

REPORT ON GEOCENTRIC AND GEODETIC DATUMS

P. VANICEK

February 1975



**TECHNICAL REPORT
NO. 32**

PREFACE

In order to make our extensive series of technical reports more readily available, we have scanned the old master copies and produced electronic versions in Portable Document Format. The quality of the images varies depending on the quality of the originals. The images have not been converted to searchable text.

REPORT ON GEOCENTRIC AND GEODETIC DATUMS

P. Vaníček

Department of Surveying Engineering
University of New Brunswick
P.O. Box 4400
Fredericton, N.B.
Canada
E3B 5A3

© February 1975

Latest Reprinting December 1992

TABLE OF CONTENTS

	Page
Foreword	i
1) BASIC DEFINITIONS	1
1.1) Spaces and Coordinate Systems	1
1.2) Coordinate Lines and Surfaces	3
1.3) Families of Coordinate Systems	3
1.4) Positioning of Coordinate System	6
2) GEOCENTRIC COORDINATE SYSTEMS	8
2.1) Two Ways of Positioning a Geocentric Coordinate System	8
2.2) Average Terrestrial System	10
3) GEODETIC COORDINATE SYSTEMS	12
3.1) Classical Definition of a Geodetic System	12
3.2) Alternatives to the Classical Definition	15
3.3) Role of Astronomic Azimuths	17
4) DETERMINATION OF THE MUTUAL POSITION OF A GEODETIC AND THE AVERAGE TERRESTRIAL COORDINATE SYSTEMS	18
4.1) The Case of Fixed and Floating Geodetic Datums	18
4.2) Use of the Earth Gravity Field	20
4.3) Use of Coordinate Values	21
4.4) Role of the Scale Factors	23
REFERENCES	25

Foreword

The work on the here presented research was undertaken as an attempt to settle some of the ongoing arguments concerning geodetic datums. The flurry of discussions I have witnessed in the recent past convinced me that a more fundamental look at these problems was needed and that the key to the understanding lays in the domain of concepts. Hence, in this report I am trying to treat the problems on philosophical level first and deal with specifics only when necessary. Since a multitude of papers have been published on the subject in the past decade or so, most of the specifics are dealt with by merely referencing the appropriate papers.

The dilemma I was faced with was "how to use the philosophical approach without introducing too much abstract mathematics". At the end I have attempted to write about rigorous mathematical structures using non rigorous language that has proved to be a very difficult task.

I wish to acknowledge the many helpful discussions I have had over the past two years with the U.N.B. geodesy group: Drs. E.J. Krakiwsky and D.E. Wells, Messrs. C.L. Merry and D.B. Thomson; with the Surveys and Mapping employees: Mr. H.E. Jones, Brig. L.J. Harris, Messrs. T. Wray and J. Kouba; and with Dr. G. Blaha and Prof. I.I. Mueller. Also communications from other O.S.U. researchers Prof. R.H. Rapp and Mr. A. Leick, helped me in formulating some of the ideas here contained. The research was carried out under Research Contract DSS File No. SV03.23244-4-4009 with the Geodetic Survey of Canada.

Petr Vaníček
21th January, 1975.

1) BASIC DEFINITIONS

1.1) Spaces and Coordinate Systems

While in mathematics we work with abstract spaces, i.e. spaces that are accepted axiomatically, in geodesy we have to work with physical space. Physical space is the space in which there are physical objects and in which we are able to take measurements. We can say that the physical space is implied by the physical objects around us. In addition, when using the mathematical apparatus, we have to work with the mathematical abstraction of the physical space as well.

The most important property of any space, abstract or physical, is its dimensionality. The measure of dimensionality is the number of coordinates needed to describe positions of points in the space. This, of course, presupposes that there is a coordinate system defined in the space which we shall assume will be always the case. In this study, we shall be dealing exclusively with three-dimensional spaces, where point positions are uniquely given by triplets of coordinates.

A coordinate system, being an abstract concept, is introduced in the space (physical or abstract) by definition. If a Cartesian coordinate system, i.e. the usual X, Y, Z system with "straight and mutually orthogonal" coordinate axes, can be defined in a space then the space is called Euclidean. In such a space the distance d between two points (X_1, Y_1, Z_1) and (X_2, Y_2, Z_2) is given by

$$d = [(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2]^{1/2}. \quad (1)$$

Not all the spaces are Euclidean. However, in this study, we shall work only with Euclidean spaces, both abstract and physical.

In Euclidean spaces we are not only able to define the Cartesian coordinate system but also various curvilinear coordinate systems. One example of such a curvilinear system is the ellipsoidal system ϕ, λ, h , with h reckoned from the ellipsoid known in geodesy as the reference ellipsoid. Another such system is the astronomic system Φ, Λ, H , with H being for instance the orthometric height reckoned from the geoid.

Generally, curvilinear coordinates u, v, w can be expressed as functions of the Cartesian coordinates, called transformation equations from the Cartesian to the curvilinear system. They may be written as follows:

$$\begin{aligned} u &= u(X, Y, Z) \\ v &= v(X, Y, Z) \\ w &= w(X, Y, Z) \end{aligned} \quad (2)$$

where the functions may be simple or complicated.

Following up the above example, the transformation equations for the ellipsoidal system ϕ, λ, h , co-axial with the Cartesian system are fairly complicated. They are given in [Paul, 1973]. It so happens that the inverse transformations for this particular curvilinear system are much more simple. They are given by the well known expressions [Krakiwsky and Wells, 1971]:

$$\begin{aligned} X &= (N + h) \cos \phi \cos \lambda \\ Y &= (N + h) \cos \phi \sin \lambda \\ Z &= \left(N \frac{b^2}{a^2} + h \right) \sin \phi, \end{aligned} \quad (3)$$

where N is the radius of curvature of the ellipsoid (given by the two semi-axes a, b) in the prime vertical direction. The transformations

between the astronomic and Cartesian systems are much more involved still [Hotine, 1969] and require a detail knowledge of the earth gravity field.

1.2) Coordinate Lines and Surfaces

Coordinate systems have coordinate lines and coordinate surfaces. Coordinate lines are the lines on which two of the three coordinates remain constant. The coordinate lines on which the value of the two constant coordinates equals to zero are the basic coordinate lines or axes of the system. As an example, the X-axis or X-basic coordinate line of the Cartesian system is the straight line on which $Y = Z = 0$. In the ellipsoidal system, the major axis of the ellipsoid passing through the point $\phi = \lambda = 0$ is also one such basic coordinate line, equator and the zero-meridian being the other.

Coordinate surfaces are the surfaces on which one of the coordinates does not change. If the value of the constant coordinate equals to zero then we speak about a basic (reference) coordinate surface. An example of a basic coordinate surface would be, say, the X, Y-coordinate plane in the Cartesian system on which $Z = 0$. Other such examples would be the reference ellipsoid for the ellipsoidal system (on which $h = 0$) or the geoid for the astronomic system (on which $H = 0$). It is customary in geodesy to call these reference coordinate surfaces datums.

1.3) Families of Coordinate Systems

In one (Euclidean) space we can have a variety of coordinate systems defined. They can be of both kinds, Cartesian as well as curvilinear. They may create families of co-axial systems. One such family

would consist of one Cartesian and one or several curvilinear systems such that they would not only share the centre of coordinates, i.e. the point $(0, 0, 0)$ (in the ellipsoidal system the centre of the ellipsoid instead) but also the basic coordinate lines of the curvilinear systems would lay in the basic coordinate surfaces of the Cartesian system. On Fig. 1, we can see an example of one such family, often encountered in geodesy and composed of the Cartesian, spherical and ellipsoidal systems.

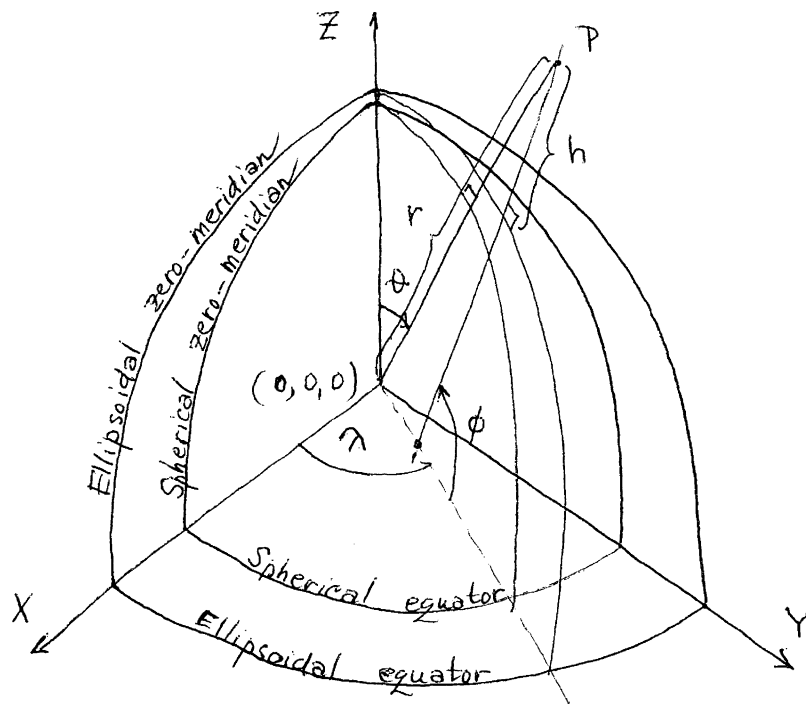


Figure 1.

Evidently, there exists no Cartesian system with which the astronomic system would be co-axial since the basic coordinate lines of the astronomic system are spatial curves. Hence the astronomic system does not belong to any family.

In geodesy, we are dealing with transformations within one family given by transformation equations of a form similar to eqs. 2. But we also work with transformations between families. For the purpose

of the inter-family transformations each family may be represented by its Cartesian system.

It is known from elementary geometry that any two Cartesian systems X, Y, Z, X', Y', Z' are related by the following transformation equations:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} - \begin{bmatrix} X'_0 \\ Y'_0 \\ Z'_0 \end{bmatrix} = R(\omega, \psi, \epsilon) \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (4)$$

where (X'_0, Y'_0, Z'_0) are the coordinates of the X, Y, Z system's origin reckoned in the X', Y', Z' system and $R(\omega, \psi, \epsilon)$ is the so-called rotation matrix composed of trigonometric functions of the three rotation angles ω (about X -axis), ψ (about Y -axis) and ϵ (about Z -axis). There are thus six independent parameters $(X'_0, Y'_0, Z'_0, \omega, \psi, \epsilon)$ to be specified for any such transformation.

If we are faced with the task of transforming from a curvilinear system A belonging to one family to a curvilinear system B belonging to another family, we have to go through the following three steps:

- (i) transform from the curvilinear system A to its Cartesian system using the transformations inverse to (2);
- (ii) transform from one Cartesian to another using equations (4);
- (iii) transform the Cartesian coordinates to curvilinear coordinates in B using eqs. (2).

Speaking about transformations in geodesy, we generally assume that the scale of all the systems belonging to one family is the same. However, it may vary from family to family. If the scale of the two

families we deal with, is different then the coordinates in transformation equations (4) have to be multiplied by the appropriate scale factors to bring them to a common scale.

1.4) Positioning of Coordinate System

In the physical space we live and measure in, there are physical objects such as the earth, survey markers, the oceans, the geometrical study of which is what geodesy is all about. These physical objects become geometrical objects (regions in space, surfaces, lines, points, networks of points) once we transfer our problems from the physical space to the abstract space.

To study the geometrical properties of these objects, such as sizes, shapes, positions, it is necessary to relate the objects to one or several coordinate systems. Since the objects constitute the physical reality around us, they must be considered the fundamental entities while the coordinate systems are just the scaffolding useful for the task to be performed. We can define the coordinate systems and position them relative to the objects any way we want.

On the other hand, in the abstract spaces we may take a coordinate system to be defined and positioned in an a priori sense. Then the abstract geometrical objects can be placed in the space and related to the so defined coordinate system.

In experimental sciences, of which geodesy is one, the first approach prevails. The coordinate values (of geometrical objects) derived from measurements are regarded as defining the position of the implied coordinate system with respect to the object. An often used alternative to define a coordinate system and its position with respect

to a physical object is to use some physical, as opposed to geometrical, properties of the object. So defined systems are known as natural coordinate systems. The astronomic coordinate system may serve as an example. It is the earth gravity field, i.e. its plumb lines and equipotential surfaces, that implies the system. Another example is the Cartesian system whose coordinate axes coincide with the principal axes of inertia (main axes of the principal ellipsoid of inertia) of a physical body [Vaníček, 1972]. We may just note that the principal axes of inertia intersect in the centre of gravity of the body.

2) GEOCENTRIC COORDINATE SYSTEMS

2.1) Two Ways of Positioning a Geocentric Coordinate System

What do we understand by the term geocentric coordinate system? Different scholars understand the term differently. Ideally, however, it is the natural Cartesian system whose axes coincide with the principle axes of inertia of a rigid earth. These axes are tied to the earth through the physical properties of the earth that manifest themselves in several different ways.

The most obvious way is the rotation (spin) of the earth that takes place around an instantaneous axis of rotation. If the earth were rigid and had no oceans and no atmosphere, the instantaneous axis would describe a circular cone around the polar principal axis of inertia with its vertex in the centre of gravity of the earth. Then we would be able to determine the position of this polar axis of inertia quite accurately from astronomic or other extraterrestrial observations. Unfortunately, the earth is not rigid and in addition it possesses oceans and atmosphere. All this causes the instantaneous axis to wobble somewhat irregularly. The main component of this polar wobble is still circular (Chandlerian) but there are also seasonal, secular and irregular components present [Vaníček, 1972].

Another manifestation of the "physics of the earth" is its gravity field. It is customary in geodesy to describe the gravity field using one or more different quantities such as gravity anomalies, geoidal heights, deflections of the vertical or disturbing potential. All these quantities are somehow related to the normal potential, i.e. an artificial potential generated by a conventionally adopted ellipsoidal body.

It has been shown [Heiskanen and Moritz, 1967; Vaníček, 1971] that if we develop any one of these quantities into an infinite series of spherical harmonics and drop out the first degree terms that the ellipsoidal body is forced to become geocentric. Its centre becomes coincident with the centre of gravity of the earth and its minor axis becomes coincident with the earth's polar axis of inertia. The ellipsoid so positioned is known as the geocentric reference ellipsoid or simply geocentric datum. Thus we can state that the geoid determined from gravity anomalies (gravimetric geoid) using the known Stokes formula - derived under the assumption that first degree spherical harmonics of the gravity anomalies equal to zero - refers to a geocentric datum. The same holds true also for the satellite derived geoid (geoidal heights) as well as for the deflections of the vertical computed from the Vening-meinesz formulae or derived from satellite orbit analysis. Inverting this view, we may say that the geoid as a known physical object represented by the mean sea level to an accuracy of a few metres, implies the geocentric datum, if the geoidal heights or the deflections are derived from gravity anomalies or satellite orbit analyses.

Obviously, the first approach, using the rotation of the earth, does not give us any information about the location of the other two principal axes of inertia of the earth. We cannot let the earth spin around an axis in the equatorial plane. The second approach is slightly more hopeful in that it is theoretically possible to determine the direction of the two equatorial axes of inertia. This determination has been attempted by a number of researchers, e.g. [Burša, 1971; Fajemirokun et al., 1973]. However, the solution is very weak because the earth's assymetry in the polar view is very slight. Thus, the product of inertia with respect to X and Y-axes is very small, of the same order of magnitude

as the noise in the gravity data.

We note here that the Geodetic Reference System 1967 [IAG, 1972] does not specify the position of the recommended geocentric reference ellipsoid with respect to any other coordinate system or any physical object. It merely stipulates, by adopting the formula for gravity potential in spherical harmonics in which the first order terms are missing, that the system is concentric with the earth and its z-axis coincides with the earth's polar principal axis of inertia.

2.2) Average Terrestrial System

Since we cannot, as yet, determine the position of the natural geocentric system of coordinates with respect to the earth, a semi-natural system, generally called Average Terrestrial, has been adopted. This system is defined so as to have the centre in the earth's centre of gravity, the Z-axis close to the earth's polar axis of inertia and the X-axis directed so that the XZ-basic coordinate plane (containing also the zero-meridian of the co-axial ellipsoidal system) passes through the Greenwich Mean Observatory. To be more specific, the Z-axis is defined as passing through the Conventional International Origin (CIO) given as an average pole indicated by the five IPMS stations functioning in the period 1900-1905 [Mueller, 1969]. The Mean Observatory is also an imaginary point adopted by convention and defined by transformations of time.

The Average Terrestrial system as defined, is subject to movement with respect to the earth via the positions of the five IPMS stations. Hence it is not convenient for solving global dynamical problems [Mather, 1974] and should be used only for stationary positioning when time variations can be regarded as being below the level of

position errors. The problem of coordinate systems for dynamic investigations is beyond the scope of this paper.

The question now arises: Can one measure the coordinates or coordinate differences in the Average Terrestrial system? The answer is no, at least not directly. It can be, however, regarded as the framework for the astronomic coordinate system. But the Average Terrestrial system is not identical with the astronomic system. The Average Terrestrial is also used as the coordinate system in which the (geocentric) coordinates, as derived from satellite observations, are expressed. In the satellite positioning, the Average Terrestrial system is implied through the gravity field as well as through the extraterrestrial observations.

From the point of view of the definition of the Average Terrestrial system's position we can see that the system can be regarded as positioned either via transformation equations from a celestial coordinate system (which in turn is defined by coordinate values of stars) or, indirectly, via satellite determined coordinates of points on the surface of the earth. Since both definitions involve use of observations contaminated by inevitable errors, there is no reason to believe that both definitions define precisely the same position. This point is elaborated on in the work of Wells and Vaníček [1975] where it is shown that it makes sense to look for actual discrepancies between the satellite implied and astronomically implied Average Terrestrial systems.

3) GEODETIC COORDINATE SYSTEMS

3.1) Classical Definition of a Geodetic System

By geodetic coordinate system we understand the family of systems of which the ellipsoidal system (ϕ, λ, h) is a member. In the past, the geodetic systems have been chosen independently for different continents, groups of countries or individual countries, leaving us with a score of systems in existence [Mueller et al., 1973]. The reason why so many independent geodetic systems were set up is that only inter-connected (terrestrial) geodetic networks could be expressed in one coordinate system, or in other words, related to one geodetic datum.

The criteria for a coordinate system were to minimise the summation of squares of either (relative) geoidal heights or (relative) deflections of the vertical in the region of validity of the system. These quantities were to be minimized to enable us to neglect them in reducing the geodetic observations made on the surface of the earth down to the ellipsoid. These criteria are generally enough to dictate not only the position of the system but also the size and shape of the geodetic reference ellipsoid, the datum. However, the shape and size of the reference ellipsoid were usually selected beforehand and the ellipsoid so selected was usually the one in vogue at that time. This preselection thus left just the position to be determined so as to conform with one of the above criteria.

The classical (standard) way of positioning the geodetic datum, and thus the geodetic coordinate system, is to specify the six necessary parameters relating the datum to a topocentric coordinate system centered at a point on the surface of the earth. This point is usually called the Origin of the geodetic network [Bomford, 1971].

The topocentric coordinate system is the local coordinate triad-tangents to coordinate lines - of the astronomic coordinate system ϕ, λ, H . This local coordinate triad and the relations of the geodetic and topocentric coordinate systems are discussed in some detail in [Vaníček and Wells, 1974].

A few pertinent points from the cited paper should, perhaps, be reiterated here. The six parameters mentioned above are usually:

$\phi_0, \lambda_0, h_0, \xi_0, \eta_0, \alpha_0$. These are:

- (i) three geodetic coordinates of the origin ϕ_0, λ_0, h_0 , where $h_0 = H_0 + N_0$, i.e. the sum of the orthometric height H_0 and the (relative) geoidal height N_0 referred to the geodetic reference ellipsoid;
- (ii) two (relative) deflection components ξ_0, η_0 , referred to the reference ellipsoid, and
- (iii) the geodetic azimuth α_0 of one geodetic line joining the origin with another point in the network. We must note here that the way the six parameters were obtained is immaterial, as long as we are willing to regard them fixed.

We also note that in the first approximation, it is not necessary to specify the three quantities N_0, ξ_0, η_0 , that link the geodetic system with the gravity field. This is possible because these quantities are not needed, in the initial stages of setting up the networks, for the geodetic computations, i.e. the computations of geodetic coordinates from geodetic measurements. We know that the effect of geoidal heights and the deflections of the vertical on geodetic observations can be neglected in the first approximation provided they can be taken as being sufficiently small.

The fact that only three quantities, $\phi_0, \lambda_0, \alpha_0$, have to be specified at the beginning is understandable. With these three quantities,

we are already able to "develop the network of points", i.e. to compute the first approximation of the geodetic coordinates of the control points from distances and angles observed on the surface of the earth, using orthometric heights of points instead of heights above reference ellipsoid for reduction of distances and forgetting about reductions of horizontal angles. This also reflects the fact that when we work in one geodetic coordinate system only we do not have to worry about its relation to any other coordinate system belonging to a different family. Once we start using other coordinate systems such as a geocentric or another geodetic (belonging to another family) we have to know the proper transformation equations.

The six positioning parameters at the origin relate the geodetic coordinate system uniquely to the astronomic system, with the geoid as its datum. Another way to look at the classical positioning of the geodetic datum is to regard it as defining the position of the geodetic coordinate system with respect to a physical object. This physical object is composed of the two survey markers of the origin ϕ_0 , λ_0 and the other point for which the geodetic azimuth α_0 has been accepted. This is, of course, not a unique definition because it leaves the geodetic system with three degrees of freedom unspecified.

We may finally observe that the described classical positioning fixes, but does not specify, the position of the geodetic coordinate system with respect to the earth. This is because the astronomic system is fixed to the earth, up to the time variations of gravity field. The positions of control points in horizontal geodetic networks are then expressed in this fixed geodetic coordinate system and whatever we do with the network does not influence the position of the coordinate system. The inevitable errors in the coordinate values originating from errors in

observations as well as from inaccurate computations and neglect of various effects [Merry and Vaníček, 1973; Thomson et al., 1974] are then interpreted as just errors in positions and nothing else. Addition of points or readjustment of coordinates have no effect on the coordinate system position either unless some of the following fundamentals for parameters ϕ_0 , λ_0 , α_0 , N_0 are changed. Errors in astronomical coordinates influence only the geoid computations and enter into the above argument only as second order errors in coordinates.

3.2) Alternatives to the Classical Definition

One alternative to the classical positioning of geodetic datums advocated by some geodesists is to select a well distributed uniquely specified fixed set of control points and declare that their coordinate values "define the position of the datum" whose size and shape had been preselected. What is meant by this is that the physical object consisting of the markers of these selected points is taken as the reference object to which the coordinate system is then related. The spirit of this definition is identical with that used in defining the CIO. It is presumably understood that in any subsequent computations (readjustments or additions) in the network the coordinate values of these selected points will not be changed.

The inherent problem with this approach is that even though the geodetic coordinate system is seemingly positioned with respect to the physical object, it is not really the case. The positioning is really done via the geometric representation of the physical object, i.e. the network of points described by the adopted coordinate values of the points. Therefore any errors in the initial determination of these coordinate values are transmitted into the position of the datum and any

corrective measures will bring about a change in the position of the coordinate system with respect to the earth.

The second alternative for positioning the geodetic datum is the one preferred by researchers either explicitly, e.g. [U.S. National Academy of Sciences, 1971; Jones, 1973], or implicitly, e.g. [Veis, 1960; Burša, 1965]; Lambeck, 1971]. In this option the whole geodetic network, i.e. all the control points indiscriminately, are considered as defining the position of the geodetic datum and thus the geodetic coordinate system. There are two problems to be faced in this alternative. In addition to the problem common with the first alternative, there is the rather unfortunate consequence that the position of the geodetic datum with respect to the earth fluctuates with any addition of points to the network, with local readjustment, etc. This leaves us with a floating datum for which the transformation equations to another coordinate system are epoch dependent.

There is, however, something to be said in favor of this last approach. In the classical definition all the errors in positions, i.e. systematic distortions as well as random errors, are associated with the network alone. In other words, only the geometrical representation of the reality is considered distorted. On the other hand, in the last approach, the errors are distributed evenly between the coordinate system and the network. This may diminish particularly the radially propagated distortions that are severe in a case such as the North-American because of the adverse location of the origin with respect to the geometrical centre of the network. The second approach can, of course, achieve the same effect if the selected points are spread out throughout the network.

3.3) Role of Astronomic Azimuths

What remains to be discussed is the role of astronomically determined azimuths in the network. These azimuths can be, to a certain degree of accuracy, converted to geodetic azimuths using the expanded Laplace equation. The role of the astronomic azimuth at the origin is adequately discussed in [Vaníček and Wells, 1974] and need not be repeated here.

The effect of introducing the Laplace azimuths determined at a number of points in the network remains unclear. Used in a simultaneous adjustment of the whole network they are likely to somehow correct the distortions provided they are properly determined and properly weighted. If, in this adjustment, the classical definition of the datum position is accepted and the parameters at the origin held fixed then the inclusion of these azimuths does not influence the position of the datum. If the classical definition is accepted but the azimuth α_0 at the origin is left free to take on a correction like any other azimuth then the inclusion of the azimuths may improve the misalignment of the geodetic system (see the next chapter).

However, if the Laplace azimuths are used in the piece-meal adjustments of subnetworks, their role is much less straightforward. It well may be that, if these azimuths are considered under such circumstances, they may add to the distortions.

4) DETERMINATION OF THE MUTUAL POSITION OF A GEODETIC AND THE AVERAGE TERRESTRIAL COORDINATE SYSTEMS.

4.1) The Case of Fixed and Floating Geodetic Datums

The question of the determination of mutual position of the geodetic and geocentric coordinate systems became important with the advent of satellites, or more precisely, with the advent of us becoming able to compute geocentric coordinates from satellite observations. As we have seen in Chapter 1, the transformation between the geodetic and geocentric systems is given by eq. (4) if we can consider the scale of both systems identical. If the scale is different, then we have to write, denoting by κ, κ' the scale factors on the two systems:

$$\kappa' \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} X'_0 \\ Y'_0 \\ Z'_0 \end{bmatrix} + \kappa R(\omega, \psi, \epsilon) \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (5)$$

Numerous variations on the same theme have been devised by various scholars, e.g. [Veis, 1960; Burša, 1962; Molodenskiĭ et al., 1962]. These variations differ by:

- (i) the selection of the point of rotation of the geodetic system - in eq. (5) it is the centre of the geodetic coordinate system;
- (ii) the interpretation of the role of the scale factor in the geodetic system. All of these transformations have subsequently been used by a number of geodesists in their investigations, e.g. [Wolf, 1963; Badekas, 1969; Krakiwsky et al., 1973; Mueller et al., 1973; Peterson, 1974], and assessed by Krakiwsky and Thomson [1974].

It is completely justified to expect the three rotations, ω, ψ, ϵ , to have some arbitrary small values if the position of the geodetic

datum is taken as defined by coordinate values of the control points. That is, if we want to regard the datum as floating or its position defined by a fixed set of points. However, if the datum is regarded as fixed with respect to the earth, i.e. positioned using the classical approach, it has been shown by Vaníček and Wells [1974] that the three rotations are restricted. The restriction is imposed by the fact that the rotation of the geodetic coordinate system (with respect to the Average Terrestrial system) has to take place around the ellipsoidal normal going through the origin of the geodetic network.

In the case of the fixed datum, the transformation equations between the two systems take the following form:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} - \begin{bmatrix} X'_0 \\ Y'_0 \\ Z'_0 \end{bmatrix} = \Delta \cdot S(\phi_0, \lambda_0) \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (6)$$

where Δ is the rotation around the normal of the origin and $S(\phi_0, \lambda_0)$ is given as

$$S(\phi_0, \lambda_0) = \begin{bmatrix} 1 & , & \sin \phi_0 & , & -\cos \phi_0 \sin \lambda_0 \\ -\sin \phi_0 & , & 1 & , & \cos \phi_0 \cos \lambda_0 \\ \cos \phi_0 \sin \lambda_0 & , & -\cos \phi_0 \cos \lambda_0 & , & 1 \end{bmatrix} \quad (7)$$

[Vaníček and Wells, 1974]. The scale factors here are left out and will be discussed later.

If we want to determine the actual mutual position of the two coordinate systems, we have to determine the six $(X'_0, Y'_0, Z'_0, \omega, \psi, \varepsilon)$ or the four $(X'_0, Y'_0, Z'_0, \Delta)$ parameters, leaving out, for the moment the scale factors. The choice between the six or four parameters depends on how we want to regard the position of the geodetic coordinate system to have been defined.

There are basically two sources of data for determining the parameters. One is the earth gravity field, in the form of either the geoidal heights or deflection components, the other is the geodetic and geocentric coordinates of some identical points. We shall deal with the first source first.

4.2) Use of the Earth Gravity Field

The gravity field is usually referred - via normal gravity - to a geocentric reference ellipsoid, as described in Chapter 2, or - via geodetic measurements - to a geodetic reference ellipsoid. In the first case, the geoidal heights and deflection components, computed using the Stoke's technique are related to a reference ellipsoid that is one of the coordinate surfaces in the Average Terrestrial family of coordinate systems. The same holds true for the geoid or deflection components derived from satellite orbit analysis as well as a combination of both satellite and terrestrial information.

Therefore, the geoid derived from astronomically determined deflections of the vertical, the astro-geodetic geoid, is referred to the geodetic datum, while the gravimetrically determined geoid is referred to the geocentric datum and so is the satellite determined geoid. Combinations of these techniques may give a "geocentric geoid" or a "geodetic geoid".

It is thus conceivable that we can take a piece of the geoidal surface related to geocentric datum and the same piece of the geoidal surface related to geodetic datum and seek such transformation parameters that would make them match as well as possible - after correcting for usually different sizes and shapes of the two ellipsoids. These transformation parameters should be identical to the parameters relating the two coordinate systems.

This idea was first proposed by Mather [1970] and subsequently used by several researchers [Vaníček and Merry, 1973; Gay, 1973].

It was shown [Merry and Vaníček, 1974] that the translation components X'_0 , Y'_0 , Z'_0 recoverable with this technique are fairly reliable, while the rotations cannot be derived with any degree of certainty.

There is one point probably worth mentioning here. When computing the astro-geodetic or astro-gravimetric geoid, the deflection at the origin, even if it had been determined, is usually treated as any other deflection within the network. This may lead someone to argue that this treatment can introduce a "tilt" to the geodetic datum since we have to view the geoid as fixed to the earth so that any change in the position of the geoid with respect to the geodetic datum has to be interpreted as a change of the position of the datum with respect to the earth. Here again, this interpretation would be valid if we wish to take the position of the datum as defined by the geoid the same way as some would have the datum defined by network coordinates.

A more attractive looking proposition is to view the deformations of the geoid, including the area immediately adjacent to the origin, as errors in the geoid determination and not couple them with the datum position. These deformations will inevitably appear as a consequence of errors in both the geodetic and astronomic coordinate values of the deflection points.

4.3) Use of Coordinate Values

In cases of using the geocentric and geodetic coordinates of an identical set of points the problem of determining the transformation parameters is seemingly much more simple. In such a case, the equations (4) and (6) become simply "observation equations" with the corresponding

triplets of coordinates (X, Y, Z) and (X', Y', Z') playing the role of observations and the six or four transformation parameters respectively being the unknowns. Because the unknowns are generally fairly small, the observation equations can be written in a linear form and an estimation technique such as the method of least squares can be used. This approach has been used in practically all the investigations referred to in this report.

However, using this technique one has to bear in mind the fact that both coordinate triplets, geodetic and geocentric, are subject to systematic and random errors. While the systematic distortions of the geodetic (terrestrial) network can become particularly severe when we leave the environment of the origin, they have a very adverse influence on the estimation of the transformation parameters. Depending on the combination of points we choose for estimating the parameters, we get very significantly different results as illustrated in [Wells and Vaníček, 1975]. In addition to this, the number of points, for which both geocentric and geodetic coordinates are known with any degree of certainty, is still fairly limited. We thus cannot yet choose a set of well distributed points around the origin which would be the most obvious way to tackle this problem.

The answer to this problem is not in weighting the coordinate values either. The distortions tend to overflow into the estimated parameters. The only way out seems to be in some realistic modeling of the network distortions as already pointed out by Krakiwsky and Thomson [1974].

Using the "geoids" appears therefore, preferable even though only the translation components can be determined. The astro-geodetic and astro-gravimetric geoid is less influenced by the distortions and,

what is very important, one can select the patch of the geoidal surface that is to be transformed, so as to contain only the appropriate environment of the origin where the distortions are likely to be significantly smaller. The errors in the "geocentric geoid" are unrelated to the origin and can probably be more readily accepted as being random.

4.4) Role of the Scale Factors

Finally, we shall try to clarify the role of the scale factors. To begin with, we may realise that we have again two choices. When talking about the scale of a coordinate system we can either say that the scale of the coordinate system is given independently and that the distance measurements provide us only with the approximate values for the distances, or we can regard the scale of the coordinate system as determined by observing (measuring) distances. In other words, the scale coming from a standard (via the measuring instrument) can be either seen as resulting in a scale distortion (with respect to another standard) of the network without any influence on the coordinate system, or the mean scale distortion can be interpreted as influencing the scale of the system. The first approach corresponds, philosophically speaking, to the U.S. Academy of Sciences vision of the datum positioning, the second reminds us of the classical approach.

To illustrate the point, let us rewrite eq. (5) in vector notation:

$$\kappa' (\vec{r}_i' - \vec{r}_0') = \kappa R(\omega, \psi, \epsilon) \vec{r}_i \quad (5a)$$

This equation obviously describes the case when each coordinate system has the same scale as the respective networks have. We can now write \vec{r}_i as

$$\vec{r}_i = \vec{r}_i^* + \Delta \vec{r}_i \quad (8)$$

where \vec{r}^* is the radius-vector of the origin of the geodetic network in the geodetic system and $\Delta\vec{r}_i$ is the vector connecting the origin with the control point "i". Substituting eq. (8) to eq. (5) we get:

$$\kappa'(\vec{r}_i' - \vec{r}_0') = \kappa R(\omega, \psi, \varepsilon)\vec{r}^* + \kappa R(\omega, \psi, \varepsilon)\Delta\vec{r}_i. \quad (9)$$

If we interpret the scale κ as influencing the network only and thus regard it as a scale distortion, we would write:

$$\kappa'(\vec{r}_i' - \vec{r}_0') = R(\omega, \psi, \varepsilon)\vec{r}^* + \kappa R(\omega, \psi, \varepsilon)\Delta\vec{r}_i. \quad (10)$$

From the computational point of view, both scale factors cannot be determined because there is no other "standard" scale implied anywhere. It thus makes sense to solve only for the ratio $\lambda = \kappa/\kappa'$. Again two kinds of transformation equations can be used:

$$\vec{r}_i' - \vec{r}_0' = \lambda R(\omega, \psi, \varepsilon)\Delta\vec{r}_i \quad (11)$$

or

$$\vec{r}_i' - \vec{r}_0' = R(\omega, \psi, \varepsilon)\vec{r}^* + \lambda R(\omega, \psi, \varepsilon)\Delta\vec{r}_i. \quad (12)$$

The interpretation of the first equation is evident. The second equation is to be understood as assigning the same scale to both coordinate system as well as the network of the "geocentric points" and allowing, at the same time, for a different scale (scale distortion) in the geodetic network.

REFERENCES

- Badekas, J., 1969: Investigations Related to the Establishment of a World Geodetic System, O.S.U. Dept. of Geod. Sci. Tech. Rep. No. 124.
- Bomford, G., 1971: Geodesy, Oxford.
- Burša, M., 1962: The Theory of the Determination of the Nonparallelism of the Minor Axis of the Reference Ellipsoid, Polar Axis of Inertia of the Earth, and Initial Astronomical and Geodetic Meridians from Observations of Artificial Earth Satellites, Studia Geoph. et Geod. No. 6.
- Burša, M., 1965: Determination of the Direction of the Minor Axis of the Reference Ellipsoid and the Plane of the Initial Geodetic Meridian from the Artificial Satellite Data, Studia Geoph. et Geodetica, Vol. 9(1).
- Fajemirokun, F.A., F.D. Hotter and I.I. Mueller, 1973: Orientation of the Earth by Numerical Integration, Symp. on Comp. Methods in Geodesy, Oxford.
- Gay, R.H., 1973: Estimation of Parameters Relating Astro-Geoids to Global Reference Systems, 54th Ann. Meet. AGU, Washington.
- Heiskanen, W.A. and H. Moritz, 1967: Physical Geodesy, Freeman.
- Hotine, M., 1969: Mathematical Geodesy, ESSA Monogr. 2.
- International Association of Geodesy, 1972: Geodetic Reference System 1967, IAG Special Pub. No. 3.
- Jones, H.E., 1973: Geodetic Datums in Canada, Can. Surveyor, Vol. 27, No. 3.
- Krakiwsky, E.J. and D.E. Wells, 1971: Coordinate Systems in Geodesy, U.N.B. D.S.E. Lecture Notes No. 16.
- Krakiwsky, E.J., D.E. Wells and D.B. Thomson, 1973: Geodetic Control From Doppler Satellite Observations For Lines Under 200 km., Can. Surv. Vol. 27, No. 2.
- Krakiwsky, E.J. and D.B. Thomson, 1974: Mathematical Models for the Combination of Terrestrial and Satellite Networks, Can. Surv. Vol. 28, No. 5.
- Lambeck, K., 1971: The Relation of Some Geodetic Datums to a Global Geocentric Reference System, Bull. Géod. No. 99.
- Mather, R.S., 1970: The Geocentric Orientation Vector for the Australian Geodetic Datum, Geoph. J.R. Astro. Soc., 22.

- Mather, R.S., 1974: Time Variations in Geodetic Coordinates, Can. Surv. Vol. 28, No. 5.
- Merry, C.L. and P. Vaníček, 1973: Horizontal Control and the Geoid in Canada, Can. Surv., Vol. 27, No. 1.
- Merry, C.L. and P. Vaníček, 1974: The Geoid and Translation Components, Can. Surv. Vol. 28, No. 1.
- Molodenskij, M., V. Yeremeyev and M. Yurkina, 1962: Methods for Study of the External Gravitational Field and Figure of the Earth, Israel Prog. for Sci. Transl., Jerusalem.
- Mueller, I.I., 1969: Spherical and Practical Astronomy, Hungar.
- Mueller, I.I., M. Kumar, J.P. Reilly, N. Saxena and T. Solar, 1973: Global Satellite Triangulation for the National Geodetic Satellite Program, O.S.U. Dept. of Geod. Sc., Rep. No. 199.
- Paul, M.K., 1973: A Note on Computation of Geodetic Coordinates from Geodetic (Cartesian) Coordinates, Bull. Géod. No. 108.
- Peterson, A.E., 1974: Combined Adjustment of Canadian Triangulation and the 1973 Doppler Satellite Network, Can. Surv. Vol. 28, No. 5.
- Thomson, D.B., M. Nassar and C.L. Merry, 1974: Distortions of Canadian Geodetic Networks Due to the Neglect of Deflections of the Vertical and Geoidal Heights, Can. Surv. Vol. 28, No. 5.
- U.S. National Academy of Sciences, 1971: North American Datum.
- Vaníček, P., 1971: Physical Geodesy I, U.N.B., D.S.E. Lecture Notes No. 21.
- Vaníček, P., 1972: Earth-Pole Wobble, U.N.B., D.S.E. Lecture Notes No. 25.
- Vaníček, P. and C.L. Merry, 1973: Determination of the Geoid from Deflections of the Vertical Using a Least-Squares Surface Fitting Technique, Bull. Géod. No. 109.
- Vaníček, P. and D.E. Wells, 1974: Positioning of Horizontal Geodetic Datums, Can. Surv. Vol. 28, No. 5.
- Veis, G., 1960: Geodetic Uses of Artificial Satellites, Smithsonian Cont. to Astrophysics, Vol. 3, No. 9.
- Wells, D.E. and P. Vaníček, 1975: Investigation of Position of Various Satellite and Geodetic Coordinate Systems with Respect to the Average Terrestrial System (in preparation).
- Wolf, H., 1963: Geometric Connection and Re-orientation of Three-Dimensional Triangulation Nets, Bull. Géod. No. 68.