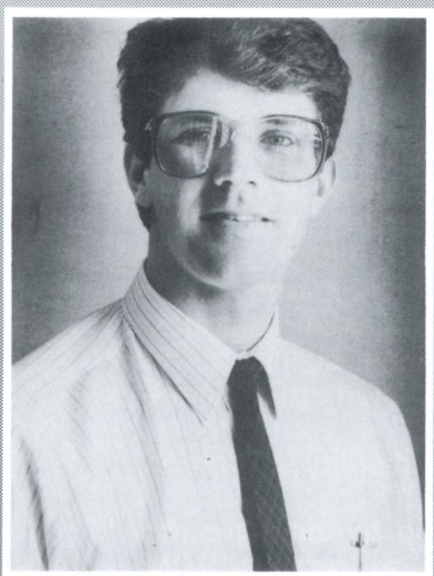
The Geodesy Corner

DATUMS AND ELLIPSOIDS - ARE THEY THE SAME?

BY JAMES FERGUSON



Datums, ellipsoids, spheroids, reference surfaces, surfaces of revolution, oblate spheroids of revolution, the geoid - do these terms sound familiar? Along with a host of others, the foregoing terms have been used to describe the basis for relating survey measurements to the earth, and compare measurements to each other. But do all these terms mean the same? In short, the answer is no, but some of the terminology does describe the same thing. As we will discover, the distinct elements are intrinsically linked, and can be related to each other using some fundamental principles.

Before we get too far, let us simplify the terminology, and decide on three items that will guide us through this discussion - a datum, an ellipsoid and the geoid.

A datum is a system defined and used to relate and present coordinate information in a logical and repeatable fashion. It has an origin, orientation, and is propagated or extended through many types of survey methods. In general, datums are used to orient large systems of geodetic control such as triangulation and trilateration chains, and in a more modern context, satellite survey networks.

A simple example of a "user defined" datum is the land surveyor's use of a local coordinate system to control a subdivision. It will have an origin with a known coordinate (10,000, 10,000 for example), and will have orientation in the form of a baseline with a known bearing or azimuth. It will probably also have some type of vertical information, with the top of a fire hydrant or manhole cover as the initial reference elevation. Hence we can speak about horizontal datums and vertical datums.

An ellipsoid is a three dimensional mathematical figure, used to best represent the figure of the earth over a specific geographic area. Some ellipsoids attempt to be a "best fit" over the whole earth, while others are best suited to an area, say the size and location of North America. Throughout the world there are many definitions of these "football" shaped figures, and some of the more common ones are the Clarke 1866 ellipsoid, the WGS84 ellipsoid and the GRS80 ellipsoid. Ellipsoids also have an origin and orientation, and several parameters to describe their size and shape.

The geoid is described as the best approximation of mean sea level over the whole earth [Vanicek and Krakiwsky 1982]. Since mean sea level is not a smooth regular surface as an ellipsoid is, neither is the geoid. It is a mathematical figure however, and has an undulating pattern related to the terrain of a particular geographic area.

In order to further our discussion let us take one particular datum, relate it to an ellipsoid, and discover how we can work with it in surveying. For the most part, all control surveys performed in Canada (and until recently North America) have been referred to two distinct datums - the North American Datum 1927 or NAD27, and the North American Vertical Datum 1929 or NAVD29. As one might guess, NAD27 is a horizontal datum, and NAVD29 is a vertical datum. We will talk about NAVD29 and height, in a future article.

NAD27 is related mathematically to the Clarke 1866 ellipsoid. That is, we are using a mathematical figure (Clarke 1866) to approximate the surface of the geoid so that we can relate earth based measurements to a common system. It is important to note that in some places the ellipsoid will be below the geoid, and in other places it will be above the geoid. Furthermore, the earth's surface topography undulates above and below the geoid.

The origin of the Clarke 1866 ellipsoid is somewhere near the centre of mass of the earth (theoretically known but unnecessary here), and its orientation can be described by a cartesian coordinate system (X,Y,Z). The Z axis is very nearly parallel to the axis of rotation of the earth, the X axis is pointed in the direction of the Greenwich meridian, and the Y axis is perpendicular to the X axis. The two principal parameters used to describe the ellipsoid are its semi-major and semi-minor axes, designated by "a" and "b" respectively. These two parameters are subsequently used to determine the flattening of the ellipsoid "f", its eccentricity "e", and other desired quantities.

The NAD27 datum has its origin at a place known as Meades Ranch in Kansas, U.S.A. An observed latitude and longitude at this point furnish the origin, and an azimuth mark from the origin supplies the orientation for the system. The actual NAD27 framework was the culmination of an adjustment done in 1927, utilizing all appropriate triangulation and trilateration networks surveyed up to that time. Subsequent adjustments have been done nationally and internationally using additional survey information, but the system is still referred to as NAD27.

Any measurement of distance, azimuth, height, latitude or longitude performed on the earth's surface is affected by the earth's gravity field. That is, the levelling of the instrument using a level bubble of vial is directly related to the force and direction of gravity at the survey station. It is common practice to call these earth based measurements astronomical quantities. The line perpendicular to the level of the instrument is known as the plumb line, and plumb line is naturally parallel to the gravity vector at the station. If the plumb line is extended through the reference ellipsoid, one would find that it is not perpendicular to the survey of the ellipsoid. Hence we discover angular differences known as the deflections of the vertical. Practically speaking, these deflections translate into the difference between astronomic and geodetic positions - where astronomic measurements are actual earth based measurements, and geodetic quantities are those that are related to the reference ellipsoid. The difference is known as the Laplace correction.

Using the basic parameters of the reference ellipsoid ("a" and "b"), two other important quantities must be determined in order to interrelate different forms of coordinates; the radius of curvature of the ellipse in the prime vertical (along a line of longitude), and the radius of curvature along the meridian (along a parallel of latitude). Both of these quantities change as a function of the latitude of the station in question, and are fundamental in relating earth based measurements to the ellipsoid (and vice-versa), via complex mathematical formulae.

How can the above theory be applied to a common survey scenario? Since a large proportion of the work a land surveyor does is over a relatively small area, it can be performed locally in a user defined plane system. In this scenario the survey area is treated as a flat surface, and all measurements are related to a local origin. When the coordinate system is required, such as the Universal Transverse Mercator (UTM) map projection, a common reference system is needed. In fact, our UTM example is convenient because it is based on the Clark 1866 reference ellipsoid. This means that any new survey can be transformed into UTM coordinates as long as several existing stations having UTM coordinates are occupied in the same survey. An adjustment procedure then produces coordinate information in a common system, relative to the existing control datum (UTM on NAD27). A word of caution here, if the existing control has been integrated into a later adjustment (ie. MAY 1976 adjustment used by Ontario's MNR) then it is essential that all the existing coordinate information be taken from the same adjustment. Disastrous results would occur if control coordinates were mixed between different adjustments, even though they are referenced to the same ellipsoid.

There are many applications in which the knowledge of a survey datum, either horizontal or vertical, and its corresponding reference ellipsoid are essential to the production of useful coordinate information. In addition, there are many aids to the surveyor for the computation of such systems. Software has been written to allow a surveyor to: traverse on an ellipse using a known starting coordinate, perform transformations between mapping plane coordinates and ellipsoidal coordinates, adjust networks using a combination of existing control and new survey data, and to perform a variety of other computations too numerous to mention here.

Further reading on the topic of datums and ellipsoids can be found in several sources - Vanicek and Krakiwsky [1982] is a good one, as is Ivan Mueller's Spherical and Practical Astronomy [1969]. Additional information can be found in a number of traditional surveying textbooks. Actual coordinate information can be obtained from the Geodetic Survey of Canada (Data Services Branch), as well as the Ministry of Natural Resources in Ontario.

Next time in the Geodesy Corner - "Azimuths - Grid, Astronomic and Geodetic".

The Geodesy Corner - Author Profile

James Ferguson graduated from the University of Toronto with a B.Sc. in Survey Science. He is a member of the A.O.L.S., holding a Certificate of Registration in the field of Geodesy. James works with the Ottawa based firm GEOsurv Inc.