307	-92	127
PITCAIRN ASTRO 1967 Pitcairn Island	International	-251	-0.14192702	1	185	165	42
PROVISIONAL SOUTH CHILEAN 1963 South Chile (near 53°S)	International	-251	-0.14192702	2	16	196	93

* Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.

** WGS 84 minus local geodetic system

Table 7.5   (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	$\Delta a(m)$	$\Delta f \times 10^4$		$\Delta X(m)$	$\Delta Y(m)$	$\Delta Z(m)$
PROVISIONAL SOUTH AMERICAN 1956 Mean Value (Bolivia, Chile, Colombia, Ecuador, Guyana, Peru, and Venezuela)	International	-251	-0.14192702	63	-288	175	-376
PUERTO RICO Puerto Rico and Virgin Islands	Clarke 1866	-69.4	-0.37264639	11	11	72	-101
QATAR NATIONAL Qatar	International	-251	-0.14192702	3	-128	-283	22
QORNOQ South Greenland	International	-251	-0.14192702	2	164	138	-189
ROME 1940 Sardinia Island	International	-251	-0.14192702	1	-225	-65	9

*   Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.
**   WGS 84 minus local geodetic system

Table 7.5 (Cont'd)

Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	Δf × 10⁴		ΔX(m)	ΔY(m)	ΔZ(m)
SANTA BRAZ Saint Miguel, Santa Maria Islands (Azores)	International	-251	-0.14192702	2	-203	141	53
SANTO (DOS) Espirito Santo Island	International	-251	-0.14192702	1	170	42	84
SAPPER HILL 1943 East Falkland Island	International	-251	-0.14192702	1	-355	16	74
SOUTH AMERICAN 1969 Mean Value (Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Guyana, Paraguay, Peru, Venezuela, and Trinidad and Tobago)	South American 1969	-23	-0.00081204	84	-57	1	-41
SOUTH ASIA Singapore	Modified Fischer 1960	-18	0.00480795	1	7	-10	-26

* Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.

** WGS 84 minus local geodetic system

Table 7.5 (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	Δa(m)	Δf × 10⁴		ΔX(m)	ΔY(m)	ΔZ(m)
SOUTHEAST BASE Porto Santo and Madeira Islands	International	-251	-0.14192702	2	-499	-249	314
SOUTHWEST BASE Azores (Pico and Terceira Islands)	International	-251	-0.14192702	5	-104	167	-38
TIMBALAI 1948 Brunei and East Malaysia (Sarawak and Sabah)	Everest	860.655	0.28361368	8	-689	691	-46
TOKYO Mean Value (Japan, Korea, and Okinawa)	Bessel 1841	739.845	0.10037483	13	-128	481	664
TRISTAN ASTRO 1968 Tristan da Cunha	International	-251	-0.14192702	1	-632	438	-609

* Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.

** WGS 84 minus local geodetic system

Table 7.5   (Cont'd)
Transformation Parameters
- Local Geodetic Systems to WGS 84 -

Local Geodetic Systems*	Reference Ellipsoids and Parameter Differences**			Number of Doppler Stations Used to Determine Transformation Parameters	Transformation Parameters**		
	Name	$\Delta a(m)$	$\Delta f \times 10^4$		$\Delta X(m)$	$\Delta Y(m)$	$\Delta Z(m)$
VITI LEVU 1916 Viti Levu Island (Fiji Islands)	Clarke 1880	-112.145	-0.54750714	1	51	391	-36
WAKE-ENIWETOK 1960 Marshall Islands	Hough	-133	-0.14192702	7	101	52	-39
ZANDERIJ Suriname	International	-251	-0.14192702	5	-265	120	-358

*   Geoid heights computed using spherical harmonic expansion and WGS 84 EGM coefficient set (n=m=180), then referenced to the ellipsoid and orientation associated with each of the local geodetic systems.

**   WGS 84 minus local geodetic system

Table 7.6

Local Geodetic System-to-WGS 84
Datum Transformation Multiple Regression Equations ($\Delta\phi$, $\Delta\lambda$, ΔH)
-- North American Datum 1927 (NAD 27)* to WGS 84 --

$$\Delta\phi" = - 0.16984 - 0.76173\ U + 0.09585\ V + 1.09919\ U^2 - 4.57801\ U^3 - 1.13239\ U^2V + 0.49831\ V^3$$
$$- 0.98399\ U^3V + 0.12415\ UV^3 + 0.11450\ V^4 + 27.05396\ U^5 + 2.03449\ U^4V + 0.73357\ U^2V^3$$
$$- 0.37548\ V^5 - 0.14197\ V^6 - 59.96555\ U^7 + 0.07439\ V^7 - 4.76082\ U^8 + 0.03385\ V^8 + 49.04320\ U^9$$
$$- 1.30575\ U^6V^3 - 0.07653\ U^3V^9 + 0.08646\ U^4V^9$$

$$\Delta\lambda" = - 0.88437 + 2.05061\ V + 0.26361\ U^2 - 0.76804\ UV + 0.13374\ V^2 - 1.31974\ U^3 - 0.52162\ U^2V$$
$$- 1.05853\ UV^2 - 0.49211\ U^2V^2 + 2.17204\ UV^3 - 0.06004\ V^4 + 0.30139\ U^4V + 1.88585\ UV^4$$
$$- 0.81162\ UV^5 - 0.05183\ V^6 - 0.96723\ UV^6 - 0.12948\ U^3V^5 + 3.41827\ U^9 - 0.44507\ U^8V$$
$$+ 0.18882\ UV^8 - 0.01444\ V^9 + 0.04794\ UV^9 - 0.59013\ U^9V^3$$

$$\Delta H_m = - 36.526 + 3.900\ U - 4.723\ V - 21.553\ U^2 + 7.294\ UV + 8.886\ V^2 - 8.440\ U^2V - 2.930\ UV^2$$
$$+ 56.937\ U^4 - 58.756\ U^3V - 4.061\ V^4 + 4.447\ U^4V + 4.903\ U^2V^3 - 55.873\ U^6 + 212.005\ U^5V$$
$$+ 3.081\ V^6 - 254.511\ U^7V - 0.756\ V^8 + 30.654\ U^8V + 0.122\ UV^9$$

* Contiguous United States (CONUS)

Table 7.6 (Cont'd)

Local Geodetic System-to-WGS 84
Datum Transformation Multiple Regression Equations ($\Delta\phi$, $\Delta\lambda$, ΔH)
- North American Datum 1927 (NAD 27)* to WGS 84 -

In the preceding equations:

U = K (ϕ - 37)

V = K (λ - 265)

K = 0.05235988

ϕ = geodetic latitude, local geodetic system, in degrees and decimal part of a degree; positive north (0° to 90°), negative south (0° to -90°)

λ = geodetic longitude, local geodetic system, in degrees and decimal part of a degree; positive east from 0° to 360°

The preceding equations reproduced Doppler-derived WGS 84 geodetic coordinates at 447, 425, and 428 comparison points to the following root-mean-square (RMS) differences, respectively:

ϕ: ±1.3 m; λ: ±1.3 m;   H (geodetic height): ±1.2 m

Test Case:

Input data for NAD 27
test point:

ϕ = 34°47'08.833"N $\Delta\phi$ = 0.356"

λ = 273°25'07.825"E $\Delta\lambda$ = 0.080"

ΔH = -38.06 m

_____

*Contiguous United States (CONUS)