THE ANALYSIS OF THE EARTH'S GRAVITY FIELD

R. S. MATHER

JANUARY 1971

SCHOOL OF SURVEYING,
UNIVERSITY OF NEW SOUTH WALES,
KENSINGTON, N.S.W. AUSTRALIA.

THE ANALYSIS

OF THE

EARTH'S GRAVITY FIELD

BY

R.S. MATHER

January 1971

School of Surveying
University of New South Wales
P.O.Box 1
Kensington, N.S.W. 2033
Australia

PRINTED AND BOUND BY RM PRINTING PTY. LTD., SYDNEY

TABLE OF CONTENTS

Table of contents (ctd) - p.3

PREFACE

The contents of this monograph form the basis for a course of lectures in the subject *Dynamic Geodesy* given as a part of the Master of Surveying Science degree in the School of Surveying, University of New South Wales.

The development is biased towards the determination, in practice, of the best representation possible for an incompletely surveyed gravity anomaly field at the surface of the earth. It is also intended to be completely self contained on the understanding that the reader has a reasonable knowledge of basic mathematical techniques.

The relevant topics in satellite geodesy are developed with the intention of providing the reader with a good understanding of the techniques involved and the sources of likely weakness in the results which are to be combined with surface gravity data to provide the continuous surface anomaly data set. It is not intended to provide an introduction to satellite geodesy which can be obtained from one or other of the references dealing with the subject and listed in the bibliography.

The writer will be grateful if any errors are brought to his attention.

A GUIDE TO NOTATION

1. COMMONLY USED NOTATION

Conventions

y_α = The variables y_1 and y_2.

y_i = The variables y_1, y_2 and y_3, unless otherwise specified.

$y \overset{(1.31)}{=} bx$ ≡ By the use of equation 1.31, y can be shown to be equal to bx.

$\dot{x}$ = dx/dt.

ss = Scale subscript. If subscript out of range, subtract maximum possible value. E.g., if i > 3, subtract 3.

Summation convention

$$y_\alpha z_\alpha = \sum_{\alpha=1}^{2} \alpha\ \alpha = y_1 z_1 + y_2 z_2.$$

$$y_i z_i = \sum_{i=1}^{3} y_i z_i = y_1 z_1 + y_2 z_2 + y_3 z_3.$$

The range of subscripts will be specified in instances where it exceeds the above values.

$y_\alpha = A + Bx_\alpha$ ≡ There are two equations obtained on assigning the values 1 and 2 for the subscripts.

$y_i = A + Bx_i$ ≡ There are three equations obtained on assigning the values 1, 2 & 3 for the subscripts.

Vector representation

Italics, with one exception will be used solely for the representation of vectors in equations.

$X_i = X_i i$ ≡ The vector with components X_1, X_2 & X_3 along the rectangular Cartesian co-ordinate axes which are represented by the unit vectors *1*, *2* & *3*.

The exception is the case

$i = \sqrt{(-1)}$.

Such use is clearly defined in the text.

Other notation

RS 1967= Reference System 1967.

2. SOME FREQUENTLY USED SYMBOLS

A	= Azimuth.
a	= Equatorial radius of the reference ellipsoid.
b	= Polar radius of the reference ellipsoid.
C	= Coefficient of spherical/surface harmonic series.
c	= Centre-to-focus distance for the family of confocal ellipsoids.
d	= The separation vector.
d	= The modulus of the separation vector.
dX	= A differential increment in X.
dS	= An element of surface area.
$d\sigma$	= The element of surface area on a unit sphere.
E	= Eccentric anomaly.
e	= Eccentricity of the meridian ellipse.
f	= Flattening of meridian ellipse; true anomaly(section 8).
g	= The value of observed gravity at the surface of the earth.
h	= Orthometric elevation.
h_d	= Height anomaly.
h_i	= Linearisation parameter,
h_n	= Normal elevation.
h_s	= Spheroidal elevation.
I(X)	= The non-real part of the complex expression X.
I_a	= Inertia tensor of order a.
i	= Inclination of orbital ellipse.
i	= $\sqrt{(-1)}$.
k	= Gravitational constant $\equiv$ $6.673\times10^{-8}cm^3gm^{-1}sec^{-2}$.
M(X)	= The mean value of X.
M	= The mass of the earth; Mean anomaly(section 8).
m	= Mass(section 1 & 2);order of spherical harmonic(as subscript).
m	= $a\omega^2/\gamma_e$.
m'	= $a^3\omega^2/kM$.
n	= The degree of a spherical harmonic (as subscript).
n_α	= The mean motion of a satellite.
$P_{nm}(\mu)$	= Associated Legendre function (section 3.7).
$Q_{nm}(\mu)$	= Associated Legendre function of the second kind (section 3.4).
$q_2(\alpha)$	= Legendre function of the second kind.
R(X)	= The real part of the complex expression X.
R	= The Radius of the earth, assumed to be a sphere.
r	= Distance between the element of surface area dS ($d\sigma$) and the point of computation P; distance from geocentre; a spherical co-ordinate.
S	= Coefficient of spherical/surface harmonic series.
t	= Time.
U	= Potential of the reference system (sphero-potential).
u	= Reduced (parametric) latitude defining position on the meridian ellipse.

SOME FREQUENTLY USED SYMBOLS (ctd)

W = Potential of the earth (or Geopotential).

w = Weight coefficient.

X_i = A geocentric Cartesian co-ordinate system.

x_i = A general Cartesian co-ordinate system.

α = One of the spheroidal set of co-ordinates. $\alpha = \sin^{-1} e$.

β = The second of the spheroidal co-ordinate set $\tan \beta = \cos u$.

γ = Normal gravity; the global mean value of normal gravity.

Δx = A small change in x.

Δg = The gravity anomaly at the surface of the earth which, to the order of the flattening, is the free air anomaly.

δ_{ik} = Kronecker delta, $\delta_{ik} = 0$ if $i \neq k$; $\delta_{ik} = 1$ if $i = k$.

ζ = Deflection of the vertical.

θ = Co-latitude; Greenwich Sidereal Time (G.S.T. in section 8).

λ = Longitude, positive east.

μ = $\sin \phi = \cos \theta$; Also $\mu = kM$ (in section 8).

ν = Radius of curvature of ellipsoid in the prime vertical normal section.

ξ_i = Set of curvilinear parameters, on occasion defining the separation vector.

$\prod_{i=1}^{3} X_i = X_1 X_2 X_3$.

ρ = Radius of curvature of ellipsoid in the meridian normal section; density of gravitating material.

$\sum_{i=1}^{4} X_i = X_1 + X_2 + X_3 + X_4$.

σ = Surface area on a sphere of unit radius.

σ_n = Degree variance of gravity anomalies.

ϕ = Latitude, positive north.

ψ = Angular distance on a sphere between the element of surface area dS and the point of computation P.

Ω = Right ascension of ascending node.

ω = Angular velocity of rotation of the earth; the argument of perigee (section 8)

N = Unit normal vector

$$\nabla = \sum_{i=1}^{3} \frac{\partial}{\partial x_i} i .$$

3. TERMINOLOGY

The term *reference ellipsoid* is used to describe an oblate spheroid, obtained by rotating an ellipse about its minor axis. The term spheroid is used in this sense in section 5.

4. SUBSCRIPTS

Subscripts which are not indices, are introduced with the intention of improving the comprehension of concepts by keeping the number of basic variable names down to a minimum and hence simplifying the *written* form of the equations.

a	Astronomically determined values; astro-geodetic values.
c	Correction to the free air geoid.
e	Value at the equator; referring to the earth ellipsoid.
f	Free air geoid values.
g	Gravimetric values; geocentric values
h	Constants defining the family of hyperboloids of revolution constituting the spheroidal system of co-ordinates.
o	Values at the geoid or spheroid of reference
p	Evaluated at the fixed point P.
r	Rotational component; radial component in section 8.
s	Values referred to the reference ellipsoid.
θ	Component perpendicular to the direction of radius.

THE ANALYSIS OF THE EARTH'S GRAVITY FIELD

BY

R.S. MATHER

1. THE REQUIREMENTS OF PHYSICAL GEODESY

1.1 Introduction

The term *gravimetric geodesy* is used to differentiate that section of physical geodesy which deals with the definition of the physical surface of the earth in relation to a reference system by the use of observed values of the acceleration due to gravity (g). This excludes astro-geodetic methods which are based on a knowledge of the direction of the vertical instead of measured values of g. Such a classification must necessarily be simplistic. Geodesy is the study of techniques for the establishment of three dimensional co-ordinates in earth space, which is the Euclidian space that has the same rotational and galactic motion as the earth. True earth space co-ordinates are essentially four-dimensional, but for short term practical purposes, the variations in the fourth co-ordinate(time) can be assumed to be negligible.

Gravimetric geodesy cannot be considered to be a true geometrical technique if considered independently of the other branches of geodesy as it is concerned with properties of the earth's gravitational field which is a consequence of a distribution of mass given by an equation of the form

$$dm = dm\ (x_i) \qquad \ldots(1.1)$$

that is an invariant in earth space in the short term, the space being specified with reference to an x_i axis system. The term invariant is used to describe a quantity whose magnitude is independent of the reference frame adopted. This mass distribution gives rise to a Newtonian gravitational potential or geopotential W which is also an earth space invariant, being a scalar (tensor of zero order). g itself is a first order tensor (i.e., a vector) and is obtained from the gravitational potential. This vector field is incapable of definition unless the mass distribution at equation 1.1 is completely specified, this being physically impossible. The current practice which is adopted in gravimetric geodesy to surmount this obstacle calls for the definition of a model for this mass distribution and the computation of the earth space vector between equivalent points on the true and reference systems being called the separation vector d.

As the characteristics of the mass distribution on the reference system are specified, it is possible to convert the zero order gravitational tensor or spheropotential U and its associated first order tensor to earth space co-ordinates. Gravimetric geodesy concentrates on the determination of the separation vector which, together with the characteristics of the reference model would specify earth space position *if* the zero order gravitational tensor or geopotential of at least one point on the surface of the earth were known. This information cannot be obtained by the techniques of gravimetric geodesy alone. On the other hand, these techniques provide a method for the determination of earth space co-ordinates which is totally independent of the vagaries of the earth's atmosphere.

The capacity of gravimetric geodesy to define earth space position is therefore dependent on the following points.

(i) Values of g must be defined at "all points" on the earth's surface

(ii) The parameters adopted for the definition of the reference model and hence the values of normal gravity γ at all points exterior to the bounding surface must bear an adequate resemblance to reality

(iii) A value should be available, at least in theory, for the geopotential at a point on the physical surface of the earth.

The third point will not be directly covered in the present development

1.2 The data requirements for computations in gravimetric geodesy

The separation vector d between equivalent points as shown in figure 1.1 can be completely represented by three curvilinear components ξ_i. It could also be referred to a local rectangular Cartesian co-ordinate system x_i, when the modulus d of d can be given by

$$d^2 = h_i^2 d\xi_i^2 \qquad \ldots(1.2)$$

where h_i are the associated linearization parameters given by

$$h_i = h_i(x_j) \qquad \ldots(1.3),$$

x_j being the co-ordinates of Q with respect to a reference co-ordinate system. The required quantities ξ_i are given by (e.g, Heiskanen &

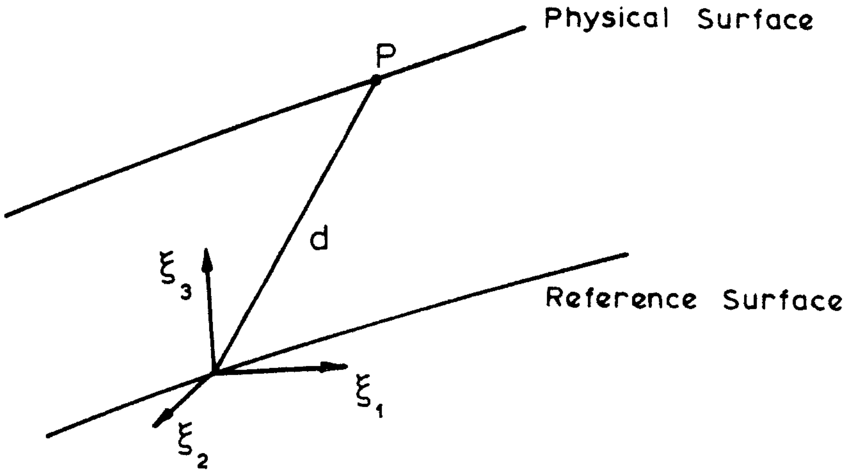


FIG. 1.1
The separation vector

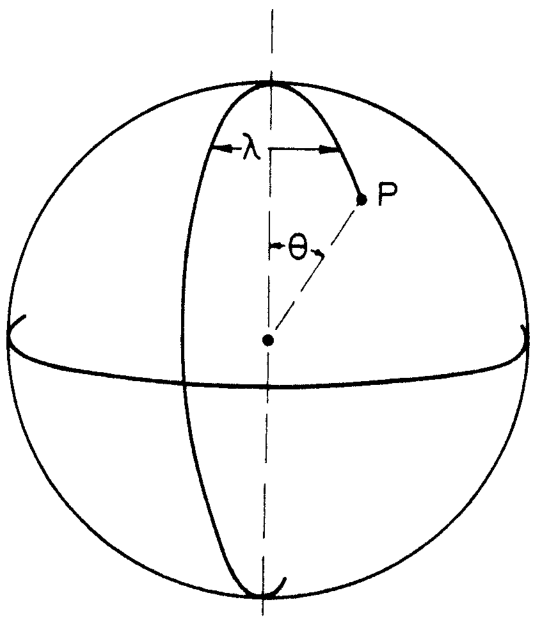


FIG. 1.2
The spherical system of co-ordinates

Moritz 1967,p.121; Moritz 1968,p.326; Mather 1970b,p.56) equations of the form

$$\xi_{ip} = K \iint f(r) \left(\Delta g + f(h,h_p)\right) ds \qquad \ldots(1.4),$$

where the subscript $_p$ refers to evaluation at the general point P on the surface S of the earth, h being the elevation of the topography at the element of surface area dS which is at a distance r from P. In the above equation $f(x) \equiv$ a function of x. The gravity anomaly Δg at the surface of the earth is given by

$$\Delta g = g - \gamma \qquad \ldots(1.5),$$

where g is the value of gravity observed at the physical surface and γ the value of normal gravity at the associated normal surface. Such pairs of surfaces can be afforded by any of the following systems.

(i) The physical surface of the earth and the telluroid (e.g., Heiskanen & Moritz 1967,p.292 ; Mather 1968b,p.517).

(ii) The geoid and the reference ellipsoid (e.g., ibid, p.522).

(iii) Any geop and its associated spherop (e.g., Moritz 1966).

The value of the gravity anomaly is thus dependent on the system being used in the definition of the separation vector, though the differences are in the third significant figure in non-mountainous areas. These discrepancies arise in "reducing" the values of gravity observed at the physical surface of the earth to the appropriate physical surface, the reductions Δg_c being of the form

$$\Delta g_c = \iint f(h,h_p)\ f(r)\ dS \qquad \ldots(1.6).$$

In all instances of the solution of the boundary value problem in physical geodesy it is therefore necessary to define the value of gravity and hence, the gravity anomaly, at all points on the earth's surface.

The surface gravity data available at the present time hardly meets the requirements of a complete coverage, there being considerable extents of the oceans which have not been adequately surveyed. Fortunately, this does not pose insurmountable problems in effecting determinations of the separation vector from a practical point of view as f(r) is primarily of the form

$$f(r) \doteqdot r^{-1} \quad \text{or} \quad r^{-2}$$

which means that distant zones have significantly smaller influences per unit area on computations than closer regions. Consequently, the former can be represented by mean values over larger areas than near ones when evaluating integrals of the type at 1.4 by quadratures. For an example of subdivisions for practical use, see (Mather 1969,p.501).

The following points are implicit in the adoption of the above procedures.

(a) f(r) in equation 1.4 is a linear function over the area being represented by the mean to the required order of accuracy.

(b) The pattern of correlation between variations of individual readings in the square from the mean, and position with respect to the point of computation P, as exists over a significant number of such squares should be negligible.

For an analytical treatment see (Mather 1968a, pp.156-163). The procedure could be extended further by replacing the set of discrete values Δg which are functions of position on the surface S by convenient continuous functions. As each of the three systems at (i) to (iii) on page 4 is an ellipsoid of revolution or a sphere to order f, while the last two are ellipsoids to order f^2, the most obvious choice of parameters for the definition of such continuous functions are sets of either spherical or spheroidal co-ordinates.

1.3 Suitable sets of co-ordinates

Problems in gravimetric geodesy, as in any other mathematical subject, are soluble with minimum complexity if an appropriate reference frame is adopted. The earth's gravitational field is well represented by a system of co-ordinates of the type ξ_i, where ξ_3 defines a surface which is preferably one of equal potential, while ξ_1 and ξ_2 define position on the surface

$$\xi_3 = \text{constant}.$$

Such a system has been studied in detail by Hotine(1969,p.69 et seq.). Its adoption as a reference frame poses certain problems as the definition of position in earth space by a family of geometrically irregular equipotential surfaces involves a quantity which is non-linear and not without dimension as the ξ_3 co-ordinate. If this quantity had the dimensions of potential, changes $\Delta\xi_3$ in ξ_3 could be converted to a linear equivalent $h_3\Delta\xi_3$ only if the mass distribution contributing to the potential were known, either implicitly or explicitly. It is therefore preferable to define general curvilinear co-ordinates with purely geometrical significance and introduce physical characteristics at a subsequent stage.

(i) Spherical co-ordinates

The position of the general point P in such a system is defined by one of the concentric spheres of radius r and centred at the

the origin O which passes through P. The exact position of P on the sphere is fixed by two surface parameters λ and ϕ, as shown in figure 1.2. Thus

$$\xi_1 = \lambda \quad ; \quad \xi_2 = \phi \quad ; \quad \xi_3 = r \qquad \ldots(1.7).$$

ξ_1 and ξ_2 will generally be assumed to be surface parameters while ξ_3 defines the family of surfaces extending through the space.

(ii) Spheroidal co-ordinates

In the following development, it will be assumed that the reader is either familiar with the co-ordinate geometry of the ellipse or has access to a standard text on the subject. An oblate spheroid (ellipsoid of revolution) is obtained by rotating the ellipse

$$\frac{x^2}{a^2} + \frac{x_3^2}{b^2} = 1 \qquad \ldots(1.8)$$

about its minor axis Ox_3 in figure 1.3. The equatorial and polar radii (a,b) of the meridian ellipse are related to the eccentricity and the flattening f by the relations

$$b = a[1-e^2]^{\frac{1}{2}} \qquad \ldots(1.9)$$

$$\frac{a-b}{a} = f \qquad \ldots(1.10)$$

$$e^2 = 2f - f^2 \qquad \ldots(1.11).$$

The direct extension of the principles adopted in section 1.3*i* for the spherical case to the spheroidal one is not possible as two parameters (a,e) are required to define the oblate spheroid, while only one (r) is required to define the sphere.

The space can, however, be totally defined by a family of confocal oblate spheroids. Such a system is scaled by the constant distance c between the origin O and the common focus F, where

$$c = ae \qquad \ldots(1.12)$$

If OB is the minor axis of the ellipse, B being on the ellipse, let

$$O\hat{B}F = \alpha .$$

As the sum of the distances from any point on the ellipse to the two foci is 2a, and as B is symmetrical with respect to the two foci,

$$BF = a.$$

The use of equation 1.12 gives

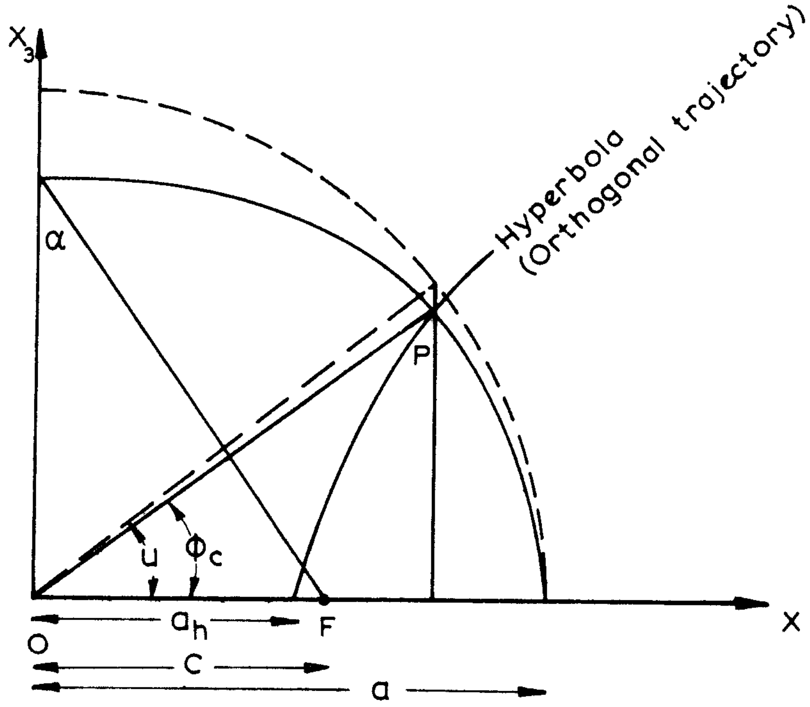


FIG. 1. 3

The oblate spheroidal system of co-ordinates

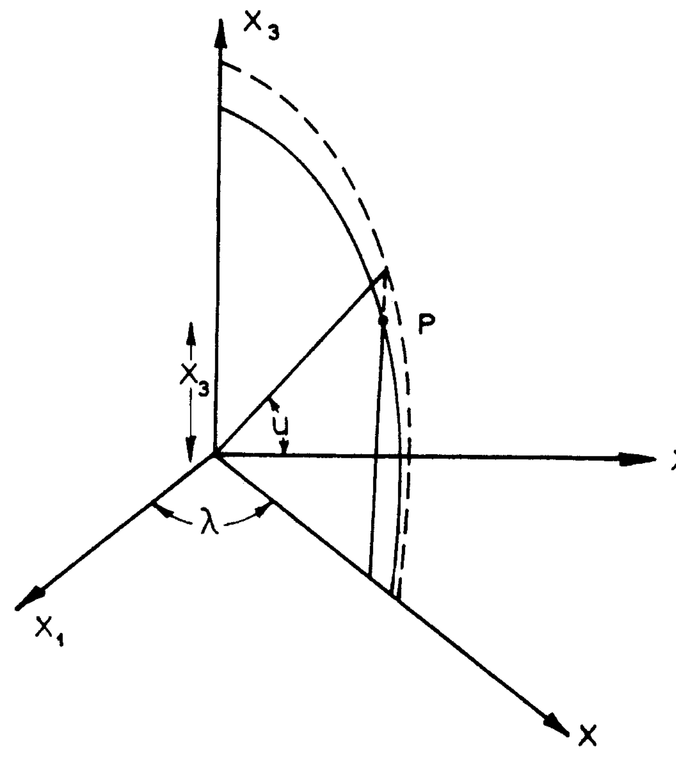


FIG. 1. 4

The three dimensional cartesian system
and oblate spheroidal co-ordinates

$$\sin \alpha = \frac{c}{a} = e \qquad \ldots(1.13).$$

Thus a family of confocal spheroids can be completely defined by the parameter α just as a family of concentric spheres can by the variable radius r. Position on the surface of the confocal spheroids can be defined by two surface parameters as in the case of the spherical system. These parameters are

$$\xi_1 = \lambda \quad ; \quad \xi_2 = u \ ,$$

where u is the parametric(or reduced) latitude defining position in the meridian ellipse, while λ has the same significance as in the previous case. The use of the parametric latitude has the advantage that

$$x = a \cos u \quad ; \quad x_3 = b \sin u \qquad \ldots(1.14).$$

Thus the adoption of the set of curvilinear co-ordinates

$$\xi_1 = \lambda \quad ; \quad \xi_2 = u \quad ; \quad \xi_3 = \alpha \qquad \ldots(1.15)$$

permits a complete definition of all points in the space. The equivalent three dimensional Cartesian co-ordinates with respect to an x_i axis system, where the plane of zero longitude λ is defined by the x_1x_3 plane and the equator by the x_1x_2 plane, are given by

$$x_1 = a \cos u \cos \lambda \ ,$$

$$x_2 = a \cos u \sin \lambda \ ,$$

and

$$x_3 = b \sin u \ .$$

Complete definition is obtained on replacing a and b in terms of α, using equations 1.12 and 1.13, when

$$a = c\, e^{-1} = c \operatorname{cosec} \alpha \qquad \ldots(1.16)$$

and

$$b = a[1 - e^2]^{\frac{1}{2}} = c \operatorname{cosec} \alpha(1 - \sin^2\alpha)^{\frac{1}{2}} = c \cot \alpha .. (1.17).$$

Thus the complete set of expressions for three dimensional Cartesian co-ordinates in terms of spheroidal co-ordinates is

$$\begin{aligned} x_1 &= c \operatorname{cosec} \alpha \cos u \cos \lambda \\ x_2 &= c \operatorname{cosec} \alpha \cos u \sin \lambda \qquad \ldots(1.18). \\ x_3 &= c \cot \alpha \sin u \end{aligned}$$

It can be seen from equation 1.18 that the elimination of λ and u gives

$$\sin^2 u + \cos^2 u = \frac{x_1^2 + x_2^2}{c^2 \operatorname{cosec}^2 \alpha} + \frac{x_3^2}{c^2 \cot^2 \alpha} = 1 \qquad \ldots(1.19)$$

which is the equation of a family of confocal oblate spheroids whose equatorial and polar radii are given by

$$a = c \operatorname{cosec} \alpha \quad ; \quad b = c \cot \alpha .$$

Similarly, the elimination of λ and α gives

$$\operatorname{cosec}^2\alpha - \cot^2\alpha = \frac{x_1^2 + x_2^2}{c^2 \cos^2 u} - \frac{x_3^2}{c^2 \sin^2 u} = 1 \qquad \ldots(1.20),$$

which is the equation of a family of confocal hyperboloids of revolution, the parameters defining the basic hyperbola being

$$a_h = c \cos u \quad ; \quad b_h = c \sin u \qquad \ldots(1.21).$$

The elimination of u and α gives

$$\tan \lambda = \frac{x_2}{x_1} \qquad \ldots(1.22).$$

Thus position as defined on a spheroidal system of co-ordinates can be visualised as being fixed as follows:-

(a) location in a plane given by the longitude λ;

and (b) position in the plane being specified by the intersection of an ellipse ($\alpha = \sin^{-1} e$) and a hyperbola ($u = \sec^{-1} e_h$), as the eccentricity e_h of the hyperbola

$$\frac{x^2}{a_h^2} - \frac{x_3^2}{b_h^2} = 1$$

is given by $a_h^2 \, [e_h^2 - 1] = b_h^2$

and the use of equation 1.20 gives

$$e_h = \frac{c}{c \cos u} = \sec u \qquad \ldots(1.23).$$

The equation of the tangent to the hyperbola at P(a cos u, b sin u) is

$$\frac{x}{a_h^2} a \cos u - \frac{x_3}{b_h^2} b \sin u = 1$$

or

$$\frac{x}{c^2}\left(\frac{a}{a_h}\right) - \frac{x_3}{c^2}\left(\frac{b}{b_h}\right) = 1.$$

The equation of the normal to the ellipse at P is

$$a x \sec u - b x_3 \operatorname{cosec} u = a^2 - b^2 .$$

The use of equation 1.21 gives

$$a x \sec u - b x_3 \operatorname{cosec} u = c^2 = a^2 - b^2 .$$

Hence the hyperbola is a normal trajectory to the ellipse and the family of confocal hyperboloids of revolution are orthogonal to the confocal oblate spheroids, both having a common centre and focus.

(iii) The conversion of spheroidal co-ordinates to spherical co-ordinates.

If the centre of the concentric spheres is the same as the common centre of the confocal system, it can be seen from rotational symmetry that

$$\xi_3 = \lambda$$

is unchanged on transformation between systems. The co-ordinates in any meridian section on the spherical system are the equivalent geocentric co-ordinates (r, ϕ_g) as shown in figure 1.3. If the spherical co-ordinates are as defined in equation 1.7, it can be seen from a consideration of figure 1.3 and the general geometry of the ellipse that

$$x = \nu \cos \phi = a \cos u = r \cos \phi_g$$

$$x_3 = \nu[1 - e^2]\sin \phi = b \sin u = r \sin \phi_g \qquad \ldots(1.24),$$

where

$$\nu = PF = \frac{a}{[1- e^2\sin^2\phi]^{\frac{1}{2}}} \qquad \ldots(1.25).$$

Thus

$$\tan \phi_g = [1 - e^2]^{\frac{1}{2}} \tan u = [1 - e^2]\tan \phi \qquad \ldots(1.26)$$

or

$$\tan \phi_g = \cos \alpha \tan u \qquad \ldots(1.27).$$

Substitution from equation 1.24 into equation 1.8 gives

$$r = \frac{b}{1 - e^2\cos^2\phi_g} .$$

The use of equations 1.13 and 1.27 gives

$$r = c \cot \alpha \left(1 - \frac{\sin^2\alpha}{1 + \cos^2\alpha \tan^2 u}\right)^{-1} \qquad \ldots(1.28).$$

As $\sin \alpha$ is a small quantity,

$$r = c \cot \alpha \left\{1 + \frac{\sin^2\alpha}{1+\cos^2\alpha \tan^2 u} + \frac{\sin^4\alpha}{[1+\cos^2\alpha \tan^2 u]^2} + \frac{\sin^6\alpha}{[1+\cos^2\alpha \tan^2 u]^3} + o\{f^4\}\right\} \quad \ldots(1.29).$$

To summarise,

$$\begin{array}{lcl} \xi_1 & & \text{is unaltered on conversion} \\ \xi_2 & = & \phi_g = \tan^{-1}[\cos\alpha \tan u] \\ \xi_3 & = & r = c \cot \alpha \left(1 - \dfrac{\sin^2\alpha}{1+\cos^2\alpha \tan^2 u}\right)^{-1} \end{array} \quad \ldots(1.30).$$

1.4 A review of determinations of the geoid to date.

The determinations of the geoid in the past have nearly always been evaluations of *co-geoids* which are obtained by the use of a particular anomaly (isostatic, free air, etc.) in Stokes' integral. Note that the term "gravity anomaly" has been avoided. In the ensuing development, this term will refer to the quantity defined by equation 1.5 and has meaning only if the physical and reference systems being related are specified. Thus the gravity anomaly for the physical surface - telluroid system(Mather 1968b,p.526) is different from that for the geoid-spheroid system(ibid,p.523). The free air anomaly is an excellent approximation for the gravity anomaly on the former system.

Gravimetric geodesy was totally pre-occupied with the Stokesian problem till the end of the Second World War. The original problem, as specified by Stokes(1849), afforded a solution in the case where the physical surface was a *bounding equipotential*, exterior to all matter, with the same potential, mass and volume as the bounding equipotential of the reference system. This gave rise to considerable efforts on the part of geodesists in the late nineteenth and early twentieth centuries to concentrate on the removal of mass exterior to the geoid in order that the correct boundary conditions were created for the use of Stokes' integral.

The anomaly which found most favour was that based on the principle of isostasy, being initially used by Hayford in the U.S. in 1909(Heiskanen & Vening Meinesz 1958,p.129). The use of isostatic anomalies was vigorously championed by the Finnish geodesist Heiskanen who founded an active group of researchers in physical geodesy at Helsinki in the nineteen thirties under the auspices of the International

Association of Geodesy and called the Isostatic Institute. He adopted a crustal model initially proposed by the British astronomer Airy which is now known to agree satisfactorily with independent evidence provided by seismology and the principal aim of the institute was the completion of all isostatic reductions necessary at all points on the earth's surface based on this crustal model(ibid,p.176 et seq.). At the same time, progress of varying degrees was made in gravimetric geodesy at the U.S. Coast & Geodetic Survey in the U.S.A., by the Survey of India under de Graaff-Hunter and by Jeffreys at Cambridge.

The problem facing all research groups was the paucity of gravity measurements at sea and in the southern hemisphere. During the period preceding the second world war, the only means available for the measurement of gravity at sea were the Vening Meinesz pendulums(ibid, p.115). It was indeed very tempting to accept as valid, the arguments of the Helsinki school in such a situation. These went as follows.

(i) The earth is in general isostatic equilibrium.

(ii) Hence the isostatic anomaly in the general case should tend to the value zero.

(iii) Thus all unsurveyed areas could be represented by zero anomaly values without significantly affecting the solution.

These arguments were, to the writer's mind, rather surprisingly supported by Vening Meinesz despite his interest in the mantle and the possibility of convection(ibid,p.72), as they implied that hydrostatic equilibrium prevailed beneath a certain depth of compensation. This interest in the mantle was probably a consequence of his observations of large scale departures from isostatic equilibrium over many parts of the earth.

Two gravimetric geoids were computed on the basis of these ideas. The first pioneering attempt was a free air co-geoid by Hirvonen in 1934 which is chiefly of historical interest. The second was an isostatic co-geoid based on four times as much gravity information as that available to Hirvonen(Heiskanen 1964,p.4). The solutions applied Stokes' integral by quadratures to an appropriately divided earth.

A different technique was adopted by Jeffreys who questioned the assumptions made by the Finnish school in the representation of un-surveyed areas. He reasoned (correctly, as subsequent events have shown)that the assumption of the concept of hydrostatic equilibrium was not warranted in view of the numerous deviations from it. He proposed the harmonic analysis of all available gravity data using a surface harmonic series of the form

$$\Delta g = \sum_{j=1}^{k} C_j [f(\phi,\lambda)]_j \qquad \ldots(1.31)$$

from which values could be obtained by least squares for the coefficients C_j and hence definition obtained for gravity anomalies at all points on the earth's surface(Jeffreys 1962,pp.185 et seq.).

It should be noted that Jeffreys' technique for the representation of the earth's gravity field is different in principle to the approach of the Helsinki school. Jeffreys recognises the fact that the earth's grvitational field is a continuous one and that the gravity value at any point is dependent on the distribution of values elsewhere. In principle, it makes no assumptions about either the nature of the field or the internal constitution of the earth. The Helsinki approach, on the other hand, assumes the assumes the existence of isostatic compensation and its gravitational consequences as the dominant contribution to the deviation of observed gravity from that of the standard reference model. This is, in fact,so but other factors make significant contributions to this deviation with far reaching geodetic consequences.

While the foundations of the analytical approach to the representation of g in unsurveyed areas were postulated by Jeffreys in 1941 (Jeffreys 1941), it was the Soviet geodesist Zhongolovich who produced the first geoid map using this technique(Zhongolovich 1952). The geoid undulations so obtained had magnitudes which were twice as large as those obtained using the zero anomaly approach for unsurveyed areas. Comparative geoid maps are given in (Heiskanen 1964). Zhongolovich's solution unfortunately included certain low degree harmonics in his analysis whose omission could be interpreted geometrically. Nevertheless there is little doubt that his sample was far too limited for an accurate solution, modern determinations indicating that the truth lies between the above extremes.

Kaula(1959) went a step further when he used a more comprehensive data set in adopting analytical techniques for the determination of the free air geoid obtained by the use of free air anomalies in Stokes' integral. His major refinement was the use of Markov processes to estimate the area means used in the harmonic analysis. The increased accuracy was possible primarily as a result of the improved gravity coverage available which enabled Kaula to use smaller basic units (1^{o} squares in contrast to 10^{o} squares used by both Jeffreys and Zhongolovich)and hence obtain greater accuracy.

In addition, Kaula introduced the concept of combining estimates of the values of harmonic coefficients obtained from satellite orbital analysis with surface gravity values to obtain an improved solution. Looking back from 1970, there is no doubt that Kaula's combined solution matches the astro-geodetic solution for the geoid in the Australian region much more closely than either his pure surface gravity solution or the isostatic co-geoid(ibid,pp.108-115). For a gravimetric solution of geodetic accuracy for the Australian geoid, see (Mather 1970b,p.111).

Kaula's 1959 determination only used values of a few selected low degree(i.e., long wave) harmonics. Significant advances in dynamic satellite geodesy enabled Kaula to make yet another development. He produced the first acceptable combined solution for the representation of the earth's gravitational field by fitting a harmonic representation to (12,12) the the available surface gravity data and obtaining as a by-product, improved values for the harmonic coefficients using the theory outlined in section 7.3(Kaula 1966a,p.5311). This can be visualised as an analytical technique for establishing values of gravity in unsurveyed areas. The resulting data set has been used to obtain geoidal solutions of geodetic accuracy(e.g., Mather 1969).

Thus physical geodesy in the seventies relies heavily on these analytical techniques for the definition of the earth's gravitational fieldas complete surface gravity coverage is still some years away. The theoretical problems involved are many. Most of these resolve themselves for solutions of the order of the flattening (i.e., ±30 cm) and the results obtained are quite adequate for practical purposes (e.g., Mather 1970b). The present development attempts to outline the problems involved in extending techniques to obtain solutions correct to order f^2.

It should be noted that the estimation of the precision of solutions to the order of the flattening does not mean that geoid determinations have errors of this order of magnitude. In the event of a complete representation of the gravity field becoming available, the residuals between solutions correct to orders f and f^2 will be position dependent. The shape of geops is dependent on two contributory factors(e.g., Moritz 1968, Mather 1970b).

(i) The general shape and mass distribution of the earth. These are long wave effects and best described as spherical or spheroidal.

(ii) The *local* deviations of the earth topography from the regular reference shapes(i.e., topographical effects).

The analytical techniques covered in the present development deal entirely with the spherical and/or spheroidal effects,

Summary

The same distribution of mass in earth space which defines the geops and the values of observed gravity (g) at the physical surface also cause perturbations in the instantaneous orbital elements defining the inertial space motion of a near earth satellite. It is conventional to analyse these perturbations using spherical functions from which an estimate could be obtained for the coefficients of a harmonic series defining the earth's gravitational field. The low degree terms of such a harmonic series can be used to define values for unsurveyed ocean areas by field extension between the available gravity anomaly data at the earth's surface.

2. INTRODUCTORY THEOREMS

2.1 General conventions

i. Reference systems

The space being considered can be defined by either of the following Cartesian systems:-

(i) x_i at the general point P;

and or (ii) X_i at the centre of mass C of the system.

Note that P may be at C but such use precludes the utilisation of any special properties of the centre of mass. Alternately, the space may be defined by a family of surfaces and a system of surface co-ordinates

$$x_\alpha$$

on a surface

$$x_3 = \text{constant}.$$

The guide to notation gives an explanation of the significance of subscripts.

ii. Vector representation

The general vector F in the space can be fully represented in terms of the three components along triply orthogonal Cartesian axes in the space and the unit vectors i along the x_i axes. It will be assumed in the following development that the rules of vector manipulation are known (e.g., Jeffreys & Jeffreys 1962, pp.57 et seq.). They are summarised here for ease of reference.

(i) If F_{1i} and F_{2i} are the three components of the vectors F_1 and F_2, their moduli F_i are given by

$$F_i = [\sum_{j=1}^{3} F_{ij}^2]^{\frac{1}{2}} \qquad \ldots(2.1)$$

More specifically, the vectors F_j can be represented by the equations

$$F_j = \sum_{i=1}^{3} F_{ji}\,i \qquad \ldots(2.2)$$

(ii) The scalar product of the vectors F_1 and F_2 is given by

$$F_1 . F_2 = F_1 F_2 \cos\theta = F_{1i}F_{2i} \left(= \sum_{i=1}^{3} F_{1i}F_{2i} \right) \qquad \ldots(2.3),$$

where θ is the angle between the vectors in the plane containing the vectors. If F_1 and F_2 are unit vectors, then F_{ij} can be considered to be direction cosines.

(iii) The vector product F_3 of the vectors F_1 and F_2 is given by

$$F_3 = F_1 \times F_2$$

where

$$F = \begin{vmatrix} F_{1(i+1)} & F_{1(i+2)} \\ F_{2(i+1)} & F_{2(i+2)} \end{vmatrix} \quad , \text{ ss} \qquad \ldots(2.4),$$

the term ss referring to the fact that subscripts take all possible values, being scaled when the value exceeds 3 by subtracting 3. Also

$$F_3 = F_1 F_2 \sin \theta \qquad \ldots(2.5)$$

(iv) Vectors are invariants in a given space. Hence the magnitude and direction of a vector are unchanged on change of reference frames. For a complete treatment see (Hotine 1969,p.3).

2.2 The Laplacian

Scalars in a space are invariants having a magnitude ϕ whose variations are a function of position alone. The operator ∇ is the vector defined by the relation

$$\nabla = \sum_{i=1}^{3} \frac{\partial}{\partial x_i} \, i \qquad \ldots(2.6).$$

The operator is of significance in defining the divergence of a vector F which is an invariant of the space, being given by the scalar product

$$\nabla . F = \sum_{i=1}^{3} \frac{\partial F_i}{\partial x_i} \qquad \ldots(2.7)$$

An important case is that of the vector F which is the gradient of the scalar ϕ given by

$$F = \nabla\phi \quad ,$$

when

$$\nabla . F = \nabla^2\phi = \sum_{i=1}^{3} \frac{\partial^2\phi}{\partial x_i^2} \qquad \ldots(2.8)$$

is called the Laplacian of ϕ. Both the Laplacian and the divergence of the general vector are invariants in the space. The gradient $\nabla\phi$ of the scalar ϕ is however, a vector.

2.3 The significance of $\nabla\phi$

Consider the surface given by the equation

$$\phi = C \quad ,$$

where C is a constant. This locus could also be expressed by the vector R given by

$$R = \sum_{i=1}^{3} x_i \, i$$

as shown in figure 2.1. If ds is an element of length wholly on the surface ,

$$\frac{d\phi}{ds} = 0$$

or

$$\sum_{i=1}^{3} \frac{\partial\phi}{\partial x_i} \frac{dx_i}{ds} = 0 = \sum_{i=1}^{3} \frac{\partial\phi}{\partial x_i} i \; . \; \sum_{i=1}^{3} \frac{dx_i}{ds} i \; .$$

This is equivalent to $\nabla\phi \, . \, \frac{dR}{ds} = 0.$

As dR/ds lies entirely in the surface $\phi = C$, the gradient $\nabla\phi$ of ϕ is therefore normal to the surface $\phi = C$.

If ϕ is potential, $\nabla\phi$ is therefore oriented along the normal to the equipotential surface $\phi = C$.

2.4 The scalar 1/r

If r is the modulus of the vector R described in section 2.3, then

$$r = \sum_{i=1}^{3} x_i i.$$

Hence

$$\frac{1}{r} = \left(\sum_{i=1}^{3} x_i^2 \right)^{-\frac{1}{2}}$$

and

$$\nabla\frac{1}{r} = \sum_{i=1}^{3} - \frac{2x_i}{2r^3} i = - \frac{1}{r^3} R \qquad \ldots(2.9)$$

Further,

$$\nabla^2\frac{1}{r} = -\nabla.\left(\frac{1}{r^3} R\right)$$

$$= -\left(R.\nabla \frac{1}{r^3} + \frac{1}{r^3}\nabla.R \right) \; .$$

In the case when $r \neq 0$, $\nabla.R = 3$ and

$$R.\nabla \frac{1}{r^3} = - \frac{3}{r^5}(x_1^2 + x_2^2 + x_3^2) = - \frac{3}{r^3} \; .$$

Hence

$$\nabla^2\frac{1}{r} = 0 \qquad \ldots(2.10).$$

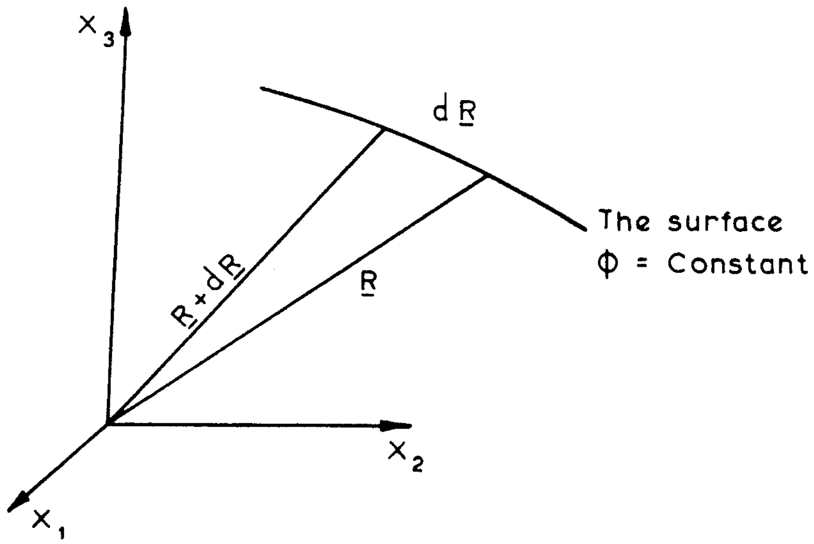


FIG. 2.1

The position vector and equipotential surfaces

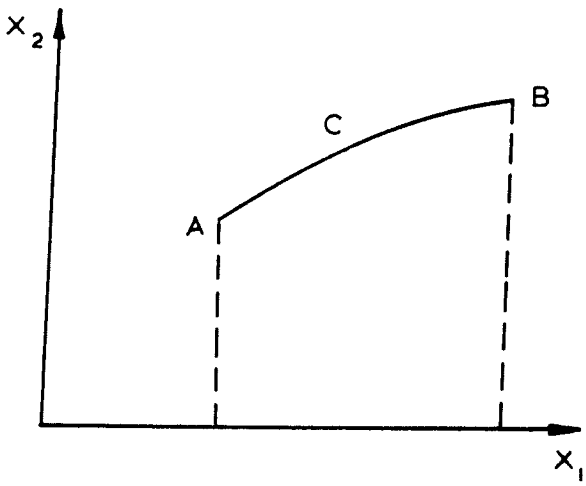


FIG. 2.2

Line integrals

1/r is said to satisfy Laplace's equation.

2.5 Surface and volume integrals

i. Basic definitions

Surfaces and volumes in a given Euclidian space are defined in terms of the loci of points in the space. Such points are specified by the set of values taken by three variables x_i. Curves, surfaces and volumes are defined by continuous sequences of points in three dimensions. The general theory of continuity in n dimensions is available in any standard text on the subject(e.g., ibid, p.176). The main conclusions may be summarised as follows.

(i) A function

$$f(P) \equiv f(x_i, i=1,m)$$

is continuous in m dimensions at a point P if the value f(Q) at the general point Q in the neighbourhood of P satisfies the inequality

$$|f(Q) - f(P)| < \varepsilon$$

where ε is a positive rational number.

(ii) The function f(P) is said to be differentiable if

$$[f(x_i+dx_i, i=1,m) - f(x_i, i=1,m) - \sum_{i=1}^{m} dx_i \frac{\partial f}{\partial x_i}] \leqq \varepsilon [\sum_{i=1}^{m} dx_i^2]^{\frac{1}{2}};$$

where $\sum_{i=1}^{m} dx_i^2 < \delta^2$,

δ being a positive rational number dependent on the location of P. This assumes that the partial derivatives $\partial f/\partial x_i$,(i=1,m) exist.

(iii) The mere existence of partial derivatives does not imply that a function is differentiable. A function f(P) is differentiable at P only if the partial derivatives $\partial f/\partial x_i$ are *continuous* in the neighbourhood of P.

(iv) If a function f(P) is continuous in the region around P, its gradient at P is a vector provided f(P) is differentiable in the region.

(v) If F_i are the components of a vector function in a region, a sufficient condition that $\partial F_i/\partial x_i$ shall be a second order tensor is that F_i be differentiable.

ii. Line integrals

Let C be a continuous curve in two dimensions defined by the x_1x_2 plane, joining two points A and B in figure 2.2. Let $f(x_\alpha)$ be a function which is defined at every point on C between A and B. If the arc AB is divided into n segments of length ds_k(k=1,n) and if $Q(x_{\alpha k})$ is a point in the k-th segment, then

$$\lim_{ds_k \to 0} \left(\sum_{k=1}^{n} f(x_{\alpha k})\, ds_k \right) = \int_C f(x_\alpha)\, ds \qquad \ldots (2.11)$$

is called a line integral. The continuity of $f(x_\alpha)$ is sufficient to define the existence of this line integral. The equation of the curve can be expressed in parametric form by the relation

$$x_\alpha = x_\alpha(s)$$

and the concept extended without loss of generality to three dimensions.

iii. Surface and volume integrals

Similar concepts can be defined for surface and volume integrals. The surface integral

$$\iint_S F(x_i)\, dS = \lim_{\substack{dS_k \to 0 \\ n \to \infty}} \left(\sum_{k=1}^{n} [F(x_i)]_k\, dS_k \right) \qquad \ldots (2.12),$$

where the subscript $_k$ refers to values at the k-th element of surface area, exists if $F(x_i)$ is bounded, i.e., it is continuous on the surface S enclosing the region V.

Similarly, the volume integral

$$\iiint_V F(x_i)\, dV = \lim_{\substack{dV_k \to 0 \\ n \to \infty}} \left(\sum_{k=1}^{n} [F(x_i)]_k\, dV_k \right) \qquad \ldots (2.13)$$

exists if $F(x_i)$ is bounded in the region V enclosed by the surface S. For a detailed treatment of these integrals as double and triple integrals together with the problems associated with change of variables, see (ibid, p.180).

2.6 The divergence theorem

Also known as Green's Lemma/ Gauss' theorem / Ostrogradskii's theorem.

Let S be a surface enclosing a volume V and let the vector *F*, given by an equation similar to 2.2 be defined at every point in V and on S. If S is such that no straight line parallel to any one of the Cartesian axes defining the space , intersects it in more than two points, the divergence theorem states that

$$\iiint_V \nabla.F \ dV = \iint_S N.F \ dS \qquad ...(2.14),$$

where ∇ is defined by equation 2.6, while N is the unit vector in the direction of the outward normal to the surface S.

Proof

The scalar product

$$\nabla.F = \sum_{i=1}^{3} \frac{\partial F_i}{\partial x_i}$$

must exist everywhere within the surface S, when the proposition to be proved can be written as

$$\iint_S N. \sum_{i=1}^{3} F_i i \ dS = \iiint_V \sum_{i=1}^{3} \frac{\partial F_i}{\partial x_i} dV \qquad ...(2.15).$$

From the definition implicit in equation 2.12, equation 2.13 can be considered as three separate integrals. For example, when i=3, as shown in figure 2.3, S being a double valued surface over its projection on the x_1x_2 plane , the element of surface area of such a projection is composed of two parts. The first is the lower element dS_1, with outward unit normal N_1 and an upper element dS_2 (normal N_2) is the second part. The four lines parallel to the x_3 axis which bound the elements of surface area dS_1 and dS_2 also enclose a certain volume within the surface. If the element of volume dV is given by

$$dV = \prod_{i=1}^{3} dx_i,$$

the terms obtained in the volume integral, when i=3 in equation 2.13 give

$$\iiint_{x_3(S_1)}^{x_3(S_2)} \frac{\partial F_3}{\partial x_3} dx_1 dx_2 dx_3$$

which, on integration with respect to x_3 gives

$$\iint dx_1 dx_2 (F_{3(S_2)} - F_{3(S_1)}),$$

where the subscripts $_{S_1}$ and $_{S_2}$ refer to values at the elements of surface area dS_1 and dS_2. From figure 2.3,

$$dx_1 dx_2 = +N_2.3 \ dS_2 = - N_1.3 \ dS_1.$$

Hence

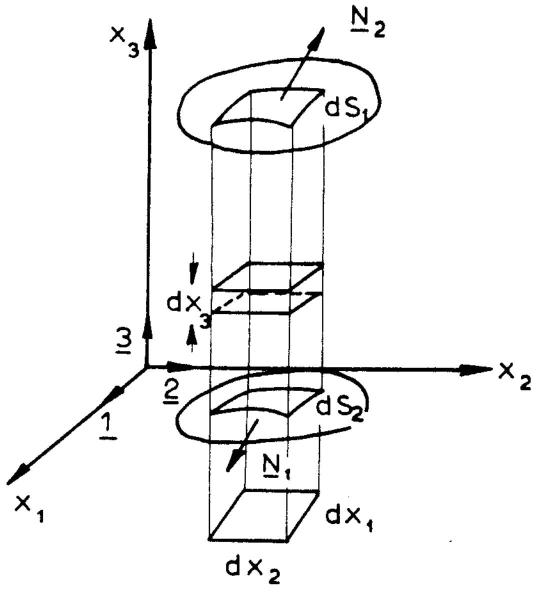


FIG. 2.3
The Divergence Theorem

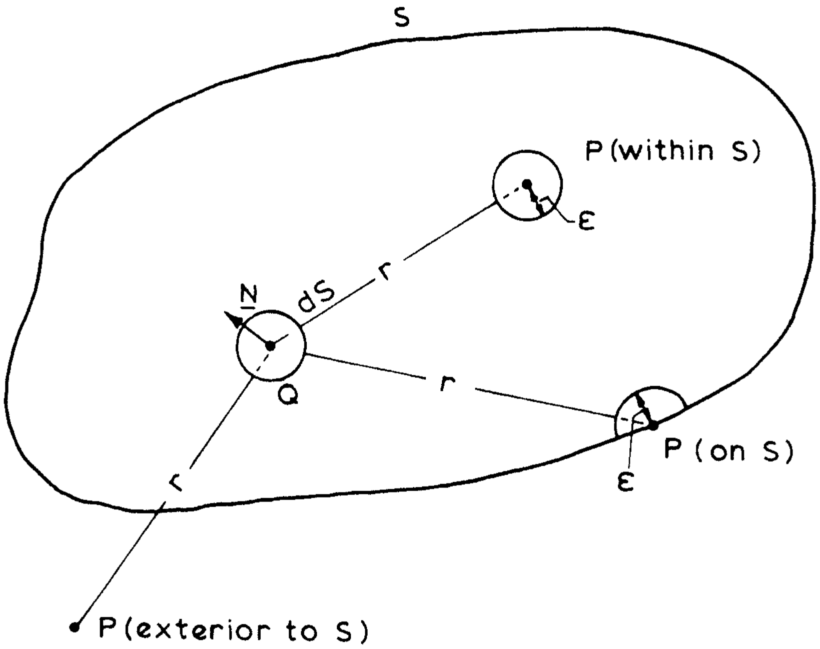


FIG. 2.4
Gauss' Theorem

$$\iiint \frac{\partial F_3}{\partial x_3} \, dV = \sum \left(\iint F_{3(S_2)} \, 3.N_2 \, dS_2 + \iint F_{3(S_1)} \, 3 \, .N_1 \, dS_1 \right) ,$$

which gives

$$\iiint \frac{\partial F_3}{\partial x_3} \, dV = \iint F_3 \, N.3 \, dS$$

as dS_1 and dS_2 make up the surface S. As this expression was derived without assigning any specific properties to the x_3 axis, the three equations

$$\iiint \frac{\partial F_i}{\partial x_i} \, dV = \iint F_i \, N.i \, dS$$

hold. Summation of these relations gives the original equation

$$\iiint \nabla.F \, dV = \iint N.F \, dS \qquad ...(2.14).$$

Notes

(i) The divergence theorem is applicable only if $\nabla.F$ exists everywhere within the volume V and $(N.F)$, which is the scalar normal derivative of the vector F, exists everywhere on S.

(ii) It is obvious from the proof that the divergence theorem also applies to those surfaces which are intersected by any line parallel to one of the co-ordinate axis an even number of times. This makes it possible to consider the case of the volume V enclosed between two surfaces S and S', when the above proof can be extended to obtain the relation

$$\iiint \nabla.F \, dV = \iint_S N.F \, dS + \iint_{S'} N.F \, dS' \qquad ...(2.16),$$

where the normal N is always outward with respect to the volume enclosed.

2.7 Green's theorem

Consider the application of the divergence theorem to the vector

$$F = U_1 \nabla U_2 ,$$

where U_1 and U_2 are scalars , when the following relations are obtained.

$$\iiint_V \nabla.(U_1 \nabla U_2) dV = \iint_S N.(U_1 \nabla U_2) dS$$

or

$$\iiint (\nabla U_1 . \nabla U_2 + U_1 \nabla^2 U_2)\, dV = \iint U_1\, N.\nabla\, U_2\, dS.$$

Similarly, in the case when

$$F = U_2 \ \nabla U_1,$$

$$\iiint (\nabla U_2 . \nabla U_1 + U_2 \nabla^2 U_1)\, dV = \iint U_2\, \nabla.N\, U_1\, dS.$$

On differencing the above equations,

$$\iiint (U_1 \nabla^2 U_2 - U_2 \nabla^2 U_1)\, dV = \iint (U_1\, \nabla.N U_2 - U_2\, \nabla.N U_1)\, dS \quad ...(2.17).$$

This result is known as Green's theorem and applies under the same conditions as the divergence theorem.

2.8 Gauss' theorem

Let S be a surface enclosing a volume V, as shown in figure 2.4. If the vector F is given by

$$F = \nabla \frac{1}{r} \overset{(2.9)}{=} - \frac{1}{r^3} R \ ,$$

where

$$R = \sum_{i=1}^{3} x_i i \ ,$$

the origin being at P in figure 2.4, the application of the divergence theorem to F gives

$$\iiint \nabla . \left(- \frac{R}{r^3}\right) dV = - \iint \frac{R}{r^3} . N \ dS.$$

If $r \neq 0$, P is outside S and the volume integral becomes zero as shown in the derivation of equation 2.10. Therefore

$$\iint \frac{R}{r^3}\, dS = 0 \qquad \text{if P outside S} \qquad ...(2.18).$$

r equals zero only if the origin P lies either on the surface S or in the volume V as shown in figure 2.4. In either of these cases, an instance occurs when 1/r is indeterminate. In the former, this instance can be excluded by considering the volume V' which arises as a consequence of removing a hemisphere of radius ε and centred on P, from the volume V. The application of equation 2.14 to the volume V' gives

$$\iiint_{V'} \nabla . \left(\frac{R}{r^3}\right) dV' = \iint_S N . \frac{R}{r^3} dS + \iint_{S'} N' . \frac{R}{r^3} dS' \qquad ...(2.19).$$

As

$$N' = -\frac{R}{\varepsilon} \quad ; \quad \text{and} \quad R.R = \varepsilon^2$$

at all points on S', the second integral on the right evaluates as

$$\frac{1}{\varepsilon^2}\iint dS' = -2\pi \quad ,$$

as S' is a hemisphere as P is on S. The volume integral on the left of equation 2.19 is zero for the same reasons which governed the evaluation of equation 2.18. Thus

$$\iint N \cdot \frac{R}{r^3}\, dS = 2\pi \qquad \text{if P is on S.}$$

Similarly,

$$\iint N.\frac{R}{r^3}\, dS = 4\pi \qquad \text{if P is within S.}$$

Summarising,

$$\iint N.\frac{R}{r^3}\, dS = \begin{cases} 0 & \text{if P exterior to S} \\ 2\pi & \text{if P is on S} \\ 4\pi & \text{if P is within S} \end{cases} \qquad ...(2.20).$$

This result is known as Gauss' theorem.

2.9 An extension of Gauss' theorem

Consider the function

$$F = \nabla \sum_{i=1}^{n} \frac{a_i}{r_i} \qquad ...(2.21),$$

where a_i is a constant associated with the points Q_i which are a distance r_i from some fixed point P. P could either be on the surface when it would be associated with an element of surface area dS or within it when it would coincide with an element of volume dV. Then

$$F = -\sum_{i=1}^{n} \frac{a_i}{r_i^3}\, x_i\, i = -\sum_{i=1}^{n} \frac{a_i}{r_i^3}\, R_i \ .$$

The application of divergence theorem to this vector gives

$$\iiint \nabla.\left(\sum_{i=1}^{n} \frac{a_i}{r_i^3} R_i\right) dV = \iint N \cdot \sum_{i=1}^{n} \frac{a_i}{r_i^3} R_i\, dS.$$

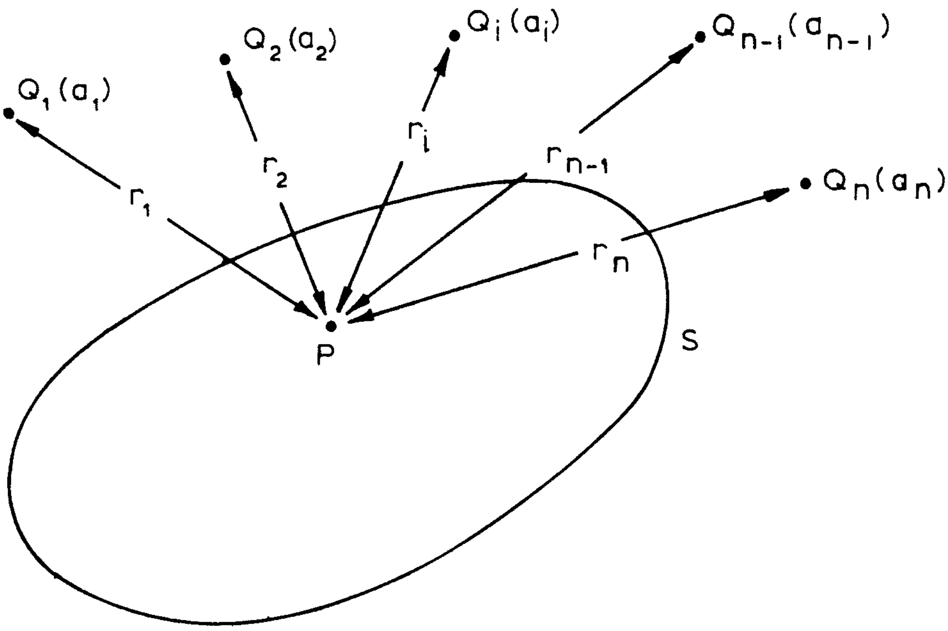


FIG. 2.5
Extension to Gauss' theorem

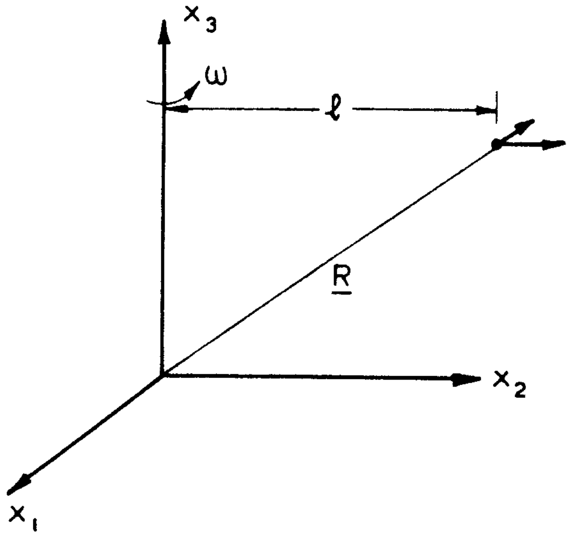


FIG. 2.6
Rotational potential

The volume integral will be zero provided no r_i is zero for reasons given in section 2.4. This occurs when all the points Q_i are exterior to the surface S and P is within S as shown in figure 2.5. If on the other hand, all the Q_i's are on S, they have to be excluded from the volume prior to integration by as many hemispheres as there are Q_i's on S, when the use of equation 2.20 gives

$$\iint N \cdot \sum_{i=1}^{n} \frac{a_i}{r_i^3} R_i \, dS = 2\pi \sum_{i=1}^{n} a_i .$$

Alternately, if all the Q_i's are within S,

$$\iint N \cdot \sum_{i=1}^{n} \frac{a_i}{r_i^3} R_i \, dS = 4\pi \sum_{i=1}^{n} a_i .$$

To summarise,

$$\iint N \cdot \sum_{i=1}^{n} \frac{a_i}{r_i^3} R_i \, dS = \begin{cases} 0 & \text{if all } Q_i\text{'s exterior to S} \\ 2\pi \sum_{i=1}^{n} a_i & \text{if all } Q_i\text{'s on S} \\ 4\pi \sum_{i=1}^{n} a_i & \text{if all } Q_i\text{'s are within S} \end{cases} \quad ..(2.22)$$

2.10 The potential due to Newtonian gravitation

The potential ϕ_p at a point P due to Newtonian gravitation as a consequence of the distribution of a set of masses dm_i at points Q_i(i=1,n) is given by

$$\phi_p = k \sum_{i=1}^{n} \frac{dm_i}{r_i} \quad ...(2.23),$$

where r_i is the distance PQ_i and k the gravitational constant. The quantity ϕ has the same characteristics as the summation whose divergence was equated to F in equation 2.21. On studying the properties of the gravitational potential with respect to the general surface S which encloses the volume V, it can be seen that equation 2.22 could be written as

$$\iint N \cdot \nabla\phi \, dS = -k \iint \sum_{i=1}^{n} \frac{dm_i}{r_i^3} R_i \, dS = - \begin{cases} 0 & \text{if all } Q_i\text{'s exterior to S} \\ 2\pi k \sum_{i=1}^{n} dm_i & \text{if all } Q_i\text{'s on S.} \\ 4\pi k \sum_{i=1}^{n} dm_i & \text{if all } Q_i\text{'s within S} \end{cases}$$

The use of the divergence theorem also gives

$$\iint N.\nabla\phi \quad dS \;=\; \iiint \nabla^2\phi \quad dV.$$

Hence if ρ is the variable density of matter giving rise to the gravitational potential, and if the matter dm_i is contained within the volume dV,

$$\iiint \nabla^2\phi \quad dV \;=\; -\begin{cases} 0 & \text{if no matter exists within S} \\ 4\pi k\iiint \rho \; dV & \text{if all matter is within S} \end{cases} \qquad ..(2.24).$$

As $\nabla^2\phi$ is zero at all points not coincident with matter, the equation

$$\nabla^2\phi = 0 \qquad ...(2.25),$$

called Laplace's equation, is said to hold under these conditions. The use of the result embodied in equation 2.24 shows that the equation

$$\nabla^2\phi \;=\; -4\pi k\rho \qquad ...(2.26)$$

is satisfied at all points occupied by matter. This result is known as Poisson's equation and relates the Laplacian of the potential to the density of matter at the point.

For comments on the relevance of the Newtonian theory, see (Hotine 1969,p.147).

2.11 The gravitational potential of a rotating body

Points on the surface of a body rotating with angular velocity ω are subject to a centrifugal force whose modulus is

$$\ell\omega^2 \; ,$$

where ℓ is the perpendicular distance from the axis of rotation to the point. This force is directed radially away from the axis of rotation (e.g., Humphreys 1945,p.216). If the x_3 axis is taken to be coincident with the axis of rotation, the centrifugal force can be expressed by the vector

$$F \;=\; \ell\omega^2 \sum_{i=1}^{2} \frac{x_i}{\ell}\, i \;=\; \omega^2 \sum_{i=1}^{2} x_i i \qquad ...(2.27),$$

where

$$\ell \;=\; (x_1^2 + x_2^2)^{\frac{1}{2}}.$$

The rotational potential ϕ_r is thus related to F by the equation

$$F = \nabla\phi_r = \omega^2 \sum_{i=1}^{2} x_i i .$$

This is equivalent to

$$\phi_r = \tfrac{1}{2} \sum_{i=1}^{2} x_i^2 \omega^2 \qquad \ldots(2.27).$$

The Laplacian of the rotational potential is given by

$$\nabla^2\phi_r = \omega^2 \sum_{i=1}^{2} \frac{\partial}{\partial x_i}\{x_i\} = 2\omega^2 \qquad \ldots(2.28).$$

The total potential V of a rotating body is therefore given by

$$V = \phi + \phi_r .$$

In addition,

$$\nabla^2 V = \begin{cases} 2\omega^2 & \text{at a point unoccupied by matter} \\ -4\pi k\rho + 2\omega^2 & \text{at a point occupied by matter of density} \end{cases} \qquad \ldots(2.29) .$$

The gravitational potential of the earth satisfies equation 2.29 at all points which rotate with it. This potential is commonly called the *geopotential* and surfaces of equal geopotential are called *geops*.

Laplace's equation

$$\nabla^2\phi = 0$$

is of considerable interest as it is a second order differential equation defining gravitational potential at all points in space which

(a) do not rotate with the gravitating body;

and (b) are unoccupied by matter.

Such conditions are satisfied by the external gravitational field of the earth in regions affecting the orbits of near earth satellites. The sections which follow deal with classes of functions which satisfy Laplace's equation and have characteristic properties which are useful in the analysis of data distributed over the surface of the earth.

2.12 The solution of Laplace's equation in general curvilinear co-ordinates

Special types of curvilinear co-ordinates have been dealt with in section 1.3. Attention will be confined in this section to systems of curvilinear co-ordinates which are orthogonal. Position with reference to these curvilinear co-ordinate systems are completely defined by three parameters ξ_i. The co-ordinates are orthogonal in that the linear equivalent ds_i corresponding to a change $d\xi_i$ in ξ_i and given by

$$ds_i = h_i d\xi_i \qquad ..(2.30),$$

where h_i is the associated linearization parameter, form a triply orthogonal system of displacements in the limit. The spherical and spheroidal system of co-ordinates described in section 1.3 are examples of such systems. The general element of volume dV in the space can be defined by the equation

$$dV = \prod_{i=1}^{3} h_i d\xi_i \qquad ...(2.31).$$

If the scalar U(e.g., gravitational potential) satisfies Laplace's equation at any point in the space, it follows from equation 2.25 that the space is unoccupied by matter. Consider the application of the divergence theorem to the vector ∇U over an infinitely small element of volume dV in this space which is defined by the system of curvilinear co-ordinates specified above. If the element of volume is small enough for $\nabla^2 U$ to be assumed to be a constant over the region,

$$\nabla^2 U \prod_{i=1}^{3} h_i d\xi_i = \iint \nabla . N \; U \quad ds \qquad ...(2.32),$$

where three parallel pairs of triply orthogonal surfaces comprise the surface S which encloses the volume dV as shown in figure 2.7. Further, each pair of surfaces is defined by parametric values of the form

$$\xi_{io} \pm \tfrac{1}{2} d\xi_i .$$

Further, the scalar product $\nabla . N$ is the normal derivative.

Hence

$$\nabla . N = \sum_{i=1}^{3} \frac{1}{h_i} \frac{\partial}{\partial \xi_i} \qquad ...(2.33).$$

The contributions to the surface integral in equation 2.32 by pairs of such surfaces is obtained by the use of a Taylor series as

$$[\nabla . N \; U \; ds]_{\xi_i = \xi_{io}} + \tfrac{1}{2} d\xi_i \frac{d}{d\xi_i} [\nabla . N U \; ds] + [\nabla . N' U \; ds]_{\xi_i = \xi_{io}} -$$

$$\tfrac{1}{2} d\xi_i \frac{d}{d\xi_i} [\nabla . N' U \; ds] \quad ,$$

where

$$dS = h_{i+1} h_{i+2} d\xi_{i+1} d\xi_{i+2} \qquad , \text{ ss } *$$

This reduces to

---

* see "Guide to Notation".

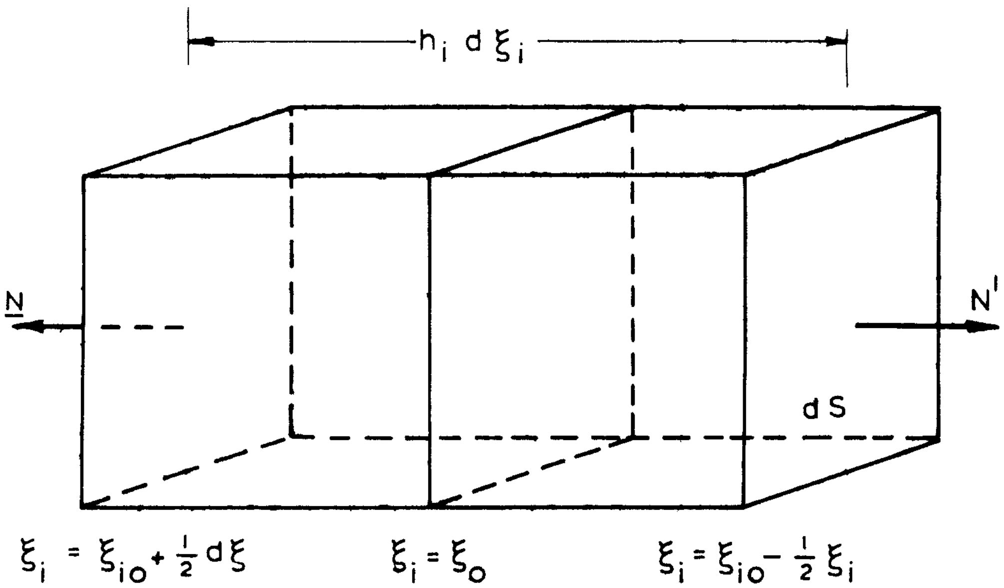


FIG. 2.7
Laplace's equation in curvilinear co-ordinates

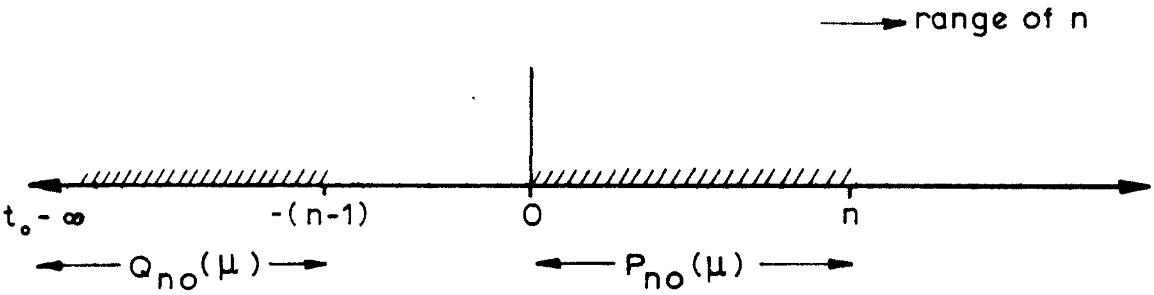


FIG. 3.1

Range of n for Legendre functions

$$d\xi_i \frac{d}{d\xi_i} \quad \nabla.NU\ ds = \left(\prod_{i=1}^{3} d\xi_i\right) \frac{d}{d\xi_i}\left(\frac{h_{i+1}h_{i+2}}{h_i}\frac{\partial U}{\partial \xi_i}\right) \quad , \quad ss$$

as $$N = -N'$$

The use of this result in equation 2.32 gives

$$\nabla^2 U \prod_{i=1}^{3} h_i d\xi_i = \left(\prod_{i=1}^{3} d\xi_i\right)\left[\sum_{i=1}^{3} \frac{d}{d\xi_i}\left(\frac{h_{i+1}h_{i+2}}{h_i}\frac{\partial U}{\partial \xi_i}\right)\right] , \quad ss$$

or

$$\nabla^2 U = \frac{1}{h_1 h_2 h_3} \sum_{i=1}^{3} \frac{d}{d\xi_i}\left(\frac{h_{i+1}h_{i+2}}{h_i}\frac{\partial U}{\partial \xi_i}\right) \overset{(2.24)}{=} 0 \quad , \quad ss.$$

Thus Laplace's equation can be expressed in terms of gereral orthogonal curvilinear co-ordinates ξ_i and their associated linearisation parameters h_i by the equation

$$\sum_{i=1}^{3} \frac{d}{d\xi_i}\left(\frac{h_{i+1}h_{i+2}}{h_i}\frac{\partial U}{\partial \xi_i}\right) = 0 \quad , \quad ss \qquad \ldots(2.34)$$

3. THE SOLUTION OF LAPLACE'S EQUATION IN SPHERICAL CO-ORDINATES

3.1 Introduction

Gravimetric geodesy is primarily concerned with the gravity field exterior to the earth. This space can be *completely* defined by the spherical system of co-ordinates described in section 1.3. The validity of an adopted model should not be confused with instances where physical characteristics are attributed to any one of the parameters. Further, it should be noted that Laplace's equation as expressed in equation 2.34 is only satisfied at points not occupied by matter. None of these problems will however intrude at this stage of the development if only the following conditions are imposed on the solution.

(i) The region applied to is exterior to all matter.

(ii) Positions in this volume are defined on a spherical co-ordinate system.

The parameters adopted for the spherical system are given in equation 1.7 and illustrated in figure 1.2, being the radius r, the co-latitude θ and the longitude λ. It can easily be verified that

$$\begin{aligned} \xi_1 &= \theta \;; & \xi_2 &= \lambda \;; & \xi_3 &= r \\ h_1 &= r \;; & h_2 &= r\sin\theta ; & h_3 &= 1 \end{aligned} \qquad \ldots(3.1).$$

for such a system of co-ordinates. This reference system can be visualised as being composed of a series of concentric spheres defined by ξ_3, while position on these spheres is defined by the *surface parameters* ξ_1 and ξ_2.

3.2 The solution in spherical co-ordinates

The substitution of the relations at 3.1 into equation 2.34 gives

$$\frac{d}{d\theta}[\sin\theta\frac{\partial U}{\partial\theta}] + \frac{d}{d\lambda}[\frac{1}{\sin\theta}\frac{\partial U}{\partial\lambda}] + \frac{d}{dr}[r^2\sin\theta\frac{\partial U}{\partial r}] = 0.$$

As r, θ and λ form an orthogonal system, differentiation and expansion of the above equation gives

$$\cos\theta\frac{\partial U}{\partial\theta} + \sin\theta\frac{\partial^2 U}{\partial\theta^2} + \frac{1}{\sin\theta}\frac{\partial^2 U}{\partial\lambda^2} + 2r\sin\theta\frac{\partial U}{\partial r} + r^2\sin\theta\frac{\partial^2 U}{\partial r^2} = 0 \qquad ....(3.2).$$

Let

$$U = U(r,\theta,\lambda)$$

be a solution of equation 3.2. Further, consider a solution which can be expressed in the form

$$U(r,\theta,\lambda) = R(r)\ S(\theta,\lambda) \qquad ...(3.3),$$

where $S(\theta,\lambda)$ is a function of the surface co-ordinates θ and λ alone. Substitution of equation 3.3 into 3.2 gives

$$\cos\theta\ R(r)\frac{\partial S}{\partial\theta} + \sin\theta\ R(r)\frac{\partial^2 S}{\partial\theta^2} + \frac{1}{\sin\theta}R(r)\frac{\partial^2 S}{\partial\lambda^2} + 2r\sin\theta\ S(\theta,\lambda)\frac{\partial R}{\partial r} +$$

$$r^2\sin\theta\ S(\theta,\lambda)\frac{\partial^2 R}{\partial r^2} = 0$$

Re-arrangement of terms and separation of variables gives

$$\frac{1}{R(r)}\left(r^2\frac{\partial^2 R}{\partial r^2} + 2r\frac{\partial R}{\partial r}\right) = -\frac{1}{S(\theta,\lambda)}\left(\frac{\partial^2 S}{\partial\theta^2} + \cot\theta\frac{\partial S}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2 S}{\partial\lambda^2}\right).$$

The two second order differential expressions given above have to be satisfied simultaneously. Two differential equations, one in r and the other in (θ,λ), can be obtained from the above by putting each side equal to a common constant. The adoption of the term $n(n+1)$ as this constant is of convenience as it enables a simple solution to be obtained for the second order differential equation in r (Hobson 1965,p.9). The two equations which result are

$$r^2\frac{\partial^2 R}{\partial r^2} + 2r\frac{\partial R}{\partial r} - n(n+1)\ R(r) = 0 \qquad ...(3.4)$$

and

$$\frac{\partial^2 S}{\partial \theta^2} + \cot\theta \frac{\partial S}{\partial \theta} + \frac{1}{\sin^2\theta}\frac{\partial^2 S}{\partial \lambda^2} + n(n+1)\ S(\theta,\lambda) \qquad \ldots(3.5).$$

In a similar manner, $S(\theta,\lambda)$ can be considered as being the product of two functions $F(\lambda)$ and $G(\theta)$. The adoption of a relation of the form

$$S(\theta,\lambda) = F(\lambda)\ G(\theta) \qquad \ldots(3.6)$$

transforms equation 3.5 into

$$F(\lambda)\frac{\partial^2 G}{\partial \theta^2} + \cot\theta\ F(\lambda)\ \frac{\partial G}{\partial \theta} + \frac{1}{\sin^2\theta}\ G(\theta)\frac{\partial^2 F}{\partial \lambda^2} + n(n+1)\ F(\lambda)\ G(\theta) = 0.$$

The separation of variables and re-arrangement of terms gives

$$-\frac{1}{F(\lambda)}\left(\frac{\partial^2 F}{\partial \lambda^2}\right) = \frac{1}{G(\theta)}\left(\sin^2\theta\ \frac{\partial^2 G}{\partial \theta^2} + \cos\theta\sin\theta\frac{\partial G}{\partial \theta} + n(n+1)\sin^2\theta\ G(\theta)\right).$$

A repetition of the procedure used above gives the second order differential equations in θ and λ if each side is put equal to a constant, a convenient form for which is m^2. The resulting equations are

$$\frac{\partial^2 F}{\partial \lambda^2} + m^2\ F(\lambda) = 0 \qquad \ldots(3.7)$$

and

$$\sin^2\theta\ \frac{\partial^2 G}{\partial \theta^2} + \cos\theta\sin\theta\frac{\partial G}{\partial \theta} + [n(n+1)\sin^2\theta - m^2]\ G(\theta) = 0 \qquad ..(3.8).$$

The solutions $R(r)$, $F(\lambda)$ and $G(\theta)$ of equations 3.4, 3.7 and 3.8, on combination through equations 3.3 and 3.6 give the solution of Laplace's equation $U(r,\theta,\lambda)$ in spherical co-ordinates. Equations 3.4 and 3.7 are standard second order differential equations and the solutions can be directly verified as

$$R(r) = A\,r^n + \frac{B}{r^{n+1}} \qquad \ldots(3.9)$$

and

$$F(\lambda) = C\cos m\lambda + D\sin m\lambda \qquad \ldots(3.10)$$

respectively, where A, B, C and D are constants of integration.

The solution of equation 3.8 is more complex. The introduction of the change of variable

$$\mu = \cos\theta$$

gives

$$\frac{dG}{d\theta} = \frac{dG}{d\mu}\ \frac{d\mu}{d\theta} = -\sin\theta\ \frac{dG}{d\mu}$$

and

$$\frac{d^2G}{d\theta^2} = \sin^2\theta \frac{d^2G}{d\mu^2} - \cos\theta \frac{dG}{d\mu} = (1-\mu^2)\frac{d^2G}{d\mu^2} - \mu\frac{dG}{d\mu}.$$

Thus equation 3.8 becomes

$$(1-\mu^2)\left((1-\mu^2)\frac{d^2G}{d\mu^2} - \mu\frac{dG}{d\mu}\right) - \mu(1-\mu^2)\frac{dG}{d\mu} + (1-\mu^2)\left(n(n+1) - \frac{m^2}{1-\mu^2}\right)G(\mu) = 0$$

or

$$(1-\mu^2)\frac{d^2G}{d\mu^2} - 2\mu\frac{dG}{d\mu} + [n(n+1) - \frac{m^2}{1-\mu^2}]\,G(\mu) = 0 \quad \text{if } \mu^2 \neq \pm 1 \quad ..(3.11)$$

Note

In the case where m=0, the above equation reduces to a form known as Legendre's equation.

Thus U can be expressed by a relation of the form

$$U = \left(Ar^n + \frac{B}{r^{n+1}}\right)[C\cos m\lambda + D\sin m\lambda]\;G(\mu) \qquad ....(3.12).$$

The constant A equals zero as $U \to 0$ when $r \to \infty$. Thus equation 3.12 can be written more generally as

$$U = \frac{1}{r^{n+1}}[\,C_{nm}\cos m\lambda + S_{nm}\sin m\lambda]\;G_{m}(\mu) \qquad ...(3.13),$$

which is a solution of Laplace's equation, C_{nm} and S_{nm} being constants whose values depend on n and m which, in turn, are not functions of the co-ordinates but are constants, independent of position. Thus an infinite number of solutions are possible depending on the combination of values adopted for n and m. It is still necessary to obtain an expression for $G_{nm}(\mu)$. This is done in stages. In the first, m is taken as equal to zero when a solution is obtained for Legendre's equation. In the second, the solution for Legendre's equation is extended to the case when $m \neq 0$.

3.3 The solution of Legendre's equation

Legendre's equation is given from equation 3.11 as

$$(1-\mu^2)\frac{d^2G_n}{d\mu^2} - 2\mu\frac{dG_n}{d\mu} + n(n+1)\;G_n(\mu) = 0 \qquad ...(3.14),$$

where μ as obtained in section 3.2 lies within the range $-1 \leq \mu \leq 1$. However, all possible values will be admitted for μ in order that the requirements of future sections can be met. A solution of the form

$$G_n(\mu) = \sum_{i=-\infty}^{\infty} a_i\mu^i$$

is adopted for $G_n(\mu)$. Successive differentiations give

$$\frac{dG_n}{d\mu} = \sum_{i=-\infty}^{\infty} i\, a_i\, \mu^{i-1}$$

and

$$\frac{d^2G_n}{d\mu^2} = \sum_{i=-\infty}^{\infty} i(i-1) a_i\, \mu^{i-2}.$$

Thus equation 3.14 can be written as

$$\sum_{i=-\infty}^{\infty} \left\{ i(i-1)\mu^{i-2} + \{n(n+1)-[i(i-1) + 2i]\}\mu^i \right\} a_i = 0.$$

This equation will be satisfied if all the coefficients of μ^i are each equal to zero, i.e.,

$$[n(n+1) - i(i+1)]a_i + (i+1)(i+2)a_{i+2} = 0$$

or

$$a_i = \frac{(i+1)(i+2)}{i(i+1)-n(n+1)}\, a_{i+2} \qquad \ldots(3.15).$$

It can also be seen that if any a_i is zero, all a_i's for smaller values of i in increments of two are also zero. This happens when $i = -1$ or $i = -2$ for a non-zero coefficient a_{i+2}. Thus all coefficients for $i < 0$ are zero even when coefficients for $i > 0$ are non-zero. Also

$$a_{i+2} = \frac{(i-n)(i+n+1)}{(i+1)(i+2)}\, a_i \qquad \ldots(3.16).$$

Again, $a_{i+2} = 0$ when $i = n$ or $i = -(n+1)$ even when $a_i \neq 0$. Thus all coefficients a_i are zero for $i > n$. In addition, all a_i's are zero when i is greater than $-(n+1)$. Thus equations 3.15 and 3.16 specify two series as illustrated in figure 3.1 The range of indices in each series are

(a) $0 < i < n$

and

(b) $-\infty < i < -(n+1)$.

More specifically, the first series is of the type

$$G_n(\mu) = \sum_{i=0}^{n} a_i\, \mu^i \qquad \ldots(3.17),$$

while the second is of the form

$$G_n(\mu) = \sum_{i=n+1}^{\infty} a_i\, \frac{1}{\mu^i} \qquad \ldots(3.18).$$

The first series is convergent for $n \to \infty$ when $|\mu| < 1$ while the second does so when $|\mu| > 1$. The relationship between successive coefficients in the series are obtained through equation 3.15, in the case of the expression at 3.17 which is of direct relevance to the solution in spherical co-ordinates. When

When $i=n-2$,

$$a_{n-2} = -\frac{n(n-1)}{2(2n-1)}a_n.$$

When $i=n-4$,

$$a_{n-4} = -\frac{(n-2)(n-3)}{4(2n-3)}a_{n-2} = \frac{n(n-1)(n-2)(n-3)}{2^2 2!(2n-1)(2n-3)}a_n.$$

Also, when $i=n-6$,

$$a_{n-6} = -\frac{(n-4)(n-5)}{6(2n-5)}a_{n-4} = -\frac{n(n-1)(n-2)(n-3)(n-4)(n-5)}{2^3 3!(2n-1)(2n-3)(2n-5)}a_n.$$

Thus the series

$$G_n(\mu) = a_n\left(\mu^n - \frac{n(n-1)}{2(2n-1)}\mu^{n-2} + \dots + (-1)^r\frac{n(n-1)(n-2)\dots(n-2r+1)}{2^r r!(2n-1)(2n-3)\dots(2n-2r+1)}\mu^{n-2r}\right.$$

is a solution of Legendre's equation. The value of $G_n(\mu)$ when

$$a_n = \frac{1.3.5\ \dots\ (2n-1)}{n!} = \frac{(2n)!}{2^n(n!)^2}$$

is designated as $P_{no}(\mu)$ (or $P_n(\mu)$ in most texts) and called a *Legendre function*. The final expression can be written as

$$P_{no}(\mu) = \frac{(2n)!}{2^n(n!)^2}\sum_{r=0}^{t}(-1)^r\frac{n(n-1)(n-2)\dots(n-2r+1)}{2^r r!(2n-1)(2n-3)\dots(2n-2r+1)}\mu^{n-2r}$$

$$= \frac{1}{2^n}\sum_{r=0}^{t}(-1)^r\frac{(2n-2r)!}{r!(n-2r)!(n-r)!}\mu^{n-2r} \qquad \dots(3.19),$$

where

$$t = \begin{cases} \tfrac{1}{2}n & \text{if } n \text{ even} \\ \tfrac{1}{2}(n-1) & \text{if } n \text{ odd.} \end{cases}$$

Notes

(i) The upper limits of equation 3.19 can easily be verified against the conclusions following equation 3.15.

(ii) It should be noted that Legendre functions are not only a solution of Legendre's equation but also of Laplace's equation.

(iii) In addition, it is equally important to realise that these functions are not the *only* solution of these equations.

(iv) Legendre functions are convergent only if $|\mu| \leq 1$.

3.4 Legendre functions of the second kind

Legendre functions of the second kind are series of the form expressed by equation 3.18. The non zero coefficient with the highest index is obtained when $i = -(n+1)$. The terms comprising the series are obtained by the application of equation 3.15.
When $i = -(n+3)$,

$$a_{-(n+3)} = \frac{(n+1)(n+2)}{2(2n+3)} a_{-(n+1)}.$$

Again, when $i = -(n+5)$,

$$a_{-(n+5)} = \frac{(n+3)(n+4)}{4(2\ +5)} a_{-(n+3)} = \frac{(n+1)(n+2)(n+3)(n+4)}{2^2 2!(2n+3)(2n+5)} a_{-(n+1)},$$

and, in general when $i = -(n+2r+1)$, an extension of this technique gives

$$a_{-(n+2r+1)} = \frac{(n+1)(n+2)(n+3)(n+4)\ldots(n+2r)}{2^r r!(2n+3)(2n+5)\ldots\ldots(2n+2r+1)} a_{-(n+1)}.$$

Thus

$$G_n(\mu) = a_{-(n+1)}\left\{\sum_{r=0}^{\infty} \frac{(n+1)(n+2)\ \ldots\ldots(n+2r)}{2^r r!(2n+3)(2n+5)\ldots(2n+2r+1)} \frac{1}{\mu^{n+2r+1}}\right\}.$$

When $a_{-(n+1)}$ is assigned the value

$$a_{-(n+1)} = \frac{n!}{1.3.5\ \ldots(2n-1)(2n+1)} = \frac{2^n(n!)^2}{(2n+1)!},$$

$G_n(\mu)$ is designated as $Q_{no}(\mu)$ [or $Q_n(\mu)$ in most texts], being given by

$$G_n(\mu) = Q_{no}(\mu)$$

$$= \frac{2^n(n!)^2}{(2n+1)!} \sum_{r=0}^{\infty} \frac{(n+1)(n+2)(n+3)\ldots(n+2r)}{2^r r!(2n+3)(2n+5)\ldots(2n+2r+1)} \frac{1}{\mu^{n+2r+1}}$$

$$= 2^n \sum_{r=0}^{\infty} \frac{(n+r)!(n+2r)!}{r!(2n+2r+1)!} \frac{1}{\mu^{n+2r+1}} \qquad \ldots(3.20).$$

$Q_{no}(\mu)$ is also a solution of Legendre's equation and is called a *Legendre function of the second kind.*

Notes

(i) The complete solution of Legendre's equation is therefore

$$G_n(\mu) = A\ P_{no}(\mu) + B\ Q_{no}(\mu) \qquad \ldots(3.21)$$

where

(a) $B = 0$ for a convergent series when $|\mu| < 1$;

and (b) $A = 0$ for this same condition if $|\mu| > 1$.

(ii) Legendre functions of the second kind[$Q_{no}(\mu)$] are needed for a development when the solution for Legendre's equation is required for a parameter which takes values greater then 1.

3.5 Alternate methods for generating Legendre functions

i. The function r^{-1}

Consider a fixed point P at a distance R_p from an origin O and a variable point Q whose distance from O is R. If QP = r, and the angle $Q\hat{O}P = \psi$, the application of the cosine formula to triangle POQ in figure 3.2 gives

$$r = [R^2 + R_p^2 - 2RR_p \cos\psi]^{\frac{1}{2}}.$$

If $R < R_p$,

$$\frac{1}{r} = \frac{1}{R_p}[1 + t^2 - 2t\mu]^{-\frac{1}{2}} \qquad \ldots(3.22),$$

where

$$t = \frac{R}{R_p} < 1 \qquad \text{and} \qquad \mu = \cos\psi \qquad \ldots(3.23).$$

Equation 3.22 can be expanded by the use of the binomial theorem to give

$$\frac{1}{r} = \frac{1}{R_p}[1 - t(2\mu - t)]^{-\frac{1}{2}}$$

$$= \frac{1}{R_p}\left\{1 + \tfrac{1}{2}t(2\mu - t) + \frac{1.3}{2^2\,2!}[t(2\mu-t)]^2 + \ldots + \frac{1.3.5\ldots(2n-1)}{2^n\,n!}[t(2\mu-t)]^n + \ldots\right\} \qquad ..(3.24)$$

The above series is convergent if

$$|t(2\mu - t)| < 1.$$

In general, this is a double series and will be absolutely convergent if, in addition, $(2|\mu| + t) < 1$ (Hobson 1965,p.15). The n-th term in the expansion on the right in equation 3.24 is

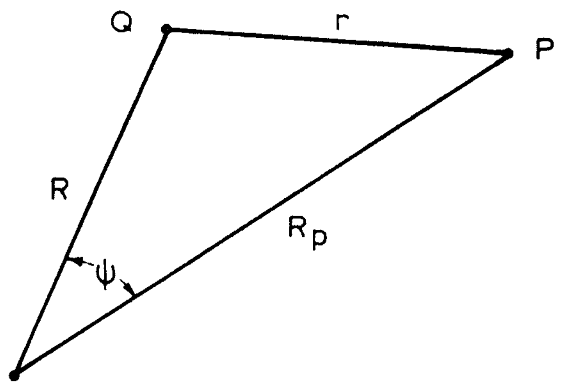


FIG. 3.2

The function r^{-1}

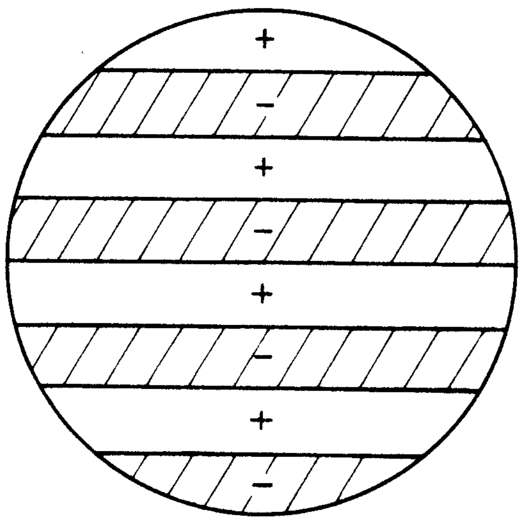


FIG. 4.1

The Zonal harmonic $P_{70}(\mu)$

$$\frac{(2n)!}{2^{2n}(n!)^2} t^n \Big((2\mu)^n - (2\mu)^{n-1} + \frac{n(n-1)}{2!}(2\mu)^{n-2} - \frac{n(n-1)(n-2)}{3!}(2\mu)^{n-3} + ..$$

$$...+(-1)^r \frac{n(n-1)\ ...\ (n-r+1)}{r!}(2\mu)^{n-r} t^r + \ .....\ +(-1)^n t^n \Big) .. (3.25).$$

Contributions to the coefficient of t^n in a series of the form

$$\frac{1}{r} = \frac{1}{R_p} \sum_{n=0}^{\infty} K_n t^n$$

are obtained from the successive terms in the series on the right of equation 3.24. E.g., the coefficient of t^n can only be contributed to by terms whose index v is less than n. The contribution of this term is obtained from the inner series given at 3,25 when $r = n-v$, is

$$\frac{(2v)!}{2^{2v}(v!)^2} (-1)^{n-v} \frac{v(v-1)\ ....(v-\overline{n-v}+1)}{(n-v)!} (2\mu)^{2v-n}.$$

The nature of the terms obtained are easily verified. When $v = n$, the contribution is

$$\frac{(2n)!}{2^n(n!)^2} \mu^n.$$

When $v = n-1$, it is

$$\frac{(2n-2)!}{2^{2n-2}[(n-1)!]^2} (-1)^1 \frac{(n-1)}{1!} 2^{n-2} \mu^{n-2} = \frac{(2n)!}{2^n(n!)^2}(-1)^1 \frac{n(n-1)}{2(2n-1)} \mu^{n-2}.$$

The contribution obtained when $v=n-2$ is

$$\frac{(2n-4)!}{2^{2n-4}[(n-2)!]^2} (-1)^2 \frac{(n-2)(n-3)}{2!} 2^{n-4} \mu^{n-4}$$

$$= \frac{(2n)!}{2^n(n!)^2} (-1)^2 \frac{n(n-1)(n-2)(n-3)}{2^2 2!(2n-1)(2n-3)} \mu^{n-4},$$

the gereral term being given by $v = n-r$ as

$$\frac{(2n-2r)!}{2^{2n-2r}[(n-r)!]^2} (-1)^r \frac{(n-2r+1)\ .....(n-r)(n-r-1)}{r!} \mu\mu^{n-2r}$$

$$= \frac{(2n)!}{2^n(n!)^2}(-1)^r \frac{n(n-1)\ .....\ (n-2r+1)}{2^r r!(2n-1)(2n-3)\ ...(2n-2r+1)} \mu^{n-2r}.$$

On summing all the contributory terms, it can be seen that the coefficient of t^n is the series

$$\frac{(2n)!}{2^n(n!)^2} \sum_{r=0}^{k} (-1)^r \frac{n(n-1)(n-2)\ldots(n-2r+1)}{2^r\, r!(2n-1)(2n-3)\ldots(2n-2r+1)} \mu^{n-2r}$$

$$\overset{(3.19)}{=} P_{no}(\mu).$$

Thus

$$\frac{1}{r} = \frac{1}{R_p} \sum_{n=0}^{\infty} t^n P_{no}(\mu) \qquad \ldots(3.26),$$

the summation commencing at $n=0$ to cover the first term in equation 3.24, as $t^o=1$ and $P_{oo}(\mu)=1$.

Notes

(i) Equation 3.26 is of considerable significance in studying the external gravitational field of the earth.

(ii) It has been shown by (Ecker 1970) that the series at 3.26 is convergent outside a sphere of radius R. He obtained this result by generalising equation 3.26 to the form

$$F = \sum_{n=0}^{\infty} \frac{A_n}{R_p^{n+1}} P_{no}(\mu)$$

when the "sphere of convergence" is one of radius

$$\lim_{n\to\infty} \sqrt[n]{|A_n|} \qquad \ldots(3.27).$$

ii. Rodrigues' formula for Legendre functions

Consider the function

$$F = \frac{d^n}{d\mu^n} [(\mu^2-1)^n]$$

$$= \frac{d^n}{d\mu^n}[\mu^{2n} - n\mu^{2n-2} + \frac{n(n-1)}{2!}\mu^{2n-4} - \frac{n(n-1)(n-2)}{3!}\mu^{2n-6} + \ldots$$

$$\ldots +(-1)^n] .$$

$$F = 2n(2n-1)(2n-2)\ldots(n+1)\mu^n - \frac{n(2n-2)(2n-3)\ldots(n-1)}{2!}\mu^{n-2} +$$

$$\ldots +(-1)^r \frac{n(n-1)\ldots(n-r+1)}{r!}(2n-2r)(2n-2r-1)\ldots(n-2r+1)\mu^{n-2r} +$$

$$\ldots$$

$$= \frac{(2n)!}{n!}\left\{\mu^n - \frac{n(n-1)}{2(2n+1)}\mu^{n-2} + \ldots +(-1)^r \frac{n(n-1)\ldots(n-2r+1)}{2^r r!(2n-1)(2n-3)\ldots(2n-2r+1)}\mu^{n-2r}\right\} \ldots(3.28).$$

A comparison of equations 3.19 and 3.28 shows that

$$F = 2^n n! P_{no}(\mu)$$

or

$$P_{no}(\mu) = \frac{1}{2^n n!} \frac{d^n}{d\mu^n}[(\mu^2-1)^n] \qquad \ldots(3.29).$$

Equation 3.29 is known as Rodrigues' formula.

iii. A recurrence formula for Legendre functions

Let

$$u = [1 - 2t\mu + t^2]^{-\frac{1}{2}} = \frac{1}{r} \quad .$$

Then

$$\frac{\partial u}{\partial t} = -\frac{1}{2} \frac{-2\mu + 2t}{[1-2\mu t+t^2]^{3/2}}$$

or

$$[1 - 2t\mu + t^2]\frac{\partial u}{\partial t} = u(\mu - t).$$

From equation 3.26,

$$u = \sum_{n=0}^{\infty} t^n P_{no}(\mu)$$

and

$$\frac{\partial u}{\partial t} = \sum_{n=0}^{\infty} n\, t^{n-1} P_{no}(\mu).$$

Therefore,

$$\sum_{n=0}^{\infty} \left\{ n\, t^{n-1} P_{no}(\mu)[1-2t\mu+t^2] + t^n P_{no}(\mu)[t-\mu] \right\} = 0.$$

This relation holds if the coefficients of each degree in t are separately equal to zero. The coefficient of t^{n-1} is

$$nP_{no}(\mu) - 2\mu(n-1)P_{(n-1)o}(\mu) + (n-2)P_{(n-2)o}(\mu) + P_{(n-2)0}(\mu) - \mu P_{(n-1)o}(\mu) = 0$$

or

$$P_{no}(\mu) - \mu(2n-1)P_{(n-1)o}(\mu) + (n-1)P_{(n-2)o}(\mu) = 0 \qquad \ldots(3.30).$$

This recurrence formula provides a convenient method for the *numerical* evaluation of Legendre functions provided they are evaluated as a series.

iv. Conclusion

The Legendre functions $P_{no}(\mu)$, which are the solution of Legendre's equation

$$(1-\mu^2)\frac{d^2G_n}{d\mu^2} - 2\mu\frac{dG_n}{d\mu} + n(n+1)G_n(\mu) = 0$$

are only a special instance of the solution of the general differential equation

$$(1-\mu^2)\frac{d^2G}{d\mu^2} - 2\mu\frac{dG}{d\mu} + \left(n(n+1) - \frac{m^2}{1-\mu^2}\right)G(\mu) = 0 \quad \ldots(3.11),$$

where $G(\mu)$ is taken to represent the term $G_{nm}(\mu)$ given in equation 3.13. This term completes the solution for gravitational potential through Laplace's equation

$$U = \frac{1}{r^{n+1}} \quad G(\mu)[\,C_{nm}\cos m\lambda + S_{nm}\sin m\lambda\,] \;.$$

It is therefore necessary to obtain the solution of the general differential equation 3.11.

3.6 Associated Legendre functions

Assume a solution for equation 3.11 of the form

$$G(\mu) = [1-\mu^2]^{m/2}\, H(\mu) \qquad \ldots(3.31).$$

Then

$$\frac{dG}{d\mu} = -\tfrac{1}{2}m[1-\mu^2]^{\frac{1}{2}m-1}2\mu\, H(\mu) + [1-\mu^2]^{\frac{1}{2}m}\frac{dH}{d\mu}$$

and

$$\frac{d^2G}{d\mu^2} = \tfrac{1}{2}m(\tfrac{1}{2}m-1)[1-\mu^2]^{\frac{1}{2}m-2}4\mu^2\, H(\mu) - m[1-\mu^2]^{\frac{1}{2}m-1}H(\mu) -$$

$$\tfrac{1}{2}m[1-\mu^2]^{\frac{1}{2}m-1}2\mu\frac{dH}{d\mu} - \tfrac{1}{2}m[1-\mu^2]^{\frac{1}{2}m-1}2\mu\frac{dH}{d\mu} + [1-\mu^2]^{\frac{1}{2}m}\frac{d^2H}{d\mu^2}$$

$$= [1-\mu^2]^{\frac{1}{2}m}\frac{d^2H}{d\mu^2} - 2m\mu[1-\mu^2]^{\frac{1}{2}m-1}\frac{dH}{d\mu} + \left(m(m-2)\mu^2[1-\mu^2]^{\frac{1}{2}m-2} -\right.$$

$$\left. m\mu[1-\mu^2]^{\frac{1}{2}m-1}\right)H(\mu).$$

The substitution of these expressions into equation 3.11 gives

$$[1-\mu^2]^{\frac{1}{2}m}\left\{[1-\mu^2]\frac{d^2H}{d\mu^2} - 2\mu(m+1)\frac{dH}{d\mu} + \left(\frac{m(m-2)}{1-\mu^2} + \frac{2\mu^2 m}{1-\mu^2} - m + n(n+1) - \frac{m^2}{1-\mu^2}\right)H(\mu)\right\}=0,$$

which can be written as

$$[1-\mu^2]\frac{d^2H}{d\mu^2} - 2\mu\,(m+1)\frac{dH}{d\mu} + [\,(n-m\,)(n+m+1)\,]H(\mu) = 0 \qquad \ldots(3.32)$$

on simplifying the coefficient of $H(\mu)$ in the following steps.

$$n(n+1) - \frac{m^2[1-\mu^2]}{1-\mu^2} - m = n^2 + n - m^2 - m$$

$$= (n-m)(n+m+1).$$

Consider Legendre's equation at 3.14. Differentiation with respect to μ gives

$$[1-\mu^2]\frac{d^{2+1}G_n}{d\mu^{2+1}} - 2\mu\frac{d^2G_n}{d\mu^2} - 2\frac{dG}{d\mu}n - 2\mu\frac{d^{1+1}G_n}{d\mu^{1+1}} + n(n+1)\frac{dG_n}{d\mu} = 0$$

or

$$[1-\mu^2]\frac{d^{2+1}G_n}{d\mu^{2+1}} - 2(1+1)\mu\frac{d^{1+1}G_n}{d\mu^{1+1}} + [n(n+1)-2]\frac{dG_n}{d\mu} = 0.$$

A further differentiation gives

$$[1-\mu^2]\frac{d^{2+2}G_n}{d\mu^{2+2}} - 2\mu\frac{d^{2+1}G_n}{d\mu^{2+1}} - 2(1+1)\frac{d^{1+1}G_n}{d\mu^{1+1}} - 2(1+1)\mu\frac{d^{1+2}G_n}{d\mu^{1+2}} +$$

$$[n(n+1) - 2]\frac{d^{1+1}G_n}{d\mu^{1+1}} = 0$$

or

$$[1-\mu^2]\frac{d^{2+2}G_n}{d\mu^{2+2}} - 2(1+2)\mu\frac{d^{1+2}G_n}{d\mu^{1+2}} + [n(n+1) - 2(2+1)]\frac{d^{0+2}G_n}{d\mu^{0+2}} = 0.$$

Successive differentiation m times gives

$$[1-\mu^2]\frac{d^{2+m}G_n}{d\mu^{2+m}} - 2(m+1)\mu\frac{d^{1+m}G_n}{d\mu^{1+m}} + [n(n+1)-m(m+1)]\frac{d^{0+m}G_n}{d\mu^{0+m}} = 0 \quad \ldots(3.33).$$

Hence $d^mG_n/d\mu^m$ satisfies equation 3.32. This implies that

$$H = \frac{d^mG_n}{d\mu^m}.$$

Therefore

$$G(\mu) = [1 - \mu^2]^{\frac{1}{2}m} \frac{d^m}{d\mu^m}\{P_{no}(\mu)\} \quad \ldots(3.34)$$

is a solution of equation 3.11. A more complete solution is obtained from equation 3.21 when

$$G(\mu) = A\,[1-\mu^2]^{\frac{1}{2}m}\frac{d^m}{d\mu^m}\{P_{no}(\mu)\} + B[\mu^2-1]^{\frac{1}{2}m}\frac{d^m}{d\mu^m}\{Q_{no}(\mu)\} \quad \ldots.(3.35).$$

Notes

(i) The series at 3.35 is not a convergent one unless at least one of either A or B is equal to zero.

If $A = 0$, the series converges when $|\mu| > 1$.

If $B = 0$, convergence occurs when $|\mu| < 1$.

$$P_{nm}(\mu) = [1-\mu^2]^{\frac{1}{2}m} \frac{d^m}{d\mu^m}\{P_{no}(\mu)\}$$

$$\overset{(3.29)}{=} \frac{1}{2^n\, n!} [1-\mu^2]^{\frac{1}{2}m} \frac{d^{m+n}}{d\mu^{m+n}}[(\mu^2-1)^n] \qquad \ldots(3.36)$$

is called an *associated Legendre function of the first kind.*

3.7 The function $p_{nm}(\mu)$

It can be seen from a combination of equations 3.19 and 3.36 that

$$P_{nm}(\mu) = \frac{(2n)!}{2^n(n!)^2}[1-\mu^2]^{\frac{1}{2}m}\frac{d^m}{d\mu^m}\left\{\mu^n + \sum_{r=1}^{k}(-1)^r \frac{n(n-1)\ \ldots(n-2r+1)}{2^r r!(2n-1)(2n-3)\ldots(2n-2r+1)} \times \mu^{n-2r}\right\}$$

$$= \frac{(2n)!}{2^n(n!)^2}[1-\mu^2]^{\frac{1}{2}m}\left\{ n(n-1)\ldots(n-m+1)\mu^{n-m} + \right.$$

$$\left.\sum_{r=1}^{k'}(-1)^r(n-2r)(n-2r-1)\ldots(n-m-2r+1)\frac{n(n-1)(n-2)\ldots(n-2r+1)}{2^r r!(2n-1)(2n-3)\ldots(2n-2r+1)}\mu^{n-m-2r}\right\};$$

where k' is different from k as the number of terms in the series depends on m. Obviously m cannot be greater than n as the derivative will be zero for all terms. In addition, there will be m fewer terms in the new series. A Legendre series has r in the range

$$0 \leq r \leq k \qquad \text{where} \qquad k = \begin{cases} \frac{n}{2} & \text{if } n \text{ even} \\ \frac{n-1}{2} & \text{if } n \text{ odd.} \end{cases}$$

The associated Legendre series will therefore have r in the range

$$0 \leq r \leq k' \qquad \text{where} \qquad k' = \begin{cases} \frac{n-m}{2} & \text{if } (n-m) \text{ even} \\ \frac{n-m-1}{2} & \text{if } (n-m) \text{ odd.} \end{cases}$$

Thus

$$P_{nm}(\mu) = \frac{(2n)!}{2^n(n!)^2}[1-\mu^2]^{\frac{1}{2}m}\frac{n!}{(n-m)!}\left\{\mu^{n-m} + \right.$$

$$\left.\sum_{r=1}^{k'}(-1)^r\frac{(n-m)(n-m-1)\ldots\ldots(n-m-2r+1)}{2^r r!(2n-1)(2n-3)\ldots(2n-2r+1)}\mu^{n-m-2r}\right\} \qquad \ldots(3.37).$$

It is convenient, as will be seen in section 4.4 to define the associated Legendre function $p_{nm}(\mu)$ which has desirable properties for surface integration over a sphere. $p_{nm}(\mu)$ is given by

$$p_{nm}(\mu) = \frac{(n-m)!}{n!} P_{nm}(\mu)$$

$$= \frac{(2n)!}{2^n (n!)^2}[1-\mu^2]^{\frac{1}{2}m}\left(\mu^{n-m} + \sum_{r=1}^{k'} (-1)^r \frac{(n-m)!}{(n-m-2r)!} \times \frac{(2n-2r)!}{(2n)!} \times \frac{n!}{(n-r)!} \times \mu^{n-m-2r}\right)$$

$$= \frac{(n-m)!}{2^n\, n!} [1-\mu^2]^{\frac{1}{2}m} \sum_{r=0}^{k'} (-1)^r \frac{(2n-2r)!}{r!(n-m-2r)!(n-r)!} \mu^{n-m-2r} \quad ..(3.38).$$

Similarly,

$$P_{nm}(\mu) = \frac{n!}{(n-m)!} p_{nm}(\mu)$$

$$= \frac{1}{2^n} [1-\mu^2]^{\frac{1}{2}m} \sum_{r=0}^{k'} (-1)^r \frac{(2n-2r)!}{r!(n-m-2r)!(n-r)!} \mu^{n-m-2r} \quad ...(3.39),$$

where

$$k' = \begin{cases} \frac{n-m}{2} & \text{if } (n-m) \text{ even} \\ \frac{n-m-1}{2} & \text{if } (n-m) \text{ odd.} \end{cases}$$

Note

As $m \leqq n$, $p_{nm}(\mu) \leqq P_{nm}(\mu)$,

the equality holding when $m = 0$. Thus the function $p_{nm}(\mu)$ has smaller magnitudes than $P_{nm}(\mu)$.

3.8 Some evaluations of Legendre functions

Lower degree spherical harmonic functions can be evaluated with relative ease from equations 3.29 and 3.38. They can also be expressed in terms of rectangular Cartesian co-ordinates in three dimensions when the direct consideration of figure 1.2 gives

$$x_1 = r \sin\theta \cos\lambda$$
$$x_2 = r \sin\theta \sin\lambda$$
$$x_3 = r \cos\theta$$

The harmonics up to (4,4) are listed in table 3.1.

n	m	K_{nm} *	$P_{nm}(\cos\theta)$	Equivalent polynomial
0	0	1	1	1
1	0	$\sqrt{3}$	$\cos\theta$	x_3
1	1	$\sqrt{3}$	$\sin\theta$	x_1, x_2
2	0	$\sqrt{5}$	$\frac{1}{2}[3\cos^2\theta-1]$	$\frac{1}{2}[2x_3^2-x_1^2+x_2^2]$
2	1	$\sqrt{(\frac{5}{3})}$	$3\sin\theta\cos\theta$	$3x_1x_3$, $3x_2x_3$
2	2	$\sqrt{(\frac{5}{12})}$	$3\sin^2\theta$	$3[x_1^2-x_2^2]$, $6x_1x_2$
3	0	$\sqrt{7}$	$\cos\theta[5\cos^2\theta-3]$	$x_3[2x_3^2-3x_1^2-3x_2^2]$
3	1	$\sqrt{(7/6)}$	$\frac{3}{2}\sin\theta[5\cos^2\theta-1]$	$\frac{3}{2}x_1[4x_3^2-x_1^2-x_2^2]$, $\frac{3}{2}x_2[4x_3^2-x_1^2-x_2^2]$
3	2	$\sqrt{(7/60)}$	$15\sin^2\theta\cos\theta$	$15x_3[x_1^2-x_2^2]$, $30x_1x_2x_3$
3	3	$\sqrt{(7/360)}$	$15\sin^3\theta$	$15[x_1^3-3x_1x_2^2]$, $15[x_1^2x_2-x_2^3]$
4	0	$\sqrt{9}$	$\frac{1}{8}[35\cos^4\theta-30\cos^2\theta+3]$	$\frac{1}{8}[8x_3^4-24x_3^2(x_1^2+x_2^2)+3(x_1^2+x_2^2)^2]$
4	1	$\sqrt{(\frac{9}{10})}$	$\frac{5}{2}[7\cos^3\theta-3\cos\theta]\sin\theta$	$\frac{5}{2}x_1[4x_3^3-3x_3(x_1^2+x_2^2)]$, $\frac{5}{2}x_2[4x_3-3x_3(x_1+x_2)]$
4	2	$\sqrt{(\frac{9}{180})}$	$\frac{15}{2}[7\cos^2\theta-1]\sin^2\theta$	$\frac{15}{2}(x_1^2-x_2^2)[6x_3^2-x_1^2-x_2^2]$, $\frac{30}{2}x_1x_2[6x_3-x_1-x_2]$
4	3	$\sqrt{(\frac{9}{2520})}$	$105\cos\theta\sin^3\theta$	$105x_3[x_1^3-3x_1x_2^2]$, $105x_3[3x_1^2x_2-x_2^3]$
4	4	$\sqrt{(\frac{9}{20160})}$	$105\sin^4\theta$	$105[x_1^4-6x_1^2x_2^2+x_2^4]$, $420x_1x_2[x_1^2+x_2^2]$

Table 3.1

Legendre functions with normalisation coefficients and equivalent polynomials to (4,4)

* Normalisation factor which is the coefficient of $P_{nm}(\mu)$ in equation 4.32

4. SPHERICAL HARMONICS

4.1 The general solution of Laplace's equation in spherical co-ordinates

It follows from equations 3.13, 3.35 and 3.39 that the expression

$$U = \frac{1}{r^{n+1}} P_{nm}(\mu)[C_{nm}\cos m\lambda + S_{nm}\sin m\lambda] \qquad \ldots(4.1),$$

where $\mu = \cos\theta$ and m,n are constants such that $m \leq n$ and $0 \leq n < \infty$ when $|\mu| \leq 1$, is a solution for Laplace's equation for potential in space unoccupied by matter. As n and m are arbitrary integral constants with only the restrictions defined above, a complete solution for U is given by

$$U = \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \sum_{m=0}^{n} P_{nm}(\mu)[C_{nm}\cos m\lambda + S_{nm}\sin m\lambda] \quad \ldots(4.2).$$

Equation 4.2 is said to express the potential at points external to a gravitating mass like the earth, in an infinite series of *Spherical Harmonics*, the term being attributed to Kelvin (Hobson 1965, p.119). The spherical harmonic generated for each value of n, e.g.,

$$U_n = \frac{1}{r^{n+1}} S_n$$
$$= \frac{1}{r^{n+1}} \sum_{m=0}^{n} P_{nm}(\mu)[C_{nm}\cos m\lambda + S_{nm}\sin m\lambda] \quad \ldots(4.3)$$

is called a solid harmonic or a spherical harmonic of *degree* n. The series of terms

$$S_n = \sum_{m=0}^{n} S_{nm} = \sum_{m=0}^{n} P_{nm}(\mu)[C_{nm}\cos m\lambda + S_{nm}\sin m\lambda] \quad \ldots(4.4)$$

is called a *surface harmonic*.

The value of the index m is called the *order* of the term S_{nm} which contributes to the surface harmonic S_n. A solid harmonic takes values in all regions of the space satisfying the conditions for Laplace's equation. A surface harmonic on the other hand, represents variations on a sphere whose radius is defined by r. Thus the set of surface harmonics

$$\sum_{n=0}^{\infty} S_n = \sum_{n=0}^{\infty} \sum_{m=0}^{n} S_{nm}$$

can represent the variations of any continuous function which takes values on the surface of a sphere. A surface harmonic of degree n has (2n+1) terms. For a proof of this statement see (ibid,p.122).

4.2 Orthogonality relationships between surface harmonics

Consider two surface harmonics S_n of degree n and S_ℓ of degree ℓ. The solid harmonics V_n and V_ℓ, given by

$$V_n = r^n S_n \qquad \text{and} \qquad V_\ell = r^\ell S_\ell$$

satisfy Laplace's equation in regions not occupied by matter, as can be seen from equation 3.12. Hence the application of Green's theorem, given in equation 2.7, to the scalars V_n and V_ℓ in the space V bounded by the spherical surface S and in which Laplace's equation is satisfied, gives

$$\iiint_V \left(V_n \nabla^2 V_\ell - V_\ell \nabla^2 V_n\right) dv = \iint_S \left(V_n \frac{\partial V_\ell}{\partial r} - V_\ell \frac{\partial V_n}{\partial r}\right) ds \overset{(2.24)}{=} 0.$$

On completing the differentiation and evaluation at $r = a$, where a is the radius of the sphere ,

$$\iint_S \left(V_n \ell\, r^{\ell-1} S_\ell - V_\ell n\, r^{n-1} S_n\right) ds = a^{\ell+n-1} \iint_S -S_n S_\ell (n-\ell)\, ds$$

$$= -a^{\ell+n-1}(n-\ell) \iint S_n S_\ell ds = 0.$$

Thus

$$\iint S_n S_\ell\, ds = 0 \qquad \text{if } n \neq \ell \qquad \ldots(4.5),$$

on noting that the integration is performed over the surface of a sphere. Equation 4.5 defines the orthogonal property of surface harmonics. This is a most important relationship as it enables surface harmonic functions to be eminently suitable for the analysis of data distributed on the surface of a sphere. This characteristic also permits functional manipulation of continuous representations of such data. The case when $n = \ell$ is dealt with in section 4.6.

4.3 Surface integration of a Legendre function

Consider an integral of the type shown at 4.5 when the surface harmonics are the specific functions

$$S_n = C_{no}\, P_{no}(\mu) \qquad \text{and} \qquad S_\ell = C'_{\ell o} P_{\ell o}(\mu)$$

respectively. Equation 4.5 will be satisfied if $n \neq \ell$.

In the instance when $n = \ell$, the surface integral I over the surface of the sphere which is considered to be of unit radius becomes

$$I = C_{no}\, C'_{no} \int_0^{\pi}\int_0^{\pi} [P_{no}(\mu)]^2 \sin\theta\, d\theta\, d\lambda \qquad .$$

As $\mu = \cos\theta$, $d\mu = -\sin\theta\, d\theta$. Change of variable gives

$$I = C_{no} C'_{no} \int_0^{2\pi} \int_0^{\pi} [P_{no}(\mu)]^2 \, d\mu \, d\lambda .$$

Partial integration, together with the use of equation 3.29 gives

$$= \frac{2\pi C_{no} C'_{no}}{2^{2n} (n!)^2} \int_{-1}^{1} \frac{d^n}{d\mu^n}[(\mu^2-1)^n] \frac{d^n}{d\mu^n}[(\mu^2-1)^n] \, d\mu .$$

On using the formula for integration by parts which can be written as

$$\int_{-1}^{1} u \frac{dv}{d\mu} \, d\mu = [u\, v]_{-1}^{1} - \int_{-1}^{1} v \frac{du}{d\mu} \, d\mu ,$$

the term to be integrated in the above equation(I') becomes

$$I' = \left[\frac{d^{n-1}}{d\mu^{n-1}}[(\mu^2-1)^n] \frac{d^n}{d\mu^n}[(\mu^2-1)^n] \right]_{-1}^{1} - \int_{-1}^{1} \frac{d^{n-1}}{d\mu^{n-1}}[(\mu^2-1)^n] \frac{d^{n+1}}{d\mu^{n+1}}[(\mu^2-1)^n] \times d\mu \quad ..(4.6).$$

The integrated term will always have a factor (μ^2-1) as a factor and will therefore be zero at both the upper and lower limits of the integration. Successive integration n times gives

$$I' = (-1)^n \int_{-1}^{1} (\mu^2-1)^n \frac{d^{2n}}{d\mu^{2n}}[(\mu^2-1)^n] \, d\mu .$$

As

$$\frac{d^{2n}}{d\mu^{2n}}[(\mu^2-1)^n] = (2n)! ,$$

and

$$\int_{-1}^{1} (\mu^2-1)^n d\mu = \int_{-1}^{1} (\mu-1)^n (\mu+1)^n d\mu =$$

$$= \left[(\mu-1)^n \frac{1}{n+1} (\mu+1)^{n+1} \right]_{-1}^{1} - \frac{n}{n+1} \int_{-1}^{1} (\mu-1)^{n-1} (\mu+1)^{n+1} \, d\mu$$

$$= (-1)^n \frac{(n!)^2}{(2n)!} \frac{2^{2n+1}}{2n+1}$$

due to the integrated term in the intermediate stage being zero at the limits while the integral, on partial integration a further n times gives

$$\int_{-1}^{1} (\mu^2-1)^n d\mu = (-1)^n \frac{(n!)^2}{(2n)!} \int_{-1}^{1} (\mu+1)^{2n} d\mu = (-1)^n \frac{(n!)^2}{(2n)!(2n+1)} [(\mu+1)^{2n+1}]_{-1}^{1} ,$$

I' is therefore given by

$$I' = \frac{(n!)^2}{2n+1} 2^{2n+1}$$

and

$$I = 2\pi \quad C_{no} C'_{no} \quad \frac{2}{2n+1} \qquad \ldots(4.7).$$

An important corollary is the result

$$\int_{-1}^{1} [p_{no}(\mu)]^2 d\mu = \frac{2}{2n+1} \qquad \ldots(4.8).$$

4.4 A property of $p_{nm}(\mu)$

The form of the associated Legendre function $p_{nm}(\mu)$ lends itself to satisfying the important relation

$$p_{nm}(\mu) = (-1)^m \, p_{n(-m)}(\mu) \qquad \ldots(4.9).$$

This is obtained from a consideration of the function

$$U = [x_3 + i\, x_1]^n \quad ,$$

where $i = \sqrt{-1}$, can be shown to satisfy Laplace's equation $\nabla^2 U = 0$, as a consideration of figure 1.2 gives

$$\nabla U = n(x_3 + ix_1)^{n-1}[i\ 1 + 3]$$

and

$$\nabla^2 U = n(n-1)(x_3 + ix_1)^{n-2}[\,-1 + 1] = 0.$$

The conversion of the Cartesian co-ordinates to spherical co-ordinates gives

$$U = (x_3 + ix_1)^n = r^n[\cos\theta + i\sin\theta\cos\lambda]^n$$

$$= r^n[\mu + i(1-\mu^2)^{\frac{1}{2}}\cos\lambda]^n \qquad \ldots(4.10),$$

where $|\mu| < 1$ and $i(1-\mu^2)^{\frac{1}{2}} = (\mu^2-1)^{\frac{1}{2}}$, U being a solution of Laplace's equation. The expression on the right of equation 4.4 can be replaced by a function of the term

$$\left(\mu + (\mu^2-1)^{\frac{1}{2}} e^{i\lambda}\right)^2 - 1.$$

As

$$e^{im\lambda} = \cos m\lambda + i\sin m\lambda$$

and

$$\left(\mu + (\mu^2-1)^{\frac{1}{2}} e^{i\lambda}\right)^2 - 1 = \mu^2 + 2\mu(\mu^2-1)^{\frac{1}{2}} e^{i\lambda} + (\mu^2-1)e^{2i\lambda} - 1$$

$$= \mu^2[1+\cos 2\lambda] + i\mu^2\sin 2\lambda + 2\mu(\mu^2-1)^{\frac{1}{2}} e^{i\lambda} - (e^{2i\lambda} + 1)$$

$$= 2\mu^2\cos\lambda[\cos\lambda + i\sin\lambda] + 2\mu(\mu^2-1)^{\frac{1}{2}}e^{i\lambda} - (e^{2i\lambda} + e^{i\lambda}e^{-i\lambda})$$

$$= \frac{2}{e^{-i\lambda}}\left(\mu^2\cos\lambda + \mu(\mu^2-1)^{\frac{1}{2}} - \frac{e^{i\lambda} + e^{-i\lambda}}{2}\right)$$

$$= \frac{2}{e^{-i\lambda}}\left((\mu^2-1)\cos\lambda + \mu(\mu^2-1)^{\frac{1}{2}}\right),$$

it follows that

$$r^n\left(\mu + (\mu^2-1)^{\frac{1}{2}}\cos\lambda\right)^n = \frac{e^{-in\lambda}r^n}{2^n(\mu^2-1)^{\frac{1}{2}n}}\left(\left(\mu + (\mu^2-1)^{\frac{1}{2}}e^{i\lambda}\right)^2 - 1\right)^n \quad \ldots(4.11).$$

The term within the bracket on the right and raised to the power n can be treated as a function of the form

$$f(x+h) = [(\mu+h)^2-1]^n$$

by Taylor's theorem, when

$$f(x+h) = \sum_{n=0} \frac{1}{n!} h^n \frac{d^n}{dx^n}[f(x)].$$

On putting

$$h = (\mu^2-1)^{\frac{1}{2}}e^{i\lambda}, \qquad \text{equation 4.11 can be written}$$

as

$$r^n\left(\mu + (\mu^2-1)^{\frac{1}{2}}\cos\lambda\right)^n = \frac{r^n e^{-in\lambda}}{2^n(\mu^2-1)^{\frac{1}{2}n}}\sum_{p=0}^{2n}\frac{(\mu^2-1)^p}{p!}e^{ip\lambda}\frac{d^p}{d\mu^p}[(\mu^2-1)^n]$$

....(4.12).

The upper limit of the summation includes all non-zero terms, the higher differential coefficients being zero. The terms on the right of equation 4.12 can be considered to be made up of three sets of quantities:-

i The terms obtained when $p = n$

The relevant term is

$$r^n\left(\frac{1}{2^n n!}\frac{d^n}{d\mu^n}[(\mu^2-1)^n]\right) \overset{(3.29)}{=} r^n P_{no}(\mu).$$

ii The n terms which occur when $n+1 \leqq p \leqq 2n$

$$\text{Put } m = p - n \qquad \text{or} \qquad p = n + m.$$

The relevant terms form the series

$$r^n\left(\frac{1}{2^n}\sum_{m=1}^{n}\frac{1}{(n+m)!}(\mu^2-1)^{\frac{1}{2}m}e^{im\lambda}\frac{d^{n+m}}{d\mu^{n+m}}[(\mu^2-1)^n]\right)$$

$$= r^n\left(\frac{1}{2^n}\sum_{m=1}^{n}\frac{1}{(n+m)!}(\mu^2-1)^{\frac{1}{2}m}\frac{d^{n+m}}{d\mu^{n+m}}[(\mu^2-1)^n](\cos m\lambda + i\sin m\lambda)\right)$$

iii The n terms which occur in the range $0 \leq p \leq (n-1)$

Replace p by m according to $m = p + n$ or $p = m - n$. The series so formed is

$$r^n\left(\frac{1}{2^n}\sum_{m=1}^{n}\frac{1}{(n-m)!}(\mu^2-1)^{-\frac{1}{2}m}\frac{d^{n-m}}{d\mu^{n-m}}[(\mu^2-1)^n](\cos m\lambda - i\sin m\lambda)\right).$$

Thus

$$\left(\mu+(\mu^2-1)^{\frac{1}{2}}\cos\lambda\right)^n = p_{no}(\mu) + \frac{1}{2^n}\sum_{m=1}^{n}\left(\frac{1}{(n-m)!}(\mu^2-1)^{-\frac{1}{2}m}\frac{d^{n-m}}{d\mu^{n-m}}[(\mu^2-1)^n]e^{-im\lambda} + \frac{1}{(n+m)!}(\mu^2-1)^{\frac{1}{2}m}\frac{d^{n+m}}{d\mu^{n+m}}[(\mu^2-1)^n]e^{im\lambda}\right).$$

As the expression on the right should not contain terms in sin $m\lambda$, all coefficients of sin $m\lambda$ should be zero. Thus

$$\frac{1}{(n-m)!}(\mu^2-1)^{-\frac{1}{2}m}\frac{d^{n-m}}{d\mu^{n-m}}[(\mu^2-1)^n] = \frac{1}{(n+m)!}(\mu^2-1)^{\frac{1}{2}m}\frac{d^{n+m}}{d\mu^{n+m}}[(\mu^2-1)^n] \quad ,m=1,n. \qquad ...(4.13)$$

$p_{nm}(\mu)$ is given by equations 3.36 and 3.38 as

$$p_{nm}(\mu) = \frac{(n-m)!}{2^n(n!)^2}(1-\mu^2)^{\frac{1}{2}m}\frac{d^{m+n}}{d\mu^{m+n}}[(\mu^2-1)^n] \qquad ...(4.14),$$

while $p_{n(-m)}(\mu)$ is given by

$$p_{n(-m)}(\mu) = \frac{(n+m)!}{2^n(n!)^2}(1-\mu^2)^{-\frac{1}{2}m}\frac{d^{m-n}}{d\mu^{m-n}}[(\mu^2-1)^n] \qquad ...(4.15).$$

Equation 4.13 relates the quantities defined in equations 4.14 and 4.15. As

$$(\mu^2-1)^{\frac{1}{2}m} = (-1)^{\frac{1}{2}m}(1-\mu^2)^{\frac{1}{2}m} \quad \text{and} \quad (\mu^2-1)^{-\frac{1}{2}m} = (-1)^{-\frac{1}{2}m}(1-\mu^2)^{-\frac{1}{2}m},$$

appropriate substitution into equation (4.13) gives

$$p_{nm}(\mu) = (-1)^m p_{n(-m)}(\mu) \qquad ...(4.9).$$

4.5 The integral $\int_{-1}^{1}[P_{nm}(\mu)]^2 d\mu$

The relation derived in section 4.4 provides a means of evaluating the above integral.

$$\int_{-1}^{1} [P_{nm}(\mu)]^2 d\mu \overset{(4.8)}{=} (-1)^m \int_{-1}^{1} P_{nm}(\mu)\, P_{n-m}(\mu)\, d\mu$$

$$\overset{(4.14),(4.15)}{=} \frac{(n-m)!\,(n+m)!}{(n!)^4\, 2^{2n}} (-1)^m \int_{-1}^{1} \frac{d^{n+m}}{d\mu^{n+m}}[(\mu^2-1)^n] \frac{d^{n-m}}{d\mu^{n-m}}[(\mu^2-1)^n]\, d\mu \; .$$

Integration m times by parts on the lines used in obtaining equation 4.6 gives

$$\int_{-1}^{1} [P_{nm}(\mu)]^2 d\mu = \frac{(n-m)!\,(n+m)!}{(n!)^4\, 2^{2n}} \int_{-1}^{1} \frac{d^n}{d\mu^n}\left((\mu^2-1)^n\right) d\mu$$

$$\overset{(3.29)}{=} \frac{(n-m)!\,(n+m)!}{(n!)^2} \int_{-1}^{1} [P_{no}(\mu)]^2 d\mu$$

$$\overset{(4.8)}{=} \frac{(n-m)!\,(n+m)!}{(n!)^2} \cdot \frac{2}{2n+1} \qquad \ldots(4.16).$$

Note

$$\int_{0}^{4\pi} [P_{nm}(\mu)]^2 d\sigma = \frac{4\pi}{2n+1} \frac{(n-m)!\,(n+m)!}{(n!)^2} \qquad \ldots(4.17).$$

4.6 The surface integration of products of harmonics

i. General principles

The results of the integration of the products of surface harmonics over a surface is of considerable importance in physical geodesy. The orthogonality relationships which exist between surface harmonics of different degrees have already been established in section 4.2. In the case of products of surface harmonic series of the same degree, the following result is obtained. Let S_n and S'_n be two surface harmonics of degree n, given by the equations

$$S_n = \sum_{m=0}^{n} p_{nm}(\mu)[C_{nm}\cos m\lambda + S_{nm}\sin m\lambda]$$

and

$$S'_n = \sum_{m'=0}^{n} p_{nm'}(\mu)[C'_{nm'}\cos m'\lambda + S'_{nm'}\sin m'\lambda].$$

Then

$$I = \iint S_n S'_n ds$$

$$= \int_0^{2\pi}\int_{-1}^{1} \sum_{m=0}^{n} p_{nm}(\mu)[C_{nm}\cos m\lambda + S_{nm}\sin m\lambda] \sum_{m'=0}^{n} p_{nm'}(\mu)[C'_{nm'}\cos m'\lambda + S'_{nm'}\sin m'\lambda]\, d\mu\, d\lambda.$$

The terms on the right can be considerably simplified by observing that

$$\int_0^{2\pi} \cos m\lambda \sin m'\lambda\, d\lambda = \int_0^{2\pi} \sin m\lambda \sin m'\lambda\, d\lambda = \int_0^{2\pi} \cos m\lambda \cos m'\lambda\, d\lambda$$

$$= 0 \qquad \text{if} \quad m \neq m' \qquad \ldots(4.18)$$

When $m = m'$, the longitude dependent terms take one of three forms

$$\int_0^{2\pi} \sin^2 m\lambda\, d\lambda \;, \quad \int_0^{2\pi} \cos^2 m\lambda\, d\lambda \quad \text{or} \quad \tfrac{1}{2}\int_0^{2\pi} \sin 2m\lambda\, d\lambda .$$

The last integral is zero while the first two can be written as

$$\tfrac{1}{2}\int_0^{2\pi} [1 \pm \cos 2m\lambda]\, d\lambda = \pi \qquad \ldots(4.19).$$

This series reduces to

$$\iint S_n S'_n ds = \pi\int_{-1}^{1} \sum_{m=0}^{n}[p_{nm}(\mu)]^2 \left(C_{nm}C'_{nm} + S_{nm}S'_{nm}\right) d\mu$$

$$= \frac{2\pi}{2n+1} \frac{1}{(n!)^2} \sum_{m=0}^{n} \left(C_{nm}C'_{nm} + S_{nm}S'_{nm}\right)[(n-m)!\,(n+m)!] \,.(4.20).$$

ii. The product $\int_0^{dS=4\pi} S_n \; p_{no}(\mu)\, d\mu$

The definition of the pole of a surface harmonic must precede the evaluation of the above integral. This is defined as the radius corresponding to $\theta = 0$ in the case where

$$S_n = \sum_{m=0}^{n} p_{nm}(\mu) \; [C_{nm}\cos m\lambda + S_{nm}\sin m\lambda] \qquad \ldots(4.21)$$

where $\mu = \cos\theta$.

For a more detailed discussion see (Hobson 1965, pp.129 et seq.).

The integral

$$\int_0^{2\pi}\int_{-1}^{1} \sum_{m=0}^{n} p_{nm}(\mu)[\,C_{nm}\cos m\lambda + S_{nm}\sin m\lambda]\; p_{no}(\mu)\,d\mu\; d\lambda$$

$$\overset{(4.5)}{=} \; 2\pi\, C_{no} \int_{-1}^{1} [p_{no}(\mu)]^2\, d\mu$$

$$\overset{(4.8)}{=} \; \frac{4\pi}{2n+1}\; C_{no} \qquad \ldots(4.22).$$

A study of equation 3.36 in the light of the observations following equation 4.6 indicates that

$$p_{nm}(1) = 0 \qquad \ldots(4.23)$$

Also when μ in equation 3.32 equals 1,

$$\frac{1}{r} = \frac{1}{R_p}[1-t]^{-1} = \frac{1}{R_p}\sum_{n=0}^{\infty} t^n$$

$$\overset{(3.26)}{=} \; \frac{1}{R_p}\sum_{n=0}^{\infty} t^n p_{no}(1).$$

Thus

$$p_{no}(1) = 1 \qquad \text{for all } n \qquad \ldots(4.24).$$

Similarly, when $\mu = -1$ (i.e., $\theta = \pi$)

$$\frac{1}{r} = \frac{1}{R_p}\,[1+t]^{-1} = \frac{1}{R_p}\sum_{n=0}^{\infty} (-1)^n\, t^n$$

Hence

$$p_{no}(-1) = (-1)^n \qquad \ldots(4.25).$$

It can therefore be seen from equation 4.21 that

$$S_n(0,0) = C_{no}$$

and

$$\int_0^{dS=4\pi} S_n \, p_{no}(\mu) \, dS = \frac{4\pi}{2n+1} S_n(0,0) \qquad \ldots(4.26),$$

where $S_n(0,0)$ is the value of S_n at the pole of $p_{no}(\mu)$. This result also implies that if the pole (θ',λ') of the Legendre function is *not* that of the surface harmonic(i.e., 0,0), the relation

$$\int_0^{dS=4\pi} S_n \, p_{no}(\cos\psi) \, dS = \frac{4\pi}{2n+1} S_n(\theta',\lambda') \qquad \ldots(4.27)$$

holds, where ψ is the co-latitude with respect to (θ',λ') as pole.

iii. Numerical characteristics of the surface harmonic S_{nm} *of degree* n *and order* m.

The general associated Legendre function is given by equation 3.38 as

$$p_{nm}(\mu) = \frac{(n-m)!}{2^n \, n!}[1-\mu^2]^{\frac{1}{2}m} \sum_{r=0}^{k'} (-1)^r \frac{(2n-2r)!}{r!(n-m-2r)!(n-r)!} \mu^{n-m-2r}$$

and its surface integral over a unit sphere has been shown in equation 4.17 to satisfy

$$\int_0^{dS=4\pi} [p_{nm}(\mu)]^2 \, dS = \frac{4}{2n+1} \frac{(n-m)!(n+m)!}{(n!)^2} .$$

Consequently the magnitude of the function $p_{nm}(\mu)$ is dependent on n and m which makes it unsuitable for analysis as the numerical values taken by the coefficients C_{nm}, S_{nm} are influenced by those of $p_{nm}(\mu)$. It is therefore common practice to use *normalised* (sometimes called fully normalised) harmonics $\bar{p}_{nm}(\mu)$ (e.g., Heiskanen & Moritz 1967, p.31) which satisfy the relation

$$\iint_{\sigma} [\bar{p}_{nm}(\mu)]^2 \, dS = 4\pi .$$

It can be seen from equation 4.17 that

$$\bar{p}_{nm}(\mu) = n! \left(\frac{2n+1}{(n-m)!(n+m)!}\right)^{\frac{1}{2}} .$$

It is preferable to normalise the surface harmonic S_{nm} itself according to the relation

$$\int_0^{dS=4\pi} [\bar{S}_{nm}]^2 \, dS = 4\pi \qquad \ldots(4.28),$$

where $S_{nm} = P_{nm}(\mu) \begin{cases} \cos m\lambda \\ \sin m\lambda \end{cases}$

when the use of equations 4.17 and 4.19 gives

$$\int_0^{dS=4\pi} [S_{nm}]^2 \, dS = \frac{2\pi}{2n+1} \frac{(n-m)!\,(n+m)!}{(n!)^2} \quad , \quad m \neq 0,$$

whence

$$\bar{S}_{nm} = \left\{\frac{(2n+1)(2-\delta_{om})}{(n-m)!\,(n+m)!}\right\}^{\frac{1}{2}} S_{nm} \qquad \ldots(4.29),$$

where δ_{om} is the Kronecker delta which satisfies

$$\delta_{ik} = \begin{cases} 0 & \text{if } i \neq k \\ 1 & \text{if } i = k \end{cases} \qquad \ldots(4.30).$$

Normalised harmonics have the property expressed in equation 4.28 and comprise surface harmonic series on combination with the normalised coefficients $\bar{C}_{nm}$, $\bar{S}_{nm}$. These coefficients approach the same order of magnitude and hence lend themselves to the reliable analysis of data . Hence normalised coefficients ($\bar{C}_{nm}$, $\bar{S}_{nm}$) are related to ordinary coefficients (C_{nm},S_{nm}) by the relation

$$\bar{C}_{nm} = \frac{1}{n!} \left\{\frac{(n+m)!\,(n-m)!}{(2n+1)(2-\delta_{om})}\right\}^{\frac{1}{2}} C_{nm} \qquad \ldots(4.31).$$

If, on the other hand, the function considered was $P_{nm}(\mu)$, the use of equations 3.37 and 3.38 gives

$$\bar{P}_{nm}(\mu) = \left\{\frac{(2-\delta_{0m})(2n+1)(n-m)!}{(n+m)!}\right\}^{\frac{1}{2}} P_{nm}(\mu) \qquad \ldots(4.32)$$

and the associated coefficients are related by equations of the form

$$\bar{C}_{nm} = \left\{\frac{(n+m)!}{(2-\delta_{om})(2n+1)(n-m)!}\right\}^{\frac{1}{2}} C_{nm} \qquad \ldots(4.33).$$

4.7 The interpretation of surface harmonics

A set of spherical harmonics as described in equation 4.2 is adequate for the representation of the gravitational potential U at all points in space exterior to the gravitating body.

This definition can be considered to apply to a space defined by a family of concentric spheres of radius r when the variations in on a given sphere can be represented by the set of surface harmonics

$$U_{(r=const)} = \sum_{n=0}^{\infty} S_n = \sum_{n=0}^{\infty} \sum_{m=0}^{n} S_{nm}$$

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{n} p_{nm}(\cos\theta)[C_{nm}\cos m\lambda + S_{nm}\sin m\lambda] \quad ..(4.34),$$

where (θ,λ) are the set of surface co-ordinates defining position on the sphere. Variations with co-latitude θ in U are a consequence of the term $p_{nm}(\cos\theta)$ while those with longitude are due to the terms $\cos m\lambda$ and $\sin m\lambda$. Three distinct cases arise.

(i) When $m = 0$.
(ii) When $0 < m < n$
(iii) When $m = n$.

i. m = 0

In this case, the coefficient S_{no} is not of consequence as $\sin m\lambda = 0$ while $\cos m\lambda = 1$. Thus

$$S_{no} = C_{no}\, p_{no}(\mu)$$

where $\mu = \cos\theta$. It can be seen from equation 3.19 that the series for $p_{no}(\mu)$ has n roots, being a polynomial in μ. In addition, the results of equations 4.24 and 4.25 indicate that $p_{no}(\mu)$ takes the values +1 at $\theta = 0$ and $(-1)^n$ at $\theta = \pi$. Consequently the function varies with co-latitude being rotationally symmetric about the axis $\theta = 0$, with nodal points tracing out parallels of latitude. The nett effect is one of alternate bands of positive and negative values akin to the two banded function sin x in the range $0 < x < 2\pi$, which is positive in $0 < x < \pi$ and negative in $\pi < x < 2\pi$.

The function $p_{oo}(\mu)$ does not change in magnitude between the poles while $p_{1o}(\mu)$ has one nodal value in this range and satisfies

$$p_{1o}(1) = 1 \qquad \text{and} \qquad p_{1o}(-1) = -1$$

at the poles. Such a function is asymmetric about the equator, this being a characteristic of all odd degree Legendre functions. Similarly, it can be seen directly from equation 3.19, which is a power series in μ with an increment μ^2, reducing from a maximum index n, that even degree Legendre functions are symmetrical about the equator given by $\theta = \frac{1}{2}\pi$ as

$$\mu = \cos\theta \qquad \text{and} \qquad [\cos(\tfrac{1}{2}\pi+x)]^2 = [\cos(\tfrac{1}{2}\pi-x)]^2.$$

This class of surface harmonic is called a *zonal harmonic*, an example being illustrated in figure 4.1.

Summary

(i) A zonal harmonic of degree n has n zeros between the poles giving rise to (n+1) bands between n nodal parallels (or circles of equal θ), taking the values 1 at $\theta = 0$ and $(-1)^n$ at $\theta = \pi$.

(ii) The lower the degree of the harmonic, the wider the spacing between the nodal parallels.

(iii) Even degree zonal harmonics represent variations which are symmetrical about the equator while those of odd degree represent asymmetry with respect to it.

ii $0 < m < n$

The appropriate surface harmonic in this case is

$$S_{nm} = P_{nm}(\mu)[C_{nm}\cos m\lambda + S_{nm}\sin m\lambda] .$$

$P_{nm}(\mu)$ is a polynomial of degree (n-m) and is given by equation 3.39. This associated Legendre function takes the value zero at the poles. In all other respects it is similar to the Legendre function $P_{nm}(\mu)$ and has (n-m) zeros along parallels of latitude which are symmetrical about the equator if (n-m) is even and asymmetrical if (n-m) is odd. The function therefore has (n-m+1) bands between the poles and the (n-m) nodal lines mentioned above. In addition, the term $\cos m\lambda$ has m zeros in the range $0 < \lambda < \pi$, the nodal lines coinciding with meridians at a spacing of λ/m apart.

The resulting network of nodal lines along meridians and parallels, as illustrated in figure 4.2, has given rise to the application of the term *tesseral harmonic* to this type of surface function on a sphere.

Notes

(i) No tesseral harmonics are admissible in the analysis of a function which has no variation with longitude on the surface of a sphere.

(ii) Low degree tesseral harmonics have widely spaced nodal lines. Some of these can be related to low order inertia tensors of the earth's gravitational field if the function being represented is the geopotential.

iii. $m = n$

The limiting tesseral harmonic, on the basis of the

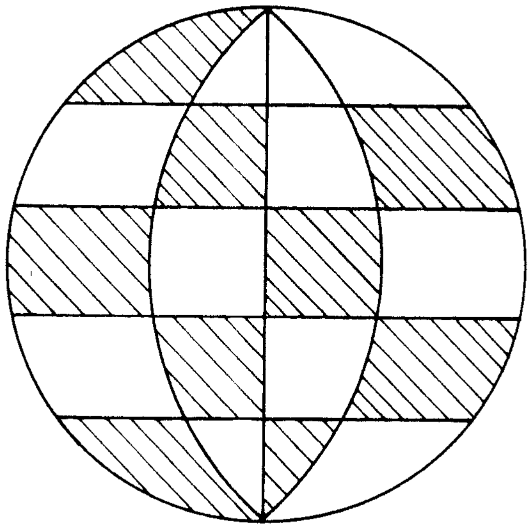


FIG. 4.2

The Tesseral harmonic $P_{84}(\mu)$

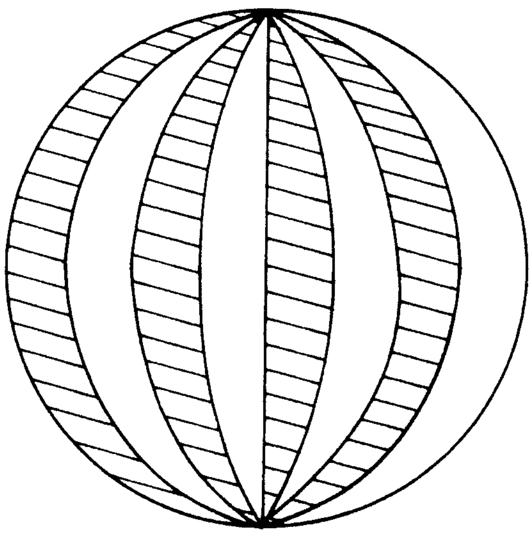


FIG. 4.3

The Sectional harmonic $P_{88}(\mu)$

above development, will have m nodal meridians in $0 < \lambda < \pi$, with zeros at the poles but not along any parallels. The resulting harmonic is called a *sectorial harmonic*, being illustrated in figure 4.3.

iv. Conclusion

Surface harmonic series provide a useful analytical method for the representation of data with a continuous distribution on the surface of a sphere for the following reasons.

(i) Surface harmonic series satisfy Laplace's equation.

(ii) The functions are capable of geometrical interpretation.

(iii) This facility can be used to relate certain characteristics of the functions represented to other physical properties.

(iv) The representation by surface harmonics involves, at least in theory, an infinite series thus restricting their consideration, in the strictest sense, to mathematical manipulation alone. In practice, significant applications are possible with the use of a limited number of terms if their geometrical significance is fully appreciated. Their usage can be considered to be analagous to fitting a curve to a continuous set of data, the only difference being the use of surface co-ordinates instead of plane rectangular reference frames. The higher the degree of the harmonic function used in the analysis of data distributed on a sphere, the more accurately will surface harmonics represent the local fluctuations of the quantity being analysed.

The usage is discussed further in section 7.

4.8 Statistical properties of surface harmonics

Consider a function x which can be represented by a set of fully normalised surface harmonics as

$$x(\theta,\lambda) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} x_{nm} ,$$

where

$$x_{nm} = \bar{P}_{nm}(\cos\theta)[\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda].$$

The value of the function at the pole is given by

$$x(0,0) = \sum_{n=0}^{\infty} \bar{C}_{no} .$$

The covariance $C(\theta)$ of the function at a distance θ from the pole and given by

$$C(\theta) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} C_{nm}(\theta)$$

is defined by the equation

$$C_{nm}(\theta) = \frac{1}{2\pi}\int_0^{2\pi} \overline{P}_{nm}(\cos\,\theta)\,[\overline{C}_{nm}\cos\,m\lambda + \overline{S}_{nm}\sin\,m\lambda]\,\sum_{n'=0}^{\infty}\overline{C}_{n'0}\,d\lambda$$

$$= \begin{cases} 0 & ,\ m \neq 0 \\ \overline{P}_{no}(\cos\,\theta)\overline{C}_{no}\displaystyle\sum_{n'=0}^{\infty}\overline{C}_{n'o} & ,\ m = 0 \end{cases} \qquad \ldots(4.35).$$

This covariance function is also known as the power spectrum (Kaula 1967, p.87), being dependent on θ alone, and has coefficients which are characteristic of the distribution of X. The general covariance function $C(\theta)$ given by

$$C(\theta) = \sum_{n=0}^{\infty}\sum_{n'=0}^{\infty}\overline{P}_{no}(\cos\,\theta)\,\overline{C}_{no}\,\overline{C}_{n'o}$$

can also be represented by the equation

$$C(\theta) = \sum_{n=0}^{\infty} c_n\,\overline{P}_{no}(\cos\,\theta)\ ,$$

where

$$c_n = \overline{C}_{no}\sum_{n'=0}^{\infty}\overline{C}_{n'o} \qquad \ldots(4.36)$$

and its magnitude can be obtained by the use of the equation

$$c_n = \frac{1}{4\pi}\iint C(\theta)\,\overline{P}_{no}(\cos\,\theta)\,ds \qquad \ldots(4.37)$$

Hence it is possible to define the coefficients in the zonal harmonic series which represent the covariance function through the equation

$$C(\theta) = \sum_{n=0}^{\infty} c_n\,\overline{P}_{no}(\cos\,\theta) = \sum_{n=0}^{\infty}\overline{P}_{no}(\cos\,\theta)\overline{C}_{no}\sum_{n'=0}^{\infty}\overline{C}_{n'o} \qquad \ldots(4.38)$$

if $C(\theta)$ is known.

The expressions derived so far refer to the covariance of the quantity with respect to the pole. The mean covariance over a sphere is the mean of the covariance at all points on the surface of the sphere each treated as the pole in turn. If c_n were treated as a function of position on a sphere, it could be represented by a surface harmonic series of the type

$$C(\theta,\lambda) = \sum_{n=0}^{\infty} c_n(\theta,\lambda) = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\overline{P}_{nm}(\mu)\,[\overline{a}_{nm}\cos\,m\lambda + \overline{b}_{nm}\sin\,m\lambda].$$

The mean value σ_n^2 of the square of the covariance function is given by the series of equations

$$M\{c_n^2\} = \sigma_n^2 = \frac{1}{4\pi}\iint c_n^2 \, ds$$

$$\overset{(4.5)}{=} \frac{1}{4\pi}\iint \sum_{m=0}^{n} \overline{P}^2_{nm}(\mu)\left(\overline{a}^2_{nm}\cos m\lambda + \overline{b}^2_{nm}\sin m\lambda\right) ds$$

$$\overset{(4.28)}{=} \sum_{m=0}^{n} \left(\overline{a}^2_{nm} + \overline{b}^2_{nm}\right) \qquad \ldots(4.39).$$

σ_n^2 is the average square of the n-th degree harmonic in the representation of the covariance function , being analogous to the variance and is hence termed the *degree variance* (Kaula 1966a,p.5304).

The magnitude of the degree variance of wave number n of a set of harmonic coefficients is a measure of the variability of the function. These power spectrum values σ_n^2 are of significance only if the coefficients $\overline{a}_{nm}, \overline{b}_{nm}$ were completely independent of one another. Consideration should otherwise be paid to moments other than the second.

Notes

(i) The application of degree variances to any function could give a measure of local fluctuations. The greater the values of σ_n^2 with increase of n, the greater the variations with short wave-length.

(ii) The degree variances remain invariant on rotation of axes and are thus a fundamental property of a given distribution of a function on a sphere(Kaula 1959,p.52). For an example of the magnitude of the degree variances of gravity anomalies, see (Kaula 1966, p.5304).

4.9 The relevance of spherical harmonic expansions in the vicinity of a gravitating body

The potential U in space external to a gravitating body satisfies Laplace's equation and hence can be completely represented by a spherical harmonic series of the type

$$U \overset{(4.2)}{=} \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \sum_{m=0}^{n} P_{nm}(\mu)\left[a_{nm}\cos m\lambda + b_{nm}\sin m\lambda\right] .$$

It is convenient, for the purpose of analysis, to express U by the relation

$$U = \frac{kM}{r} \sum_{n=0}^{\infty} \left(\frac{a_e}{r}\right)^n \sum_{m=0}^{n} P_{nm}(\mu)\left(C_{nm}\cos m\lambda + S_{nm}\sin m\lambda\right) \qquad \ldots(4.40),$$

where a_e is the equatorial radius of the gravitating body, assumed ellipsoidal, M its mass and k the gravitational constant. The expression of equation 4.2 in this form has a distinct advantage in the case of the earth as all coefficients C_{nm}, S_{nm} are dimensionless quantities. C_{oo} in such a representation has the greatest magnitude, being unity. All other coefficients, with the exception of C_{2o} are of order 10^{-6}, the latter having an order of magnitude 10^{-3}.

Two important factors should be borne in mind when using spherical harmonic functions to analyse the gravitational field of a body like the earth with its nearly oblate spheroidal symmetry.

Firstly, the ratio (a_e/r) is less than unity at exterior points. If $r = a_e + dr$, the use of the binomial theorem gives

$$\left(\frac{a_e}{r}\right)^{n+1} = \left(1 + \frac{dr}{a_e}\right)^{-(n+1)} = \sum_{i=0}^{n+1} (-1)^i \binom{n+1}{i} \left(\frac{dr}{a_e}\right)^i \qquad \ldots(4.41)$$

where

$$\binom{n+1}{i} = \frac{(n+1)!}{i!(n-i+1)!} \qquad \ldots(4.42).$$

This ratio has the characteristic of scaling down the effect of higher degree harmonics, as can be seen from a study of table 4.1 which sets out the results obtained on evaluating equation 4.41 for different values of n and dr. It is well known that the potential of any gravitating body acts as a point mass over great distances. This effect is completely represented by the first term in equation 4.40 (i.e., when $n = 0$). The effect of the other variations as represented by the other coefficients C_{nm},S_{nm} are made negligible due to the influence of the factor $(a_e/r)^n$ which is small. Consequently the effect of the higher degree harmonics in U becomes significant as this ratio approaches unity.

The characteristics exhibited by $(a_e/r)^{n+1}$, as given in table 4.1 indicate that the effect of higher degree harmonics is considerably damped with increase of elevation above the surface of the earth. Thus the motion of near earth satellites are more significantly affected by harmonics of lower degree than those where n is large. The effect of some higher degree harmonic terms is enhanced by the phenomenon of resonance. For details see section 8.3.

It can therefore be concluded that some ambiguity may exist in the values of harmonic coefficients determined from the orbital perturbations of near earth satellites due to the variable damping effect. Thus the harmonic of degree 20 will have its effect damped by 50% in relation to its value at the surface of the earth, at an elevation of 200 km, while that of degree 3 will be damped by only 12%. Higher degree harmonics as determined by this method will therefore always be subject to some uncertainty which restricts their *direct* application at the surface of the earth.

a = 6 378 160 metres

dr (km)	degree (n) of harmonic) 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.998	0.998	0.997	0.996	0.995	0.995	0.994	0.993	0.992	0.991	0.991	0.990	0.989	0.988	0.988	0.986	0.985	0.984	0.984	0.984
10	0.997	0.995	0.994	0.992	0.991	0.989	0.988	0.986	0.984	0.983	0.981	0.980	0.978	0.977	0.975	0.974	0.972	0.971	0.969	0.968
20	0.994	0.991	0.988	0.984	0.981	0.978	0.975	0.972	0.969	0.966	0.963	0.960	0.957	0.954	0.951	0.948	0.945	0.942	0.939	0.936
50	0.985	0.977	0.969	0.962	0.954	0.947	0.939	0.932	0.925	0.918	0.911	0.903	0.896	0.889	0.883	0.876	0.869	0.862	0.855	0.849
100	0.969	0.954	0.940	0.925	0.911	0.897	0.883	0.869	0.856	0.843	0.830	0.817	0.804	0.792	0.780	0.768	0.756	0.744	0.733	0.721
200	0.940	0.912	0.884	0.857	0.831	0.806	0.781	0.757	0.734	0.712	0.690	0.669	0.649	0.629	0.610	0.591	0.574	0.556	0.539	0.523
300	0.912	0.871	0.832	0.795	0.759	0.725	0.692	0.661	0.632	0.603	0.576	0.550	0.524	0.502	0.479	0.458	0.437	0.418	0.399	0.381
400	0.885	0.833	0.784	0.738	0.694	0.653	0.615	0.578	0.544	0.512	0.482	0.454	0.427	0.402	0.378	0.356	0.335	0.315	0.296	0.279
500	0.860	0.797	0.739	0.686	0.636	0.590	0.547	0.507	0.470	0.436	0.404	0.375	0.348	0.322	0.299	0.277	0.257	0.238	0.221	0.205
600	0.835	0.764	0.698	0.637	0.583	0.533	0.487	0.445	0.407	0.372	0.340	0.311	0.284	0.259	0.237	0.217	0.198	0.181	0.166	0.151
700	0.812	0.732	0.659	0.594	0.535	0.482	0.435	0.392	0.353	0.318	0.287	0.258	0.233	0.210	0.189	0.170	0.153	0.138	0.125	0.112
800	0.790	0.701	0.623	0.554	0.492	0.437	0.389	0.345	0.307	0.273	0.242	0.215	0.191	0.170	0.151	0.134	0.119	0.106	0.094	0.084
900	0.768	0.673	0.590	0.517	0.453	0.397	0.348	0.305	0.267	0.234	0.205	0.180	0.158	0.138	0.121	0.106	0.093	0.081	0.074	0.063
1000	0.747	0.646	0.558	0.483	0.417	0.361	0.312	0.270	0.233	0.201	0.174	0.151	0.130	0.113	0.097	0.084	0.073	0.063	0.054	0.050
1500	0.655	0.531	0.430	0.348	0.282	0.228	0.185	0.149	0.121	0.098	0.079	0.064	0.052	0.042	0.034	0.028	0.022	0.018	0.015	0.012
2000	0.580	0.441	0.336	0.256	0.195	0.148	0.113	0.086	0.065	0.050	0.038	0.029	0.022	0.017	0.013	0.010	0.007	0.005	0.004	0.003

Table 4.1

The damping effect of elevation on harmonics of different degrees up to n=20. Values of $(a_e/r)^{n+1}$.

The *second* question of relevance is the applicability of spherical harmonic expansions at the surface of the earth. Laplace's equation must, for all practical purposes, be considered to be valid upto the surface of the earth. Consider figure 4.4. The position of a point P on the earth's surface can be referred to a spheroid of reference whose equatorial radius is a_e and flattening f, if its spheroidal elevation h_s were known. The relation between position with reference to the spheroid and the spherical system of co-ordinates (r,θ,λ) is obtained as follows. θ is related to the co-latitude θ_g of the spheroidal geocentric latitude ϕ_g by the relation

$$\theta = \theta_g - d\theta \qquad \ldots(4.43).$$

The longitude λ remains unaltered. The geocentric distance r is related to the geocentric radius r_g to the equivalent point P_s on the spheroid by the relation

$$r = r_g \cos d\theta + h_s \cos(\phi-\phi_g-d\theta) \qquad \ldots(4.44).$$

The geodetic latitude ϕ is related to ϕ_g by the well known relation(e.g., Bomford 1962,p.496)

$$\phi - \phi_g = [f + \tfrac{1}{2}f^2]\sin 2\phi + \tfrac{1}{2}f^2 \sin 4\phi + o\{f^3\} \ldots(4.45).$$

The use of the relation preceding equation 1.28, together with equations 1.9 to 1.11, expansion by the use of the binomial theorem and slight re-arrangement of terms gives

$$r_g = \frac{a_e(1-f)}{[1-(2f-f^2)\cos^2\phi_g]^{\frac{1}{2}}} = a_e\left(1 - (f+\tfrac{3}{2}f^2)\sin^2\phi_g + \tfrac{3}{2}f^2\sin^2\phi_g + o\{f^3\}\right) \ldots(4.46).$$

The application of sine formula to triangle OPP_s in figure 4.4 gives

$$\frac{h_s}{r_g} = \frac{\sin d\theta}{\sin(\phi-\phi_g-d\theta)} \doteq \frac{d\theta}{\sin(\phi-\phi_g) - d\theta\cos(\phi-\phi_g)}$$

or

$$d\theta[\, r_g + h_s\cos(\phi-\phi_g)\,] = h_s\sin(\phi-\phi_g).$$

If terms of order $(\phi-\phi_g)^3 \doteq f^3$ are disregarded,

$$d\theta = \frac{h_s(\phi-\phi_g)}{r_g + h_s[1 - \tfrac{1}{2}(\phi-\phi_g)^2]}$$

or

$$d\theta = \frac{h_s}{r_g} f \sin 2\phi + o\{f^3\} \qquad \ldots(4.47).$$

The combination of equations 4.43 and 4.47 gives

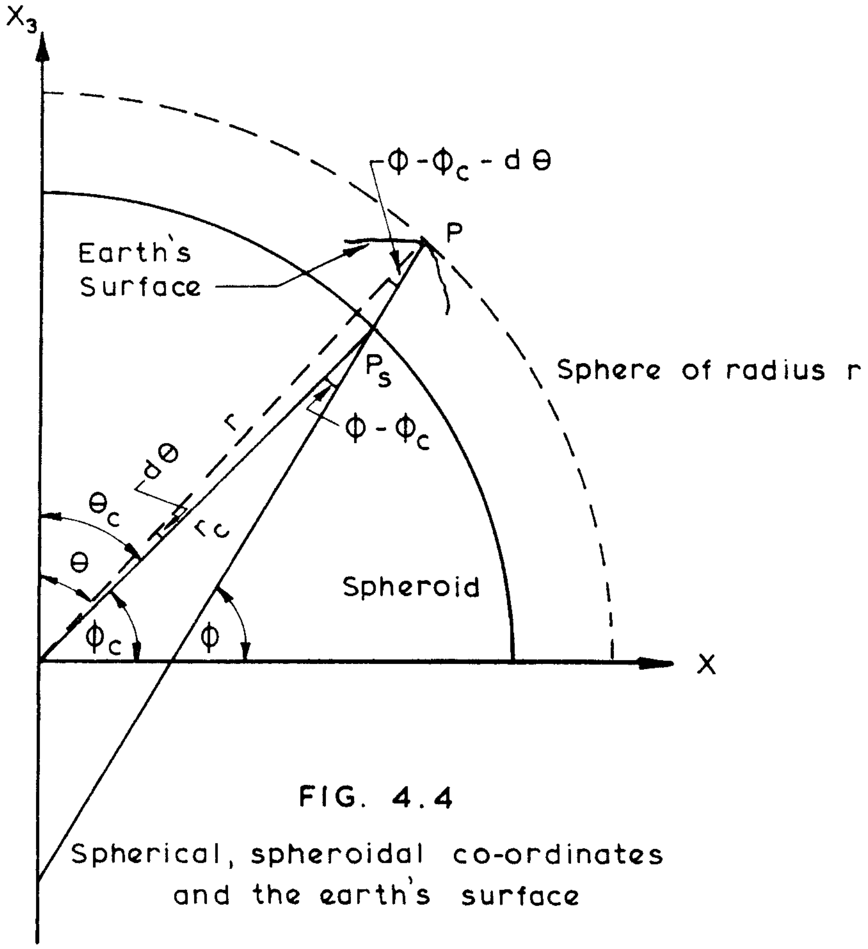


FIG. 4.4

Spherical, spheroidal co-ordinates and the earth's surface

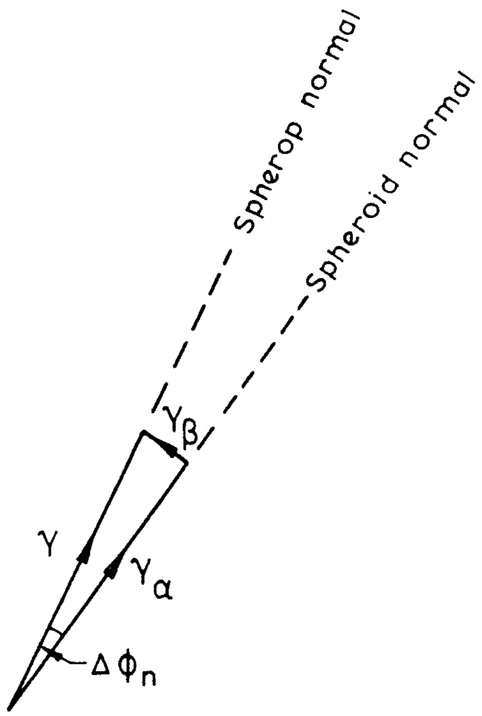


FIG. 6.1

The spheroid and spherop normals

$$\theta = \theta_g - \frac{h_s}{r_g} f \sin 2\phi = \theta_g + o\{f^2\} .$$

The use of equations 4.44 and 4.45 gives

$$r = r_g + h_s + o\{f^3\} = r_g \left(1 + \frac{h_s}{r_g} + o\{f^3\}\right) \quad \ldots(4.48).$$

On replacing r_g by a_e using equation 4.46 and θ_g by θ , equation 4.48 becomes

$$r = a_e\left(1 - (f + \tfrac{3}{2}f^2)\cos^2\theta + \tfrac{3}{2}f^2\cos^4\theta\right) 1 + \frac{h_s}{a_e}(1 + f\cos^2\theta)$$

$$= a_e\left(1 + \frac{h_s}{a_e} - \{f + \tfrac{3}{2}f^2 - \frac{h_s}{a_e}\}\cos^2\theta + \tfrac{3}{2}f^2\cos^4\theta + o\{f^3\}\right).$$

Therefore

$$\frac{a_e}{r} = 1 - \frac{h_s}{a_e} + \left(\frac{h_s}{a_e}\right)^2 + \{f + \tfrac{3}{2}f^2 - 3\frac{h_s}{a_e}f\}\cos^2\theta - \tfrac{1}{2}f^2\cos^4\theta + o\{f^3\}$$

$$= 1 + o\{f\} \quad \ldots(4.49).$$

Some important conclusions can be drawn from the above development.

(i) A spherical harmonic expansion for the potential is valid, for all practical purposes, down to the surface of the earth at all points in space exterior to the physical surface.

(ii) The ratio (a_e/r) can be taken as being unity at the surface of the earth to the order of the flattening.

(iii) The geopotential can be expressed to the order f by the series

$$U = \frac{kM}{a_e}\sum_{n=0}^{\infty}\left(1 - \frac{h_s}{a_e} + f\mu^2\right)^{n+1}\sum_{m=0}^{n} P_{nm}(\mu)[C_{nm}\cos m\lambda + S_{nm}\sin m\lambda] + o\{f^2U\} \quad \ldots(4.50)$$

at the surface of the earth,

where $\mu = \sin\phi_g$.

(iv) The order of magnitude of U is 6×10^6 kgal m, the major contribution coming from the term $n = 0$. The second largest term, on excluding those of degree one from the present discussion, is that obtained when $n = 2$ which is shown in section 6.2 to be the main ellipticity term of magnitude $f\,U \simeq 2\times10^4$ kgal m. All other terms can be seen from table 8.3 to be of magnitude f^2U .

It would therefore appear that the full expression for (a_e/r), given in equation 4.49) should be used in equation 4.50 to define an expression for U to order f^3U. Such a procedure is not necessary if the technique of linearisation is adopted, as developed in section 6.1.

No objection should therefore be raised to the representation of the geopotential in terms of spherical co-ordinates using equations 4.49 and 4.50 at all points in space exterior to the surface of the earth.

5. SPHEROIDAL HARMONICS

5.1 Laplace's equation in spheroidal co-ordinates

The spheroidal system of co-ordinates is described in section 1.3 *ii* , where any point in the space can be defined in terms of the three parameters

$$\xi_1 \overset{(1.16)}{=} u \; ; \qquad \xi_2 = \lambda \quad ; \qquad \xi_3 = \alpha \; .$$

The first parameter u(i.e., ξ_1) which defines the family of unparted hyperboloids of revolution can also be represented by the parameter β when

$$\cos u = \operatorname{sech} \beta \quad ; \quad \sin u = \tanh \beta \qquad \ldots (5.1),$$

resulting in the simplification of subsequent differentiations. These co-ordinates were shown to form a triply orthogonal system and will also satisfy Laplace's equation in general curvilinear co-ordinates which was derived in section 2.12 as

$$\sum_{i=1}^{3} \frac{d}{d\xi_i}\left(\frac{h_{i+1}h_{i+2}}{h_i} \frac{\partial U}{\partial \xi_i} \right) \overset{(2.34)}{=} 0 \quad , \text{ss.}^{(*)}$$

The linearisation parameters in this case are obtained by referring to the expressions for position on a three dimensional Cartesian reference frame which is given in section 1.3 *ii* by the equations

$$\begin{aligned} x_1 &\overset{(1.18)}{=} c \operatorname{cosec} \alpha \operatorname{sech} \beta \cos \lambda \\ x_2 &= c \operatorname{cosec} \alpha \operatorname{sech} \beta \sin \lambda \\ x_3 &= c \cot \alpha \tanh \beta \end{aligned}$$

the functions in u having been replaced by equivalent expressions in β by the use of equation 5.1. As both systems of co-ordinates are triply orthogonal, it can be seen that(e.g., Eisenhart 1960,p.447)

$$h_i^2 d\xi_i^2 = dx_i^2 \quad ^{(*)}$$

---

* See "Guide to notation" for explanation

or

$$h_j^2 = \sum_{i=1}^{3} \left(\frac{\partial x_i}{\partial \xi_j}\right)^2$$

Thus differentiation of equation 1.18 in its modified form gives

$$h_1^2 = h_\beta^2 = c^2 \left[\mathrm{cosec}^2\alpha \tanh^2\beta\ \mathrm{sech}^2\beta + \cot^2\alpha\,\mathrm{sech}^4\beta\right]$$

$$= c^2[(1 + \cot^2\alpha)\tanh^2\beta\ \mathrm{sech}^2\beta + \cot^2\alpha(1 - \tanh^2\beta)\mathrm{sech}^2\beta]$$

$$= c^2[\tanh^2\beta + \cot^2\alpha]\mathrm{sech}^2\beta \qquad \ldots(5.2a)$$

Similarly,

$$h_2^2 = h_\lambda^2 = c^2\mathrm{cosec}^2\alpha\ \mathrm{sech}^2\beta \qquad \ldots(5.2b)$$

and

$$h_3^2 = h_\alpha^2 = c^2[\mathrm{cosec}^2\alpha\ \cot^2\alpha\ \ \mathrm{sech}^2\beta + \mathrm{cosec}^4\alpha\ \ \tanh^2\beta]$$

$$= c^2[\mathrm{cosec}^2\alpha\ (1 - \tanh^2\beta)\cot^2\alpha + (1 + \cot^2\alpha)\tanh^2\beta]$$

$$= c^2\mathrm{cosec}^2\alpha[\cot^2\alpha + \tanh^2\beta] \qquad \ldots(5.3).$$

The application of these results in equation 2.34 gives

$$\frac{\partial}{\partial\beta}\left\{\frac{c\ \mathrm{cosec}\ \alpha\ \mathrm{sech}\ \beta(\cot^2\alpha + \tanh^2\beta)^{\frac{1}{2}}}{\mathrm{sech}\ \beta(\cot^2\alpha + \tanh^2\beta)^{\frac{1}{2}}}\frac{\partial U}{\partial\beta}\right\} \quad +$$

$$\frac{\partial}{\partial\lambda}\left\{\frac{c\ \mathrm{cosec}\ \alpha\ \mathrm{sech}\ \beta(\cot^2\alpha + \tanh^2\beta)}{\mathrm{cosec}\ \alpha\ \mathrm{sech}\ \beta}\ \frac{\partial U}{\partial\lambda}\right\} \quad +$$

$$\frac{\partial}{\partial\alpha}\left\{\frac{c\ \mathrm{cosec}\ \alpha\ \mathrm{sech}^2\beta(\cot^2\alpha + \tanh^2\beta)^{\frac{1}{2}}}{\mathrm{cosec}\ \alpha(\cot^2\alpha + \tanh^2\beta)^{\frac{1}{2}}}\frac{\partial U}{\partial\alpha}\right\} \quad = 0$$

which simplifies to

$$\mathrm{cosec}^2\alpha\ \frac{\partial^2 U}{\partial\beta^2} + (\cot^2\alpha + \tanh^2\beta)\frac{\partial^2 U}{\partial\lambda^2} + \mathrm{sech}^2\beta\ \frac{\partial^2 U}{\partial\alpha^2} = 0.$$

This can be written as

$$\cosh^2\beta\ \frac{\partial^2 U}{\partial\beta^2} + (\cosh^2\beta - \sin^2\alpha)\ \frac{\partial^2 U}{\partial\lambda^2} + \sin^2\alpha\ \frac{\partial^2 U}{\partial\alpha^2} = 0 \quad \ldots(5.4)$$

which expresses Laplace's equation in *oblate spheroidal co-ordinates*.

Notes

(i) The (β,λ,α) system of co-ordinates define the space as the intereection of

(a) an oblate spheroid with radii {c cosec α, c cot α} ;

(b) a confocal unparted hyperboloid of revolution with semi axes {c sech β ; c tanh β} and with a common centre;

and (c) the plane through the common axis of revolution and the point, defined by the angle λ it makes with a fixed plane.

(ii) The entire space is covered by the range of values

$$0 \leq \alpha \leq \tfrac{1}{2}\pi \quad ; \quad -\infty \leq \beta \leq \infty \quad ; \quad 0 \leq \lambda \leq 2\pi \qquad \ldots(5.5).$$

The representation is extremely unstable when $\alpha \to 0$ as will be seen in the next section.

5.2 The solution

The technique adopted for the solution of Laplace's equation in oblate spheroidal co-ordinates follows closely on that used in section 3.1. On adoption of a representation of the form

$$U = L(\alpha) \quad M(\beta) \quad N(\lambda) \qquad \ldots(5.6)$$

for U, appropriate substitution in equation 5.4 gives

$$\cosh^2\beta \; L(\alpha) \; N(\lambda)\frac{\partial^2 M}{\partial\beta^2} + (\cosh^2\beta - \sin^2\alpha)L(\alpha) \; M(\beta)\frac{\partial^2 N}{\partial\lambda^2} + \sin^2\alpha \; M(\beta) \; N(\lambda)\frac{\partial^2 L}{\partial\alpha^2} = 0.$$

The first stage of the separation of the variables gives

$$\frac{\cosh^2\beta}{\cosh^2\beta - \sin^2\alpha}\frac{1}{M(\beta)}\frac{\partial^2 M}{\partial\beta^2} + \frac{\sin^2\alpha}{\cosh^2\beta - \sin^2\alpha}\frac{1}{L(\alpha)}\frac{\partial^2 L}{\partial\alpha^2}$$

$$= -\frac{1}{N(\lambda)}\frac{\partial^2 N}{\partial\lambda^2} = m^2 \qquad \ldots(5.7),$$

the common constant being chosen for the same reasons given in the formation of equation 3.7. The differential equation governing the function $N(\lambda)$ is therefore

$$\frac{\partial^2 N}{\partial\lambda^2} + m^2 N(\lambda) = 0 \qquad \ldots(5.8).$$

The second stage gives

$$-\frac{1}{M(\beta)}\cosh^2\beta\frac{\partial^2 M}{\partial\beta^2} + m^2\cosh^2\beta = \sin^2\alpha\frac{1}{L(\alpha)}\frac{\partial^2 L}{\partial\alpha^2} + m^2\sin^2\alpha = n(n+1),$$

the relations being satisfied simultaneously if they are both equal to a common constant whose most convenient form is $n(n+1)$. The resulting differential equations are

$$\cosh^2\beta\frac{\partial^2 M}{\partial\beta^2} + M(\beta)\left[-m^2\cosh^2\beta + n(n+1)\right] = 0 \qquad \ldots(5.9)$$

and

$$\sin^2\alpha\frac{\partial^2 L}{\partial\alpha^2} + L(\alpha)\left[m^2\sin^2\alpha - n(n+1)\right] = 0 \qquad \ldots 5.10)$$

Both equations 5.9 and 5.10 can be transposed into the form

$$(1-p^2)\frac{d^2G}{dp^2} \quad - 2p\,\frac{dG}{dp} \quad + \left(n(n+1) \; - \; \frac{m^2}{1-p^2}\right) G(p) \; = \; 0$$

which is equation 3.11, whose solution is developed in sections 3.3 and 3.6. Equation 5.10 is transformed by adopting the change of variable to ω defined by the relation

$$\omega \; = i \; \cot \, \alpha$$

when

$$\frac{d}{d\alpha} \; = - \; i \operatorname{cosec}^2\alpha \, \frac{d}{d\omega} \; = - \, i(1 \; -\omega^2)\frac{d}{d\omega}$$

and

$$\frac{d^2}{d\alpha^2} = 2 \; i \operatorname{cosec}^2\alpha \; \cot \; \alpha \; \frac{d}{d\omega} \; - \operatorname{cosec}^4\alpha \; \frac{d^2}{d\omega^2} \quad .$$

As

$$\sin^2\alpha \; = \frac{1}{\operatorname{cosec}^2\alpha} \; = \frac{1}{1 \, + \, \cot^2\alpha} \; = \frac{1}{1 \, - \, \omega^2} \quad ,$$

equation 5.10 becomes

$$\frac{1}{1-\omega^2}\left(2\omega(1-\omega^2)\;\frac{dL}{d\omega} \; - \; (1-\omega^2)^2\;\frac{d^2L}{d\omega^2}\right) \; + \; L(\omega)\left(\frac{m}{1-\omega^2} \; - \; n(n+1)\right) \; = \; 0$$

as $i\cot\,\alpha \; = \omega,$ the equation reducing further to

$$(1-\omega^2)\frac{d^2L}{d\omega^2} \quad - \; 2\omega \; \frac{dL}{d\omega} \quad + \; L(\omega)\left(n(n+1) \; - \; \frac{m}{1-\omega^2}\right) \quad = \; 0$$

which is identical in form to equation 3.11. Hence its solution is given by equation 3.35. As $\cot\,\alpha > 1$, the required solution in this case is

$$L(\alpha) \; = \; B \quad (1+\cot^2\alpha)^{\frac{1}{2}m} \; \frac{d^m}{d(i\cot\,\alpha)^m} \; \{Q_n\,(i\cot\;\alpha)\} = B\,Q_{nm}(i\cot\;\alpha)$$

...(5.11).

The appropriate change of variable in the case of equation 5.9 is

$$v \; = \; \tanh \; \beta$$

when

$$\frac{d}{d\beta} \; = \; \operatorname{sech}^2\beta \; \frac{d}{dv}$$

and

$$\frac{d^2}{d\beta^2} \; = \; -2\operatorname{sech}^2\beta \; \tanh \; \beta \; \frac{d}{dv} \; + \; (\operatorname{sech}^2\beta)^2 \; \frac{d^2}{dv^2} \quad .$$

Equation 5.9 therefore becomes

$$\frac{1}{1-v^2}\left((1-v^2)^2 \; \frac{d^2M}{dv^2} \; - \; 2v(1-v^2) \; \frac{dM}{dv}\right) + \; M(v)\left(\; n(n+1) \; - \; \frac{m^2}{1-v^2}\right) \; = \; 0$$

or

$$(1-v^2) \; \frac{d^2M}{dv^2} \; - \; 2v \; \frac{dM}{dv} \; + \; M(v)\left(\; n(n+1) \; - \; \frac{m^2}{1-v^2}\right) \; = \; 0$$

whose solution from equation 3.35 is

$$M(\beta) = A\, P_{nm}(\tanh \beta) \qquad \ldots(5.12).$$

Thus

$$U_{nm} = P_{nm}(\tanh \beta)\, Q_{nm}(i\cot \alpha)\left(C'_{nm}\cos m\lambda + S'_{nm}\sin m\lambda\right) \quad ..(5.13)$$

is a solution of Laplace's equation in spheroidal co-ordinates. A *complete* solution is given by

$$U = \sum_{n=0}^{\infty} \sum_{m=0}^{n} P_{nm}(\tanh \beta)\, Q_{nm}(i\cot \alpha)\left(C'_{nm}\cos m\lambda + S'_{nm}\sin m\lambda\right)$$

$$\ldots(5.14).$$

At this stage, it is in order to revert to the variable $\xi_1 = u$ when the use of equation 5.1 gives

$$U = \sum_{n=0}^{\infty} \sum_{m=0}^{n} P_{nm}(\sin u)\, Q_{nm}(i\cot \alpha)\left\{C'_{nm}\cos m\lambda + S'_{nm}\sin m\lambda\right\}.(5.15)$$

which is a complete solution of Laplace's equation in oblate spheroidal co-ordinates.

Notes

(i) The incorporation of the term i^{n+m+1} in the coefficients according to

$$C'_{nm} = i^{n+m+1}\, C_{nm} \qquad \ldots(5.16)$$

with a similar expression for S'_{nm} will result in C_{nm}, S_{nm} being real numbers for real values of the geopotential U.

5.3 The evaluation of $Q_{no}(i \cot \alpha)$

The evaluation of $Q_{no}(i\cot \alpha)$ is of importance in section 5.4. The general function $Q_{no}(\mu)$ is most simply evaluated from $P_{no}(\mu)$. This is done by considering Legendre's equation given in equation 3.14 as

$$(1-\mu^2)\,\frac{d^2G_n}{d\mu^2} - 2\mu\,\frac{dG_n}{d\mu} + n(n+1)\,G_n(\mu) = 0$$

and whose general solution is

$$G_n(\mu) \overset{(3.21)}{=} A\, P_{no}(\mu) + B\, Q_{no}(\mu).$$

$P_{no}(\mu)$ is a solution of equation 3.21 when $B = 0$ and $A = 1$. Also, if $|\mu| \neq 1$, consider the following solution of Legendre's equation.

$$G_n(\mu) = R\ P_{no}(\mu) \qquad \ldots(5.17)$$

where R is a function of μ. Differentiation of equation 5.17 with respect to μ gives

$$\frac{dG_n}{d\mu} = \frac{dR}{d\mu} P_{no}(\mu) + R \frac{dP_{no}(\mu)}{d\mu} .$$

A further differentiation gives

$$\frac{d^2G_n}{d\mu^2} = \frac{d^2R}{d\mu^2} P_{no}(\mu) + 2 \frac{dR}{d\mu} \frac{dP_{no}(\mu)}{d\mu} + R \frac{d^2P_{no}(\mu)}{d\mu^2}$$

Substitution in equation 3.14 after division by $(1-\mu^2)$ gives

$$R \left\{ \frac{d^2P_{no}(\mu)}{d\mu^2} - \frac{2}{1-\mu^2} \frac{dP_{no}(\mu)}{d\mu} + \frac{n(n+1)}{1-\mu^2} P_{no}(\mu) \right\} +$$

$$\frac{d^2R}{d\mu^2} P_{no}(\mu) + 2 \frac{dR}{d\mu} \frac{dP_{no}(\mu)}{d\mu} - \frac{2}{1-\mu^2} \frac{dR}{d\mu} P_{no}(\mu) = 0 \quad ..(5.18).$$

The term within the bracket is zero as $P_{no}(\mu)$ is a solution of Legendre's equation. Equation 5.18 is simplified further by making the substitution

$$\frac{dR}{d\mu} = R'$$

when this equation becomes

$$\frac{dR'}{d\mu} P_{no}(\mu) + 2R' \left\{ \frac{dP_{no}(\mu)}{d\mu} - \frac{\mu}{1-\mu^2} P_{no}(\mu) \right\} = 0.$$

The separation of variables enables the above equation to be written as

$$\frac{1}{R'} \frac{dR'}{d\mu} + 2 \left\{ \frac{1}{P_{no}(\mu)} \frac{dP_{no}(\mu)}{d\mu} - \frac{\mu}{1-\mu^2} \right\} = 0$$

which, on integration with respect to μ gives

$$\log R' + \log[P_{no}(\mu)]^2 + \log(1-\mu^2) = C$$

where C is the constant of integration. Alternately,

$$R' = \frac{B}{[P_{no}(\mu)]^2 (1-\mu^2)} .$$

Further integration with respect to μ gives

$$R = A + B \int \frac{d\mu}{[P_{no}(\mu)]^2 (1-\mu^2)} ,$$

A and B being constants of integration. Hence the solution of Legendre's equation given by equation 5.17 takes the form

$$G_n(\mu) = A\ P_{no}(\mu) + B\ P_{no}(\mu)\int \frac{d\mu}{[P_{no}(\mu)]^2(1-\mu^2)} \qquad \ldots(5.19).$$

The comparison of equation 5.19 with equation 3.21 gives

$$Q_{no}(\mu) = C\ P_{no}(\mu) \int \frac{d\mu}{[P_{no}(\mu)]^2(1-\mu^2)} \qquad \ldots(5.20)$$

where C is a constant. The limits of integration and the range of μ are obtained by studying

$$F = \frac{1}{[P_{no}(\mu)]^2(1-\mu^2)}$$

This expression is a rational fraction which becomes indeterminate at

(a) ± 1;

and (b) all the roots of $P_{no}(\mu)$ which are real and lie in the range $-1 < \mu < 1$.

Hence if $\mu^2 > 1$,

$$\int_\mu^\infty F\,d\mu$$

is finite and determinate and contains no constant term. It is conventional to adopt the value -1 for C when equation 5.20 as

$$Q_{no}(\mu) = P_{no}(\mu)\int_\mu^\infty \frac{d\mu}{[P_{no}(\mu)]^2(\mu^2-1)} \qquad \ldots(5.21),$$

where $\mu > 1$.

When $n = 0$, $P_{oo}(\mu) = 1$

and

$$Q_{oo}(\mu) = \int_\mu^\infty \frac{d\mu}{\mu^2-1} = \tfrac{1}{2}\log\left(\frac{\mu+1}{\mu-1}\right) \qquad \ldots(5.22).$$

When $n = 1$, $P_{10}(\mu) = \mu$

and

$$Q_{10}(\mu) = \mu\int_\mu^\infty \frac{d\mu}{(\mu^2-1)\mu^2} = -\mu\int_\mu^\infty \frac{d\mu}{\mu^2} + \mu\int_\mu^\infty \frac{d\mu}{\mu^2-1}$$

$$= \mu\left|\frac{1}{\mu}\right|_\mu^\infty + \tfrac{1}{2}\mu\log\left(\frac{\mu+1}{\mu-1}\right)$$

$$= \tfrac{1}{2}\mu\log\left(\frac{\mu+1}{\mu-1}\right) - 1 \qquad \ldots(5.23).$$

The direct evaluation of terms of higher degree by the use of equation 5.21 is comparatively difficult. These are more simply derived by the use of the recurrence relation for Legendre functions of the second kind given by (Hobson 1965, p.67)

$$n\,Q_{no}(\mu) - (2n-1)\,\mu\,Q_{n-1\,o}(\mu) + (n-1)\,Q_{n-2\,o}(\mu) = 0.$$

Thus

$$2Q_{20}(\mu) = 3\mu\left(\tfrac{1}{2}\mu\,\log\left(\frac{\mu+1}{\mu-1}\right) - 1\right) - \tfrac{1}{2}\,\log\left(\frac{\mu+1}{\mu-1}\right)$$

or

$$Q_{20}(\mu) = -\frac{3}{2}\,\mu + \frac{1}{4}(3\mu^2-1)\log\left(\frac{\mu+1}{\mu-1}\right) \qquad \ldots(5.24).$$

If $q_n(\alpha)$ is defined by the equation

$$q_n(\alpha) = i^{n+1}\,Q_{no}(i\cot\alpha),$$

the expressions for $q_n(\alpha)$ when $n \le 2$ are given by the following series of equations.

$$q_o(\alpha) = Q_{oo}(i\cot\alpha) = \tfrac{1}{2}i\log\left(\frac{i\cot\alpha+1}{i\cot\alpha-1}\right) = \tfrac{1}{2}i\,\log\left(\frac{\cos\alpha - i\sin\alpha}{\cos\alpha + i\sin\alpha}\right)$$

$$= \tfrac{1}{2}\,i\log\left(\frac{e^{-i\alpha}}{e^{i\alpha}}\right) = \alpha \qquad \ldots.(5.25)$$

The use of the same principles on equation 5.23 gives

$$q_1(\alpha) = i^2\,Q_{10}(i\cot\alpha) = 1 - \tfrac{1}{2}i\cot\alpha\,(-2i\alpha)$$

$$= 1 - \alpha\cot\alpha \qquad \ldots(5.26).$$

Similarly, the use of equation 5.24 gives

$$q_2(\alpha) = i^3\,Q_{20}(i\cot\alpha) = -i\left[-\frac{3}{2}\,i\cot\alpha + \tfrac{1}{4}(-3\cot^2\alpha - 1) - 2i\alpha\right]$$

$$= -\frac{3}{2}\cot\alpha + \tfrac{1}{2}\alpha(1+3\cot^2\alpha) = \tfrac{1}{2}\left[\alpha(3\cot^2\alpha+1) - 3\cot\alpha\right]$$

$$\ldots(5.27).$$

An alternate expression of interest is obtained by the use of the change of variable

$$x = \tan\alpha$$

when

$$\alpha = \tan^{-1}x = \sum_{s=1}^{\infty}(-1)^{s+1}\frac{x^{2s-1}}{2s-1}\ .$$

Thus

$$q_2(\alpha) = \frac{1}{2x^2}\left(\sum_{s=1}^{\infty}(-1)^{s+1}\frac{x^{2s-1}}{2\ -1}(3+x^2) - 3x\right)$$

$$= \frac{1}{2x^2}\left(\left(x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \frac{x^{11}}{11} + \ldots\right)(3+x^2) - 3x\right)$$

$$= \frac{1}{2x^2}\left(x^3(1-1) + x^5\left(\frac{3}{5} - \frac{1}{3}\right) + x^7\left(-\frac{3}{7} + \frac{1}{5}\right) + x^9\left(\frac{1}{3} - \frac{1}{7}\right) + x^{11}\left(-\frac{3}{11} + \frac{1}{9}\right)\right)$$

$$= \tfrac{1}{2}\left(\frac{4}{3.5}x^3 - \frac{8}{5.7}x^5 + \frac{12}{7.9}x^7 - \frac{16}{9.11}x^9 + \ldots\right).$$

Thus

$$q_2(\alpha) = 2\left(\frac{1}{3.5}\tan^3\alpha - \frac{2}{5.7}\tan^5\alpha + \frac{3}{7.9}\tan^7\alpha - \frac{4}{9.11}\tan^9\alpha\right) \quad ..(5.28).$$

Note

As

$$\sin\alpha \overset{(1.13)}{=} e$$

and $e \approx 8 \times 10^{-2}$ for a spheroid with the same dimensions as the earth, the term in $\tan^9\alpha \approx 10^{-10}$ and can be ignored.

6. THE REFERENCE SYSTEM IN GRAVIMETRIC GEODESY

6.1 Introduction

The exterior gravitational field of the earth can be expressed in terms of either a spherical harmonic series of the type at equation 4.40 or a spheroidal harmonic series of the type at equation 5.15, as both are solutions of Laplace's equation. The major contribution to U is from the term obtained when $n = 0$ as explained in section 4.9, being of the order of magnitude 6×10^6 kgal m(kilogal metres = $cm^2 sec^{-2} \times 10^{-5}$). The second largest contribution is due to the departure of the earth from a sphere, i.e., the ellipticity term whose principal contribution is through the harmonic ($n=2, m=0$), being of magnitude 2×10^4 kgal m. The contribution of all the other terms combined is unlikely to exceed 100 - 200 kgal m. Thus a considerable increase in accuracy can be achieved without a related increase in the complexity of formulae by merely removing the effect of the two terms of significant magnitude.

This is done by introducing the concept of a reference ellipsoid (or spheroid of reference) and superimposing its gravitational field on the space. The potential due to this reference system (spheropotential) U can be compared with the geopotential W at every point in the space. The introduction of such a concept with the adoption of well chosen parameters for the definition of the gravitational field of the reference ellipsoid results in the quantity V_d, given by the equation

$$V_d = W - U \qquad \ldots(6.1)$$

having an order of magnitude less than 10^2 kgal m, as the effect of the terms ($n=0$) and ($n=2, m=0$) described above would have been removed. The quantity V_d is known as the *disturbing potential* (also potential disturbance; potential anomaly).

The primary function of the reference system is the *linearisation* of the magnitude of quantities on the physical system. In addition, the former can be used as a reference frame to define the geometry of the space. The gravitational field of a reference ellipsoid an be represented at all points in the space unoccupied by matter by the use of either the appropriate spherical harmonic or spheroidal harmonic series. There is little to choose between the two representations which are dealt with in section 6.2, though the spheroidal harmonic series lends itself with comparative elegance to representations of ellipsoids of revolution by closed expressions. This is offset by the fact that the representation of the geopotential by a series of spheroidal harmonics is cumbersome and subject to much the same limitations as a spherical harmonic series as the surface of the earth deviates significantly from the surface of an ellipsoid of revolution.

6.2 The gravitational potential due to an ellipsoid of revolution

An ellipsoid of revolution with a distribution of mass which is symmetrical about both the rotation axis and the equator produces a gravitational field with similar characteristics of symmetry. Any harmonic expressions for the potential of such a system should therefore satisfy the following conditions.

(i) All longitude dependent terms should make zero contribution to such a series, i.e., $m = 0$;

and (ii) The remaining zonal terms should have zero coefficients for odd degree terms which make asymmetric contributions about the equator.

Thus the possible harmonic series for the representation of the gravitational potential of an ellipsoid of revolution are

(a)

$$U_g = \sum_{n=0}^{\infty} i^{2n+1} \; Q_{2n\,o}(i \cot \alpha) \; P_{2n\,o}(\sin u) a_{2n\,o} \quad \ldots(6.2)$$

for a spheroidal harmonic series and

(b)

$$U_g = \sum_{n=0}^{\infty} \frac{1}{r^{2n+1}} P_{2n\,o}(\sin \phi_g) \; C_{2n\,o} \quad \ldots(6.3)$$

in the case of a spherical harmonic series.

i. The spheroidal harmonic series

Let the reference ellipsoid rotate with the same angular velocity ω as the earth giving rise to a rotational potential U_r defined by equations 2.27 and 1.18 as

$$U_r = \tfrac{1}{2} c^2 \operatorname{cosec}^2\alpha \cos^2 u \ \omega^2 \qquad \ldots(6.4).$$

As

$$\tfrac{1}{2} \cos^2 u = \tfrac{1}{2}(1 - \sin^2 u),$$

$P_{20}(\sin u)$ is given from section 3.8 as

$$P_{20}(\sin u) = \frac{3}{2}(\sin^2 u - 1) \qquad \ldots(6.5).$$

Thus

$$U_r = \frac{1}{3} c^2 \operatorname{cosec}^2\alpha[1 - P_{20}(\sin u)]\omega^2 \qquad \ldots(6.6).$$

In addition, the spheroid approximates to order f^2 to the geoid which is the geop $W = W_o$ corresponding to Mean Sea Level. It is therefore convenient to make it an equipotential surface $U = U_o$. In addition, the potential should tend to zero at very great distances from the earth, approaching this value through expressions of the form

$$U = \frac{kM}{r}$$

for large values of r which are equivalent to small values of α. It can therefore be seen from equations 1.13 and 1.16 that the range of values for α from the surface of the ellipsoid to ∞ is

$$4^\circ 42' > \alpha > 0.$$

Thus the total gravitational potential U is given by

$$U = U_g + U_r$$

$$\overset{(6.2)(6.6)}{=} \sum_{n=0}^{\infty} q_{2n}(\alpha) P_{2n\,o}(\sin u) a_{2no} + \frac{1}{3} c^2 \operatorname{cosec}^2\alpha[1-P_{20}(\sin u)]\omega^2,$$

where $q_{2n}(\alpha)$ is given by equations 5.25 to 5.28 for $2n \leq 2$. If this expression is to represent an equipotential surface when $\alpha = \alpha_o$, where α_o is the parameter defining the reference ellipsoid whose potential at the surface is U_o, then U_o should be independent of terms in u. On putting $\mu = \sin u$, the above equation becomes

$$U_o = q_o(\alpha) P_{00}(\mu) a_{oo} + q_2(\alpha) P_{20}(\mu) a_{20} + q_4(\alpha) P_{40}(\mu) a_{40} + \ldots +$$

$$\frac{1}{3} c^2 \operatorname{cosec}^2\alpha_o [1-P_{20}(\mu)]\omega^2 \quad \ldots(6.8).$$

It therefore follows that

$$a_{40} = a_{60} = a_{80} = \ldots\ldots = 0$$

and the coefficient of $P_{20}(\mu)$ satisfies

$$q_2(\alpha_0)\, a_{20} - \tfrac{1}{3}c^2 \operatorname{cosec}^2\alpha_0\, \omega^2 = 0$$

or

$$a_{20} = \frac{1}{3}\,\frac{c^2 \operatorname{cosec}^2\alpha_0\, \omega^2}{q_2(\alpha_0)}$$

As

$$c \operatorname{cosec} \alpha_0 \overset{(1.16)}{=} a\ ,$$

the potential of the reference ellipsoid is therefore given by

$$U_o = a_{00}\,\alpha + \tfrac{1}{3}a\,\omega^2 \qquad \ldots(6.9).$$

The potential U at any other point (u,λ,α) exterior to the ellipsoid of reference and which rotates with it is given by

$$U = a_{00}\,\alpha + \frac{1}{3}\,\frac{c^2 \operatorname{cosec}^2\alpha_0\,\omega^2}{q_2(\alpha_0)}\, q_2(\alpha) P_{20}(\sin u) + \tfrac{1}{2}c^2 \operatorname{cosec}^2\alpha\, \cos^2 u\, \omega^2$$

$$\ldots(6.10).$$

a_{00} is evaluated by the consideration of the gravitational potential as expressed by equation 6.9 when $r \to \infty$ in the space which does not rotate with the ellipsoid. As

$$r \overset{(1.18)}{=} \left(\sum_{i=1}^{3} x_i^2\right)^{\frac{1}{2}} = c\,[\operatorname{cosec}^2\alpha \cos^2 u + \cot^2\alpha \sin^2 u]^{\frac{1}{2}}.$$

As $r \to \infty$, $\cot\alpha$, $\operatorname{cosec}\alpha \to 0$ and

$$\cot^2\alpha = \frac{1}{\alpha^2} - \frac{2}{3} + o\{\alpha^2\}; \quad \operatorname{cosec}^2\alpha = \frac{1}{\alpha^2} + \frac{1}{3} + o\{\alpha^2\} .$$

Therefore

$$r = c\,[\frac{1}{\alpha^2} - \frac{2}{3}\sin^2 u + \frac{1}{3}\cos^2 u + o\{\alpha^2\}]^{\frac{1}{2}} \to \frac{c}{\alpha} \quad \text{as } \alpha \to 0$$

$$\ldots(6.11).$$

Also $q_2(\alpha)$ can be seen from equation 5.28 to tend to zero as $\alpha \to 0$. Hence

$$U \to \frac{kM}{r} = \frac{kM}{c} \qquad \text{as } r \to \infty ,$$

where M is the total mass of the reference system and k the gravitational constant. Hence

$$a_{00} = \frac{kM}{c} .$$

The final expression for U_o in terms of a and the eccentricity e of the reference ellipsoid is obtained by the use of equations 1.12 and 1.13 as

$$U = \frac{kM}{a}\left(\frac{\sin^{-1}e}{e}\right) + \frac{1}{3}\frac{a^2\ \omega^2}{q_2(\alpha_o)}\, q_2(\alpha)\ P_{20}(\sin u) +$$

$$\tfrac{1}{2}\ c^2 \operatorname{cosec}^2\alpha\ \cos^2 u \qquad \ldots(6.12)$$

while that for U_o is

$$U_o = \frac{kM}{a}\left(\frac{\sin^{-1}e}{e}\right) + \frac{1}{3}\, a^2\ \omega^2 \qquad \ldots(6.13)$$

Notes

(i) The potential U at the general point (u,λ,α) is given by equation 6.11.

(ii) It can be seen that the different confocal ellipsoids are *not* equipotential surfaces of the reference system (spherops) as the surface

$$U = \text{constant}$$

defined by equation 6.11 is not independent of u. Thus surfaces along which no change of U occur with change of u are not surfaces of

$$\alpha = \text{constant}.$$

In fact, the deviation of equipotential surfaces from the family of confocal ellipsoids increases with the difference of α from α_o.

(iii) The surface of theearth does not correspond with either a spherop or an ellipsoid. Consequently the use of spheroidal harmonics does not, in any way, facilitate the analysis of data distributed on the surface of the earth.

ii. The spherical harmonic series

The potential U_o on the reference ellipsoid is obtained from equation 6.3 in this case on using equation 6.7, as

$$U_o = U + U_r$$

$$= \frac{C'_{00}}{r} + \frac{1}{r^3} P_{20}(\mu) C'_{20} + \frac{1}{r^5} P_{40}(\mu)\ C'_{40} + \ldots + U_r \qquad \ldots(6.14)$$

where the rotational potential U_r is given by

$$U_r = \tfrac{1}{2}\ r^2\ \omega^2 \sin^2\theta_g = \tfrac{1}{2}\ a^2\ \omega^2 \left(\frac{r}{a}\right)^2 \sin^2\theta_g\ ,$$

a being the equatorial radius of the reference ellipsoid, r and θ_g defining geocentric distance and latitude respectively as in figure 6.1 for a point on the surface. The use of equation 4.49 when $h_s = 0$ together with the relation

$$m' = \frac{a^3\,\omega^2}{kM} \doteqdot \frac{\text{centrifugal acceleration at equator}}{\text{equatorial gravity}} \qquad \ldots(6.15)$$

gives

$$U_r = \tfrac{1}{2}\,\frac{kM}{a}\,m'\left(1-(f+\tfrac{3}{2}f^2)\cos\theta_g + \tfrac{3}{2}f^2\cos^2\theta_g\right)^2[1-\cos^2\theta_g]\ .$$

As $\omega \doteqdot 0.729\times 10^{-4}\ \sec^{-1}$, simple verification of orders of magnitude for a reference ellipsoid which fits the geoid (i.e., $a \simeq 6.37\times 10^{6}$ m ; $f^{-1} \simeq 3\times 10^{2}$) shows that

$$\ldots(6.16).$$

Hence

$$U_r = \tfrac{1}{2}\,\frac{kM}{a}\,m'\left(1-(1+2f)\cos^2\theta_g + 2f\cos^4\theta_g + o\{f^3\}\right) \quad (6.17)$$

Further, a study of table 8.3 shows that

$$C_{00} \simeq 1\ ; \qquad C_{20} \simeq 10^{-3}\ ; \qquad C_{40},\ C_{60},\ \text{etc} \simeq 10^{-6}.$$

On following the lines adopted in the formulation of equation 4.40, equation 6.14 becomes

$$U_o = \frac{kM}{a}\left(\left(\frac{a}{r}\right)\left(1+\left(\frac{a}{r}\right)^2 C_{20}\,P_{20}(\mu) + \left(\frac{a}{r}\right)^4 C_{40}P_{40}(\mu) + \ldots\right) + \tfrac{1}{2}m'[1-(1+2f)\cos^2\theta_g + 2f\cos^4\theta_g]\right)$$

$$\overset{(4.49)\,(\text{table } 3.1)}{=} \frac{kM}{a}\Big([1+(f+\tfrac{3}{2}f^2)\cos^2\theta_g - \tfrac{1}{2}f^2\cos^4\theta_g]\big(1 + 2f\cos^2\theta_g\big)C_{20}[\tfrac{3}{2}\cos^2\theta_g-\tfrac{1}{2}] + \tfrac{1}{8}[35\cos^4\theta_g - 30\cos^2\theta_g+3]C_{40} + \ldots\Big) +$$

$$\tfrac{1}{2}m' - \tfrac{1}{2}m'(1+2f)\cos^2\theta_g + m'f\cos^4\theta_g + o\{f^3\}\Big)$$

$$= \frac{kM}{a}\Big([1-\tfrac{1}{2}C_{20}+\tfrac{1}{2}m'+\tfrac{3}{8}C_{40}] + [f+\tfrac{3}{2}f^2+\tfrac{1}{2}C_{20}(3-2f-f) - \tfrac{15}{4}C_{40} - \tfrac{1}{2}m'(1+2f)]\times$$

$$\cos^2\theta_g + [-\tfrac{1}{2}f + C_{20}(3f+\tfrac{3}{2}f) + \tfrac{35}{8}C_{40} + m'f]\cos^4\theta_g + o\{f^3\}\Big)$$

Thus

$$U_o = \frac{kM}{a}\Big([1-\tfrac{1}{2}C_{20}+\tfrac{3}{8}C_{40}+\tfrac{1}{2}m'] + [f-\tfrac{1}{2}m'+\tfrac{3}{2}f^2-m'f+\tfrac{3}{2}C_{20}(1-f)-\tfrac{15}{4}C_{40}]\cos^2\theta_g +$$

$$[m'f - \tfrac{1}{2}f^2 + \tfrac{9}{2}C_{20}f + \tfrac{35}{8}C_{40}]\cos^4\theta_g + o\{f^3\}\Big) \qquad \ldots(6.18).$$

As U_o, being the value of the potential of the reference system on the surface of the ellipsoid, is constant on the ellipsoid, it must be independent of θ_g in equation 6.18. Thus

$$U_o = \frac{kM}{a}[1-\tfrac{1}{2}C_{20}+\tfrac{3}{8}C_{40}+\tfrac{1}{2}m' + o\{f^3\}].$$

In addition, the coefficients of $\cos^2\theta_g$ and $\cos^4\theta_g$ should also be equal to zero. Thus

$$f - \tfrac{1}{2}m' + \frac{3}{2}f^2 - m'f + \frac{3}{2}C_{20}(1-f) - \frac{15}{4}C_{40} = 0 \quad \ldots(6.18)$$

and

$$m'f + \frac{9}{2}C_{20}f - \tfrac{1}{2}f^2 + \frac{35}{8}C_{40} = 0 \quad \ldots(6.19).$$

The elimination of C_{40} between the two equations gives

$$\frac{7}{2}f - \frac{7}{4}m' + f^2[\frac{21}{4} - \frac{3}{2}] + m'f[3 - \frac{7}{2}] + C_{20}[\frac{21}{4}(1-f) + \frac{27}{2}f] = 0$$

or

$$\tfrac{1}{4}C_{20}(21 + 33f) = -[\frac{7}{2}f - \frac{7}{4}m' + \frac{15}{4}f^2 - \tfrac{1}{2}m'f].$$

Thus

$$C_{20} = \frac{4}{3\times7}[1+\frac{11}{7}f]^{-1}(\frac{7}{2}[\tfrac{1}{2}m' - f - \frac{15}{14}f^2 + \frac{1}{7}m'f])$$

$$= \frac{2}{3}[\tfrac{1}{2}m - f - \frac{9}{14}m'f + \tfrac{1}{2}f^2]$$

$$= \frac{1}{3}m' - \frac{2}{3}f - \frac{3}{7}m'f + \frac{1}{3}f^2 + o\{f^3\} \quad \ldots(6.20).$$

The use of equation 6.20 in 6.19 gives

$$C_{40} = \frac{8}{35}[\tfrac{1}{2}f^2 - m'f - \frac{9}{2}f(\frac{1}{3}m' - \frac{2}{3}f)]$$

$$= \frac{4}{5}f^2 - \frac{4}{7}m'f + o\{f^3\} \quad \ldots(6.21).$$

The terms of order f^3 in equation 6.20 are dependent on the assumption that terms of degree 6 do not make a significant contribution to the result to required order of accuracy. For a result correct to order e^3, see (Cook 1959).

The spheropotential U_p at any point P exterior to the spherop $U = U_o$ is given by

$$U_p = \frac{kM}{r}\left\{1 + \left(\frac{a}{r}\right)^2 P_{20}(\mu)C_{20} + \left(\frac{a}{r}\right)^4 P_{40}(\mu)C_{40} + \ldots\right\} \quad \ldots(6.22).$$

6.3 The formula for normal gravity on an ellipsoid of reference

The value of gravity on the surface of an ellipsoid of reference whose bounding surface is the equipotential $U = U_o$ is called *normal gravity* (γ). γ varies in magnitude with position on the surface of the ellipsoid. The formula for normal gravity can be obtained from either spherical or spheroidal harmonics as in the case of spheropotential. The latter are preferred for reasons given at the end of section 6.2.

If the reference ellipsoid is defined by $\alpha = \alpha_o$, the value of normal gravity is defined on the basis of the development in section 2.3 as the quantity

$$\gamma = -\nabla.N\ U$$

where N is the unit vector along the outward normal to the ellipsoid which is an equipotential surface. It can also be seen from section 1.3 *ii* that the normal derivative is a function of the variable α alone *in the immediate vicinity of the reference ellipsoid.* Consider two ellipsoids E_1 and E_2, corresponding to the parameters α and $\alpha+\Delta\alpha$ respectively where $\Delta\alpha \to 0$. If $P\{(u,\lambda,\alpha) \equiv (x_1,x_2,x_3)\}$ is a point on E_1 while $Q\{(u,\lambda,\alpha+\Delta\alpha) \equiv (x_i+\Delta x_i)\}$ is on E_2, the elemental length PQ can be assumed to be normal to both ellipsoids under limiting conditions. Thus

$$\Delta x_i = \frac{\partial x_i}{\partial \alpha}\Delta\alpha \quad + o\{(\Delta\alpha)^2\}.$$

The changes Δx_i are related to $\Delta\alpha$ through relations which are obtained by the differentiation of equation 1.18 when

$$\begin{aligned}\Delta x_1 &= -c\operatorname{cosec}\alpha\cot\alpha\cos u\cos\lambda\ \Delta\alpha\\ \Delta x_2 &= -c\operatorname{cosec}\alpha\cot\alpha\cos u\sin\lambda\ \Delta\alpha \qquad \ldots(6.23).\\ \Delta x_3 &= -c\operatorname{cosec}^2\alpha\sin u\ \ \Delta\alpha\end{aligned}$$

The element of length Δn along the normal is given by

$$\Delta n = \lim_{\Delta\alpha\to 0} PQ = \Big(\sum_{i=1}^{3}(\Delta x_i)^2\Big)^{\frac{1}{2}}$$

$$\begin{aligned}&= -c\operatorname{cosec}\alpha\ \Delta\alpha[\cot^2\alpha\cos^2 u + \operatorname{cosec}^2\alpha\sin^2 u]^{\frac{1}{2}}\\ &= -c\operatorname{cosec}\alpha\ \Delta\alpha[\cot^2\alpha + \sin^2 u]^{\frac{1}{2}}\\ &= -c\operatorname{cosec}^2\alpha\ \Delta\alpha[1 - \sin^2\alpha\cos^2 u]^{\frac{1}{2}} \qquad \ldots(6.24),\end{aligned}$$

the negative sign allowing for the decrease in α with an increment along the outward normal. The value of normal gravity γ_o on the reference ellipsoid is thus given by

$$\begin{aligned}\gamma_o &= -\left(\nabla.N\ U\right)_{U=U_o} = -\left(\frac{\partial U}{\partial n}\right)_{U=U_o}\\ &= -\left(\frac{\partial U}{\partial \alpha}\right)_{U=U_o}\left(\frac{\Delta\alpha}{\Delta n}\right)_{U=U_o} = \frac{\sin\alpha_o}{c[\cot^2\alpha_o + \sin^2 u]^{\frac{1}{2}}}\left(\frac{\partial U}{\partial \alpha}\right)_{\alpha=\alpha_o} \quad ..(6.25).\end{aligned}$$

The differentiation of equation 6.12 gives

$$\frac{\partial U}{\partial \alpha} = \frac{kM}{c} + \frac{1}{3}\frac{a^2\,\omega^2}{q_2(\alpha_o)}q_2'(\alpha)P_{20}(\sin u) - c^2\operatorname{cosec}^2\alpha\cot\alpha\cos^2 u\ \omega^2 \quad ..(6.26).$$

Evaluation at $\alpha = \alpha_o$ and the definition of F by the relation

$$F = \frac{q_2'(\alpha_o) \sin \alpha_o}{3q_2(\alpha_o)} \qquad \ldots(6.27)$$

gives

$$\gamma_o = \left(\frac{kM}{c} - a^2\omega^2 \cot \alpha \cos^2 u + a^2\omega^2 F \operatorname{cosec} \alpha_o P_{20}(\sin u) \right) \times$$

$$\left(\frac{\sin \alpha_o}{c \operatorname{cosec} \alpha_o (1-\sin^2\alpha_o \cos u)^{\frac{1}{2}}} \right)$$

as $\quad c \overset{(1.16)}{=} a \sin \alpha_o .$

Let m be given by the relation

$$m = \frac{a\,\omega^2}{\gamma_e} \qquad \ldots(6.28),$$

where γ_e is the value of normal gravity at the equator. It should be noted that m differs from m' defined in equation 6.15 by a quantity of order f^2. The use of this relation in the expression for γ_o gives

$$\gamma_o = \left(\frac{kM}{a^2} - m\,\gamma_e \cos \alpha_o \cos^2 u + m\,\gamma_e F\, P_{20}(\sin u) \right) \frac{1}{[1-\sin^2\alpha_o \cos^2 u]^{\frac{1}{2}}}$$

$$\ldots(6.29).$$

Equation 6.29 can be evaluated at the equator when $\gamma_o = \gamma_e$, $u = 0$ and $P_{20}(\sin u) = -\frac{1}{2}$. The equation then reduces to

$$\gamma_e = \left(\frac{KM}{a^2} - m\,\gamma_e \cos \alpha_o - \tfrac{1}{2} m\,\gamma_e F \right) \frac{1}{\cos \alpha_o} \qquad \ldots(6.30)$$

It follows that

$$\frac{kM}{a^2} = \gamma_e \left(\cos \alpha_o (1 + m) + \tfrac{1}{2} m F \right) \overset{(1.17)}{=} \gamma_e \left((1 - f)(1 + m) + \tfrac{1}{2} m F \right)$$

$$\ldots.(6.31)$$

The use of equation 6.31 in 6.29 gives

$$\gamma_o = \gamma_e \left(\cos \alpha_o (1+m \sin^2 u) + m F \{ \tfrac{1}{2} + P_{20}(\sin u) \} \right) \frac{1}{[1-\sin^2\alpha_o \cos^2 u]^{\frac{1}{2}}}$$

$$= \gamma_e \left(1 + m \{ 1 + \frac{3}{2} F \sec \alpha_o \} \sin^2 u \right) \frac{1}{[1+\tan^2\alpha_o \sin^2 u]^{\frac{1}{2}}} \qquad \ldots(6.32).$$

The parametric latitude u is replaced by the geodetic latitude ϕ by the use of equation 1.14 as follows. x and x_3 are given by the relations

$$x = a \cos u = \nu \cos \phi \; ; \text{ and } x_3 = b \sin u = \nu (1-f)^2 \sin \phi \; ,$$

where ν is PF in figure 1.3, the second equality in the first expression following from the figure itself while a proof of that in the second expression is available in (Mather 1970a, Appendix A). Thus

$$\frac{b}{a}\tan u = (1-f)^2 \tan\phi$$

or

$$\tan u = (1-f)\tan\phi \overset{(1.17)}{=} \cos\alpha_o \tan\phi \quad \ldots(6.33).$$

Thus

$$\sin u = \frac{\cos\alpha_o \sin\phi}{[\cos^2\alpha_o \sin^2\phi + \cos^2\phi]^{\frac{1}{2}}} = \frac{\cos\alpha_o \sin\phi}{[1-\sin^2\phi\,\sin^2\alpha_o]^{\frac{1}{2}}} \quad ..(6.34).$$

Thus equation 6.32 becomes

$$\gamma_o = \frac{\gamma_e}{\left(1 + \dfrac{\tan^2\alpha_o \cos^2\alpha_o \sin^2\phi}{1-\sin^2\alpha_o \sin^2\phi}\right)^{\frac{1}{2}}}\left\{1 + m\left(1 + \frac{3}{2}F\sec\alpha_o\right)\frac{\cos^2\alpha_o \sin^2\phi}{1-\sin^2\alpha_o\sin^2\phi}\right\}$$

which can be written as

$$\gamma_o = \gamma_e[1-\sin^2\alpha_o\sin^2\phi]^{\frac{1}{2}}\left\{1 + m\left(1 + \frac{3}{2}F\sec\alpha_o\right)\frac{\cos^2\alpha_o\sin^2\phi}{1-\sin^2\alpha_o\sin^2\phi}\right\} \quad \ldots(6.35).$$

Notes

(i) Equation 6.35 is a closed expression and free from approximations.

(ii) A pre-requisite for its solution is a knowledge of the value of m which is dependent on γ_e through equation 6.28. Hence an iterative procedure has to be adopted for the computation of γ_e which is obtained as special instance of equation 6.35 when $\phi = 0$.

The evaluation of F

Equation 6.35 can also be expressed for most practical purposesas a power series in f. This calls for the evaluation of F which is obtained from equation 6.27 as

$$F = \frac{q_2'(\alpha_o)\sin\alpha_o}{3q_2(\alpha_o)}$$

where

$$q_2(\alpha) \overset{(5.28)}{=} 2\left(\frac{1}{3.5}\tan^3\alpha - \frac{2}{5.7}\tan^5\alpha + \frac{3}{7.9}\tan^7\alpha - \frac{4}{9.11}\tan^9\alpha + \frac{5}{11.13}\tan^{11}\alpha + ..\right).$$

As

$$q_2'(\alpha) = 2\sec^2\alpha\left(\frac{1}{5}\tan^2\alpha - \frac{2}{7}\tan^4\alpha + \frac{3}{9}\tan^6\alpha - \frac{4}{11}\tan^8\alpha + \ldots\right),$$

$$F = \frac{\frac{2}{5}\sec\alpha_o\tan^3\alpha_o\left(1 - \{\frac{2.5}{7}\tan^2\alpha_o - \frac{3.5}{9}\tan^4\alpha_o + \frac{4.5}{11}\tan^6\alpha_o\}\right)}{\frac{2}{5}\tan^3\alpha_o\left(1 - 3.5\{\frac{2}{5.7}\tan^2\alpha_o - \frac{3}{7.9}\tan^4\alpha_o + \frac{4}{9.11}\tan^6\alpha_o\}\right)}$$

Thus

$$F = \sec\alpha_o\left(1 - \tan^2\alpha_o\left[\frac{2.5}{7} - \frac{2.3}{7}\right] + \tan^4\alpha_o\left[\frac{3.5}{9} + \frac{5}{7} - \frac{2.5.6}{49} + \frac{36}{49}\right]\right)$$

$$= \sec\alpha_o\left(1 - \frac{4}{7}\tan^2\alpha_o + \frac{68}{147}\tan^4\alpha_o\ \left[-\frac{4612}{11319}\tan^6\alpha_o\right]\right)$$

As

$$\sec\alpha_o \overset{(1.17)}{=} \frac{1}{1-f} \quad \text{and} \quad \tan^2\alpha_o = \frac{2f-f^2}{(1-f)^2} = 2f\left(1+\frac{3}{2}f + 2f^2 + o\{f^3\}\right)$$

and as F is always multiplied by $m(\simeq f)$ in the expression for γ_o, a solution which includes all terms of f or greater is obtained by approximating F by the expression

$$F = (1+f)\ \left[1 - \frac{4}{7}2f\right] = 1 - \frac{1}{7}f + o\{f^2\} \qquad \ldots(6.36)$$

The value of normal gravity at the pole (γ_p) is given by

$$\gamma_p = \gamma_e(1-f)\left(1 + m\left[1 + \frac{3}{2}\left(1 - \frac{1}{7}f\right)(1+f)\right] + o\{f^3\}\right)$$

as $\cos\alpha_o = 1 - f$.

Defining the dynamic flattening β by the relation

$$\beta = \frac{\gamma_p - \gamma_e}{\gamma_e}$$

$$= (1-f)\left(1 + m\{1 + \frac{3}{2}(1 + \frac{6}{7}f)\}\right) - 1$$

$$= \frac{5}{2}m - f - \frac{5}{2}mf + \frac{9}{7}mf,$$

it follows that

$$\beta = \frac{\gamma_p - \gamma_e}{\gamma_e} = \frac{5}{2}m - f - \frac{17}{14}mf + o\{f^3\} \qquad \ldots(6.37).$$

This equation is also known as *Clairaut's theorem*. In addition,

$$\beta = \cos\alpha_o\{1 + m(1 + \frac{3}{2}F\sec\alpha_o)\} - 1.$$

Substitution into equation 6.35, where

$$m\ (1 + \frac{3}{2}F\sec\alpha_o) = (1+\beta)\sec\alpha_o - 1$$

gives

$$\gamma_o = \gamma_e[1-(2f-f^2)\sin^2\phi]^{\frac{1}{2}}\left(1 + \{(1+\beta)(1-f)^{-1} - 1\}\ \frac{(1-f)^2\sin^2\phi}{1-(2f-f^2)\sin^2\phi}\right)$$

$$= \gamma_e[1-(2f-f^2)\sin^2\phi]^{-\frac{1}{2}}\left(1-(2f-f^2)\sin^2\phi + (\beta+f+\beta f+f^2)(1-2f)\sin^2\phi\right).$$

The grouping of terms after expansion using the binomial theorem gives

$$\gamma_o = \gamma_e\left(1 + [\beta+f -\beta f -f^2-2f+f^2+\tfrac{1}{2}(2f-f^2)]\sin^2\phi +[\tfrac{3}{2}f^2+\beta f-f^2]\sin^4\phi + o\{f^3\}\right).$$

As

$$\sin^4\phi = \sin^2\phi(1-\cos^2\phi) = \sin^2\phi - \tfrac{1}{4}\sin^2 2\phi ,$$

$$\gamma_o = \gamma_e\left(1 + [\beta- \beta f-\tfrac{1}{2}f^2+\beta f+\tfrac{1}{2}f^2]\sin^2\phi + \tfrac{1}{4}[\beta f+\tfrac{1}{2}f^2]\sin^2 2\phi\right) .$$
$$= \gamma_e[1 + \beta\sin^2\phi + \beta_2\sin^2 2\phi + o\{f^3\}] \qquad ...(6.38),$$

where

$$\beta_2 = \tfrac{1}{4}[(\tfrac{5}{2}m -f)f + \tfrac{1}{2}f^2] = \tfrac{1}{4}[\tfrac{5}{2}mf - \tfrac{1}{2}f^2] \qquad ...(6.39).$$

Equation 6.38 is the required series form of the *formula for normal gravity.*

Notes

(i) The most commonly used formula for normal gravity, given by equation 6.38, was that provided by the International Gravity Formula provided by the International Ellipsoid, defined by the parameters(e.g., Heiskanen & Moritz 1967,p.79)

$$\{a = 6\ 378\ 388 \text{ m} ; \quad f^{-1} = 297.0\} \qquad ...(6.40)$$

and the value of equatorial gravity given by

$$\gamma_e = 978\ 049.0 \text{ mgal}$$

The spheroid was based on the isostatic investigations of the U.S. geodesist Hayford in 1909, while the value of γ_e was obtained by Heiskanen in 1928 by a global analysis of the available isostatic anomalies. The adoption of these parameters is equivalent to the other constants of the reference system taking the following values.

$$\{ kM = 3.986\ 329\times10^{20}\text{cm}^3\text{sec}^{-2};\ \beta = 5.2884\times10^{-3} ;\ \beta_2 = 5.9\times10^{-6}\} \qquad ...(6.41).$$

The value of γ_e was based on the so-called Potsdam datum which was approximately 14 mgal too high.

(ii) The value of the flattening obtained from isostasy is consistent with estimates obtained from other studies based on the theory of the earth's interior being in hydrostatic equilibrium (Jeffreys 1962,p.152). The results obtained from the orbital perturbations of near earth satellites are significantly different from the above value, indicating that hydrostatic equilibrium cannot be assumed to prevail at the shallow depths of compensation implied in the isostatic assumption.

(iii) The following set of parameters was adopted at the general assembly of the International Association of Geodesy (IAG) at Lucerne in 1967 in place of the International Gravity Formula (IAG Resolutions 1967,p.367)

$$\{a = 6\ 378\ 160\ \text{m}\ ; \quad C_{20} = -1082.7\ 10^{-6}\ ; \quad kM = 3.986\ 03\ 10^{20}\text{cm}^3\text{sec}^{-2}\} \quad \ldots(6.42)$$

The equivalent flattening of the reference figure, known as *Reference Ellipsoid 1967* is approximately 1/298.25. The resulting model was called *Reference system 1967(RS 1967)* but cannot still be said to be in common use at the time of writing. Its use in geodetic work must be combined with a change of -14 mgal in all observed values of gravity based on the old Potsdam datum (Mather 1968c,p.344).

The formula for normal gravity on RS 1967 is given by (Bursa 1969; Moritz 1969)

$$\gamma_o = 978\ 031.8\ [1 + 5.3024\times10^{-3}\sin^2\phi - 5.9\times10^{-6}\sin^2 2\phi]..(6.43).$$

(iv) For a development of formulae for normal gravity using a spherical harmonic development see (Heiskanen & Vening Meinesz 1958, p.47 et seq.).

6.4 Gravity at an external point rotating with the ellipsoid

This section will consider closed expressions for the value of gravity at points rotating with the reference ellipsoid. The principles adopted are the same as those used in section 6.3. Let the general point P have co-ordinates $\{(\beta+\Delta\beta,\lambda,\alpha+\Delta\alpha)\equiv(x_i+\Delta x_i)\}$ being equivalent to the point $P_o\{(\beta,\lambda,\alpha)\equiv(x_i)\}$ on the ellipsoid, there being no change in the rotationally symmetric co-ordinate λ. The co-ordinate u has been replaced by β as differentiation is in involved in the definition of normal derivatives. Equation 5.1 applies. The differentiation of equation 1.18 with respect to β gives

$$\begin{aligned}\Delta x_1 &= -c\ \text{cosec}\ \alpha\ \text{sech}\ \beta\ \tanh\ \beta\ \cos\ \lambda\ \Delta\beta \\ \Delta x_2 &= -c\ \text{cosec}\ \alpha\ \text{sech}\ \beta\ \tanh\ \beta\ \sin\ \lambda\ \Delta\beta \quad \ldots(6.44). \\ \Delta x_3 &= c\ \cot\ \alpha\ \text{sech}^2\beta\ \Delta\beta\end{aligned}$$

Thus

$$\frac{\partial n}{\partial \beta} = c\ \text{sech}\ \beta\ [\text{cosec}^2\alpha\ \tanh^2\beta + \cot^2\alpha\ \text{sech}^2\beta]^{\frac{1}{2}}$$

$$\overset{(5.1)}{=} c\ \cos\ u\ \text{cosec}\ \alpha[1 - \sin^2\alpha\ \cos^2 u]^{\frac{1}{2}} \quad \ldots(6.45).$$

Thus normal gravity γ at the general point P exterior to the ellipsoid can be considered to have two components γ_α and γ_β.

(a) The first component γ_α is given by

$$\gamma_\alpha = -\left(\frac{\partial U}{\partial \alpha} \cdot \frac{\partial \alpha}{\partial n} \right)$$

$$\overset{(6.25)\downarrow(6.26)}{=} \frac{\sin^2\alpha}{c\,[1-\sin^2\alpha\,\cos^2 u]^{\frac{1}{2}}} \left(\frac{kM}{c} + \frac{1}{3}\frac{a^2\,\omega^2}{q_2(\alpha_o)} q_2'(\alpha) P_{20}(\sin u) - c^2 \operatorname{cosec}^2\alpha \cot\alpha \cos^2 u\, \omega^2 \right)$$

$$= \frac{1}{[1-\sin^2\alpha\,\cos^2 u]^{\frac{1}{2}}} \left(\frac{kM}{a^2} \left(\frac{a}{a_p}\right)^2 + m\,\gamma_e \left(\frac{a}{a_p}\right) \frac{1}{3} \frac{\sin\alpha}{q_2(\alpha_o)} q_2'(\alpha) P_{20}(\sin u) - m\,\gamma_e \left(\frac{a_p}{a}\right) \cos\alpha \cos^2 u \right) \quad \ldots(6.46),$$

where m, kM and γ_e have the same significance as in section 6.3.

(b) The second component γ_β is given by the differentiation of equation 6.12 according to the relation

$$\gamma_\beta = -\left(\frac{\partial U}{\partial \beta} \cdot \frac{\partial \beta}{\partial n} \right)$$

$$= \frac{\sin\alpha}{c\cos u\,[1-\sin^2\alpha\cos^2 u]^{\frac{1}{2}}} \left(-c^2 \operatorname{cosec}^2\alpha\, \operatorname{sech}^2\beta \tanh\beta\, \omega^2 + \frac{1}{3}\frac{a^2\,\omega^2}{q_2(\alpha_o)} q_2(\alpha) \times \frac{3}{2}\, 2\tanh\beta\, \operatorname{sech}^2\beta \right)$$

$$\overset{(5.1)\downarrow(6.28)}{=} \frac{m\gamma_e \cos u \sin u}{[1-\sin^2\alpha\cos^2 u]^{\frac{1}{2}}} \left(-\left(\frac{a_p}{a}\right) + \left(\frac{a}{a_p}\right) \frac{q_2(\alpha)}{q_2(\alpha_o)} \right) \quad \ldots(6.47)$$

where γ_β is directed in the direction of decrease of β with n, i.e., the pole., and a_p is the equatorial radius of the ellipsoid through P.

Notes

(i) The first term in equation 6.46 makes the major contribution, being the gravitational attraction of a sphere with the same mass and at an external point. The second term is that due to the rotation while the third is due to the ellipticity. The total acceleration due to gravity is given by

$$\gamma = [\gamma_\alpha^2 + \gamma_\beta^2]^{\frac{1}{2}} \quad \ldots(6.48)$$

and acts in a direction which is not along the normal to the spheroid at P_s but along that to the spherop at P as shown in figure 6.1. Thus the direction of the spherop normal at P differs from that at the equivalent point on the spheroid by an angle $\Delta\phi_n$ given by

$$\tan\Delta\phi_n = \frac{\gamma_\beta}{\gamma_\alpha} \quad \ldots(6.49),$$

where the magnitude of γ_β is dependent on both the ellipticity and the rotational characteristics of the ellipsoid.

(ii) *Series expansions*

The change of variables which occur between the general point P and the equivalent point on the reference ellipsoid are in

(a) α from α to α_o

(b) u from u to u_o

or ϕ from ϕ to ϕ_o.

These changes occur on moving along the trajectory orthogonal to the familiy of spherops defining the resultant gravitational field exterior to the reference ellipsoid. Position exterior to the latter for a given value of ϕ can be expressed in terms of one of three variables. They are:-

1. The difference in potential U between the values at the reference ellipsoid U_o and at P(U). Thus

$$\Delta U = U_o - U = \Delta U(\phi,h) = \Delta U(\phi,\Delta\alpha)$$

where $\Delta\alpha$ is the change in α between the ellipsoid(α_o) and P(α)

2. $\Delta\alpha$ is given by

$$\Delta\alpha = \alpha_o - \alpha \qquad \ldots(6.50)$$

3. The spheroidal elevation h given by

$$h = h(\phi, \Delta U) = h(\phi, \Delta\alpha).$$

6.5 The change in spheropotential ΔU with Δα

The potential U at an exterior point $P(u,\lambda,\alpha_o-\Delta\alpha)$ is given from equations 6.12 and 6.50 as

$$U = \frac{kM}{c}(\alpha_o-\Delta\alpha) + \tfrac{1}{2}a_p^2\omega^2\cos^2 u + \frac{1}{6}a^2\omega^2\frac{q_2(\alpha_o-\Delta\alpha)}{q_2(\alpha_o)}[3\sin^2 u - 1]$$

$$\overset{(6.28)}{=} \frac{kM}{c}(\alpha_o-\Delta\alpha) + {}_e ma \left(\tfrac{1}{2}\left(\frac{a_p}{a}\right)^2(1-\sin^2 u) + \frac{1}{6}\frac{q_2(\alpha_o-\Delta\alpha)}{q_2(\alpha_o)}\beta\sin^2 u - 1\right)$$

The term γ_e a is obtained from equation 6.30 as

$$a\gamma_e \doteq \frac{kM}{a(1-f)}\left(1 - m(1-f) - \tfrac{1}{2}m(1-\tfrac{1}{7}f)\right)$$

$$= \frac{kM}{b}[1 - \tfrac{3}{2}m + o\{f^2\}] = \frac{kM}{a}[1 - \tfrac{3}{2}m + f + o\{f^2\}]$$

or, more precisely,

$$\gamma_e a = \frac{kM}{b}[1 - (1-\tfrac{3}{2}m) - \tfrac{1}{2}m(1-\tfrac{3}{2}m+f)(1-\tfrac{1}{7}f)]$$

$$= \frac{kM}{b}[1 - \tfrac{3}{2}m + \tfrac{9}{4}m^2 - \tfrac{3}{7}mf + o\{f^3\}] \qquad \ldots(6.51),$$

as $\cos\alpha_o = 1 - f$.

As equation 6.13 can be written as

$$U_o = \frac{kM}{c}\alpha_o + \frac{1}{3}\gamma_e m a \quad ,$$

it follows that

$$\Delta U = U_o - U = \frac{kM}{c}\Delta\alpha + \gamma_e m a\left(\frac{1}{3} - \tfrac{1}{2}\left(\frac{a_p}{a}\right)^2(1-\sin^2 u) + \frac{1}{6}\frac{q_2(\alpha_o-\Delta\alpha)}{q_2(\alpha_o)} - \sin^2 u\left(\tfrac{1}{2}\frac{q_2(\alpha_o-\Delta\alpha)}{q_2(\alpha_o)} - \tfrac{1}{2}\left(\frac{a_p}{a}\right)^2\right)\right) \quad ..(6.52).$$

It should be noted that

$$\frac{\Delta\alpha}{\alpha} \simeq \frac{\Delta U}{U} \simeq 10^{-3}.$$

As $\alpha \simeq 10^{-1}$, $\Delta\alpha \simeq 10^{-4}$ and terms of $(\Delta\alpha)^2$ which are of order 10^{-8} will be treated as negligible. In addition,

$$\gamma_e a \simeq \frac{kM}{c}\alpha$$

and $m\,\Delta\alpha \simeq 10^{-7}$.

The term $(a_p/a)^2$ is evaluated as follows, noting that it is always multiplied by m.

$$\frac{a_p}{a} = \frac{c\,\mathrm{cosec}(\alpha_o-\Delta\alpha)}{c\,\mathrm{cosec}\,\alpha_o} = \frac{\sin\alpha_o}{\sin\alpha_o - \Delta\alpha\cos\alpha_o - \tfrac{1}{2}(\Delta\alpha)^2\sin\alpha_o}$$

$$= 1 + \Delta\alpha\cot\alpha_o + \tfrac{1}{2}(\Delta\alpha)^2(1 + 2\cot^2\alpha_o) + o\{(\Delta\alpha\cot\alpha_o)^3\} \quad .$$

Thus

$$\left(\frac{a_p}{a}\right)^2 = 1 + 2\,\Delta\alpha\cot\alpha_o + (\Delta\alpha)^2(1 + 3\cot^2\alpha_o) + o\{(\Delta\alpha\cot\alpha_o)^3\} \quad .$$

A consideration of equation 5.28 gives

$$q_2(\alpha_o-\Delta\alpha) = \frac{2}{15}\tan^3(\alpha_o-\Delta\alpha)\left(1 - 15\left[\frac{2}{35}\tan^2(\alpha_0-\Delta\alpha) - \frac{1}{21}\tan^4(\alpha_o-\Delta\alpha)\right]\right).$$

As

$$\tan(\alpha_o-\Delta\alpha) = \tan\alpha_o - \Delta\alpha\sec^2\alpha_o + o\{10^{-9}\} \quad ,$$

$$\tan^2(\alpha_o-\Delta\alpha) = \tan^2\alpha_o(1 - 2\,\Delta\alpha\sec^2\alpha_o\cot\alpha_o + o\{10^{-8}\})$$

and

$$\tan^3(\alpha_o-\Delta\alpha) = \tan^3\alpha_o(1 - 3\,\Delta\alpha\sec^2\alpha_o\cot\alpha_o + o\{10^{-8}\}),$$

$$\frac{q_2(\alpha_o-\Delta\alpha)}{q_2(\alpha_o)} = [1 - 3\,\Delta\alpha\sec^2\alpha_o\cot\alpha_o][1 - \frac{6}{7}\tan^2\alpha_o(1 - 2\,\Delta\alpha\sec^2\alpha_o\cot\alpha_o)]\times [1 + \frac{6}{7}\tan^2\alpha_o]$$

$$= 1 - \Delta\alpha\cot\alpha_o\sec^2\alpha_o(3 - \frac{12}{7}\tan^2\alpha_o)$$

$$= 1 - \Delta\alpha\cot\alpha_o[\,3 + \frac{9}{7}\tan^2\alpha_o] + o\{(\Delta\alpha\cot\alpha_o)^2\} \quad ...(6.55).$$

The use of equations 6.53 and 6.55 in equation 6.52 gives

$$\Delta U = \frac{kM}{c}\Delta\alpha + \gamma_e ma\left[\left(\frac{1}{3} - \tfrac{1}{2} - \Delta\alpha\cot\alpha_c - \tfrac{1}{2}(\Delta\alpha)^2(1+3\cot^2\alpha_o + \right.\right.$$

$$\frac{1}{6} - \frac{1}{6}\Delta\alpha\cot\alpha_o(3 + \frac{9}{7}\tan^2\alpha_o) + \frac{1}{6}(\Delta\alpha)^2\sec^2\alpha_o(6+3\cot^2\alpha_o - \frac{36}{7}\sec^2\alpha_o)\Big) -$$

$$\sin^2 u\,\Big(\tfrac{1}{2} - \tfrac{1}{2}\Delta\alpha\cot\alpha_o(3+\frac{9}{7}\tan^2\alpha_o) + \tfrac{1}{2}(\Delta\alpha)^2\sec^2\alpha_o(6+3\cot^2\alpha_o -$$

$$\left.\left.\frac{36}{7}\sec^2\alpha_o) - \tfrac{1}{2} - \Delta\alpha\cot\alpha_o - \tfrac{1}{2}(\Delta\alpha)^2(1+3\cot^2\alpha_o)\right)\right]$$

The use of equation 6.51 together with the relation

$$c = b\tan\alpha$$

gives

$$\Delta U = \frac{kM}{b}\Delta\alpha\cot\alpha_o\left[1 + m(1-\frac{3}{2}m)\left(-1-\frac{1}{6}(3+\frac{9}{7}\tan^2\alpha_o) - \Delta\alpha\cot\alpha_o\right.\right. -$$

$$\left.\left.\sin^2 u(-\frac{5}{2} - \frac{9}{14}\tan^2\alpha_o)\right)\right]$$

which, on rounding off to order f^3 gives

$$\Delta U = \frac{kM}{b}\Delta\alpha\cot\alpha_o\left[\left(1 - \frac{3}{2}m + \frac{9}{4}m^2 - \frac{3}{14}m\tan^2\alpha_o - m\Delta\alpha\cot\alpha_o\right) + \right.$$

$$\left.\sin^2 u\left(\frac{5}{2}m + \frac{9}{14}m\tan^2\alpha_o - \frac{15}{4}m^2\right) + o\{f^3\}\right] \quad \ldots(6.56)$$

The solution is completed on replacing u by ϕ using equation 6.34 when

$$\sin^2 u = \frac{\cos^2\alpha\sin^2\phi}{1-\sin^2\alpha\sin^2\phi}$$

$$= \cos^2\alpha\sin^2\phi[1+\sin^2\alpha\sin^2\phi] + o\{f^2\}$$

$$= \cos^2\alpha_o(1+2\Delta\alpha\tan\alpha_o)\sin^2\phi\,\{1+\sin^2\alpha_o(1-2\Delta\alpha\cot\alpha_o)\sin^2\phi\}$$

$$= \sin^2\phi\cos^2\alpha_o(1+2\Delta\alpha\tan\alpha_o) + \sin^4\phi\cos^2\alpha_o\sin^2\alpha_o$$

$$= \sin^2\phi\,(1-2f) + 2f\sin^4\phi + o\{f^2\} \quad \ldots(6.57).$$

Thus, on replacing $\tan\alpha_o$ by $2f + o\{f^2\}$,

$$\Delta U = \frac{kM}{b}\Delta\alpha\cot\alpha_o\left[1-\frac{3}{2}m + \frac{9}{4}m^2 - \frac{3}{7}mf - m\Delta\alpha\cot\alpha_o + \right.$$

$$\left.\sin^2\phi[\frac{5}{2}m + \frac{9}{7}mf - \frac{15}{4}m^2 - 5mf] + 5mf\sin^4\phi + o\{f^3\}\right]$$

$$= \frac{kM}{b}\Delta\alpha\cot\alpha_o\left[1-\frac{3}{2}m + \frac{9}{4}m^2 - \frac{3}{7}mf + \sin^2\phi[\frac{5}{2}m - \frac{15}{4}m^2 - \frac{26}{7}mf] - \right.$$

$$\left. m\Delta\alpha\cot\alpha_o + 5mf\sin^4\phi + o\{f^3\}\right].$$

The term $\Delta\alpha \cot \alpha_o$ within the bracket is multiplied by m and hence its magnitude is only approximately required for evaluation therein. This is given by

$$\Delta\alpha \cot \alpha_o = b \frac{\Delta U}{kM} .$$

Its more precise value is obtained by inverting the expression for when

$$\Delta\alpha \cot \alpha_o = \frac{b}{kM} \Delta U \left\{1 + \frac{3}{2}m + m^2\left(\frac{9}{4} - \frac{9}{4}\right) + \frac{3}{7}mf + m\, b\frac{\Delta U}{kM} - \sin^2\phi\left(\frac{5}{2}m - \frac{15}{4}m^2 + \frac{15}{2}m^2 - \frac{26}{7}mf\right) + \sin^4\phi\left(\frac{25}{4}m^2 - 5mf\right)\right\}$$

$$= \frac{b}{kM} \Delta U \left\{1 + \frac{3}{2}m + \frac{3}{7}mf + mb\frac{\Delta U}{kM} - \sin^2\phi\left(\frac{5}{2}m + \frac{15}{4}m^2 - \frac{26}{7}mf\right) + \left(\frac{25}{4}m^2 - 5mf\right)\sin^4\phi + o\{10^{-8}\}\right. \quad ..(6.58) .$$

Notes

(i) Terms of order $(\Delta\alpha)^2$ have been ignored in the above development as they are of order 10^{-8}.

(ii) For a development correct to order $(\Delta\alpha)^2$, see (Hirvonen 1960, pp.17 et seq.)

(iii) Equation 6.56 gives the change in geopotential for a change $\Delta\alpha$ in α. Note that $(kM/b) \simeq U$.

(iv) Thus the term $(m\, b\Delta U/kM) \simeq f^2$

6.6 The evaluation of the angle $\Delta\phi_n$ between the spheroid and associated spherop normals at the surface of the earth

The angle $\Delta\phi_n$ defines the displacement between the normals to spherop and the member of the family of confocal spheroids which passes through the relevant point at the surface of the earth, being obtained from equations 6.49, 6.46 and 6.47. The use of equations 6.51 and 6.54 in equation 6.47 gives

$$\gamma_\beta = \frac{m\,\gamma_e \cos u \sin u}{(1-\sin^2\alpha \cos^2 u)^{\frac{1}{2}}} \left\{-1 - \Delta\alpha \cot \alpha_o + (1 - \Delta\alpha \cot \alpha_o)(1 - 3\,\Delta\alpha \cot \alpha_o + o\{(\Delta\alpha \cot \alpha_o)^2\}\right\}$$

$$= \frac{m\,\gamma_e \cos u \sin u}{(1-\sin^2\alpha \cos^2 u)^{\frac{1}{2}}} \Delta\alpha \cot \alpha_o \left(-5 + o\{\Delta\alpha \cot \alpha_o\}\right) \quad ..(6.59) .$$

As $\Delta\alpha \cot \alpha_o \simeq 10^{-3}$, $\gamma_\beta \simeq 10^{-6}\, \gamma_e$.

Thus $|\tan \Delta\phi_n| \simeq |\gamma_\beta/\gamma_\alpha| \simeq 10^{-6}$ or $(\Delta\phi_n)^3 \simeq 10^{-16}$.

Hence

$$\Delta\phi_n = \frac{\gamma_\beta}{\gamma_\alpha} + o\{10^{-16}\}$$

and the expression for γ_α is only required to order $\Delta\alpha \cot \alpha_o$. This is obtained from equations 6.31 and 6.36, together with equation 6.51 as

$$\gamma_\alpha = \frac{\gamma_e}{(1-\sin^2\alpha \cos^2 u)^{\frac{1}{2}}}\left(1 + \tfrac{3}{2}m - f + \tfrac{1}{2}m(3\sin^2 u - 1) - m(1-\sin^2 u) + o\{f^2\}\right)$$

$$= \frac{\gamma_e}{(1-\sin^2\alpha \cos^2 u)^{\frac{1}{2}}}\left(1 - f + \tfrac{5}{2}m \sin^2 u + o\{f^2\}\right).$$

It follows from equation 6.59 that the outward spherop normal is displaced towards the pole in relation to the spheroid normal. Adopting this as the magnitude for $\Delta\phi_n$,

$$\Delta\phi_n = m \cos u \sin u \, (\Delta\alpha \cot \alpha_o)\left(5 + o\{f\}\right) \qquad \ldots(6.60).$$

Proceeding on lines similar to those used in the derivation of equation 6.57,

$$\sin u = \sin\phi \left(1 - f + f \sin^2\phi + o\{f^2\}\right)$$

and

$$\cos u = \cos\phi \left(1 + f \sin^2\phi + o\{f^2\}\right).$$

Thus,

$$\Delta\phi_n = m \sin\phi \cos\phi \, (\Delta\alpha \cot \alpha_o)\left(5 + o\{f\}\right) \qquad \ldots(6.61).$$

6.7 Changes $\Delta\gamma$ in normal gravity with potential (ΔU)

Strict expressions for the required changes can be obtained by operating on equations 6.36, 6.46 and 6.47, adopting principles similar to those used in sections 6.5 and 6.6. These results can however be obtained with comparative ease by the use of relations derived in the above mentioned sections along with development using Taylor series.

The spheropotential U at any point P can be expressed by the series

$$U = U_o + h\frac{\partial U}{\partial h} + \tfrac{1}{2} h \frac{\partial^2 U}{\partial h^2} + \ldots \qquad \ldots(6.62),$$

where h is the spheroidal elevation and U_o the value of the potential on the reference ellipsoid. It can be seen from section 2.3 that the gravity vector is normal to the local equipotential. It can also be seen from equations 6.49 and 6.60 that the change in direction between spherop and ellipsoid normals at equivalent points is of order (i.e., 2 arc sec or less). Thus

$$\frac{\partial U}{\partial h} = -\gamma + o\{10^{-10}\} \qquad \ldots(6.63)$$

and

$$U = U_o - U = \gamma h - \tfrac{1}{2} h^2 \frac{\partial \gamma}{\partial h} + o\{10^{-10}\} \quad \ldots(6.64).$$

The quantity $(\partial\gamma/\partial h)$ is known as the vertical gradient of normal gravity. It is most conveniently evaluated by the application of Laplace's equation at the surface of a spherop

$$U(x_i) = U_c . \qquad \ldots(6.65).$$

The curvatures c_α in the principal sections of the spherop are given by

$$c_\alpha = \frac{1}{\rho_\alpha} = -\frac{\left(d^2x_3/dx_\alpha^2\right)}{\left\{1+\left(\frac{dx_3}{dx_\alpha}\right)^2\right\}^{3/2}} \qquad \ldots.(6.66),$$

where ρ_α are the principal radii of normal curvature with reference to a triply orthogonal local Cartesian system with the positive direction of the x_3 axis coincident with the local outward normal. In practice, the x_1x_2 plane coincides with the tangent plane at the local point and the x_1 and x_2 axes can be taken as being oriented north and east respectively.

As the equation of the equipotential surface is given by equation 6.65, differentiation with respect to x_α along the surface in the immediate vicinity of the origin gives

$$\frac{\partial U}{\partial x_\alpha} + \frac{\partial U}{\partial x_3}\frac{dx_3}{dx_\alpha} = 0$$

Repetition of the procedure results in

$$\frac{\partial^2 U}{\partial x_\alpha^2} + \frac{\partial^2 U}{\partial x_3 \partial x_\alpha}\frac{dx_3}{dx_\alpha} + \frac{\partial^2 U}{\partial x_\alpha \partial x_3}\frac{dx_3}{dx_\alpha} + \frac{\partial U}{\partial x_3}\frac{d^2 x_3}{dx_\alpha^2} + \frac{\partial^2 U}{\partial x_3^2}\frac{dx_3}{dx_\alpha} = 0.$$

As the differentiation is performed along an equipotential surface with the x_3 axis oriented along the normal,

$$\frac{dx_3}{dx_\alpha} = 0$$

Thus

$$\frac{\partial^2 U}{\partial x_\alpha^2} = -\frac{\partial U}{\partial x_3}\frac{d^2 x_3}{dx_\alpha^2} .$$

If ρ and ν are the principal radii of normal curvature in the case of the reference ellipsoid, being in the meridian and prime vertical sections respectively, the use of equations 6.63 and 6.66 gives

$$\frac{\partial^2 U}{\partial x_1^2} + \frac{\partial^2 U}{\partial x_2^2} = -\gamma\left(\frac{1}{\rho} + \frac{1}{\nu}\right) .$$

The use of equations 2.8 and 2.29 which define Laplace's equation in the case of a rotating body, gives

$$\nabla^2 U = \sum_{i=1}^{3} \frac{\partial^2 U}{\partial x_i^2} = 2\omega^2 \qquad \ldots(6.67).$$

As

$$\frac{\partial^2 U}{\partial x_3^2} = -\frac{\partial \gamma}{\partial h} ,$$

appropriate substitution in equation 6.67 gives

$$\frac{\partial \gamma}{\partial h} = -\gamma\left[\frac{1}{\rho} + \frac{1}{\nu}\right] - 2\omega^2 \qquad \ldots(6.68).$$

This expression, attributed to Bruns, gives the vertical gradient of normal gravity. On using the well known relations(e.g., ibid, appendix A)

$$\rho = \frac{a(1-f)^2}{[1-(2f-f^2)\sin^2\phi]^{3/2}} \quad \text{and } \nu = \frac{a}{[1-(2f-f^2)\sin^2\phi]^{\frac{1}{2}}} ,$$

$$\frac{1}{\rho} + \frac{1}{\nu} = \frac{1}{a}(1-f)^{-2}[1-(2f-f^2)\sin^2\phi]^{\frac{1}{2}}[1-(2f-f^2)\sin^2\phi + (1-f)^2]$$

$$= \frac{2}{a}\Big([1+2f][1-f\sin^2\phi\ ][1-f-f\sin^2\phi]\Big) + o\{f^2\}$$

$$= \frac{2}{a}\,[1 + f - 2f\sin^2\phi + o\{f^2\}] .$$

Also

$$\omega^2 = \frac{m\,\gamma_e}{a} \overset{(6.38)}{\doteq} \frac{m\gamma}{a}\,[1 - \sin^2\phi] = \frac{m\gamma}{a}[1+(\tfrac{5}{2}m-f)\sin^2\phi] .$$

The combination of the above results gives

$$\frac{\partial \gamma}{\partial h} = -2\,\frac{\gamma}{a}\,[1 + m + f - 2f\sin^2\phi + o\{f^2\}] \qquad \ldots(6.69).$$

The use of this result in equation 6.62 gives

$$\Delta U = U_o - U = \gamma\, h - \frac{\gamma\, h^2}{a}[1 + m + f - 2f\sin^2\phi] + o\{f^3\}$$

$$= \gamma\, h\,\Big(1 - \frac{h}{a}\,[1 + m + f - 2f\sin^2\phi] + o\{f^3\}\Big)$$

Thus h is given by

$$h = \frac{\Delta U}{\gamma}\left(1 + \frac{h}{a}[1 + m + f - 2f\sin^2\phi\,] + \left(\frac{h}{a}\right)^2 + o\{f^3\}\right)$$

$$= \frac{\Delta U}{\gamma}\left(1 + \frac{\Delta U}{a\gamma}[1 + m + f - 2f\sin^2\phi\,] + \left(\frac{\Delta U}{a\gamma}\right)^2 + o\{f^3\}\right) \quad ..(6.70).$$

Notes

(i) The relation between spheroidal elevation h and geopotential ΔU is best studied by defining the latter in kgal m (kilogal metres) which are given by

$$1 \text{ kgal m} = 10^{-5} \times \text{cm}^2\text{sec}^{-2}.$$

Thus

$$\Delta U^{(\text{kgal m})} = h^{(m)} \times \gamma^{(\text{kgal})}$$
$$= 0.98 \times h^{(m)}$$

or

$$\Delta U^{(\text{kgal m})} = 98\%\ h^{(m)} \qquad ...(6.71).$$

Thus spheroidal evelations are approximately 2% greater than the difference in geopotential in kgal m. Equation 6.70 is correct to order f^3, being correct to 0.1 mm in all circumstances on the earth's surface. Terms of order f in this equation reach a magnitude of 1 mm when $h \simeq 25$ m, while those of order f^2 achieve this magnitude when $h \simeq 1.5$ km ...(6.72)

(ii) The geometrical significance of the spheroidal elevation has not been specified. In the context of equation (6.62), it is merely a parameter defining position on normals to successive equipotential surfaces and not the distance along the ellipsoid normal which passes through the surface point. As the maximum displacement is of order f^2, this is equivalent to linear differences of order 10^{-9}, which is of the order of 10^{-2} mm even for a 10 km elevation and such differences can be neglected. Thus linear displacements along the curve normal to successive equipotential surfaces are, for all practical purposes, equal in length to lengths along the relevant ellipsoid normal.

6.8 Normal gravity γ_p at a surface point

The value of normal gravity γ_p at a surface point can also be defined by a Taylor series of the form

$$\gamma_p = \gamma_o + h\frac{\partial\gamma}{\partial h} + \tfrac{1}{2}h^2\frac{\partial^2\gamma}{\partial h^2} + \ldots \qquad ...(6.73).$$

The contribution of the third term is extremely small and can be obtained from a consideration of the first term in equation 6.46, when

$$\frac{\partial\gamma}{\partial h} = -2\,\frac{kM}{a_p^3} + o\{f\frac{\partial\gamma}{\partial h}\} \qquad \text{and} \qquad \frac{\partial^2\gamma}{\partial h^2} = 6\,\frac{kM}{a_p^4} + o\{f\frac{\partial^2\gamma}{\partial h^2}\}$$

or

$$\frac{\partial^2\gamma}{\partial h^2} = 6\,\frac{\gamma}{a^2} + o\{f\frac{\partial^2\gamma}{\partial h^2}\} \qquad \ldots(6.74).$$

The combination of equations 6.74, 6.69 and 6.73 gives

$$\Delta\gamma = \gamma_p - \gamma_o = -2\gamma\left(\frac{h}{a}\,[1 + f + m - 2f\sin^2\phi] - \frac{3}{2}\left(\frac{h}{a}\right)^2 + o\{f^3\}\right) \qquad \ldots(6.75).$$

Notes

(i) The quantity $\Delta\gamma$ is commonly known as the free air reduction.

(ii) The spheroidal elevation is a derived rather than an observed quantity except when expensive astro-geodetic procedures are resorted to. It is therefore relevant to define the free air reduction in terms of the difference in potential ΔU which, in practice, is a measured quantity.

The incorporation of equation 6.70 in 6.75 gives

$$\Delta\gamma = -2\gamma\left(\frac{\Delta U}{a}[1 + f + m - 2f\sin^2\phi + \frac{\Delta U}{a\gamma}\,] - \frac{3}{2}\left(\frac{\Delta U}{a\gamma}\right)^2\right)$$

$$= -\frac{2\,\Delta U}{a}[1 + f + m - 2f\sin^2\phi - \tfrac{1}{2}\frac{\Delta U}{a\gamma} \qquad + o\{f^2\}]\ldots(6.76).$$

(iii) A development, correct to order f^3, is given by Hirvonen(1960,pp.17-24).

(iv) The magnitude of the principal term in equation 6.75 is

$$\frac{2\gamma h}{a} \simeq \frac{2\times10^6}{6\times10^6}\,h^{(m)} \simeq 0.3\,h^{(m)}\ \text{mgal}.$$

It is necessary to use the terms of order f^2 if the correction $\Delta\gamma$ is required to the nearest tenth-mgal in the case where $h > 1$ km. It is also necessary to know the elevation to the nearest 1/3 of a metre.

(v) It is conventional to adopt the correction

$$c = 0.3086\,h^{(m)}\ \text{mgal} \qquad \ldots(6.77)$$

as the free air reduction for practical use.

(vi) Laplace's equation does not hold in regions occupied by matter, in which case Poisson's equation holds. This expression is given in the case of a rotating body by equation 2.29. Equation 6.68 takes the form

$$\frac{\partial\gamma}{\partial h} = -\left(\frac{1}{\rho}+\frac{1}{\nu}\right) - 2\omega^2 + 4\pi k\rho' \qquad \ldots(6.78)$$

in this case, ρ' being the density of matter at the point where the gradient is being evaluated.

Summary

(i) The angle between the normal to the confocal spheroid and the equipotential surface of the reference system at a point in space exterior to the reference ellipsoid is given by

$$\Delta\phi_n = \frac{5}{2}m \sin 2\phi \quad \frac{h}{a} + o\{f^3\} \qquad \ldots(6.79),$$

as

$$\Delta\alpha \cot \alpha_o \overset{(6.58)}{\doteq} \frac{\Delta U}{a\gamma} + o\{f^2\} \overset{(6.70)}{\doteq} \frac{h}{a} + o\{f^2\}\ . \qquad \ldots(6.80).$$

(ii) Given the difference in geopotential ΔU between a point P_s on the reference ellipsoid and the general point P,

(a) the spheroidal height h_s is given by

$$h_s \overset{(6.70)}{\doteq} \frac{\Delta U}{\gamma}\left[1 + \frac{\Delta U}{a\gamma}(1 + f + m - 2f \sin^2\phi) + \left(\frac{\Delta U}{a\gamma}\right)^2 + o\{f^2\}\right];$$

(b) the difference in normal gravity $\Delta\gamma$ is given by

$$\Delta\gamma \overset{(6.76)}{\doteq} -2\frac{\Delta U}{a}\left(1 + f + m - 2f \sin^2\phi - \tfrac{1}{2}\frac{\Delta U}{a\gamma} + o\{f^2\}\right)^1$$

$$\overset{(6.75)}{\doteq} -2\frac{\gamma h}{a}\left(1 + f + m - 2f \sin^2\phi - \frac{3}{2}\frac{h}{a} + o\{f^2\}\right).$$

(iii) The above equations define both the position and gravitational characteristics of all points in the space exterior to the reference ellipsoid and rotating with the same angular velocity, provided the spheroidal elevations are not in excess of 20 km.

The system described in this section affords a reference model for studying the gravity field of the earth provided the difference of geopotential in relation to the geoid were known at all surface points.

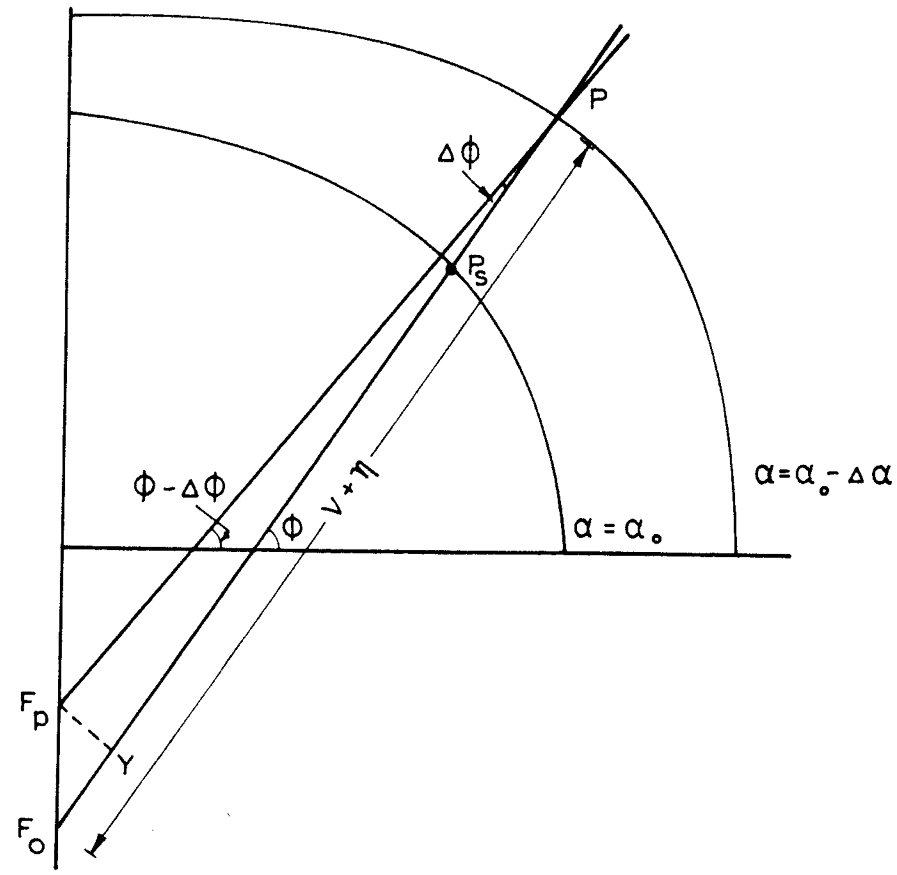


FIG. 6.2
Oblate spheroidal co-ordinates and elevations

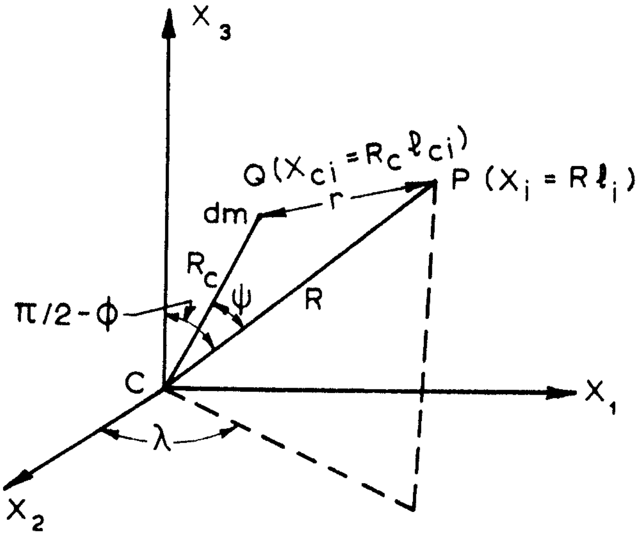


FIG. 7.1
The geocentric and associated co-ordinates

7. THE EXTERIOR GRAVITY FIELD OF THE EARTH

7.1 Introduction

It has already been shown in sections 4.1 and 4.9 that the gravity field of the earth exterior to it can be represented fully by a spherical harmonic series of the form

$$W = \frac{kM}{r} \sum_{n=0}^{\infty} \left(\frac{a_e}{r}\right)^n \sum_{m=0}^{n} P_{nm}(\mu)\,[\, C_{nm}\cos m\lambda + S_{nm}\sin m\lambda\,] \qquad \ldots(7.1)$$

where a_e is the equatorial radius of the earth, $\mu = \cos\theta$ and (r,θ,λ) define the set of spherical co-ordinates. A study of table 8.3 shows that the coefficients C_{nm}, S_{nm} are all of magnitude 10^{-6} except C_{00} which is unity and C_{20} which is of order 10^{-3}. Thus the contributions of the former terms to the magnitude of W will each be of the order of 10 kgal m or less, while those of C_{20} & C_{00} are of the order of 10^4 kgal m and 6×10^6 kgal m respectively.

It is therefore desirable to remove the effects of the terms C_{00} and C_{20} in order that the contributions of all terms to the function being studied are of the same order of magnitude. If the earth were a sphere to order f^2, all terms other than C_{00} would make contributions of order f^2. The earth is not a rotational ellipsoid to order f due to departures of the topography from this shape. It would therefore appear that all values of C_{nm},S_{nm} should also be of order f like C_{20}. This is not so in the case of the exterior field due to the existence of isostatic compensation.

The deviations of the mass distribution of the earth from the isostatic model influence the orbits of near earth artificial satellites. As the resulting coefficients defining the earth's exterior gravitational field are of order f^2, it must be assumed that the gravity anomalies based on such a density model must also be of the same order of magnitude. The effect of the higher degree harmonics which are of short wave length, is rapidly damped out as expected from purely empirical considerations, due to the term $(a_e/r)^n$, as shown in table 4.1

Thus the harmonic coefficients determined from the analysis of the orbital perturbations of near earth satellites cannot necessarily be considered to be applicable at the physical surface of the earth as the effect of the higher degree harmonics, which are damped out by the factor $(a_e/r)^n$ at the relatively higher altitudes relevant to satellite orbits, re-assert their effect at lower altitudes and hence, any evaluation of W based purely on the analysis of satellite orbital perturbations provide an over-smoothened representation of the gravitational field at the earth's surface.

Low degree harmonics, on the other hand, are less effected by the departures of (a_e/r) from unity and hence more reliably determined from satellite orbital analysis than those of higher degree. These low degree harmonics afford a convenient means of estimating values of gravity for the representation of the unsurveyed regions of the world. The techniques currently adopted for doing so are developed in section 7.3. The higher the degree of harmonics determined from such analysis, the more accurate will be the field extension. The upper limit to the degree of the harmonic which can be adequately determined in this manner is governed by the lowest possible altitude at which a satellite can orbit the earth without suffering significant drag effects as a consequence of friction between the satellite and the atmosphere. This fixes the minimum altitude of geodetically useful satellites at 200 km. It can be seen from table 4.1 that harmonics of degree less than 6 can be determined with confidence from satellite orbital analysis alone as the elliptical orbits have apogee height of over 1000 km in this case. Analyses up to degree 20 should be possible if an appropriate variety of orbital sizes and inclinations are included in the solution with appropriate computational precautions. Harmonic coefficients up to degree 8 and zonal harmonics of higher degree seem to be capable of reliable determination at the time of writing though combined solutions up to degree 16 have been obtained using techniques outlined in section 7.3.

Computational stability is obtained by resorting to the procedure of linearisation. The principles involved are as follows. A model is adopted for the function being represented in order that the residual variations between the real magnitudes of the function and those predicted from the adopted model are small, preferably giving rise to terms of the same order of magnitude in the representation of the linearised function.

In the case of the geopotential W represented by equation 7.1, linearisation can be effected to a limited degree by adopting a spherical approximation for the earth. The potential U of a sphere with the same mass as the earth at an external point P, distant r from the centre of the sphere, is given by

$$U = \frac{kM}{r} .$$

Thus the *disturbing potential* V_d obtained by linearisation is

$$V_d = W - U$$

$$= \frac{kM}{r} \sum_{n=2}^{\infty} \left(\frac{a_e}{r}\right)^n \sum_{m=0}^{n} P_{nm}(\sin\phi)[C_{nm}\cos m\lambda + S_{nm}\sin m\lambda] \qquad \ldots(7.2)$$

if the centre of the sphere were at the centre of mass of the earth. This is developed further in section 7.2. It should however be noted that C_{20} is still three orders of magnitude larger than the other coefficients. Its effect could be eliminated by adopting a spheroidal model of reference with symmetry about the minor axis. The resulting spheropotential U can be expressed by equations of the form

$$U = \frac{kM}{r} \sum_{n=0}^{\infty} \left(\frac{a_e}{r}\right)^{2n} P_{(2n)0}(\sin\phi)\, C_{(2n)0} \qquad \ldots(7.3).$$

The coefficients which result from the use of equation 7.3 in defining 7.2 will now be of the same order of magnitude (10^{-6}).

The spherical reference model will however be used when studying the dynamic effects of the earth's gravitational field on the orbits of near earth satellites as the use of such a model with a uniform density distribution will result in a satellite orbital plane which remains fixed in inertial space (i.e., space which does not rotate with the earth) if not influenced by any other factors.

7.2 The dynamic significance of selected low degree harmonics of the earth's gravitational field

The disturbing potential V_{dp} at the general point P in earth space exterior to the physical surface, is given by

$$V_{dp} = W_p - U_p \qquad \ldots(7.4),$$

the subscript $_p$ referring to evaluation *at P*. W_p is given by

$$W_p = k\iiint \frac{dm}{r}$$

where r is the distance of the element of mass dm at Q in figure 7.1 from P. If R is the distance of $P(X_i)$ from the centre of mass C of the earth, which is chosen as the origin of the Cartesian co-ordinate system X_i, this location being purely a matter of convenience,

$$X_i = R\,\ell_i$$
$$X_{ci} = R_c\,\ell_{ci} \qquad \ldots(7.5),$$

where X_{ci} are the co-ordinates of Q and R_c the distance of Q from C. If ρ is the density of matter within the volume dV,

$$dm = \rho\, dV.$$

The quantities ℓ_i, ℓ_{ci} in equation 7.5 are the direction cosines of the lines concerned. W_p can then be expressed without approximation using equations 3.26 as

$$W_p = k \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \iiint r_c^n \, p_{no}(\cos\psi)\, dV \quad \ldots(7.6),$$

where $Q\hat{C}P = \psi$.

The geopotential W_p can also be expressed by the spherical harmonic series

$$W_p = \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \sum_{m=0}^{n} p_{nm}(\sin\phi)[\, a_{nm}\cos m\lambda + b_{nm}\sin m\lambda] \quad \ldots(7.7)$$

as discussed in section 7.1. If equations 7.6 and 7.7 are identical, the contributions of low degree harmonics for the same value of n must be equal. The following results are obtained on equating the appropriate sets of terms in the two series.

When $n = 0$

$p_{00}(\cos\psi) = 1$. Hence

$$a_{00} = k \iiint \rho\, dV = kM \quad \ldots(7.8),$$

where M is the mass of the earth and called the inertia tensor of zero order.

When $n = 1$

$$p_{10}(\cos\psi) = \cos\psi = \sum_{i=1}^{3} \ell_i\, \ell_{ci}.$$

The appropriate term from equation 7.6 is

$$k \iiint \rho\, R_c \cos\psi\, dV = k \sum_{i=1}^{3} \ell_i \iiint X_{ci}\, \rho\, dV$$

$$= k \sum_{j=1}^{3} \ell_i\, M\, \overline{X}_i = 0 \quad \ldots(7.9),$$

$\overline{X}_i$ being the co-ordinates of the centre of mass of the earth. The last equality in 7.9 holds as the origin of the co-ordinate system has been defined as coinciding with the earth's centre of mass. If $\overline{X}_i \neq 0$, the origin will no longer be situated at the centre of mass of the earth. The appropriate terms from equation 7.7 are obtained from table 3.1 as

$$a_{10}\sin\phi + a_{11}\cos\phi\cos\lambda + b_{11}\cos\phi\sin\lambda$$

$$= a_{10}\,\ell_3 + a_{11}\,\ell_1 + b_{11}\,\ell_2 .$$

Thus

$$\begin{aligned} a_{10} &= I_{13} = k\iiint X_{c3}\,\rho\,dV \\ a_{11} &= I_{11} = k\iiint X_{c1}\,\rho\,dV \\ b_{11} &= I_{12} = k\iiint X_{c2}\,\rho\,dV \end{aligned} \qquad \ldots(7.10).$$

I_{1i} are the first order inertia tensors of the earth (e.g., Hotine 1969,p.156). The computation of the disturbing potential V_d using an ellipsoid containing the same mass as the earth, symmetrically distributed with respect to both the rotation axis and the equatorial plane, and the same first order inertia tensors as the earth would result in the elimination of both zero and first degree harmonics in V_d.

When $n = 2$.

The kernel of the second degree term in equation 7.6 is

$$R_c^2\,P_{20}(\cos\psi) = R_c^2\left(\frac{3}{2}\cos^2\psi - \frac{1}{2}\right) = \tfrac{1}{2}R_c^2\left(3\Big(\sum_{i=1}^{3}\ell_i\ell_{ci}\Big)^2 - 1\right)$$

$$= \tfrac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\ell_i\ell_j[3X_{ci}X_{cj} - \delta_{ij}R_c^2] ,$$

where δ_{ij} is the Kronecker delta. As $\ell_i\ell_j$ is independent of the volume integration, the resulting second degree term from equation 7.6 can therefore be written as

$$\tfrac{1}{2}k\sum_{i=1}^{3}\sum_{j=1}^{3}\ell_i\ell_j\iiint[3X_{ci}X_{cj} - \delta_{ij}R_c^2]\,\rho\,dV$$

$$= -\frac{3}{2}k\sum_{i=1}^{3}\sum_{j=1}^{3}\ell_i\ell_j\iiint[\delta_{ij}R_c^2 - X_{ci}X_{cj}]\rho dV + k\sum_{i=1}^{3}\sum_{j=1}^{3}\ell_i\ell_j\iiint\delta_{ij}R_c^2\,\rho dV$$

$$\ldots(7.11).$$

The second integral on the right is the scalar form of the second order inertia tensor I_2, given by (e.g., ibid, p.165)

$$I_2 = \iiint \sum_{i=1}^{3} X_{ci} \quad \rho \, dV = \tfrac{1}{2} \sum_{i=1}^{3} I_{2ii} \qquad \ldots(7.12),$$

where I_{2ii} are the moments of inertia of the earth about the X_i axes, given by

$$I_{2ii} = \iiint (X^2_{i+1} + X^2_{i+2}) \, \rho \, dV \quad , \text{ ss } \ldots(7.13).$$

The following equation is obtained on equating second degree terms between equations 7.6 and 7.7 on expressing equation 7.11 in matrix notation.

$$\sum_{m=0}^{2} P_{2m}(\sin\phi)[a_{2m}\cos m\lambda + b_{2m}\sin m\lambda] = -\tfrac{3}{2}k\,\{\mathbb{L}^T \mathbb{I} \mathbb{L}\} + kI_2,$$

where

$$\mathbb{I} = \begin{vmatrix} \iiint [R_c^2 - X_{c1}^2]\rho \, dV & -\iiint X_{c1}X_{c2}\,\rho dV & -\iiint X_{c1}X_{c3}\,\rho dV \\ -\iiint X_{c2}X_{c1}\,\rho \, dV & \iiint [R_c^2 - X_{c2}^2]\,\rho \, dV & -\iiint X_{c2}X_{c3}\,\rho \, dV \\ -\iiint X_{c3}X_{c1}\,\rho \, dV & -\iiint X_{c3}X_{c2}\,\rho \, dV & \iiint [R_c^2 - X_{c3}^2]\rho dV \end{vmatrix}$$

$$\ldots(7.14)$$

and

$$\mathbb{L} = \begin{vmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \end{vmatrix} = \begin{vmatrix} \cos\phi \cos\lambda \\ \cos\phi \sin\lambda \\ \sin\phi \end{vmatrix} \qquad \ldots(7.15),$$

the elements of being the components of the unit vector in the direction CP. $\mathbb{L}^T\mathbb{I}\mathbb{L}$ is the moment of inertia about the axis CP. If the non-diagonal elements of the array $\mathbb{I}$ were zero,

$$\mathbb{I} = I_2$$

as the elements along the diagonal are the moments of inertia I_{2ii} defined in equation 7.12. As the array $\mathbb{I}$ is symmetrical,

$$\mathbb{L}^T\mathbb{I}\mathbb{L} = I_{211}\cos^2\phi\cos^2\lambda + I_{222}\cos^2\phi\sin^2\lambda + I_{233}\sin^2\phi -$$

$$2I_{212}\cos^2\phi\sin\lambda\cos\lambda - 2I_{213}\sin\phi\cos\phi\cos\lambda -$$

$$2I_{223}\sin\phi\cos\phi\cos\lambda\ ,$$

where

$$I_{2ij} = \iiint X_{ci} X_{cj} \rho \, dV, \qquad i \neq j.$$

Thus

$$\sum_{m=0}^{2} P_{2m}(\sin\phi)[a_{2m}\cos m\lambda + b_{2m}\sin m\lambda] = -k\Big(I_{211}(\tfrac{3}{2}\cos^2\phi\cos^2\lambda - \tfrac{1}{2}) +$$

$$I_{222}(\tfrac{3}{2}\cos^2\phi\sin^2\lambda - \tfrac{1}{2}) + I_{233}(\tfrac{3}{2}\sin^2\phi - \tfrac{1}{2}) - 3I_{212}\cos^2\phi\sin\lambda\cos\lambda -$$

$$3I_{213}\sin\phi\cos\phi\cos\lambda - 3I_{223}\sin\phi\cos\phi\sin\lambda\Big)$$

$$= -k\Big(I_{211}\{\tfrac{3}{4}(1-\sin^2\phi)(1+\cos 2\lambda) - \tfrac{1}{2}\} + I_{222}\{\tfrac{3}{4}(1-\sin^2\phi)(1-\cos 2\lambda) - \tfrac{1}{2}\} +$$

$$I_{233}\,P_{20}(\sin\phi) - 2P_{22}(\sin\phi)\sin\lambda\cos\lambda - 2P_{21}(\sin\phi)\big(I_{213}\cos\lambda + I_{223}\sin\lambda\big)\Big)$$

$$= k\Big(P_{20}(\sin\phi)[\tfrac{1}{2}(I_{211}+I_{222}) - I_{233}] + P_{21}(\sin\phi)[2I_{213}\cos\lambda +$$

$$2I_{223}\sin\lambda] + P_{22}(\sin\phi)[I_{212}\sin 2\lambda + \tfrac{1}{2}(I_{222}-I_{211})\cos 2\lambda]\Big),$$

where the values of $P_{21}(\sin\phi)$ and $P_{22}(\sin\phi)$ are obtained from table 3.1 and equation 3.38.

Thus,

$$\left.\begin{aligned} a_{20} &= k\,(\tfrac{1}{2}\{I_{211} + I_{222}\} - I_{233}) & & \\ a_{21} &= 2k\,I_{213} & b_{21} &= 2k\,I_{223} \\ a_{22} &= \tfrac{1}{2}k\{I_{222} - I_{211}\} & b_{22} &= k\,I_{212} \end{aligned}\right\} \quad ..(7.15)$$

A study of equations 7.13 and 7.15 shows that $\underline{L}^T I \underline{L}$ can, in effect, be diagonalised if

$$\phi = \tfrac{1}{2}\pi$$

when the direction CP coincides with the X_3 axis. Such an axis is called a *principal axis of inertia*. It can be shown (e.g., Hotine 1969, p.166) that products of inertia, using any pair of mutually perpendicular axes which are also orthogonal to a principal axis, are zero. Thus

$$I_{213} = I_{223} = 0 \qquad \ldots(7.17)$$

if the X_3 axis is a principal axis of inertia.

Notes

(i) Any spherical harmonic series used for the representation of the disturbing potential will have the coefficients a_{00}, a_{10}, a_{11}, b_{11}, a_{21}, b_{21} equal to zero.

$$a_{00} = 0$$

if the mass of the reference system, assumed to be an ellipsoid of revolution, equals that of the earth.

$$a_{10} = a_{11} = b_{11} = 0$$

if the centre of mass of the reference system is coincident with that of the earth.

$$a_{21} = b_{21} = 0$$

if the minor axis of the reference ellipsoid coincides with the axis of rotation of the earth.

(ii) If the gravitating system defined by equations 7.6 and 7.7 was a rotation ellipsoid with a symmetrical mass distribution,

$$W_p = \sum_{n=0}^{\infty} \frac{1}{r^{2n+1}} P_{(2n)0} C_{(2n)0} ,$$

when the use of equations 7.8, 7.11 and 7.12 gives

$$W_p = \frac{kM}{r} + \frac{k}{2r^3}\left(I_{211} + I_{222} + I_{233} - 3\{L^T I L\}\right) + o\{\tfrac{1}{r^4}\} \qquad \ldots(7.18).$$

This equation is known as *MacCullagh's formula*.

(iii) The coefficients a_{nm}, b_{nm} in the case where is the potential due to a reference sphere, are related to the coefficients C_{nm}, S_{nm} in equation 7.2 by the expressions

$$C_{nm} = \frac{a_{nm}}{kM\, a_e^n} .$$

with a similar relation for S_{nm}. The relation of greatest importance is that obtained when $(n = 2, m = 0)$ and $(n = 2, m = 2)$. The use of equation 7.15 gives

$$C_{20} = \frac{\tfrac{1}{2}[I_{211} + I_{222}] - I_{233}}{M a^2} \qquad \ldots(7.19)$$

and

$$C_{22} = \frac{\tfrac{1}{2}[I_{222} - I_{211}]}{M a^2} \qquad \ldots.(7.20).$$

If the earth were an ellipsoid of revolution with a symmetrical mass distribution, $I_{211} = I_{222}$ and

$$C_{20} = \frac{I_{211} - I_{233}}{M a^2} \overset{(6.20)}{=} \tfrac{1}{3}m' - \tfrac{2}{3}f - \tfrac{3}{7}m'f + \tfrac{1}{3}f^2 \qquad \ldots(7.21)$$

and $C_{22} = 0$.

Hence C_{22} is a measure of any ellipticity of the earth's equator. A study of table 8.3 shows that the magnitude of C_{22}, while large in comparison to most of the other coefficients, nevertheless is of the same order of magnitude and no significant improvement is achieved by introducing the concept of a tri-axial ellipsoid as the reference model.

7.3 The combination of data obtained from satellite orbits and surface gravimetry

The earth's gravitational field can be completely defined by gravity measurements at the surface of the earth. This is not a practical possibility at the present time (1970). The gravity data available at present consists of measurements at discrete points which are irregularly spaced on the surface of the earth. The major concentration of readings occur in the continental regions of the earth while the oceans are inadequately surveyed. Thus no gravimetric solutions for the geoid will be possible unless some method is available for the prediction of values of gravity for the representation of unsurveyed areas.

The simplest technique is to use the low degree harmonics of the earth's gravitational field as obtained from the analysis of the orbital perturbations of artificial earth satellites for this purpose. This technique was initially used by Kaula (1966a). All available surface gravity data is assembled in the form of area means for some standard block (e.g., $5^{o}\times 5^{o}$ or 300 n.mi. squares). It is preferable to sub-divide each of these basic squares into a smaller unit, attempt to represent each of these by a single reading, and then compute the area mean for the basic square.

The gravity anomaly at the surface of the earth Δg which is commonly called the free air anomaly, is given by

$$\Delta g = g - \gamma_o + \Delta\gamma \qquad \ldots(7.22),$$

where g is the value of observed gravity, γ_o is the value of normal gravity on the reference ellipsoid, given by equation 6.38, and $\Delta\gamma$ the free air reduction given by equation 6.76. Let the values of the coefficients $(\overline{C}_{nm}, \overline{S}_{nm})$ of the series representing the geopotential, given by equation 4.40, as obtained by satellite orbital analysis be $(\overline{C}_{nm_s}, \overline{S}_{nm_s})$, where

$$\overline{C}_{nm} = \overline{C}_{nm_s} + dC_{nm} \; ; \quad \overline{S}_{nm} = \overline{S}_{nm_s} + dS_{nm} \qquad \ldots(7.23).$$

The disturbing potential can be expressed by an equation of the form

$$V_d \overset{(4.50)}{=} \frac{kM}{r} \sum_{n=2}^{\infty} \left(\frac{a_e}{r}\right)^n \sum_{m=0}^{n} \bar{P}_{nm}(\sin\phi)[\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda] + o\{fV_d\} \quad \ldots(7.24)$$

at the surface of the earth to the order of the flattening. The disturbing potential V_{dp} at a point P, having a difference in geopotential W with reference to the geoid ($W=W_o$), is given by

$$V_{dp} = W_p - U_p \quad \ldots(7.25)$$

where W_p is the potential on the reference system at P, the geoid and the spheroid being assumed to have the same potential. Let Q in figure 7.2 be a point on the spherop $U = W_p$ and on the vertical through P. Let γ_q be the value of normal gravity at Q. If $PQ=h_d$, the use of equations 6.62 and 6.63 gives

$$U_p = W_p + h_d(-\gamma_q)$$

or

$$V_{dp} = \gamma_q h_d + o\{fV_d\} \quad \ldots(7.26).$$

The differentiation of equation 7.25 with respect to elevation h gives

$$\left(\frac{\partial V_d}{\partial h}\right) = \left(\frac{\partial W}{\partial h}\right)_p - \left(\frac{\partial U}{\partial h}\right)_p = -g + \left(\gamma_q + h_d \frac{\partial \gamma}{\partial h}\right) .$$

The use of equations 6.69 and 7.26 gives

$$\frac{\partial V_d}{\partial h} = -\Delta g - 2\frac{V_d}{R} + o\{f \frac{\partial V_d}{\partial h}\} \quad \ldots(7.27),$$

where R, to the required order of accuracy, is the mean radius of the earth. The differentiation of equation 7.24 with respect to h(i.e., r) gives

$$\frac{\partial V_d}{\partial h} = \frac{kM}{r^2} \sum_{n=2}^{\infty} -(n+1) \left(\frac{a_e}{r}\right)^n \sum_{m=0}^{n} \bar{P}_{nm}(\sin\phi)[\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda] + o\{f\frac{\partial V_d}{\partial h}\} \quad ..(7.28).$$

The use of this result in equation 7.27 along with equation 7.24 gives

$$\Delta g = \gamma \sum_{n=2}^{\infty} (n-1) \left(\frac{a_e}{r}\right)^n \sum_{m=0}^{n} \bar{P}_{nm}(\sin\phi)[\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda] + o\{f\ \Delta g\} \quad \ldots(7.29),$$

where γ is the mean value of gravity over the earth. Equation 7.29 applies at the physical surface of the earth to the order of the flattening. As this surface is a sphere to this same order,

$$\left(\frac{a_e}{r}\right)^n = 1 + o\{f\} \quad .$$

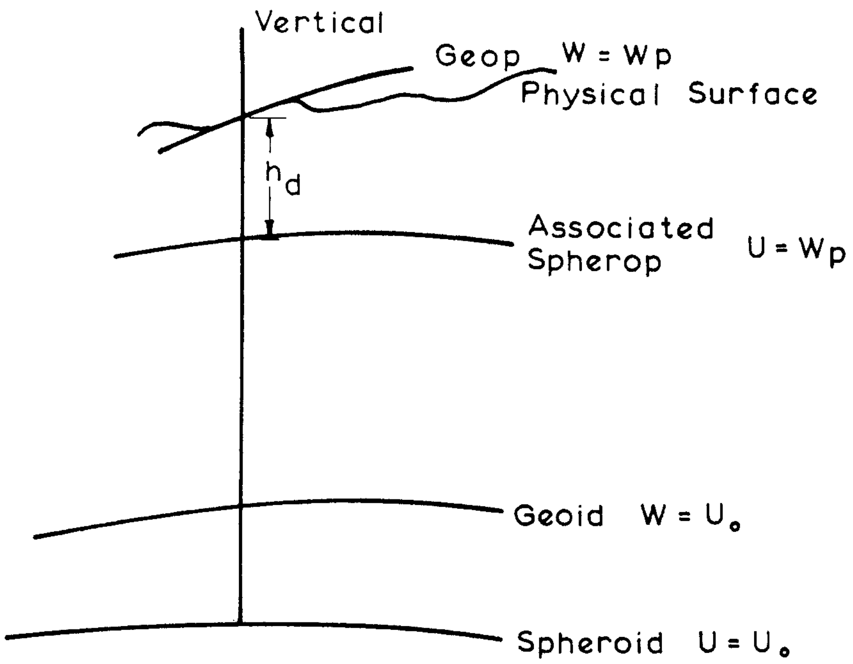


FIG. 7.2
The height anomaly on a perfect reference system

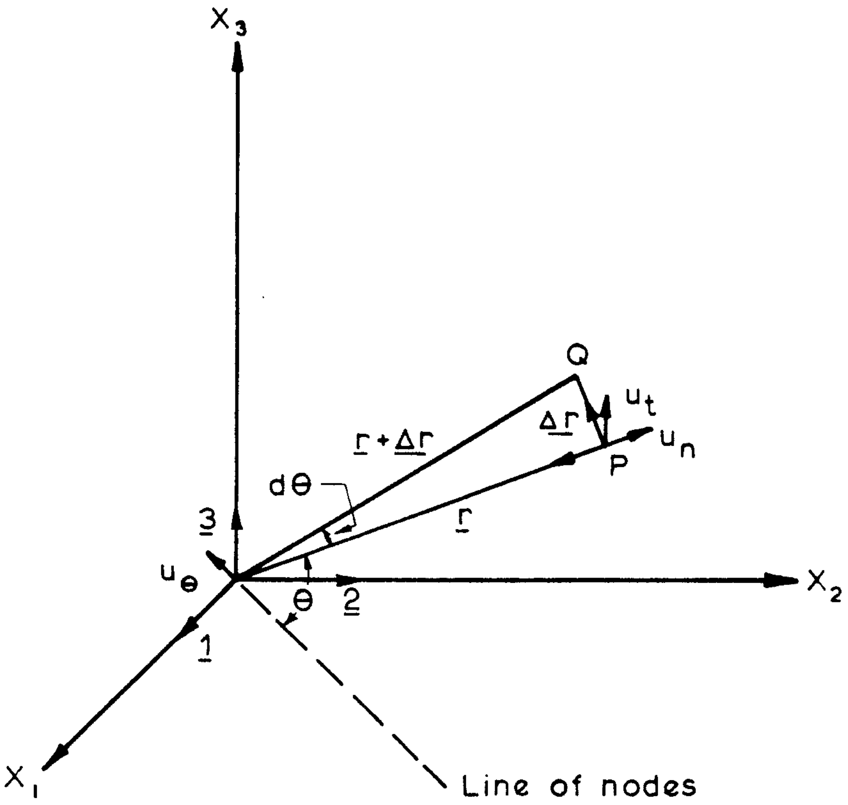


FIG. 8.1
Unit vectors in the orbital plane

Observation equations can therefore be set up for every basic unit of surface area where observed area means are available. These take the form

$$\Delta g_o + v = \gamma \sum_{n=2}^{\infty} (n-1) \sum_{m=0}^{n} \bar{P}_{nm}(\sin \phi)[(\bar{C}_{nm_S}+dC_{nm})\cos m\lambda + (\bar{S}_{nm_S}+dS_{nm})\sin m\lambda] \quad \ldots(7.30).$$

This relation can also be written as

$$v_i = K_i + \gamma \sum_{n=2}^{\infty} (n-1) \sum_{m=0}^{n} \bar{P}_{nm}(\sin \phi_i)[dC_{nm}\cos m\lambda_i + dS_{nm}\sin m\lambda_i] \quad \ldots(7.31),$$

where the terms K_i are given by

$$K_i = \Delta g_{oi} - \gamma \sum_{n=2}^{\infty} (n-1) \sum_{m=0}^{n} \bar{P}_{nm}(\sin \phi_i)[\bar{C}_{nm_S}\cos m\lambda_i + \bar{S}_{nm_S}\sin m\lambda_i] \quad \ldots(7.32),$$

the index $_i$ referring to the particular element of surface area on the spherical model of the earth. While table 4.1 indicates a considerable damping effect of elevation on the coefficients of higher degree, there is no limit to the maximum value of n to which the analysis is carried out except limits imposed by the computer storage available. As it is generally held that the smallest unit of surface area which does not give rise to area means which are correlated with position, is the $30^o \times 30^0$ square (Hirvonen 1956), it follows that the minimum possible analysis which is valid is one which included all terms up to and including degree 6. Practical analyses have been carried out using the above techniques to (14,14) (Kaula 1966a; Rapp 1968). Such an analysis would include 219 coefficients, if the 6 inadmissible coefficients described in section 7.2 are excluded. The total number of coefficients q is given by

$$q = (n_m + 1)^2 - 6 \quad \ldots(7.33),$$

where n_m is the highest degree to which the analysis is carried out.

The quantities v_i defined in equation 7.31 would be normally distributed if no systematic errors existed in the values of Δg_{oi}. This is a not unreasonable assumption under the circumstances. In the case of squares with no observed values of gravity, v_i can be estimated as having a magnitude approaching ±25 mgal, as can be seen from a study of table 9.2. Such squares should have as little influence as possible on any solutions for the values of (dC_{nm}, dS_{nm}). This effect is achieved by the introduction of appropriate weight coefficients and allowing for any systematic effects in the values of Δg_{oi} in order that the v_i's take values which are as small as possible, prior to solution.

Equations 7.30 to 7.32 form a set of observation equations of the form

$$V = A X - K \qquad \ldots(7.34)$$

where

$$V = \begin{vmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{vmatrix} ; \; K = \begin{vmatrix} K_1 \\ K_2 \\ \vdots \\ K_p \end{vmatrix} ; \; A = \begin{vmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1q} \\ a_{21} & a_{22} & \cdots\cdots & a_{2q} \\ \vdots & & & \\ a_{p1} & a_{p2} & \cdots\cdots & a_{pq} \end{vmatrix} ; \; X = \begin{vmatrix} dC_{20} \\ dS_{20} \\ dC_{22} \\ \vdots \\ dS_{n_m n_m} \end{vmatrix}$$

where X is a (p,1) array, p being given by equation 7.33. Equation 7.34 is a standard block of observation equations which is solved by minimising the sum of the weighted squares of the residuals v_i. This condition can be expressed in matrix notation by the equation

$$\Phi = V^T W_v V = \text{minimum} \qquad \ldots(7.35)$$

where W_v is the diagonal matrix of weight coefficients of the elements forming the array V and given by

$$W_v = \begin{vmatrix} w_{11} & 0 & 0 & 0 & \cdots\cdots & 0 \\ 0 & w_{22} & 0 & 0 & \cdots\cdots & 0 \\ 0 & 0 & w_{33} & 0 & \cdots\cdots & 0 \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ 0 & 0 & 0 & 0 & \cdots\cdots & w_{pp} \end{vmatrix} \qquad \ldots(7.36)$$

w_{ii} is the function of the number of readings in the square considered. It is also a function of the variability of readings within a square for limited extents (Mather 1967,p.135). This suggests a weight coefficient of the type

$$w_{ii} = c\, n_i \qquad \ldots(7.37),$$

where c is a constant and n_i is the number of readings available within the square. Kaula (1966a,p.5312) suggests an expression of the form

$$w_{ii} = \frac{n_i + 1}{M\{\Delta g^2\}} ,$$

where $M\{\Delta g^2\}$ is the mean value of Δg^2 over the earth. An expression of the same form was given by Mather(1967,p.135). Rapp (1968,p.5) uses

$$w_{ii} = \frac{\cos\phi}{m}$$

where the factor cos ϕ is used as $5^{o} \times 5^{o}$ squares were used in the analysis with attendant variation in the actual surface area with latitude. m^2 is the square of the estimate of error in the value adopted for Δg_i.

Values are obtained for the V's by differentiating the equations at 7.35 with respect to the quantities defined in the array X. The resulting equations can be expressed in matrix form by the relation

$$A^T W_v A X - A^T W_v K = 0,$$

where A^T is the transpose of A,
the solution of which is

$$X = (A^T W_v A)^{-1} A^T W K \qquad \ldots(7.38).$$

The use of these results in equation 7.34 gives the corrections to the adopted means Δg_i. An additional by-product of this solution is the set of corrections dS_{nm}, dS_{nm} to the coefficients of the spherical harmonic series. Hence the technique defines both a complete set of gravity anomalies at the surface of the earth as well as a spherical harmonic series for the representation of gravity anomalies *at the surface of the earth*. Such a representation is obviously valid only to the order of the flattening, in view of the departures of the earth's surface from a sphere when equation 7.29 will no longer hold. Such a restriction is not an embarassment of significance as it is unlikely that this order of accuracy can ever be exceeded if prediction forms part of the necessary processes in physical geodesy.

An alternate method of solution due to Kaula is as follows. The limiting case of equation 7.29 when $r \to a_e$ can be written as

$$g = \sum_{n=2}^{\infty} g_n = \gamma \sum_{n=2}^{\infty} (n-1) \sum_{m=0}^{n} \bar{P}_{nm}(\sin\phi)[\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda] \ldots(7.39),$$

where

$$g_n = \sum_{m=0}^{n} \bar{P}_{nm}(\sin\phi)[\bar{g}_{1nm}\cos m\lambda + \bar{g}_{2nm}\sin m\lambda] \ldots(7.40).$$

Thus

$$g_{1nm}/g_{2nm} = (n-1)\,\bar{C}_{nm}/\bar{S}_{nm} \qquad \ldots(7.41).$$

Any single coefficient in the above series when multiplied by its allied harmonic functions, satisfies surface integrals of the type

$$\iint \bar{g}_{1nm}\,\bar{P}_{nm}(\sin\phi)\cos m\lambda\,\bar{P}_{nm}(\sin\phi)\cos m\lambda\; d\sigma \overset{(4.28)}{=} 4\pi\bar{g}_{1nm} \ldots(7.42)$$

on using equation 4.29 and the definition of fully normalised harmonics. The use of equations 7.41, 7.42 and the orthogonal property of surface harmonics, given in equation 4.18, therefore provides the relation

$$\iint \sum_{n=2}^{\infty} g_{1nm} \; \overline{P}_{nm}(\sin\phi)\; \cos m\lambda \;\; d\sigma = 4\pi\, \overline{C}_{nm}\, \gamma(n-1)$$

or

$$\Delta g\; \overline{P}_{nm}(\sin\phi)\cos m\lambda \;\; d\sigma \;=\; 4\pi\, \overline{C}_{nm}\, \gamma(n-1)\,.$$

Thus the available surface gravity anomalies in the form of area means Δg_{oi} can be combined with harmonic coefficients $\{C_{nm_s},\, S_{nm_s}\}$ obtained from orbital perturbations alone using observation equations of the form

$$4\pi\gamma(n-1)\left(\,\overline{C}_{nm_s} + dC_{nm}\right) = \sum_i [\Delta g_{oi} + v_i]\overline{P}_{nm}(\sin\phi)\;\cos m\lambda\; d\sigma \quad ..(7.43).$$

The resulting observation equations can be represented in matrix notation by the equation

$$A_v V + A_x X - K_k = 0 \qquad \ldots(7.44)$$

whose solution can be effected in one of two ways. If it is assumed that the corrections $\{dC_{nm}, dS_{nm}\}$ are normally distributed, a least squares determination could be obtained from the relation

$$\Phi = V^T W_v V + X^T W_x X - L^T(A_v V + A_x X - K_k) = \text{minimum} \qquad \ldots(7.45),$$

L being a column array of Lagrangian multipliers and W_x is the weight matrix for the correction to the coefficients.

If such a premise is invalid, equation (7.43) can be solved using the technique outlined in equations 7.34 to 7.38. In the case under consideration, equation 7.44 is solved in the usual manner by minimising Φ with respect to the elements of the V and X arrays. The resulting equations can be conveniently expressed as two separate sets of equations. The corrections to the area means are given by the relation

$$W_v V - A_v^T L = 0 \qquad \ldots(7.46).$$

Similarly, the differentiation of Φ with respect to the corrections to the harmonic coefficients gives

$$W_x X - A_x^T L = 0 \qquad \ldots(7.47).$$

The substitution of these results in equation 7.44 gives

$$A_v (W_v^{-1} A_v^T L) + A_x (W_x^{-1} A_x^T L) - K_k = 0$$

or

$$L = \left(A_v W_v^{-1} A_v^T + A_x W_x^{-1} A_x^T \right)^{-1} K_k \qquad \ldots(7.48).$$

The use of the solution at 7.48 gives, in turn, the elements in the arrays V and X on substitution in equations 7.46 and 7.47. The adoption of this technique is complicated by the fact that harmonic coefficients determined from satellite orbital analysis are all based on truncated versions of equation 7.2. It may therefore be preferable to use the elements of the array without attributing properties to it based on the individual elements being representative of normally distributed populations. and free from correlation effects. Such an assumption has greater validity in the case of observed area means whose departures from the true value are a function of the available sample and not of position. This compounded by the fact that most poorly surveyed regions are oceanic.

7.4 The problems associated with the determination of low degree coefficients from orbital analysis.

A near earth satellite can be considered to be a particle moving primarily under the influence of the earth's gravitational field. Such a particle will be shown in the next section to describe an ellipse if the only force acting on it were the earth's mass, assumed to be concentrated at a point. Such gravitational characteristics are also exhibited by a sphere containing a symmetrical mass distribution. This condition is not satisfied in practice due to the low altitudes of satellites above the surface of the earth when the equatorial bulge causes not only the precession of the orbital plane but also changes in the position of the orbit with respect to the earth which can be related to the earth's gravitational field if the satellite's minimum height is great enough to avoid atmospheric drag. Such a statement assumes that luni-solar attraction, relativistic and earth tide effects, electromagnetic effects and radiation pressure are of negligible magnitude. These topics are discussed at length by Kaula (1962). His conclusions can be summarised as follows.

1. *Luni-solar attraction*

The luni-solar attraction is generally very small though it can be allowed for by the adoption of a model similar to that used for the representation of the earth's gravitational field.

2. *Relativistic effects*

These arise as a consequence of the departure of gravitation from the Newtonian model. Its effect is largest on the secular motion of perigee (i.e., the position in the orbit which is closest to the earth). The motion itself is of the order of 10^{-5} to 10^{-6}

times that due to the principal term C_{20} and is therefore approximately three orders of magnitude smaller than of the other terms C_{nm}, S_{nm}. The effect is fortunately less for satellites of geodetic interest (i.e., $200 < h^{(km)} < 1000$), as its magnitude increases with elevation h above the earth.

3. Tidal effects

The gravitational effects of earth tides are of the order of 10^{-1} mgal(Heiskanen & Vening Meinesz 1958,p.119). This figure is one hundred times smaller than the average gravity anomaly at the surface of the earth and hence should have effects of similar comparative magnitude on the orbits of near earth satellites. This figure is in agreement with that estimated by Kaula who puts the effect of earth tides as being an order smaller than the luni-solar effect itself (Kaula 1962,p.230). He also estimates the effect of oceanic tides at f^3 or almost three orders of magnitude smaller than the general coefficients C_{nm}, S_{nm}.

4. Electromagnetic effects and radiation pressure

These are a consequence of the bombardment of the satellites by electrons causing a transfer of charge and momentum to the satellite. The circumstances are listed by Kaula as being significant at higher altitudes when the satellite leaves the protective atmosphere of the earth (ibid,p.235). These effects are not noticeable at the elevations of satellites of geodetic interest.

Radiation pressure effects are of significance as most observations are made in the earth's shadow. It is however reported as being too small to be identified in the case of normal satellite orbits.

5. Drag

This is a consequence of the earth's atmosphere and is of significance at low altitudes when friction lessens apogee heights, causing the orbit to decay(i.e., the equatorial radius of the orbital ellipse decreases with time). Air drag cannot be satisfactorily allowed for due to departures of the actual density of the atmosphere from any model that could be adopted (Kaula 1963,p.533). Current practice calls for the definition of some parameters expressing the major perturbations due to drag from an analysis of the orbital data itself. These are then used to calculate other second order effects. This procedure enables satellites which come as close as 200 km from the earth's surface to be used in analyses of the earth's gravitational field (Kaula 1966b,p.56).

It is in this context that attempts are made to evaluate the low degree harmonics of the earth's gravitational field from the orbital perturbations of near earth satellites. The section which follows develops the theory for the determination of the coefficients

C_{nm}, S_{nm} from the tracking of satellites. As these perturbations are functions of the parameters defining the orbit, it is necessary to

(a) establish the principles of central force motion;

(b) choose an appropriate set of parameters for the definition of the instantaneous elliptical orbit;

(c) define a set of relations governing the rate of change of these parameters, called orbital elements, in relation to the disturbing potential;

(d) relate equation 7.24 to the orbital elements.

Most of these problems have been considered in celestial mechanics when developing the theory of the motion of satellites (moons) around planets. The major portion of the theoretical development with special relevance to near earth satellites is adapted from a treatment by Kaula (1961; 1966b).

8. THE USE OF THE ORBITAL PERTURBATIONS OF NEAR EARTH SATELLITES TO DEFINE THE LOW DEGREE HARMONICS OF THE GRAVITY FIELD

8.1 The motion of a near earth satellite

i. Central force motion

Consider a particle moving the the gravitational field of a massive body (mass M) in inertial space which does not rotate with the body and is related to a three dimensional Cartesian co-ordinate system X_i. If the velocity vector v of the particle is given by

$$v = \sum_{i=1}^{3} \frac{dX_i}{dt} i = \sum_{i=1}^{3} \dot{X}_i i \qquad \ldots(8.1),$$

where i has the same definition as in section 2, being unit vectors along the X_i axes, its acceleration vector is given by

$$\dot{v} = \frac{d}{dt} v = \sum_{i=1}^{3} \ddot{X}_i i \qquad \ldots(8.2).$$

$\dot{v}$ can also be related to the unit vectors u_t, u_n which are tangential and normal to the instantaneous motion of the satellite in its orbital plane, as shown in figure 8.1. Such a plane contains both the centre of mass of the body and the position P of the satellite. In this case,

$$\dot{v} = \dot{v}\, u_t + \frac{v^2}{r} u_n \qquad \ldots(8.3),$$

where

$$\dot{\mathrm{v}} = \left(\sum_{i=1}^{3} \ddot{X}_i^2 \right)^{1/2}$$

and

$$\mathrm{v}^2 = \sum_{i=1}^{3} \dot{X}_i^2 \qquad \ldots(8.4),$$

r being the distance CP given by

$$r = \left(\sum_{i=1}^{3} X_i^2 \right)^{1/2} \qquad \ldots(8.5)$$

and the position vector r of the satellite being given by

$$\mathit{r} = \sum_{i=1}^{3} X_i \, \mathit{i} \qquad \ldots(8.6).$$

Th orbit is governed by Newton's second law of motion which, in the case of a gravitational field of potential U and with no other impressed forces, is given by

$$\mathit{F} = - \sum_{i=1}^{3} \frac{\partial U}{\partial X_i} \mathit{i} = \dot{\mathit{v}} \qquad \ldots(8.7)$$

if the particle is assumed to have unit mass. F in equation 8.7 is the gravitational attraction. This motion can also be defined in terms of u_r and u_θ which are unit position and normal vectors in the orbital plane as shown in figure 8.1. By the definition of central force motion,

$$\mathit{F} = \mathrm{F} \; u_r \qquad \ldots(8.8)$$

The vector product

$$\mathit{r} \times \dot{\mathit{v}} \overset{(8.7),(8.8)}{=} \mathit{r} \times \mathrm{F} \, u_r = 0 \qquad \ldots(8.9),$$

the second equality holding as it is the product of co-planar vectors(e.g., see Jeffreys & Jeffreys 1962,p.67). The velocity vector v is also related to the unit vectors u_r and u_θ by the relations

$$\mathit{v} \overset{(8.1)}{=} \dot{\mathit{r}} = \frac{d}{dt}(r u_r) = \dot{r} \, u_r + \frac{d}{dt} u_r = \dot{r} u_r + r\dot{\theta} \, u_\theta \qquad \ldots(8.10).$$

The element of vector area $\Delta \mathit{A}$ is given from basic vector concepts as

$$\mathit{A} = \tfrac{1}{2}\mathit{r} \times [\mathit{r} + \Delta \mathit{r}] = \tfrac{1}{2}\mathit{r} \times \Delta \mathit{r} \qquad \ldots(8.11).$$

Differentiation with respect to time t gives

$$\frac{\Delta \mathit{A}}{\Delta t} = \tfrac{1}{2}\mathit{r} \times \frac{\Delta \mathit{r}}{\Delta t} \underset{\Delta t \to 0}{\overset{(8.10)}{=}} \tfrac{1}{2}\mathit{r} \times \mathit{v} = \tfrac{1}{2} r u_r \times (\dot{r} \, u_r + r\dot{\theta} \; u_\theta).$$

Thus

$$\lim_{\Delta t \to 0} \frac{\Delta \mathit{A}}{\Delta t} = \tfrac{1}{2} r^2 \, \dot{\theta} u_r \times u_\theta = \tfrac{1}{2} r^2 \dot{\theta} \; u_A \qquad \ldots(8.12),$$

where u_A is the unit vector normal to the instantaneous orbital plane. The quantity $(\Delta A/\Delta t)$ is the rate of change of area in the orbital ellipse which, on further differentiation with respect to t gives

$$\frac{d}{dt}\dot{A} \overset{(8.12)}{=} \tfrac{1}{2}\frac{d}{dt}(r \times v) = \tfrac{1}{2}[\dot{r}\times v + r\times\dot{v}] \overset{(8.9)}{=} \tfrac{1}{2}\dot{r} \times v \overset{(8.10)}{=} 0 \qquad \ldots(8.13)$$

in the case of central force motion as $\dot{r}$ and v are coplanar vectors. Thus

$$\tfrac{1}{2}r^2\dot{\theta} = \dot{A} = \tfrac{1}{2}h \text{ (say)} \qquad \ldots(8.14),$$

as u_A is the unit vector perpendicular to the plane of motion and $\dot{A}$ is the scalar rate of change of orbital area with respect to time which, from equation 8.13, is a constant, equal to $\tfrac{1}{2}h$ (say).

The integration of equation 8.14 with respect to time gives

$$A = \tfrac{1}{2}h\,t + C \qquad \ldots(8.15), \qquad C$$

being the constant of integration.

Notes

(i) Equation 8.15 is well known as Kepler's second law of planetary motion. Thus the position vector of a particle moving under the influence of a central force alone, sweeps out equal areas in a given interval of time.

(ii) The constant h defined in equation 8.14 is called the *constant of areas*.

(iii) As the oribit of a satellite is an ellipse, as will be shown in the following development, it moves faster as it approaches the attracting body.

Equation 8.10 expresses the velocity vector in terms of components which are radial and normal to the position vector. Differentiation of this equation with respect to t gives

$$\dot{v} = \ddot{r}u_r + \dot{r}\frac{d}{dt}u_r + \dot{r}\dot{\theta}\,u + r\ddot{\theta}\,u_\theta + r\dot{\theta}\frac{d}{dt}u_\theta .$$

As a direct consideration of figure 8.1 shows that

$$\frac{d}{dt}u_\theta = -\dot{\theta}\,u_r,$$

$$\dot{v} \overset{(8.10)}{=} \ddot{r}\,u_r + \dot{r}\dot{\theta}\,u_\theta + \dot{r}\dot{\theta}\,u_\theta + r\ddot{\theta}\,u_\theta + r\dot{\theta}^2\,u_r$$

$$= [\ddot{r} - r\dot{\theta}^2]\,u_r + [2\dot{r}\dot{\theta} + r\ddot{\theta}]\,u_\theta \qquad \ldots(8.16).$$

The central force controlling the motion is purely gravitational, being given by

$$F = -\frac{\mu}{r^2}\,u_r \overset{(8.7)}{=} \dot{v} \qquad \ldots(8.17),$$

where

$$\mu = kM \quad .$$

It has been assumed that the central force is due to a gravitating mass or set of masses whose distribution is compatible with representation by a point mass. This would, for example, apply in the case of a perfectly symmetrical mass distribution within a sphere. The use of equation 8.17 in equation 8.16 gives the following relations.

$$\left.\begin{aligned} 2\dot{r}\dot{\theta} + r\ddot{\theta} &= 0 \\ \ddot{r} - r\dot{\theta}^2 &= -\frac{\mu}{r^2} \end{aligned}\right\} \quad \ldots(8.18).$$

The use of equation 8.14 and the introduction of a change of variable defined by

$$u = \frac{1}{r} \qquad \text{gives} \qquad \dot{\theta} = h\,u^2 \qquad \ldots(8.19).$$

In addition,

$$\dot{r} = \frac{d}{dt}\left(\frac{1}{u}\right) = -\frac{1}{u^2}\dot{u} = -\frac{1}{u^2}\frac{du}{d\theta}\dot{\theta} \overset{(8.19)}{=} -h\frac{du}{d\theta}$$

and

$$\ddot{r} = -h\frac{d^2u}{d\theta^2}\dot{\theta} = -h^2u^2\frac{d^2u}{d\theta^2} \quad .$$

Thus equation 8.18 becomes

$$-\mu u^2 = -h^2u^2\frac{d^2u}{d\theta^2} - \frac{1}{u}h^2u^4$$

which can be written as

$$\frac{d^2u}{d\theta^2} + u = \frac{\mu}{h^2} \qquad \ldots(8.19),$$

being the differential equation of central force motion. Its solution is of the form

$$u = k_1 + k_2\cos\theta$$

when

$$\frac{du}{d\theta} = -k_2\sin\theta \quad ; \qquad \frac{d^2u}{d\theta^2} = -k_2\cos\theta$$

or

$$\frac{d^2u}{d\theta^2} + u = k_1 \quad .$$

Comparison with equation 8.19 shows that

$$k_1 = \frac{\mu}{h^2} \quad .$$

The final solution is therefore

$$u = \frac{\mu}{h^2} + k_2\cos\theta$$

or

$$r = \frac{h^2/\mu}{1 + \frac{k_2h}{\mu}\cos\theta} \qquad \ldots(8.20).$$

This can be shown to be the equation of an ellipse.

Consider the general point P on an ellipse in relation to its focus C and the centre O with respect to an $x\,x_3$ axis system. If the equatorial radius is a and the angle SCA is f and called the *true anomaly* in figure 8.2, being the angle θ in the previous development, it is possible to define position on the ellipse in relation to the focus using the angle f. Af AS'P is the auxiliary circle of radius a where S'SH is perpendicular to OCA and if the angle S'OA equals E, the *eccentric anomaly*, the latter is related to f through the standard relations

$$r \cos f \overset{(1.12)}{=} a \cos E - ae \qquad \ldots(8.21)$$

$$r \sin f = x_3 \overset{(1.14)}{=} b \sin E \qquad \ldots(8.22),$$

where E can be recognised as being the parametric latitude for an ellipse and e its eccentricity, the x axis coinciding with the major axis and the x_3 axis perpendicular to it. On squaring and adding equations 8.21 and 8.22,

$$r^2 = a^2\cos^2E - 2a^2e\cos E + a^2e^2 + a^2(1-e^2)\sin^2E$$

$$= a^2[1 - 2e\cos E + e^2\cos^2E] \quad .$$

Thus

$$r = a[1 - e\cos E] \qquad \ldots(8.23).$$

The use of equation 8.21 in 8.23 gives

$$r = a - e[r\cos f + ae] = a(1-e^2) - e\,r\cos f \quad .$$

Hence

$$r = \frac{a(1-e^2)}{1 + e\cos f} \qquad \ldots(8.24).$$

The comparison of equations 8.20 and 8.24 shows that the motion of a satellite under the influence of a central force is an ellipse with the centre of mass of the gravitating body at its focus. In addition, it can be seen that

$$\frac{h^2}{\mu} = a(1-e^2) \quad \text{or} \quad h = [\mu a(1-e^2)]^{\frac{1}{2}} \qquad \ldots(8.25).$$

The ratio

$$\frac{\text{Area SHA}}{\text{Area S'HA}} = \frac{\text{Area SCP - Triangle SCH}}{\text{Area S'OA- Triangle S'OH}}$$

$$\overset{(8.15)}{=} \frac{\frac{1}{2}h\,(t-T) - \frac{1}{2}r^2\sin f\cos f}{\frac{1}{2}a^2E - \frac{1}{2}a^2\sin E\cos E} \quad ,$$

where (t-T) is the interval of time between perigee (equatorial) passage at time T and the instant considered t. Thus

$$\frac{\text{Area SHA}}{\text{Area S'HA}} \overset{(8.25)}{=} \frac{[\mu a(1-e^2)]^{\frac{1}{2}}(t-T) - r^2\sin f\cos f}{a^2[E - \sin E\cos E]} \qquad \ldots(8.26).$$

As $\quad x = a\cos E \;,$

$$\frac{\text{Area SHA}}{\text{Area S'HA}} = \frac{\int_a^x x_3 dx}{\int_a^x \sin E\, dx} = \frac{\int_0^E b \sin E\, d(a \cos E)}{\int_0^E a \sin E\, d(a \cos E)}$$

$$= \frac{\int_0^E a\, b \sin^2 E\, dE}{\int_0^E a^2 \sin^2 E\, dE} = \frac{b}{a} \overset{(1.17)}{=} (1-e^2)^{\frac{1}{2}} \quad \ldots(8.27)$$

On defining the mean motion n_α, given by

$$n_\alpha = \left(\frac{\mu}{a^3}\right)^{\frac{1}{2}} \quad \ldots(8.28),$$

the combination of equations 8.26 and 8.27 gives

$$(1-e^2)^{\frac{1}{2}} \overset{(8.21),(8.22)}{=} \frac{a^2(1-e^2)^{\frac{1}{2}}[\, n_\alpha(t-T) - \sin E\, (\cos E - e)]}{a^2[E - \sin E \cos E]}$$

or

$$n_\alpha(t-T) = E - e \sin E \quad \ldots(8.29).$$

Notes

(i) Equation 8.29 is known as Kepler's equation.

(ii) Consider the quantity M defined by

$$M = E - e \sin E \overset{(8.29)}{=} n_\alpha(t-T) \quad \ldots(8.30)$$

and called the *mean anomaly*. It can be seen from equation 8.30 that the position of a satellite in its elliptical orbit can be uniquely defined by the true anomaly f which, from Kepler's second law (equation 8.15), has an irregular motion with time. It can also be defined by the eccentric anomaly E which, again, varies irregularly with time (equation 8.29). The mean anomaly M, on the other hand, represents a mean motion along the auxiliary circle (equation 8.30), n_α representing an angular velocity or rate at which the satellite completes a single orbit. This concept is very useful in studying the characteristics of the earth's gravitational field for periods longer than that occupied by a single revolution in the orbit.

(iii) For near circular orbits, $M \to E$ as $e \to 0$.

ii. The orbital plane, the celestial sphere and earth space Cartesian co-ordinate systems

A three dimensional Cartesian co-ordinate system X_i is adopted as an earth space reference frame , the X_3 axis being coincident with the earth's rotation axis. The X_I axis in figure 8.3

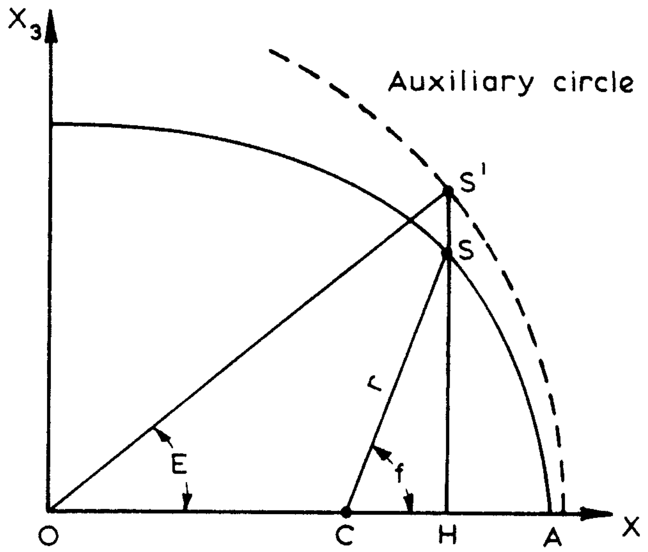
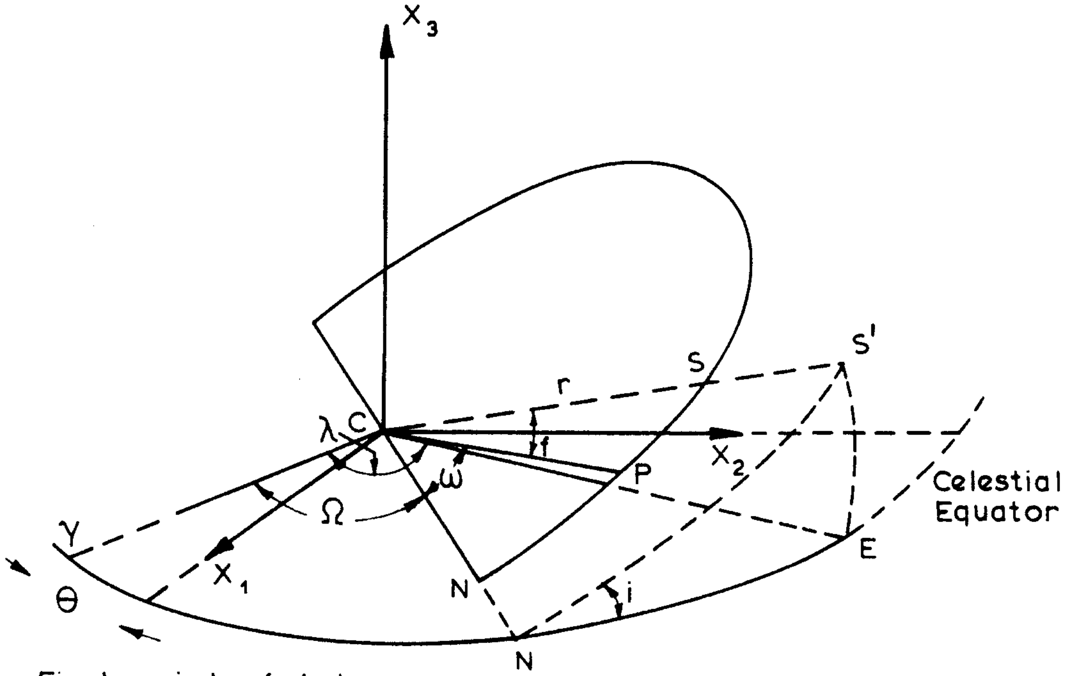


FIG. 8.2

Satellite in an elliptical orbit



γ = First point of Aries
P = Perigee

FIG. 8.3

The orbital plane, the celestial sphere and inertial space is shown to be coincident with the line of intersection of the meridian at Greenwich and the equator. If the orbital plane is inclined at an angle i to the equatorial plane, $S'\hat{N}'E'$ in triangle $S'N'E'$ on the celestial sphere is i, where E' is the projection of the foot of the perpendicular E from the satellite's position S onto the equatorial plane, S' being the same relation to S. The angle ECX_1 is equal to the longitude λ. The line CN in which the plane of the orbit intersects that of the equator, defined by the X_1X_2 plane, is called the *line of nodes* and $N\hat{C}\gamma$ is called the right ascension Ω of the ascending node, where γ is the first point of Aries.

If P is the closest point in the orbit to the earth which is, by definition, on the major axis of the orbital ellipse, $P\hat{C}N$ is called the *argument of perigee* ω, while $S\hat{C}P = f$ is the true anomaly with the same significance as in figure 8.2. The spherical triangle S'E'N' in figure 8.3 on the celestial sphere is right angled at E' and the great circle arc lengths comprising its sides are given by

$$N'S' = \omega + f \quad ; \quad S'E' = \phi \quad ; \quad N'E' = \lambda - \Omega$$

if the Greenwich Sidereal time θ is zero. (λ, ϕ) are the geographical co-ordinates of the satellite such that

$0 < \lambda < 2\pi$, positive east and $-\frac{1}{2}\pi \leq \phi \leq \frac{1}{2}\pi$, positive north.

Let

$$u = \omega + f \qquad \ldots(8.31).$$

The application of Napier's rules of circular parts to triangle S'N'E' gives

Sin(any part)
=Π[tan(adjacent parts)]
=Π[cos(opposite parts)].

$\frac{1}{2}\pi - u$

$\frac{1}{2}\pi - i$ $\qquad$ $\frac{1}{2}\pi - N'\hat{S}'E'$

$\lambda - \Omega$ $\qquad$ ϕ

Hence,

$$\left.\begin{aligned}
\cos u &= \cos\phi\cos(\lambda-\Omega) && \ldots a\\
\cos i &= \cot u\tan(\Lambda-\Omega) && \ldots b\\
\sin(\lambda-\Omega) &= \tan\phi\cot i && \ldots c\\
\sin\phi &= \sin u\sin i && \ldots d\\
\text{and}\quad \sin(\lambda-\Omega) &= \sin u\cos i && \ldots e
\end{aligned}\right\} \quad (8.32).$$

Thus the orbital parameters are related to the earth space co-ordinate system by the equations

$$\left.\begin{aligned}
X_1 &= r[\cos u\cos\Omega - \sin u\cos i\sin\Omega]\\
X_2 &= r[\cos u\sin\Omega + \sin u\cos i\cos\Omega]\\
X_3 &= r\sin u\sin i
\end{aligned}\right\} \quad \ldots(8.33),$$

where r is given by equations 8.23 and 8.24. The true anomaly f is related to the eccentric anomaly E through equation 8.21 which can be transposed as follows.

$$a \cos E = r \cos f - ae \overset{(8.24)}{=} \frac{(1-e^2)\cos f + ae(1+e\cos f)}{1+e\cos f} .$$

Thus

$$\cos E = \frac{e + \cos f}{1 + e\cos f} \qquad \ldots(8.34).$$

As

$$1 + \cos E = \frac{(1+e)(1+\cos f)}{1 + e\cos f} = 2\cos^2 \tfrac{1}{2}E \qquad \ldots(8.35)$$

and

$$1 - \cos E = \frac{(1-e)(1-\cos f)}{1 + e\cos} = 2\sin^2 \tfrac{1}{2}E ,$$

it follows that

$$\frac{1 - \cos E}{1 + \cos E} = \frac{\sin^2 \frac{1}{2}E}{\cos^2 \frac{1}{2}E} = \left(\frac{1-e}{1+e}\right)\frac{\sin^2 \frac{1}{2}f}{\cos^2 \frac{1}{2}f}$$

and

$$\tan \tfrac{1}{2}E = \left(\frac{1-e}{1+e}\right)^{\frac{1}{2}} \tan \tfrac{1}{2}f \qquad \ldots(8.36).$$

Also,

$$\sec^2 \tfrac{1}{2}E \, \frac{dE}{2} = \left(\frac{1-e}{1+e}\right)^{\frac{1}{2}} \sec^2 \tfrac{1}{2}f \, \frac{df}{2} .$$

As

$$\sec^2 \tfrac{1}{2}f = \frac{2}{1+\cos f} ,$$

the use of equation 8.35 gives

$$dE = \left(\frac{1-e}{1+e}\right)^{\frac{1}{2}} \frac{(1+e)(1+\cos f)}{2(1+e\cos f)} \frac{2}{1+\cos f} df = \frac{(1-e^2)^{\frac{1}{2}}}{1+e\cos f} df \qquad \ldots(8.37).$$

Similarly, the use of equations 8.23 and 8.24 gives

$$1 + e\cos f = \frac{(1-e^2)}{1-e\cos E} .$$

Thus

$$df \overset{(8.37)}{=} \frac{(1-e^2)}{1-e\cos E} \frac{1}{(1-e^2)^{\frac{1}{2}}} dE = \frac{(1-e^2)^{\frac{1}{2}}}{1-e\cos E} dE \qquad \ldots(8.38).$$

iii Motion in a perturbed gravitational field

It is no longer possible to consider the motion of a satellite under a central force alone when a satellite orbits in the gravitational field of a body with small departures from symmetry. The total gravitational force vector F can be considered to be the resultant of two vectors, one of which is defined in equation 8.17 while the other is the disturbing attraction represented by the vector R. Thus

$$F = \ddot{r} = -\frac{\mu}{r^3} r + R \qquad \ldots(8.39).$$

If this attraction is obtained from the disturbing potential V_d, it follows that

$$\ddot{r} + \frac{\mu}{r^3} r = \nabla V_d \qquad \ldots(8.40),$$

where ∇ is defined by equation 2.6.

Notes

(i) The position vector r can be defined by equation 8.6. It can also be related through the equations at 8.33 to the orbital elements $c_i(i=1,6)$ defining the Keplerian orbital ellipse where

$$c_i = \{a, e, f(\text{or } E \text{ or } M), \Omega, \omega, i\} \qquad \ldots(8.41).$$

These elements are called the Keplerian elements.

(ii) The existence of the disturbing potential, given by

$$W = \frac{\mu}{r} + V_d \qquad \ldots(8.42),$$

results in c_i no longer being independent of time . In contrast, only f(or E or M) change with time in the case of pure central force motion. This is of significance in formulating expressions for accelerations and their relations to gravitational forces.

(iii) The consequence is a set of changes dc_i in the Keplerian elements c_i which are functions of the disturbing potential V_d. The next section develops formulae establishing the required relations between these quantities.

(iv) The orbital period of an artificial earth satellite is dependent on its elevation above the earth. An estimate of this quantity can be obtained from the mean motion defined in equation 8.28. Thus if the equatorial radius a of the orbit is approximately 6700 km, the period of revolution is 90 minutes. Table 8.1 gives some idea of the relation between orbital period and the equatorial radius.

Equatorial radius a (km)	Orbital period T (min)
6700	90
7200	100
10600	180
12800	240
16800	360 (6 hr)
26600	720 (12 hr)
42300	1440 (24 hr)

Table 8.1

The period of a single orbit

The apparent motion of the satellite is made more complex by the rotation of the earth from west to east. Consequently every succeeding pass of a satellite will appear a further 23 ° westward if its period is 90 minutes. Thus the position of the observer with respect to the orbital plane is restored every 24 hours if *the latter remains fixed in inertial space*. It does not follow that the satellite orbital plane remains fixed in such a space with the passage of time in the case of a gravitating body like the earth.

8.2 The disturbing potential in terms of the Keplerian elements

i. Lagrange's equations

The problem may be stated as follows. The "disturbed" gravity field is due to departures of the earth from a sphere containing the same mass as the earth in a symmetrical distribution. The resulting potential discrepancies are represented by the disturbing potential V_d in equation 8.42. The existence of these deviations results in the orbital motion no longer being the Keplerian ellipse. The motion can still be conveniently represented by the concept of an *instantaneous* orbit defined by the set of parameters c_i whose values are time dependent. Such an orbit is not the true path of the satellite during a complete revolution but it has the important property that the rate of change of position with respect to time, in the limit, produces no change in the orbital elements as it is the essence of the definition of the instantaneous orbit. This does *not* apply to the second derivatives which define the changes between successive instantaneous orbits.

Let the position vector r be given by

$$r = r\{c_i,\ i=1,6),t\} \qquad ..(8.43),$$

where the c_i are given by equation 8.41. The velocity $\dot{r}$ of the satellite in inertial space is

$$\dot{r} = \frac{dr}{dt} + \sum_{i=1}^{6} \frac{\partial r}{\partial c_i} \dot{c}_i$$

where $\partial r/\partial t$ is its velocity in the instantaneous orbit which is equal to the velocity in inertial space. Thus

$$\sum_{i=1}^{6} \frac{\partial r}{\partial c_i} \dot{c}_i = 0 \qquad ...(8.44).$$

Further differentiation gives

$$\ddot{r} = \frac{\partial^2 r}{\partial t^2} + \sum_{i=1}^{6} \frac{\partial^2 r}{\partial t\, \partial c_i}\dot{c}_i + o\{\frac{\partial r}{\partial c_i}\ddot{c}_i\} \qquad \ldots(8.45).$$

As the intention is to effect a change of variables and relate the changes in the orbital elements to the disturbing potential, the combination of equations 8.40 and 8.45 gives

$$\frac{\partial^2 r}{\partial t^2} + \sum_{i=1}^{6} \frac{\partial \dot{r}}{\partial c_i}\dot{c}_i + \frac{\mu}{r^3} r = \nabla V_d.$$

For a non-perturbed field when no changes occur in the orbital elements with time, equation 8.40 becomes

$$\ddot{r} + \frac{\mu}{r^3} r = 0 \qquad \ldots(8.46)$$

which is equation 8.17. Thus

$$\sum_{i=1}^{6} \frac{\partial \dot{r}}{\partial c_i}\dot{c}_i = \nabla V_d \qquad \ldots(8.47).$$

Equations 8.44 and 8.47 completely define orbital perturbations in terms of the deviations V_d of the total geopotential from that of a sphere with the same mass and symmetrically distributed. The scalar multiplication of equation 8.44 by $\partial\dot{r}/\partial c_j$ and of equation 8.47 by $\partial r/\partial c_j$ and differencing the results gives

$$\sum_{i=1}^{6} \left(\frac{\partial r}{\partial c_j} \cdot \frac{\partial \dot{r}}{\partial c_i} - \frac{\partial \dot{r}}{\partial c_j} \cdot \frac{\partial r}{\partial c_i} \right) c_i = \frac{\partial r}{\partial c_j} \cdot \nabla V_d \text{ , } j=1,6 \qquad ..(8.48).$$

The six first order differential equations comprising the above set replace the three second order differential equations implied in equation 8.40. These equations and the subsequent treatment are due to Lagrange who then proceeded to devise an elegant means for solving them to obtain expressions of more direct application. He first defined a system of notation for the expression of equation 8.48 as

$$\sum_{i=1}^{6} [c_j, c_i]\, \dot{c}_i = \sum_{i=1}^{3} \frac{\partial V_d}{\partial X_i} \frac{\partial X_i}{\partial c_j} = \frac{\partial V_d}{\partial c_j} \text{ , } j=1,6 \qquad ..(8.49),$$

where $[c_j, c_i]$ are called Lagrange's brackets and given by

$$[c_j, c_i] = \sum_{k=1}^{3} J\left(\frac{X_k, \dot{X}_k}{c_j, c_i} \right) \qquad \ldots(8.50),$$

and the Jacobian J by the relation

$$J\left(\frac{X_k,\dot{X}_k}{c_j,c_i}\right) = \begin{vmatrix} \frac{\partial X_k}{\partial c_j} & \frac{\partial \dot{X}_k}{\partial c_j} \\ \frac{\partial X_k}{\partial c_i} & \frac{\partial \dot{X}_k}{\partial c_i} \end{vmatrix} \qquad \ldots(8.51).$$

Notes

These Lagrangian brackets have the following properties.

(i)

$$[c_i,c_i] = 0 \qquad \ldots(8.52)$$

as

$$J\left(\frac{X_k,\dot{X}_k}{c_i,c_i}\right) = 0.$$

(ii)

$$[c_j,c_i] \overset{(8.51)}{=} -[c_i,c_j] \qquad \ldots(8.53).$$

(iii) The Lagrangian brackets are independent with time. This is proved as follows. Consider the variation of

$$J = J\left(\frac{X,\dot{X}}{p,q}\right) \qquad \text{with time.}$$

$$\frac{\partial J}{\partial t} \overset{(8.51)}{=} \left(\frac{\partial^2 X}{\partial t\,\partial p}\frac{\partial \dot{X}}{\partial q} + \frac{\partial X}{\partial p}\frac{\partial^2 \dot{X}}{\partial t\,\partial q}\right) - \frac{\partial^2 X}{\partial t\,\partial q}\frac{\partial \dot{X}}{\partial p} - \frac{\partial X}{\partial q}\frac{\partial^2 \dot{X}}{\partial t\,\partial p}$$

which can be written as

$$\frac{\partial J}{\partial t} = \left(\frac{\partial \dot{X}}{\partial p}\frac{\partial \dot{X}}{\partial q} - \frac{\partial \dot{X}}{\partial q}\frac{\partial \dot{X}}{\partial p}\right) + \frac{\partial X}{\partial p}\frac{\partial}{\partial q}(\ddot{X}) - \frac{\partial X}{\partial q}\frac{\partial}{\partial p}(\ddot{X})$$

$$\overset{(8.2)}{=} 0 + \left(\frac{\partial X}{\partial p}\frac{\partial}{\partial q}\left(\frac{\partial V}{\partial X}\right) - \frac{\partial X}{\partial q}\frac{\partial}{\partial p}\left(\frac{\partial V}{\partial X}\right)\right)\dot{X}$$

$$= \left(\frac{\partial}{\partial X}\left(\frac{\partial V}{\partial q}\right)\frac{\partial X}{\partial p} - \frac{\partial}{\partial X}\left(\frac{\partial V}{\partial p}\right)\frac{\partial X}{\partial q}\right)\dot{X} \qquad = \left(\frac{\partial^2 V}{\partial p\,\partial q} - \frac{\partial^2 V}{\partial p\,\partial q}\right)\dot{X} =$$

$$= 0 .$$

Thus

$$\frac{\partial}{\partial t}[c_j,c_i] = 0 \qquad \ldots(8.54).$$

This relation is of importance as it enables the evaluation of the Lagrangian brackets in that position in the orbit which lends itself to ease of computation. Thus the Lagrangian brackets which need evaluation are the following 15:-

$[c_1,c_2]$, $[c_1,c_3]$, $[c_1,c_4]$, $[c_1,c_5]$, $[c_1,c_6]$, $[c_2,c_3]$, $[c_2,c_4]$, $[c_2,c_5]$, $[c_2,c_6]$, $[c_3,c_4]$, $[c_3,c_5]$, $[c_3,c_6]$, $[c_4,c_5]$, $[c_4,c_6]$ and $[c_5,c_6]$.

ii. The evaluation of Lagrange's brackets

The Lagrangian brackets are evaluated by converting the reference frame from the three dimensional inertial system implied in section 8.2*(i)* to one in two dimensions defining the orbital plane. Let the inertial axis system X_i be defined as in figure 8.3 and section 8.1*(ii)*, the directions along the axes being specified by the unit vectors i as in section 8.1*(i)*. If l_i are unit vectors along the x_i axis system, where the x_1 axis is along the major axis of the instantaneous orbital ellipse, the x_2 axis perpendicular to it in the orbital plane (i.e., $f = \frac{1}{2}\pi$) and the x_3 axis normal to the orbital plane (i.e., $l_3 = l_1 \times l_2$, with the origin at the focus, it can be seen that

$$l_j = A_{ji}\, i_i \qquad \ldots(8.55),$$

where the elements of $\mathbf{A}_{ji}$ constitute the array A obtained from figure 8.4 as

$$A = \begin{vmatrix} \cos\omega\cos\Omega - \sin\omega\cos i\sin\Omega & \cos\omega\sin\Omega + \sin\omega\cos i\cos\Omega & \sin\omega\sin i \\ -\sin\omega\cos\Omega - \cos\omega\cos i\sin\Omega & -\sin\omega\sin\Omega + \cos\omega\cos i\cos\Omega & \cos\omega\sin i \\ \sin i\sin\Omega & -\sin i\cos\Omega & \cos i \end{vmatrix}$$

$$= |A_{ij}| \qquad \ldots(8.56).$$

The transformation between the X_i and x_i co-ordinate systems is therefore dependent on only 3 of the c_i's(i.e., ω,Ω,i) which define the location of the instantaneous orbital ellipse in inertial space. The three other c_i's (a,e,M) define the position of the satellite in an orbit of specified size and are independent of the co-ordinate transformation defined by equations 8.55 and 8.56. On defining these two types of Keplerian elements as

$$c_{1i} = \{a,e,M\} \quad \text{and} \quad c_{2i} = \{\omega,\Omega,i\} \qquad \ldots(8.57),$$

there are three types of Lagrangian brackets to be evaluated. These are $[c_{1j},c_{1i}]$, $[c_{1j},c_{2i}]$ and $[c_{2j},c_{2i}]$. The position vector, lying entirely in the orbital plane, can be represented by the equation

$$r = \sum_{i=1}^{3} X_i i = \sum_{\alpha=1}^{2} x_\alpha l_\alpha \qquad \ldots(8.58).$$

The velocity vector can be represented as

$$\dot{r} = \sum_{i=1}^{3} \dot{X}_i i \overset{(8.58)}{=} \sum_{\alpha=1}^{2} \dot{x}_\alpha l_\alpha \qquad \ldots(8.59).$$

on retaining the instantaneous orbit as the frame of reference. Thus

$$[c_{1j},c_{1i}] \overset{(8.50)\downarrow(8.51)}{=} \sum_{k=1}^{3} \begin{vmatrix} \frac{\partial X_k}{\partial c_{1i}} & \frac{\partial \dot{X}_k}{\partial c_{1i}} \\ \frac{\partial X_k}{\partial c_{1j}} & \frac{\partial \dot{X}_k}{\partial c_{1j}} \end{vmatrix} \qquad \ldots(8.60).$$

In addition,

$$\frac{\partial X_k}{\partial c_{1i}} = \sum_{\alpha=1}^{2} \frac{\partial X_k}{\partial x_\alpha} \frac{dx_\alpha}{dc_{1i}} \quad , \ k=1,3 \qquad \ldots(8.61)$$

and

$$\frac{\partial \dot{X}_k}{\partial c_{1i}} = \sum_{\alpha=1}^{2} \frac{\partial \dot{X}_k}{\partial \dot{x}_\alpha} \frac{dx_\alpha}{dc_{1i}} \quad , \ k=1,3 \qquad \ldots(8.62),$$

where

$$\frac{\partial X_k}{\partial x_\alpha} \overset{(8.56)}{=} A_{\alpha k} = \frac{\partial \dot{X}_k}{\partial \dot{x}_\alpha} \quad , \ k=1,3; \ \alpha=1,2 \ .$$

Thus

$$J\left(\frac{X_k,\dot{X}_k}{c_{1j},c_{1i}}\right) = \begin{vmatrix} \sum_{\alpha=1}^{2} A_{\alpha k} \frac{\partial x_\alpha}{\partial c_{1j}} & \sum_{\alpha=1}^{2} A_{\alpha k} \frac{\partial \dot{x}_\alpha}{\partial c_{1j}} \\ \sum_{\alpha=1}^{2} A_{\alpha k} \frac{\partial x_\alpha}{\partial c_{1i}} & \sum_{\alpha=1}^{2} A_{\alpha k} \frac{\partial \dot{x}_\alpha}{\partial c_{1i}} \end{vmatrix} \quad , \ k=1,3 \qquad \ldots(8.63).$$

As

$$\sum_{k=1}^{3} A_{\alpha k}^2 = 1 \qquad \ldots(8.64) \qquad \text{and} \qquad \sum_{k=1}^{3} A_{1k}A_{2k} = 0 \qquad \ldots(8.65),$$

$$\begin{aligned} [c_{1j},c_{1i}] \overset{(8.60)\downarrow(8.63)}{=} \sum_{k=1}^{3} \Big(& \frac{\partial x_1}{\partial c_{1j}} \frac{\partial \dot{x}_1}{\partial c_{1i}} A_{1k}^2 + \frac{\partial x_1}{\partial c_{1j}} \frac{\partial \dot{x}_2}{\partial c_{1i}} A_{1k}A_{2k} + \frac{\partial x_2}{\partial c_{1j}} \frac{\partial \dot{x}_1}{\partial c_{1i}} A_{2k}A_{1k} + \\ & \frac{\partial x_2}{\partial c_{1j}} \frac{\partial \dot{x}_2}{\partial c_{1i}} A_{2k}^2 - \frac{\partial x_1}{\partial c_{1i}} \frac{\partial \dot{x}_1}{\partial c_{1j}} A_{1k} - \frac{\partial x_1}{\partial c_{1j}} \frac{\partial \dot{x}_2}{\partial c_{1i}} A_{1k}A_{2k} - \\ & A_{2k}A_{1k} \frac{\partial x_2}{\partial c_{1j}} \frac{\partial \dot{x}_1}{\partial c_{1i}} - \frac{\partial x_2}{\partial c_{1i}} \frac{\partial \dot{x}_2}{\partial c_{1j}} A_{2k}^2 \Big) \\ \overset{(8.64)\downarrow(8.65)}{=} & \sum_{\alpha=1}^{2} \begin{vmatrix} \frac{\partial x_\alpha}{\partial c_{1j}} & \frac{\partial x_\alpha}{\partial c_{1j}} \\ \frac{\partial x_\alpha}{\partial c_{1i}} & \frac{\partial x_\alpha}{\partial c_{1i}} \end{vmatrix} = \sum_{\alpha=1}^{2} J\left(\frac{x_\alpha,\dot{x}_\alpha}{c_j,c_i} \right) \qquad \ldots(8.66). \end{aligned}$$

Equation 8.66 expresses Lagrange's brackets in the case of the Keplerian elements c_{1i} in terms of a plane rectangular Cartesian system in the orbital plane without attributing any special properties to the latter. The required expressions are obtained on evaluation near perigee when

$$\sin E = E - \frac{E^3}{6} + o\{E^5\} \quad ; \quad \cos E = 1 - \tfrac{1}{2}E^2 + o\{E^4\} \ .$$

Thus

$$M \overset{(8.30)}{=} E(1-e) \overset{(8.29)}{=} n\,(t-T)$$

or

$$E = \frac{n_\alpha (t-T)}{1-e} \qquad \ldots(8.67).$$

In addition it can be seen from figure 8.2 that

$$x_1 = r \cos f \overset{(8.21)}{=} a(\cos E - e)$$

while

$$x_2 = r \sin f \overset{(8.22)}{=} a(1-e^2)^{\frac{1}{2}} \sin E \quad .$$

The use of equation 8.67, together with the replacement of n_α by equation 8.28 gives

$$x_1 = a\left[1 - \frac{\mu(t-T)^2}{2a^2(1-e)^2} - e\right] \quad ; \quad x_2 = \left(\frac{\mu(1+e)}{a(1-e)}\right)^{\frac{1}{2}}(t-T) \qquad \ldots(8.68).$$

Differentiation with respect to t gives

$$\dot{x}_1 = -\frac{\mu(t-T)}{a^2(1-e)^2} \quad ; \quad \dot{x}_2 = \left(\frac{\mu(1+e)}{a(1-e)}\right)^{\frac{1}{2}} \qquad \ldots(8.69).$$

As

$$c_{11} = a \quad ; \quad c_{12} = e \quad ; \quad c_{13} = M^* = M + n_\alpha T = n_\alpha t,$$

where the form adopted for c_{13} is of convenience as $\frac{\partial M^*}{\partial t} = n_\alpha$, the Lagrangian brackets can be evaluated as follows.

$$[a,e] \overset{(8.66)}{=} \lim_{t\to T}\left(\frac{\partial x_1}{\partial a}\frac{\partial \dot{x}_1}{\partial e} + \frac{\partial x_2}{\partial a}\frac{\partial \dot{x}_2}{\partial e} - \frac{\partial x_1}{\partial e}\frac{\partial \dot{x}_1}{\partial a} - \frac{\partial x_2}{\partial e}\frac{\partial \dot{x}_2}{\partial a} \right)$$

$$= \left((1-e)\times 0 + 0\times \frac{\partial \dot{x}_2}{\partial e} - (-a)\times 0 - 0\times \frac{\partial \dot{x}_2}{\partial a} \right)$$

$$= 0 \qquad \ldots(8.70)$$

The principles adopted in the evaluation of the Lagrangian brackets at perigee, when $t = T$, are as follows. All terms containing (t-T) will be zero and only those terms which are *not* multiplied by the term (t-T) need be considered. Thus

$$[a,M^*] = \lim_{t\to T}\left(\frac{\partial x_1}{\partial a}\frac{\partial \dot{x}_1}{\partial t}\frac{\partial t}{\partial M^*} + \frac{\partial x_2}{\partial a}\frac{\partial \dot{x}_2}{\partial t}\frac{\partial t}{\partial M^*} - \frac{\partial x_1}{\partial t}\frac{\partial \dot{x}_1}{\partial a}\frac{\partial t}{\partial M^*} - \frac{\partial x_2}{\partial t}\frac{\partial \dot{x}_2}{\partial a}\frac{\partial t}{\partial M^*} \right)$$

$$= (1-e)\left(\frac{\mu}{a^2(1-e)^2}\right)\frac{1}{n_\alpha} + 0 - 0 - \left(\frac{\mu(1+e)}{a(1-e)}\right)^{\frac{1}{2}}\frac{1}{n_\alpha}\times\left(\frac{\mu(1+e)}{a^3(1-e)}\right)^{\frac{1}{2}}\times(-\tfrac{1}{2})$$

$$= \frac{1}{n_\alpha}\left(-\frac{\mu}{a^2(1-e)} + \tfrac{1}{2}\frac{\mu(1+e)}{a^2(1-e)} \right) = -\tfrac{1}{2}n_\alpha a \qquad \ldots(8.71).$$

Similarly,

$$[e,M^*] = 0 \ .$$

The terms $[c_{2j}, c_{2i}]$ cannot be conveniently treated in this manner and are best defined by the use of equations 8.56, 1.31 and 1.32 in relation to the inertial reference frame(e.g., Kaula 1966b,p.28).

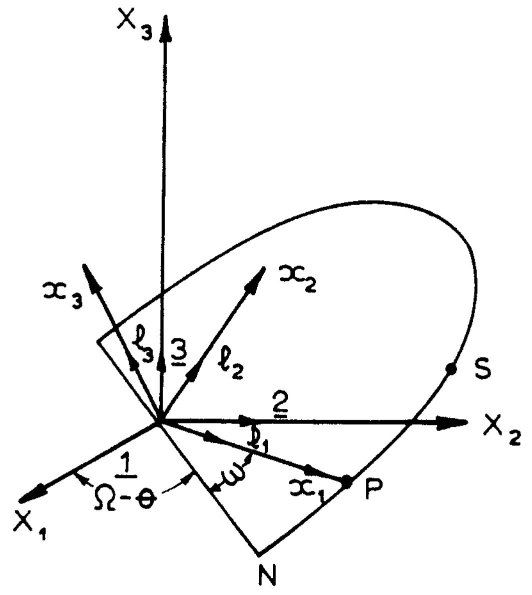


FIG. 8.4
Cartesian axes and the orbital plane.

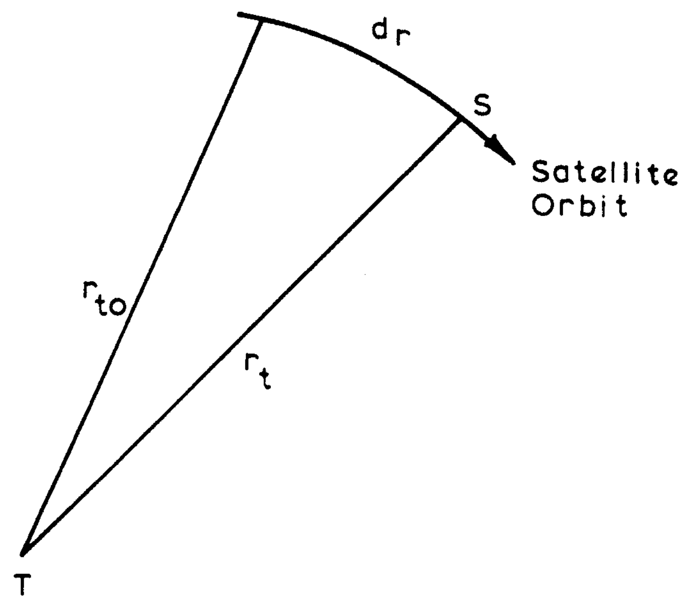


FIG. 8.5
The Doppler System

The remaining non-zero Lagrange's brackets are

$$[i,\Omega] = n_\alpha a^2(1-e^2)^{\frac{1}{2}}\sin i \quad ; \quad [\Omega,a] = \tfrac{1}{2}n_\alpha a \cos i\,(1-e^2)^{\frac{1}{2}}$$

$$[\omega,a] = \tfrac{1}{2}n_\alpha a\,(1-e^2)^{\frac{1}{2}} \quad ; \quad [e,\Omega] = n_\alpha \frac{a^2 e \cos i}{(1-e^2)^{\frac{1}{2}}}$$

$$[e,\omega] = n_\alpha \frac{a^2 e}{(1-e^2)^{\frac{1}{2}}} \qquad ..(8.72).$$

iii. Lagrange's planetary equations in terms of the orbital parameters

Equation 8.49 can be written as follows on noting that

$$[c_i,c_i] \overset{(8.52)}{=} 0 \quad , \quad [c_i,c_j] \overset{(8.53)}{=} -[c_j,c_i]$$

and choosing

$c_1 = a$; $c_2 = e$; $c_3 = M^*$; $c_4 = i$, $c_5 = \omega$; $c_6 = \Omega$.

When $j = 1$

$$[a,a]\dot{a} + [a,e]\dot{e} + [a,M^*]\dot{M}^* + [a,i]\frac{di}{dt} + [a,\omega]\dot{\omega} + [a,\Omega]\dot{\Omega} = \frac{\partial V}{\partial a}d$$

or

$$-\tfrac{1}{2}n_\alpha a\,\dot{M}^* - \tfrac{1}{2}n_\alpha a(1-e^2)^{\frac{1}{2}}\dot{\omega} - \tfrac{1}{2}n_\alpha a\cos i(1-e^2)^{\frac{1}{2}}\dot{\Omega} = \frac{\partial V}{\partial a}d \qquad ...(8.73a).$$

Similarly, when $j = 2$,

$$\frac{n_\alpha a^2 e}{(1-e^2)^{\frac{1}{2}}}\dot{\omega} + \frac{n_\alpha a^2 e \cos i}{(1-e^2)^{\frac{1}{2}}}\dot{\Omega} = \frac{\partial V}{\partial e}d \qquad ...(8.73b);$$

when $j = 3$,

$$\tfrac{1}{2}n_\alpha a\,\dot{a} = \frac{\partial V}{\partial M^*}d \quad .(8.73c);$$

when $j = 4$,

$$n_\alpha a^2(1-e^2)^{\frac{1}{2}}\sin i\,\dot{\Omega} = \frac{\partial V}{\partial i}d.(8.73d);$$

when $j = 5$,

$$\tfrac{1}{2}n_\alpha a(1-e^2)^{\frac{1}{2}}\dot{a} - \frac{n_\alpha a^2 e}{(1-e^2)^{\frac{1}{2}}}\dot{e} = \frac{\partial V}{\partial \omega}d..(8.73d);$$

when $j = 6$,

$$\tfrac{1}{2}n_\alpha a\cos i(1-e^2)^{\frac{1}{2}} - \frac{n_\alpha a^2\, e \cos i}{(1-e^2)^{\frac{1}{2}}}\dot{e} \qquad -n_\alpha a^2(1-e^2)^{\frac{1}{2}}\sin i\,\frac{di}{dt} = \frac{\partial V}{\partial \Omega}d \qquad ....(8.73f).$$

The changes in the orbital elements with time are obtained from the above equations on successive evaluations as follows.

$$\frac{da}{dt} \overset{(8.73c)}{=} \frac{2}{n_\alpha a}\frac{\partial V}{\partial M}d \qquad ...(8.74a).$$

$$\frac{de}{dt} \overset{(8.73c)(8.74a)}{=} \frac{1-e^2}{n_\alpha a^2 e}\frac{\partial V}{\partial M}d - \frac{(1-e^2)^{\frac{1}{2}}}{n_\alpha a^2 e}\frac{\partial V}{\partial \omega}d \qquad ...(8.74b).$$

$$\frac{di}{dt} \overset{(8.73f)(8.74a)(8.74b)}{=} \frac{1}{n_\alpha a^2(1-e^2)^{\frac{1}{2}}\sin i}\left\{\cos i(1-e^2)^{\frac{1}{2}}\frac{\partial V_d}{\partial M} - \right.$$

$$\left. - \frac{\cos i}{(1-e^2)^{\frac{1}{2}}}\left((1-e^2)\frac{\partial V_d}{\partial M} - (1-e^2)^{\frac{1}{2}}\right)\frac{\partial V_d}{\partial \omega} - \frac{\partial V_d}{\partial \Omega}\right\}$$

$$= \frac{\cot i}{n_\alpha a^2(1-e^2)^{\frac{1}{2}}}\frac{\partial V_d}{\partial \omega} - \frac{\operatorname{cosec} i}{n_\alpha a^2(1-e^2)^{\frac{1}{2}}}\frac{\partial V_d}{\partial \Omega} \quad \ldots(8.74c).$$

$$\frac{d\Omega}{dt} \overset{(8.73d)}{=} \frac{\operatorname{cosec} i}{n_\alpha a\,(1-e^2)^{\frac{1}{2}}}\frac{\partial V_d}{\partial i} \quad \ldots(8.74d).$$

$$\frac{d\omega}{dt} \overset{(8.73b),(8.74d)}{=} \frac{(1-e^2)^{\frac{1}{2}}}{n_\alpha a^2 e}\left(\frac{\partial V_d}{\partial e} - \frac{e\cot i}{1-e^2}\frac{\partial V_d}{\partial i}\right)$$

$$= \frac{(1-e^2)^{\frac{1}{2}}}{n_\alpha a^2 e}\frac{\partial V_d}{\partial e} - \frac{\cot i}{n_\alpha a^2(1-e^2)^{\frac{1}{2}}}\frac{\partial V_d}{\partial i} \quad \ldots(8.74e)$$

and

$$\frac{\partial M^*}{\partial t} = \left(\frac{2}{a n_\alpha} - \frac{1-e^2}{2ae}\frac{\partial V_d}{\partial e} + \frac{\cot i}{2a}\frac{\partial V_d}{\partial i} - \frac{\cot i}{2a}\frac{\partial V_d}{\partial i} - \frac{\partial V_d}{\partial a}\right)$$

$$= -\frac{1-e^2}{n_\alpha a e}\frac{\partial V_d}{\partial e} - \frac{2}{n_\alpha a}\frac{\partial V_d}{\partial a}\ .$$

As

$$\frac{\partial M^*}{\partial t} = \frac{\partial M}{\partial t} - n_\alpha\ ,$$

$$\frac{\partial M}{\partial t} = n_\alpha - \frac{1-e^2}{n_\alpha a e}\frac{\partial V_d}{\partial e} - \frac{2}{n_\alpha a}\frac{\partial V_d}{\partial a} \quad \ldots(8.74f).$$

Notes

(i) Equations 8.74a - e are called *Lagrange's planetary equations* and express variations in the orbital elements as a consequence of the effect of the disturbing potential V_d in respect to the orbital elements.

(ii) The term $\partial V_d/\partial M$ has been written for what strictly should $\partial V_d/\partial M^*$ in equations 8.74a and 8.74b. This is in order as the rates of change of both M and M* with respect to time are equally affected by changes in V_d.

(iii) If the changes in the instantaneous orbital parameters with time can be measured, these can be related through equations 8.74a to e to the perturbations of the earth's gravitational field from that of a sphere containing a symmetrical mass distribution.

(iv) The disturbing potential V_d can be expressed by the spherical harmonic series

$$V_d = \frac{\mu}{r}\sum_{n=2}^{\infty}\left(\frac{a_e}{r}\right)^n \sum_{m=0}^{n}\bar{P}_{nm}(\sin\phi)[\bar{C}_{nm}\cos m\lambda + \bar{S}_{nm}\sin m\lambda] \quad \ldots(8.75)$$

from equation 7.24, where

$$\bar{P}_{nm}(\sin\phi) \overset{(3.38),(4.29)}{=} \left[\frac{(2n+1)(n-m)!(2-\delta_{0m})}{(n+m)!}\right]^{\frac{1}{2}}\frac{\cos^m}{2^n}\sum_{t=0}^{k}\frac{(-1)^t(2n-2t)!\sin^{n-m-2t}\phi}{(n-m-2t)!\,t!\,(n-t)!}$$

$$\ldots(8.76),$$

where

$k = \text{Integer}\left(\frac{n-m}{2}\right)$ and δ_{om} is the Kronecker delta.

The problem therefore reduces to one which seeks to establish the relations governing the transformation

$$V_d(r,\phi,\lambda) \rightarrow V_d(a,e,i,M,\omega,\Omega) \qquad \ldots(8.77)$$

iii. The conversion of surface harmonic terms to expressions in the Keplerian elements

The development given in this section is due to Kaula (1961). The position of the satellite can be represented on a spherical co-ordinate system by the parameters (r,ϕ,λ) while this same definition can be achieved by the use of the system afforded by the Keplerian elements (a,e,i,M,ω,Ω). From normal celestial sphere considerations in figure 8.3,

$$\lambda = \text{Right Ascension} - \text{Greenwich Sidereal Time} = R - \theta$$
$$= (R - \Omega) - (\Omega - \theta),$$

where R is the right ascension of the satellite and θ the Greenwich sidereal time. Thus the terms $\cos m\lambda$ and $\sin m\lambda$ in equation 8.75 can be replaced by the relations

$$\left.\begin{aligned} \cos m\lambda &= \cos[m(R-\Omega) + m(\Omega-\theta)] \\ &= \cos[m(R-\Omega)]\cos[m(\Omega-\theta)] - \sin[m(R-\Omega)]\sin[m(\Omega-\theta)] \\ \text{and} \quad & \\ \sin m\lambda &= \sin[m(R-\Omega)]\cos[m(\Omega-\theta)] + \cos[m(R-\Omega)]\sin[m(\Omega-\theta)] \end{aligned}\right\} \ldots(8.78).$$

In addition,

$$\cos mX = \mathcal{R}(e^{imX}) = \mathcal{R}([\cos X + i\sin X]^m)$$
$$= \mathcal{R}\left(\sum_{s=0}^{m} \binom{m}{s} i^s \cos^{m-s}X \sin^s X\right) \qquad \ldots(8.79),$$

where

$$\left.\begin{aligned} \mathcal{R}(X) &= \text{Real part of } X \quad \text{and} \quad i = \sqrt{(-1)} \\ \text{and} \quad \binom{m}{s} &= \frac{m(m-1)(m-2)\ \ldots\ (m-s+1)}{s!} = \frac{m!}{s!(m-s)!} \end{aligned}\right\} \ldots(8.80),$$

while

$$\sin mX = -\mathcal{R}(i\, e^{imX})$$
$$= \mathcal{R}\left(\sum_{s=0}^{m} \binom{m}{s} i^{s-1} \cos^{m-s}X \sin^s X\right) \qquad \ldots(8.81).$$

Thus,

$$\cos[m(R-\Omega)] \overset{(8.33)}{=} R\left(\sum_{s=0}^{m} \binom{m}{s} i^s \frac{\cos^{m-s}u}{\cos^{m-s}\phi} \frac{\sin^s u \cos^s i}{\cos^s \phi} \right)$$

$$= R\left(\sum_{s=0}^{m} \binom{m}{s} i^s \frac{\cos^{m-s}u \sin^s u \cos^s i}{\cos^m \phi} \right) \quad \ldots(8.82)$$

and

$$\sin[m(R-\Omega)] = R\left(\sum_{s=0}^{m} \binom{m}{s} i^{s-1} \frac{\cos^{m-s}u \sin^s u \cos^s i}{\cos^m \phi} \right) \ldots(8.83).$$

Thus the equations at 8.78 become

$$\cos m\lambda = R\left(\sum_{s=0}^{m} i^s \binom{m}{s} \frac{\cos^{m-s}u \sin^s u \cos^s i}{\cos^m \phi} \cos[m(\Omega-\theta)] + i\sin[m(\Omega-\theta)] \right)$$

$$\ldots(8.84).$$

and

$$\sin m\lambda = R\left(\sum_{s=0}^{m} i^s \binom{m}{s} \frac{\cos^{m-s}u \sin^s u \cos^s i}{\cos^m \phi} \sin[m(\Omega-\theta)] - i\cos[m(\Omega-\theta)] \right)$$

The use of equations 8.84 and 8.32 transform equation 8.75 to give the expression for the disturbing potential as

$$V_d = \frac{\mu}{r} \sum_{n=2}^{\infty} \left(\frac{a_e}{r}\right)^n \sum_{m=0}^{n} \frac{K_{nm}}{2^n} \sum_{t=0}^{k} \frac{(-1)^t (2n-2t)!}{(n-m-2t)!\ t!\ (n-t)!} \sin^{n-m-2t} i \sin^{n-m-2t} u \times$$

$$R\left(\sum_{s=0}^{m} \binom{m}{s} i^s \sin^s u \cos^{m-s} u \cos^s i \left([\bar{C}_{nm} - i\bar{S}_{nm}] \cos\{m(\Omega-\theta)\} + \right.\right.$$

$$\left.\left. [\bar{S}_{nm} + i\bar{C}_{nm}] \sin\{m(\Omega-\theta)\} \right) \right)$$

where

$$K_{nm} \overset{(8.76)}{=} \left(\frac{(2n+1)(n-m)!(2-\delta_{om})}{(n+m)!} \right)^{\frac{1}{2}} \quad \ldots(8.85).$$

As

$$\sin^a X \cos^b X = \left[-\frac{i}{2}(e^{iX} - e^{-iX}) \right]^a \left[\tfrac{1}{2}(e^{iX} + e^{-iX}) \right]^b$$

$$= \frac{(-1)^a i^a}{2^{a+b}} \sum_{c=0}^{a} (-1)^c \binom{a}{c} e^{i(a-c)X} e^{-icX} \times \sum_{d=0}^{b} \binom{b}{d} e^{i(b-d)X} e^{-idX}$$

$$= \frac{(-1)^a i^a}{2^{a+b}} \sum_{c=0}^{a} (-1)^c \sum_{d=0}^{b} \binom{a}{c} \binom{b}{d} e^{i(a+b-2c-2d)X}$$

$$= \frac{(-1)^a i^a}{2^{a+b}} \sum_{c=0}^{a} (-1)^c \sum_{d=0}^{b} \binom{a}{c} \binom{b}{d} [\cos(a+b-2c-2d)X + i\sin(a+b-2c-2d)X]$$

$$\ldots(8.86),$$

the earlier equation, on the combination of all indices of sin u and the application of equation 8.86, becomes

$$V_d = \frac{\mu}{r}\sum_{n=2}^{\infty}\left(\frac{a_e}{r}\right)^n \sum_{m=0}^{n} \frac{K_{nm}}{2^n}\sum_{t=0}^{k}\frac{(-1)^t(2n-2t)!}{(n-m-2t)!\,t!\,(n-t)!}\sin^{n-m-2t} i \times$$

$$R\Big(\sum_{s=0}^{m}\binom{m}{s} i^s \cos^s i\,\frac{(-i)^{n-m-2t+s}}{2^{n-2t}}\sum_{c=0}^{n-m-2t+s}(-1)^c\sum_{d=0}^{m-s}\binom{n-m-2t+s}{c}\binom{m-s}{d}\times$$

$$\Big([\bar{C}_{nm} - i\,\bar{S}_{nm}]\cos\{m(\Omega-\theta)\} + [\bar{S}_{nm} + i\bar{C}_{nm}]\sin\{m(\Omega-\theta)\}\Big)\times$$

$$\Big(\cos(n-2t-2c-2d)u + i\sin(n-2t-2c-2d)u\Big)\Big) \qquad \ldots(8.87).$$

The function whose real part has to be ascertained in equation 8.87, consists of a series of terms, one of whose constituents is a set of trigonometrical functions which are made up of products of the arguments $m(\Omega-\theta)$ and $(n-2t-2c-2d)$. The real terms are those which have terms with i raised to an even power as coefficients. The combination of these terms with terms which are powers of (-1) gives

$$(-1)^{n-m-2t+s+t}\times\ i^{n-m-2t+s+s} = \left.\begin{matrix}(-1)^{\frac{1}{2}(n-m)-t+s} & \text{if } (n-m) \text{ even}\\ i(-1)^{\frac{1}{2}(n-m-1)-t+s} & \text{if } (n-m) \text{ odd}\end{matrix}\right\}(-1)^{n-m-t+s}$$

$$= \begin{cases}(-1)^{[3(n-m)/2-2t+2s]}\\ (-i)(-1)^{[3(n-m-1)/2-2t+2s]}\end{cases} = \begin{cases}(-1)^{-\frac{1}{2}(n-m)} & \text{if } (n-m) \text{ even}\\ (-i)(-1)^{-\frac{1}{2}(n-m-1)} & \text{if } (n-m) \text{ odd}\end{cases}$$

$$= \begin{cases}(-1)^{-k} & \text{if } (n-m) \text{ even}\\ (-i)(-1)^{-k} & \text{if } (n-m) \text{ odd}\end{cases} \qquad \ldots(8.88)$$

as $(-1)^{-2t}, (-1)^{2s}, (-1)^{4(n-m)/2}$ and $(-1)^{4(n-m-1)/2}$ are always equal to +1.

In addition, on defining the index p by the relation

$$p = t + c + d \qquad \ldots(8.89)$$

and as

$$\cos X \cos Y = \tfrac{1}{2}[\cos(X+Y) + \cos(X-Y)] \;;$$
$$\sin X \sin Y = \tfrac{1}{2}[\cos(X-Y) - \cos(X+Y)] \;;$$
$$\sin X \cos Y = \tfrac{1}{2}[\sin(X+Y) + \sin(X-Y)] \;;$$
$$\cos X \sin Y = \tfrac{1}{2}[\sin(X+Y) - \sin(X-Y)] \;;$$

it follows that

$$\Big([\bar{C}_{nm} - i\bar{S}_{nm}]\cos\{m(\Omega-\theta)\} + [\bar{S}_{nm} + i\,\bar{C}_{nm}]\sin\{m(\Omega-\theta)\}\Big)\times$$
$$[\cos(n-2p)u + i\sin(n-2p)u]$$

can be expressed as

$$\tfrac{1}{2}\left([\bar{C}_{nm}-i\bar{S}_{nm}]-[\bar{S}_{nm}+i\bar{C}_{nm}]\right)\cos\{(n-2p)u+m(\Omega-\theta)\}+$$
$$\tfrac{1}{2}\left([\bar{C}_{nm}-i\bar{S}_{nm}]+[\bar{S}_{nm}+i\bar{C}_{nm}]\right)\cos\{(n-2p)u-m(\Omega-\theta)\}+$$
$$\tfrac{1}{2}\left([\bar{S}_{nm}+i\bar{C}_{nm}]+[\bar{C}_{nm}-i\bar{S}_{nm}]\right)\sin\{(n-2p)u+m(\Omega-\theta)\}+$$
$$\tfrac{1}{2}\left([\bar{S}_{nm}+i\bar{C}_{nm}]+[\bar{C}_{nm}-i\bar{S}_{nm}]\right)\sin\{(n-2p)u+m(\Omega-\theta)\}$$

$$=\left(\bar{C}_{nm}-i\bar{S}_{nm}\right)[\cos\{(n-2p)u+m(\Omega-\theta)\}]+\left(\bar{S}_{nm}+i\bar{C}_{nm}\right)[\sin\{(n-2p)u+m(\Omega-\theta)\}] \quad ..(8.90).$$

Thus equation 8.87 becomes

$$V_d = \frac{\mu}{r}\sum_{n=2}^{\infty}\left(\frac{a_e}{r}\right)^n\sum_{m=0}^{n}K_{nm}\sum_{t=0}^{k}\frac{(2n-2t)!}{(n-m-2t)!\ t!\ (n-t)!}\sin^{n-m-2t} i \ \times$$

$$\sum_{s=0}^{m}\binom{m}{s}\frac{\cos^{s} i}{2^{2n-2t}}\sum_{c=0}^{n-m-2t+s}(-1)^{c-k}\sum_{d=0}^{m-s}\binom{n-m-2t+s}{c}\binom{m-s}{p-t-c}\times$$

$$\left(\begin{bmatrix}\bar{C}_{nm}\\ -\bar{S}_{nm}\end{bmatrix}_{(n-m)\,odd}^{(n-m)\,even}\cos\{(n-2p)u+m(\Omega-\theta)\}+\begin{bmatrix}\bar{S}_{nm}\\ \bar{C}_{nm}\end{bmatrix}_{(n-m)\,odd}^{(n-m)\,even}\sin\{(n-2p)u+m(\Omega-\theta)\}\right)$$

...(8.91).

Notes

(i) Equation 8.91 is simplified further by using p, defined in equation 8.89 as the sixth index instead of d. The maximum value p can take (p_{max}) is given from equation 8.89 as

$$p_{max} = \frac{n-m}{2} + [n-m-2\frac{n-m}{2}+m] + (m-m) = \frac{n+m}{2},$$

which is a maximum when $m = n$. Thus $p_{max} = n$.

(ii) The limts of c are obtained as follows.

As $0 \le p-t-c \le m-s$,

$c \ge (p-t)-(m-s)$, i.e., $(p-t) \ge (m-s) \ge 0$.

as (m-s) is always greater than 0. It also follows, for this same reason that $(p-t) \ge c$. Thus

$$p \ge t,$$

the equality holding when $m = 0$.

In addition,

$$0 \le c \le (n-m-2t+s).$$

Thus the range of c is given by

$(p-t)<(m-s)$	0	$\le c \le$ $n-m-2t+s$	if $n-m-2t+s \le (p-t)$
$(p-t)>(m-s)$	$p-t-m+s$	$p-t$	if $n-m-2t+s \ge (p-t)$

...(8.92).

The range of c in each case commences at the greater of the lower indices and terminated at the smaller of the upper values, being the shortest range of values in each case.

(iii) The maximum value of t on using d as the sixth index was t_{max} given by

$$t_{max} = k = \text{Integer}[\frac{n-m}{2}] .$$

In addition, it is also shown in note (ii) above that $t \leq p$. This means that non-zero terms occur when the series in t is summed so that

$$t_{max} = \begin{cases} k = \text{Integer}[\frac{n-m}{2}] \\ p \end{cases} \text{, whicherever is less}$$

...(8.93).

Kaula (1966b,p.34) expresses the disturbing potential by the equation

$$V_d = \sum_{n=2}^{\infty} \sum_{m=0}^{n} V_{d_{nm}}$$

where

$$V_{d_{nm}} = \frac{\mu}{r} \left(\frac{a_e}{r}\right)^n K_{nm} \sum_{p=0}^{n} F_{nmp}(i) \left\{ \begin{bmatrix} \bar{C}_{nm} \\ -\bar{S}_{nm} \end{bmatrix}_{(n-m)\,\text{odd}}^{(n-m)\,\text{even}} \cos\{(n-2p)u+m(\Omega-\theta)\} + \right.$$

$$\left. \begin{vmatrix} \bar{S}_{nm} \\ \bar{C}_{nm} \end{vmatrix}_{(n-m)\,\text{odd}}^{(n-m)\,\text{even}} \sin\{(n-2p)u+m(\Omega-\theta)\} \right\} \qquad ...(8.94),$$

where

$$F_{nmp}(i) = \sum_{t=0}^{t_{max}} \frac{(2n-2t)!}{2^{2n-2t}\, t!\,(n-t)!\,(n-m-2t)!} \sin^{n-m-2t} i$$

$$\sum_{s=0}^{m} \binom{m}{s} \cos^s i \sum_{c} (-1)^{c-k} \binom{n-m-2t+s}{c} \binom{m-s}{p-t-c}$$

...(8.95).

Notes

(i) Direct evaluation of the function $F_{nmp}(i)$ shows that the possible values of m, p and the expression for $F_{nmp}(i)$ are as follows for $n = 2$.

m	p	$F_{2mp}(i)$
0	0	$-\frac{3}{8}\sin^2 i$
0	1	$\frac{3}{4}\sin^2 i - \frac{1}{2}$
0	2	$-\frac{3}{8}\sin^2 i$
1	0	$\frac{3}{4}\sin i(1+\cos i)$
1	1	$-\frac{3}{2}\sin i \cos i$
1	2	$-\frac{3}{4}\sin i(1-\cos i)$
2	0	$\frac{3}{4}(1+\cos i)^2$
2	1	$\frac{3}{2}\sin^2 i$
2	2	$\frac{3}{4}(1-\cos i)^2$

Expressions up to and including n = 4 are given in by Kaula (ibid,p.34).

The second stage in the development is the replacement of r and f in equation 8.94 by a, e and M. The mean anomaly M is introduced as the analysis seeks to study long period variations. It has the facility of affording a simple means for the averaging of the effect of a single orbit using a relation of the form

$$M\{X\} = \frac{1}{2\pi}\int^{\pi} X \quad dX \quad .$$

The use of equations 8.24, 8.30 and 8.37 gives

$$dM = dE(1-e\cos E) \overset{(8.37)}{=} (1-e\cos E)\ \frac{(1-e^2)^{\frac{1}{2}}}{1+e\cos f}\ df$$

$$\overset{(8.23)\ (8.24)}{=} \frac{r^2}{a^2(1-e^2)^{\frac{1}{2}}}\quad df \qquad \ldots(8.96).$$

In addition,

$$\frac{1}{r^{n-1}} \overset{(8.24)}{=} \left(\frac{1+e\cos f}{a(1-e^2)}\right)^{n-1} = \frac{1}{a^{n-1}(1-e^2)^{n-1}} \sum_{b=0}^{n-1}\binom{n-1}{b} e^b \cos^b f$$

$$\overset{(8.56)}{=} \frac{1}{a^{n-1}(1-e^2)^{n-1}} \quad \sum_{b=0}^{n-1}\binom{n-1}{b}\frac{e^b}{2^b}\sum_{d=0}^{b}\binom{b}{d}\cos(b-2d)f \qquad ..(8.97).$$

Thus the long period effect (i.e., $M > 2\pi$) of the relevant terms is given by expressions of the form

$$M\{X\} = M\{\frac{1}{r^{n+1}}\left|\begin{matrix}\sin\\ \cos\end{matrix}\right|[(n-2p)(\omega+f) + m(\Omega-\theta)]\}$$

$$= \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{r^{n+1}}\left|\begin{matrix}\sin\\ \cos\end{matrix}\right|[(n-2p)(\omega+f) + m(\Omega-\theta)]\right) dM$$

$$\overset{(8.97)}{=} \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{a^{n+1}}\frac{1}{(1-e^2)^{n-\frac{1}{2}}}\sum_{b=0}^{n\ 1}\binom{n-1}{b}\frac{e^b}{2^b}\sum_{d=0}^{b}\binom{b}{d}\cos(b-2d)f \quad\times\right.$$

$$\left.\left|\begin{matrix}\sin\\ \cos\end{matrix}\right|[(n-2p)(\omega+f) + m(\Omega-\theta)]\right)df \ .$$

As

$$Y = \cos[(b-2d)f]\left|\begin{matrix}\sin\\ \cos\end{matrix}\right|[(n-2p)(\omega+f) + m(\Omega-\theta)]$$

$$= \tfrac{1}{2}\quad\left|\begin{matrix}\sin\\ \cos\end{matrix}\right|[(n-2p)\omega + (n-2p+b-2d)f + m(\Omega-\theta)] \ +$$

$$\left|\begin{matrix}\sin\\ \cos\end{matrix}\right|[(n-2p)\omega + (n-2p-b+2d)f + m(\Omega-\theta)] \qquad \ldots(8.98),$$

and as long period effects are not contributed to by terms in f(or M), the coefficient of f in equation 8.98 should be zero. This occurs when

$$n - 2p \pm (b - 2d) = 0.$$

Thus

$$b = 2d \pm (n-2p) = 2d + n - 2p' \text{ , where} \begin{cases} p'=p & \text{when} \quad p \leq \tfrac{1}{2}n \\ p'=n-p & \text{when} \quad p \geq \tfrac{1}{2}n \end{cases} \quad \ldots(8.99).$$

Thus only two terms in the inner d series satisfy the condition for long period variation, being symmetrically placed coefficients in the binomial expansion of $(1+e\cos f)^{n-1}$ for each value of b. $M\{X\}$ can be written as

$$M\{X\} = \frac{1}{a^{n+1}} G_{np(2p-n)}(e) \left| \begin{matrix} \sin \\ \cos \end{matrix} \right| [(n-2p)\omega + m(\Omega-\theta)] \quad ..(8.100),$$

where

$$G_{np(2p-n)}(e) = \frac{1}{(1-e)^{n-\frac{1}{2}}} \sum_{d=0}^{p'-1} \binom{n-1}{2d+n-2p'} \binom{2d+n-2p'}{d} \left(\frac{e}{2}\right)^{2d+n-2p'} \quad ..(8.101).$$

The solution for the disturbing potential in the general case can therefore be written as (ibid,p.37)

$$V_d = \sum_{n=2}^{\infty} \sum_{m=0}^{n} V_{d_{nm}}$$

where

$$V_{d_{nm}} = \frac{\mu}{a} \left(\frac{a}{r}\right)^n K_{nm} \sum_{p=0}^{n} F_{nmp}(i) \sum_{q=-\infty}^{\infty} G_{npq}(e)\; S_{nmpq}(\omega,M,\Omega,\theta) \quad \ldots(8.102)$$

and $S_{nmpq}(\omega,M,\Omega,\theta)$ is given by

$$S_{nmpq}(\omega,M,\Omega,\theta) \overset{(8.94),(8.98)}{=} \left| \begin{matrix} \bar{C}_{nm} \\ -\bar{S}_{nm} \end{matrix} \right| \begin{matrix} (n-m)\text{ even} \\ \\ (n-m)\text{ odd} \end{matrix} \cos[(n-2p)\omega + M(n-2p+q) + m(\Omega-\theta)] +$$

$$\left| \begin{matrix} \bar{S}_{nm} \\ \bar{C}_{nm} \end{matrix} \right| \begin{matrix} (n-m)\text{ even} \\ \\ (n-m)\text{ odd} \end{matrix} \sin[(n-2p)\omega + (n-2p+q)M + m(\Omega-\theta)] \quad \ldots(8.103).$$

Notes

(i)

The condition for long period effects is given by the equation

$$n - 2p + q = 0.$$

(ii) The long term contributions from the term $G_{npq}(e)$ come from the quantities $G_{np(2p-n)}(e)$. The values of these long term contributions for values of $n \leq 4$ are given in table 8.2. A more complete table is give by Kaula (ibid,p.38) after Cayley. A study of table 8.2 shows that the major contributions occur when $q = 0$ and n is even. Contributions an order of magnitude smaller occur when $q = 0$ and n is odd.

Thus the infinite summation in equation 8.102 is not necessary in the case of practical solutions when only a few terms (usually two for every value of n) make significant contributions as the function is a power series in the eccentricity which is a very

n	p	q = 2p − n	p'	$G_{np(2p-n)}(e)$
2	0	-2	0	0
2	1	0	1	$(1-e^2)^{-3/2}$
2	2	2	0	0
3	0	-3	0	0
3	1	-1	1	$(1-e^2)^{-5/2}$
3	2	1	1	$(1-e^2)^{-5/2}$
3	3	3	0	0
4	0	-4	0	0
4	1	-2	1	$\frac{3e^2}{4}(1-e^2)^{-7/2}$
4	2	0	2	$[1+3e^2/2](1-e^2)^{-7/2}$
4	3	2	1	$\frac{3e^2}{4}(1-e^2)^{-7/2}$
4	4	4	0	0

Table 8.2

The function $G_{np(2p-n)}(e)$

small quantity. It should also be noted that the resulting terms are symmetric about the zero value of the index q.

8.3 The evaluation of the coefficients of the harmonic series

Techniques for the observation of near earth satellites are beyond the scope of the present development. Details are given in (Mueller 1964,pp.235 et seq.; Gaposchkin 1966,pp.77 et seq.). Such observations give the instantaneous co-ordinates of the satellite on the celestial sphere which can then be converted to the appropriate set of Keplerian elements. A practical solution of this problem is afforded by the Differential Orbit Improvement (DOI) program developed by the Smithsonian Astrophysical Observatory (Gaposchkin 1967,pp.97 et seq). Lagrange's equations are based on the assumption that the second derivatives of the variations of the orbital elements with time are negligible quantities. The time interval between successive observations on the same satellite could be as long as several days. Hence these equations can only be applied successfully for the analysis of those effects which satisfy the conditions governing their formation. Variations in the magnitude of the orbital elements are the consequence of a number of effects, some of short period (e.g., 24 hours or less), others of long period (e.g., 100 days) as well as secular variations which are linear with time.

These long period and secular variations are related to the disturbing potential V_d as expressed in equation 8.102 in the following manner.

Linear perturbations

These variations satisfy the conditions under which Lagrange's equations, given at 8.74, were derived. If the disturbing potential is given by equation 8.102 and if

$$\frac{dF}{di}_{nmp} = F'_{nmp}(i) \quad \ldots(8.104)$$

and

$$\frac{dS}{d\alpha}_{nmpq} = S'_{nmpq}(\Omega,\omega,M,\theta) \quad ,..(8.105),$$

where

$$\alpha = (n-2p)\omega + (n-2p+q)M + m(\Omega-\theta) \quad \ldots(8.106)$$

and

$$\frac{dG}{de}_{npq} = G'_{nmp}(e) \quad \ldots(8.107).$$

The equations at 8.74 can be written as

$$\frac{da}{dt}_{nmpq} = \frac{2}{n_\alpha a}\left(\frac{ae}{a}\right)^n K_{nm} F_{nmp} G_{npq} S'_{nmpq}(n-2p+q) \quad ..(8.108a).$$

$$\frac{de}{dt}_{nmpq} = \frac{\mu}{n_\alpha a^3 e}\left(\frac{a_e}{a}\right)^n K_{nm} F_{nmp} G_{npq} S'_{nmpq}\left((1-e^2)(n-2p+q) - (1-e^2)^{\frac{1}{2}}(n-2p)\right) \quad ..(8.108b).$$

$$\frac{di}{dt}_{nmpq} = \frac{\mu}{n_\alpha a^3(1-e^2)^{\frac{1}{2}}}\left(\frac{a_e}{a}\right)^n \operatorname{cosec} i\, K_{nm} F_{nmp} G_{npq} S'_{nmpq}\left((n-2p)\cos i - m\right) \quad \ldots(8.108c)$$

$$\frac{d\Omega}{dt}_{nmpq} = \frac{\mu \operatorname{cosec} i}{n_\alpha a^3(1-e^2)^{\frac{1}{2}}}\left(\frac{a_e}{a}\right)^n K_{nm} F'_{nmp} G_{npq} S_{nmpq} \quad \ldots(8.108d).$$

$$\frac{d\omega}{dt}_{nmpq} = \frac{\mu}{n_\alpha a^3}\left(\frac{a_e}{a}\right)^n K_{nm} S_{nmpq}\left(\frac{(1-e^2)^{\frac{1}{2}}}{e} F_{nmp} G'_{npq} - \frac{\cot i}{(1-e^2)^{\frac{1}{2}}} F'_{nmp} G_{npq}\right) \quad \ldots(8.108e).$$

$$\frac{dM}{dt}_{nmpq} = n_\alpha + \frac{\mu}{n_\alpha a^3}\left(\frac{a_e}{a}\right)^n K_{nm} F_{nmp} S_{nmpq}\left(2(n+1)G_{npq} - \frac{1-e^2}{e} G'_{npq}\right) \quad \ldots(8.108f).$$

Notes

(i) Equations 8.108a - f give the contributions for single terms in the total expressions for the changes in the orbital elements with time.

(ii) As the condition for long period variations is

$$n - 2p + q = 0,$$

it can be seen that no long period variations occur in a. Similarly the term (n-2p+q) in equation 8.108b) does not contribute to de_{nmpq}/dt.

Equations 8.108 a-f can be integrated with respect to time when only S_{nmpq} and S'_{nmpq} can be considered to require integration. Let

$$\int S_{nmpq}\, dt = \int S_{nmpq} \frac{dt}{d\alpha} d\alpha = \frac{I\{S_{nmpq}\}}{\dot{\alpha}}$$

where

$$I\{S_{nmpq}\} = \int S_{nmpq}\, d\alpha \qquad \ldots(8.109)$$

and

$$\dot{\alpha} = (n-2p)\dot{\omega} + (n-2p+q)\dot{M} + m(\dot{\Omega}-\dot{\theta}) \qquad \ldots(8.110).$$

Similarly

$$\int S'_{nmpq}\, d\alpha = \frac{S_{nmpq}}{\dot{\alpha}} \qquad \ldots(8.111).$$

Thus integration of the equations at 8.108 gives the following expressions for the variations in the orbital elements.

$$\Delta a_{nmpq} = \frac{2\mu}{n_\alpha a^2}\left(\frac{a_e}{a}\right)^n K_{nm} \frac{F_{nmp} G_{npq} S_{nmpq}(n-2p+q)}{\dot{\alpha}} \qquad ..(8.112a).$$

$$\Delta e_{nmpq} = \frac{\mu}{n_\alpha a^3}\left(\frac{a_e}{a}\right)^n K_{nm} \frac{F_{nmp} G_{npq} S_{nmpq}}{\dot{\alpha}}[(1-e^2)(n-2p+q) - (1-e^2)^{\frac{1}{2}}(n-2p)] \qquad ..(8.112b).$$

$$\Delta i_{nmpq} = \frac{\mu}{n_\alpha a^3(1-e^2)^{\frac{1}{2}}}\left(\frac{a_e}{a}\right)^n K_{nm} \operatorname{cosec} i \frac{F_{nmp} G_{npq} S_{nmpq}}{\dot{\alpha}}[(n-2p)\cos i - m] \qquad \ldots(8.112c)$$

$$\Delta\Omega_{nmpq} = \frac{\mu \operatorname{cosec} i}{n_\alpha a^3(1-e^2)^{\frac{1}{2}}}\left(\frac{a_e}{a}\right)^n \frac{K_{nm} F'_{nmp} G_{npq} I\{S_{nmpq}\}}{\dot{\alpha}} \qquad \ldots(8.112d).$$

$$\Delta\omega_{nmpq} = \frac{\mu}{n_\alpha a^3}\left(\frac{a_e}{a}\right)^n K_{nm} \frac{I\{S_{nmpq}\}}{\dot{\alpha}}\left(\frac{(1-e^2)^{\frac{1}{2}}}{e} F_{nmp} G'_{npq} - \frac{\cot i}{(1-e^2)^{\frac{1}{2}}} F'_{nmp} G_{npq}\right) \qquad \ldots(8.112e).$$

$$\Delta M_{nmpq} = \frac{\mu}{n_\alpha a^3}\left(\frac{a_e}{a}\right)^n \frac{K_{nm} F_{nmp} I\, S_{nmpq}}{\dot{\alpha}}\left[2(n+1)G_{npq} - \frac{1-e^2}{e} G'_{npq}\right] \qquad \ldots(8.112f).$$

Notes

(i) The total linear perturbation of any orbital element Δc_i is given by an equation of the form

$$\Delta c_i = \sum_{n=2}^{\infty} \sum_{m=0}^{n} \sum_{p=0}^{n} \sum_{q=-\infty}^{\infty} \Delta c_{i_{nmpq}} .$$

(ii) It can be seen that the largest perturbations occur when $\dot{\alpha}$ is smallest, i.e., when

$$(n-2p)\dot{\omega} + (n-2p+q)\dot{M} + m(\dot{\Omega}-\dot{\theta}) \rightarrow 0.$$

The condition

$$n - 2p + q = 0$$

has already been specified as the relation governing long period effects.

(iii) $\dot{\theta}$ = 1 cycle per day,

being the angular velocity of rotation of the earth. The quantities $\dot{\Omega}, \dot{\omega}$ which are the rates of rotation of the line of nodes (a regressive motion) and the argument of perigee, are both observed to be of the order of 10^{-2} cycles per day.

Consequently, if only those variations over periods in excess of a single orbit of the satellite is considered, the dominant variation is that due to the term $m(\dot{\Omega}-\dot{\theta})$, which cancels itself out m times a day. Thus harmonics which cause

$$m(\dot{\Omega}-\dot{\theta}) \rightarrow 0$$

give rise to comparatively large variations in the orbital elements when the satellite has an integral number of revolutions per day. This effect is known as *resonance*.

(iv) As pointed out earlier, a study of table 8.3 shows that C_{20} is three orders larger than all other coefficients C_{nm}, S_{nm}. It is therefore necessary to consider the perturbation of second order due to the term $[C_{20}]^2$ which has the same order of magnitude as the general coefficients. For a treatment, see (Kaula 1966b,pp.41 et seq).

The evaluation of zonal harmonics

These harmonics are obtained when $m = 0$.
and can give rise to

(a) short period variations, i.e., with periods less than a single revolution of the satellite which is of the order of 90 for the usual case;

(b) long period variations, i.e., periods governed by the case when

$$n - 2p + q = 0,$$

which could be as long as 100 days;

and (c) secular variations, i.e., changes which are linear with time and hence should be independent of trigonometrical terms.

On the adoption of the condition for long period variation $q = 2p - n$ in the case where $m = 0$, equation 8.102 becomes

$$V_{d_{no}} = \frac{\mu}{a}\left(\frac{a_e}{a}\right)^n K_{nm}\, \overline{C}_{no}\, F_{nop}(i)\, G_{np(2p-n)}(e) \begin{cases} \cos(n-2p)\omega & , n \text{ even} \\ \sin(n-2p)\omega & , n \text{ odd}, \end{cases}$$

as $\overline{S}_{no} = 0$.

The variations of longer period are obtained as $n - 2p \rightarrow 0$. It can be seen from table 8.2 that two different results can be obtained depending on whether n is even (when $(n-2p) = 0$) or odd when $(n-2p) = \pm 1$ giving two terms, when the contributions of longest period and greatest magnitude are given by

$$V_{d_{no}} = \frac{\mu}{a}\left(\frac{a_e}{a}\right)^n K_{nm}\,\bar{C}_{no}\,F_{nop}(i)\,G_{np(2p-n)}(e)\begin{cases}1 & ,\ n \text{ even}\\ 2\sin\omega & ,\ n \text{ odd}\end{cases} \quad \ldots(8.113).$$

As the variations in a, i and e are relatively small over small periods, their values can be treated as constants when evaluating equation 8.113. Hence secular variations are due entirely to even degree zonal harmonics while the dominant long period effects are due to odd degree zonal harmonics. This can be studied with cmmparative ease by putting $n - 2p = j$ when it can be seen from equation 8.113 and table 8.2 that j is even when n is even and odd when n is odd.

The most interesting zonal harmonics are obtained when n takes the values 2, 3. When $n = 2$, the secular variation is given by $j = 0$ or $p = 1$ and $q = 2p - n = 0$, when the equations at 8.108 take the form

$$\Delta a_{2010} = 0 \quad \ldots(8.114a),$$
$$\Delta e_{2010} = 0 \quad \ldots(8.114b),$$
$$\Delta i_{2010} = 0 \quad \ldots(8.114c),$$

$$\Delta\Omega_{2010} = \frac{\mu \operatorname{cosec} i}{n_\alpha a^3 (1-e^2)^{\frac{1}{2}}}\left(\frac{a_e}{a}\right)^2 \sqrt{5}\,\frac{\partial F_{201}}{\partial i}\,G_{210}\bar{C}_{20}\Delta t,$$

which, on using equations 8.28, 8.95 and table 8.2, gives

$$\Delta\Omega_{2010} = \frac{3\,n_\alpha}{2(1-e^2)^2}\left(\frac{a_e}{a}\right)^2\sqrt{5}\,\cos i\,\bar{C}_{20}\,\Delta t \quad \ldots(8.114d).$$

Similarly,

$$\Delta\omega_{2010} = n_\alpha\left(\frac{a_e}{a}\right)^2\sqrt{5}\left\{\frac{(1-e^2)^{\frac{1}{2}}}{e}F_{201}(i)\,\frac{\partial G_{210}}{\partial e} - \frac{\cot i}{(1-e^2)^{\frac{1}{2}}}\,\frac{\partial F_{201}}{\partial i}G_{210}\right\}\times \bar{C}_{20}\Delta t$$

$$= \frac{n_\alpha}{(1-e^2)^2}\left(\frac{a_e}{a}\right)^2\sqrt{5}\left(3\left\{\tfrac{3}{4}\sin^2 i - \tfrac{1}{2}\right\} - \tfrac{3}{2}\sin i\cos i\cot i\right)\bar{C}_{20}\Delta t$$

$$= \frac{3\,n_\alpha}{4(1-e^2)^2}\left(\frac{a_e}{a}\right)^2\sqrt{5}(1 - 5\cos^2 i)\,\bar{C}_{20}\,\Delta t \quad \ldots(8.114e)$$

and

$$\Delta M_{2010} = \left\{n_\alpha + n_\alpha\,F_{201}(i)\left(\frac{a_e}{a}\right)^2\sqrt{5}\left(6\,G_{201} - \frac{1-e^2}{e}\,\frac{\partial G_{201}}{\partial e}\right)\bar{C}_{20}\right\}\Delta t$$

$$= n_\alpha\,\Delta t\left\{1 + \left\{\tfrac{3}{4}\sin^2 i - \tfrac{1}{2}\right\}\left(\frac{a_e}{a}\right)^2\left(\frac{6}{(1-e^2)^{3/2}} - \frac{1-e^2}{e}\,\frac{3\,e}{(1-e^2)^{5/2}}\right)\bar{C}_{20}\right\}$$

$$= n_\alpha\,\Delta t\left\{1 + \frac{3\,a_e^2}{4(1-e^2)^{3/2}a^2}(3\cos^2 i - 1)\sqrt{5}\,\bar{C}_{20}\right\} \quad \ldots(8.114f).$$

As $\bar{C}_{20}$ is seen from table 8.3 to be $\simeq 4.84 \times 10^{-3}$, and as normal near earth satellite orbits which are of geodetic use have rations $(a_e/a) \doteqdot 1.1$, e being of order 10^{-2}, it can be seen that

$$\frac{d\omega}{dt}_{2010} \doteq 3.5(5\cos^2 i - 1) \text{ degrees per day;}$$

$$\frac{d\Omega}{dt}_{2010} \doteq -6.7\cos i \text{ degrees per day;}$$

$$\frac{dM}{dt}_{2010} \doteq 14.4 + 9\times10^{-3}(3\cos^2 i - 1) \text{ revolutions per day.}$$

Thus the line of nodes can be seen to regress, while perigee advances if $i < 63^\circ 26'$ or regresses if this is not so. Graphical representation of these equations is given by Mueller(1964,pp.201-2).

When $n = 3$,
two significant terms occur, as can be seen from table 8.2. These are
$n = 3; m = 0$; $(n-2p) = \pm 1$, i.e., $p = 2$ or 1 and $q = (2p-n) = 1$ or -1.
The resulting perturbation which is no longer secular , has a wavelength 2π in ω for the above subscripts. In such a case

$$\Delta a_{3011} = \Delta a_{301(-1)} = \Delta a_{3021} = \Delta a_{302(-1)} = 0$$

Table 8.2 shows that only $G_{31(-1)}$ and G_{321} are non-zero.
Hence

$$\Delta e_{30} = \Delta e_{301(-1)} + \Delta e_{3021} = n_\alpha \left(\frac{a_e}{a}\right)^3 \sqrt{7}\,\bar{C}_{30} \left(\frac{F_{301}(i)\,G_{31(-1)}(e)\sin\omega}{\dot{\omega}} \times \left(-(1-e^2)^{\frac{1}{2}}\right) + \frac{F_{302}(i)\,G_{321}(e)\sin(-\omega)}{-\dot{\omega}} (1-e^2)^{\frac{1}{2}} \right).$$

As $F_{301} = \frac{15}{16}\sin^3 i - \frac{3}{4}\sin i = \frac{3}{4}\sin i\left(\frac{5}{4}\sin^2 i - 1\right)$

and $F_{302} = - F_{301}$,
while

$$G_{31(-1)} = G_{321} = e(1-e^2)^{-5/2}$$

from table 8.2,

$$\Delta e_{30} = \frac{3n_\alpha \; a_e^3 \sqrt{7}\,\bar{C}_{30}(1 - \frac{5}{4}\sin^2 i)}{2(1-e^2)^2} \frac{\sin\omega}{\dot{\omega}} .$$

Similar expressions can be obtained for the other orbital elements. It should be noted that higher degree zonal harmonics have, in addition to the term of wavelength 2π, those of shorter period. For example,the harmonic of degree 6 will give rise to secular variations ($p=3$) as well as periodic variations of wavelength $\frac{1}{2}(2\pi)$ when $p=2$ and $\frac{1}{4}(2\pi)$ when $p=1$. Similarly, the zonal of degree 9 will have harmonics of wavelength 2π ($p=4$), $\frac{1}{3}\times 2\pi$ ($p=3$), $\frac{1}{5}\times 2\pi$ ($p=2$) and $\frac{1}{7}\times 2\pi$ ($p=1$).

The long period terms in the tracking of the perturbations of satellite orbits are therefore analysed by a series of terms which are harmonic in the argument of perigee . For detailed formulae see (Mueller 1964,pp.191 et seq.; Gaposchkin 1966,p.129 et seq.).

The evaluation of tesseral harmonics

Tesseral harmonics are relatively more difficult to determine than zonals as the effect of such terms is cancelled m times per day. A special case does arise however in the case of those orbits where

$$\dot{\alpha} = (n-2p)\dot{\omega} + (n-2p+q)\dot{M} + m(\dot{\Omega}-\dot{\theta}) \rightarrow 0$$

when the tesseral harmonic could, under the right conditions, make significant contributions to long period perturbations due to resonance effects which are due to both the interaction between the slower changing elements $\dot{\Omega}$ and $\dot{\omega}$, together with that between the more rapidly changing quantities $\dot{M}$ and $\dot{\theta}$. Thus the conditions for the existence of resonance could be expressed as

$$(n-2p)\dot{\omega} \rightarrow m\dot{\Omega} \qquad \text{and} \qquad (n-2p+q)\dot{M} - m\dot{\theta} \rightarrow C .$$

The latter effect is the dominant one and the major contributions are made by the terms obtained when $q = 0$, as can be seen from table 8.2. Thus resonance occurs when

$$\frac{m}{n-2p} \rightarrow \frac{\dot{\theta}}{\dot{M}} = \text{Number of revolutions per day.}$$

Integral terms are obtained when

$$m = n - 2p \qquad \text{or} \qquad p = \frac{n-m}{2} \qquad \ldots(8.115)$$

and hence when $(n-m)$ is even. Further, since the effect of all harmonics is scaled by the term $(a_e/a)^n$, the terms which dominate would be those with small values of n. For example, in the case of a satellite with an orbital period of 90 minutes (i.e., 14 orbits per day), the principal resonance effect is obtained when

$$p = \frac{n-m}{2} = 0,\ 1 \text{ or } 2 \qquad \ldots(8.116)$$

with secondary effects when $\frac{n-m}{2} = 1,\ 2 \text{ or } 3.$

The effect of tesseral harmonics on long period effects can therefore be considered to be negligible except in instances where resonance occurs. In the case of a 24 hour satellite, the dominant resonant terms are C_{22}, S_{22} with lesser contributions from C_{42}, S_{42} and C_{44}, S_{44}.

Resonant terms cannot be evaluated through equation 8.111 as $\dot{\alpha} = 0$. Their effect is most easily obtained from the Lagrangian equations expressed in terms of the general series for V_d in terms of the Keplerian elements c_i through equation 8.108. The differentiation of these equations with respect to t all give expressions of the form

$$\ddot{c}_{i_{nmpq}} = o\{\bar{C}_{nm}\,\dot{\alpha}\} + o\{\bar{C}^2_{nm}\} \qquad \ldots(8.117),$$

except that for $\ddot{M}$ which is the only Keplerian element to exhibit a significant acceleration, given by

$$\ddot{M}_{nmpq} = -\frac{3}{2}\frac{n_\alpha}{a}\dot{a}_{nmpq} + o\{\overline{C}_{nm}\dot{\alpha}\} + o\{\overline{C}^{\,2}_{nm}\} \quad \ldots(8.118).$$

The use of equation 8.108a gives

$$\ddot{M}_{nmpq} = -3(n-2p+q)n_\alpha \left(\frac{a_e}{a}\right)^n F_{nmp}(i)\, G_{npq}(e)\, S'_{nmpq}(\Omega,\omega,\theta,M) \quad \ldots(8.119).$$

Douglas(1967,figure 10) gives examples of $\ddot{M}$ for different values of n,m,p,q in the cases of 24 hour orbits (e=0.8) and a 2.4 hour orbit (e=0.2) and shows that the resulting along track variations for the former is about the same as the latter, in angular units, though due to different harmonics. The largest angular accelerations are of the order of 4×10^{-9} radians per 180 minutes which, if held constant over a significant period, can accumulate to approximately 7° in 100 days. In comparison, secular variations are of the order of 360° in 100 days. The magnitude of the resonance effect is however, a detectable quantity, though the separation of the various constituent resonant effects is a complex problem. It is doubtful whether any but the dominant resonant terms can be obtained for any particular period.

A harmonic of geophysical significance is C_{22} which can be seen from equation 7.15 to be related to the difference between the moments of inertia with respect to two mutually perpendicular axes in the equatorial plane. The resonant effect is due to

$n=2,\ m=2,\ p=0, q=0$

and can be obtained from the study of orbital perturbations of 24 hour satellites. The longitude of such a satellite is given by

$$\lambda = \omega + M + \Omega - \theta \quad \ldots(8.120)$$

as the orbit is equatorial. The acceleration of the satellite in longitude $\ddot{\lambda}_c$ is computed according to the relation

$$\ddot{\lambda}_c = \ddot{M} \overset{(8.117),(8.118)}{=} 3\, n_\alpha \sum_{\substack{n-m\\ \text{even}}} m \left(\frac{a_e}{a}\right)^n F_{nmp}(i)\, G_{npo}(e)$$

$$\left(C_{nm}\sin\{m(\Omega+M+\omega-\theta)\} - S_{nm}\cos\{m(\Omega+M+\omega-\theta)\}\right) \quad \ldots(8.121),$$

where $p = \frac{1}{2}(n-m)$.

This computed value $\ddot{\lambda}_c$ of the acceleration is compared with the observed acceleration $\ddot{\lambda}_o$ which is obtained from the variation in the Keplerian elements computed over the period and the coefficients deduced from the resulting observation equations. For a detailed treatment see (Kaula 1966,pp.49 et seq; Wagner 1967,pp.161 et seq; Douglas 1967,pp.189 et seq).

Tesseral harmonics from ranging to satellites

This method for the determination of the tesseral harmonics calls for the ranging to be performed over periods which are short enough to enable an accurate sample of the longitude dependant variations of the earth's gravitational field to be obtained. Various systems are available for ranging to satellites. The earlier interferometric and Secor methods are described in (Mueller 1964, p.269 et seq). The results obtained have been of dubious value in studying the earth's gravitational field (e.g., Kaula 1965,p.3) when compared with the results obtained by the use of Doppler techniques which are at present used on a world wide basis in the US Navy Satellite Doppler system known as *Tranet* which provides a global navigation system (Yionoulis 1967).

The Doppler principle is based on a continuous unmodulated transmission of mixed frequency f_o from a satellite which, on being received at a ground station, exhibits a frequency change df due to the velocity of the satellite relative to the latter. The received frequency f is given by

$$df = f - f_o = \frac{f_o \dot{r}_t}{c} ,$$

where r_t is the station-to-satellite distance, $\dot{r}_t$ the radial velocity and c the velocity of electromagnetic waves. Thus

$$\dot{r}_t = \frac{c\, df}{f_o} \qquad \ldots(8.122).$$

$\dot{r}_t = 0$ when the satellite is at its closest point, as shown in figure 8.5, when the orbital velocity is perpendicular to the topocentric distance vector. The distance r_o corresponding to this case is the required range to the satellite, being related to the value r_t of the satellite-to-station distance at any other instant, an interval of time t later, by the relation

$$r_t = [r_o^2 + v^2t^2]^{\frac{1}{2}} \qquad \ldots(8.123) ,$$

where v is the along-track velocity of the satellite. Differentiation with respect to time gives

$$2 r_t \dot{r}_t = 2 v^2 t .$$

On squaring and replacing r_t from equation 8.123,

$$\dot{r}_t^2 \{r_o^2 + v^2t^2\} = v^4t^2$$

or

$$\frac{r_o^2 + v^2t^2}{v^4t^2} = \frac{1}{\dot{r}_t^2}$$

which can be written as

$$\frac{r_o^2}{v^4} + \frac{t^2}{v^2} \overset{(8.122)}{=} \frac{t^2}{\left(\frac{c\, df}{f_o}\right)^2}$$

being an equation of the form

$$\left(\frac{t}{df}\right)^2 = \frac{c^2}{f_o^2}\frac{r_o^2}{V^4} + \frac{c^2}{f_o^2 V^2}t^2 \qquad \ldots(8.123).$$

This represents a straight line on treating the observed quantities $(t/df)^2$ and t^2 as the variables when the intercept gives the required minimum distance r_o.

These distances r_o are usually obtained from about 200 readings of df and t (ibid,p.128) and grouped into 24 - 48 hour time spans. A "best fit" orbit is determined by adjusting the values of r_o as determined at a number of observing stations onto a model provided by tracking station co-ordinates based on previous solutions, together with an adequate representation of all other forces acting on the satellite. The result is a mean Keplerian ellipse representative of a 24 hour span. The tesseral harmonics are studied by the analysis of data available within this period.

A local three dimensional Cartesian co-ordinate system x_i is chosen at each observing station associated with the determination of the mean orbit, the x_3 axis being coincident with the geocentric radius to the satellite at its position of minimum topocentric distance, the x_1 axis defining the direction of motion of the satellite at right angles to x_3, the x_2 axis completing a right-handed mutually perpendicular system. The departures of individual observations from the mean orbit are used to study the short period orbital perturbations. The general principles are outlined in (Guier & Newton 1965; Yionoulis 1967). The solution can be summarised as follows.

The discrepancies between observations and the model of the mean orbit are due to

(i) errors in the model ;

(ii) errors in the measurement of the Doppler effect df;

(iii) the perturbations in the gravity field;

and (iv) the variable effect of errors in the model used to represent the earth's atmosphere.

The satellite's differential position components (i.e., between two adjacent positions in orbit) can be expressed by three displacements δr, $\delta \ell$, δz radially, tangentially and perpendicular to the plane of the orbit. These coincide with the unit vectors u_n, u_t and u_θ in figure 8.1. Consequently equation 8.39 can be expressed in the form

$$\begin{aligned} \ddot{\delta r} - 3n_\alpha\,\delta r - 2n_\alpha\,\dot{\delta \ell} &= F_r + o\{f\,\delta r\} \\ \ddot{\delta \ell} + 2n_\alpha\,\dot{\delta r} &= F_\ell + o\{f\,\delta r\} \ldots(8.124), \\ \ddot{\delta z} + n_\alpha\,\delta z &= F_z + o\{f\,\delta r\} \end{aligned}$$

where F_r, F_ℓ and F_z are the respective forces acting on the satellite. Significant perturbations occur due to those harmonics which have wavelengths corresponding to the orbital frequency of the satellite as the station position was based on a mean orbit and the station-to-satellite vectors at the time of closest approach.

The errors arising from the departure of the mean orbit from the instantaneous orbit are eliminated by the adoption of a model relative to some arbitrary epoch t_o of the form

$$\delta r = A_1 + A_2[n_\alpha(t-t_o)] + A_3 \cos[n_\alpha(t-t_o)] + A_4 \sin[n_\alpha(t-t_0)] \quad \ldots(8.125),$$

with similar expressions for $\delta\ell$, δz, noting that one or more of the quantities A which define the model, may be zero. Normally A_1 and A_2 are put equal to zero for studies of a single pass. This procedure enables the positional perturbations of the satellite to be expressed in terms of the topocentric Cartesian co-ordinate system x_i. The slant range vector TS in figure 8.5 can also be expressed in this co-ordinate system. Once these have been allowed for, the residuals are next separated into components which are either symmetrical or asymmetrical with respect to the minimum range r_o.

The principle contribution to the symmetric terms are errors in the station co-ordinates and the parameters defining the orbit. The minimisation of symmetrical contributions will therefore provide one relationship for improving station positions and satellite orbital parameters. A second set of relations is provided by the asymmetric terms as they are dependent on slant range only. The model so defined is related to the representation of the disturbing potential given by equation 7.24 by adopting a relation of the form

$$V_{d_{nm}} = \bar{C}_{nm} \Re(\psi_{nm}) + \bar{S}_{nm} \Im(\psi_{nm}) \quad ,$$

where $\Im(X)$ refers to the non-real part of the function X, ψ_{nm} being given by the relations

$$\psi_{nm} = U_{nm} + i\, V_{nm} \quad ,$$

$$U_{nm} = \frac{1}{r^{n+1}} \bar{P}_{nm}(\sin\phi) \cos m\lambda$$

and

$$V_{nm} = \frac{1}{r^{n+1}} \bar{P}_{nm}(\sin\phi) \sin m\lambda \quad .$$

The solution which is obtained from the expanded forms of the above relations follows lines similar to those given in detail in section 8.2*(iii)*. The final expressions obtained relate the parameters defining the earth's gravitational field to the displacements δr, $\delta\ell$, δz and hence the model adopted for the definition of the station-satellite system. The minimisation of the data residuals gives

(i) improved station co-ordinates;

(ii) the harmonic coefficients $\overline{C}_{nm}$, $\overline{S}_{nm}$;

and (iii) six orbital parameters associated with each span of data.

It is necessary to have observations from as many tracking stations as possible in order that the harmonic coefficients may be reliably determined. It should be noted that the Doppler system can operate 24 hours a day and hence affords a means for the determination of tesseral harmonics without recourse to the study of resonance effects.

Notes

(i) Table 8.3 gives values of normalised harmonic coefficients to (8,8) in the representation of V_d through equation 7.24. The two solutions given are those by Kaula(1966a,p.5311) and Rapp (1968,p.11), obtained by the combination of satellite data and surface gravimetry as described in section 7.3. Kaula's data is referred to the International Gravity Formula which eliminates the major contribution due to the $\overline{C}_{20}$ term

(ii) It has been indicated by Kaula (1966c,p.4379) that a good rule of thumb for fixing the magnitude of the normalised coefficients upto degree 15 is

$$\overline{C}_{nm}, \overline{S}_{nm} \doteq \frac{10^{-5}}{n^2} .$$

(iii) In practice, it is only those harmonics of long wavelength which are of interest in geometrical solutions from gravity at the surface of the earth. The accuracy of these terms is an important factor in deciding whether gravity data can be successfully used for earth space studies (Mather 1970b,p.117). It is fortunate that these critical harmonics are the ones most easily determined from satellite orbital analysis as the dominant resonant effects occur as a result of the appropriate low degree term in view of the fact that the strength of the contribution declines as the inverse power of n.

(iv) The resonant orbit of greatest interest is the synchronous one having a period of 24 hours, the dominant effect being obtained when

$$p = 2 \quad ; \quad q = 0$$

as

$$G_{220}(e) = 1 - \frac{5}{2} e^2 + o\{e^4\}$$

and

$$\left(\frac{a_e}{a}\right) \doteq \frac{1}{6.6}$$

It is therefore resonable to conclude that if satellites with suitable periods and a variety of orbital inclinations were available, it would be possible to obtain those long wave-length harmonics which are of use in the extension of the surface gravity field to unsurveyed areas using the techniques outlined in section 7.3,from the perturbations

Order	0		1		2		3		4		5		6		7		8	
Coeff	Kaula	Rapp	Kaula	Rapp	Kaula	Rapp	Kaula	Rapp	Kaula	Rapp	Kaula	Rapp	Kaula	Rapp	Kaula	Rapp	Kaula	Rapp
C_{1m}	0.000	*0.000*	0.000	*0.000*														
S_{1m}	0.000	*0.000*	0.000	*0.000*														
C_{2m}	4.472	*-484.18*	0.018	*0.000*	2.404	*2.371*												
S_{2m}			0.054	*0.000*	-1.341	*0.055*												
C_{3m}	0.913	*0.937*	1.794	*1.855*	0.819	*0.713*	0.569	*0.633*										
S_{3m}			0.240	*0.243*	-0.731	*-0.589*	1.442	*1.523*										
C_{4m}	0.248	*0.552*	-0.565	*-0.551*	0.355	*0.297*	0.933	*0.873*	-0.040	*0.094*								
S_{4m}			-0.425	*-0.447*	0.550	*0.584*	-0.203	*-0.198*	0.314	*0.274*								
C_{5m}	0.047	*0.050*	-0.046	*-0.082*	0.474	*0.523*	-0.329	*-0.356*	-0.010	*-0.049*	0.082	*0.086*						
S_{5m}			-0.005	*-0.063*	-0.291	*-0.213*	-0.011	*0.027*	-0.117	*0.074*	-0.478	*-0.569*						
C_{6m}	-0.144	*-0.137*	-0.091	*-0.065*	0.005	*0.028*	0.099	*-0.054*	-0.158	*-0.025*	-0.127	*-0.291*	-0.042	*-0.009*				
S_{6m}			0.179	*-0.019*	-0.304	*-0.284*	0.075	*0.060*	-0.449	*-0.412*	-0.608	*-0.451*	-0.281	*-0.184*				
C_{7m}	0.150	*0.070*	0.062	*0.120*	0.306	*0.307*	0.017	*0.180*	-0.263	*-0.193*	0.033	*0.070*	-0.172	*-0.165*	0.030	*0.068*		
S_{7m}			0.087	*0.106*	0.039	*0.137*	-0.017	*0.009*	0.059	*-0.091*	-0.017	*0.036*	0.230	*0.092*	0.057	*-0.030*		
C_{8m}	-0.113	*0.046*	0.022	*-0.043*	0.071	*0.040*	0.037	*-0.004*	0.005	*-0.092*	-0.090	*-0.063*	0.033	*-0.107*	0.072	*0.026*	-0.065	*-0.104*
S_{8m}			-0.083	*0.025*	0.005	*0.098*	-0.038	*0.038*	-0.028	*0.019*	-0.225	*0.082*	0.197	*0.287*	0.003	*0.030*	-0.108	*-0.017*

Table 8.3

Normalised harmonic coefficients of the earth's gravitational potential V_d upto (8,8) observed in the orbits.

9. VARIATIONS IN GRAVITY AT THE SURFACE OF THE EARTH

9.1 Long wave variations

The most comprehensive studies of the earth's gravity field on a global scale are those by Kaula (1959; 1963; 1966a). Other investigations of significance were made earlier by de Graaff Hunter (1935) and Hirvonen (1956). The primary interest in studying the characteristics of the earth's gravity field is for the purpose of predicting values for the representation of those regions which have not been surveyed. There are numerous statistical functions which are available for the assessment of the variability of the field. The initial studies were due to de Graaff-Hunter who introduced the concept of the *error of representation* for a region.

If the values of the gravity anomaly Δg_i are known at n points in a region, the error of representation ε_s was defined by the relation (de Graaff-Hunter 1935,p.407)

$$\varepsilon_s = M\{\Delta g_i - M\{\Delta g_i\}\} \qquad \ldots(9.1),$$

where

$M\{X\}$ = the mean value of X.

He also obtained a rule of thumb formula for the error of representation from a study of the data available to him, as

$$\varepsilon_s = \pm\, 11.3 \sqrt{s} \text{ mgal} \qquad \ldots(9.2)$$

for a square whose side was of length s°.

Hirvonen (1956)suggested the alternate formula for ε_s given by

$$\varepsilon_s^2 = \frac{M\{(\Delta g_i - M\{\Delta g_i\})^2\}}{n} \qquad \ldots(9.3).$$

The quantity $M\{\Delta g_i\}$ in equations 9.1 to 9.3 is the mean value of the observed gravity anomalies Δg_i within a specified region while the outer mean is taken over all such available regions.

The error of representation can be interpreted as follows. A single observed gravity anomaly is likely to represent the mean value for a specified region with a standard deviation equal to the error of representation ε_s. Table 9.1 gives Hirvonen's values

Square size (s)	Error of representation (ε_s)± mgal Eurasia	South Australia
0.1°× 0.1°	2.8	3.0
0.5°× 0.5°	9.0	10.1
1°× 1°	12.7	13.5
2°× 2°	17.6	17.7
5°× 5°	23.1	-
10°× 10°	24.9	-
30°× 30°	26.6	-
World wide	28.0	-

Table 9.1

The error of representation ε_s

for ε_s based on a sample of data obtained principally from Eurasia (ibid,p.3) while the values for South Australia were obtained by Mather (1967,p.131). The square size s is the basic unit of area at the surface of the earth, bounded by meridians and parallels.

Kaula's first study of significance dealt with the analysis of the free air anomaly field for the harmonic coefficients upto (8,8). One of the functions he used in the analysis was similar to the function G_s , called the root mean square(rms) anomaly by Hirvonen and defined for a set of squares of a given size by the relation

$$G_s = M\{(M\{\Delta g_i\})^2\} \qquad \ldots(9.4)$$

G_s is obtained by determining the mean anomaly of all gravity readings in each of a series of regions of a given square size s^o on the surface of the earth and computing the mean square of these values. G_o is the rms of individual gravity anomalies, while $G_{0.1}, G_{\frac{1}{2}}, G_1$ etc. represent the variability between the area means of 0.1°, ½° and 1° squares as the rms value. The comparison between figures obtained by Hirvonen (1956,p.3) and Kaula (1959, pp.59-65) is given in table 9.2.

The following conclusions can be drawn in the light of the information contained in tables 9.1 and 9.2, bearing in mind that continuous gravity coverage at the surface of the earth is unlikely to be a practical possibility.

(i) The larger the area a single reading has to represent, the greater the value of ε_s and consequently the weaker the representation. As errors in the measurement of gravity are of

Square size (s)	Root Mean Square anomaly (G_s) ± mgal Hirvonen	Kaula
0°	28.0	34.7
1°	25.0	27.8
2°	21.8	22.7
5°	15.8	18.7
10°	12.8	15.7
30°	8.7	9.2

Table 9.2

The mean value of the rms anomaly for squares of different sizes

the order of 0.1 - 0.2 mgal on land, using gravimeters and as the average gravity station elevation is seldom as accurate as one would wish, a desirable value of ε_s would therefore be ±1 mgal, on the basis that the gravity station elevation is correct to ±3 m. The use of equation 9.2 indicates that this would call for a representation of the gravity field at the corners of a 0.01° square grid. The limited amount of data available at present indicate that a 0.1° grid is a practical possibility , this being equivalent to an ε_s value of ±3 mgal. Such a system could be used in an appropriate manner such that errors which arise in the computation of the separation vector can be kept down to less than ±5 cm (Mather 1968c, Mather 1969,p.501).

(ii) The area means of larger regions have a smaller variability than those of smaller sub-divisions.

(iii) The rms anomalies obtained by Kaula are systematically larger than Hirvonen's values. This is due to Kaula having access to a much larger data set covering both continental and oceanic regions while Hirvonen's data was essentially continental.

Another interesting quantity computed by Kaula in the course of his studies of surface gravity was the degree variance of free air anomalies (Kaula 1959,pp.62&91) given by equation 4.39. It can be shown from equations 4.37 and 4.39, on writing

$$dS = d\psi \sin\psi ,$$

where ψ is the angular distance over which the covariance of gravity anomalies is to be estimated, that

$$\int_0^{\pi} \sigma_n^2 \, [p_{no}(\cos\psi)]^2 \sin\psi \, d\psi = \sum_{\psi=0}^{\pi} C(\psi) \, p_{no}(\cos\psi) \sin\psi \, \Delta\psi \quad ..(9.5).$$

On evaluation of the integral on the left of equation 9.5 using equation 4.8, it can be seen that

$$\sum_{i=1}^{k} C(\psi) \quad P_{no}(\cos\psi) \sin\psi \, \Delta\psi \; = \; \frac{2\,\sigma_n^2}{2n+1} \qquad \ldots(9.6),$$

where $k = \pi/\Delta\psi$ and

$$\psi = (i-\tfrac{1}{2})\Delta\psi \quad \text{and} \quad \sin[(i-\tfrac{1}{2}\Delta\psi)]\Delta\psi = \cos[(i-1)\Delta\psi] - \cos[i\Delta\psi] \qquad \ldots(9.7).$$

Equation 9.6 affords a practical means for the evaluation of σ_n^2. The values obtained by Kaula (ibid,p.62) are given in column 2 of table 9.3.

Degree n	σ_n^2 (mgal^2) from covariance	after harmonic analysis	from satellites alone	from combined data sets
2	7.3	1.1	7	7.1
3	43.6	11.1	31	30.4
4	29.8	13.9	17	16.2
5	10.5	6.6	16	12.3
6	24.2	10.4	27	14.5
7	2.8	4.2	48	9.4
8	22.7	5.6	32	6.7

Table 9.3

Degree variances of surface gravity anomalies

Column three gives values obtained direct from equation 4.39, while column four gives results obtained from satellites (Kaula 1965, table 2), the latter values being a composite set obtained from different solutions. Thus covariance analysis provides reasonable assessments of degree variances. The same cannot be said for the values obtained from the coefficients defined after spherical harmonic analysis of surface gravity data (column 3) following the use of Markov analysis for the prediction of values used in the representation of unsurveyed areas, when the degree variances are found to be *too small*. This indicates over-smoothing in the field extension process and is borne out by the fact that all geoids produced by the spherical harmonic analysis of surface gravity data alone are over-smoothened(e.g., Kaula 1959,p.108).

The fifth column in table 9.3 gives the degree variances for a combined data set obtained by the use of the principles outlined in section 7.3 and obtained by Rapp in the course of a (14,14) analysis (Rapp 1968,p.30). A comparison of tables 9.2 and 9.3 indicates that an analysis to (8,8) is approximately equivalent to representation by $30^\circ \times 30^\circ$ area means. This could also have been inferred from section 4.7 as such units of surface area correspond to (6,6) analysis, if it is assumed that adjacent

$30^\circ \times 30^\circ$ area means are uncorrelated. This is thought to be the case (e.g., Hirvonen 1956,p.1).

It must therefore be concluded that any combined data set referred to above, even if obtained from a spherical harmonic analysis to (14,14) where the total of the degree variances approach 150, have a variability equivalent to that of a set of $10^\circ \times 10^\circ$ square means. It therefore follows that the values of smaller area means (e.g., $5^\circ \times 5^\circ$ squares) obtained from such a data set have position dependent deviations from true surface values as a consequence of oversmoothing. This correlation of error with position is restricted to extents less than the basic area (a ten degree square in the case of a 14,14 analysis). Such effects are not of consequence in the representation of distant zones but can cause significant error when the data set is used to define the earth's gravity field in the close vicinity of any point whose gravitational characteristics are being investigated.

Such a procedure was used for a series of calculations of the geoid for Australia and the results proved to be of adequate accuracy when the combined data set was used to represent regions which were more than 20° away from the point of computation(Mather 1969; Mather 1970c).

9.2 Local variations in the gravity field

The gravity field over limited regions cannot be satisfactorily studied by the use of spherical harmonic functions alone due to the excessive number of higher degree terms required to provide an adequate representation. If this is not achieved (i.e., n too low), the errors in adjacent values will be correlated to the extent that any quantity computed from this data set will be significantly incorrect. No local variations can ever be adequately mapped unless sufficient observational data is available. Thus local fields cannot be defined satisfactorily by prediction methods alone , as the free air anomaly field is affected not only by variations in the topography but also by anomalies beneath the earth's surface, especially at very small depths, as frequently occurs in Australia, thus making prediction a very complex task.

It is well known that gravity values in mountainous regions undergo a two-fold variation. The first is the decrease in observed gravity with elevation as a consequence of the increased distance from the centre of mass, being given by equation 6.75. The second is an increase due to the occurrence of the mass of the mountain directly below it. This can be very crudely represented by the Bouguer reduction c_b (e.g., Heiskanen & Moritz 1967,p.130) which represents the topography as an infinite slab of thickness equal to that of the elevation of the station. Its magnitude, on assuming a density of 2.67 gm cm^{-3} for the material of the slab, is given by

$$c_b = 2\pi k\rho h = 0.1118\, h^{(m)} \text{ mgal .}$$

This model is improved in practice by allowing for the departure of the earth's topography from an infinite slab. The resulting correction c_t is called the *terrain correction* and has the effect of always reducing the magnitude of c_b. c_t is, for all practical purposes, zero if the surrounding topography within 150 km of the point considered, is planar. Its magnitude is dependent entirely on the variability of topographical elevations in the vicinity of the point, the local regions making larger contributions, all other factors being equal. c_t can be calculated by formulae of the type given by Mather (1968b,p.521). These effects can be as large as 50 mgal (Heiskanen & Vening Meinesz 1958,p.154). (note :- $c_b > 300$ mgal for $h \simeq 3$ km).

Observed gravity is also influenced by the gravitational effect of the mass deficiency due to the existence of isostatic compensation beneath elevated continental regions (ibid,pp.124 et seq) which is generally held to satisfy the model postulated initially by Airy.

Gravity anomalies at the surface of the earth(i.e., the so called "free air" anomalies) are thus correlated with topography as the decreases in observed gravity with increase of elevation do not equal the reverse effect due to the close topography. Consequently the correlation is a positive one, the general variations over very limited regions being expected to be of the form (Uotila 1960)

$$\Delta g = \Delta g_o + 0.1118\,\Delta h^{(m)} \quad \text{mgal}$$

where Δh is the difference in elevation between the point at which the gravity anomaly Δg_o is known and that at which Δg is observed. This positive correlation between surface gravity anomalies and elevation is clearly illustrated by Kaula in his study of surface gravity (Kaula 1959,p.5), the results being reproduced in table 9.4. He confined his analysis to $1^o \times 1^o$ square means and classified them according to elevation. The only serious departures from correlated behaviour occur at continental margins where not only do large topographical mass changes occur but observed gravity is also affected by the rapid changes in the depth of the crust-mantle boundary.

Kaula also established a transition count of the rate of change of the $1^o \times 1^o$ mean surface gravity anomaly with height as

$h(\times 10^{-3})$ * ft	% of earth's surface	Mean anomaly (mgal)	% observed
< -12	6	- 49.2	2
-11	8	- 14.6	5
-10	10	- 12.5	10
-9	12	- 9.8	7
-8,-7	16	- 8.3	10
-6,-5	7	- 3.6	10
-4,-3	3	- 1.2	25
-2,-1	4	12.1	23
0	11	3.7	30
1	6	0.9	51
2, 3	10	4.1	21
4, 5	3	16.9	24
> 6	4	40.2	20

* h 0.62d,
where d=depths

Table 9.4

Correlation between one degree area means of the free air anomaly and elevation (after Kaula)

$$M\{(\frac{\delta\Delta g}{\delta h})_{1^o}\} = +0.00851 \text{ mgal ft}^{-1} = +0.028 \text{ mgal met}^{-1}.$$

Such correlation should be recognised as providing broad indications of overall trends. Kaula also shows sufficient figures to illustrate the existence of a variety of gravity anomaly values at any given elevation and/or depth.

Thus predicted values of the gravity anomaly field can only be used, with any degree of confidence, for the representation of regions distant from the area in which space vectors are to be defined. The near earth satellite should therefore be considered to be a measuring instrument which affords a means for the definition of the long wavelength variations of the earth's gravity anomaly field. These harmonic terms, on combination with all available surface gravity data, provide a representation of the gravity anomaly field which is equivalent to large area means (approximately ten degree squares). The resulting data set can be used successfully for the representation of the effect of distant regions in geodetic computations till such time as an adequate surface gravity coverage becomes available. This cannot be done in the case of regions within 20° of any point of computation as the application of this data to smaller intervals gives rise to systematic errors which seriously affect any space vector determinations.

Gravimetric computations cannot be considered to be free from large scale systematic error unless the field within 20° of the point of computation is either completely represented by surface gravimetry or alternately, the gaps exist by design and not chance (Mather 1967,p.133). This calls for a planned gravity survey where stations are sited after a study of the topography of the region. Under such circumstances it is possible to keep systematic errors to within acceptable limits.

9.3 Some probable sources of density anomalies and geoidal undulations

The factors which cause the gravity field of the earth to deviate from that of a reference ellipsoid containing a symmetrical distribution of matter are many, some of which are well known while others are the subject of speculation. The nature of the geoid itself can be illustrated clearly by Mather's study of the free air geoid for Australia (Mather 1970c) which shows the geoid referred to a geocentric reference ellipsoid (figure 7) and also to the local astro-geodetic datum (figure 9 in reference) after allowing for the translation of axes. In the latter case, the datum is approximately parallel to the mean geoid slope across the region. Consequently, the near uniform grade of approximately 5 arc sec exhibited by the geoid across the continent in figure 7 is converted to a chart of highs and lows in the latter figure. Some of these maxima and minima are correlated with topography to the extent of a few metres while others are not. It should be noted that Australia is a relatively flat region with mountainous regions confined to the south-east. In these latter regions, the geoid does appear to rise by about 5 metres due to the effect of the mountains alone whereas other maxima of similar magnitude occur in other regions which exhibit no correlation with topography.

It can therefore be concluded that mass anomalies do occur in the earth's crust, very close to the surface, having magnitudes as great as the total gravitational effect of topography and its isostatic compensation. The Australian investigation referred to above showed that mass anomalies of sufficient extent to influence the gravity field over 10° of the earth's surface are by no means rare. These results appear to be in agreement with the observations of others who conclude from correlation studies between seismic and gravity data that compensation is not complete in the upper mantle and that the crust and upper mantle are regions of strong density anomalies and stresses (Toksoz,Arkani-Hamed & Knight 1969,p.3768).

The long wave effects of surface gravity anomalies due to harmonics less than or equal to 6, as computed from satellite orbital analysis have no correlation with topography. It may therefore be concluded that some form of isostasy prevails with a compensation which may not be complete, in view of the numerous deviations which occur from the isostatic model (Heiskanen & Vening Meinesz 1958,pp.311 et seq). These long wave effects due to harmonics of degree less than six, have been attributed by Hide and Malin to possible undulations of the liquid core-solid mantle interface (Hide & Malin 1970,p.605). The existence of such undulations of the order of 4 km is a distinct possibility on the basis of recent seismic studies of the lower mantle which indicate lateral density variations of the order of 4×10^{-2} gm cm^{-3}. These are reported to be an order of magnitude larger than anomalies detected by the orbital perturbations of artificial earth satellites. It is estimated that a 1 km undulation in the core mantle boundary is equivalent to 0.45×10^{6} gm cm^{-3} (Toksoz,Arkani-Hamed & Knight 1969,p.3765) and hence can completely account for mass anomalies in the lower mantle if the undulation were of the order of 4 km.

Hide and Malin also produce evidence for such undulations. They note that the removal of the secular westward drift of the geomagnetic field by 0.23° longitude per year gives a residual field which is significantly correlated with the low degree harmonics of the earth's gravitational field if the former is rotated eastward through 160°(Hide & Malin 1970,p.606). They further argue that if this secular drift is due to strong magneto-hydrodynamic interactions between motions in the liquid core and the undulations of the core-mantle boundary, such an occurrence should have taken place ($160^{\circ}/0.23^{\circ}$ $\doteqdot$) 600 years ago.

The second theory advanced for the long wavelength harmonics in the earth's gravitational field is the existence of convection currents in the mantle (e.g., Vening Meinesz 1964,p.57). The possibility of such currents existing is indicated by the phenomenon described as sea floor spreading (e.g.,Heirtzler 1968) together with some indication of the non-permanence of earth space co-ordinates. Munk and MacDonald (1960,p.250) have examined the possibilities from the viewpoint of position of the earth's pole in relation to earth space and reach the conclusion that the position of the pole would adjust to the location indicated by the principal axis of greatest moment of inertia, defined by the mass inhomogeneities in the mantle at the time. Evidence for the position of the pole in the past is based on paleomagnetism and therefore assumes some degree of coincidence between the magnetic and rotational poles.

The convection currents, if they exist, are as a consequence of convection cells which are not necessarily symmetrical about the rotation axis (Goldeich & Toomre 1969,p.2556), will result in the earth's gravitational field exhibiting characteristics which are possibly random. Consequently, low degree harmonics need not be confined to the zonal terms, etc. It has been observed that the residual in $\bar{C}_{20}$ after the removal of the effect of the flattening based on the best fitting ellipsoid of revolution ($\bar{C}_{20}$ = -484.2×10^{-6}) is'

$$d\bar{C}_{20} = -4.7$$

which is of the same order of magnitude as

$$\bar{C}_{22} \overset{(7.20)}{=} \frac{1}{4}\,\frac{I_{222} - I_{211}}{M a^2} \overset{\text{table } 9.3}{=} 2.4.$$

This leads to the possibility that some form of convection can be the mechanism for the occurrence of non-zero low degree harmonics which cannot be due to the topography and its isostatic compensation. One would expect such convection to be correlated with heat flow at the surface. Such is not the case as there is no doubt surface heat flow measurements exhibit no significant low degree harmonics, indicating that whatever flow occurs is due to either near surface crustal or upper mantle effects. In addition, researchers have been unable to establish any correlation between surface gravity anomalies and heat flow data (Toksoz et al 1969, pp.3765-8).

Thus the main conclusions to be drawn are as follows.

Low degree harmonics in the representation of the gravity anomaly field are due to mass anomalies in the mantle. Such anomalies may extend down to the liquid core-mantle interface which could have undulations as great as 4 km. It is also likely that partial compensation occurs and the long wave harmonics are merely the resul-tant anomaly. Any evidence for such harmonics being due to convection currents in the upper mantle is principally based on the observational evidence of sea floor spreading and not on heat flow data analysis.

Higher degree harmonics are a consequence of local and regional variations which are either crustal or due to the topography The magnitude of the effect of underground density anomalies are as great as those due to mountains on observed gravity, with similar extents of regionality. Such harmonics cannot be picked up by satellite orbital analysis as their magnitude is damped by the effect of whatever isostatic compensation may occur.

School of Surveying
University of New South Wales
Kensington, N.S.W. 2033
AUSTRALIA

REFERENCES

Ayres, F. 1962 *Matrices* Schaum, New York.

Bomford, G. 1962 *Geodesy* Oxford Univ. Press.

Brouwer, D & Clemence, G.M. 1961 *Methods of Celestial Mechanics* Academic Press, New York.

Bursa, M. 1969 Sur certaines relations entre les parametres de L'Ellipsoid Terrestre et du Champ de Gravite, en particulier par rapport au Systeme de Reference *Publ. Institut Geodesique de Recherches*, Praha.

Cook, A.H. 1959 The external Gravity Field of a Rotating Spheroid to order of e^3. *Geophys.J.R.astr.Soc.* 2,199-213.

de Graaff-Hunter, J. 1935 The Figure of the Earth from Gravity Observations and the Precision Obtainable *Proc.R.Soc.*A,234,397-431.

Douglas, B.C. 1967 Determination of Earth Potential Harmonics with Resonant Satellite Orbits *Les Methodes Dynamiques du Geodesie par Satellites* ,189-194, Centre Nationale Etudes Spatiale, Paris.

Ecker, E. 1970 *The Convergence of a series of Zonal Harmonics* Publ. Technische Universitat Berlin,Lehrstuhl fur Hohere Geodasie und Astronomie.

Eisenhart, L.P. 1960 *A Treatise on the Differential Geometry of Curves and Surfaces* Dover, New York.

Gaposchkin, E.M. 1966 Orbit Determination in *Geodetic Parameters for a 1966 Smithsonian Institution Standard Earth* Vol.1, 77-183, S.A.O. Spec. Rep. 200.

Gaposchkin, E.M. 1967 Determination of the Earth's gravity field by the use of optical observations in *Les Methodes Dynamiques du Geodesie par Satellites*, 91-126, Centre Nationale Etudes Spatiale, Paris.

Goldreich,P & Toomre, A. 1969 Some Remarks on Polar Wandering *J.geophys. Res.* 74, 2555-2567.

Guier, W.H. & Newton, R.R. 1965 The Earth's Gravity Field as Deduced from the Doppler Tracking of Five Satellites *J.geophys.Res.* 70,4613-4626.

Heiskanen, W.A. 1965 Present problems of Physical Geodesy *Suomal.Tiedeakat.Toim III* A,82.

Heiskanen, W.A. & Moritz, H. 1967 *Physical Geodesy* Freeman, London.

Heiskanen, W.A. & Vening Meinesz, F.A. 1958 *The Earth and its Gravity Field* McGraw-Hill, New York.

Hide, R & Malin, S.C. 1970 Novel Correlations between Global Features of the Earth's Gravitational and Magnetic Fields *Nature* 225,605-609.

Hirvonen, R.A. 1956 The Precision of the Gravimetric Determination of the Geoid *Trans.Am.geophys.Un.* 37,1-8.

Hirvonen, R.A. 1960 New theory of the Gravimetric Geodesy *Suomal.Tiedeakat.Toim III* A,56.

Hobson, E.W. 1965 *The Theory of Spherical and Ellipsoidal Harmonics* Chelsea, New York.

Hotine, M. 1969 *Mathematical Geodesy* ESSA Monograph No 2.

References (ctd)

Humphreys, D. 1945 *Intermediate Mechanics* Longmans, London.

I.A.G. Resolutions 1967 Resolutions adopted at the General Assembly, Lucerne, 1967 *Bull. geodes.* 86,367-383.

Jeffreys, H. 1941 The Determination of the Earth's Gravitational Field *Mon.Not.R.astr.Soc.* Geophys.Supl. 5,1-22.

Jeffreys, H. 1943 The Determination of the Earth's Gravitational Field (Second paper) *Mon.Not.R.astr.Soc.* Geophys.Supl. 5,55-66.

Jeffreys, H. 1948 The Figure of the earth and the Moon *Mon.Not.R.astr.Soc.* Geophys.Supl.5, 219-247.

Jeffreys, H. 1962 *The Earth* Cambridge University Press.

Jeffreys, H. & Jeffreys, B.S. 1962 *Methods of Mathematical Physics* Cambridge University Press.

Kaula, W.M. 1959 *The Statistical and Harmonic Analysis of Gravity* U.S.Army Map Service Tech. Rep. 24.

Kaula, W.M. 1961 Analysis of Gravitational and Geometric aspects of the Geodetic Utilisation of Satellites *Geophys.J.R.astr.Soc.* 5,104-133.

Kaula, W.M. 1962 Celestial Geodesy *Advances in Geophysics* 9, 191-293.

Kaula, W.M. 1963 Determination of the Earth's Gravitational Field *Rev.geophys.* 5, 83-107.

Kaula, W.M. 1965 Comparison and Combination of Satellite with other Results for Geodetic Parameters , Institute of Geophysics & Planetary Physics, Univ. of California, Los Angeles, Publication No 431.

Kaula, W.M. 1966a Tests and Combination of Satellite Determinations of the Gravity Field with Gravimetry *J.geophys.Res.* 71,5303-5316.

Kaula, W.M. 1966b *Theory of Satellite Geodesy* Blaisdell, Waltham Mass.

Kaula, W.M. 1966c Tesseral Harmonics of the Earth's Gravitational Field from Camera Tracking of Satellites *J.geophys.Res.*71,4377-4388.

Kaula, W.M. 1967 Theory of the Statistical Analysis of Data distributed on a sphere *Rev.geophys.*5,83-107.

Lambert, W.D. 1961 The Gravity Field of an Ellipsoid of Revolution as a Level Surface *Suomal.Tiedeakat.Toim.III* A,57.

Mather, R.S. 1967 Extersion of the Gravity Field in South Australia *Ost.Z.f.VermessWes.* 25,126-138.

Mather, R.S. 1968a *The Free Air Geoid in South Australia and its relation to the equipotential surfaces of the Earth's Gravitational Field* Univ. of NSW UNISURV Rep. 6.

Mather, R.S. 1968b The Free Air Geoid as a Solution of the Boundary Value Problem *Geophys.J.R.astr.Soc.* 16,515-530.

Mather, R.S. 1968c The Formula for Normal Gravity in Geodetic Calculations *Surv. Rev.* 150,341-348.

Mather, R.S. 1969 The Free Air Geoid for Australia. *Geophys.J.R.astr.Soc.* 18,499-515.

Mather, R.S. 1970a *The Theory and Geodetic Use of some Common Projections* Dept. of Surveying, Univ. of N.S.W.

Mather, R.S. 1970b *The Australian Geodetic Datum in Earth Space* Univ. of N.S.W. UNISURV Rep.19.

References (ctd)

Mather, R.S. 1970c The Geocentric Orientation Vector for the Australian Geodetic Datum *Geophys.J.R.astr.Soc.* 20 (in press).

Moritz, H. 1966 *Linear Solutions of the Geodetic Boundary Value Problem* Ohio State Univ. Dept. of Geod.Sci. Rep. 79.

Moritz, H. 1968 On the Computation of the Deflection of the Vertical *Boll.geofis.teor.appl.* 40,326-330.

Moritz, H. 1969 *Preliminary Computations for the Geodetic Reference System 1967* Symposium on European Triangulation, Paris.

Mueller, I.I. 1964 *Introduction to Satellite Geodesy* Ungar, New York.

Rapp, R.H. 1968 *Comparison of two methods for the Combination of Satellite and Gravimetric Data* Ohio State Univ. Dept. of Geod.Sci. Rep.113.

Stokes, G.G. 1849 On the variations of Gravity at the Surface of the Earth *Trans.Camb.Phil.Soc.* 8,672-695.

Toksoz, M.N., Arkani-Hamed, J & Knight, C.A. 1969 Geophysical Data and Long Wave Heterogeneities of the Earth's Mantle *J.geophys.Res.* 74,3751-3770.

Vening Meinesz, F.A. 1964 *The Earth's Crust and Mantle* Elsevier,Amsterdam.

Uotila, U.A. 1960 Investigations into the Gravity Field and Shape of the Earth *Suomal.Tiedeakat.Toim.III* A,55.

Wagner, C.A. 1967 The Use of Resonant Librating Orbits in Satellite Geodesy in *Les Methodes Dynamiques du Geodesie par Satellites*, 161-187, Centre National D'Etudes Spatiales, Paris.

Yionoulis, S.M. 1967 The Use of Satellite Doppler Data for Geodetic Purposes in *Les Methodes Dynamiques du Geodesie par Satellites*. 127-160, Centre National D'Etudes Spatiales, Paris.

Zhongolovich, I.D. 1952 *The External Gravity Field of the Earth and the Fundamental Constants involved* (in Russian), USSR Acad. Sci., Inst. Theor. Astron. Vol 3.

END