IN-46 CR
45160
33P.

Final Technical Report

NASA Grant NAG5-360

DEVELOPMENT OF HIGH ACCURACY AND RESOLUTION GEOID AND GRAVITY MAPS

For the Period 1 August 1983 - 31 March 1985

E. M. Gaposchkin

MIT Lincoln Laboratory
Lexington, MA 02173

(NASA-CR-179978) DEVELOPMENT OF HIGH ACCURACY AND RESOLUTION GEOID AND GRAVITY MAPS Final Technical Report, 1 Aug. 1983 - 31 Mar. 1985 (Massachusetts Inst. of Tech.) 33 p CSCL 55O G3/46

N87-13880

Unclas 43902

Satellite to Satellite Tracking to Map the Geoid

E.M.Gaposchkin

M.I.T. Lincoln Laboratory
Lexington Mass. 02173

Abstract

Precision Satellite to Satellite Tracking can be used to obtain high precision and resolution maps of the geoid. A method is demonstrated to use data in a limited region to map the geopotential at the satellite altitude. An inverse method is used to downward continue the potential to the earths surface. The method is designed for both satellites in the same low orbit.

Introduction

With the general acceptance of the need for improved global geoid and gravity data for both Solid Earth Geophysics and Physical Oceanography investigations, the fact that this data can only be provided with a close earth satellite measurement system, and the realization that at least one satellite measurement system is now able to provide observations with the required accuracy (NAS/NRC 1979 report), attention must be focused on the details of extracting the geoid and gravity data from these observations in an efficient way that will also be suitable for these investigations. The view taken here is that the Gravity Research Mission (GRM), a Satellite to Satellite Tracking Mission (SST) will satisfy these objectives. Generally we assume that the SST mission will have two surface force compensated satellites in similar low altitude orbits (160 to 180 km) separated by 300 to 1000 km and a tracking system with an accuracy of 1 micron/sec (1μ/sec = 10**-6 m/sec) for 10 second averages. This analysis concentrates on this generic mission.

The concept of the SST mission is simply to calculate the gravitatinal force acting on a spacecraft from changes in its measured velocity. The satellite itself is the sensor, and its velocity, the observable. This is pictured simplistically as follows. Consider a satellite in orbit around the earth approaching a region of excess mass. As the satellite approaches, it is accelerated toward the mass, and after passing the mass, it is decelerated. By measuring the time history of the velocity variation, an estimate of the magnitude and position of the mass excess can be deduced. Of course the actual situation is much more complex, for several reasons: the structure of the earth's mass distribution is very complicated, other forces act on the satellite, only one component of the satellite velocity is measured from another satellite, and the observations contain errors. In the SST concept, the second satellite could be very high, say in geosynchronous orbit (the high-low configuration). Alternatively, the second satellite could be in nearly the same orbit trailing the first, low satellite (the low-low configuration). In the low-low case, the second satellite would experience similar velocity changes, but at a later time.

SST tracking has been realized in a number of missions. The earliest example of doppler tracking a satellite from a point not on the body being studied is the mapping of the lunar gravity field by means of lunar orbiters tracked from the earth (Muller and Sjogren, 1968). Since this remarkable success, there have been notable doppler tracking experiments for earth satellites: the tracking of Geos-3 from ATS-6 geosynchronous satellite, the tracking of the Apollo Command Service Module (CSM) from the ATS-6, and the tracking of the Apollo CSM from a small satellite trailing it in the same orbit. The first two are examples of the high-low configuration, and the last is an example of the low-low configuration.

The classical method of determination of geopotential information using perturbation methods, and spherical harmonic representation could be employed (Gaposchkin, Kaula, Guire & Newton, Balmino, Lerch et al etc). However, the increased resolution makes this approach unattractive. For example, assuming a desired resolution of 1 degree, i.e. a spherical harmonic expansion to degree and order 360, requires the simultaneous recovery of almost 130000 coefficients, in addition to the orbital parameters. Direct recovery of gravity anomalies from SST data has also been investigated (eg Hajela 1977). Some view the problem as an estimation problem in a finite-dimension vector space of parameters (Vonbun et al., 1978; Schwarz, 1970). Others have treated the problem as mapping the acceleration field at the satellites altitude (Muller and Sjorgren, 1968; Marsh etal. 1977). Studies have also been done on mapping the acceleration or potential at the earth's surface (Rummel, 1975, Rummel et al, 1978). Some consideration has also been given to use of fourier techniques (Colombo, Kaula), and they show promise of significant economies in computer resources. The development presented here takes the view, that we can use the data directly to recover the desired quantity. It is analgous to an approach used by Muller & Sjogren 1968, and Sjogren 1976.

We propose that the SST mission be viewed as a mapping mission. That is, the result will be maps of the geoid or gravity, as contrasted to determination of spherical harmonic or fourier coefficients. Of course (digital) maps can be used to derive these coefficients. As a mapping mission, the computation can be done area by area. This avoids the enormous computer costs of simultaneously recovering all the geopotential information and the trajectories.

The process of mapping the SST signal to a geoid would have a number of steps:

1) With a chosen geopotential model, compute the range rate residuals from the best trajectory available. This need only be done once.

2) With a suitable algorithm, map the differential velocity (δv) to the anomalous potential (δT) at the satellite altitude.

3) Regularize the potential along the satellite tracks to the potential on a sphere near the satellite altitude.

4) Downward continue the potential from the reference sphere at the satellite altitude to the ellipsoid.

5) Correct the SST signal with the (revised) anomalous potential, and repeat from step 2.

The approach has a number of advantages. First, it separates the inferance of the geopotential from the downward continuation problem. One can evaluate the geopotential at satellite altitude independently from its use on the ellipsoid. For some uses, e.g. precision orbit computation, the result can be used directly. Second, the mapping can be done locally. One can develop a map of the geopotential for a region, without solving the complete problem. This results in more manageable computations, and control on the ingrediants of the computation. Third it does not require costly recalculation of trajectories using the improved potential at each iteration. It allows proceeding arc by arc, and comparison of the restults at each step with other data.

What follows is the formulation of an analytical solution of the problem of calculating satellite perturbations caused by the anomalous potential. Such a solution provides insight into the proposed measurement with significant computational economies. It gives a direct way to calculate analytically partial derivatives of the observable (velocity) the respect to the desired end product (the geoid). The analytical solution also allows the kernel function in the generalized inversion to be calculated directly. Furthermore we have investigated the use of generalized inverse (Backus and Gilbert, 1967,1968, 1970; Parker, 1975) to convert the observed velocity measurements to a mapping of surface values of geoid height. The ill-posed boundary value problem and the unstable downward continuation problem are addressed in a way that the additional assumptions used to obtain the solutions are clearly identified and an error estimate can be found.

The Range Rate Equation

We can begin with a geometrical discussion of the range equation. In general the distance between two satellites (Δ) with positions r_i, ϕ_i, λ_i, i=1,2 can be written

$$\Delta^2 = r_1^2 + r_2^2 - 2 r_1 r_2 \cos\psi \qquad (1$$

where the (geocentric) angle ψ between them is given by

$$\cos\psi = \sin\phi_1 \sin\phi_2 + \cos\phi_1 \cos\phi_2 \cos(\lambda_1 - \lambda_2) \qquad (2$$

The range rate equation is easily found to be

$$\dot{\Delta} = \frac{(r_1 - r_2\cos\psi)\dot{r}_1 + (r_2 - r_1\cos\psi)\dot{r}_2 + r_1 r_2 \sin\psi\,\dot{\psi}}{\Delta} \qquad (3$$

Equations 1,2, and 3 are valid for any satellite positions, and are not restricted to the low-low configuration. We proceed by assuming a reference orbit ($r_{oi}, \phi_{oi}, \lambda_{oi}$, i=1,2) and potential U and determine corrections ($\delta r_i, \delta\phi_i, \delta\lambda_i$, i=1,2) to the orbit. The dynamical equations enter when relating these corrections to corrections to the potential (T). When dealing with perturbations of a reference orbit we have

$$\begin{aligned} \Delta &= \Delta_o + \delta\Delta \\ \dot{\Delta} &= \dot{\Delta}_o + \delta\Delta \\ r_i &= r_{oi} + \delta r_i \end{aligned} \qquad (4$$

Considering $\delta\Delta, \delta r_1, \delta r_2, \ldots.$ to be quantities of first order we have

$$\Delta = \frac{(r_{o1} - r_{o2}\cos\psi_o)\delta\dot{r}_1 + (r_{o2} - r_{o1}\cos\psi_o)\delta r_2 + r_{o1} r_{o2}\sin\psi_o\,\delta\psi}{\Delta} \qquad (5$$

with

$$-\sin\psi_0\,\delta\dot{\psi} = (\cos\phi_1 \sin\phi_2 - \sin\phi_1 \cos\phi_2 \cos(\lambda_1-\lambda_2))\,\delta\dot{\phi}_1 + (\sin\phi_1 \cos\phi_2 - \cos\phi_1 \sin\phi_2 \cos(\lambda_1-\lambda_2))\,\delta\dot{\phi}_2 - \cos\phi_1 \cos\phi_2 \sin(\lambda_1-\lambda_2)(\delta\dot{\lambda}_1-\delta\dot{\lambda}_2) \quad (6$$

This range rate equation is valid for any combination of high and low satellites. The perturbations $\delta r, \delta\phi$, and $\delta\lambda$ are just those calculated by the theory given in the next section when it is applies in a reference frame defined by the earth's equator and prime meridian. Therefore this theory is applicable to any combination of satellites. If we do make the assumptions of two low small eccentricity satellites, then we obtain the usual relation ($r_i = R + \delta r_i$,i=1,2)

$$\delta\dot{\Delta} = \frac{R(1-\cos\psi)(\delta\dot{r}_1+\delta\dot{r}_2) + R^2\sin\psi\ \delta\dot{\psi}}{\Delta} \quad (7$$

Proceeding to treat the two coplanar low satellite case, it is convenient to use the perturbation solution in a reference frame defined by the mean orbital plane of the two satellites and its equator crossing. In this case the perturbations in radius are the same, and the separation angle becomes the difference in true anomaly of the two satellites which is the same as the difference in longitude ($\delta\theta$) of the theory. In this case (ϕ_i =0,i=1,2) , $\psi = \lambda_1 - \lambda_2 = \theta_1 - \theta_2$ <<1 and we have

$$\delta\dot{\Delta} = R\,\delta\dot{\psi} = R(\delta\dot{\lambda}_1 - \delta\dot{\lambda}_2) = R(\delta\dot{\theta}_1 - \delta\dot{\theta}_2) \quad (8$$

The solution (eq 33) specialized for small eccentricity with R=a is

$$a\,\delta\dot{\theta} = \frac{T}{na} - 2n\ \delta r(T) \quad (9$$

where $\delta r(T)$ is given in equation 27. We now have the range rate equation

$$\delta \dot{\Delta} = \frac{T(1)}{na} - \frac{T(2)}{na} - 2n\left(\delta r(1) - \delta r(2) \right) \qquad (10$$

which becomes our observation equation. It will be shown that the term $n\,\delta r$ is smaller than T/na. Then the first terms provide the desired mapping of the observable ($\delta\dot{\Delta}$) to the potential difference T(1)-T(2). Now δr is a rather complicated integral function of T, depending on the past motion of the satellite. It is critical to recognize that it is the complex dependence of δr on T that mandantes an iterative proceedure to obtain T. This arises due to the fact that $\partial T/\partial r$ occurs in the expression for $\delta \dot{r}$.

To summarize the discussion to this point, we have presented an iterative scheme that uses residuals in satellite to satellite range rate ($\delta\dot{\Delta}$) with respect to a reference trajectory and geopotential model (Uo). The scheme proceeds in two stages. A number of crossing tracks are regularized by imposing that the crossing points, analytically continued to the same reference surface (a sphere), have the same value of potential. This computation is analagous to crossing arc analysis used with satellite to sea surface altimeter data. This regularization cannot be done perfectly due to the δr terms in equation 9, which require a_priori knowledge of the potential to be computed. Therefore the iteration proceeds by using the potential (T_{n-1}) calculated in iteration n-1 to calculate δr and $\partial T/\partial r$ for iteration n. The resulting potential (T) can then be analytically continued to another surface using methods based on inverse theory.

Theory of Satellite Motion

In discussing the recovery of gravity anomalies (g) or geoid heights (N) with the SST approach, it is convenient, if possible, to have analytical formulae relating satellite position and velocity (the state vector) to the desired quantity. Such formulae also enable the sensitivity, or partial derivatives, of gravity anomalies and geoid heights to be more easily obtained; otherwise one would have to resort to costly numerical methods, with consequent loss of generality and insight.

For discussion purposes, we start with the theorem of conservation of energy of the satellite orbit. Of course, drag and radiation pressure perturbations would also have to be taken into account if the following relationships were used for actual data. One can obtain the same result for circular orbits by using the conservation of angular momentum. Let the total potential be represented by Ψ , which, for convenience, can be separated into a reference potential U and an anomalous potential T, i.e.

$$\Psi = U + T \tag{11}$$

In standard physical geodesy notation, U is the normal potential corresponding to that of a reference ellipsoid; however here we prefer to view it as a reference or adopted potential, with T being the remaining (unmodeled) part that we seek. Some confusion is bound to occur because the anomalous potential, or disturbing function, is generally denoted by R in celestial mechanics and by T in Physical Geodesy. If we write the kinetic energy as $mv^2/2$, with the vector component of velocity along track (v_u), cross track (v_w) and radial (v_r), then

$$\tfrac{1}{2}(v_u^2 + v_w^2 + v_r^2) + \Psi = \text{constant} \tag{12}$$

For a satellite with small eccentricity, we can treat the along track velocity and the unperturbed velocity vo, which gives v_u = vo + δv_u , $v_w = \delta v_w$, and $v_r = \delta v_r$. Therefore, to first order in small quantities,

$$\delta v_u = T/vo \tag{13}$$

This formula was first derived by Wolf (1969) and subsequently used by many others. For small eccentricity satellites, equation (13) gives the change in along track velocity as a function of the potential in space (i.e. position and time) to about 10%, the errors arising from the change in radial distance owing to T, interacting with U.

Equation 13, a direct mapping of the potential into the along track velocity, could be used as a first approximation for inverting observations of δv_u to determine T. However, significant perturbations in δv_w and δv_r also contain information about the disturbing potential T, so to obtain general formulae for the three components, we must solve the dynamical equations. This is done by using established proceedures in planetary theory for special perturbations.

Following the treatment in Brouwer and Clemence (1961), we can write the equations of motion as

$$\frac{d^2x}{dt^2} + \mu \frac{x}{r^3} = \frac{\partial T}{\partial x}$$

$$\frac{d^2y}{dt^2} + \mu \frac{y}{r^3} = \frac{\partial T}{\partial y} \tag{14}$$

$$\frac{d^2z}{dt^2} + \mu \frac{z}{r^3} = \frac{\partial T}{\partial z}$$

where $\mu = GM = 3.986\times10^{20}\ cm^3\ sec^{-2}$ and

$$r^2 = x^2 + y^2 + z^2 \tag{15}$$

For the moment, we assume T excludes only the central force term from Ψ. If the coordinates are r, θ, and ϕ, where θ is the longitude from the reference (x y plane) equator crossing in the reference plane, and ϕ is the latitude, then the equivalent differential equations of motion are

$$\frac{d^2 r}{dt^2} - r\cos^2\theta\left(\frac{d\theta}{dt}\right)^2 - r\left(\frac{d\phi}{dt}\right)^2 + \frac{\mu}{r^2} = \frac{\partial T}{\partial r}$$

$$\frac{d}{dt}\left(r^2\cos^2\phi\ \frac{d\theta}{dt}\right) = \frac{\partial T}{\partial \theta} \qquad (16$$

$$\frac{d}{dt}\left(r^2\frac{d\phi}{dt}\right) + r^2\sin\phi\cos\phi\left(\frac{d\theta}{dt}\right)^2 = \frac{\partial T}{\partial \phi}$$

Now these equations are most commonly used with respect to the earths equator. However, they are valid for any inertial frame. For our purposes it is convenient to adopt the orbital plane, at the beginning of the track, as the reference equator. In this frame, θ becomes the along track coordinate and ϕ the cross track coordinate. The origin, zero point, of course is arbitrary, and we chose the beginning of the data for convenience. We now must express the potential in this orbital frame.

To obtain an expression for the potential (T), and its gradiants, as a function of position (and time), we begin by assuming the potential on the earth's surface at r=R. Since $T(R, \phi, \lambda)$ given on a sphere is the Dirichlet boundary condition for Laplace's equation external to the sphere r=R, we can use the basic results from potential theory to obtain T at any point in space. Therefore, we can expand T in terms of orthonormal base functions (associated Legendre functions) as

$$T(R,\phi,\lambda) = \frac{GM}{R}\sum_{l,m}\left(\bar{C}_{lm}\cos m\lambda + \bar{S}_{lm}\sin m\lambda\right)\bar{P}_{lm}(\sin\phi) \qquad (17$$

which, using the properties of solutions of Laplace's equation in spherical coordinaes, can be upward continued as

$$T(r,\phi,\lambda) = \frac{GM}{r}\sum_{l,m}\left(\frac{R}{r}\right)^{l}\left(\bar{C}_{lm}\cos m\lambda + \bar{S}_{lm}\sin m\lambda\right)\bar{P}_{lm}(\sin\phi) \qquad (18$$

Alternatively, since T is harmonic in space, we can use Poisson's integral to obtain

$$T(r,\phi,\lambda) = \frac{R(r^2-R^2)}{4\pi}\int_0^{2\pi}\int_{-\pi/2}^{\pi/2} \frac{T(R,\phi',\lambda')}{\rho^3} \cos\phi' \, d\phi' \, d\lambda' \quad (19$$

where ρ is the distance between the integration point R, ϕ', λ' and the sample point r, ϕ, λ ; it is often written as

$$\rho^2 = r^2 + R^2 - 2Rr \cos\psi \quad (20$$

ψ being the central angle, which can be expressed as

$$\cos\psi = \sin\phi \sin\phi' + \cos\phi \cos\phi' \cos(\lambda - \lambda') \quad (21$$

We recall that the relation between the anomalous potential T and the geoid height N by Brun's formula is

$$N = T/\gamma \quad (22$$

where γ is the normal gravity on the reference surface. Furthermore, the relation between gravity anomalies Δg and T is given by

$$\Delta g = -\frac{\partial T}{\partial r} - \frac{2}{r} T \quad (23$$

and we can formally obtain N via Stokes' equation

$$N = \frac{R}{4\pi\gamma}\int S(\psi)\,\Delta g \, d\sigma \quad (24$$

and Molodenskii's equation

$$\Delta g_i = -\frac{\gamma N_i}{R} - \frac{\gamma}{16\pi R^3}\int \frac{N - N_i}{\sin^3(\psi/2)}\, d\sigma \qquad (25$$

Both equations (18) and (19) are suitable for use with the equations of motion (16). The spherical harmonic form (18) is equivalent to classical perturbation analysis and thus is not of particular interest here. However, equation (19) will allow mapping of the velocity directly to a surface potential.

Following the approach in Brouwer and Clemence (1961) we can write the perturbations in r, θ , and ϕ , as follows. Let

$$X = \frac{r^3}{\mu p}\left[2\int dT + r\frac{\partial T}{\partial r}\right]$$

$$Y = \frac{r^2}{\mu p}\frac{\partial T}{\partial \theta} \qquad (26$$

$$Z = \frac{r^3}{\mu p}\frac{1}{r}\frac{\partial T}{\partial r}$$

where $p=a(1 - e^2)$, a is the satellites semi major axis, e the eccentricity, and r corresponds to the unperturbed motion. We have

$$\delta r = \int X \sin[(f) - f]\, df$$

$$\delta \theta = \int\left[\int Y\, df - 2\frac{\delta r}{r}\right] df \qquad (27$$

$$\delta \phi = \int Z \sin[(f) - f]\, df$$

In equation (26) we interpret

$$
\begin{aligned}
dT &= \frac{\partial T}{\partial r} dr + \frac{\partial T}{\partial \theta} d\theta + \frac{\partial T}{\partial \phi} d\phi \\
&= \frac{\partial T}{\partial x} dx + \frac{\partial T}{\partial y} dy + \frac{\partial T}{\partial z} dz \\
&= \left(\underline{\nabla} T \cdot \underline{v}\right) dt
\end{aligned}
\tag{28}
$$

Therefore, for example the integral in the first part of equation (27) could be obtained by expressing r and T in terms of the true anomaly f, and it becomes

$$
\int dT = \int \frac{\partial T}{\partial f} df \tag{29}
$$

Other choices for the variable of integration are possible; for example, we could use the mean anomaly M, or the eccentric anomaly E. In equation (27) we have written them as integrals in f, although the same choices obtain. The form, using the true anomaly (f) is the one implemented in the following analysis. Note that the notation "(f)", due to Hansen, means that the variable is held fixed in the integration but takes its proper value when the perturbation is evaluated. These expressions then have the mathematical form of a convolution.

Equations (26) and (27) are quite general and depend on finding a once differentiable potential function T. The observable velocity can, of course, be found from the appropriate linear combination of derivatives of equation (27):

$$
\begin{aligned}
v_r &= \frac{d\delta r}{dt} = \frac{n}{(1-e^2)^{3/2}}\left(\frac{a}{r}\right)^3 \int X \cos[(f)-f]\, df \\
v_u &= r\frac{d\delta\theta}{dt} = \frac{n}{(1-e^2)^{3/2}}\left(\frac{a}{r}\right)^3 \left[r\int Y\, df - 2\delta r \right] \\
v_w &= r\frac{d\delta\phi}{dt} = \frac{n}{(1-e^2)^{3/2}}\left(\frac{a}{r}\right)^3 \int Z \cos[(f)-f]\, df
\end{aligned}
\tag{30}
$$

As a mater of interest, these expressions can be simplified for very small eccentricity:

$$X = \frac{1}{n^2 a}\left[2\int dT + a\frac{\partial T}{\partial r}\right]$$

$$Y = \frac{1}{n^2 a^2}\frac{\partial T}{\partial M} \qquad (31$$

$$Z = \frac{1}{n^2 a^2}\frac{\partial T}{\partial Z}$$

and

$$v_r = n \int X \cos[(M) - M]\, dM$$

$$v_u = n\left[a\int Y\, dM - 2\delta r\right] \qquad (32$$

$$v_w = na \int Z \cos[(M) - M]\, dM$$

We can rederive equation (13) from these equations.

$$v = n\left(a\frac{1}{n^2 a^2}T - 2\delta r\right) = \frac{1}{na}T - 2n\delta r \qquad (33$$

where we can make the identification vo = na. The second term can be written:

$$-n\delta r = -\frac{GM}{na^3}\delta r = -\frac{GM}{naa^2}\delta r = -\frac{GM}{v_o}\frac{\delta r}{a^2} = \frac{1}{v_o}\frac{d}{dr}\left(\frac{GM}{r}\right)\Bigg|_{r=a}\delta r \qquad (34$$

which is simply the change in velocity due to the perturbation δr interacting with the central force term, as was pointed out above.

This development can be employed in a number of ways. The particular computer implementation made the following choices. The perturbations were numerically computed using the Poisson integral (19) form of the potential. The reference trajectory was given and the gradiants were evaluated for each data point. Then a simple Simpsons Rule quadrature was used to calculate the integrals for each data point. Of course, the inital values of the integral were assumed zero. Any constant of integration is applied during the regularization, when all the track crossings are adjusted, and to have zero mean. The gradiants were computed in an earth fixed system, and then transformed to the adopted orbital plane, along track, cross track, and radial as follows. In the terrestrial frame we have:

$$\frac{\partial T}{\partial r} = \frac{R}{4\pi}\int\left[5R^2 r - r^3 - R\left(r^2+3R^2\right)\cos\psi'\right]\frac{T(R,\phi',\lambda')}{\rho^5}\,d\sigma' \quad (35$$

The gradiant vector can be expressed in the inertial frame with

$$\bar{F} = R_3\left(-\frac{\pi}{2}-\lambda-\theta\right)\,R_1\left(-\frac{\pi}{2}+\phi\right)\begin{bmatrix}\partial T/r\cos\phi\,\partial\lambda \\ \partial T/r\,\partial\phi \\ \partial T/\partial r\end{bmatrix} \quad (36$$

where $R_1(x)$, $R_2(x)$, and $R_3(x)$ are three dimensional rotations about the x,y and z axes respectively, and θ is the siderial angle for the time in question (Gaposchkin 1973). The along track, cross track and radial components are given by

$$\begin{aligned} F_u &= \hat{u}_u \cdot \bar{F} \\ F_w &= \hat{u}_w \cdot \bar{F} \\ F_r &= \hat{u}_r \cdot \bar{F} \end{aligned} \tag{37}$$

where the along track, radial and cross track unit vectors are

$$\begin{aligned} \hat{u}_u &= \bar{v} / |v| \\ \hat{u}_r &= \bar{r} / |r| \\ \hat{u}_w &= \hat{u}_r \times \hat{u}_u / |\hat{u}_r \times \hat{u}_u| \end{aligned} \tag{38}$$

The last important detail is the implementation of the Poisson integral. It is formally defined as an integral over the sphere. We assume that only the nearby area make nonzero contributions to the integral. This is based on the notion that since we are using a relatively high degree and order reference potential, (e.g. l=m=12) distant zones formally average to zero. Therefore, to limit the quadrature, the Poisson integral only includes surface elements less than 15 times the minimum distance from the evaluation point and the reference sphere, i.e. $\rho < 15(r-R)$.

Downward Continuation and Inverse Theory.

The two issues concerning the fundamental process of converting potential measurements at satellite altitudes to potential, i.e. geoid heights, at the earth's surface are the solution of a boundary value problem in potential theory with incomplete data on an undefined surface, and the downward continuation of the potential to the earth's surface in a manner that keeps the errors within bounds. The approach analysed here address both problems in one step. We will use generalized inverse methods (Bakus and Gilbert, 1967, 1968, 1970; Parker, 1977). This approach provides a means to incorporate error analysis, and to provide an arbitrary amount of smoothing.

Briefly, observations of δv are averages over some integration time, a property of the instrumentation. This is typically on the order of 10 sec. These observations are given on a smooth curve, but the curve is not a suitable boundary for solution of Laplace's equation. Therefore, we must somehow express the observed δv as if if were on a boundary. Also, the boundary values are given at a finite number of points, immediately precluding a complete solution. The limited number of points requires some additional assumptions about the potential field, either in terms of smoothness or as a_priori information. In addition, the observations will leave errors. The analytical continuation of the potential field is a stability issue (Bullard and Cooper, 1948): Any short wavelength error in the potential at satellite altitudes will grow arbitrarily large when downward continued, and the chosen method must control this error build up in a well defined way.

One common practice is to limit the vector space of the representation by performing a least squares solution (Vonbun et al. 1978). This has the advantage of a definite computation but does not identify the errors in the recovered anomalies as a function of wavenumber.

In the study performed, the following outline of spectral expansion inverse theory, we follow Parker(1977). Consider that we can write T at a point P_i as

$$T(P_i) = \int_{\sigma} G(P_i, P_Q)\, T(P_Q)\, d\sigma_Q \qquad (39$$

This is exactly the form of the Poisson integral. In equation (39) P_i is the point where the observation is made, and P_Q are on the desired surface, which can be taken as the earth's surface, and the element of integration d is on that surface; G(P ,P)is now a kernel function. We can treat the observation errors by using

$$T_i' = T_i/\sigma_i \tag{40}$$

where σ_i is the standard error of the observation T. The matrix with elements is defined as

$$\Gamma_{ij} = \int G(P_i, P_Q) G(P_j, P_Q)\, d\sigma \tag{41}$$

For the case where the observation points P_i and P_j are at the same geometric distance a, the elements can be shown to depend only on a and the angular distance between P_i and P_j. This corresponds to the low low case with circular orbits. The positive definite matrix can be diagonalized with an orthogonal matrix with

$$\Xi^T \Gamma \Xi = \Lambda \tag{42}$$

where

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \cdots \lambda_N), \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 \geq \lambda_5 \geq > > > \geq \geq \lambda_N > 0 \tag{43}$$

The orthogonal matrix Ξ can be calculated by standard numerical techniques due to Jacobi. We can now define a set of orthogonal base functions,

$$\psi_i(P) = \frac{1}{\sqrt{\lambda_i}} \sum_j \Xi_{ji} G(P_j, P) \tag{44}$$

and write

$$T(P) = \sum_{i=1}^{N} a_i \psi_i(P) \qquad (45$$

where

$$a_i = \frac{1}{\sqrt{\lambda_i}} \sum_j \Xi_{ji} T(P_j) \qquad (46$$

The standard error of a_i is $1/\sqrt{\lambda_i}$. Then the contribution of $a_i \psi_i$ to T can be explicitly determined, and only those base functions that are significant need be included. The choice of the set of base functions, however, is somewhat a matter of taste. Generally, the function becomes more oscillatory as i increases, i.e. for high wavenumbers. From this, we can use the size of to select the degree of smoothing. The economy of the inverse method hinges on being able to find an analytical expression for equation (41), which will be described next.

To be useful, the computation of should be fast enough so that an arbitarily large number of observations could be used. With n observations, or observation points, one has n(n+1)/2 values of to evaluate. Since is defined as a double integral over the sphere, evaluating this function could become very costly, using strictly numerical methods. Therefore we have begun by developing an algorithm to make this calculation simpler. The details are given in Lane and Gaposchkin (1986). We want to compute

$$\Gamma_{ij} = \int_0^{\pi}\int_0^{2\pi} \left[\frac{r_i^2 - R^2}{4\pi \rho_i^3}\right] \left[\frac{r_j^2 - R^2}{4\pi \rho_j^3}\right] R^2 \sin\phi \, d\lambda \, d\phi \qquad (47$$

where

$$\rho_i^2 = r_i^2 + R^2 - 2 r_i R \cos\psi_{iy} \ , \ \rho_j^2 = r_j^2 + R^2 - 2 r_j R \cos\psi_{jy} \qquad (48$$

where ψ_{iy} is the spherical angle between (ϕ_i, λ_i) and $y=(\phi, \lambda)$ for arbitrary i,j. We may rewrite $\cos(\psi_{jy})$ as

$$\cos\psi_{jy} = \cos\psi_{ij}\cos\psi_{iy} + \sin\psi_{ij}\sin\psi_{iy}\cos\alpha \tag{49}$$

where α is the angle between the arcs ψ_{ij} and ψ_{iy}. Now as ϕ and λ range over the whole sphere, α and ψ_{iy} will also range over the whole sphere with $0 \leqslant \alpha \leqslant 2\pi$ and $0 \leqslant \psi_{iy} \leqslant \pi$. Hence, integrating the double kernel with respect to α and ψ_{iy} will yield an equivalent integral, and we can write

$$\Gamma_{ij} = \int_0^{\pi}\int_0^{2\pi} \left[\frac{r_i^2 - R^2}{4\pi \rho_i^3}\right]\left[\frac{r_j^2 - R}{4\pi \rho_j^3}\right] \sin\psi \, d\alpha \, d\psi \tag{50}$$

This can be rewritten as

$$\Gamma_{ij} = \int_0^{\pi} b_1(\psi) \int_0^{2\pi} \frac{d\alpha}{(1 + b_2(\psi)\cos\alpha)^{3/2}} \, d\psi \tag{51}$$

with

$$\begin{aligned}
& b_1 = c_1 \sin\psi \,/\, (c_2 a_2(\psi) + a_1(\psi) a_2(\psi))^{3/2} \\
& c_1 = (r_i^2 - R^2)(r_j^2 - R^2)/16\pi^2 \\
& c_2 = r_i^2 + R^2 \qquad c_5 = -2 r_j R \\
& c_3 = r_j^2 + R^2 \qquad c_6 = \cos\psi_{ij} \\
& c_4 = -2 r_i R \qquad c_7 = \sin\psi_{ij} \\
& a_1(\psi) = c_4 \cos\psi \quad , \quad a_2(\psi) = c_3 + c_5 c_6 \cos\psi
\end{aligned} \tag{52}$$

and

$$\begin{aligned}
b_2 &= a_3(\psi) / a_2(\psi) \\
&= \frac{-2 r_j R \sin\psi_{ij} \sin\psi}{r_j^2 + R^2 - 2 r_j R \cos\psi_{ij} \cos\psi}
\end{aligned}$$

(53

So we can write the desired function as

$$\Gamma_{ij} = \int_0^{\pi} b_1(\psi)\, Y(\psi)\, d\psi \qquad (54$$

where

$$Y(\psi) = \frac{4}{1+b_2\sqrt{1-b_2}} \int_0^{\pi/2} \sqrt{1 + \frac{2b_2}{1-b_2}\sin^2 y}\; dy \qquad (55$$

which can be written as an elliptic integral of the second kind as

$$Y(\psi) = \frac{4}{1+b_2\sqrt{1-b_2}} E(\bar{k}) \qquad (56$$

where

$$\bar{k}^2 = -\frac{2b_2(\psi)}{1-b_2(\psi)} \qquad (57$$

Rapidly convergent series for Elliptic Integrals are readily available. We have reduced the calculation of the double integral, to a single quadrature, which can be effectively done with standard proceedures such as a Romberg extrapolation. The accuracy of calculation of these matrix elements depends on the accuracy of the quadrature used, and the user must balance speed with required accuracy.

A further simplification and speed up is possible for the special case where $r_i = r_j$. In this case the function depends only on the angular distance between the points. One could precompute these functions for a number of values of the separation and interpolate in this table.

To investigate the effectiveness of this inverse method for the downward continuation problem, a sample calculation was performed. The details of the calculation are given in Lane, and Gaposchkin 1985). The experiment consisted of selecting a grid of 24 mass points on a 1 deg x 1 deg grid at the earths surface. The mass points were selected to have zero average. They were arranged to have equal magnitude and alternate in sign. The center point of the 5 x 5 grid was zero. The potential was computed, exactly, at 121 grid points at an altitude of 212 km, on a 1 deg x 1 deg, 11 x 11 grid. The potential was continued back down to an intermediate integration surface using this inverse method. The potential at this surface was recovered with an accuracy of 4%.

In this experiment we had the advantage of knowing the true value of the potential by simply computing the potential from the mass points in the same way the observed data were computed. Hence, we were able to compute a measure of the performance of the inverse theory. For each test point (R,y) we computed a kind of performance index, the standard deviation divided by the average value of the actual potential over the range of sample points

$$S^2 = \frac{\sum_i \left[\hat{T}(y_i) - T(y_i)\right]^2}{\sum_i T^2(y_i)} \qquad (58$$

It is clear that the inverse method is computationally expensive, and so we give some information about the execution time of the routines. The computations were done of a Harris H800 computer, and all programming was done in FORTRAN. Table 1 lists the time involved for execution of principal steps for different values of N, the number of sample points.

Table 1.
Computation time for Geopotential Inverse

N	CPU for	CPU for Eigenvectors and Eigenvalues
20	18.471 sec	0.417 sec
50	113.708 sec	6.309 sec
100	449.639 sec	48.136 sec

From table 1 we can see that the computation of the eigenvalues and eigenvectors is slightly less than an n**3 problem. This is consistent with the computation of the eigenvalues and eigenvectors being equivalent to computation of the inverse of . The computation of itself appears to be slightly less than an n**2 problem, but the computation of takes much longer. However, the computation time will change if one changes the sphere of integration and the sphere at which the sample points are taken, or if one changes the configuration of the sample points. For example the CPU time involved for computing where the 121 sample points are at 424 km height and the sphere of integration is 212 km is only 13.523 sec: more than 30 times faster than the lower configuration! The computations in the remainder of the routine are linear, and so if one needs a satisfactory upper bound on the total CPU time involved for the 212 km height, 100 point configuration, then one can estimate it from (1/3)(n+n**2+n**3)529 sec, where n=N/100.

The numerical experiments conducted are described in detail in Lane and Gaposchkin (1986). In partucular the spectral structure of the inversion. Therefore we briefly summarize the results in table 2. This test case concerns the 121 sample points on a sphere of r=1.1ae, (height =637.8134 km) and and 25 test points on a sphere of R=1.05ae (height = 318.906).

Table 2
Downward Continuation with 121 sample points

	Radius	km	S
Sample Points	1.1 ae	637.813	0.019
Test Points	1.05 ae	318.907	
Sample Points	1.033ae	212.604	0.04

Test Points 1.017ae 106.302

Not only are these results very satisfactory, but the size of the program needed is much smaller than for a similar computation in a global field.

GEOS-3/ATS-6 Tracking data analysis

To test some of the algorithms for computing potential from SST data, the existing set of data from the GEOS-3 to ATS-6 tracking experiment was analysed. The daa was provided by J. Marsh (GSFC) and is described in Marsh et al 1977. The data was provided in two files. For each of the observed 179 revolutions, Marsh computed an orbit based on all available tracking data, using the GEM10 gravity field, truncated at and complete through degree and order 12 as a reference. The residuals from SST observations were then computed. We interpret those residuals to be completely due to the anomalous potential above degree and order 12. Marsh then used a smoothing algorithm to obtain a smoothed range rate observation, and its rate or range acceleration. These three quantities, observed range rate, smoothed range rate, and smoothed range acceleration, were then given as a time tagged time series for each pass. In addition, for each pass, the state vectors for GEOS-3 and ATS-6 at a time near the beginning of the pass were given. The state vectors were presented as position and velocity in an inertial reference system. These state vectors allowed us to compute an approximate position for both satellites for each observation point by numerical integration. This provided the position and velocity used in evaluating the integrals for peturbations due to he anomalous potential.

The GEOS-3 ATS-6 experiment was not ideal from our standpoint for a number of reasons. First, the data is not of 1 micron/sec limited by system parameters, and refraction effects. Second, the GEOS-3 satellite is at 800km altitude. Both of these factors limit the additional geopotential information to be gained from analysis of the data. In addition, the coverage was a fraction of one hemisphere. Also, the satellites were not surfaceforce compensated. Marsh took every possible step to account for nongravitational effects. Yet we believe there are still very important nongravitational effects in the reduced range rate residuals. Most important is, the High low configuration only measures a component of the along track velocity. When the lower satellite is just beneath the high satellite, it measures none. On the otherhand, the radial acceleration, which can be derived from numerical differentiation of the velocity data, directly measures the gravity acceleration. For the data taken at or near the sub satellite point, successful recovery of the gravity anomaly at satellite altitude has been demonstrated (Marsh et al, 1977). However, for data away from the sub satellite point, a significant component of the along track velocity is detectable, and the question is, how to extract this from the other velocity components.

In any event, we proceed to set up the following analysis. Having solved the forward problem, in that, given an anomalous potental T, we can compute the velocity in all components, these can be projected on the line of sight from ATS-6 to GEOS-3. To obtain, a correction to the anomalous potential, we assume that the observed residual velocity on any iteration, is all along track, and is therefore a measure of the potential at that point in space.

Summary and Conclusions

A relatively efficient computation has been developed for the kernal function used in the Geophysical Inverse Problem.

Geophysical Inverse Theory seems a viable method to solve the downward continuation problem, and map the geopotential from satellite altitudes to the Earth's surface.

Geophysical Inverse Theory seems able to recover the geopotential at the Earth's surface with an accuracy of a few percent, in the absence of noise.

The Inverse problem seems appropriately broken up into local networks, which can be solved separately, leading to significantly simplification in computing requirements.

The analytical development for computing perturbations in velocity have been developed and are shown to be have sufficient accuracy for the analysis of SST data.

Future Studies

For the Geophysical Inverse Method

The trade-off of size of the sample point space and the solution space.

The influence of errors in the sample point space, both position and measurement errors.

The benefit of sample points at different altitudes, in improving the solution, and eliminating possible nonuniqueness.

Formal understanding of the annihilator, and possible tests for nonuniqueness.

Most important, it is essential to complete this analysis by merging the Satellite perturbation analysis with the downward continuation calculation, and to iterate on this process until convergence. Each part is demonstrated to be effective.

References

Backus, G.,and Gilbert, F., 1967. Numerical applicatins of a formalism for geophysical inverse problems. Geophys. Journ. Roy Astron. Soc., Vol 13, pp247-276.

Backus, G.,and Gilbert, F., 1968, The resolving power of gross earth data. Geophys. Journ. Roy. Astron. Soc., Vol 16, pp169-205.

Backus, G., and Gilbert, F., 1970, Uniqueness in the inversion of inaccurate gross earth data. Phil. Trans. Roy, Soc. London, Ser. A. Vol 226, pp343-344.

Brouwer, D., and Clemence, G.M., 1961, Methods of Celestial Mechanics, Academic Press, New York, 598pp.

Bullard, E.C., and Cooper, R.I.B., 1948, The deterimination of the masses necessary to produce a given gravitational field. Proc. Roy. Soc. London, Ser. A, Vol 194, pp332-347.

Colombo, O.L., Global Geopotential Modelling from Satellite to Satellite Tracking, The Ohio State Univ., Dept. of Geod. Sci., Report No. 317, 1981.

Gaposchkin, E.M., 1973, Satellite Dynamics. In 1973 Smithsonian Standard Earth (III), ed. by E.M. Gaposchkin, Smithsonian Astrophysical Obs. Spec. Rep. No. 353, pp85-192.

Gaposchkin, E.M., 1980, Averaging on the Surface of a Sphere, Journ. Geophys. Res. Vol 85, No B6, pp3187-3193,June 10 1980.

Giacaglia, G.E.O., and Lundquist, C.A., 1972, Sampling functions for geophysics. Smithsonian Astrophysical Obs. Spec. Rep. No. 344, 93pp.

Guier, W.H., and Newton, R.R., The Earth's Gravity Field as Deduced from the Doppler Tracking of Five Satellites, Journ. Geophys. REs., vol 70, pp4613-4626, 1965.

Hajela, D.P., 1977, Recovery of 5 deg Mean Gravity Anomalies in local areas from ATS-6/GEOS-3 Satellite to Satellite range rate observations, Report 259, Dep. of Geod. Sci., Ohio State Univ., Columbus, 1978.

Heiskanen, W.A., and Moritz, H., 1967, Physical Geodesy, W.H. Freeman and Co., San Francisco, 364pp.

Kaula, W.M., 1966, Theory of Satellite Geodesy, Blaisdell Publishing Co., Waltham Mass.

Kaula, W.M., 1967, Theory of statistical analysis of data distributed over a sphere. Rev. of Geophys., Vol 5, pp38-107.

Kaula, W.M., 1970, The Terrestrial Environment: Solid-Earth and Ocean Physics. Report of a study at Williamstown Mass., NASA CR 1579, Massachusetts Institute of Technology, Cambridge, Mass., 152pp.

Kaula, W.M., 1983, Inference of Variations in the Gravity Field from Satellite to Satellite Range-Rate, Journ. Geophys. Res.

Lane, M.T., and Gaposchkin, E.M., The Estimation of Geopotentials by Way of Geophysical Inverse Theory, TR 735, M.I.T. Lincoln Laboratory, Lexington Mass. 02173, 31pp, 27 January 1986.

Lerch, F.L., Klosko, S.M., Patel, G.B., 1982, A Refined Gravity Model from Lageos (GEM-L2), Geophys. Res. Letters, p. 1263.

Lerch, F.L., Klosko, S.M., Laubscher, R.E., and Wagner, C.A. Gravity Field Improvement using GEOS-3 (GEM 9 and 10), Journ. Geophys. REs., 84(B8), pp3897-3916, 1979.

Lerch, F.L., Wagner, C.A., Klosko, S.M., Belott, R.P., Laubscher, R.E., amd Taylor, W.A., Gravoity Field Model Improvement using GEOS-3 Altimetry (GEM-10A and 10B), Journ. Mar. GEod.,1981.

Marsh J.G., Marsh, B.D., Conrad, T.D., Wells, W.T., and Williamson, R.G., Gravity Anomalies Near the East Pacific Rise with Wavelengths Shorter than 3300 km Recovered from GEOS-3/ATS-6 Satellite to Satellite Doppler Tracking Data, NASA TM 79553, Goddard Space Flight Center, Greenbelt Md. 20771, 269pp, December 1977.

Marsh J.G., Marsh B.D., Williamson, R.G., and Wells, W.T., The Gravity Field in the Central Pacific From Satellite to SAtellite Tracking, Journ. Geophys. Res., Vol 86, No B5, pp3979-3997, May 10, 1981.

Moritz, H., 1974, Some first accuracy estimates for applicatins of aerial gradiometry, Ohio State Univ., Dept. Geod. Sci. Rep. No. 209, 72pp.

Moritz, H., 1975, Integral formulas and collocation. Ohio State Univ., Dept. Geod. Sci. Rep. No. 234, 63pp.

Moritz, H., 1978, Least Squares Collocation, Rev. Geophys. Space Phys., Vol 16, pp421-430.

Muller, P.M., and Sjogren, W.L., Lunar Mass Concentrations: Mascons, Science, Vol. 161, pp680-684, 1968.

National Academy of Sciences/National Research Council, 1979. Report of a dedicated gravitational satellite mission. Panel on Gravity and Seal Level of the Committee on Geodesy, 53pp.

Parker, R.L., 1977, Understanding Inverse Theory, Ann. Rev. Earth Planet. Science, Vol. 5, pp35-64.

Rummel, R., 1975, Downward continuatin of gravity informatin from satellite to satellite tracking or satellite gradiometry in local areas. Ohio State Univ., Dept. Geod. Sci. Rep. No. 221, 50pp.

Schwarz, C.R., 1970, Gravity field refinement by satellite to satellite doppler tracking, Ohio State Univ., Dept. Geod. Sci. Rep. No. 147, 141pp.

Sjogren, W.L., Laing, P.A., Liu, A.S., and Wimberly, R.N., Earth Gravity Field Variations from GEOS-3/ATS-6 Satellite to Satellite Radio Tracking, EO Transactions AGU, Vol 57, No. 4, pp234, 1976.

Sjogren, W.L., Wimberly, R.M., 1978. GEOS-3 Experiment Report (1977): Analysis of SST Doppler Data for the Determinatin of Earth Gravity Field Variations, 27pp. JPL Report 900-884, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.

Wolf, M., 1969, Direct measurements of the earth's gravitatinal potential using a satellite pair. Journ. Geophys. Res., Vol. 74, pp5295-5300.

Wong, L., Matthews, E., and Downs, W. Analysis of Satellite to Satellite Tracking (SST) and Altimetry Data from GEOS-C. NASA Contractor REport 14138, Wallops Flight Center, Wallops

Island, Virginia 23337, 75pp, April 1978.