A GUIDE TO THE MATHEMATICS OF MAP PROJECTIONS

R.E. Deakin
School of Mathematical and Geospatial Sciences
RMIT University, GPO Box 2476V, MELBOURNE VIC 3001
email: rod.deakin@rmit.edu.au

Presented at the Victorian Tasmanian Survey Conference *Across the Strait,* Launceston Tasmania
April 15-17, 2004

ABSTRACT

In Australia, large-scale topographic mapping and survey coordination is based on rectangular grids overlaying conformal map projections; e.g., the Australian Map Grid (AMG) and Map Grid Australia (MGA) used Australia wide and the Integrated Survey Grid (ISG) used in New South Wales overlay *Transverse Mercator* projections and VICGRID, sometimes used in Victoria, overlays a *Lambert Conformal Conic* projection. As spatial data experts, surveyors require a sound understanding of projections, grids and associated formulae; this paper provides a brief history of geodesy and the shape of the Earth, information on geodetic datums, some theory of projections and a detailed development of the formulae for the Transverse Mercator projection of the ellipsoid that should enhance the practical knowledge of surveyors.

INTRODUCTION

In Australia, large-scale topographic mapping and survey coordination is based on rectangular grids overlaying conformal map projections; e.g., the Australian Map Grid (AMG) and Map Grid Australia (MGA) used Australia wide, the Integrated Survey Grid (ISG) used in New South Wales and VICGRID sometimes used in Victoria. The AMG and MGA are grids superimposed over *Universal Transverse Mercator* (UTM) projections, the ISG overlays a *Transverse Mercator* (TM) projection and VICGRID overlays a *Lambert Conformal Conic* projection with two standard parallels. Other projections and grids have been or are being used for mapping and coordination in Australia but their use has either been superseded, or is local in extent. Only the TM and UTM projections are considered in this paper.

A sound knowledge of map projections and grids requires an understanding of the mathematical nature of projections and the size and shape of the Earth, since in our context; a projection is a mathematical transformation of coordinates on a reference surface approximating the Earth to coordinates on a projection plane. The reference surface is an ellipsoid (a surface of revolution created by rotating an ellipse about its minor axis) representing the "mathematical" figure of the Earth. This paper gives a brief history of the determination of the size and shape of the Earth, geometry and formulae of ellipsoids, information on geodetic datums and coordinate systems in use in Australia and an outline of the mathematical theory of map projections.

The main body of the paper is a detailed derivation of the formulae for a TM projection of the ellipsoid giving X, Y coordinates, grid convergence and point scale factor. In Australia, these equations are commonly referred to as *Redfearn's* formulae, published by J.C.B Redfearn of the Hydrographic Department of the British Admiralty in the *Empire Survey Review* (now *Survey Review*) in 1948 (Redfearn 1948). Redfearn noted in his five-page paper that: "…formulae of the projection itself have been given by various writers, from Gauss, Schreiber and Jordan to Hristow, Tardi, Lee Hotine and others – not, it is to be regretted, with complete agreement in all cases." Redfearn's formulae, accurate anywhere within zones of 10°–12° extent in longitude, removed this "disagreement" between previous published formulae and are regarded as the definitive TM formulae. Redfearn provided no method of derivation but mentioned techniques demonstrated by Lee and Hotine in previous issues of the Empire Survey Review. In 1952, the American mathematician Paul D. Thomas published a detailed derivation of the TM formulae in *Conformal Projections in Geodesy and Cartography,* Special Publication No. 251 of the Coast and Geodetic Survey, U.S. Department of Commerce (Thomas 1952); Thomas' work can be regarded as the definitive derivation of the TM formulae.

Surveying and geodesy textbooks, with the notable exception of G.B. Lauf's *Geodesy and Map Projections* (Lauf 1983), often have only an "outline" of the mathematics of the TM projection and a statement of formula – if there is any mention of the projection at all. Consequently, the mathematics of the TM projection is not well known; not that it was ever a "hot" topic of conversation amongst surveyors, or indeed students, who see it as masses of calculus saddled with turgid algebra. Help is at hand though: in the form of mathematical computer packages such as MAPLE® that relieve the interested student of the drudgery (some find it a beauty) of mathematical manipulation. In this paper, the method of derivation follows that of Thomas (1952) and Lauf (1983) but all the TM formulae (coordinates, grid convergence and point scale factor) were obtained using MAPLE; reducing the work from pages of algebra to, in some cases, half a dozen computer commands.

Whilst this paper does not provide any instructions or commands specific to mathematical computer packages in the derivation of formulae, it is hoped that it will be of some use to those who wish to demonstrate (to students) the power of mathematical computer packages versus the traditional methods of solution. In addition, it is hoped that this paper will supplement the excellent technical manuals available to practitioners in Australia: *The Australian Geodetic Datum Technical Manual,* Special Publication 10 (NMC 1986), *Geocentric Datum of Australia Technical Manual* – Version 2.2 (ICSM 2002) and *The Map Grid of Australia 1994* – A Simplified Computation Manual (Land Vic 2003). The latter two publications are available online via the Internet with links to additional information sources and Microsoft® *Excel* spreadsheets for computations.

A BRIEF HISTORY OF GEODESY AND THE ELLIPSOIDAL SHAPE OF THE EARTH

Geodesy is the scientific study of the size and shape of the Earth. Since ancient times, philosophers and scientists have attempted to determine its shape and size. Ancient methods ranged from comparisons with other heavenly bodies (Pythagoras, 6th century BC) to the measurement of the incident angles of rays of sunlight at selected points on the Earth's surface (Eratosthenes 3rd century BC). Later techniques involved astronomical observations and measured lengths of meridian arcs, e.g. the French Academy of Sciences expeditions in the 1700's verifying Newton's theoretical deduction of an ellipsoidal Earth based on his Universal Law of

Gravitation. The latest methods rely on gravimetric observations and satellite observations. The following is an edited extract from the Encyclopaedia Britannica (Britannica® CD 99) outlining some of the historical determinations of the Earth's shape.

Spherical era

Credit for the idea that the Earth is spherical is usually given to Pythagoras (flourished 6th century BC) and his school, who reasoned that, because the Moon and the Sun are spherical, the Earth is too. Notable among other Greek philosophers, Hipparchus (2nd century BC) and Aristotle (4th century BC) came to the same conclusion. Aristotle devoted a part of his book *De caelo* (*On the Heavens*) to the defence of the doctrine. He also estimated the circumference of the Earth at about 400,000 stadia. Since the Greek stadium varied in length locally from 154 to 215 metres, the accuracy of his estimate cannot be established. This seems to be the first scientific attempt to estimate the size of the Earth. Eratosthenes (3rd century BC), however, is considered to be one of the founders of geodesy because he was the first to describe and apply a scientific measuring technique for determining the size of the Earth. He used a simple principle of estimating the size of a great circle passing through the North and South poles. Knowing the length of an arc and the size of the corresponding central angle that it subtends, one can obtain the radius of the sphere from the simple proportion that length of arc to size of the great circle (or circumference, $2\pi R$, in which R is the Earth's radius) equals central angle to the angle subtended by the whole circumference (360°). In order to determine the central angle α, Eratosthenes selected the city of Syene (modern Aswan on the Nile) because there the Sun in midsummer shone at noon vertically into a well. He assumed that all sunrays reaching the Earth were parallel to one another, and he observed that the sunrays at Alexandria at the same time (midsummer at noontime) were not vertical but lay at an angle 1/50 of a complete revolution of the Earth away from the zenith. Probably using data obtained by surveyors (official pacers), he estimated the distance between Alexandria and Syene to be 5,000 stadia. From the above equation Eratosthenes obtained, for the length of a great circle, $50 \times 5{,}000 = 250{,}000$ stadia, which, using a plausible contemporary value for the stadium (185 metres), is 46,250,000 metres. The result is about 15 percent too large in comparison to modern measurements, but his result was extremely good considering the assumptions and the equipment with which the observations were made.

Ellipsoidal era

The period from Eratosthenes to Picard (the French scientist who, in the late 1600's, measured a short meridian arc by triangulation in the vicinity of Paris) can be called the spherical era of geodesy. Newton and the Dutch mathematician and scientist Christiaan Huygens began a new ellipsoidal era. In Ptolemaic astronomy it had seemed natural to assume that the Earth was an exact sphere with a centre that, in turn, all too easily became regarded as the centre of the entire universe. However, with growing conviction that the Copernican system is true – the Earth moves around the Sun and rotates about its own axis – and with the advance in mechanical knowledge due chiefly to Newton and Huygens, it seemed natural to conceive of the Earth as an oblate spheroid. In one of the many brilliant analyses in his *Principia,* published in 1687, Newton deduced the Earth's shape theoretically and found that the equatorial semi-axis would be 1/230 longer than the polar semi-axis (true value about 1/300). Experimental evidence supporting this idea emerged in 1672 as the result of a French expedition to Guiana. The members of the expedition found that a pendulum clock that kept accurate time in Paris lost 2 1/2 minutes a day at Cayenne near the Equator. At that time no one knew how to interpret the observation, but Newton's theory that gravity must be stronger at the poles (because of closer proximity to the Earth's centre) than at the Equator was a logical explanation. It is possible to determine whether or not the Earth is an oblate spheroid by measuring the length of an arc corresponding to a geodetic latitude difference at two places along the meridian (the ellipse passing through the poles) at different latitudes. The French astronomer Gian Domenico Cassini and his son Jacques Cassini made such measurements of arc in France by continuing the arc of Picard north to Dunkirk and south to the boundary of Spain. Surprisingly, the result of that experiment (published in 1720) showed the length of a meridian degree north of Paris to be 111,017 metres, or 265 metres shorter than one south of Paris (111,282 metres). This suggested that the Earth is a prolate spheroid, not flattened at the poles but elongated, with the equatorial axis shorter than the polar axis. This was completely at odds with Newton's conclusions. In order to settle the controversy caused by Newton's theoretical derivations and the measurements of Cassini, the French Academy of Sciences sent two expeditions, one to Peru led by Pierre Bouguer and Charles-Marie de La Condamine to measure the length of a meridian degree in 1735 and another to Lapland in 1736 under Pierre-Louis Moreau de Maupertuis to make similar measurements. Both parties determined the length of the arcs using the method of triangulation. Only one baseline, 14.3 kilometres long, was measured in Lapland, and two baselines, 12.2 and 10.3 kilometres long, were used in Peru. Astronomic observations for latitude determinations from which the size of the angles was computed were made using instruments with zenith sectors having radii up to four metres. The expedition to Lapland returned in 1737, and Maupertuis reported that the length of one degree of the meridian in Lapland was 57,437.9 toises. (The toise was an old unit of length equal to 1.949 metres.) This result, when compared to the corresponding value of 57,060 toises near Paris, proved that the Earth was flattened at the poles. Later, large errors were found in the measurements, but they were in the "right direction." After the expedition returned from Peru in 1743, Bouguer and La Condamine could not agree on one common interpretation of the observations, mainly because of the use of two baselines and the lack of suitable computing techniques. The mean values of the two lengths calculated by them gave the length of the degree as 56,753 toises, which confirmed the earlier finding of the flattening of the Earth. As a combined result of both expeditions, these values have been reported in the literature: semi-major axis, a = 6,397,300 metres, flattening, f = 1/216.8. Almost simultaneously with the observations in South America, the French mathematical physicist Alexis-Claude Clairaut deduced the relationship between the variation in gravity between the Equator and the poles and the flattening. Clairaut's ideal Earth contained no lateral variations in density and was covered by an ocean, so that the external shape was an equipotential of its own attraction and rotational acceleration. Clairaut's result, accurate only to the first order in f, clearly showed the relationship between the variations of gravity at sea level and the flattening. Later workers, particularly Friedrich R. Helmert of Germany, extended the expression to include higher order terms, and gravimetric methods of determining *f* continued to be used, along with arc methods, up to the time when Earth-orbiting satellites were employed to make precise measurements. Numerous arc measurements were subsequently made, one of which was the historic French measurement used for definition of a unit of length. In 1791 the French National Assembly adopted the new length unit, called the metre and defined as 1:10,000,000 part of the meridian quadrant from the Equator to the pole along the meridian that runs through Paris. For this purpose a new and more accurate arc measurement was carried out between Dunkirk and Barcelona in 1792-98 by Delambre and Méchain. These measurements combined with those from the Peruvian expedition yielded a value of 6,376,428 metres for the semi-major axis and 1/311.5 for the flattening, which made the metre 0.02 percent "too short" from the intended definition. The length of the semi-major axis, a, and flattening, f, continued to be determined by the arc method but with modification for the next 200 years. Gradually instruments and methods improved, and the results became more accurate. Interpretation was made easier through introduction of the statistical method of least squares.

For those with an interest in the history of geodesy the book *The Measure of All Things* (Alder 2002) has an interesting account of the determination of the size of the Earth and the definition of the metre by the French scientists Delambre and Méchain in the 1790's. A concise treatment of the history of geodesy, with a technical flavour, can be found in G.B. Lauf's book *Geodesy and Map Projections* (Lauf 1983).

GEOMETRY OF ELLIPSOIDAL REFERENCE SURFACE OF THE EARTH

An ellipsoid, a surface of revolution created by rotating an ellipse about its minor axis, is regarded as the simplest mathematical surface that is the closest approximation to the actual size and shape of the Earth. The size and shape of an ellipsoid can be defined by specifying pairs of geometric constants; (i) semi-major axis a and flattening f, or (ii) semi-major axis a and eccentricity squared e^2 or (iii) semi-major axis a and semi-minor axis b. Figure 1 shows an ellipsoid with major and minor axes $2a$ and $2b$ respectively and the following relationships will be useful.

Ellipsoid Relationships and Formulae

Referring to Figure 1:

(i) O is the centre of the ellipsoid, $OEMG$ is the equatorial plane of the ellipsoid (the reference plane for latitudes), ONG is the Greenwich meridian plane (the reference plane for longitudes) and ONM is the meridian plane of P.

(ii) $OE = OG = OM = a$ and $ON = b$ are semi-major and semi-minor axes of the ellipsoid respectively.

(iii) PH is the ellipsoidal normal and $PQ = h$ is the ellipsoidal height. Q is the projection of P onto the ellipsoid via the normal.

(iv) $HQ = \nu$ (nu) and $CQ = \rho$ (rho) are the radii of curvature in the prime vertical and meridian planes of P respectively.

(v) The normal PH intersects the equatorial plane at D and the angle PDM is the latitude ϕ (phi) of P.

(vi) The longitude λ (lambda) of P is the angle MOG, i.e., the angle between the meridian plane of P and the Greenwich meridian.

(vii) $OH = \nu e^2 \sin\phi$, $DH = \nu e^2$ and $DQ = \nu\left(1 - e^2\right)$

(viii) The x,y,z Cartesian axes are shown with the z-axis passing through the North pole. The x-y plane is the Earth's equatorial plane and the x-z plane is the Greenwich meridian plane. The x-axis passes through the intersection of the Greenwich meridian and the equator and the y-axis is advanced 90° eastwards along the equator. The longitude of P is the angular measure between the Greenwich meridian plane and the meridian plane passing through P and the latitude is the angular measure between the equatorial plane and the normal to the datum surface passing through P.

Longitude is measured 0° to 180° positive east and negative west of the Greenwich meridian and latitude is measured 0° to 90° positive north, and negative south of the equator.

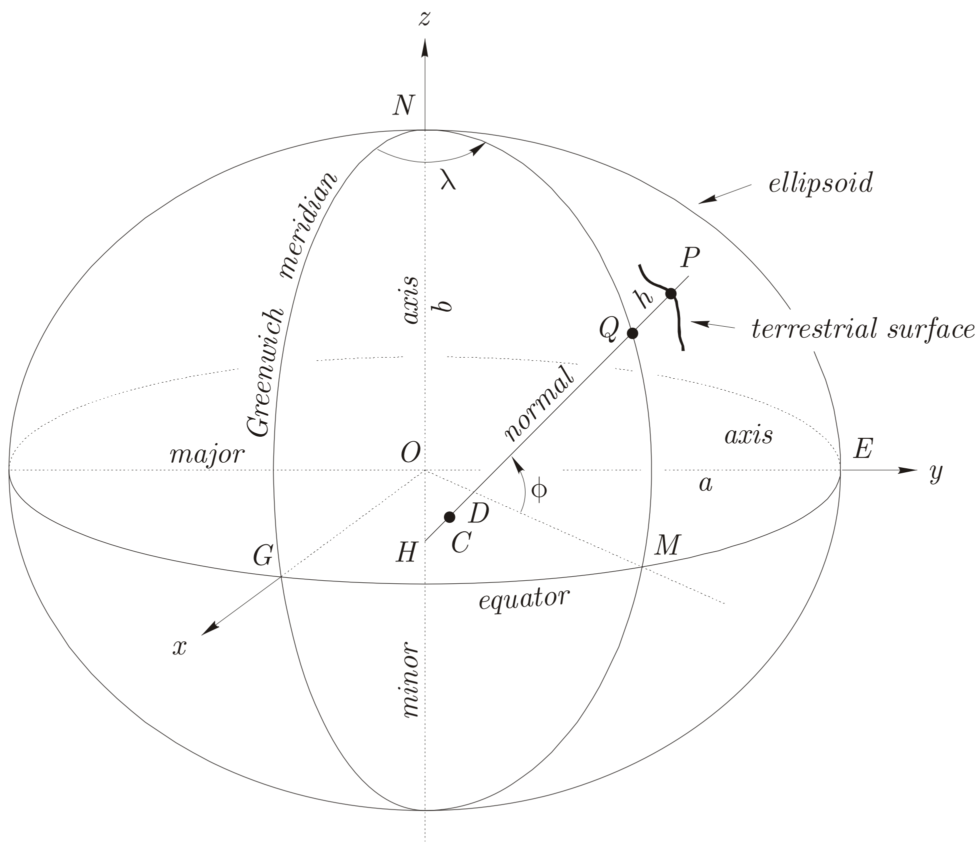

Figure 1. Reference ellipsoid

The following relationships between geometric constants of the ellipsoid are also of use

flattening	$f = \dfrac{a-b}{a}$	(1)
semi-minor axis	$b = a(1-f)$	(2)
first eccentricity squared	$e^2 = \dfrac{a^2-b^2}{a^2} = f(2-f)$	(3)

The radii of curvature of the prime vertical plane (ν) and the meridian plane (ρ) are

$$\nu = \frac{a}{\left(1-e^2\sin^2\phi\right)^{\frac{1}{2}}} \tag{4}$$

$$\rho = \frac{a\left(1-e^2\right)}{\left(1-e^2\sin^2\phi\right)^{\frac{3}{2}}} \tag{5}$$

The x,y,z Cartesian coordinates of P are given by

$$\begin{aligned} x &= (\nu + h)\cos\phi\cos\lambda \\ y &= (\nu + h)\cos\phi\sin\lambda \\ z &= \left[\nu\left(1 - e^2\right) + h\right]\sin\phi \end{aligned} \tag{6}$$

The meridian arc length between two points ϕ_1 and ϕ_2 is given by the integral

$$\text{meridian arc length} = \int_{\phi_1}^{\phi_2} \rho\, d\phi = \int_{\phi_1}^{\phi_2} \frac{a\left(1 - e^2\right)}{\left(1 - e^2\sin^2\phi\right)^{3/2}}\, d\phi$$

This integral can only be evaluated by an expansion in series, followed by term-by-term integration. It is usual to set the lower limit of integration to zero to give the meridian distance m from the equator to a point of latitude ϕ (radians)

$$\begin{aligned} m = a\Bigg\{&\left(1 - \frac{1}{4}e^2 - \frac{3}{64}e^4 - \frac{5}{256}e^6 + \cdots\right)\phi - \frac{3}{8}\left(e^2 + \frac{1}{4}e^4 + \frac{15}{128}e^6 + \cdots\right)\sin 2\phi \\ &+ \frac{15}{256}\left(e^4 + \frac{3}{4}e^6 + \cdots\right)\sin 4\phi - \frac{35}{3072}\left(e^6 + \cdots\right)\sin 6\phi + \cdots\Bigg\} \end{aligned} \tag{7}$$

An alternative is *Helmert's formula* in n that has a faster rate of convergence

$$\begin{aligned} m = a(1-n)\left(1-n^2\right)\Bigg\{&\left(1 + \frac{9}{4}n^2 + \frac{225}{64}n^4 + \cdots\right)\phi - \left(\frac{3}{2}n + \frac{45}{16}n^3 + \cdots\right)\sin 2\phi \\ &+ \frac{1}{2}\left(\frac{15}{8}n^2 + \frac{105}{32}n^4 + \cdots\right)\sin 4\phi - \frac{1}{3}\left(\frac{35}{16}n^3 + \cdots\right)\sin 6\phi \\ &+ \frac{1}{4}\left(\frac{315}{128}n^4 + \cdots\right)\sin 8\phi + \cdots\Bigg\} \end{aligned} \tag{8}$$

where $n = \dfrac{a-b}{a+b} = \dfrac{f}{2-f}$. Substituting $\phi = 90^\circ = \pi/2$ radians into (8) gives Q, the quadrant length from the equator to the pole

$$Q = a(1-n)\left(1-n^2\right)\left\{1 + \frac{9}{4}n^2 + \frac{225}{64}n^4 + \cdots\right\}\frac{\pi}{2} \tag{9}$$

The inverse formula, i.e., the latitude ϕ (radians) given a meridian distance m is

$$\begin{aligned} \phi = \sigma &+ \left(\frac{3}{2}n - \frac{27}{32}n^3 + \cdots\right)\sin 2\sigma + \left(\frac{21}{16}n^2 - \frac{55}{32}n^4 + \cdots\right)\sin 4\sigma \\ &+ \left(\frac{151}{96}n^3 + \cdots\right)\sin 6\sigma + \left(\frac{1097}{512}n^4 + \cdots\right)\sin 8\sigma + \cdots \end{aligned} \tag{10}$$

where $\sigma = \dfrac{m}{Q}\dfrac{\pi}{2}$ radians

The derivation of equations (7) to (10) is given in Lauf (1983, pp. 35-8) and for computation purposes in Australia the part of the infinite series for m given by equation (7) has been limited to the terms shown (ICSM 2003, p.5-19). The error introduced by this truncation is approximately 0.00003 m (Lauf 1983).

Geometric Parameters of Some Selected Ellipsoids

Date	Name	a (metres)	$1/f$
1830	Airy	6377563.396	299.324964600
1830	Everest (India)	6377276.345	300.801700000
1880	Clarke	6378249.145	293.465000000
1924	International	6378388 (exact)	297.0 (exact)
1966	Australian National Spheroid (ANS)	6378160 (exact)	298.25 (exact)
1967	Geodetic Reference System (GRS67)	6378160 (exact)	298.247167427
1980	Geodetic Reference System (GRS80)	6378137 (exact)	298.257222101
1984	World Geodetic System (WGS84)	6378137 (exact)	298.257223563

Table 1. Geometric constants of selected ellipsoids.
From Appendix A1, Technical Report, Department of Defense World Geodetic System 1984 (NIMA 2000)

Prior to 1967 the geometric constants of various ellipsoids were determined from analysis of arc measurements and or astronomic observations in various regions of the Earth, the resulting parameters reflecting the size and shape of "best fit" ellipsoids for those regions; the *International Ellipsoid* of 1924 was adopted by the International Association of Geodesy (at its general assembly in Madrid in 1924) as a best fit of the entire Earth. In 1967 the *International Astronomic Union* (IAU) and the *International Union of Geodesy and Geophysics* (IUGG) defined a set of four physical parameters for the *Geodetic Reference System 1967* based on the theory of a geocentric equipotential ellipsoid. These were: *a,* the equatorial radius of the Earth, *GM,* the geocentric gravitational constant (the product of the Universal Gravitational Constant G and the mass of the Earth *M,* including the atmosphere), J_2 , the dynamical form factor of the Earth and ω, the angular velocity of the Earth's rotation. The geometric constants e^2 and f of an ellipsoid (the *normal ellipsoid*) can be derived from these defining parameters as well as the gravitational potential of the ellipsoid and the value of gravity on the ellipsoid (*normal gravity*).

The *Geodetic Reference System 1980* (GRS80), adopted by the XVII General Assembly of the IUGG in Canberra, December 1979 is the current best estimate with $a = 6378137$ m, $GM = 3986005 \times 10^8$ m^{3}s^{-2}, $J_2 = 108263 \times 10^{-8}$ and $\omega = 7292115 \times 10^{-11}$ rad s^{-1} (BG 1988). The *World Geodetic System 1984* (WGS84), the datum for the Global Positioning System (GPS), is based on the GRS80, except that the dynamical form factor of the Earth is expressed in a modified form, causing very small differences between derived constants of the GRS80 and WGS84 ellipsoids (NIMA 2000). These differences can be regarded as negligible for all practical purposes (a difference of 0.0001 m in the semi-minor axes). The *Geocentric Datum of Australia* (GDA) uses the GRS80 ellipsoid as its reference ellipsoid.

GEODETIC DATUMS AND COORDINATE SYSTEMS IN AUSTRALIA

A map projection is the mathematical transformation of coordinates on one surface, in our case the ellipsoidal reference surface of the Earth, to coordinates on the projection plane. Points P on the Earth's terrestrial surface are related to the ellipsoid via normals passing through those points (see Figure 1 where P is referenced as Q on the surface of the ellipsoid) and have *geodetic coordinates* ϕ, λ, h (latitude, longitude, ellipsoidal height). The third coordinate *h*, <u>plays no part in any map projection</u> and we are only interested in ϕ and λ, *curvilinear coordinates* of the reference surface for P (and Q), i.e., in any map projection we are only transforming points on the ellipsoid to points on the projection.

Before any sensible mapping (or coordination) of a region can take place a *Geodetic Datum* must be established; which in its simplest form consists of two "actions"; (i) a definition of the size and shape of a suitable reference ellipsoid, and (ii) the location of the ellipsoid's centre and orientation of its minor axis with respect to the Earth's centre of mass and rotational axis.

In Australia, mapping and coordination is related to two geodetic datums, the *Australian Geodetic Datum* (AGD) in use since 1966 and the more modern *Geocentric Datum of Australia* (GDA), in use since the late 1990's. These two geodetic datums have different ellipsoids; the AGD uses the Australian National Spheroid (ANS) and the GDA uses the ellipsoid of the GRS80 (see Table 1) and the centres of these ellipsoids are at different locations. The centre of the GDA ellipsoid can be assumed to be at the Earth's centre of mass (or geocentre, hence the term geocentric in the datum name) whilst the centre of the ANS is displaced from the geocentre by approximately $\delta x = +133$ m, $\delta y = +48$ m, $\delta z = -148$ m where $x_{AGD} = x_{GDA} + \delta x$ and similarly for y and z (Appendix B, NIMA 200, with GDA replacing WGS84). The minor axis of the GDA's ellipsoid is considered to be coincident with the Earth's rotational axis and the minor axis of the AGD's ellipsoid is considered to be parallel with the Earth's rotational axis.

Having two geodetic datums leads to the interesting (and often confusing) fact that a single point can have two sets of geodetic coordinates (ϕ, λ). Figure 2, showing zOy meridian sections of the AGD and GDA ellipsoids (greatly exaggerated), hopefully explains this situation.

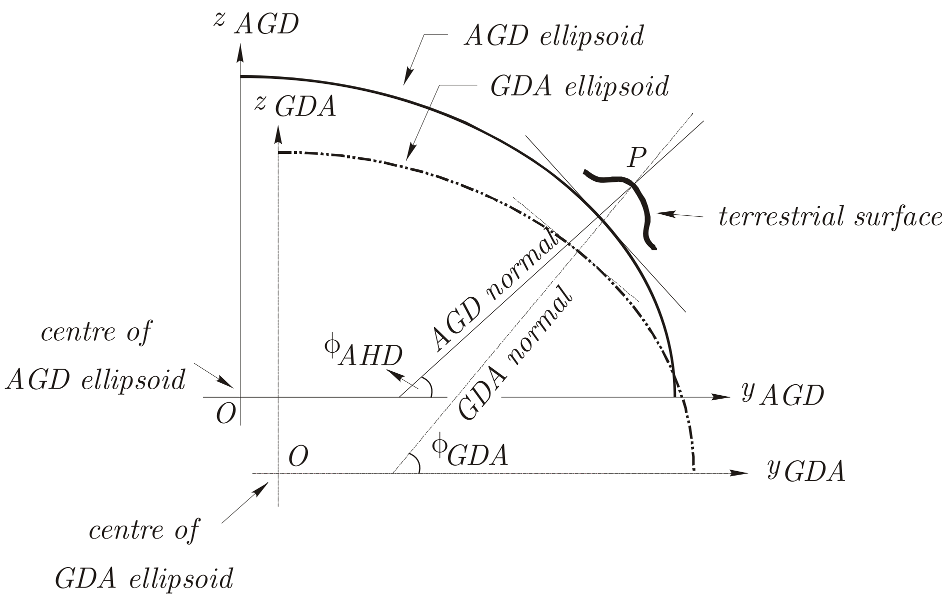

Figure 2. Sections of AGD and GDA ellipsoids showing two latitudes for the single point P

Some confusion also arises from the fact that there have been two "realizations" of the AGD, i.e., there was an initial adjustment of the national geodetic network in 1966 leading to an AGD coordinate set (latitudes and longitudes) followed by a subsequent re-adjustment of the network, that had been improved by additional measurements and stations, leading to another AGD coordinate set. The two sets were designated AGD66 and AGD84. The following extract from a paper titled 'Transforming Cartesian coordinates X,Y,Z to geographical coordinates ϕ, λ, h' published in the *Australian Surveyor* (Gerdan & Deakin 1999) gives an explanation of these two data sets and the AGD.

> In 1966, under the direction of the National Mapping Council (NMC) all geodetic surveys in Australia were recomputed and adjusted on the then new AGD, an astronomically derived topocentric datum having a physical origin near the centroid of the geodetic network and fixing an ellipsoid of revolution, the Australian National Spheroid (ANS), with respect to the Earth's rotational axis. The national adjustment yielded an homogeneous set of geographical coordinates (latitudes and longitudes) for the geodetic network. At the same time, the NMC defined a system of rectangular grid coordinates (eastings and northings) known as the Australian Map Grid (AMG), based on a Universal Transverse Mercator (UTM) projection of AGD latitudes and longitudes.
>
> After 1966 there were several readjustments of the national geodetic network, densified and strengthened by the inclusion of improved measurements, each readjustment referred to as a Geodetic Model of Australia (GMA). In 1984 the NMC, recognizing the eventual need for Australia to convert to a geocentric datum, adopted the latest readjustment at the time, GMA82, as an interim step in this process. This geographical coordinate set was defined as AGD84 with AMG84 grid coordinates, and to avoid confusion, earlier coordinate sets derived from the 1966 adjustment were defined as AGD66 and AMG66. Both ADG66 and AGD84 coordinates have a common datum (defined in 1966) excepting that AGD84 coordinates were derived from an adjustment, which more correctly allowed for the separation between the geoid and the ANS over Australia (NMC 1986).
>
> In 1988, the NMC was superseded by the Intergovernmental Committee on Surveying and Mapping (ICSM), representing the mapping organizations of the States and Territories of the Commonwealth of Australia and New Zealand. The GDA was adopted by the ICSM in November 1994 in response to anticipated demand by major users of GPS technology such as the Australian Defence Force, the International Civil Aviation Organization, the International Hydrographic Organization and the International Association of Geodesy (Steed 1996). The new datum is primarily based on the coordinates of eight geologically stable sites across Australia with permanent GPS tracking facilities known as the Australian Fiducial Network (AFN), supplemented by a network of seventy survey stations (covering Australia

at approximately 500km intervals) which together form the Australian National Network (ANN). Geocentric Cartesian coordinates of these stations were derived from an adjustment of precise GPS observations obtained from – (i) a two week global observation period in 1992 conducted by the International GPS Geodynamics Service at approximately two hundred sites around the world (including all the AFN sites) and (ii) ICSM campaigns in 1992, '93 and '94 linking all AFN and ANN sites. These coordinates are related to the International Earth Rotation Service (IERS) Terrestrial Reference Frame for 1992 (ITRF92) at epoch 1994.0 [The epoch 1994.0 (1st Jan. 1994) reflects the fact that monitoring stations used by IERS are moving with respect to each other due to earth crustal motion; the epoch date indicating the datum is ITRF92 adjusted for station motion in the intervening period]. The ICSM has defined GDA94 coordinates as latitudes and longitudes related to the ellipsoid of the Geodetic Reference System 1980 (GRS80) [BG 1988] and Map Grid Australia 1994 (MGA94) grid coordinates as a UTM projection of those latitudes and longitudes.

SOME MAP PROJECTION THEORY

A map projection is the mathematical transformation of coordinates on a datum surface to coordinates on a projection surface. In all the map projections we will be dealing with, the datum surface is an ellipsoid representing the Earth and on this surface, there are imaginary sets of reference curves, or parametric curves, that we use to coordinate points. We know these parametric curves as parallels of latitude ϕ and meridians of longitude λ and along these curves one of the parameters, ϕ or λ are constant. Points on the datum surface having particular values of ϕ and λ are said to have curvilinear coordinates that we commonly call geographical or geodetic coordinates. Points on the datum surface can also have *x,y,z* Cartesian coordinates and there are mathematical connections between the curvilinear and Cartesian that we call functional relationships and write as

$$\begin{aligned} x &= f_1(\phi,\lambda) = \nu\cos\phi\cos\lambda \\ y &= f_2(\phi,\lambda) = \nu\cos\phi\sin\lambda \\ z &= f_3(\phi,\lambda) = \nu\left(1-e^2\right)\sin\phi \end{aligned} \tag{11}$$

Figure 3(a) shows a datum surface representing the Earth with meridians and parallels (the ϕ,λ parametric curves) and the continental outlines.

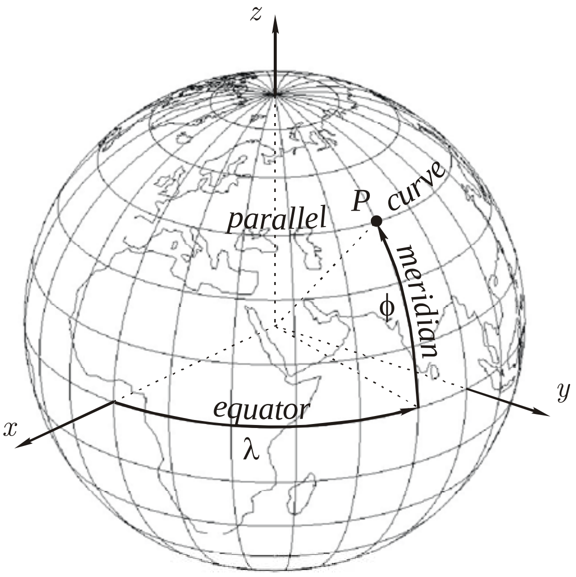

Figure 3(a)

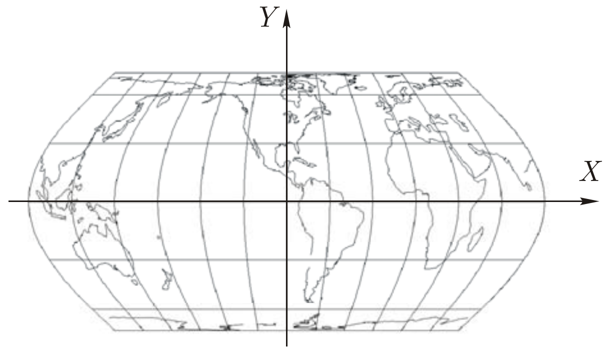

Figure 3(b)

Figure 3(b) shows the projection surface, which we commonly refer to as the map projection. In this case the projection is a *modified Sinusoidal projection,* and as in all cases we will deal with in this paper, the projection surface is a plane. [In general, the projection surface may be another curved 3D surface and we use this general concept in the theoretical development that follows]. On the projection there are sets of parametric curves, say U, V curves that are the projected meridians and parallels and points on the projection surface have U, V curvilinear coordinates. These coordinates are related to another 3D Cartesian coordinate system X, Y, Z and the two systems are related by another set of functional relationships

$$\begin{aligned} X &= F_1(U,V) \\ Y &= F_2(U,V) \\ Z &= F_3(U,V) = 0 \end{aligned} \tag{12}$$

In the case of a plane projection surface $Z = 0$ and we would like to establish the connections between the curvilinear coordinates ϕ, λ on the datum surface and X, Y Cartesian coordinates of the projection plane, i.e., we wish to find the functional relationships

$$\begin{aligned} X &= g_1(\phi,\lambda) \\ Y &= g_2(\phi,\lambda) \end{aligned} \tag{13}$$

We call these functional relationships the projection equations and they can be derived from an understanding of distortions and scale factors that measure the distortions. Inspection of the map projection, Figure 3(b), reveals distortions that we see as misshapen continental outlines (Antarctica), points projected as lines (the north and south poles) and straight lines projected as curves (the meridians). Every map projection has distortions of one sort or another and we would like to quantify these distortions. It turns out that distortions can be related to scale factors where scale is the ratio of elemental distances on the datum surface and the projection surface, and a knowledge of scale factors allow us to "uncover" the projection equations by enforcing scale conditions and particular geometric constraints.

The elemental distance *ds* on the datum surface

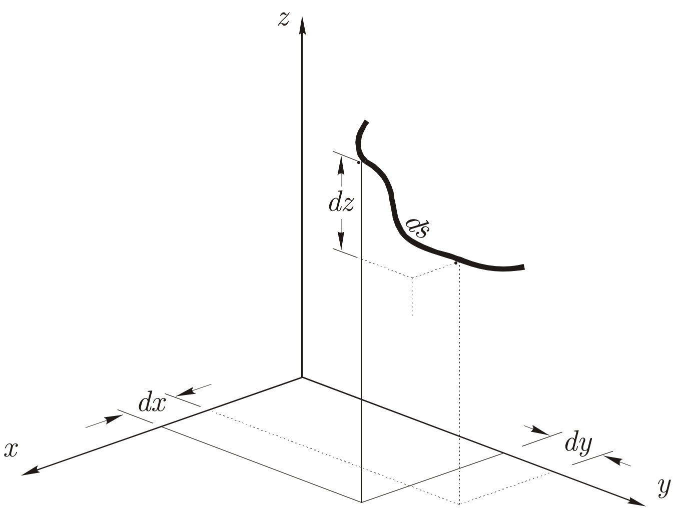

Figure 4. The elemental distance ds

From differential geometry, the square of the length of a differentially small part of a curve on the datum surface is

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{14}$$

From the functional relationships of (11) the total differentials are

$$\begin{aligned} dx &= \frac{\partial x}{\partial \phi} d\phi + \frac{\partial x}{\partial \lambda} d\lambda \\ dy &= \frac{\partial y}{\partial \phi} d\phi + \frac{\partial y}{\partial \lambda} d\lambda \\ dz &= \frac{\partial z}{\partial \phi} d\phi + \frac{\partial z}{\partial \lambda} d\lambda \end{aligned} \tag{15}$$

Substituting equations (15) into equation (14) gathering terms and simplifying gives

$$ds^2 = e\,d\phi^2 + 2f\,d\phi\,d\lambda + g\,d\lambda^2 \tag{16}$$

The coefficients of $d\phi^2$, $d\phi\,d\lambda$ and $d\lambda^2$ are called the Gaussian Fundamental Quantities and are invariably indicated in the map projection literature by *e, f* and *g* or *E, F* and *G*. In this paper, lower case letters *e, f* and *g* relate to the datum surface and uppercase letters *E, F* and *G* relate to the projection surface

$$\begin{aligned}
e &= \left(\frac{\partial x}{\partial \phi}\right)^2 + \left(\frac{\partial y}{\partial \phi}\right)^2 + \left(\frac{\partial z}{\partial \phi}\right)^2 \\
f &= \frac{\partial x}{\partial \phi}\frac{\partial x}{\partial \lambda} + \frac{\partial y}{\partial \phi}\frac{\partial y}{\partial \lambda} + \frac{\partial z}{\partial \phi}\frac{\partial z}{\partial \lambda} \\
g &= \left(\frac{\partial x}{\partial \lambda}\right)^2 + \left(\frac{\partial y}{\partial \lambda}\right)^2 + \left(\frac{\partial z}{\partial \lambda}\right)^2
\end{aligned} \tag{17}$$

Every surface having curvilinear coordinates also has Gaussian Fundamental Quantities, for the ellipsoid with parallels and meridians ϕ, λ these quantities can be determined from equations (11) and (17) as

$$e = \rho^2, \quad f = 0, \quad g = \nu^2 \cos^2 \phi \tag{18}$$

The elemental rectangle on the datum surface (the ellipsoid)

In general, the elemental distance *ds* on the ellipsoid may be shown as the diagonal of a differentially small rectangle

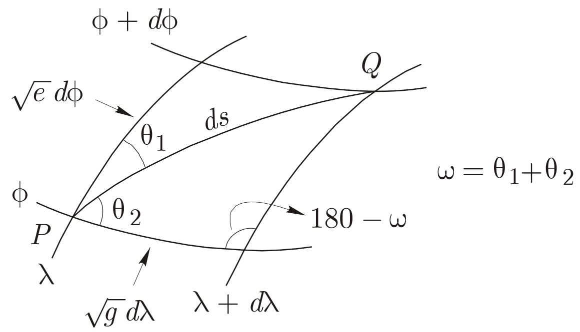

Figure 5. The elemental rectangle

Figure 5 shows two differentially close points P and Q on the datum surface. The parametric curves ϕ and λ pass through P and the curves $\phi + d\phi$ and $\lambda + d\lambda$ pass through Q. The distance PQ is the elemental distance ds. The elemental rectangle formed by the curves may be regarded as a plane figure whose opposite sides are parallel straight lines enclosing a differentially small area da. The angle between the parametric curves ϕ and λ is equal to $\omega = \theta_1 + \theta_2 = 90°$

The elemental distances along parametric curves on the ellipsoid

The elemental distances along the ϕ and λ curves can be obtained from equation (16) considering the fact that along the ϕ-curve, ϕ is a constant value, hence $d\phi = 0$ and along the λ-curve, λ is a constant and $d\lambda = 0$, hence the elemental distance along the λ-curve (a meridian) is

$$\begin{aligned} ds_\lambda^2 &= e\,d\phi^2 + 2f\,d\phi\,d\lambda + g\,d\lambda^2 \\ &= e\,d\phi^2 + 2f\,d\phi\,(0) + g\,(0)^2 \\ &= e\,d\phi^2 \end{aligned}$$

and $$ds_\lambda = \sqrt{e}\,d\phi = \rho\,d\phi \tag{19}$$

Similarly, the elemental distance along the ϕ-curve (a parallel) is

$$ds_\phi = \sqrt{g}\,d\lambda = \nu\cos\phi\,d\lambda \tag{20}$$

The angle ω between parametric curves on the datum surface

The elemental rectangle can be regarded as a plane within its infinitely small area and from the cosine rule for plane trigonometry, and bearing in mind that $\cos(180 - x) = -\cos x$

$$\begin{aligned} ds^2 &= e\,d\phi^2 + g\,d\lambda^2 - 2\left(\sqrt{e}\,d\phi\right)\left(\sqrt{g}\,d\lambda\right)\cos(180 - \omega) \\ &= e\,d\phi^2 + g\,d\lambda^2 + 2\sqrt{eg}\,d\phi\,d\lambda\cos\omega \end{aligned} \tag{21}$$

Equating (21) and (16) gives an expression for the angle ω, the angle between the parametric curves

$$\cos\omega = \frac{f}{\sqrt{eg}} \tag{22}$$

Thus, we may say: if the parametric curves on the datum surface intersect at right angles (i.e., they are an orthogonal system of curves) then $\omega = 90^\circ$ and $\cos\omega = 0$. This implies that $f = 0$. For the ellipsoid, where the parametric curves are the orthogonal network of meridians and parallels $f = 0$, see equations (18).

Elemental quantities on the projection surface

Using similar developments as we used for the datum surface, the following relationships for the projection surface may be derived.

The elemental distance dS on the projection surface

$$dS^2 = dX^2 + dY^2 = E\,d\phi^2 + 2F\,d\phi\,d\lambda + G\,d\lambda^2 \tag{23}$$

where Cartesian coordinates X, Y are functions of ϕ, λ and the Gaussian Fundamental Quantities for the projection surface are E, F and G

$$\begin{aligned} E &= \left(\frac{\partial X}{\partial \phi}\right)^2 + \left(\frac{\partial Y}{\partial \phi}\right)^2 \\ F &= \frac{\partial X}{\partial \phi}\frac{\partial X}{\partial \lambda} + \frac{\partial Y}{\partial \phi}\frac{\partial Y}{\partial \lambda} \\ G &= \left(\frac{\partial X}{\partial \lambda}\right)^2 + \left(\frac{\partial Y}{\partial \lambda}\right)^2 \end{aligned} \tag{24}$$

The angle Ω between the parametric curves on the projection surface (the projected meridians and parallels)

$$\cos\Omega = \frac{F}{\sqrt{EG}} \tag{25}$$

Scale Factor

Knowledge of scale factors is fundamental in understanding map projections and deriving projection equations. Using certain scale factors, or scale relationships, we may create map projections with certain useful properties. For example, map projections that preserve angles at a point are known as conformal, i.e., an angle between two lines on the datum surface is transformed into the same angle between the complimentary lines on the projection. Conformal projections have the unique property that the scale factor is the same in every direction at a point on the projection. Therefore, we may derive the equations for a conformal map projection by enforcing a particular scale relationship.

The equation for (linear) scale factor m is defined as the ratio of elemental distances dS on the projection and ds on the datum surface

$$\text{scale factor } m = \frac{\text{elemental distance on PROJECTION SURFACE}}{\text{elemental distance on DATUM SURFACE}} = \frac{dS}{ds}$$

or
$$m^2 = \frac{dS^2}{ds^2} = \frac{E\,d\phi^2 + 2F\,d\phi\,d\lambda + G\,d\lambda^2}{e\,d\phi^2 + 2f\,d\phi\,d\lambda + g\,d\lambda^2} \tag{26}$$

Dividing numerator and denominator of (26) by $d\lambda^2$ gives

$$m^2 = \frac{E\left(\dfrac{d\phi}{d\lambda}\right)^2 + 2F\dfrac{d\phi}{d\lambda} + G}{e\left(\dfrac{d\phi}{d\lambda}\right)^2 + 2f\dfrac{d\phi}{d\lambda} + g} \tag{27}$$

Inspection of this equation shows that in general the scale factor at a point depends directly on the term $d\phi/d\lambda$ since for the datum and projection surfaces e, f, g and E, F, G are constant for a particular point. Referring to Figure 5, $d\phi/d\lambda$ is the ratio between elemental changes $d\phi$ and $d\lambda$, and for any curve on the datum surface this ratio will vary according to the azimuth α of the curve. If the parametric curves on the surface intersect at right angles, as meridians of longitude and parallels of latitude do, then we can express this as

$$\tan\alpha = \frac{\sqrt{g}\,d\lambda}{\sqrt{e}\,d\phi}$$

where α is a positive clockwise angle measured from the λ-curve ($\alpha = \theta_1$). This equation may be rearranged to give expressions for the ratio $d\phi/d\lambda$

$$\frac{d\phi}{d\lambda} = \frac{\sqrt{g}}{\sqrt{e}\tan\alpha} \qquad \text{and} \qquad \left(\frac{d\phi}{d\lambda}\right)^2 = \frac{g}{e\tan^2\alpha}$$

Substituting these expressions into equation (27) and simplifying using trigonometric relationships gives

$$m^2 = \frac{\left(\frac{E}{e}\right)\cos^2\alpha + 2\left(\frac{F}{f}\right)\frac{f}{\sqrt{eg}}\sin\alpha\cos\alpha + \left(\frac{G}{g}\right)\sin^2\alpha}{1 + 2\frac{f}{\sqrt{eg}}\sin\alpha\cos\alpha}$$

and since $f = 0$ (parametric curves on the surface intersecting at right angles)

$$m^2 = \left(\frac{E}{e}\right)\cos^2\alpha + \frac{F}{\sqrt{eg}}2\sin\alpha\cos\alpha + \left(\frac{G}{g}\right)\sin^2\alpha \tag{28}$$

Important results from the equation for scale factor

1. Scale factor varies everywhere on the map projection. This fact can be deduced from equation (28) when it is realized that the Gaussian Fundamental Quantities are functions of the curvilinear coordinates ϕ, λ of the datum surface. Therefore, as points vary across the datum surface their complimentary points on the projection will have a varying scale factor.

2. When $\frac{E}{e} = \frac{G}{g}$ and $F = 0$ the scale factor is independent of direction, i.e., m is the same value in every direction about a point on the projection. Such projections are known as <u>CONFORMAL</u>. We can verify this by substituting a constant $K = \frac{E}{e} = \frac{G}{g}$ and $F = 0$ into (28) giving

   $$m^2 = K\cos^2\alpha + K\sin^2\alpha = K\left(\cos^2\alpha + \sin^2\alpha\right) = K$$

   Note that when $F = 0$, the parametric curves on the projection (i.e., the projected meridians and parallels) intersect at right angles.

Conformal projections have the property that shape is preserved. By this we mean that an object on the datum surface, say a square, is transformed into a square on the projection surface although it may be enlarged or reduced by a constant amount. Preservation of shape also means that angles at a point are preserved. By this we mean that an angle between two lines radiating from a point on the datum surface will be identical to the angle between the two projected lines on the projection surface. There is one minor drawback: these properties only hold true for differentially small areas since the relationships have been established from the differential ratio $m^2 = dS^2 / ds^2$. Nevertheless, these properties make conformal projections the most appropriate for topographic mapping; since measurements in the field, corrected to the ellipsoid (the datum surface), need little or no further correction and can be added to a conformal map. This fact becomes more obvious when we consider the size of the Earth and any practical mapping area we might be working on. Consider a 1:100,000 Topographic map sheet used in Australia. This map series is based on a conformal projection (UTM) of latitudes and longitudes of points related to the ellipsoid and cover 0° 30' of latitude and longitude. This equates roughly to 2,461,581,000 m^2 of the Earth's surface. The surface area of the Australian National Spheroid, a reasonable approximation to the Earth, is 5.1006927×10^{14} m^2, which means the map sheet is 0.000483% of the Earth's surface. Thus the entire map sheet can be regarded as an extremely small (almost differentially small) portion of the Earth's surface.

3. Consider the case where the datum surface is an ellipsoid with meridians and parallels as the parametric curves and two points P and Q an elemental distance ds apart. When Q is on the meridian passing through P then α, the azimuth of line PQ on the datum surface is 0° or 180° and $\cos\alpha = 1$ and $\sin\alpha = 0$ and

The meridian scale factor h $$h = \frac{\sqrt{E}}{\sqrt{e}} \tag{29}$$

Similarly, when Q is on the parallel passing through P then

The parallel scale factor k $$k = \frac{\sqrt{G}}{\sqrt{g}} \tag{30}$$

This leads to the common definition of a conformal projection:

When $f = F = 0$ and $h = k$ the projection is conformal

CYLINDRICAL MAP PROJECTIONS

In elementary texts on map projections, the projection surfaces are often described as developable surfaces, such as the cylinder (cylindrical projections) and the cone (conical projections), or a plane (azimuthal projections). These surfaces are imagined as enveloping or touching the datum surface and by some means, usually geometric, the meridians, parallels and features are projected onto these surfaces. In the case of the cylinder, it is cut and laid flat (developed). If the axis of the cylinder coincides with the axis of the Earth, the projection is said to be normal aspect, if the axis lies in the plane of the equator the projection is known as transverse and in any other orientation it is known as oblique. [It is usual that the descriptor "normal" is implied in the name of a projection, but for different orientations, the words "transverse" or "oblique" are added to the name.] This simplified approach is not adequate for developing a general theory of projections (which as we can see is quite mathematical) but is useful for describing characteristics of certain projections. In the case of cylindrical projections, some characteristics are a common feature:

(i) Meridians of longitude and parallels of latitude form an orthogonal network of straight parallel lines.
(ii) Meridians are equally spaced straight parallel lines intersecting parallels at right angles.
(iii) Parallels, in general, are unequally spaced straight parallel lines but are symmetric about the equator.

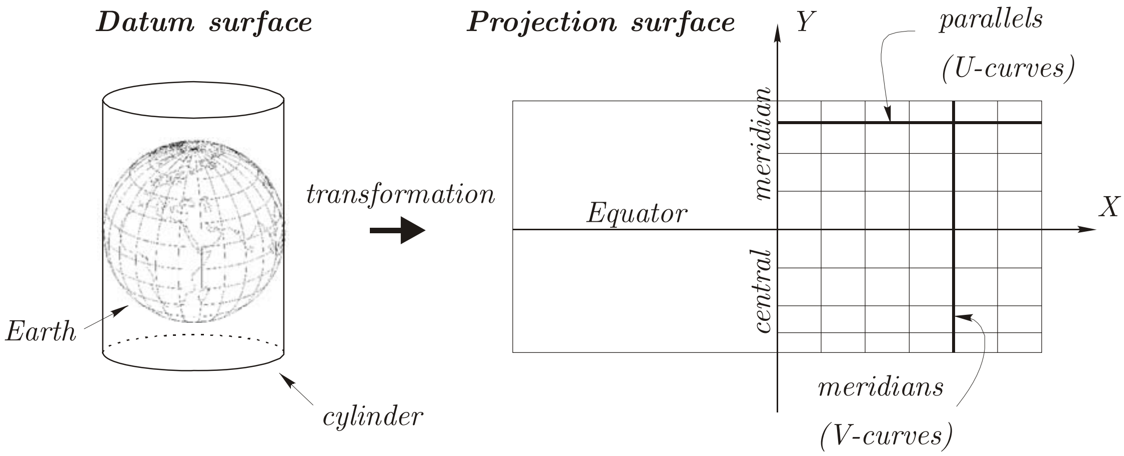

Figure 6. Schematic diagram of normal aspect cylindrical projection

Mercator's projection (normal aspect cylindrical conformal)

The equations for Mercator's projection (of the ellipsoid) are derived in the following manner.

Since the parametric curves on the ellipsoid and the projection are both orthogonal nets, i.e., $f = F = 0$ and $X = f_1(\lambda)$ and $Y = f_2(\phi)$ the Gaussian Fundamental Quantities E and G of the projection surface are

$$E = \left(\frac{\partial X}{\partial \phi}\right)^2 + \left(\frac{\partial Y}{\partial \phi}\right)^2 = \left(\frac{\partial Y}{\partial \phi}\right)^2$$

$$G = \left(\frac{\partial X}{\partial \lambda}\right)^2 + \left(\frac{\partial Y}{\partial \lambda}\right)^2 = \left(\frac{\partial X}{\partial \lambda}\right)^2$$

The Gaussian Fundamental Quantities e and g of the datum surface, given by equation (18), are

$$e = \rho^2, \quad g = \nu^2 \cos^2 \phi$$

The projection is to be conformal and the scale condition to be enforced is

$$h = k \quad \text{or} \quad \frac{\sqrt{E}}{\sqrt{e}} = \frac{\sqrt{G}}{\sqrt{g}}$$

Substituting expressions for E, G, e and g and rearranging gives the scale condition in the form of a differential equation

$$\frac{dY}{d\phi} = \frac{\rho}{\nu \cos\phi} \frac{dX}{d\lambda} \tag{31}$$

To simplify this equation, we can enforce a particular scale condition: that the scale along the equator be unity. Since k is the scale factor along a parallel, we may denote the scale factor along the equator as k_0 where

$$k_0 = \frac{dX}{\nu_0 \cos\phi_0 \, d\lambda} = 1$$

Now since $\phi_0 = 0^\circ$ and $\cos\phi_0 = 1$ and $\nu_0 = a$ [see equation (4)] this particular scale condition gives rise to the differential equation

$$dX = a \, d\lambda \tag{32}$$

Substituting equation (32) into (31) and rearranging gives

$$dY = a \frac{\rho}{\nu \cos\phi} d\phi \tag{33}$$

Integrating equations (32) and (33) gives the projection equations for Mercator's projection of the ellipsoid (Lauf 1983).

$$\begin{aligned} X &= a(\lambda - \lambda_0) \\ Y &= a\left\{ \frac{1}{2}\ln\left(\frac{1+\sin\phi}{1-\sin\phi} \right) - \frac{e}{2}\ln\left(\frac{1+e\sin\phi}{1-e\sin\phi} \right) \right\} \\ &= a\left\{ \ln\left(\tan\left(\frac{\pi}{4} + \frac{\phi}{2} \right) \right) - \frac{e}{2}\ln\left(\frac{1+e\sin\phi}{1-e\sin\phi} \right) \right\} \\ &= a\ln\left\{ \tan\left(\frac{\pi}{4} + \frac{\phi}{2} \right)\left(\frac{1-e\sin\phi}{1+e\sin\phi} \right)^{e/2} \right\} \end{aligned} \tag{34}$$

where a and e are the semi-major axis and eccentricity of the ellipsoid respectively, ln is the natural logarithm and λ_0 is the longitude of the central meridian of the projection.

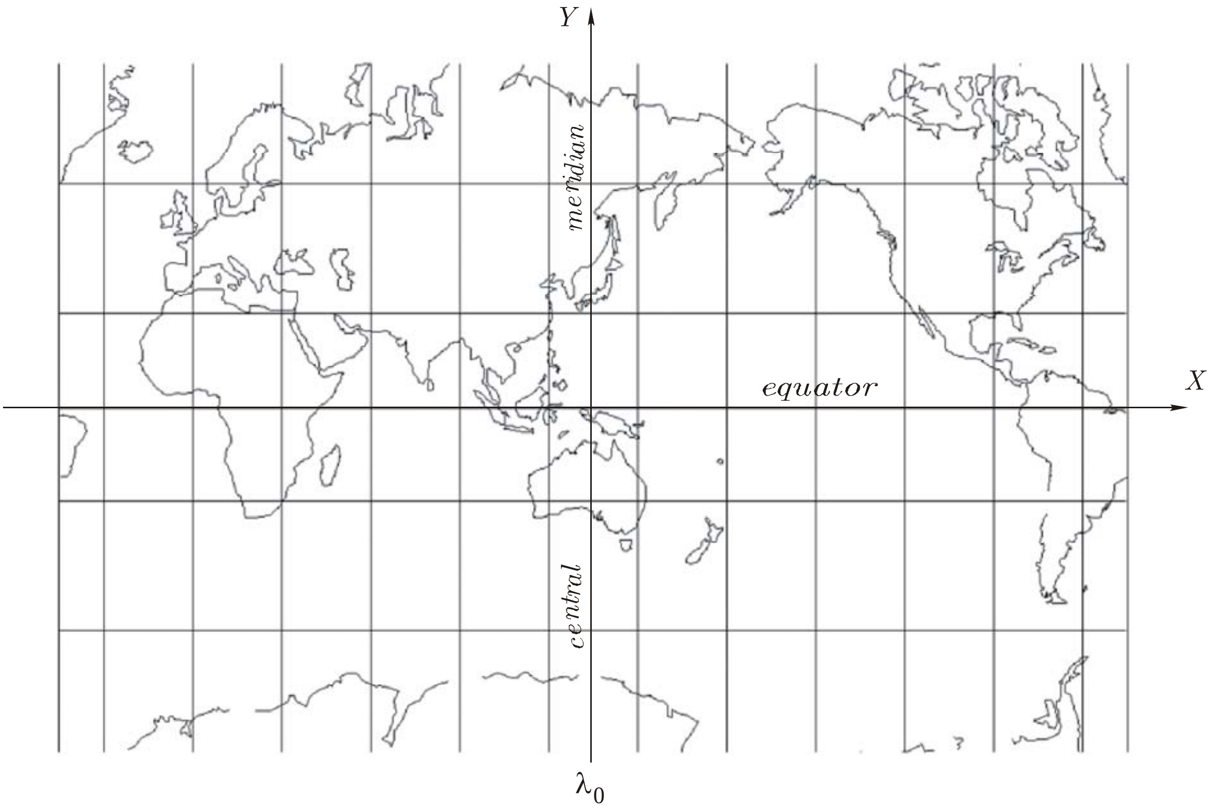

Figure 7. Mercator's projection (cylindrical conformal)
graticule interval 30°, central meridian 135°

TRANSVERSE MERCATOR AND UNIVERSAL TRANSVERSE MERCATOR PROJECTION

Mercator's projection has low scale error in a small latitude band close to the equator but increasingly larger scale errors in higher latitudes regions. By rotating the imaginary cylinder touching the Earth (see Figure 6) by 90° the central line of the projection, which is the equator in the normal aspect form, becomes a central meridian (having constant scale factor) in the transverse form and the poles lay on this line. The meridians and parallels are complex curves (intersecting everywhere at right angles), excepting the equator and the central meridian that are projected as straight lines intersecting at right angles.

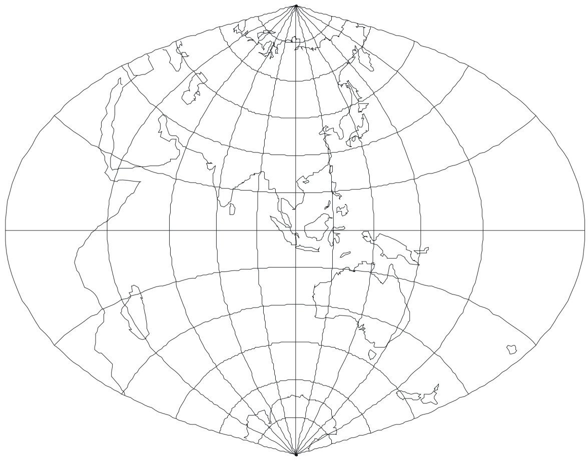

Figure 8. Transverse Mercator projection
graticule interval 15°, central meridian 105°

As in the Mercator projection, the Transverse Mercator (TM) projection has low scale error in a small longitude band about the central meridian but increasingly larger scale errors as the longitude difference from the central meridian increases. Because of this limitation, the TM projection is only used to map small bands of longitude (generally less than 3° to 4° either side of a central meridian).

The TM projection in its spherical form was invented by the mathematician and cartographer Johann Heinrich Lambert (1728-77) and was the third of seven new projections which he described in his work *Beiträge*[1] (Lambert 1772). The ellipsoidal form was developed by C.F. Gauss (1777-1855) in 1822 and L. Kr ger published studies in 1912 and 1919 providing formulae for the ellipsoid; in Europe the projection is sometimes called the *Gauss Conformal* or the *Gauss-Krüger.* The name *Transverse Mercator,* now in common usage, was first applied by the French map projection compiler Germain (Snyder 1987).

The Universal Transverse Mercator (UTM) projection and associated grid were adopted by the U.S. Army in 1947 for designating rectangular coordinates on large-scale military maps of the entire world. The UTM is the TM projection of the ellipsoid with specific parameters, such as numbered zones with designated central meridians, a defined central meridian scale factor, false origin locations in northern

[1] *Beiträge* means Contributions

and southern hemispheres, etc. All formulae derived in this paper for the TM projection are applicable to the UTM projection (Snyder 1987).

The equations for the TM projection of the ellipsoid are derived from a principle of conformal mapping developed by Gauss, i.e., conformal transformations from the ellipsoid to the plane can be represented by the complex expression

$$Y + iX = f(q + i\omega) \tag{35}$$

Providing that q and ω are *isometric* parameters and the complex function $f(q+i\omega)$ is *analytic.* In equation (35) X, Y are Cartesian coordinates on the projection plane, $i = \sqrt{-1}$ (the imaginary number), q is the isometric latitude on the ellipsoid and $\omega = \lambda - \lambda_0$ is a longitude difference (on the ellipsoid) from a central meridian. The left-hand side of (35) is a complex number (or variable) containing two parts, the *real part,* consisting of the parameter Y and the *imaginary part* consisting of the parameter X. The right-hand-side of (35) is a complex function, i.e., a function of real and imaginary parameters q and ω respectively. The word *isometric* means "of equal measure" and the parameters q and ω in the complex variable on the right-hand-side of (35) are isometric parameters related to the parameters ϕ and λ. The complex function $f(q+i\omega)$ is *analytic* if it is everywhere differentiable and we may think of an analytic function as one that describes a smooth surface having no holes, edges or discontinuities.
A necessary and sufficient condition for $f(q+i\omega)$ to be analytic is that the *Cauchy-Riemann* equations are satisfied, i.e., (Sokolnikoff & Redheffer 1966)

$$\frac{\partial Y}{\partial q} = \frac{\partial X}{\partial \omega} \qquad \text{and} \qquad \frac{\partial Y}{\partial \omega} = -\frac{\partial X}{\partial q} \tag{36}$$

Isometric parameters of the ellipsoid

Isometric means "of equal measure" and we may think of isometric parameters q and ω on the ellipsoid in the following way. Imagine you are standing on the surface of the Earth (an ellipsoid) at the equator and you measure out a metre north ds_λ and also a metre east ds_ϕ. Both of these equal lengths on the Earth would represent almost equal angular changes in latitude $d\phi$ and longitude $d\lambda$. Now imagine that you are close to the North Pole; a metre in the north direction will represent the same angular change $d\phi$ as it did at the equator, but a metre in the east direction would represent a much greater change in longitude, i.e., equal north and east linear measures near the pole do not correspond to equal angular measures. What we require is a variable angular measure along a meridian of longitude; we call this quantity the isometric latitude and it can be determined in the following manner.

Consider the elemental rectangle in Figure 5 and equations (19) and (20); we can see that the elemental distances ds_λ and ds_ϕ are not equal for equal angular differentials $d\phi$ and $d\lambda$. Thus the ϕ, λ curvilinear coordinate system of parametric curves is not an isometric system. We can create an isometric system q, λ by writing an expression for the elemental distance ds on the ellipsoid as (see equations (16) and (18))

$$\begin{aligned} ds^2 &= \rho^2 d\phi^2 + \nu^2 \cos^2\phi\, d\lambda^2 \\ &= \nu^2 \cos^2\phi \left\{ \left(\frac{\rho\, d\phi}{\nu \cos\phi} \right)^2 + d\lambda^2 \right\} \\ &= \nu^2 \cos^2\phi \left(dq^2 + d\lambda^2 \right) \end{aligned} \tag{37}$$

q is known as the isometric latitude defined by the differential relationship

$$dq = \frac{\rho}{\nu \cos\phi} d\phi \tag{38}$$

and the new curvilinear coordinate system (q, λ) is an isometric system with isometric parameters. We can see this from equation (37), where the elemental distances along the parametric curves q and λ are $ds_\lambda = \nu \cos\phi\, dq$ and $ds_q = \nu \cos\phi\, d\lambda$, i.e., the elemental distances are equal for equal angular differentials dq and $d\lambda$.

TM projection equations

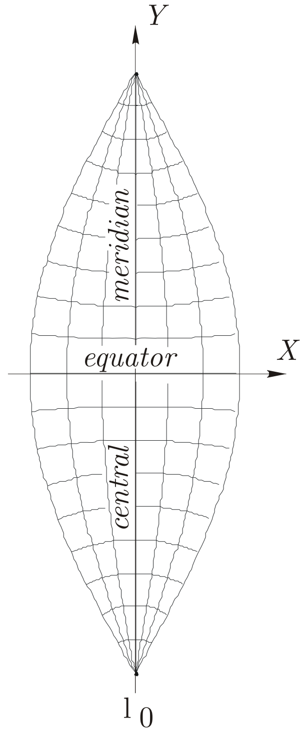

Figure 9. TM projection

To establish the projection equations the function $f(q + i\omega)$ of equation (35) must be determined. To do this, two conditions are enforced:

(i) the Y-axis shall represent a meridian and

(ii) the scale factor along that meridian (the Y-axis) is constant.

The first condition demands that when $X = 0,\ Y = f(q)$, i.e., Y is a function of the isometric latitude q only and hence $\omega = 0$. This means that the Y-axis is the *central meridian* λ_0 and is the origin of longitude differences $\omega = \lambda - \lambda_0$.

The second condition demands that when $X = 0,\ Y = k_0\, m$, where k_0 is the central meridian scale factor and m is the meridian distance on the ellipsoid from the equator to the point. But when $X = 0,\ Y = f(q)$ hence,

$$f(q) = k_0\, m \tag{39}$$

is the necessary condition.

Equation (35), the complex "mapping" equation, can be approximated (on the right-hand-side) by a power series of ever smaller terms using Taylor's theorem. Consider a point P having isometric coordinates q, ω linked to an approximate location q_0, ω_0 by very small corrections $\delta q, \delta\omega$ such that $q = q_0 + \delta q$ and $\omega = \omega_0 + \delta\omega$; equation (35) becomes

$$\begin{aligned} Y + iX &= f(q + i\omega) \\ &= f\left\{(q_0 + \delta q) + i(\omega_0 + \delta\omega)\right\} \\ &= f\left\{(q_0 + i\omega_0) + (\delta q + i\delta\omega)\right\} \\ &= f(z_0 + \delta z) = f(z) \end{aligned}$$

The complex function $f(z)$ can be approximated by a Taylor's series (a power series)

$$f(z) = f(z_0) + \delta z\, f^{(1)}(z_0) + \frac{(\delta z)^2}{2!} f^{(2)}(z_0) + \frac{(\delta z)^3}{3!} f^{(3)}(z_0) + \cdots$$

where $f^{(1)}(z_0)$, $f^{(2)}(z_0)$, etc are first, second and higher order derivatives of the function $f(z)$ evaluated at the approximate location z_0. Choosing, as an approximate location, a point on the central meridian having the same isometric latitude as *P*, then $\delta q = 0$ (since $q = q_0 + \delta q$ and $q_0 = q$) and $\delta\omega = \omega$ (since $\omega = \omega_0 + \delta\omega$ and $\omega_0 = 0$), hence $z_0 = q_0 + i\omega_0 = q$ and $\delta z = \delta q + i\delta\omega = i\omega$. The complex function $f(z) = f(q + i\omega)$ can then be written as

$$f(q + i\omega) = f(q) \qquad + i\omega \frac{d}{dq} f(q) + \frac{(i\omega)^2}{2!} \frac{d^2}{dq^2} f(q) + \frac{(i\omega)^3}{3!} \frac{d^3}{dq^3} f(q) + \cdots$$

$\frac{d}{dq} f(q)$, $\frac{d^2}{dq^2} f(q)$, etc are first, second and higher order derivatives of the function $f(q)$. Noting that $i^2 = -1,\ i^3 = -i,\ i^4 = 1$, etc and $f(q) = k_0 m$, the complex mapping equation (35) may be written as

$$\begin{aligned} Y + iX = f(q + i\omega) = k_0 \Bigg\{ m\ & + i\omega \frac{dm}{dq} - \frac{\omega^2}{2!} \frac{d^2 m}{dq^2} - i\frac{\omega^3}{3!} \frac{d^3 m}{dq^3} + \frac{\omega^4}{4!} \frac{d^4 m}{dq^4} \\ & + i\frac{\omega^5}{5!} \frac{d^5 m}{dq^5} - \frac{\omega^6}{6!} \frac{d^6 m}{dq^6} - i\frac{\omega^7}{7!} \frac{d^7 m}{dq^7} + \frac{\omega^8}{8!} \frac{d^8 m}{dq^8} + \cdots \Bigg\} \end{aligned} \tag{40}$$

Equating the real and imaginary parts of equation (40) gives the projection equations in series form

$$\begin{aligned} X &= k_0 \left\{ \omega \frac{dm}{dq} - \frac{\omega^3}{3!} \frac{d^3 m}{dq^3} + \frac{\omega^5}{5!} \frac{d^5 m}{dq^5} - \frac{\omega^7}{7!} \frac{d^7 m}{dq^7} + \cdots \right\} \\ Y &= k_0 \left\{ m - \frac{\omega^2}{2!} \frac{d^2 m}{dq^2} + \frac{\omega^4}{4!} \frac{d^4 m}{dq^4} - \frac{\omega^6}{6!} \frac{d^6 m}{dq^6} + \frac{\omega^8}{8!} \frac{d^8 m}{dq^8} - \cdots \right\} \end{aligned} \tag{41}$$

To verify that the function $f(q + i\omega)$ given by equations (40) and (41) is analytic the derivatives are

$$\begin{aligned} \frac{\partial X}{\partial q} &= k_0 \left\{ \omega \frac{d^2 m}{dq^2} - \frac{\omega^3}{3!} \frac{d^4 m}{dq^4} + \frac{\omega^5}{5!} \frac{d^6 m}{dq^6} - \frac{\omega^7}{7!} \frac{d^8 m}{dq^8} + \cdots \right\} \\ \frac{\partial X}{\partial \omega} &= k_0 \left\{ \frac{dm}{dq} - \frac{\omega^2}{2!} \frac{d^3 m}{dq^3} + \frac{\omega^4}{4!} \frac{d^5 m}{dq^5} - \frac{\omega^6}{6!} \frac{d^7 m}{dq^7} + \cdots \right\} \end{aligned}$$

$$\frac{\partial Y}{\partial q} = k_0 \left\{ \frac{dm}{dq} - \frac{\omega^2}{2!}\frac{d^3m}{dq^3} + \frac{\omega^4}{4!}\frac{d^5m}{dq^5} - \frac{\omega^6}{6!}\frac{d^7m}{dq^7} + \cdots \right\}$$

$$\frac{\partial Y}{\partial \omega} = k_0 \left\{ -\omega \frac{d^2m}{dq^2} + \frac{\omega^3}{3!}\frac{d^4m}{dq^4} - \frac{\omega^5}{5!}\frac{d^6m}{dq^6} + \frac{\omega^7}{7!}\frac{d^8m}{dq^8} - \cdots \right\} \tag{42}$$

and these derivatives satisfy the Cauchy-Riemann equations

$$\frac{\partial Y}{\partial q} = \frac{\partial X}{\partial \omega} \qquad \text{and} \qquad \frac{\partial Y}{\partial \omega} = -\frac{\partial X}{\partial q}$$

Hence, the Cartesian coordinates X and Y given by equations (41) are a conformal transformation of the isometric parameters q and $\omega = \lambda - \lambda_0$ on the ellipsoid. It only remains for the derivatives $\frac{dm}{dq}, \frac{d^2m}{dq^2}$, etc to be evaluated for the projection coordinates to be fully defined.

The successive derivatives are obtained by considering the following:

(i) From the definition of isometric latitude $dq = \frac{\rho\, d\phi}{\nu \cos\phi} = \frac{d\phi}{V^2 \cos\phi}$ giving

$$\frac{d\phi}{dq} = V^2 \cos\phi$$

where $V^2 = \frac{\nu}{\rho} = 1 + e'^2 \cos^2\phi$ and $e'^2 = \frac{a^2 - b^2}{b^2} = \frac{e^2}{1 - e^2}$ is the second eccentricity squared

(ii) From the elemental rectangle (Figure 5) the meridian distance m is a function of the latitude ϕ, i.e., $dm = \rho\, d\phi$ and

$$\frac{dm}{d\phi} = \rho$$

(iii) From the *chain rule* for differentiation

$$\frac{dm}{dq} = \frac{dm}{d\phi}\frac{d\phi}{dq}$$

and the higher order derivatives are obtained by

$$\frac{d^2m}{dq^2} = \frac{d}{dq}\left(\frac{dm}{dq}\right), \quad \frac{d^3m}{dq^3} = \frac{d}{dq}\left(\frac{d^2m}{dq^2}\right), \quad \text{etc}$$

When evaluating the derivatives it is convenient to make the substitutions

$$\eta^2 = e'^2\cos^2\phi \quad \text{and} \quad t = \tan\phi \quad \text{hence} \quad V^2 = \frac{\nu}{\rho} = 1+\eta^2$$

The variables V^2, η^2, t and ν are all functions of the latitude ϕ and the differentiations given in (iii) above will, at some stage, require the following differentials for simplification

$$\frac{dV}{d\phi} = -\frac{\eta^2 t}{V}, \quad \frac{d\eta}{d\phi} = -\eta\, t, \quad \frac{dt}{d\phi} = 1+t^2, \quad \frac{d\nu}{d\phi} = \frac{\nu\,\eta^2 t}{V^2}$$

By repeated applications of the chain rule and algebra the derivates are found:

1st derivative

$$\begin{aligned}\frac{dm}{dq} &= \frac{dm}{d\phi}\frac{d\phi}{dq}\\ &= \rho\, V^2\cos\phi\\ &= \nu\cos\phi \quad \text{since } V^2 = \frac{\nu}{\rho}\end{aligned}$$

2nd derivative

$$\begin{aligned}\frac{d^2m}{dq^2} &= \frac{d}{dq}\left(\frac{dm}{dq}\right) = \frac{d}{d\phi}\nu\cos\phi\,\frac{d\phi}{dq} \qquad \text{(chain rule)}\\ &= \left\{\rho\left(-\sin\phi\right) + \cos\phi\,\frac{d\nu}{d\phi}\right\}V^2\cos\phi\\ &= \left\{-\rho\sin\phi + \cos\phi\,\frac{\nu\eta^2 t}{V^2}\right\}V^2\cos\phi\\ &= -\nu\, V^2\cos\phi\sin\phi + \nu\eta^2 t\cos^2\phi\end{aligned}$$

and the 2nd derivative becomes

$$\begin{aligned}\frac{d^2m}{dq^2} &= \nu\cos\phi\sin\phi\left(-V^2+\eta^2\right) \qquad \text{since } t = \tan\phi = \frac{\sin\phi}{\cos\phi}\\ &= -\nu\cos\phi\sin\phi \qquad \text{since } \mathrm{V}^2 = 1+\eta^2\end{aligned}$$

3rd derivative

$$\begin{aligned}\frac{d^3m}{dq^3} &= \frac{d}{dq}\left(\frac{d^2m}{dq^2}\right) = \frac{d}{d\phi}(-\nu\cos\phi\sin\phi)\frac{d\phi}{dq} \qquad\qquad \text{(chain rule)}\\ &= \left\{-\nu\cos^2\phi + \nu\sin^2\phi - \cos\phi\sin\phi\frac{\nu\eta^2 t}{V^2}\right\}V^2\cos\phi\\ &= \nu\cos^3\phi\left\{-\left(1+\eta^2\right) + t^2\left(1+\eta^2\right) - \eta^2t^2\right\}\\ &= -\nu\cos^3\phi\left(1+\eta^2-t^2\right)\end{aligned}$$

Higher order derivatives are found in a similar manner but with an almost exponential increase in algebra.

$$\frac{d^4m}{dq^4} = \nu\cos^3\phi\sin\phi\left(5 - t^2 + 9\eta^2 + 4\eta^4\right)$$

$$\frac{d^5m}{dq^5} = \nu\cos^5\phi\left(5 - 18t^2 + t^4 + 14\eta^2 + 13\eta^4 + 4\eta^6\right)$$

$$\begin{aligned}\frac{d^6m}{dq^6} = \nu\cos^5\phi\sin\phi\big(&-61 + 58t^2 - t^4 - 270\eta^2 - 445\eta^4 - 324\eta^6 - 88\eta^8\\ &+330t^2\eta^2 + 680t^2\eta^4 + 600t^2\eta^6 + 192t^2\eta^8\big)\end{aligned}$$

$$\begin{aligned}\frac{d^7m}{dq^7} = \nu\cos^7\phi\Big(&-61 + 479t^2 - 179t^4 + t^6 - 331\eta^2 - 715\eta^4 - 769\eta^6\\ &-412\eta^8 - 88\eta^{10} + 3298t^2\eta^2 + 8655t^2\eta^4 + 10964t^2\eta^6\\ &+6760t^2\eta^8 + 1632t^2\eta^{10} - 1771t^4\eta^2 - 6080t^4\eta^4\\ &-9480t^4\eta^6 - 6912t^4\eta^8 - 1920t^4\eta^{10}\Big)\end{aligned}$$

$$\begin{aligned}\frac{d^8m}{dq^8} = \nu\cos^7\phi\sin\phi\big(&1385 - 3111t^2 + 543t^4 - t^6 + 10899\eta^2 + 34419\eta^4\\ &+56385\eta^6 + 50856\eta^8 + 24048\eta^{10} + 4672\eta^{12}\\ &-32802t^2\eta^2 - 129087t^2\eta^4 - 252084t^2\eta^6 - 263088t^2\eta^8\\ &-140928t^2\eta^{10} - 30528t^2\eta^{12} + 9219t^4\eta^2 + 49644t^4\eta^4\\ &+121800t^4\eta^6 + 151872t^4\eta^8 + 94080t^4\eta^{10} + 23040t^4\eta^{12}\big)\end{aligned} \tag{43}$$

The derivatives given in equations (43) can be found embedded in equations (288), page 96 of *Conformal Projections in Geodesy and Cartography* (Thomas 1952) and the method of derivation outlined above is given in *Geodesy and Map Projections* (Lauf 1983).

Substituting these derivatives into equations (41) give expressions for the X and Y coordinates of a TM projection but it is useful to note that the coefficients of these derivatives may be quite small, e.g., for a TM zone 12° wide in longitude $\omega = 6^{\circ} = 0.104720$ radians the coefficients of the 7th and 8th derivatives are $\dfrac{\omega^7}{5040} = 2.740121\,\mathrm{E}^{-11}$ and $\dfrac{\omega^8}{40320} = 3.586810\,\mathrm{E}^{-13}$ respectively. Using these coefficient values, all the terms in the 7th and 8th derivatives involving powers of η $(\eta^2, \eta^4,$ etc) and powers of t and η combined $\left(t^2\eta^2, t^2\eta^4, \text{ etc}\right)$ were calculated, summed and then multiplied by the coefficients, for latitudes in one-degree intervals from the equator ($\phi = 0^\circ$) to $\phi = 75^\circ$. The maximum values amount to "errors" of 0.00040 metres at the equator for an X-coordinate and 0.00003 metres at $\phi = 14^\circ$ for a Y-coordinate. Hence all these terms in the 7th and 8th derivatives may be neglected without introducing any appreciable error in the coordinates. For the development of subsequent formulae the 7th and 8th derivatives are defined as equal to:

$$
\begin{aligned}
\frac{d^7 m}{dq^7} &\triangleq \nu \cos^7\phi\left(-61 + 479t^2 - 179t^4 + t^6\right) \\
\frac{d^8 m}{dq^8} &\triangleq \nu \cos^7\phi \sin\phi\left(1385 - 3111t^2 + 543t^4 - t^6\right)
\end{aligned}
\tag{44}
$$

A further simplification of the terms involving powers of η in the 3rd, 4th, 5th and 6th derivatives can be made with the substitution

$$
\psi = \frac{\nu}{\rho} = V^2 = 1 + \eta^2 \tag{45}
$$

that leads to expressions for the powers of η

$$
\begin{aligned}
\eta^2 &= \psi - 1 & \eta^6 &= \psi^3 - 3\psi^2 + 3\psi - 1 \\
\eta^4 &= \psi^2 - 2\psi + 1 & \eta^8 &= \psi^4 - 4\psi^3 + 6\psi^2 - 4\psi + 1
\end{aligned}
\tag{46}
$$

Substituting these into the derivatives and gathering powers of ψ gives the usual expressions for the X and Y coordinates of a TM projection

$$X = k_0 \left\{ \nu\omega\cos\phi + \nu\frac{\omega^3}{6}\cos^3\phi\left(\psi - t^2\right) \right.$$
$$+ \ \nu\frac{\omega^5}{120}\cos^5\phi\left[4\psi^3\left(1-6t^2\right)+\psi^2\left(1+8t^2\right)-\psi\left(2t^2\right)+t^4\right]$$
$$\left. + \ \nu\frac{\omega^7}{5040}\cos^7\phi\left(61-479t^2+179t^4-t^6\right)\right\} \tag{47}$$

$$Y = k_0 \left\{ m + \nu\sin\phi\frac{\omega^2}{2}\cos\phi + \nu\sin\phi\frac{\omega^4}{24}\cos^3\phi\left(4\psi^2+\psi-t^2\right) \right.$$
$$+ \nu\sin\phi\frac{\omega^6}{720}\cos^5\phi\begin{bmatrix}8\psi^4\left(11-24t^2\right)-28\psi^3\left(1-6t^2\right)\\ +\ \psi^2\left(1-32t^2\right)-\psi\left(2t^2\right)+t^4\end{bmatrix}$$
$$\left. + \nu\sin\phi\frac{\omega^8}{40320}\cos^7\phi\left(1385-3111t^2+543t^4-t^6\right)\right\} \tag{48}$$

These formula, commonly known in Australia as *Redfearn's* formula were published by J.C.B. Redfearn of the Hydrographic Department of the British Admiralty in the *Empire Survey Review* (now *Survey Review)* in 1948, (Redfearn 1948), who claimed "no special mathematical qualifications except, perhaps, that of sticking to what seemed at times to be a particularly tough spot of work." *Redfearn's* formula, equations (47) and (48), and equation (7) for meridian distance m were adopted by the National Mapping Council as "exact, and not the opening terms of an infinite series" for the purposes of computing AMG coordinates in Australia (NMC 1986).

The practical limits of use of Redfearn's formulae can be determined by calculating the maximum values of the 4th terms in equations (47) and (48) for one-degree intervals of latitude from the equator ($\phi = 0^\circ$) to $\phi = 75^\circ$ on the edge of a TM zone 12° wide in longitude, i.e., $\omega = \lambda - \lambda_0 = 6^\circ$. These were found to be at the equator for the X-coordinate and at approximately $\phi = 16^\circ$ for the Y-coordinate. Note that a UTM zone, by definition, is 6° wide in longitude, i.e., $\omega = \lambda - \lambda_0 = 3^\circ$.

```
TM Coordinates
ANS (a = 6378160, f = 1/298.25)

φ = 0° ω = 6°
X-coordinate (k0 = 1)
1st term =   667919.353314
2nd term =     1228.986723
3rd term =        3.410334
4th term =        0.010661
Sum    X =   669151.761032
```

```
φ = 16° ω = 6°
Y-coordinate (k0 = 1)
m        =   1769649.337353
1st term =      9268.572180
2nd term =        38.929264
3rd term =         0.152707
4th term =         0.000542
Sum    Y =   1778956.992047
```

Inspection of these values would seem to indicate that the "missing terms" in the truncated series would, in all likelihood, be at least an order of magnitude less than the 4th terms. We could be fairly confident that Redfearn's formulae are accurate to at least 1 mm on the edge of a TM zone 12° wide in longitude.

For a TM zone 6° wide in longitude, i.e., $\omega = \lambda - \lambda_0 = 3^\circ$ the maximum 4th term values are

```
TM Coordinates
ANS (a = 6378160, f = 1/298.25)

φ = 0° ω = 3°
X-coordinate (k0 = 1)
1st term =   333959.676657
2nd term =      153.623340
3rd term =        0.106573
4th term =        0.000083
Sum    X =   334113.406654
```

```
φ = 16° ω = 3°
Y-coordinate (k0 = 1)
m        =   1769649.337353
1st term =      2317.143045
2nd term =         2.433079
3rd term =         0.002386
4th term =         0.000002
Sum    Y =   1771968.915865
```

East and North Coordinates, False Origins and True Origins

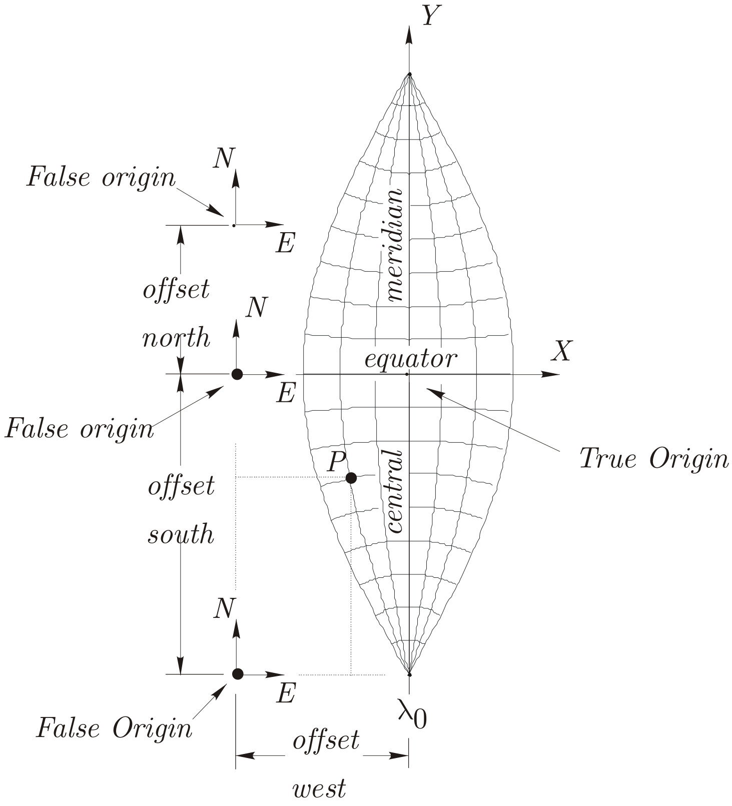

Figure 10. True and False origins of a TM projection

Redfearn's formula give X and Y coordinates of P relative to the *True Origin,* which is located at the intersection of the Equator and the Central Meridian. For P in the southern hemisphere and west of the Central Meridian, these coordinates will both be negative. To make all coordinates in a zone positive quantities, a new rectangular East-North (E,N) coordinate system is introduced with its origin, known as the *False Origin,* offset from the True Origin. East and North coordinates are then given by

$$E = X + \text{ west offset}$$
$$N = Y \begin{Bmatrix} -\text{ north offset} \\ +\text{ south offset} \end{Bmatrix} \tag{49}$$

Note: UTM zones are divided into Northern and Southern hemisphere portions. For a northern zone, the west offset is 500,000 metres and the north offset is zero, i.e., the False Origin lays on the Equator, 500,000 metres west of the True Origin. For a southern zone the west offset is 500,000 metres and the south offset is 10,000,000 metres.

Grid Convergence γ on a TM projection

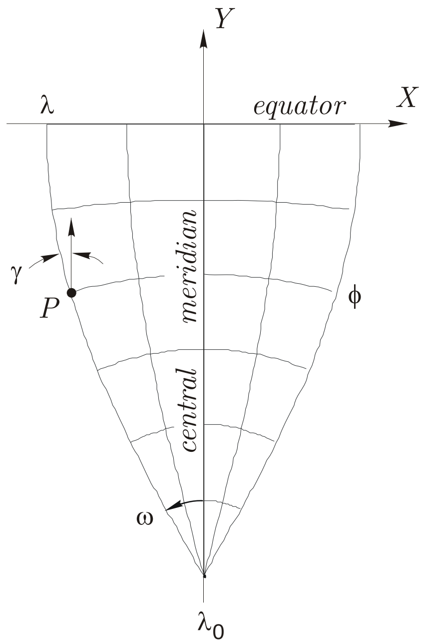

Figure 12. Grid Convergence

Grid convergence γ (gamma) at a point P is the angle between *True North,* the direction of the projected meridian through P, and *Grid North,* the direction of a line through P that is parallel with the central meridian.

Grid convergence is defined by the differential relationship

$$\tan\gamma = -\frac{dX}{dY} \tag{50}$$

As X and Y are both functions of the isometric latitude q and the longitude difference $\omega = \lambda - \lambda_0$ we may use the *Total Differential* theorem of calculus and write

$$\begin{aligned} dX &= \frac{\partial X}{\partial q}dq + \frac{\partial X}{\partial \omega}d\omega \\ dY &= \frac{\partial Y}{\partial q}dq + \frac{\partial Y}{\partial \omega}d\omega \end{aligned} \tag{51}$$

We can evaluate equation (50) using the differentials of (51) at a point of our choosing or in a direction of our choice. Choosing a direction along a meridian where $\omega = \lambda - \lambda_0$ is constant, hence $d\omega = 0$ in equations (51), give expressions for dX and dY, that when substituted into equation (50) give

$$\tan\gamma = -\frac{dX}{dY} = -\frac{\dfrac{\partial X}{\partial q}dq}{\dfrac{\partial Y}{\partial q}dq} = -\frac{\partial X}{\partial q}\left[\frac{\partial Y}{\partial q}\right]^{-1} \tag{52}$$

The partial derivatives $\partial X/\partial q$ and $\partial Y/\partial q$ are given in equations (42)

$$\begin{aligned}\frac{\partial X}{\partial q} &= -k_0\left\{\omega\frac{d^2m}{dq^2}-\frac{\omega^3}{3!}\frac{d^4m}{dq^4}+\frac{\omega^5}{5!}\frac{d^6m}{dq^6}-\frac{\omega^7}{7!}\frac{d^8m}{dq^8}+\cdots\right\}\\ \frac{\partial Y}{\partial q} &= k_0\left\{\frac{dm}{dq}-\frac{\omega^2}{2!}\frac{d^3m}{dq^3}+\frac{\omega^4}{4!}\frac{d^5m}{dq^5}-\frac{\omega^6}{6!}\frac{d^7m}{dq^7}+\cdots\right\}\end{aligned} \tag{53}$$

The differentials $\frac{dm}{dq}, \frac{d^2m}{dq^2}$ etc have already been evaluated and given in equations (43) [1st to 6th derivatives] and (44) [7th and 8th derivatives] hence

$$-\left(\frac{\partial X}{\partial q}\right) = -k_0\left\{\begin{aligned}&-\omega\nu\cos\phi\sin\phi\\ &-\frac{\omega^3}{6}\nu\cos^3\phi\sin\phi\left(5-t^2+9\eta^2+4\eta^4\right)\\ &+\frac{\omega^5}{120}\nu\cos^5\phi\sin\phi\left(-61+58t^2-\cdots\right)\\ &-\frac{\omega^7}{5040}\nu\cos^7\phi\sin\phi\left(1385-3111t^2-\cdots\right)\end{aligned}\right\} \tag{54}$$

$$\frac{\partial Y}{\partial q} = k_0\left\{\begin{aligned}&\nu\cos\phi\\ &+\frac{\omega^2}{2}\nu\cos^3\phi\left(1+\eta^2-t^2\right)\\ &+\frac{\omega^4}{24}\nu\cos^5\phi\left(5-18t^2+\cdots\right)\\ &-\frac{\omega^6}{720}\nu\cos^7\phi\left(-61+479t^2-\cdots\right)\end{aligned}\right\} \tag{55}$$

With common factors of $-\omega\nu\cos\phi\sin\phi$ and $\nu\cos\phi$ taken outside the braces the derivatives can be written as

$$\begin{aligned}-\left(\frac{\partial X}{\partial q}\right) &= k_0\omega\nu\cos\phi\sin\phi\left\{1+D\frac{\omega^2}{6}\cos^2\phi-E\frac{\omega^4}{120}\cos^4\phi+F\frac{\omega^6}{5040}\cos^6\phi\right\}\\ \frac{\partial Y}{\partial q} &= k_0\nu\cos\phi\left\{1+A\frac{\omega^2}{2}\cos^2\phi+B\frac{\omega^4}{24}\cos^4\phi-C\frac{\omega^6}{720}\cos^6\phi\right\}\end{aligned}$$

A,B,C,D,E,F are the respective coefficients in t and η from equations (43) and (44) noting that C and F contain no powers of η since they were shown to be negligible in the derivation of the projection equations. Substituting these expressions into equation (52) with $y=\omega^2\cos^2\phi$ gives

$$\tan\gamma = \omega\sin\phi\left\{1+D\frac{y}{6}-E\frac{y^2}{120}+F\frac{y^3}{5040}\right\}\left\{1+A\frac{y}{2}+B\frac{y^2}{24}-C\frac{y^3}{720}\right\}^{-1} \tag{56}$$

The reciprocal of the 2nd term in the product can be expanded into a sum of terms by using a special case of the Binomial series

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + x^4 - \cdots$$

Expanding equation (56) and ignoring powers greater than y^3 gives

$$\tan\gamma = \omega\sin\phi\left\{\begin{array}{l} 1 + (D-3A)\dfrac{y}{6} + \left(30A^2 - E - 5B - 10AD\right)\dfrac{y^2}{120} \\ + \left(F + 7C + 210AB + 21AE - 35BD + 210DA^2 - 630A^3\right)\dfrac{y^3}{5040} \end{array}\right\}$$

With the values for A,B,C,D,E,F from equations (43), (44) and simplifying gives

$$\tan\gamma = \omega\sin\phi\left\{\begin{array}{l} 1 + \dfrac{\omega^2\cos^2\phi}{3}\left(1 + t^2 + 3\eta^2 + 2\eta^4\right) \\ + \dfrac{\omega^4\cos^4\phi}{15}\left(\begin{array}{l} 2 + 4t^2 + 2t^4 + 15\eta^2 + 35\eta^4 + 33\eta^6 \\ + 11\eta^8 - 40t^2\eta^4 - 60t^2\eta^6 - 24t^2\eta^8 \end{array}\right) \\ + \dfrac{\omega^6\cos^6\phi}{315}\left(17 + 51t^2 + 51t^4 + 17t^6\right) \end{array}\right\} \tag{57}$$

Note that all coefficients involving η have been neglected in the last term of equation (57). This equation is essentially the same as Thomas 1952, p. 97, eqn 294).

To find an expression for γ we may use the series expansion for arc tan x

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$$

Hence

$$\gamma = \tan\gamma - \frac{\tan^3\gamma}{3} + \frac{\tan^5\gamma}{5} - \frac{\tan^7\gamma}{7} + \cdots \tag{58}$$

where

$$\tan\gamma = \omega\sin\phi\left(1 + a\frac{\omega^2\cos^2\phi}{3} + b\frac{\omega^4\cos^4\phi}{15} + c\frac{\omega^6\cos^6\phi}{315}\right) \tag{59}$$

and, from (57)

$$\begin{aligned} a &= \left(1 + t^2 + 3\eta^2 + 2\eta^4\right) \\ b &= \left(2 + 4t^2 + 2t^4 + 15\eta^2 + \cdots\right) \\ c &= \left(17 + 51t^2 + 51t^4 + 17t^6\right) \end{aligned}$$

Ignoring terms greater than ω^7 in any expansions we may write

$$
\begin{aligned}
-\frac{\tan^3\gamma}{3} &= -\frac{\omega^3\sin^3\phi}{3}\left(1+3a\frac{\omega^2\cos^2\phi}{3}+3b\frac{\omega^4\cos^4\phi}{15}+3a^2\frac{\omega^4\cos^4\phi}{9}\right)\\
\frac{\tan^5\gamma}{5} &= \frac{\omega^5\sin^5\phi}{5}\left(1+5a\frac{\omega^2\cos^2\phi}{3}\right)\\
-\frac{\tan^7\gamma}{7} &= -\frac{\omega^7\sin^7\phi}{7}
\end{aligned}
\tag{60}
$$

Remembering that $t^2=\tan^2\phi=\dfrac{\sin^2\phi}{\cos^2\phi}$ we may write $\sin^3\phi=\sin\phi\left(t^2\cos^2\phi\right)$ and equations (60) and (59) may be substituted into (58) and simplified to give

$$
\gamma=\omega\sin\phi\left\{\begin{aligned}&1+\frac{\omega^2\cos^2\phi}{3}\left(a-t^2\right)+\frac{\omega^4\cos^4\phi}{15}\left(b-5at^2+3t^4\right)\\&+\frac{\omega^6\cos^6\phi}{315}\left(c-21bt^2-35a^2t^2+105at^4-45t^6\right)\end{aligned}\right\}
$$

Substituting the values for $a,b,c,$ simplifying and ignoring all terms containing η in the last term in the braces above gives (Thomas 1952, p. 98, eqn 298)

$$
\gamma=\omega\sin\phi\left\{\begin{aligned}&1+\frac{\omega^2\cos^2\phi}{3}\left(1+3\eta^2+2\eta^4\right)\\&+\frac{\omega^4\cos^4\phi}{15}\left(\begin{aligned}&2-t^2+15\eta^2+35\eta^4+33\eta^6+11\eta^8\\&-15t^2\eta^2-50t^2\eta^4-60t^2\eta^6-24t^2\eta^8\end{aligned}\right)\\&+\frac{\omega^6\cos^6\phi}{315}\left(17-26t^2+2t^4\right)\end{aligned}\right\}
\tag{61}
$$

With the substitution $\eta^2=\psi-1$ [see equations (45) and (46)] into equation (61) and gathering powers of ψ, we have the usual expression for the grid convergence on a TM projection

$$
\gamma=\omega\sin\phi\left\{\begin{aligned}&1+\frac{\omega^2\cos^2\phi}{3}\left(2\psi^2-\psi\right)\\&+\frac{\omega^4\cos^4\phi}{15}\left[\begin{aligned}&\psi^4\left(11-24t^2\right)-\psi^3\left(11-36t^2\right)\\&+2\psi^2\left(1-7t^2\right)+\psi t^2\end{aligned}\right]\\&+\frac{\omega^6\cos^6\phi}{315}\left(17-26t^2+2t^4\right)\end{aligned}\right\}
\tag{62}
$$

The derivation of equation (62) shown here follows Thomas (1952). Lauf (1983) derives a smaller number of terms and Redfearn (1948) published equations (61) and (62) but without any derivation.

Equation (62) is used in Australia for computing grid convergence on a UTM projection (NMC 1986, ICSM 2003) but with a minus sign attached. This is to ensure that grid convergence is a positive quantity west of the central meridian (where $\omega = \lambda - \lambda_0$ is negative) and a negative quantity east of the central meridian (where $\omega = \lambda - \lambda_0$ is positive).

Point Scale Factor k on a TM projection

From the definition of scale factor given by equation (26) and with $f = F = 0$ (since the projection is conformal), $e = \rho^2$, $g = \nu^2 \cos^2\phi$ (Gaussian Fundamental Quantities for the ellipsoid) and $E = \left(\dfrac{dX}{d\phi}\right)^2$, $F = \left(\dfrac{dY}{d\lambda}\right)^2$ (Gaussian Fundamental Quantities for the projection where the parametric curves are functions of ϕ and λ only) we have

$$k^2 = \frac{dS^2}{ds^2} = \frac{E\,d\phi^2 + G\,d\lambda^2}{e\,d\phi^2 + g\,d\lambda^2} = \frac{(dX)^2 + (dY)^2}{\rho^2\,d\phi^2 + \nu^2\cos^2\phi\,d\lambda^2} \tag{63}$$

Using the expression for isometric latitude given by equation (38) we may write

$$d\phi^2 = \frac{\nu^2\cos^2\phi}{\rho^2}$$

and the expression for scale factor (in terms of isometric parameters) becomes

$$k^2 = \frac{(dX)^2 + (dY)^2}{\nu^2\cos^2\phi\left(dq^2 + d\lambda^2\right)} \tag{64}$$

where, as before

$$dX = \frac{\partial X}{\partial q}dq + \frac{\partial X}{\partial \omega}d\omega \quad \text{and} \quad dY = \frac{\partial Y}{\partial q}dq + \frac{\partial Y}{\partial \omega}d\omega \tag{65}$$

Since the projection is conformal, the scale factor is the same in any direction and equations (65) can be evaluated in a direction of our choice, say in a direction along a meridian where $\omega = \lambda - \lambda_0$ is constant hence $d\omega = 0$, and $dX = \left(\partial X/\partial q\right) dq$, $dY = \left(\partial Y/\partial q\right) dq$. Substituting these expressions into (64) gives

$$k^2 = \frac{1}{\nu^2 \cos^2 \phi} \left\{ \left(\frac{\partial X}{\partial q} \right)^2 + \left(\frac{\partial Y}{\partial q} \right)^2 \right\} \tag{66}$$

Now the derivatives $-\partial X/\partial q$, $\partial Y/\partial q$ are given in equations (54) and (55) and substituting these into (66) gives, after some simplification

$$k^2 = k_0^2 \left\{ \begin{aligned} & 1 + \omega^2 \cos^2 \phi \left(1 + \eta^2\right) \\ & + \frac{\omega^4 \cos^4 \phi}{3} \begin{pmatrix} 2 - t^2 + 5\eta^2 + 4\eta^4 + \eta^6 \\ -7t^2n^2 - 12t^2n^4 - 6t^2n^6 \end{pmatrix} \\ & + \frac{\omega^6 \cos^6 \phi}{45} \left(17 - 26t^2 + 2t^4\right) \end{aligned} \right\} \tag{67}$$

Equation (67) is in the general form $k^2 = k_0^2 (1 + x)$ and taking the square root of both sides of the equation gives the scale factor $k = k_0 \sqrt{(1 + x)} = k_0 (1 + x)^{1/2}$. Using a special case of the Binomial series, the square root $(1 + x)^{1/2}$ is given by

$$(1 + x)^{1/2} = 1 + \frac{1}{2} x - \frac{1}{2 \cdot 4} x^2 + \frac{1 \cdot 3}{2 \cdot 4 \cdot 6} x^3 - \cdots$$

and after some algebra

$$k = k_0 \left\{ \begin{aligned} & 1 + \frac{\omega^2 \cos^2 \phi}{2} \left(1 + \eta^2\right) \\ & + \frac{\omega^4 \cos^4 \phi}{24} \begin{pmatrix} 5 - 4t^2 + 14\eta^2 + 13\eta^4 + 4\eta^6 \\ -28t^2n^2 - 48t^2n^4 - 24t^2n^6 \end{pmatrix} \\ & + \frac{\omega^6 \cos^6 \phi}{720} \left(61 - 148t^2 + 16t^4\right) \end{aligned} \right\} \tag{68}$$

With the substitution $\eta^2 = \psi - 1$ [see equations (45) and (46)] into equation (68) and gathering powers of ψ , we have the usual expression for the point scale factor on a TM projection

$$k = k_0 \left\{ \begin{array}{l} 1 + \dfrac{\omega^2 \cos^2 \phi}{2}(\psi) \\ + \dfrac{\omega^4 \cos^4 \phi}{24}\left\{4\psi^3\left(1 - 6t^2\right) + \psi^2\left(1 + 24t^2\right) - 4\psi\left(t^2\right)\right\} \\ + \dfrac{\omega^6 \cos^6 \phi}{720}\left(61 - 148t^2 + 16t^4\right) \end{array} \right\} \tag{69}$$

SUMMARY

This paper has provided a detailed outline of the derivation of Redfearn's equations. Some of the calculus (and algebra) has been shown but mostly it is hidden from view. In the past, this all would have been done using pencil and paper, employing skills developed by practice. Now, with the aid of mathematical computer packages such as MAPLE, the hard work of algebra can be reduced to a series of computer commands, revealing more clearly the connection between theories and working formulae. All the results in this paper (the TM formulae for coordinates, grid convergence and point scale factor) have been developed using MAPLE, but the method of derivation is one that would have been applied in the past. It is hoped that this paper may be of some use to those who wish to demonstrate the power of mathematical computer packages versus the traditional methods of solution.

In addition, it is hoped that this paper will supplement the excellent technical manuals available to practitioners in Australia.

REFERENCES

Alder, Ken, 2002. *The Measure of All Things,* Little, Brown, London.

BG, 1988. The Geodesist's Handbook 1988, *Bulletin Géodésique,* Vol. 62, No. 3.

Britannica® CD 99. *Encyclopaedia Britannica CD 99 Multimedia Edition,* International Version 1999, © 1994-1999 Encyclopaedia Britannica, Inc., Chicago, Illinois.

Gerdan, G.P. and R.E. Deakin, 1999. 'Transforming Cartesian coordinates X,Y,Z to geographical coordinates φ,λ,h', *The Australian Surveyor,* Vol. 44, No. 1, pp. 55-63.

ICSM, 2003. *Geocentric Datum of Australia Technical Manual* – Version 2.2, Intergovernmental Committee on Surveying & Mapping (ICSM), February 2002, available online at:
www.icsm.gov.au/icsm/gda/gdatm/index.html
(last accessed 25th February, 2004)

Lambert, J.H., 1772. *Beiträge zum Gebrauche der Mathematik und deren Anwendung* (Contributions for the Use of Mathematics and their Application), Part III, section 6, Anmerkungen und Zusätze zur Entwerfung der Land-und Himmelscharten, Berlin: Translated and introduced by W.R. Tobler, University of Michigan, 1972. (Citation from Snyder, 1987)

Land Vic, 2003. *The Map Grid of Australia 1994* – A Simplified Computation Manual, Land Victoria, Department of Sustainability and Environment, State Government of Victoria, 2003, available online at:
www.land.vic.gov.au/4A2569E6000982D8/frameset/geodesy?OpenDocume
nt
(last accessed 25th February, 2004)

Lauf, G.B., 1983. *Geodesy and Map Projections,* 2nd edn, TAFE Publications Unit, Collingwood, Victoria, Australia.

NIMA, 2000. *Department of Defense World Geodetic System 1984,* Technical Report TR8350.2, 3rd edn, Amendment 1, 3 January 2000, National Imagery and Mapping Agency.

NMC, 1986. *The Australian Geodetic Datum Technical Manual,* Special Publication 10, National Mapping Council of Australia, Canberra.

Redfearn, J.C.B., 1948. 'Transverse Mercator formula', *Empire Survey Review,* No. 69, Vol. IX, pp. 318-22.

Snyder, J.P., 1987. *Map Projections – A Working Manual,* US Geological Survey Professional Paper 1395, United States Department of Interior, Washington.

Sokolnikoff, I.S. and Redheffer, R.M., 1966. *Mathematics of Physics and Engineering,* 2nd edn, International Student Edition, McGraw-Hill, Inc., London.

Steed, J, (1996). 'The Geocentric Datum of Australia', *Azimuth,* March 1996, pp. 24-26 and April 1996, p. 23.

Thomas, P.D., 1952. *Conformal Projections in Geodesy and Cartography,* Special Publication No. 251, Coast and Geodetic Survey, United States Department of Commerce, Washington.