CHAPTER 1

The Figure of the Earth and the reference surfaces used in surveying and mapping

The precise shape of the earth is usually referred to as a 'geoid', a term which conveys nothing beyond earth-shaped.

G. P. Kellaway, Map Projections, 1946

Introduction

Geodesy is the science concerned with the study of the shape and size of the earth in the geometrical sense and with the study of certain physical phenomena, such as gravity, in seeking explanation of fine irregularities in the earth's shape. The subject is intimately linked with surveying and cartography. A major part of the evidence about the shape and size of the earth is based upon surveys. Indeed in some European languages the word 'geodesy' is practically equivalent to English usage of the word 'surveying'. Knowledge about the earth's size and shape is indispensable if we are to make maps of its surface. Put in the simplest form, it is necessary to know the size of the earth in order to make maps of it at known scale.

We know that the earth is a nearly spherical planet upon which are superimposed the surface irregularities created by land and sea, highland and lowland, mountains and valleys. However these topographical irregularities represent little more than a roughening of the surface. Since the radius of the earth is about 6371 km and since the major relief features do not rise more than 9 km above or fall more than 11 km below sea level, they are relatively less important than, say, the seam on a cricket ball or the indentations on the surface of a golf ball. For example, if the earth is drawn to scale as a circle of radius 6 cm, which is almost as large as the width of this page can accommodate, the variation in line thickness of the circumference which would show the entire height range from Mount Everest to the Mariana Trench at the same scale is less than 0.2 mm.

The idea that the earth is a sphere dates from the Greek geometers of
the sixth century BC. The first serious attempt to measure the size of this sphere was the classic experiment carried out by Eratosthenes in the third century BC.

Towards the end of the seventeenth century, Newton demonstrated that the concept of a truly spherical earth was inadequate to explain the equilibrium of ocean surface. He argued that because the earth is a rotating planet, the forces created by its own rotation would tend to force any liquids on the surface towards the equator. He showed, by means of a simple theoretical model, that hydrostatic equilibrium would be maintained if the equatorial axis of the earth were longer than the polar axis. This is equivalent to the statement that the body is flattened towards the poles.

**The ellipsoid of rotation or spheroid**

The three-dimensional body which corresponds is called an *ellipsoid of rotation*, which may be represented in section by means of an ellipse, as shown in Fig. 1.01 and elsewhere. The amount of polar flattening may be expressed by

\[
f = \frac{(a - b)}{a}
\]

(1.01)

where \(a\) and \(b\) are the lengths of the major and minor semi-axes of the ellipse. The value of \(f\), which is also known as the *ellipticity or compression* of the body, is always expressed as a fraction. For the earth this value is close to 1/298. We now know that the difference in length between the two semi-axes is approximately 11.5 km, or the polar axis is about 23 km shorter than the equatorial axis. It is interesting to reflect that this difference is about the same order of magnitude as the total relief variation on the earth. Thus at the approximate scale of 1/100 000 000 which represents the earth by a circle of radius 6 cm, the amount of polar flattening is also about 0.2 mm. Since 0.2 mm is also the width or gauge of line used for fine linear detail on maps, it follows that at very small scales the ellipticity of the earth is about the width of the line used to draw the elliptical section, and is therefore negligible. This is an important conclusion from the cartographic viewpoint because it permits the assumption that the earth can be regarded as truly spherical for certain purposes. We examine the validity of this assumption elsewhere (pp. 20–26). However, we must also note that any attempt to represent the terrestrial ellipsoid diagrammatically by a recognisable ellipse must involve considerable exaggeration. This, in turn, leads to possible misinterpretation of some of the illustrations depicting the geometry of the ellipsoid.

Since the ellipsoid of rotation approximates so closely to the sphere it may be called a *spheroid*. Since the flattening occurs at the poles rather than the equator, the figure may be further defined as an *oblate spheroid*. 
In the literature of surveying and cartography no real distinction can be made between the use of the two words 'ellipsoid' and 'spheroid'. Both are used indiscriminately.

**Measurement of the earth's figure**

Eight kinds of evidence have been used to determine the shape and size of the earth. These are:

- measurement of astro-geodetic arcs on the earth’s surface,
- measurement of variations in gravity at the earth’s surface,
- measurement of small perturbations of the moon’s orbit,
- measurement of the motion of the earth’s axis of rotation relative to the stars,
- measurements of the earth’s gravity field from the orbits of artificial satellites,
- measurement of very long astro-geodetic arcs derived from worldwide triangulation networks,
Certain of the methods are only of value in determining the parameter \( f \). The purely astronomical methods, which are the third and fourth in this list, are now only of historical interest. By far the most important modern method of determination is that of radar altimetry, which has been used since 1973.

**Astro-geodetic arc measurement**

This is the classic method which has been used to measure both the size and shape of the earth. It is based upon comparison of the **angular distance** between two points on the earth's surface and the **linear distance** between them. The first may be determined by making astronomical observations at the two places; the second by using the precise methods of surveying referred to as **geodetic** or **first order survey**. The **radii of curvature** of the earth may be determined from these data and finally the lengths of the semi-axes of the ellipsoid can be calculated.

If the earth were a true sphere its radius would be easily calculated, for it is a fundamental property of a sphere that all points on the surface are equidistant from its centre, i.e. it has constant radius. This is why it is possible to illustrate any section passing through the centre of a sphere by means of a circle as in Fig. 1.02. If there are two points, \( A \) and \( B \), on the surface of the sphere with centre \( O \), the angular distance between the points is the angle \( \angle AOB \) measured at the centre and the arc distance

![Figure 1.02](image)

**Fig. 1.02** A sphere in section, illustrating the relationship between angular distance and arc distance for all parts of the surface. \( AT \) represents a tangent to the circumference at \( A \).
between them is the shorter part of the circumference passing through the points. The relationship between these two measurements can be determined from

\[ \text{arc length } AB = R \cdot z \]  

(1.02)

where \( R \) is the radius of the sphere and \( z \) is the angle \( AOB \) expressed in radians. For example, if \( z = 10^\circ = 0.17453 \) radians and \( R = 6371 \) km, the arc distance \( AB = 1111.9 \) km. This is constant for all values of \( z = 10^\circ \) on this sphere irrespective of where the arc is situated. The converse argument is used to derive the radius from the measured length of the arc and an angular measurement. Thus, if astronomical observations made at both \( A \) and \( B \) showed that they lie \( 10^\circ \) apart and survey has established that the distance between them on the surface is 1111.9 km from equation (1.02)

\[ R = \frac{1111.9}{0.17453} \]

\[ = 6371 \text{ km} \]

The radius of the sphere has been defined as the line \( OA \). A further property of the sphere, which may be proved from the elementary plane geometry of the circle, is that when a tangent meets a circle at the point \( A \), the normal or perpendicular to that tangent passes through the centre of the circle. Thus on the sphere, \( OA \) is perpendicular to any tangent at \( A \) and if a series of tangents are drawn through \( A \) in any other directions than the section illustrated, these all lie in the same tangent plane.

This is important in defining the radii of curvature of an ellipsoid which are lines perpendicular to the tangent plane at any point on the curved surface. They are not represented by straight lines joining points on the surface to the geometrical centre of the body. Thus at some point \( A \) on the surface of the ellipsoid, we may imagine the tangent plane. In Fig. 1.03 the normal to this tangent plane is \( AQ'Q \). A further difficulty in defining the geometry of the ellipsoid is that two separate radii may be distinguished. One of these is the radius of the arc \( NAE \); the other is the radius of the arc which is perpendicular to \( NAE \) at \( A \). The radii are represented in Fig. 1.03 by the lines \( AQ' \) and \( AQ \) respectively. Thus both arcs occupy the same position in space but have different lengths. Moreover the line \( AQ'Q \) does not pass through the geometrical centre of the ellipse, \( O \), except where the normal to the surface forms either \( NO \) or \( EUN \), which are the semi-axes of the figure. It follows that the radii of an ellipsoid are variable quantities. Two separate radii may be defined for each point on the surface and both of these vary with position of the point. It follows, therefore, that the linear distance corresponding to a given angular distance varies with latitude. For example, the angle \( z = 10^\circ \) between the points \( A \) and \( B \) near the equator represents an arc distance
of approximately 1105.6 km on the surface of the terrestrial ellipsoid, whereas the same angle between the points $A'$ and $B'$ near the poles corresponds to about 1169.9 km. In other words the arc distance corresponding to a given angle increases polewards. This relationship is shown on Fig. 1.04, but care must be taken in the interpretation of the diagram. The ellipse is shown with exaggerated compression and the directions of the radii of curvature are shown as the normals to the tangents at the four points. These must be produced to give the points of intersection at $K$ and $M'$ to show that $AKB = 10^\circ = A'M'B'$. The reader should avoid making the implied comparison with Fig. 1.02, which suggests that the radii of the ellipse are the lines $AK$, $BK$ etc., and hence the fallacious interpretation of them as being much greater or less than $OA$ or $OB$ in Fig. 1.02.

This preliminary excursion into the geometrical properties of the sphere and ellipsoid, which are examined in greater detail in Chapter 3, has been made to indicate the kind of evidence to be obtained from astro-geodetic arc measurement. The variation in arc length with latitude was one of the first important pieces of evidence to be obtained which supported Newton's theoretical gravitational model. It was obtained from the measurement of two arcs, in Peru and Lapland, by the French during the early part of the eighteenth century.

The period of greatest activity in this field of geodesy occurred during the nineteenth and early twentieth centuries. Figure 1.05 illustrates those
The sphere has been supported as one of the poles, with the geodetic system of ellipsoids, and the geodetic system of ellipsoids is shown in the diagram. The poles are the points where the ellipsoids touch the sphere.

The small differences in the size and ellipticity of the earth's figure are shown in Table 1.01. The differences in the size and ellipticity of the earth's figure are shown in the figure of the earth. The differences in the size and ellipticity of the earth's figure are shown in the figure of the earth. The differences in the size and ellipticity of the earth's figure are shown in the figure of the earth. The differences in the size and ellipticity of the earth's figure are shown in the figure of the earth.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.

The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation. The preferred method of measurement was by means of triangulation.
Fig. 1.05 World map showing the location of the main astro-geodetic arcs used for determination of the figure of the earth during the classical period of arc measurement before 1914. The map is based upon the Aitoff–Wagner projection (No. 39a in Appendix I) which is a member of the polyconic group of projections with equidistant spacing of the parallels along each meridian. The geographical poles are represented by curves.
whether there is any justification for regarding these as deserving separate recognition. The figures have been determined by modern methods of using tracking of artificial satellites and by direct measurement of the height of the sea surface using radar altimetry. Associated with these developments also came the methods of position fixing by measuring doppler frequency changes and therefore distances between craft and groups of artificial satellites. The systems known as the Navy Navigation System (Transit or Navstar) and the Global Positioning System (GPS) both require the motions of such a satellite (or its ephemeris) to be referred to a specific Figure of the Earth. Conversely, the correct figure must be used with a particular navigation system to achieve the expected accuracy.

**Gravity measurements**

Newton arrived at the conclusion that the earth was an ellipsoid from the theoretical consideration of the forces created by the earth's mass and rotation (see page 64). Consequently the second important line of evidence concerning the shape of the earth has been from the study of variation in gravity.

In the absolute sense gravity varies with latitude; and it was early recognised that pendulum clocks which kept good time in Europe tended to lose time near the equator.

Gravity also affects the observations made during astro-geodetic arc measurement. It is this relative aspect of gravity which is particularly important in geodesy. In order to make observations in survey and astronomy it is necessary to align the instruments to a common datum. This datum is provided by the tangent plane to the earth's curved surface at the point of observation. This plane is geometrically important and also has a physical significance because it is defined by the spirit bubble mounted on a theodolite which is adjusted by means of its footscrews until the bubble is stationary in the centre of its run. The normal to this tangent plane is defined by the plumb-line which is used to set the instrument precisely over the point from which the observations are to be made. In short, we use gravity to determine both the horizontal plane of reference and the direction of the vertical. These adjustments are normal survey practice and are especially important in geodetic measurements. Supposedly horizontal angles observed by a theodolite which is not level contain errors which consequently deform the shapes of the triangles which have been observed. This, in turn, leads to errors in the computed distances between points and therefore to error in the computed positions of the stations. Precise determination of the horizontal plane of reference is an even more vital requirement in field astronomy because position is determined from measurements of vertical angles (or the altitudes) of stars. The datum for these measurements is the horizontal...
<table>
<thead>
<tr>
<th>Date</th>
<th>Name</th>
<th>Lengths of semi-axes (m)</th>
<th>Ellipticity (f)</th>
<th>Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Major (a)</td>
<td>Minor (b)</td>
<td></td>
</tr>
<tr>
<td>1687</td>
<td>Newton</td>
<td>—</td>
<td>—</td>
<td>1/231</td>
</tr>
<tr>
<td>1738</td>
<td>Maupertuis</td>
<td>6 397 300</td>
<td>—</td>
<td>1/191</td>
</tr>
<tr>
<td>1749</td>
<td>Bouguer</td>
<td>—</td>
<td>—</td>
<td>1/266</td>
</tr>
<tr>
<td>1799</td>
<td>Commission des Poids et Mésures</td>
<td>6 375 738.7</td>
<td>—</td>
<td>1/334.29</td>
</tr>
<tr>
<td>1810</td>
<td>Delambre</td>
<td>6 376 428</td>
<td>6 355 863</td>
<td>1/311.5</td>
</tr>
<tr>
<td>1817</td>
<td>Plessis</td>
<td>6 376 523</td>
<td>6 356 075.413</td>
<td>1/308.65</td>
</tr>
<tr>
<td>1830</td>
<td>Everest</td>
<td>6 377 276.345</td>
<td>6 356 078.963</td>
<td>1/300.8017</td>
</tr>
<tr>
<td>1841</td>
<td>Bessel</td>
<td>6 377 397.155</td>
<td>6 356 078.963</td>
<td>1/299.1528</td>
</tr>
<tr>
<td>1849</td>
<td>Airy</td>
<td>6 377 563.396</td>
<td>6 356 256.9</td>
<td>1/299.3249</td>
</tr>
<tr>
<td>1858</td>
<td>Clarke</td>
<td>6 378 294</td>
<td>6 356 618</td>
<td>1/294.261</td>
</tr>
<tr>
<td>1866</td>
<td>Clarke</td>
<td>6 378 206.4</td>
<td>6 356 762</td>
<td>1/298.97866982</td>
</tr>
<tr>
<td>1876</td>
<td>Andrae</td>
<td>6 377 104.43</td>
<td>6 356 514.9</td>
<td>1/300.0</td>
</tr>
<tr>
<td>1880</td>
<td>Clarke</td>
<td>6 378 249.17</td>
<td>6 356 818</td>
<td>1/293.465</td>
</tr>
<tr>
<td>1907</td>
<td>Helmert</td>
<td>6 378 200</td>
<td>6 356 911.946</td>
<td>1/298.30</td>
</tr>
<tr>
<td>1909</td>
<td>Hayford = International (1924)</td>
<td>6 378 388</td>
<td>6 356 911.946</td>
<td>1/297.0</td>
</tr>
<tr>
<td>Year</td>
<td>Datum/Reference System</td>
<td>Ellipsoid (Major Axis, flattening)</td>
<td>Conformal (Major Axis, flattening)</td>
<td>Note</td>
</tr>
<tr>
<td>------</td>
<td>------------------------</td>
<td>-----------------------------------</td>
<td>-----------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1876</td>
<td>Andræ</td>
<td>6377104.43 6356762.00</td>
<td>1/300.00</td>
<td>Denmark, Scania†</td>
</tr>
<tr>
<td>1880</td>
<td>Clarke</td>
<td>6378249.17 6356514.99</td>
<td>1/293.465</td>
<td>France, most of the African continent, Egypt†</td>
</tr>
<tr>
<td>1907</td>
<td>Helmert</td>
<td>6378200 6356818.00</td>
<td>1/298.30</td>
<td>Whole world excluding North America, Africa and a few other small areas</td>
</tr>
<tr>
<td>1909</td>
<td>Hayford = International (1924)</td>
<td>6378388 6356911.946</td>
<td>1/297.00</td>
<td></td>
</tr>
<tr>
<td>1940</td>
<td>Krasovsky</td>
<td>6378245 6356863.019</td>
<td>1/298.3</td>
<td>USSR and all other communist countries</td>
</tr>
<tr>
<td>1956</td>
<td>Army Map Service</td>
<td>6378270.0 6356794.343</td>
<td>1/297.0</td>
<td></td>
</tr>
<tr>
<td>1960</td>
<td>Fischer (Mercury Datum)</td>
<td>6378166.0 6356784.284</td>
<td>1/298.3</td>
<td></td>
</tr>
<tr>
<td>1960</td>
<td>World Geodetic System (WGS60)</td>
<td>6378165 6356783.00</td>
<td>1/298.3</td>
<td>Australia†</td>
</tr>
<tr>
<td>1961</td>
<td>Kaula</td>
<td>6378163</td>
<td>1/298.24</td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>International Astronomical Union (IAU65)</td>
<td>6378160 6356774.719</td>
<td>1/298.25</td>
<td>Australia</td>
</tr>
<tr>
<td>1965</td>
<td>Applied Physics Laboratory (APL 4.5)</td>
<td>6378137</td>
<td>1/298.25</td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>Naval Weapons Laboratory (NWLL9D)</td>
<td>6378145 6356759.8</td>
<td>1/298.25</td>
<td></td>
</tr>
<tr>
<td>1966</td>
<td>World Geodetic System (WGS66)</td>
<td>6378145 6356760.00</td>
<td>1/298.25</td>
<td></td>
</tr>
<tr>
<td>1967</td>
<td>Geodetic Reference System (GRS67) or International Astronomical Union (IAU68)</td>
<td>6378160.00 6356774.161</td>
<td>1/298.2472</td>
<td></td>
</tr>
<tr>
<td>1968</td>
<td>Fischer (Modified Mercury Datum)</td>
<td>6378150 6356768.955</td>
<td>1/298.3</td>
<td></td>
</tr>
<tr>
<td>1972</td>
<td>World Geodetic System (WGS72)</td>
<td>6378135 6356750.52</td>
<td>1/298.26</td>
<td>North America (NAD 83)</td>
</tr>
<tr>
<td>1979</td>
<td>Lerch * et al.*</td>
<td>6378139</td>
<td>1/298.257</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>Geodetic Reference System (GRS80) or New International</td>
<td>6378137.00 6356752.3141</td>
<td>1/298-257222101</td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>World Geodetic System</td>
<td>6378137.00</td>
<td>1/298.257223563</td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>Engelis</td>
<td>6378136.05</td>
<td>1/298.2566</td>
<td></td>
</tr>
</tbody>
</table>

The principal figures used for topographic mapping are tabulated in **Bold type**.

Obsolete figures formerly used for some mapping are denoted †.

There are numerous inconsistencies in use for different map series by certain national survey organisations. In military mapping other doctrines prevail. For example NATO policy for use with the UTM system (see pp. 357—360) favours use of the International Spheroid almost everywhere except North America, Africa and parts of Asia. Soviet policy has been to adopt the Krasovsky figure for all Warsaw Pact mapping.
plane indicated by the spirit bubble, or an artificial horizon formed by a liquid such as a dish of mercury which takes a horizontal position through gravitational attraction. The consequence of a slight inclination of either plane of reference leads to incorrect measurement of the vertical angle, and therefore to the determination of an incorrect astronomical position for the instrument.

**The geoid**

If the height of each observation station is reduced to sea-level, then by virtue of the fact that the instruments have always been carefully levelled, this is equivalent to stating that the observations have all been reduced to the same *equipotential surface* where the spirit bubble is always at rest. This surface is known as the geoid. It can be likened to the surface of an imaginary world ocean without land, waves, swell, tides or currents.

If the earth were such a homogeneous body, then from classical gravitational theory the surface of the geoid would coincide everywhere with the surface of an ellipsoid of rotation. However, this is not so. The geological history of the earth has led to irregular distribution of crustal rocks having different densities. The denser rocks exert their own attraction upon a spirit bubble, although this is small compared with the main gravitational component. Thus an instrument may appear to be level because the spirit bubble is at rest in the centre of its run, but the plumb-line is not normal to the spheroid for it is deflected slightly towards the areas of greater rock density. Since the amount of deflection varies from place to place it follows that the geoid has an undulating surface. Figure 1.06 illustrates how these undulations occur. Since all the observations have been made with reference to the geoid, additional measurements of the *gravity anomalies* which are present can be used to correct for and increase knowledge about the location of the undulations of its surface. Stokes first demonstrated these principles in 1849 and methods of correcting for anomalies have been used since 1855, when Pratt attempted to account for discrepancies in the position of Kalianpur observed in the astro-geodetic arc measured by the Great Trigonometrical Survey of India. The attempts to explain this and similar inconsistencies in other arc measurements led to the formulation of the different theories of *isostasy*, which have been a major preoccupation of geodesists and also revolutionised early theories about structural geology.

It follows that the increasing refinement of determination of the Figure of the Earth, characterised by the small variation in $f$, obtained after 1900, is largely owing to the increasing availability of gravity data and the methods of employing these to adjust the astro-geodetic observations. By the late 1940s sufficient information about gravity anomalies had been collected to attempt the compilation of maps showing the undulations of
The Figure of the Earth

The relationship between the geoid and reference spheroid, indicating the deflection of the perpendicular to the geoid and resulting undulations in the surface of the geoid.

The contribution of satellite geodesy

The first artificial satellite had been launched a month earlier. This heralded a major step forward in advancement of knowledge about the earth’s true shape and size, and moreover removed the dependence upon the slow acquisition of terrestrial measurements. The principal reason for this advance was that artificial satellites overcame a fundamental difficulty in deciphering the earth’s gravity field, namely that because terrestrial measurements were confined within it, this made it impossible to make any external measurement of the forces. This difficulty had long been
realised; indeed, attempts had been made to employ the moon, as our natural satellite. However, the attempted measurements were somewhat insensitive because of the distance between the earth and the moon.

**The evidence of satellite tracking**

The earliest, and some of the most significant, information was obtained within a year or two, simply from observation of the changing orbits of the early Sputnik and Vanguard satellites. Satellite tracking has yielded much information about the gravity potential of the earth, and led eventually to remarkably detailed mapping of the geoid throughout the world (Figs 1.10 and 1.11). The second use of satellites has been to provide survey beacons which have been located high enough above the earth’s surface to be simultaneously visible from places which are hundreds, or even thousands, of miles apart. Consequently these may be used to create unified and world-wide networks of geodetic stations (Fig. 1.07). This made it possible to compare astro-geodetic arcs for much greater distances on the earth’s surface than had ever been accomplished in classic geodesy.

If the earth were spherical, and of homogeneous density, the orbit of a satellite would be an ellipse fixed in shape and size, and with its plane in a fixed direction in space. Any departure of the earth from a spherical form causes changes in the gravitational forces acting upon the satellite, and therefore upon its orbit. The main effect of the earth’s ellipticity upon a satellite orbit is to make the plane of the orbit rotate about the earth’s axis in the direction opposite to the satellite’s motion, while leaving the inclination of the orbit to the equator virtually constant. This phenomenon is known as the **precession of the nodes** (Fig. 1.08). The rate of precession can be measured with extraordinarily high precision using quite simple equipment because the movement is regular and therefore it can be allowed to accumulate over long periods and therefore many orbits between observations. The value of ellipticity, obtained only a year or so after the first artificial satellite had been launched, was \( f = 1/298.24 \), or practically the same as that determined by Helmert in 1907 and Krasovsky in 1940.

Study of the variations in gravity potential with latitude has led to the evaluation of a series of numerical coefficients, called J-harmonics, which describe a sequence of increasingly elaborate geometrical figures. The \( J_2 \) coefficient, which defines the ellipticity of the spheroid, is by far the most important of these, but some of the other coefficients are not wholly insignificant. They indicate that the earth is somewhat asymmetrical in section, for the North Pole lies about 10 m further from the equator than can be accounted for by ellipticity of 1/298.24, but the South Pole lies about 30 m nearer the equator than this amount of compression suggests. The resulting meridional section (Fig. 1.09) has been likened to the shape
The Figure of the Earth

The shape of the Earth is somewhat

somewhat

on our earth.

noon.

noon.
Angular momentum of satellite about the polar axis remains constant

![Diagram](image)

**Fig. 1.08** Diagrammatic representation of the precession of the nodes. The equator-wards force, resulting from the earth's equatorial bulge, causes an artificial satellite to cross the equator on a different meridian at each successive orbit.

**Fig. 1.09** Inferred meridional section of the earth based upon the calculation of variations in gravity potential with latitude but excluding any variation with longitude. The diagram indicates the departure (in metres) of this section (full line) from an ellipse with compression 1/298.24 (broken line).
of a pear. However, despite much publicity of this conclusion in the early
days of satellite tracking, not too much importance should be placed on
it, for the pear shape is an average value of the undulations of the geoid
determined with reference to latitude and ignoring any variations in
longitude. More important to modern concepts of geodesy were the
attempts to produce a contour map of the height of the geoid for the
whole world. A study first undertaken by Kaula in 1961 produced the
world map illustrated in the first edition of this book. Events have pro-
ceeded so quickly that much more detailed geoid contours are now
available, as illustrated in Fig. 1.11.

It is also important to appreciate that in classical geodesy the arc
measurements were self-contained and isolated from one another by
whole continents and oceans. Consequently the results of these arc
measurements were fitted to a comparatively small portion of the
spheroid, and it was impossible to relate the results precisely to the axis
of rotation and the true centre of the earth. Thus a particular Figure of
the Earth would not be referred to the true axis of rotation but to a
parallel axis which was displaced from the true axis by a small but
unknown amount, as illustrated in Fig. 1.10. For the creation of reliable
satellite navigation systems the ephemeris of each satellite has to be
known more precisely. This includes knowledge about the true position
of the earth's centre. Consequently there has been a revolution in the
concept of how the earth's figure should be defined, and a variety of new
figures have emerged from these data. Modern determinations of the

![Figure 1.10](image_url)

**Fig. 1.10** A comparison between the earth's figure based upon an equipotential
ellipsoid, having a geocentric origin to the \(X, Y, Z\) cartesian coordinate system,
and a figure derived from classical methods of geodesy in which the centre is
offset from its true position.
earth's figure from the time of GRS67 onwards are truly geocentric and based upon the theory of an *equipotential ellipsoid*. Consequently the modern trend is to describe new figures initially in geophysical terms and only later derive the various parameters to which we are accustomed. See, for example, the detailed description of the IUGG specification for GRS80 by Moritz (1980a).

**Global triangulation schemes**

A vital stage in satellite geodesy was therefore the accomplishment of various world-wide control surveys. The period of greatest activity in this field was in the late 1960s, during which time the whole task of providing a world-wide geodetic control network was accomplished. A variety of different techniques were employed by the different branches of the US administration involved in this renaissance of geodesy. One system favoured the use of large satellites, like the PAGEOS satellite which was a balloon that became inflated when in orbit, and therefore large enough to be simultaneously photographed against the background of stars by several BC-4 ballistic cameras. Because of the designation of the camera this project is now commonly referred to as the *BC-4 Triangulation*. The ANNA satellite contained a brilliant flashing light bright enough to be identifiable as a beacon in space. The third idea was to use electronic distance measurement to track a comparatively small reflecting satellite. This was exemplified by the SECOR system used to establish an equatorial control network round the world. Later still came the application of even more sophisticated methods of distance measurement using lasers and doppler, resulting in much greater accuracy in the methods of satellite tracking. Indeed the roles were reversed; for the positions of many modern satellites are now determined so accurately that distance measurements from clusters of them are now used to locate positions on the earth’s surface. This has been developed through the various satellite navigation systems to the Global Positioning System (GPS) which promises to offer the world-wide ability to fix position with an accuracy equivalent to conventional geodetic surveys.

**Satellite altimetry**

Satellite altimeters directly measure the distance between a satellite and the instantaneous sea surface. By accurately determining the satellite orbit with respect to positions on the earth’s surface it is possible to estimate the height of the sea surface above the reference ellipsoid. Therefore the construction of contours for the surface of the geoid can be used to estimate the deflection of the vertical at sea. The first experiments in radar altimetry in space were made from SKYLAB, launched in
November 1973. Two later satellites have so far been equipped with radar altimeters, first GEOS-3 and secondly SEASAT.

We have already likened the equipotential surface of the geoid to that of an imaginary planetary ocean. The question which naturally arises is whether the actual surface of the ocean is anything like the idealised surface, and what corrections can be applied to the natural disturbances caused by ocean currents, tides and other surface displacements in order to describe the geoid.

The GEOS-3 mission was designed to improve knowledge of the earth's gravitational field, the size and shape of the terrestrial geoid, deep ocean tides, sea state, current structure, crustal structure, solid earth dynamics, and remote sensing technology. The GEOS-3 altimeter was designed to provide the means for establishing the feasibility for directly measuring some of these variables. In every respect the altimeter far exceeded its expectations. For example, although the system was designed for a 1-year lifetime, the satellite was still operational after more than 3½ years in orbit. In addition, the altimeter showed that it was capable of providing valid measurements over land and ice. Neither of these capabilities had been predicted prior to launch.

The second reason for the success of altimetric measurements is the speed with which the information may be collected by satellite compared with conventional marine gravity measurements. A research ship on a cruise to the Antarctic might be away for 6 months, but only a small proportion of that time will be spent making observations in the intended working area. By contrast a satellite will not only make the journey 14 times in one day, storing its results and transmitting them to a convenient ground station, but will also sense all the other oceans several times in the same day.

The principal limitation in the use of GEOS-3 altimetry was the restricted cover of the world's oceans which could be sampled. These data were largely confined to the North Atlantic, the Gulf of Mexico, North Pacific and the Bering Sea. The restricted cover was owing to the small number and the location of ground stations capable of receiving signals from the satellite.

An important method of analysis of the altimeter records is the study of those crossover points where the height of a point on the sea surface has been measured when the satellite has occupied different orbits (Marsh et al., 1982a,b) allowing the precision of the surfaces to approach that of the measurements themselves (25 cm for GEOS-3, better than 10 cm for SEASAT). Analysis of the sea height residuals at the crossing points of the satellite arcs provides information about the long-term variability of sea height in these regions.

A more sensitive radar altimeter was fitted aboard SEASAT, which was launched on 27 June 1978; the network of receiving stations had also
been much extended. The satellite operated successfully until 10 October 1978, when a power failure brought transmission to a stop. A mission overview has been given by Lame and Born (1982), who have shown that, despite its short lifetime, SEASAT acquired a wealth of data on sea-surface winds and temperature, ocean wave heights, internal waves, atmospheric water content, sea ice, topography of the ocean surface and shape of the marine geoid. Analysis of the output from the radar altimeter was one of the most important aspects because most of the world's oceans were sampled; therefore better estimations were obtained for the slope of the marine geoid for the world as a whole.

Concurrent with these developments, various attempts were made to produce increasingly more sophisticated models of the geoid. These have been conventionally named after the American laboratories which have undertaken the study; notably the Smithsonian model earths, labelled SAO, after the Smithsonian Astrophysical Observatory and the GEM, or Goddard Earth Models, after the Goddard Space Center operated by NASA.

The choice of a suitable reference surface for mapping

Because we now know that the geoid is a complicated body, we must enquire how it should be described mathematically for the practical purposes of mapping. Since there is no merit to be gained from increasing the mathematical complexity of a solution beyond defining those irregularities which have practical significance, it is desirable to consider the possibility of using various reference surfaces which describe the shape and size of the earth adequately for different purposes. The variations illustrated by the contour pattern in Fig. 1.11 may amount to only a few metres but they are of considerable importance to the study of dynamic geodesy and some branches of geophysics. For work in these fields there are cogent reasons for defining as a reference surface a triaxial ellipsoid in which the observed undulations along the equator may also be fitted to an ellipse. However, these variations in the geoid are practically negligible for most other kinds of survey and in cartography.

Thus we may simplify the problem and consider three different ways in which we may define the shape and size of the earth for different purposes in surveying and mapping. These are:

1. a plane which is tangential to the earth at some point;
2. a perfect sphere of suitable radius;
3. an ellipsoid of rotation of suitable dimensions and ellipticity.

They are listed in ascending order of refinement. Thus a suitable ellipsoid fits the shape of the geoid better than does a perfect sphere of equivalent size. The sphere, in turn, is a better approximation of the curved surface
than is a plane. On the other hand, the list is in order of increasing mathematical difficulty. The formulae needed to define position, to determine the relationships between distances and angles on a plane are simpler than are those for the curved surface of a sphere. These, in turn, are simpler than the corresponding formulae for an ellipsoid. Bearing in mind the desirability of using the simplest reference figure which is compatible with accuracy of representation, it follows that we should inspect the properties of each kind of reference surface to discover when it should be used.

**The plane reference surface**

At first sight it may seem to be a retrograde step to assume that the earth is a plane. However, it is a very useful assumption because it is so simple to use. For a start we can avoid the whole problem of map projection transformations which are the preoccupation of this book. Figure 1.02 indicates that near a point $A$ on the curved surface of the earth, the tangent to the curved surface also lies close to it. The tangent plane and the curved surface only diverge from one another as one moves away from $A$. It may therefore be argued that if we only need to make a survey of a small area around $A$, it is reasonable to assume that we are making the measurements on the tangent plane. The survey can be computed by the methods of plane trigonometry (it is then called plane surveying). Plotting of the map can be done simply by converting ground dimensions to the required map scale. The crux of the argument is the definition of what is represented by the immediate vicinity around the point $A$. It implies that the plane assumption should be confined to the preparation of maps of small areas, but it still remains necessary to define what we mean by a small area. We defer quantitative consideration of this problem until Chapter 15 (pp. 310–335) because it is desirable to consider this assumption together with the kinds of map projections which are used by surveyors, and which are also important in large-scale cartography of small areas.

**The spherical assumption**

We have already commented upon the fact that, at a scale of 1/100 000 000, the lengths of the two axes of the spheroid differ by about the width of the lines needed to draw them. This implies that the main use of the spherical assumption will occur in the preparation and use of comparatively small format maps showing large parts of the earth's surface such as maps of the world, a hemisphere, a continent or even a large country, such as appear in atlases. The question to be answered is:
increasing to determine the simpler turn, are in mind compatible except the it should

the earth so simple projection figure 1.02 earth, the lane and away a survey making puted by trajectories).

the figure of the earth.

scale of by about the main use of the earth's or even a woman is:

'These are used for the sphere and spheroid. He used the Clarke 1866 Figure of the Earth, which

...What is the approximate maximum scale at which the spherical assumption can be justified?'

This subject was tackled theoretically by Driencourt in 1932, and his work has been reproduced more recently by Richardus and Adler (1972). Therefore we need not reproduce the detailed mathematical argument here. Driencourt showed that the largest errors occur in lines which are orientated east or west from a point, and that the maximum linear displacement, \( \Delta t \) is directed northwards or southwards. He calculated the following results for a line of length \( s \) (km). The following table shows that the discrepancy \( \Delta t \) at a distance of approximately 100 km from the central point, does not exceed 1 mm or \( 10^{-8} \). At a distance of 1000 km from the point the proportion \( \Delta t/s \) is approximately \( 10^{-5} \), which is about three times the present precision of electronic distance measurement.

Tobler (1964) also investigated the problem from the point of view of mapping the United States of America. He calculated the distances and bearings between 200 randomly selected places in the USA for both the sphere and spheroid. He used the Clarke 1866 Figure of the Earth, which

---

**Fig. 1.12** The geometry of the reference surfaces; a comparison between the spheroidal (a) and spherical (b) surfaces showing corresponding observations. A line of length \( s \) is measured from the point \( O \) along a bearing \( \theta \). On the spheroid this bearing traces the arc \( t \) at the distance \( s \) from \( O \); on the sphere the corresponding arc is \( t \). The difference \( \Delta t = t - t \) is a measure of the discrepancies which occur if the earth is assumed to be spherical. The amount varies with the size of the angle \( \theta \) and with the distance \( s \). (Source: Driencourt and Laborde, 1932.)
was that still in use for North America at the time. For the spherical assumption he chose as the radius \( R = 6378.206 \) km, which is the equatorial radius for the Clarke 1866 figure. The results are given in Table 1.03.

If we assume 0.2 mm to be the smallest linear distance which can be measured on a map without special magnification, and if we take Tobler’s average difference in distance as being equal to this, then the largest scale at which the USA might be represented by a projection of the sphere is 1/370,000. However, the spread of the results, characterised by the values for the standard deviation and the two extremes, indicates that it would be optimistic to use the spherical assumption at such a large scale and imagine that no errors in mapping would arise from this cause. The figures suggest that, strictly speaking, the spherical assumption ought to be confined to use for maps of scale 1/15,000,000 or smaller, which is about the scale at which 7.8 km is represented by 0.2 mm. In practical cartography, however, the limit of using the spherical assumption is usually taken to be a scale of 1/5,000,000 or thereabouts. Using Tobler’s data it can be argued that at this scale about two-thirds of the points lie within 1 mm of the spheroidal position if mapped on a sphere. We shall see later that this discrepancy is small compared with the displacements which are inherent in the process of representing a large country at a small scale on a plane map.

A third approach has been adopted by Snyder (1987a), who has applied the same distortion theory which we shall investigate in the study of plane map projections to the projection of the spheroid to the sphere. This gives rise to a series of values for particular scales and distortion characteristics.
which are introduced in Chapter 5. The numerical characteristics thus obtained may be used to determine the maximum scale at which the distortion cannot be recognised on a map.

The spheroidal assumption

Obviously the spheroid fits the shape of the geoid more closely than does a sphere. Consequently this is the reference surface which ought to be employed in surveying. This is because the survey of a country is first computed to determine the positions of the control points in their natural dimensions or, as it were, for a map of scale 1/1. Consequently the small discrepancies in position (or closing errors) may be expressed to the nearest millimetre or less on the ground and not absorbed by scale reduction as would happen if the results of a survey were first plotted on a sheet of paper. In order to appreciate the quality and precision of the work it is desirable to make these computations with respect to a particular reference spheroid rather than risking the introduction of errors arising from assuming a flat or spherical earth. At the later stage of producing topographical and other map series, extending throughout an entire country, continuity of information across boundaries of adjacent map sheets is important. Hence it is desirable to use the reference ellipsoid as the basis of such maps. It is also used for the compilation of large-scale navigation charts and small-scale charts to the approximate limit of 1/4000000–1/5000000.

Table 1.01, on page 10, indicated that about 15 different reference ellipsoids may be encountered in world mapping, and about six of them are in common use. From the point of view of practical cartographic work the correct spheroid for use should always be clearly stated in the mapping specification. From the point of view of evaluating existing topographical or other maps as source documents for compilation, references such as the United Nations’ summaries on the status of world topographic mapping (United Nations, 1970, 1976, 1979) and the national survey reports provide the information which is needed. In an analysis of the UN data Brandenberger and Gosh (1985) have estimated that nearly 93% of the earth’s land area has been mapped on only four of the classical figures. These are:

- International spheroid 28.3%
- Krasovsky spheroid 25%
- Bessel spheroid 19.9%
- Clarke 1880 spheroid 19.4%.

Originally a particular spheroid was selected by the national survey because the parameters of the figure fitted the observed data better than any other. A typical example of this was the use of the Airy spheroid for
Great Britain, for this had been derived from astro-geodetic distances obtained during the original Primary Triangulation of the country. In the days before digital computing, once a national survey had been computed using a particular reference figure it would have been extremely inconvenient and costly to convert the positions of many hundreds or even thousands of control points to another spheroid. It was done in the USSR when the decision was taken in 1942 to transform the entire control network from the Bessel spheroid to the newly described Krasovsky figure, but that was a practically unique example. It follows that usually a national survey continued to be based upon a particular figure long after the original reasons for its choice had ceased to be valid.

This argument carries less weight today than before digital computing became commonplace. It is interesting to note in this context that probably the first major use of digital computing in geodesy and surveying was the work undertaken by the US Army Map Service shortly after World War II, when they accomplished the formidable task of reducing the national surveys of western Europe to a common datum on the International Spheroid. This is known as the European Datum, 1950, or ED50. This network had hitherto been based upon a multiplicity of different points of origin, reference spheroids, units of measure and projections. We shall also refer to the change in the North American Datum from NAD 27 into NAD 83 during the 1980s, which amongst other changes includes that from the Clarke 1866 figure to GRS 80.

Nevertheless the use of different figures still remains. It arises partly from historical accident, partly from inertia and partly for reasons of national prestige. Sometimes it also happens that the chosen spheroid fits the shape of the geoid in that country better than any of the others.

Finally the continuity of use is important. Indeed Chovitz (1981) has argued that this continuity is at least as important as the formal accuracy of recording the length of the major semi-axis and flattening. Some of the better-known figures, such as Airy, Everest and the three useful Clarke determinations, have been slightly modified on many occasions for use in different places or for different purposes. Typical examples include retaining the original value for the semi-axis, a, but using it with a slightly different (rounded) value for f. Other changes have been enforced by the discrepancies introduced to the dimensions of the semi-axes through converting from British Standard into metric units or vice-versa. For example, Strasser (1975) has shown how US legislation concerning the definition of the metre has created numerous difficulties in reconciling different versions of the Clarke 1866 figure. Sometimes we know enough about the history of a survey to understand where discrepancies have arisen. More often it may be extremely difficult to reconcile these so that mistakes are sometimes made in choosing the correct version of Everest or Airy.