Where in the World are We?

Version 1.7

Andrew Jones
Resource Information Group
Department for Environment, Heritage and Aboriginal Affairs
in cooperation with the
South Australian Spatial Information Committee

Government of South Australia

Adelaide, South Australia
August 1999

Resource Information

RI Resource Information

Contents

Resource Information

Where in the World are We?

Since the dawn of civilisation, man has found it necessary to measure and map his domain. Examples abound throughout history, and include:

- the mapping of land holdings in the valley of the Nile (Ancient Egypt),
- the recording of journeys of global exploration (Columbus, Magellan and others),
- the recording of topographic information for military purposes.

The process continues today. The management of the world's natural and economic resources has become increasingly dependent on the availability of accurate and consistent geographic information. The methods for storing this data have changed radically in recent years, paper maps giving way to computer-based storage, and manual drafting to digital production techniques. However, the underlying principals for ensuring spatial compatibility and consistency among data have remained the same throughout.

The foundation of any geographically-based dataset is a spatial reference system. It is the mechanism through which grids can be placed on maps and navigation can be reliably achieved. A spatial reference system allows a location to be unambiguously identified through a set of coordinates (usually latitude and longitude or easting and northing). It also allows distances and areas to be reliably calculated.

The objective of this booklet is to provide a brief introduction to the development of spatial reference systems. It is intended for casual users of maps and navigation devices who wish to know more about the coordinates that they use. It addresses some of the technical issues surrounding spatial reference systems, presenting formulae when appropriate. It also discusses the benefits of a spatial reference system to the community in general.

Every attempt has been made to present the material in non-technical terms. However, some reference to the science and terminology of Geodesy and Geodetic Datums has been both inevitable and necessary. It is hoped that the associated explanation is sufficiently clear to the reader.

A little bit of Background

The Need for a Reference Model

The Earth is a very complex shape. Its surface is disturbed by mountain ranges and deep oceans. In order to map its geography, a reference model is needed which will allow such topographic irregularities to be recorded. The model needs to be simple so that it is easy to use. In addition:

- It needs to include a coordinate system which will allow the positions of features to be uniquely identified,
- It needs to be readily associated with the physical world so that its use is intuitive.

A Flat Earth

If the area being mapped is small (for example, a 10 Km square), a suitable reference model can be provided by a simple three dimensional (3D) framework (see Diagram 1). The coordinate axes of the framework are arranged such that:

- The horizontal axes (N and E) are aligned in the directions of North and East,
- The horizontal plane is defined as being coincident with sea level, a physically identifiable surface,
- The height axis (H) is perpendicular to the horizontal plane and coincides with the direction of gravity (the line along which an object will fall if you drop it).

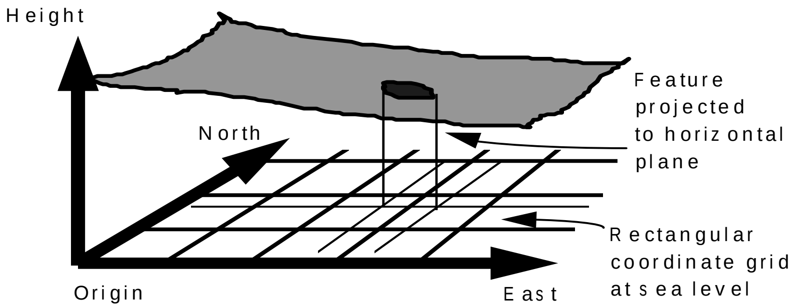

Diagram 1: *Local 3D Coordinate System*

The positions of individual features are projected vertically to the horizontal plane, allowing their positions relative to one-another to be clearly (and mathematically) identified. Furthermore, the height of each feature relative to sea level is provided by the vertical distance from the horizontal plane.

The orientation to North (the direction to the pole), together with the adoption of sea level as the reference for heights, provide the necessary association with the real world to make the reference system intuitively useable.

A Curved Earth

The 'flat earth' approach breaks down rapidly when larger areas are to be measured or mapped. There are two complicating factors, being:

- the curvature of the Earth,
- the composition of the Earth's interior.

Curvature of the Earth

The curvature of the Earth is the most obvious problem. It forces the replacement of the 'flat earth' model with a 'curved earth' model. The selection of a curved reference surface is an important consideration as it needs to satisfy two criteria. Not only must it closely represent the shape of the Earth, it must also be mathematically simple to use.

At first glance, the most appropriate figure would seem to be a sphere. It is geometrically very simple. Moreover, it obviously approximates the general shape of the Earth.

Appearances can be deceptive!

In fact, the Earth is poorly approximated by a sphere, being significantly wider at the equator than between the poles. For this reason, a sphere is rarely used as a reference surface.

A spheroid (also referred to as an ellipsoid) provides a better option. A spheroid is the figure generated by rotating an ellipse about its minor (shorter) axis (See Diagram 2). It has the advantages of accommodating the bulge at the equator while remaining relatively simple mathematically. For these reasons, a spheroid is the figure usually chosen to represent the shape of the Earth.

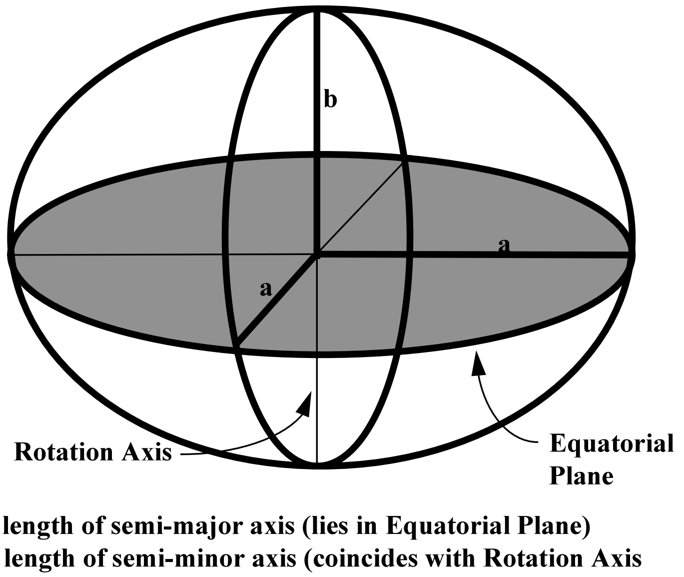

Diagram 2: *Spheroid or Ellipsoid*

Composition of the Earth's Interior

The second complicating factor is less obvious. It relates to the Earth's composition.

The composition of the Earth is not uniform. It varies from place to place. There are variations in the density and distribution of the different rock types. There are also distribution irregularities caused by mountain ranges and ocean depths. Together, the variations lead to anomalies in the Earth's gravity field. This, in turn, causes irregularities in the surface of sea level which is shaped by the gravity field. You can't see them but they are there.

It is possible to mathematically formulate a model of the Earth's sea-level surface. However, the model is very complex and is not suitable for recording the geographic positions of features. Sea level does not, for example, coincide with a mathematically simple surface such as a sphere or spheroid (see Diagram 3).

Why is this important?

In the small-area 'flat earth' model, the horizontal plane containing the N and E axes was able to be positioned such that it coincided with sea level. It is not possible to do this with a 'curved earth model' due to the irregularity of the sea-level surface. Instead, a best-fit of the spheroid to the sea-level surface has to be performed.

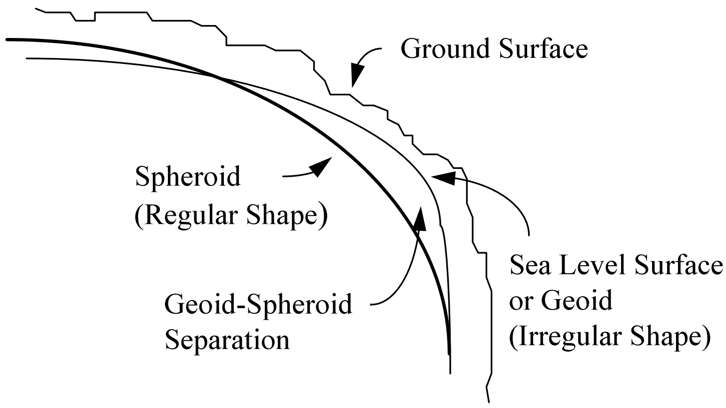

Diagram 3: *Relationship Between Spheroid and Sea Level*

Implications for Users

What are the implications of the lack of coincidence between the spheroid and the sea level surface to the user?

There are several. They are illustrated in Diagram 4.

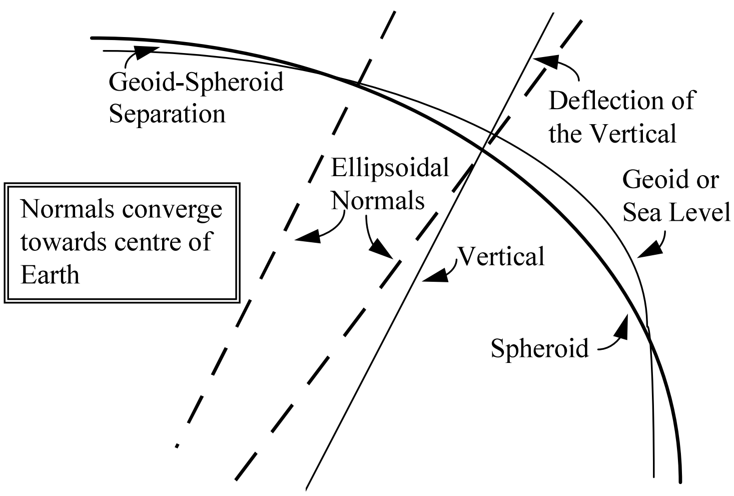

Diagram 4: *The Vertical and the Ellipsoidal Normal*

The Deflection of the Vertical

First of all, there is the difference between the vertical (the line perpendicular to the sea-level surface) and the ellipsoidal normal (the line perpendicular to the spheroid).

The vertical is coincident with the direction of gravity at a point. It is the line along which an object falls when it is dropped. The vertical is very important to measurements taken by conventional surveying instruments (such as theodolites and levels). These instruments are set-up such that their rotation axes are either coincident with, or perpendicular to, the vertical. Consequently, all angles are measured relative to the vertical.

The ellipsoidal normal, on the other hand, is the line along which a feature on the earth's surface is projected down to the spheroid. It is the line which is used in computations involving observations at the feature.

To summarise:

- the vertical is the line associated with measurements, while
- the ellipsoidal normal is the line associated with computations.

In our 'flat-earth' model, the two lines are considered coincident. In a 'curved earth' model, they generally are not. Accordingly, in order to use an angle measurement in a computation process, the measurement should first be corrected for the 'tilt' between two lines.

The tilt is referred to as the Deflection of the Vertical. It is described by two small angles, these being the component tilts in the northerly and easterly directions.

The Geoidal Undulation

There are two reference surfaces which are commonly used as a basis for height values. They are the sea-level surface (or geoid) and the spheroid. In our 'flat earth' model, the two surfaces are considered coincident. In a 'curved earth' model, this is rarely the case.

In most parts of the world, distance above sea-level has been the traditional mechanism for measuring height. This has been due to:

- a preference for a physically identifiable surface as a reference,
- the importance of sea level to economic activity (for example shipping),
- the linkage between sea level, the earth's gravity field, and the conventional instrumentation used to measure height differences,
- the importance of gravity-related heights to water-flow problems.

The Australian Height Datum (AHD) is an example of such a system.

Spheroidal heights (heights relative to the spheroid) are now becoming increasingly popular. Until recently, they were more difficult to determine than sea-level heights. However, satellite surveying and satellite navigation receivers have reversed this situation, producing spheroidal heights as part of their output.

The distance between the geoid and the spheroid is referred to as the geoid-spheroid separation or geoidal undulation (See Diagrams 3 and 4). If it is known, a sea-level height can be converted to a spheroidal height and vice-versa. There are a number of methods for calculating the separation. All are mathematically complex and are beyond the scope of this discussion.

Convergence of the Ellipsoidal Normals

In our 'flat earth' model, the lines projecting surface features to the horizontal plane were parallel. In a 'curved earth' model, the ellipsoidal normals converge towards the centre of the spheroid (See Diagram 4).

Consequently, a distance measured on the surface of the earth has to be shortened before it can be used in computations on the spheroid. The amount of the shortening will depend on the height of the measurement above the spheroid. It is approximately one millimetre per kilometre for every 6.3 metres of height.

Geodesy and Geodetic Datums

Geodesy

It is now time to define the term Geodesy. Essentially, Geodesy is the branch of science concerned with the determination of the size and shape of the Earth. Its range of contributing activities is vast. They include the processing of survey measurements on the curved surface of the Earth as well as the analysis of gravity measurements.

A 'curved-earth' reference model of the world is referred to as a geodetic datum. The characteristics of a geodetic datum have been discussed in the preceding sections. However, to summarise:

- A geodetic datum is a simplified mathematical representation of the size and shape of the earth.
- It usually takes the form of a spheroid, this being an ellipse rotated about its minor (shorter) axis.
- A geodetic datum is vital to all activities involving spatial data. The spheroid provides a mathematically simple surface for performing surveying and navigation computations over a wide area. It also provides a reference surface on which to base mapping and Geographic Information Systems (GIS).
- The surface of the spheroid is positioned such that it is a best-fit to the Earth's sea level surface (ie to the geoid). An exact fit to sea level is not possible due to anomalies in the Earth's gravity field. Gravity anomalies (caused by variations in the density and distribution of the earth's mass) cause irregularities in the sea level surface. This renders sea level unsuitable as a horizontal reference surface for mapping activities.

· Sea level is widely used as the reference surface for the measurement of height. The contours on a map will usually show height above sea level. However their position will be mapped on the spheroid.

Types of Geodetic Datum

Geodetic datums are usually classified into two categories. These are known as local geodetic datums and geocentric datums.

A Local Geodetic Datum is a datum which best approximates the size and shape of a particular part of the earth's sea-level surface. Invariably, the centre of its spheroid will not coincide with the Earth's centre of mass.

Until very recently, most country's spatial information systems were based on local geodetic datums.

The Australian Geodetic Datum (AGD) is an example of a local datum. Its spheroid is a good approximation to the size and shape of the sea-level surface in the region of Australia but a poor approximation in other parts of the world (see Diagram 5).

A Geocentric Datum is one which best approximates the size and shape of the Earth as a whole. The centre of its spheroid coincides with the Earth's centre of mass (see Diagram 6).

Geocentric datums do not seek to be a good approximation to any particular part of the Earth. Rather, their application lies in projects or undertakings which have global application.

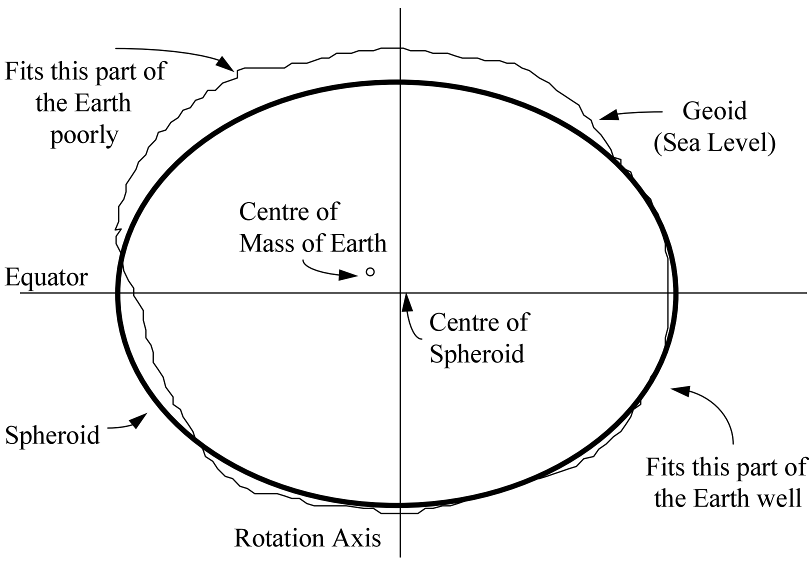

Diagram 5: *Local Geodetic Datum*

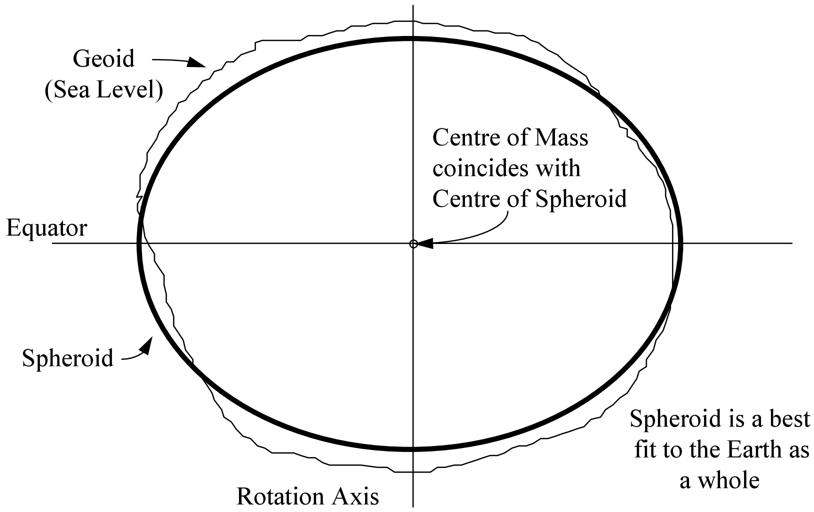

Diagram 6: *Geocentric Geodetic Datum*

The Global Positioning System (GPS), which is operated by the United States Department of Defence, utilises a geocentric datum to express its positions because of its global extent. The Russian GLONASS satellite navigation system also uses a geocentric datum. However it is a different datum to that used by GPS.

Identifying Position on a Geodetic Datum

Two coordinate systems are implicitly associated with a geodetic datum.

A Geodetic coordinate system is the system naturally associated with the spheroid. It allows positions on the earth's surface to be described in terms of latitude, longitude and spheroidal height (See Diagrams 7, 8 and 9).

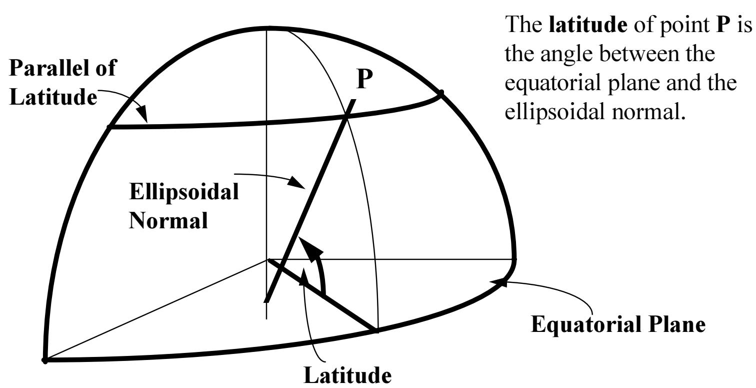

Diagram 7: *Geodetic Latitude*

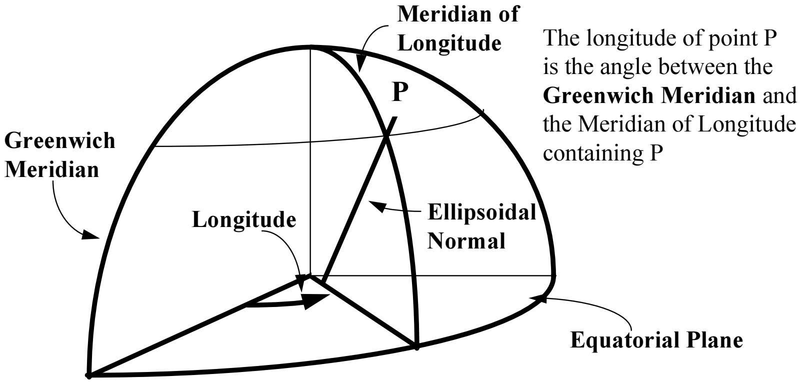

Diagram 8: *Geodetic Longitude*

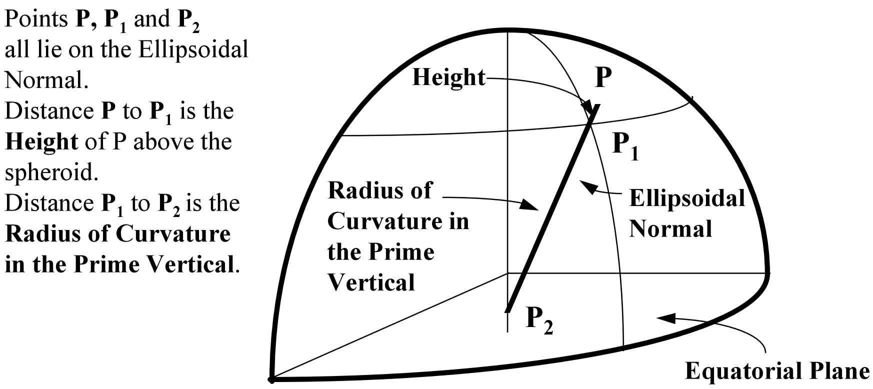

Diagram 9: *Spheroidal Height*

An X,Y,Z Cartesian coordinate system (Diagram 10) may also be associated with the datum, such that:

- the positive X axis lies in the equatorial plane and passes through 0^0 longitude,
- the positive Y axis lies in the equatorial plane and passes through 90^0 East longitude,
- the positive Z axis is parallel to the earth's rotation axis and passes through 90^0 North latitude.

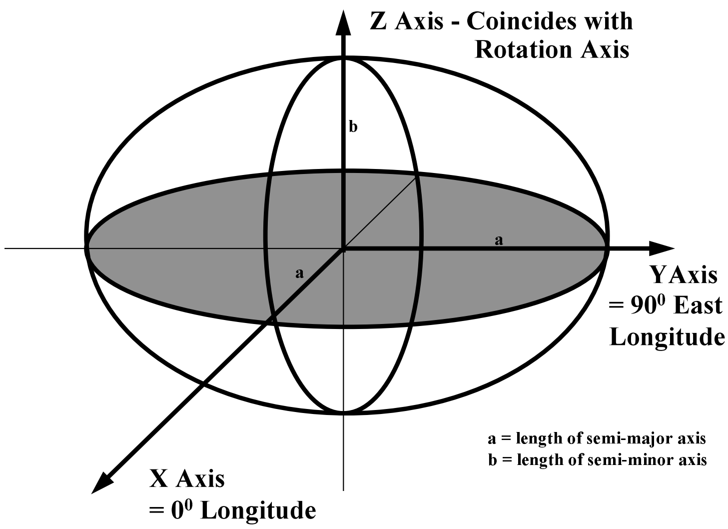

Diagram 10: *Cartesian Coordinate System*

A rigorous (or near-rigorous) conversion of coordinates between the two systems is easily achievable (see Diagram 11).

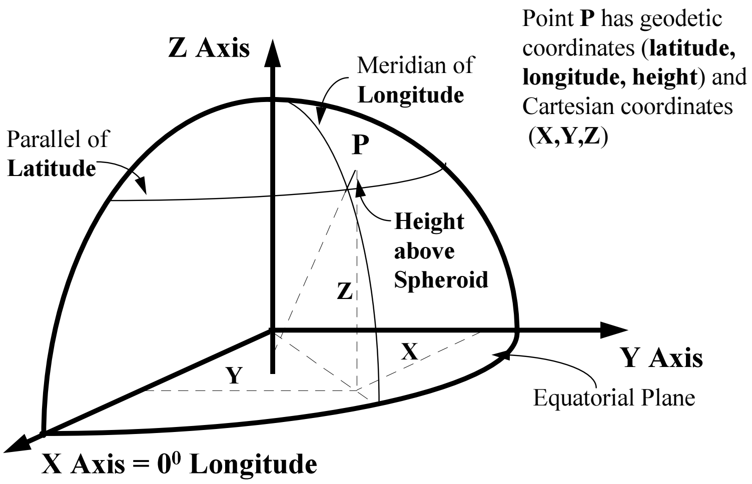

Diagram 11: Geodetic and Cartesian Coordinates

For the mathematically inclined, the formulae are included as Inserts 1 and 2.

Insert 1: Conversion - Geographic to Cartesian

The formulae for converting latitude, longitude and spheroidal height to X, Y, Z are:

$$X = (N + h)\cos\varphi\cos\lambda$$
$$Y = (N + h)\cos\varphi\sin\lambda$$
$$Z = ((b^2/a^2)N + h)\sin\varphi$$

where

X, Y, Z	*are the Cartesian coordinates of the point*
φ, λ	*are the latitude, longitude of the point*
h	*is the height of the point above the spheroid along the ellipsoidal normal*
a, b	*are the lengths of the semi-major and semi-minor axes of the spheroid*
N	*is the radius of curvature in the prime vertical at the point.*

$$= \frac{a^2}{\sqrt{a^2\cos^2\varphi + b^2\sin^2\varphi}}$$

Insert 2: Conversion - Cartesian to Geographic

The formulae for converting X, Y , Z to latitude, longitude and spheroidal height are:

$$\varphi = \arctan((Z + e'^2\, b \sin^3\theta) / (p - e^2\, a \cos^3\theta))$$

$$\lambda = \arctan(Y / X)$$

$$h = (p / \cos\varphi) - N$$

where

X, Y, Z	are the Cartesian coordinates of the point
φ, λ	are the latitude, longitude of the point
h	is the height of the point above the spheroid
a, b	are the lengths of the semi-major and semi-minor axes of the spheroid
N	is the radius of curvature in the prime vertical

$$= \frac{a^2}{\sqrt{a^2\cos^2\varphi + b^2\sin^2\varphi}}$$

e^2	is the eccentricity squared $= (a^2 - b^2) / a^2$
e'^2	is the second eccentricity squared $= (a^2 - b^2) / b^2$
θ	is an auxiliary quantity $= \arctan(Z\, a / p\, b)$
p	$= \sqrt{X^2 + Y^2}$

Map Projection Coordinates

A third coordinate system is that provided by a map projection. A map projection is used to enable the curved surface of the spheroid to be represented on a flat sheet of paper (in other words, a map). The projection process results in the map's spatial representation being distorted. (Imagine the stretching and tearing that you would have to do to a basketball to make its curved surface lie flat on the ground.) The magnitude of the distortion can be calculated, allowing corrections to be made when necessary.

A rectangular grid coordinate system (similar to our 'flat earth' grid) is associated with every map projection. Map projection coordinates are described in terms of Easting and Northing, being distances to the East and North of an origin. They are usually expressed in units of metres or feet.

There are a large number of map projections available for use. Each can be implemented on any spheroid. One of the most commonly used is the Transverse Mercator Projection, known as the Gauss-Kruger System in some parts of the world. The formula for converting latitude and longitude to grid easting and northing are too complex to be quoted in this document. However they can be found in any reference book on map projections.

Insert 3: Universal Transverse Mercator (UTM)

Universal Transverse Mercator (UTM) is a global implementation of the Transverse Mercator projection. It divides the earth into 60 zones, each being bounded by meridians of longitude (extending from the North Pole to the South Pole). Imagine an orange being comprised of 60 segments. Each segment would be equivalent to a UTM zone.

Each UTM zone is restricted to a width of six degrees of longitude. This is to avoid distortions becoming too large. The meridian at the zone's center is referred to as the Central Meridian.

The point of intersection between the Equator and the Central Meridian, is assigned the following values, ensuring that all coordinates within the zone are greater than zero,

East	*500,000.000 metres*
North	*0.000 metres (Northern Hemisphere) or*
	10,000,000.000 metres (Southern Hemisphere)

Under the UTM system, each East and North coordinate pair could refer to one of sixty points on earth. The same coordinate pair exists in each of the sixty zones. Consequently, the zone number must be quoted with the East and North values. The zone number is effectively a third coordinate.

The Australian Map Grid is a UTM projection. It lies in UTM Zones 47 to 58.

Geodetic Networks

A Geodetic Survey is an activity concerned with the determination of latitude, longitude and height for points on the surface of the Earth. The activity involves the acquisition of surveying measurements (for example, angles, distances, height differences, observations to artificial satellites) and their subsequent processing to produce coordinates.

The geometry of the measurements in a geodetic survey is known as a Geodetic Network. The principal geodetic network covering a country or continent is referred to a Primary Geodetic Network (or Primary Network for short). The vertices in the primary network are commonly known as Trigonometric (Trig) or Control Stations (see Diagram 12).

Trig Stations are monumented by ground marks. The marks normally consist of brass plaques set in concrete. Traditionally they have been located on hill tops, and have frequently been witnessed by some form of cairn or survey beacon. The beacons serve to aid the location of stations as well as to protect them from possible destruction. In addition, they provide a landmark for distant users.

Prior to the late 1950s' horizontal geodetic networks were mostly established by means of triangulation. The methodology was to use theodolites at each trig station to measure the angles between every other visible station in the network. One or two short distances were also measured to scale the network. Orientation (or bearing) was determined from astronomic observations.

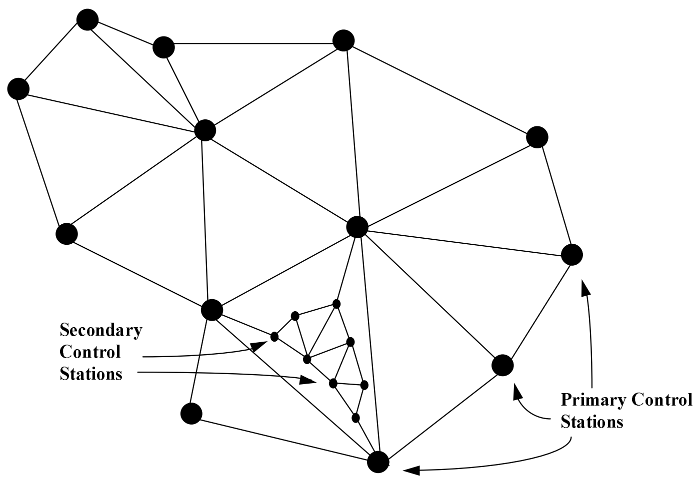

Diagram 12: Geodetic Network

More recently, measurement systems based on artificial earth satellites have become available. The GPS system is the best known of these. The system, which usually requires satellite receivers to be simultaneously positioned at each end of an observed line, has removed the need for network stations to be intervisible. Consequently, trig stations no longer need to be placed on the tops of hills; they can be established in more accessible locations.

Secondary and Tertiary Geodetic Networks are often created following the establishment of a Primary Network. The difference between these and the Primary Network is that their relative accuracy is lower and their station spacing is closer. Stations in a Primary Network will typically be 25 to 50 kilometres apart. Stations in a Tertiary Network may be as close as 200 metres.

Geodetic Datum Definition

The process through which a geodetic datum is formulated to provide a best-fit to the physical earth is referred to as datum definition.

Datum definition can be a long and complex undertaking. Several experimental definitions might be required before a final datum can be declared. Each trial might involve the processing of thousands of survey measurements. The data set might include satellite, gravity and astronomic observations, as well as angle (triangulation) and distance (trilateration) measurements.

In the past, definitions of local datums have depended heavily on astronomic observations. Positional astronomy was used to determine latitude and longitude for a selection of trig stations distributed over the area of interest. A corresponding set of latitude and longitude values were computed using angle and distance measurements (the geodetic network), the computations being done on the surface of the nominated spheroid starting from an origin station.

The two sets of coordinates would inevitably differ due to irregularities in the earth's gravity field (the differences being the Deflections of the Vertical). The differences were analysed to determine whether a revised solution, maybe using a different spheroid or origin station, would produce smaller deflections. When the analysts were satisfied that the smallest differences had been achieved, the datum would be defined by declaring the latitude, longitude and spheroidal height of the origin station as well as the parameters defining the ellipsoid.

The definition of a geocentric datum is a more complex process, involving measurements over the entire globe. The process generates coordinates for all participating stations simultaneously.

The Problems of Multiple Datums

A Multitude of Datums

A geodetic datum is a mathematical concept. It is therefore theoretically possible (although certainly not desirable) to define an infinite number of datums to cover any area. Ideally, a single datum should be used in a country or region so that all data can be referenced to a common coordinate system. In practice it is frequently necessary to deal with two or more datums for technical or political reasons (for example, a local datum for existing mapping and a newer geocentric datum for satellite navigation).

It is important to understand that the coordinate values for a point are dependent on the datum being used. The latitude, longitude and height of a point defined on Datum 1 will almost certainly be different to its latitude, longitude and height defined on Datum 2. The differences may be a consequence of:

- the ellipsoids being different shapes,
- the centres of the ellipsoids being displaced, possibly by hundreds of metres,
- the Cartesian coordinate axes of the two datums not being parallel or being subject to a scale difference.

IT IS VITAL THAT USERS BE FULLY AWARE OF THE DATUM(S) THAT THEY ARE USING. TRYING TO DIRECTLY COMBINE COORDINATES FROM TWO DIFFERENT DATUMS WILL PRODUCE WRONG NUMBERS AND CAN RESULT IN CATASTROPHIC OUTCOMES.

The Australian Geodetic Datum

At present, there are two geodetic coordinate sets in common use throughout Australia. Both have been computed on the Australian Geodetic Datum (a local datum). They are referred to as:

· The Australian Geodetic Datum 1966 (AGD66)

· The Australian Geodetic Datum 1984 (AGD84)

The coordinate sets have been derived using the same spheroid (the Australian National Spheroid) and the same origin coordinate values (see Insert 4). They differ because the AGD84 values were computed using a much improved data set. Effectively, AGD84 is an updated, but distinctly different, version of AGD66. Care should be taken not to mix the two sets of coordinates during mapping operations.

Insert 4: Definition of the Australian Geodetic Datum

The origin station for the Australian Geodetic Datum (AGD) is the Johnston Geodetic Station, sometimes referred to as the Johnston Origin. Its coordinate values, determined by astronomy, are:

Geodetic Latitude	*25^{0} 56′ 54.5515″ South*
Geodetic Longitude	*133^{0} 12′ 30.0771″ East*
Spheroidal Height	*571.2 Metres*

The deflections of the vertical are defined to be zero at this station. The separation between sea-level and the spheroid was defined to be zero in 1966 and +4.9 metres in 1984.

The Geocentric Datum of Australia (GDA)

By the year 2000, Australia will adopt a geocentric datum as its principle national datum. Its coordinate set will be referred to as the Geocentric Datum of Australia 1994 (GDA94)

Its adoption will allow closer integration with international coordinate frameworks and navigation systems. In particular, the datum will coincide almost exactly with the datum used to support the Global Positioning System (WGS84). This will allow GPS-derived coordinates to be used directly with GDA94 in almost all circumstances.

All Australian data sets, including revised hard copy maps, will be progressively converted to the new coordinate system. GDA94 coordinates have already been computed for every trig station in the Australian Primary Geodetic Network. The computations involved a much larger data set than was available at the time of AGD84.

An Important Warning !

THERE ARE MAJOR DIFFERENCES BETWEEN THE GDA94 AND AGD66/84 COORDINATE SYSTEMS.

Remember!

- The AGD coordinates use a local datum, while the GDA coordinates use a geocentric datum,
- The two datums use different shaped ellipsoids (See Insert 7).

As a result, the AGD and GDA coordinates for the same point differ by approximately 200 metres (between 120 and 180 metres in both the easterly and northerly directions).

Note that this is the case for map grid coordinates as well as geodetic coordinates. The map grid coordinates for a point are directly related to the geodetic coordinates of the same point. If the geodetic coordinates change due to the adoption of a different datum and/or spheroid, the map grid coordinates will also change. THE ADOPTION OF A NEW DATUM CHANGES EVERYTHING.

Transforming Between Datums

Coordinates can be converted from one datum to another if the relationship between the two is known. The relationship is described by two components, being:

- a set of formulae which describe the mathematics of the transformation process,
- a set of parameters, referred to as transformation parameters.

The transformation parameters, which are substituted into the formulae, identify the relationship between the particular datums in question. They are calculated by comparing coordinate sets from the two datums. Consider the typical situation where we have:

- Coordinate Set 1 which has been computed from Network 1 on Datum 1,
- Coordinate Set 2 which has been computed from Network 2 on Datum 2.

The transformation parameters are derived by analysing survey control stations which have coordinates on both datums. The minimum number of stations required for this process depends on the transformation method being proposed. However, there is no maximum limit, and in general, as many stations as possible will be used.

Usually, the two coordinate sets will not coincide exactly. In other words, if the coordinates from Network 1 were transformed onto Datum 2 and compared with the coordinates from Network 2, there will usually be differences between the corresponding values. This is a consequence of the coordinate sets being derived from different measurement sets or network geometries (for example, Network 1 may have been measured by triangulation and Network 2 by GPS). One network will appear distorted relative to the other.

The differences between the transformed Datum 1 coordinates and the network Datum 2 coordinates are known as residuals. Their magnitude provides an indication of the quality of the two networks, as well as an indication of the accuracy of any subsequently transformed coordinates. An exact transformation (in other words, zero residuals) will only be achieved if the coordinates on both datums were computed using identical measurement data sets. This is seldom the case.

Transformation parameters are commonly generated by government mapping organisations. They are freely available to users and can be used in conjunction with GPS receivers and other spatial computer software.

Common Transformation Models

If the XYZ coordinate axes of two datums are known to be parallel and identically scaled, a three parameter transformation can be derived to represent their relationship (see Diagram 13 and Insert 5).

If the coordinate axes are not parallel and identically scaled, a seven parameter transformation can be derived (see Diagram 14 and Insert 6).

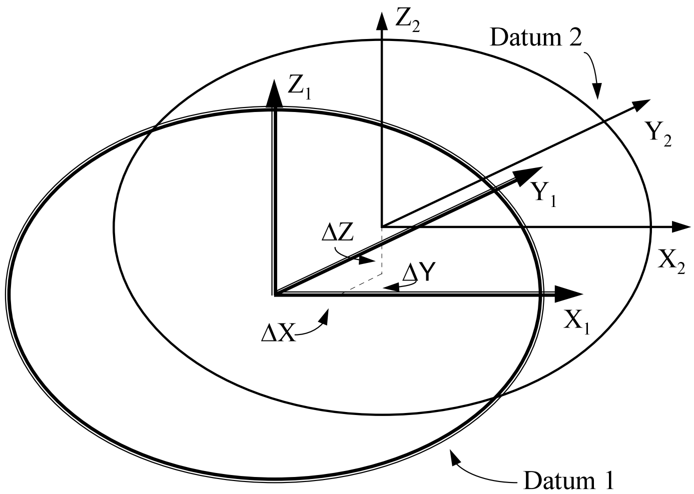

Diagram 13: Three-Parameter Transformation

Insert 5: Three-Parameter Transformation Formulae

$$X_1 = X_2 + \Delta X$$
$$Y_1 = Y_2 + \Delta Y$$
$$Z_1 = Z_2 + \Delta Z$$

where

X_1, Y_1, Z_1 = *Cartesian Coordinates of Datum 1*

X_2, Y_2, Z_2 = *Cartesian Coordinates of Datum 2*

$\Delta X, \Delta Y, \Delta Z$ = *The difference between the centres of the two spheroids.*

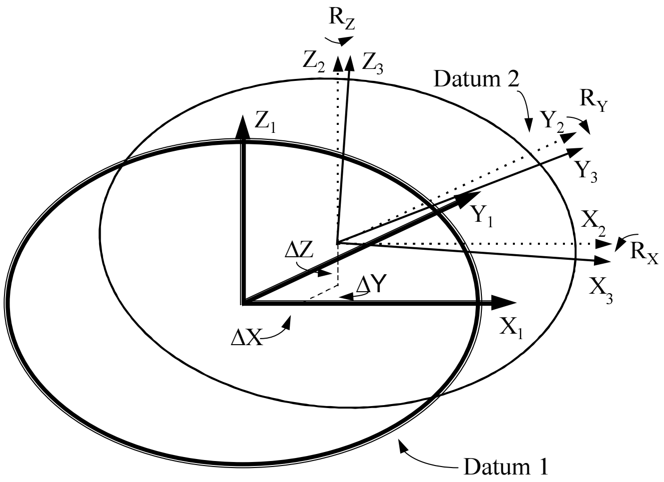

Diagram 14: Seven-Parameter Transformation

Insert 6: Seven Parameter Transformation Formulae (Bursa-Wolf Model)

$$\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + S_c \begin{bmatrix} 1 & R_Z & -R_Y \\ -R_Z & 1 & R_X \\ R_Y & -R_X & 1 \end{bmatrix} \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix}$$

where

X_1, Y_1, Z_1 = *Cartesian Coordinates of Datum 1,*
X_2, Y_2, Z_2 = *Cartesian Coordinates of Datum 2,*
$\Delta X, \Delta Y, \Delta Z$ = *The difference between the centres of the two spheroids,*
R_X, R_Y, R_Z = *The rotations around the three coordinate axes,*
S_C = *The scale difference between the coordinate systems.*

Rotations are positive anticlockwise about the axes of Datum 2 coordinate system when viewing the origin from the positive axes.

Converting Coordinates Between Australian Datums

Several sets of transformation parameters are being prepared to relate the AGD66 and AGD84 coordinates sets to the GDA94 datum. Parameters suitable for low and medium accuracy data sets, have already been computed. They are listed in Inserts 8, 9 and 10 below.

The sequence for transforming coordinates from Datum ‘1’ to Datums ‘2’ or ‘3’ (see Diagrams 13 and 14) is as follows:

1. [If necessary] Transform map grid coordinates (eg. UTM) on Datum ‘1’ to geodetic coordinates on Datum ‘1’ using the appropriate map projection formulae.
2. Convert the geodetic coordinates on Datum ‘1’ to XYZ Cartesian values on Datum ‘1’ using the formulae in Insert 1.

3. Convert the XYZ Cartesian values on Datum '1' to XYZ Cartesian values on Datum '2' or '3' using the formulae in Insert 5 or 6 respectively.
4. Convert the XYZ Cartesian values on Datum '2' or '3' to geodetic coordinates on Datum '2' or '3' using the formulae in Insert 2.
5. [If necessary] Convert the geodetic coordinates on Datum '2' or '3' to map grid coordinates on Datum '2' or '3' using the appropriate map projection formulae.

Note: Datums '1' and '2'/ '3' may use different reference ellipsoids. It may therefore be necessary to use different values for the semi-major and semi-minor axes in steps 1-2 and 4-5 (See Insert 7 below).

Insert 7: Spheroid Parameters for Australian Datums

This table lists the spheroidal parameters for the principal Australian datums. Note that the spheroids are defined by the semi-major axis and the flattening. The semi-minor axis is derived using the formula $b = a(1 - f)$.

Datum	*Australian Geodetic Datum (AGD)*	*Geocentric Datum of Australia (GDA)*
Spheroid	*Australian National Spheroid (ANS)*	*Geodetic Reference System 1980 (GRS80)*
Semi-Major Axis (a)	*6,378,160.0 metres*	*6,378,137.0 metres*
Flattening (f)	*1 / 298.25*	*1 / 298.257222101*
Semi-Minor Axis (b)	*6,356,774.719 metres*	*6,356,752.3141 metres*

Insert 8: Transformation from AGD66 to GDA94 – Three Parameters

Three Parameter Transformation

The three parameters for transforming AGD66 coordinates to GDA94 are:

ΔX *-127.756 metres*

ΔY *-52.263 metres*

ΔZ *152.893 metres*

- *The parameters are the values endorsed by the Inter-Governmental Council on Surveying and Mapping (ICSM) in November 1997.*
- ***Due to distortions in the AGD66 coordinate set, the accuracy of transformed coordinates computed using these parameters will be no better than FIVE METRES.***
- *To transform from GDA94 to AGD66, reverse the signs on all parameters.*

Insert 9: Transformation from AGD84 to GDA94 – Three Parameters

Three Parameter Transformation

The three parameters for transforming AGD84 coordinates to GDA94 are:

ΔX	*-128.494 metres*
ΔY	*-53.040 metres*
ΔZ	*153.441 metres*

- *The parameters are the values endorsed by the Inter-Governmental Council on Surveying and Mapping (ICSM) in November 1997.*
- ***Due to distortions in the AGD84 coordinate set, the accuracy of transformed coordinates computed using these parameters will be no better than FIVE METRES.***
- *To transform from GDA94 to AGD84, reverse the signs on all parameters.*

Insert 10: Transformation from AGD84 to GDA94 – Seven Parameters

Seven Parameter Transformation

The seven parameters for transforming AGD84 coordinates to GDA94 are:

ΔX	*-117.763 metres*
ΔY	*-51.510 metres*
ΔZ	*139.061 metres*
R_X	*-0.292 seconds*
R_Y	*-0.443 seconds*
R_Z	*-0.277 seconds*
S_C	*-0.191 parts per million*

- *The parameters are the values endorsed by the Inter-Governmental Council on Surveying and Mapping (ICSM) in November 1997.*
- ***Due to distortions in the AGD84 coordinate set, the accuracy of transformed coordinates computed using these parameters will be no better than ONE METRE.***
- *To transform from GDA94 to AGD84, reverse the signs on all parameters.*
- *The rotations must be converted to radians prior to inclusion in the formula.*

The Value of a Geodetic Reference System

(This chapter draws heavily on 'The Use and Value of a Geodetic Reference System', published by Earl F. Epstein and Thomas D. Duchesneau at the University of Maine in 1984.)

A geodetic network, and its associated spatial reference system, is a fundamental component of a nation's infrastructure. Rather than being an end in itself, it derives its value from being an input to other production processes. The utility of the network is determined by identifying the products which are dependent upon the network's unique properties. The unique property of the geodetic network is its ability to integrate multiple geographically-dependent data sources into a single geographic reference frame.

The concept of a Land Information System (LIS) is illustrated in Diagram 15. Typically the system will be multi-layered, each layer comprising data which relates to a particular theme. For example, one layer may represent the roads in an area. A second layer may illustrate the distribution of a particular plant or animal. Further layers may contain the noise contours surrounding an aircraft flight path, or the location of electricity transmission lines.

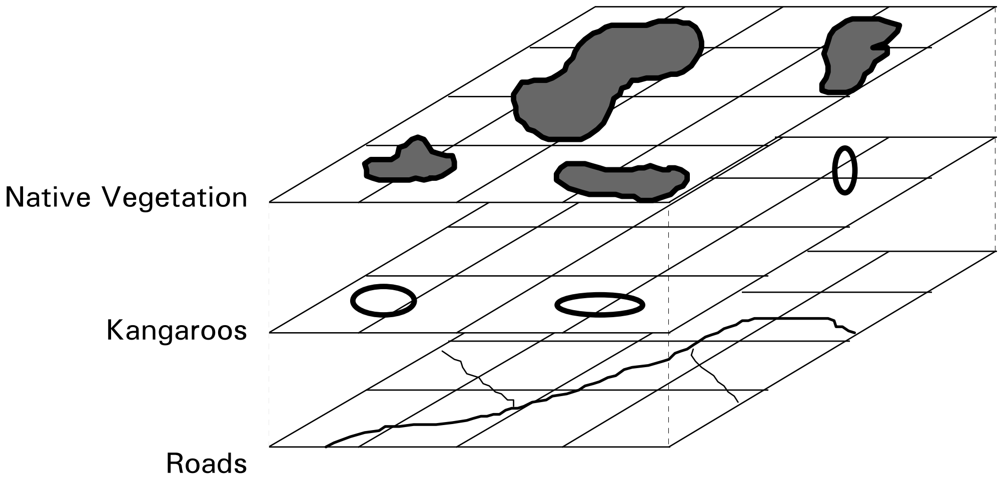

Diagram 15: Layers in a Land Information System

Modern Land Information Systems are computer based. The data for these layers may be stored on a single computer. Alternatively, they may be stored on the computers of the custodial authorities, and accessed by networking or other means.

In order to function effectively, the LIS must possess one <u>essential</u> attribute. It must have the ability to geographically relate and inter-relate the data in its layers. Consider, for example, an LIS layer representing the distribution of native vegetation in a region. The utility of the layer will be governed by its ability to accurately represent the position and size of one pocket of vegetation relative to other such pockets. A spatial referencing system which permits the definition of position and extent in terms of coordinates has great application in this situation.

However, the adoption of such a system becomes even more important when the data in two or more layers are to be combined. The requirement to inter-relate the data from different themes necessitates the use of a common coordinate system in each. Suppose it was desired to inter-relate the native vegetation data with data describing the distribution of kangaroos in the same region. If the positions of the kangaroo concentrations and vegetation pockets were represented in different and independent coordinate systems, the inter-relation process would be very difficult. However, if the same geographic referencing system were used for both data sets, the required spatial relationships could be determined very quickly and very efficiently.

The role of the spatial framework in an LIS, therefore, is to provide the medium through which data sets can be inter-related geographically. The geodetic network provides survey marks whose positions are accurately determined in terms of a single national coordinate system. If the positions of the items in each layer of the LIS are described in terms of this system (either through direct measurement to the survey marks, or from the grid on a map), the data sets will be capable of efficient integration.

The demand for universal compatibility of geographically related data most often arises from secondary and tertiary users of information. The authorities which generate data sets for a specific purpose (primary users) tend to have little interest in the needs of those who wish to inter-relate the data sets at a later time. For example, organisations which compile data regarding soils and geology (for agriculture and mining) may not consider the needs of hydrologists who are interested in correlating that data with water run-off. The hydrologist, as a secondary or tertiary user of the data, clearly requires a means to render the data sets compatible.

In the absence of a common spatial framework, an alternative medium would be needed to facilitate the inter-relation process. For example, it might be decided to undertake measurements to relate the elements of each data set to easily identifiable physical features. The data sets could then be inter-related through these common features. However, this process would also necessitate the acquisition of measurements between the physical features themselves, in order to establish their relative positions in some local coordinate system. This immediately raises questions regarding efficiency.

The duplication of the measurement process by every organisation needing to inter-relate data sets would result in a clear economic cost being incurred by the community. It must follow therefore that the avoidance of that cost through the availability of a common spatial framework represents a source of economic benefit.

This is the crux of the argument regarding the value of a common spatial framework and geodetic network. If a common geographical referencing system were not available, alternative means would have to be found to enable data sets to be inter-related. The existence of a permanently monumented geodetic network effectively means that the bulk of the required measurements only ever have to be done once.

Connecting the individual data sets to local control stations allows the positions of the data items to be expressed in a common coordinate system, avoiding the waste which would result from separately inter-relating each pair of data sets. Clearly the avoided costs would be considerable, effectively representing the costs of remeasuring the network several times over without gaining the benefit of universal consistency.

An investigation into the value of a geodetic network was undertaken in the United States of America (Epstein E.F. and Duchesneau T.D, 'The Use and Value of a Geodetic Reference System', University of Maine at Orono, April 1984).

It endeavoured to quantify the avoided costs resulting from network availability through an analysis of case studies. The case studies included projects involving -

- Land use and development plans,
- Watershed and related water studies,
- Construction of capital works, in particular highway construction.

Each of the cases was characterised by a frequent need for accurate and compatible data.

The study concluded that the ratio of benefits to costs flowing from the network lay in the range 1.7 to 4.5. Furthermore, these figures were considered to be conservative due to the non-availability of certain data relevant to the study.

A Final Word

The common spatial framework provided by a geodetic datum and network is a major resource to the Australian community. The network of primary, secondary and tertiary control stations provides the physical infrastructure through which Australia's geographical referencing system is established and maintained. This, in turn, enables the positions of all earth-related information to be expressed in a common coordinate system, thus offering large efficiencies when integrating dissimilar or spatially separated data.

The need to know location and position is so pervasive throughout the activities of Australian society, that it is almost impossible to fully appreciate the extent to which the geographic referencing system is used. Indeed, the demand for universal compatibility among data sets by those who are not part of the survey industry is extensive and appears to be growing. When viewed in this context, it must be concluded that an increasing number of projects will owe both their economic and technical viability to the spatial infrastructure defined by Australia's geodetic network.

Acknowledgments

This paper has been prepared from training material and presentations which I have gathered or developed over a number of years.

The formulae in Inserts 1 and 2 were taken from 'GPS Theory and Practice' by B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins. (Springer-Verlag, 1992).

Information regarding the definition of the Australian Geodetic Datum in Insert 4 was taken from 'The Australian Map Grid Technical Manual' (Special Publication 7, National Mapping Council of Australia, Australian Government Publishing Service, 1972).

The transformation parameters in Inserts 7, 8 and 9 were derived by the Australian Land Information Group (AUSLIG) in 1997 and were endorsed by the Inter-Governmental Council on Surveying and Mapping (ICSM) in November 1997.

The chapter entitled 'The Value of a Geodetic Reference System' draws heavily on work published by Earl F. Epstein and Thomas D. Duchesneau at the University of Maine in 1984.

The author wishes to thank all reviewers of the manuscript, in particular:

- Geoff Sandford, Resource Information Group, Department for Environment, Heritage and Aboriginal Affairs,
- Brian Burbidge, ICSM Executive Officer, AUSLIG,
- Rod Menzies, Chief Surveyor, Department of Urban Services, ACT
- Charles Zahra, Land Information Centre, Bathurst, New South Wales,
- Graeme Rush, Department of Environment and Natural Resources, Queensland
- Alan Deck, Hydrographic Service, Royal Australian Navy
- Gordon Homes, Hydrographic Service, Royal Australian Navy
- Wayne Jones, Hydrographic Service, Royal Australian Navy

Finally, the author wishes to thank Dr Tom Stubbs, Director Resource Information, Department of Environment, Heritage and Aboriginal Affairs for his support, and the Honourable Dorothy Kotz, M.P., Minister for Environment and Heritage.