Reports of the Department of Geodetic Science

Report No. 280

# APPLICATIONS OF GEODESY TO GEODYNAMICS

An International Symposium

Proceedings of the Ninth Geodesy/Solid Earth
and Ocean Physics (GEOP)
Research Conference

October 2-5, 1978
Columbus, Ohio

Edited by

Ivan I. Mueller



The Ohio State University
Department of Geodetic Science
1958 Neil Avenue
Columbus, Ohio 43210

ACKNOWLEDGEMENTS

The Ninth GEOP Research Conference was financially sponsored by the following organizations:

NASA Headquarters
Washington, D.C. 20546

NOS/NOAA
Rockville, Maryland 20852

U.S. Geological Survey
Reston, Virginia 22092

Defense Mapping Agency
Washington, D.C. 20305

Department of Geodetic Science, OSU
Columbus, Ohio 43210

The financial arrangements were handled by NASA under Contract No. NASW-3194 (Technical Director: James P. Murphy, Code ERG-2, NASA Headquarters, Washington, D.C. 20546), OSURF Project No. 711022 (Supervisor: Ivan I. Mueller, Department of Geodetic Science, The Ohio State University). The above assistance is gratefully acknowledged.

Scientific sponsors of the Conference were

American Geophysical Union
COSPAR
International Association of Geodesy/IUGG

The scientific program committee consisted of the chairmen of the eight panels: G.M.R. Winkler, P.L. Bender, C.A. Whitten, P. Vaníček, R.J. Anderle, U.A. Uotila, J.C. Harrison, W.M. Kaula, and Ivan I. Mueller, the convenor of the Conference, as chairman.

Administrative arrangements were handled by the staff of the Department of Geodetic Science as follows:

| | |
|---|---|
| Conference Secretary | Irene Tesfai |
| Registration | Erricos Pavlis |
| Transportation | Christopher Jekeli |
| Facilities | Charles Challstrom |

| | |
|---|---|
| Meals | John Hannah |
| Desk | Roberta Canegali, Liza Marshall |

Their dedicated service contributed significantly to the success of the Conference and was especially appreciated.

This report was produced directly from author-prepared typescripts by the Editorial and Printing Services staff of the OSU Research Foundation. Their help is gratefully acknowledged.

TABLE OF CONTENTS AND PROGRAM

SESSION 4 — VERTICAL CRUSTAL MOVEMENTS (TERRESTRIAL TECHNIQUES)
Petr Vaníček, Chairman

# Introduction

This international symposium, sponsored by IAG and COSPAR, and co-sponsored by AGU, DMA, NASA, NOAA, USGS and the Department of Geodetic Science of The Ohio State University, took place in the Fawcett Center for Tomorrow, Columbus, Ohio, October 2-5, 1978. It was attended by 210 persons from 16 countries (53 non-U.S. participants).

The meeting was divided into eight scientific sessions, each dealing with a geodynamic phenomenon to which geodesy can be applied. The last session was devoted to general geodynamic applications.

Harold L. Enarson, President of OSU, welcomed the participants on behalf of the University. Special welcome was extended to members of the IAU Working Group on Earth Rotation and to the NAS/NRC Committee on Geodesy. Both of these groups scheduled meetings during the Conference. Dr. Enarson was followed by Ivan I. Mueller, the convenor of the symposium, who reflected on the successful previous GEOP Conferences held during 1972 and 1974 (see Department of Geodetic Science Report No. 231). These were the first serious efforts to organize topical conferences of an interdisciplinary nature. The AGU soon followed with the Chapman Conferences. He mentioned that the main purpose of this symposium is to review the developments since 1974 and the future in geodetic (observational and analytical) techniques in detecting and monitoring geodynamic phenomena. He expressed his satisfaction with the apparent interdisciplinary mixture of the audience, in which solid earth geophysicists, oceanographers, astronomers, space technologists, and geodesists were present. This is a good indication that communication problems between these disciplines are on the decrease.

After explaining the structure and the conduct of the Conference, he expressed thanks to the program committee, which consisted of the session chairmen, and to the OSU Conference staff, and opened the scientific portion of the symposium.

To obtain full benefit of the Proceedings of this Conference, it is recommended that the reader start with the contribution of W.M. Kaula, presented during the last session.

# Progress in Geophysical Aspects of the Rotation of the Earth

Kurt Lambeck
Research School of Earth Sciences, Australian National University
Canberra 2600, Australia

Introduction. Rochester (1973) reviewed the subject of the Earth's rotation at the second Geodesy, Solid Earth and Ocean Physics Research Conference, emphasizing the developments made since the publication of Munk and MacDonald's (1960) book, "The Rotation of the Earth". In his introduction, Rochester stressed the many aspects of the geophysical causes and consequences of the Earth's rotation and how the subject has drawn the attention of scientists working in fields ranging from astronomy to palaeontology. This multiplicity renders a completely satisfactory review of the subject unlikely and I limit myself here to some of the more pertinent results obtained since Rochester's review, results which were already foreshadowed in the Conference Report by Kaula et al. (1973). I have attempted a more complete review elsewhere (Lambeck, 1978a).

Any discussion of the Earth's variable rotation is conveniently separated into three parts: (i) the motion of the rotation axis in space, precession and nutation, (ii) the motion of the rotation axis relative to the Earth, polar motion, and (iii) the rate of rotation about this axis, or changes in the length of day. I follow this separation in the following review with emphasis on the last two aspects. I do not discuss the methods nor results of the observation process (see Guinot, 1978; Aardoom, 1978; Lambeck, 1978b).

## Polar motion

### Low frequency variations

The schematic polar motion power spectrum is illustrated in figure 1. The power near zero frequency (not shown) is a consequence of a secular drift in the pole position; in a westerly direction at a rate of 0".002-0".003 $yr^{-1}$. This motion appears to be real and probably a consequence of an exchange of mass between the Greenland ice sheet and the oceans. If the observed secular rise in sea level of about 1 mm $yr^{-1}$ is due to the melting of this ice cap, then the rate of melting must be of the order of 10 cm $yr^{-1}$ over the entire sheet and this would result in the observed secular pole shift. The pole shift is quite insensitive to the melting of the Antarctic ice (Lambeck, 1978a). There is a suggestion that, superimposed on this drift, a decade scale wobble occurs, mainly an oscillation at right angles to the direction of drift (Figure 2). Markowitz (1970) finds a period of about 24 years, Vincente and Currie (1976) suggest 30 years. The reality of this motion remains obscure. In particular, a comparison of recent annual mean pole positions obtained by the Bureau International de l'Heure (BIH) and the five station International Latitude Service (ILS), indicate differences that are quite similar to this wobble (figure 3). This strongly suggests that it is a consequence of the observing process and not of a real excitation of the rotation axis direction.

Evidence for polar wander over geologic time remains obscure. Goldreich and Toomre (1969) concluded that if "continental drift" occurred, a large scale wandering of the axis of rotation is inevitable and that, for N plates moving at an average velocity $\sigma$, the pole moves at a rate of $\sigma\sqrt{N}$. But the more recent paleomagnetic studies do not require such large scale "absolute" motion of the pole to explain the apparent pole paths for the various parts of the world (McElhinny, 1973). This does not imply that the lower mantle viscosity is, after all, very high but is a consequence of the continents not being isostatically compensated in the Goldreich and Toomre model. The actual change in the inertia tensor due to plate tectonics is a second order effect rather than a first order one (Lambeck, 1978a,c).

### Chandler wobble

The interesting part of the polar motion spectrum remains the Chandler wobble, centered at a period of about 14 months. This is the Eulerian precession for the non-rigid Earth. The three questions associated with this motion concern its period, dissipation and excitation. Smith (1977) has re-evaluated the lengthening of the wobble period due to the mantle elasticity and core using the normal mode approach suggested by

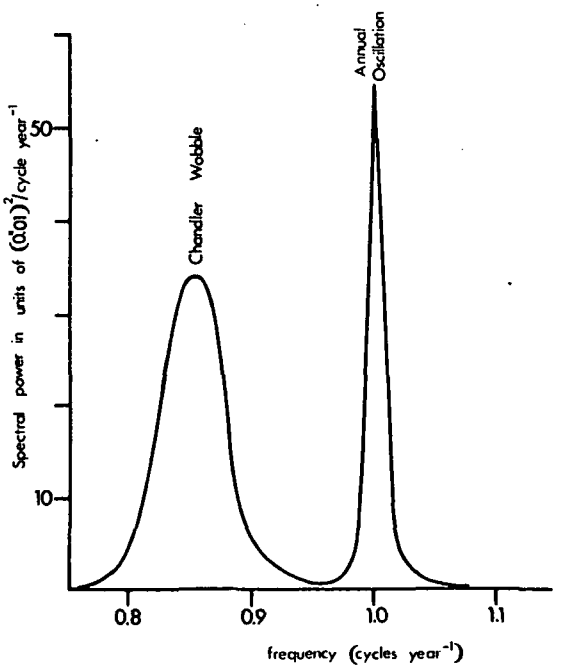



Fig. 1. Schematic Power spectrum of polar motion.

Proc. of the 9th GEOP Conference. *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

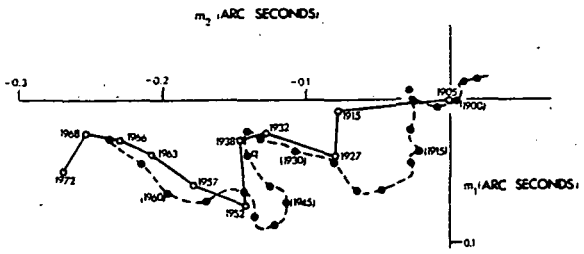



Fig. 2. Motion of the mean pole according to A. Stoyko. The pole positions present running means based on six year intervals. Positions every 3 years are indicated. The motion of the mean pole according to W. Markowitz is indicated by the open circles.

Gilbert (1971). In this method, the rotational deformations are expressed by the normal-mode eigen functions which themselves can be verified from studies of the Earth's free oscillations. Smith's treatment also places fewer constraints on the nature of the permissible core motion than do the earlier Jeffreys-Vicente and Molodensky theories. Sasao et al. (1977, 1978) discusses these theories in more detail, in particular, they introduce viscous coupling between the core and mantle. Smith has considered core models with different degrees of stratification, expressed by the Brunt-Vaisala frequency $N_{B-V}$, but only in the case of a strong stable stratification is the wobble period affected by a significant amount (Table 1). Seismic data remains inadequate to

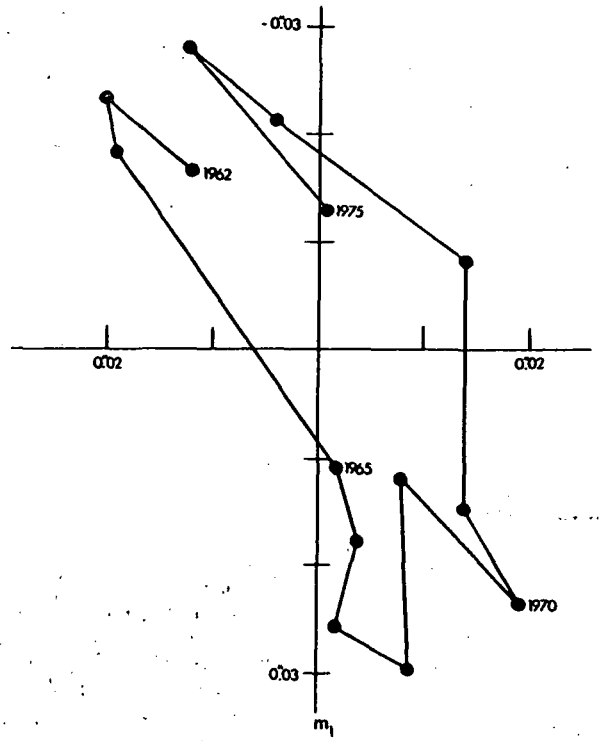



Fig. 3. Annual mean differences between the B.I.H. and ILS pole positions from 1962 to 1975.

TABLE 1. Summary of theoretical contributions to the Chandler wobble period.

| | Model 1 ($N^2_{BV}=0$) | | Model 2 ($N^2_{BV}=3.38\times10^{-7}$) | |
|---|---|---|---|---|
| Mantle and Core (Smith, 1977) | 403.6 | | 405.2 | |
| Mantle anelasticity | | | | |
| $Q_m = 300$ } Constant Q model | | 3.9 | | |
| $Q_m = 600$ } | | 1.8 | | |
| frequency dependent Q model | | 7.6 | | |
| Ocean equilibrium theory | | 27.4 | | |
| Non equilibrium ocean tide | | | | |
| $Q_w = 20$; $\varepsilon = 22°$ | | 25.2 | | |
| $Q_w = 40$; $\varepsilon = 11°$ | | 26.8 | | |
| $Q_w = 80$; $\varepsilon = 6°$ | | 27.2 | | |
| Total period | $Q_m$=300 | $Q_m$=600 | $Q_m$=300 | $Q_m$=600 |
| $Q_w = 20$ | 432.7 | 430.6 | 434.3 | 432.2 |
| $Q_w = 40$ | 434.3 | 432.2 | 435.9 | 433.8 |
| $Q_w = 80$ | 434.7 | 432.6 | 436.3 | 434.2 |
| $Q_w = \infty$ | 434.9 | 432.8 | 436.5 | 434.6 |
| Observed period (Jeffreys, 1968) | | 434.3 $\pm$ 2.2 | | |

TABLE 2.

| Period | 12 hrs | 24 hrs | 6 m | 12 m | 430 days |
|---|---|---|---|---|---|
| Q = 10 | 42.3 | 44.5 | 60.6 | 61.6 | 62.8 |
| 50 | 9.7 | 10.3 | 14.8 | 15.2 | 15.4 |
| 100 | 5.0 | 5.2 | 7.6 | 7.8 | 7.9 |
| 200 | 2.6 | 2.7 | 3.8 | 3.9 | 4.0 |
| 400 | 1.3 | 1.4 | 2.0 | 2.0 | 2.0 |
| 600 | 0.93 | 1.0 | 1.4 | 1.4 | 1.4 |
| 1000 | 0.60 | 0.63 | 0.80 | 0.80 | 0.80 |
| 600 | 0.5 | 0.7 | 4.1 | 5.2 | 5.5 |

Percentage changes in $k_2$ due to dispersion as a function of the tidal frequence. The first 7 lines are based on the assumption that Q is independent of frequency from about 1 sec periods to the period in question. The last line is based on the assumption that Q is proportional to $(\text{frequency})^{1/3}$.

determine the extent to which the core may be stratified.

In most recent calculations of Love-numbers and wobble period, the elastic parameters K,μ have been assumed to be frequency independent and values deduced from body waves or free oscillations have been used. This is at variance with the current interpretation given to the observation that the mantle Q is nearly constant over a wide range of seismic frequencies (Anderson et al. 1977; Kanamori and Anderson, 1977; Liu et al. 1976), and dispersion effects may be of consequence. In particular, the real parts of Love numbers become frequency dependent. Table 2 summarizes some results for a homogeneous incompressible Earth model and Table 1 gives corrections to the wobble period (see also Dahlen, 1978; Lambeck, 1978a). Very recently, Anderson and Minster (1978) suggested that Q may be proportional to $(\text{frequency})^{1/3}$, based on an observation that "high temperature background" may reflect as important dissipation mechanism. Table 2 also summarizes the resulting frequency dependent Love numbers in this case.

Dahlen (1976) has developed the equilibrium pole tide theory to allow for self attraction and the yielding of the Earth under the variable tidal load. But whether or not this tide follows an equilibrium theory remains obscure. Work by Miller and Wunsch (1973) and Currie (1975) confirm that, while the pole tide is an ocean-wide phenomena, available evidence is quite inadequate to map its global characteristics. Especially the evidence for any lag of the tide behind the forcing function remains unsatisfactory (Hosoyama, 1976). Lambeck (1978a) has modelled the consequences of dissipation of the pole tide on the rotation in a manner similar to that used for dissipation in the lunar tides (Lambeck, 1977) and the modification of the wobble period may be of the order of 0.5 days (Table 1). This remains below the noise level of the astronomical estimates of the wobble period (Jeffreys, 1968).

As the wobble period depends on the rate of rotation and, to a lesser degree on the ocean-land distribution, the period will have changed during geological time. In particular it could have approached that of the annual frequency (Cannon, 1974) but the geological consequences of such a resonance do not appear to be very severe (Lambeck, 1975a).

Excitation mechanisms of the wobble have continued to receive attention since Rochester's 1973 review. The then current state of the seismic excitation hypothesis was reviewed in Kaula et al. (1973). At that time, differences of opinion existed on (i) the correct treatment of the core-mantle interface in the theoretical evaluation of the earthquake induced changes in the inertia tensor, (ii) on the appropriate moment-magnitude relationship, (iii) on the most satisfactory manner of evaluating the cumulative effects of seismic activity and, (iv) on the observational evidence for discontinuities in the curvature of the pole path. The theoretical aspects now appear to be resolved (Mansinha et al. 1978) and all recent estimates of the pole shift due to an earthquake of given source parameters are in agreement (Table 3). O'Connell and Dziewonski (1976), using the free oscillation

TABLE 3.

| Event | Wobble excitation Magnitude (0".01) | Direction | Author |
|---|---|---|---|
| Alaska | | | |
| 1964 | 0.72 | 201° | M.L. Smith (1977) |
| " | 0.73 | 202° | Dahlen (1973) |
| " | 1.11 | 203° | O'Connell and Dziewonski (1976) |
| Chile | | | |
| 1960 (1) | 2.12 | 114° | M.L. Smith (1977) |
| (2) | 2.80 | 118° | |
| " (1) | 2.56 | 109° | O'Connell and Dziewonski (1976) |
| " (1) | 2.2 | 101° | Mansinha et al. (1977) |

Comparison of estimates of the shift of the inertia axis due to the Alaskan (1964) and Chile (1960). For the Chile event (1) refers to the main shock while (2) refers to the precursor.

formulation, evaluated the cumulative seismic excitation function from 1900 to 1970 and conclude that it is adequate to maintain the wobble against damping, although Kanamori (1976, 1977a) argues that their seismic moments are overestimated. It is now clear that no single moment-magnitude relationship describes all earthquakes, and that the magnitudes, determined from 20-100s period seismic waves, may be an inadequate measure of the overall seismic moment. In particular, there may be a considerable low frequency or aseismic contribution to the moment. Evidence for this comes from observations of (i) precursors (Kanamori and Cipar, 1974; Kanamori and Anderson, 1975; Dziewonski and Gilbert, 1975; Thatcher, 1974), (ii) tsunami earthquakes (Kanamori, 1972), (iii) studies of aftershock areas (Kanamori, 1977a; Stuart and Johnston, 1974) and, (iv) discrepancies between seismic slip and plate motions (Chen and Molnar, 1977; Kanamori, 1977b). If the aseismic slip occurs over time intervals that are short compared to the Chandler wobble period, they may contribute significantly to the excitation of this wobble. This suggests that, since it is not well understood which particular earthquakes are associated with significant aseismic slip, a more useful measure for comparing the wobble with seismic activity than seismic moments of the largest events, is the frequency (N) of earthquakes above a certain magnitude. This is shown in figure 4 together with the elastic energy $E_e$ released by large earthquakes. Both N and $E_e$ show trends that are comparable to the fluctuations in the wobble amplitude.

The other excitation mechanism that has recently been revised is the variability in the atmospheric mass distribution (Wilson and Haubrich, 1976). But the meteorological data is neither sufficiently reliable nor complete to permit an unambiguous interpretation. About the Chandler frequency, Wilson and Haubrich estimates an average power in the meteorological excitation spectrum that is about one sixth of that required to maintain the wobble. Part of the missing power may be a consequence of an absence of surface pressure data over central Asia, of an overestimation of the astronomic spectrum due to noise in the data and to hydrological fractors, but the evidence remains tenuous. In particular, differences found between the results by Wilson and Haubrich, Siderenkov (1973) and Jochmann (1976) for the annual atmospheric excitation function points to an unsatisfactory situation (Lambeck, 1978a).

Possibly a combination of seismic and atmospheric processes contribute to the excitation although the combined excitation (Figure 5) is not much better than either one alone (see also Wilson and Haubrich, 1977). Inversion of the wobble data, after removal of the atmospheric contribution, for the seismic moments has not led to conclusive results, and it is not yet possible to infer parameters defining large seismic or aseismic events from past Chandler wobble data.

The third Chandler wobble question concerns its energy sink. Oceans are a probable sink but observational evidence for the departure of the pole tide from equilibrium remains totally inadequate. Wunsch (1974) has modelled dissipation of the pole tide in the North Sea and concludes that, by extrapolating to the world's shallow seas, the oceans may just provide an adequate sink. In the pole tide problem one requires the lag of the 2,1 harmonic in the tide expansion to describe the total rate of dissipation and a value of 5° is adequate to explain the observed wobble Q. But observed values for the lag are very variable and unreliable.

Significant dissipation in the mantle is generally ruled out because estimates for the wobble Q are less than the corresponding seismic Q's. Stacey (1977) has stressed, however, that a direct comparison of seismic and wobble Q's is not valid since the two are defined differently. The former is defined as being proportional to $\Delta E/E_e$ where $\Delta E$ is the amount of energy dissipated in one cycle of the motion and $E_e$ is the peak elastic energy stored in the cycle (see, for example, O'Connell and Budiansky, 1978). The

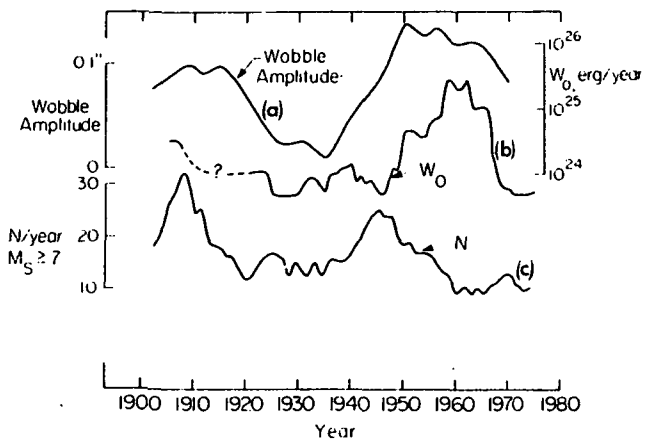



Fig. 4. Comparison of the amplitude of the Chandler wobble (curve a), the elastic energy realeased by earthquakes (5-year running means, curve b) and the annual number of earthquakes of $M_S \geq 7.0$ (5-year running means, curve c). From Kanamori (1977a).

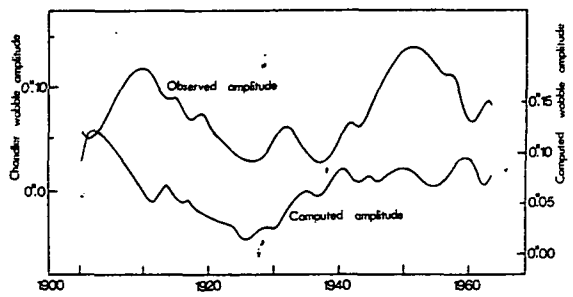



Fig. 5. Fluctuations in the Chandler wobble amplitude as observed (upper curve) and as computed from the seismic and atmospheric data. Note that the vertical scales of the two curves have been displaced.

wobble Q, however, is usually defined as being proportional to $\Delta E/E_k$, where $E_k$ is the kinetic energy of the rotational motion. But $E_e$ is about 15% of $E_k$ (Merriam and Lambeck. 1978) and a wobble Q of, say 100, implies a seismic Q of about 15. The lower Q values at the Chandler wobble frequency than at seismic frequencies may not be wholly unexpected if Anderson and Minster's Q model is valid.

Annual frequency

The comparison of meteorological and astronomical estimates of the seasonal polar motion remains unsatisfactory for several reasons, including (i) incomplete meteorological, oceanic and hydrological data, (ii) global year to year variability in these data, (iii) poor wobble data at the annual frequency prior to the introduction of the B.I.H. polar motion results. Siderenkov (1973), Jochmann (1976), Wilson and Haubrich (1976) have discussed the atmospheric contributions but significant differences between the results exist. The contribution from ground water storage is important but only very approximately known. Seasonal variations in the ocean volume also are not insignificant. The unsatisfactory situation is discussed in more detail by Lambeck (1978a).

Length of day

Astronomers observe the integrated amount by which the Earth is slow or fast relative to a uniform time scale. This gives the difference between Universal Time and Atomic Time, or UT-AT. Of greater geophysical interest is the proportional change in the length of day (1.o.d), or

$$m_3 = \frac{\omega_3 - \Omega}{\Omega} = -\frac{\Delta(1.o.d)}{1.o.d} = -\frac{d}{dt}(UT-AT),$$

where $\omega_3$ is the instantaneous rotation velocity and $\Omega$ is a nominal or mean velocity. Figure 6 illustrates the schematic spectrum of $m_3$. It consists mainly of a continuum upon which a number of periodic phenomena are superimposed.

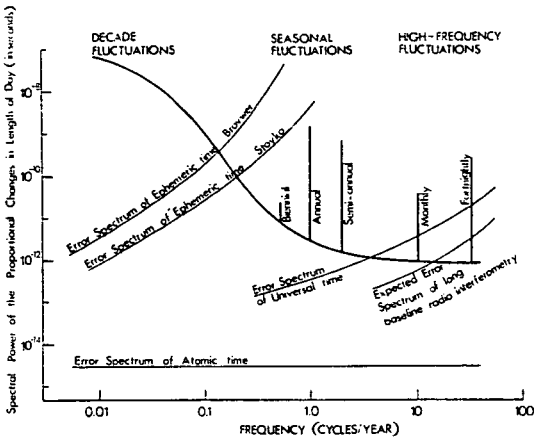



Fig. 6. Schematic power spectrum of the proportional changes in length of day.

Secular acceleration

The secular acceleration of the Earth's rotation, $\dot{\Omega}$, is mainly a consequence of the dissipation of energy of the Moon and Sun raised ocean tides (see Lambeck, 1977, for a review of the problem). A further contribution may come from slow changes in the mass distribution within the Earth or from non-tidal torques acting on the mantle. Denote these two contributions to $\dot{\Omega}$ by $\dot{\Omega}_T$ for the tidal part and $\dot{\Omega}_{NT}$ for the non-tidal part. Associated with $\dot{\Omega}_T$ is an acceleration $\dot{n}$ of the Moon in longitude and, as astronomical observations give $\dot{n}$ and $\dot{\Omega}$, the tidal contribution to the spin can be separated from the non-tidal part. This separation has been evaluated in detail by Lambeck (1975b, 1977). Table 4 summarizes recent results. The astronomical estimates of $\dot{\Omega}$ and $\dot{n}$ come mainly from the analysis of discrepancies between computed and recorded eclipse paths. Most of these records, made during the last three millenia, have been revised by Newton (1970, 1972), Stephenson (1972), Muller and

TABLE 4.

| | Astronomical Estimates | Tidal Estimates | Satellite Estimates |
|---|---|---|---|
| $\dot{n}_{\alpha}$ ($10^{-23}s^{-2}$) | -1.35±0.10 | -1.49±0.5 | -1.33±0.25 |
| $\dot{\Omega}_T\|_{n_{\alpha}}$ ($10^{-22}s^{-2}$) | -5.48 | -6.05 | -5.40 |
| $\delta\dot{\Omega}_T$ ($10^{-22}s^{-2}$) | -1.47 | -1.47 | -1.47 |
| $\dot{\Omega}_T$ ($10^{-22}s^{-2}$) | -6.95 | -7.52 | -6.87 |
| $\frac{dE}{dt}$ ($10^{19}$ergs $s^{-1}$) | -3.94±0.30 | -4.26±0.45 | -3.90±0.70 |

Recent astronomical, tidal and satellite estimates of the tidal and non-tidal accelerations of the Earth and Moon. $\delta\dot{\Omega}_T$ is the contribution to $\dot{\Omega}_T$ from tidal terms other than those contributing to $\dot{n}$.

Stephenson (1975), and Muller (1975, 1976), with the consequence that we have not only a more extensive data set but also more confidence in the reliability of the results. Muller (1975, 1976) estimates that a principal uncertainty now is due to the constants used in the lunar theory. Telescope observations give $\dot{n}$ but not $\dot{\Omega}$ since they cover too short a time interval (about 170 years) to permit a separation of $\dot{\Omega}$ from long-period fluctuations (Morrison, 1978). Perturbations in the motions of artificial Earth satellites, due to the ocean and body tide potentials, also permit $\dot{n}$ to be estimated (Lambeck, 1975; Cazenave et al. 1977). Recent estimates of tidal parameters are by Cazenave and Daillet (1977), Goad and Douglas (1978) and Felsentreger et al. (1978). The satellite deduced results for $\dot{n}$ are comparable with those obtained from the astronomical observations. Lunar laser ranging also gives a comparable results for $\dot{n}$ (Williams et al., 1978).

At present, $\dot{\Omega}_T \simeq 53.1\dot{n}$ (Lambeck, 1977) and this value has been used to determine the tidal acceleration $\dot{\Omega}_T$. The difference between this value and the observed acceleration is small and uncertain and Muller (1976) attributes most of it to a change in the gravitational constant. He finds, for different cosmologies, $\dot{G}$ of the order 2-5 parts in $10^{11}$ year$^{-1}$ from an analysis of historical records of astronomical events, telescope observations and tidal theory.

There has never been a shortage of geophysical explanations for the non-tidal acceleration but this quantity now appears to be sufficiently small to conclude that there have not been significant changes in the Earth's moment of inertia during the past 3000 years. Improvements in the value of $\dot{\Omega}_{NT}$ will come only if $\dot{\Omega}$ can be determined with greater accuracy and this will only come about if the record of eclipse observations can be extended. The method used by Sawyer and Stephenson (1970), to locate the valuable Ugarit record of 1375 BC, leaves hope that new observations may still be found in, for example, the Sumarian records. The Chinese records of the court astronomers have been fully explored for astronomical references by Stephenson (1972) but, according to C.P. Fitzgerald, provincial records may provide valuable additional accounts of eclipses back to about 200 AD. Furthermore, ancient Chinese astronomical records of occulatations and conjunctions are many (e.g. Ho, 1966) and these have not yet been fully investigated for purposes of determining $\dot{n}$ and $\dot{\Omega}$.

The paleontological evidence for the Earth's past rotation has been reviewed by Lambeck (1978a, d) and Scrutton (1978). Since 1973 there have been few new observations of the frequencies of growth rhythms and the interval has been characterized by caution not always evident in the earlier studies (Clarke, 1974; and papers in the volume edited by Rosenberg and Runcorn, 1975). Lambeck (1978d) has estimated $\dot{\Omega}$ and $\dot{n}$ from the better documented coral and bivalve growth rhythms since the Ordovician and finds that $\dot{\Omega}$ has, on the average, been about 75% of the present value and that there is no evidence in the data for a significant non-tidal acceleration during the last $4\times10^8$ years (Table 5). The results

TABLE 5.

| | $\dot{\Omega}$ $(10^{-22}s^{-2})$ | $\dot{n}_{\ell}$ $(10^{-23}s^{-2})$ |
|---|---|---|
| Coral Data | -6.3±0.7 | -1.5±0.4 |
| Bivalve Data | -5.9±0.6 | -1.3±0.3 |
| Combined Data | -5.2±0.2 | -1.2±0.2 |
| With constraint that $\dot{\Omega}_{NT}\simeq 0$ | -5.2±0.2 | -1.0±0.1 |

Summary of estimates of the Paleorotation of Moon and Earth.

and some geophysical consequences are discussed further in Lambeck (1978c).

## Long period fluctuations

Evidence for the long period (from about 10 to 300 years) comes from the telescope observations made since the eighteenth century and, according to F.R. Stephenson, there is some hope that timed eclipse observations may aid in extending the record further back into time. Morrison (1972) has transformed the data since 1820 into a uniform system and is now carrying out a complete revision of the data. Figure 7 illustrates the results for $m_3$ and its derivative $\dot{m}_3$ (see Lambeck, 1978a). The causes of these changes remain obscure. Electromagnetic coupling of core motions to the mantle, possibly reinforced by topographic coupling, is the main contender. In particular, they can explain the changes occurring over periods of several decades and longer (Yukutake, 1973; Watanabe and Yukutake, 1975). The main unknowns in the theory are the strength of the torroidal field at the core-mantle boundary and the value

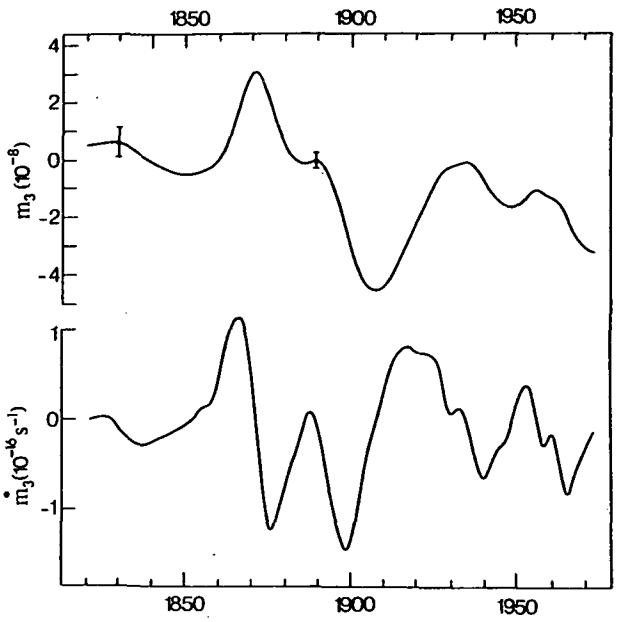



Fig. 7. $m_3$ and $dm_3/dt$ based on telescope observations from 1820 to 1970.

for the lower-mantle electrical conductivity $\delta_m$. Values of $2\text{-}3\times10^2$ ohm $m^{-1}$ are required to explain the decade fluctuations but several recent investigators has suggested that it may be much higher than this (Kolomiyseva, 1972; Braginski and Nikolaichik, 1973; Alldredge, 1977; Stacey et al. 1978). The observational evidence, however, remains inadequate (see, for example, Courtillot and LeMouel, 1977). An argument by Hide (private communication, July 1978) also opposes these high conductivity values, on the grounds that his proposed method of estimating the core radius from magnetic data (Hide, 1978) would otherwise not work, and his good agreement with the seismic estimate for the core radius would have to be considered as fortuitous.

Topographic coupling was proposed by Hide (1969) when electromagnetic coupling appeared inadequate. Hide (1977) has reviewed the state of the mechanism. To estimate the total torque on the mantle exerted by the flow past the irregular core-mantle boundary, requires the form of the topography and the flow-field. Even their power spectra are unknown although, from considerations of the Earth's gravity field, it does not appear that the height of the irregularities can exceed a few hundred meters (Lambeck, 1976). The mechanism, however, does not necessarily require large values for the topography since the drag coefficient is dependent on the topography itself and high topography does not necessarily imply strong topographic coupling. Moffatt (1978) has investigated the dependence of the drag coefficient on the strength of the magnetic field within the core and concludes that this coefficient may be significantly greater than unity, thereby enhancing the effectiveness of the topographic coupling mechanism without requiring large amplitude topography.

Variations on the decade time scale in the atmospheric and oceanic mass redistributions have been evaluated by Lambeck and Cazenave (1976) but, while they find variations that exhibit similar trends to those observed in length of day, the magnitudes are inadequate. In particular, periods of an accelerating Earth, as occurred from 1840 to 1870 and again from 1900 to 1940 are associated with increasing strength in the zonal wind circulation and with increasing surface temperatures, while periods of deceleration occur when there is a general decrease in the strength of the zonal wind circulation and surface temperatures decrease (Figure 8). These correlations require further study but they do raise the possibility that length of day changes and the global climatic changes, over intervals of a few decades, may have a common origin.

## Seasonal and higher frequency variations

The Earth's rotation appears to exhibit considerable fluctuation in speed due to its apparent sensitivity to changes in the zonal wind circulation. Thus the astronomical observations may provide useful measures of the strengths and frequencies of the global wind circulation. The atmospheric induced changes in length of day appear to be important for periods of several years down to a few days and, as such, they need to be known if other excitation mechanisms are investigated. Examples of such excitations include the study of changes in length of day caused by large earthquakes and aseismic motion or by core-mantle coupling. Lambeck and Cazenave (1977) have reviewed the subject. The seasonal changes in the length of day are clearly of atmospheric origin, due to a periodic exchange of angular momentum between the zonal wind circulation and the Earth's mantle (Lambeck and Cazenave, 1973). Changes in the inertia tensor associated with atmospheric, oceanic and hydrological mass redistributions, play only a minor role at the semi-annual frequency are tidal contributions significant. The astronomical data since 1955 have indicated that there have been some significant changes in the year-to-year amplitudes of the seasonal terms and this is indicative of a substantial change in the year-to-year circulation of the atmosphere. The significance of the year to year fluctuation in the semi-annual oscillation amplitude in length of day is less clear, since Okazaki (1975) has suggested that part of this may be a consequence of the astronomical reduction process, at least for the data prior to 1962. The length of day observations also suggest an intermittent quasi-biennial oscillation, variable in both period and in amplitude. This is a further consequence of an exchange of angular momentum between the mantle and the atmosphere. Meteorological evidence for longer period global wind cycles is minimal, particularly for winds at higher altitudes. Siderenkov (1969) has used the surface torque approach to estimate the angular momentum changes in the atmosphere and finds some evidence that length of day changes over roughly 5 year periods are a consequence, at least in part, of zonal winds. Okazaki (1977) came to a similar conclusion as did Lambeck and Cazenave (1974). Recently Rosen et al. (1976) computed the yearly

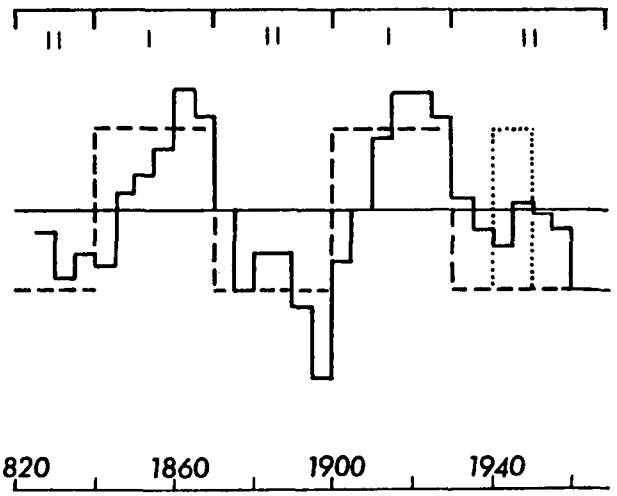



Fig. 8. Schematic representation of the circulation pattern (broken line). Periods of increasing strength of the zonal circulation are denoted by I, those of decreasing circulation are denoted by II. The astronomically observed $\dot{m}_3$, running mean values over 15 years, are defined by the solid line.

value of the zonal angular momentum in the northern hemispheric circulation for a ten year period. The fluctuations are comparable to those deduced from the l.o.d. but smaller (Figure 9). This is further evidence for an important atmospheric excitation functions for l.o.d changes over periods up to about 10 years and stresses the need once more for detailed atmospheric angular momentum calculations if one wants to study core-mantle coupling mechanisms.

At frequencies above 2 cycles per year, the length of day data shows considerable perturbation and much of this can be attributed to the zonal winds (Lambeck and Cazenave, 1974; Okazaki, 1977). Studies of the global atmospheric circulation spectrum show a broad continuum from about 8 months to a few days (Mitchell, 1976) with peaks rising above the average near the semi-annual and fortnightly frequencies. The circulation patterns become more regional at the higher frequencies and it is not clear at what frequency they cease to have inconsequential effects on the rotation. Abrupt aperiodic changes in circulation patterns are often observed and may cause some of the irregular week-to-week changes in rotation occasionally reported (e.g. Guinot, 1968). Here, however, both the astronomical and meteorological data become less reliable and a more detailed and precise compilation of the zonal winds is required for all latitudes and altitudes.

## Precession and Nutation

Revisions of E.W. Woolards rigid body nutation theory have been carried out by Kinoshito (1976) and Murray (1978) but the interest of the subject lies mainly in the departures from this theory due to the Earth's non-rigid response to the solar and lunar torques. The symposium proceedings edited by Federov et al. (1978) discuss various aspects of these motions.

The most complete solution for the forced nutations of realistic Earth models is by Sasao et al. (1977, 1978) although the theory by Smith (1977) could be readily extended to include these forced terms. Sasao et al's theoretical

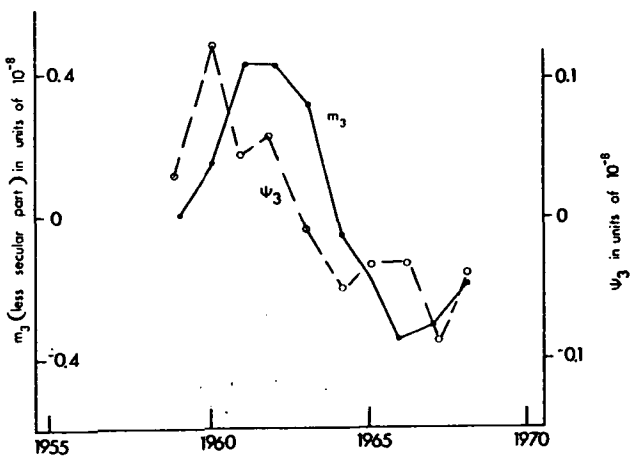



Fig. 9. Estimates of the zonal angular momentum in the northern hemispheric circulation (from Rosen et al. 1976) and changes in length of day.

TABLE 6.

| | Period in mean solar days | | | | |
|---|---|---|---|---|---|
| | -365.3 | -6798 | 6798 | 182.6 | 13.7 |
| Observed* amplitude | 1.336 | 0.9962 | 1.0030 | 1.034 | 1.031 |
| Theoretical* amplitude | 1.283 | 0.9965 | 1.0030 | 1.034 | 1.031 |

* based on the mean of values given by Sasao et al.

Summary of the theoretical nutation amplitudes (Sasao et al. 1977) and a comparison with the observed periods. The amplitudes are normalized by the rigid body value.

nutuation periods are generally in satisfactory agreement with the observed values (Table 6) although the latter are subject to some uncertainty.

A second nutuation has received considerable attention in the recent literature. This is the nearly diurnal wobble, or what Toomre (1974) calls the principal core nutation. This motion, already discussed in the last century, received only intermittent attention until N.A. Popov claimed to have detected it in the astronomical observations. The motion, a free wobble due to the Earth's liquid core, is a retrograde motion of the rotation axis about the axis of figure and is accompanied by a very much larger motion of the rotation axis in space with a period of roughly 400 days (Toomre, 1974). Yatskiv (1972) and Rochester et al. (1974) summarized the subsequent papers dealing with the search for this wobble in the latitude observatios while Rochester et al. and Capitaine (1975) have searched for it in the declination observations without success. The absence of evidence of this motion in space then indicates that the corresponding wobble amplitude becomes totally insignificant. Yatskiv et al. (1975) object to this conclusion but their argument is obscure.

Calculations of the period of the nearly diurnal wobble have most recently been made for realistic Earth models by Smith (1977) and Sasao et al. (1977, 1978) and the results are not very model dependent. This appears to make the whole problem rather academic.

## Some further problems

### Secular decrease in obliquity

Astronomical observations have revealed a "secular" change in the obliquity, most of which is really a long period perturbation in the Earth's motion due to the gravitational attractions of the planets. A small discrepancy between the observed and the gravitational results has been explained by frictional coupling between the core and mantle during the precession period (Aoki, 1969; Kakuta and Aoki, 1972). A revision of the astronomical data by Duncombe and van Flandern (1976) now points to an insignificant discrepancy and the theoretical study of Rochester (1976) also indicates an insignificant role of core mantle coupling during the

orbital evolution of the Earth-Moon system. This conclusion can also be deduced from the studies by Sasao et al. (1977, 1978).

Evolution of the Earth-Moon system

Kaula and Harris (1975) have given a recent review of the dynamical evolution of the Earth-Moon system and of the possible constraints that this may place on the origin of the Moon. These constraints do not appear to be important. The main problem remains the inclination problem: How to get the Moon into an equatorial orbit when it was close to the Earth. The rate of change of the orbital inclination is the sum of perturbations due mainly to three tidal terms ($M_2$, $K_1$, $O_1$), with two tending to decrease ($M_2$ and $K_1$) and one ($O_1$) tending to increase the inclination with time (Lambeck, 1975) and situations can be contrived in which the overall sign if dI/dt changes simply by introducing an appropriate frequency dependent Q law. Rubincam (1975) discusses one possibility, namely a Maxwell Earth in which Q is proportional to frequency. Anderson and Minster's (1978) suggestion of Q proportional to (frequency)$^{1/3}$ may do the same thing, if the tidal frequencies at a given moment are such as to accentuate the $O_1$ relative to $M_2$ and $K_1$ tides.

References


Aardoom, L., Laser tracking, (This volume), 1978.

Alldredge, L.R., Deep mantle conductivity. J.Geophys.Res., 82, 5427-5431, 1977.

Anderson, D.L. and J.B. Minster, The frequency dependence of Q in the Earth and implications for mantle rheology and Chandler wobble, (in press), 1978.

Anderson, D.L., H. Kanamori, R.S. Hart and H.P. Liu, The Earth as a seismic absorption band, Science, 196, 1104-1106, 1977.

Aoki, S., Friction between mantle and core of the Earth as a cause of the secular change in obliquity, Astron.J., 74, 284p, 1969.

Braginski, S.I. and V.V. Nikolaichik, Estimation of electrical conductivity of the lower mantle of the earth by electromagnetic signal delay, Phys.Solid Earth, 601-602.

Brosche, P. and J. Sundermann, Effect of oceanic tides on the rotation of the Earth, in Scientific Applications of Lunar Laser Ranging (ed. J.D. Mullholland) 133- 141, D. Reidel Publ.Co., Dordrecht, 1977.

Cannon, W.H., The Chandler annual resonance and its possible geophysical significance, Phys. Earth Planet.Int., 9, 83-90, 1974.

Capitaine, N., Total effect of any nearly-diurnal wobble of the Earth's axis of rotation in latitude and time observations, Geophys.J., 43, 573-588, 1975.

Cazenave, A. and S. Daillet, Determination de la maree oceanique $M_2$ avec STARLETTE. Internal Rpet.Groupe Rech.Geod.Spat., Centre Nationale d'Etudes Spatiales, Toulouse, 39 pp, 1977.

Cazenave, A., S. Daillet and K. Lambeck, Tidal studies from the perturbations in satellite orbits, Trans.Roy.Soc.Lond., A284, 595-606, 1977.

Chen, W.P. and P. Molnar, Seismic moments of major earthquakes and the average rate of slip in Central Asia, J.Geophys.Res., 82, 2945-2969, 1977.

Clark, G.R., Growth lines in invertebrate skeletons, Ann.Revs.Earth Planet.Sci., 2, 77-99, 1974.

Courtillot, V. and J-L Le Mouel, On the long period variations of the Earth's magnetic field from 2 months to 20 years, J.Geophys. Res., 81, 2941-2950, 1976.

Currie, R.G., The spectrum of sea level from 4 to 40 years, Geophys.J., 46. 513-520, 1976.

Dahlen, F.A., The passive influence of the oceans upon the rotation of the Earth, Geophys.J., 46, 363-406, 1976.

Dahlen, F.A., The period of the Chandler wobble. in Federov et al. 1978.

Duncombe, R.L. and T.C. Van Flandern, The secular variation of the obliquity of the ecliptic, Astron.J., 81, 1976.

Dziewonski, A.M. and F. Gilbert, Temporal variation of the seismic moment tensor and the evidence of precursive compression for two deep earthquake, Nature, 247, 185-188, 1974.

Federov, E.P., M.L. Smith and P.L. Bender, editors, Nutation and the Earth rotation. D. Reidel Publ.Co., Dordrecht (in press), 1978.

Felsentreger, T.L., J.G. Marsch and R.G. Williamson, Tidal perturbations on the satellite 1967-92A, J.Geophys.Res., 83, 1837-1842, 1978.

Gilbert, F., Excitation of the normal modes of the Earth by earthquake sources, Geophys.J., 22, 223-226, 1971.

Goad, C.C. and B.C. Douglas, Lunar tidal acceleration obtained from satellite-derived ocean tide parameters, J.Geophys.Res., 83, 2306-2310, 1978.

Goldreich, P. and A. Toomre, Some remarks on polar wandering, J.Geophys.Res., 74, 2555-2567, 1969.

Guinot, B., Work of the Bureau International d l'Heure on the rotation of the Earth, in L. Mansinha et al. (1970), 54-62, 1970b.

Guinot, B., Services, (This volume), 1978.

Hide, R., Interaction between the earth's liquid core and solid mantle, Nature, 222, 1055-1056, 1969.

Hide, R., Towards a theory of irregular variations in the length of day and core-mantle coupling, Phil.Trans.Roy.Soc.Lond., A284, 547-554, 1977.

Hide, R., How to locate the electrically conducting fluid core of a planet from external observations, Nature, 271, 640-641, 1978.

Ho, P-Y., The Astronomical Chapters of the Chin Shu, Mouton & Co., Paris, 271 pp, 1966.

Hosoyama, K., I. Naito and N. Sato, Tidal admittance of the pole tide, J.Phys.Earth Japan, 24,

43-50, 1976.
Jeffreys, H., The variation of latitude, Monthly Notices Roy.Astron.Soc., 141, 255-268, 1968).
Jochmann, H., Der einfluss von luftmassenbewegungen in der Atmosphäre auf die polbewegung, Veröff Zentalinstituts Physik Erde, Akad.Wiss.D.D.R., 35, 1-38, 1976.
Kakuta, C. and S. Aoki, The excess secular change in the obliquity of the ecliptic and its relation to the internal motion of the Earth, IAU Symposium 48, D. Reidel Publ.Co., Dordrecht, 192-195, 1972.
Kanamori, H., Mechanism of Tsunami earthquakes, Phys.Earth Planet.Interiors, 6, 346-359, 1972.
Kanamori, H., Are earthquakes a major cause of the Chandler wobble, Nature, 262, 254-255, 1976.
Kanamori, H., The energy release in great earthquakes, J.Geophys.Res., 82, 2981-2987, 1977a.
Kanamori, H., Seismic and aseismic slip along subduction zones and their tectonic implications, in Island Arcs, Deep Sea Trenches, and Back-Arc Basins, Vol.1, (ed. M. Talwani and W.C. Pitman), 163-174, AG, Washington, 1977b.
Kanamori, H. and D.L. Anderson, Theoretical basis of some empirical relations in seismology, Bull.Seis.Soc.Am., 65, 1073-1095, 1975b.
Kanamori, H. and D.L. Anderson, Importance of physical dispersion in surface wave and free oscillation problems, Revs.Geophys.Space Phys., 15, 105-112, 1977.
Kanamori, H. and J.J. Cipar, Focal process of the great Chilean earthquake May 22, 1960, Phys. Earth Planet.Int., 9, 128-136, 1974.
Kaula, W.M. and A.W. Harris, Dynamics of lunar origin and orbital evolution, Rev.Geophys. Space Phys., 13, 363-371, 1975.
Kaula, W.M., K. Lambeck, W. Markowitz, et al., The rotation of the Earth and polar motion, EOS, Trans.Am.Geophys.Union, 54, 792-798, 1973.
Kinoshito, H., Formulas for precession, Smithson. Astrophys.Obs.Spec.Rept. 364, 25pp, 1975.
Kolomiytseva, G.I., Distribution of electric conductivity in the mantle of the Earth, according to data on secular geomagnetic field variations, Geomagnetism and Aeronomy, 12, 938-941, 1972.
Lambeck, K., Effects of tidal dissipation in the oceans on the Moon's orbit and the Earth's rotation, J.Geophys.Res., 80, 2917-2925, 1975a.
Lambeck, K., The Chandler annual resonance, Phys. Earth Planet.Int., 11, 166-168, 1975b.
Lambeck, K., Lateral density anomalies in the mantle, J.Geophys.Res., 81, 6333-6340, 1976.
Lambeck, K., Tidal dissipation in the oceans astronomical, geophysical and oceanographic consequences, Phil.Trans.Roy.Soc.Lond.A287, 545-594, 1977.
Lambeck, K., The Earth's Variable Rotation: Geophysical causes and consequences, Cambridge University Press, (in press), 1978a.
Lambeck, K., Methods and geophysical applications of satellite geodesy, Rept.Prog.Phys. (in press), 1978b.
Lambeck, K., The history of the Earth's rotation, in The Earth: Its origin, Structure and Evolution, (ed. M.W. McElhinny), Academic Press, 1978c.
Lambeck, K., The Earth's Paleorotation, in Tidal Friction and the Rotation of the Earth, (ed. P. Brosche and J. Sunderman), Springer-Verlag, Berlin, 1978d.
Lambeck, K. and A. Cazenave, The Earth's rotation and atmospheric circulation - I, Seasonal variations, Geophys.J., 32, 79-93, 1973.
Lambeck, K. and A. Cazenave, The Earth's rotation and atmospheric circulation - II, The Continuum, Geophys.J., 38, 49-61, 1974.
Lambeck, K. and A. Cazenave, Long term variations in length of day and climatic change, Geophys. J., 46, 555-573, 1976.
Lambeck, K. and A. Cazenave, The Earth's variable rate of rotation: a discussion of some meteorological and oceanic causes and consequences, Phil.Trans.Roy.Soc.Lond.,A284, 495-506, 1977.
Liu, H-P., D.L. Anderson and H. Kanamori, Velocity dispersion due to anelasticity; implications for seismology and mantle composition, Geophys. J., 47, 41-58, 1976.
Mansinha, L., D.E. Smylie and A.E. Beck, (editors) Earthquake Displacement Fields and the Rotation of the Earth, D. Reidel Publ.Co., Dordrecht, 308 pp, 1970.
Mansinha, L., D.E. Smylie and C.H. Chapman, Seismic excitation of the Chandler wobble revisited, Submitted to Geophys.J., 1978.
Markowitz, W., Sudden changes in rotational acceleration of the Earth and secular motion of the pole, in Mansinha, Smylie and Beck, 69-81, 1970.
McElhinny, M.W., Palaeomagnetism and Plate Tectonics, Cambridge University Press, 358 pp, 1973.
Merriam, J.B. and K. Lambeck, Comments on the Chandler wobble Q, Submitted to Geophysical Journal, Royal Astron.Soc., 1978.
Miller, S.P. and C. Wunsch, The Pole Tide, Nature, 246, 97-102, 1973.
Mitchell, J., An overview of climatic variability and its causal mechanisms, Quaternary Res., 6, 481-491, 1976.
Moffatt, H.K., Topographic coupling at the core-mantle interface, Geophys.Astrophys.Fluid Dyn., 9, 279-288, 1978.
Morrison, L.V., Tidal deceleration of the Earth's rotation deduced from astronomical observations in the period AD 1600 to the present, in Tidal Friction and the Earth's Rotation, (ed. P. Brosche and J. Sundermann), Springer-Verlag, 1978.
Muller, P.M., An Analysis of the Ancient Astronomical Observations with the Implications for Geophysics and Cosmology, Thesis, University of Newcastle.
Muller, P.M., Determination of the cosmological rate of change of G and the tidal accelerations of Earth and Moon from ancient and modern astronomical data, Report SP 43-36, Jet Propulsion Laboratory, Pasadena, California, 24 pp, 1976.
Muller, P.M. and F.R. Stephenson, The acceleration of the Earth and Moon from early astronomical observations, in Rosenberg and Runcorn, 459-534, 1975.
Munk, W.H. and G.J.F. MacDonald, The Rotation of

the Earth, University Press, Cambridge, 323 pp, 1960.
Murray, C.A., On the precession and nutation of the Earth's axis of figure, Mon.Not.R.Astron. Soc., 183, 677-685, 1978.
Newton, R.R., Ancient Astronomical Observations and the Accelerations of the Earth and Moon, Johns Hopkins University Press, Baltimore, 309 pp, 1970.
Newton, R.R., Medieval Chronicles and the Rotation of the Earth, Johns Hopkins University Press, Baltimore, 825 pp, 1972.
O'Connell, R.J. and B. Budiansky, Measures of attentuation in dissipative media, Geophys. Res.Lett., 5, 5-8, 1978.
O'Connell, R.J. and A.M. Dziewonski, Excitation of the Chandler wobble by large earthquakes, Nature, 262, 259-262, 1976.
Okazaki, S., On the amplitude changes of seasonal components in the rate of rotation of the Earth, Publ.Astron.Soc.Japan, 27, 367-378, 1975.
Okazaki, S., On a relation between irregular variations of the Earth's rotation and anomalous changes of the atmospheric circulation, Publ. Astron.Soc.Japan, 29, 619-626, 1977.
Rochester, M.G., The Earth's rotation, EOS.Trans. Amer.Geophys.Union, 54, 769-780, 1973.
Rochester, M.G., The secular decrease of obliquity due to dissipative core-mantle coupling, Geophys.J., 46, 109-126, 1976.
Rochester, M.G., O.G. Jensen and D.E. Smylie, A search for the Earth's nearly diurnal free wobble, Geophys.J., 38, 349-363, 1974.
Rosen, R.D., M-F. Wu and J.P. Peixoto, Observational study of the interannual variability in certain features of the general circulation, J.Geophys.Res., 81, 6383-6389, 1976.
Rosenberg, G.D. and S.K. Runcorn, (ed). Growth Rhythms and the History of the Earth's Rotation, John Wiley and Sons, 559 pp, 1975.
Rubincam, D.P., Tidal friction and the early history of the Moon's orbit, J.Geophys.Res., 80, 1537-1548, 1975.
Sasao, T., I. Okamoto and S. Sakai, Dissipative core-mantle coupling and nutational motion of the Earth, Publ.Astron.Soc.Japan, 20, 83-105, 1977.
Sasao, T., S. Okubo and M. Saito, A simple model on dynamical effects of stratified fluid core upon nutational motion of the Earth, 1977, in Federov et al. 1978.
Sawyer, J.F.A. and F.R. Stephenson, Literary and astronomical evidence for a total eclipse of the Sun observed in Ancient Ugarit on 3 May 1375 BC, Bull.School Oriental African Studies, 33, 467-489, 1970.
Scrutton, T., Periodic growth features in fossil organisms and the length of the day and month, in Tidal Friction and the Rotation of the Earth, (ed. P. Brosche and J. Sunderman), Springer-Verlag, (in press), 1978.
Shapiro, I.I., VLBI (this volume), 1978.
Siderenkov, N.S., The influence of atmospheric circulations on the Earth's rotational velocity, Sov.Astron., 12, 706-714, 1969.
Siderenkov, N.S., The inertia tensor of the atmosphere, the annual variations of its components and the variations of the Earth's rotation, Izv.,Atmos.Ocean Phys., 9, 339-351, 1973.
Smith, M.L., Wobble and nutation of the Earth, Geophys.J., 50, 103-140, 1977.
Stacey, F.D., Physics of the Earth, second edition, John Wiley and Sons, New York, 414 pp, 1977.
Stacey, F.D., H.W.S. McQueen, D.E. Smylie, M.G. Rochester and D. Conley, Spectral character of the geomagnetic field and the electrical conductivity of the lower mantle, (in preparation), 1978.
Stephenson, F.R., Some Geophysical, Astrophysical and Chronological Deductions from Early Astronomical Records, Thesis, University of Newcastle, 1972.
Stuart, W.D. and M.J.S. Johnston, Tectonic implications of anomaous tilt before central California earthquakes, (Abstract). Trans.Am. Geophys.Union, 58, 1196 pp, 1974.
Thatcher, W., Strain release mechanism of the 1906 San Francisco earthquake, Science, 184, 1283-1285, 1974.
Toomre, A., On the nearly diurnal wobble of the Earth, Geophys.J., 38, 335-348.
Vicente, R.O. and R.G. Currie, Maximum entropy spectrum of long period polar motion, Geophys. J., 46, 67-73, 1976.
Watanabe, H. and T. Yukutake, Electromagnetic core-mantle coupling associated with changes in the geomagnetic dipole field, J.Geomag. Geoelectri., 27, 153-175, 1975.
Williams, J.G., W.S. Sinclair and C.F. Yoder, Tidal acceleration of the Moon, (in preparation) 1978.
Wilson, C.R. and R.A. Haubrich, Atmospheric contributions to the excitation of the Earth's wobble, 1901-1970, Geophys.J., 46, 745-760, 1976.
Wilson, C.R. and R.A. Haubrich, Earthquakes, weather and wobble, Geophys.Res.Lett., 4, 283-284, 1977.
Wunsch, C., Dynamics of the pole tide and the damping of the Chandler wobble, Geophys.J., 39, 539-550, 1974.
Yatskiv, Ya.S., On the comparison of diurnal nutation derived from separate series of latitude and time observations, in Melchior and Yumi, 200-205, 1972.
Yatskiv, Ya.S., Y. Wako and Y. Kaneko, Study of the Nearly Diurnal Free Nutation Based on Latitude Observations of the ILS Stations (I), Publ.int.Latit.Obs.Mizusawa, 10, 1, 1975.
Yukutake, T., Fluctuations in the Earth's rate of rotation related to changes in the geomagnetic dipole field, J.Geomag.Geoelectr., 25, 195-212, 1973.

# Rotation of the Earth and Polar Motion, Services

Bernard Guinot
Bureau International de l'Heure
61, Avenue de l'Observatoire, Paris 75014, France

Abstract. The need of a continuous monitoring of the polar motion appeared at the end of the 19th century, and was at the origin of one of the oldest international projects : the establishment of the International Latitude Service. This service still operates, but other organizations now determine the polar motion, using the astrometric measurements of latitude and time and also Doppler observations of artificial satellites. On the other hand, since the advent of atomic clocks in 1955, the universal time has become a measure of the rotation of the Earth, which is also currently required and which must be evaluated by a Service. The services providing polar motion and universal time data will be described, the precision and accuracy of these data will be estimated.

## Introduction

The full description of the rotation of the Earth in space is traditionally given by the motion of the rotation axis with respect to the Earth (polar motion), and in space (luni-solar precession, nutation), and by the angular position around the rotation axis (universal time UT1). Precession and nutation can be fairly well modeled, and require only occasional improvements of their representation ; polar motion and universal time are still unpredictable and require continuous monitoring.

While for several decades, since 1900, the International Latitude Service (ILS) was the only source of the pole coordinates x and y, the development of new astrometric instruments led to the organization of a new service, the International Polar Motion Service (IPMS) in 1962. But the advent of atomic clocks in 1955 made obsolete the division of the work between polar motion and universal time. The Bureau International de l'Heure (BIH), in charge of universal time, began in 1955 to determine its own set of coordinates of the pole needed in the evaluation of UT1. On the other hand, the successful recovery of the pole coordinates using Doppler observations of Transit satellites led national organizations of the USA to determine routinely these coordinates.

Thus the user has the choice among several sets of pole coordinates (but there is only one for UT1), which, of course, differ. This is often considered as a nuisance. But it has, at least, one important advantage : it is a warning against too much faith in the published results. We have too many examples of analyses and interpretations where the limitations of the ILS data where ignored.

## Methods

The astrometric methods refer to the directions of the plumb-lines of the observatories. We will assume that the Earth is rigid and that these directions are fixed within the Earth. [This is not true ; some motions due to the luni-solar attraction can be modeled, but others due to various local causes and to the plate motions are not sufficiently known to correct the observations.] Figure 1 shows on an auxiliary sphere of unit radius the directions of the instantaneous North pole P, of a reference fixed pole $P_o$ and of the zenith Z of a station. The origin of the astronomical longitudes is a fixed point O on the equator of $P_o$. We can see that the astronomical latitude $\varphi$ and longitude L vary with the coordinates of the pole x and y. The universal time UT1 is simply linked to the angular motion in space of the PO meridian around P ; it is therefore dependent on x and y. The fundamental equations used in classical services are, for each station i, UTC being the worldwide time reference,

$$x(t)\cos L_{o,i} + y(t)\sin L_{o,i} = \varphi_{t,i} - \varphi_{o,i} \quad (1)$$

$$[-x(t)\sin L_{o,i} + y(t)\cos L_{o,i}]\tan\varphi_{o,i} + [UT1-UTC]_t = [UTO_i - UTC]_t \quad (2)$$

where $\varphi_{o,i}$ is the initial fixed latitude. $L_{o,i}$, the initial longitude, does not explicitly appear, but it is used in deriving $UTO_i$, which is therefore computed assuming x = y = 0.

Except for the ILS, the computations of $\varphi_{t,i}$, $[UTO_i - UTC]_t$ are made by the contributing observatories themselves, using the values of astronomical constants recommended by the International Astronomical Union. The role of the central services is thus to combine the equations (1) and (2). There is no standard procedure to accomplish that ; the main choices are related to

- the weighting factors, according to the quality of the observations,
- the choice of initial latitudes and longitudes,

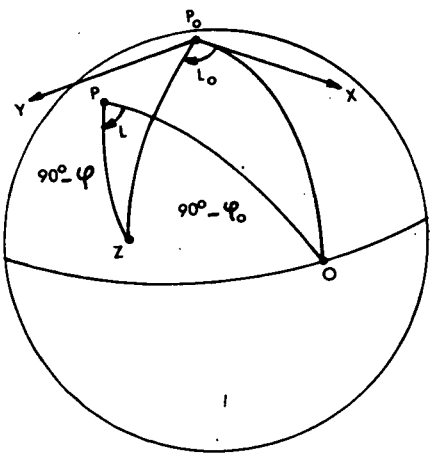



Fig. 1. Variation of astronomical latitude and longitude with the coordinates of the pole.

Proc. of the 9th GEOP Conference. *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

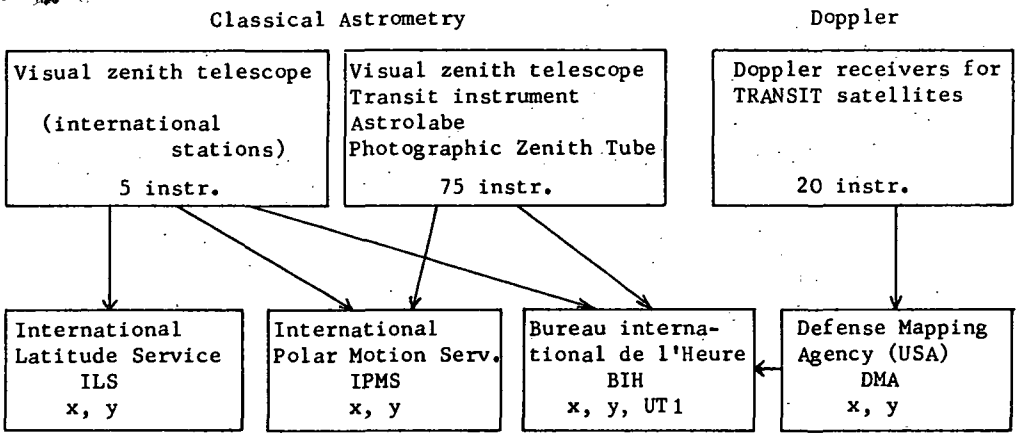



Fig. 2. Data flow to the services.

- the way of removing some systematic errors,
- the averaging time,
- the smoothing techniques on observational data and evaluated results.

The Doppler determinations of polar motion is an example of the satellite techniques for the study of Earth rotation, which are discussed by Aardoom [1978] at this Conference. Therefore we will only point out that in these techniques, instead of referring the observations directly to the quasi-non-rotating directions of stars, one uses intermediary objects, the directions of which are computed in the non-rotating reference frame. Providing that the motion of the object can be correctly modeled, UT1, x and y can be derived from the observations. In practice, a spurious drift of UT1 cannot be avoided, but the coordinates of the pole can be obtained to a large extent free from systematic errors and drifts.

Although astrometry only measures angles, we will use the meter as the unit for polar motion, assuming that the radius of the auxiliary sphere is the polar radius of the Earth.

Figure 2 shows the general organization of the services in 1978.

## Precision and accuracy of the results

A matter of importance is the degree of confidence the user may have in the results published by the Services. This problem is not easily solved. In most cases, the data are given without any information on their precision. In some cases a standard deviation is given, computed from the internal consistency of the data contributing to the determination of a raw value over a given averaging time $\tau$. But even when this information is given, it is far from being sufficient : it is well known that the errors are not a white noise and that the standard deviation does not vary proportionally to $1/\sqrt{\tau}$ .

Let us suppose that $A_o$, $A_1$,...,$A_n$ are measured quantities obtained at instants $t_o$, $t_o + \tau$,..., $t_o + n\tau$, by averaging over intervals $\tau$. If $a_o$, $a_1$,..., $a_n$ are the random errors of these quantities, it is possible to characterize the random noise by the pair variance (or Allan variance)

$$\sigma_a^2(\tau) = \text{mean of } \frac{(a_{i+1} - a_i)^2}{2} .$$

This function is represented by stability curves (Fig. 3), as it is customary for the characterization of the stability of oscillators [Barnes et al., 1971] . Thus instead of speaking loosely of the precision of the results, we can speak of their stability.

In general, for small values of $\tau$, $\sigma_a(\tau)$ follows a law in $1/\sqrt{\tau}$, as in the case of white noise, but for larger values of $\tau$, $\sigma_a(\tau)$ reaches a minimum which is called the flicker floor. For still larger values of $\tau$, $\sigma_a(\tau)$ generally increases. In this latter domain it is sometimes difficult to make the distinction

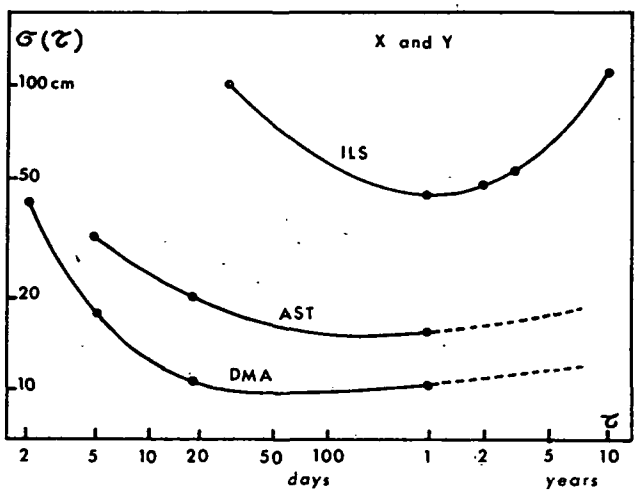



Fig. 3. Precision of the pole coordinates as a function of the averaging time $\tau$. $\sigma^2(\tau)$ is the pair variance. AST stands for global astrometric services : IPMS and BIH special solution for astrometry alone. The current BIH results including DMA have about the same precision as DMA.

between random and systematic errors. For instance, for the Earth rotation parameters, the increase of $\sigma_a(\tau)$ with $\tau$ can be the consequence of changes in the network of observing stations, changes of programs and methods in the stations, and/or plate motions.

In the following, the stability curves will be given. One must be aware that they represent an estimation. For values of $\tau$ smaller than a month, we can assume that the observational noise is larger than the true noise of the observed quantities themselves, and the use of a high-pass filter gives fairly reliable values of the random errors a. But for larger values of $\tau$, one has to make less reasonable assumptions. The "three-corner-hat method" is not available because there are only two truly independent series of data, the astrometric and the Doppler.

The accuracy of a series of results expresses the degree of conformity with an adopted standard. In our case, the standard is not well defined. Let us take the motion of the pole. In the case of astrometry, we only define the direction of the pole relative to the direction of the zeniths. On account of plate motions, the reference to the zeniths is vague. The satellite method refers the position of the pole to the Earth. Similarly, the plate motions displace the stations and also the plate on which the pole moves. Nevertheless, some types of errors can be recognized, such as annual terms due to wrong positions of stars, or drifts due to erroneous proper motions of stars. An attempt to identify these sources of errors will be made.

## Services, Series of Results

### International Latitude Service (ILS)

The ILS was born from the recognition of the existence of polar motion in 1880-1890. This recognition is not a sudden discovery. Its complicated history [see Sevarlic, 1957] began with Euler's theoretical work, but in spite of many attempts to measure polar motion, this motion was not undoubtedly found in the observations until Künstner's work in 1884-86. Further campaigns of observations, the analysis by Chandler, Nyren, Marcuse and others, the famous interpretation of the Chandler period by Newcomb (in 1892), showed the necessity of continuous monitoring of the pole. The original proposal of Fergola, presented to the International Association of Geodesy (IAG) in 1883 by Schiaparelli, to organize systematic determinations of latitude was reintroduced in 1889 by Förster, then followed by the creation of the ILS (IAG meetings of 1895, 1896 and 1899).

The fundamental idea, in organizing the ILS, was to remove the systematic errors due to the poorly known star positions by using stations on a common parallel, with identical instruments and programs. The adopted latitude measurements, according to the Horrebow-Talcott method, consists of meridian measurements of zenith distances, the divided circles and refraction uncertainties being avoided by the use of pairs of stars at nearly equal distances, North and South : only differential zenith distances are measured, which requires well calibrated micrometer screws. The stations, located on the 39°8'N parallel are now :

| | | |
|---|---|---|
| Mizusawa, Japan, | longitude | 141° E |
| Kitab, URSS, | " | 67° E |
| Carloforte, Italy, | " | 8° E |
| Gaithersburg, USA, | " | 77° W |
| Ukiah USA, | " | 125° W |

Initially, the declination errors were removed by the so-called chain method. Two or more groups of stars are observed every night. Assuming that the latitude does not vary during the night, the difference between group results represents the contribution of declination errors. As only night observations are possible, a full year is required for a complete evaluation of the declinations, which permits to refer all the observed latitude to the same standard before solving equations (1). Next year, the cycle is resumed and the star positions can be improved. That explains that no definitive results can be published, as long as the same stars are observed (normally during 6 to 12 years, because after some time precession makes the list of stars obsolete).

This procedure was improved in 1922 by Kimura, who gave the ILS its present form. The Kimura z term of equation (3)

$$x(t)\cos L_{qi} + y(t)\sin L_{o,i} + z(t) = \varphi_{t,i} - \varphi_{o,i} \qquad (3)$$

contains all the non-polar effects common in the observed latitudes of the stations.

A further improvement was the "latitude control method" [Markowitz, 1961], which eliminates the influence of wrong micrometer screw calibrations, and which is now currently used as a check.

With the Kimura method, the results of the ILS could be made available within short delays. But only Yumi, in 1962, took advantage of this possibility when he started to publish the Monthly Notes of the IPMS, giving the ILS results with a delay of about 3 months. In addition detailed results are given in the IPMS Annual Reports with a delay of the order of two years. In the past, it happened that the volumes of definitive results were published more than 20 years after the observations.

The ILS obtains the raw values of x and y monthly. It is customary to give smoothed values at 0.05 year intervals.

According to its organization, the ILS should give accurate results if there were no local errors : the fixed network of stations and the z-term ensure in that case a perfect geometrical solution to the polar motion determination. That is why, in 1967, the International Astronomical Union (IAU) and the International Union of Geodesy and Geophysics (UGGI) defined the "Conventional International Origin" (CIO) for the pole by giving conventional values of the five initial latitudes $\varphi_{o,i}$ of the present stations. However, after solving equations (3), it can be seen that the local residuals are not random. In particular, they exhibit annual components, showing that there are local seasonal effects (on in-

struments, refraction...), and that therefore the annual path of the ILS pole is not accurate. Another source of systematic uncertainties lies in the tectonics. The stations are poorly established in active areas, so that we cannot be sure that the observed drift of the pole toward Canada, at a rate of 10 cm/year, is real.

The precision of the ILS coordinates is schematically shown by figure 3. The long-term noise is due to changes of observers, of reference latitudes, of methods, of instruments. The figure gives a very rough estimate, since the noise in the results is not stationary. Not only the number, but also the quality of the stations changed. After the enthusiasm of the beginning, there has been a period of lack of interest, until the development of the Earth sciences in 1950-60. Presently the Service suffers from the lack of observers and from the discouragement in view of the better results obtained by IPMS and BIH and of the expectations from new techniques.

Nevertheless, the 80-year series of the ILS is invaluable. It can be improved by using modern techniques of computation, better initial star coordinates and astronomical constants, a more homogeneous set of initial latitudes. The enormous work of revision was undertaken by Yumi and his staff, and is nearly achieved. In the mean-time Vicente and Yumi [1969] published an homogeneous set of coordinates of the pole since 1900, based on ILS results.

International Polar Motion Service (IPMS)

The ILS organization has the drawback of not allowing the use of data from outside stations. Excellent series of latitude measurements obtained by the observatories of Greenwich, Pulkovo, Washington, from 1912 to 1952 have been very little used in the investigation of polar motion. This situation could no longer be accepted when in 1950-60 many new instruments, photographic zenith tubes, astrolabes, improved visual zenith telescopes for the Horrebow-Talcott method, photoelectric transit instruments, came into use. Although many astronomers were reluctant to combine the data from these various sources, in 1962, the ILS was reorganized into the IPMS with the task of deriving polar motion from latitude and universal time data of all astronomical instruments (UGGI and UAI decisions in 1960 and 1961). The IPMS was located at the International Latitude Observatory in Mizusawa, Japan. S. Yumi became director in 1962 after the death of T. Hattori.

Although the ILS formally disappeared, it is advantageous to keep the ILS designation, as we did, for the particular pole coordinates set based on the international stations-a continuation of the former ILS.

The IPMS began to publish the coordinates of the pole based on all latitude measurements, on a current basis, with the issue for January 1975 of the Monthly Notes of the IPMS. In this issue the coordinates for 1974 were also given. Solutions starting from 1962, based on latitude alone, and on latitude and time were given in the IPMS Annual Reports for 1972 and 1974.

The number of stations contributing to the IPMS work is fairly stable and of the order of 80. For instance, in the IPMS Annual Report for 1975 we see that 21 stations participate with latitude only, 26 with UTO only, 29 with both UTO and latitude.

The method of computation is based on the monthly averages of latitude and UTO data [S. Yumi, 1976]. It requires a coherent set of initial latitudes and longitudes. As the quality of the contributing observations is much different for the various stations, the weighting procedure is important. The weights are computed from the scattering of individual measurements. Several solutions were attempted, using equations (1) and (2), and also similar equations with auxiliary unknowns to take into account the errors in the development of the celestial nutation. The raw results are monthly values of x, y (and also UT1-UTC, although the IPMS does not currently publish these results). The smoothed results are given at 0.05 year intervals.

In the IPMS, the work is based on latitude and UTO data forwarded by the observatories. In these observatories, the only way of taking into account the star position errors is to use the chain method, or similar methods. It is well known that these methods do not eliminate spurious annual terms on account to apparent variations of latitude and UTO during the night, due to temperature effects and refraction. Therefore the IPMS annual term of the pole motion cannot be accurate (a remark which also applies to the BIH work). Similarly the local drifts due to erroneous proper motions contribute to a spurious drift of the IPMS (and BIH) pole. Nevertheless, these effects are reduced by the averaging procedure. The peak-to-peak spurious annual variation is probably less than 1 m (of the same order as for ILS) and the spurious drift of the pole less than 3 cm/year (smaller than for ILS).

The estimated stability curve is given by figure 3. The long term unstability is mainly due to changes in the station list and in the programs.

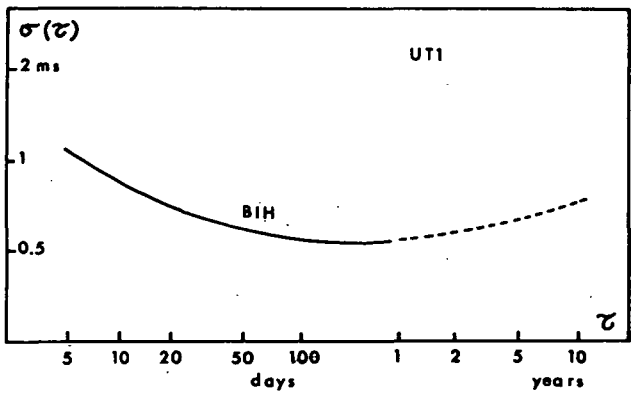



Fig. 4. Precision of UT1 obtained by BIH. $\sigma^2(\tau)$ is the pair variance, as a function of the averaging time $\tau$.

Bureau International de l'Heure (BIH)

The BIH was created in 1912 and has been located at the Paris Observatory since that date. Its present director is B. Guinot (since 1965).

The initial task of the BIH was to unify time by publishing the time of emission (in universal time or at that time mean solar time) of radio time signals. With the improvement of time comparison methods and the advent of operational atomic time standards in 1955, the universal time measurements could be referred to a common very uniform time scale, presently designated by International Atomic Time TAI, or to its sister time scale UTC. Thus, besides establishing TAI and UTC, the traditional task of the BIH evolved in giving the difference between universal time and TAI or UTC - a measure of the irregularity of the Earth rotation. On the other hand the increase of the observational precision necessitated taking into account the polar motion (IAU recommendation, 1955), leading to the definition of UT1. As the coordinates of the pole were not available from the ILS in due time, N. Stoyko, formerly in charge of the BIH, began to establish his own set of coordinates from the latitude data of an increasing number of participating observatories. In 1965, in order to reduce the delays of availability of the results, Guinot began to solve equations (1) and (2) simultaneously. Thus, the overlap of functions with IPMS appears to be complete ; but this was unavoidable, because the scientific unions did not make a comprehensive evaluation of the situation in 1962 when creating the IPMS.

There have been many changes in the BIH work since its creation, due to the evolution of techniques. The following information refers to the present situation. The BIH uses all data on Earth rotation from any source, as soon as their systematic errors can be sufficiently well modeled. Until 1972.0, only astrometric data (the same as for IPMS) were used. Since 1972, it was possible to use the pole determinations by satellite observations from the DMA (see below). The computational methods were described by Guinot and Feissel [1968] and with more details by Feissel [1972] . They consist first in establishing a "system" by determination of the initial latitudes and longitudes, and the annual systematic corrections to data. Then, equations (1) and (2), with appropriate weights, are solved for individual measurements of $\varphi_{t,i}$ and $[UTO_i - UTC]_t$. It is thus possible to adopt a short averaging time.

Raw five-day values of x, y, UT1-UTC have been available since 1967. In 1972.0, the satellite data began to contribute to these raw values ; they presently receive about half the total weight for polar motion. Nevertheless, for comparison purposes a solution from astrometry only is continued ; it is based on the same data as IPMS work. Smoothed five-day values are published filtered with a cut-off at about 30 days ; they extend in an homogeneous series since 1962.

The current results for month m are published in BIH Circular D at the beginning of month m + 2, with provisional raw values. Improved results for the year a are published towards June of year a + 1, in an Annual Report. In addition, the BIH operates a rapid service, giving with lower precision the results of week w on Thursday of week w + 1, under a contract with the Jet Propulsion Laboratory (USA).

The BIH computations are devised in order to keep the spurious annual term as constant as possible. The satellite data are introduced after annual corrections, which express them in the initial BIH system. Thus the annual systematic error is due to the annual errors of astrometric measurements at initial epoch, in 1968 ; for polar motion, it is of the same order as for IPMS (1m peak-to-peak) ; for UT1 it can be of the order of 3 ms, peak-to-peak. The spurious drift of the BIH pole should be less than 3 cm/year. The spurious drift of UT1 is mainly due to the drifts of the star catalog equinoxes ; a drift of 1 ms/year appears possible.

The precision of the BIH results is given by the stability curve of figure 3 and 4.

The BIH is involved in the EROLD project (Earth rotation by Lunar Distances) and in the MEDOC experiment (Polar motion by Doppler observations of satellites). A revision of past data was recently undertaken by Feissel.

The BIH, as the IPMS, sends its current results free of charge to about 800 addresses. These data are reproduced in several national publications.

Defense Mapping Agency (DMA)

The computation of the pole position based on Doppler observations of Transit satellites originated at the Naval Weapons Laboratory, USA, (now the Naval Surface Weapons Center), as a consequence of the research of Anderle and Beuglass [1970]. A service was organized under the name of Dahlgren Polar Monitoring Service, giving the coordinates of the pole since 1969. In April 1970, the responsibility for the computation of the orbits of the Transit satellite, and therefore of the pole motion, was transferred to the Topographic Center of Defense Mapping Agency, without changes of the computation programs.

Many improvements took place, which are reported, together with a comprehensive bibliography by Oesterwinter [1978]. The most important one appeared in the 1972 results, where the polar motion, instead of being derived from station residuals, was computed directly in the least square solution together with the orbit parameters. Other improvements resulted from better gravity field models, from better station coordinates, from the increase of observing stations (about 20 presently).

The reduction techniques were described by Anderle [1973]. It is reminded that 48-hour time spans are used, leading to two-day raw values of the pole coordinates for each satellite. These values, with their standard deviations, are made available within a few days in DMA reports. Five-day averages are also given weekly in the USNO Series 7 Circulars.

The possibility of systematic errors due to the model of forces was investigated by Taton [1972] and by Bowman and Leroy [1976]. Errors

with periods 5 to 6 days and 11 days were found, but they can be easily filtered. Longer term errors may exist, but are probably small when compared to the errors due to changes in the station network and plate motions. The size of seasonal effects, if any, cannot be estimated, but appears to be much smaller than for astrometry.

The precision curve is given by figure 3. A source of long term, noise could be the successive refinements of the model of forces : these effects are now very small.

Thus, the DMA maintains a polar motion service which is not officially recognized by scientific unions, but is essential, as being more accurate, more precise and more rapid than the official ones. The possibility of managing a Doppler network by scientific organizations is being tested in the MEDOC project [Guinot and Nouel, 1976 ; Nouel and Gambis, 1978].

## Other series of results

As prior to 1962 no latitude series other than those of northern ILS stations contributed to the official work on polar motion, Fedorov and his collaborators at the Kiev observatory computed the polar motion from all known measurements. They were able to get some coordinates starting from 1846, and high precision coordinates from 1890.0 to 1969.0 Fedorov et al., [1972].

The Gosstandard of USSR computes a solution for UT1 based on the 21 instruments for universal time measurements operating in USSR and neighbouring countries. This solution is based on an original optimum estimation ; it is published in the bulletins "Vcemirnoe Vremja" (Series E) of the USSR Gosstandard.

## Conclusion

Although their organization is not optimum, the services have produced uninterrupted series of pole coordinates and UT1 values, with good homogeneity. They have proved their ability to issue routinely the data with the short delays needed in modern research. The causes of this success are worth considering when preparing new services.

The work of the visual observers was (and is still) essential. One must be reminded that the same person has often to work more than six hours in the middle of the night, at outside temperatures, ignoring holidays and week-ends. Even with new automatized instruments the importance of the attendance should not be under-estimated, in service operation.

The astronomical instruments are very reliable. The Doppler receivers can be easily replaced in case of failure. Are we sure that other techniques, with sophisticated devices, will not suffer from interruptions, especially when based on a small number of stations ?

The operation of a service is quite different from an experiment. It has to run continuously, at the highest level of quality, in spite of many difficulties : vacations, illnesses, pregnancies, strikes, computer failures... Although requiring good scientists and technicians, this metrological work is not considered as attractive ; it is not easily supported by national and international organizations, which prefer to grant new projects.

## Acknowledgments

The IPMS and BIH are services of the Federation of Astronomical and Geophysical Services (FAGS), a member of the International Council of Scientific Unions (ICSU). The allocation of FAGS to the services, based on ICSU and UNESCO funds, although essential, only represents 5 to 10% of the real cost of the central bureaus. The remaining is supported mainly by the host countries, and also by many establishments of outside countries, private fundations, and other international organizations.

The observation stations, the Doppler service are entirely supported by national sources.

It would be much too long to give a list of all contributors, they will recognize themselves, and accept our thanks in the name of the scientific community.

## References

Aardoom, L., Proceed. GEOP 9 (1978), this volume.
Anderle, R.J., Geophysical Surveys, 1, 147, 1973.
Anderle, R.J., and L.K. Beuglass, Bull. géodésique, 96, 125, 1970.
Barnes, J.A., et al., IEEE Trans. on Instrum. and Meas., IM-20, N° 2, 105, May 1971.
Bowman, B.R., and C.F. Leroy, Proceed. int. geodetic symp., p. 141, Las Cruces, oct. 1976.
Fedorov, E.P., et al., The motion of the pole of the Earth from 1890.0 to 1969.0, Ed. : V.A. Bulkina, published "Naukova Dumka", Kiev, 1972.
Feissel, M., BIH Annual Report for 1971, part E, 1972.
Guinot, B., and M. Feissel, Proceed. UAI Symp. N° 32, p. 63, D. Reidel Pub. Co., 1968.
Guinot, B., and F. Nouel, Proceed. int. geodetic Symp., p. 159, Las Cruces, Oct. 1976.
Markowitz, W., Bull. géod., 59, 29, 1961.
Nouel, F., and D. Gambis, Proceed. IAU Symp. 82 (Cadiz, May 1978), in press.
Oesterwinter, C., Proceed. IAU Symp. 82 (Cadiz, May 1978), in press.
Ševarlić, B.M., Publ. Obs. astron. Belgrade, 5, 1957.
Taton, N., Détermination du mouvement du pôle à partir d'observations de satellites artificiels, Thèse, Université de Paris VI, 1972.
Vicente, R.O., and S. Yumi, Publ. intern. Latitude Obs. Mizusawa, 7, N° 1, 41.
Yumi, S., Annual Report of the IPMS for the year 1974, 1976.

# Earth Rotation and Polar Motion from Laser Ranging to the Moon and Artificial Satellites

L. Aardoom
Department of Geodesy, Delft University of Technology
Delft, The Netherlands

Abstract. Earth-based laser ranging to artificial satellites and to the moon is considered as a technique for monitoring the Earth's polar motion and diurnal rotation. The kinematics of Earth rotation as related to laser ranging is outlined. The current status of laser ranging as regards its measuring capabilities is reviewed. Artificial satellite laser ranging has recently yielded pole position to better than 0.02 arcseconds with 5 days averaging as the best result. In recent years single-station lunar laser ranging has produced UT0-values to better than 1 msec. The relative merits of artificial satellite and lunar laser ranging are pointed out. It appears that multi-station combined artificial satellite and lunar laser ranging is likely to ultimately meet a 0.002 arcseconds in pole position and 0.1 msec in UT1 daily precision requirement.

## Introduction

Although the most intriguing implications of the Earth orientation phenomenon lie in its dynamics and the geophysical effects involved, the present contribution will focus on its kinematics. This is done in relation to the use of Earth-based laser ranging to the moon and to artificial satellites as a technique to monitor polar motion and diurnal rotation. It should be pointed out, however, that such monitoring can only be performed efficiently if a dedicated program of measurement is based on a profound qualitative understanding of the phenomenon.

Earth orientation is involved where measurements connect positions of objects, related to an extra-terrestrial frame of reference, to terrestrial objects. The geometric relationships established by the measurements depend on the orientation of a conventional Earth-fixed frame of reference with respect to the extra-terrestrial one. Such situation arises in space geodesy in general and in satellite geodesy in particular, the moon conveniently being considered as a satellite. Until the advent of precise tracking means, including laser ranging, the time-dependent orientation of the Earth was numerically modelled in terms of the current theories of precession and nutation and the orientation parameters as provided by the Bureau International de l'Heure (BIH); see Veis [1963]. Noting the inherent precision of modern tracking techniques and some of their further advantageous characteristics it was realized that certain constituents of Earth orientation should be rather modelled in terms of unknown solution parameters in order to exploit the qualities of the measurement data. One of these techniques is laser ranging to the moon and to artificial satellites.

Before reviewing the current status of this technique and its foreseen development the geometry and kinematics of Earth orientation as pertaining to laser ranging will be presented. The potentialities to monitor Earth orientation by laser ranging will be assessed and results obtained from both lunar and artificial satellite ranging will be quoted. These concern the diurnal rotation of the Earth and polar motion.

## Earth orientation as related to laser ranging

A meaningful approach to Earth orientation requires the operational definition of at least two frames of reference: one to which the motion of the Earth is sufficiently well modelled (the rectangular Cartesian conventioned "Earth-fixed" frame: $x_1$, $x_2$, $x_3$) and another rectangular Cartesian frame ($z_1$, $z_2$, $z_3$) with respect to which the orientation is to be monitored. The definition of reference frames in a contemporary context is not always straightforward as demonstrated by Kolaczek and Weiffenbach [1974]. Of particular relevance is the problem of fixing a reference frame to a deformable Earth. The problem of Earth orientation is essentially the problem of finding the relative orientation of two reference frames, the components of unit vectors referred to both reference frames being given as time-dependent quantities. The basic problem is thus to find both sets of components, in this case from the laser ranging.

In the classical optical stellar approach to monitor Earth orientation, the x-system is defined by a set of conventionally adopted latitudes $\phi$ and longitudes $\lambda$ of participating observatories, these observatories measuring the time-dependent direction components of the observatory verticals with respect to a fundamental stellar reference frame adopted as z-frame. This z-frame is in fact defined by conventional positions and proper motions of fundamental stars and considered inertial. Satellite ranging is, however, what Newton [1974] called, a "blind" technique and the relation to the fundamental z-frame is less straightforward. The main complication of the satellite approach is that the orbits of the satellites with respect to which the Earth's orientation is to be monitored by Earth-based measurements, have to be monitored themselves by means of such measurements. This holds in particular for artificial satellites somewhat less for the moon. The critical issue here is that satellite orbits referenced to the z-frame cannot be determined from Earth-based tracking without the à priori involvement of the Earth's orientation. This involvement is however restricted so that Earth-based determined satellite orbits are at least in certain respects independent of à priori knowledge about the Earth's orientation, and Earth-based tracking data may thus contain signatures of Earth-orientation.

Of crucial importance to any practical approach to Earth orientation is the operationality of re-

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

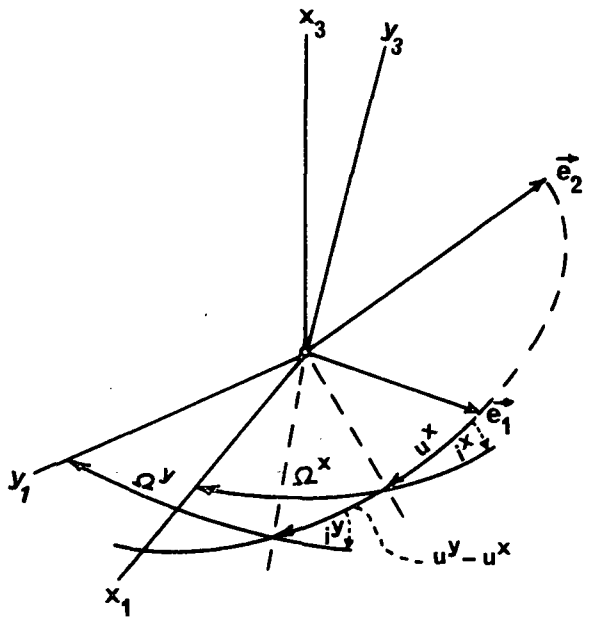



Fig. 1. Definition of reference frames by angular Keplerian orbital elements u, i and Ω.

ference frame definitions. The orientation of a right-handed rectangular Cartesian frame is uniquely defined by two orthogonal unit vectors $\vec{e}_1$ and $\vec{e}_2$. Any frame obtained by rotation from this basic frame, can then be specified by angles $u^x$, $i^x$ and $\Omega^x$; see Figure 1. So is a quasi Earth-fixed $x_1$, $x_2$, $x_3$-frame, in which laser tracking locations and their modelled motions due to tides, tectonics etc, are specified. Likewise any $y_1$, $y_2$ $y_3$-frame with respect to which the orientation of the x-frame is to be described: $u^y$, $i^y$, $\Omega^y$.

Given $u^y$, $i^y$, $\Omega^y$, the differences

$$\Delta u = u^x - u^y$$
$$\Delta i = i^x - i^y$$
$$\Delta\Omega = \Omega^x - \Omega^y$$

uniquely describe the rotation from the y-frame to the x-frame.

Now u, i and Ω respectively can be identified with the angular Keplerian orbital elements of a satellite: argument of latitude ω + f, inclination and argument of the ascending node, f being the true anomaly and ω the argument of perigee. Thus the use of satellites to monitor Earth orientation seems rather obvious if $u^y$, $i^y$, $\Omega^y$ and Δu, Δi, ΔΩ can to sufficient accuracy be determined as time varying quantities. The procedure to obtain such quantities is a complicated one, to be summarized as follows. (see Figure 2).

Osculating Keplerian elements of the satellite $\Omega^x(t)$, $i^x(t)$, $\omega^x(t)$, $a^x(t)$, $e^x(t)$ and $f^x(t)$ are obtained from a state vector $\vec{x}(t)$; $\dot{\vec{x}}(t)$, determined in the x-frame from tracking data, e.g. laser ranging, in a purely kinematic way.

An inertial y-frame can be defined by adopting values $u^y(0)$, $i^y(0)$, $\Omega^y(0)$ for $u^y$, $i^y$, $\Omega^y$ at a se-

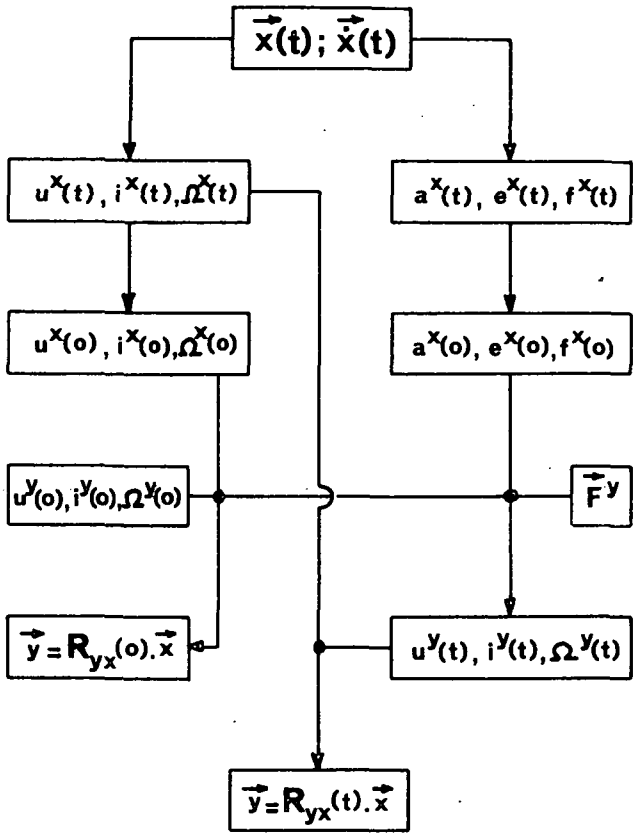



Fig. 2. Orientation of an Earth-fixed x-frame relative to an inertial y-frame by means of perturbed angular Keplerian orbital elements. $\vec{F}^y$ is the field of perturbing forces as described on the y-frame, a is the satellite orbit's semi-major axis, e its eccentricity and f its true anomaly.

lected epoch $t_0$. Together with $a^x(0)$, $e^x(0)$ and $f^x(0)$ these provide the initial state vector at $t_0$ of the satellite in the y-frame.

The two sets of angular elements

$$u^x(0),\ i^x(0),\ \Omega^x(0)$$
$$u^y(0),\ i^y(0),\ \Omega^y(0)$$

together define the relative orientation of the x- and y-frames at $t_0$: $\vec{y} = R_{yx}(0) \cdot \vec{x}$.

To track the time-dependent orientation $R_{yx}(t)$ of the x-frame relative to the y-frame, angular elements

$$u^y(t),\ i^y(t),\ \Omega^y(t)$$

are required in addition to

$$u^x(t),\ i^x(t),\ \Omega^x(t)$$

as obtained by measurement in the x-frame. Elements

$$u^y(t),\ i^y(t),\ \Omega^y(t)$$

can be extrapolated from the initial state vector at $t_0$ if the complete field of forces acting in the y-frame ($\vec{F}^y$) is assumed. In this extrapolation the y-frame may be considered inertial. It should be noted however that this inertial frame does in general not coincide with the fundamental stellar z-frame of reference, which was also considered inertial, but which deviates from the y-frame by a time-invariant relative orientation

$$\vec{z} = R_{zy} \cdot \vec{y},$$

the elements of which remain as yet unspecified.

The force field as referred to the y-frame will consist of two classes of contributions:

- forces depending on the time-dependent relative orientation $R_{yx}(t)$ of the x- and y-frames, e.g. non-central terrestrial gravitation including solid-Earth and ocean tides;
- forces independent of this orientation, e.g. solar radiation pressure and luni-solar gravitation.

The first class of forces poses a theoretical complication because the relative orientation $R_{yx}(t)$ to be derived has to be known in advance in order to extrapolate the initial angular elements

$$u^y(0),\ i^y(0),\ \Omega^y(0) \quad \text{at } t_0$$

to obtain instantaneous values:

$$u^y(t),\ i^y(t),\ \Omega^y(t) \quad \text{at } t.$$

It should be noted that $R_{yx}$-dependent inertial forces acting in the x-frame do not interfere, provided the procedure to obtain angular elements $u^x(t)$, $i^x(t)$, $\Omega^x(t)$ is indeed a purely kinematic one.

To obtain the $R_{yx}$-dependent force contribution with respect to the y-frame poses a complication leading to the introduction of what Lambeck [1971] called "dynamic perturbations". Although these perturbations may not be entirely negligible with future precise laser ranging to artificial satellites, this complication is disregarded here. In doing so a unique latent opportunity to verify the coincidence of the adopted $x_3$-axis of maximum inertia [Melchior, 1972] is likewise disregarded. On the other hand it seems unlikely that even advanced tracking precision would permit this verification in a foreseeable future [Gaposchkin, 1972; Kolaczek and Weiffenbach, 1974]. Nevertheless in a detailed discussion the formal non-coincidence should not be overlooked:

$$\vec{x} \approx R_{x\xi} \cdot \vec{\xi},$$

the $\xi$-system being the axis-of-figure (maximum inertia) system. Hence:

$$\vec{z} = R_{zy} \cdot R_{yx} \cdot R_{x\xi} \cdot \vec{\xi} = R_{z\xi} \cdot \vec{\xi}$$

$R_{z\xi}$, thus defined includes:

a. the unknown small and virtually invariable deviation of the adopted quasi Earth-fixed x-frame from the axis-of-figure $\xi$-frame;
b. polar motion relative to the $x_3$-axis, which does not necessarily coincide with the CIO, but will be close to it;
c. the Earth's diurnal rotation as measured by the siderial angle $\theta$;
d. the small deviation termed "sway" of the Earth's instantaneous rotation axis from the direction or the total angular momentum vector; McClure [1973] found that for a deformable Earth this deviation can approach 0.01 second of arc.
e. luni-solar forced nutation and precession of the Earth's total angular momentum vector;
f. the relative orientation of the inertial y-frame to the quasi-inertial fundamental stellar z-frame.

Of these only items (b) and (c) pertain directly to the present subject. Ignoring for the sake of simplicity of the present treatment the possible deviations between the $\xi_3$-axis and the CIO and between the CIO and the $x_3$-axis and also sway, we can summarize:

$$\vec{y} = R_{yz} \cdot P \cdot N \cdot R \cdot \vec{x} = R_{yx} \cdot \vec{x},$$

R, N and P being rotation matrices describing combined polar motion and diurnal rotation, forced nutation and precession respectively, in a form as presented by Veis [1963]. Writing

$$P' = R_{yz} \cdot P,$$

P' may be considered as describing precession with respect to the arbitrarily defined inertial y-frame, leading to the conclusion that satellite techniques yield the compound rotation

$$\vec{y} = P' \cdot N \cdot R \cdot \vec{x},$$

precession P' relative to the y-frame. To comply in practice with convention and to conveniently separate R from precession and nutation, when assuming the latter two, the initial angular orbital elements

$$u^y(0),\ i^y(0),\ \Omega^y(0)$$

are selected so, that the y-frame coincides with the z-frame, although because of incomplete knowledge about R(0) at epoch, this can be realized only approximately, even if N(0) and P(0) are assumed. If realized, $R_{yz}$ is unity and P' = P, so that:

$$P \cdot N \cdot R = R_{zx}$$

and consequently:

$$R(t) = N^* \cdot P^* \cdot R_{zx}$$

If not fully realized, R(t), thus derived, will be biased by an unknown, but contstant relative orientation.

A sequence of R(t) for a sequence of instants t will provide the time-dependent orientation of the Earth's rotation axis with respect to the Earth-fixed x-frame ($\psi^x$, $\lambda^x$) and the siderial angle $\theta$, see Figures 3 and 4. Finally, to comply with IPMS and BIH convention, pole position is given as $x_p$ and $y_p$.

In practice [Lambeck, 1971] $u^x$, $i^x$, $\Omega^x$ may be

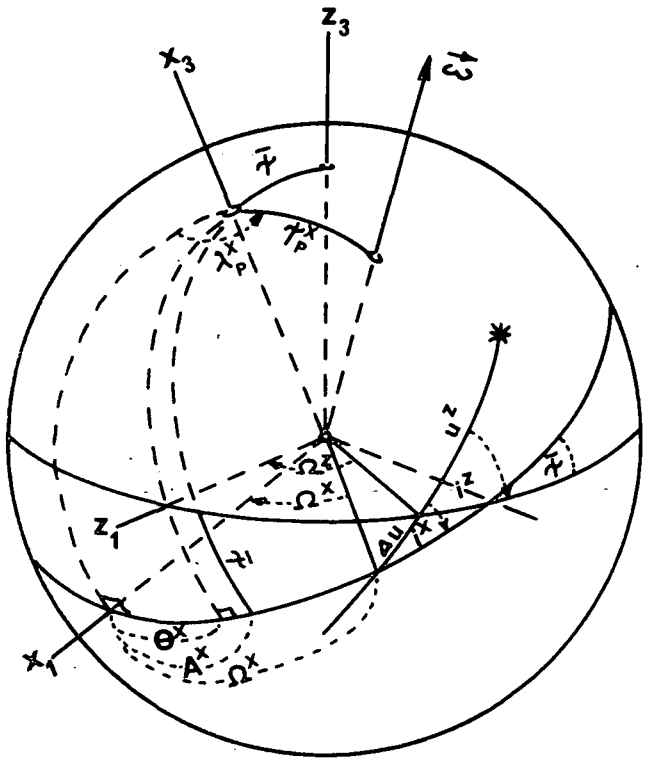



Fig. 3. Geometry of Earth orientation with respect to stellar z-frame, the orientation to be obtained from satellite orbits. $\lambda_p^x$ and $\psi_p^x$ specify the direction of the Earth's rotation axis relative to an Earth-fixed x-frame.

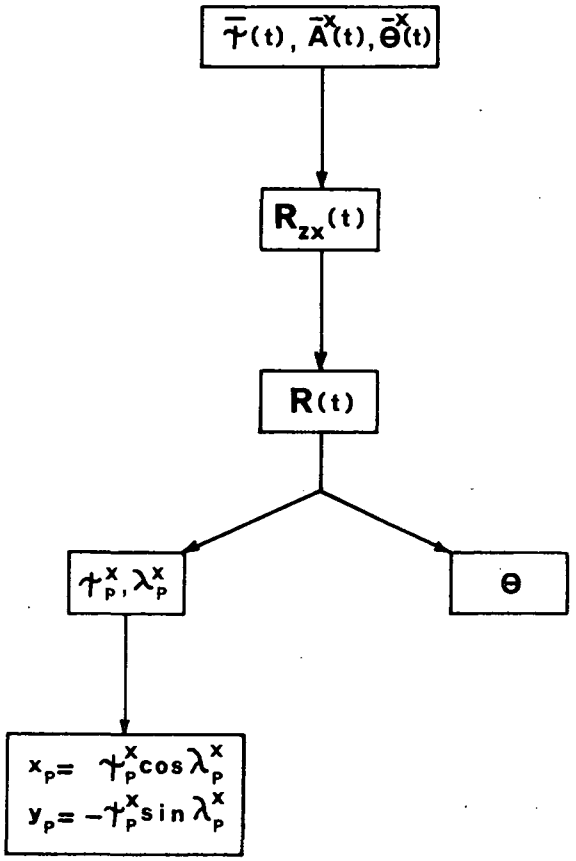



Fig. 4. From the relative orientation of Earth-fixed x- and stellar z-frames to pole position $(x_p, y_p)$ and siderial angle $\Theta$.

transformed to $u^{z'}$, $i^{z'}$, $\Omega^{z'}$ as referred to an intermediate quasi inertial z'-frame, which approximates the z-frame:

$$\vec{z} = D(t) \cdot \vec{z}'$$

This is performed by approximating $R(t)$ by $\tilde{R}(t)$, replacing $\theta$ by an approximate value $\tilde{\theta}$ and equating $x_p$ and $y_p$ to zero:

$$\tilde{R}(t) = \begin{bmatrix} \cos\tilde{\theta} & -\sin\tilde{\theta} & 0 \\ \sin\tilde{\theta} & \cos\tilde{\theta} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Solving then for $D(t)$, rather than for $R(t)$ directly, yields $\psi_p^x$, $\lambda_p^x$ and $\Delta\tilde{\theta} = \theta - \tilde{\theta}$; see Figure 5.

This general approach is valid for both artificial satellites and the moon although it is primarily tuned in to the use of artificial satellites, not restricted however to laser ranging. Before pointing out special features of the lunar laser ranging case some general conclusions may be drawn:

- because the x-frame does not necessarily coincide with the conventional terrestrial reference frame as defined by the CIO and the conventional zero meridian of Greenwich, there may appear constant biases in all three determined rotatio-

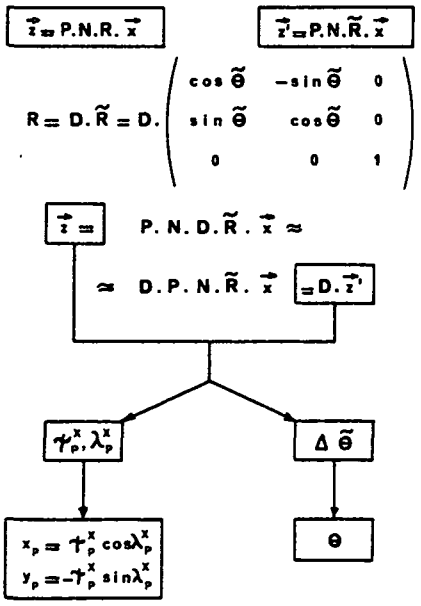



Fig. 5. Pole position $(x_p, y_p)$ and siderial angle $(\Theta)$ from the relative orientation of the quasi-inertial z'-frame relative to the inertial z-frame. $\Theta$ is an approximated siderial angle.

nal parameters $x_p$, $y_p$, $\theta$ as derived in the classical optical stellar way;
- due to the well-known longitude-ambiguity in artificial satellite orbits such bias is so likely in $\theta$ that artificial satellite techniques are expected not to provide UT1, but rather the rotation rate $\omega$ of the Earth;
- additional not necessarily constant biases may be caused by erroneously modelled lunar-solar precession and mutation (P- and N-matrices);
- a detailed force model and a scrupulous application of that model are required.

The moon is just another satellite, kinematically differing from the others by:
- its larger mean orbital radius;
- its relatedly longer period of revolution, in fact much longer than that of the Earth's diurnal rotation;
- its more stable center of mass orbit, less effected by the Earth's graviational potential and other terrestrial perturbing forces;
- its substantial size, which requires the selenodetic coordinates of the retroreflector sites and the lunar rotation to be modelled or solved for;
- there is only one moon, hence one lunar center-of-mass orbit but the moon carries several geometrically distinct retroreflectors.

These differences entail marked differences concerning tracking-operational and data analysis aspects. These are indicated in the next two paragraphs respectively.

## The status of laser ranging

The technique itself is assumed known and will be reviewed only as regards its main characteristics pertaining to Earth orientation determination. Detailed and up to date technical information on availabe instrumentation and current developments can be obtained from Pearlman and Hamal [1978].

Relevant system characteristics are:
- maximum range capacity;
- ranging precision;
- day-light ranging capability;
- repetition rate.

Speaking in general terms these characteristics are not mutually independent.

In the present context, systems can as regards their maximum range capibility be classified in broadly three catagories:
- those of lunar ranging capability;
- those able to range at Lageos;
- those restricted to closer artificial satellites, like Beacon Explorer-C, Starlette and Geos-3.

It should be noted that the operational aspects of lunar laser ranging (LLR) are quite different from those of artificial satellite laser ranging (SLR), and a system of lunar range capability will not necessarily be able to range at artificial satellites.

The most advanced SLR systems attain between 5 and 10 cm single shot precision, expressed as a distance standard deviation; 2 to 3 cm is foreseen for the next few years. LLR "normal observation point" precisions (see next paragraph) reach or are expected to reach the same level.

Day-light ranging capability is of crucial importance in order to provide a continuous record of short period Earth orientation phenomena, such resolution requiring averaging times of a fraction of a day. It is to be understood that such short averaging time is a prerequisite of future precise Earth rotation monitoring systems. Many of the operational and most of the planned SLR devices have day-light capability on the closer satellites, few on Lageos. LLR features guiding difficulties when ranging is attempted close to new moon. Because of that, the only routinely operational LLR system (at the McDonald Observatory, Fort Davis, Texas) cannot effectively range within 3 days off new moon.

The repetition rates of most SLR systems exceed 0.1 pps. The McDonald LLR performs 1 pulse per 3 seconds, but the McDonald team has made it practice to compress the data from each 5 to 20 minutes run into a single "normal point", three runs being attempted per day, leading to an equal number of daily "normal points", except around new moon.

Of paramount importance for Earth orientation work, will be station siting, considering average atmospheric conditions and the need of global deployment, the first requirement in view of data continuity, the second for "Earth-fixed" (x-) reference frame definition.

Although only part of the available facilities has been used in dedicated programs of polar motion and diurnal rotation studies, encouraging preliminary results have been obtained from both SLR and LLR.

## Review of work accomplished

In the past eight years or so several types of contribution have been made to the determination of the Earth's diurnal rotation and polar motion:
- theoretical modelling;
- feasibility analysis of simulated data;
- dedicated data taking;
- analysis of actual data.

Only the latter of these will be reviewed, separately for SLR and LLR, arbitrarily, in this order. In the present context this review can only be sketchy and cannot satisfactory reflect the amount of effort spent by contributing individual or groups of investigators.

Early attempts at NASA's Goddard Space Flight Center (GSFC) to detect polar motion from artificial laser ranging were initiated in 1970 and demonstrated the capability. [Smith et al, 1972a]. First preliminary results [Smith et al, 1972b] indicated that using the data of a single 30 cm precision ranging station the variation in latitude of that station could be derived to 0.03 seconds of arc with a time resolution of 6 hours. Considering the latitude of that station at Greenbelt, MD ($39^{\circ}$N) and the inclination of the single Beacon Explorer-C satellite used ($41^{\circ}$) this satellite could be tracked near apex and i' (briefly written instead of $i^z$ ) be accurately determined to about 0.001 second of arc, the accuracy of $\Delta i = i' - i$ limited by that of i (briefly, instead of $i^z$), because of gravitational field uncertainties. What could be derived from a single station was the variation of latitude, thus polar motion projected onto the Greenbelt-meridian; not the corresponding

component of polar deviation from the CIO, simply not because the Greenbelt-latitude is not known in the conventional BIH-frame. The method employed to obtain such single station polar motion results has been described to some detail in [Dunn et al, 1974] and [Kolenkiewicz et al, 1977]. It is a special version of the general approach outlined before, the angular element analysed being the inclination i, the pertinent kinematic equation reading [Lambeck, 1971]:

$$i' - i \simeq -\psi_p \sin(\lambda_N - \lambda_p) ,$$

where $\lambda_N$ stands for the longitude of the satellite's ascending node. Using up to four consecutive station passes of the satellite within a time span of about 6 hours, short arc osculating inclinations i' were obtained (see Figure 6) and compared with a reference orbit, extended over the entire period of the experiment. Later this so-called "max-lat" approach was abandoned and replaced by a more flexible approach yielding both variation of latitude and length of day information [Dunn et al, 1977]. During a 3-week period in 1970 a second ranging station operated from Seneca, N.Y., 400 km North of Greenbelt, yielding also a small number of 4-pass arcs. This enabled two independent determinations of both solution parameters.

The single-station approach as just described, has several limitations:

- of polar motion only the meridional component can be measured;
- it is difficult to separate this component from dynamical perturbations of i, these to be modelled over the full time span of the investigation;
- it is impossible to separate polar motion from precession and nutation;
- the use of several satellites or interrupted arcs of the same satellite may cause discontinuities and inconsistencies in the pole path as measured.

These can completely or to a substantial degree be overcome by deploying a multi-station network. Such network enables to determine a complete set of osculating angular orbital elements from a day or less of laser tracking. A minimally perturbed reference orbit $u^z(t)$, $i^z(t)$, $\Omega^z(t)$ valid only for the same period of a day or less, say, will suffice to yield a sufficiently detailed $R_{zx}(t)$- or $R_{xz}$-record to derive the direction of the orientation axis for the observation period. Because of the shortness of this period incomplete modelling of precession (P) and nutation (N) and of acting forces will practically not interfere. To be more precise as concerns precession and nutation an error in P(0) and N(0) at $t_o$ would cause a biased reference orbit, but because of the shortness of the tracking period this bias will remain constant throughout and persist as such in $P^*(t)$ and $N^*(t)$.

Polar motion results from multi-station laser tracking of artificial satellites were reported recently from two sources by Smith et al [1978b] and Schutz et al [1978b]. These results announced a break-through in polar motion determination from laser ranging.

Smith et al [1978b] reported preliminary results obtained from Lageos tracking at a total of seven GSFC- and Smithsonian Astrophysical Observatory (SAO)-stations, four of which in the U.S.A., two in South-America and one in Australia. The four GSFC-stations in the U.S.A. claim a 10 cm single shot ranging precision, the SAO-stations in South-America and Australia 1 m. First a consistent set of station coordinates was obtained from this data to define an x-frame of reference [Smith et al, 1978a]. Subsequently an iterative procedure of fitting 5-day arcs to three interlinking 30-day reference orbits yielded average pole positions, relative to the adopted x-frame, for each of the eighteen 5-day periods of October, November and December 1976. The formal standard error in the $x_p$-component of polar position is 0.003 arcseconds, 0.002 arcseconds in $y_p$. The authors believe however that a precision between 0.01 and 0.02 arcseconds is more realistic. From this experience and previous simulations [Kolenkiewics et al, 1977] it was concluded that with more stations ranging Lageos at the 10 cm level and improved modelling, in particular of solid-Earth and ocean tidal effects, ultimately daily 5-cm pole position and UT1 values to about 0.2 msec consistent over 2 to 3 months are feasible. Recently GSFC issued a first bulletin of preliminary 5-day pole positions as obtained from Lageos ranging; the period covered is May-December 1976. Standard deviations given there range on the average from about 0.01 to about 0.03 arcseconds.

Schutz et al [1978b] reported pole positions obtained from laser tracking of Geos-3, spanning a roughly one-month period in early 1976. Data was used from three GSFC 5 to 10 cm single shot precision systems located at Greenbelt, MD, Bermuda and Grand Turk. The analysis was based on a previously obtained set of station coordinates de-

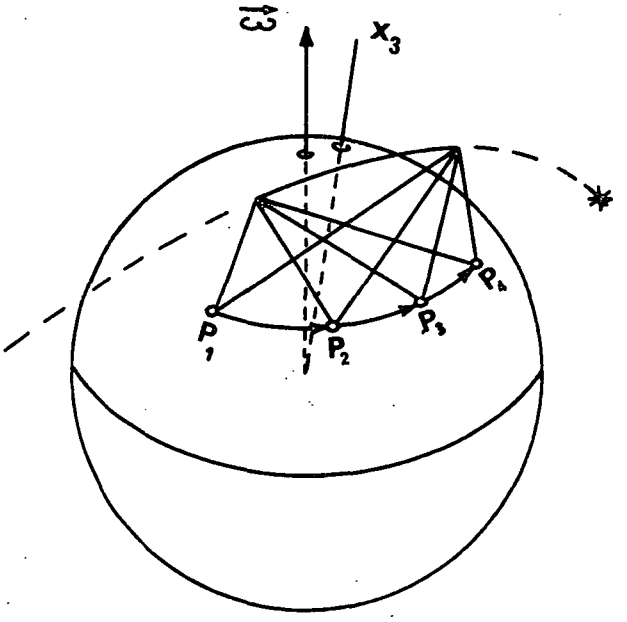



Fig. 6. A single station P (rotated to respective positions $P_1$, $P_2$, $P_3$ and $P_4$) is able to track four consecutive passes of an artificial satellite near apex and thus derive its orbital inclination.

fining an x-system and the GEM-10 geopotential model. The data span was devided into sixteen 2-day arcs. Special care was taken to model upper atmosphere effects. Weighted least squares straight line fits to the determined $x_p$ and $y_p$ values yielded estimated $x_p$ and $y_p$ standard deviations smaller than 0.07 arcseconds. The results indicate that a determination of both pole position components is possible from a regional station network, that as expected, the $y_p$-component is more accurately recovered than the $x_p$-component and that additional model improvements may be required for further analysis of Geos-3 data. The method used has been documented in more detail in connection with polar motion analysis of earlier Geos-3 data [Schutz et al, 1978a]. Although at first glance not impressive the authors consider the results from the 1976 data analysis of particular significance because this data was not used in the GEM-10 geopotential and station coordinates solution.

The LLR experiment has been outlined by Bender et al [1973]. A total of five retroflector arrays has been deposited on widely dispersed sites on the moon by the US and USSR space programs and subsequently laser ranged. Although returns have been announced by nine stations in five countries [Silverberg, 1978], LLR on a routine basis has in essence so far been a single-station operation from the McDonald Observatory. This restricted the yield of practical results as regards Earth orientation, like this was restricted in single-station SLR. Nevertheless, like single-station SLR demonstrated a capability to measure polar motion, so demonstrated the single-station LLR a capability to measure UTO as pertaining to that station.

The lunar orbit differs most markedly from artificial satellite orbits in its much larger semimajor axis, which entails a 200 to 400 times longer period of revolution and a superior orbital covarage from a single observing station. These features have consequencies for both the data acquisition and the data analysis. Apart from a gap around new moon as caused by guiding problems which seem difficult to overcome, a geographically well selected LLR station can observe the moon daily, more or less evenly distributed over a full revolution. As compared to SLR the longer period and the high degree of predictability of the lunar orbit entail distinct differences in observation data handling. A LLR "normal point" covers about 5-15 minutes of ranging, a portion of a lunation too short to construct orbital elements. Such normal point may be regarded as an equivalent photon return representing the observation run. Over the past few years an average of 25-30 normal points have been produced at McDonald per lunation [Mulholland, 1978]. When solving for Earth orientation a normal point is fitted to a retroreflector reference orbit, obtained by taking into account the selenodetic retroreflector coordinates and the physical librations of the moon. The range residuals are then analysed in an attempt to recover Earth orientation signatures. This procedure follows the general pattern as outlined before and seems to deviate from SLR data analyses to recover Earth orientation, only in so far dictated by the kinematics of the lunar orbit and the ensuing data acquisition strategy.

Range residuals have been analysed by several authors to study various aspects of the Earth-moon system, including the Earth's diurnal rotation in terms of UTO; Stolz et al [1976], Harris and Williams [1977], Shelus et al [1977], King et al [1978], Calame [1978]. Single-station operation implies that as regards the direction of the rotation axis, only the projection of its change (polar motion) with respect to an Earth-fixed x-frame, onto the station's meridian can be measured. This, in turn, implies that not UT1, but only UTO can be derived, since the transverse component of pole position, required to correct UTO in order to obtain UT1, cannot be determined from a single station. Therefore Earth orientation results from LLR have up till now only been obtained in terms of UTO-corrections valid for McDonald. This nevertheless is a unique contribution, because it is a field in which reliable results from SLR are not expected in the near future, and if so, these will measure the rotation rate $\omega$, rather than the angular position of the Earth-fixed frame in terms of the siderial angle as obtained from UT1.

First UTO-values derived from LLR were reported by Stolz et al [1976]. Using diurnal variation of range residuals they obtained, scattered over a roughly 5 years' time span, 194 single day values with a medium standard deviation of 0.7 msec if allowance is made for the uncertainty of the current lunar ephemeris. Only days with data well distributed in time were included in the analysis.

Rather than selecting the best isolated observation days of a data span, Shelus et al [1977] obtained UTO results for a complete lunation on a daily basis, accepting all observations, as one would have to do in a routinely operating Earth rotation monitoring service. The results were obviously less precise than those of Stolz et al [1976], the standard deviation relative to the corresponding BIH data amounting to 2.6 msec in the most realistic case of analysis, but they may provide a more realistic measure of what can be expected on a continuing near real-time basis from a single station operating under normal conditions. The results obtained by Stolz et al [1976], on the other hand, demonstrate what can be done on a daily averaging basis under the best conditions. It should possibly not be overlooked that although Shelus et al [1977] did not select the best days of data within a lunation, they seem to have selected the best LLR lunation at McDonald.

King et al [1978] analysed the normal points on four retroreflectors taken at McDonald between October 1970 and November 1975 and solved for corrections to UTO in terms of 126 "tabular points" defining a continuous, piecewise linear function of time spanning the 5-year interval. The solution was a phased one for a total of 166 parameters, only the UTO parameters being obtained in the second phase. After removing a constant difference, the standard deviation of the tabular points as compared to BIH values is 2.1 msec. The standard deviation of a comparison of the tabular points with the results obtained by Stolz et al [1976], who analysed most of the same data, is 0.8 msec.

Calame [1978] proceeded along several lines, varying the selection of parameters solved for simultaneously with UTO, data selection criteria and averaging time (1, 2 or 5 days). Using data through January 1978, the estimated UTO standard

deviation ranges from about 1 to 3 msec.

A major break-through as regards Earth orientation from LLR is expected from the deployment of the proposed multi-station EROLD-network [e.g. Mulholland and Calame, 1978]. Such network would enable the separation of all three components of Earth orientation at a level of precision assessed by Stolz and Larden [1977]: better than measuring accuracy with an averaging time of two days. The BIH's involvement in EROLD has been outlined by Calame et al [1976].

Summarizing one could say that both SLR and LLR have already demonstrated some of their capabilities and defaults in single-station operation to measure Earth orientation. Multi-station results are available from SLR. Multi-station LLR is expected in the near future. Up till now SLR and LLR have been complementary in that SLR provided mainly polar motion, while LLR yielded information on diurnal rotation. In multi-station operation one would expect both SLR and LLR to provide complete three-parameter Earth-orientation, comprising both polar motion and diurnal rotation. Short averaging times are more likely to be achieved with SLR. On the other hand, the lunar orbit offers a more stable long term reference, in particular for the determination of UT1 and $\omega$. Considering this, SLR and LLR should be considered complementary techniques, rather than competitive. It is important to note here that at least some of the planned LLR stations will also have Lageos-capability. Alternatively Silverberg [1978] envisions SLR stations with occasional LLR capability. In such integrated arrangements the SLR could be used to track polar motion on a daily basis, the LLR from one or more of the stations being then used to measure UT1 applying the SLR polar coordinates to correct UTO as measured. Ultimately Earth orientation to equivalent measurement precision or better with averaging times of one day or shorter is expected from integrated LR.

## Concluding remarks

Figure 7 depicts types of Earth orientation results obtained or expected from laser ranging, classified according to modes of operation. As pointed out in the preceding paragraphs, only multi-station operation will be able to provide both polar motion and UT1 results. Ultimately both SLR and LLR should be able to do so independently. Up till the present time, only SLR has demonstrated multi-station operation and yielded polar motion in this mode. Once in multi-station operation, as foreseen in the EROLD campaign, LLR is more likely than SLR to provide long-term consistency of results in particular as regards UT1. On the other hand SLR may offer shorter averaging times. Considering moreover practical operational constraints, there is strong tendency to foresee SLR and LLR as operating in a complementary rather than in a competitive way when monitoring Earth orientation, allowing both techniques to contribute on their strong points. Instruments having both SLR and LLR capability seem advantageous in this respect. A future Earth orientation

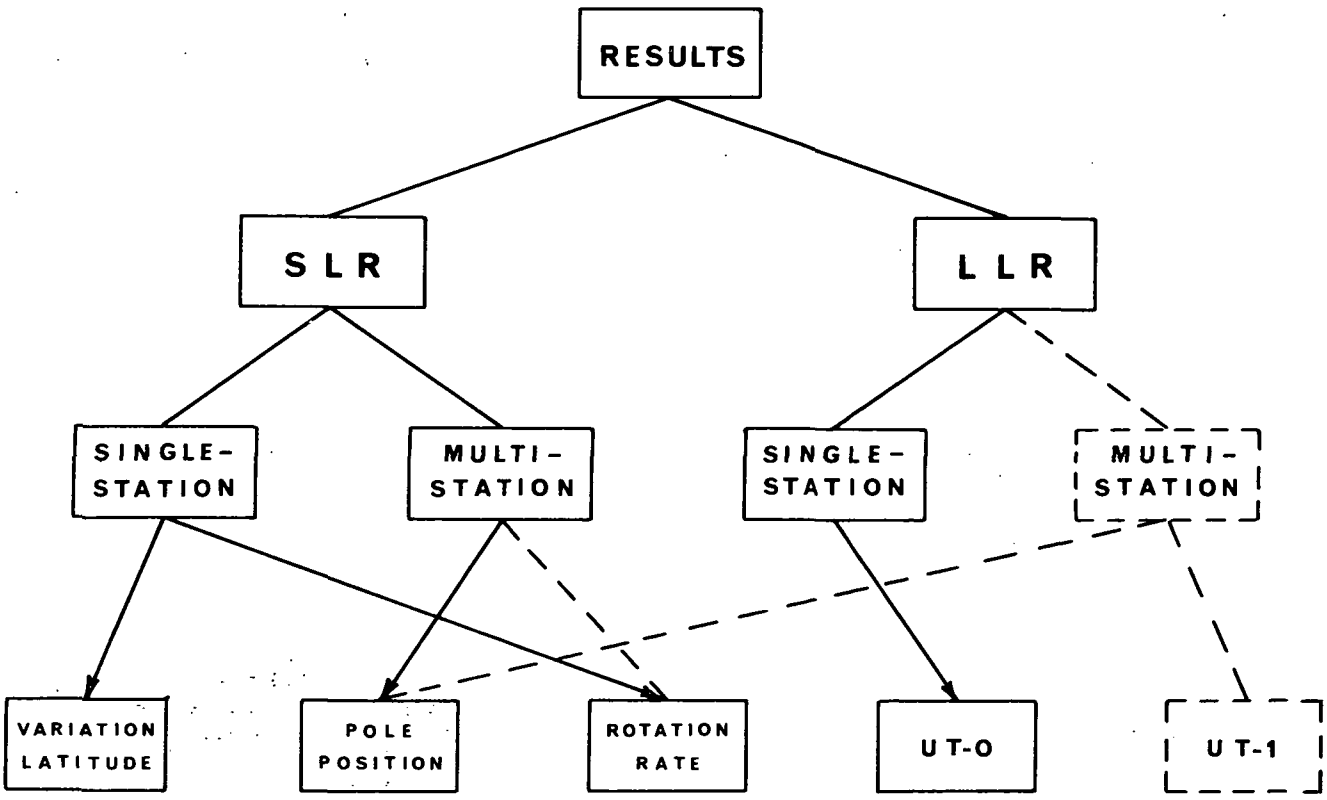



Fig. 7. Classification of Earth orientation results. Solid lines indicate demonstrated capabilities, dashed lines indicate potentialities.

service is supposed to provide pole position to 0.002 arcsec and UT1 to 0.1 msec with averaging times of one day or shorter. When extrapolating current experience into the future by means of simulation [e.g. Stolz and Larden, 1977; Smith et al, 1978b] it seems that the above requirements can be ultimately met by multi-station combined SLR and LLR. Additional requirements to be considered for a service bureau type of operation are continuity of results, quasi real time availability of results and cost-effectiveness.

Acknowledgement. The author is indebted to Dr. O. Calame (Centre d'Etudes et de Recherches Géodynamique et Astronomique, Grasse, France) for helpful information and criticism concerning the lunar laser ranging contribution to Earth orientation.

References

Bender, P.L., D.G. Currie, R.H. Dicke, D.H. Eckhardt, J.E. Faller, W.M. Kaula, J.D. Mulholland, H.H. Plotkin, S.K. Poultney, E.C. Silverberg, D.T. Wilkinson, J.G. Williams and C.O. Alley, The lunar laser ranging experiment, Science, 182, 229-238, 1973.

Calame, O., F. Chollet and B. Guinot, Participation of the Bureau International de l'Heure in the EROLD campaign, COSPAR Information Bulletin No. 77, 43-47, 1976.

Calame, O., Preliminary UTO results from EROLD data, IAU Symposium 82, Cadiz, Spain, May 1978.

Dunn, P.J., D.E. Smith and R. Kolenkiewicz, Techniques for the analysis of geodynamic effects using laser data, The use of artificial satellites for geodesy and geodynamics (G. Veis, ed.), Athens, Greece, 563-575, 1974.

Dunn, P.J., D.E. Smith and R. Kolenkiewicz, The measurement of latitude, time, and height variations at a single laser tracking station, J. Geophys. Res., 82, 895-897, 1977

Gaposchkin, E.M., Pole position studied with artificial Earth Satellites, Rotation of the Earth (P. Melchior and S. Yumi, ed.), Reidel Publ. Co., 128-130, 1972.

Harris, A.W. and J.G. Williams, Earth rotation study using lunar laser ranging data, Scientific applications of lunar laser ranging (J.D. Mulholland, ed.) Reidel Publ. Co., 179-190, 1977.

King, R.W., C.C. Councilman III and I.I. Shapiro, Universal Time: results from lunar laser ranging, J. Geoph. Res., 83, 3377-3381, 1978.

Kolaczek, B. and G.C. Weiffenbach (ed.), On reference coordinate systems for Earth dynamics, Proceedings IAU Colloquium No. 26, Toruń, Poland, 1974.

Kolenkiewicz, R., D.E. Smith, D.P. Rubincam, P.J. Dunn and M.H. Torrence, Polar motion and Earth tides from laser tracking, Phil. Trans. Royal Soc. of London, A 284, 485-494, 1977.

Lambeck, K., Determination of the Earth's pole of rotation from laser range observations to satellites, Bulletin Géodésique, 101, 263-281, 1971.

McClure, P., Diurnal polar motion, X-592-73-259, NASA/GSFC, 1973.

Melchior, P., Past and future of research methods in problems of the Earth's rotation, Rotation of the Earth (P. Melchior and S. Yumi, ed.), Reidel Publ. Co. XI-XXIII, 1972.

Mulholland, J.D., Is lunar ranging a viable component in a next-generation Earth rotation service?, IAU Symposium 82, Cadiz, Spain, May 1978.

Mulholland, J.D. and O. Calame, Earth Rotation from lunar distances: basis and current status, Precise time and time interval no. 9, NASA TM 78104, 1978.

Newton, R,R., Coordinates used in range or range rate systems and their extension to a dynamic Earth, On reference coordinate systems for Earth dynamics (B. Kolaczek and G. Weiffenbach, ed.), Warsaw, Poland, 181-200, 1974.

Pearlman, M.R. and K. Hamal (ed.), Proceedings Workshop on laser ranging instrumentation, Lagonissi, Greece, May 1978.

Schutz, B.E., B.D. Tapley, J. Ries and R. Eanes, Polar motion results from Geos-3 laser ranging Institute of Advanced study in Orbital Mechanics Tech. Rep. 78-4, The Univ. of Texas at Austin, 1978a.

Schutz, B.E., B.D. Tapley and J. Ries, Polar motion from laser range measurements of Geos-3, IAU Symposium No. 82, Cadiz, Spain, May 1978b

Shelus, P.J., S.W. Evans and J.D. Mulholland, Earth rotation as inferred from McDonald Observatory lunar laser observations during October 1975, Scientific applications of lunar laser ranging, (J.D. Mulholland, ed.), Reidel Publ. Co., 191-200, 1977.

Silverberg, E.C., On the effective use of lunar ranging for the determination of the Earth's orientation, IAU Symposium 82, Cadiz, Spain, May 1978

Smith, D.E., R. Kolenkiewicz and P.J. Dunn, Geodetic studies by laser ranging to satellites, The use of artificial satellites for geodesy, (S.W. Henriksen, A. Mancini and B.H. Chovitz, ed.), AGU Geophysical monograph 15, 187-196, 1972a.

Smith, D.E., R. Kolenkiewicz, P.J. Dunn, H.H. Plotkin and T.S. Johnson, Polar motion from laser tracking of artificial satellites, X-553-72-247, NASA/GSFC, 1972b.

Smith, D.E., R. Kolenkiewicz, P.J. Dunn and M.H. Torrence, Determination of station coordinates from Lageos, Second international symposium on the use of artificial satellites for geodesy and geodynamics, Athens, Greece, May-June 1978a.

Smith, D.E., R. Kolenkiewicz, P.J. Dunn and M.H. Torrence, Determination of polar motion and Earth rotation form laser tracking of satellites, IAU symposium No. 82, Cadiz, Spain, May 1978b.

Stolz, A., P.L. Bender, J.E. Faller, E.C. Silverberg, J.D. Mulholland, P.J. Shelus, J.G. Williams, W.E. Carter, D.G. Currie and W.M. Kaula, Earth rotation measured by laser ranging Science, 193, 997-999, 1976.

Stolz, A. and D. Larden, Accuracy obtainable for Universal Time and polar motion during the EROLD campaign, Scientific applications of lunar laser ranging (J.D. Mulholland, e.d), Reidel Publ. Co., 201-216, 1977.

Veis, G., Precise aspects of terrestrial and celestial reference frames, The use of artificial satellites for geodesy (G. Veis, ed.), North-Holland Publ. Co., 201-216, 1963.

# Principles of Very-Long-Baseline Interferometry

Irwin I. Shapiro
Department of Earth and Planetary Sciences
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Abstract. We present the basic principles of very-long-baseline interferometry as related to its use in the determination of vector baselines, polar motion, and earth rotation.

## Introduction

Ten years ago, almost to the day, the first successful bandwidth-synthesis VLBI measurements were made. It is thus appropriate to now review the principles underlying the technique. The review will be restricted to aspects relevant to geodetic applications that involve observations of extragalactic radio sources. In such applications, arrays of two or more radio telescopes observe any given source simultaneously. From sets of observations of a suite of such sources, one can obtain the desired geodetic information: baseline-vector, polar-motion, and earth-rotation parameters.*

We shall first describe briefly the instrumentation used in these observations and then discuss the basic observables and their simplest interpretation. Finally, we consider some complications of the interpretation due to the various geophysical "signals" and non-geophysical "noise" that affect the observables.

## Instrumentation

A VLBI system consists of an array of at least two antennas that observe the same radio source simultaneously. A direct electrical connection is not maintained between the antennas, thus allowing them to be separated by thousands of kilometers. The local-oscillator signals, used at each antenna to convert the radio-frequency signals from the source to the video (low-frequency) band, are derived from a frequency standard at the site. These standards are sufficiently stable that the relative phases of the signals from the source received at the two antennas are preserved.

The video signals are recorded on magnetic tape at each site, with the reference time for the recordings being derived from the same standard as is used to govern the local-oscillator signals. The tape recordings are then transported to a common center where those recorded simultaneously are cross-correlated to obtain the basic VLBI observables.

## Observables

The basic observables in geodetic VLBI experiments are (i) the difference in the times of arrival at two antennas of a signal from a source; and (ii) the rate of change of this time difference. The measurement of the time-of-arrival difference can be of two types: the phase-delay difference or the group-delay difference. The phase-delay-difference observable (hereinafter "phase delay"), being based on measurements of phase, can be obtained very precisely, but usually ambiguously, due to the inability to resolve the "$2\pi n$" problem. The group-delay-difference observable (hereinafter "group delay")is usually determined with less precision than is the phase delay. But the group delay, being determined by the rate of change of phase delay with frequency, is usually unambiguous. To obtain reasonable accuracy in the measurement of group delay it is necessary to make phase-delay measurements over a wide band of frequencies simultaneously, or nearly simultaneously, since the uncertainty in the group-delay measurement is inversely proportional to this bandwidth. In practice, only a relatively narrow band of frequencies can be recorded. But this band can be split up into narrower bands which are spread over a very wide band. This technique is called bandwidth synthesis. In the newest, Mark III, system (see below), 28 narrow bands, each 2 MHz wide, are being distributed over a total band of up to 400 MHz. If the error in the measurement of phase for any one band is $\sigma(\phi)$, then the error $\sigma(\tau_g)$ in the measurement of the group delay will be given by $\sigma(\tau_g) \sim \sigma(\phi)/\Delta f$, where $\Delta f$ is the rms spread of the center frequencies of the individual bands about their mean. Because these individual bands do not cover the entire spanned band, the estimate of the group delay, too, could be ambiguous. However, a proper choice of the spacing of the individual bands, as explained below, can insure that any inherent ambiguity can be eliminated reliably.

The difference of the phases of the

*Several groups are currently engaged in such geodetic applications of VLBI: (1) A group from the Goddard Space Flight Center, the Haystack Observatory, the Massachusetts Institute of Technology, and the National Geodetic Survey; (2) a Jet Propulsion Laboratory - National Geodetic Survey group; and (3) a Canadian-British collaboration. A European consortium is also being organized for similar purposes.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

signals that would be received at two antennas as a function of frequency exhibits curvature due primarily to the effects of the earth's ionosphere: the lower the frequency, the shorter the phase delay. The actual shape of the phase vs. frequency curve for VLBI observations will depend on the relative amounts of plasma between the source and the two sites. But with appropriate spacings of the narrow bands, one can "connect" unambiguously the phase at one band with those at all the other bands. For example, bands spaced in accord with a geometric series allows use of a bootstrap technique to connect phases first between the closest bands and then between the more widely separated bands through use of the characteristics of the curve established by the connection between the closest bands. Of course, the "absolute" phase would still be uncertain by multiples of $2\pi$, but the relative phases between bands would be freed from any such ambiguity. Only the relative phases affect the group delay which is equal to the slope of the curve of phase delay vs. (angular) frequency.

## Information Content

We now consider the information content of the observables under simplified assumptions. In particular, let us ignore the propagation medium and assume that the earth is rigid and rotates with a constant, known, angular velocity. We may then write the expression for the delay observable as a function of time as:

$$\tau(t) = \frac{1}{c}\vec{B}(t)\cdot\hat{s} + \tau_o^{cl} + \dot{\tau}_o^{cl}\,[t-t_o] \quad , \qquad (1)$$

where $\vec{B}$ is the baseline vector connecting a pair of antennas, $\hat{s}$ is a unit vector in the direction of the source, and $\tau_o^{cl}$ and $\dot{\tau}_o^{cl}$ are, respectively, the offsets in epoch and rate of the clock at one site with respect to those at the other site. This equation represents a diurnal sinusoid added to a straight line. The first term in the equation contributes the diurnal sinusoid due to the rotation of the baseline vector in inertial space. The slope of the straight line is due to the clock-rate offset and the intercept is due to a combination of the clock-epoch offset and the product of the polar components of $\vec{B}$ and $\hat{s}$. Clearly this curve can be specified by four parameters: the intercept and slope of the straight line, and the amplitude and phase of the sinusoid. (We do not consider the period of the sinusoid, since that is given by assumption.) Thus four measurements of the delay suffice, in principle, to determine $\tau(t)$; any additional measurements will be redundant. But how many unknown parameters are there for this situation? Naively, one would conclude that the baseline vector contributes three, the source two, and the clocks two, for a total of seven. However, the origin of right ascension of our system is arbitrary; only the origin of declination is fixed by the assumption of a known angular velocity for the earth. Since the right ascension of the source can be used to define this arbitrary origin, the number of unknown parameters is only six. Nonetheless, a unique solution cannot be obtained for these six parameters from observations of a single source. Observations of each additional source adds two unknowns: the coordinates of the source on the plane of the sky. But such observations also can be used to determine three additional parameters: the intercept of the straight line and the amplitude and phase of the sinusoid appropriate for the additional source. The slope of the straight line provides no new information since it is determined solely by the clock-rate offset. It is clear that with four observations of one source and three each of two more sources, a useful solution for all of the parameters of this simple model can in general be determined. (The measurements of delay rates simultaneously do not reduce the requirement for observations of three sources.) The accuracy of the determination of these ten parameters will depend, of course, not only on the accuracy of the measurements of delay (and delay rate), but also on the baseline, the distribution of the sources in the sky, and the distribution of the observations in time.

## Complications

The situation actually encountered with VLBI is, of course, far more complicated than outlined in the previous section. We can conveniently divide these complications into two categories: signals and noise. Here signals refer to those effects on the observables which are of geophysical interest, and noise refers to those of no intrinsic interest. (This point of view, needless to say, is a rather parochial one since one person's noise is often another's livelihood.)

We shall consider precession, nutation, solid-earth tides, crustal motions, variations in UT1, and polar motion to be signals. On the other hand, clock instabilities and uncertainties in our knowledge of source characteristics and of the propagation medium shall be considered as noise. We discuss each set in turn.

### Signals

Precession and Nutation. Changes in the direction in space of the spin axis of the earth with periods long compared

to a day are sensed with VLBI through the corresponding changes in the coordinates of the radio sources. These changes will, however, preserve the arclengths between sources. At present, estimates of the precession constant made from VLBI measurements have an uncertainty of a few tenths of an arcsecond per century, several-fold larger than the uncertainty associated with the presently accepted value based on optical observations. The VLBI estimate is consistent with the optical one to within twice the formal standard error of the former. No estimates of any of the nutation terms have yet been made.

Solid-Earth Tides. The semi-diurnal solid-earth tide imparts a distinctive signature to the VLBI observable, since almost all other effects introduce a diurnal signature. The maximum magnitude of this effect on the observable has been slightly greater than one nanosecond. Thus estimates of the local values of the vertical and horizontal Love numbers, $\ell$ and $h$, can be obtained from VLBI data. Current estimates agree, to within their uncertainty of about 0.05, with the "expected" values.

Crustal Motion. Changes in crustal configuration can be sensed by long-term changes in baseline lengths; in addition, significant changes in the corresponding baseline directions for an array of antennas would signify crustal motions provided that these changes were incompatible with a rigid rotation of the array. No measurements of crustal motions have yet been obtained from VLBI data, but there is every reason to believe that such motions will be detected within a few years.

UT1 and Polar Motion. Variations in the rate of rotation of the earth and in the position of the axis of figure with respect to the axis of rotation affect the directions of baseline vectors. As with the arclengths between sources with respect to precession and nutation, the lengths of baselines are unaffected by such variations in the rate of rotation of the earth and in the position of the pole.

We should stress that the VLBI observables, for an arbitrary baseline, have no sensitivity to the "initial" orientation of the earth and direction (in space) of its axis of figure. Only changes in these quantities can be detected. In addition, any "common-mode" errors in the epoch settings of the clocks at the antenna sites will be indistinguishable from corresponding changes in the orientation of the earth about its spin axis (UT1). Finally, note that VLBI data obtained for one baseline are sensitive to only two independent combinations of the three parameters needed to specify changes in the position (in space) of the axis of figure of the earth and in the orientation of the earth about this axis: Such changes affect only the direction of the baseline which is described by only two independent parameters. Consider, as an example, a wholly north-south baseline. With such a configuration, the VLBI data would have no sensitivity to a rotation of the earth about the pole. For an east-west baseline, on the other hand, polar motion in a direction along the meridian of the midpoint of the baseline would not affect the VLBI observables. In both these cases, the baselines would undergo parallel displacements which cannot be detected from observations of sources "at infinity." Two baselines, or at least three antennas, are needed in a VLBI array to detect all three components of the changes in UT1 and pole position. These baselines must, of course, not be parallel.

Estimates of UT1 and polar motion from VLBI data now have accuracies comparable to those of other techniques, such as the classical optical methods, the Doppler tracking of satellites (for polar motion), and the laser ranging to retroflectors on the moon (for UT1). It is expected that the accuracy of the VLBI estimates will improve nearly tenfold within the next five years due primarily to the introduction and use of the new, Mark III, VLBI system. A prototype of this system has already been tested successfully; five copies of the complete system are currently under construction for placement at suitably distributed antennas. It will be important to check the improved determinations of polar motion and UT1 through redundancy and through comparison with the results from other improved techniques in order to assess the accuracy of these determinations.

## Noise

Clock Instabilities. The two parameters, for clock epoch and rate offsets, do not provide an adequate representation of the relative behavior of the clocks at any two antennas of an interferometric array over the period of many hours needed to determine the baseline vector, source positions, etc. This statement applies to the current field units of all atomic clocks, including the hydrogen-maser frequency standards.

A number of possibilities exists to minimize the impact of these clock instabilities on the accuracy with which geophysical information can be extracted from VLBI data. First, one can use higher-order polynomials to represent the relative clock behavior; here the point of "diminishing returns" sets in at about the sixth order. Second, the clock performance, especially of hydrogen-maser standards, can be improved to match that achieved in the laboratory. The sensiti-

vity of the maser standards to environmental effects can also be reduced to minimize the introduction of diurnal signatures into the VLBI data. Third, one can reduce the effects of long-term drifts in the relative clock behavior by using clock stars. Thus, one can make observations repeatedly, say every hour, of some suitable source and use these observations to correct for the relative clock drifts. To be suitable, this source should be visible from both sites for a large fraction of the diurnal cycle and should yield a reasonably large correlated flux density so that accurate delay observations are possible to make.

Source Characteristics. The radio sources affect the determination of geophysically interesting quantities through the strength of their radio emissions, their distribution on the sky, and the accuracy with which we can determine their positions. These positions, in turn, depend on the structure and internal kinematics of the regions of radio emission in each object.

At present, the entries in the catalog of known, and potentially-usable, extragalactic radio sources number in the hundreds. Positions of a few dozen of those with the strongest emissions are now being determined routinely with an estimated accuracy of about 0".02, except for the declinations of sources that lie near the equatorial plane. The accurate determination of the declination of those sources requires the use of interferometers with baselines that possess large components in the north-south direction. Few such baselines have so far been available for extensive sets of measurements.

Aside from the examination of the characteristics of the postfit residuals, the main method for assessment of the accuracy of source-position determinations is the comparison of results obtained with different equipment. Such comparisons, made several years ago and based on data obtained with somewhat less advanced VLBI systems, showed agreement to within about 0".05 rms. (Note that a 0".001 error in source position corresponds approximately to a two-centimeter error in length for a 4,000-km baseline.)

Most extragalactic radio sources are not "points" when viewed on the scale of milliarcseconds. Rather, they exhibit complicated structure. This structure in their brightness can be mapped and a suitable feature in the map, or the overall center of brightness, can be used as a reference point. There are, however, technical difficulties in the determination of unambiguous brightness maps. These difficulties are being overcome and reliable maps on the scale of tenths of a milliarcsecond are now being obtained for some of the radio sources.

There is yet a further difficulty in the use of extragalactic radio sources: most are not static. Dramatic changes have been observed in the brightness structure of some of these sources at the level of tenths of a milliarcsecond in angular resolution on a time scale of a few months. Thus, to enable positions of extragalactic radio sources to be used effectively as a reference system at the level of milliarcsecond accuracy for geophysical applications of VLBI, one must monitor the brightness distributions of these sources as a function of time and, perhaps, as a function of radio frequency as well.

Propagation Medium. In regard to any substantial effects on VLBI data, the propagation medium can be considered to be composed of two components: the ionosphere and the troposphere. The effects of the ionosphere can, and will, be virtually eliminated by observing simultaneously in two widely separated radio frequency bands ($\simeq$2GHz and $\simeq$8GHz). The Mark III VLBI system is equipped for such dual-band observations. Moreover, enough suitable sources exist to allow effective use of the dual-band technique.

The troposphere is in effect non-dispersive at radio frequencies and is therefore a more troublesome contributor of noise. The troposphere can also be decomposed into two components: wet and dry. For the latter, the assumption of hydrostatic equilibrium is a very good one; measurements at each site of surface pressure combined with a good model of the atmosphere, then allows a good estimate to be made of the phase delay added in the zenith direction by the dry component: about 7.5 nanoseconds (equivalent to an increase in path length of about 2.3 m). It is widely thought that the error in this estimate can be kept at the 0.1% level or perhaps below. Mapping to other zenith angles, however, will increase the error somewhat since the "slant" atmospheric path length cannot be determined so accurately from the pressure measurement at the antenna site. The situation with the wet component is more difficult. The water-vapor in the atmosphere is not in hydrostatic equilibrium and is quite variable in amount. Although the total effect on the path length of radio waves is, on average, only about 7% of that of the dry component, the wet component cannot be modeled accurately. Various simple techniques have been used to try to ameliorate this problem. Such techniques involve various combinations of model atmospheres and mapping functions with or without dependence on surface measurements of temperature, pressure, and dew point, and with or without parameters that can be estimated for each site from short, ~8 hr, spans of data. Unfortunately, these techniques may well be defi-

cient, especially for long baselines, in removing the effects of the atmosphere on the estimates of the "vertical" component of the baseline with which they are highly correlated.

The technique which has elicited the greatest expectations for providing the solution to the wet-component problem is based on the use of radiometer measurements at each site of the brightness temperature of the atmosphere at and near the ≈23GHz line in the spectrum of water-vapor emission. Studies indicate that this brightness temperature can be related with reasonably high accuracy to the excess path length attributable to the water-vapor content of the atmosphere. However, to date, almost all VLBI results have been obtained without the benefit of water-vapor radiometer measurements.

Atmospheric effects thus loom as the limiting factor in the accuracy achievable with VLBI in the determination of geophysical quantities. What will that limit be? An assessment based on theory alone is unlikely to be accurate. Measurements are clearly called for. Series of VLBI experiments should be made with supplementary water-vapor radiometer measurements, under a variety of local weather conditions, and for various baseline lengths. For short baselines, up to several kilometers in length, independent determination of the baseline vector can usually be made, with some effort, at the millimeter level of accuracy by means of conventional survey techniques. For long baselines, up to several thousand kilometers in length, independent means of verification at the relevant level of accuracy seem to be limited to laser ranging to artificial satellites or to the moon; such verification, however, will not be easy nor inexpensive for a number of practical reasons. The repeatability and consistency of VLBI results themselves may well have to provide the main standards. Since suitably accurate, and independent, estimates of UT1 and polar motion may not be available, repeated checks on the individual components of the vector baseline will likely require multi-site experiments, say with four or more separated antennas, to reduce the confusion between UT1 and pole position changes on the one hand and changes in baseline direction on the other. The many antenna sites serve to over-determine UT1 and polar motion, with the redundancy providing the meaningful check on the consistency of the estimates of some of the baseline components. For some combination of the precision and the time spanned by the sets of measurements, one must be concerned also about the genuine changes in baselines expected from plate tectonics; of course, detection of such changes are a major purpose of the measurements.

At present, checks on the repeatability of baseline determinations have involved primarily two-element interferometers. For short baselines, of the order of one kilometer in length, repeatability has been obtained at the five millimeter level in all components, and verified later by the results of a conventional survey. For long baselines, of the order of several thousand kilometers in length, only repeatability in baseline length has been meaningful; here the spread about the mean in a recent series of a dozen sets of measurements was under five centimeters. The source positions used in the analysis of each of these experiments were fixed in accord with the results from the ensemble of experiments; errors in these positions tended therefore not to have a serious effect on the repeatability of the determinations of baseline length.

## Conclusion

The future of VLBI as applied to geodetic and geophysical problems, especially to the determination of UT1 and polar motion, looks quite bright. Although the last ten years have been devoted almost exclusively to the development of VLBI, the next ten should yield significant results.

## Acknowledgment

This work was supported in part by the National Science Foundation and in part by the National Aeronautics and Space Administration.

## Bibliography

Cannon, W. H., et al., 1979, J. Geophys. Res., 84, in press.

Cohen, M. H., et al., 1977, Nature, 268, 405.

Clark, T. A., et al., 1976, Astron. J., 81, 599.

Ong, K. M., et al., 1976, J. Geophys. Res., 81, 3587.

Robertson, D. S., et al., 1979, Proc. of IAU Colloq. No. 82, in press.

Rogers, A. E. E., et al., 1978, J. Geophys. Res., 83, 325.

Shapiro, I. I., et al., 1974, Science, 186, 920.

Thomas, J. B., et al., 1976, J. Geophys. Res., 81, 995.

# Polar Wander Analysis from Paleomagnetic Data

Rob Van der Voo
Department of Geology and Mineralogy, University of Michigan
Ann Arbor, Michigan 48109

Abstract. Utilizing marine magnetic anomalies and paleomagnetic pole positions, paleogeographic maps have been constructed for three time intervals back to the Early Cretaceous. From the maps lithospheric plate motions have been calculated and these global displacement fields have been analyzed to determine best-fitting rigid rotations , which then could be ascribed to true polar wander. The values so obtained are no larger than a few degrees and are within the magnitude of the uncertainties involved.

## Introduction

Paleomagnetic pole positions determined for a given continental unit can be connected in a temporal sequence; the result is called an apparent polar wander path for that continent. Assuming the ancient paleomagnetic field to have been, on average, dipolar, co-axial, and roughly geocentric, the apparent polar wander paths describe the movement of the pole with respect to the continent held fixed. In reality, the apparent polar wander paths may be due to either or both of two causes: (i) the motion of the plate relative to a fixed rotation axis or (ii) the motion of the rotation axis relative to the lithosphere as a whole. It is commonly accepted that at least some of the differences in the apparent polar wander paths of different continents are due to the relative motions of the continents; on the other hand, it has been suggested a long time ago (Creer et al.,1957) that the comparable length of the apparent polar wander paths of the individual continents imply a common cause, true polar wander.

Jurdy and Van der Voo (1974, 1975a, 1975b) have discussed previous attempts to separate the effects of continental drift (plate motions) from true polar wander and have proposed a new method to achieve this. I will briefly summarize their method here and present the results of their analysis.

## Paleogeographic Maps and Paleomagnetic Poles

Maps have been constructed showing the estimated positions of the continents and plate boundaries for the Early Tertiary, the Late Cretaceous, and the Early Cretaceous. Two of these maps are shown here (Figures 1 and 2). The relative positions of the plates have been determined from marine magnetic anomalies, using published data (referenced in Jurdy and Van der Voo, 1974, 1975). All paleomagnetic pole determinations which passed minimum reliability criteria (such as a significant number of samples, demagnetization experiments, well-determined ages, etc.) have been used by combining them in approximate time intervals of 30 million years for each map (e.g., Early Tertiary: 40 to 70 my BP; Early Cretaceous: 100 to 130 my BP). Since the continents were rotated in order to obtain relative ancient positions, the paleomagnetic poles were rotated with

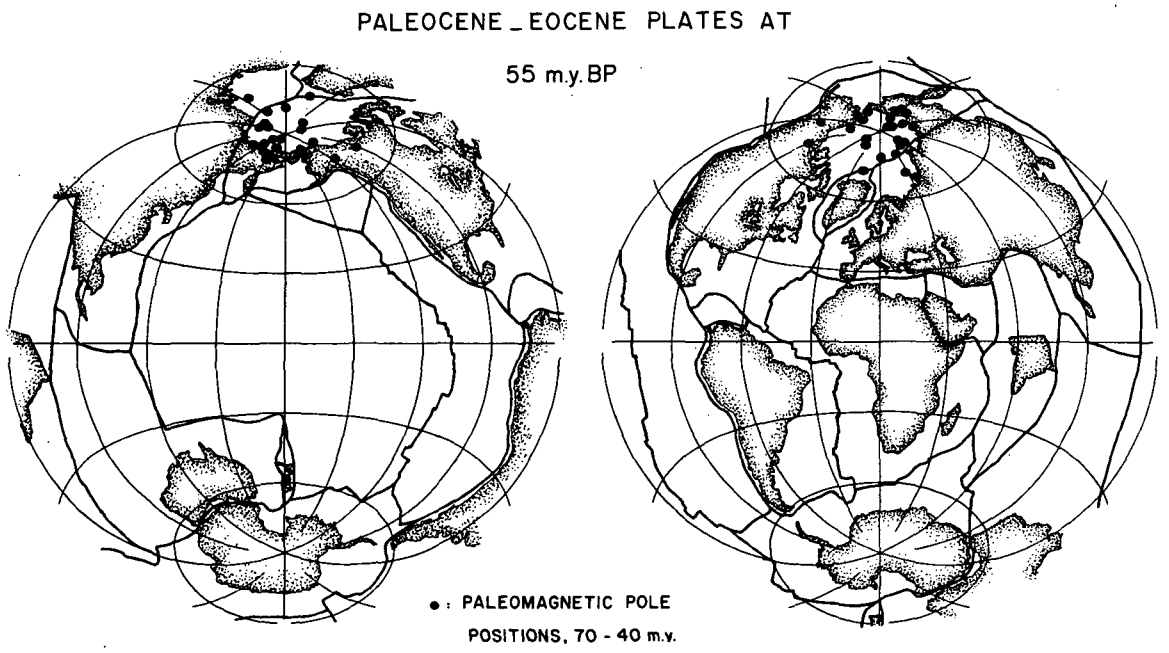



Fig.1. Paleogeographic map of the plates and plate boundaries for the Early Tertiary. The paleomagnetic poles used for the determination of the mean are plotted. The longitudes are arbitrary.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

the continents. It must be noted that the grouping of the poles is significantly better for all periods for the paleo-configurations than for the present-day positions of the plates.

## The Polar Wander Analysis

Knowing the plate positions at the present and at these earlier times, Jurdy and Van der Voo constructed for each of the three time intervals, a displacement field which describes the plate motions during an interval. True polar wander has been defined (e.g., Munk and MacDonald, 1960) as a bodily shift of the earth relative to its spin axis and, if it had occurred, it would appear as a special kind of surface displacement field: a pure rigid rotation about some axis. If no relative motions between plates had occurred but such a rigid rotation of the lithosphere had taken place, it would be observed as true polar wander and the apparent polar wander paths of all continents would coincide. If, on the other hand, no true polar wander but only plate motions relative to each other had occurred, the apparent polar wander paths of different plates would diverge in the geological past. The method of Jurdy and Van der Voo can therefore be best summarized as finding a best-fitting rotation of the entire lithosphere, and it is this rotation that must be attributed to true polar wander. It is crucial for an understanding of the method to realize that in the case of no true polar wander a summation or integration of the displacements of all the plates over the entire surface of the earth would yield a zero global average. In more mathematical terms, such plate motions do not contribute to a first-degree displacement field ( a rotation of the entire lithosphere) but instead make up the higher-degree displacement fields such as hemispherical twists or zonal rotations in which the motions are opposite to each other.

The plate configurations were evaluated in a coordinate framework fixed by the paleomagnetic data of the reassembled plates. The mean magnetic poles were used to fix the earth's polar axis and to define a Cartesian coordinate system in which latitudes but not longitudes can be determined. The axes lying in the equatorial plane thus must be positioned arbitrarily, and we chose to hold the longitudinal coordinates of North America more or less fixed. This arbitrary positioning of the equatorial axes implies an indeterminancy in the longitudinal displacements and indeed allows a rigid rotation of the entire lithosphere. However, this is a rotation precisely about the polar axis and such a rotation is irrelevant for true polar wander, since it leaves the position of the magnetic polar axis relative to the lithosphere as a whole unchanged. Consequently, if one decomposes a rigid rotation about a geocentric axis cutting a sphere at arbitrary latitude and longitude, into three component rotations about each of the coordinate axes, respectively, one finds that only rotations about the two equatorial axes will contribute to true polar wander and that the sum of the squares of these rotations is independent of the arbitrary positioning of the equatorial axes.

The plate reconstructions gave Jurdy and Van der Voo (1974, 1975a) three time intervals over which the displacement field could be evaluated. They developed a mathematical method to find the rigid rotation which best-fitted in a least-squares sense a set of observed displacements on a sphere. Denoting the observed displacement field by ($\bar{F}$) and a rigid rotation, corresponding to true polar wander, by a displacement field

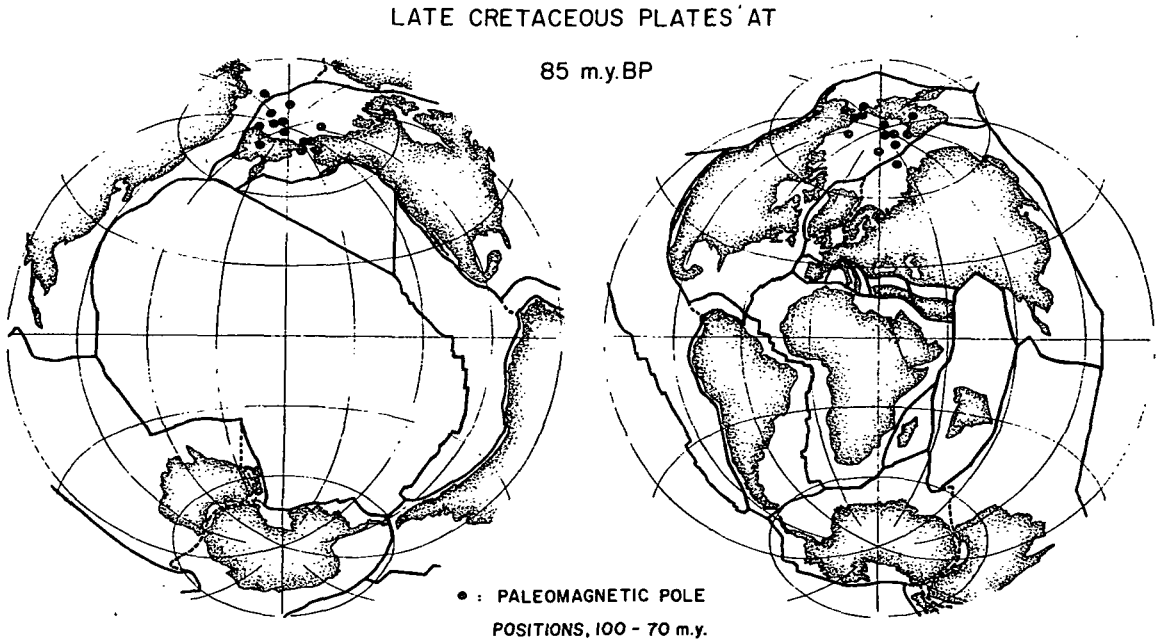



Fig.2. Paleogeographic map of the plates and plate boundaries for the Late Cretaceous. The paleomagnetic poles used for the determination of the mean are plotted. The longitudes are arbitrary.

($\bar{G}$), one can determine how much of the observed field $\bar{F}$ can be accounted for by a rigid rotation of the lithosphere $\bar{G}$, by minimizing the integral

$$\int (\bar{F} - \bar{G})^2 ds$$

where ds is an element of surface, and the integration is done over the entire surface of the earth. Further mathematical details have been presented by Jurdy and Van der Voo (1974).

The analysis of the displacement fields for the three time intervals considered yielded very small rigid rotations $\bar{G}$, an indication of little or no true polar wander. The calculated values are presented in Table 1, where the amount of true polar wander is the length of arc along which the pole moves relative to whole lithosphere or vice versa. The cumulative polar wander, found by adding the motion over longer time is given in Table 2, along with the uncertainty ($\alpha_{95}$, Fisher, 1953) in the mean paleomagnetic pole used for the plate reconstructions.

The resultant polar wander is quite small, even for the entire time interval from the present to the Early Cretaceous. This is due to the differences in the direction of movement (Table 1) over the three time intervals. The calculated amounts of true polar wander were generally within the uncertainty of the mean paleomagnetic pole and thus must be considered insignificant. Only for the earliest time interval (Early Cretaceous) did the amount of true polar wander exceed the value of the associated $\alpha_{95}$, but just for this early period additional uncertainties exist in the relative positioning of the plates. Since a primary requirement of the method is a reliable displacement field all over the earth, including the oceans, poor constraints on some Early Cretaceous plate positions (e.g., East and West Antarctica, the southern part of the Pacific ocean etc.) result in large uncertainties of the polar wander analysis (see also Jurdy, 1978). When better Mesozoic sea-floor spreading data become available, it is conceivable that more reliable determinations can also be carried out for the earlier Mesozoic periods.

TABLE 1. Calculated amounts of polar wander

| Time interval | Amount of polar wander | Direction relative to North America longitudes |
|---|---|---|
| Present to 55 my | 2.0° | 142°W |
| 55 to 85 my BP | 3.2° | 177°W |
| 85 to 115 my BP | 7.7° | 11°W |

TABLE 2. Cumulative amounts of polar wander

| Time interval | Amount of polar wander | Direction relative to North America longitudes | $\alpha_{95}$ |
|---|---|---|---|
| Present to 55my | 2.0° | 142°W | 4.1° |
| Present to 85my | 5.0° | 164°W | 5.0° |
| Present to 115my | 4.9° | 50°W | 4.7° |

References


Creer, K.M., E.Irving, and S.K.Runcorn, Geophysical interpretation of paleomagnetic directions from Great Britain, Phil.Trans.Roy.Soc. London, Ser.A., 250, 144-155, 1957.

Fisher, R.A., Dispersion on a sphere, Proc.Roy. Soc. London, Ser.A.,217,295-305,1953.

Jurdy, D.M., An alternative model for Early Tertiary absolute plate motions, Geology, 6, 469-472, 1978.

Jurdy, D.M., and R.Van der Voo, A method for the separation of true polar wander and continental drift, including results for the last 55 my, J.Geophys.Res.,79, 2945-2952, 1974.

Jurdy, D.M., and R.Van der Voo, True polar wander since the Early Cretaceous, Science,187,1193-1196, 1975a.

Jurdy, D.M., and R.Van der Voo, Reply (to D.P. McKenzie), J.Geophys.Res.,80,3373-3374,1975b.

Munk, W.H., and G.J.F.MacDonald, The rotation of the Earth, Cambridge University Press, London, 323 pp, 1960.

Page Intentionally Left Blank

N79 21466

# Present Status and Future Prospects for NASA's Program in Geodynamics

Edward A. Flinn
Geodynamics Branch, Office of Space and Terrestrial Applications
NASA Headquarters, Washington, D.C. 20546

Since 1972 NASA has been developing systems for precise determination of polar motion, earth rotation, and position on the earth's surface. The two observational methods are laser ranging to the moon and to artificial satellites, and very long baseline microwave interferometry (VLBI).

Lunar laser ranging has been done at the University of Texas' McDonald Observatory at Fort Davis, Texas, since 1969. A second lunar laser observatory is under development by the University of Hawaii on Haleakala, Maui (Hawaii). Similar observatories are also being constructed in Japan, Australia, and in Europe. Satellite laser ranging facilities capable of ranging to Lageos (6000 km altitude) are in operation at Haleakala, several sites in Europe, and at facilities operated by the Smithsonian Astrophysical Observatory (Mt. Hopkins, Arizona; Natal, Brazil; Arequipa, Peru; and Canberra, Australia).

Several radio telescopes are now equipped for geodetic VLBI observations: Haystack Observatory (Massachusetts), Owens Valley Radio Observatory (California), Onsala (Sweden), Bonn (West Germany), the Deep Space Network Station at Goldstone (California), and the Harvard antenna at Fort Davis (Texas). The National Geodetic Survey plans to establish a network of three dedicated VLBI observatories at Westford (Massachusetts), Richmond (Florida), and the Fort Davis antenna, in order to measure polar motion and earth rotation.

The range accuracy now being attained routinely at McDonald Observatory is about 6 cm. Repeated measurements of the baseline length between Haystack and OVRO agree to within 4 cm.

NASA has constructed eight mobile laser ranging facilities (Moblas) some of which have been used for the San Andreas Fault Experiment since 1972, in which the distances between two sites in California have been measured every two years. Another transportable laser ranging facility is being developed by the University of Texas at Austin; this is a small and highly mobile station, to be completed in 1980, that can move from site to site very rapidly and which should achieve an accuracy of better than 3 cm.

Two mobile VLBI facilities have been constructed by the Jet Propulsion Laboratory, one with a 9 m antenna and one with a 4 m antenna. The larger of these (ARIES) has been used to measure position differences between several sites in California and one of the fixed stations (Goldstone and OVRO).

A plan for a NASA geodynamics program has been written, and copies may be obtained on request from the address above. The document describes the activities NASA intends to carry out to apply space technology to research in earth dynamics. These activities may be categorized as being global, regional, and local in nature. On a global level, NASA will cooperate with other Federal agencies and with scientists in other countries to establish observatories for measurement of polar motion and earth rotation, to establish a global reference frame, and to serve as base stations for the mobile stations.

On a smaller scale, mobile stations will be used to establish a reference frame for the North American, Pacific, and Australia plates. A major part of the program will be to operate mobile stations in tectonically active regions in order to observe crustal movements related to tectonic processes. Initially the program will concentrate on Western North America, where a network of about thirty sites between the Colorado Plateau and the Pacific will be occupied by mobile stations about three times a year. The sites will be about 50 km apart near the San Andreas Fault, and up to 500 km apart toward the east.

Later, beginning in about 1982, the mobile stations will begin to operate in other active areas. Promising candidates are Central and South America, New Zealand, the Samoa-Fiji-New Caledonia area, and Japan. NASA scientists are involved in informal discussions with scientists in these areas, and it is hoped that agreements on cooperative programs will be developed within the next few years. The European Space Agency is formulating plans for a geodynamics program, and it is anticipated that a program similar to that under development by NASA will take shape in the near future.

The problem of measuring vertical and horizontal movements on a local scale (less than 50 km, say) is an important one. It appears that space methods may be capable of augmenting ground-based methods at this scale, and may have advantages in terms of economy and rapidity of coverage. Candidate methods include spaceborne laser ranging to retroreflectors on the ground, and the use of radio sources in artificial satellites to measure position differences in several ways. Specifically, the possibility appears promising to make use of the Global Positioning System satellites for this purpose. NASA intends to support studies of several systems for measuring local crustal movements.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

Page Intentionally Left Blank

# Mobile Satellite Ranging

Eric C. Silverberg
The University of Texas, McDonald Observatory
Austin, Texas 78712

## Introduction

For more than a decade, dozens of laser ranging stations have been active throughout the globe monitoring the orbits of about 15 retro-reflector-equipped artificial satellites. As has been reported by several publications, including Pearlman et al. [1975], McGunigal et al. [1975], and Masevitsch and Hamel [1975], the lasers are an operationally acceptable manner with which to closely monitor the orbital parameters of a wide variety of objects. Almost from the beginning, the possibility of establishing geodetic benchmarks by the use of laser satellite tracking was recognized, leading to a number of convincing demonstrations over selected baselines [Smith et al., 1978]. Even though these experiments were conducted at quasi-fixed locations, owing to the lack of truly mobile facilities, and used low level orbiting satellites, they showed great promise for the use of this technique over continental distances. With the relatively recent launch of the Lageos target into an unusually stable orbital configuration [Pearlman et al., 1976], and with the continuing development of improved techniques for laser ranging, we are now in a position to realize a much wider utilization of this geodetic application. One can expect the current emphasis being placed on the measurement of crustal motions (and on earth dynamics in general) to fuel an evolution in the field of satellite ranging, from the use of the versatile fixed station devices, which served so well until now, to the procurement of specialized, highly mobile instrumentation devoted to the global establishment of geodetic benchmarks.

The main thrust of the following article will be a brief review of the constraints which have limited satellite ranging hardware and an outline of the steps which are underway to improve the status of the equipment in this area. In addition, some suggestions will be presented for the utilization of newer instruments and for possible future R & D work in this area.

## Current Status

Most of the three dozen laser ranging stations now in existence were designed as fixed observatories. A small subset, consisting of eight NASA stations and two French stations [McGunigal et al., 1978 and Gaignebet, 1978], was configured to be routinely transported between various locations. Several other satellite stations can also be classified as transportable [Pearlman et al., 1978; Wilson, 1978; and Hamel, 1978]; but, to this author's knowledge, they are not expected to move between various locals with any frequency. While none of these existing ranging units is considered "mobile", their operation at many different locations has demonstrated the geodetic benchmark capability and has provided a test-bench for many operational concepts. The operating parameters which were used by these first generation transportable systems are summarized in Table 1. The two French systems generally represent the smallest and largest of the relevant parameters, while the NASA/Goddard system parameters lie somewhat between.

Each of the current stations is sufficiently versatile to permit ranging either the bright, very fast moving, low targets, or the much higher NTS2 satellite, which gives the weakest return of the current lot. Such flexibility, however, is not without cost. The current systems must routinely cope with over five orders of magnitude in target strength. Until very narrow pulse lasers recently became available, high accuracies dictated the use of high radar signals, sometimes

TABLE 1. A Summary Of The Operating Parameters For The Transportable Satellite Ranging Stations

| | Smallest | Largest |
|---|---|---|
| Aperture | 0.36 m (CNES I) | 1.0 m (CNES II) |
| Energy per Pulse | 0.25 J (Moblas II-VIII) | 2 J (CNES II) |
| Firing Rate | 0.5 PPS (CNES II) | 3 PPS (Moblas) |
| Pulse Width | 2 nsec (CNES II) | 12 nsec (CNES I) |
| Signal Strength on Lageos | 5 photoelectrons/shot | 80 photoelectrons/shot |
| Single Shot Uncertainty | $\sim$ 4 cm (RMS) | $\sim$ 75 cm (RMS) |
| Size (critical components) | 1 van approximately 2.3 x 3.5 x 12 m | 3 vans approximately 2.3 x 3.5 x 12 m |

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

exceeding hundreds of photoelectrons per shot. Such signal strengths call for moderate aperture receivers and strong lasers. The energy in the output laser beam was sufficiently strong, in some instances, to warrant the use of active radar to monitor the airspace above stations. The lower but brighter objects, on the other hand, dictated fast track rates and frequent update of the orbital predictions. Furthermore, with multiple tracking tasks involved, considerable data handling capability was required. Lastly, and most telling, these multipurpose stations have been expensive to build and expensive to operate, predicating their use to cost-shared tracking programs, and preventing great utility in many of the specific areas of interest such as crustal dynamics.

Any laser contribution to a more aggressive attack on earth dynamics requires a system which is optimized solely for providing accurate benchmarks. Newer laser systems must be far more mobile than the current compliment. However, unlike the first generation units, these future systems need only concern themselves with obtaining accurate data on the Lageos satellite, which is the optimum target from which to infer crustal motions. This specialization is a significant advantage from the hardware standpoint for, while Lageos is weaker than most other satellites, it still provides a good signal from a moderately slow, highly predictable orbit. Bender et al. [1978] did propose, therefore, that a highly mobile ranging station be developed specifically for use on the Lageos satellite, taking advantage of low signal techniques, such as used for lunar ranging, and designed to respond to problems of crustal strain monitoring.

## Mobile Station Design

The University of Texas McDonald Observatory has been working for the past nine months in cooperation with NASA and the Joint Institute for Laboratory Astrophysics to develop the required laser station. The basic design goals which have been chosen for this effort are the following:

1. The station will be highly mobile, able to be driven routinely from place to place in a time scale of less than a few days.
2. The station will be air transportable.
3. The station should range the Lageos target with a normal point accuracy better than 2 centimeters for a three-minute average. And,
4. The station must be eye-safe and present no hazard to overflying aircraft or ground-based observers.

This set of constraints, while relatively simple, leaves little flexibility in the choice of station parameters. There is so little flexibility, in fact, that until the availability of mode-locked laser systems, with pulsewidths of less than 1 nanosecond, the entire subset of requirements was impractical. In particular, the requirement for eye safety is very restrictive to the choice of system parameters. It is thought that the eye damage threshold for short pulse lasers at visible wavelengths is approximately $5 \times 10^{-6}$ joules $cm^{-2}$, for a distended pupil [Taboada and Gibbons, 1978]. In other words, the system can transmit no more than 3.5 millijoules of energy on each laser shot provided that the beam is dispersed over a 30 cm transmitting aperture. It is easy to show that an appreciable size receiver is required before an average of even one photoelectron per shot is expected from the Lageos target, even if such an energy is transmitted with a beam divergence of only 10 arc seconds. When this fact is coupled with the desire to maintain a very compact instrument, several conclusions are immediately evident.

1. The transmitting and receiving functions of this station must be contained in a single aperture in order to prevent having to transport two moderately large, high quality telescopes. (The single aperture telescope should be a coude instrument to accommodate the connection to the laser light source.)
2. The transportable laser station will have to operate on the Lageos target with signal levels at or below one photoelectron per shot.
3. The system will have to operate with a relatively narrow beam divergence, requiring an excellent prediction and pointing system in order to acquire the target. And,
4. The system should employ a subnanosecond (mode-locked) laser.

It is also interesting to note that the use of a single aperture system which is restricted to a specific output density, implies a return signal proportional to the fourth power of the aperture. This puts a great premium on the packaging of the telescope, since 40 cm aperture optics will produce almost ten times the return of a 30 cm aperture instrument.

The design parameters chosen for the University of Texas station (TLRS) are shown in Table 2. Owing to the availability of high repetition rate lasers, the system operates well into the single photoelectron regime. Even so, the design is relatively conservative, with a predicted return of some 20 photoelectrons per minute at worst zenith distances using a conventional photomultiplier, and some three times higher with newer detectors which are now on the market. Single shot uncertainties from each detected photon from the Lageos target will be on the order of 5 to 7 cm. These parameters forecast a worst case ranging precision for a one minute normal point of approximately 2 cm, with the best case accuracy limited to about 1 cm by the atmospheric corrections [Gardner, 1977] and the measurement of system biases.

Figure 1 gives two views of the system as it is now being constructed. These drawings differ somewhat from previous presentations; but, with construction in progress on all major components, is likely to remain relatively accurate. The system has been sized such that it will load without disassembly through the side door of a Boeing 747 air freighter, thus permitting movement between continents on commercial airlines. The system is designed to be able to operate without site preparation on any relatively level, firm surface. Mount orientation will be determined by means of an electronic level and an azimuth marker, or by the observation of stellar

sources. We rely on an automatic, internal calibration method not requiring outside targets. Sufficient computing capability is carried with the station such that daily communication with the home facility is not required. Among the other innovations being tested by this prototype is a mode-locked laser which transmits several narrow pulses on each shot, so as to save power and complexity of system. At present, the items pacing the construction of this unit include the delivery of the van, which will hold the major equipment, and the accurate telescope gearing. We expect the unit to be assembled in February of 1979, after which an extensive program of testing will begin.

## Program Expectations

Although the University of Texas transportable station is, in many respects, a prototype, it is expected to be able to play a significant role in crustal dynamics studies. Modelling studies by Bender and Goad [1978] indicate that a site occupation of less than a week should be sufficient to accurately determine any location relative to other Lageos observatories. Even presuming longer stays during an initial campaign and moderate weather statistics, it is obvious that large numbers of sites are possible, provided that the logistics of each move are kept reasonable. Bender and Goad claim that the accuracy of each determination should be on the order of 4-5 cm, presuming reasonable monitoring by other stations and will improve measurably over the next few years as additional studies are completed on the satellite's orbit. Scaler baselines to specific nearby fixed laser stations also observing the same target may be somewhat better determined. It is expected that the sites would be revisited on a bi-annual schedule so as to maintain their status as fundamental benchmarks. If some of these sites are shared with systems sensitive to "inertial" reference frames, such as lunar ranging or VLBI, one could monitor long-term crustal motions in a fixed coordinate system.

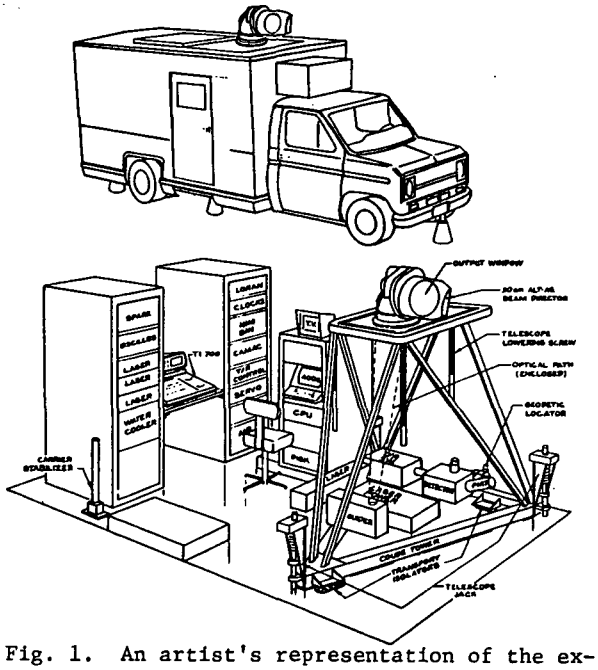


Fig. 1. An artist's representation of the exterior and interior of the highly transportable laser ranging system.

Since the use of this station, even if only partially successful, represents a large increment in the global capability for establishing high accuracy benchmarks, there are a large number of potential programs which could be chosen. It is hoped that conferences such as GEOP can be instrumental in setting priorities in this area. Since this station will enjoy a temporary advantage in mobility, it is the personal opinion of this author that it should concentrate on those relatively uncharted plate boundaries which have reasonably suitable weather statistics. (Poor weather areas, like the Aleutians, would, by necessity, need to be done by radio techniques.) Programs along the poorly known fault zones could not only take advantage of the station's relative transportability, but also its ability to monitor the character of the strain field around each benchmark by using the system as a very powerful, pulsed geodimeter. In the latter application, single, disposable one-inch retroreflectors would be permanently placed on distant mountain tops around each occupied benchmark. These lines would be monitored by the so-called "ratio" method, first conceived by Robertson [1972], and could greatly increase the impact of this station at remote sites. Provided that the ratio method proves accurate to a few parts in $10^7$, as has been indicated by preliminary NGS work around McDonald Observatory [Carter, 1978], and to the extent that this method would be applicable over the much longer lines which would be accessible by this pulsed system, this added facility could prove very useful in "uncharted" areas, especially since the horizontal monitoring could be done at very little additional cost.

Figure 2 shows a hypothetical program which could be conducted with a combination of these capabilities in the complex Central American re-

TABLE 2. Design Parameters For The Highly Transportable Lageos Ranging Station

| | |
|---|---|
| Aperture | 0.30 m (common trans./ receive apert.) |
| Energy per Pulse | 3.5 mj @ 5322 Å |
| Firing Rate | 10 PPS nominal |
| Beam Divergence | 20 arc sec (nominal) |
| Pulse Width | 100 psec |
| Nominal Signal Strength | 0.06 photoelectrons/ shot |
| Single Shot Uncertainty | ± 7 cms (RMS) |
| Size Critical Pkg. | Two single chassis vehicles, 2.4 m (width) x 2.97 (height) x 6.5 (length) |

gion, where the North American, Caribbean and Cocos plates intersect. This example is shown to indicate the scale at which the two methods might be applied and to show how the powerful combination of capabilities could monitor crustal strain in logistically difficult areas. The eight occupied sites are located near paved roads at selected plate locations. The eight would find the overall character of the motions, act as base sites for horizontal ranging into the inaccessible regions, and set scale for the ratioed lines. All of the horizontal lines outside of Guatemala terminate near unimproved roads for the one-time retro deployment, while inside Guatemala the proximity and inaccessibility would warrant a short helicopter campaign. Needless to say, however, the many possible facets of this system warrant a great deal of testing before we dare plan for any ambitious program. Specific recommendations must await a later time when the results of preliminary tests of the station capability are available.

## Future Expectations

The design of the TLRS has been extremely cautious, permitting the staff to bring into operation as few new concepts as possible during the early operations. A number of improvements which might be added to this or subsequent stations with little effort are: 1) two color ranging so as to permit some additional atmospheric information; 2) the addition of advanced detectors which have both higher quantum efficiency as well as lower pulse jitter; 3) the addition of a computer-readable TV guide system so as to permit fully automatic mount modelling; and, 4) the addition of software to permit on-site orbital updates so that the system may, in fact, bootstrap its orbital predictions without requiring frequent updates from larger analysis centers. In addition, we certainly expect much effort at other institutions, particularly Goddard Space Flight Center, to optimize the packaging for such systems, and to improve the ranging accuracy [McGunigal, 1978].

Recently, Wilson and several other authors [1978] attempted to determine the limitations on the compaction of portable laser systems. The major thrust of these discussions was that the higher cross sections of the lower satellite targets, such as Starlette, makes it possible to

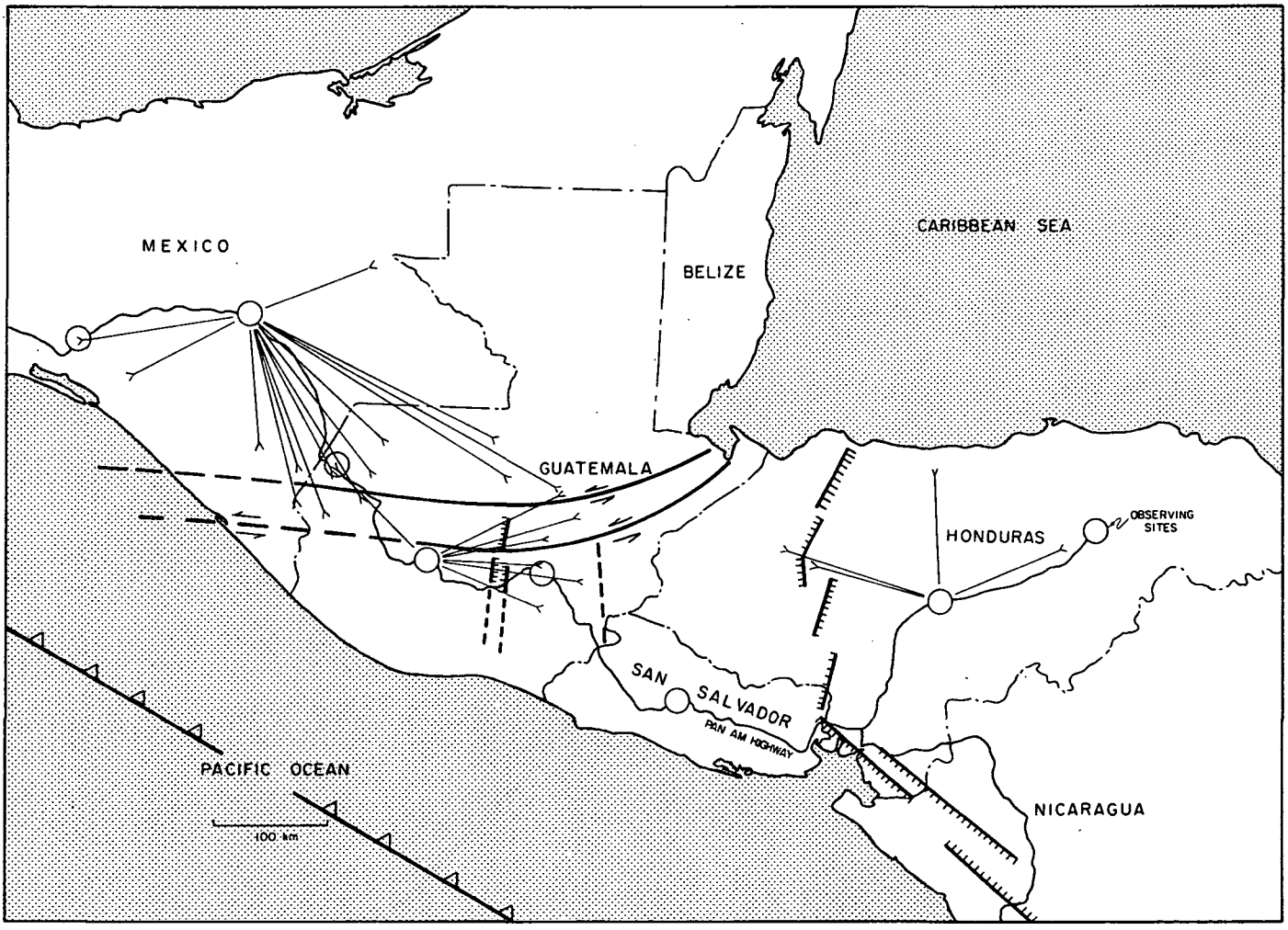



Fig. 2. A possible TLRS campaign for regional strain analysis illustrating the station's ability to use both satellite observing and long baseline terrestrial ranging in a combined program.

package a geodetic laser system small enough to be transported by a single jeeplike vehicle. Since it uses lower satellites the smaller system would only be practical for measuring the shorter (<1000 Km) baselines. However, the system could, of course, also act as a long baseline horizontal geodimeter and might reduce the cost of large operational tracking deployments such as required for the Seasat instrument. In those areas where the horizontal "ratio" method permits the measurement of a large number of unoccupied baselines for little cost, it would be unnecessary to use more highly mobile systems to permit the occupied measurement of these terrestrial lines, particularly if the smaller systems were unable to track the Lageos target which would provide the best overall control. On the other hand, there are many areas where horizontal work does not appear practical -- many which would be more easily studied by a system significantly more mobile than the TLRS. Thus, the success or failure of the concepts to be tested by the TLRS will certainly have a marked impression in the development and possible deployment of more advanced laser ground systems. Most certainly, the degree of success by totally different techniques will also play a major role in these developments.

In summary, it should be clear that the wide application of laser satellite ranging to the field of crustal dynamics is promising, but untried. Systems development in this area is, and will remain for at least a year or two, in a high state of flux. It seems likely that this method can be an important contributor to many problems in this area, but the exact nature which this interaction will take is not clear at this time. The recent technological advances have opened up new avenues which we have only recently started to explore. It is clear, furthermore, that a large number of other possibilities including space borne lasers [Smith, 1978], extra-galactic radio receivers [MacDoran, 1978], and satellite based radio systems [Counselman and Shapiro, 1978] are also available. If we are to properly adjust to the discoveries of the next one to two years with intelligent approaches to the acquisition of meaningful geophysical data, we must, for the moment, retain a flexible, responsive approach, as free from inertia as possible.

Acknowledgements. The TLRS development is supported by NASA Contract NASW-2974.

## References

Bender, P. L., and C. C. Goad, Probable Lageos contributions to a worldwide geodynamics control network, submitted for publication, Aug., 1978.

Bender, P. L., J. E. Faller, J. Levine, S. Moody, M. R. Pearlman, and E. C. Silverberg, Possible high-mobility lageos ranging station, Tectonophysics, in press, 1978.

Carter, W. E., T. Vincenty, Survey of the McDonald Observatory radial line scheme by relative lateration techniques, Technical Report NOS74 NGS9, U.S. Department of Commerce, Washington, D. C., 1978.

Counselman, C. C., and I. I. Shapiro, Compact VLBI terminal for geodetic applications, These proceedings.

Gaignebet, J., The CNES satellite ranging systems, in Proceedings of the Third Workshop on Laser Ranging Instrumentation, edited by K. Hamel and M. Pearlman, Lagonissi, in press, 1978.

Gardner, C. S., Correction of laser tracking data for the effects of horizontal refractivity gradients, Applied Optics, 16(9), 2427, 1977.

Hamel, K., Consideration for a transportable system, in Proceedings of the Third Workshop on Laser Ranging Instrumentation, edited by K. Hamel and M. Pearlman, Lagonissi, in press, 1978.

Masevitsch, A. G., and K. Hamel, Interkosmos laser radar network, in Proceedings of the Second Workshop on Laser Tracking Instrumentation, edited by G. Weiffenback and K. Hamel, Faculty of Nuclear Science and Physical Engineering, Technical University of Prague, Prague, 1975.

MacDoran, P. F., Mobile long baseline interferometry, These proceedings.

McGunigal, T. E., W. J. Carrion, L. O. Caudill, C. R. Grant, T. S. Johnson, D. A. Premo, P. L. Spadin, and G. C. Winston, Satellite laser ranging work at the Goddard Space Flight Center, WESCON Technical Papers, 19, Section 9/2, pp. 1-8, 1975, (Preprint X-723-75-172, Goddard Space Flight Center).

McGunigal, T. E., Laser ranging work at the Goddard Space Flight Center - An update, in Proceedings of the Third Workshop on Laser Ranging Instrumentation, edited by K. Hamel and M. Pearlman, Lagonissi, in press, 1978, (Preprint X-723-78-22, Goddard Space Flight Center).

Pearlman, M. R., C. G. Lehr, N. W. Lanham, J. Wohn, The Smithsonian satellite ranging laser system, in Proceedings of the Second Workshop on Laser Tracking Instrumentation, edited by G. Wieffenbach and K. Hamel, Faculty of Nuclear Science and Physical Engineering, Technical University of Prague, Prague, 1975.

Pearlman, M. R., J. M. Thorp, D. A. Arnold, and F. O. Vonbun, Lageos orbital acquisition and initial assessment, Preprint No. 563, Center for Astrophysics, Smithsonian Astrophysical Observatory, Cambridge, 1976.

Pearlman, M. R., N. W. Lanham, J. Wohn, J. M. Thorp, E. Imbier, F. D. Young, J. Latimer, and I. G. Campbell, A report on the Smithsonian Astrophysical Observatory laser ranging systems, in Proceedings of the Third Workshop on Laser Ranging Instrumentation, edited by K. Hamel and M. Pearlman, Lagonissi, in press, 1978, (Preprint No. 983, Center for Astrophysics, Smithsonian Astrophysical Observatory).

Robertson, K. D., The use of line pairs in trilateration and traverse, Survey Review, 21 (165), 290, 1972.

Smith, D. E., Spaceborne ranging system, These proceedings.

Smith, D. E., R. Kolenkiewicz, P. J. Dunn, and M. Torrence, The measurement of fault motion by satellite laser ranging, Tectonophysics, in press, 1978.

Taboada, J., and W. D. Gibbons, Retinal tissue damage induced by single ultrashot 1060 mm laser light pulses, Applied Optics, 17(18), A217, 1978.

Wilson, P., K. Nottarp, and H. Seiger, The short pulse laser-ranging system installed in Wettzell, in Proceedings of the Third Workshop on Laser Ranging Instrumentation, edited by K. Hamel and M. Pearlman, Lagonissi, in press, 1978.

Wilson, P., E. Silverberg, R. Schutz, I. Malevich, and S. Ramsden, A proposal for the design and application of a high-mobility, low-cost satellite laser ranging system, in Proceedings of the European Workshop on Space Oceanography, Navigation and Geodynamics, edited by S. Hieber and T. D. Guyenne, European Space Agency, SP-137, pp. 111-117, 1978.

N79 21468

# Mobile Radio Interferometric Geodetic Systems

P. F. MacDoran, A. E. Niell, K. M. Ong, G. M. Resch, D.D. Morabito, E. S. Claflin, T. G. Lockhart
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, California 91103

SUMMARY

The mobile VLBI geodetic system called ARIES for Astronomical Radio Interferometric Earth Surveying is currently operating in a proof of concept mode that is adapted to also acquire geophysically significant data. These geophysical data are acquired by careful choice of the transportable antenna's location using the advice and counsel of the Caltech Seismological Laboratory. Figure I shows the fixed base stations at Goldstone and Owens Valley and portable antenna sites in southern California.

Initially, the ARIES 9m stations began S-band (2.3 GHz) experiments on a 307 m baseline near the Goldstone 64m MARS station (DSS 14) and demonstrated a 3 cm accuracy relative to a direct line of sight geodetic survey (Ong, et al 1976). That first experiment sequence demonstrated absolute three-dimensional accuracy (within 3 cm) in a limited case; that is, if the technique could not achieve the absolutely correct vector on 307 m, it certainly would fail on longer baselines where the initial goal was 10 cm. Such success was, of course, only a necessary condition since over such a short baseline the factors of troposphere and ionosphere were self-cancelling and the system was insensitive to radio source positional uncertainties, universal time and polar motion effects.

The next step was to operate on a longer baseline so the parameters that were previously self-cancelling or of low sensitivity could now be sensed, and the ability to calibrate for their effects could be demonstrated.

In August 1974, the ARIES station was moved to JPL in Pasadena, a location logistically favorable, accessible to conventional geodetic systems to check on claimed results and close to the San Gabriel fault.

The measurement accuracy for the 180 km JPL/Goldstone (DSS 14) baseline was estimated in June 1975 at approximately 8 cm. ARIES then made its measurements available to the National Geodetic Survey in advance of NGS geodetic connection of JPL to the TCT, Trans Continental Transverse, the most accurate large scale horizontal control network in the U.S. The results of the comparison of the ARIES derived baseline length with that of the NGS agreed within 13 cm, in good agreement with the 8 cm accuracy estimates of both techniques.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

While the ARIES accuracy results on 180 km were gratifying, the conventional geodesy accuracy estimates of 8 cm may have been overstated and a more conclusive test situation was desired. What was needed was as long a distance as could be found (hopefully also of geophysical interest) which a laser distance measuring device could make in a single direct line of sight. Such a pair of locations was found, Malibu (MAL), on the Santa Monica/Malibu fault and Palos Verdes (PV) Peninsula near the Inglewood/Newport fault. Both the MAL and PV sites are abandoned NIKE air defense bases with roads and their perimeter fences still intact. By observing on approximately 380 km baselines from ARIES 9m station to the 40m telescope of the Caltech Owens Valley Radio Observatory (OVRO), the three dimensional position of PV and MAL was derived relative to the OVRO telescope. By differencing the components of the positions of the two sites, the vector was obtained from Malibu to Palos Verdes, a 42 km path across Santa Monica Bay in southern California, with an estimated accuracy of better than 10 cm. In May 1977, the Santa Monica Bay experiment results were again provided in the blind to the National Geodetic Survey which then directly measured both the intersite distance and azimuth by conventional first-order horizontal geodetic control methods. The two techniques differ by $6 \pm 10$ cm in baseline length and $0.5 \pm 1.2$ arc sec in azimuth (corresponding to $10 \pm 20$ cm). The details of the Santa Monica Bay experiments are contained in (Niell, et al, 1979).

The ARIES Santa Monica Bay accuracy demonstrations were a prerequisite to the Sea Slope Experiment of the NGS to study the apparent differences between oceanographic and geodetic leveling determinations of the sea surface along the Pacific Coast. With joint NGS and NASA OSTA support, ARIES developed the relative geometric positions of tide gages at La Jolla and San Francisco, California where a 65 cm sea slope to the south was indicated by geodetic leveling. ARIES acquired data at San Francisco (SF) in June 1977 and La Jolla (LJ) in July 1977 spanning a total of 42 days. Using the 40m telescope at OVRO as the primary base station, 6 cm or better, three dimensional accuracies were obtained over these 380 to 500 km baselines. A secondary base station, the 26m Venus antenna at Goldstone, also participated. Based upon a triangle baseline vector closure criterion, individual baseline determinations indicated 6 cm accuracy with a one part in $10^7$ closure. Simultaneous S-band (2.3 GHz) and X-band

(8.4 GHz) observations at SF to OVRO processed as separate experiments yielded identical baseline vector results within the 2 to 5 cm uncertainty estimates imposed by signal to noise limits. The San Francisco/La Jolla experiments also utilized dual channel water vapor radiometers at all stations and played an essential role in calibrations of local vertical relative positioning. Hydrogen maser frequency references were made available by the Goddard Space Flight Center for use at all stations. The final combination of ARIES geometric tide gage positional data with the geoid between La Jolla and San Francisco remains an analysis task for the NGS in order to directly compare with spirit leveling observations, to resolve the sea slope controversy.

Five measurement sessions have now been conducted between the JPL Pasadena site and Goldstone since August 1974 for a total of eight measurements to Goldstone and five to OVRO. The precision in the east-west coordinate appears to be approximately 3 cm when measurements are taken closely in time as in the Sea Slope experiment previously discussed. For the JPL/Goldstone baseline, a 12-15 cm westward movement of JPL site relative to Goldstone has occurred with most of the displacement taking place in the last half of 1975. For 1976 to 1978 relative stability is indicated using both OVRO and Goldstone base stations. For all baselines, no vertical displacements appear significant in the presence of the 6 to 10 cm system noise imposed mainly because water vapor radiometer tropospheric calibrations are not yet consistently available.

A 4m high mobility ARIES station has now begun initial field demonstrations. Early experiment results indicate an inadequate signal to noise margin which limited the quasar catalog resulting in a biased solution. Telecommunications upgrades to the 4m antenna are dual Mark II recording (8 Mb/sec) and a traveling wave maser first stage receiver amplifier. With these improvements, the 4m station will probably out perform the 9m station while providing a two site per week site measurement yield compared to the one site per month yield of the 9m system. At present the 4m antenna is undergoing side-by-side tests with the 9m antenna at the JPL site. Following successful side-by-side tests, the 4m station will demonstrate its high mobility and accuracy by repeating the Malibu/Palos Verdes experiment but this time to acquire the data at both sites in one week.

The satellites of the NAVSTAR Global Positioning System (GPS) offer an important new geodetic resource making possible a highly accurate portable radio geodetic system. A concept called SERIES (Satellite Emission Radio Interferometric Earth Surveying) makes use of GPS radio transmissions without any satellite modifications. By employing the technique of VLBI and its calibration methods, 0.5 to 4 cm three dimensional baseline accuracy can be achieved over distances of 3 to 300 km respectively, with only 2 hours of on-site data acquisition. The use of quasar referenced ARIES Mobile VLBI to establish a sparse fundamental control grid will provide a basis for making SERIES GPS measurements traceable to the time-invariant quasar directions. Using four SERIES stations deployed at previously established ARIES sites, will allow the GPS satellite apparent positions to be determined. These apparent positions then serve as calibrations for other SERIES stations at unknown locations to determine their positions in a manner traceable to the quasars. Because this proposed radio interferometric configuration accomplishes its signal detection by cross-correlation, there is no dependence upon knowledge of the GPS transmitted waveform which might be encrypted. Since GPS radio signal strengths are $10^5$ stronger than quasar signals, a great reduction in telecommunications sophistication is possible which will result in an order of magnitude less cost for a SERIES GPS station-compared to a quasar based mobile VLBI system. The virtually all-weather capability of SERIES offers cost-effective geodetic monitoring at less than $1,000 per site.

Details of the SERIES-GPS system will be published in the Bulletin Géodésique in early 1979.

REFERENCES


1. Ong, K. M., P. F. MacDoran, J. B. Thomas, H. F. Fliegel, L. J. Skjerve, D. J. Spitzmesser, P. D. Batelaan, S. R. Paine, M. G. Newsted, A Demonstration of a Transportable Radio Interferometric Surveying System with 3-cm Accuracy on a 307 m Baseline, Journal of Geophysical Research, Vol. 81, No. 20, July 10, 1976.
2. Niell, A. E., K. M. Ong, P. F. MacDoran, G. M. Resch, D. W. Fite, L. J. Skjerve, D. J. Spitzmesser, D. D. Morabito, L. Tanida, E. S. Claflin, B. B. Johnson, M. G. Newsted, A. Banisch and J. F. Dracup, Comparison of a Radio Interferometric Differential Baseline Measurement with Conventional Geodesy, Tectonophysics, 52 (1979), 532.


---


This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration, Office of Space and Terrestrial Applications.

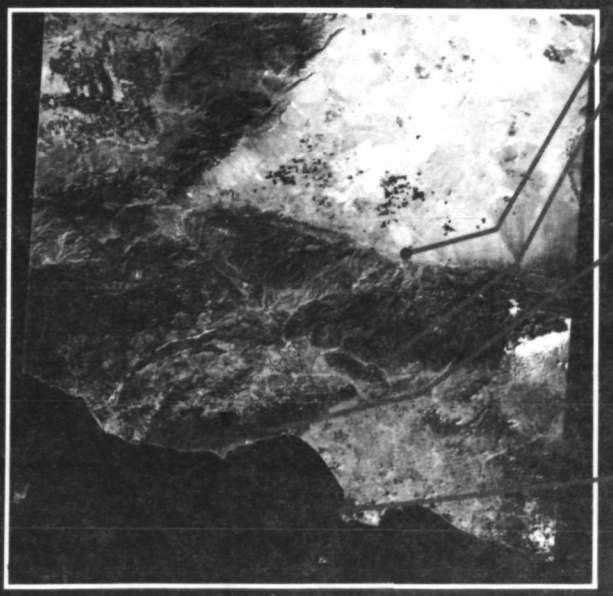
# PROJECT ARIES NETWORK


OWENS VALLEY

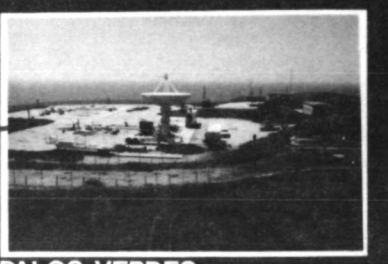
GOLDSTONE


PEARBLOSSOM

JPL

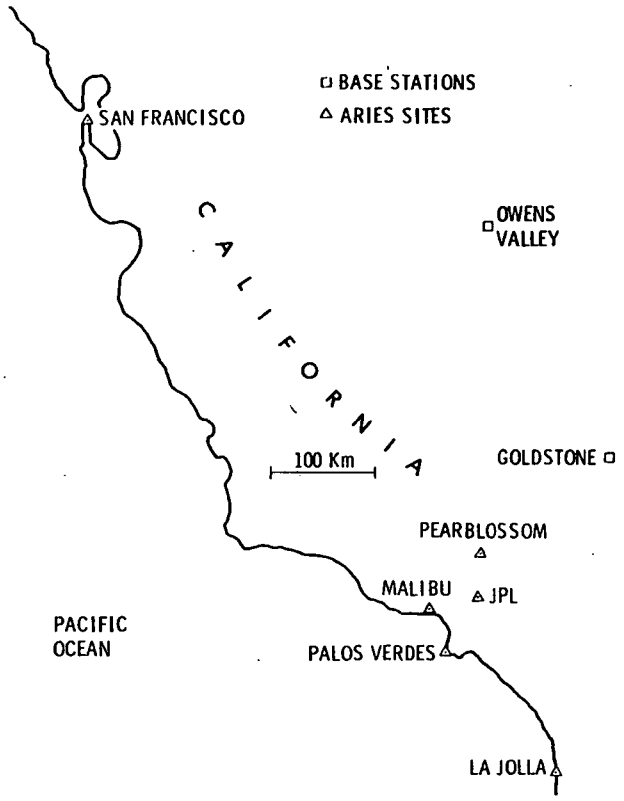


MALIBU



PALOS VERDES

## ARIES NETWORK




## ARIES/CONVENTIONAL GEODESY INTERCOMPARISON

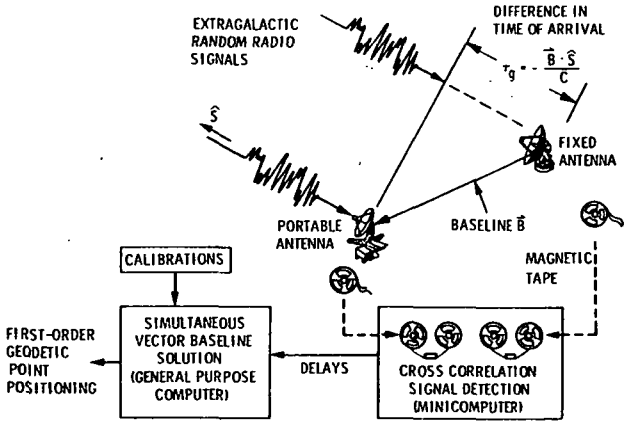
ARIES $|(\overrightarrow{OVRO\text{-}PV}) - (\overrightarrow{OVRO\text{-}MAL})| = 41\,573.902 \pm 0.086$ m

NGS* $(PV\text{-}MAL) = 41\,573.844 \pm 0.06$ m

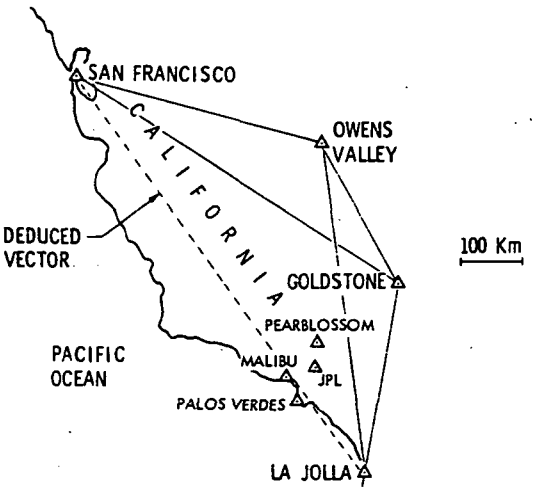
ARIES - NGS = $0.058 \pm 0.105$ m AND AGREEMENT WITHIN 0.5 ARC SEC RELATIVE TO GEODETIC NETWORK LOCATIONS IMPLYING ~ 6 CM THREE DIMENSIONAL ACCURACY IS BEING ACHIEVED BY ARIES ON ~ 400 KM BASELINES.

*RESULTS AS OF 9/16/77 FROM J. DRACUP, CHIEF
HORIZONTAL CONTROL BRANCH
NATIONAL GEODETIC SURVEY

## ARIES GEODETIC SYSTEM




## ARIES SEA SLOPE I NETWORK



## ARIES SEA SLOPE I RESULTS

OVRO, PRIME BASE STATION

| | X, m | Y, m | Z, m |
|---|---|---|---|
| OVRO-SF | 298 106.79 ± 0.03 | -220 735.63 ± 0.04 | -49 774.66 ± 0.05 |
| LJ-OVRO | -45 849.77 ± 0.03 | -289 087.63 ± 0.04 | -397 130.42 ± 0.06 |
| SF-LJ | -252 257.02 ± 0.05 | 509 823.26 ± 0.06 | 446 905.08 ± 0.08 |

LENGTH SF-LJ = 723 379.23 ± 0.03 m

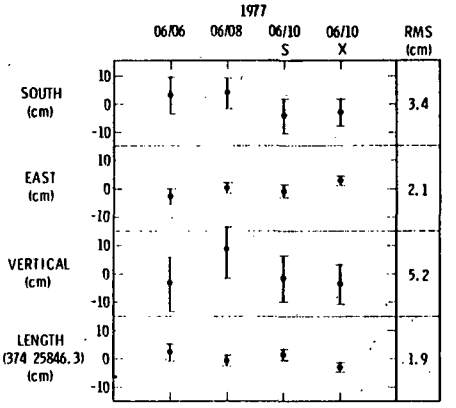
GOLDSTONE (DSS 13) SECONDARY BASE STATION

| | | | |
|---|---|---|---|
| SF-LJ | -252 257.12 ± 0.04 | 509 823.07 ± 0.06 | 446 905.17 ± 0.09 |

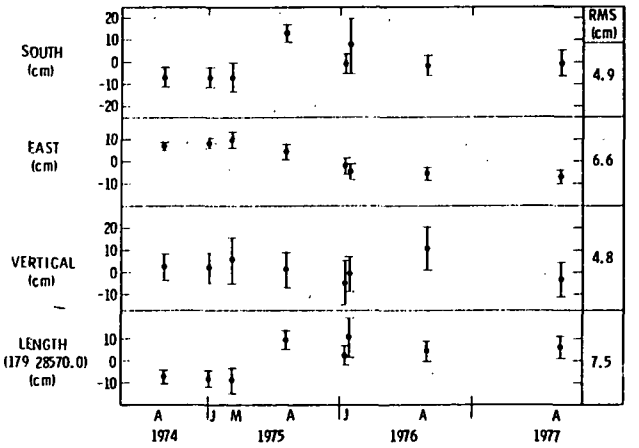
LENGTH = 723 379.19 ± 0.05m

$(\overrightarrow{SF\text{-}LJ})_{ov} - (\overrightarrow{SF\text{-}LJ})_{13} = \begin{pmatrix} 10\text{cm} \\ 19 \\ -9 \end{pmatrix}_{\text{geocentric}} \Rightarrow \begin{pmatrix} -6\text{ cm} \\ -1 \\ -22 \end{pmatrix}_{\text{local}}$

### OVRO/ARIES (SAN FRANCISCO) BASELINE VECTOR
RELATIVE LOCAL COORDINATES AT S.F.; OVRO ASSUMED FIXED




### GOLDSTONE (DSS-14)/ARIES (JPL) BASELINE VECTOR
RELATIVE LOCAL COORDINATES AT JPL; DSS-14 ASSUMED FIXED



Page Intentionally Left Blank

# Application of Global Positioning System to Determination of Tectonic Plate Movements and Crustal Deformations

R. J. Anderle
Naval Surface Weapons Center
Dahlgren, Virginia 22448

The NAVSTAR Global Positioning System is intended to provide observers with instantaneous measurements of position and velocity in three dimensions to 10 m accuracy in 1985. In order to obtain position in three dimensions, "pseudo-range" measurements are computed to four satellites on the basis of the difference in the time of transmission and the time of receipt of signals transmitted by the satellite. Since the time of receipt of the signals is biased due to the error in the observer's clock, the computed ranges are called "pseudo-ranges" and the measurements are made to four satellites so that the error in the observer's clock as well as the three components of position can be computed. In order to test the concept in 1979, six satellites are being launched into orbits which will allow an observer in the Southeastern part of the United States to receive simultaneous data from four satellites for several hours each day. Both test and operational satellites will be in circular twelve hour orbits at an inclination of 63 degrees. The operational system will have eight satellites in each of three orbit planes to provide world-wide continuous coverage.

A number of proposals have been made to utilize the GPS satellites to obtain positions to a few centimeters accuracy for use in crustal motion research. In each instance, orbit and satellite clock errors are minimized by computing the relative position of ground sites. Because of the high altitude of the satellites, these errors will not be significant even for stations separated by hundreds and perhaps thousands of kilometers, depending on whether predicted, fitted, or specially computed ephemerides are used (Anderle, in press). MacDoran (in press) proposes to make measurements on the spread spectrum signal broadcast by the satellite. Shapiro (private communication) would make measurements on an additional signal source he proposes to place aboard the satellites. Measurements could also be made on side-tone signals if these were provided by the satellite. In this report I will address use of the signals transmitted by the satellite and recorded by prototype equipment which is now being tested. Each of these approaches has certain advantages and disadvantages:

| Proponent | Signal | Advantages | Disadvantages |
|---|---|---|---|
| Mac Doran | Spread Spectrum | Precision, Code not required | Untested, Antenna size |
| Shapiro | Unknown | Precision, low cost | Untested, special satellite hardware |
| Easton | Sidetone | Low cost, navigation option | Special satellite hardware |
| Anderle | Recovered code | Equipment developed | Precision or antenna size |
| Anderle | Reconstructed carrier | Equipment developed, precision | Oscillator error |

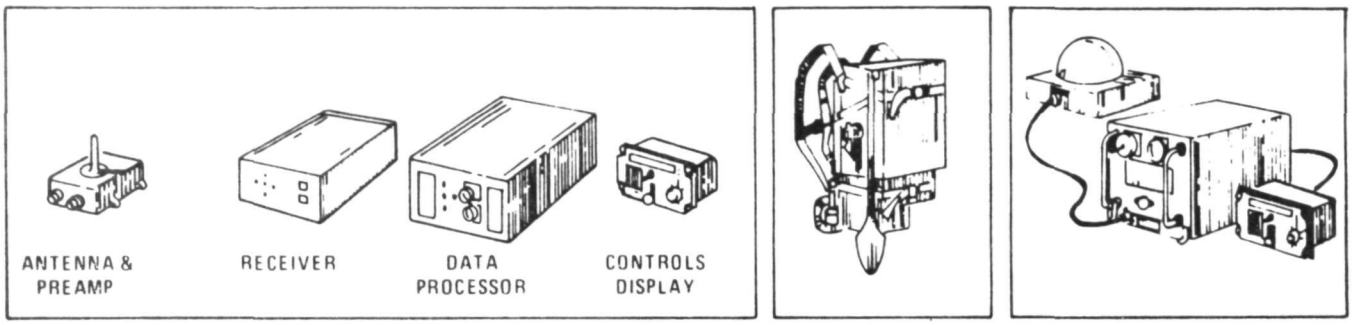
The satellite transmits on two L band frequencies centered at 1575 ($L_1$) and 1227 ($l_2$) Mhz in order to provide a capability for correction for first order ionospheric effects. The carriers are fully suppressed by a bi-phase pseudo-random modulation at a 10 megabit rate. This "P code" is supplemented by a lower bit rate, shorter code sequence, "C/A code" on the $L_1$ signal which permits rapid acquisition of the signal. The code is also generated by the receiver, where it is shifted in time to obtain maximum correlation with the received signal. The time of maximum correlation is recorded as the time of receipt of the signal. Subtraction from broadcast time of transmission of the signal gives the apparant travel time to the signal, and multiplication of the difference by the velocity of light yields the pseudo range to the satellite. Once the correlation of the codes is made, the carrier signal is reconstructed and Doppler measurements are made.

Eight receiver channels would be required to maintain continuous lock on the $L_1$ signal from four satellites as well as the ($L_1$-$L_2$) signals. To reduce receiver size, power, weight and cost, fewer channels are used. Since the initial search for the signal correlation can take tens of seconds, five channels is the minimum number used when high data rates are desired; four of the channels maintain continuous lock on the $L_1$ signals from four satellites while the fifth channel rapidly sequences through the ($L_1$-$L_2$) signals. Simpler receivers have two channels which lock on the $L_1$ and $L_2$ signals from a sateliite, then switch to the next satellite, or a single channel which sequences through the $L_1$ then $L_2$ signals on each satellite in turn. The lowest cost receiver uses a single channel which sequences through only the $L_1$ C/A signal on each satellite. The primary output of these receivers, depicted in figure 1, is the range obtained by correlation of the pseudo random noise signal. Doppler is counted to obtain velocity and reduce the noise of the PRN measurements. The NAVSTAR Geodetic Receiver (NGR) built by the Naval Surface Weapons Center and Stanford Telecommunications Incorporated (figure 2) has two channels and records range at $L_1$ and Doppler at $L_1$ and $L_2$, sequencing through satellites according to computer commands.

The measurement precision depends on antenna gain and the receiver hardware. Examples of pre-

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

# PHASE I USER EQUIPMENT




| | X-SET HDWE<br>• MAGNAVOX<br>• TEXAS INST | HI AJ<br>• COLLINS | Y SET<br>• MAGNAVOX | MANPACK<br>• MAGNAVOX<br>• TEXAS INST | Z-SET<br>• MAGNAVOX |
|---|---|---|---|---|---|
| • | 4 CHANNEL CONTINUOUS TRACKING | 4 CHANNEL CONTINUOUS TRK IMU AIDED | 1 CHANNEL SEQUENTIAL TRACKING | 1 OR 2 CHANNEL SEQUENTIAL TRACKING | 1 CHANNEL SEQUENTIAL TRACKING |
| • | $L_1$ $L_2$ P OR C/A | $L_1$ $L_2$ P OR C A | $L_1$ $L_2$ P OR C A | $L_1$ $L_2$ P OR C A | $L_1$ C/A ONLY |
| • | HI PERFORMANCE | HI AJ | MED PERFORMANCE LOWER COST | LOWEST SIZE AND WEIGHT | LOWEST COST |
| • | TIME TO FIRST FIX 180 SEC<br>J/S (PRECISION TRK) 43 dB | 150 SEC<br>70 dB (?) | 300 SEC<br>40 dB | 240 SEC<br>40 dB | 300 SEC<br>10 dB |
| • | SIZE 9700 IN $^3$<br>WEIGHT 235 LB<br>POWER 860 W | 13500 IN<br>335 LB<br>1000 W | 9700 IN $^3$<br>220 LB<br>800 W | 1000 IN $^3$<br>30 LB<br>30 W | 1100 IN $^3$<br>28 LB<br>75 W |

FIGURE 1

cision for two antenna gains are given below based on the equations given in figure 3 for monitor stations (similar to X sets) used to obtain data for orbit computations; the precision is also given for the NGR.

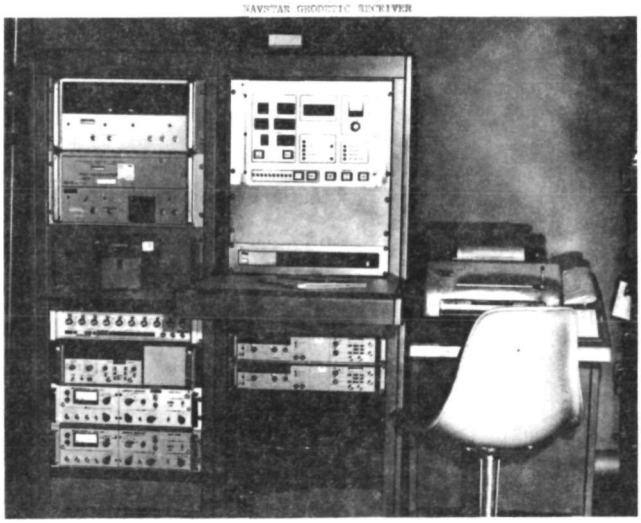



Table 1

Sample Measurement Precisions

| Equipment | Monitor Station (6 sec. average) | NGR |
|---|---|---|
| Range Precision | | |
| 0 db antenna | 36 cm | 65 cm |
| 21 db antenna | 0.4 cm | 5 cm |
| Doppler Precision | | |
| 0 db antenna | 0.14 cm | 0.17 cm |
| 21 db antenna | 0.08 cm | 0.01 cm |

The above figures are based upon a receiver power of 163 dbw, which is 3 dbw poorer than that observed for the first two satellites launched. The zero db antenna is a small stub which does not have to be aimed. A 21 db antenna is equivalent to a one meter dish which must track each satellite, although the required aiming accuracy is only about 10 degrees. If the measurement precision at each of the L-band frequencies is

the same, the precision of the range corrected for ionospheric refraction will be a factor of four worse than that fundamental precision after application of the correction algorithm. However, by tracking the ($L_1$-$L_2$) the refraction correction can be obtained without significant degradation of the basic precision, as shown in figure 3. The precision inferred from monitor station data which has geen aggregated to 15 minute intervals is currently 20 cm. This is higher than expected for aggregated data, according to table 1; however, the value of 20 cm was obtained as the residuals of fit to five to ten hours of data. Figure 5 shows that oscillator variations integrate to range errors of 50 to 100 cm after 5 to 10 hours, which easily accounts for the result obtained. Hermann (private communication) analyzed 15 minute segments of data obtained on two satellites during an initial demonstration of the NGR. He found the noise level of the range measurements matched the expected 65 cm for one satellite. The noise of frequency measurements on each satellite at $L_1$ matched that expected for the frequency counter used in the demonstration, although the equipment was only capable of validating the precision of the phase measurements to the 20 cm level. Frequency measurements at $L_2$ were considerably worse than expected, probably due to interferring signals encountered in the area.

The difference in positions of two stations based on pseudo-range measurements to four satellites is equivalent in information content to an inter-ferometric determination of relative station positions. In such solutions, the precision of the position determination is a factor of two to three worse than the precision of the range measurements due to the Geometric Dilution of Precision (GDOP) arising from the simultaneous solution for position and time bias. Thus the expected precision for solutions for relative station positions based on X-set data over a six hour time span would be $\sqrt{2}$x3x35/$\sqrt{3600}$ = 2.5 cm, where $\sqrt{2}$ arises from differencing data from two stations, 3 is a GDOP, 36 cm is the precision of six second data, and 3600 is the number of observations in a six hour interval. About the same

Precision of Monitor Station Data*

$$\text{Range Variance, } \sigma_r^2 = \frac{B_c}{S/N_o}\left(\frac{c}{2\pi f}\right)\left[\frac{1}{2T_1}+\frac{M}{16T_v}+\frac{M}{32T_vT_1}\sum_{i=1}^{16T_v/M}\left(\frac{4T_1-1}{4T_1}\right)^{i-1}\right]+\frac{q^2}{12}$$

$$\text{Variance in Refraction Data, } \sigma^2_{M_1-M_2} = \frac{25}{T_{12}\,S/N_o}\left(\frac{c}{2\pi f}\right)^2+\frac{q^2}{12}$$

$$\text{Range Difference Variation, } \sigma^2_r = \frac{2B_L}{S/N_o}\left(\frac{c}{2\pi f_c}\right)^2+\frac{q_c^2}{12}+\left(\frac{\Delta t}{T_I}\right)^2$$

where

$B_c$ = code channel bandwidth ~ 25 hz

$S/N_o$ = signal to noise ratio in hz ~ $10^{db/10}$hz

c = velocity of light
f = code frequency ~10.23 mhz
$T_1$= $L_1$ code loop time constant ~ 10 sec

M = number of operating channels
$T_v$ = nominal sampling interval ~ 10 sec

q = c/64 f = 1/64th of a code chip
$T_{12}$ = min ($T_{12}$, $T_{Ts/3}$), where $T_2$ = $L_1$-$L_2$ code loop constant ~160 sec, $T_{Ts}$ = time slot of MS tracking sequence ~ 60 sec

$B_L$ = costas loop bandwidth for Doppler ~ 3.33 hz

$f_c$ = $L_1$ carrier frequency ~1575.42 mhz

$q_c$ = c/64 $f_c$ = 1/64th of a carrier cycle

t = integrated doppler measurement interval ~ 6 sec
$T_I$ = time interval per unit change in range due to ionosphere ~197 sec/m

*interface control document 12436, MCS/NSWC

FIGURE 3

level of accuracy was obtained by Goad (private communication) in a computer simulation which considered the geometric and other factors in the solution in a more realistic fashion.

Because of the higher precision of phase measurements and the successful use of such measurements of Navy Navigation Satellites in geodetic solutions, the Naval Surface Weapons Center is exploring the use of such data from GPS satellites for positioning. In this approach phase measurements and integral Doppler counts are made over successive time intervals during the passage of a satellite between the start and end of the interval can be computed. For the lower altitude, polar, Navy satellites, the range difference observations obtained during one pass of the satellite over the station provide a determination of station latitude and the range to the satellite at the time of closest approach of the satellite to the station. Data from two passes at different longitudes are required to resolve the two slant range measurements into station longitude and height to acceptable precision. For the higher altitude GPS satellites, a three-dimensional solution for station position can be obtained from observations made on a single pass of the satellite over the station, but not to the precision possible using data from two passes. The range differences observed during a pass can be treated in two ways: the range differences over successive time intervals can be assumed to be uncorrelated, or, if the cycle count is not reset at the time of readout, the range differences can be accumulated to provide a series of range measurements subject to a range bias common to the observation set. Simulations by Anderle (in press) indicate that biased range data during a pass yields a precision in the determination of two components of station position which is about equal to the precision of the range difference measurement. Uncorrelated range difference data provides a precision which is a factor of three or more worse than the precision of measurement. Due to the high measurement precision possible, this would still be a highly useful data type which would permit observations of two or more satellites sequentially with a two channel receiver. However, range difference data, treated as either correlated or uncorrelated, set severe requirements on the oscillator used to make the measurements. Figure 4 (Fell, private communication) shows that a Cesium oscillator will produce systematic errors which would reach 20 cm during a fit of mean frequency offset to eight hours of range difference data. The figure is a sample result for one sequence of data. Fell also gives the root mean square of the error for a large set of sequences of such data in figure 5 as a function of time within the pass; the rms error ranges from 10 to 15 cm, with the maximum error in the center and at the ends of the pass. The effect of this systematic error

Precision of NAVSTAR Geodetic Receiver*

$$\text{Range Variance, } \sigma_r^2 = \frac{c^2 B_L(\text{DLL})}{\left(S/N_o\right) f^2} \left[ .905 + \frac{1.612\ B_I}{S/N_o} \right]$$

$$\text{Doppler Variance, } \sigma_r^2 = \left(\frac{c}{2\ f_c}\right)^2 \frac{B_L(\text{PLL})}{S/N_o} \left[ 1 + \frac{B_I}{2\ S/N_o} \right]$$

where

$c$ = velocity of light

$B_L$(DLL) = one-sided loop noise bandwidth = 2 hz

$B_L$(PLL) = one-sided phase loop bandwidth = 16 hz

$S/N_o$ = signal to noise ratio in hz ($10^{db/10}$)

$f$ = code frequency~10.23 mhz

$B_I$ = IF bandwidth = $10^3$ hz

$f_c$ = carrier frequency ~ 1574.42 mhz

* Stanford Telecommunications Incorporated

FIGURE 4

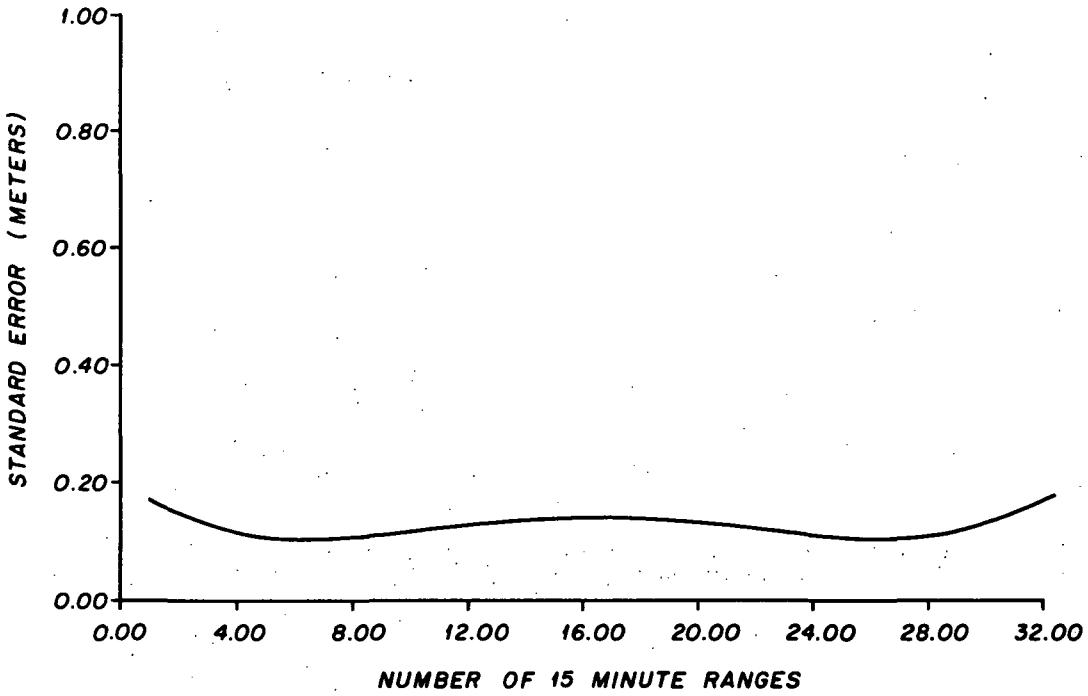



FIGURE 5

on station position is being computed. But it will clearly dominate the effects of the measurement error.

To summarize, pseudo-range measurements to four GPS satellites based on correlation of the pseudo random code transmissions from the satellites can be used to determine the relative position of ground stations which are separated by several hundred kilometers to a precision at the centimeter level. This precision is attained within 12 hours using a small antenna or much more rapidly with a directional antenna. Carrier signal measurements during the course of passage of two satellites over a pair of stations would also yield centimeter precision in the relative position, but oscillator instabilities would limit the accuracy by an as yet undetermined amount. Measurement precisions of code and carrier signals have generally been consistent with test conditions, but have not been tested to design levels yet. The accuracy of solutions based on either type of data would be limited by unmodeled tropospheric refraction effects which would reach five centimeters at low elevation angles for widely separated stations.

References

Anderle, R. J., Geodetic Applications of the NAVSTAR Global Positioning System, Proceedings of the Second International Symposium on Problems Related to the Redefinition of North American Networks, April 1978, Washington, D.C. (in press).

Page Intentionally Left Blank

# Spaceborne Ranging System

David E. Smith
Goddard Space Flight Center
Greenbelt, Maryland 20771

Abstract. A spaceborne laser ranging system is described that could survey a large network of ground reflectors, and provide, their relative locations to a precision of ± 1 cm. This performance is believed realizable for networks covering up to $10^6$ square kilometers and from only a few days of observations. This system could be used to monitor crustal movements in many areas of the world and has the potential to provide an almost real-time system for detecting precursory ground motions before large earthquakes.

## Introduction

During the last decade or so improvements in our ability to measure precise positions on the earth's surface by conventional ground based techniques has made us more aware of the existence of crustal motions taking place at the centimeter level. This has become particularly apparent for seismic regions such as California where the existence of sizable geodetic motion, both horizontal and vertical, have been observed (e.g., Castle et al. 1974) and the majority of earth scientists now seem ready to accept that motions at the centimeter level are probably occurring in many regions, not only those considered seismically active or undergoing well-documented subsidence. The extension of conventional surveying at the centimeter level into larger areas presents difficulties from a point of view of accuracy and the frequency with which re-surveys of the areas can be reasonably accomplished. With only limited knowledge of the time scales on which motions along plate boundaries and plate interiors are occurring it seems highly desirable to be able to make these geodetic measurements almost synoptically, or at least on an "as required" basis. Ground deformation preceding large earthquakes is known to occur (Scholz et al., 1973) but little seems to be known about the time scale or the magnitude of the deformation nor, of course, precisely where it will occur. Thus techniques capable of monitoring large areas on a scale that could detect possible precursory motion and able to re-survey the area quickly in order to confirm the initial observation could play a very important role in the detection and understanding of this type of crustal phenomenon; at least in so far as indicating where intensive ground based measurements might be concentrated.

Space techniques involving mobile ground based systems, such as very long baseline interferometry (VLBI) and laser ranging to satellites and the moon can be expected to provide a capability close to the required accuracy in the next several years but it is not clear if the measurement and re-measurement of hundreds, or perhaps thousands, of points by these methods is practical in an operational role. For these reasons it was proposed several years ago (Mueller et al., 1975; Vonbun et al., 1977; Smithsonian Astrophysical Observatory, 1977), that putting a laser system into space and putting simple unmanned reflector systems on the ground might have a distinct advantage when the scale of monitoring ground motions grew to a level of becoming a global problem or the frequency of re-surveys coupled with areal size become too large for the mobile space techniques. Furthermore, it seemed that the spaceborne laser would probably have a greater potential accuracy than these other techniques for distances up to a thousand or more kilometers if the altitude and configuration of the satellite were carefully chosen.

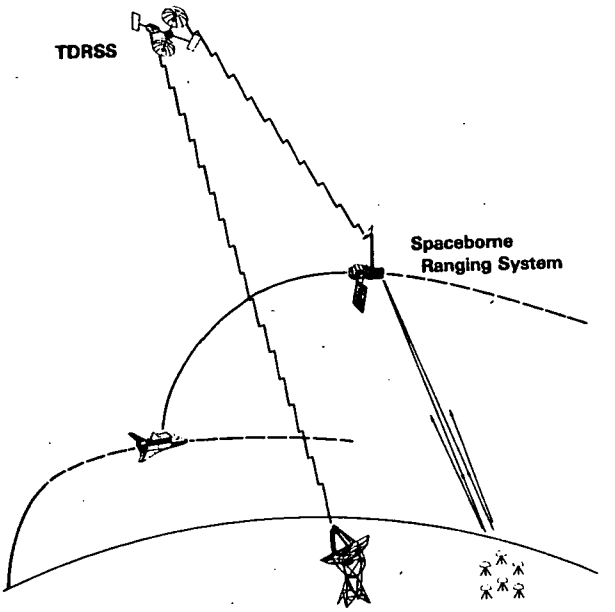



Fig. 1. Spaceborne ranging system: general concept.

In this paper a spaceborne laser ranging system is described and the results of performance simulations presented that show the capability of such a space mission in providing precision geodetic positions on a global, near real-time basis.

## Spacecraft System

The Spaceborne Ranging System consists of an orbiting spacecraft carrying a pulsed laser distance measurement system that sequentially measures the distance to a number of retro-reflector arrays on the ground. Figure 1 shows the general concept of the system. Launched by the Space Shuttle into a low altitude orbit the spacecraft is subsequently lifted into a higher orbit by its own propulsion system. Once in orbit the spacecraft ranges to the corner reflectors on the

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

arth's surface, as it passes overhead. The measurements are stored on the spacecraft and subsequently relayed to a high altitude relay satellite, the Tracking and Data Relay Satellite System, for re-transmission to a ground terminal. The measurement objective of the system is a relative position uncertainty in the locations of the reflectors of ± 1 cm precision or better for separations of reflectors as large as 1200 to 1500 km. The specifications of the spaceborne and ground reflector system are such that they should meet this measurement accuracy for a mission to be launched in the early to mid-1980's.

The proposed laser system (M.W. Fitzmaurice, private communication) would consist of a Nd YAG laser with a 200 picosecond pulse length and a repetition rate of 10 pulses/sec. The rms range uncertainty of a single pulse at 5 to 10 photoelectrons is expected to be 1 to 2 cm with a bias of a few millimeters.

The ground targets (P.O. Minott, private communication) will consist of a small corner cube

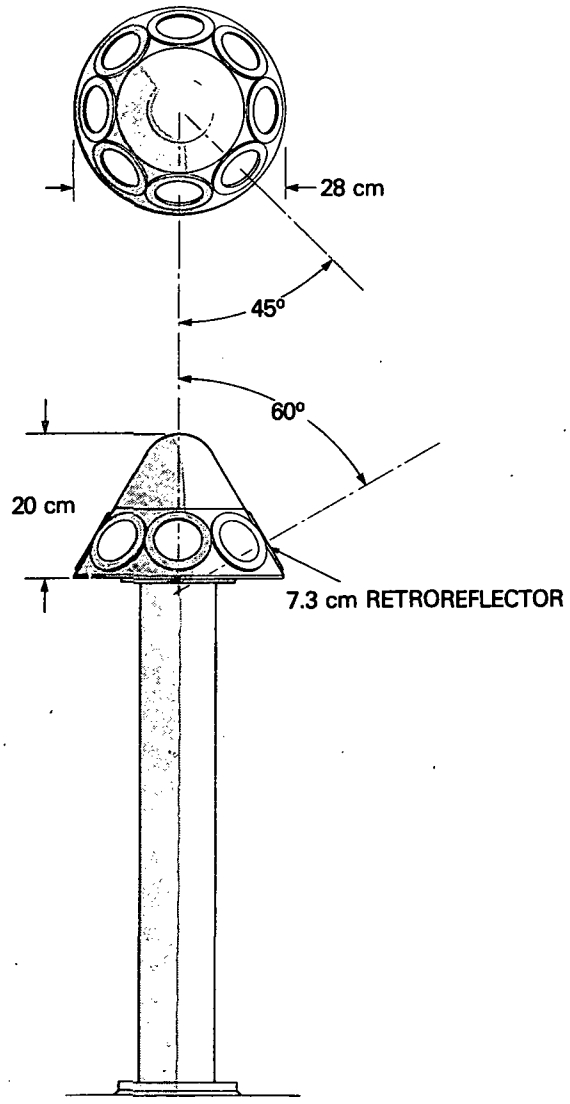



Fig. 2. Ground reflector array.

array of retroreflectors mountable on a permanent pillar or a surveyors tripod. A design for these targets is shown in Figure 2. Eight corner cubes are mounted in a cone approximately 28 cm in diameter at the base and 20 cm high. Each retroreflector consists of a glass prism, 73 mm in diameter and the proposed housing is constructed of aluminum or moulded plastic. Estimated weight of the reflector array is 3 to 5 kg. Details of the laser system and ground reflector system are shown in Table 1.

As the first reflector of a ground network comes into view of the spacecraft an acquisition procedure is initiated that is expected to take 10 to 15 seconds. The procedure consists of a search for the reflector based on a priori knowledge of the reflectors location and the position of the spacecraft. After the successful acquisition of the first reflector the laser makes 20 to 30 range measurements in a 2 to 3 second period and then swings on to the next reflector taking less than 0.5 seconds for this operation even for the most widely separated reflectors. The laser dwells 2 to 3 seconds on the second reflector making range measurements and then moves to the next. No acquisition time is expected to be necessary for the second and subsequent reflectors because the relative location of the spacecraft and the ground network will be updated from the corrected a priori positions obtained during the acquisition of the first reflector.

On any particular pass of the spacecraft over the network, the spaceborne ranging system will range to a given reflector three times, each for a 2 to 3 second period; once at low to medium elevation on the approach, once at high elevation and once at medium to low elevation on the way out. For a thousand kilometer altitude orbit a pass of the satellite over the network will last about 10 minutes which indicates that about 50 reflector arrays could be surveyed on every pass over the region assuming a 2 to 3 second dwell time and a

TABLE 1. System Specifications

Laser

- Nd YAG
- 0.2 nanosecond pulse width
- 10 pulses/sec
- 1-2 cm rms uncertainty (single pulse) at 5-10 photo electrons
- Lifetime - $5 \times 10^6$ pulses
- Beam divergence - 10 arcseconds

Pointing

- 2-3 arcsecond accuracy

Retroreflectors

- Number - 8
- Size - cone, 28 cm diameter base, 20 cm high
- Weight - approx. 3 to 5 kg
- Materials - glass prisms, 73 mm diameter
  housing - moulded plastic or aluminum
- Initial acquisition - 15 seconds
- Stay time - 2-3 seconds
- Transfer time - 0.5 seconds (max) between reflectors

few tenths of a second transfer time between reflectors. The lowest elevation at which measurements are expected to be made is about 20 degrees due to uncertainties in being able to account accurately for atmospheric refraction. Figure 3 shows the sequence of events as the system passes over a network.

Simulated Survey

A simulation has been performed of a survey of the State of California by the Spaceborne Ranging System (W.D. Kahn and T.S. Englar, private communication). In the simulation approximately 150 corner cube arrays are distributed over California at a separation of about 50 km. Using the system described in the previous section, the simulation estimates the accuracy and precision with which the relative positions of the reflector arrays can be obtained in the presence of noise and bias on the laser system, perturbations of the spacecraft motion and errors in the refraction calculations.

The orbit of the satellite is assumed to be circular at 1000 km altitude and 50 degree inclination to the equator. A medium inclination orbit was chosen because it provides ground tracks across California in almost orthogonal directions (southwest to northeast and northwest to southeast) and thus provides a strong geometric distribution of range measurements. In contrast, a polar orbit provides only north to south or south to north tracks and these provide strong geodetic ties in the north-south direction but only weak control in the east-west direction.

The simulation has been conducted for a six-day observation period assuming 50% cloud cover that effectively reduces the number of successfully observed tracks over the area from 18 to 9. The data on these tracks is simulated at an effective rate of 10 pulses/sec. with a noise of 2 cm and a bias of 0.3 cm. The effect of errors in the gravity field on the motion of the satellite were accounted for in the simulation by adopting the GEM 10 covariance model out to degree and order 22. GEM 10 is a model of the gravity field derived from satellite tracking and surface gravity data (Lerch et al., 1977). The effects of solar radiation pressure and air drag on the satellite were assumed to be in error by a constant percentage in the estimation of their effect on the solution. The effect of atmospheric refraction was estimated through a two parameter (pressure and temperature) model developed by Gardner (1977). In this model the temperature and pressure are assumed known at a limited number of locations in the region and are used to develop an atmospheric model of the whole region from which the temperature and pressure at each of the reflectors can be estimated.

Table 2 shows the accuracy and precision of the baselines obtainable between reflectors over various distances and summarizes the error models and assumptions made in the simulation. The accuracy is the estimated total error in the baseline; the precision is the ability to repeat the measurement. Thus the noise in the precision is approximately 1.44 times that of the accuracy while the systematic errors are generally smaller in the precision than the accuracy because of the correlations.

The general trend in Table 2 is for all the error sources to increase as the intersite distance increases. The gravity model error dominates the accuracy while the solar radiation pressure and drag effects on the orbit are negligible in both the accuracy and precision and over all distances. The noise level in the

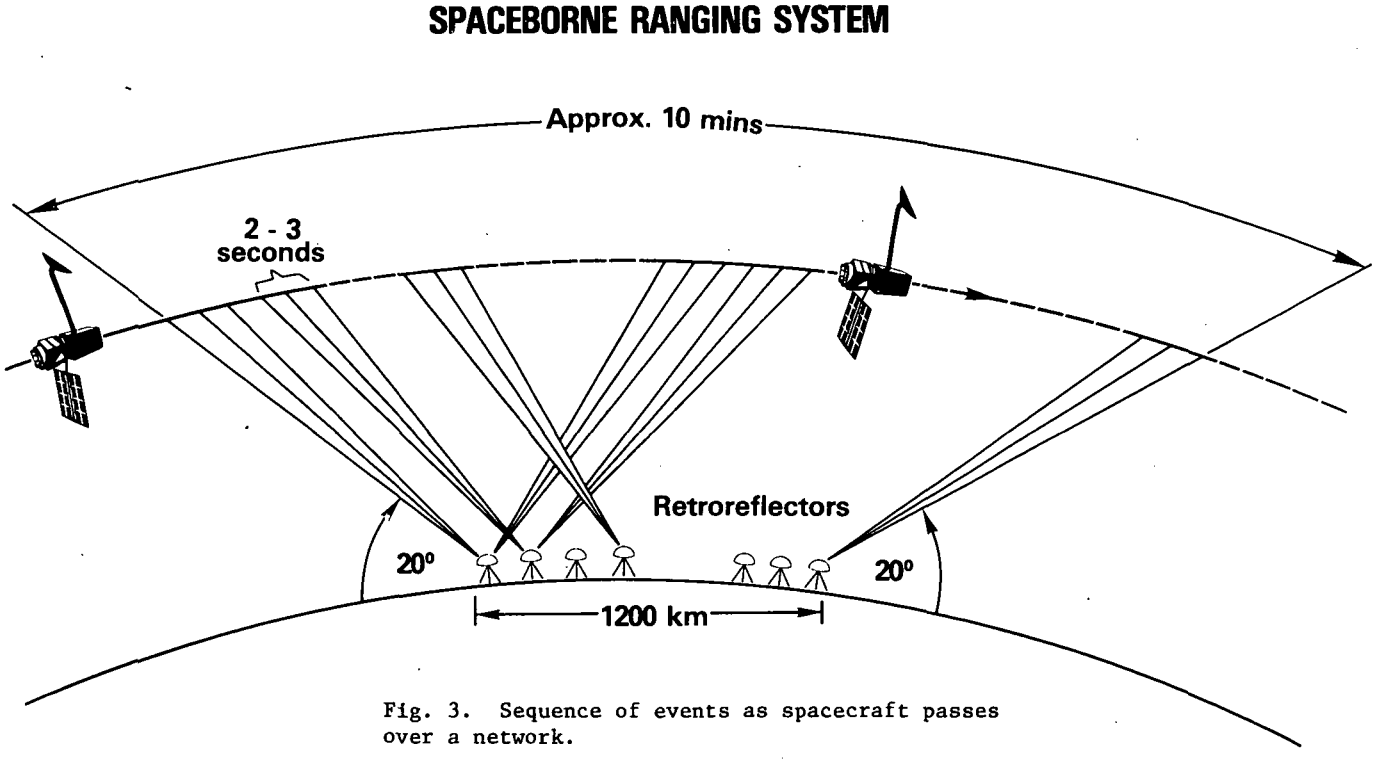



Fig. 3. Sequence of events as spacecraft passes over a network.

precision is larger than in the accuracy but the gravity contribution is significantly reduced in the precision columns showing that a large segment of the gravity error is of a systematic nature that is repeated from one simulation (or observation) to the next. The atmospheric refraction is larger in the precision than the accuracy showing that much of this error has the property of noise. Similarly the laser system bias behaves like noise in the precision columns because the bias on one track is assumed independent of any other track.

The conclusions to be drawn from Table 2 are that a survey of the complete State of California can be accomplished with a precision of 1 cm and an accuracy of 3 cm in about 6 days. Making an allowance for an increase in our knowledge of the gravity field over the next several years suggests an overall accuracy in the mid 1980's of 1 cm as a reasonable objective.

After the gravity field and data noise, the largest source of error is the atmospheric refraction. Alternative ways of making the refraction correction are presently being considered. One possibility is that the laser system operate in two colors separated enough in frequency that the total atmospheric refraction can be derived from the delay between the two pulses. This is technically difficult but might be possible in the next few years. Another possibility is to limit the observations to greater than 30 degrees elevation. This reduces the refraction error by about 40% but decreases the amount of data thereby increasing the noise contribution in the analysis. At the present time, with the noise on the data a larger contribution to the total error than the expected atmospheric refraction it seems that restricting the elevation may not be desirable.

Although no mention has been made so far of the ability to measure relative height changes the Spaceborne Ranging System does have this capability. At this stage it appears that all error sources affect the vertical more than the horizontal (baseline) but that similar accuracies can be achieved in height if the network is constrained at three perimeter reflectors. That is, the system could measure changes of height within the network but not of the whole network itself or of possible rotations of the network. This aspect is continuing to be investigated.

## Applications in Geophysics

Perhaps the single most important application of a spaceborne ranging system is in its potential for monitoring crustal motions in seismic zones. The capability of providing high accuracy geodetic measurements in a very short period of time, i.e., a geodetic snapshot, could have significant impact on our ability to study pre- and post-seismic motions. With this system it would be possible to establish in a region such as California, a network of reflector systems that could be monitored routinely and would provide warning of any motion at the centimeter level. Monitoring of an active

TABLE 2. CALIFORNIA SIMULATION: BASELINES

| | ACCURACY | | | | | | | PRECISION | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DISTANCE (km) | NOISE (cm) | GRAVITY (cm) | ATM. REFR (cm) | MEAS BIAS (cm) | SOL.RAD PRESS (cm) | DRAG (cm) | RSS (cm) | NOISE (cm) | GRAVITY (cm) | ATM. REFR (cm) | MEAS BIAS (cm) | SOL.RAD PRESS (cm) | DRAG (cm) | RSS (cm) |
| 150 | .18 | .54 | .10 | .01 | .00 | .00 | .58 | .25 | .10 | .14 | .01 | .00 | .00 | .30 |
| 300 | .20 | 1.06 | .12 | .01 | .00 | .00 | 1.09 | .28 | .19 | .17 | .02 | .00 | .00 | .38 |
| 600 | .24 | 1.65 | .16 | .02 | .01 | .00 | 1.68 | .34 | .24 | .23 | .04 | .01 | .00 | .48 |
| 900 | .27 | 2.21 | .18 | .03 | .01 | .00 | 2.23 | .38 | .33 | .26 | .05 | .01 | .00 | .57 |
| 1300 | .31 | 3.09 | .25 | .05 | .02 | .00 | 3.11 | .44 | .49 | .35 | .07 | .03 | .00 | .75 |

Assumptions:

- Orbit
  - Mean Altitude: 1000 km
  - Inclination: 50°
- Observation Period: 6 days
- Cloud Cover: 50%
- No. Retroreflectors: 150 @ 50 km spacing
- Measurement Noise: ± 2cm single pulse, 10 pulses/sec
- Measurement Bias: 0.3 cm
- Gravity Uncertainty: Gem 10 covariances ($\ell$,m=22)
- Atmospheric Error Model: 2 parameter
  - Pressure Noise: ± 1.0 mbar
  - Bias: 0.33 mbar
  - Temp. Noise: ± 1.4°C
- Radiation Pressure: 33% error
- Atmospheric Drag: 22% error

region might be accomplished by regular surveying of the area every three months, say, until such time as evidence of motion is detected and then increasing the frequency of the re-surveys down to once per week, and perhaps even more frequently. The major advantage of the spaceborne system is that it can monitor a large enough region that it can identify a sub-area within the survey network of potential activity that can then be monitored more accuractely by ground based methods.

In principle, the separation of ground reflector systems can be as close as a few hundred meters but in reality it would seem that the most advantageous separation is likely to be greater than 20 km where the accuracy of the space system appears to start to become superior. In this respect, it should be remembered that the reflector systems can be located in almost any kind of terrain and not necessarily at the most accessible locations and that line of sight visibility between reflectors is not a requirement, in contrast to ground based surveying.

One of the more interesting possibilities that this system could make possible is the "capture" of a very large earthquake, say magnitude 7.5 and above. By suitably instrumenting with reflectors a number of regions of the world which historically have had large earthquakes and regularly monitoring these regions over many years, we could eventually expect to observe the geodetic development, occurrence, and relaxation after a major shock. Such an experiment is impossible at the present time because of the magnitude of the task but a spaceborne system of the kind described could probably simultaneously monitor about ten areas the size of California distributed around the globe. The limitations to such an extensive activity are primarily with the weather, which might limit the choices of location of some reflectors, and to the lifetime of the laser. If we consider monitoring 10 regions around the world, each containing 100 reflectors and we plan 10 complete surveys of each area per year, then we have $10^4$ site determinations per year, which is probably the limit that can be expected from a single laser unit. Thus, for a 2 or 3 year mission lifetime the spacecraft must contain a multiple laser system.

Another area in which this system could be of potential assistance is in routine geodetic surveying. Although the accuracy of a centimeter or two may not always be required, a large area the size of the United States could be surveyed in a matter of weeks with the main problem being the deployment of the reflector array.

Many other applications of the system have been suggested including subsidence monitoring (Kahn and Vonbun, 1977), volcano monitoring, measuring the changing thickness of the ice caps etc., and these have been discussed in the report (University of Texas, 1978) of a Workshop on the Spaceborne Laser held at the University of Texas, Austin.

## Conclusion

The general concept of a spaceborne ranging system has been described that would have a geodetic capability of $\pm$ 1 cm in relative positioning of a network of ground reflectors separated from 20 to 1200 km, or more. Such a system appears technically feasible for launch in the early to mid-1980's and could have a major impact on our ability to observe the precursory geodetic motions believed to occur before large earthquakes. Indeed, established on a global scale, with survey areas around all major seismic zones the spaceborne system could provide the first real probability for "capturing" a magnitude 7.5, and above, earthquake.

Simulations of a survey of a large area indicate that the largest error sources in the recovered reflector positions are the gravity field and atmospheric refraction, the latter being smaller but less well understood and itself liable to error. The technological consideration that presently poses the greatest challenge appears to be the development of a laser system that can operate for two or more years in a spacecraft and provide the order of $2 \times 10^7$ pulses. In the timescale proposed for the spaceborne laser this development appears realizable.

As in the case for all optical systems, a consideration in its operation is the weather. In the simulations that have been performed a 50% cloud cover has not significantly affected the quality of the results but has extended the time over which data must be collected, e.g., from 3 days to 6 days. For general surveying work, however, where there is little or no time restriction for acquiring the data, it is probable that weather may be a minor consideration.

Acknowledgements. The author would like to thank Drs. M.W. Fitzmaurice and P.O. Minott for providing information on the spaceborne laser and the ground reflector system, and Mr. W.D. Kahn and Dr. T.S. Englar for allowing me to make use of their simulation studies. In addition, I would like to thank Dr. F.O. Vonbun, Dr. H.H. Plotkin and Mr. D. Premo who have played a major part in developing the concept of the spaceborne ranging system.

## References

Castle, R.O., J.N. Alt, J.C. Savage and E.I. Balazs, Elevation changes preceding the San Fernando earthquake of February 9, 1971, Geology, 2, 61-66, 1974.

Gardner, C.S., Comparison between the refraction covariance model and ray tracing, Radio Research Laboratory Publication No. 486, College of Engineering, University of Illinois, Urbana, Illinois, 1977.

Kahn, W.D. and F.O. Vonbun, Detectability of land subsidence from space, Proceedings of the Second International Symposium on Land Subsidence, International Association of Hydrological Sciences, 1977.

Mueller, I.I., B.H.W. van Gelder and M. Kumar, Error analysis for the proposed close grid geodynamic satellite measurement system (CLOGEOS), Department of Geodetic Science, Report No. 230, Ohio State University, 1975.

Scholz, C.H., L.R. Sykes and Y.P. Aggarwal, Earthquake prediction: A physical basis, Science, 181, 803-810, 1973.

Smithsonian Astrophysical Observatory, Staff, Study of a close-grid geodynamic measurement system, Smithsonian Astrophysical Observatory, Reports in Geoastronomy, No. 5, 1977.

University of Texas, Report of spaceborne geodynamics ranging system workshop, Dept. of Aerospace Engineering, University of Texas, Austin, 1978.

Vonbun, F.O., W.D. Kahn, P.D. Argentiero and D.W. Koch, Spaceborne earth applications ranging system (SPEAR), J. Spacecraft and Rockets, 14, 492-495, 1977.

# Miniature Interferometer Terminals for Earth Surveying

Charles C. Counselman III and Irwin I. Shapiro
Department of Earth and Planetary Sciences
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Abstract. A system of miniature radio interferometer terminals is proposed for the measurement of vector baselines with uncertainties ranging from the millimeter to the centimeter level for baseline lengths ranging, respectively, from a few to a few hundred kilometers. Each terminal would have no moving parts, could be packaged in a volume of less than 0.1 $m^3$, and would operate unattended. These units would receive radio signals from low-power (<10 w) transmitters on Earth-orbiting satellites. The baselines between units could be determined virtually instantaneously and monitored continuously as long as at least four satellites were visible simultaneously. Acquisition of the satellite signals by each terminal would require about one minute, but about one second of signal integration, and the collection of only a few kilobits of data from two receiving units would suffice to determine a baseline. Different baseline lengths, weather conditions, and desired accuracies would, in general, dictate different integration times.

The system proposed here could be used to monitor the regional accumulation and release of strain preceding, following, and even during earthquakes. The terminals could be deployed in arrays of various dimensions and densities. Their use could also include monitoring variations in transcontinental and intercontinental baselines, but with reduced accuracy. Comparisons with other systems proposed for extensive measurements of regional baseline vectors appear to favor this interferometric approach.

## I. Introduction

The technique of very-long-baseline interferometry (VLBI) is only a decade old and barely approaching adolescence. Nonetheless this radio interferometric method has seen broad application, especially in astronomy. In geodetic applications, the demonstrated level of repeatability of baseline-length determinations ranges from ~3 mm for ~1 km distances (Rogers et al., 1978) to ~3 cm for transcontinental distances (Robertson et al., 1979). This combination of precision and range should make VLBI a very powerful technique for monitoring the time dependence of regional and continental baselines. Yet it is still not widely used for this purpose. Why? A principal reason has been cost. Applications of VLBI to geodesy have hitherto involved observations of the random, weak, radio signals received from distant, extragalactic, sources. The achievement of useful signal-to-noise ratios with these sources has dictated the use of large diameter antennas, expensive atomic frequency standards, and wideband tape-recording and correlating systems.

In contrast, only very small, simple, and inexpensive ground equipment is required to utilize the relatively strong, precisely controlled, radio signals that can be transmitted from Earth satellites. Nonetheless, although several methods have employed satellite signals to determine baselines, none of these methods has yet achieved the measurement precision demonstrated with VLBI. Why not? What is the "secret ingredient" of the interferometric technique? Basically it is the use of differencing. Interferometry, per se, involves the differencing of the phases of signals received at the two ends of a baseline. With properly designed equipment, the inherent "2π" ambiguity in these phases of radio signals from a given source can be eliminated, and advantage taken of precise phase measurements of the signals from several sources, to determine a baseline with an uncertainty equal to a small fraction of the wavelength of the radio signals. Further, this baseline determination does not depend on the signals from any source having any particular temporal regularity.

The baseline vectors determined by radio interferometric techniques can be related to the best known approximation to an inertial frame: the positions in the sky of compact, extragalactic, radio sources. Of course, when the baseline vector is determined from interferometric observations of radio signals from satellites, an extra step is required to relate the positions of the satellites to those distant radio sources. Again, the technique used is interferometric but, here, use of the full panoply of the conventional VLBI armamentarium is required.

A system that combines the advantages of VLBI with the benefits of strong satellite signals could open a new era in geodesy. We describe a relatively simple system here. It would employ compact ground equipment with no moving parts and low-power radio transmitters on a set of satellites. We dub this combination the Mighty MITES system, MITES being an acronym for Miniature Interferometer Terminals for Earth Surveying (or, to be presumptuous and facetious simultaneously, the Massachusetts Institute of Technology Engineering Success). Our system appears

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280. The Ohio State Univ., Columbus, Ohio 43210.

to be potentially more efficient than any other so far proposed for the three-dimensional monitoring of crustal strain accumulation and release over distances ranging from tens of meters to hundreds of kilometers. This system could also be applied to a variety of other surveying and navigation problems on land, sea, and air, and in space.

In the remainder of this paper, we describe a preliminary design of the proposed system, the operation of the system, the limits on the achievable accuracy, an inexpensive method to demonstrate the feasibility of the system, some of its possible applications, and, finally, a comparison of the system with other space geodetic methods.

## II. System Description

In this section we discuss, in turn, the basic concepts underlying our proposed system, our preliminary thoughts on the design of the relevant satellite and ground equipment, and some possible design modifications.

### A. Basic Concepts

Here we outline briefly some of the concepts upon which our proposed system is based.

1. Observable. The basic quantity that would be measured with this system is the interferometric phase -- the difference between the phases of radio signals from a single satellite received at any two terminals. These phases would each be obtained simultaneously for a set of different radio frequencies ("tones") covering a wide, ~1 octave, band so that (i) the effect of the ionosphere could be virtually eliminated, and (ii) the inherent "$2\pi$" ambiguity of the phase observable could be resolved. Measurements of the frequencies of the received signals would be made concurrently to enable conversion of the phase differences to corresponding delay differences.

2. Baseline Determination. These delay differences could be interpreted in a standard manner to determine the components of the baseline vector. For observations of a given satellite at a given instant, the interferometric phase delay can be expressed approximately as

$$\tau \approx \frac{1}{c}\vec{B}\cdot\hat{s} + \tau^{cl} , \qquad (1)$$

where $\vec{B}$ is the baseline vector, $\hat{s}$ is a unit vector in the direction of the satellite being observed, and $\tau^{cl}$ represents the difference between the epoch settings of the clocks at the two terminals (i.e., their departure from synchronism). For the purpose of this simplified explanation, we assume $B \ll r_s$, where $r_s$ is the minimum distance of a satellite from a terminal, and we suppress the effect of the difference in rate of the clocks at the two terminals (see Section III). Equation (1) shows that at any instant, the observed delay contains information on the projection of the baseline vector along the direction to the satellite, and on the clock-synchronization error. To determine all three components of the baseline vector simultaneously with the synchronization error at any instant, the signals from at least four satellites must be observed and the resultant four linear equations solved for the four unknowns. For a unique solution, the satellites cannot all appear to lie on the same circle in the sky. If they do, then the component of $\vec{B}$ in the direction of the center of this circle cannot be separated from $\tau^{cl}$. (See also Section III.A.)

3. Reference System. We assume that the positions of the four (or more) satellites observed are known with respect to a well-defined coordinate system. Different coordinate systems could be utilized, depending on the application. We consider one example (see, also, Section V): a nearly inertial coordinate system with an orientation defined by the directions of extragalactic radio sources. Such a system could be used if three (or more) base stations made continual differential interferometric phase observations of the satellites with respect to those extragalactic radio sources that appear in the same part of the sky (Counselman et al., 1972; Preston et al.,1972). Given that the vector positions of the base stations with respect to the directions of the extragalactic sources were already known, virtually instantaneous measurements of the differential interferometric phases between a satellite and an extragalactic source from observations from each of two independent baselines suffice to determine the satellite's direction in this nearly inertial frame. Thus, the vector baselines between our compact ground terminals could be related to this frame through their coordinates in the satellite frame. The base stations would of course require conventional VLBI instrumentation to observe the extragalactic radio sources. As long as the direction of the comparison extragalactic source is neither too far from that of the satellite nor too near that of the sun, the observations of the satellite need be made only of the highest transmitted tone.

We defer until Section IV discussion of the limits on achievable accuracy in the determination of the baseline vector with respect to a reference frame.

4. Global Positioning System. The satellites of the NAVSTAR Global Positioning System (GPS) (Parkinson, 1976), scheduled for launch during the next few years, will have orbits almost perfectly suited to our proposed VLBI system. In the planned steady-state configuration, there will be 24 such satellites, eight distributed in each of three orbital planes, spaced by 120° in nodal longitude. Within each plane the satellites will be approximately evenly distributed around a nearly circular orbit of ~63° inclination, ~20,000-km altitude, and ~12-hour period. As a result, at any place on the surface of the Earth at least six satellites will always be visible, and of these at least four will be suitably distributed (i.e., not all on or near any single circle in the sky).

The expected useful lifetime of a GPS satellite is five years with a maximum of about seven set by the amount of stored gas for the attitude-control system. Thus, in principle at least, there will be continual opportunities to modify the payload to accommodate our proposed system.

B. Satellite Equipment

We now describe the equipment we propose to place on each satellite. First, we discuss the factors governing the choice of subsystems and then each of the subsystems. For definiteness, we consider the design in the context of the GPS satellites.

1. Requirements. What special requirements must be satisfied by the equipment aboard the satellites for our proposed system to be operated successfully? First of all, the satellite must transmit a sufficiently strong and stable signal. The signal must also be structured to allow the user to resolve the "2π" ambiguity in the interferometric phase. In effect, the same cycle of the signal must be identified at both ends of the baseline. The usual method of making such an identification in VLBI is to combine observations made simultaneously at several different frequencies, or tones, $f_i$ (i=1, 2,...). The respective ambiguity spacings in delay, equal to $1/f_i$, are different, and the overall delay ambiguity can be resolved if, within the range of possible delays allowed by a priori information, there is a unique delay that is consistent with the phase observations at all frequencies. The task of resolving the phase ambiguity is complicated, however, by the dispersion of the propagation medium and by possible interference from reflected signals (see Section VII). The set of frequencies used must be chosen with these complications in mind.

Since, for the range of frequencies of possible interest, the Earth's ionosphere introduces a delay proportional to the inverse square of the frequency of the signal, it would appear desirable to employ the highest possible frequencies with the proposed system. However, we must also consider our paramount desire for a simple system. This desire to avoid, for example, high-gain antennas and high transmitted power levels places an upper bound on the usable frequency. Since, for given antenna directivities and transmitted power, the received power is inversely proportional to the square of the frequency, sufficiently high signal-to-noise ratios can only be obtained simply at low frequencies. The optimum range of frequencies is approximately 1-2 GHz, limited on the low end by the ionosphere and on the high end by signal-to-noise ratio considerations. An approximately one octave spread in the frequencies of the tones is important for the elimination of 2π ambiguities in the face of both the ionosphere and the possible interference from reflected signals.

Within this one-octave range of frequencies there are several wide bands allocated under existing international regulations for purposes that would encompass those of the proposed system (International Telecommunications Union, 1968). This system requires relatively little spectrum space -- no more than ten narrow bands, each at most 100-200 kHz wide, within the ~1-to-2 GHz range. The precise placement of these bands within this range is quite flexible and could be adapted to existing constraints. Basically, we require one band near the low-frequency limit, one near the high-frequency limit, and the others distributed between these limits. The inter-band spacings would range from a minimum of ~1 MHz to a maximum of ~400 MHz. The spacings could be in approximately geometric progression but do not have to vary monotonically with frequency (see Section III for further discussion).

2. Transmitter. To satisfy these requirements, each satellite could be equipped to transmit an unmodulated circularly polarized continuous wave (CW) with ~1 watt of power in each of up to ten bands. To avoid having to contend with interference between signals received from different satellites, each of these bands could be subdivided into contiguous, nonoverlapping "channels". Since each GPS satellite could share a channel with that other satellite spaced 180 deg apart in longitude, the number of channels need only be 12. Setting the width of each channel equal to $8\times10^{-6}$ of its center frequency would allow for a tolerance on transmitted frequency of ±1 part in $10^6$, as well as for the extremes of Doppler shift, of ±3 parts in $10^6$, that could be

observed on Earth for signals from satellites in orbits of the GPS type. Thus, the total width of the band of 12 channels at ~1 GHz would be $\lesssim$100 kHz, and at ~2 GHz would be $\lesssim$200 kHz. Should less of the spectrum be available, modifications are possible to compensate, as discussed in Section II.D. Here we proceed on the assumption that no modifications will be required.

Transmitters are available commercially which yield 1 w of radio-frequency output for 3.6 w of d.c. input power. Such devices have dimensions and mass of about 5 cm x 5 cm x 10 cm and 0.9 kg, respectively. Thus, for a total of ten bands, 36 w of DC power and a volume of 0.0025 $m^3$ would be required. These transmitters have an adequate long-term frequency stability of 1 part in $10^6$ and a more than adequate spectral purity of ~1 part in $10^{10}$. The overall volume and mass could undoubtedly be lowered for use in space, especially if ten transmitters were packaged together. The total DC power requirements could also be lowered by some combination of an insignificant sacrifice in spectral purity and a use of a single oscillator to generate multiple tones through modulation which, itself, requires little power. However, to maintain high reliability, one would prefer not to eliminate too much redundancy.

3. Power Combiner. The outputs of the separate transmitters would be combined before being fed to an antenna. Design of such a combiner would be straightforward, involving mainly a set of resonant, low-loss circuits.

4. Antenna. We envision each satellite equipped with a circularly-polarized antenna of modest directivity ($30° \lesssim$ beamwidth $\lesssim 60°$), which illuminates the entire visible portion of the Earth approximately uniformly. Such an antenna would have at least 10 db gain, but would be physically small and would not require precise pointing. A suitable antenna would also have a bandwidth sufficiently broad to yield this gain over the span from 1 to 2 GHz. (Variations in the transmitting antenna's gain, as well as in the receiving antenna's effective aperture, at the frequencies of the various tone transmissions, could be compensated by suitable adjustments in the relative transmitted powers.) Types of antennas that might satisfy these requirements include cavity-backed spirals and conical spirals. The sizes and masses of such antennas are quite modest. For example, a commercially available cavity-backed spiral for these frequencies has a gain of nearly 10 db and is about 12 cm in diameter and 10 cm deep, with a mass of about 1 kg.

These characteristics of the satellite equipment are summarized in Table 1 and an overall block design is shown in Figure 1. This equipment would, of course, require integration into the GPS configuration. Questions of location, mass, volume, and heat generation and dissipation all must be addressed. The only positional requirement is that the antenna must face earthward; the other major requirement is, of course, for the appropriate supply of direct current. Any requirement for uplink and downlink telemetry would be very modest, because the system is so simple.

## C. Ground Terminal

The most important feature of the transmitter described in Section II.A is that it enables a relatively simple ground terminal to be used. Each such ground terminal would consist of an antenna, a receiver, a frequency and time standard, digital counting and timing circuitry, a calibration signal generator, atmospheric sensors, control logic, a system to store and/or telemeter data, and a power supply. We discuss these components, each in turn, after first describing some of the requirements that the ground terminal must meet.

1. Requirements. The ground terminal must have an antenna matched in circular polarization to the satellite antenna, but with little directivity. We wish to avoid the complexity and expense of re-pointing the receiving antenna, by either mechanical or electrical means. Further, the antenna must receive signals simultaneously from widely separated directions in the sky, with elevation angles as low as 10°. The receiving antenna must also operate over an approximately 2-to-1 range of frequency, corresponding to a wavelength range of approximately 15 to 30 cm.

With any receiving-antenna design, it is also necessary to ensure that the phase of the signal received directly from a satellite is not altered significantly (see Section VII) and in an unknown way through interference that could arise from waves reflected or scattered to the antenna from the ground. The significance of such phase errors depends, of course, on the desired geodetic accuracy. In Section IV.C, we show that the errors in baseline determination would generally be comparable to the equivalent path-length errors in the measurements of the interferometric phase delays. Thus, for example, to achieve one centimeter baseline accuracy, would require phase errors under 24 deg for $f \simeq 2$ GHz. Equivalently, the reflected signal would have to be at least 7 db weaker than the di-

rectly received signal. For one millimeter geodetic accuracy, the field strength of the reflected and scattered signals, and/or the antenna's sensitivity to such signals, would have to be attenuated tenfold (20 db in power) further. On the other hand, if geodetic precision, rather than accuracy, were the goal, as in some applications, one could tolerate a more cluttered and uneven environment for the receiving antenna because the distribution of azimuths and elevations of the GPS satellites as viewed from any given location repeats approximately with a period of 12 hours, so that the signal-reflection pattern and the consequent electrical phase variations would also be repeated periodically.

2. Antenna. The requirement that the phase of the signals received directly from the satellite not be altered significantly by signals reflected or scattered into the antenna could be satisfied by ensuring that the receiving antenna has an unobstructed view of the sky above 10° elevation, and that no large, elevated planar reflectors such as building walls or fences are situated so as to reflect satellite signals toward the antenna. The antenna also would have to be placed above a ~30 to 90 cm diameter metal "ground screen" of sheet metal or stiff wire mesh placed flush with the ground. The exact diameter needed and the restrictions on the local terrain would depend on the desired accuracy. The screen would also serve to reduce pickup of thermal "noise" radiated from the ground. (This method of siting an antenna is commonly used for high-accuracy measurements of microwave antenna radiation patterns).

The antenna would be mounted in the center of the ground screen and could have various forms. We describe one possible antenna design here. Mechanically somewhat complicated, but with apparently more than adequate electrical properties, this antenna would consist of a simple stack of crossed pairs of horizontal, half-wavelength dipoles. They would be made of metal rod or tubing with each pair cut for one of the, at most, ten frequency bands. Of course, the closely-spaced bands could be served by a single pair of dipoles. Each dipole pair would be placed three-eighths of a wavelength above the ground screen. The two orthogonal dipoles comprising each pair would be fed in phase quadrature for circular polarization. Each pair could be connected to a separate, narrowband, receiver preamplifier if desired; more than one pair could also be connected efficiently to a single broadband preamplifier.

With a height of $3\lambda/8$ above a horizontal ground plane, the dipole pair has greater than unity gain relative to a circularly polarized "isotropic" antenna at all elevation angles above about 20°, and its gain is about -5 db at 10° elevation, although the response approaches zero at the horizon. There would also be negligible ohmic loss with this antenna.

A crossed-dipole stack would have to be specially designed and tuned to the MITES system frequencies. However, this task is rather simple, and the cost of replicating the resultant antenna in moderately large quantity would probably be under $50 per unit.

Finally, we note that the phase-shift characteristics of the antenna itself, in conjunction with its ground screen, would be unimportant if identical antennas were used for all terminals; but, in any event, the antennas could be calibrated on a test range.

3. Receiver. The receiver could have a simple, uncooled, transistor radio-frequency (RF) front-end amplifier, or possibly two or more such amplifiers, each tuned to a different portion of the 1-to-2 GHz range spanned by the transmitted signals. In either case the total volume and mass of the front end would be under 100 $cm^3$ and 0.1 kg, respectively. A system noise temperature of 200°K could be easily achieved and would be sufficiently low, given the characteristics of the transmitted signals and the antennas described above. No difficult gain or phase stability demands would need to be met, because the output of the final amplifying stage of the receiver would be hard-limited or "clipped", and because delays of the signals through the receiver would be separately monitored, as discussed in Subsection 6.

After passage through the front-end amplifier, the signals received in each of the 100-to-200-kHz-wide bands would be mixed with a fixed-frequency local oscillator(LO) signal, derived by coherent multiplication from the master oscillator, to convert them to an intermediate-frequency (IF) band centered at ~200 kHz. The value of 200 kHz was chosen to minimize the requirements on time resolution in the subsequent measurement of the phase of the IF signals (see Subsection 8). No lower value could be chosen because of the desire to keep the ratio of the frequencies of the highest and the lowest IF signals below three in order to prevent interference from the odd harmonics generated by clipping. The down conversion to IF might be accomplished in a single stage with a quadrature-phasing, single-sideband mixer of the type described by Rogers (1971), or depending on considerations of dynamic range, stability, etc., two or more conversion stages might be used.

4. Time and Frequency Standard. Since observations would be made of at least four satellites simultaneously, the departure of the clock at each terminal from synchronism in epoch (and rate) with the clock at every other terminal can be determined from the observations as explained in Section II.A.2. A high order of stability is therefore not required of the time and frequency standard at each terminal. The primary requirement is for sufficient short-term frequency stability, of ~1 part in $10^{10}$, to maintain phase coherence during each, approximately one second, period of coherent integration. Long-term stability of ~1 part in $10^{6}$ is desired to facilitate acquisition of the satellite signals. The need for the ~1-sec observation periods of different terminals to overlap substantially implies that time must be kept at each terminal to an accuracy better than about 0.2 seconds, unless the observations are made continuously. In the latter case, interpolation between successive observations could be used to "match" time tags. Otherwise, drift in epoch error would have to be monitored and the clock in each terminal would have to be reset periodically, via telemetry if necessary. (See, also, Section III.A.)

These requirements on frequency and time are readily met by a compact, commercially-available, crystal oscillator. We shall henceforth refer to this as the "master" oscillator.

5. Digital Counting and Timing Circuitry. The digital counting and timing circuitry comprise the heart of the terminal and we describe them in some detail here. For convenience in description, assume that the received signals have been converted to the IF band, centered at ~200 kHz, as mentioned in Subsection 3. Each of these IF signals, one for each of the up to ten RF bands, would be bandpass filtered to approximately 100 to 300 kHz and symmetrically limited or "clipped" as mentioned earlier, to obtain a two-level signal which would be switched via "logic" circuits to selected inputs of a set of approximately 20 identical modules (see Subsection 8). Each such signal could be directable to the input of any module, according to the plan discussed below, although a more efficient arrangement is possible. The function of each module would be to measure the phase and the frequency of one signal from one satellite.

The IF input to a module would be fed to a second-order phase-locked loop whose bandwidth under the expected conditions of +10 db or greater signal-to-noise ratio (SNR) could be switched to values of either 7.5 Hz or 30 Hz, and whose tracking range in either case would span the IF band of from 100 to 300 kHz. (Note that the SNR is inversely proportional to the bandwidth of the loop and that the power received by the crossed-dipole antenna will be inversely proportional to the square of the radio frequency. Hence, under the assumption that the effective radiated power from each satellite is the same for all tones, the loop bandwidth for reception of the 2-GHz signals must be fourfold smaller than the loop bandwidth for the 1-GHz signals in order to maintain the same value of SNR.) The loop would be locked to, and would track, the first suitable satellite signals encountered in this band. The loop SNR would be at least +10 db, and the corresponding root-mean-square (rms) random phase error in the loop output would be less than 13°, given (i) the effective transmitted power of ~10 w (1 w transmitted, coupled with 10 db of antenna gain); (ii) the worst-case satellite elevation angle of 10°; (iii) the crossed-dipole receiving antenna; (iv) the receiving system noise temperature of 200°K; (v) the 2-db clipping loss; and (vi) the 30-Hz and 7.5-Hz loop bandwidths for the 1-GHz and 2-GHz signals, respectively. The dynamic tracking error due to the time-rate-of-change of the input signal frequency would be well under 1°, as would the static phase error, even with the use of the narrower loop bandwidth for the highest-frequency signals. The loop output would be fed to digital circuits which would perform two functions: (i) a continuously accumulating count of the integral cycles of the locked oscillator; and (ii) an accurate measurement, modulo one cycle, of the oscillator phase relative to that of the receiver clock. The instantaneous value of the accumulated cycle-count would be sampled nondestructively at selected times, according to the clock. The phase measurement would also be made at selected times, with ~100 ns time resolution, equivalent to ~0.02 cycle phase resolution at the ~200 kHz IF. This measurement could be made with a digital start-stop counter with a "clock" rate of 10 MHz. The count would be initiated by a command from the ground terminal clock, and would be halted appropriately by the output of the locked oscillator, for example, by the next occurring positive-going transition. Averaging would be performed by timing a succession of such time intervals (i.e., by restarting and restopping the count) over a time span of approximately one second. By virtue of this averaging, the rms random noise of the equivalent phase measurement could be reduced to less than 5°, the equivalent of 2mm of electrical path length at a frequency of 2 GHz. (The equivalent noise bandwidth of the measurement would be ~2 Hz instead of ~1 Hz because of the image "fold over": for

the 2-GHz signals, the SNR in this bandwidth would be >18 db.)

The phase-locked-loop portion of each module could be implemented with low-power, analog, integrated circuits and low-power Schottky transistor-transistor logic (TTL) circuits. All of this circuitry combined would consume less than 0.5 w. Continuous cycle-counting could be performed by a standard, low-power, "large-scale integrated" (LSI) circuit that also would consume about 0.5 w of power. The 100-ns resolution timing could be implemented with standard TTL circuits that would use approximately 2 w. However, power would have to be supplied to the latter circuitry only for the ~1 second interval during which the fractional-cycle phase measurement was being made. For many applications, these measurements might be repeated infrequently, at several-minute intervals, for example. In such cases, the average power consumed by each module would remain at the ~1 w level required to sustain the phase-locked-loop and cycle-accumulating circuits. Inasmuch as the complete receiver would require about 20 such modules, low power consumption per module is an important advantage.

The feasibility of building a module with the characteristics described has been demonstrated by the differential Doppler receiver developed by H. F. Hinteregger and one of us (C.C.C.) and applied to differenced VLBI observations of the signals transmitted from the Apollo lunar-surface experiments packages (Counselman and Hinteregger, 1973; Counselman et al., 1972; see also Figure 3 in Counselman, 1976).

6. Calibration Signal Generator. Signals from satellites would undergo delays in the RF amplifier(s), the mixers, and the IF amplifiers of the receiver. These delays could be calibrated by injecting a suitable signal of low power directly into the RF input of the receiver. The calibration signal could consist of a periodic train of pulses, each $\lesssim$20 ps in duration and all derived directly from the master crystal oscillator. A pulse repetition frequency of $\lesssim$100 kHz would be chosen so that at least two of the harmonics of this calibration signal would appear within the passband of each IF amplifier. Two harmonics are required to monitor the group delay as well as the phase delay variations of the receiver. The group delay -- the variation of phase with frequency -- is important to monitor because the signals received from different satellites have different frequencies.

The phases of the calibration-signal harmonics would be measured in exactly the same way as would the phases of the satellite signals. This method has been used successfully in the calibration of other geodetic VLBI receiver systems, at the millimeter level of accuracy (Whitney et al., 1976; Rogers et al., 1978). When combined with the results from just one set of measurements of the phase-shift-vs.-frequency characteristics of the IF portions of the receiver, carried out when these are originally built, the information from the calibration signals should serve to account for all receiver phase variations to within an uncertainty of under 1 deg. The calibration signals would not have to be monitored continuously, but could be checked, for example, before and after each satellite tracking period as discussed below.

7. Atmospheric Sensors. The delay of the radio signals introduced by the neutral atmosphere must be modeled accurately, as discussed in Section IV.A. Thus we would include a suitably compact, electronically-readable barometer, thermometer, and hygrometer in each interferometer terminal. Such instruments are available commercially and are incorporated, for example, in some compact satellite Doppler-tracking receivers. If the data from a particular set of terminals are to be analyzed at a central location, it may be feasible and more economical to omit the sensors from the terminals and to interpolate from weather-station records to obtain the values of the atmospheric parameters (see Section IV.A).

8. Control Logic. All of the control functions to be carried out in the ground terminal could be automated straightforwardly through the use of an integrated-circuit microprocessor. Routine functions would include collection of data including the atmospheric information, the cycle-counts, the time-increments, and the received signal and noise levels. Less routine functions would include the determination of the proper phase calibration of the signals received in the different bands. The switching of the phase-locked loops to the appropriate frequencies for acquisition of the satellite signals would also be automated. For signal acquisition, at least two alternative strategies could be followed, each based on the controller's ability to (i) monitor the frequency of the oscillator in each phase-locked loop through cycle-counting; (ii) slew each such oscillator upward or downward in frequency; and (iii) note the amplitude of the coherently-detected signal from each phase detector.

If the controller were in two-way communication with a central processing station, this station could instruct the controller regarding the times and frequencies at which new satellites could be acquired. But it would also be possible,

and often preferable, for the controller to operate automatically. At preset intervals, each of the 12 frequency channels in the lowest-frequency (~1 GHz) band would be swept by a phase-locked loop searching for a signal. (Note that it is vastly more efficient to conduct the search at the lowest frequency because (i) the loop bandwidth is wider; (ii) the loop time constant is shorter; and (iii) the spectrum to be searched is smaller. Since the first two factors each vary with the square and the third with the first power of the radio frequency, we conclude that the search time depends on the fifth power of the radio frequency!) For the signal strengths and bandwidths given previously, approximately 10 seconds would be required for one loop to search completely one of these 8-kHz-wide channels. Of course, different loops would be programmed to search different channels simultaneously. Once a new satellite had thus been acquired, others of the ~20 loops could be assigned by the controller to lock onto the signals transmitted in the other bands by the same satellite. Less searching would be required in this step, since the Doppler shifts would be known.

Whenever a satellite is acquired, the phases of all of the sinusoidal signals that it transmits would have to be measured simultaneously at least once, to later enable the interferometric phase ambiguities of the signals to be properly resolved. This task would require the simultaneous commitment of up to ten phase-locked loops to that one satellite. However, once this task, which would take only a few seconds, had been accomplished, only two loops would be needed to track the highest and lowest frequency signals from the satellite. Thus, after elimination of the $2\pi$ ambiguities, it would be necessary to maintain a continuous count of the cycles received from these two signals only to remove ionospheric effects. In the steady state, simultaneous tracking of, say, six satellites would require twelve loops, and eight loops would remain for other tasks. These tasks would include the periodic monitoring of the phases and amplitudes of the calibration signals in all of the IF bands, and the searching for new satellites. To reach the steady-state condition after five of the six satellites have already been acquired will require up to 20 of the phase-locked loops -- ten for the steady state tracking of the first five satellites and up to ten loops to enable the proper phase connection to be made for the signals from the sixth satellite. The total time to reach a steady state would, on average, be about one minute.

The control logic must also contain means to recognize any lock on a spurious signal and to reject this signal; in such a scheme, advantage must be taken of the known relation between the frequencies of the various tones transmitted by each satellite.

9. Data Storage and Telemetry. We estimate that about 3500 bits of data would need to be stored at, or transmitted by, each terminal for the initial determination of a baseline. This total includes allowances of 33 bits for the integer-cycle count, 9 bits for the fractional-cycle phase measurement, 12 bits for the measured value of frequency, and at least 1 error-flag bit to indicate loss of lock for each of the up to 10 signals received from each of 6 satellites, plus a few hundred bits for time tags and other labels, and for data on the atmospheric parameters. This quantity of data could be transmitted over an ordinary telephone line in 2-3 seconds. After the initial signal-acquisition and phase-ambiguity-resolution operations had been completed, and when only two phase-locked loops were tracking each satellite, a baseline redetermination could be performed with an additional 1 kilobit of data from each terminal. Data from a large number of observations could also be stored within the terminal in inexpensive, solid-state memory devices, and could be collected later, for example by "dial-up" telephone. With the addition to the terminal of a 1-watt microwave transmitter, the data could be relayed from a remote to a central location through a satellite repeater. Alternatively, if a telephone line were availa-

TABLE 1. Possible Characteristics of Satellite Equipment for Use with the Mighty MITES System

| Equipment | Description/ Characteristics |
|---|---|
| TRANSMITTER | |
| Frequencies | ≤10 tones spaced between 1 and 2 GHz |
| Polarization | circular |
| Power | ~1 w per tone |
| Total DC Input | < 35 w |
| Total Volume | ~0.0025 $m^3$ |
| Total Mass | ≤10 kg |
| ANTENNA | |
| Type | cavity-backed spiral |
| Bandwidth | ~1 to 2 GHz |
| Beamwidth | ~30 to 60 deg |
| Gain | ≤10 db |
| Dimensions | 12 cm dia x 10 cm deep |
| Mass | ~ 1 kg |

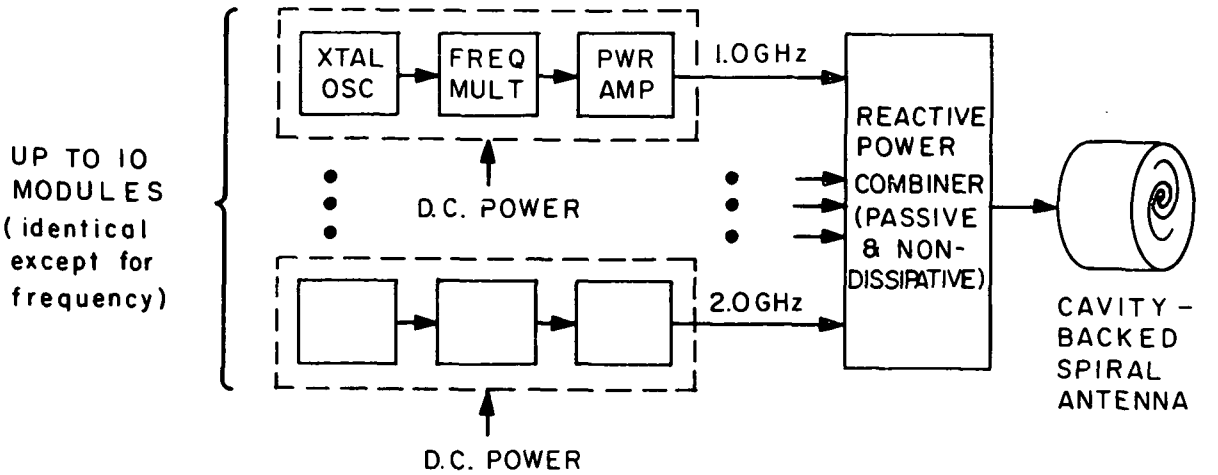



Fig. 1. Block diagram of radio transmitting equipment to be placed on satellite, for use by Mighty MITES.

ble within ~50 km, a UHF or VHF radio link and a telephone "patch" could be used.

10. Power Supply. The complete terminal as described would require about 50 watts of power. Thus, it could not operate continuously and unattended for long periods without a steady power source. For applications in which observations every second were not required, the average power consumption could be kept to about 1 watt, since most of the electronics, save the master crystal oscillator and the clock, would not have to be powered except during observations. In these applications, power might be drawn from batteries charged by an array of solar cells.

These characteristics of the proposed interferometer ground terminal are summarized in Table 2. See also Figure 2.

D. Possible Design Modifications

The particular design we described was motivated mostly by the desire for simplicity and reliability. However, many modifications of the design of the Mighty MITES system could be made to accommodate different requirements or constraints. For example, if necessary, the bandwidth required for the system could be reduced substantially with perhaps negligible degradation of performance. It is likely that a simple analysis, yet to be performed, would show that satellites separated by only 90° in orbital longitude could be assigned the same nominal transmitter frequencies without risk of mutual interference, because of the difference between the Doppler shifts of their signals as received on the Earth. If so, the number of channels, and the total bandwidth, could be halved. It might

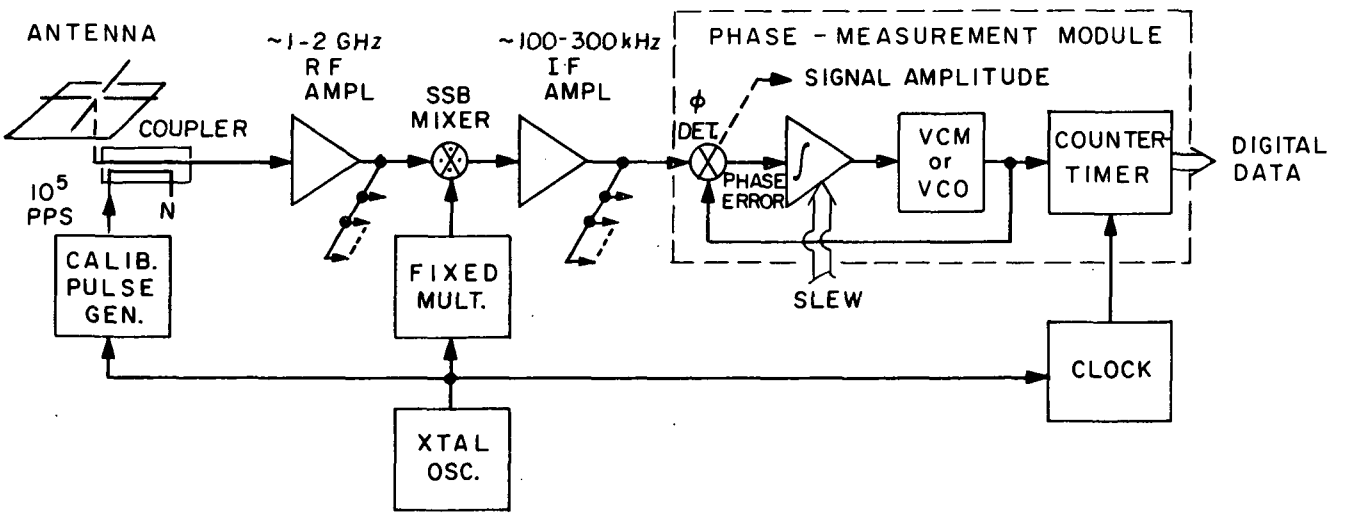



Fig. 2. Block diagram of radio signal receiving and processing portions of Mighty MITES ground terminal. One RF-to-IF converter and one IF amplifier are required for each of the up to ten RF bands. A total of 20 phase-measurement modules are required, as explained in the text.

even be feasible to reduce the number of channels in each band to one; that is, all 24 satellites might have the same nominal transmitter frequencies. Typically, about eight satellites are simultaneously visible at a given ground location,

TABLE 2. Possible Characteristics of Ground Equipment for the Mighty MITES System

| Equipment | Description/ Characteristics |
|---|---|
| ANTENNA | |
| Type | Crossed dipoles stacked $3\lambda/8$ above ~30 to 90 cm diameter ground screen |
| Polarization | Circular |
| Frequencies | Selected coverage from 1 to 2 GHz to match transmitted tones |
| Gain | >1 relative to circular "isotropic" antenna for elevation angles ≥20 deg; -5 db at elevation angle of 10 deg |
| RECEIVER SYSTEM | |
| Pre-amplifier(s) | Uncooled transistor(s) |
| Frequency Range | 1 to 2 GHz |
| IF Amplifiers | ≤10 units |
| Center Frequency | 200 kHz |
| Bandwidth | 200 kHz |
| Time and Frequency Standard | Crystal oscillator |
| Short-term Stability | ~1 part in $10^{10}$ |
| Long-term Stability | ~1 part in $10^{6}$ |
| Digital Counting and Timing Circuitry | ~20 units |
| Phase-Locked Loop Bandwidths | 7.5 and 30 Hz |
| Time Resolution | 100 nsec |
| RMS Noise in Phase Measurement | <5° for one second integration in worst case |
| Calibration Signal Generator | |
| Pulse Duration | ≲20 psec |
| Pulse Repetition Frequency | ≲100 kHz |
| Atmospheric Sensors | Barometer, thermometer, hygrometer |
| Controller | Integrated circuit microprocessor |

TABLE 2. (continued)

| Equipment | Description/ Characteristics |
|---|---|
| Receiver System (continued) | |
| Data Storage or Telemetry | |
| Capacity | 1000-3500 bits per baseline determination |
| Medium | Solid-state memory, cassette recorder, radio link and/or telephone line |
| Total Power | |
| Peak | ~50 w |
| Average | ~1 w |
| Total Mass | ~15 kg |
| Total Volume | ~0.1 $m^3$ |

but the Doppler-shift and random, transmitter-frequency, differences would cause the eight received frequencies to be scattered throughout the width of the band. Two received signals would have frequencies differing by less than, say, 5 Hz probably less than 5% of the time. It should be possible to design a receiver to make accurate phase measurements on unmodulated signals which have similar strengths and frequencies that differ by as little as 5 Hz; the receiver could also give an indication of when valid measurements were not obtained due to insufficient frequency separation between two signals. However, this approach might introduce considerable complication into the receiver. A more efficient approach, if it were desired to have many or all satellites sharing a single channel, might be to modulate the signal transmitted by each satellite with a unique, but narrowband (≪1 kHz) "code" signal which could be detected in the receiver and used to discriminate between the transmissions from the different satellites. The latter approach is in fact used with the standard GPS navigational signals, except that the codes employed are rather complex and the modulation bandwidth is over 1 MHz. (The high GPS code complexity and modulation bandwidth are required for other purposes that are not essential to our system.)

## III. System Operation and Costs

In this section, we recapitulate briefly the operation of the Mighty MITES system with some more detailed emphasis on the technique for elimination of the $2\pi$ ambiguity in the phase-delay measurements. We also comment very briefly on the costs of the system.

### A. Operation

Each satellite would transmit a set of up to ten different tones. Each tone from each satellite would lie in a given radio-frequency (RF) band. These bands, up to ten in all with each 100 to 200 kHz wide, would be distributed suitably between ~1 GHz and 2 GHz. The ground terminal antenna, being approximately omnidirectional for elevation angles above 10 deg and sensitive to signals in each of these bands, would pick up the tones from those satellites within view. The incoming signals in each band would then be amplified and converted to an intermediate frequency (IF) band, centered at ~200 kHz, by mixing the amplified signals with a local oscillator signal derived by coherent multiplication from the master crystal oscillator (Section II.C.4). The resultant IF signals would then be bandpass filtered, to select signals within a 100 to 300 kHz band, and clipped to obtain a two-level signal that could be handled easily by digital circuitry. A subset of the ~20 modules (Section II.C.8) that contain phase-locked loops would be automatically directed to search for signals originating in the 12 different channels in the lowest RF band (Section II.B.2), since the efficiency of the search would be proportional to the inverse fifth power of the RF. (While the terminal is tracking signals from a given set of channels in this band, further search would be confined to the remaining channels.) After acquisition of a signal in the lowest band, the terminal would seek to acquire the signals in each of the corresponding channels in the up to nine higher-frequency bands. The known separations between the frequencies in the various bands would enable this search to be conducted very efficiently. Should the search fail in a sufficiently large number of bands, the originally-acquired signal would be assumed to be spurious and a new search of the channel in the lowest RF band would be undertaken picking up from the frequency of the spurious signal. Only ten seconds would be required for a module to search completely each ~8 kHz wide channel (Section II.B.2) in the lowest RF band as the loop bandwidth would be ~30 Hz (Section II.C.5), and its time constant ~0.03s. Any module would also be able to track any signal in the total 200 kHz-wide IF band. The signal-to-noise ratio for the satellite signals received would be sufficient for these operations to be carried out for elevation angles as low as ~10 deg.

The outputs from the set of loops which had been locked to the up to ten tones transmitted by a satellite, would be analyzed by digital circuits to (i) accumulate a continuous count of the number of integral cycles made by the locked oscillator in the loop; and (ii) measure, modulo one cycle, the phase of this signal, relative to the clock in the terminal which is controlled by the master crystal oscillator. These measurements, combined, would constitute the basic one to be made by each ground terminal: the total phase of each of the up to ten tones received from each of the satellites being tracked. These phases would also be calibrated for the effects of delays within the receiver (Section II.C.6). The resultant total phases would be suitably averaged, say over one second, and time tagged according to the clock in the terminal. After initial acquisition, it might be sufficient for some applications to continue to monitor only the phases of the signals in the highest and lowest frequency channels for each satellite (see Section II.C.8).

In one mode of operation, these averaged data would be transmitted, along with auxiliary information including the measured values of the frequencies of the signals (Section II.C.9), from each terminal to some central location. At this central location, the data from any pair of terminals could be analyzed to (i) remove the $2\pi$ ambiguity and, simultaneously, the effect of the ionosphere; and (ii) estimate the components of the baseline vector between the locations of the two terminals. To perform this task, additional information is needed, namely the positions of the satellites as functions of time. Crude or refined, a priori, information on terminal locations could also be used.

The radio frequencies of the tones would be distributed between ~1 and 2 GHz in a manner designed to facilitate removal of the $2\pi$ ambiguities in the interferometric phase delays. This distribution would involve spacings in frequency that are nearly in geometric progression, starting from a minimum spacing of about 1 MHz. (A strictly geometric progression would not be used because (i) a priori knowledge, e.g. of the ionosphere, would be incorporated in the choice of the first few spacings; (ii) the total spread would be limited by other constraints; and (iii) we would wish to be conservative.)

The phases of each of the signals measured by each terminal could be converted to delay via use of the measured value of the radio frequency. The $2\pi$ ambiguity and the ionospheric effects would be eliminated from the resulting set of up to ten interferometric phase delays at each measurement epoch by a "bootstrap" algorithm in which the ambiguity is eliminated first between the delays that result from measurements at the closely-spaced pairs of frequencies. These frequencies would be placed at the high, 2 GHz, end of the total band since the uncertainty in the ionospheric effects would be lowest there. Apart from the effects of noise, the shape of the ambiguity-free phase-delay vs. frequency curve is accurately known and, of course, taken into account in the algorithm. We omit the detailed presentation of this fairly straightforward algorithm to spare the reader the superficial complications engendered by symbols being festooned with the three sets of subscripts and superscripts necessary to distinguish the different satellites, ground terminals, and tones.

For continuing observations of the same satellites from the same terminals, the ambiguity and ionospheric elimination algorithms could be largely bypassed after the initial elimination. Only the interferometric phase delays for the highest and lowest tones need be followed continuously; these could be combined very simply to remove the ionospheric effect. Similarly, for terminals closely-enough spaced for ionospheric effects to be negligible, the ambiguity-removal algorithm could be simplified somewhat.

The corrected interferometric phase delays for the satellites tracked from a pair of terminals would be analyzed by, say, a standard least-squares algorithm for each epoch to determine the vector baseline, as outlined in Section II.A.2. Variations in clock behavior at either terminal over the signal integration interval would not affect the baseline result since each satellite is observed at the same times from a given terminal. The effect of such variations would therefore cancel because of the linearity of Equation (1), which would cover the same time interval for the observations of each satellite. However, a "common-mode" error in the epochs of the clocks at the two terminals would affect the baseline result because the assumed positions used for the satellites would be incorrect. Time tags on the interferometric data accurate to a millisecond would reduce this error to a tolerable level. Such clock accuracy could easily be maintained for terminals in two-way communication with a central processor. It would also be possible to obtain this information directly from the signals transmitted by the GPS satellites in their normal mode of operation. Alternatively, with observations of at least five satellites simultaneously, one could solve for this common-mode epoch error. In general, at least six satellites would be tracked simultaneously and thus some redundancy would be retained.

### B. Costs

It is obviously premature to discuss in detail the costs of a Mighty MITES system. However, a few general statements can be made. The ground terminal, if replicated in reasonably large numbers, should not cost more than a few thousand dollars per unit which would make the capital equipment costs negligible compared to the overall costs of the system which would likely be dominated by labor costs. Because of their potential simplicity and reliability, and ability to be operated unattended, the Mighty MITES could, however, allow the labor costs to be reduced relative to those for many other monitoring systems (see Section VII).

The cost of the equipment to be added onto the GPS satellites, if amortized over a large number of ground terminals, might not add significantly to the overall systems costs. Neither would the cost of a central computer be significant since most of the computations are carried out in the ground terminals. However, the development costs for both the ground-based and the satellite parts of the system, despite their intended simplicity, would likely be of the order of a million dollars.

## IV. Limits on Attainable Accuracy

The accuracy attainable in baseline determinations will be limited primarily by errors in knowledge of (i) the propagation medium and (ii) the positions of the satellites. The effects of both of these sources of error increase, albeit differently, as the length of the baseline increases; we discuss each in turn. Finally, we discuss the less important limits due to the relative geometric configuration of terminals and satellites, and due to instrumental effects.

### A. Propagation Medium

The propagation medium contains, in effect, two components: the ionosphere and the atmosphere. The influence of the former can be virtually eliminated, as described in Section III, by utilizing its dispersive nature. The tropospheric

effect on the electrical path length of the radio signals is harder to determine because the neutral atmosphere is nearly non-dispersive throughout the radio band of frequencies. This effect, typically about 8 ns at the zenith, is variable by about 1 ns, due mainly to variations in the amount of water vapor in the atmosphere. If ground-level measurements of atmospheric pressure, temperature, and dew point are used to calculate the atmospheric zenith delay, the rms error in the result may be reduced to approximately 2 to 3 cm (Murray and Marini, 1976). However, for baselines of a few kilometers or less in length, and for sites at nearly equal elevations above sea level, the atmospheric delays introduced at the two ends of the baseline tend to cancel to a high degree. Our limited experience with such short baselines shows, for example, that tropospheric effects can be lowered to the millimeter level (Rogers et al., 1978) and, in appropriate climates, to the tenth millimeter level (Elsmore and Ryle, 1976) for baselines of length up to five kilometers. For long baselines, a series of interferometric measurements by our group (Robertson et al., 1979) involving the Haystack and Owens Valley antennas, separated by nearly 4,000 km, demonstrated that a dozen baseline length determinations, distributed over a one and a half year period, show repeatability at the three centimeter level with the use of only surface measurements of atmospheric parameters. Higher accuracy results might be obtainable through use of water-vapor radiometers (see, for example, Schaper et al., 1970 and Moran and Penfield, 1976) to monitor the water-vapor content above each terminal; perhaps the contribution of the troposphere to the uncertainty in baseline-vector determination could thereby be reduced to the centimeter level in all three components even for baselines of transcontinental dimensions. However, the efficacy of water-vapor radiometers for this purpose has yet to be established reliably in VLBI experiments under various climatic conditions. Moreover, at the present stage of technology, the water-vapor radiometers would be much larger, and more expensive, than our proposed ground terminals -- the use of such radiometers would be similar to the tail wagging the dog.

One can also take advantage of averaging. Since a baseline determination can be made, on average, once per second, one can afford for most applications to average the results for many minutes or longer. Further, with more than four satellites observed simultaneously, another form of redundancy is possible. For both types, one can use the level of stability of the data and of the results as an indication of the accuracy of the baseline determination.

B. Satellite Positions

Errors in our knowledge of satellite positions are muted in their effects on baseline determination by the ratio of the baseline length to the satellite altitude. Thus for baseline lengths of a few hundred kilometers, the sensitivity to satellite position errors is reduced by a factor of nearly two hundred for satellites of the GPS type. For example, to achieve "instantaneous" accuracies of 2 cm over a 100-km baseline would require satellite position errors smaller than 4 m. (Note that for a baseline short compared to the satellite altitude, the interferometric delay is sensitive primarily to the direction of the satellite, not to its altitude). For transcontinental baselines, the immunity factor is under ten and satellite position errors become of correspondingly greater importance. But, as discussed earlier, effects of satellite position errors might be reducible to the level of the atmospheric errors if base stations were equipped to tie the satellite positions continuously over the required periods to those of extragalactic objects through differential VLBI observations. Averaging the baseline results over long periods of time, and, hence, over many satellites, would be acceptable for many applications and would tend to reduce the effects of satellite position errors.

C. Geometry

We now consider the effects on the accuracy of baseline determination of the geometric distribution of the satellites. Given the uncertainty in the measurement of interferometric phase delay, the determination of the corresponding uncertainty in the estimate of the baseline vector is non-trivial. The task is complicated largely because of the unknown, systematic, effects introduced by the neutral atmosphere. However, one pertinent question can be answered by an elementary analysis: What is the purely geometrical multiplying factor in the conversion of phase-delay uncertainty to baseline-component uncertainty? To obtain an answer, we used the following simplified expression for the interferometric phase delay obtained from observations of a satellite by two terminals (see, however, section III.A):

$$\tau_{12} \cong \frac{1}{c}\{|\vec{s}-\vec{r}_2| - |\vec{s}-\vec{r}_1|\} + \tau^{cl}$$

$$\simeq \frac{1}{c}\{(\vec{r}_2-\vec{r}_1)\cdot\hat{s} + \frac{1}{4s}[2(r_2^2-r_1^2) - (\vec{r}_2\cdot\hat{s})^2 + (\vec{r}_1\cdot\hat{s})^2]\} + \tau^{cl}; \qquad (2)$$

$$r \ll s,$$

where $\vec{r}_i$ (i=1,2) and $\vec{s}$ are vectors, from the center of the Earth to the terminals and to the satellite, respectively. The analysis was based on the GPS configuration, with two satellites in each of the three orbital planes assumed to be crossing the equator at t = 0. Baselines with $|\vec{r}_2-\vec{r}_1| \ll s$ were considered as an illustration; under these conditions, the term in brackets in Equation (2) can be ignored. However, the effects of parallax, due to the finite altitude of the satellites, must be considered and so $\hat{s}$ was calculated for the vector from the baseline midpoint to the satellite. Observations were assumed to be made of all satellites whose elevation angles, as viewed from the terminals, exceeded 10 deg.

From such observations, at each instant, the three components of the baseline vector and the epoch offset of the clock at one terminal with respect to that of the other, could be estimated and the standard errors in these estimates determined. The results from the analysis and related information on the geometry are presented in Figures 3 and 4. Figure 3 is based on the terminals being placed at a north latitude of 40 deg and at a longitude coincident at t = 0 with an ascending node of one of the orbital planes of the satellites. In Figure 3a, we show the elevation angles as functions of time for all satellites visible from the terminals. Figure 3b shows the results of the error analysis: the standard deviations in the estimates of the vertical and the two horizontal components of the baseline as functions of time. The standard deviations are given in units of the standard error in the determination of the interferometric phase delay. For either of the two horizontal components, the geometric multiplication factor is never more than 1.2 with its average value being about 0.8. The multiplication factor is larger for the vertical component whose uncertainty depends more importantly on the total spread of the elevation angles. Thus, as can be seen in the figure, the multiplication factor is relatively large when the spread is small, and vice versa. For all three components, the standard errors as functions of time are discontinuous when a satellite passes either inside or outside the allowable elevation angle limit. We also note in the figure that if the elevation angle cutoff at t = 6 hr were lowered by only 0.3 deg, the multiplication factor for the vertical component would drop nearly twofold, the total spread in the elevation angles then included being increased thereby from about 30 to 50 deg. Figure 4 contains the corresponding results for observations from terminals placed at the same longitude, but on the equator.

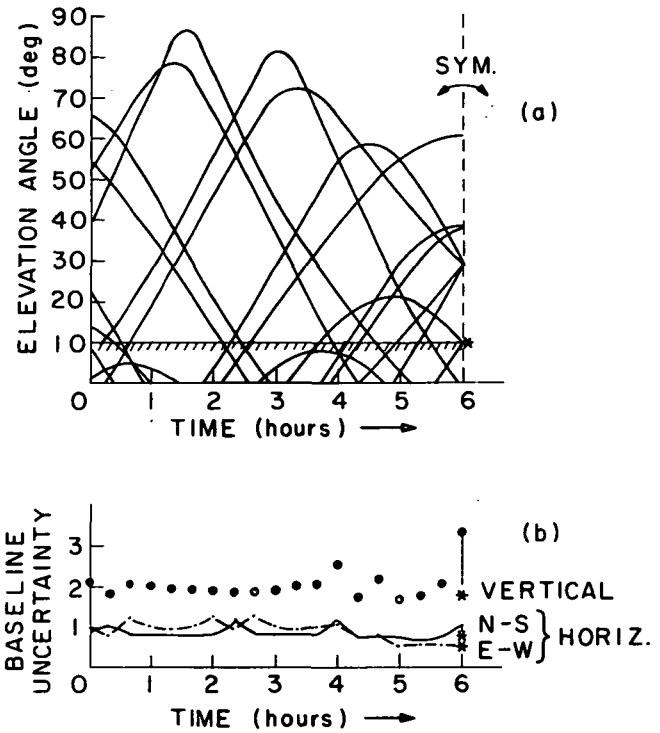



Fig. 3. Error analysis for Mighty MITES system.

(a) Elevation angles of satellites as seen from ground terminals at 40°N. latitude. Observations are made of all satellites visible above 10° elevation, except at t=6 hours, when the two satellites with elevations of 9.7°, indicated by the asterisk (*), are also observed.

(b) Standard errors of estimates of baseline vector components, computed at time-intervals of 20 minutes. Points plotted at these intervals have been connected by continuous lines for clarity, although actual "curves" have discontinuities when satellites cross 10° elevation limit. Unit of baseline uncertainty is the standard error in measurement of interferometric phase delay, converted to equivalent path length. At time=6 hours, results are shown both for 10° (regular symbols) and for 9.7° (asterisks) elevation-angle cutoffs.

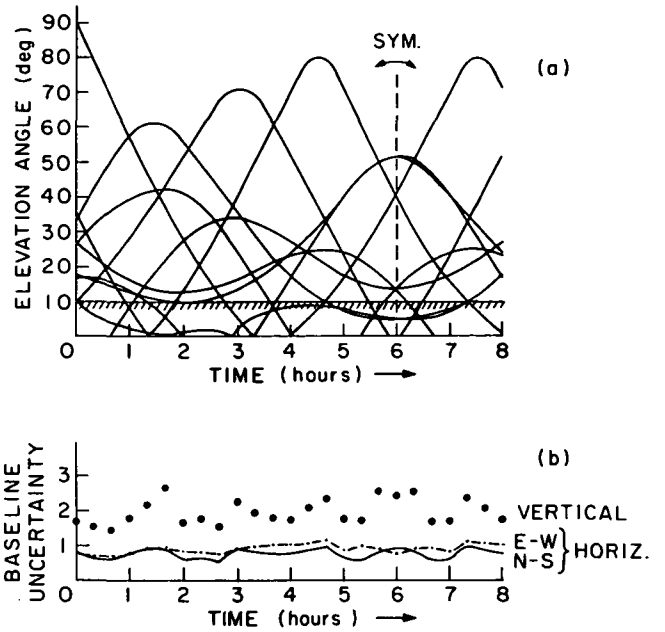



Fig. 4 (a) and (b). Same as Fig. 3, except that terminals are at the equator.

The above error analysis, as mentioned, ignores the effects of the troposphere and of the uncertainties in the orbits of the satellites. To obtain an indication of the magnitude of the effects of the troposphere, we repeated the analysis, but with the standard error for each observation equal to the cosecant of the elevation angle and with the elevation angle "cutoff" maintained at 10 deg. The results are summarized in Figure 5. Here we see that the multiplication factors are larger, as expected, reaching up to 2.0 and 2.5, respectively, for the horizontal components for terminals at 40 deg and 0 deg latitude. The corresponding factors for the vertical component are both about 5.0, which serves to emphasize the importance of the observations being widely distributed in elevation angle, and, especially, being extended to high elevation angles.

In a further refinement, we could solve in addition for the common mode epoch error (see Section III.A).

D. Instrument

For very closely spaced terminals the accuracy attainable in baseline determination could be limited by instrumental effects. For example, with a 100-meter spacing of terminals, of some interest for surveying and for monitoring the effects of earthquakes (see Section VI), the contribution to the error in baseline determination from the troposphere in good weather should be under 1 millimeter and from the satellite-position uncertainties under 0.05 millimeters. Any instrumental effects of order 1 millimeter could, under these circumstances, degrade the system performance. Thus, it would be especially important for such applications to minimize the effects of ground reflections, as discussed in Section II.C. With proper design and calibration, other sources of systematic errors in the instrument could probably be kept to a level of about 0.3 psec, equivalent to 0.1 mm in distance. The random errors, due to signal-to-noise limitations (Section II.C), could be reduced to this negligible level after less than ten minutes of integration. Should the ground reflections be the dominant contribution to the error, this fact would be disclosed in the slow variations with time of the estimate of the baseline, caused by the changing directions of the signals. These variations would also tend to repeat with a 12-hour period, as indicated earlier.

A potential problem with current geodetic applications of VLBI concerns the determination of any changes with time of the length of the signal path through the antenna system. With the large structures now in use, one must account for deformations associated with changing winds, temperature, and gravity loads (see, for discussion, Rogers et al., 1978). However, for a miniature terminal with a dipole-type antenna such changes are negligible at the millimeter level due to the small size and rigidity of the antennas.

V. Demonstration Experiment

Could the feasibility of this system be demonstrated without the use of four or more satellites? The answer is yes, and the means are quite economical, too. A transmitter could, for example, be flown on each of four, small, private planes. More planes could be used for redundancy and, in addition or in replacement, one or more transmitters could be placed at suitable ground locations. For purposes of discussion, we assume that only four planes would be used. The planes could fly at altitudes of about 2 to 3 kilometers and each could hold a tight, nearly circular, pattern of radius about 0.5 kilometers. The horizontal separation between planes could be about 5 to 6 kilometers to ensure good geometry without the use of excessively low elevation angles. The four planes could each be tracked in three dimensions by a simple set of four or more ground terminals, separated by, say, 50 to 100 meters and arranged suitably to ensure good geometrical relations. These ground terminals would be connected with phase-

stable links, as in conventional radio interferometry, to eliminate clock epoch offsets as unknowns. Thus, the three, simultaneous, independent, measurements of interferometric phase for the signals from each plane would suffice to determine the three coordinates of the plane with respect to the reference frame defined by the four ground terminals. The relative surface locations of these ground terminals could easily be determined with the use of classical techniques to within a centimeter. Two, additional, "test" terminals could be placed near the center of the array of four terminals, about 10 or more meters apart from each other to reduce interference to a negligible level, and their vector separation determined from the signals received from the transmitters on the four planes. The epoch of the clock at one terminal would have to be known quite accurately a priori, or, alternatively, a fifth plane could be employed. The accuracy of the determination of the separation of the test terminals could be checked easily to the millimeter level with the use of simple, classical, geodetic techniques. The accuracy of the VLBI determination should be at the centimeter level, or below, depending on the accuracy of the location of the planes and on the relative separation of the various terminals. The direct tropospheric effects on this accuracy will be negligible, because of the small size of the array.

As an additional test, very low elevation angles could be used and the results compared as a function of terminal separation. It might also prove useful, with perhaps higher flying craft, to test extensively the effects of the troposphere on baseline determination in various locations and weather conditions.

Either of these simulations can be used to test just about all aspects of the proposed system, save for the technique to eliminate the effects of the ionosphere on the phase-delay measurements. Although only few may see the need for an actual test of this technique, because of our firmly based knowledge of the dispersive nature of a plasma, such a test is possible to perform. For example, one could use conventional VLBI observations of extragalactic radio sources; the center frequencies of the spectral regions sampled would be matched to the frequencies of the tones proposed for the satellite transmitters. The sampled regions of the spectrum would not be so narrow as for the tones because of the necessity to attain adequate signal-to-noise ratios. However this difference is inessential. The main difficulty with this test involves the need to develop suitable feed systems for the large antennas to cover the ~1 to 2 GHz range effectively. Although observations of more than one source simultaneously will probably not be feasible in this test, the results from interleaved observations of a pair of sources can be used in much the same way as proposed for the Mighty MITES System.

## VI. Applications

There are a large number of potential applications of this interferometric system. We discuss some of these briefly.

### A. Earthquake Monitoring

A primary use of the Mighty MITES would be to monitor the regional accumulation and release of strain. Their use would be most effective in regions around a fault such as the San Andreas where the terminals could be distributed widely, and yet densely, in the immediate vicinity of the fault and the data collected routinely. In effect, a seismic array could be set up, capable of measuring displacements with millimeter to centimeter precision in the 0 to 1 Hz part of the spectrum. The extraordinary time resolution of these measurements of baseline vectors and the high accuracy, especially for short baselines, should allow a very complete geometric characterization to be made of the crustal movements during actual earthquakes. Moreover, during earthquakes, any errors in our knowledge of satellite positions, which change slowly with time, would not significantly degrade the accuracy with which the changes in baseline vectors could be determined with the terminals.* These Mighty MITES should be sufficiently inexpensive to allow a dense net to be set up, and their ability to operate unattended should also cut down dramatically on the cost of the overall monitoring system, as indicated in Section III. B.

Compact terminals could probably be designed to operate under rather extreme temperature and wind conditions; thus, such terminals could be placed at otherwise inaccessible locations, for example at the many Asian sites of geophysical interest where plates seem to be colliding and fracturing. Due to the use of low radio frequencies in our system, the signals could even penetrate through substantial snow and ice cover. However,

* Note, however, that during and immediately preceding and following earthquakes insufficient time would be available for averaging out short-term meteorological effects.

signal reflections at the various interfaces might be devastating. In addition, under heavy snow or ice conditions, it might be necessary to use, say, radioisotope power to melt the snow and ice, and to operate the receiver and the transmitter, the latter to convey the data via a suitable satellite link to a data collection and processing center. Perhaps an acceptable global environmental impact statement could be adopted that would allow use of radioisotopes for such purposes.

In any event, the terminals could be built to survive winters and operate during the remainder of the year in normal fashion with, say, solar cells providing a source of power.

B. Land Surveying

These terminals could also be used for more conventional surveying. For example, for local surveying, over distances of a few kilometers or less, two or more terminals could be employed. The surveyor would merely have to set the terminals down at the positions whose vector separations are desired and, after the minute or so required for signal acquisition, baseline results could be obtained once per second. The surveyor would also require satellite-position information, a small microprocessor, and either a radio or a line link to each terminal to be able to determine these baselines. The needed computations could easily be done in real time and would thus allow the surveyor to complete his tasks virtually as fast as he could place the terminals at the desired positions. Of course, observations could be continued by the surveyor as long as desired and the changes, or fluctuations, in the baseline vectors, as well as their running averages, could be monitored. Averaging largely removes the effects of short-term atmospheric fluctuations, caused, for example, by passing clouds.

Semi-permanent, self-contained, local arrays of terminals could also be used for a variety of purposes. As examples, we mention the monitoring of local crustal movements in the vicinity of nuclear plants, wells, pipelines, dams, mines, and rocket launching sites, and even the monitoring of oil rigs in the sea. In some cases, combination with gravity monitoring will yield a more powerful set of data for the determination of both crustal motions and changes in the mass underlying the area.

For purposes of both two-way communication and supply of power, a direct electrical link to each terminal might be more suitable for fixed installations designed for long-term monitoring. If direct links are not convenient, small cassette recorders could be used to store the data which could be collected intermittently.

C. Navigation

A number of navigational uses can be envisioned for the Mighty MITES. For example, they could be used aboard airplanes, boats, or land rovers. There should be no difficulty in the acquisition of satellite signals even aboard moving vehicles, provided the acceleration is not too high. Thus, a "one tenth g" acceleration could easily be tracked by the receiver. For significantly higher accelerations, a change in the design of the signal lock-on system might be required. The information on changes in vehicle position would also have to be transmitted to a central processor. Alternatively, the matching interferometric information on the signals received at some known, fixed site or sites could be transmitted to all vehicles from a central location and the determination of position carried out aboard each vehicle.

D. Spacecraft Tracking

The concepts involved in the Mighty MITES System could also be used in space applications. For example, with some modifications of the system to account for orbital motion, the vector separation between any two satellites could be monitored continuously provided both view the same five GPS satellites simultaneously. In fact, the GPS satellites themselves, if so equipped, could monitor their own vector separations! (However, one would have to take care in the frequency allocations to insure that the transmitted signals did not interfere with those received.) The reference system for the direction determinations would be an ensemble average; this (changing) system could be related to the extragalactic radio source reference frame through the previously-described differential VLBI measurements.

Another possibility is to use a set of these compact terminals at fixed locations to track any satellite equipped with an appropriate transmitter. With each terminal augmented by a suitable frequency standard, the accuracy achievable with this system (see, for example, Preston et al., 1972) might well be competitive with that now obtainable from laser tracking. In addition, the VLBI system would have the advantages of being all weather and having far less expensive ground equipment. (Even the transmitter may be less expensive than the satellite-borne optical corner reflectors.) Satellite programs such as SEASAT and LANDSAT, some of whose scientific results depend

importantly on the continuous determination of orbits with high accuracy, might especially benefit from use of this VLBI tracking system. The GPS satellites could, of course, also be tracked in this mode.

E. Gravity Anomalies

One possibly useful application of these concepts may be in the measurement of gravity anomalies on the Earth. Consider a pair of "drag-free" satellites separated by, say, 200 to 300 km, and traveling in an orbit of that altitude. If each satellite were equipped with an appropriately modified Mighty MITES terminal to observe the GPS signals from the GPS satellites, sensitivity would exist to anomalies of the strength of 1 mgal or greater, averaged over areas of about 200 km in linear extent. The limit on accuracy of this novel tracking system would be set by the (high-frequency) uncertainties in the knowledge of the orbits of the GPS satellites. Signal-to-noise ratios and instrumental calibrations should easily be obtained which do not limit the system sensitivity above the stated spatial and temporal resolutions. There would be no contribution to the uncertainty by the troposphere, unless the GPS satellites were replaced in whole or in part by ground terminals.

Another variant of this general scheme would be to place a tone-transmitter on each of the low-orbit satellites and to place receivers on the GPS satellites. Three-dimensional tracking of the low-orbit satellites could then also be accomplished in the GPS-defined reference frame, given suitable frequency standards aboard the GPS satellites. Again only high-frequency "noise" in the orbits of the GPS satellites would limit the sensitivity of the system to gravity anomalies.

F. Miscellaneous

One can also envision the Mighty MITES system being used for virtually instantaneous time transfer over transcontinental and intercontinental distances at about the few nanosecond level of accuracy. Here the accuracy would be limited primarily by the uncertainty in the knowledge of the positions of the GPS satellites.

Other possible applications include monitoring the motion of icebergs and, for example, ice topography in locations such as Greenland and Antarctica. Of course, during severe weather conditions, the operation of the system would be subject to the same constraints as mentioned in Section VI.A.

We should emphasize that all of these potential applications must be considered in far greater detail before their utility can be assessed properly.

## VII. Comparison with Other Space Systems

How does this VLBI system compare with other "space" systems proposed for geodetic applications? We examine briefly five such possible systems as examples: the GPS, with and without modification of the satellites; a conventional VLBI system; a satellite Doppler positioning system; a ground-based laser system; and a spaceborne laser system.

A. Global Positioning System

The various options originally considered for GPS ground terminals were not intended to yield accuracy in position determination better than about 1 meter. A crucial question is: can the already planned GPS signals be utilized in other ways to yield substantially higher accuracies in position determination? Various possibilities can be envisioned. One possibility (MacDoran, 1978) would be to receive the GPS signals interferometrically as if they were random noise, as from extragalactic radio sources. Because of the strength of the GPS signals, even when treated as random noise, a 1-meter diameter, transportable, antenna can be used with conventional VLBI receiving and data processing techniques. The estimated precision in baseline determination (MacDoran, 1978) is 2 cm from 1.5 hours of data collection for baselines $\leq$300 km in length. This system promises comparable accuracy to that claimed for the Mighty MITES system, but at the sacrifice of cost, simplicity, and time resolution.

Another possibility is to utilize knowledge of the pseudo-random noise GPS codes so that the effective noise bandwidth at the receiver is of the order of, say, ten Hertz instead of the order of ten Megahertz. One may then consider rapid determination of the interferometric group delay, at both the ~1.2 and the ~1.6 GHz GPS bands, with accuracy sufficient to determine the ionospheric contribution and to eliminate the $2\pi$ ambiguities in interferometric phase delays. The two phase delays, one at each of the two frequency bands, could then be combined to refine further the knowledge of the ionospheric contribution and to yield an accurate vacuum-equivalent phase delay. (The requirements on group-delay accuracy are somewhat ameliorated for sites closely-enough spaced for the ionospheric effects on the interferometric observable to be negligible.) If this technique were viable, then the same accuracy in base-

line determination could be achieved as with the Mighty MITES system. However, one important caveat which bears on the cost must be considered: the achievement of sufficient group delay accuracy to eliminate the phase delay ambiguity may entail considerable complication with such a narrow (≈1 %) fractional bandwidth available at each GPS band. Although signal-to-noise ratios appear sufficient, systematic effects on the frequency dependence of the phase delays could be serious. As an oversimplified illustration, consider ground reflections of the signal that introduce an (erroneous) contribution to the slope of the phase of the signal as a function of frequency across the band. A difference in phase of only 1°, in opposite directions, at both ends of a 10 MHz band, would introduce an error in group delay equivalent to a displacement of about 15 cm -- a large fraction of the ~20 cm wavelength at 1.6 GHz. To be more quantitative about the possible sources of such reflected signals, we note that the root-mean-square(rms) phase error due to, say, isotropic scattering by objects near the antenna is given approximately by

$$\langle\phi^2\rangle^{1/2} \simeq 10\ \frac{\sigma^{1/2}}{R}\text{deg},$$

where σ is the cross section of the object, R is its distance from the antenna, and both are measured in compatible units. We assume here that the receiving antenna is isotropic. Thus, an object such as a tree with a 2 $m^2$ effective cross section, situated 15 m from the receiving antenna, would lead to a 1 deg rms phase error. To cut this figure to 0.1 deg would require the gain of the antenna to be 20 db less in the direction of the scattered radiation than in the direction of the directly received signals. A simple antenna with such high directivity could not receive signals from all satellites simultaneously; hence the effects of instabilities of the frequency standard could not be eliminated simply by subtraction, as in the Mighty MITES system. A more stable, and presumably more expensive, standard would be needed. Alternatively, one could employ a phased-array antenna in either of two modes. In the simpler, but still rather elaborate mode, the pointing of this antenna could be cycled from satellite to satellite with the phase-locked loops that track the signals from any given satellite being gated in synchronism with this cycle. (The period of the cycle should be short compared with the time constants of the phase-locked loops.) In the second mode, the array could observe all visible satellites simultaneously at the cost of multiple amplifiers and phase shifters for combining appropriately the outputs from the different elements of the array. We conclude that this problem of elimination of the 2π ambiguity in phase with the planned GPS signals may require a rather expansive system to insure sufficient suppression of the effects of ground reflections (see also Section II.C).

In regard to systems, like ours, that involve use of modified GPS signals, one might enquire about non-interferometric schemes as well. We considered several such possibilities, for example a system in which the terminals received and transponded a ranging code to each satellite. None of these schemes looked as attractive as our interferometric system. Each had a serious flaw that seemed very costly to overcome.

### B. Conventional VLBI System

A relatively large, transportable, conventional VLBI system that is used to observe extragalactic radio sources has several advantages over our compact terminal. First, since the observed signals are from extragalactic sources whose positions are virtually static, no orbit determinations are required and loss of accuracy in baseline determination with increasing baseline length is much more modest; second, the addition of water-vapor radiometers to allow possible achievement of higher accuracies in the estimate of tropospheric effects, is a relatively smaller complication for the larger system. The disadvantages of the larger VLBI system are partly economic: compact units, operating unattended, should prove far more cost effective with virtually unlimited numbers of such terminals usable simultaneously; and part technical: the temporal resolution achievable with the compact units should be at least three orders of magnitude better than with the larger system. Furthermore, determinations of baselines between all sites of interest simultaneously, which would be feasible only with compact terminals, would obviate the need for knowledge of polar motion and earth-rotation variations to "tie" all inter-site vectors to a common reference system.

### C. Satellite Doppler System

The satellite Doppler system, which is used to determine positions on Earth from the Doppler shift of signals transmitted by satellites, has one distinct advantage over any VLBI system: each ground unit can be used independently to determine the position of any point on the Earth's surface in a suitably defined

Earth-fixed reference system. However, this system has the apparent disadvantage that much more time of the order of many hours and perhaps days, may be required to determine the position of the receiver with accuracy comparable to that achievable between sites with VLBI; in fact, the attainable accuracy in relative position determination may still be several-fold worse than that attainable with a VLBI system.

D. Ground-Based Laser System

Another system that could be used for monitoring crustal motions entails the use of highly-transportable lasers to observe the LAGEOS satellite from various sites (Bender et al., 1978). It is estimated that each such laser system would require several days per site to obtain a vector position with uncertainty at the centimeter level. Although the accuracy achievable with this system may be comparable to that expected with the proposed Mighty MITES system, the cost of each laser system will probably be more than a hundredfold greater and will require, on average, about a thousandfold longer time per position determination. The laser system would therefore suffer, relative to the Mighty MITES system, from the same difficulties discussed in connection with the conventional VLBI system. The main advantage of this laser system, as with every optical system, is its relative immunity to the effects of the water vapor in the troposphere (see Subsection E). Of course, optical systems are also freed from the complications entailed in the removal of ionospheric effects.

E. Spaceborne Laser System

The spaceborne laser system, which would measure the echo delays of signals reflected from arrays of optical corner cubes on the ground (Smith, 1978), has one main advantage with respect to the VLBI system: The laser signals are about 40 times less sensitive than radio signals to the effects of atmospheric water vapor on the propagation times of the signals. The Mighty MITES system, on the other hand, has several distinct advantages over the laser system. First, with the Mighty MITES, baselines could be monitored continuously whereas, with lasers, the monitoring would be only intermittent, due both to weather interruption and to lack of continuous satellite coverage; measurements coincident with earthquakes, for example, might therefore be missed with the laser system. Second, the Mighty MITES system could be used effectively for long as well as for short baselines (although at

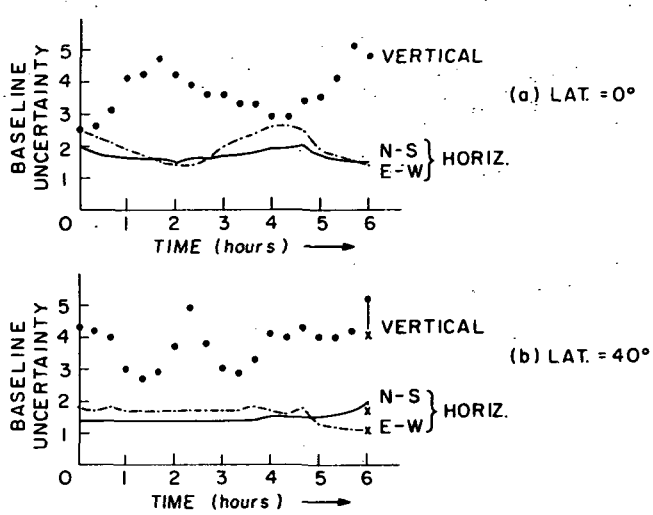



Fig. 5 (a) and (b). Same as Figs. 4 (b) and 3 (b), respectively, except that here the standard error in the measurement of the interferometric phase delay, instead of being unity, is equal to the cosecant of the elevation angle of the satellite.

greater expense because of the expected need to monitor the satellite orbits with larger VLBI systems), whereas the laser system, which may require a low orbit for the laser, would probably be considerably less effective in the determination of transcontinental baselines. Third, the spaceborne part of the Mighty MITES system is well within the state of the art and should have a lifetime of at least several years. By contrast, a laser system which can transmit pulses unattended and unmaintained at the required rates for such a period has yet to be demonstrated. However, only one satellite at a time is required for the spaceborne laser system whereas many are required for the Mighty MITES system. (On the other hand, the requirements for pointing are far more severe for the laser than for the radio system.) With regard to the lifetime of the ground-based parts of the laser and radio systems, it is not clear which has the advantage: one must balance the maintenance of the electronic equipment, which may be capable of unattended performance for several years, against the erosive and obscuring effects on the corner cubes of wind-blown dirt and rainborne dust.

The Mighty MITES receiver and electronics package could be placed underground, with only the antenna exposed to view. It is thus unclear whether corner cubes, which would be totally exposed above ground, or the radio terminals would be a greater attraction for souvenir hunters.

References


Bender, P. L., et al., Tectonophysics (in press), 1978.

Counselman, C. C., et al., Science, 178, 607, 1972.

Counselman, C. C., and H. Hinteregger, Proc. I.E.E.E., 61, 478, 1973.

Counselman, C. C., Annual Review of Astronomy and Astrophysics, 14, 197, 1976.

Elsmore, B., and M. Ryle, Mon. Not. Roy. Astron. Soc., 174, 411, 1976.

International Telecommunications Union, "Radio Regulations" (International Telecommunications Union, Geneva, 1968).

King, R. W., et al., J. Geophys. Res., 81, 6251, 1976.

MacDoran, P. F., presented at the International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics, Lagonissi, Greece, 29 May - 6 June, 1978.

Moran, J. M., and H. Penfield, Final Report, Contract NAS5-20975, 1976.

Murray, C. W., and J. W. Marini, GSFC Technical Memorandum, 1976.

Parkinson, B. W., "NAVSTAR Global Positioning Systems (GPS)", presented at the 1976 National Telecommunications Conference in Dallas, Texas (I.E.E.E. Catalog Number 76 CH 1149-4 CSCB).

Preston, R., et al., Science, 178, 407, 1972.

Robertson, D. S., et al., Recent Results of Radio Interferometric Determinations of a Transcontinental Baseline, Polar Motion, and Earth Rotation, in Proc. of IAU Symposium No. 82 (in press), 1979.

Rogers, A. E. E., Proc. I.E.E.E., 59, 1617, 1971.

Rogers, A. E. E., et al., J. Geophys. Res., 83, 325, 1978.

Schaper, L. W., et al., Proc. I.E.E.E., 58, 272, 1970.

Smith, D. E., these Proceedings, 1978.

Whitney, A. R., et al., Radio Science, 11, 421, 1976.

Page Intentionally Left Blank

# Monitoring Regional Crustal Deformation with Horizontal Geodetic Data

R. A. Snay and J. G. Gergen
National Geodetic Survey
National Ocean Survey, NOAA
Rockville, Maryland 20852

Abstract. The National Ocean Survey is developing an automated system to derive parameters of horizontal crustal motion from existing geodetic data by the process of least-squares estimation. The estimated parameters will describe crustal motion as a function of geographic position. The system will first be tested in the Imperial Valley region of southern California, using data from 8 individual field projects spanning four decades of time.

## Introduction

Global models for tectonic activity hypothesize the existence of rigid plates rotating with constant velocity. In contrast, local crustal motion as observed by geodetic and geophysical instrumentation varies from nearly continuous creep to the stop and go process associated with large earthquakes. To better understand the transition from global to local phenomena, the National Geodetic Survey of the National Ocean Survey is performing several studies of the geodetic data of the past 100 years on a regional level. These studies are designed to establish the pattern of horizontal crustal motion in areas from 100 to 300 km in diameter. This paper describes one of the techniques being used and some preliminary results obtained from a pilot study of the Imperial Valley area in southern California.

The participation of the National Geodetic Survey in crustal motion study is required by the forthcoming redefinition of the North American Horizontal Datum. A new adjustment of the entire U.S. control network will accompany this redefinition, and new positions will be published in 1983 for all stations of the control network. The following arguments are presented to support the geodetic community's need for a better understanding of crustal motion.

1. In any adjustment which incorporates data from different epochs, especially in an area of crustal movement, the observations need to be reduced to a model of the earth which allows geodetic positions to vary with time.
2. A model for crustal motion is needed for predicting the changes in position of published stations. To the extent feasible, parameters of motion could be published in 1983 along with station positions in a fashion similar to star catalogs.
3. A better understanding of crustal motion will help to better define requirements for reobserving disrupted sections of the control network.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280. The Ohio State Univ., Columbus, Ohio 43210.

## The Technique

In most studies of crustal motion the usual technique is to compare different surveys of the same geodetic network, two epochs at a time. The technique to be discussed here will differ out of necessity. The geographic areas of study will generally be larger than the area covered by any one field project, and in most cases the various field projects will only partially overlap one another since most of them were observed to establish geodetic control where it previously did not exist, not for crustal motion study. The basic technique is to estimate parameters describing crustal motion by a simultaneous least-squares adjustment of all the pertinent geodetic data. This is accomplished by introducing into the adjustment a mathematical model which describes station positions as a function of time. The following paragraph describes the model which was used in the Imperial Valley pilot test.

In the model the region of study can consist of one or more subregions. Existing fault lines will usually provide the boundaries between subregions. The latitude $\phi_t$ and the longitude $\lambda_t$ of a geodetic station in the $i^{th}$ subregion at time t are given by the formulas.

$$\phi_t = \phi_{t_o} + f_{i,1}\,(t-t_o) + f_{i,3}\,(t-t_o)^2$$

$$\lambda_t = \lambda_{t_o} + f_{i,2}\,(t-t_o) + f_{i,4}\,(t-t_o)^2$$

Here $t_o$ is a fixed time of reference, and $(\phi_{t_o},\lambda_{t_o})$ are the geodetic coordinates of the station at time $t_o$. Each $f_{i,j}$ for $1\le j\le 4$ is a function over the variables $\phi_{t_o}$ and $\lambda_{t_o}$. In the first applications of the technique these functions will be of the form

$$f_{i,j}(\phi_{t_o},\lambda_{t_o}) = b_{i,j,1} + b_{i,j,2}\,\phi_{t_o} + b_{i,j,3}\lambda_{t_o} + b_{i,j,4}\phi^2_{t_o} + \cdots$$

Note that this models the motion as a continuous function of time and a discontinuous function of position with the discontinuities occurring along the boundaries between subregions. Existing horizontal survey data in the form of directions, distances, and azimuths can be input into the adjustment process to obtain the least-squares estimates for the unknown coordinates $(\phi_{t_o},\lambda_{t_o})$ and the unknown coefficients $b_{i,j,k}$.

This technique has several advantages over the standard technique of directly comparing two sets of measurements of the same quantities. It allows for the linking together of neighboring field projects into a single data set even though several years might exist between the times when the individual field projects were observed. It

allows for the rigorous inclusion of astronomic azimuths which may have been observed separately and the inclusion of data at stations which have been destroyed. Additionally, the model provides a built-in mechanism for interpolating the values of velocity and acceleration over the entire region of study. Finally, the model generalizes the information contained in the data. This last point can be considered a disadvantage as well as an advantage. It is an advantage in so much as the concern is toward regional trends as opposed to local details. For example, local movement phenomena like hillside creep will be smoothed-out. On the other hand, it is a disadvantage in that unmodeled variations in regional motion will also be smoothed-out. For example, motion across an unmodeled, yet active, fault will be interpreted as a continuous function of position. For this reason the standard technique of directly comparing two sets of measurements over the same quantities will never be fully abandoned. Instead it will provide the standards by which to evaluate the accuracy of the mathematical model.

Although information is lost in the process of modeling the motion, a model is desired to provide a clear overall picture. When the inadequacies of a model are identified, the model can be refined. The study of the Imperial Valley data was conducted as a pilot test to evaluate the above model, identify its inadequacies, and suggest how the model might be changed to more accurately reflect the physical situation.

## The Data

Of the seismic areas in the United States, the Imperial Valley and its immediate surroundings represent the most frequently observed part of the national network over any geographic area of comparable size. The basic network was observed in 1934. After the El Centro earthquake of May, 1940 (Richter scale magnitude = 7.1) the network was reobserved in 1941 to determine a new set of positions for several geodetic stations. Observations of the basic network were performed in 1954-55 and again in 1967 for the specific purpose of studying post-seismic activity in the area. In the period 1974-76 two field projects along the southern extent of the network were observed as part of the Transcontinental Traverse. This study also includes two minor field projects in the area, a 1942 triangulation survey around the southern half of the Salton Sea and a 1959 highway traverse survey extending parallel and about 10 km north of the Mexico-California border. Figure 1 shows the essential part of the network which was studied and its relationship to the fault system in the area.

Different parts of this data were previously investigated. The results of these investigations can be used to evaluate the performance of the model. Displacement vectors for the region were reported by Meade [1948] for the 1934-1941 period, Whitten [1956] for the 1941-55 period, and Gergen [1978] for the 1941-1976 period. Miller et al. [1970] published dis-

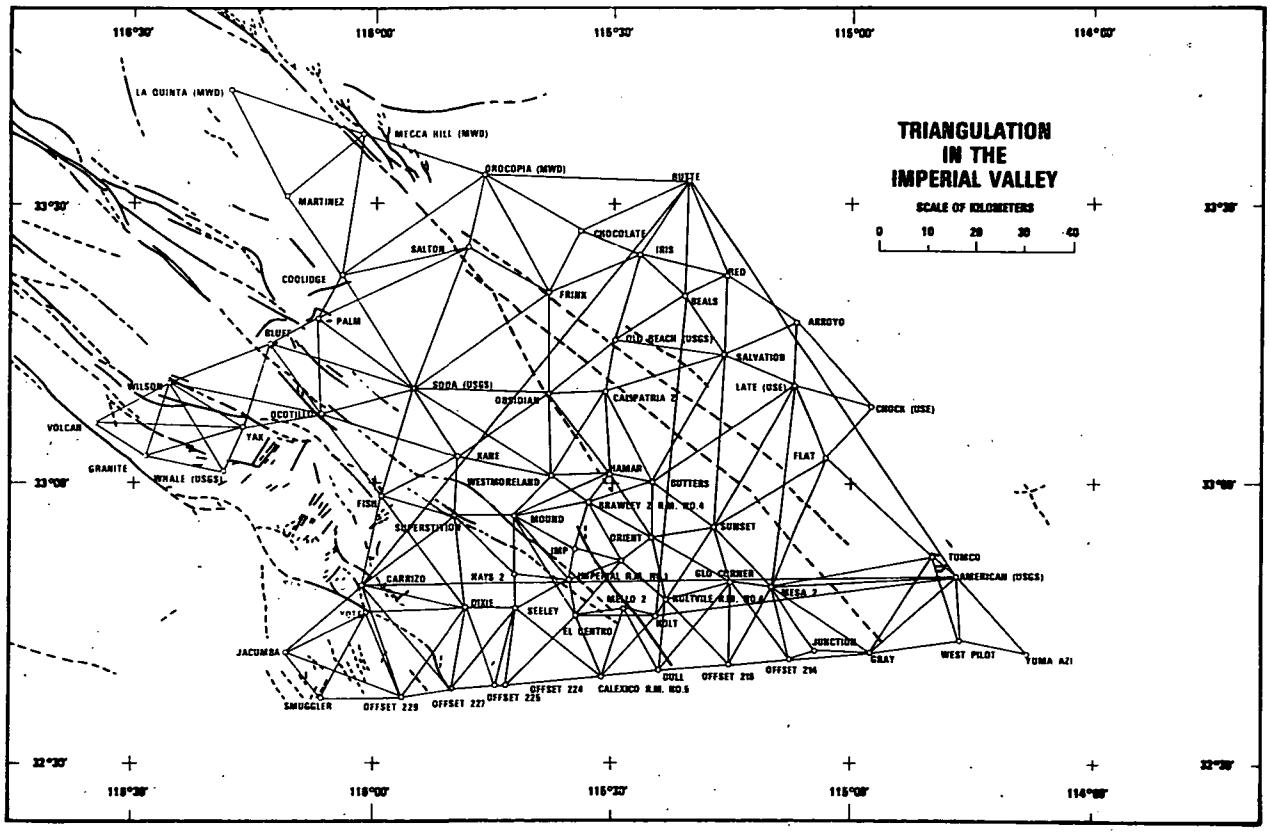



Fig. 1. Triangulation in the Imperial Valley and its relationship to the fault system.

placement vectors and strain components for various periods using appropriate subsets of the 1934-41-55-67 data. A geophysical interpretation of the data was given by Scholz and Fitch [1969] by comparing the 1941-55 data to the dislocation model of Chinnery [1961]. Interpretations of the data were also performed by Savage and Burford [1970], Barker [1976], and Thatcher [1978]. Each of these last three papers analyzed the strain components derived from the 1941-55-67 data by the method of Frank [1966].

## Experiments with the Model

The model as programmed for this pilot test allows the user to partition the region of study into three subregions and to solve for 15 coefficients (all terms in $\phi$ and $\lambda$ up to degree four) for each of the four polynomials associated with a subregion. This gives the user a total of 180 parameters with which to describe the motion. The experiments are to determine the appropriateness of these parameters. However, it was first necessary to establish the location of active faults. This was accomplished with geological maps and an adjustment of the data to a model which does not include any time parameters. The geological maps located the known faults. The residuals obtained from the adjustment identified which of these faults were the most active. In some instances the maps were ambiguous as to the location of a station relative to the fault. Station B in figure 2 illustrates the problem. These ambiguities were resolved by assuming one sense for the relative motion between opposite sides of the fault and checking whether the angle $\alpha$ at B measured clockwise from A to C is increasing or decreasing with time. In figure 2 right-lateral motion is assumed. Thus, if B were to the left of the fault, then the angle $\alpha$ would decrease in size with time. If B were to the right of the fault, the angle would increase. Note that the situation is reversed if left-lateral motion is assumed.

Once the location of the active faults were incorporated into the model, the data was readjusted several times. The first readjustment revealed an inadequacy of the model in that it could not accommodate large discontinuities in motion as a function of time such as those which occurred along the Imperial fault as a result of the 1940 El Centro earthquake. To continue with the test all pre-1940 data was removed from the data set except for a 1935 astronomic azimuth near station YOTE (see figure 1). This azimuth was retained for better orientation control over time. In retaining the azimuth it is assumed that the distance between the location of this observation and that of the earthquake is sufficiently large that the orientation of the observed line did not change discontinuously at the time of the earthquake. Future mathematical models need a feature to accommodate the relatively instantaneous shifts in position associated with major earthquakes.

With the pre-1940 data removed, the remaining data revealed a velocity pattern corresponding to a general shrinking of the network, indicating a problem with scale. The 1959 highway traverse, the 1967 observations, and the two Transcontinental Traverse projects 1974-76 all have sufficient electronic distance measurements to render good scale control for these epochs. However, closer examination of the 27 observed distances of 1959 indicated that they were too long by an estimated 13 parts per million. These observations were accordingly rescaled for the purpose of this test. Further investigation of these distances is being undertaken.

An adjustment with the rescaled distances revealed a velocity pattern corresponding to a rotation about the fixed station, indicating an orientation problem. The data includes 13 astronomic azimuth observations, one observed in 1935, one in 1967, and the remaining 11 as part of the two Transcontinental Traverse projects 1974-76. These azimuths are not suspected to contain any serious non-random error. Instead the results indicate a case of modeling observational errors as movement. The model is a second degree polynomial in time and the above azimuths essentially represent three epochs, i.e., three points on the graph of network orientation versus time, the minimum required to determine a second degree polynomial. Hence all error in these three epochs of orientation is absorbed into the model. For the first two epochs, 1935 and 1967, the orientation is determined by a single observation. It is not unreasonable for an astronomic azimuth to be in error by one arc second which corresponds to 0.485 meters at distance of 100 km from the fixed station. Further evidence of this effect was revealed by the high correlation coefficients between the estimated parameters which are linear in time and the corresponding parameters which involve the second power of time. Consequently, the data does not allow for the solution of an acceleration term. One way to overcome this weakness in the data would be to enlarge the network so as to include additional orientation control from nearby projects. There is a limit, however, to the effectiveness

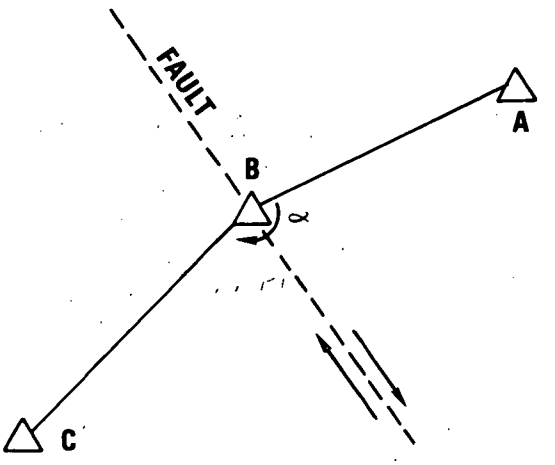



Fig. 2. The position of station B relative to the fault can be established by the direction of change in the angle $\alpha$ with time.

of this technique. The more distant the observations are from the location of interest, the smaller is their relative information content.

With the elimination of all terms involving the second power of time, a solution involving 16 time parameters was obtained. In this solution the polynomial in the $i^{th}$ area is of the form

$$f_{i,j}(\phi_{t_o},\lambda_{t_o}) = b_{i,j,1} + b_{i,j,2}(\phi_{t_o}-\bar{\phi}) + b_{i,j,3}(\lambda_{t_o}-\bar{\lambda})$$

where $1\leq i\leq 3$, $1\leq j\leq 2$ and $(\bar{\phi},\bar{\lambda})$ are the coordinates assigned to the station ORIENT. The constraints $b_{1,1,1} = b_{1,2,1} = 0$ were imposed. This corresponds to the assumption that station ORIENT in subregion 1 did not move with time. The estimated values for several parameters in this solution were below the estimated values of their standard errors. Hence, additional constraints need to be imposed to compensate for the inadequacies of the data. Some experimental adjustments were performed constraining different combinations of weakly determined parameters to specific values. Figure 3 illustrates the velocity vectors relative to station ORIENT obtained in one of these experiments. Here the five constraints $b_{1,2,2} = b_{2,1,1} = b_{3,1,3} = b_{3,2,1} = b_{3,2,3} = 0$ were imposed in addition to fixing station ORIENT. The heavy wavy lines in figure 3 correspond the subregion boundaries input to the solution and dividing the region into three subregions. The error ellipses in figure 3 indicate the 95% confidence limits for the velocity vectors. Note that error ellipses are relative to the origin and depend on the choice of constraints.

Statistical analysis in the form of an F-test indicate that the 11 parameter solution of figure 3 is overconstrained relative to the 16 parameter solution at the 0.01 significance level. A few more adjustments were attempted to find the optimum set of constraints utilizing the statistical concept of fixing a parameter whenever there is insufficient information to significantly estimate its value. Sometimes more realistic constraints can be derived from the physical theory itself. Both the 16 and 11 parameter solutions result in nonsymmetric strain matrices for each subregion. Physically this corresponds to a rotation of the network with time. The average rotation of the network obtained from these solutions is of the order of 0.1 ($10^{-6}$) radians/yr. Since the estimated standard error in astronomic azimuths is 5.3 ($10^{-6}$) radians [Strange and Pettey, 1977], this rotation is probably only noise in the 13 observed azimuths. If it is physically plausible that the overall network does not rotate with

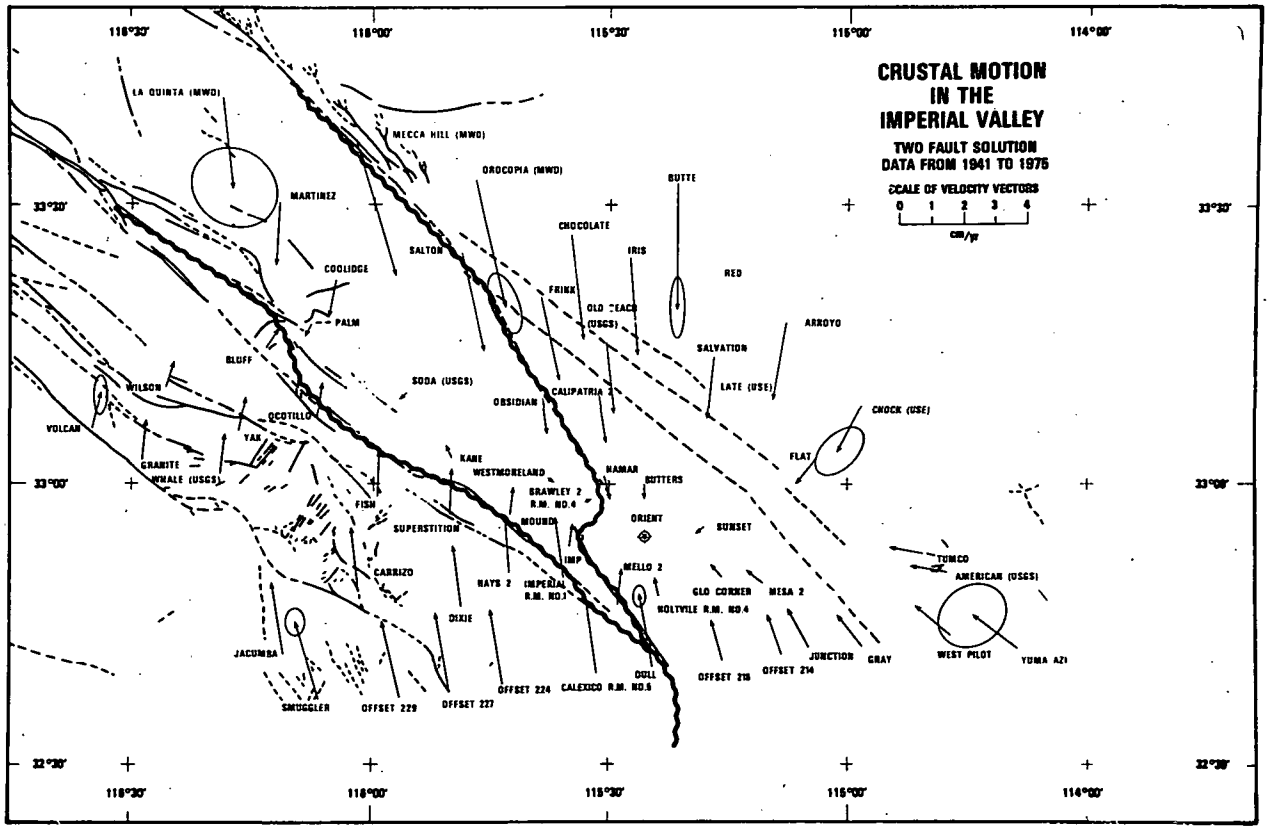



Fig. 3. The velocity vectors relative to station ORIENT as obtained from an 11 parameter solution of the 1941-76 data. The ellipses are the 95% confidence limits, and the heavy wavy lines represent the subregional boundaries supplied in the solution.

time, then a more realistic type of constraint is to restrict some or all of the strain matrices to a symmetric form. Experiments with this type of constraint could not be performed in time to include them in this report.

## Evaluation of the Model

The results of the various solutions translate into north-south contraction in the Imperial Valley in addition to right-lateral strike slip motion along the faults. The interpretation of north-south contraction was at first questioned as a possible result of poor scale control in the data. However, the same interpretation was obtained [Savage et al., 1978] from trilateration networks observed by the U.S. Geological Survey between 1972 and 1978. The data used by Savage et al. and the data of this study have no observations in common. Possible corroboration of the north-south contraction hypothesis is also provided by the observed subsidence in the area to the southeast of the Salton Sea (vicinity of station CALIPATRIA in figure 1) relative to the area just north of the Mexico-California boundary [Reese, 1977]. This subsidence is based on three epochs of leveling data spanning the period 1972-76.

To further check the accuracy of the model, the 16 parameter solution is compared in table 1 to the results obtained by Miller et al. [1970] and Savage et al. [1978]. In general the velocity vectors of this study have a more north-south trend than those of the other two studies. However, the difference is not statistically significant. Since a symmetric strain matrix was

TABLE 1. Comparison of velocities relative to OLD BEACH deduced from 1941-67 triangulation [Miller et al., 1970], 1972-78 trilateration [Savage et al., 1978], and the 1941-76 triangulation and trilateration of this study.

| Station | Epoch | $\mu_1$ (East) (mm/yr) | $\mu_2$ (North) (mm/yr) |
|---|---|---|---|
| BUTTE | 41-67 | 16 | 2 |
| | 72-78 | 3 ± 9 | -2 ± 12 |
| | 41-76 | 2 ± 8 | -15 ± 5 |
| OROCOPIA | 41-67 | 1 | 1 |
| | 72-78 | 6 ± 4 | -4 ± 4 |
| | 41-76 | 4 ± 9 | -8 ± 8 |
| ALAMO | 41-67 | 0 | 3 |
| | 72-78 | 1 ± 2 | -1 ± 3 |
| | 41-76 | 0 ± 3 | 5 ± 3 |
| SODA | 41-67 | -16 | 9 |
| | 72-78 | -3 ± 3 | 13 ± 4 |
| | 41-76 | -10 ± 5 | 18 ± 10 |
| KANE | 41-67 | -20 | 19 |
| | 72-78 | -7 ± 5 | 16 ± 4 |
| | 41-76 | -10 ± 7 | 23 ± 8 |
| FISH | 41-67 | -15 | 26 |
| | 72-78 | -12 ± 5 | 24 ± 4 |
| | 41-76 | -11 ± 10 | 33 ± 12 |
| DIXIE | 41-67 | -11 | 33 |
| | 72-78 | -22 ± 9 | 27 ± 4 |
| | 41-76 | -13 ± 14 | 35 ± 9 |
| OFFSET 225 | 41-67 | -5 | 43 |
| | 72-78 | -31 ± 14 | 22 ± 10 |
| | 41-76 | -15 ± 18 | 38 ± 9 |
| CARRIZO | 41-67 | -15 | 33 |
| | 72-78 | -19 ± 8 | 29 ± 5 |
| | 41-76 | -13 ± 14 | 36 ± 13 |
| OFFSET 229 | 41-67 | -16 | 35 |
| | 72-78 | -24 ± 14 | 33 ± 5 |
| | 41-76 | -16 ± 19 | 40 ± 12 |

assumed by Savage et al., the standard errors associated with their values are in general smaller than those of this study.

In table 1 the numbers representing the solution of this study would change by varying the number of parameters which are estimated. Some variations of the model that deserve investigation are the inclusion of the Sand Hill fault which runs from station GRAY northwesterly to station FRINK (figure 1), and the inclusion of a fault line in the vicinity of station OFFSET 227. In addition to varying the allowable coefficients of this model, it is desirable to refine the entire mathematical model so that the estimated parameters correspond closer to physically observable quantities like the elasticity of the crust or the depth of faulting. On the other hand, even if the model were perfect the results of this study could differ from the results of Miller and Savage. The results of Miller are based only on the 1941 and 1967 networks with the assumption that three stations were fixed in time. The results of Savage represent different data over a significantly shorter time span and were obtained by assuming one station and one azimuth fixed in time. Finally, recall that the model is unable to extract acceleration information from this particular data set. Thus it is impossible to check Thatcher's [1978] result that the average velocity across the extent of the fault zone decreased from 82 ± 11 mm/yr for the 1941-54 period to 23 ± 15 mm/yr for the 1954-67 period.

## Conclusion

The pilot test in the Imperial Valley demonstrated the advantages of fitting a model to the data. In particular, data from several sources can be assimilated. Observations of scale, orientation, and triangulation which could not be used directly by the technique of comparing two sets of observations over the same quantities have been included in a single data set. In the same manner the model will allow for the merger of classical geodetic observations with data derived from radio interferometry, creepmeters, Doppler, and satellite lasar-ranging observations. However, before embarking on such an ambitious project, a model is sought which corresponds more closely to physical reality. This pilot test was a preliminary step in constructing such a model. It provides a departing point for future models and it reveals to some extent the information content of classical geodetic data.

## References


Barker, T. G., 1976: Quasi-static motions near the San Andreas fault zone. Geophys. J. Royal Astron. Soc., 45, 689-705.

Chinnery, M. A., 1961: The deformation of the ground around surface faults. Bull. Seismol. Soc. Amer., 51, 355-372.

Frank, F. C., 1966: The deduction of earth strains from survey data. Bull. Seismol. Soc. Amer., 56, 35-42.

Gergen, J. G., 1978: Horizontal displacements in the earth's crust in the vicinity of El Centro, California (abstract). EOS Trans. Am. Geophys. Union, 59, 242.

Meade, B. K., 1948: Earthquake investigation in the vicinity of El Centro, California; horizontal movement. Trans. Amer. Geophys. Union, 29, 27-31.

Miller, R. W., A. J. Pope, H. S. Stettner and J. L. David, 1970: Crustal movement investigations - triangulation; Imperial Valley. Operational Data Report CGS DR-10, Coast and Geodetic Survey, U.S. Dept. of Commerce, Rockville, Md.

Reese, S. M., Jr., 1977: Preliminary results of the 1976-77 Imperial Valley leveling surveys (internal report), National Geodetic Survey, Rockville, Md.

Savage, J. C., and R. O. Burford, 1970: Accumulation of tectonic strain in California. Bull. Seismol. Soc. Amer., 60, 1877-1896.

Savage, J. C., W. H. Prescott, M. Lisowki and N. King, 1978: Deformation across the Salton Trough, California, 1973-77 (submitted to J. Geophys. Res.).

Scholz, C. H., and T. J. Fitch, 1969: Strain accumulation along the San Andreas Fault. J. Geophys. Res., 74, 6649-6666.

Strange, W. E., and J. E. Pettey, 1977: The new adjustment of the North American datum: geodetic astronomy, Am. Congress of Surveying and Mapping Bulletin, 57, 19.

Thatcher, W., 1978: Horizontal crustal deformation from historic geodetic measurements in southern California (submitted to J. Geophys. Res.).

Whitten, C. A., 1956: Crustal movement in California and Nevada. Trans. Amer. Geophys. Union, 37, 393-398.

# Strain Patterns and Strain Accumulation Along Plate Margins

J. C. Savage
U.S. Geological Survey, 345 Middlefield Road
Menlo Park, California 94025

Abstract. Observations of strain accumulation along plate margins in Japan, New Zealand, and the United States indicate that: 1) a typical maximum rate of secular strain accumulation is on the order of 0.3 ppm/a, 2) a substantial part of the strain accumulation process can be attributed to slip at depth on the major plate boundary faults, and 3) some plastic deformation in a zone 100 km or more in width is apparently involved in the strain accumulation process.

## Introduction

Repeated geodetic surveys along plate margins show gradual changes imposed by the motion of plates as well as abrupt changes occurring at the time of great earthquakes. Secular changes in angles within a triangulation network may amount to 6 arcseconds/century or more, and 10-km lines in a trilateration network may change length at a rate of 0.3 m/century or more. Probably the best way to display the changes observed in geodetic networks is by means of the inferred strain field. The strain field is preferred to the displacement field for two reasons: 1) Strain is calculated from the local changes and does not involve accumulation and propagation of errors across the network. 2) The displacement field is generally ambiguous to the extent that the relative translation and rotation of the two surveys is uncertain. The general procedure for calculating strain is to treat a network as a whole or some subsection of it as subject to uniform strain, and then to find the uniform strain field that best accounts for the observed changes in angles (triangulation) or length (trilateration). Because scale in triangulation is somewhat less certain than angles, it is probably best to calculate only the shear components $\gamma_1 = e_{xx} - e_{yy}$ and $\gamma_2 = 2e_{xy}$ directly from the angle changes rather than attempt to calculate the complete surface strain field. (Notice that the shear components are given in engineering shear, twice the tensor shear.) Schemes for calculating strain from triangulation data have been described by Frank (1966) and Prescott (1976). For trilateration data a scheme similar to that used by Scholz and Fitch (1969) is recommended. The data should be sufficiently redundant such that not only can the strain components be determined but also reasonable estimates of the standard errors in those components.

The mode of strain accumulation along a plate boundary depends to a large extent upon the nature of the plate boundary. In the simplest model the plate boundary is thought of as a zone several hundred kilometers wide that accommodates relative plate motion by continuously distributed deformation (Figure 1). At a convergent boundary

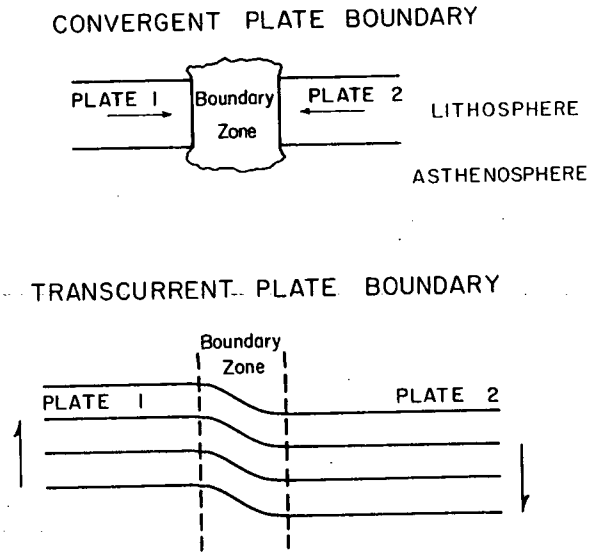



Fig. 1. Plastic boundary zone at the plate margin. The upper figure shows a vertical section across a convergent boundary, and the lower figure shows a plan view of a transcurrent boundary.

this zone would be squeezed horizontally and perhaps thickened vertically, whereas at a transcurrent boundary the zone would undergo distributed shear parallel to the boundary. More elaborate representations of plate-boundary accommodation generally involve gradual slip at depth on great plate-boundary faults with occasional abrupt seismic slip on the upper reaches of the fault. Such representations are based upon dislocation models of faulting (Chinnery, 1961; Freund and Barnett, 1976). A dislocation model for a convergent plate boundary is shown in Figure 2 where the strain released by a major thrust earthquake at a plate boundary is shown. The strain release in that figure is calculated for a constant reverse slip of 1 m over the entire width W of a two-dimensional fault (infinite length perpendicular to the plane of the paper in Figure 2). Because the actual slip on a fault is not constant but presumably varies smoothly, the actual strain profile would be a somewhat smoothed version of the profile in Figure 2. For this reason the short interval of contraction (negative strain) in Figure 2 probably would not be observed. Because W is typically 100 km or so and the slip in a major earthquake perhaps 5 m, the strain release may exceed 20 μstrain over distances of several hundred kilometers. In the period between great earthquakes, an amount of strain equal to that released coseismically must accumulate. Thus, the strain accumulation rate should on the

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

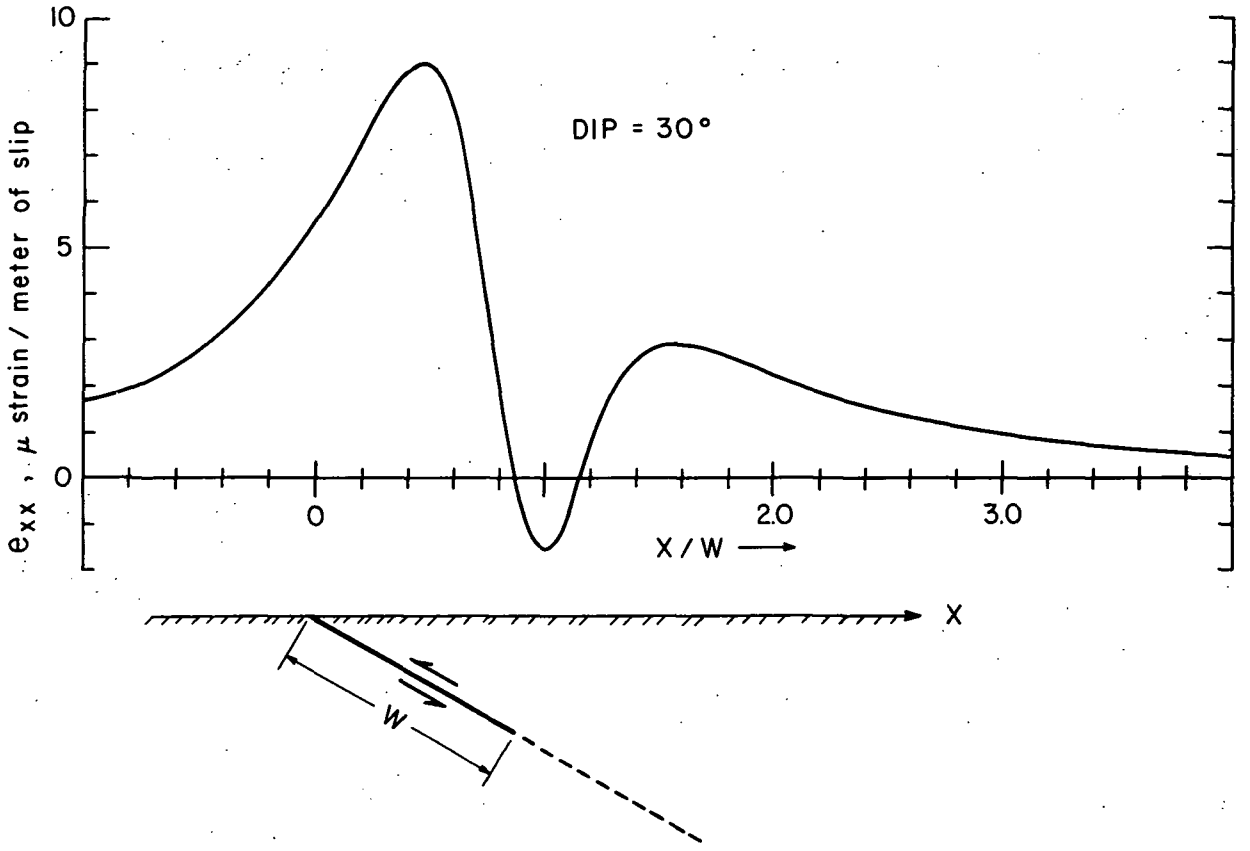



Fig. 2. Dislocation model of a convergent plate margin. The upper sketch shows the strain release by slip on the thrust fault shown in the lower sketch.

average be equal to (but opposite in sign from) the quotient strain release divided by time between earthquakes. This accumulation is presumably due to slip on the deeper sections of the fault that did not slip at the time of the earthquake. Figure 3 shows the equivalent model for strain accumulation at a transcurrent plate boundary. In this figure the strain accumulation (top center) has been calculated for slip at depth on the fault (top left). The earthquake is represented by the abrupt transformation from the configuration shown at the top of the figure to that at the bottom. A striking difference between the two models (Figures 2 and 3) is that strain accumulation is concentrated very close to the plate boundary in the transcurrent model whereas it is broadly distributed at the convergent boundary.

Japan

Figure 4 shows the total shear component $\gamma = (\gamma_1^2 + \gamma_2^2)^{1/2}$ accumulated in Japan during the interval 1900-60. In that figure the magnitude of the total shear component is represented by the length of the bar symbol and the direction of the bar indicates the strike of the plane of maximum shear. (If the plane of maximum left-lateral shear is shown, a solid bar is used; if the plane of maximum right-lateral shear is shown, a dashed bar is used. In either case one symbol may be replaced by the other symbol drawn perpendicular to the first.) The direction of the axis of maximum contraction is 45° counterclockwise from the solid bars and 45° clockwise from the dashed bars. Recall that Figure 4 shows the net strain accumulation during the 1900-60.

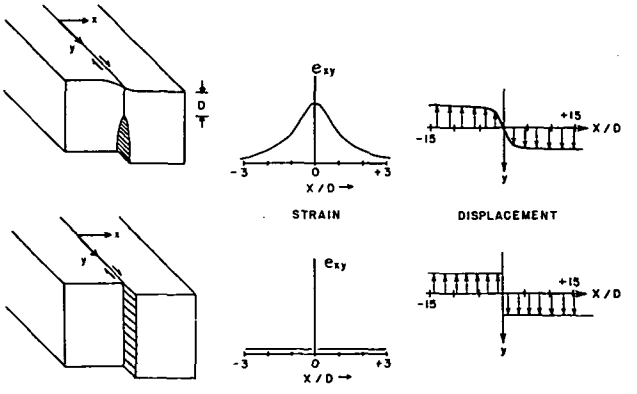



Fig. 3. Dislocation model of a transcurrent plate margin. Strain (top center) accumulates in response to slip at depth on the fault (top left). An earthquake releases the strain (lower sketches).

Several major earthquakes occurred in Japan during that period, and the strain released in those seismic events locally dominates the total accumulation of strain. This circumstance has introduced some complexity into the patterns shown in Figure 4.

We will discuss here only the south coast of Japan between longitudes 133°E and 140°E, the region of interaction with the Philippine Sea plate. That plate underthrusts Japan in a generally northwest direction resulting in a uniaxial northwest-southeast compression (solid bars north-south or dashed bars east-west). Stress release by a major earthquake results in a local uniaxial northwest-southeast extension (solid bars east-west and dashed bars north-south). With this background, it is easy to interpret the strain accumulation along the south coast of Japan in Figure 4. Only in the Tokai district (longitude 138°E) and perhaps extreme western Shikoku (longitude 133°E) is the expected northwest-southeast compression apparent. Elsewhere along the south coast there is a northwest-southeast extension corresponding to strain relief by the great 1946 Nankaido (longitude 134°E), 1944 Tonankai (longitude 136°E), and 1923 Kanto (longitude 140°E) earthquakes. It is not hard to see why the Japanese are presently concerned about an earthquake hazard in the Tokai district. Although the strain measurements are quite consistent with the dislocation model of Figure 2, the possibility that some of the accommodation occurs by anelastic deformation distributed over a broad boundary zone (Figure 1) is not excluded. Measurements of strain accumulation and release over several earthquake cycles could identify the relative contribution of continuous anelastic deformation and discrete dislocation motion.

New Zealand

As a second example we consider the deformation in New Zealand as described by Walcott (1978a). New Zealand lies along the Pacific-Indian plate boundary about 1500 km north of the present pole of relative plate rotation. Because of the proximity of the pole of rotation, the relative motion of the Pacific and Indian plates changes from approximately normal convergence at the rate of 50 mm/a (Pacific plate being underthrust) on North Island to highly oblique convergence at the rate of 40 mm/a (Indian plate being underthrust) on South Island.

The rate of strain accumulation in New Zealand as determined from triangulation data by Walcott (1978a) is shown in Figure 5. (Notice that in Figure 5 the bar indicates the direction of the axis of greatest contraction rather than the

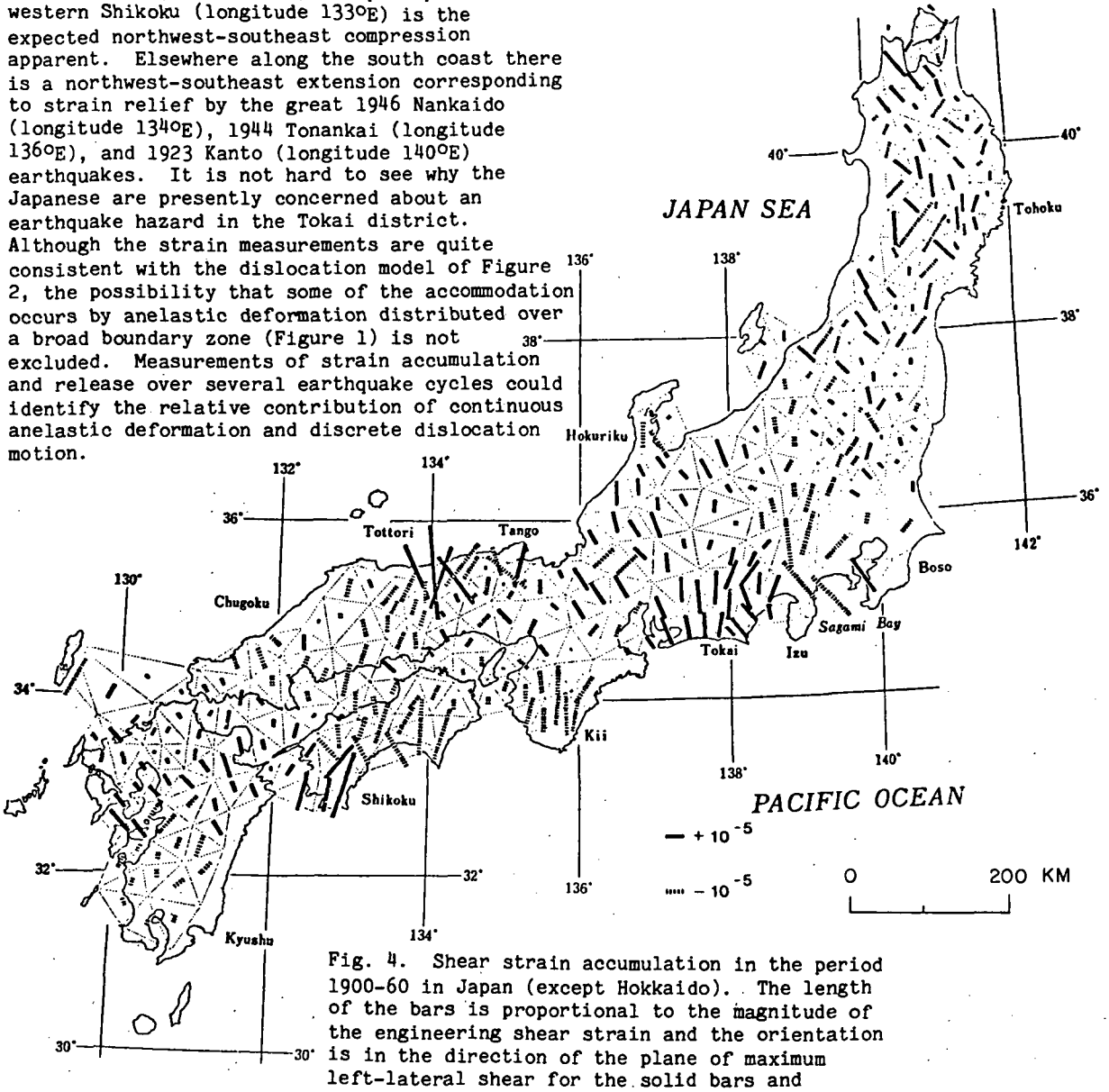



Fig. 4. Shear strain accumulation in the period 1900-60 in Japan (except Hokkaido). The length of the bars is proportional to the magnitude of the engineering shear strain and the orientation is in the direction of the plane of maximum left-lateral shear for the solid bars and right-lateral shear for the dashed bars. (From Harada and Kassai, 1971).

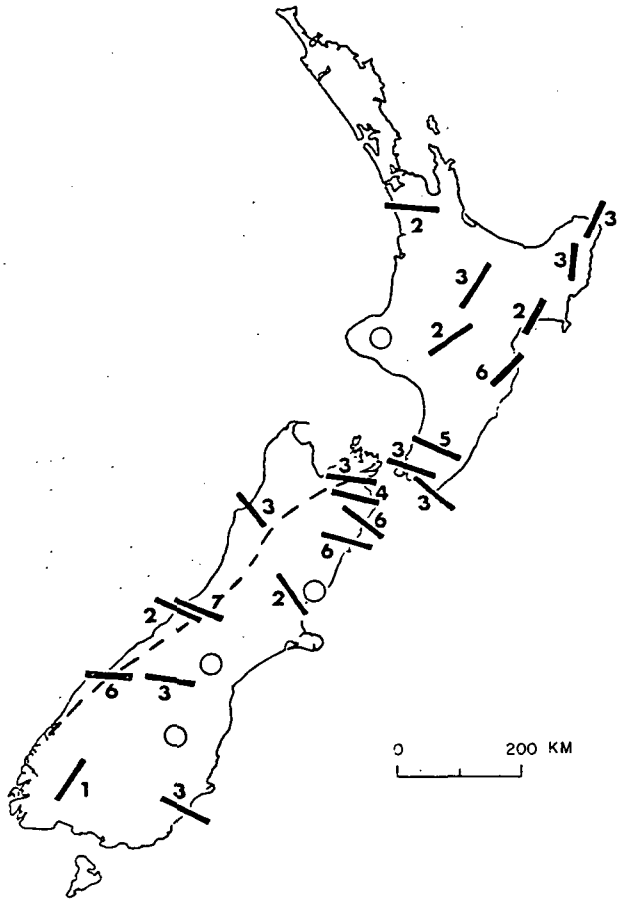



Fig. 5. Rate of shear strain accumulation in New Zealand as given by Walcott (1978a). The bars are in the direction of maximum compression, and the magnitude of the engineering shear strain in units of $10^{-7}$/a is given by the number beside the bar. Open circles indicate regions where no significant strain rate was observed. The dashed line running the length of the South Island shows the position of the Alpine fault.

strike of the plane of maximum shear as in Figure 4). These strain rates represent the average rate of accumulation over periods generally in excess of 60 years and in some cases may include strain release associated with major earthquakes, as in northeastern New Zealand, site of the 1931 Hawke's Bay earthquake (Walcott, 1978b). There the shear strain (Figure 5) shows a northwest-southeast extension typical of strain release rather than compression expected for a convergent boundary. In central New Zealand the compression axis is approximately normal to the plate boundary as expected for convergence of the Pacific and Indian plates. In the southern part of New Zealand the compression axis has rotated counterclockwise so that the relative plate motion is oblique to the plate boundary along the Alpine fault. The strain accumulation rates are also quite consistent with the plate tectonics model. As shown in Figure 5 the strain accumulation appears to average about 0.3 μstrain/a across a zone perhaps 200-km wide, implying relative motion of about 60 mm/a. Such continuously distributed deformation is, of course, implied by the models in Figure 1. The observations on North Island where plate convergence is dominant could also be explained by the model of Figure 2 which provides for a broad distribution of strain. Superficially at least, the observations of a broad distribution of strain on South Island would appear to exclude the transcurrent boundary model of Figure 3. However, there is some evidence that the strain in South Island is appreciably concentrated along the Alpine fault. For example, the two adjacent strain determinations of 0.2 and 0.7 μstrain/a (2 and 7 in units of the figure) on the central west coast of South Island in Figure 5 represent measurements solely on the western fault block (0.2 μstrain/a) and spanning the Alpine fault (0.7 μstrain/a), respectively. The appreciably higher strain rate in the latter case indicates that a major part of the plate motion is accommodated by slip on the Alpine fault in that area.

## United States

The San Andreas fault in California is presumed to define the transcurrent boundary between the Pacific and North American plates. Measurements of strain accumulation in central California, where the San Andreas fault is reasonably straight and closely parallel to the direction of relative plate motion, are shown in Figure 6. Because the measurements shown in that figure are based upon trilateration surveys, the complete strain tensor including dilatation may be calculated. On the average the strain is a shear parallel to the fault and of magnitude about 0.3 μstrain/a engineering shear. (The tensor shear is half as large.) The breadth of the shear zone is somewhat greater than might be expected from the model of Figure 3, perhaps partly because accommodation of plate movement is distributed over several subparallel faults. However, even the sum of the motion on the faults seems to be perhaps 30 percent less than the anticipated plate motion; whether the missing 30 percent of the motion is accommodated on offshore faults or distributed over a broad anelastic zone is not known.

Along a 100-km-long straight section of the San Andreas fault south of the Gavilan net (Figure 6), the plate motion is particularly simple. No appreciable strain accumulation is measured on either side of the San Andreas fault, and the entire relative motion appears to be accommodated by steady slip on the fault as shown in the lower diagrams of Figure 3. The relative plate movement measured directly over a zone a few kilometers in width is about 32 mm/a (Savage and Burford, 1973), a value substantially below the 55 mm/a estimated for relative plate motion averaged over the last 4 Ma. The discrepancy may indicate that another fault system, perhaps offshore, accounts for part of the relative plate motion or alternatively that a very broad plastic zone (Figure 1, lower) may be involved.

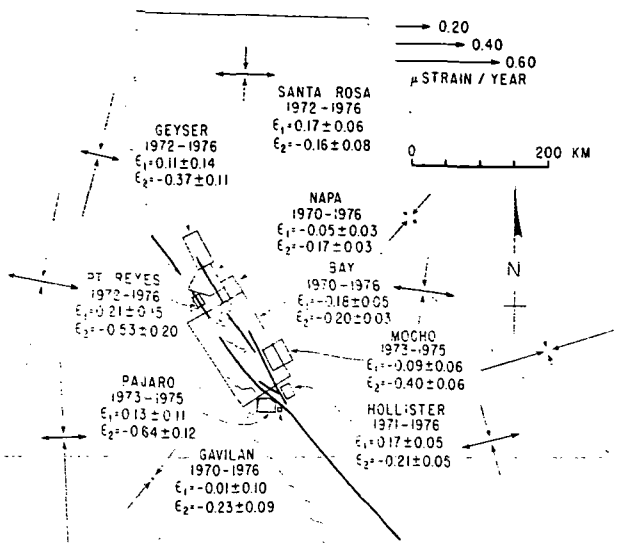



Fig 6. Average principal tensor strain accumulation rates along the San Andreas fault system (heavy lines) in central California for the period 1970-76. The strain rates shown are averages over the areas shown in the associated polygon. The units are strain/a; $e_1$ denotes the most tensile strain rate, and $e_2$ the most compressive strain rate. (From Prescott et al., 1978).

The plate boundary is somewhat more complicated in southern California. There the San Andreas fault exhibits a major bend so that an element of plate convergence is introduced. In spite of the bend in the plate boundary, the strain accumulation pattern is remarkably simple: essentially a uniaxial north-south contraction over a broad region as shown in Figure 7. The origin of this strain field is as yet unexplained. It may be a short-term aberration in the secular strain accumulation as the period of measurement spans only five years.

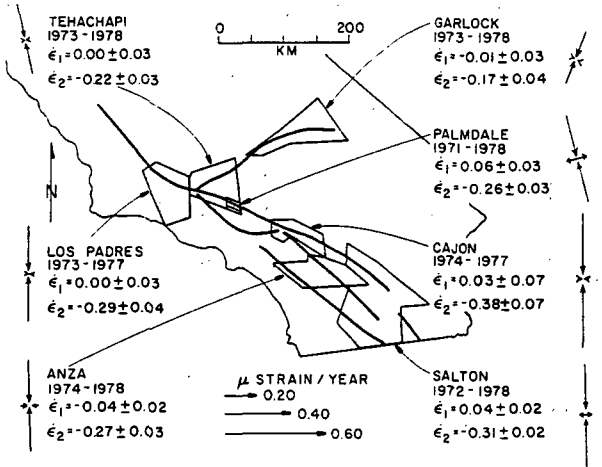



Fig. 7. Average principal tensor strain accumulation rates along the San Andreas fault system (heavy lines) in southern California for the period 1971-78. The strain rates shown are averages over the area shown in the associated polygon. The units are strain/a; $e_1$ denotes the most tensile strain rate, and $e_2$ the most compressive strain rate (From Savage et al., 1978).

## References


Chinnery, M. A., The deformation of the ground around surface faults, Bull. Seismol. Soc. Am., 51, 355-372, 1961.

Frank, F. C., Deduction of earth strain from survey data, Bull. Seismol. Soc. Am., 56, 35-42, 1966.

Freund, L. B., and D. M. Barnett, A two dimensional analysis of surface deformation due to dip slip faulting, Bull. Seismol. Soc. Am., 66, 667-675, 1976.

Harada, T., and A. Kassai, Horizontal strain of the crust in Japan for the last 60 years (in Japanese), J. Geodetic Soc. Japan, 17, 4-7, 1971.

Prescott, W. H., An extension of Frank's method for obtaining crustal shear strains from survey data, Bull. Seismol. Soc. Am., 66, 1847-1853, 1976.

Prescott, W. H., J. C. Savage, and W. T. Kinoshita, Strain accumulation rates in the western United States between 1970 and 1978, J. Geophys. Res., (in press), 1978.

Savage, J. C., and R. O. Burford, Geodetic determination of relative plate motion in central California, J. Geophys. Res., 78, 832-845, 1973.

Savage, J. C., W. H. Prescott, M. Lisowski, and N. King, Strain in southern California: measured uniaxial north-south regional contraction, Science, (in press), 1978.

Scholz, C. H., and T. J. Fitch, Strain accumulation along the San Andreas fault, J. Geophys. Res., 74, 6649-6666, 1969.

Walcott, R. I., Present tectonics and late Cenozoic evolution of New Zealand, Geophys. J. Roy. Astr. Soc., 52, 137-164, 1978a.

Walcott, R. I., Geodetic strains and large earthquakes in the axial tectonic belt of North Island, New Zealand, J. Geophys. Res., 83, 4419-4429, 1978b.

Page Intentionally Left Blank

# Multiple Wavelength Geodesy

Judah Levine *
Time and Frequency Division, National Bureau of Standards
Boulder, Colorado 80302

Abstract. We are constructing an apparatus which should be able to measure baselines up to 50 km long with a fractional uncertainty of about $5 \times 10^{-8}$. The instrument will measure both the optical length and the required correction due to the refractivity of the atmosphere using three wavelengths transmitted in one direction over the path to an active receiver. The three wavelengths are 632.8 nm, 441.6 nm and 3.7 cm. The two endpoint instruments are synchronized using subsidiary return transmissions at 632.8 nm and another telemetry signal. The one-way nature of the system allows an increase in range over existing round-trip systems.

Geodetic measurements have important bearing on many areas of tectonophysics such as plate tectonics and earthquake processes. In particular precise geodetic measurements are needed to understand how strain fields near plate boundaries change with time and why there are stresses in the interiors of plates large enough to cause isolated zones of seismicity.

These questions are being addressed using fixed geophysical instruments and portable instruments. Although one instrument might be used for all geodetic measurements, there may be some advantage to considering different techniques for the portable and fixed measurement programs.

In the case of a permanent observatory, weight and size are second order considerations, and emphasis should be placed on accuracy and sensitivity.

Berger and Levine (1974) have estimated the power spectrum of the random fluctuations in strain over a wide frequency range. They concluded that the power spectrum is inversely proportional to the square of the frequency, and that the power at 1 Hz is approximately $8 \times 10^{-28}$ $(\Delta L/L)^2/Hz$.

The important point is that they used two very different instruments located 2000 km apart in very different geologies. The two spectra are essentially identical in spite of these differences. This suggests that both instruments are limited by the same processes within the earth and that their power spectra represent a lower bound to what might be obtained using other techniques.

Their results do not support the widely-held beliefs that installations in mines or tunnels present insuperable problems and that long (800 meter) baselines are *a priori* better than shorter (30 meter) ones. The instruments appear to be limited by pier tilt, local effects, and, possibly, by drift in the wavelength stabilizer. These are correctable, at least in principle, so that it is not unreasonable to expect that a carefully designed instrument of moderate length (of order 100 meters) could be installed a meter or two below ground level, and that such an instrument would be limited by earth noise for periods shorter than a few years. Furthermore, readout schemes are possible which would obviate the need to operate the instrument continuously.

A strainmeter of this type would have only two weaknesses. First, it would be impractical to move to a new site since a significant fraction of the cost of the system comes from site preparation and installation. Second, it might be influenced by local effects such as tilting of the piers or shears within the ground in the immediate vicinity of the piers.

These local effects were not a problem in the Poorman Mine because it was a hard-rock site far below ground level, but it is possible that they will be the dominant problem at surface (or near surface) sites even if the material appears to be competent.

Berger has studied these problems extensively at the Piñon Flat Observatory. He finds significant correlations between rainfall and changes in strain rate and an anomalous shear in the top few meters of the ground under the pier (Berger, 1978). He concludes that the only way to deal with this problem is to reference the endpoints to deeper, presumably more stable, rock.

These sorts of difficulties will limit measurements made with instruments of any length. Anomalous pier motions on the order of millimeters represent fractional changes of parts in $10^8$ even over 50 km baselines, so that such effects will make significant contributions to the error budget of any instrument now in operation or under construction. It is very important that these effects be studied in detail. It would be especially useful to compare measurements made in the same area by instruments using very different length baselines.

Laser strainmeters are clearly unsuitable if the instrument must be portable. In this case we must use an instrument capable of measuring distances through the atmosphere. There are many instruments which have been designed to perform such measurements (Slater and Huggett, 1976). Since all such instruments effectively measure the transit time for some electromagnetic signal, it is necessary to know the actual speed of light in the atmosphere in order to derive the distance.

The refractivity (the deviation of the index of refraction from unity) of the atmosphere is about $3 \times 10^{-4}$ for visible wavelengths. Since it is desirable to measure distances with a fractional uncertainty of less than $1 \times 10^{-7}$, the atmospheric correction is very important. The refractivity is a function of atmospheric tem-

*Fellow, Joint Institute for Laboratory Astrophysics of the National Bureau of Standards and the University of Colorado.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280. The Ohio State Univ., Columbus, Ohio 43210.

perature, pressure and humidity so that making the corrections using meteorological data is cumbersome but possible. For optimum results the atmosphere must be sampled along the path, usually using sensors carried by an airplane (Savage and Prescott, 1973).

Methods which determine the refractivity of the atmosphere by measuring its dispersion (i.e., the apparent difference in distance between measurements made using different wavelengths) have been proposed for some years (Bender and Owens, 1965).

The use of two optical wavelengths, for example, enables one to determine the dry-air contribution to the refractivity, but does not correct for the refractivity due to water vapor, since the refractivity of water vapor is almost non-dispersive across the visible spectrum. The addition of a third measurement at a microwave frequency allows nearly perfect determination of the index except for a small term which can almost always be determined using temperature measurements at the endpoints. An analysis by Thayer (1967) of such a three-wavelength system suggests that baselines up to 50 km long could be measured with fractional uncertainties of $3 \times 10^{-8}$, and that the main limitation on the accuracy of such a measurement would come from the difference in the paths traversed by the two optical wavelengths due to the vertical gradient in the refractivity of the atmosphere.

Instruments which measure the refractivity of the atmosphere using multiple wavelength methods have been described by Slater and Huggett (1976) and by others (Wood, 1971). The one shortcoming of these instruments is that they can measure only rather short baselines (up to 15 km). These instruments are limited by spreading and attenuation of the beam in the atmosphere and by an inability to totally distinguish between a true return signal and light scattered backwards from the exit optics of the transmitter.

If these systems are converted to one-way operation by replacing the retroreflector by an active receiver and a second transmitter, then a large increase in returned signal will result. The increase in range realized by conversion to one-way operation is at least a factor of two (if the signal-to-noise ratio is limited primarily by attenuation) and may be considerably more than that.

The simplified schematic diagram of such an instrument is shown in Figure 1. Light from the source is sent through the local modulator, traverses the path, and then goes through the far-end modulator before detection. After being twice modulated the light at the detector will show variation at the difference frequency $f_b = f_1 - f_2$ (the other, higher, frequency components are not detectable). If two different optical wavelengths are sent through the system simultaneously, the two signals at the far end will show a phase difference proportional to the refractivity of the atmosphere. If a third light source sends light backwards through the system, it will arrive at the original end with a phase proportional to the transit time along the path and the various oscillator offsets. If a secondary link is used to transmit synchronizing information between the two ends then both the dispersion and the distance may be determined. This link is most conveniently done using a microwave carrier, in which case the phase shift of the microwave carrier provides information

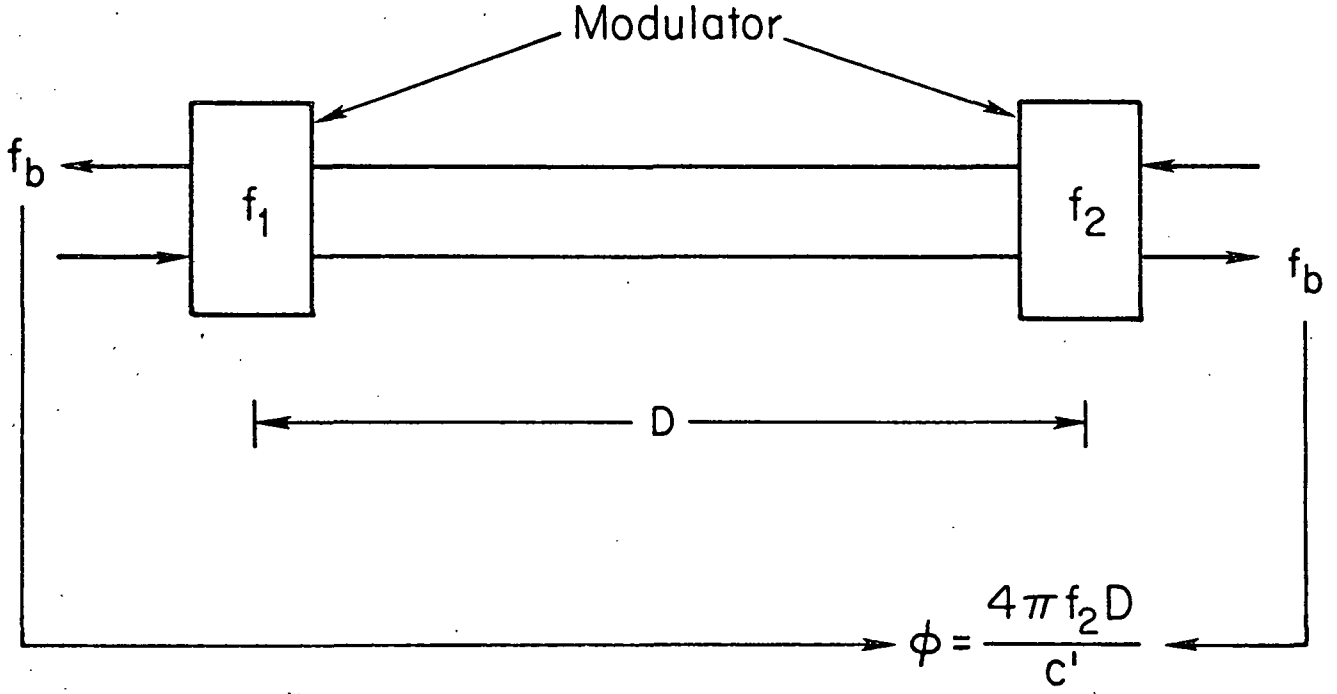



Fig. 1. Principle of two-way optical distance measurement. Note that the measured phase depends only on the frequency of the oscillator at the end where the measurements are made.

about the contribution of atmospheric water vapor to the refractivity.

Figure 2 shows a block diagram of the complete instrument. The dispersion measurements are made using 441.6 nm, 632.8 nm and 8.1 GHz. The red (632.8 nm) signal is sent back along the path for the distance measurement.

Figure 3 shows the calculated range of the instrument, and compares it with a retroreflecting instrument using similar technology. The lower three curves are for the retroreflecting instrument and the upper ones for the two-way design. Both systems are assumed to use 5 mW lasers with 1% overall detection efficiency. The range is defined to be the distance at which shot noise limitations allow a measurement with an uncertainty at $1 \times 10^{-8}$ in 10 seconds. We feel that an angular beam spread of 10 arcseconds is probably realistic for a typical atmosphere.

An instrument designed according to these principles is currently under construction. This instrument can be easily broken down into pieces, the heaviest of which weighs about 60 kg so that it can be assembled by two people at any location accessible by a four-wheel drive vehicle or a helicopter.

It is possible to use the principles of the multiple wavelength system to construct a three wavelength refractometer. Such an instrument would measure the index of refraction of the path but would not measure the distance.

Although such an instrument would use essentially the same hardware as the distance-measuring instrument, the electronics are measurably simpler and no return transmissions are necessary.

Such an instrument might prove useful as an adjunct to existing geodetic instruments.

This work is supported in part by NASA Grant No. NSG-7344 through the University of Colorado.

References

Bender, P.L. and J.C. Owens, Correction of optical distance measurements for the fluctuating atmospheric index of refraction, J. Geophys. Res., 70, 2461-2, 1965.

Berger, J., private communication, 1978.

Berger, J. and J. Levine, The spectrum of earth strain from $10^{-8}$ to $10^{2}$ Hz, J. Geophys. Res., 79, 1210-4, 1974.

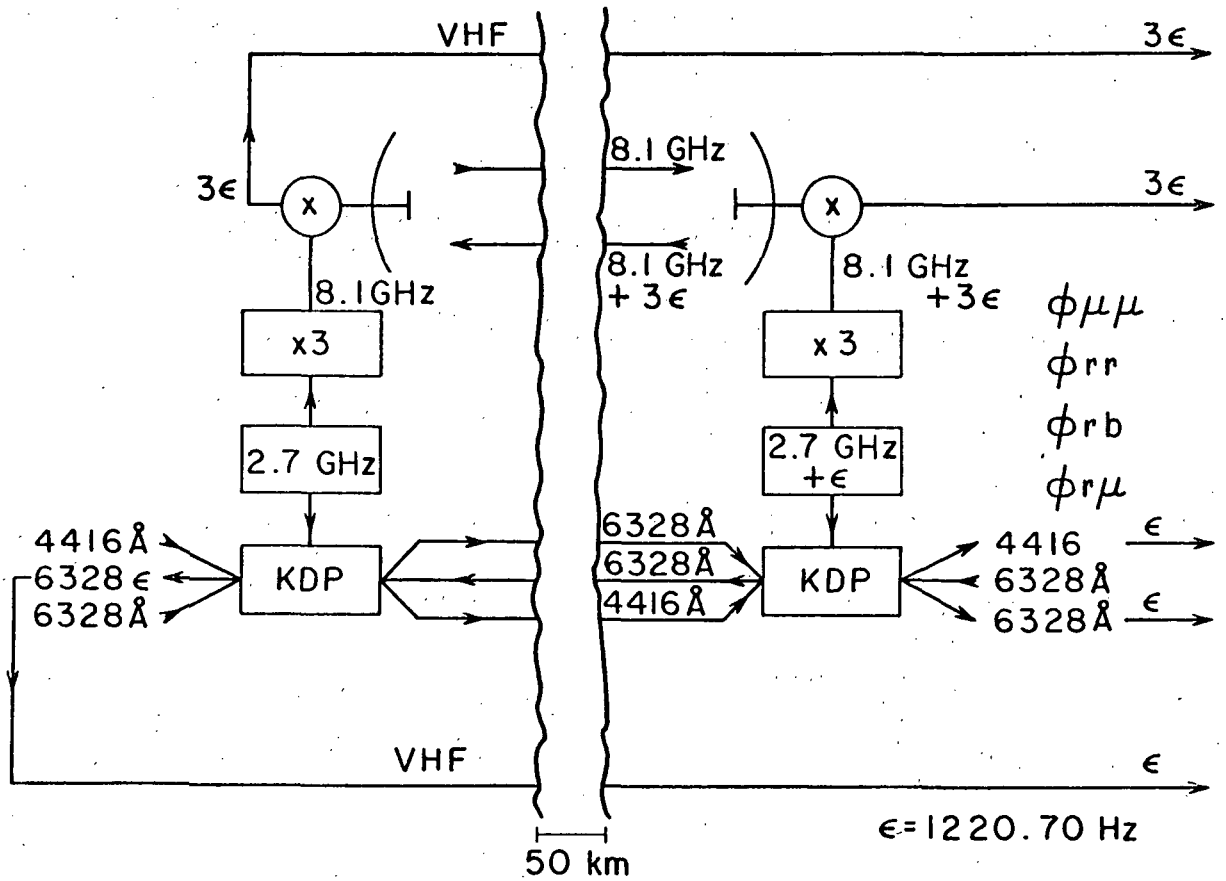



Fig. 2. Block diagram of the prototype system. "UHF" is a low bandwidth telemetry channel. The various phase measurements are identified with subscripts r for red (632.8 nm), b for blue (441.6 nm) and μ for microwave (8.1 GHz).

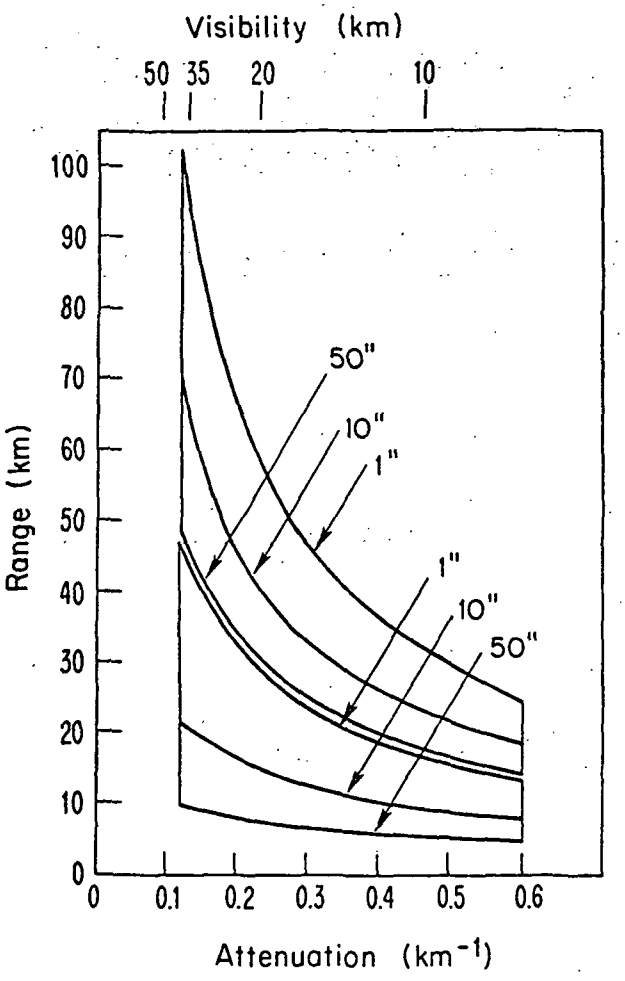



Fig. 3. Calculated range of one-way (upper curves) and retroreflector EDM instruments as a function of optical beam spread and atmospheric attenuation.

Savage, J.C. and W.H. Prescott, Precision of geodolite distance measurements for determining fault movements, J. Geophys. Res., 78, 6001-8, 1973.

Slater, C.R. and G.R. Huggett, A multiwavelength distance measuring instrument for geophysical experiments, J. Geophys. Res., 81, 6299-6305, 1976.

Thayer, G.C., Atmospheric effects of multiple frequency range measurements, Technical Report IER 56-ITSA 53, Environmental Science Services Administration, Rockville, MD, 1967.

Wood, L.E., Progress in electronic surveying, U. S. National Report, EOS, 52, IUGG17-21, 1971.

# Problems and Advances in Monitoring Horizontal Strain

Michele Caputo
Istituto di Fisica, Università degli Studi
Roma, Italy

Abstract. After a brief description of the modern instrumentation used in Geodesy for the detection of the deformations of the crust of the Earth, follow a list of problems and a discussion concerning the needs for the survey of the physical quantities of interest in geodesy, geology, geophysics and engineering such as the strain invariants, the optimal network of baselines and the accuracy. An analytic method is also given for the computation of the effect of a source of dilatation in a spherical Earth.

## Introduction.

From the observation of the stress drop in the earth's crust caused by earthquakes in several regions of the world we have learned that the most frequent stress drops are those at low level (Caputo, Console., (1978)); but we have also learned that stress drops of less than 0.4 bar are almost never found in earthquakes recorded with usual seismographs. The corresponding strain released amounts to several parts of $10^{-6}$. This is a quantity which is measurable with instruments of various type, including portable geodetic instruments, with an accuracy better than 10%. This in turn implies that we can try to monitor the accumulation of strain which preceedes the earthquakes. The same accuracy is sufficient for monitoring the deformation of large structures such dams, bridges, buildings, tunnels, mineshafts and others. To determine this accumulation of strain one should observe directly in the rock. This in turn can be achieved with several techniques some of which are of geodetic interest. Other geodetic techniques are of lesser accuracy but are of interest because they can be applied in some circumstances when the more accurate techniques are not possible because of logistic difficulties; they will be discussed in a separate presentation.

In the following we shall first describe the geodetic techniques used in the past, and discuss for future needs. Some advances will also be presented.

## Geodetic measurements.

The basic concept in the geodetic measurement of horizontal deformation is that the distances between selected points are measured at different times and then compared. The instrumentation is based on electromagnetic waves which are used in two ways.

One way is to observe the motion of the interference fringes of a monochromatic light source caused by the relative motion of two bench marks connected with the light source and a reflector. This method allows to monitor continuously slow variations of distances with a resolution of the order of $10^{-10}$. The distance between the bench marks is limited by the fact that the light beem should stay in the vacuum; distances of almost 1 km have been used. There should not be basic difficulties in making this instrument portable or to operate it between bench marks connected to the rock such as in natural caves in mine shafts, railway of road tunnels, or in tunnels drilled for this purpose. Since the instrument does not give any absolute measurement of distance but it monitors its variation the need to make it portable is for saving time in assembling it and to make its use easier. A good example for this type of instruments is that of Wyatt and Berger (1978). A method of much lesser accuracy, at most one part in $10^{-7}$, but of much easier and faster use, is that of measuring distances between several points using laser instruments with one or more colors. To increase the accuracy of the instruments with one color Slater and Huggett (1976) introduced the technique of using a radial set of baselines, of almost instantaneous measurements and of comparing the ratios of the lengths of the baselines at different times to detect relative variations of length. The method reduces the errors due to the variation of the atmospheric parameters and is applicable with the hypothesis that the longterm changes in the index of refraction over the baselines are proportional to the index of refraction at each line; the proportionality coefficient must be the same for all the lines in the array during the specific survey. The method of the ratio is useful because it requires to record the atmospheric parameters during the survey only at one end point of the baseline. The precision of the method is about one or two parts in $10^{-7}$.

This method has been succesfully used by measuring the atmospheric parameters at both end point of all baselines for very long baselines (of the order of 100 km); the precision of the method is reported of 2 parts in $10^{-7}$ (Carter and Vincenty (1978)). Another method to minimize the errors due to the non uniform and constant atmosphere has been introduced by Parm (1973, 1975); he used base lines over flat terrain and made the observations only when the wind was almost parallel to the baseline.

These methods do not give the absolute measure of the baseline with the same accuracy obtained measuring the average atmospheric parameters along the baseline by means of an aircraft (Savage (1975) and Savage and Prescott (1973)) although the va-

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

riations of the length of the baseline is obtained perhaps with greater accuracy.

However the best results in measuring the variations of the length of baseline and its absolute length are obtained with a multiwave length geodimeter (Hugget et al. (1977), Slater and Hugget (1976)) with an accuracy of $10^{-7}$ in the variation of the length and in the absolute value of the length of the baseline. It seems therefore that the accuracy of the instrumentation would allow to obtain the necessary accuracy on the surface of the earth.

The computation of the strain invariants.

Problems which are still partly or completely unsolved are those connected with logistics and cost. But the most important problem to solve is the meaning of the information which is now being collected on the surface of the earth in connection with what is happening below it and which we really want to know. Strictly speaking the problem which I should discuss here is two dimensional but in reality the problem is intrinsically threedimensional and one should never forget it. Therefore I shall try to set it in its proper space, the partition line between horizontal and vertical deformations will be set later where and how it will seem most proper.

The stress accumulated in the Earth's crust probably has uniform distribution over fairly large scale at least in the tectonic processes. This circumstance allows to observe all the components of the stress and therefore the stress invariants even when the baselines of the trilateration are large.

If one succeeds in measuring the displacement vectors

$\bar{u}_0(u_{10},u_{20},u_{30}), \bar{u}_1(u_{11},u_{21},u_{31}), \bar{u}_2(u_{12},u_{22},u_{32}), \bar{u}_3(u_{13},u_{23},u_{33}), P_j(x_{ij},x_{2j},x_{3j}), j=0,3,$

($x_{ij}$ are cartesian orthogonal coordinates with the $x_3$ axis normal to the surface of the Earth), assuming that the displacement is a linear function of the coordinates, one may write the following system of 9 equations

$$u_{ji}-u_{jo} = \frac{du_i}{dx_k}(x_{ki}-x_{ko}), i,j,k,=1,2,3 \qquad (1)$$

in the 9 unknown $\frac{du_i}{dx_j}$ which will give the components of the strain and therefore the invariants. It may be noted that on the surface of the Earth one may reasonably assume

$$\frac{du_3}{dx_1} = -\frac{du_1}{dx_3}, \frac{du_3}{dx_2} = -\frac{du_2}{dx_3}, \frac{du_3}{dx_3} = \frac{si}{si-1}\left(\frac{du_1}{dx_1}+\frac{dn_2}{dx_2}\right)$$

where si is the Poisson ratio. It is obvious that if one observes only horizontal displacement and one estimates the invariants with the hypothesis that $\frac{du_1}{dx_3} = \frac{du_2}{dx_3} = 0$, with few exceptional cases (e.g. see Caputo 1978), the results could be misleading. The cavity effects and the topographic and thermal effects (Harrison 1976, Harrison and Herbst 1977) give classic examples of wrong interpretation of strain observed in wrong places or better they give examples of wrong extrapolations of data, as it would be in our case.

Since we can measure displacements on the surface of the Earth the question is: how may we observe the displacement vector in points under the surface of the Earth? An estimate of the vertical component $\frac{du_3}{dx_3}$ at depth could be made by measuring directly the dilatation d and then subtracting the surface components computed with measurements on the surface of the earth. The Sacks Everston (Sacks et al. 1971) dilatometer is a good example of an instrument which can measure dilatation with an accuracy of $10^{-7}$. But the additional knowledge of d does not allow to compute the maximum shear strain and its direction which are of the greatest importance. The real problem is to estimate

$\frac{du_3}{dx_2}, \frac{du_3}{dx_1}, \frac{du_2}{dx_3}$ and $\frac{du_1}{dx_3}$, levellings allow to estimate the last two and therefore to retrieve the first two, but only on the surface.
Operating in a mine shaft could be an experiment to try, however, in this case, the displacement at the end of the baseline could be determined only by means of an open poligon with an accuracy which would hardly be acceptable and in any way would be by far inferior to that of the surface measurements. The use of a road or railway tunnel could allow to determine the displacement vector with a closed poligon, but in most cases the tunnel would not be located in a useful position. The problem is open and it is one of the most important.

A problem which I think we should also consider is that of the most suitable distribution of the baselines or of the bench marks for determining the distribution of strains on the surface of the Earth in the case of the hypothesis $du_1/dx_3 = du_2/dx_3 = 0$. In this case the system (1) is reduced to 2 linear systems of 2 equations each namely

$$\begin{aligned}
u_{11} - u_{10} &= \frac{du_1}{dx_1}(x_{11}-x_{10}) + \frac{du_1}{dx_2}(x_{21}-x_{20}) \\
u_{12} - u_{10} &= \frac{du_1}{dx_1}(x_{12}-x_{10}) + \frac{du_1}{dx_2}(x_{22}-x_{20}) \\
u_{21} - u_{20} &= \frac{du_2}{dx_1}(x_{12}-x_{10}) + \frac{du_2}{dx_2}(x_{21}-x_{20}) \\
u_{22} - u_{20} &= \frac{du_2}{dx_1}(x_{12}-x_{10}) + \frac{du_2}{dx_2}(x_{22}-x_{20})
\end{aligned} \qquad (2)$$

The matrix of the coefficients in both systems is

$$\begin{matrix} x_{11} - x_{10} & x_{21} - x_{20} \\ x_{12} - x_{10} & x_{22} - x_{20} \end{matrix} \qquad (3)$$

which should be different from zero. Therefore the three points $P_0, P_1, P_2$ should not be in a straight line.

Accuracy of the determination of strain.

Given the m.s.e.s $(x_1) = s(x_2)$ in the determination of the coordinates of the points of the network in all the surveys, which in turn give the displacement vectors, the m.s.e. of the invariants are in the first approximation

$$s(d) = \frac{s}{\sin z} \sqrt{2}\left(\frac{1}{d_1^2} + \frac{1}{d_2^2}\right)^{\frac{1}{2}} = 2\bar{s}$$

$$s(a) = \bar{s} \qquad (4)$$

$$s(S) = 2\sqrt{2}\ \bar{s}$$

Where S is the maximum shear strain, a is its direction, z is the angle $P_1 P_0 P_2$ and $d_1$ and $d_2$ are the length of $P_0 P_1$ and $P_0 P_2$. If the strain field is uniform then the best s(a) and s(S) are obtained for $z = \pi/2$ and for $d_1$ and $d_2$ very large. This implies that when one computes the strain invariants one loses at least a factor 3 in the accuracy.

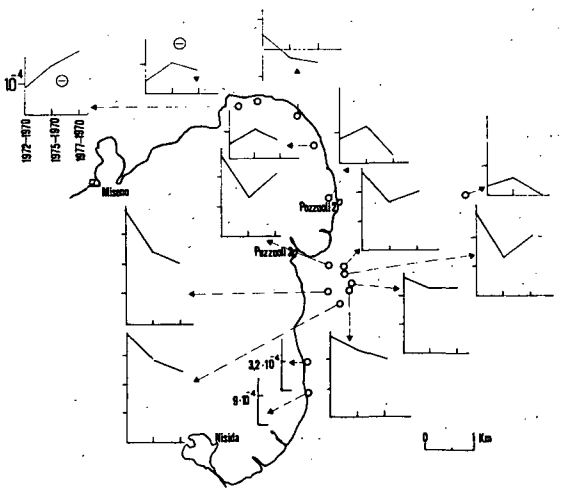



Fig. 1a - Dilatation as function of time in the period 1970-1977 in the Phlegrean Fields near Naples. Each diagram gives the dilatation in the point connected by the dashed arrow. In the top left diagram are indicated the units of dilatation and the time interval in which it has been accumulated.

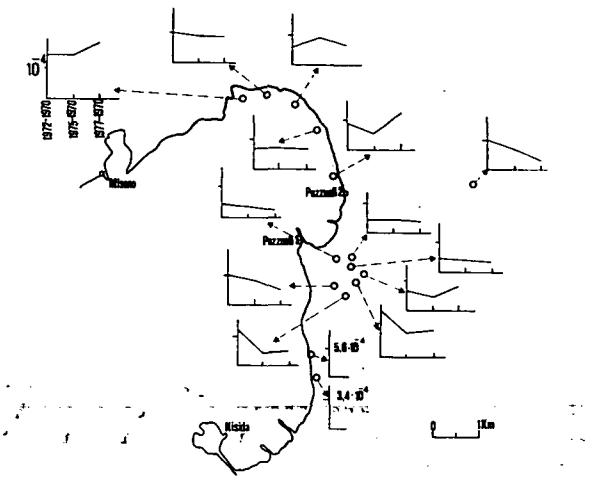



Fig. 1b - Dilatation as function of time in the period 1970-1977 in the Phlegrean Fields near Naples. Each diagram gives the maximum shear in the point connected by the dashed arrow. In the top left diagram are indicated the units of dilatation and the time interval in which it has been accumulated.

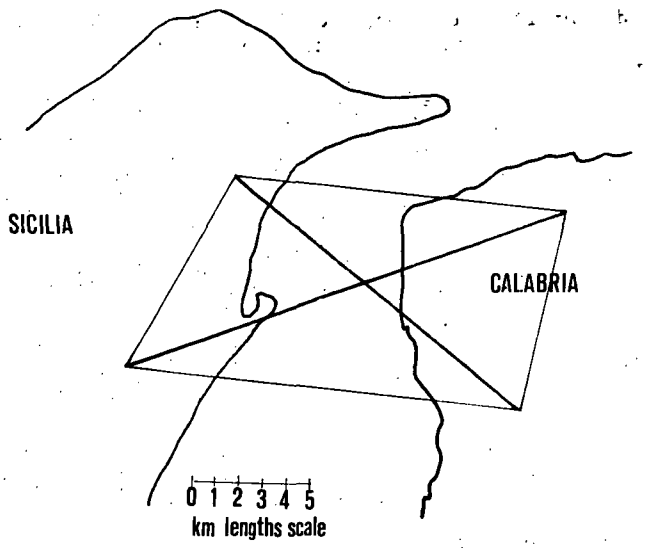



Fig. 1c - Network of baselines for the measurement of displacements across the Messina Straits.

When the strain field is not uniform formulae (4) set a limit to the accuracy of the determination of the invariants. It is the indetermination principle of the strain fields observations, which is expressed by

$$l\epsilon > \bar{s}$$

where $\epsilon$ is the strain observed in the path $l$. An almost complete example of a determination of a strain field is that of the Phlegrean Fields near Naples (Caputo 1978) shown in Fig. 1a, 1b.

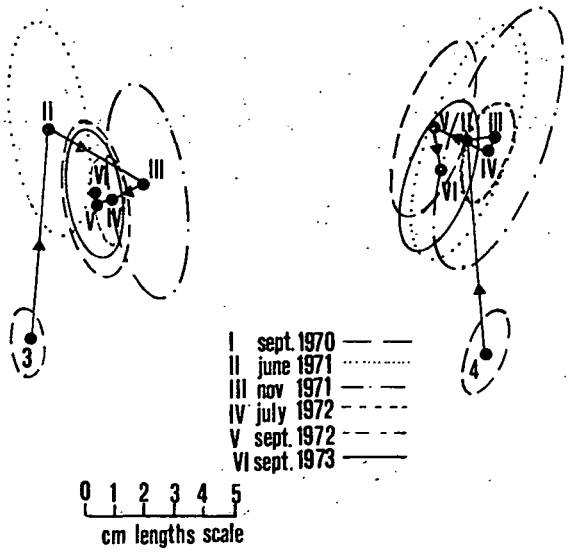



Fig. 1d - Displacement reported in the two points located North in the network of Fig. 1 .

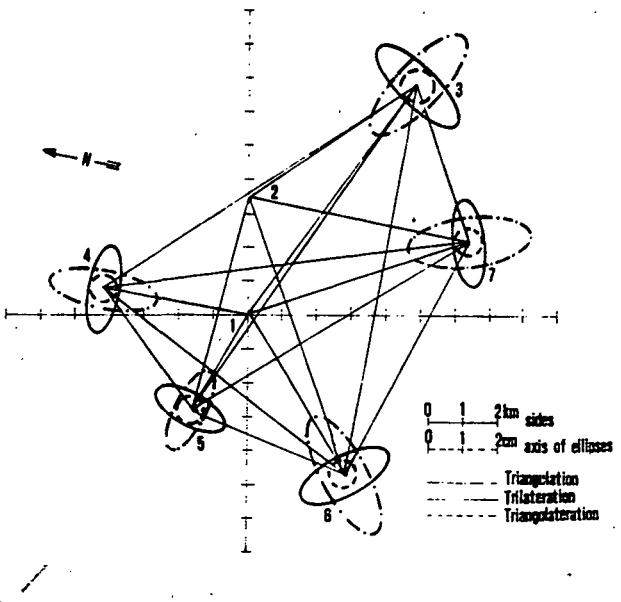



Fig. 1e - Network of baselines for the measuremnt of strain in the Ancona area.

The accuracy which one may achieve with standard portable instrumentation and procedures in a triangular network with 5 km baselines is $s \simeq 1$ cm and one obtains $s(d) = 6 \cdot 10^{-6}$; if one uses the mekometer then $s \simeq 2$ mm and the baselines are 2 km therefore $s(d) = 3 \cdot 10^{-6}$; by using a multiwavelength geodimeter with baselines of 5 km $s(d)$ would be $3 \cdot 10^{-7}$.

It is desiderable to develop numerical methods and network designs in order to improve the accuracy in the determination of the invariants.

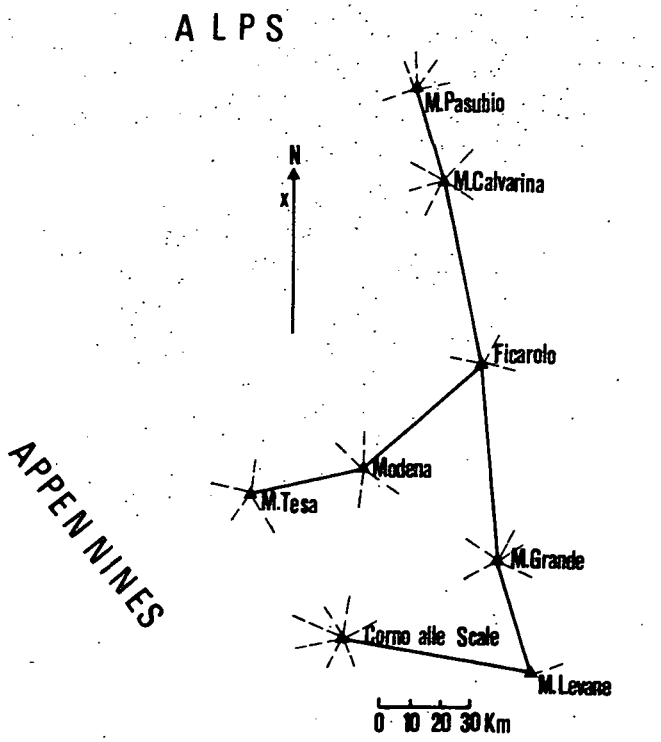



Fig. 1f - Network of baselines for the measuremnts of displacements between the Appennines and the Alps.

With reference to the network design let us consider the limiting case of a network of infinite size in two dimension. If the distribution of the strain is unknown obviously one should consider the most symmetric distribution of bench marks. According to the principle of the Pflastersatz the most simple distributions of points which have properties of symmetry are the three indicated in the Fig.2

When one considers the baselines needed to determine the invariants of strain one sees that the networks (c) of Fig.2 is not convenient because of the length of the baselines which should also be $\sqrt{3}\,l$ (with $l$ the minimum distance between points).

Since the two networks of points (a) and (b) have no condition equation the ratio of the numbers of measured baselines to the number of points where the strain is measured is the same. Also, to the first order $s(d) = \frac{1}{l}$ for the network (b), $\Delta s(d) = \frac{1}{l}$ for (a); that is they are almost the same. However the distribution of the points where the invariatns are measured is better in (a), because in (b) some points have the invariants which are direct interpolation of others.

We may also note that it may be needed to reinforce the strength of the networks which we have considered here and for which, it is immediately seen that the condition equations are non existing. If the baselines are short enough one may consider to measure some of the angles or some

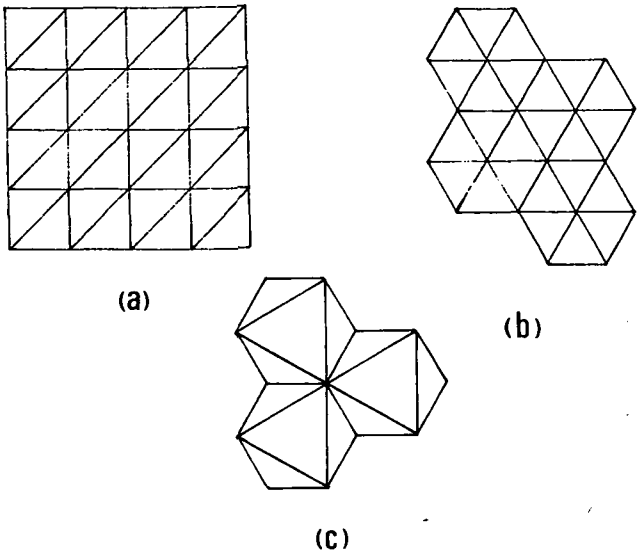



Fig. 2 - Possible networks of baselines formed with regular poligons.

more baselines between the existing points. For instance in the network which we used across the Messina Straits (Caputo et al. (1974)), when we measured only the baselines we had only one condition equation; we had more condition equations when (at times) we measured also the angles; the decision to adopt this procedure was due to logistic problems. In the case of the network of the Ancona area (Baldi et al. (1976)) we decided instead to add more points to the network and we had 9 condition equation. The method of adding more baselines is often preferred to the measurement of angles because one obtains the strains in a larger number of points and therefore obtain a higer resolution in the distribution of strain; also with the modern instrumentation, in networks with baselines of 10 km or longer, the measurement of angles is almost one order of magnitude less accurate. Before closing I must call the attention to the set of problems which I raised and which we should eventually discuss with the many more which probably will be raised after my presentation. These problems are

a) the need to have a threedimensional description of the strain fields
b) the need to have available a multiwave length geodimeter
c) the need to monitor strain on long baselines with an interferometer easy to install and to operate
d) the need to measure on bench marks which are rigidly connected to the medium to observe
e) to establish networks which are optimal for the final objective
f) to select a test ground where to make an experiment to satisfy all these requirement and to do the experiment. (This is a proposal for those who are interested).
g) to develop mathematical models for the displacement fields caused by internal perturbation of the Earth elastic field.
h) the use of theodolites, in conjunction with electromagnetic distance equipment with accuracy $\bar{s}$, is of use only if the length of the baseline l is $l < 8 \cdot 10^5 \bar{s}$ in which case its error would be $2\ \bar{s}$.
i) preference to interferometers when investigating slow rate accumulation of strain.
j) need to adjust the observed displacement fields with the Navier equations.

The displacement field caused by a source of dilatation in a spherical Earth.

It has been established that pryor to some earthquakes there is a deformation of the Earth's crust measurable on its surface. According to the dilatancy model in some cases this deformations caused by a dilatation in a threedimensional portion of the Earth's crust. It is therefore needed to have a threedimensional geometric elastic model of this phenomenon for the interpretation of the geodetic data.

A solution for a flat Earth model and a point source of dilatation has been given by Sacks et al (1975); we shall extend here this solution to a spherical Earth model $S_o$ of radius $r_o$ and a point source of dilatation $P_1$ at a distance h from the center of the Earth.

The method consists in finding first the field $\bar{u}$ generated by the dilatation source in $P_1$ in an infinite space assuming the origin of the polar coordinate system in $P_1$; then we shall move the origin of the coordinates in the center O of $S_o$ and finally add to $\bar{u}$ a vector $\bar{v}$ which renders the component of the stress applied to $S_o$ nil on $S_o$ as it naturally is. In a more sofisticated solution one may render $\bar{t}$ such to balance the weight caused by the topography.

We shall assume that the body forces caused by the source and its perturbation are negligible; in the first approximation this is acceptable when the source dimensions and the dilatation are limited to realistic cases. We assume also that the gravity field of the Earth has negligible effect with respect to the other fields playing the major role in this phenomenon. Solutions taking into account all those effetcs would obviously be more accurate; the solution presented here is considered only a first approximation of the spherical Earth.

The displacement caused in the space by a generic source if dilatation in P with the condition of convergence at infinity is

$$u_r = dV\, r^{-2},\quad u_z = 0 \qquad (5)$$

where dV is the change in volume of any volume which includes $P_1$. It is immediately verified that the dilatation of this displacement field is nil at all points different from $P_1$.

The formule for the mentioned transformation of coordinates, with reference to the figure 3 are

$$r' = \left[r^2+h^2-2rh\cos z\right]^{\frac{1}{2}} = \frac{h \sin z}{\sin a} \qquad (6)$$

and substituting in (1) we obtain

$$u_r = dV\left[r^2+h^2-2rh\cos z\right]^{-\frac{1}{2}} \cos a$$
$$u_z = dV\left[r^2+h^2-2\,r\,h\cos z\right]^{-\frac{1}{2}} \sin a \qquad (7)$$

where sin a and cos a are obtained from (6):

$$\sin a = h \sin z \left[r^2+h^2-2\,r\,h\cos z\right]^{-\frac{1}{2}}$$
$$\cos a = \left[1-\frac{h^2\sin^2 z}{r^2+h^2-2\,rh\cos z}\right]^{\frac{1}{2}} \qquad (8)$$

$$u_r = \frac{dV(r - h\cos z)}{\left[r^2+h^2-2r\,h\cos z\right]^{3/2}}$$
$$u_z = \frac{dV\,h\sin z}{\left[r^2+h^2-2r\,h\cos z\right]^{3/2}} \qquad (9)$$

We shall compute now the solution $\bar{v}$ of the Navier equations which is to be added to $\bar{u}$ in order to obtain the final solution $\bar{u} + \bar{v}$ such that the stress components $t_r$ applied to $S_o$ are nil on $S_o$.
According to the formulae of Caputo (1961) the general solutions of the Navier equations in the axisymmetrical case and homogeneous Earth of this case are

$$v_r = \sum_1^\infty n\,(A_{1,n}\,r^{n+1}+A_{3,n}\,r^{n-1})\,\frac{2n+1}{2}\,P_n(\cos z)+$$
$$+A_{10}\,r \qquad (10)$$
$$v_z = \sum_1^\infty (B_{1,n}\,r^{n+1}+A_{3,n}\,r^{n-1})\frac{2n+1}{2n(n+1)}\,\frac{dPn}{dz}$$
$$B_{3,n} = \frac{A_{3,n}}{n}$$
$$B_{1,0} = -\frac{3g+5}{2}\,,\quad g = \frac{1}{m}$$

l and m are the elastic parameters.

The components of the stress of the solution $\bar{v}$ are

$$t_{rr} = \sum_2^\infty n\left\{\left\{\left[(n+3)l+2(n+1)m\right]A_{1,n}-ln(n+1)B_{1,n}\right\}r^n+\right.$$
$$+\left\{\left[(n)\,l+2(n-1)m\right]A_{3,n}-n(n+1)l\,B_{3,n}\right\}r^{n-2}\Big]Pn$$
$$+(3l\,A_{1,0}+2\,m\,B_{1,0})\,P_1$$
$$t_{rz} = m\sum_1^\infty n\left[(n\,B_{1,n}+A_{1,n})\,r^n+\left[(n-2)\,B_{3,n}+\right.\right.$$
$$\left.+A_{3,n}\right]\,r^{n-2}\Big]\frac{dPn}{dz} \qquad (11)$$

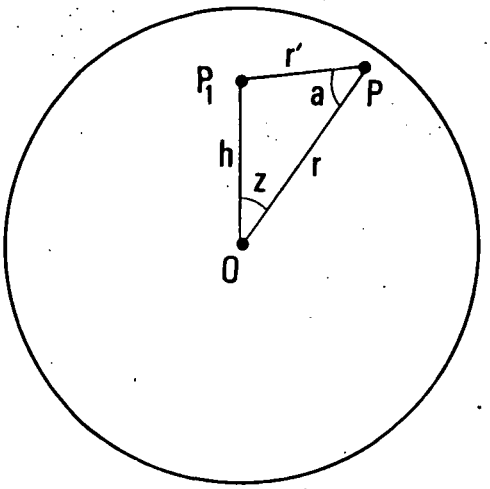



Fig. 3 - Coordinate system used in formulae (5) through (12).

If we require that the components of the stress $\bar{t}_r$ of the solution $\bar{u} + \bar{v}$ applied to $S_o$ are nil on $S_o$ the conditions are

$$\sum_2^\infty n\left\{\left\{\left[(n+3)l+2m(n+1)\right]\bar{A}_{1,n}-ln(n+1)\,\bar{B}_{1,n}\right\}r_o^n+\right.$$
$$+\left.(n-1)(l+2m)\bar{A}_{3,n}-l\,n(n+1)\,\bar{B}_{3,n}\right\}r_o^{n-2}\Big]Pn$$
$$+3\,l\,\bar{A}_{1,0}+2m\,\bar{B}_{1,0}+\frac{1}{r_o^2}\left[\frac{d}{dr}(r^2\,u_r)\right]_{r=r_o}+$$
$$+\frac{1}{r_o\sin z}\,\frac{d}{dz}(u_z\sin z)_{r=r_o}+2m\left[\frac{d\,u_r}{dr}\right]_{r=r_o}=0 \qquad (12)$$
$$\sum_1^\infty n\left[(\bar{A}_{1,n}+n\bar{B}_{1,n})r_o^n+(\bar{A}_{3,n}+(n-2)\bar{B}_{3,n})r_o^{n-2}\right]x$$

$$x \frac{dP_n}{dz} + \left[\frac{du_z}{dr} - \frac{u_z}{r} + \frac{1}{r}\frac{du_r}{dz}\right]_{r=r_o} = 0$$

where $\bar{A}_{i,n}$ and $\bar{B}_{i,n}$ are the unknown.

An expansion of $u_r$ and $u_z$, appearing in (12), in series of Legendre polinomials gives finally the values of $\bar{A}_{i,n}$ and $\bar{B}_{i,n}$ to sobstitute in (10) to obtain the final solution.

The extension of this result to a prestressed, Earth is almost obvious. What we really need is to prepare tables and figures with the computations of the effects of a dilatation in $P_1$ according to the formulae and to observe the phenomena on the real Earth.

The extension of the results of this section to a layered Earth is also trivial (Caputo 1961).

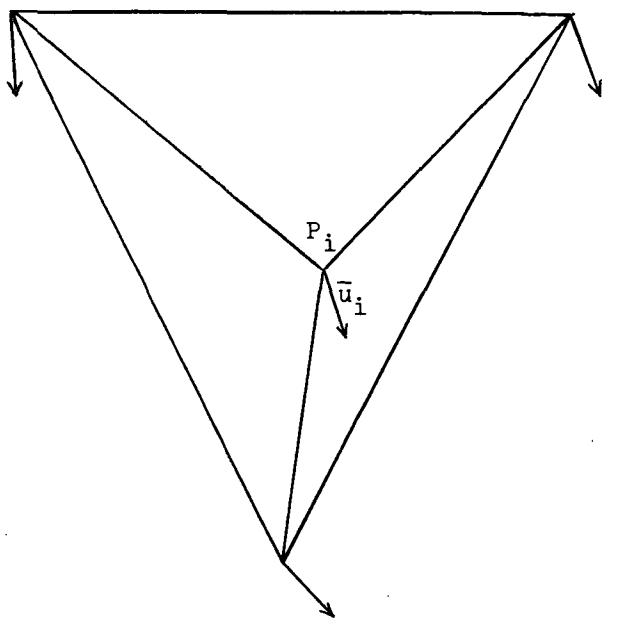



Fig. 4 - The four points $P_i$ should not belong to a single plane.

References.


Baldi P., Caputo M., Postpischl D., Boschi E., Unguendoli M., Geodetic network for crustal movements in the Ancona area, Boll.Geod.Sc.Affini, XXXV, 4, 494-507, 1976.

Berger J. A note on thermoelastic strains and tilts J.Geophys.Res., 80, 274-277, 1975.

Caputo M., Strain and other geodetic observations in the Phlegrean Fields near Naples in the last 2000 years. Geophys. J.R.astr.Soc. In press.

Caputo M., Deformation of a layered earth by an axially symmetric surface mass distribution J.Geophys.Res.,66,5,1479-1483,1961.

Caputo M., Console R., Model and observed statistics of California earthquakes parameters Annali di Geofisica, in press 1977.

Caputo M., Folloni G., Modern views regarding slow horizontal movements of the soil and feasability of measuring them, Geologia Applicata e Idrogeologia, X, I, 289-307, 1975.

Caputo M., Folloni G., Pieri L., Unguendoli M., Geodimetric control across the Straits of Messina, Geophys.J.R. astr. Soc., 38, 1-8, 1974.

Carter W.E., and Vincenty T., Survey of the McDonald Observatory radial line scheme by relative lateration techniques, Technical Report NOS 73 NGS-8,National Oceanic and Atmospheric Administration, 1978.

Hugget, G.R., Slater L.E., and Langbein J., Fault slip episodes near Hollister, California: initial results using a multiwavelength distance-measuring instrument, J.Geophys.Res., 82, 3361-3368 (1977).

Harrison J.C., Cavity and topographis effects in tilts and strain measurements, J.Geophys.Res., 81, 2, 319-328, 1976.

Harrison J.C., Herbst K., Thermoelastic strains and tilts revisited, preprint 1977.

Parm.T., Observing procedure and metereological factors for electronic distance measurements, paper presented at International Symposium on Electronic Distance Measurements, Int.Ass.of Geod., Lagos, 1975.

Parm T., High precision traverse for scale determination of satellite and stellar triangulation and for controlling the first order triangulation, paper presented at General Assembly, Int.Ass. of Geod., Grenoble, 1975.

Sacks I.S., Suyehiro S., Everston D.W., and Yamakishi Y., Sacks-Everston strainmeter, its installation in Japan and some preliminary results concerning strain steps. Pap.in Meteor. and Geophys., 22, 3-4, 195-208, 1971.

Sacks I.S., Snoke J.A., Yamakishi Y., and Suyehiro S., Borehle Strainmeters.Long-term stability and sensitivity to dilatancy, Carnegie Inst.Wash. Year Book,74, 1975.

Savage J.C., and Prescott W.H., Precision of geodolite distance measurements for determining fault movements, J.Geophys. Res., 78,6001-6008, 1973.

Savage J.C. and Prescott W.H., Geodimeter measurements of strain during the Southern California uplift, Preprint, 1978.

Slater L.E., Can the ratios of single wavelength EDM data improve the resolution of small changes in baseline length? A comparison with multiwavelength EDM data. Preprint, 1978.

Slater L.E. and Huggett G.R., A multiwavelength distance-measuring instrument for geophysical experiments, J.Geophys.Res., 81, 6299-6306, 1976.

Wyatt F., Berger J., Correlation of short based tilt measurements and long base strain measurements. Abstract, EOS, 59, 4, 331, 1978.

Page Intentionally Left Blank

# Detection of Horizontal Crustal Movements by Photogrammetric Methods

Luigi Solaini
Institute of Surveying, Technical University, Milano, Italy

The photogrammetric methods can be applied only if the horizontal displacements of the ground are of the order of several centimeters, because the standard error of the points so determined cannot, at the present time, be smaller than two centimeters.

The procedure to be adopted, which has already been applied with satisfactory results for the densification of geodetic nets, is the aerial analytical triangulation; therefore we shall here deal only about this method.

The taking can be executed with wide-angle cameras (f = 15 cm), or normal-angle cameras (f = = 30 cm) and format 23 x 23 cm. Since the average scale of photos must be chosen in relation with the desired accuracy and then is independent of the principal distance of the cameras, the number of photos is constant, whereas the flying height varies according to the focal length. We believe it advisable to use normal-angle cameras, because the lens have less distortion, better definition and are less influenced by the atmospheric refraction. Furthermore the small value of the ratio b/H does not remarkably influence the planimetric accuracy.

The scale of photograms must be large, because the measure error of the comparator is practically constant. However, an upper limit to the scale is given by the image motion (I.M.) which, at present, is not corrected and reaches sensible values if the photos are taken at low altitude.

Let us take into consideration the scales 1:3300 and 1:5000 which correspond to flying heights of 1000 m and 1500 m with f = 30 cm (or 500 m and 750 m with f = 15 cm) and let us compute the values of the I.M.

If the plane speed is 250 km/h which corresponds to about 70 m/sec and the exposure time is 0.0014 sec, the I.M. is 30 $\mu$ m for the lower flying height and is therefore larger than that generally accepted. At the higher height it reduces down to 20 $\mu$ m. This simple computation, together with the reduced number of the photos to be examined suggest, in our opinion, a flight 1500 m high, with focal length of 30 cm. At this height the effect of the atmospheric refraction is of a few $\mu$ m in the image plane and therefore the errors due to its approximate evaluation are negletable.

It is advisable to execute the strips with a minimum 60% forward and side overlap. The 60% side overlap is extremely useful in these types of high precision photogrammetry even when only the planimetric position is wanted because, by doing so, the image of the same point appears on several photograms.

It is also strongly advisable to plan two different sets of strips in two perpendicular directions both because the number of images of the same point is still increasing and because it looks like that some systematic errors depend on the direction of the flight.

If we consider a 5 x 5 km$^2$ block, photo-scale 1:5000 and 60% forward and side overlap, we will require 10 strips of 10 photograms each, that is to say a total of 200 photos, 100 in each flying direction.

We are therefore dealing with a small block which does not entail any computational difficulty. Even if the linear dimensions of the block become twice as much, the photograms to be measured go up to 800, which is still a relatively moderate number.

The control net must consist of a series of points, whose three coordinates have to be known, located along the block perimeter, at a distance from each other not larger than four times the base length, that is to say that in a flight taken from an altitude of 1500 m with a normal-angle camera, they must not be farther than 1.8 km. However, in the case of smaller blocks (i.e. 5 x x 5 km$^2$) it is better to choose the ground points nearer to each other so that their number will be sufficient for the adjustment. It is not required, but advisable, to determine some extra points in the central part of the block, provided they are situated in areas that are not liable to move, just as the control points in the perimetral zone should be.

The ground control net should be determined with the utmost accuracy, but it is hard to make the standard error of the coordinates go lower than the error of the points determined by means of the analytical triangulation. However, this is not very important because the successive computations will make use of the same points and of the same set of weights. The most important fact is that the ground known points be located in stable positions.

The signalization is on the contrary quite important. All the ground points, that is control and new points, must be signalized. It is generally accepted that the best type of signal is a light disk, yellow for instance, surrounded with a black square tar paper. The size of the disk has to be chosen in relation to the size of the

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

comparator measuring mark; if the diameter of the latter is 40 μm, the disk image can be about 50 μm so that, in the above mentioned hypothesis of 1:5000 scale photos, the actual diameter can be 25 ÷ 30 cm wide.

The measures will have to be executed with a high precision monocomparator such as, for instance, the Ascorecord by Zeiss Jena, the PK-1 by Zeiss Oberkochen, the Mann, and will have to be repeated in two perpendicular directions. Furthermore it is advisable to repeat each reading from two to four times.

The machine coordinates are transformed in the plate coordinate system by means of a rototranslation and a scale variation computed on the basis of the measured coordinates of the fiducials of the camera. We can remark in this instance, that it would be possible to use a grid camera, but recent experiments have shown that the increase in precision is very small.

Furthermore, the transformed coordinates are corrected for lens distorsion and comparator errors; as far as the latter is concerned, it must be noticed that the Ascorecord and the PK-1 have errors not larger than 1 μm. The refraction correction can be given to the plate coordinates or during the triangulation computations. In any case, it doesn't go beyond 3 μm.

The best method for the adjustment of block triangulation is no doubt the bundle method which, in the scheme we have here proposed, is very rigid, since each point can be measured on several photos.

Nowadays it is certain that the adjustment by the method of least squares is not completely satisfactory since it doesn't take into account the small systematic errors that are extremely harmful in the error propagation. The most extensively applied procedure for their correction consists in giving the plate coordinates some further corrections in form of polynomial functions of the coordinates themselves with constant coefficients (additional parameters) for all the strips or, at least, for those having the same direction. The conditions those parameters must satisfy are essentially two: to be significant and uncorrelated, both among themselves and the parameters of external orientation.

Some researches have been carried out in this direction and some expressions have been proposed; we can mention, among those which have proved successful, the ones given by Brown and by Ebner.

In the most recent experimental essays the standard error that can be obtained in each of the planimetric coordinates is not larger than 4 μm at the photo scale. If the photos have a 1:5000 scale, this figure corresponds to an error of 2 cm in the ground. This is a result that only few years ago no one would have dared to dream of.

Of course the standard error of the difference between coordinates referring to different times goes up to 2.8 cm, so that displacements of the order of 5 cm that groups of points have undergone in time, can be detected. However the most important thing is that the displacements be similar in value and direction; in other words a judgement cannot be made on the displacement of a single point, but one the average displacement of groups of suitably chosen points.

References

Slama, C.C., High precision analytical Photogrammetry using a special reseau geodetic lens cone, Presented paper to the Symposium of ISP, Commission III, Moscow, 1978

Brown, D.C., Accuracies of analytical triangulation in applications to cadastral surveying, Surveying and Mapping, 3, 281-302, 1973

Ebner, H., Self calibration block adjustment, Invited paper to the XIII Congress of the ISP, Commission III, Helsinki, 1976

Abstract. The use of analytical aerial triangulation is suggested, by particular rules, to determine horizontal displacements of soil. This method is suitable provided the coordinate variations be higher than 5 cm.

# Local Precision Nets for Monitoring Movements of Faults and Large Engineering Structures

Heinz G. Henneberg
Escuela de Geodesia, Universidad del Zulia
Maracaibo, Venezuela

Foreword. Although the title of the paper indicates a general presentation, I am exposing only some special cases from our country, Venezuela. Naturally, these special cases may contribute to general conclusions later on. It was a splendid idea from Dr. Charles Whitten to include in this presentation man-made big structures. As I see it now, in discussing crustal deformations, big structures cannot be excluded because they are affected too, as well as the geodetic nets which should control the structure. There exists a certain public interest in the stability and security of structures and we may further this interest for more support for our programs. We should see "local deformation and movement measurements" as a program which includes deformations and movements of structures like bridges, dams, tunnels etc., earth surface deformations as subsidences and horzontal components caused by water- and oil withdrawals and the natural neo tectonic movements and deformations; that is because sometimes there could be a superposition of components coming from different sources.

Abstract. Along Bocono Fault were installed local high precision geodetic nets to observe the possible horizontal crustal deformations and movements. In the fault area there are few big structures which are also included in the mentioned investigation. In the near future, measurements shall be extended to other sites of Bocono Fault and also to the El Pilar Fault. In the same way and by similar methods high precision geodetic nets are applied in Venezuela to observe the behavior of big structures , as bridges and large dams and of earth surface deformations due to industrial activities. This presentation gives some general and detailed information about the measurements and net installations at the following sites and the mentioned structures: (a) Bocono Fault and dam at Mitisus, (b) Bocono Fault at Mucubaji, (c) Bocono Fault and tunnel at Yacambu, (d) faults and dams at Uribante Caparo, (e) Guri Dam, (f) Maracaibo-Lake-Bridge (g) Orinoco River Bridge, (h) oilfield of Tia Juana, (i) Socuy-Tule dams.

Introduction. As a result of geological investigations the Bocono- and the El Pilar Fault separate the Caribbean Plate from the South American as is shown in figure 1. Between the ecological department of IVIC (Instituto Venezolano de Investigaciones Cientificas) and the geodetic department of Zulia University was established a geodetic investigation program to control the behavior of fault sites and of some big structures located in the fault area (fig.4). This program is supported principally by CADAFE (Empresa de Energia Electrica del Estado Venezolano), by Conicit (Consejo Nacional de Investigaciones Cientificas y Tecnologicas - Venezuelan research counsel), and by the engineering faculty of Zulia University. The observations of bridge behavior are made through a cooperation program with the MTC (Ministerio de Transporte y Comunicaciones), the Socuy-Tule dams with MARN (Ministerio del Ambiente y Recursos Naturales), the Guri Dam with EDELCA (Electrificacion del Caroni) as part of CVG (Corporacion venezolano de Guayana), the oilfield measurements with Maraven (Venezuelan oil company) and the Yacambu tunnel measurements with TRANARG (Venezuelan geodetic company). In figures 2 and 3 we see the sites of faults and structures, subjects of geodetic investigations. In figure 5 are shown in a schematic representation the geodetic net installations at Tia Juana oilfield, Socuy-Tule dams, Maracaibo-Lake-Bridge, Guri Dam, Uribante-Caparo project and Orinoco Bridge.

## Mucubaji

In the Mucubaji Fault area were installed two local precision nets: The principal net with 10.5 km maximum extention and the secondary net as a small extention net with maximum side lengths of 1.8 km (fig.6). The principal net covers laterally the total fault area and the observation stations are anchored in solid rock, as is shown in fig.6 (bottom). The small net extends around the visible fault trace and is located on moraines in consolidated soil or on individual rock sites. The foundation of the observation stations as shown in fig. 6 varies according to the soil conditions. All observation stations have the forced centering device for all types of instruments and targets. The two nets were measured as combined trilateration-triangulation. In the principal net only five distances could be measured in the first net determination due to topographic and weather conditions at that time. Two distances were measured several times with the Geodimeter Model 8 Laser and three distances several times with Tellurometer 1000. (It should be mentioned that the laser instrument can only be operated in two distances of the principal net due to its own weight and the very difficult accessibility of the stations). Angular measurements were made with T3 and DKM 3 theodolites and as targets were used specially constructed metal plates. The distances in the small net were measured with Mekometers, ELDI 2 and ELDI 3 instruments and angles were measured with T3 and DKM 3 instruments. The Mekometers were run by personnel from Prof. Linkwitz's institute at Stuttgart University in a cooperation program.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics*, October 2-5, 1978. Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

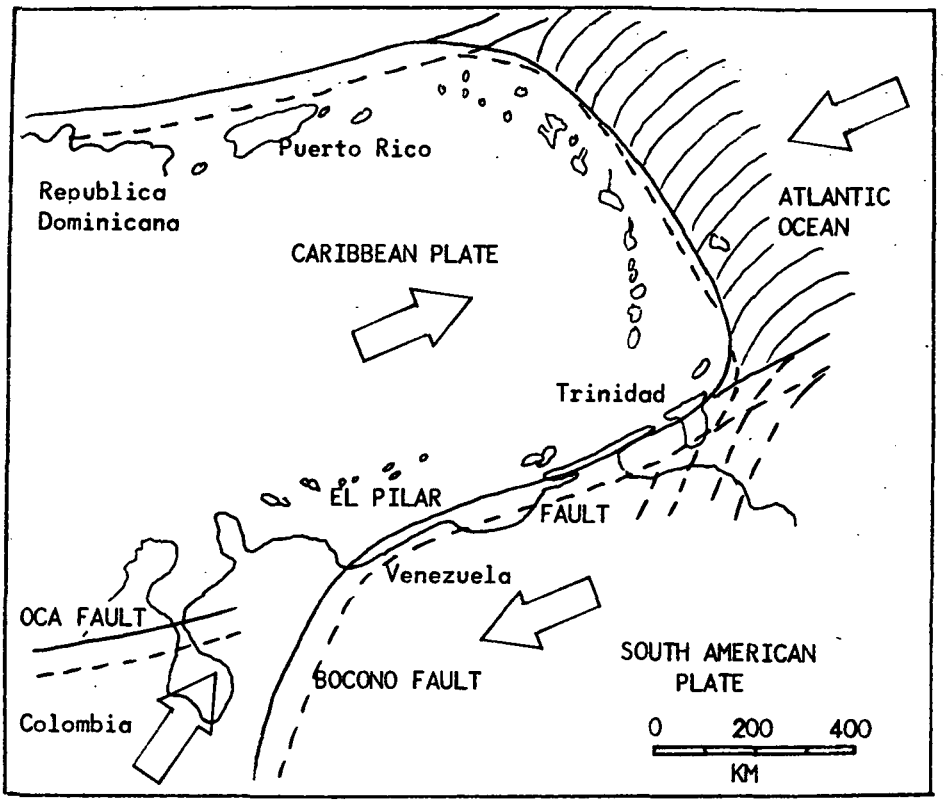



Fig. 1 Schematic representation of tectonic relations in the Caribbean area. The arrows show relative movements. The Caribbean Plate is seperated from the South American Plate by the Bocono and the Pilar Fault. (Molnar and Sykes 1969; Schubert 1970a).

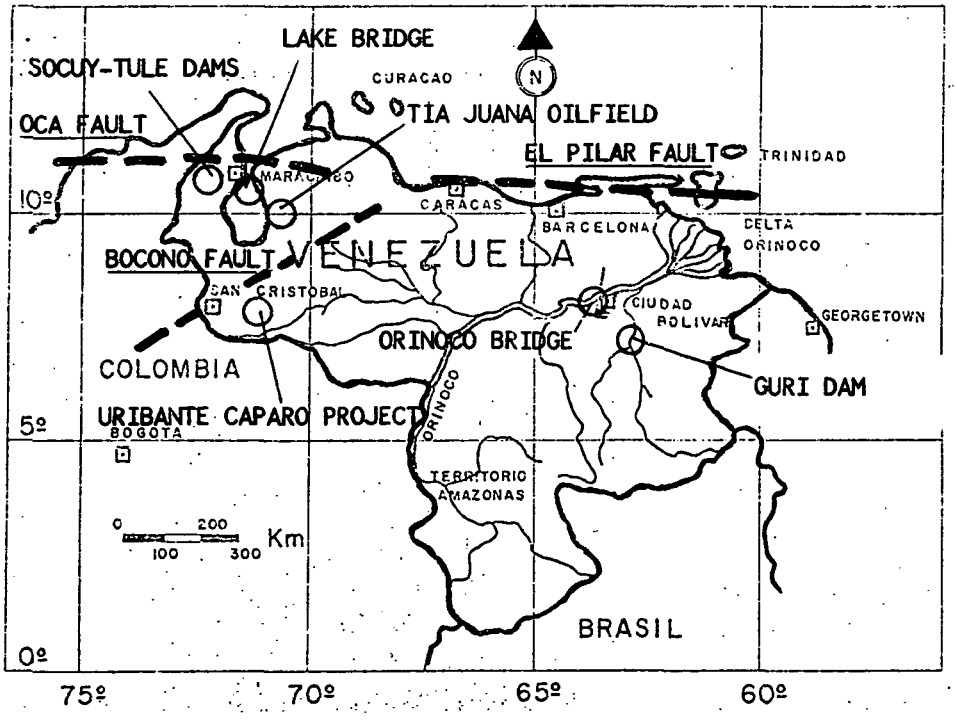



Fig. 2 Map of Venezuela, showing the principal faults and the sites of Socuy-Tule Dams, Lake Bridge, Tia Juana oilfield, Orinoco Bridge, Guri Dam and Uribante-Caparo project.

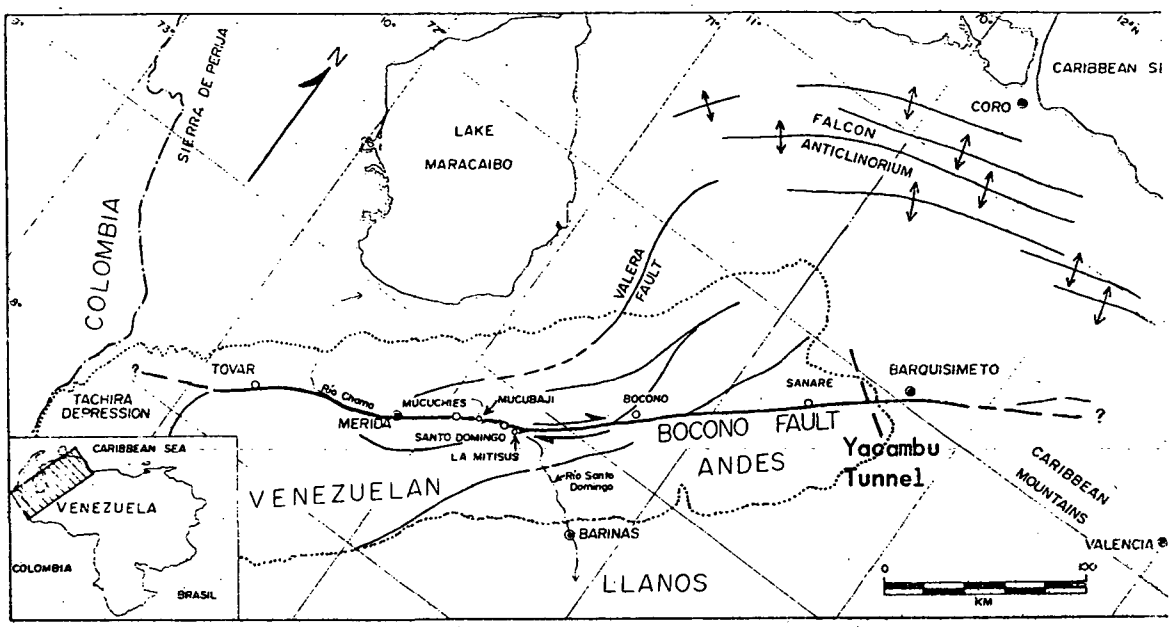



Fig. 3 Bocono Fault system in the Venezuelan Andes mountains showing the study sites Mucubaji, Mitisus and Yacambu.

Measurements started in spring 1975 and lasted about one year in the principal net. The small net was measured three times: 1975, 1976 and 1977. The 1975 and 1976 measurements are combined trilateration-triangulation with Mekometers and T3/DKM 3, and in 1977 only distances were measured with ELDI 2 and ELDI 3. In the principal net the free net work adjustments gave an accuracy range from ± 0.8 to ± 2.1 cm in coordenates (see table 1). In the secondary net the free net work adjustments gave accuracy ranges in coordenates: 1975 from ± 0.8 to ± 1.6 mm -Mekometer with angles-, 1976 from ± 0.7 to ± 1.4 mm -Mekometer with angles-, 1977 from ± 1.5 to ± 2.8 mm -only ELDI length measurements-, (see table 2).

TABLE 1. Computacion Results In The Principal Net At Mucubaji

| P.No. | x(m) | ±Mx(cm) | y(m) | ±My(cm) |
|---|---|---|---|---|
| 1 | 13579,780 | 0,8 | 7632,521 | 1,5 |
| 2 | 15378,922 | 1,2 | 13350,643 | 0,9 |
| 3 | 15343,666 | 0,9 | 17040,829 | 0,9 |
| 4 | 14664,429 | 0,9 | 17826,734 | 0,9 |
| 5 | 10000,001 | 1,1 | 17446,142 | 1,2 |
| 6 | 10583,620 | 1,0 | 14584,284 | 1,1 |
| 7 | 10259,229 | 1,1 | 10787,580 | 1,8 |
| 8 | 10000,006 | 0,9 | 10000,011 | 2,1 |

TABLE 2 Computation Results Of The Small Mucubaji Network

| P.No. | Year | x (m) | ±Mx (mm) | y (m) | ±My (mm) |
|---|---|---|---|---|---|
| 9 | 1975 | 9593,748 | 1,5 | 9843,714 | 0,8 |
| | 1976 | ,744 | 1,3 | ,714 | 0,9 |
| | 1977 | ,747 | 2,6 | ,717 | 2,0 |
| 10 | 1975 | 10000,002 | 1,0 | 10000,000 | 0,8 |
| | 1976 | ,005 | 1,0 | ,001 | 0.9 |
| | 1977 | ,005 | 1,5 | ,006 | 2,5 |
| 11 | 1975 | 10444,733 | 1,0 | 10099,985 | 0,9 |
| | 1976 | ,736 | 1,1 | ,986 | 0.8 |
| | 1977 | ,735 | 1,6 | ,983 | 1,9 |
| 12 | 1975 | 10671,059 | 1,5 | 10169,278 | 0,9 |
| | 1976 | ,056 | 1,4 | ,280 | 0,9 |
| | 1977 | ,056 | 2,3 | ,278 | 1,9 |
| 13 | 1975 | 10519,561 | 1,2 | 11445,089 | 0,8 |
| | 1976 | ,559 | 0,8 | ,089 | 0,7 |
| | 1977 | ,555 | 2,8 | ,086 | 2,1 |
| 14 | 1975 | 10400,740 | 1,6 | 11315,878 | 1,0 |
| | 1976 | ,741 | 0,9 | ,878 | 0,7 |
| | 1977 | ,742 | 2,5 | ,875 | 2,2 |
| 15 | 1975 | 10000.000 | 1,1 | 10985,165 | 0,8 |
| | 1976 | ,003 | 1,2 | ,163 | 0,9 |
| | 1977 | ,001 | 2,0 | ,166 | 1,7 |
| 16 | 1975 | 9465,427 | 1,0 | 10560,885 | 1,0 |
| | 1976 | ,426 | 1.0 | ,885 | 1,0 |
| | 1977 | ,429 | 2,5 | ,884 | 1,9 |

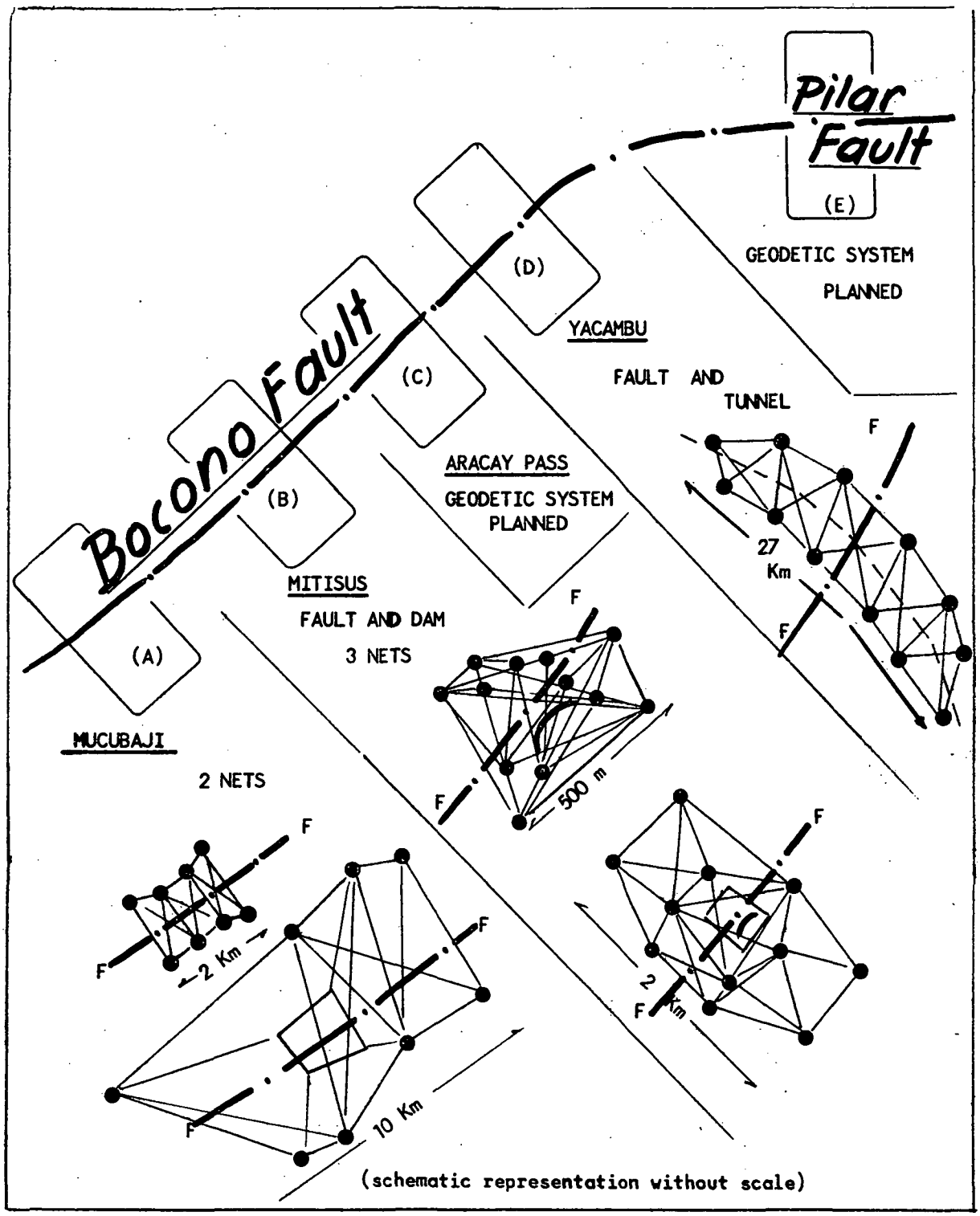



Fig. 4 Schematic representation of geodetic systems installed and planned along Bocono and Pilar Fault.

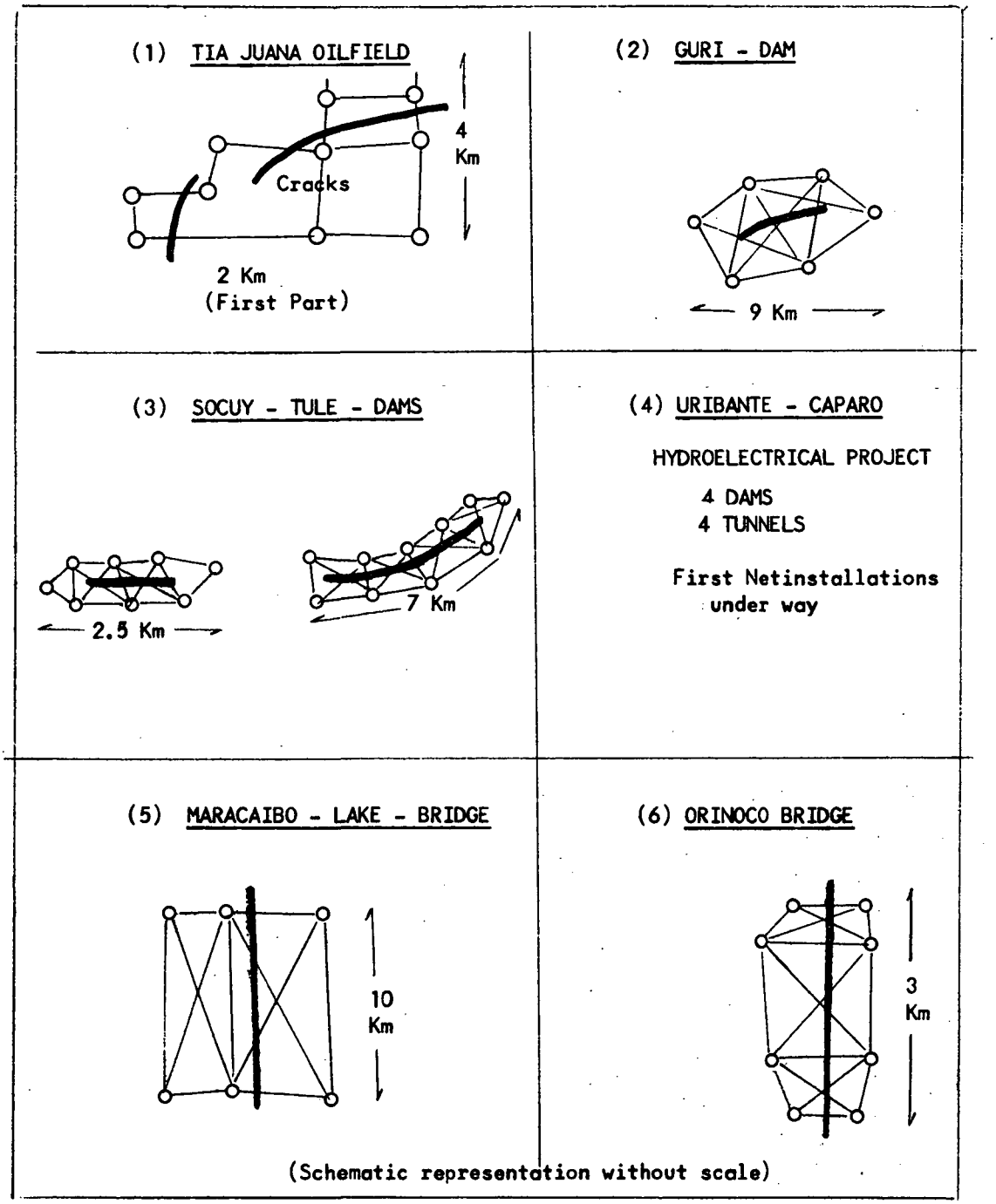



Fig. 5 Schematic representation of geodetic stystems to investigate movements of structures and earth surface.

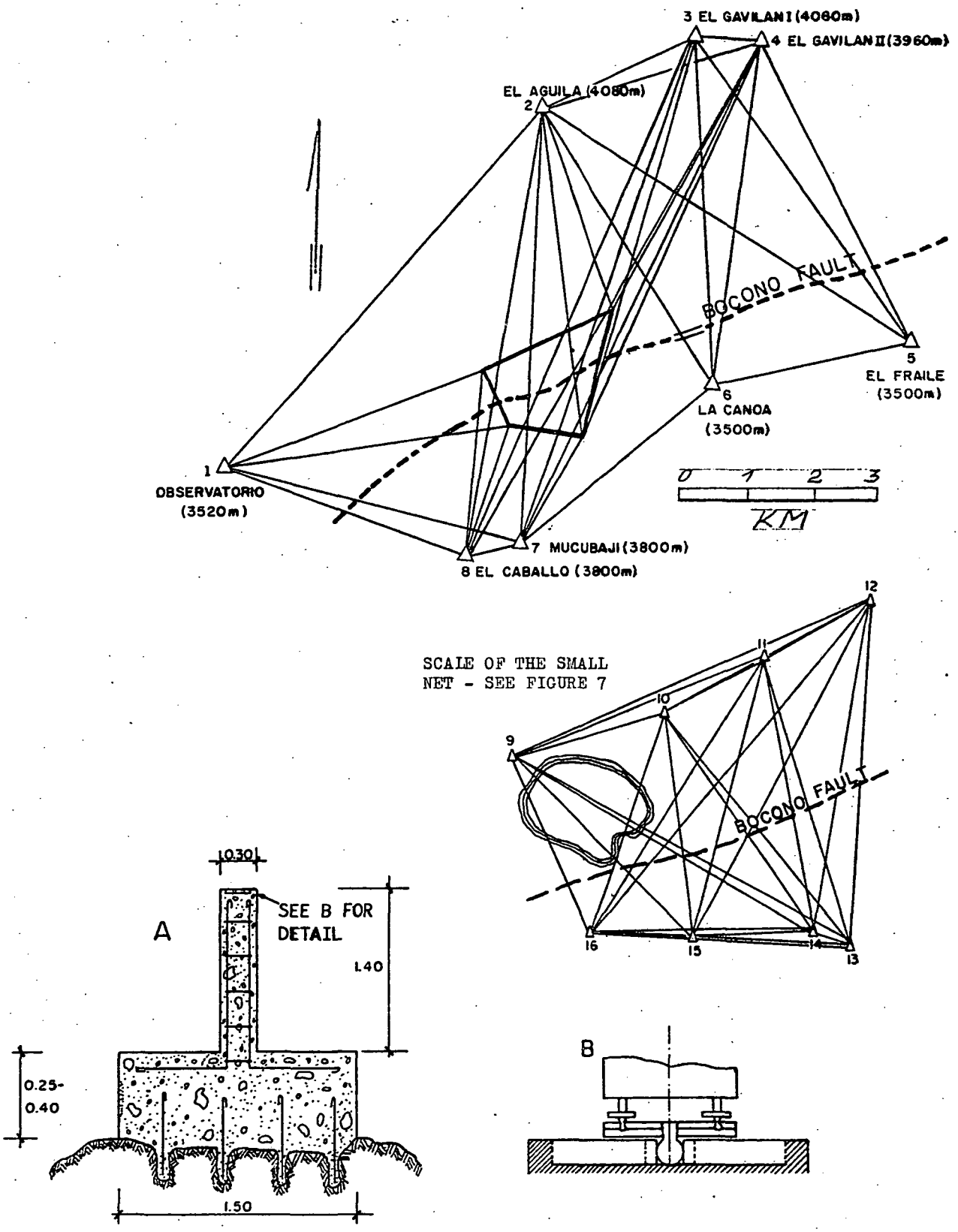



Fig. 6 Principal (large) and secondary (small) net at Mucubaji site including type of structure of observation stations.

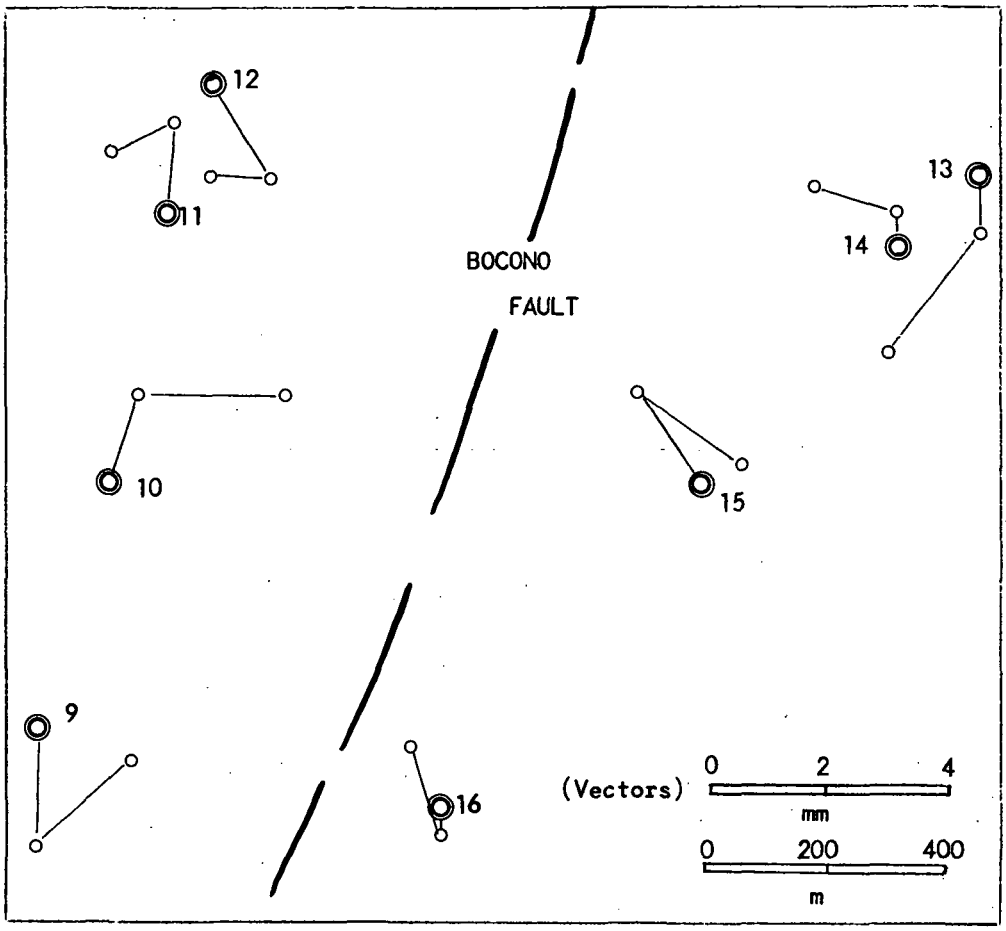



Fig. 7 Difference vectors as result of coordentate determination in the small Mucubaji net 1975 (initial), 1976, 1977.

La Mitisus

At "La Mitisus" (figure 3) we have three precision nets: 1. the principal dam system, 2. the "Pentagon" as security system of the dam system (the two systems are shown in figure 9) 3. the fault system as shown in figure 8, where the "Pentagon" mentioned before, constitutes the center part. The observation stations, constructed at each point, are similar to those installed at Mucubaji and as demonstrated in figure 6. Due to the different precisions, the measurements and computations are independent in the dam system and in the fault system (see tables 4 and 5). There are used two local coordenate systems: The fault system has $F_b$ as origin and $F_b - F_d$ as main direction; the dam system is computed in UTM with local coordenates. The two systems are connected in between through the "Pentagon" which belongs to the two coordenate systems. Measurements started in September 1973 at the dam site to observe the deformation of the structure and its surroundings which are situated in the Bocono Fault zone. The measurements for the dam are made through triangulations, using T 2 and sometimes T 3 instruments. The accuracies after adjustments after each triangulation vary between 0.3 to 1.4 mm in the coordenates. In this way the movement vectors of the damcrest are shown with high significance, as we can see in figure 10, where the movement vectors indicate the horizontal crest components of the dam deformation between low- and high water in the reservoir. In the figure the dashed line represents the dam axis. The dam is a doble curved concrete structure, 70m high and 210m long. The powerhouse lies about 16 km from the dam at a height difference of approx. 900m below. The energy capacity is of 240 MW. The fault system was measured first in 1974/5 and shall be observed again in 1979. The coordenates and accuracies are shown in table 4. The scale of the fault system was obtained through a trilateration in the Pentagon, applying a HP 3800 (see table 3). Later, T3 measurements completed the fault system observations.

Yacambu

The "Yacambu project" consists of two big structures: a doble-arc concrete dam of 150m height and a 24 km tunnel, which transports the water of the Yacambu river through the mountain chain of the Andes from south to north. The tunnel crosses Bocono Fault at Km 15 from the south entrance. Along the curved tunnel pro-

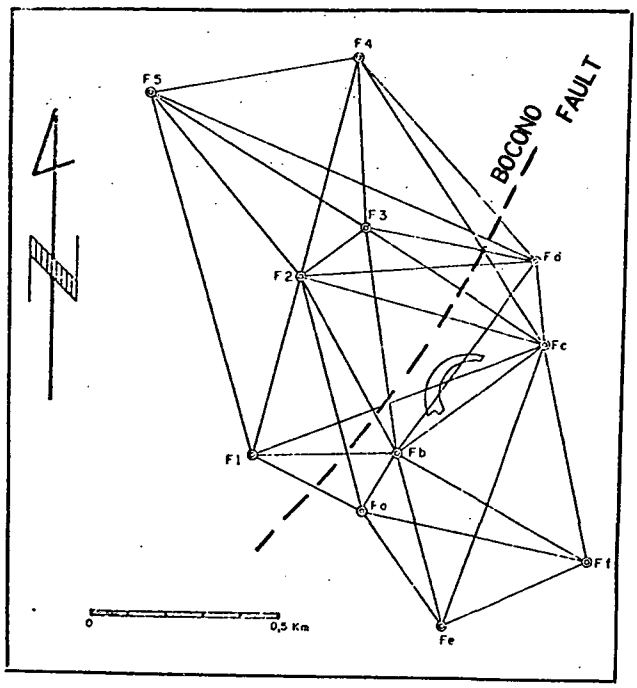



Fig. 8 Fault net at Mitisus

TABLE 3 Double Distance Measurements With HP 3800 In The Pentagon

| D | 1 | 2 | M | D - Hor. |
|---|---|---|---|---|
| Fb -Fd | 656,624 | ,627 | ,625$_5$ | 656,606 |
| Fb -Fc | 490,599 | ,604 | ,6015 | 487,530 |
| Fb -F2 | 509,092 | ,100 | ,096 | 508,827 |
| Fb -F3 | 580,804 | ,806 | ,805 | 580,749 |
| F2 -Fc | 667,522 | ,527 | ,5245 | 666,427 |
| F2 -Fd | 627,080 | ,087 | ,0835 | 626,978 |
| F2 -F3 | 202,222 | ,220 | ,221 | 202.044 |
| F3 -Fc | 527,926 | ,927 | ,9265 | 571,018 |
| F3 -Fd | 465,020 | ,023 | ,0215 | 465,012 |
| Fd -Fc | 266,246 | ,245 | ,2455 | 261,551 |

TABLE 4 Coordenates And Accuracies In The Fault Network.

| | x | ± Mx(mm) | y | ± My(mm) |
|---|---|---|---|---|
| Fa | 9818,471 | 1,9 | 10023,178 | 1,3 |
| Fb | 10000,000 | 0 | 10000,000 | 0 |
| Fc | 10457,216 | 2,8 | 10169,259 | 2,6 |
| Fd | 10656,609 | 1,7 | 10000,000 | 0 |
| Fe | 9800,972 | 2,7 | 10449,179 | 3,2 |
| Ff | 10097,675 | 2,9 | 10625,525 | 2,9 |
| F1 | 9741,673 | 2,3 | 9654,599 | 2,2 |
| F2 | 10226,119 | 2,1 | 9544,175 | 1,4 |
| F3 | 10420,469 | 2,0 | 9599,408 | 1,4 |
| F4 | 10904,227 | 2,9 | 9101,358 | 2,8 |
| F5 | 10383,998 | 1,7 | 8909,294 | 3,3 |

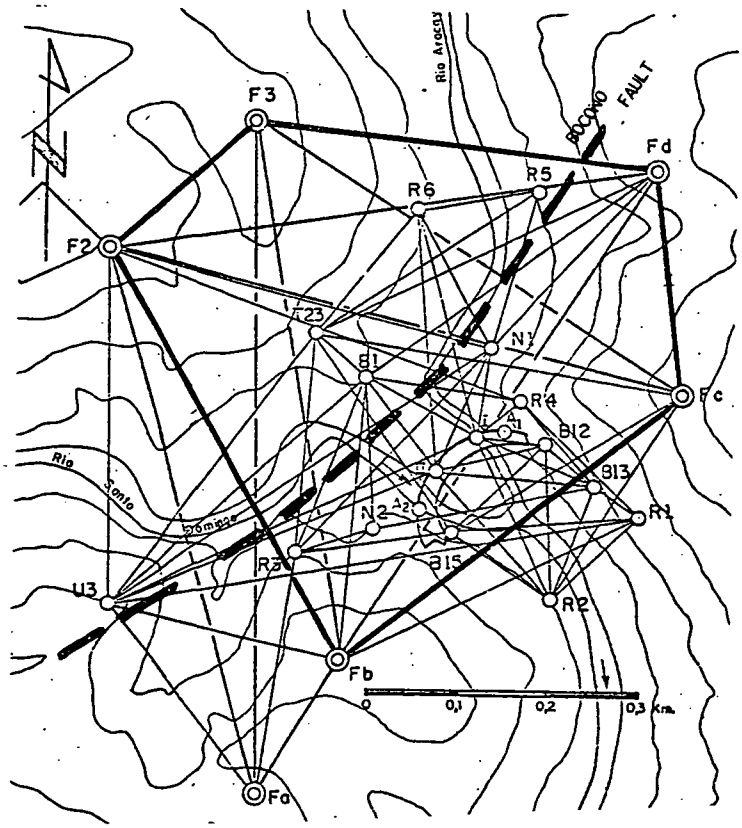



Fig. 9 Santo Domingo Dam triangulation with security system (pentagon F2, F3, Fb, Fc, Fd as part of the fault net).

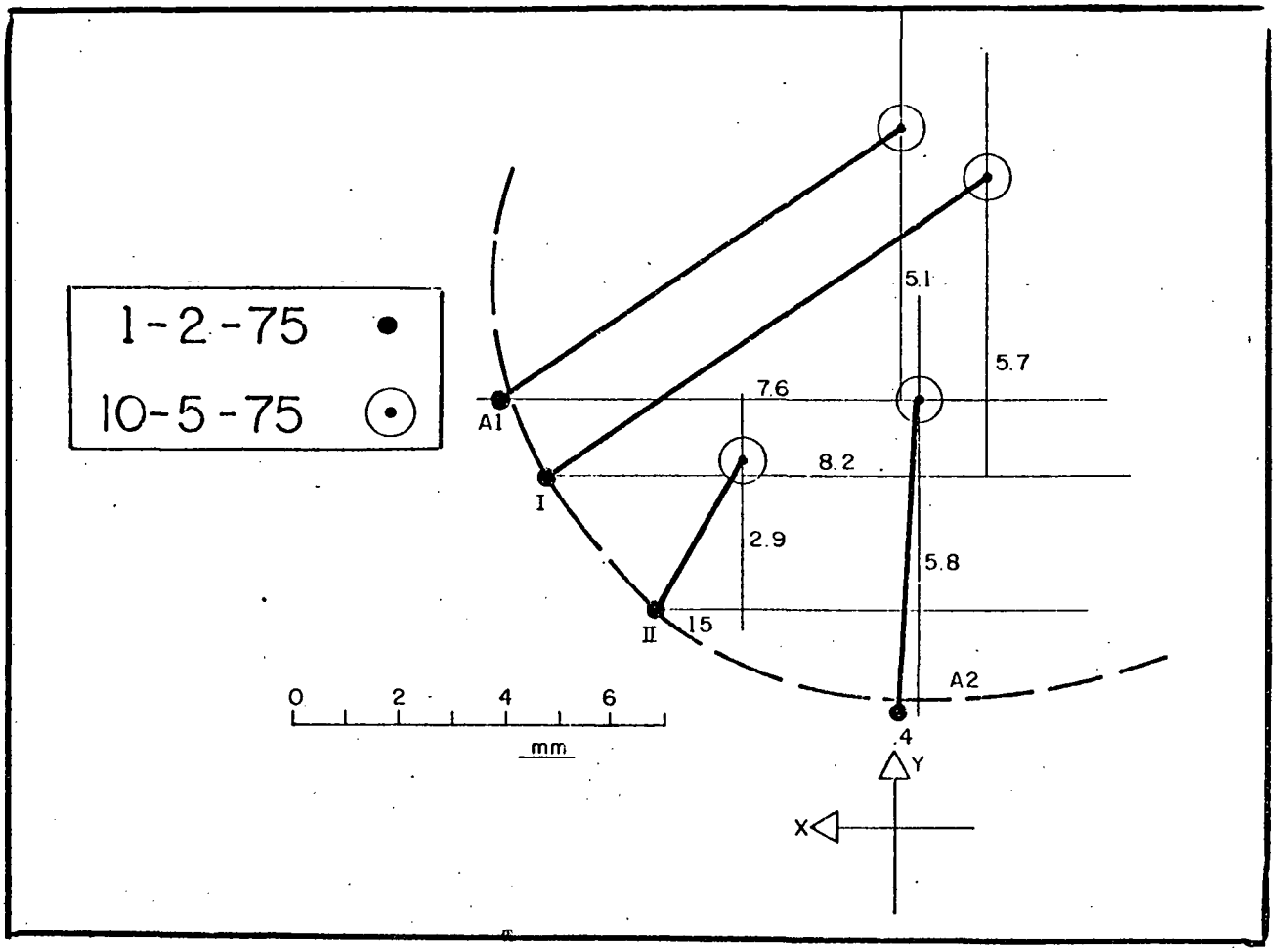



Fig. 10 Movement vectors of the dam crest between February and May 1975 (low- and high water in the reservoir).

jection a typical combined trilateration-triangulation was installed. The narrow band of the net work has the advantage of a high precision traverse: strong lateral rigidy by angular measurements and strong longitudinal rigidy through the distances (figure 11). For the distance measurements was used an "Electrotape". There were measured 45 distances and 76 angles, having thus 121 observation equations for 34 unknowns. For the final computations, shown in this paper, the "free network adjustment method" was used. Measurements were made at two epochs: 1973 and 1975. Between the two measurements strong coordenate differences showed a change of sign, where the network passes Bocono Fault (table 6). The question arose, if these coordenate differences were significant enough in comparison with the coordenate accuracies to determine a possible crustal deformation, due to Bocono Fault. As shown in table 6, the maximum value in coordenate changes is 52mm in x-direction at point 6, 47mm in y-direction at point 10. The maximum difference

TABLE 5 Coordenate Accuracies of Three Independant Determinations Of 2 Fault - points(N1,N2) And 2 Dampoints(I,II) In The Dam Network (in mm)

| | 1 | | 2 | | 3 | |
|---|---|---|---|---|---|---|
| Station | ± Mx | ± My | ± mx | ± My | ± Mx | ± My |
| N1 | 0,9 | 0,9 | 1,3 | 1,4 | 0,6 | 0,7 |
| N2 | 0,4 | 0,6 | 0,8 | 1,2 | 0,4 | 0,6 |
| I | 0,4 | 0,6 | 0,1 | 0,8 | 0,3 | 0,5 |
| II | 0,4 | 0,4 | 0,6 | 0,6 | 0,4 | 0,4 |

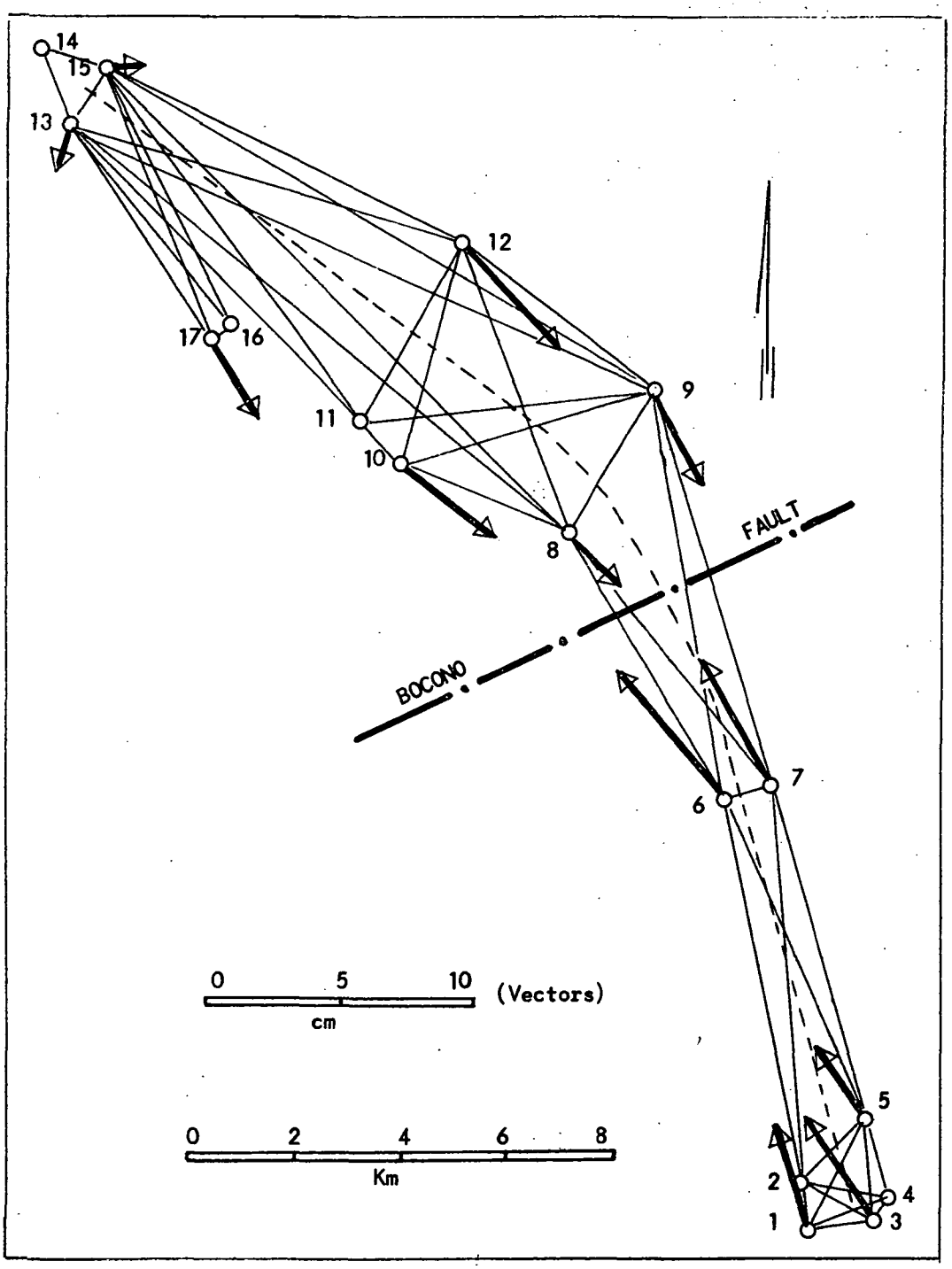



Fig. 11 Combined triangulation-trilateration for Yacambu tunnel and difference vectors 1973 - 1975

vector exists atpoint 6 with 63 mm. In the mentioned point the accuracies in the coordenates by single determination are 20 to 25 mm. We see that the coordenate accuracies show errors still too high for real significant movement. Nevertheless, at points 6 and 7, where we have strong difference vectors south of Bocono Fault, we may come to a certain expression about the existence of movement. At these points the cofactors of the variance - covariance matrix are (values from 1973): point 6 - Qxx = 0.14, Qyy = 0.11, Qxy = - 0.0037, point 7 - Qxx = 0.13, Qyy = 0.13, Qxy = - 0.0026. The low covariance values in comparison with the variance values show very little correlation. At these points the computation of the accuracy values for the difference vectors gave: point 6 - Difference vector D = 63mm, MD = ± 32mm, point 7 - difference vector D = 56mm, MD = 31mm. Thus, we can say, that the accuracy values expressed in "mean square errors" have the half size in comparison with the vectors. This is true only for points 6 and 7. As we may deduct from table 6, in the other points the relation is not so good. Therefore, I incline to say that it is still too early to speak of significant movements in the fault area. But it is still interesting to outline: 1. the change of sign of the coordenate differences at Bocono Fault trace, 2. the directions of vectors, showing all toward Bocono Fault trace, 3. increasing values of vectors approaching Bocono Fault, specially in points 6, 7, 10, 12. Therefore, we may interprete these results as an obvious trend situation, showing possible compression in the fault surroundings. A trend analysis, as used sometimes in gravity measurement computations, may be very useful for the interpretation of the existing results. Further measurements shall bring more information.

## Uribante - Caparo - Project

The Uribante - Caparo project is one of the largest hydroelectrical enterprises on the continent. Several net work installations are under way at this moment to investigate the behavior of dam- and tunnel structures, their surroundings and the existence of fault traces in the area. The net installations are very similar to those discussed before. Definite information about measurements, instrumentation and results shall be given at later events. The project consists in the interconnection of four large dams from Uribante River at 1104 m above sea level down to Caparo River at 277 a.s.l. The interconnections are realised through tunnel constructions. The dimensions of the structures are: La Honda Dam - L= 450m, H= 118m; Las Cuevas Dam - L= 740m, H= 106m; Borde Seco Dam- L= 340m, H= 108m; La Vueltosa Dam - L= 430m, H= 118m; Uribante Doradas Tunnel - L= 7884m; Doradas Camburito Tunnel - L= 4484m; Agua Linda Doradas Tunnel - L= 5500m; Camburito Caparo Tunnel - L= 640m. The total hydroelectrical capacity of the final project is 2.260 MW. The project is situated in the Tachira, Merida y Barinas states of Venezuela in the west, south of Maracaibo Lake, near San Cristobal, capital of Tachira. The total length of influenced area of dams, tunnels and reservoirs comprises about 80 km.

## Guri

Guri Dam is the biggest hydroelectrical installation in Venezuela and shall be at completion one of the biggest on the continent. At this moment the energy capacity is of 2.100 MW, at the final stage of construction will be about 9.000 MW. The dam is situated on Caroni River in the south eastern part of Venezuela (see figure 2). The Caroni River contributes to the Orinoco River from the south and is the second river of the country. The hydroelectrical generation is principally used for the development of the industrial area, situated near Ciudad Guayana at the entrance of Caroni in the Orinoco River. The act ual length and maximum height of Guri Dam are: L= 690m, H= 106m (main dam), the final length and max. height are: L= 6-7 km, H= 160m (see figure 12). The first geodetic measurements in situ started in 1956. 1950 Cartografia Nacional (national survey) completed the Caroni triangulation. From that time on several local net works were used for construction surveys and behavior measurements. Actually, precision measurements through triangulations and levellings are applied to determine the behavior of the dam structure. In figure 13, we see the results in a three dimentional representation of the movement of point 19, situated on the right side rock fill dam. This movement is the consequence of the consolidation deformation of the dam, as we can note from the down-slowing motion of the different components. The movements as functions of time have already an asymtotic approach. For the final project development of Guri all existing geodetic network shall vanish because of the extention of the new reservoir at 54 m uplift of lake level. Therefore, a new observation system was planned and is actually in execution as shown in figure 12. This network is a combined triangulation - trilateration system and in the observation points will be concrete pillars as shown in Mucubaji and Mitisus.
This geodetic program is going to work in combination with geophysical measurements, specially with the microseismological net already prepared in the dam and reservoir area. The extention of the observation net is planned, as shown in figure 12, to the surroundings of the future lake.

## Maracaibo Lake Bridge

Maracaibo Lake Bridge has a length of 9 km and is situated in the southern part of Maracaibo city. The bridge spans Maracaibo Lake at one of the narrowest parts. The maximum height of the bridge table is 50m above lake level at the bridge site. The bridge is a pre-

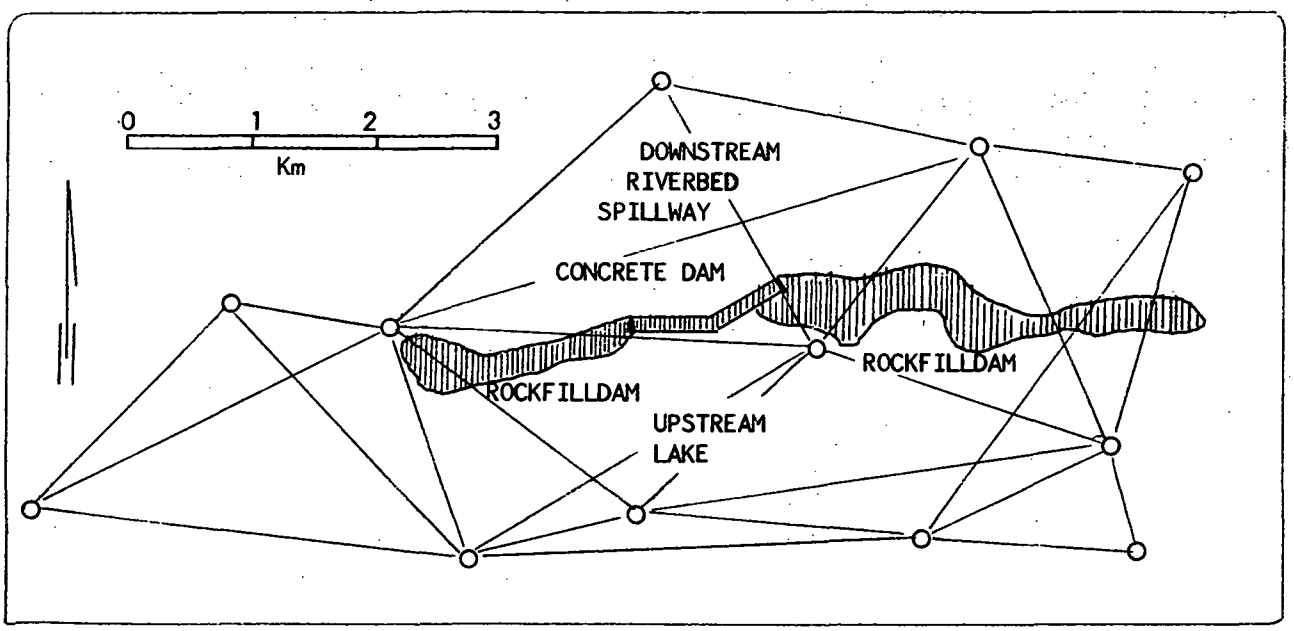



Fig. 12 Geodetic observation system for the final construction stage of Guri Dam

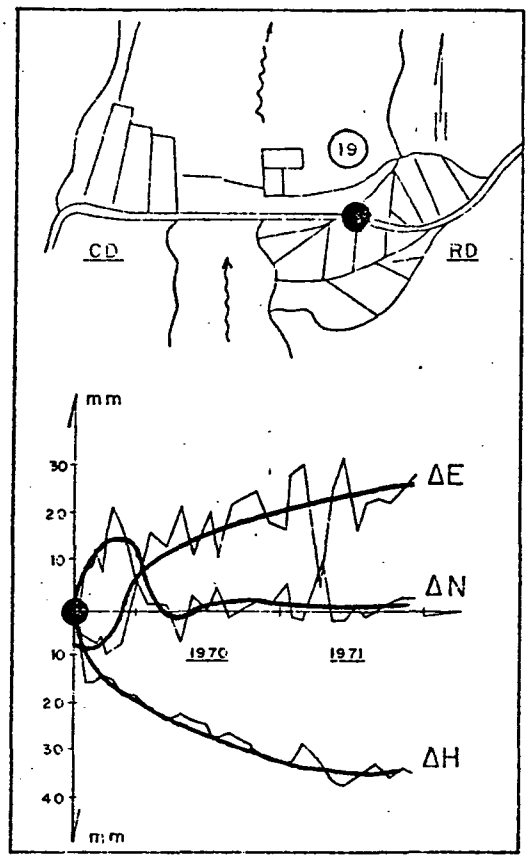



Fig. 13 Accumulated 3 dimensional movements of point 19 located in the actual right side rockfill dam of Guri.

TABLE 6 Results Of The 1973 And 1975 Net - work Computations At Yacambu

| P. | 1975 - 1973 x (mm) | y (mm) | D (mm) | 1973 ±Mx (mm) | ±My (mm) | 1975 ±Mx (mm) | ±My (mm) |
|---|---|---|---|---|---|---|---|
| 1 | 44 | -14 | 46 | 27 | 24 | 29 | 26 |
| 2 | 35 | -15 | 38 | 27 | 22 | 29 | 23 |
| 3 | 45 | -28 | 53 | 26 | 22 | 28 | 23 |
| 4 | 41 | -31 | 51 | 26 | 21 | 28 | 23 |
| 5 | 26 | -27 | 37 | 27 | 17 | 29 | 19 |
| 6 | 52 | -36 | 63 | 22 | 20 | 24 | 23 |
| 7 | 49 | -27 | 56 | 21 | 21 | 23 | 22 |
| 8 | -21 | 20 | 29 | 19 | 20 | 22 | 20 |
| 9 | -37 | 11 | 38 | 22 | 18 | 24 | 21 |
| 10 | -31 | 47 | 56 | 22 | 24 | 24 | 25 |
| 11 | -24 | 40 | 47 | 20 | 22 | 21 | 25 |
| 12 | -43 | 33 | 54 | 21 | 19 | 24 | 20 |
| 13 | -19 | 9 | 21 | 28 | 16 | 31 | 18 |
| 14 | 7 | 4 | 8 | 31 | 25 | 34 | 27 |
| 15 | -1 | 12 | 12 | 25 | 19 | 28 | 21 |
| 16 | - | - | - | - | - | - | - |
| 17 | -30 | 19 | 35 | 36 | 32 | 36 | 33 |

TABLE 7 Height Of Maracaibo Observation Towers And Visuals Over Lakelevel

| Station | Construction Height | Visual over Lake Level |
|---|---|---|
| Camacho | 11,50 m | 12,00 m |
| San Francisco | 11,50 m | 12,50 m |
| Palmarejo | 11,50 m | 17,00 m |
| Iguana Sur | 9,50 m | 14,00 m |
| Iguana Norte | 9,60 m | 11,50 m |
| Redonda | 9,50 m | 18,00 m |
| Isla de Aves | 7,50 m | 12,00 m |
| Punta Piedras N | 7,50 m | 20,00 m |
| Manzanillo | 7,50 m | 27,50 m |

stressed concrete structure, where beams of 46.6 m length were prefabricated on shore and used later on in the construction process. For the horizontal control of the bridge was installed a triangulation, as we see in fig. 14, with error ellipses in the outer stations. The ellipse values are in mm. At each net point was erected an observation tower of 7.50 m to 11.50 m height (see figure 16). The forced centering device is the same as used in the other networks described before and shown in figure 6. The different construction heights of observation towers and the height of visuals above lake level is to be seen in table 7. The network is a local system, with coordenate origin at bridge pillar 1: $x_o$ = 10 000 000, $y_o$ = 10 000 000. The following instruments were used for the net installation: Tellurometers, T3 and DKM3. It was necessary to undergo an extensive research program to investigate the strong atmospheric perturbances in the bridge site and lake area. Horizontal- and vertical refraction investigations showed periodical influences as functions of time in optical measurements. For electromagnetic measurements with Tellurometer it was necessary to determine the refraction index distribution over the lake area to obtain minimum perturbanced values in the side measurements.

## Orinoco River Bridge

Orinoco River Bridge is a suspension bridge of L = 1272m with a main span of 712m. The total bridge length with connected concrete spans is L = 1679m. The height of bridge towers is 140m above cero river level, and of the bridge table 68m. The maximum water depth at bridge site from cero water level is 45m and the high water range from cero water level is about 20m. Due to this big range of water level we distinguish two characteristic streambeds: the low water bed and the high water bed (see figure 15). The level change is periodical: low water in March, high water at the end of August. The medium level change per day is 10 cm. These special conditions were essential for the network planning. The network, as seen in figure 15, has two parts: the main system, surrounding the suspension bridge and located around the low water bed (figure 16), and the security system on higher parts which are islands in the high water zone, and in the north two stations outside the high water bed. The security system has the principal purpose to serve as a possibility to reproduce local net deformations of the main net. In the net points were used concrete towers from 3 to 12 m height, as shown in figure 16, with forced centering device as mentioned before. The coordenate system is a local one as applied at Maracaibo Bridge. The point accuracies are shown in figure 16 by error ellipses (max. and min. values 2.6mm, 0.9mm).

## Tia Juana Oilfield

As a consequence of the withdrawal of oil from the approx. 500m deep soil strata, a considerable subsidence of the ground surface of the oilfields has occured. These subsidence values are now about 4 meters at the center of the three main oilfields (figures 17, 18): a. Pueblo Viejo - Bachaquero, b. Lagunillas, c. Tia Juana. The 1974 - 1976 rate of movement has been at: a. Pueblo Viejo - Bachaquero - 17 cm (max), b. Lagunillas - 39 cm (max), Tia Juana - 33 cm (max). The subsided parts on land are protected by dyke constructions permanently under supervision. The height of the dykes increases according to the subsidence of the area involved. Figure 18 shows the subsidence cone at Tia Juana Oilfield. Caused by this subsidence, horizontal cracks of width up to 1m have appeared at the edges of the cone as shown in figure 19 where the dashed lines show the location of these cracks. The subsidence measurements started at Lagunilla site in June 1926, at Tia Juana site in November 1937 and at Pueblo Viejo-Bachaquero site in April 1938. At this moment, measurements of the total area are carried out every two years by the "Departamento de Topografia de Maraven". Before the nationalization of the Venezuelan oilindustry the measurements were made principally by "Shell Oil Company of Venezuela". Between Maraven and the geodetic department of the University of Zulia a new, more complete measurement program was establishd, which included a "geodetic-geophysical study of the subsidence area". The additional investigations to be carried out are: 1. Horizontal control by high precision geodetic network, 2. gravity measurements, 3. seismic measurements and registration, 4. study of the Z-component of the geomagnetic field. The horizontal control is carried out to observe horizontal components of the subsidence movement. The observations are made through a high precision traverse of various interconnected quadrilateral figures of the Hollister type, in which are measured the outer sides with the electro-optical method (Zeiss ELDI 2) and the inner and outer angles at the traverse points with Wild T3 and Kern DKM3 theodolites. The diagonal lines of the Hollister figures are not mea - sured, but computed afterwards (figure 19). All measurements are made from reinforced concrete columns with forced centering device. The accuracy of adjusted measurements gives a mean square error in coordenates of about $\pm$ 3 mm (table 8).

## Socuy - Tule Dam System

The two dams, Socuy and Tule, are located in the north-west of Maracaibo (figure 2). They are constructed as rockfill and earth dams and are interconnected through a tunnel and an open canal. The rivers controlled by these dams are the Socuy and the Cachiri River which are coming from the mountain range between Venezuela and Colombia. The dimensions of

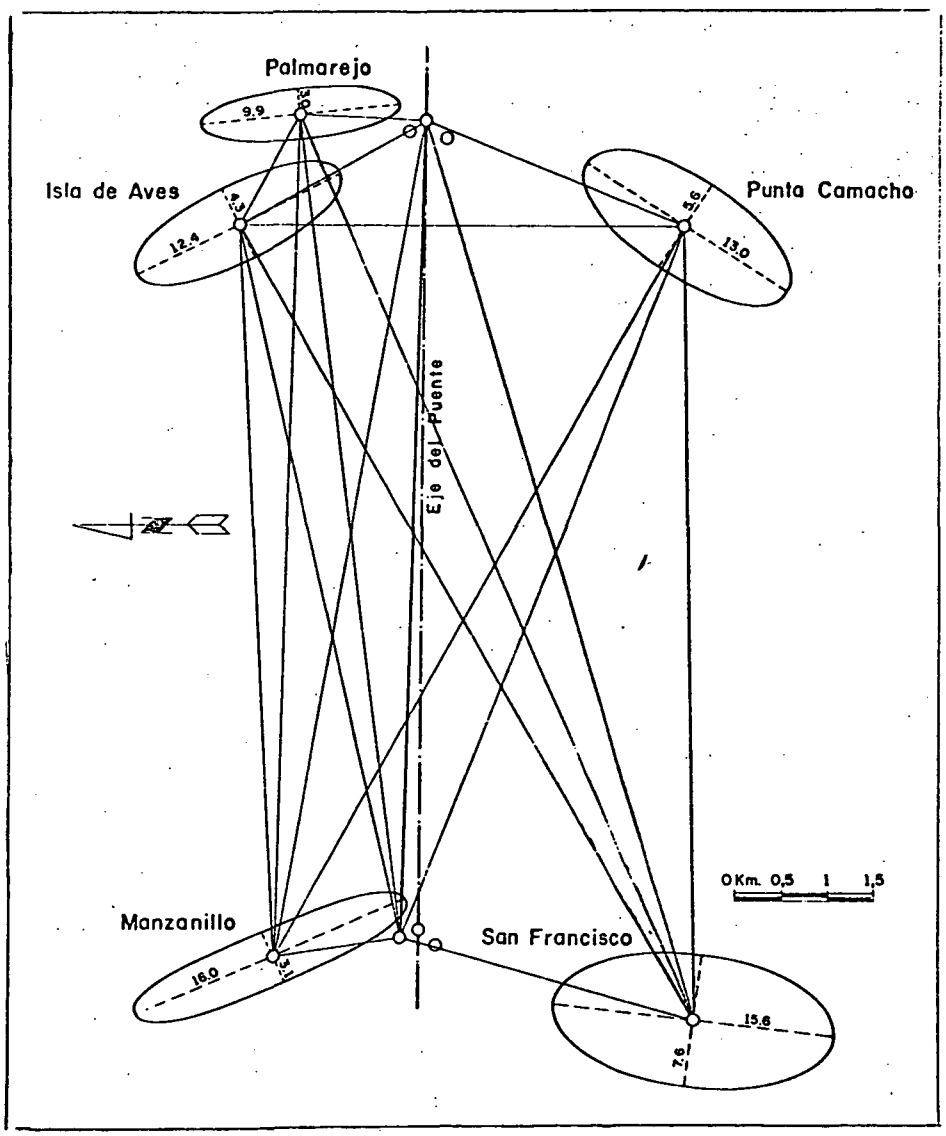



Fig. 14 Maracaibo Lake Bridge triangulation with error ellipses.

TABLE 8 Coordenate Accuracies In The High Precision Traverse

| Station | ± Mx (mm) | ± My (mm) |
|---|---|---|
| U 1 | 3,0 | 2,6 |
| U 2 | 2,7 | 2,6 |
| U 3 | 2,3 | 2,9 |
| U 4 | 2,1 | 2,8 |
| U 8 | 1,9 | 2,5 |
| U 9 | 2,7 | 2,4 |
| U 10 | 3,2 | 2,6 |
| U 11 | 3,3 | 3,1 |
| U 12 | 2,8 | 3,3 |

TABLE 9 Coordenate Accuracies Socuy Net

| Station | ± Mx (mm) | ± My (mm) |
|---|---|---|
| 1 | 1,7 | 0,7 |
| 2 | 1,1 | 0,5 |
| 3 | 1,1 | 0,5 |
| 4 | 2,4 | 1,1 |
| 5 | 1,7 | 0,4 |
| 6 | 1,6 | 0,5 |
| 7 | 1,7 | 0,7 |
| 8 | 2,8 | 1,4 |
| 9 | 2,7 | 1,0 |
| 10 | 2,7 | 1,0 |
| 11 | 0,9 | 0,5 |
| 12 | 1,0 | 0,6 |
| 13 | 1,7 | 0,7 |
| 14 | 0,9 | 0,5 |

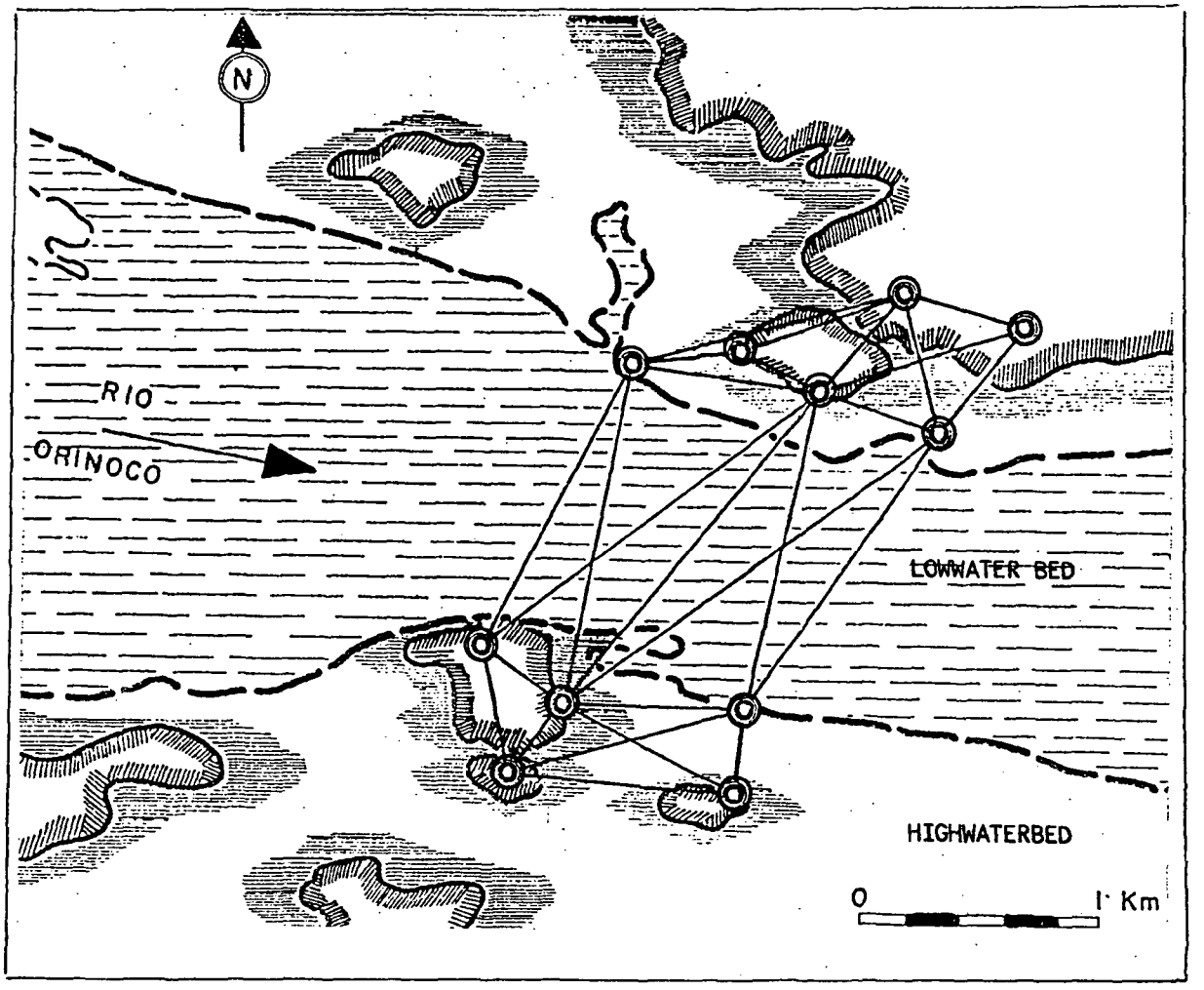



Fig. 15 Location of Orinoco Bridge triangulation.

Socuy dam are: main dam - L= 1300m, H= 40m; side dam - L= 1600m, H= 14m, and the dimensions of Tule dam: L= 5200m, H= 20m. The tunnel length is 1035 m and the length of the canal is 5240 m. For this large extention complex there were located several local geodetic systems. We distinguish four different control systems: Network 1 with ground fixed stations, as combined triangulation-trilateration, serves as controlnet for the surroundings of Socuy dam and lake. Network 2, as shown in figure 20 (1), with pillar observation stations similar to those at Mucubaji, only with different foundations, serves as observation net to investigate the deformation of Socuy main dam. Network 3, as shown in figure 20 (2), with pillar observation stations is the observation net to investigate deformations of Tule dam. Network 4, with ground fixed stations, is a high precision traverse to control the tunnel and canal and to connect the different dam networks -see figure 20 (3)-. Networks 2 and 3, too, are combined trilateration - triangulations. Distances were measured in net 1 with DI 10 and ELDI 3, in the other nets only with ELDI 3. All nets were computed as free networks and combined together through coordenate transformations and later on transformed also to the national system - figure 20 (4)-. Table 9 shows the results of the Socuy adjustment: instruments - ELDI 3, DKM 3; observation equations: 296, unknowns:28

Acknowledgments. Part of the research work described in this paper was supported by CADAFE (Venezuelan electrical power institution), CONICIT (Venezuelan research counsel), Maraven (Venezuelan oil company), MARN and MTC (Ministry of Natural Resources and Ministry of Transportation - former Ministry of Public Works), CVG (Venezuelan Guayana Corporation). The author thanks the following persons und institutions for their cooperation: K. Baer, C. Schubert, K. Linkwitz, H. Boettinger, H. Arp, J. Fischer, H. Drewes, A. Sandoval, Cartografia Nacional, Tranarg company, Carl Zeiss-Oberkochen, students and personnel of the geodetic department and all other who helped directly and indirectly.

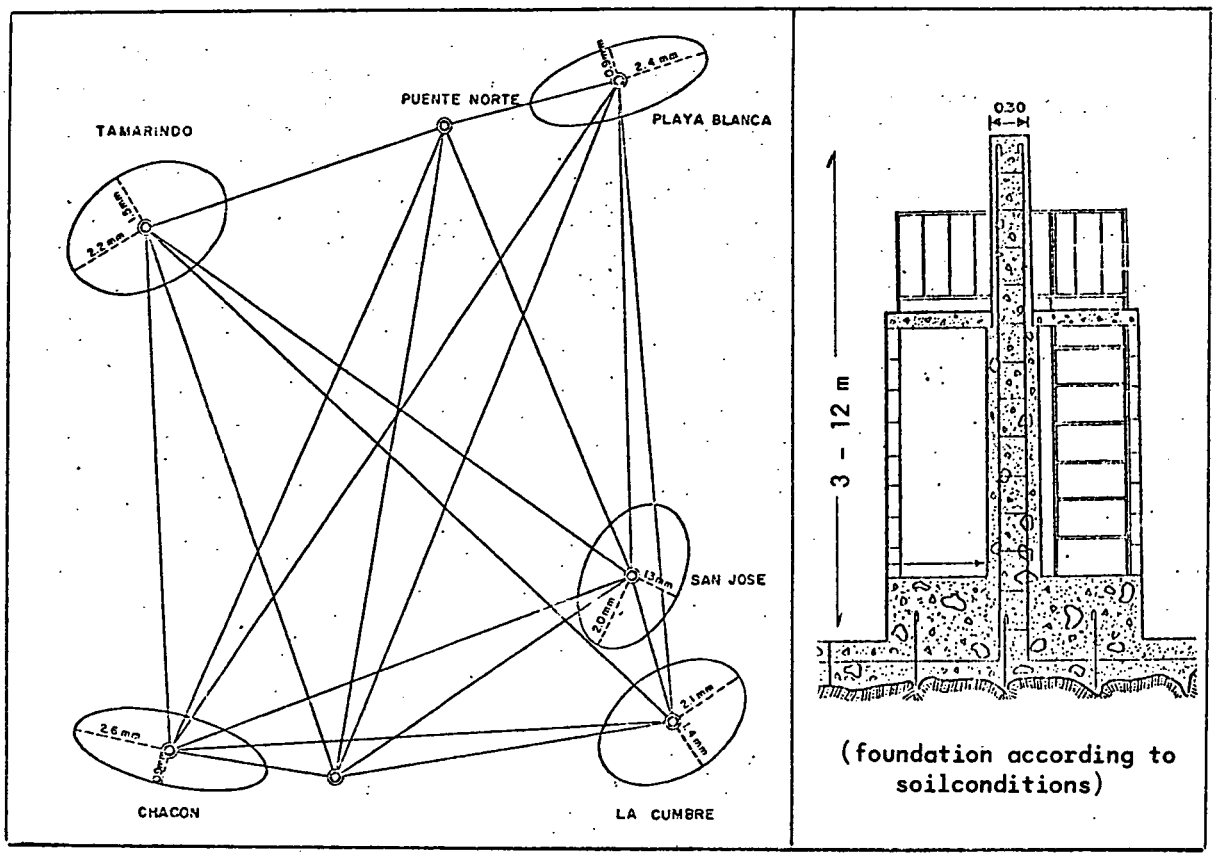



Fig. 16 left: Main net of Orinoco Bridge triangulation with error ellipses. right: Construction type of observation towers for Maracaibo and Orinoco Bridge triangulation.

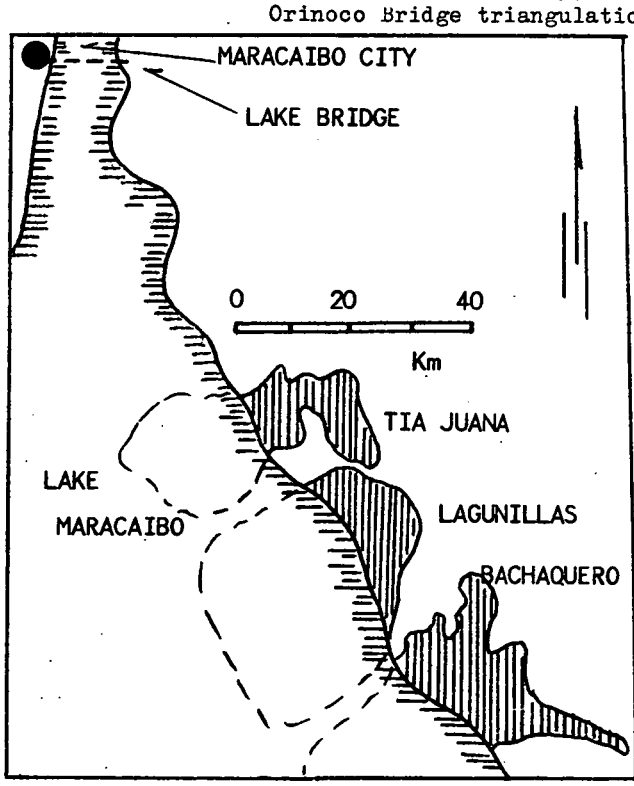



Fig. 17 Location of the main oilfields on the east shore of Maracaibo Lake.

References

Molnar, P., and Sykes, L. R., Tectonics of the Caribbean and Middle America regions from local mechanisms and seismicity, Geol. Soc. America Bull., v. 80, p. 1639 - 1684, 1969.

Schubert, C., Venezuela y la "Nueva Tectonica Global", Acta cient.venezolana, v. 21, p. 13-16, 1970a.

Schubert, C., Glaciation of the Sierra de Santo Domingo, Venezuelan Andes, Quaternaria, v. 13, p. 225-246, 1970b.

Schubert, C., and Henneberg, H. G., Geological and geodetic investigations on the movement along the Bocono Fault, Venezuelan Andes, Tectonophysics, v. 29, p. 199-207, 1975.

Henneberg, H. G., Results of geodetic deformation measurements on large scale structures including some application in geology, Int. Symp. on Deformation Measurements by Geod. Methods, Krakow - Poland, 1975.

Nuñez, O., and Escogido, D., Subsidence in the Bolivar Coast, Second Int. Symp. on Land Subsidence in Anaheim, 1976

Henneberg, H. G., Geodaetische Arbeiten an Talsperrenbauten in Venezuela, Veroeff. d. Geod. Inst. T. H. Aachen, No. 23, 1977

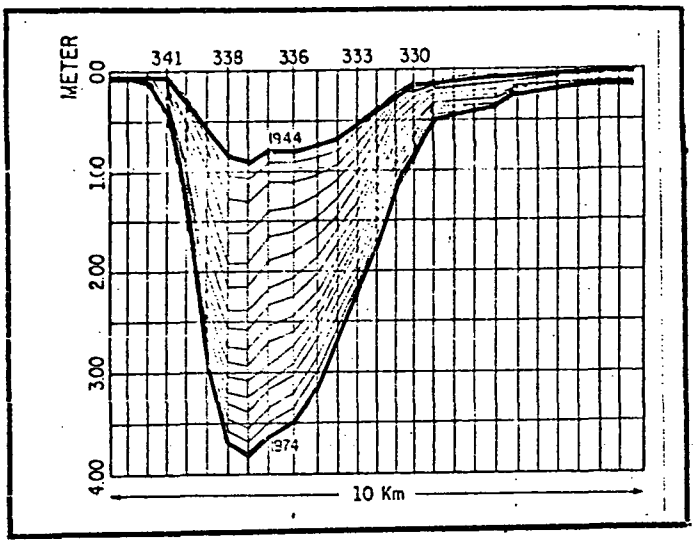



Fig. 18 Subsidence cone of Tia Juana oilfield.

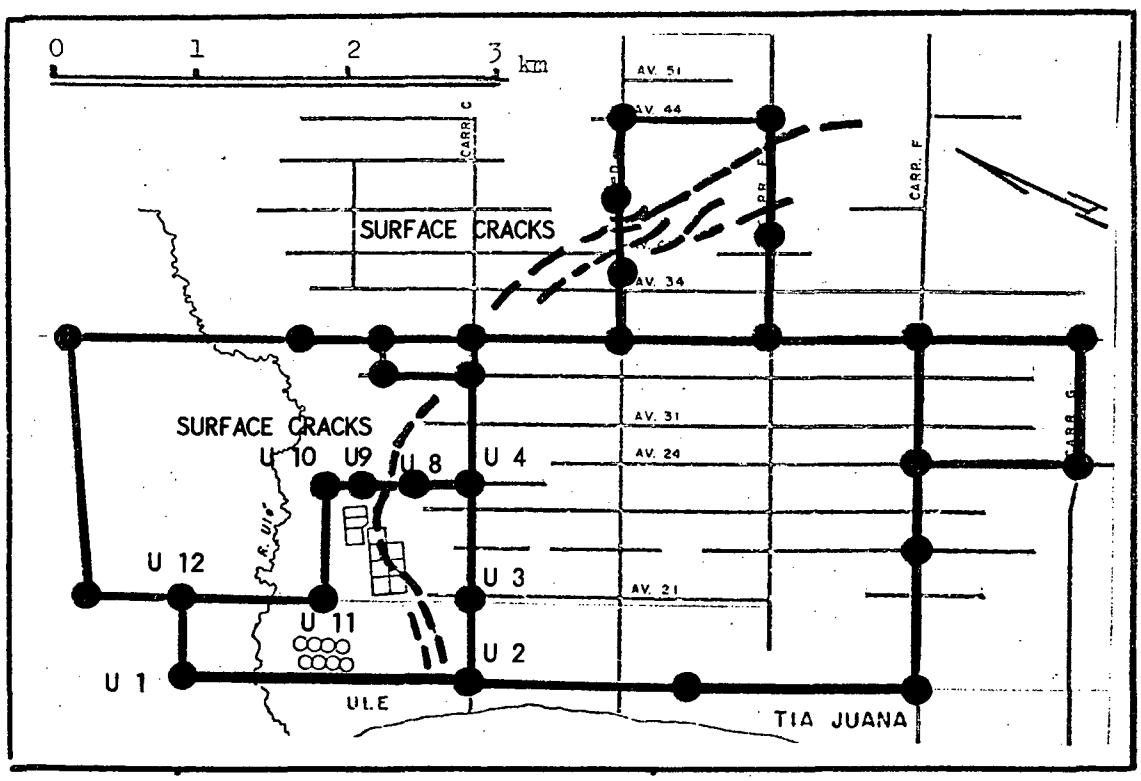



Fig. 19 High precision traverse to investigate the horizontal movements of surface cracks at Tia Juana oilfield.

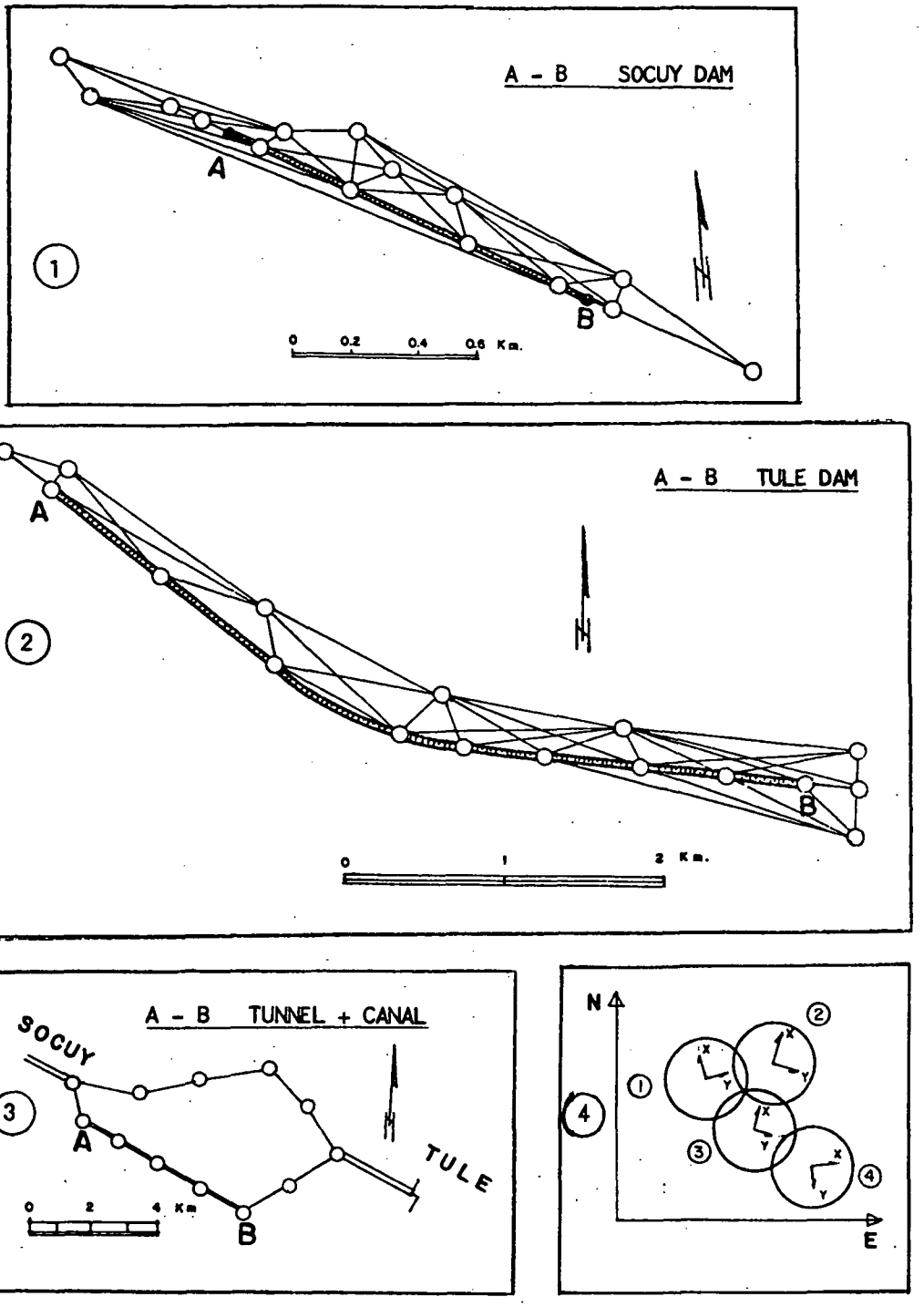



Fig. 20 (1) Socuy Dam - triangulation-trilateration, (2) Tule Dam - triangulation-trilateration, (3) Precision traverse for Socuy tunnel and canal and to combine the Socuy and Tule nets, (4) Schematic representation for the combination and transformation of the different coordenate systems at Socuy-Tule.

# Geodetic Measurements at Sea Floor Spreading Centers

F. N. Spiess
The Marine Physical Laboratory of the Scripps Institution of Oceanography
University of California, San Diego, San Diego, California 92152

Abstract. While space and terrestrial systems can provide the surface motion boundary conditions needed to constrain theories of tectonic driving forces and plate behavior on the large scale and in detail at some transform fault boundaries, the problems of new crust emplacement and its initial deformation as it is incorporated into the steady state plate motion at intermediate and fast spreading centers (5 cm per year full rate or greater) can only be attacked in the deep ocean floor. Advances in three areas of knowledge and experience make it feasible to design, deploy and use acoustic systems in that environment now.

The first area of advance is in our knowledge of the details of spreading center morphology and its implications. We now know that these areas typically consist of a band from a few hundred meters to a km. wide in which there are extensive pillow lava and sheet flows interspersed with occasional small conical peaks all completely devoid of any trace of sediment. Flanking this zone on either side is a region two to four times as wide in which there is extensive fissuring and some faulting with gradual buildup of sediment cover as one goes away from the central strip. Beyond that there is usually little evidence of tectonic activity, except for some additional fault displacement. The clear implication is that the transition from no lateral motion up to full plate speed takes place in a region usually less than 10 km wide and must involve strain rates of 10 microstrain per year or greater. How the strain buildup is distributed within this zone is the major question which sea floor geodetic measurements can answer.

The second area in which knowledge has been built up over the recent past is based on a growing quantity of horizontal temperature and salinity profiles made in these rise crest regions. These show that the small scale random inhomogeneities in sound propagation speed will not introduce errors of more than one or two cm over 10 km except for paths passing within 100 to 200 m. of an active hydrothermal vent. Since such vents are usually at least hundreds of meters apart, any system which is intelligently installed, incorporates sound velocity measurements (or their equivalent) in its survey techniques, and allows for spatial averaging will not be prevented by the environment from achieving cm. accuracy.

Finally, our operational experience with acoustic transponders in the less demanding geological mapping context, plus recently developed precision transponder designs, make it possible to build systems which will yield round trip travel time measurement accuracies of 10 microsec., corresponding to distance uncertainties of less than a cm. Transponder lifetimes of five years are available today, as well as procedures for replacing the units in a given network without loss of precision overall.

A network of 8 or more precision transponder units mounted on the sea floor and interrogated periodically from an instrument package towed near bottom through the area to provide the necessary spatial averaging could thus provide, today, a practical system for observing the pattern of buildup of strain at intermediate and fast spreading centers.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

Introduction. Geodetic measurements can make significant contributions to the field of plate tectonics by providing the surface, kinematic boundary conditions which theories of driving forces and plate behaviour must satisfy. Terrestrial methods seem particularly adapted to delineation of strain patterns associated with transform faults as they traverse continental environments, particularly in the San Andreas region, but also in South America and Eurasia. Space techniques can provide larger scale observations of gross motions within and between plates and at present appear to be most applicable across trench zones and for determination of plate deformability, particularly when used in conjunction with detailed data from other sources at plate boundaries. The particular role of undersea methods is for detailed work at the spreading centers, since only short sections of these are available for terrestrial examination and even these are slow spreading and more diffuse than the typical rise crest environments.

As yet only the terrestrial techniques have achieved the centimeter precision which these problems require, and a number of the other papers in this meeting have been devoted to descriptions of land and satellite systems now under development to reach these goals. The purpose of this paper is to display the fact that undersea techniques, well matched to the rise crest strain measurement problem, can be assembled today in such a manner as to meet this challenge.

The advances which make this assertion possible have occurred in three areas over the last decade, primarily in conjunction with fine scale geological studies of the rise and ridge crests using deeply towed instrument systems (Spiess & Mudie, 1970; Spiess et al, 1976; Ballard et al 1975) complemented in some instances by submersible observations (Mid Atlantic Ridge, Galapagos Spreading Center). The three areas of advance are:

1. Delineation of the detailed morphology of a variety of typical spreading centers
2. Extensive observations of the structure of the immediately overlying water at several representative sites.
3. Successful operational use of acoustic techniques in a large number of deep sea sites in a geological mapping context (precisions ~1-2m) and development of equipment designs capable of pushing these

capabilities to cm. accuracy.
Each of these three topics will be discussed below, followed by a description of a system configuration appropriate to the problem.

The special questions to which the rise crest data would be relevant can be visualized by anticipating a possible set of results. It seems quite likely that the large scale satellite measurements between points well into the interiors of two separating plates will show a steady drift at a rate comparable to that deduced from magnetic anomaly patters. It is equally likely that detailed measurements of the pattern of buildup of strain very close to the crest will show substantial irregularities with space and time, at least on the scale of meters and years. Anticipating the morphological data to be presented below, it appears likely that the "acceleration" from close to zero strain rate at the rise center to nearly full steady state plate velocity may take place in less than 10 km. If so then we should see relatively steady separation going on at the edge of that zone, while a description of what is happening now within the zone at a variety of sites and over times requiring some patience, may give us direct evidence as to the sequence of events and mechanisms involved in creation & stabilization of new crust; shedding light on the relative importance of the actual insertion of new material, of subsequent fissuring, of hydrothermal activity, of large scale faulting, etc. Determining the pattern of strain buildup across this zone is thus the problem we address.

Spreading Center Morphology

The major consideration which bears on the feasibility and form of a sea floor strain measurement system is the region it must span both across

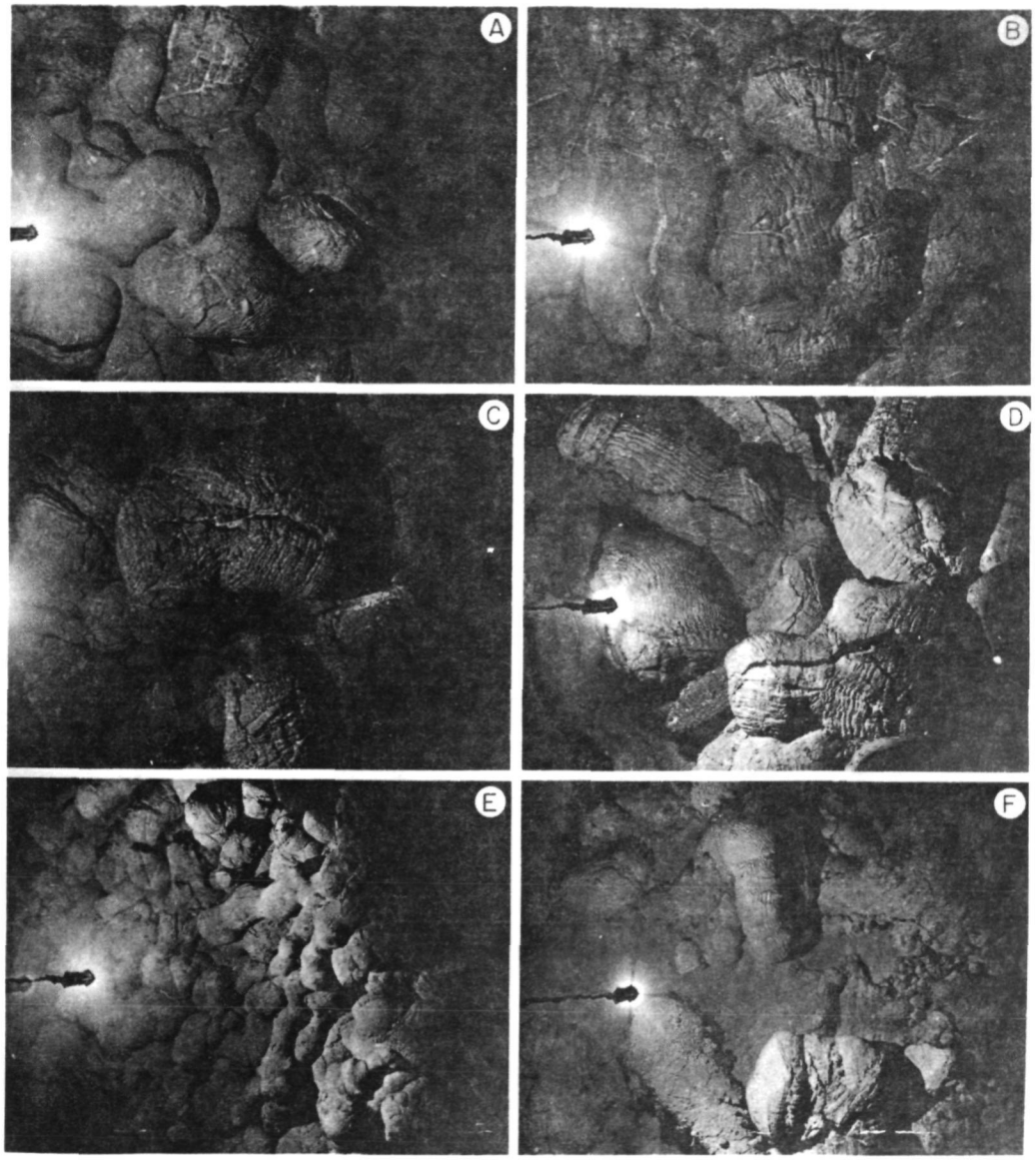



Fig. 1 - Fresh pillow lavas on East Pacific Rise Crest at 9°N.

and along the strike of the central zone. A variety of observations support the notion that the emplacement of new material and the acceleration of that material up to nearly full steady state plate speed takes place in a marginal strip not more than 5 to 10 km wide. The first piece of evidence in that regard has been with us since very early in the development of the field. As the magnetic anomaly patterns were recognized in many oceanic regions and as time scales were assembled it became clear that there were some distinct polarity epochs only a few km across and thus only a few hundred thousand years in duration. Quite consistently observed, for example, is the JARAMILLO normal event which occurred about 900,000 years ago with a duration of about 60,000 years. Widespread observation of this strip, which is only 2 km wide in intermediate (3 cm/yr half rate) spreading rate settings, carries a clear implication that the startup process must take place primarily over lateral distances certainly no larger than the width of that zone.

Near bottom magnetometer observations of the boundaries of these zones (summarized by Klitgord, 1975) show that they are quite sharp - complete reversals take place in strips ranging from hundreds of meters to a little over one km wide. A recent near bottom survey carried out at the Brunhes/Matuyama boundary near 21°N on the East Pacific Rise (Macdonald, et al, in prep.) covered a zone about 6 km along (and across) strike and showed undulations in the reversal line with an along strike scale of 0.5 to 1.5 km and across strike amplitude of 200 km. All this clearly indicates a highly localized process.

Moving in to the actual spreading centers it is quite clear from direct visual and photographic evidence that there is a highly localized (usually the order of a km or less) central zone in which most of the new material is being introduced. It is a region characterized by pillow and sheet flow lava forms, with fresh glassy facets and completely lacking any signs of sediment cover (Fig. 1) As one moves away from this zone, at least on the intermediate and fast spreading regions surveyed on the East Pacific Rise and Galapagos spreading centers (Larson & Spiess, 1969; Klitgord & Mudie, 1974; Normark, 1976; Lonsdale, 1977; Larson, 1971), there is an immediate flanking region a few km wide in which there is extensive fissuring and the beginnings of normal faulting. In this region the sediment cover also begins to build up in a fairly steady manner, although local irregularities preclude an assertion that this evidence demonstrates steady state motion on a time scale as short as 1,000 years.

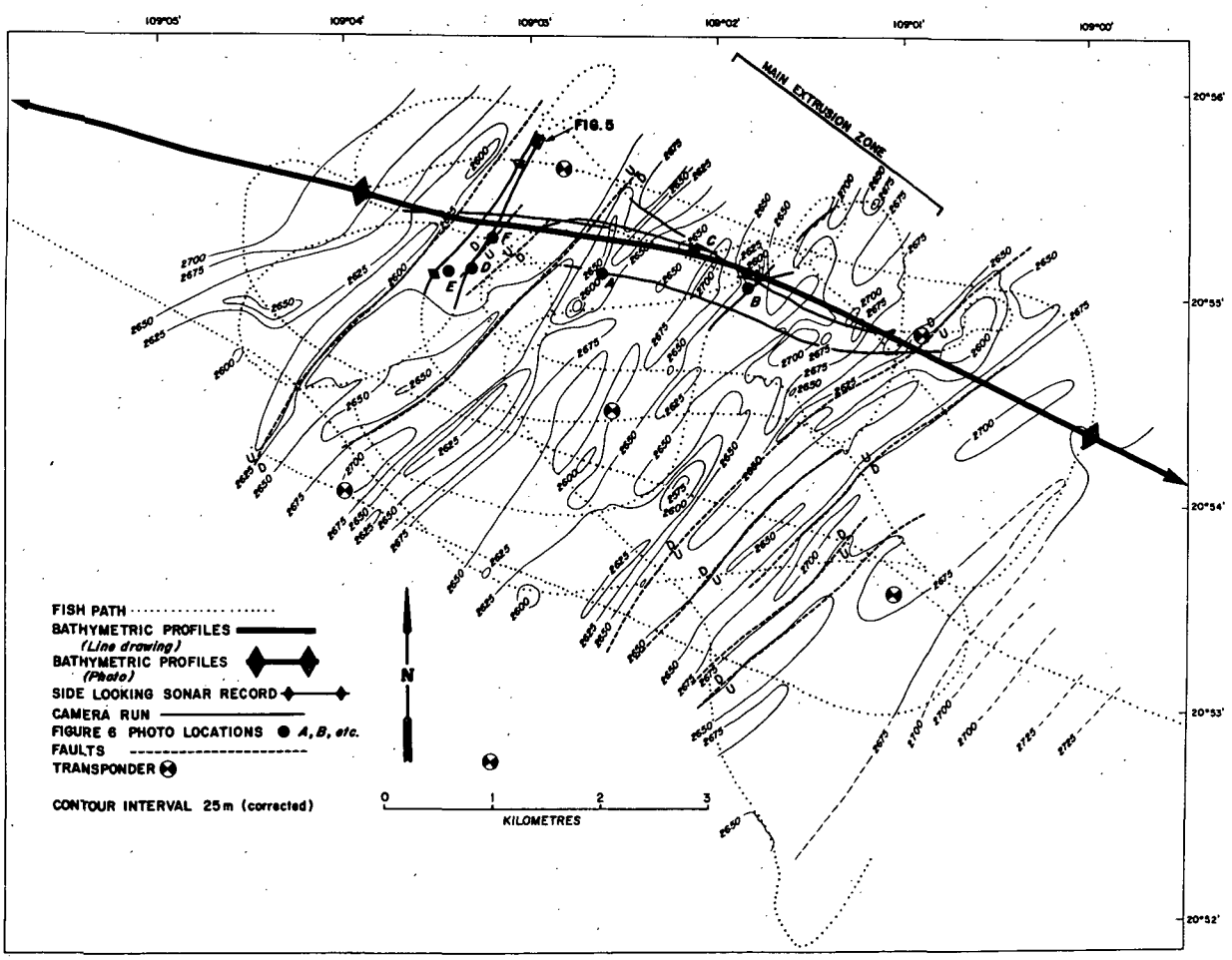



Fig. 2 - Fine scale topographic chart of East Pacific Rise Crest at 21°N. Full spreading rate 6 cm/yr.

Using data (Fig. 2) from the East Pacific Rise crest at 21°N (Normark, 1976) where the full spreading rate is 6 cm/yr, the total separation appears to be generated in a 6 km wide strip for an overall predicted strain rate of 10 microstrain per year. It seems likely that this will not build up in a spatially uniform manner, thus there should be intervals within this band which will show 2 or 3 times this value over distances of the order of a km or so. In a faster spreading environment (12 cm/yr full rate) at 9°N the region in which activity appears to be concentrated is somewhat more diffuse (Fig. 3), but still is less than 10 km for an overall rate of 1.2 microstrain per year and again an implication of local rates within the zone which must exceed this.

This all implies a requirement to determine distances between reference points a km apart to within about a cm in order to obtain useful data from an annual re-survey cycle. It also implies that, unless there is substantial deformation taking place within the plate, one should in fact see a very continuous displacement generated between the extreme edges of this zone.

A number of possible models can be postulated based on the evidence in hand, ranging from a primarily tensional picture across the entire zone with lava flows filling gaps in the center, to the other extreme in which local dike intrusions produce compression in their immediate vicinity and tension in adjacent along-strike zones. A network of measurement sites with one to three km spacing, extending across the entire zone and along

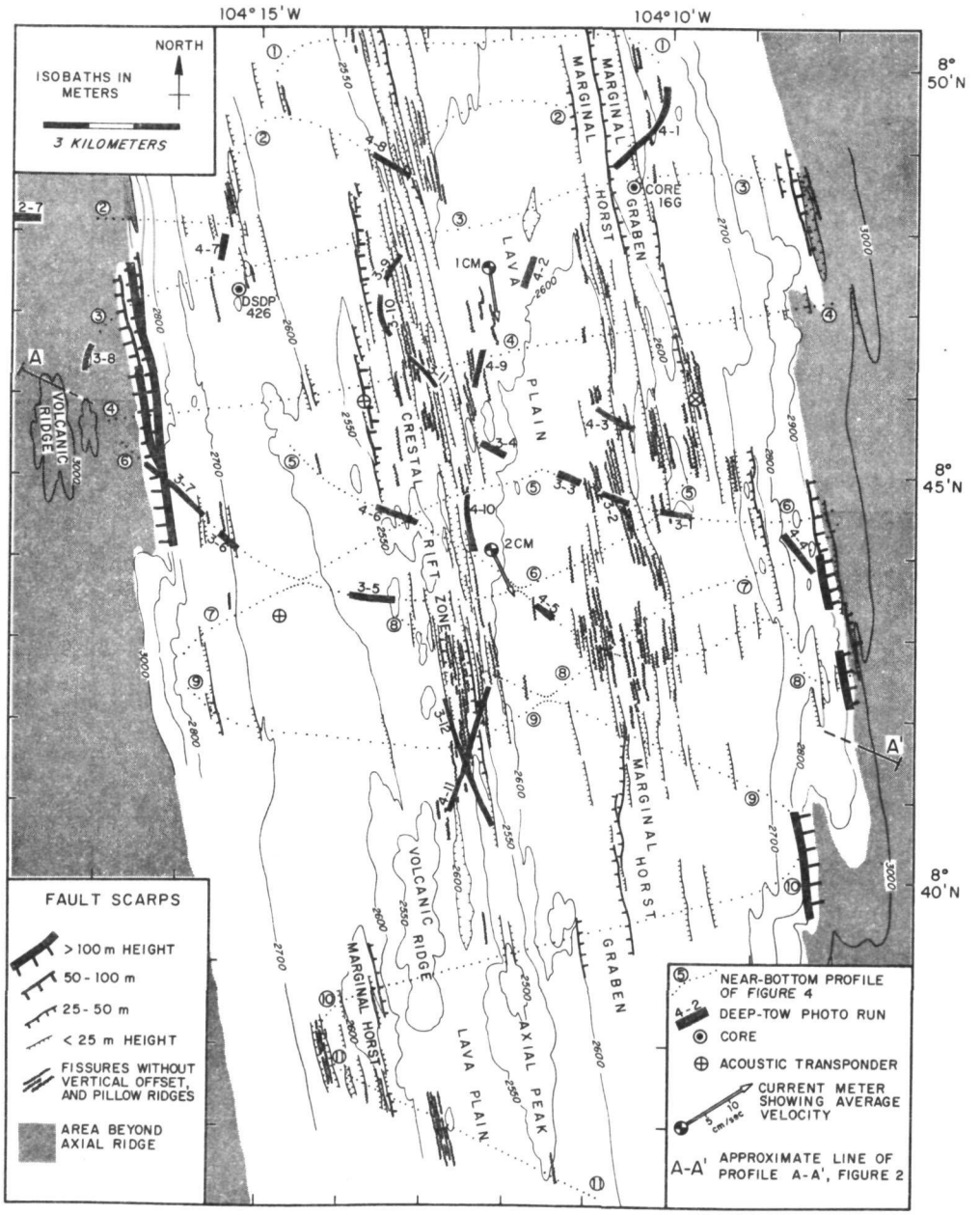



Fig. 3 - Fine scale topography at 9°N on East Pacific Rise Crest. Full spreading rate 12 cm/yr.

strike for several km would provide the necessary information to distinguish among the alternatives and place strong boundary conditions on the development of models for plate structure and driving forces.

Water Column Considerations

The two major concerns in the overlying water column are first, the attenuation of whatever radiation might be used for measurement of distance or angle and second, the nature of the index of refraction or energy propagation speed variations with time and space. In sea water the attenuation lengths for electromagnetic radiation are very short, generally only a few wavelengths, except as one approaches the optical regime. Even here, however, in the clearest water one can expect 1/e in 20 meters, making the km dimensions across the highly irregular central zone very difficult to achieve even with lasers or other optical systems. Acoustic energy is attenuated as well, with absorption increasing with frequency. The values are such, however, that a useful regime for km transmission with adequate bandwidth exists from a few tens of kiloherz to hundreds, with an optimum in the 20 to 100 kHz range strongly controlled by the nature of the background noise spectrum. The acoustic approach thus seems to be the most practical manner in which to attack the geodetic problem (Spiess, 1966).

Given a sound propagation speed of about 1500 m/sec this implies wavelengths of the order of a few cm. In turn this means that devices for accurate angle measurement must be quite large and confirms one's intuitive impression that distance rather than angle measurements are more appropriate to this situation.

The environmental factors that enter are thus the motion of the water and the irregularities in the sound velocity field. Currents of the order of 15 cm/sec are occasionally observed at rise crests, varying with tidal periods. While such velocitiies would produce appreciable effects on one way paths (a part in $10^{-4}$), when two way paths are used their influence drops to $\sim 10^{-8}$ which is far below the required $10^{-5}$.

The major limitation is the variability of sound velocity in the water due to temperature variations and, to a lesser extent, chemical effects. 0.003 C over the entire path length will give a change in sound speed of a part in $10^{-5}$. Such changes can occur over long times and large areas at many places near the deep sea floor. In addition there are hydrothermal effects at intermediate and fast spreading centers which can locally produce larger anomalies. It will thus be essential to couple into any acoustic measurement program of geodetic type a capability to determine the sound speed locally to the desired accuracy at the time of each survey. Laboratory measurements of sound speed to $10^{-6}$ have been made (Greenspan 1972) and, although today's oceanic sound velocity meters are usually good only to $10^{-4}$, it appears that $10^{-5}$ is an achievable goal and such an instrument is under construction in the Marine Physical Laboratory at the present time (F.H. Fisher & E.D. Squier, private communic.).

With an in situ measurement capability of this type, one is still left with residual small scale variations in space and time which must be treated statistically. During several recent rise crest surveys (Galapagos spreading center, East Pacific Rise at 21°N and 9°N) we have made continuous near bottom temperature records over the central 10 km of the terrain (Crane, 1977; Crane and Normark, 1977). Typically, except within a hundred meters of the few active hydrothermal vents in the Galapagos area, the fluctuations are less than 5 mdeg, with scales of the order of 100 m. A number of authors have treated the problem of fluctuations in travel times as related to spatial variations of propagation velocity. One form is given (Chernov, 1960) roughly by:

$$\Delta t = \frac{\sqrt{2AR}}{C_o}\sqrt{\overline{\mu^2}} \qquad (1)$$

Here R is the one way path length, $C_o$ the average sound speed, A the characteristic length scale of the fluctuations and $\mu = \frac{C_o}{C} - 1$. Converting this to a length error, E, recalling that we are using round trip travel times, gives

$$E = \sqrt{\tfrac{1}{2}}\, AR\sqrt{\overline{\mu^2}} \qquad (2)$$

The error thus grows as $\sqrt{R}$. Using A = 100 m and $\sqrt{\overline{\mu^2}} = 10^{-5}$, the uncertainty ranges from a little over 2 mm at 1 km to 7 mm at 10 km, all well below the desired cm. precision. An additional safety factor can be introduced, as will be described in the section below, by arranging the system layout to provide substantial averaging over both space and time.

Technological Advances

Over the last ten years there has been a major step forward in ability to determine positions of devices in the deep sea through the introduction of acoustic transponders. Systems for this purpose are now in use on a regular basis by several research and operational groups. Most of the transponders are of the general type described some years ago by McGehee and Boegeman (1966 ), receiving a signal at one frequency in the 7 to 15 kHz range, recognizing it with some threshold device and replying at a second frequency. Power supplies giving a capability of 5 years in the ready receiving mode and $10^6$ output pulses are now in existence.

With a set of these units moored to the sea floor one can measure the acoustic travel times corresponding to the ranges between the vehicle or instrument package and several of the transponders and make a determination of its coordinates relative to that array. In our system the vertical coordinate of the vehicle is determined either by echo sounding on the sea surface or by a precision measurement of hydrostatic pressure.

While a number of operational computational approaches have been used the most sophisticated to be implemented to date is an adaptation of a method well known in geodesy. In this we start the survey using a set of rough approximate coordinates for the transponders and compute successive vehicle positions relative to these. In cases for which more than two simultaneous ranges are

...ilable a least squares criterion is used to determine the position and an error measure is calculated at the same time. The raw range data are stored and after the vehicle has traversed the area a number of times 50 or more well placed multi-transponder positions are selected and used in iterative fashion to determine an improved set of transponder and vehicle coordinates.

A carefully made survey for geological purposes result is in position uncertainties of the order of a couple of meters. This is quite adequate for the navigational need but obviously differs by a factor of 100 from the cm. uncertainties which represent both the required geodetic accuracy and one which is attainable relative to environmental limitations. The necessary improvements can be achieved by including more detail in the computational methods and using a more sophisticated type of transponder.

Current transponder designs involve a delay of a millisec or more in the recognition circuitry and lack of control of this parameter results in a contribution of the order of a meter to the timing uncertainty. We have designed and made in-water tests of a new approach which provides a very accurately controlled time delay (a few microsec) in the pulse retransmission process (Spiess, et al, 1978). Using this approach and pulses in the 30 to 40 kHz range, with signal to noise ratios in excess of 15 db, one can easily achieve resolution to 10 $\mu$sec, or less than 1 cm, on each transmission.

On the computational side the principal complications are that one must compensate for motion of the survey package while the sound pulse is in transit for several seconds and take into account the substantial variation of sound speed with water depth. These are aspects which merely dictate that care is exercised in developing the detailed computer programs to be used.

With these improved transponders and computational methods one can begin to gather the data necessary to determine strain change distributions at rise crests over prolonged periods. The most logical approach is to install a network of 10 or more units in the area, with spacings along and across the crest ranging from one to three km with the exact layout dictated by the detailed morphology of the particular site. Each resurvey operation (done on a time scale of months to one or two years, depending on the spreading rate) would be carried out by towing the survey package close to the sea floor through the area, ranging on the transponders and measuring the sound velocity. By using many multiple transponder fixes and calculating the network geometry from the moving vehicle one proyides means for spatial averaging over the sound speed inhomogeneities as well as smoothing other random errors. The network elements can be replaced without precision sea floor operations by installing the new units close to the old prior to a survey operation and removing the old ones for refurbishing after the entire augmented network geometry has been determined.

It thus appears feasible, given the narrow zones of activity and the rapid spreading rates, to match underwater acoustic techniques to the problem of determining the strain buildup patterns in these areas on time scales ranging from months to decades.

References


Ballard, R. D., W. B. Bryan, J. R. Heirtzler, G. Keller, J. G. Moore, and Tj. van Andel, Manned submersible observations in the FAMOUS area: Mid-Atlantic Ridge, Science, 190, 103-108, 1975.

Chernov, L. A., Wave propagation in a random medium, McGraw-Hill, New York, 1960.

Crane, K., Hydrothermal activity and near-axis structure at mid-ocean spreading centers, Ph.D. Thesis, Univ. of Calif., San Diego, 1977.

Crane, K., and W. R. Normark, Hydrothermal activity and crestal structure of the East Pacific Rise, J. Geophys. Res., 82(33), 5336-5348, 1977.

Greenspan, M., Acoustic properties of liquids, American Institute of Physics Handbook, 3rd Edition, McGraw-Hill, section 3e, 1972.

Klitgord, K. D., and J. D. Mudie, The Galapagos spreading centre: a near-bottom geophysical survey, Geophys. J. R. Astron. Soc, 38, 563, 1974.

Klitgord, K. D., Near-bottom geophysical surveys and their implications on the crustal generation process, sea-floor spreading history of the Pacific, and the geomagnetic time scale: 0 to 6 m.y.b.p., Ph.D. Thesis, Univ. of Calif., San Diego, 1974.

Larson, R. L., Near-bottom geologic studies of the East Pacific Rise crest, Bull. Geol. Soc. Amer., 82, 823-841, 1971.

Larson, R. L., and F. N. Spiess, East Pacific Rise crest: a near-bottom geophysical profile, Science, 163, 68, 1969.

Lonsdale, P., Structural geomorphology of a fast-spreading rise crest: the East Pacific Rise crest near 3°25'S, Mar. Geophys. Res., 3, 251-293, 1977.

McGehee, M. S., and D. E. Boegeman, MPL Acoustic Transponder, Reviews of Scientific Instruments, 37, 1450-1455, 1966.

Normark, W. R., Delineation of the main extrusion zone of the East Pacific Rise at 21°N, Geology, 4, 681-685, 1976.

Spiess, F. N., Underwater acoustic positioning: Applications, Proc. of First Mar. Geod. Symp., Columbus, Ohio, 28-30 September 1966, 93-101, 1966.

Spiess, F. N., and J. D. Mudie, Small scale topographic features, The Sea, 4, part I, A. E. Maxwell, Ed., Wiley-Interscience, New York, 205-250, 1970.

Spiess, F. N., C. D. Lowenstein, D. E. Boegeman and J. D. Mudie, Fine scale mapping near the deep sea floor, Oceans '76, 2nd Annual combined Conf. MTS-IEEE, 1976.

Spiess, F. N., C. Lowenstein, D. E. Boegeman and V. Pavlicek, Precision transponder and method of communication therewith, Pat. Applic. 885893, 13 March 1978.

# Sea Level Data and Techniques for Detecting Vertical Crustal Movements

G. W. Lennon
School of Earth Sciences, Flinders University
Bedford Park, South Australia 5042

Abstract. In accordance with the wishes of the Panel Chairman an attempt is made to survey "problems, requirements, and the outlook for the future" in the study of sea level time series so as to determine the relative movement of land and sea levels. The basic aim is to eliminate from the record the contributions from whatever marine dynamic phenomena respond to treatment, allowing the secular element to be identified with optimum clarity. Nevertheless the concept of sea level perturbation varies according to regional experience. The recent work of the Permanent Service for Mean Sea Level helps to eliminate geodetic noise from the series and makes it possible, perhaps, to treat the global mean sea level data bank so as to define eustatic changes in ocean volume which, in the present context, may be regarded as the final goal, allowing the identification of vertical crustal motion itself.

## Scale Lengths

This survey is not concerned with geological time scales, although sea level evidence for crustal motion can be relevant in this sense. Neither is it intended to discuss archaeological time scales since interpretation here is limited to locations, such as are found in the Mediterranean or in Scandinavia, which experience a small tidal range combined with low rates of accretion and erosion. The aim is to consider the period of instrumental history where continuous or near-continuous observed time series of sea level are available for interpretation. The interest is therefore focussed upon recent time, within the last hundred years, and often upon the last two decades. With regard to spatial scales it should be noted that observed sea levels are ideally referred to a fixed mark local to the instrumental site. Although attempts are invariably made to select a stable bench mark, it is clearly a matter for survey procedures to determine whether a particular bench mark is regionally representative. Again, if any attempts are to be made to determine absolute sea level or sea level topography, rather than variations at a point, then much greater reliance must be placed upon survey evidence and the well-known controversy over apparent sea level slopes becomes a matter of anxiety. However, where secular variations of level, determined at a number of points, have the appearance of spatial coherency then these reinforce each other and in some cases justify interpolation within the observing network.

## Limitations

The restrictions of the exercise are numerous, not least being the obvious lack of control over the data. The material is historic and in large measure obtained by instruments designed and installed for operational rather than scientific purposes and somewhat inadequately maintained. Location, in harbours or estuaries, is often inconvenient and in another context [e.g. Lennon, 1971] reference is made to the several imperfections of the standard instrument. In fact it is in many ways surprising that the search for a secular signal which may be as small as one mm per year can be pursued with success.

In the context of this Symposium it is relevant to address a particular problem which does not appear in the associated literature. It is only during the last decade that satellite orbital analysis and laser altimetry have been combined with gravimetric surveys to determine the variations of the gravitational field with longitude. As experience has been gained with these techniques and as they have become more precise, there has been developed within the last few years a detailed determination of the geoid. However as yet we have no conception of the stability of the geoid in time and space. What we do know is that the geoid demonstrates significant relief ranging through almost 200m from a depression south of India to a peak in the vicinity of Papua New Guinea. By definition the geoid describes an equipotential surface and we can expect the ocean surface to conform with the geoid relief. The point which must be made however is that if one bears in mind the above-mentioned apparent gradient over 200m, it will require only a small time variation of the geoid at a point to seriously perturb the observed secular variations of sea level at such a station. The nature of the problem is such that little can be stated at this time although some thought has been given to geoid stability. In a personal communication Kaula considers that the most likely source of change lies in the phenomenon of glacial rebound in the higher latitudes and figures of approximately 0.3mm per year have been postulated. Though small, such a change nevertheless is significant in terms of secular variations of sea level.

## Techniques

The common procedure is first to apply a tidal filter so as to remove tides of diurnal and higher species. Tides of fortnightly, monthly, semi-diurnal, annual and longer period tend to be more intricately entangled with contributions from atmosphere and other sources and require sophisticated treatment. In the initial stages therefore they are allowed to remain in the data. It is the attention paid to such perturbations and the design of procedures for their elimination from the record which determines the accuracy with which the secular variation of sea level can be establ-

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

ished. It is also relevant that one should be assured that none of the perturbing parameters is itself influenced by secular change so that the total secular signal may be ascribed to the basic sea level phenomenon. In fact there has been little anxiety on this score.

The very slow change in the data base predetermines that the discipline proceeds at a pedestrian pace and in fact the major interest lies in the sophistication of the procedures utilized in the elimination of the noise in so far as this is possible. It is interesting that individual workers place stress upon phenomena of which they are aware in their own environment and there has been no attempt as yet to apply a comprehensive model of long period sea level phenomena. For example in North Europe the emphasis has been upon barometric pressure and associated wind stress phenomena with the contributions of storm surges clearly in mind. In contrast, in N.America the absence of large shallow water areas has resulted in lower priority being given to such features and the emphasis has been placed upon steric phenomena resulting from variations in water density and upon the more general features of marine dynamics.

## Simple Regression

The European school is perhaps best exemplified by work during the sixties [Lennon, 1966 and Rossiter, 1967]. Here some four thousand station years of annual mean sea level were processed in a regression exercise which for each station took the form

$$Z_y = \sum_{p=0}^{p=3} a_p Y^p + \sum_{r=1} b_r B_r + c_1 \mathrm{CosN} + c_2 \mathrm{SinN} + \phi_y \quad (1)$$

where $Z_y$ is annual mean sea level for year Y referred to a working datum, $B_r$ is annual mean value of air pressure at station r for year Y, N is the mean longitude of the Moon's ascending node for year Y, $\phi_y$ is the contribution to $Z_y$ from all other causes considered inexplicable in this context and therefore regarded as residual noise. The coefficients a,b & c are determined in the regression procedures.

Some comments are appropriate. It can be seen that the term $\Sigma a_p Y^p$ is the primary objective of the exercise and does not constrain the secular variation to a linear form. The terms containing N were intended to reproduce the lunar nodal tide with a period of 18.61 solar years. The stations, r, for which observed barometric pressures were obtained, were selected in triplets arranged as far as possible in an equilateral manner around a specific sea area which in turn was considered to be a potential contributor to mean sea level at the station e.g. the North Atlantic, the North Sea, the Baltic etc. It was assumed that such a treatment would reproduce contributions of static barometric pressure and also all directional components of wind stress. Tentative trials were conducted combined with a study of the resultant, $\phi_y$, so as to select an optimum set of triplets for each station.

Figure 1 illustrates the treatment of the sea level time series for Newlyn and allows a subjective assessment of the success of the technique. In fact the exercise did show spatial coherency making it possible to produce the regional pattern of figure 2 and its indication of glacial rebound in the Baltic and especially in the Gulf of Bothnia.

What was surprising in this exercise was the apparent success of the barometric triplets in modelling the wind stress phenomena. The latter conceptually are of short duration in the range twelve to thirty hours, whereas the associated variables in the regression model were annual means. To succeed then the barometric terms must represent long term averages of many events, yet even here one is presented with a difficulty. Wind stress is a non-linear phenomenon varying approximately as the square of wind speed. Although in equation (1) one can conceive of barometric gradients being represented, there

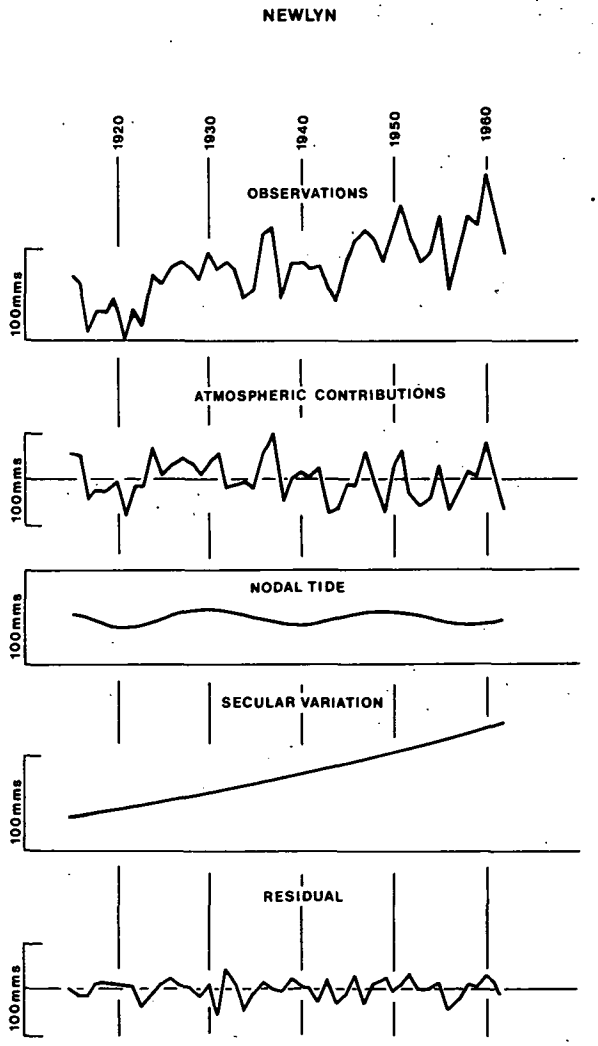



Fig. 1. Annual mean sea level at Newlyn and its components revealed by regression analysis.

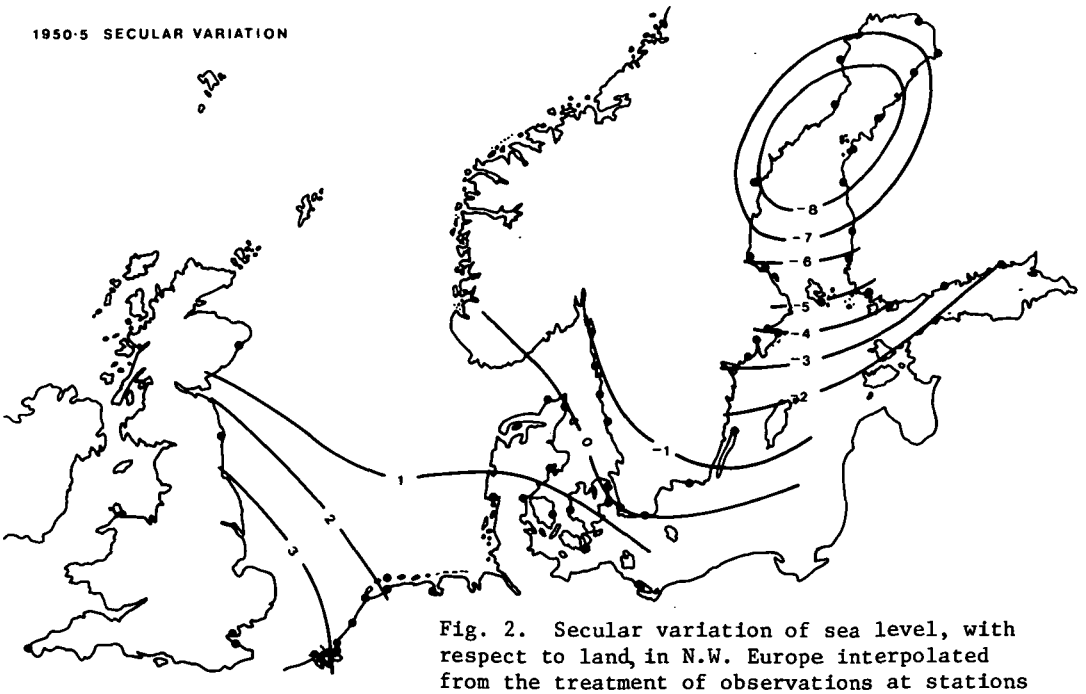



Fig. 2. Secular variation of sea level, with respect to land, in N.W. Europe interpolated from the treatment of observations at stations marked by solid circles.

is no attempt to reproduce a non-linear wind stress. Again one may comment that there is no simple solution in the time scales used since the square of the mean wind, possibly derived from annual mean pressures, cannot be equated with the mean of the squared wind from a succession of discrete events such as occur in nature. The success of the procedure outlined above is puzzling and has initiated much thought as to the manner in which techniques may be improved.

Complex Regression

As one attempts to provide more rigorous treatment of the meteorological perturbations yet another problem emerges with increasing significance, namely that of the high degree of intercorrelation of the meteorological variables. More sophisticated techniques also call for longer data bases, only forty-seven annual mean levels for Newlyn in the earlier exercise, for example, being a serious restriction. A notable attempt to face these and other problems has been made by Thompson of the Institute of Oceanographic Sciences in the U.K. [Thompson, 1978].

Here an interesting exercise is based upon monthly values of mean sea level and this approach inevitably carries the premium of careful attention to other long period tides in particular the solar annual and solar semi-annual tides. At the expense of this complication a useful extension of the available sea level time series is effected. However the basic innovation is the treatment of the North Atlantic barometric pressure field expressed in terms of monthly mean pressures from a five by five grid at 5° spacing in latitude and 10° spacing in longitude. These pressure series clearly demonstrate a high degree of intercorrelation but treatment by the Principal Component Analysis effectively replaces the set of correlated monthly air pressures by a set of uncorrelated eigenvector coefficients. Two further attractive advantages emerge as a bonus. In the first place a highly convergent system is produced in that the first eigenvector accounts for 54% of the total variance of the standardised pressure fields in this particular application, while the first eight account for 99% of the total variance. In the second place the eigenvectors assume a physical significance as modes of variation over the grid showing progressively detailed structure as the order of the eigenvectors increases. In figure 3 the respective patterns of the eigenvectors are shown as positive and negative areas so as to illustrate this point.

This preparatory analysis set the scene for the treatment of some eighteen years of monthly mean sea levels at Newlyn. Incorporating also trend, and tides, a stepwise regression procedure automatically selected an optimal subset of eigenvector coefficients with a satisfying result. Over 90% of the variance of the original mean sea level series was explained and insight obtained into some large scale marine dynamics of the region.

The relevance to the present survey lies largely in the fact that the secular variation of sea level at Newlyn computed from a sea level record, eighteen years in duration from 1957 to 1974, emerged as $1.42 \pm 0.19$ mm per year. The earlier exercise had used forty seven years of data up to 1962 and produced a trend of $2.15 \pm 0.14$ mm per year. The first indicat-

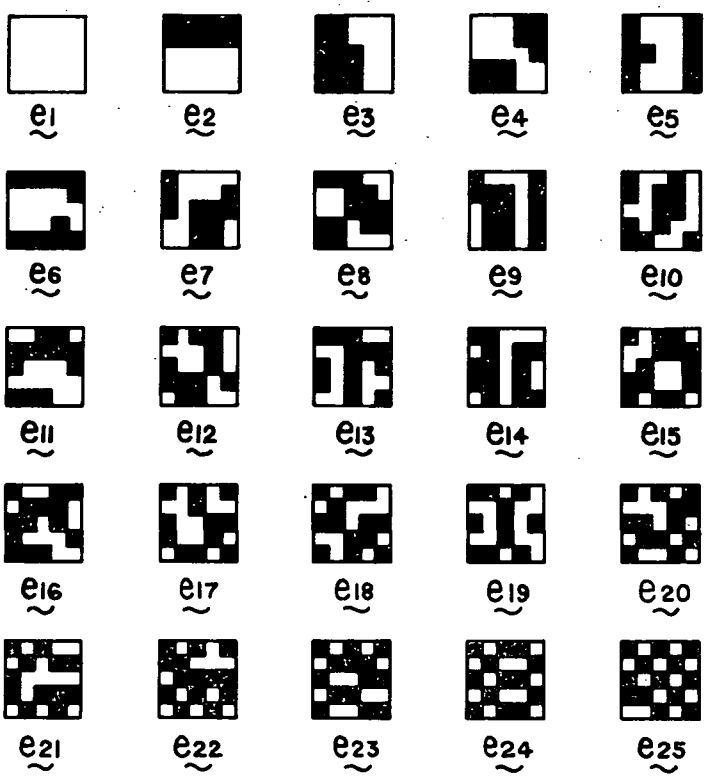



Fig. 3. The physical features of the eigenvectors representing air pressure variation over a North Atlantic grid.

ion is therefore given that the rate of rise of sea level with respect to the land may be decelerating with time at this station and certainly a powerful tool has been created which, depending upon relatively short time spans, can investigate the fine structure of secular variation in time. The same exercise gave an assessment of secular trend at Corunna over the identical period 1957 to 1974, as $0.57 \pm 0.42$ mm per year showing a clear disparity in space of crustal motion.

## Sea Level Perturbations

It has already been suggested that the European school has been influenced by the environment of a broad continental shelf on which the mechanics of wind stress have free rein. With a deep ocean aspect, as in North America, it is not unreasonable that the concentration should be on the large scale phenomena of ocean circulation, on the western boundary currents intensified by planetary vorticity and on dynamic height phenomena associated with the thermohaline vertical structure of the oceans [e.g. Sturges, 1974 and Chew, 1977]. These phenomena are also associated with time-variations and as yet have not featured in detail in studies of secular variations.

If the accuracy with which long term trends can be extracted from the records depends upon an understanding and elimination of other perturbations, then it should be noted that our knowledge of such mechanisms is far from complete. For example in recent time certain evidence is accumulating which suggests that in the Southern Hemisphere the greater exposure of the oceans, combined with the circumpolar nature of the Southern Ocean itself, allows a range of phenomena to be generated which may be unique to southern waters. Taking examples from Australia there is evidence on both east and west coasts of thermohaline structures associated with tongues of warm water, showing anomalies of $5^{\circ}$C or more, moving southwards in late summer. Much work has been done in the Tasman Sea [e.g. Hamon and Cresswell, 1972, also Andrews and Scully-Power, 1976] on the phenomenon of associated anticyclonic eddies, somewhat similar to those recorded in connexion with the Gulf Stream in the Atlantic, but having a recorded life of the order of one year. Dynamic heights up to 80 cms, accompanying geostrophic flow in excess of 100 cms per sec. are recorded. In this and in other contexts the tendency seems to be to provide sea level perturbations with periods greater than their counterparts in the European exercise previously outlined. More particularly a study of nontidal residuals reveals evidence of energy at characteristic periods as follows:

circa two days, and here it should be noted that this corresponds to the inertial period at latitude $14.4^{\circ}$

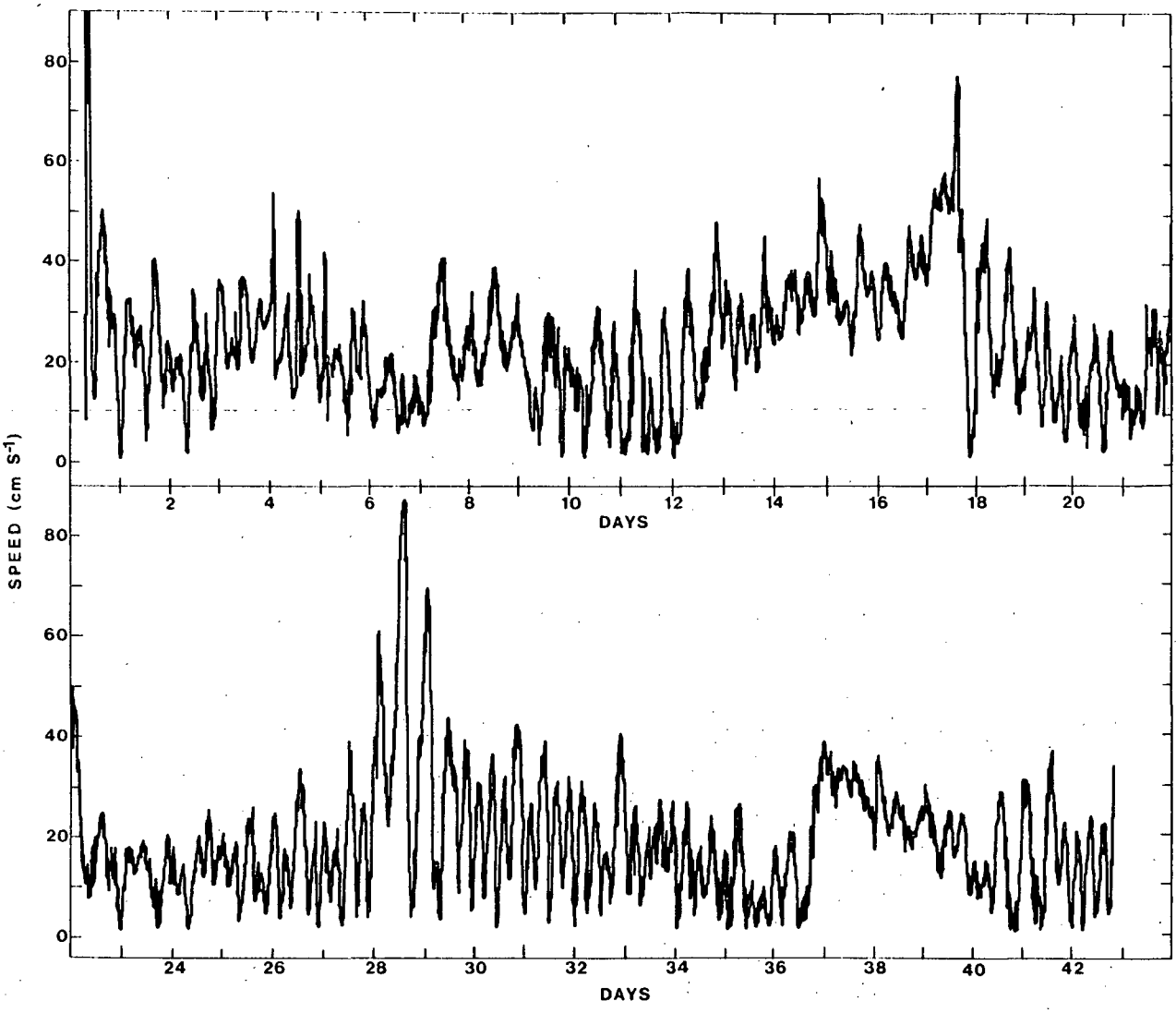



Fig. 4. An extract from a record of current speed at a typical station on the N.W. Australian Shelf.

circa four days, to which reference is increasingly appearing in the marine dynamics of equatorial and southern latitudes.

in excess of twenty days, which is most uncommon elsewhere yet waves of amplitude greater than 0.5m and apparently trapped on the Australian coastline have been traced and tracked as coherent phenomena over distances in excess of 5000 kms, from Dampier in the north-west, southwards along the west coast, across the Great Australian Bight and finally dissipating northwards along the east coast.

It is only in the last few months that commercial operations off-shore, notably on the North West Shelf, have stimulated experimental work and this is revealing exciting perturbing phenomena. The material in figure 4 has its origin in a personal communication from R.K. Steedman and Associates, Perth, contracted to Woodside Petroleum Development Limited. Here in a region of large vertical and horizontal tides, the stream record shows such tides to be superimposed upon significant disturbances. Note the two day event near day 28, the four day event commencing day 36 and the longer period feature commencing day 12. Supporting evidence confirms that these phenomena are real, common in occurrence and of considerable spatial scale. The inference may be drawn that they will interact with sea level topography in a complex manner yet their mechanisms remain obscure at this time.

In the context of this symposium the message seems to be that there is much to learn of marine phenomena which influence the temporal topography of the oceans. Any attempt to explain and eliminate these features in secular studies would be premature, but at least techniques developed for such studies must be aware of their existence. It is also clear that the dev-

eloping science of satellite altimetry should not be too confident that the problems of ocean topography are close to a final solution. It would appear that collaboration between oceanographers and space scientists offers prospects of fruitful mutual progress.

## The Mean Sea Level Data Bank

In spite of these apparent limits to the present capability of sea level studies to provide an extensive and definitive survey of crustal motion, the potential of the discipline to provide unique evidence in this context is recognised. It has also been apparent that although the discipline is limited to existing historic data with all its limitations, two positive actions may be taken. In the first place there is an obligation to ensure that long sea level time series should be continued, that instrumental developments should be monitored so as to ensure that the desirable 'absolute' character of measurement should not be degraded and that strenuous effort be expended to maintain datum stability. Indeed recommendations from the I.U.G.G. urge participating nations to set up high quality sea level monitoring stations for this purpose. In the second place it has been recognised that the quality of the existing data bank can be improved in retrospect and this is the current task of the FAGS Permanent Service for Mean Sea Level (PSMSL) operating from the Institute of Oceanographic Sciences, U.K. Although the PSMSL has worked for many years in accessing, collating, treating and publishing monthly and annual mean sea level values on a global basis, it has recently embarked upon a re-processing and re-publication programme. The point at issue is that many countries in the past adopted the practice of referring sea levels at a station to the primary national reference level, often at great distance from the marine station itself. This procedure had the unfortunate disadvantage of incorporating within the sea level time series all the uncertainties of geodetic levelling. Of late we have had the experience firstly of attempting to understand the anomalous sea level slopes which emerged when attempts were made to make spatial geodetic connexions between observed sea level at individual stations, and secondly of beginning to appreciate that geodetic procedures perhaps contain systematic error, or at least are demonstrably imperfect as evidenced by marked discrepancies between repetitive exercises. In consequence the PSMSL has revised the entire data bank with a view to the production at each station of a sea level record which is referred to a local fixed mark (Revised Local Reference). The revised series then are, hopefully, homogeneous with respect to datum. This arrangement attempts to delineate the respective responsibilities of oceanography and geodesy and, given a stable local mark, ensures a more representative sea level record. The work is now well advanced with the publication in 1976 of data from 250 stations in Europe, Africa and India. This was followed one year later with a publication of similar content incorporating stations from North and South America. The global coverage will be completed in 1978/9 with the publication of data from Japan, the Phillipines, Australasia and the Pacific Islands [Lennon and Spencer, 1976, 1977]. Albeit bearing in mind cautionary remarks expressed earlier, this updated data bank offers a new prospect for the global study of the mean ocean level trend in an attempt to identify eustatic changes. Although of great scientific interest in its own right, it is relevant in the present context to note that the programme has the potential to remove the final perturbation in sea level trends. Hitherto the subject has dealt with the relative movements of land and sea levels. If it is possible to identify the secular trend of ocean levels then the identification of crustal motion becomes a more exact science.

## Conclusion

The treatment of sea level time series to identify vertical crustal motion requires the elimination from the record of the effects of marine dynamic phenomena which in this context must be treated as noise.

Unfortunately these perturbing phenomena are not in many cases fully understood and perhaps this is particularly true of the southern hemisphere.

Despite the hazards and acknowledging the imperfections of historic observations, experience has shown, in certain specific cases, that it is possible to establish with some certainty the secular trends in the relative levels of land and sea surface at discrete stations.

Where vertical crustal motion is large with respect to the eustatic rise of sea level, say, of order 5mm per year and above, positive achievement is entirely possible.

Recent work of PSMSL offers to improve this situation by providing higher quality sea level data for treatment and also by making possible the identification of the common eustatic element.

Meanwhile techniques have been developed whereby the general interaction of atmosphere and ocean may be reproduced so that it is now feasible to construct a global model to attain this end in such a manner that not only eustasy, but also seasonal variations, with their inherent evidence of large scale ocean dynamics, and long period tides will emerge together with the evidence of crustal motion. PSMSL is currently planning such a treatment.

Nevertheless it is salutary to note that the stability of the apparent topography of the geoid itself is still uncertain and may present the ultimate obstacle to a definitive solution.

## References

Andrews, J.C., and P. Scully-Power, The structure of an East Australian current anticyclonic eddy, J. Phys. Ocean., 6, 756-765, 1976.

Chew, F., Advection of planetary vorticity in a western boundary current, Marine Geodesy, 1, 103-116, 1977.

Hamon, B.V., and G.R. Cresswell, Structure functions and intensities of ocean circulation off east and west Australia, Aust. J. Marine Freshwater Res., 23, 99-103, 1972.

Lennon, G.W., An investigation of secular variation of sea level in European waters, Ann. Acad. Sci. Fennicae, A.III, 90, 225-236, 1966.

Lennon, G.W., Sea level instrumentation, its limitations and the optimisation of the performance of conventional gauges in Great Britain, Int. Hydrographic Rev., XLVIII, 129-147, 1971.

Lennon, G.W., and N.E. Spencer, Monthly and annual mean heights of sea level, Permanent Service for Mean Sea Level, Natural Environment Research Council, U.K., 1976 and 1977.

Rossiter, J.R., An analysis of annual sea level variations in European waters, Geophys. J.R. astr. Soc., 12, 259-299, 1967.

Sturges, W., Sea level slope along continental boundaries, J. Geophys. Res., 79, 825-880, 1974.

Thompson, K.R., Regression models for monthly sea level, Marine Geodesy, in press.

Page Intentionally Left Blank

# Modelling of Physical Influences in Sea Level Records for Vertical Crustal Movement Detection

Edward G. Anderson
Department of Surveying Engineering, University of New Brunswick
Fredericton, N.B., Canada E3B 5A3

Abstract. The analysis of sea level time variations recorded by shore tide gauges for evidence of recent and contemporary vertical crustal movement is an established technique. However, the additional refinement of first modelling the numerous physical influences on sea level--which are of considerable importance, but additional complexity in the coastal zone--has not been fully exploited.

In this paper, attempts to specify and evaluate such physical influences are reviewed with the intention of identifying problem areas and promising approaches. For a routinely viable procedure, it is considered important to limit the input data, which would be needed on a continuing basis to model physical phenomena, to widely and regularly available data records.

An example of linear modelling based on air/water temperatures, atmospheric pressure, river discharges, geostrophic and/or local wind velocities, and including forced period terms to allow for the long period tides and Chandlerian polar motion is evaluated and applied to monthly mean sea levels recorded in Atlantic Canada. Refinement of the model to admit phase lag in the response to some of the driving phenomena is demonstrated. Spectral analysis of the residuals is employed to assess the model performance. The results and associated statistical parameters are discussed with emphasis on ellucidating the sensitivity of the technique for detection of local episodic and secular vertical crustal movements, the problem areas most critical to the type of approach, and possible further developments.

## Introduction

Vaníček [1978] has described a method whereby some of the known physical influences on sea level can be removed by a process of mathematical modelling. He concludes that, with such a model restricted to the effects of atmospheric pressure and temperature, river discharge, and three basic tidal constituents: "...we should be able to detect possible local episodic vertical movements (of duration more than 4 months) of magnitude of 10cm and more...". It is further stated that, by extending and refining the model to include additional known effects, it should be possible to see movements down to 5cm in the sea level record.

The work discussed in this paper is merely a continuation and extension of Vaníčeks approach. In particular, the basic model has been substantially retained, but refined by introducing an air-to-water heat transfer model with provision for time lag, and also by including wind stress terms based on either the observed wind vector or the geostrophic wind vector derived from the observed atmospheric pressure field.

So far, only results of preliminary and test computations are available, based on data from eight locations in the Canadian Maritime provinces (fig. 1). However, these results have been adequate to indicate some strengths and weaknesses in the basic approach and to suggest areas where immediate refinement could be advantageous. Also, some model improvements with the best potential for future development have been highlighted.

## The Problem and Proposed Approach

A detailed description of the problem, the proposed approach, and the reasons for choosing such an approach have already been set out by Vaníček [1978]. It will suffice here to provide a brief summary.

The problem is to devise a viable technique which would enable use of sea level records, derived from tide gauge observations, to detect contemporary vertical crustal movements over periods from a few months to several years. Both

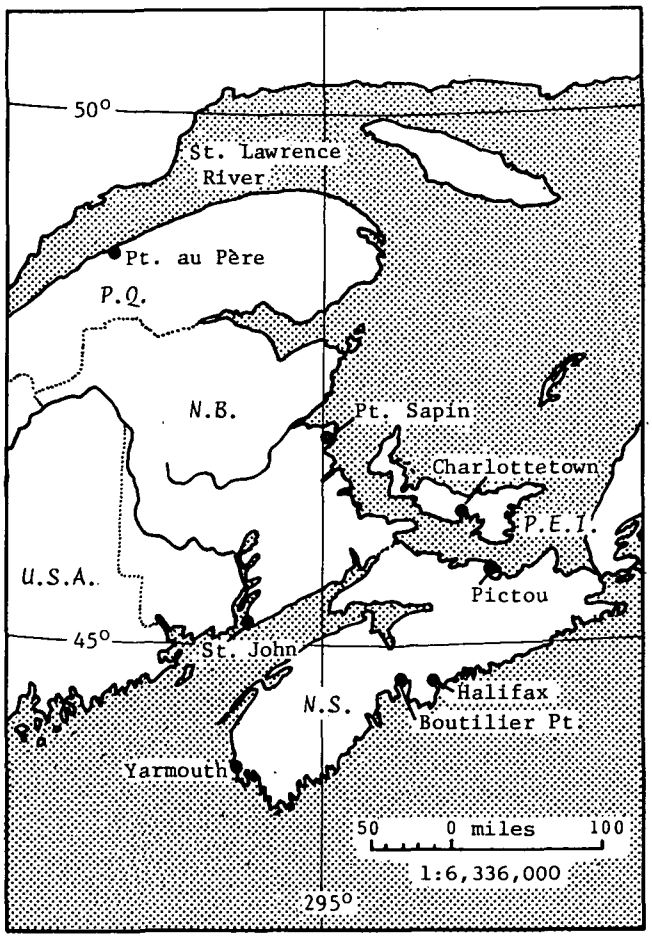



Figure 1. Map of Canadian Maritime provinces showing location of tide gauges

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280. The Ohio State Univ., Columbus, Ohio 43210.

secular movements (i.e. linear trend) and irregular or episodic movements are sought. To achieve sufficient accuracy in isolating these phenomena, a variety of other known physical influences on the sea level record must be removed. It is proposed that this be done by empirically modelling the known effects, so that much of the variance in the sea level record on that account can be eliminated.

Two critical factors are inherent in this approach. Firstly, the tide gauge environment may be substantially more complex than that which is experienced in the deep oceans. For example, esturine topography, the dynamics of coastal bathymetry, seasonal freezing of coastal waters, and dynamic river discharge influences are some of the many factors which may interact with the meteorological and tidal effects to anomalously perturb the coastal sea level record. Consequently, many of the well proven modelling techniques devised for the deep ocean environment and much of the data observed there may not necessarily be extrapolated with any validity to the tide gauge environment. Secondly, the method of modelling, to be viable for routine monitoring of vertical movements, should rely only on readily available long-term data records, collected on a regular basis. This is a significant and frustrating limitation.

## The Basic Model and Method

Basically, the model adopted (table 1) is a linear combination of the physical effects to be represented, so that the method might be referred to as a multiple linear regression analysis. However, certain components of the basic model may be represented by time series data which has been derived from additonal observations through sub-models. This is specifically so in the case of the water temperature and wind stress terms. The sub-models are not necessarily linear.

A solution for the unknown parameters, along with estimates of their variance and a correlation matrix, is obtained by the conventional least squares method. Actual implementation of this solution, however, is just a by-product of the optimum least squares spectral analysis algorithm [Wells and Vaníček, 1977] which is employed. This algorithm, based on the technique developed by Vaníček [1971], primarily provides the spectrum of the residual time series, $R(t_i)$. Two particular advantages claimed for this algorithm are: (a) suppression of the regression model components produces no movement of the inherent spectral peaks; and (b) equally spaced time series data is not mandatory. Spectral analysis of the residuals provides a supplementary quantitative assessment of the success of the regress-

The Linear Regression Model

$$S(t_i) = C_A + C_L t_i + C_P \delta P(t_i) + C_T \delta\tau(t_i) + C_D \delta D(t_i) + C_{W_t} \delta W_t(t_i) + C_{W_n} \delta W_n(t_i)$$

$$+ \sum_{j=1}^{5} A_j \cos(\omega_j t_i - \theta_j) + R(t_i)$$

Unknown Parameters

| | |
|---|---|
| $C_A$ | datum bias |
| $C_L$ | linear trend coefficient |
| $C_P$ | air pressure coefficient |
| $C_T$ | surface water temperature coefficient |
| $C_D$ | river discharge coefficient |
| $C_{W_t}$ | shoreline tangential component of wind stress coefficient |
| $C_{W_n}$ | shoreline normal component of wind stress coefficient |
| $A_j$ | amplitude of five periodic tidal components |
| $\theta_j$ | phase angle of five periodic tidal components |
| R | sea level residual time series |

Observed Time Series

(functions of time $t_i$)

| | |
|---|---|
| S | monthly mean sea level at tide gauge |
| $\delta P$ | monthly mean air pressure variation (with respect to temporal mean) |
| $\delta\tau$ | monthly mean surface water temperature variation |
| $\delta D$ | monthly mean (combined) river discharge |
| $\delta W_t$ | shoreline tangential component of observed or geostrophic wind stress |
| $\delta W_n$ | shoreline normal component of observed or geostrophic wind stress |
| $\omega_j$ | assumed constant frequency of five periodic tidal components as follows: $\omega_1$ = .6 mths. (semi-annual), $\omega_2$ = 12 mths. (annual), $\omega_3$ = 14.33 mths. (chandlerian or polar), $\omega_4$ = 8.847 yrs. (lunar perigee), $\omega_5$ = 18.613 yrs. (lunar nodal). |

Table 1. The linear regression model

ion model and provides diagnostic evidence by depicting the distribution of the remaining variance.

Refinement of the Model

Assuming the desirability of a linear model, two further possibilities remain available for improvement. The sub-models may be reformulated to better portray the physical processes or the total linear model may be extended to include additional terms, representing known or suspected physical effects. Both methods were employed, as exemplified respectively by the empirical sub-modelling of water temperature and the addition of wind stress terms to the basic model.

A Water Surface Temperature Model. While thermal variations have been cited as an important contribution to sea level changes, earlier analyses of the Maritimes data, using a simple linear term based on air temperature, produced unexpectedly small and weakly determined regression coefficients. It was suggested that this could be the result of two factors associated with the use of air, rather than water, temperatures. Clearly, the primitive model ignored time lag between air temperature variations and their consequent effect on water temperatures and thus the sea level. Also, air temperatures usually fall below freezing point during Maritime winters but not all of the gauge locations actually freeze, which suggests that a better thermal model, capable of accounting for the "winter non-linearity" should be constructed from water temperatures. Unfortunately, regular long-term records of water temperature are not readily available at most tide gauges. However, the viability of the method could be preserved by constructing an empirical sub-model for the air/water heat transfer process, based on long-term air and water surface temperature records available for the Halifax, N.S. tide gauge. The assumption might then be made, and possibly tested, that such a sub-model could be validly extrapolated to the other gauges. Accordingly, a linear heat transfer sub-model, incorporating time lag, was postulated as follows:

$$\tau(t_i) = hT(t_i + \delta t) + k, \qquad (1)$$

where $\tau(t_i)$ is the predicted water surface temperature at time $t_i$, $T(t_i + \delta t)$ is the observed air temperature at time $t_i + \delta t$, $\delta t$ is a time lag, h is a heat transfer coefficient, and k is a constant temperature bias. A least squares solution for the parameters $\delta t$, h, and k was obtained using the monthly mean air and water surface temperature series available for Halifax. Since the observed time series comprise discrete values, the admission of a time lag necessitates interpolation of the air temperature data. Though the simplicity of polynomial interpolation was attractive it was found, by experiment, to be considerably less reliable than fourier interpolation; a not entirely unexpected outcome when the "sine wave" form of the air temperatures is considered. When a fourier transformation is introduced into the model (1), the observation equations become, after linearization:

$$T(t_i+\hat{\delta t})\Delta h - [\hat{h} \sum_{j=1}^{m} j\omega A_j \sin(j\omega(t_i+\hat{\delta t}) - \theta_j)]\Delta t$$

$$+ k + [\hat{h}T(t_i+\hat{\delta t}) - \tau(t_i)] = v_i \quad (i=1,n); \qquad (2)$$

where $\delta t$ and h are estimates of the parameters $\delta t$ and h respectively, so that $\delta t = \hat{\delta t} + \Delta t$ and $h = \hat{h} + \Delta h$; $A_j$ and $\theta_j$ are respectively amplitude and phase fourier coefficients of the air temperature time series; $\omega$ is the harmonic frequency; $v_i$ are the discrepancies in the observations; m is the number of harmonics; and n is the number of observations.

Least squares estimates of the parameters and their standard deviations are listed in table 2, along with other relevant information. The estimates are apparently well determined, mainly because of the large number of degrees of freedom, and the overall RMS error of prediction--at about 1°C--is quite satisfactory. Distribution of the residuals was significantly normal. Use of these values in the sub-model (1) also requires fourier interpolation of the air temperature series. Extrapolation to other locations has not yet been attempted, as additional water surface temperature data is being sought so that the spatial coherence of the coefficients $\delta t$, h, and k can be tested.

Wind Stress. Two methods of modelling the wind stress effect on sea level have been considered. The first relies on a sub-model for the geostrophic wind component, derived from the air pressure gradient field as follows:

$$\begin{bmatrix} W_E \\ W_N \end{bmatrix} = \begin{bmatrix} (R\rho\omega \sin 2\phi)^{-1} \frac{\partial P}{\partial \lambda} \\ (2R\rho\omega \sin \phi)^{-1} \frac{\partial P}{\partial \phi} \end{bmatrix}; \qquad (3)$$

where $W_E$ and $W_N$ are the east and north components respectively of the geostrophic wind; R is the earth's radius; $\rho$ is air density; $\omega$ is the

| Coefficient | Least Squares Estimate | Standard Deviation |
|---|---|---|
| $\delta t$ (mths.) | -0.66 | 0.03 |
| h | 0.628 | 0.009 |
| k (°C) | 3.05 | 0.12 |
| Data | | |
| Span: Jan. 1927 to Nov. 1973 excluding Jul. 1933 to Jan 1946 | | |
| Number of observations, n = 412 | | |
| Number of fourier coefficients, m = 206 | | |
| Statistical analysis | | |
| RMS residual = 1.035°C | | |
| Maximum residual = 4.4°C | | |
| Variance ratio = 1.06 | | |

Table 2. Air/water heat transfer model coefficients for Halifax, N.S.

earth's angular velocity of rotation; and $\partial P/\partial\lambda$, $\partial P/\partial\phi$ are the components of the barometric pressure (P) gradient with respect to longitude ($\lambda$) and latitude ($\phi$). Even though this model may only account for part of the actual wind field, it has been introduced for use at gauge locations where suitable observed wind records are not available. The suitability of the wind records is stressed here because, in many instances, monthly mean magnitudes and directions are seperately compiled from the hourly mileages, without proper vector summations. Such data is thus not suitable for modelling wind stress effects. However, early attempts using the geostrophic wind in the Maritimes region--resolved into components tangential and normal to the local shoreline and squared--have not been particularly successful [Vaníček, 1977]. It is now suspected that these results may have suffered from computational difficulties and the present author is attempting to rectify the analyses.

As a means of verifying the significance of the wind stress effect, the second method--which relies on the observed monthly mean wind vector--has been tested at Halifax. Although the Halifax wind records are burdened with the unsuitable compilation techniques previously mentioned, it was possible to recompute, without undue effort, correct monthly vector averages since the records included monthly mean wind mileages within a number of "direction sectors". Prior to 1967 the horizon was divided into 8 such sectors of 45° each, but since that date 16 sectors have been employed. North and east components ($W_N$ and $W_E$) of the observed monthly mean wind vector were thus easily derived. These components are employed in the wind stress sub-model thus:

$$\begin{bmatrix} W_t \\ W_n \end{bmatrix} = \begin{bmatrix} (W_N \cos\alpha + W_E \sin\alpha)^2 \\ (W_E \cos\alpha - W_N \sin\alpha)\,|\,(W_E \cos\alpha - W_N \sin\alpha)\,| \end{bmatrix}; \quad (4)$$

where $W_t$ and $W_n$ are, respectively, tangential and normal to shoreline components of the wind stress and $\alpha$ is the shoreline azimuth.

## Results

Results for four of the eight gauge locations are displayed in table 3 and figures 2 and 3. The analyses of the remaining gauge records (i.e. Point Sapin, Pictou, Boutilier Point, and Yarmouth) are not included here because the usefulness of the results is considerable diminished by the short data spans--all less than 9 years. In some cases, these additional results provide evidence of spatial coherence of the phenomena studied. This, as well as any significant disagreements with the results presented, will be mentioned in the relevant discussion below.

### Sea Level Data

The first four lines of table 3 quantitatively describe the observed sea level time series $S(t_i)$ for the selected tide gauges and they are depicted graphically in figure 2. The disjoint nature of the Pointe au Père record should be noted.

| Location of Tide Gauge | Pointe au Père | Charlottetown | St. John | Halifax |
|---|---|---|---|---|
| Number of observations | 422 | 370 | 480 | 564 |
| Number of years | 47 | 31 | 47 | 47 |
| Epoch | 1927-1974 | 1943-1974 | 1927-1974 | 1927-1974 |
| Number of gaps in data | 7 | 2 | 5 | 0 |
| Datum bias, $C_A$ (cm) | 27.6 ± 0.9 | 158.6 ± 0.4 | 426.4 ± 0.5 | 110.1 ± 0.3 |
| Linear trend $C_L$ (cm/cent.) | 7.6 ± 2.3 | 27.0 ± 2.2 | 37.6 ± 1.6 | 43.4 ± 1.1 |
| Tide: Amplitude $A_j$ (cm) | | | | |
| Semi-annual | 2.2 ± 0.4 | 0.9 ± 0.3 | 2.1 ± 0.3 | 0.9 ± 0.2 |
| Annual | 1.5 ± 0.5 | 4.6 ± 0.4 | 1.4 ± 0.7 | 4.0 ± 0.3 |
| Chandlerian | 0.7 ± 0.4 | 0.7 ± 0.3 | 1.3 ± 0.3 | 0.6 ± 0.2 |
| Lunar perigee | 1.6 ± 0.4 | 0.7 ± 0.3 | 0.3 ± 0.3 | 0.2 ± 0.2 |
| Lunar nodal | 0.5 ± 0.4 | 1.6 ± 0.3 | 1.2 ± 0.3 | 0.3 ± 0.2 |
| Tide: Ascending node date, $N_j$ (yrs) | | | | |
| Semi-annual | 1926.81 ± 0.02 | 1926.83 ± 0.02 | 1926.81 ± 0.01 | 1926.81 ± 0.02 |
| Annual | 1926.39 ± 0.05 | 1926.64 ± 0.04 | 1926.80 ± 0.04 | 1926.69 ± 0.02 |
| Chandlerian | 1926.66 ± 0.10 | 1926.49 ± 0.07 | 1926.56 ± 0.04 | 1926.52 ± 0.06 |
| Lunar perigee | 1932.73 ± 0.38 | 1932.80 ± 0.55 | 1930.28 ± 1.13 | 1926.59 ± 1.39 |
| Lunar nodal | 1936.50 ± 3.00 | 1930.16 ± 0.52 | 1933.54 ± 0.70 | 1928.21 ± 2.31 |
| Temperature, $C_T$ (cm/°C) | -0.15 ± 0.09 | -0.033 ± 0.087 | 0.22 ± 0.07 | 0.004 ± 0.076 |
| Air pressure, $C_P$ (cm/mbar) | -0.74 ± 0.11 | -0.90 ± 0.06 | -0.51 ± 0.07 | -0.63 ± 0.05 |
| River discharge, $C_D$ (cm/m$^2$s$^{-1}$) | 0.0006 ± 0.0001 | --- | 0.004 ± 0.0004 | --- |
| Tangential wind stress, $C_{W_t}$ (cm/m$^2$s$^{-2}$) | --- | --- | --- | -0.04 ± 0.06 |
| Normal wind stress, $C_{W_n}$ (cm/m$^2$s$^{-2}$) | --- | --- | --- | 0.85 ± 0.07 |
| St. deviation original record $S(t_i)$ (cm) | 7.26 | 6.53 | 9.10 | 7.80 |
| St. deviation of residuals $R(t_i)$ (cm) | 5.96 | 3.44 | 4.19 | 3.31 |
| % reduction in variance | 33% | 72% | 79% | 82% |

Table 3. Linear regression analysis of selected tide gauge records

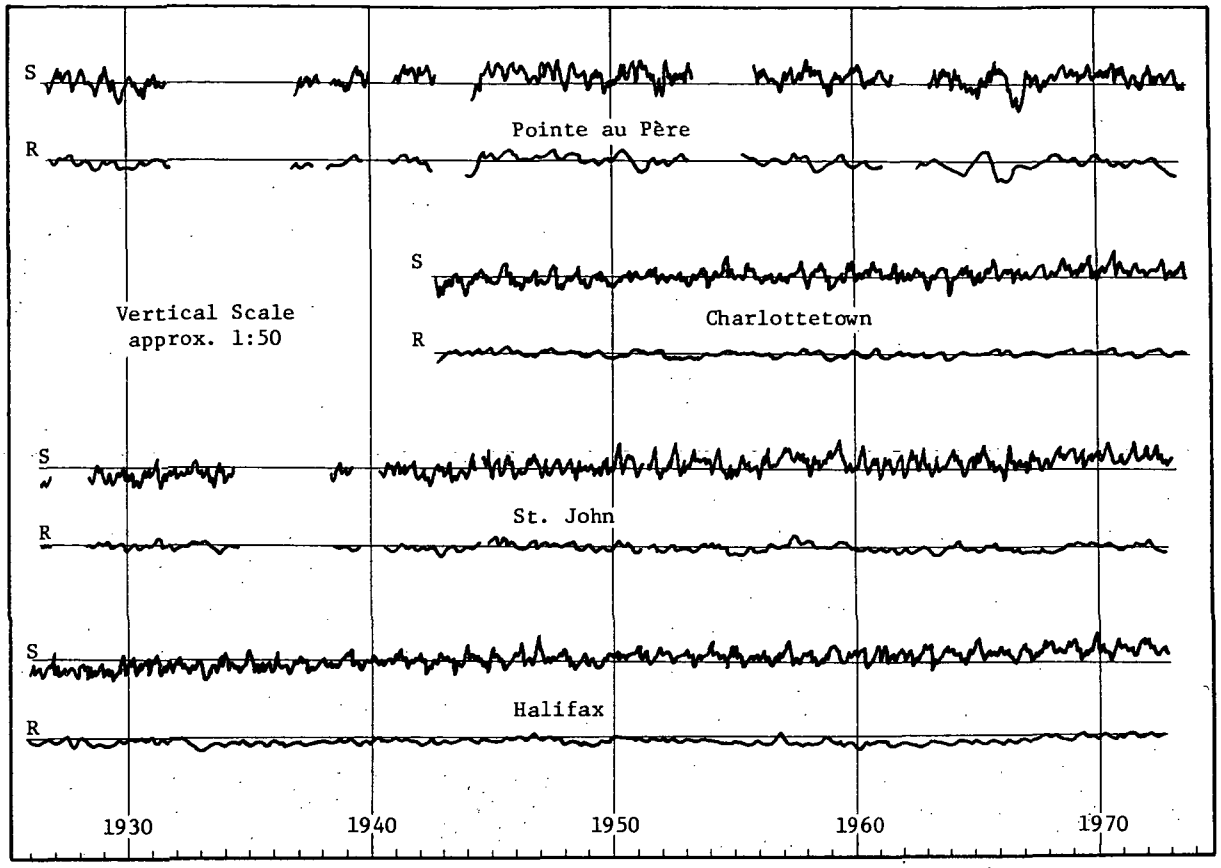



Figure 2. Original tide gauge records (S) and filtered residuals after least squares regression analysis (R) plotted on same vertical scale

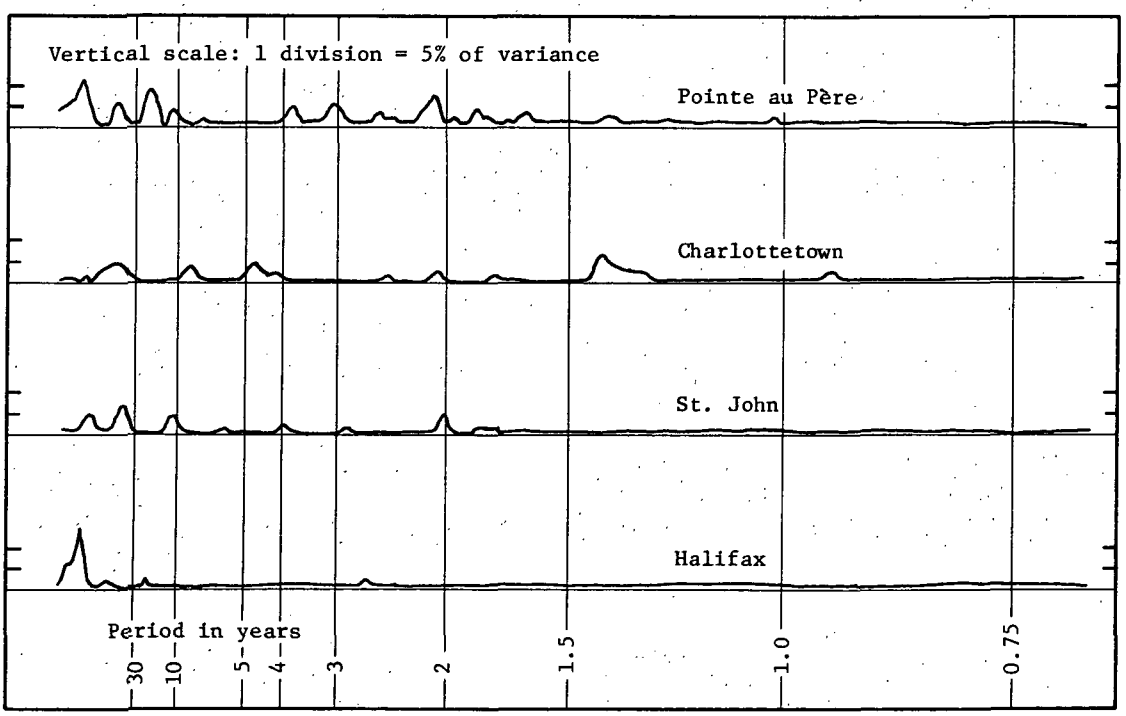



Figure 3. Spectra of unfiltered residuals

Datum Bias

The datum bias terms ($C_A$) merely reflect the arbitrary origin of the particular tide gauge scale.

Linear Trend

This quantity ($C_L$) may be expected to combine secular vertical crustal movements and eustatic sea level variations. According to the standard deviations obtained, these results are quite well determined. Generally, there appears to be a high degree of spatial coherence, which is further supported by the results for the four omitted gauges, despite their lower accuracy. Only Pt. Sapin and Pictou were exceptional; the latter displaying a negative trend. Overall, the linear trend appears to increase fairly smoothly for locations towards the south and east. Presumably this reflects a regional crustal movement phenomenon, as the eustatic effect should be similar at all locations.

Tides

Table 3 includes results for both the amplitude ($A_j$) and the ascending node date ($N_j$) for the five tidal components. Ascending node date is derived from the phase angle $\theta_j$ of the tide by the relationship:

$$N_j = B_j + (\theta_j - \frac{\pi}{2} \pm 2\pi n)/\omega_j, \qquad (5)$$

where $B_j$ is the date of the time origin used for the particular data series (usually the middle of the month immediately preceding the first month of data). The dates given in table 3 have been adjusted as necessary by an integral number of periods (n) for purposes of comparison.

Semi-annual Tide. The phase of this component is very well determined and spatially consistent. A systematic difference in amplitude is apparent between those locations affected by river discharge and those which are not. It seems likely that the greater amplitudes found for the gauges where river discharge is significant may result from poor modelling of that effect. At St. John, moderate correlation was found between the phase of the semi-annual tide and the river discharge.

Annual Tide. Amplitudes of the annual tides are not as well determined and again the spatial coherence seems to depend on the presence, or absence of river discharge. There is, however, also a strong interaction with the temperature effect. For instance, at Halifax fourier transformations of both the air and observed water surface temperature series yielded amplitude coefficients for the annual harmonic at least an order of magnitude larger than those of any other period. High correlation coefficients between both amplitude and phase of the annual tide and the temperature effect (0.60 and 0.87) respectively at Halifax) are thus not unexpected. This evidence suggests that these two effects are not being adequately isolated and, indeed, the philosophy of attempting to do so might be questionable. Ultimately, both the annual tide and the strong annual component of temperature variation are merely different manifestations of the same motivating energy source: the earth's orbital motion. Poor coherence in the phase of the annual tide may also result from the influence of the temperature effect.

Chandlerian (Polar) Tide. Good spatial consistency and satisfactory precision are achieved here. The amplitudes and phases are also quite comparable with results obtained for several east coast ports in the U.S.A. [Currie, 1975; Vaníček, 1978].

Lunar Perigee and Nodal Tides. The precision attained in evaluating these longer period tides is consistent with the total length of the time series analysed. Despite fairly weak determinations, the amplitudes agree reasonably with the empirical knowledge of these effects, and also with the estimated equilibrium tides [Rossiter, 1966].

Temperature Effect

Generally, the temperature coefficient ($C_T$) is so weakly determined as to be meaningless. Indeed, the negative results obtained at Pt. au Père and Charlottetown would be contrary to expectation based on an understanding of the physical process involved. At all locations except Halifax, air temperatures were used directly in the regression model. However, a test implementation at Halifax of the water surface temperature sub-model described above, does not noticeably improve the result. In view of this outcome, further investigation of the significance of the interaction between the annual tide component and the temperature effects is warranted. It seems possible that the lack of constraints on the amplitude and phase of the annual tide allows undue absorbtion of the data variance at that frequency--thus reducing the temperature effect--and simultaneously falsifying the determination of the tidal coefficients.

Air Pressure Effect

These coefficients ($C_p$) are well determined and are substantially in agreement with the theoretical model and empirical studies of this effect. Taking into account the maximum observed range of air pressure at the gauge locations, the coefficients obtained represent a minimum equivalent sea level variation of about 10 cm at St. John and a maximum of about 20 cm at Charlottetown.

River Discharge Effect

River discharge effects are expected to be highly variable, depending on the local estuary configuration. The coefficients ($C_D$) obtained are equivalent to maximum sea level variations of approximately 12 cm and 8 cm at Pt. au Père and St. John respectively.

## Wind Stress

Replacement of the geostrophic wind sub-model (eq. 3) by the observed wind sub-model (eq. 4) at Halifax appears to have been successful. The tangential wind stress effect is effectively negligible as anticipated. In contrast, the normal component is quite substantial, amounting to an equivalent maximum sea level variation of approximately 20 cm. Further study of this effect will be directed towards establishing an empirical relationship between the geostrophic and observed wind models, with the aim of improving the geostrophic model for use where suitable observed wind vectors are not routinely available.

## Residuals

A worthwhile reduction in variance of the sea level record is evident when the standard deviation of the residuals $R(t_i)$ is compared with that of the original data (table 3, last three lines). The same result is depicted graphically in figure 2, where the plot of residuals may be compared visually with the original record. To facilitate this visual comparison, the residuals displayed in figure 2 (but not the results in table 3) have been subjected to a low-pass filter to remove the high frequency components with a period shorter than about 4 months.

Spectra of Residuals. Figure 3 illustrates the small amount of power remaining in the spectra of the unfiltered residuals. If reasonable limits are imposed on the band-width, so as to be consistent with the data span, there are no spectral peaks which exceed about 8% of the variance. The noticeable remaining variance in the Pt. au Père residuals may relate to the disjointed data series or it may reflect the additional complexity of the St. Lawrence River environment.

## Conclusions

While the results presented here must be regarded as preliminary--and any dependent conclusions consequently tentative--the following summary will serve to highlight the strengths and weaknesses of the method and hence usefully direct further investigations towards realization of its full potential.

(1) A linear sub-model of the air/water heat transfer process, incorporating time lag, can be constructed empirically with acceptable reliability. However, it would appear that the use of water surface temperatures--either derived from such a model or actually observed--in the overall model for consequent sea level variations, does not produce a reliable or coherent thermal coefficient and does not significantly reduce the variance of the sea level record. It is suggested that a likely cause of this result may be the interaction of the highly correlated annual tide effect. Even though the coefficient for the thermal effect might actually be small--for example, Thompson [pers. comm. reported in Anderson, 1978] has obtained results which suggest that this is the case in the waters near U.K.--its weak determination and poor coherence indicate a remaining problem in the analysis.

(2) Use of properly derived observed wind vectors produces acceptable wind stress coefficients and contributes significantly to the reduction of sea level variance. The utility of a geostrophic wind sub-model remains uncertain, but development of an empirical relationship between such a model and the observed wind field would seem to be worthwhile.

(3) A very satisfactory determination of the linear trend coefficient is possible and this result appears to provide a reliable estimate of regional vertical crustal movement. This is especially significant in the context of tide gauge applications for continental levelling datums.

(4) While the signatures of the polar and longer period tides seem to have been adequately identified and separated from the sea level record, there are remaining difficulties with the annual and semi-annual tides. The treatment as free parameters of both amplitude and phase of these tides, as well as the coefficients of thermal and river discharge effects, apparently fails to adequately isolate their contributions to the sea level variation. Further study of this problem is warranted.

(5) Since modelling of wind stress and barometric pressure effects is reasonably successful, further refinement of the model, by inclusion of steric level and ocean current terms for instance, might usefully be pursued. This may even assist in better resolution of the tidal terms.

(6) Even in its present preliminary status, the modelling process provides sufficient reduction of variance in the sea level record for its utilization in identification of episodic movements at the 10 cm level. This facility may be further enhanced by investigating the differences in the residual variations between pairs of tide gauge locations.

## Acknowledgements

Sponsorship of the investigations described here was provided by the Geodetic Survey of Canada, Department of Energy, Mines and Resources under part of research contract DSS# 01SU.23244-7-4913. Some of the results used in table 3 were determined by M. N. Nassar, who also developed portions of the computer programmes.

## References

Anderson, E. G. 1978 : Report on Official Visits to Several European Institutions ... in the Period July 21 to August 6, 1978. Unpubl. report to Geodetic Survey of Canada, Aug. 1978.

Currie, R. G. 1975 : Period, $Q_p$ and Amplitude of the Pole Tide, Geophys. J. R. Astr. Soc., 43, 73-86.

Rossiter, J. R. 1966 : Long-term Variations in Sea-level, The Sea, Vol. 1 (M. N. Hill ed.), Interscience.

Thompson, K. 1978 : Institute of Oceanographic Sciences, Bidston Observatory, U.K., personal communication.

Vaníček, P. 1971 : Further Development and Properties of the Spectral Analysis by Least Squares, Astrophys. Space Sci. 12, 10-73.

Vaníček, P. 1977 : Sea Level Variations in Maritime Canada. Unpubl. status report for Geod. Surv. of Canada.

Vaníček, P. 1978 : To the Problem of Noise Reduction in Sea Level Records Used in Vertical Crustal Movement Detection. Phys. Earth Planet. Interiors (in press).

Wells, D. E., and P. Vaníček, 1977 : Least Squares Spectral Analysis, Computers and Geosci. (submitted).

# The Gravity Method and Interpretive Techniques for Detecting Vertical Crustal Movements [1]

Robert C. Jachens
U.S. Geological Survey, 345 Middlefield Road
Menlo Park, California 94025

Observations of temporal variations of gravity can be used as an inexpensive and rapid means of detecting, monitoring, and studying crustal deformation associated with many active geologic processes. Such observations can yield qualitative or semi-quantitative information on elevation changes and, when combined with independent elevation data, can yield information about changes of the subsurface density field arising from both subsurface displacements and temporal variations of the density of materials in the subsurface. The effectiveness of gravity techniques in specific cases of crustal deformation depends on the configuration of the local gravity field, the physical processes involved in the deformation, and the accuracy with which temporal gravity variations can be measured.

Assuming that the effects of earth tides, and variations in atmospheric pressure have been removed, gravity measured at a fixed point on the surface of the earth can vary with time as a result of two factors: (1) displacement of the observation point along the free-air gravity gradient and (2) variation of the subsurface density field. Generally, both factors result in gravity variations of a few microGals per centimeter of vertical displacement. The normal vertical gradient of gravity is approximately -3.09 μGal/cm whereas the actual free-air gradient typically may differ from this value by ±5% and in special situations may differ by more than 15% (Hammer, 1970). Gravity changes caused by variations in the subsurface density field accompanying deformation may enhance, subdue, or dominate gravity changes resulting from vertical displacement, and the relation between gravity change and elevation change, $\Delta g/\Delta h$, may assume a wide range of values.

Some possible relations of gravity change and elevation change have been derived on the basis of theoretical considerations. The values of $\Delta g/\Delta h$ discussed below are based on simple crustal models and are presented primarily to illustrate the possible variability in $\Delta g/\Delta h$ for different geologic processes. All values include both the effect of vertical displacement along a normal free-air gradient and the effect of changes in the subsurface density field.

In a numerical model study of a dilating sphere buried in a homogeneous elastic half-space, Rundle (1978) found $\Delta g$ proportional to $\Delta h$ and the ratio $\Delta g/\Delta h$ equal to -3.1 μGal/cm. For this model, the gravity change is equal to the free-air effect due to uplift. If the increased volume in this model were partially or completely filled with water, $\Delta g/\Delta h$ would be slightly smaller. In the same study Rundle found that $\Delta g$ caused by thrust movement on an infinitely long dipping fault buried in a homogeneous elastic half-space also was proportional to $\Delta h$. For a medium with density $\rho = 2.8$ g/cm$^3$, $\Delta g/\Delta h$ is equal to -1.9 μGal/cm or equivalent to the free-air effect due to uplift plus the gravity effect due to mass added to the vertical section, as Barnes (1966) proposed.

In a simple model of a homogeneous elastic plate of rectangular cross section and infinite length and subject to uniform horizontal compression or extension, $\Delta g$ again is proportional to $\Delta h$. In this case, the gravity changes due to changes of elevation and changes in subsurface density field are nearly equal and tend to cancel. Thus the value of $\Delta g/\Delta h$ is approximately 0 μGal/cm. Numerical calculations of deformation and gravity change resulting from surface loads applied to radially symmetric, elastic earth models (Farrell, 1972) show a $\Delta g/\Delta h$ value near -2.3 μGal/cm.

When deformation is associated with processes dominated by fluid movement such as magma movement in volcanic areas or ground-water movement in areas subject to ground-water extraction, possible values of $\Delta g/\Delta h$ cover a wide range. The wide range is due in large measure to the range of subsurface volume changes possible in response to changes in pore pressure. For example, in some areas ground-water extraction is not accompanied by any appreciable subsidence of the surface whereas in other areas the removal of ground water appears to be accompanied by almost a complete collapse of the resulting voids (Poland and Davis, 1969). Furthermore, in some areas, a certain amount of ground water can be extracted before appreciable subsidence begins (Riley, 1970). Here, the relation between $\Delta g$ and $\Delta h$ probably would be nonlinear. Analogous behavior can be expected in volcanic areas.

$\Delta g/\Delta h$ relations measured in a limited number of cases are in general agreement with the model results, although excep-

---

1/ Most of the material will be released as part of a paper in Proceedings of Conference VI, Measurements of Stress and Strain Pertinent to Earthquake Prediction, U.S. Geological Survey Open-File Report.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

tions do exist. Barnes (1966) and Oliver et al. (1975) remeasured gravity in regions that had undergone deformation associated with slip on subsurface faults. Barnes found that many observations of $\Delta g/\Delta h$ in southern Alaska fell close to a value of -1.97 μGal/cm whereas $\Delta g/\Delta h$ values at stations along a profile extending northeast from Valdez were closer to the normal free-air gradient. An anomalous value of $\Delta g/\Delta h$ was found near Anchorage, an area in which a very small gravity change was associated with nearly 1 m of subsidence. Oliver and his coworkers determined a value of $-2.15 \pm 0.26$ (s.d.) μGal/cm associated with deformation accompanying the 1971 San Fernando, California earthquake. Jachens et al. (1976) found a very good correlation between gravity changes and elevation changes that occurred during the November 1975 deflation of Kilauea Volcano, Hawaii. There, $\Delta g/\Delta h$ equaled $-1.71 \pm 0.05$ (s.e.) μGal/cm. W. E. Strange and D. G. Carroll (W. E. Strange, written comm., 1977) studied the relation between gravity change and elevation change resulting from groundwater extraction in the San Joaquin Valley of California. They found a $\Delta g/\Delta h$ value near -3.0 μGal/cm in areas where water was withdrawn from confined aquifers, but a simple relation did not exist in areas where water was extracted from unconfined aquifers. Isherwood (1977) reported gravity changes and subsidence over a producing geothermal field at The Geysers, California. He found a $\Delta g/\Delta h$ relation of about 2.5 μGal/cm, indicating that gravity changes due to loss of fluid from the subsurface were larger than those due to vertical displacement.

Measurements of gravity change and elevation change have been reported that do not fit the simple models discussed above. Kisslinger (1975) reported that $\Delta g/\Delta h$ values of -3.7 μGal/cm and near -10 μGal/cm accompanied some phases of the Matsushiro, Japan earthquake swarm. Fujii (1976) discussed repeated gravity surveys before and after the 1973 Nemuro-hento-oki earthquake in southeastern Hokkaido, Japan. Coseismic and postseismic gravity changes were as large as 400 μGal. No values were given for the associated elevation changes, but they were described as being too small to affect the observed gravity values. In another area of Japan, the Muira Peninsula, which was uplifted 1-2 m during the 1923 Kanto earthquake, repeated gravity and leveling surveys conducted since 1955 have shown the expected inverse correlation between gravity change and elevation change (Hagiwara, 1974). However, the gravity changes are about ten times larger than expected. These results may be slightly affected by ground-water fluctuations beneath the reference station at Tokyo.

On the basis of the above discussion, some general statements can be made about temporal variations of gravity as related to crustal deformation. First, gravity changes accompany most types of deformation, and a knowledge of such changes can yield information about the spatial distribution of the deformation. Second, unambiguous estimates of elevation changes are not possible on the basis of gravity data alone. Third, in some situations (for example, San Fernando, California and Hawaii), the measurements of $\Delta g$ and $\Delta h$ at a few locations may permit the determination of $\Delta g/\Delta h$ that then can be used to infer elevation changes from measured gravity changes. Finally, the wide range of possible values of $\Delta g/\Delta h$ shows that, in some cases, observed values of this relation can effectively constrain the interpretation of the causes of the deformation.

## References


Barnes, D. F., Gravity changes during the Alaska earthquake, J. Geophys. Res., 81, 451-456, 1966.

Farrell, W. E., Deformation of the earth by surface loads, Rev. Geophysics and Space Physics, 10, 761-797, 1972.

Fujii, Yoichiro, Seismic crustal movement and associated gravity change, J. Geod. Soc. Japan, 22, 308-310, 1976.

Hagiwara, Yukio, Geodesy in Japan, Geophys. Surveys, 1, 305-323, 1974.

Hammer, S., The anomalous vertical gradient of gravity, Geophysics, 35, 153-157, 1970.

Isherwood, W. F., Geothermal reservoir interpretation from change in gravity, in Kruger, P. and Ramey, H. J., Jr. (Eds.), Proceedings third workshop on geothermal reservoir engineering, Stanford University, Stanford, California, 1977.

Jachens, R., G. Eaton, P. Lipman, and R. Okamura, Temporal gravity variations associated with the November 1975 deflation of Kilauea Volcano (abs.), Amer. Geophys. Union Trans., 57, 1015-1016, 1976.

Kisslinger, G., Processes during the Matsushiro, Japan, earthquake swarm as revealed by leveling, gravity, and spring-flow observations, Geology, 3, 57-62, 1975.

Oliver, H. W., S. L. Robbins, R. B. Grannell, R. W. Alewine, and Shawn Biehler, Surface and subsurface movements determined by remeasuring gravity, in Oakshott, G. B., (Ed.), San Fernando earthquake of February 9, 1971, Calif. Div. of Mines and Geology Bull. 196, 195-211, 1975.

Poland, J. F., G. H. Davis, Land subsidence due to withdrawal of fluids: in Reviews of Engineering Geology II, Geol. Soc. of Amer., 187-269, 1969.

Riley, F. S., Analysis of borehole extensometer data from central California, International Association of Scientific Hydrology, UNESCO Pub. 89, 423-431, 1970.
Rundle, J. B., Gravity changes and the Palmdale uplift, Geophys. Research Letters, 5, 41-44, 1978.

Page Intentionally Left Blank

# Review of Canadian Experience in Precise Gravimetry

A. Lambert
Gravity and Geodynamics Division, Earth Physics Branch
Department of Energy, Mines and Resources, Ottawa, Ontario Canada K1A 0Y3

Abstract. Results of gravity observations made in Canada from 1974 to 1978 have been reviewed, in order to estimate the true accuracy of present-day gravimetry and thereby assess the potential capability of the method for detecting crustal movements. The standard error of the mean of ties is 15-20 nm/s$^2$. Inter-instrument comparisons and other tests show, however, that a more realistic estimate of D meter accuracy is 30-40 nm/s$^2$. This accuracy can only be maintained over the long term where uncertainties in gravimeter calibration curves are minimized by resetting to the same dial reading on the resurveys. A further deterioration in accuracy to 40-50 nm/s$^2$ occurs where reliance is placed on presently available D meter calibration curves. Despite the present accuracy limitations significant time variations in gravity of 100-150 nm/s$^2$ are seen over spatial scales of 10-100 kilometers in Canada over a period of several months.

## Accuracy of Gravity Measurements

A distinction must be made between short-term and long-term accuracy of gravimeters, in order to make a realistic evaluation of gravimeter performance. The standard error of the mean of a set of ten consecutive gravity ties with a LaCoste and Romberg model D gravimeter under average transport conditions is 15-20 nm/s$^2$. In the hand-carried mode this value reduces to about 10 nm/s$^2$. When hand-carried measurements of near-zero gravity differences are made, gravimeter calibration is not important and, for practical purposes, 10 nm/s$^2$ can be taken as an estimate of the accuracy of the gravity difference determination. The reliability of this short-term accuracy estimate for three model D gravimeters was tested by making hand-carried measurements of sixteen small (<400 nm/s$^2$) gravity differences between adjacent (<1 m) stations. Means of measurements of these small gravity differences over a period of two years exhibited a standard deviation of 30-40 nm/s$^2$, not 10 nm/s$^2$ as expected. Since the gravity difference between such closely spaced stations should remain constant, the presence of significant systematic instrumental effects is indicated.

In typical gravity networks where larger gravity differences are measured, uncertainties in gravity calibration could be important in estimating gravimeter accuracy. In Canadian networks the effect of the unknown structure of the gravimeter dial factor curves is minimized by always resetting to within 1000 nm/s$^2$ of the same dial readings in the resurveys. In addition, the overall scale factor for the gravimeter is controlled by one or more independent calibration ranges. Therefore, 30-40 nm/s$^2$ is a realistic estimate of gravimeter accuracy in practical networks.

This estimate of the long-term accuracy based on the measurement of near-zero differences is borne out by comparisons of time changes seen by different gravimeters throughout various precise gravity networks in Canada. European tests (Kiviniemi, 1974; Brein et al., 1977) on LaCoste and Romberg model G gravimeters also show similar inter-instrument discrepancies. Thus, the repeatability of gravity ties in the short-term tends to give an over-optimistic estimate of the long-term accuracy of spring gravimeters. The contribution of temperature, pressure, magnetic effects, mechanical "sets" and levelling errors to the long-term uncertainties in gravimetry are being investigated by many groups but, so far, no consensus on the causes has emerged.

## Gravimeter Dial Factor Curves

Dial-factor curves for LaCoste and Romberg model D gravimeters can be determined in the laboratory by a method devised by the manufacturer involving the addition and removal of a small calibrated weight equivalent to a 200 μm/s$^2$ change in gravity. Significant (>0.1%) variations in calibration factor across the instrument range are revealed by these tests (Lambert et al., 1978). These variations are probably due to nonlinearities in the lever system that is activated by the dial screw to null the gravimeter (Harrison, personal communication). Although application of such calibration curves to the instrument readings is obviously important, it has been found that significant and repeatable inter-instrument discrepancies remain. The amplitude of these discrepancies suggests the presence of further uncertainty in gravimeter calibration equivalent to ±40-50 nm/s$^2$. These results require that at least two instruments be operated simultaneously in a network to ensure that time-changes in gravity be monitored continuously at the 30-40 nm/s$^2$ level of accuracy in the event of the demise of one instrument.

## Stability of the Gravity Field

In spite of the more-conservative estimates of accuracy, there is indisputable evidence of significant relative variations of gravity with periods of a year or less. Semi-annual resurveys of local-scale (<100 km) precise gravity networks in both eastern and western Canada show variations of up to 150 nm/s$^2$ (Dragert, Liard and Lambert, 5th annual meeting, Canadian Geophysical Union, London, Ontario, 1978). These changes appear to be a combination of seasonal effects of the type discussed by Lambert and Beaumont (1977), as well as possible superimposed tectonic effects, presently under investigation. Significant effects due to ocean-tide attraction at coastal stations have also been observed (Lambert and

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

Bower, 1978). It is obviously important to understand these short period variations, if longer term trends due to crustal movements are to be delineated.

## Conclusion

There is a discrepancy between short-term precision determined from repeated gravity ties and long-term accuracy revealed by inter-instrument comparisons and other tests. An isolated gravity difference can be determined in reality with a standard error of 30-40 $nm/s^2$. In practice, the uncertainties in measured gravity differences in a network are reduced somewhat by the network structure. Useful investigations into the time variations of the gravity field and their causes can be accomplished with presently available instrumentation.

Acknowledgements. The efforts of H. Dragert and J. Liard in the collection and synthesis of the gravity data is gratefully acknowledged. R.K. McConnell suggested useful improvements to the text. Earth Physics Branch contribution no. 756.

## References

Brein, R., C. Gerstenecker, A. Kiviniemi, and L. Pettersson, Report on high precision gravimetry, Tekniska shrifter (Professional Papers), 1, National Land Survey, Gavle, Sweden, 1977.

Kiviniemi, A., High precision measurements for studying the secular variation in gravity in Finland, Publications of the Finnish Geodetic Institute, No. 78, Helsinki, Finland, 1974.

Lambert, A., and C. Beaumont, Nano variations in gravity due to seasonal ground-water movements: Implications for the gravitational detection of tectonic movements, J. Geophys. Res., 82, 297-306, 1977.

Lambert, A., and D.R. Bower, Gravity tide effects on precise gravity surveys, Proc. 8th Int. Symp. Earth Tides, Bonn (in press).

Lambert, A., J. Liard, and H. Dragert, Canadian precise gravity networks for crustal movements studies, Tectonophysics (in press).

# Investigations on Vertical Crustal Movements in the Venezuelan Andes by Gravimetric Methods

Hermann Drewes
Escuela de Ingeniería Geodésica, Universidad de Zulia
Maracaibo, Venezuela

Abstract. A precise gravimetric network has been installed in the Venezuelan Andes to study eventual gravity changes due to vertical tectonic movements. The design and the measurements of the network are described and the accuracy is estimated.

In the center of the region a local gravity network has been reobserved three times. The detected variations are discussed.

In order to obtain a genuine statement as far as possible about the significance of observed gravity changes, requirements for the procedure of monitoring precise gravity networks are pointed out.

## 1. Introduction

The tectonic plate boundary between the Caribbean and the southern part of the American plate crosses Northwestern Venezuela following the course of the Venezuelan Andes (fig. 1). Horizontal movements in this area have been detected by geological methods along the Boconó Fault (Schubert and Sifontes, 1970), vertical movements being supposed in connection with the Andes' uplift.

As the total length of the considered zone is about 600 km with elevations from 100 m to 4000 m, it would hardly be possible to control vertical movements by use of levelling methods in a short time interval. Gravimetric observations, however, are capable of detecting height changes at reobserved points because of the dependence of gravity on elevation. This method has been used in several extended tectonic active regions (e.g. Torge and Drewes, 1977).

The problem is the conversion of gravity variations into height variations, which only can be done by knowledge of the actual vertical gravity gradient along the path of the moving masses. The determination of this value is somewhat problematic, but we can measure the local free air gradient as a rough estimation. In any case we obtain at least a qualitative estimation of vertical movements.

The advantage of the gravimetric survey is the easy and rapid performance and the almost invariable accuracy in respect to the distance between the points. Therefore the gravimetric method was chosen for monitoring the vertical component of the supposed movements in the Venezuelan Andes.

## 2. Regional gravimetric network

The precise gravimetric network consists of 58 stations and covers all the Venezuelan Andes in a length of about 600 km between the Colombian border (San Antonio) and the Caribbean Sea (Puerto Cabello). It has a width of about 100 km and is formed as it were three parallel profiles, one situated in the center of the Andes and one and one on each side of the mountains in the lowlands. One reference station is situated far off the network in Maracaibo (fig. 2).

The points are in general BMBM of the first order levelling net of Venezuela, that are concrete monuments with a 1m-foundation. Some stations are situated on foundation walls of churches. The net includes 11 stations of the National Gravity Network of Venezuela. The total range of gravity is 0.85 $cm \cdot s^{-2}$.

In two sites of the investigation area, in Maracaibo and near Mérida, earth tides are recorded with an equipment of the Institute of Theoretical Geodesy of the University of Hannover / Germany (LaCoste and Romberg model G gravimeter, chart-recorder) in order to obtain actual parameters for the earth tide reduction.

## 3. Gravimetric measurements

The first gravimetric survey of the net was carried out in February / March 1978 using two LaCoste and Romberg model G gravimeters (no. 401 and no. 405). Totally there were performed 280 observations, each of those being the mean of three independent readings at one station. The instruments were carried by a station wagon.

The sequence of point observations had been planned before by a net optimisation, the target function for a free net adjustment being

$$m_p < 10 \cdot 10^{-8}\ m \cdot s^{-2}. \qquad (1)$$

($m_p$ = a/v r.m.s.e. of point gravity values)

In conjunction with the present gravity network, another net around the Lake of Maracaibo was observed. This net was installed to detect gravity changes in the oilfields near the lake, which are due to the extraction of petroleum and related subsidences (Drewes, 1978). By means of the direct connection, a fusion of both nets is possible, covering in this manner a large region of the tectonic "Maracaibo-Block".

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

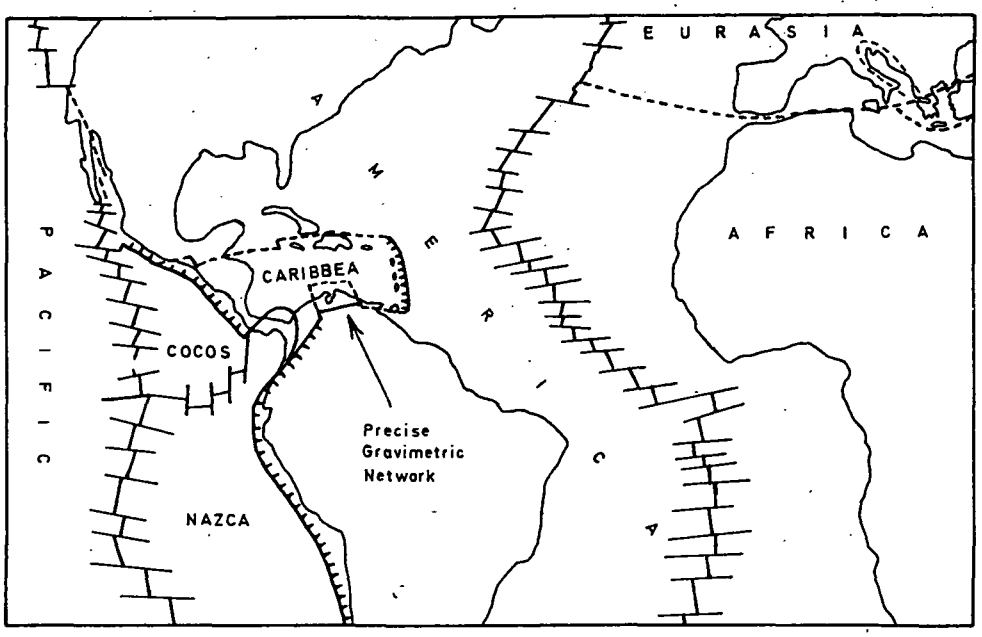



Fig. 1. Global Situation of the Precise Gravimetric Network

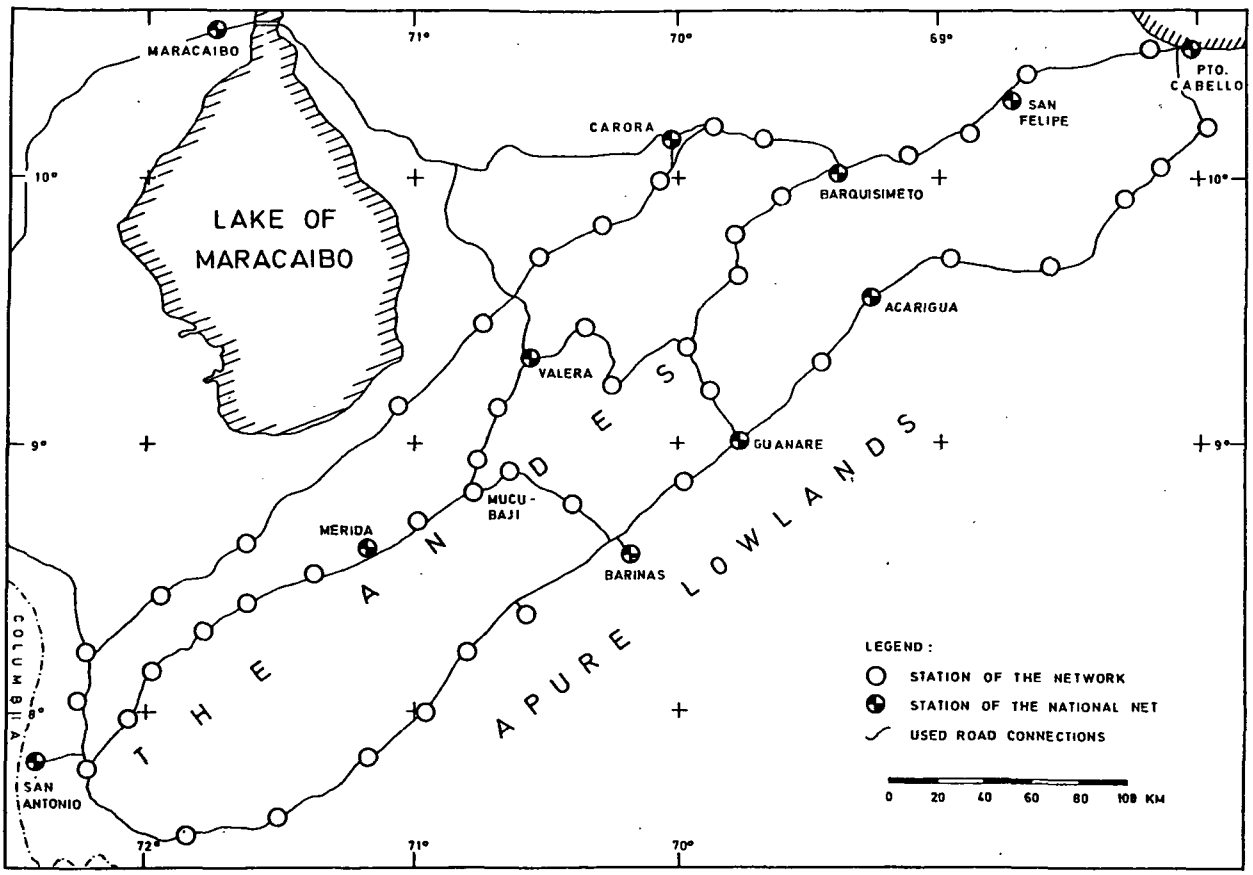



Fig.2. PRECISE GRAVIMETRIC NETWORK IN THE VENEZUELAN ANDES

## 4. Evaluation and results

After transformation of the readings into approximate mgal-scale (1 mgal = $10^{-5}$ $m \cdot s^{-2}$) by means of the manufacturer's tables, and the reduction because of earth tides by a modified Cartwright-Edden procedure with 505 partial tides (Wenzel, 1976), separate free net adjustments for each instrument have been calculated. The mathematical model for the adjustment of parameters is (Drewes, 1978, formula 14):

$$v_i = g_i - L_k - Y \cdot r_i - D \cdot t_i \mid p_i \quad (2)$$

($g_i$=gravity value of station i, $L_k$=gravity level of instrument at the begin of period k corresponding to reading 0.000, Y=scale factor, D = drift coefficient, $r_i$=transformed and reduced reading, $t_i$=time since begin of period k, $p_i$=weight of the observation).

As a result we obtain the r.m.s.e. of unit weight for both gravimeters:

$$m_o(401) = \pm 13 \cdot 10^{-8} m \cdot s^{-2}$$
$$m_o(405) = \pm 13 \cdot 10^{-8} m \cdot s^{-2}.$$

As there is no difference between the precision of the two instruments, all observations were introduced with an unit weight p=1 to a common adjustment, which also was calculated as a free net. Thereby the instrument no. 401 was fixed in its scale (Y=1). As a resulting r.m.s.e. of unit weight (now including also discrepancies between the gravimeters) we obtain

$$m_o = \pm 15 \cdot 10^{-8} m \cdot s^{-2}.$$

The average root mean square error of the point gravity values is

$$m_p = \pm 9 \cdot 10^{-8} m \cdot s^{-2},$$

fulfilling the condition (1) of optimisation. The results for the g-values of the adjustment are given in table 1.

The connection of the free network with a superior net, e.g. the National Gravity Network of Venezuela (RGNV) - in order to get an absolute orientation and scale for later comparisons - meets two principal problems:

1) The RGNV was observed in 1970 and adjusted within the Latin American Gravity Network (LAGN). In a readjustment, however, there were found gross errors of observations included in the former computation. The old and the new adjustment differ strongly.
2) As the time interval between the measurement of the RGNV and the present Andes' Network is 8 years, we cannot suppose constant gravity values. So we may not connect the present net to the gravity values of the RGNV 1970.

The main resulting crux is the determination of the actual scale factor of the instruments. The uncertainty of a roughly determined scale factor from the RGNV is

$$m_Y = \pm 6 .. 7 \cdot 10^{-5}.$$

Related to the total range of the Andes' gravity network this produces an uncertainty of $50 .. 60 \cdot 10^{-8} m \cdot s^{-2}$. This exceeds by far the internal precision of the net. The r.m.s.e. of the connected point values are also given in table 1.

## 5. Local network Mucubají

In the center of the Andes at Mucubají a small gravity network was installed to study local variations. The net is situated on both sides of the Boconó Fault, which in this region is defined by geologists within ±100 m uncertainty of position, and it is identical with a geodetic network of horizontal control (Schubert and Henneberg, 1975). Totally there are eight stations, all concrete monument with an 1 m x 1 m x 1 m foundation.

The mean topographic height of the net is about 3500 m, the mean gravity value 9.77360 $m \cdot s^{-2}$. The total gravity range is $18 \cdot 10^{-5}$ $m \cdot s^{-2}$. Because of this small difference in gravity the above mentioned problem of scale determination does not affect this network.

The first gravimetric observation took place in September 1977, the second in January 1978 and the third in August 1978. The same principle of readings and evaluation was used as described in sec. 3 and 4. The results of the free net adjustment are given in table 2, the relative gravity variations corresponding to the different periods are shown in fig. 3.

In addition to the measurements of the net, the vertical gradient of gravity has been observed in one reference station. The result is

$$\frac{dg}{dh} = 0.385 \cdot 10^{-5} \ s^{-2}.$$

At one side of the fault (north) we find no significant changes of gravity. The observed variations are always within the r.m.s.e. of determination. In the southern part, however, gravity changes exceed the threefold r.m.s.e. in several points. But there is no systematic in it. Therefore one should be careful interpreting those variations as tectonic movements. A great deal of local influences (ground-water etc.) should be considered.

To filter all the local effects from regional variations, an analysis of time series can be used (e.g. least squares prediction filtering). For this reason, however, a greater set of data in different epochs is necessary. The repetitions of gravity observations should therefore be

TABLE 1. Point Values and Errors of the Andes Gravity Network

| Station Name | Gravity Value (mgal) | r.m.s.e. Free Net | r.m.s.e. National Net | Station Name | Gravity Value (mgal) | r.m.s.e. Free Net | r.m.s.e. National Net |
|---|---|---|---|---|---|---|---|
| Chinita | 175.053 | 0.006 | 0.010 | Chachopo | -439.728 | 0.009 | 0.030 |
| Pto.Cab. | 229.299 | 0.008 | 0.012 | Mucubaji | -621.586 | 0.008 | 0.039 |
| Moron | 220.079 | 0.009 | 0.013 | Las Pie. | -237.251 | 0.008 | 0.020 |
| La Pica | 196.232 | 0.009 | 0.013 | Mucuruba | -417.334 | 0.008 | 0.028 |
| S.Felipe | 133.279 | 0.007 | 0.009 | Mérida | -248.612 | 0.010 | 0.018 |
| Chivacoa | 98.002 | 0.009 | 0.011 | Lagunil. | -162.673 | 0.009 | 0.016 |
| Yaritag. | 74.304 | 0.009 | 0.011 | S.Cruz | -69.474 | 0.009 | 0.013 |
| Barquis. | 39.610 | 0.007 | 0.008 | LaPlaya | -184.742 | 0.009 | 0.017 |
| S.Pablo | 19.210 | 0.009 | 0.011 | LaGrita | -223.115 | 0.010 | 0.018 |
| Pte.Tor. | 86.170 | 0.009 | 0.011 | M. Aura | -347.861 | 0.010 | 0.024 |
| Carora | 85.950 | 0.007 | 0.009 | S.Crist. | -130.516 | 0.007 | 0.014 |
| Sicarig. | 63.917 | 0.009 | 0.011 | S.Anton. | -57.097 | 0.009 | 0.011 |
| El Empe. | 28.146 | 0.009 | 0.011 | Valencia | 62.797 | 0.009 | 0.011 |
| Valerita | 69.827 | 0.008 | 0.010 | Carabobo | 65.583 | 0.009 | 0.011 |
| Mendoza | 32.231 | 0.008 | 0.010 | Tinaqui. | 90.555 | 0.009 | 0.011 |
| Caja Se. | 9.918 | 0.010 | 0.012 | S.Carlos | 121.069 | 0.009 | 0.011 |
| ElVigia | -3.727 | 0.010 | 0.012 | S.Rafael | 120.028 | 0.009 | 0.011 |
| Cano Am. | 4.028 | 0.010 | 0.012 | Acarigua | 92.644 | 0.007 | 0.009 |
| La Fria | 10.244 | 0.009 | 0.012 | Ospino | 88.145 | 0.009 | 0.011 |
| Colon | -92.116 | 0.007 | 0.013 | Guanare | 79.762 | 0.005 | 0.008 |
| Quibor | 12.282 | 0.009 | 0.012 | Boconoi. | 67.531 | 0.009 | 0.011 |
| El Tocu. | 26.332 | 0.009 | 0.011 | Barinas | 66.329 | 0.006 | 0.009 |
| Guarico | -70.098 | 0.009 | 0.014 | Barinit. | 5.714 | 0.008 | 0.011 |
| Biscucuy | 28.235 | 0.007 | 0.010 | Bolivia | 44.105 | 0.009 | 0.011 |
| Cimarro. | 56.069 | 0.009 | 0.011 | Socopo | 15.743 | 0.009 | 0.011 |
| Bocono | -118.134 | 0.007 | 0.014 | Capitan. | 31.803 | 0.009 | 0.011 |
| Sta.Ana | -192.994 | 0.009 | 0.018 | S.Barba. | 35.110 | 0.007 | 0.010 |
| Valera | -20.503 | 0.007 | 0.010 | Abejales | 24.822 | 0.009 | 0.011 |
| La Puer. | -264.184 | 0.008 | 0.021 | ElPinal | -30.645 | 0.009 | 0.012 |

Gravity Value (IGSN'71) of the Reference Station (Chinita) : 978160.12 mgal
Scale Factor (IGSN'71) of the Reference Gravimeter (401) : 1.0005±0.0001

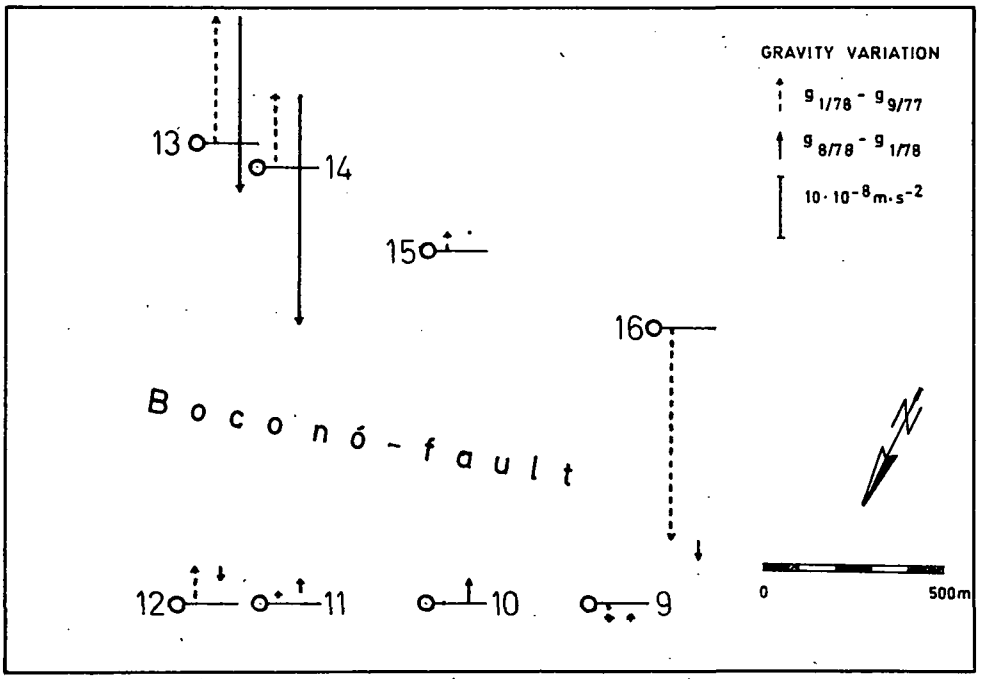



Fig. 3. Gravity Variations in the Gravimetric Network Mucubají

done in a short time interval.

## 6. Conclusions

From the analysis of the presented network we learn two principal requisites for the gravimetric control in geodynamics:

a) The internal precision of a network is often superposed by a greater uncertainty due to the insufficient determination of the scale factor. So, if we repeat the observations with another gravimeter, or if the scale factor of one gravimeter changes with time, we cannot compare the different epochs. The scale factor should therefore be determined externally with the same (or less) uncertainty as the determination of point values:

$$m_Y \leq \frac{m_P}{r_g} \quad . \qquad (3)$$

($m_Y$ = r.m.s.e. of scale factor, $m_P$ = a/v r.m.s.e. of point values in a free net adjustment, $r_g$ = gravity range of net)

b) Observed gravity variations are often superposed by local effects. To eliminate these disturbances we need short period repetitions. Precise gravity networks for geodynamics should therefore be reobserved at least as soon as variations are observable, i.e. the gravity changes increase to the order of the uncertainty of determination:

$$\Delta T \leq \frac{m_P}{\Delta g} \quad . \qquad (4)$$

($\Delta T$ = time interval of observation, $\Delta g$ = supposed or actual gravity variation)

Acknowledgement. The presented project is part of the joint investigation of the "Escuela de Ingeniería Geodésica, Universidad del Zulia", Maracaibo/Venezuela and the "Dirección de Cartografía Nacional de Venezuela", Caracas, in cooperation with the "Institut für Theoretische Geodäsie der Universität Hannover", Germany. The computations were performed at the "Centro de Cálculo Electrónico de la Universidad del Zulia". The author wants to thank all concerned persons.

TABLE 2. Local Gravity Network Mucubaji

| Station Name | Gravity ($\mu$gal, free net) Sept. 77 | Jan. 78 | Aug. 78 |
|---|---|---|---|
| Mucu 9 | -2676 ± 10 | -2679 ± 4 | -2678 ± 7 |
| Mucu 10 | -1226 ± 10 | -1226 ± 4 | -1222 ± 6 |
| Mucu 11 | 2671 ± 9 | 2673 ± 3 | 2675 ± 5 |
| Mucu 12 | 7474 ± 11 | 7480 ± 4 | 7478 ± 7 |
| Mucu 13 | -3192 ± 11 | -3171 ± 4 | -3200 ± 7 |
| Mucu 14 | 178 ± 14 | 190 ± 4 | 152 ± 7 |
| Mucu 15 | 1787 ± 14 | 1790 ± 4 | 1790 ± 7 |
| Mucu 16 | -10370 ± 9 | -10405 ± 4 | -10408 ± 7 |

## References

Drewes, H., Regional Subsidence of the Lake of Maracaibo as Determined by Repeated Gravimetric Measurements, Paper pres. to the 8th meeting of the Intern. Gravity Comm., Paris 1978.

Drewes, H., Zur Ausgleichung von Gravimeternetzen, Z. f. Vermessungswesen, 103, 485-496, 1978.

Schubert, C., and H. G. Henneberg, Geological and Geodetic Investigations on the Movements along the Boconó Fault, Venezuelan Andes, Tectonophysics, 29, 199-207, 1975.

Schubert, C., and R. S. Sifontes, Boconó Fault, Venezuelan Andes: Evidence of Postglacial Movement, Science, 170, 66-69, 1970.

Torge, W., and H. Drewes, Gravity Variation with Time in Northern Iceland 1965-1975, J. Geophysics, 43, 771-790, 1977.

Wenzel, H. G., Zur Genauigkeit von gravimetrischen Erdgezeitenbeobachtungen, Wiss. Arb. Lehrst. Geod., Phot. u. Kart. Techn. Univ. Hannover, Nr. 67, 1976.

Page Intentionally Left Blank

# The Measurement of Long Period and Secular Deformation with Deep Borehole Tiltmeters

Gerry H. Cabaniss
Air Force Geophysics Laboratory
Hanscom Air Force Base, Massachusetts 01731

Abstract. Biaxial borehole tiltmeters have been or soon will be deployed in deep boreholes by groups in W. Germany, England, the United States, and Canada. Their purpose is to measure tidal loading and premonitory earthquake phenomena as well as long-period and secular deformations. Two clusters of instruments were emplaced in fractured bedrock in eastern Massachusetts in 1970 and 1975. The intrasite agreement at tidal periods was about two percent, but there is no agreement at longer periods. A strong temperature-induced annual component ranging from 3 to 15 urads was present on instruments installed at depths of 15-20m; it was not apparent on those at 100-120m. One instrument, in continuous operation for three years at 100m, showed a net drift of 0.3 urads down to the SW, with a maximum departure of 2.0 urads from the trend. Pore pressure variations, material corrosion and creep, and local movements are apparently the limiting factors to long-term measurements.

## Deep Borehole Tiltmeter Experiments

Borehole instrumentation for the measurement of tilt (as well as strain and stress) has proved attractive for geophysical measurements because of the isolation from surface noise, such as that produced by meteorological and cultural sources, as well as the logistical advantage provided by the option of drilling at sites of interest as opposed to dependence upon mines and caves. Of even greater importance is the fact that borehole measurements should, in principle, be virtually immune to cavity effects [Harrison, 1976], which dominate the results from short-base tilt measurements made in niches carved in the walls of mines [Melchior, 1978]. Additionally, it is comparatively simple to deploy clusters of instruments to mitigate local perturbations.

Disadvantages include short baselength, cost, the lack of standardized instrumentation, and installation problems. Measurements from instruments installed at shallow depth (about 3m) in soils [Mortensen & Johnston, 1975], Lewcowicz & McConnell, 1977] have proved to be heavily contaminated by temperature and rainfall effects [Wood and King, 1977]. To avoid these the instruments must be placed at substantially greater depth. There we are confronted with groundwater problems, pore pressure changes, casing corrosion, etc. The short baselength means that very small vertical displacements and local tilting, which is not representative of the region under study, can dominate the measurements. Material creep, stress corrosion, stress release, and mineral weathering of the order of several atomic diameters, can produce significant spurious tilts. A number of long baseline instruments, which in essence measure the variations of height between two reservoirs and are free from effects of local tilt, have been developed (e.g., Beavan & Bilham [1977]). Definitive results at non-tidal periods have not yet been published.

There have been three types of deep borehole tiltmeters deployed: The Askania mechanical pendulum [Rosenbach & Jakoby, 1969], the Arthur D. Little diamagnetic suspension [Simon et al, 1968] and the servoed bubble flat [Hansen, 1968]. Two groups in West Germany, the Clausthal Technical University [Herbst, 1976] and Geophysical Institute at Kiel [J. Zschau, personal communication, 1978] have deployed a number of the Askania instruments in profiles across various geological structures. In England, the Institute for Coastal Oceanography and Tides has evaluated the Askania in a vault and is planning to deploy a Hughes sensor in a deep borehole [Baker, personal communication, 1977]. In North America, the Air Force Geophysics Laboratory (AFGL) operates two clusters of instruments in E. Massachusetts; details are presented in the next section. The University of Colorado plans to install about ten instruments in Colorado, Wyoming, and Montana [J. C. Harrison, personal communication, 1977]. The sensors are simple mechanical pendulums developed by Larry Burris of Instech, Inc.

## The AFGL Tiltmeter Arrays

AFGL operates two deep borehole tiltmeter clusters (small arrays), as well as several shallow instruments, at two sites in eastern Massachusetts. The Bedford site is located on the north side of Hanscom (Air) Field in the town of Bedford, about 14 km WNW of Boston. Another 17 km to the WSW is the Maynard array, located on the Sudbury Annex of the Army Natick Laboratory in the town of Maynard.

### Bedford Site

The Bedford array has been described in detail by Simon [1971] and Cabaniss [1974], so a brief summary will suffice. Three holes 20 cm in diameter and 100 m apart were drilled about 18 m into a foliated granitic gneiss beneath 0-1m of overburden. The holes were cased to a depth of about 6 m; so water in the fracture system was free to enter. Each instrument was emplaced at the bottom of a hole, resting either on its own 5 cm flat base or on a 10 cm blunt spike screwed into the base. The tiltmeter was wedged into position with heavy bronze weights. Alignment was made to a known direction using a set of rods rigidly coupled to the top of the instrument but which could be removed afterwards. Several analyses for the $M_2$ tidal component showed that the three instruments agreed among themselves to within 2% in amplitude and 2° in azimuth and phase, although the discrepancy between the observed and calculated tides

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

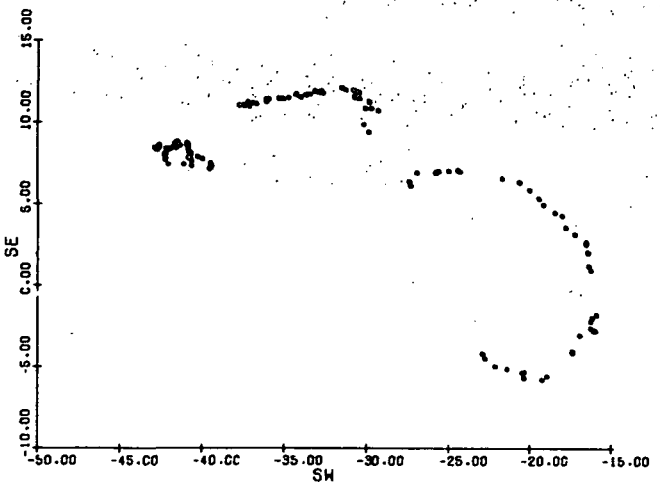



Fig. 1. Long-term tilt from Bedford Inst. 1. 1971-72 (left), 1972-73 (middle) and 1973-74 (right). Each set begins at the left in October, proceeds clockwise, and terminates in June. No connection between years is implied. Tilt is positive down to the SE and SW.

was considerably larger [McConnell, 1971; McConnell et al, 1973].

At periods greater than that of the diurnal tides, the coherence among the data from the three instruments approached zero. Although an annual term was apparent on each, its orientation and amplitude showed no apparent relationship to neighboring instruments. As an example of the annual behavior spot readings from an instrument oriented SE-SW (positive down), covering about three-quarters of three successive years are shown in Fig. 1. The 1971-72 and 1973-74 data show annual ellipses oriented in an E-W direction which can be explained by unequal temperature effects in the two orientations in response to the 0.1° C annual temperature variation at this depth. The data were interrupted in July of each year because of lightening damage; the instrument was removed and re-emplaced after repair by October. Thus any connection between the sets would be highly speculative. The differences from year to year were ascribed to slight changes in installation details [Cabaniss, 1974]. "Secular" drift has shown even greater disagreement and far exceeds the 4-15 nanoradians (down to the north) per year, as estimated from shoreline, geodetic re-leveling, and tide gage data [Cabaniss, 1974]. Those discrepancies are undoubtedly related to the problems outlined above.

## Maynard Site

The Maynard tiltmeters were installed in cased holes (16 cm diameter) with a horizontal separation of about 100 m drilled 100 m into granitic gneiss beneath 20 m of overburden. The instruments were the Hughes servoed bubbles packaged by Earth Sciences Research, Inc. They were emplaced 5-15 m above the hole bottoms, each resting on a holelock which forces wedges against the casing wall with a heavy spring [McConnell, 1978]. They were aligned by sighting on incandescent lamps with a transit device. Preliminary results from tidal analyses showed inter-instrument and monthly variations reaching 10% in amplitude [McConnell, 1978]. These have been tentatively ascribed to local variations in scale factor caused by bubble flat topography [McConnell, personal communication, 1978; Cabaniss, in preparation].

The long-term tilt variations are presented as a series of spot readings spanning a period of three years (Figure 2 and Figure 3). The two instrument components have been computationally rotated to North/East coordinates, and relevels have been removed, except those that occurred during the first eight months for Inst. 1 and two months for Inst. 3. Tides have not been removed and reach an amplitude of 0.25 urads (p-p) in the East component and 0.07 urads in the North. As at Bedford, there is little agreement among the data from the three instruments. No. 1 drifted back and forth, often in a well-defined direction; No. 2, after a stabilization period of about a year, moved in a very linear fashion until Day 740, when both instrument components drifted rapidly about the same amount but in opposite directions for about 45 days, after which the previous rate resumed. The direction of the rapid excursion was associated with that of one of the holelock wedges; so it has been surmised that casing corrosion might have been the cause of the event. Inst. 3 displayed remarkable stability over the same period. The first year was marked by decaying exponentials on both channels but the maximum rate was on the East component. The second year was characterized by roughly equal motion on both channels at the beginning but finished with accelerated movement on the North component. The last 18 months of the record are shown vectorially in Figure 4. There has been little motion during that period except for a counter-clockwise elliptical movement, equally partitioned between both instrument axes, which may have an annual periodicity. The net drift over three years on Inst. 3 was 0.3 urads, with a maximum excursion of 2.0 urads from the trend.

## Discussion

Our results show that the annual components generally decrease in amplitude with depth but are highly dependent on the installation epoch. Herbst [1976], however, reported annual ellipse amplitude (semi-major axes) of 0.9 and 0.1 urads at depths of 15 and 30 m, respectively, which can be explained by thermoelastic effects caused by the coupling of the annual temperature wave with the topography.

The long-term records from the Bedford and Maynard arrays show drifts that vary widely in both rate and direction. If it is assumed that a low-drift record is "best", Maynard #3 shows a mean rate of 0.1 urads per year to the SW, compared to the tide gage and geodetic re-leveling estimates of about 0.01 urads per year down to the north. The disturbing result so far from

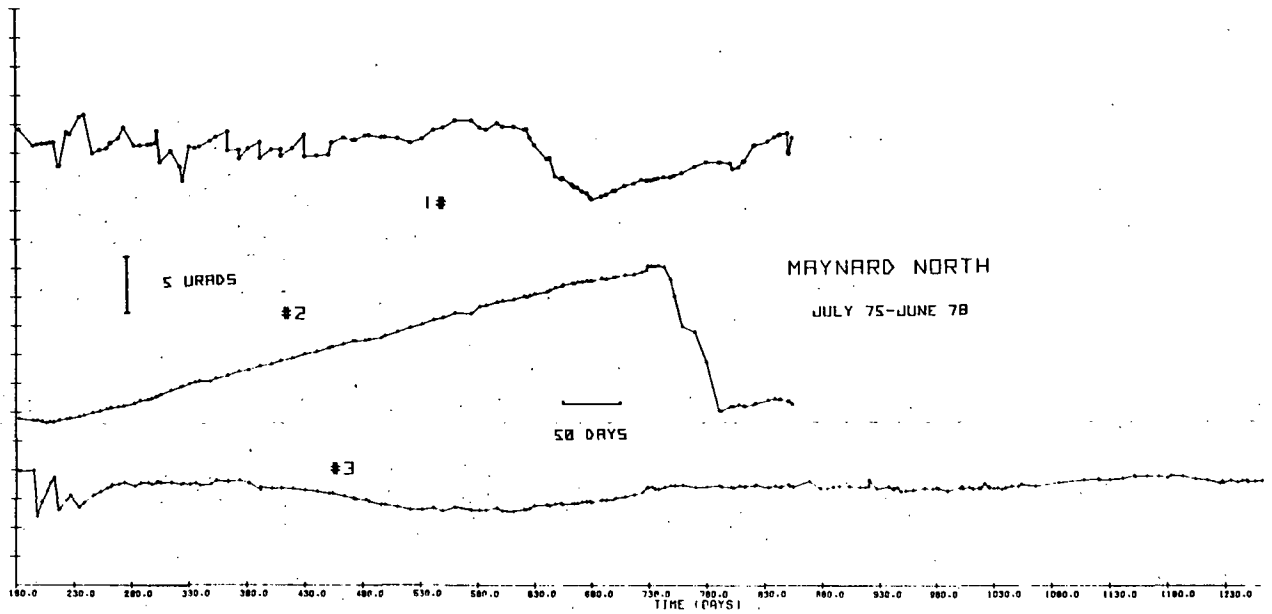



Fig. 2. North component of the long-term tilt from the Maynard array from readings taken at intervals of about 10 days. Data have been rotated from the azimuths of the instrument axes (#1, 294° - 24°; #2, 35° - 125°; #3, 2° -92°). Tilt is positive up to the north and east. The time axis is in days since 0 Jan 75.

these experiments is the wide disagreement among instruments spaced less than 100 m apart. Herbst [1976], for example, reported annual drift rates of 1.8 and 0.75 urads for instruments at depths of 15 and 30 m, respectively, spaced several meters apart. The drift directions varied by 15°. The question is whether this type of measurement is limited by the length scale of the phenomena, by the installation techniques, or by the inherent stability of the instruments. Parenthetically, no "events", including tilt steps associated with magnitude 7.7 teleseisms, have been detected on all instruments within a cluster.

The AFGL experiments suggest that improvements might be made by emplacing the instruments in sections of stainless steel casing at depths greater than 30 m at the bottoms of sealed holes. Installation in comparatively fracture-free material and at close spacing might also prove efficacious.

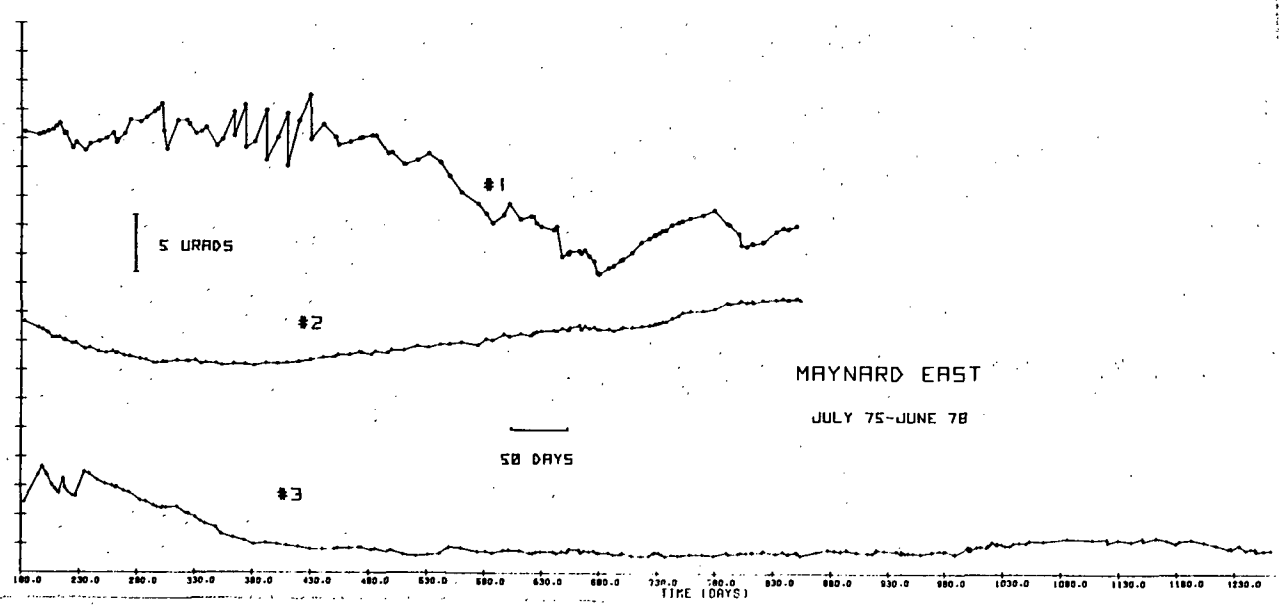



Fig. 3. East component of the long-term tilt from the Maynard Array. (See Fig. 2 legend,)

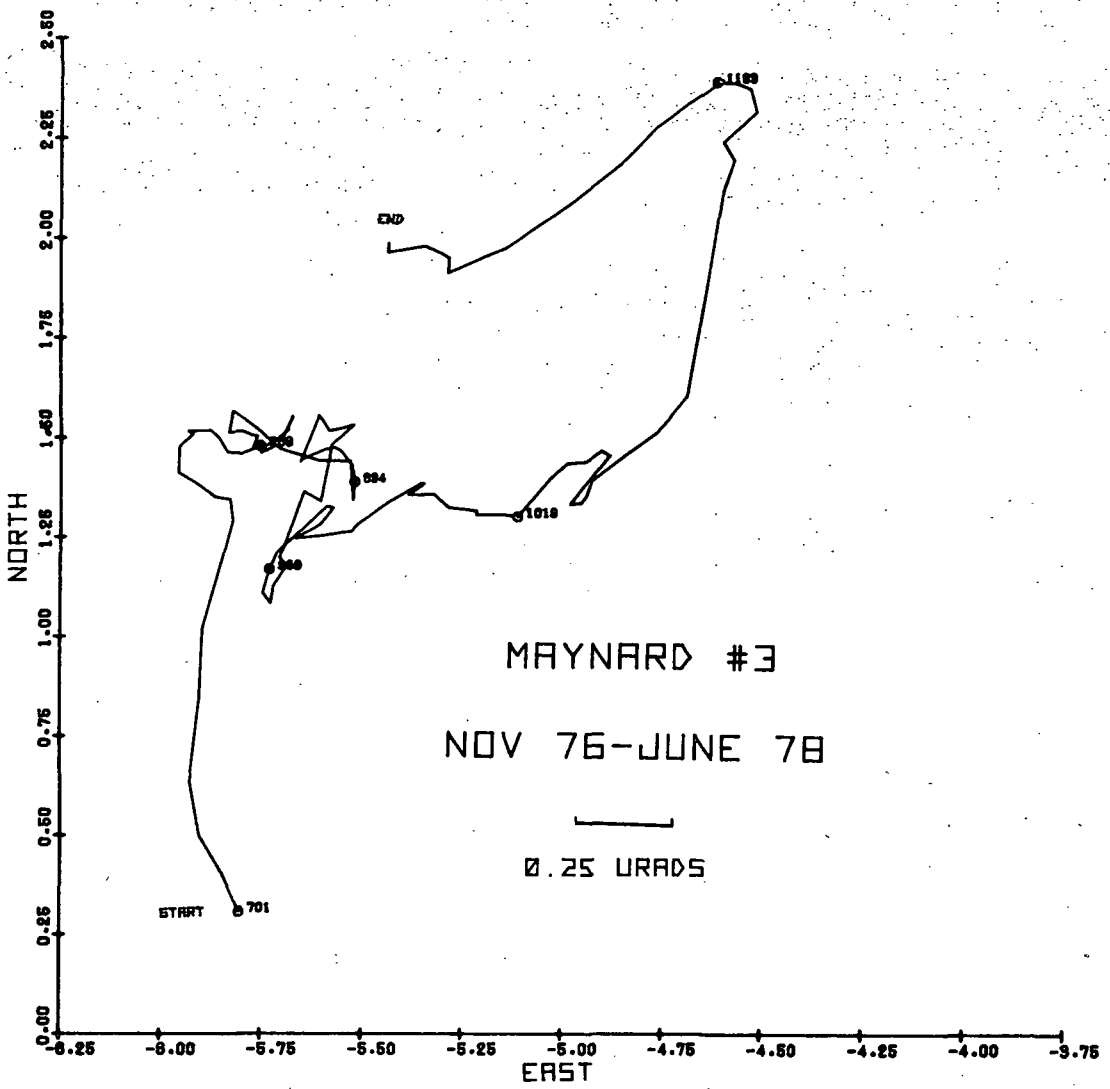



Fig. 4. Vector plot of a portion of the long-term data from Maynard Inst. 3. Symbols and numbers correspond to times indicated on Fig. 2 and 3. The last point is at Day 1250.

References


Beavan, J., and R. Bilham, Thermally induced errors in fluid tube tiltmeters, J. Geophys. Res., 82, 5699, 1977.

Cabaniss, G.H., Crustal Tilt in Coastal New England - An Experimental Study, PhD Thesis, Boston Univ., 1974.

Harrison, J.C., Cavity and topographic effects in tilt and strain measurements, J. Geophys. Res., 81, 319, 1976.

Herbst, K., Interpretation of Tilt Measurements in the Period Range Above that of the Tides, PhD Thesis, Tech. Univ. of Clausthal, W. Germany, 1976.

Lewcowicz, J., and R.K. McConnell, Jr., Preliminary results from a shallow borehole tilt array, AFGL, TR-77-0168, 1977 (NTIS A047523).

McConnell, R.K., Jr., Experimental and theoretical studies of long-period tilt of the earth's crust: Part II - Theoretical, AFCRL, 71-0388 (II), 1971 (NTIS AD729917).

McConnell, R.K., Jr., Research on the nature of ground tilts in the period range $10^3$ to $10^7$ seconds, AFGL, TR-78-0105, 1978.

McConnell, R.K., Jr., P.C. Dunn, B.D. Peatie, Research on crustal tilts in the greater Boston area, AFCRL, TR-73-0681, 1973 (NTIS AD776058).

Melchior, P., The Tides of the Solid Earth, Pergamon Press, New York, 1978.

Mortensen, C.E., and M.J.S. Johnston, The nature of surface tilt along 85 km of the San Andreas fault - Preliminary results from a 14-instrument array, Pure Appl. Geophys., 113, 237, 1975.

Rosenbach, O., and H. Jakoby, First experiences with the Askania borehole tiltmeter (earth tide pendulum), in Problems of Recent Crustal Movements of the Earth, Moscow, 1969.

Simon, I., Experimental and theoretical studies of long-period tilt of the earth's crust: Part I-Experimental, AFCRL, 71-0388 (I), 1971 (NTIS AD 729916).

Simon, I., A. G. Emslie, P. F. Strong, and R. K. McConnell, Jr., Sensitive tiltmeter utilizing a diamagnetic suspension, Rev. Sci. Inst., 39, 1666, 1968.
Wood, M. D., and N. E. King, Relation between earthquakes, weather, and soil tilt, Science, 197, 154, 1977.

Page Intentionally Left Blank

# Applicability of APT Aided-Inertial System to Crustal Movement Monitoring

J. Arnold Soltz
The Charles Stark Draper Laboratory, Inc.
Cambridge, Massachusetts 02139

Abstract. The APT system, its stage of development, hardware, and operations are described. The algorithms required to perform the real-time functions of navigation and profiling are presented. The results of computer simulations demonstrate the feasibility of APT for its primary mission: topographic mapping with an accuracy of 15 cm in the vertical. Also discussed is the suitability of modifying APT for the purpose of making vertical crustal movement measurements accurate to 2 cm in the vertical, and at least marginal feasibility is indicated.

## Introduction

The Charles Stark Draper Laboratory, Inc., under sponsorship of the U.S. Department of the Interior, Geological Survey*, is developing an airborne-instrument system capable of providing precisely located, geophysical data. This equipment, called APTS (Aerial Profiling of Terrain System), is designed to be carried by a relatively small maneuverable aircraft. It incorporates a surveying instrument package, unaffected by aircraft motion, capable of continuously providing a high-accuracy, three-coordinate position datum. A self-contained terrain-measuring device for recording the vertical position of the terrain below the aircraft is also provided for applications such as stream-valley profiling. A vidicon camera is to be boresighted with the terrain-measuring device to aid in the data processing. Performance goals call for locating an unsurveyed ground control point with respect to three or more established ground control points to an accuracy of ±15 cm in the vertical coordinate and ±60 cm in the horizontal coordinate.

The work completed to date consists of engineering analysis, system configuration design, mathematical analysis, and computer simulation, selection and specification of all major hardware components, mechanical design, thermal design and test, electronic design, subsystem interface design and specification, selection and recommendation of flight computer, analysis of flight strategy, specification and design of key flight algorithms, and, in addition, a flight test of a breadboard laser profiler. Within the time and funding resources available, the program has been organized so that a balanced effort has been maintained in the various aspects of system and subsystem design to facilitate an orderly transition into the fabrication, laboratory test, and flight trial of a test bed system.

The APT system is designed to function in a small aircraft at altitudes of 1,000 meters or less. The system is to be used for applications such as:

(1) Producing topographic maps.
(2) Testing the reliability of older published maps.
(3) Fitting specified flood magnitudes into local stream-valley geometry.
(4) Classifying public lands for waterpower potential.
(5) Mapping the earth's gravity field.
(6) Definition of temporal change in groundpoint elevation or position:
   (a) Subsidence in and around producing oil fields, geothermal reservoirs, and heavily pumped areas of continuing water withdrawal.
   (b) Open-pit and strip-mine mineral extraction and land reclamation.
   (c) Volcano inflation, faulting, landslides, and beach and slope erosion.

## APT System Configuration

The airborne-instrument package (Figure 1) consists of a three-gimbaled inertial platform to define the position of the aircraft in three coordinates. A two-axis laser tracker is mounted on the same base as the inertial platform in order to update the long-term drift of the inertial platform. Three or more surveyed retroreflectors interspersed with several unlocated retroreflectors provide ground truth. The inertial platform and laser tracker provide the high-accuracy position datum. A laser profiler, to provide line scans of the terrain, is provided for those applications requiring terrain profiling information. However, the three-coordinate reference system may be used with other systems such as scanning lasers or radars, side-looking radar, aerial cameras, and radiometers. The laser altimeter or profiler is just one of a number of equipments that may be used in conjunction with the position and attitude measurement system. An airborne digital computer accepts data from the three sensors and performs the necessary computations and data processing for alignment of the IMU, and position or velocity updating calculations from range and angles. In addition, the computer outputs data to the magnetic-tape recorder and display unit.

The initial flight test configuration will differ somewhat from Figure 1. The purpose of the initial flights is to basically prove the concept, i.e., locate the position of unsurveyed retroreflectors with respect to three surveyed retroreflectors. The basic differences in this initial flight configuration compared to the configuration depicted in Figure 1 include the fact that the profiler will be hard-mounted (the angular errors from the profiler laser beam will be computationally corrected). Interfaces for leveling the profiler platform and reading its gimbal angles are provided, however, so that a stabilized profiler platform may be added in the future. Another difference in the initial flight configuration is that the autopilot/guidance function indicated in Figure 1 will not be implemented. Finally, the indicated Kalman filter update, combining tracker and IMU data, will not be performed in real time. Tracker and IMU data will

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

* Under contract 14-08-0001-14578

be recorded in flight, and filtering of the data will be performed post-flight. A simple reset of position and velocity will be performed in real time, however, so that the acquisition time of a new retroreflector is minimized.

A mechanical schematic (Figure 2) of the system shows the geometric relationship and functions of the gimbaled subassemblies. The IMU and tracker are mounted in a common housing. The profiler-vertical camera assembly is hard mounted to the aircraft, but referenced to the IMU pitch and roll isolation axes to account for motion of the aircraft.

The IMU consists of a stable member which houses three high-performance gyros and three accelerometers along with their associated electronics. Outside the stable member are three servo-driven isolation gimbals (azimuth, elevation, and roll, in that order) isolating the instruments from aircraft rotation. The support structure surrounding the isolation gimbals is designed so that the tracker assembly mounts directly to its base, thereby providing a physical tie between the structure of the inertial reference system and the tracker pointing axes.

The IMU has a unique thermal-control system which isolates the inertial instruments from the aircraft environment. The gyro mounting surfaces are held at approximately 46.1 ± 0.05°C and the accelerometer mounting surfaces are held at approximately 43.3 ± 0.05°C. This is accomplished by mounting the inertial instruments directly to the stable member air-cooled heat exchanger and employing thermal shims and individual controllers where necessary.

An X, Y, Z coordinate frame fixed to the inner, or stable, member of the IMU provides the attitude datum. This frame is called the indicated frame, platform frame, or "p" frame. It is initially established during a ground-alignment process, consisting of leveling and gyrocompassing. During leveling, the $X_p - Y_p$ plane is established normal to the local gravity vector. During gyrocompassing, the $Y_p$ axis is established normal to the earth's rotation vector. The process of calibration and alignment of an inertial navigation system is a complex subject in its own right. A preliminary exposition of this process for the APT system is given in Reference 1.

In the IMU, the sensitive axes of the accelerometers are fixed in a known orientation with respect to the three coordinate axes of the "p" frame. The outputs of the accelerometers are integrated twice in the computer to obtain position data for navigation, and converted to latitude and longitude. The "p" frame is maintained when in flight by torquing signals proportional to the computed change in geodetic latitude and longitude. Thus, the platform is driven to maintain its attitude (except for errors) with respect to the normal gravity field of the earth and not the actual gravity field of the earth. It is this property of an IMU which allows the measuring of changes in the deflection of the vertical, a point discussed in detail in Reference 2. In an error-free IMU, one would have a completely self-contained instrument capable of giving position, velocity, and direction of the normal gravity vector in real time.

The inertial system is a low-noise datum at high frequencies, capable of providing base-motion isolation from maneuvers or wind gusts. The inertial sensors (gyros and accelerometers) suffer from long-term drift, however, and cause errors in the low-frequency portion of the spectrum. To minimize these long-term drifts, the inertial components selected are high-performance instruments developed for an Air Force missile application.

Even with the most accurate inertial systems, it is necessary to obtain updating from ground truth in order to reach the specified precision. To provide high-accuracy position updates, a laser tracker is employed. The tracker makes a vector measurement to a surveyed retroreflector; range and two angles are measured. A minimum of three retroreflectors, not in a straight line, are required in a 3- by 30-km area. The high-frequency information from the inertial system is optimally combined with the low-frequency information from the tracker to obtain an estimate much better than either. The statistical techniques employed for the optimum filtering of the data are given in the following section. It can be shown that the important system errors, including position errors and platform misalignments, are independently observable.

The laser tracker and laser profiler employ pulsed-laser rangefinding techniques. Distance is determined by measuring the round-trip time of a transmitted laser pulse. Pulsed gallium arsenide lasers are used as the transmitters for both devices. Constant fraction techniques are employed in the receiver threshold detectors to minimize timing errors due to received pulse-amplitude variations. In the profiler, pulse-selection techniques are employed to maximize the number of valid returns from terrain. The tracker contains a beamsplitter to separate the ranging and tracking functions. A four-quadrant detector is used to provide the error signals to the gimbal servos of the tracker. A precision time-interval counter is multiplexed between the profiler and tracker to provide the range measurements.

A control and display unit enables the operator to sequence the system through its various operating modes and to monitor and observe the status of the system. The system may also be sequenced automatically by the computer. Various system characteristics, including temperatures of inertial and other key components, critical voltages, and operational modes are observed through the use of the control panel, as well as by computer monitoring. Steering information for the pilot to fly the desired courses for obtaining the required data is provided.

Several military and commercial-type computers were investigated. For the APT system, the selected computer is the NORDEN PDP-11/70M, which has the speed, memory size, input/output capability, instruction repertoire, architecutre, and software support necessary for the mission.

## Equations for Navigation and Topographic Profiling

The first implementation of APT will be to prove the concept by performing only one mission, topographic profiling, from the list given in the Introduction. The same APT hardware, in the same aircraft, with the same on-board real-time software can perform the rest of the tasks by using different ground-based post-processing algorithms and

possibly different operational procedures. In this section the algorithms needed for the solution of the primary APT functions, navigation and profiling, are presented.

When the aircraft reaches the survey area, two passes are made over three previously surveyed retroreflectors, deliberately chosen or placed not in a straight line. Range and angle data are obtained by the laser tracker in these passes by sightings on each retroreflector in sequence. These data are used to update and calibrate the position and orientation of the inertial system in preparation for the profiling part of the mission. The positions of any unsurveyed retroreflectors previously placed to satisfy any recognized need for added ground-truth references are determined relative to the surveyed retroreflectors during the initial flight passes. Thereafter, each retroreflector serves as a known reference point and may be tracked to obtain updating data as needed during the profiling phase of the mission.

Figure 3 shows the flow of information through the onboard digital computer of the APT system. The input quantities shown are the profiler range; the three accelerometer velocity readouts and three gimbal angles from the IMU subsystem; and the two tracker gimbal angles and range to the retroreflector from the tracker subsystem. The rates at which each of these quantities is sampled are indicated.

The profiler subsystem data are sampled, time-tagged, and subjected to reasonableness checks to eliminate signal dropouts and returns from the tops of foliage. The slope and intercept of the best-fit straight line are computed from 40 ms of those data which pass the reasonableness checks. The effect of the curve fit is to act as a low-pass filter of the profiler data.

The IMU navigation subsystem samples accelerometers for 80 ms (16 vector samples). The vectors are averaged and time-tagged at the mid-point of the data-acquisition interval. The effect of the averaging is to provide a simple low-pass filter on the input signals, in order to suppress further any aircraft vibration not taken out by the shock-mounting of the IMU assembly. Next, the acceleration vector is rotated by a fixed matrix to compute the north, east, and down components of acceleration, $\vec{A}$, since the three orthogonally disposed accelerometers are mounted on the stable platform so as to receive equal components of gravity, $\vec{g}$, as indicated in Figure 2.

Equation (1) is the IMU navigation algorithm.

$$\ddot{\vec{R}}_{IMU} = \vec{A} + \vec{g} - 2\vec{W} \times \dot{\vec{R}}_{IMU} - \dot{\vec{W}} \times \vec{R}_{IMU} \qquad (1)$$

where $\vec{R}_{IMU}$ is the position of the aircraft. This equation is solved, in the "p" frame above for $\vec{R}_{IMU}$ and $\dot{\vec{R}}_{IMU}$, by using trapezoidal rule integration with a time step of 80 ms. The $\vec{W}$ vector is the angular velocity of the stable platform with respect to inertial space, given by Eq. (2)

$$\vec{W} = \begin{bmatrix} (\Omega + \dot{\lambda}) \cos \lambda \\ \dot{\phi} \\ -(\Omega + \dot{\lambda}) \sin \lambda \end{bmatrix} \qquad (2)$$

where $\Omega$ = sidereal earth rate.

$\dot{\lambda}$ = the longitude rate.

$\phi, \dot{\phi}$ = the latitude and latitude rate.

The $\vec{g}$ vector in Eq. (1) is a function of position $\vec{g} = \vec{g}(\vec{R}_{IMU})$ and is given by the output of a gravity model; for example, the WGS72 ellipsoid. An error in position thus generates an error in $\vec{g}$ which, through integration of Eq. (1), generates more error in position. This feedback causes the characteristic 84-minute oscillations of error in the horizontal channels, and the characteristic doubling of the error in the vertical channel every 395 seconds. It is important to note, in this context, that one must use a gravity model to separate the earth's gravitational field from aircraft accelerations. The inertial instruments obey Einstein's fundamental principle of the equivalence of gravity and acceleration in a reference coordinate system; they sense only the linear combination of the two. In what follows, we shall see that the laser tracker measurements, being referenced to the outside world, provide the mechanism for separating gravitational acceleration from nongravitational acceleration.

The data from the tracker subsystem are sampled at high rates, low-pass filtered, and the tracker. navigation equation, Eq. (3) is solved for $\vec{R}^p_{lt}$.

$$\vec{R}^b_{lt} = C^b_s \cdot C_x(\theta) \cdot C_y(\rho) \cdot \begin{bmatrix} 0 \\ 0 \\ -r \end{bmatrix} + \vec{d}^b$$

$$\vec{R}^p_{lt} = C_z(-\alpha) \cdot C_y(-\beta) \cdot C_x(-\gamma) \cdot \vec{R}^b_{lt} \qquad (3)$$

In Eq. (3) the tracker range r is rotated about the y-axis for the inner tracker gimbal angle $\rho$; then rotated about the x-axis for the outer tracker gimbal angle $\theta$; then rotated by a fixed orthogonal matrix $C^b_s$ for the alignment of the tracker assembly respect to the aircraft; then translated for the fixed distance, $\vec{d}^b$, from tracker assembly to IMU; and then rotated about the x, y, and z axes for the outer, middle and inner IMU gimbal angles, respectively. The result $R^p_{lt}$ is the distance from the retroreflector to the aircraft in platform (or "p" frame) coordinates.

The IMU provides a slowly drifting (but smooth) navigation estimate at all times. The tracker provides a noisy but nondrifiting navigation estimate for short intervals of time (30 to 60 seconds) during which a retroreflector is within a 60° cone of visibility determined by the tracker viewing hole in the floor of the aircraft. As mentioned previously, the data from each of the navigators will be recorded for post-processing on the first flight trials. The algorithms in this section

which are used to process these data, however, are the same whether done in real time or after the fact. The rest of this paper proceeds with the description of the real-time implementation of data processing.

The next step is to mix the data from the two subsystems.

$$\vec{D} = (\vec{R}_{IMU} - \vec{R}_{rr})^1 - \vec{R}^p_{lt} \qquad (4)$$

By taking the difference between the positions indicated by the two navigators, as given by Eq. (4), the aircraft motion is subtracted out and there remains only a vector which contains a linear combination of the errors in the two subsystems. The vectors $\vec{D}$ are then accumulated over a 2-second interval (25 vectors), and low-pass filtered again. The 2-second sample then represents input to a Kalman filter, or optimal estimation algorithm as described in Reference 3, Chapters 3 and 4.

A Kalman filter is specified very briefly by listing the elements of the state vector $\vec{x}$, giving the system dynamics matrix F, the system measurement matrix H, and the measurement covariance matrix R, and the spectral density matrix Q.

The state vector $\vec{x}$ consists of 20 elements. These are:

- $\vec{e}$ = a linear combination of IMU position error and retroreflector survey error (3 elements)
- $\dot{\vec{e}}$ = IMU velocity error (3 elements)
- $\vec{\psi}$ = vector of stable-platform misalignment angles (3 elements)
- $\vec{B}_a$ = a linear combination of accelerometer bias, vertical accelerometer scale-factor error, and gravity model bias error (3 elements)
- $\vec{\epsilon}$ = a linear combination of the constant component of gyro drift and the along-track gravity gradient errors (3 elements)
- $\vec{\eta}$ = a linear combination of various body-axis-fixed constant angle errors (3 elements)
- $\eta_4$ = the bias error in the inner tracker gimbal (1 element)
- $\delta r$ = the bias error in the tracker range measurement (1 element)

Other elements which may be added to the state vector include:

(1) The position coordinates of unsurveyed retroreflectors.
(2) Unknown values of gravity disturbance vectors and gradients at a point where a retroreflector is located.
(3) The gravity gradients and even rate of change of gradients averaged over the flight trajectory.

The present state of development of the software includes the above 20 elements, which suffice to account for all the APT instrumentation errors.

This state vector $\vec{x}$ satisfies a linear differential equation $\dot{\vec{x}} = F\vec{x} + \vec{\nu}$, where the system dynamics matrix F is a 20 × 20 matrix which is partitioned as follows:

$$F = \begin{bmatrix} F_1 & F_2 \\ 0 & 0 \end{bmatrix} \begin{matrix} 9 \text{ rows} \\ 11 \text{ rows} \end{matrix}$$

9 cols 11 cols

where . . .

$$F_1 = \begin{bmatrix} 0 & I & 0 \\ S_d - \dot{\underline{W}} - \underline{W}^2 & -2\underline{W} & \ddot{\underline{R}} - \underline{g} \\ 0 & 0 & -\underline{W} \end{bmatrix} \begin{matrix} 3 \text{ rows} \\ 3 \text{ rows} \\ 3 \text{ rows} \end{matrix}$$

3 cols 3 cols 3 cols

$$F_2 = \begin{bmatrix} 0 & 0 & 0 \\ I_3 & 0 & 0 \\ 0 & I_3 & 0 \end{bmatrix} \begin{matrix} 3 \text{ rows} \\ 3 \text{ rows} \\ 3 \text{ rows} \end{matrix}$$

3 cols 3 cols 5 cols

In order to build this F matrix, take $\vec{W}$ and $\dot{\vec{W}}$ from Eq. (2) and $\ddot{\vec{R}}$ from the accelerometers. The notation used is that an underscore denotes the 3 × 3 antisymmetric matrix associated with a given vector, and

$$S_d = \frac{g}{R_e} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & +2 \end{bmatrix}$$

where $R_e$ = radius of the earth
$g$ = nominal magnitude of gravity

In the differential equation $\dot{\vec{x}} = F\vec{x} + \vec{\nu}$, the (20 element) vector $\vec{\nu}$ is called the plant noise. Spectral-density plots of laboratory measurements taken from actual APTS instruments revealed that a suitable model of the statistics of $\vec{\nu}$ is zero mean, Gaussian noise with a covariance matrix, called the Spectral Density Matrix, Q.

$$\xi(\nu\nu^T) = Q = \begin{bmatrix} \sigma_p^2 & 0 & 0 & 0 \\ 0 & \sigma_v^2 & 0 & 0 \\ 0 & 0 & \sigma_a^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} 3 \text{ rows} \\ 3 \text{ rows} \\ 3 \text{ rows} \\ 11 \text{ rows} \end{matrix}$$

3 cols 3 cols 3 cols 11 cols

It is possible to show (see Reference 4 page 64) that the vector $\vec{D}$ in Eq. (4) is, to first order, linear in the elements of the state:

$$\vec{D} = \vec{e} - \underline{R^p_{lt}}\,\vec{\psi} + \underline{R^p_{lt}}\left[C_z(-\alpha)\cdot C_y(-\beta)\cdot C_x(-\gamma)\right](n_1,n_2,n_3)^t + \underline{R^p_{lt}}\left[C_z(-\alpha)\cdot C_y(-\beta)\cdot C_x(-\gamma)\right]\left[C^b_s\cdot C_x(\theta)\cdot\vec{j}\right]n_4 - \underline{\text{unit}\,(\vec{R}^p_{lt})}\,\delta r + \vec{n}_{lt}$$

i.e., $\vec{D} = \vec{H}x + \vec{n}_{lt}$

. . .where the time-varying $\vec{H}$ matrix has 3 rows and 20 columns. The vector $\vec{n}_{lt}$ is the measurement noise in the laser tracker. Since five gimbal angle measurements and one range measurement are used to compute the tracker position (see Eq. (3)); since an angle error affects the tracker position perpendicular to the line from aircraft to retro; since a range error is along that line; and under the assumption that the angle errors are uniformly distributed; then, it is concluded that

$$R \equiv \xi(\vec{n}_{lt}\cdot\vec{n}_{lt}) = \sigma^2_{range}(\vec{u}\vec{u}^t) + \frac{5}{2}r^2\sigma^2_{angle}(I - \vec{u}\vec{u}^t)$$

where $\sigma^2_{range}$ = the variance in range measurements.

$\sigma^2_{angle}$ = the variance in angle measurements.

$\vec{u}$ = unit $(\vec{R}_{lt})$

$r$ = $|\vec{R}_{lt}|$

This completes the brief description of the Kalman filter.

Referrring again to Figure 3, the Kalman filter provides estimates $\hat{\vec{x}}$ of the state and estimates P of the state covariance every 2 seconds. These estimates may be combined with the IMU data to produce a best estimate $\vec{R}_{a/c}$ of aircraft position. If the aircraft has acquired data from at least three retroreflectors not in a straight line, and the quantities $\hat{\vec{e}}$, $\dot{\hat{\vec{e}}}$, $\hat{\vec{\psi}}$ (the first nine elements of the state) are significantly different from zero (using a 3σ test based on the square root of the diagonal of the covariance matrix), then Eq. (5), is used; otherwise, the best estimate is considered to be the uncorrected IMU data.

$$\vec{R}_{a/c} = \vec{R}_{IMU} - \hat{\vec{e}} - \dot{\hat{\vec{e}}}(t - t_o) - \ddot{\hat{\vec{e}}}(t - t_o)^2/2$$

$$\vec{V}_{a/c} = \vec{V}_{IMU} - \dot{\hat{\vec{e}}} - \ddot{\hat{\vec{e}}}(t - t_o)$$

$$\vec{A}_{a/c} = \vec{A}_{IMU} - \ddot{\hat{\vec{e}}} \qquad (5)$$

where $\hat{\vec{e}}$ and $\dot{\hat{\vec{e}}}$ and $\ddot{\hat{\vec{e}}}$ = state elements obtained from $\dot{\hat{\vec{x}}} = F\hat{\vec{x}}$

$t_o$ = the time of the most recent Kalman filter update, so that $(t - t_o) < 2$ seconds

The corrected aircraft navigation data from Eq. (5) are used to build the state dynamics matrix F in the the Kalman filter, and $\vec{W}$ and $\dot{\vec{W}}$ or Eq. (2). $\vec{W}$ is the gyro torque rates used to locally level the stable platform.

Finally, the position of the aircraft $\vec{R}_{a/c}$ from Eq. (5) can be combined with the profiler data to obtain elevation and position at the nadir. Since the profiler will be located several feet from the IMU, it will be measuring the elevation of a point which is displaced horizontally from $\vec{R}_{a/c}$ and which (in general) will thus have an elevation different from the IMU, depending upon aircraft pitch and roll angles. The geometry of the aircraft gives (in locally level coordinates)

$$\vec{R}_{alt} = \vec{R}_{a/c} - (I - \underline{\psi})\cdot C_z(-\alpha)\cdot C_y(-\beta)\cdot C_x(-\gamma)\cdot\vec{p}^b$$

where $\vec{R}_{alt}$ = position of the profiler.

$\vec{R}_{a/c}$ = position of the IMU.

$C_z(-\alpha), C_y(-\beta), C_x(-\gamma)$ are the same rotation matrices, dependent on IMU gimbal angles α, β and γ that are defined in Eq. (3).

$\vec{p}^b$ = the distance of the profiler from the IMU in aircraft body-axis coordinates.

This leads to the actual survey datum:

$$(x_{alt}, y_{alt}, z_{alt} + \hat{h})^t = \vec{R}_{gr} \qquad (6)$$

where $x_{alt}, y_{alt}, z_{alt}$ are the north, east, and down components of $\vec{R}_{alt}$.

## Computer Simulations

In order to verify the APT system concepts, there is an on-going effort to exhaustively model all the dynamics and statistics of a typical APT mission on a digital computer. The strategy is: (1) to write an off-line non-real-time version of the navigation and profiling estimation equations given in the previous section, called an estimation model; (2) to write a more exhaustive, second-order simulation of the APT system error dynamics, called a truth model; and (3) to compare the outputs of the estimation model (estimates of the state $\hat{\vec{x}}$) with the inputs $\vec{x}$ to the truth model, the residuals giving a realistic estimate of the eventual APT performance. Figures 4 and 5 illustrate the use of the APT truth model and estimation model. At their current stage of development, both computer programs can give a realistic estimate of the amount of error contri-

buted by the APT instruments themselves. The programming for component of error contributed by residual gravity model errors is not yet complete, however, and so will be accounted for later in this section by another technique.

In a typical computer run, we simulated a 20-minute section of a nominal APT mission with the truth model. The simulation flew the aircraft over the trajectory indicated in Figure 6, wherein the three circles represent the region of visibility around three perfectly surveyed retroreflectors. The retros are about 10 km apart, the aircraft has a mean speed of 53 1/3 m/s, a mean altitude of 0.91 km, and the viewing hole for the tracker defines a 60° cone of visibility. In this simulation the laser accuracy, after low-pass filtering for 1 second, was 3 cm in range, 22 $\widehat{sec}$ (about 0.107 mrad) in pointing. The IMU performance was based on parameters and spectral models derived from classified military data taken on instruments identical to the APT accelerometers and gyros. The truth model program simulated the aircraft dynamics, IMU error drift, laser measurement errors, and retroreflector geometry, but not the laser profiler or the gravity disturbance vector.

The 20-state Kalman filter described previously was used as an estimation model on the data generated by the truth model. For each state the residual error was computed as a function of time, and plotted. The Kalman filter fluctuates wildly for the first part of the flight, until data from three different retroreflectors (not in a straight line) have been processed, and the filter achieves geometric sufficiency. Accordingly, Figures 7 and 8 corresponding to the vertical position and acceleration error, respectively are plotted only for the last 700 seconds of the 20-minute flight. On each graph, three curves are plotted. A middle curve representing the state-element residual estimation error is enclosed by an envelope corresponding to $\pm 1.6\ \hat{\sigma}$ (90% confidence), where $\hat{\sigma}$ is computed by the Kalman filter as the square root of the appropriate diagonal element of the estimated covariance matrix $\hat{P}$.

In Figure 7, we plot the vertical position error in cm. Since the error curve fills, but remains inside, its 90% envelope most of the time, we may say that the Kalman filter is a self-consistent model when driven by this truth model data. Qualitatively, the curve and envelope show characteristic periods of compression and expansion corresponding to the aircraft being inside or outside a retroreflector cone of visibility. When outside a cone of visibility, the envelope expands as time to the 3/2 power. 90% of the errors are less than 8 cm. This figure represents the fundamental surveying error from the APT instruments; however, if surveyed retroreflectors were placed closer together, then we might expect even better performance.

In Figure 8, the vertical acceleration error is plotted in mgals. It can be seen that the filter is still converging at the end of 20 minutes, however, 90% of the data are already within a 0.1-mgal band. What we have shown here is the ability to determine the mean gravity anomaly over the region (actually averaged over the trajectory of the aircraft) to within 0.1 mgal absolute accuracy. More sophisticated modelling will allow determination of mean gravity gradients over the region, or mean gravity gradients between any pair of retroreflectors or values of gravity disturbance vectors above each retro at the height of the aircraft, all to within an absolute accuracy of 0.1 mgal. Placing more retroreflectors will extend the region of the gravity survey beyond the simulated 30 km; placing them closer together will increase the resolution of a gravity survey better than the simulated 1 point in 30 km.

The other 18 state elements were plotted in a similar format, yielding information on navigation, horizontal retroreflector survey error, deflection of the vertical, gyro drift, APT system mechanical misalignments, IMU platform attitude error, and laser tracker errors. These results will be published in January 1979.

So far only the APT instrumentation errors have been simulated. A back-of-the-envelope computation was carried out using gravity anomaly data from the Denver area supplied by the Department of the Interior, Geological Survey, to demonstrate the feasibility of correcting for the effects of gravity anomalies. The study consisted of calculating position errors attributable to residual gravity errors along a 30-km flight path by integrating twice. It was assumed that the aircraft would fly along the path at a constant velocity of 53.3 m/s, resulting in a 600-second flight. The initial position and velocity errors were assumed to be zero and these errors were again set to zero each time it was necessary to simulate the effect of three laser tracker updates. In addition, it was assumed that measurements of gravity anomalies were known a priori, or could be estimated, at three retroreflectors located along the flight path at 13-km intervals. Based on these three measurements, a quadratic model for the gravity data was chosen and the coefficients of the model were computed. The difference between the known gravity anomalies and the gravity values obtained from the quadratic model were considered to be the residual gravity errors and thus the cause of position errors.

The result of this study was that the standard deviation associated with the vertical position errors was 6.1 cm, the maximum error was 21 cm, 90% of the vertical position errors were bounded by 12 cm.

Although this study represents only a first cut at the gravity problem, it suggests that a combination of modeling, estimation, and a priori measurements should be sufficient to account for gravity anomaly errors in the APT system.

Since the vertical error from instrumentation is within a 90% confidence interval of 8 cm, since the unsimulated gravity disturbances will be about 12 cm, and since these are uncorrelated, then

$$\text{RSS error bound} = \sqrt{8^2 + 12^2} < 15 \text{ cm}$$

## Crustal Movement Monitoring

In the preceding section, the capability of the APT system to perform real-time navigation and topographic profiling to an accuracy of 15 cm in the vertical, relative to three control points, was demonstrated in a survey area about 30 × 3 km. For those applications where the survey region has been chosen to straddle a fault line and where one plate

has moved relative to another by more than 15 cm, the APT system could be used, unmodified, by the simple expedient of placing all three control points on the same plate; differencing the profiler elevations taken on an APT survey at time $t_0$ from an APT survey taken at time $t_1$; and averaging the differences on each plate separately to remove bias errors and the effect of not repeating the aircraft path exactly. If $\overline{Z_J(t_k) - Z_J(t_0)}$ is the average of APT vertical displacement measurements of plate J at time $t_k$, then the vertical movement of plate 1 with respect to plate 0 during the time from $t_0$ to $t_1$ is:

$$\overline{[Z_1(t_1) - Z_1(t_0)]} - \overline{[Z_0(t_1) - Z_0(t_0)]}$$

There are some problems with this method: resolution, extent of survey area, and accuracy. By modifying the APT system slightly we can solve these problems.

First of all, the resolution problems can be solved by eliminating the profiler. Profiling data is taken every 40 ms (i.e., 3 meters) then has to be averaged over the aircraft trajectory to give a resolution meaningful for crustal plate dimensions. Also, profiling data will not be exactly the same in successive surveys, since the aircraft can not repeat its flight path exactly; hence, the survey area has to be profiled densely so that the averages $\overline{Z}_J(t_k)$ are representative of the area. Lastly, the laser tracker can be used as a more effective surveying device than the profiler. This can be accomplished by inverting Eq. (3) and (4) and solving for $\vec{R}_{rr}$. Instead of using measured tracker range and pointing angles with a known retroreflector position to compute the aircraft position, the aircraft position is assumed known and the rotation and translations performed to compute the location of the retroreflector. The accuracy will be as good as the IMU subsystem navigation accuracy. The aircraft needs to repeat its trajectory from survey to survey only well enough to enter the cone of visibility of the retroreflectors. Retroreflectors can be placed atop fixed monuments in a grid at whatever spacing corresponds to meaningful resolution of crustal movement measurements. If $\hat{\vec{R}}_J(t_k)$ is the APT surveyed position of retroreflector J at time $t_k$, then the movement of retro J with respect to retro I from time $t_0$ to time $t_1$ is

$$[\hat{\vec{R}}_J(t_1) - \hat{\vec{R}}_I(t_1)] - [\hat{\vec{R}}_J(t_0) - \hat{\vec{R}}_I(t_0)] \qquad (7)$$

Notice that we compute vector displacement, not just the vertical component.

The problem of survey extent was related to survey accuracy. As mentioned in the previous section in reference to Figure 7, position errors in the IMU grow as time to the 3/2 power when there are no tracker measurements. One way to increase the size of the survey area is to fly faster so that more distance can be covered for a given error bound. In a jet aircraft, speeds of 266.7 m/s, or five times the speed assumed in the previous section, may be assumed. Such speeds imply a higher safe operating altitude, say 4.5 km.

The problem of accuracy was addressed by a computer simulation of the APT system in which an aircraft was flown at a height of 4.5 km, at a speed of 266.7 m/s, repeatedly around a closed, oval traverse about 120 km long and 30 km across. At each end of the oval was a cluster of three surveyed retroreflectors about 20 km apart in a right triangle configuration. The clusters were about 100 km apart, and the complete traverse could be flown in 20 minutes. The parameters for the laser tracker were the same as for the profiling mission. If the navigation error from this simulation is observed, we get a measure of the ability of APTS to survey in retroreflectors. In this simulation we did not include the effects of unmodelled gravity errors on the system, because these may be reduced to residuals by availability of a priori gravity surveys, reduced further by the ability of APTS to estimate the gravity field, and removed through the following survey operation. If we fly the same (±500 m) trajectory at the same speed, then the gravity-induced IMU errors will be the same from survey to survey and subtract out in Eq. (7). Of course, the local gravity field may undergo temporal change also; however, the effect, which is 3.3 microgals per cm of vertical displacement is distinctly second order.

The result of the simulation was that 90% of the errors were less than 18 cm. Such a result shows that the APT could be used for measuring only very large vertical crustal displacements unless further modifications are made. One improvement would be in more sophisticated modelling to use all of the information available. There is a post-processing technique, called back smoothing, (see Reference 3), which is a refinement to the Kalman filter algorithm offering the potential of significant improvement. A forward Kalman filter forms the best estimate of the state vector given initial conditions, the models, and all the measurements up to the time of the estimate. A backward Kalman filter forms the best estimate of the state given final conditions, the models, and all the measurements which occurred after the time of the estimate. An optimally weighted average of the forward and backward filter outputs then forms the optimally smoothed estimate of the state. Under the assumption of $t^{3/2}$ error growth, it is easy to show that the peak error the optimally smoothed estimate is 1/4 the peak error in the Kalman filter estimate. The back smoothing algorithm is very cumbersome, costly of computer and programming time, and yet offers the possibility of measuring crustal movement with the APTS to within about $18 \div 4 = 4.5$ cm.

We shall now discuss modifications to the APT system hardware and improvements to mathematical modelling which would push the APT performance to its ultimate limit. What has been shown, so far, is that the errors derived only from APT instrumentation are about 2.5 times higher for measuring crustal movements ($\geq$ 18 cm before backsmoothing) than for topographic profiling. This factor occurs because the geometry of the tracker measurements is so much less favorable at the higher altitude and sparser retroreflector spacing used to measure crustal movements. This geometry affects the system through the $t^{3/2}$ error growth between retrosightings and through the degradation in tracker accuracy at higher altitude. The time between tracker position fixes could be shortened only by flying even faster, involving the choice

of a supersonic aircraft; however, it was felt that aircraft maintainance expenses for such a vehicle were not justified by the marginal gain in APT performance.

The degradation in tracker performance with increased altitude occurs for two reasons. First of all, the laser beam must pass through more of the atmosphere, and over a greater range of pressure, temperature, and humidity. Therefore the index of refraction will vary over a greater range along the ray-path, hence (through Snell's Law) there will be more atmospheric ray bending. Because of ray bending both the direction and distance measurements of the tracker are in error. If the APT system were modified so that the tracker were replaced by a two-color laser, then one could correct for the change in arc length due to raybending. The scheme is described in Reference 5. The pointing uncertainties can be reduced to 3 $\widehat{\text{sec}}$ by atmospheric modelling of $\delta\theta$ (raybending distortion) versus $\theta$ (angle of incidence) as described in Reference 4. The second effect is a purely geometric one. The distance error caused by tracker pointing errors is 1 mm per microradian for each km of altitude. The precision angle readouts, as envisioned for APTS, have $\sigma$ = 22 $\widehat{\text{sec}}$ effective pointing error. At 0.91 km altitude this corresponds to 10 cm, at 4.5 km to about 0.5 meter. Most of this error is repeatable and could be removed by careful laboratory calibration of a table of $\delta\theta$ versus $\theta$ for each precision angle resolver. If this were done the effective pointing error could be reduced by at least a factor of 4 to 5.5 $\widehat{\text{sec}}$.

In the computer simulation which estimated 18-cm accuracy for Eq. (7) (4.5 cm presumed after back smoothing) we assumed all tracker information available for the 100-km flight path (37.5 seconds flying time) between the two clusters of surveyed retros was used for survey and none for navigation. In fact, such navigation information is available, and could be used if more sophisticated modelling were performed; thus, the results of the computer simulation are unduly pessimistic. There are three major sources of navigation data not included in the simulation. First, each retro supplies a velocity measurement with which to update the drifting IMU. The 1600-Hz laser tracker pulse rate gives a complete time history of the aircraft with respect to the retro, and that retro is not moving for the $10^4$ seconds of the survey. In fact the measurement equation for a velocity update can be obtained by differentiating Eq. (3) and (4). If one places enough retroreflectors along the survey trajectory, then one simultaneously surveys-in more points for measuring crustal plate movements and supplies almost continuous velocity information to the system. The limit is reached when the retros are so close that the regions of visibility overlap, because it takes 2 or 3 seconds to move the tracker from locking on one retroreflector to the next. The effect of such a near-continuous velocity update on position accuracy should be quite dramatic, since, instead of a characteristic $t^{3/2}$ error growth caused by twice integrating random acceleration errors, one would expect $t^{1/2}$ error growth caused by a single integration of the nondrifting random residual velocity errors. In Reference 4, Chapter 3, Appendix B an estimate of less than 1 cm for t = 375 seconds is made assuming a perfect and uninterrupted velocity update.

The second major source of navigation information comes from the fact that the crustal motions of closely spaced points will be highly correlated. The correlations, if modelled, would allow the extraction of position information for navigation from redundantly placed retroreflectors.

The third source of information occurs because there is a bandwidth separation of the navigation error and APT survey error which has not been taken advantage of. So far, we have been observing navigation error as though this were the same as survey error. However, the navigation errors are pure functions of time, and the survey displacements which we are attempting to compute are pure functions of latitude and longitude. If one were to repeatedly cross over a retroreflector, then one would expect the IMU error to be statistically independent from crossing to crossing, so that the survey error could be a factor of $\sqrt{N}$ times better than the navigation error, where N is the number of times the retro has been crossed.

Unfortunately, updating the existing APT simulation program with the equations necessary to implement velocity updates and survey computations, does not fit the APT development schedule, priorities, timing, and budget. As mentioned in the previous section, the system will be programmed to perform the tasks of navigation and profiling. What can be done, however, is to use the existing simulation program to infer a best case bound on APT performance. There is already a worst case bound, 4.5 cm with back smoothing, gotten by assuming no tracker information available during an APT survey. The best case bound is gotten by using the existing computer simulation to supply position data from perfectly surveyed retros and observing navigation error. Admittedly the combined effect of velocity updates, survey correlations, bandwidth separation of survey and navigation errors, and back smoothing can not be as good as position updates, but it does set an APT performance limit.

Accordingly, another computer simulation was run, with the IMU model the same as APTS, with the laser tracker pointing accuracy improved to $\sigma$ = 5.5 $\widehat{\text{sec}}$. Sixteen perfectly surveyed retroreflectors about 16 km apart with a 60° cone of visibility lay along an oval flight trajectory as indicated in Figure 9. The aircraft was flown at 266.7 m/s at an altitude of 4.5 km. The vertical position errors for one traverse of the oval are plotted in Figure 10 in the same format used in Figure 7. In Figure 10, one can see by inspection that most of the errors lie between ±1.8 cm. This number was confirmed by plotting the points in Figure 10 in an integrated histogram format. In Figure 11, the abscissa is error bound in cm, the ordinate is the % of errors from the simulation less than or equal to the abscissa. From Figure 11 one can read off 1.8 cm as the 90% error bound.

From these computer simulations we have reached the conclusion that the APT system could survey vertical crustal movements much better than 4.5 cm relative to three control points, but cannot survey better than 1.8 cm, in a survey area of diameter 120 km (about 1.5 - 4 parts in $10^7$). The system also has the capacity to measure horizontal displacements to the same order of accuracy.

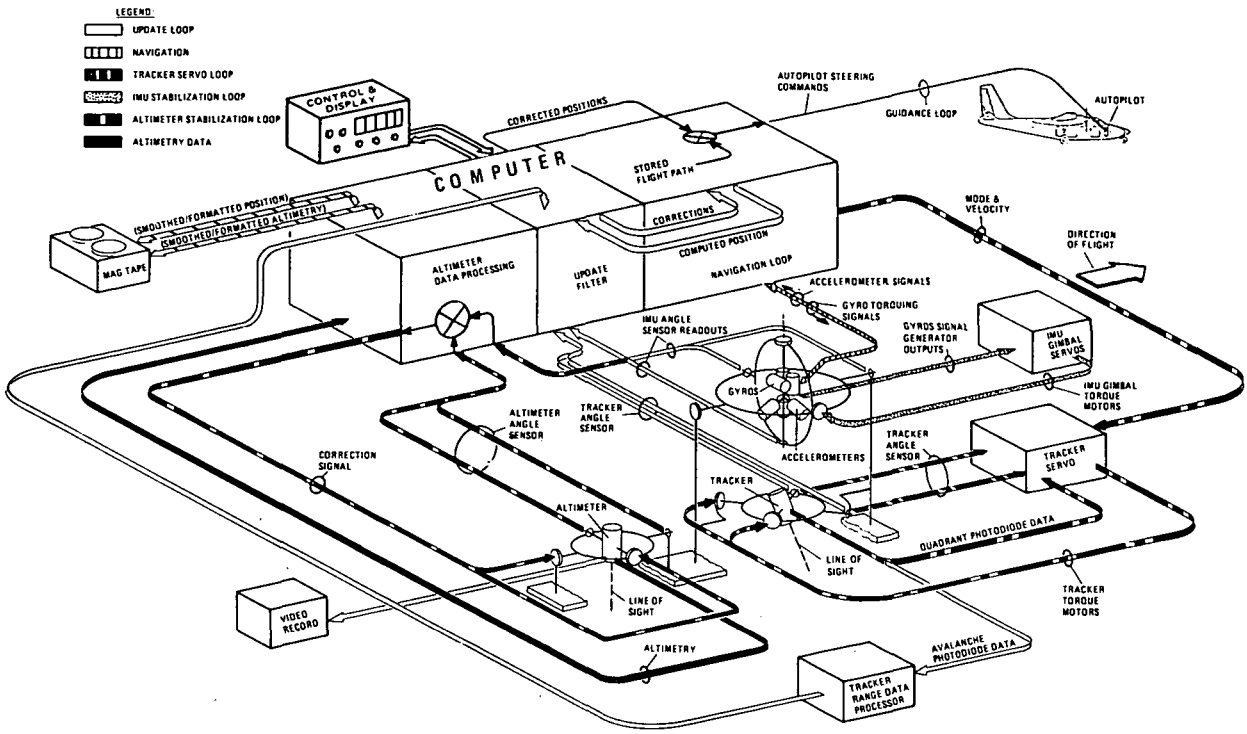



Figure 1. USGS/APTS functional diagram.

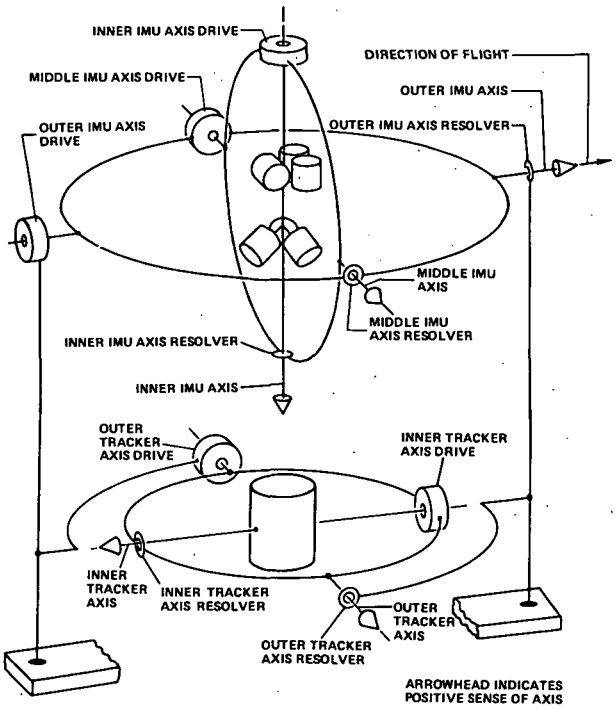



Figure 2. APTS Inertial Measurement Unit (IMU) and laser tracker schematic.

Acknowledgments. This study was made possible through the dedicated work of the CSDL-APTS technical staff. Thanks is especially due to Linda K. Lemos, John H. Barker, Glenn Mamon, and John W. Hursh who are also, to some degree, responsible for the information presented here.

Publication of this paper does not constitute approval by The Charles Stark Draper Laboratory, Inc., of the findings or conclusions contained herein. It is published for the exchange and stimulation of ideas.

References

1. Garofalo, F.J. 1978. APTS-IMU calibration and alignment. Cambridge, MA: The Charles Stark Draper Laboratory, Inc. report R-1193.
2. Schwartz, K.P. 1978. "Accuracy of vertical deflection determination by present-day inertial instrumentation." Presented at Ninth Geodesy/Solid-Earth and Ocean Physics Conference. 2-5 October 1978, at Columbus, OH.
3. Technical Staff, The Analytical Sciences Corp. 1974. Applied optimal estimation. Gelb, A., ed. Cambridge, MA: M.I.T. Press.
4. APTS-CSDL Staff. 1977. Aerial profiling of terrain system design phase, interim report. Cambridge, MA: The Charles Stark Draper Laboratory, Inc., report R-1070.
5. Earnshaw, K.B., and Hernandez, E.N. 1972. "Two-laser optical distance-measuring instrument that corrects for the atmospheric index of refraction." Applied Optics v.11, no.4: 749-54.

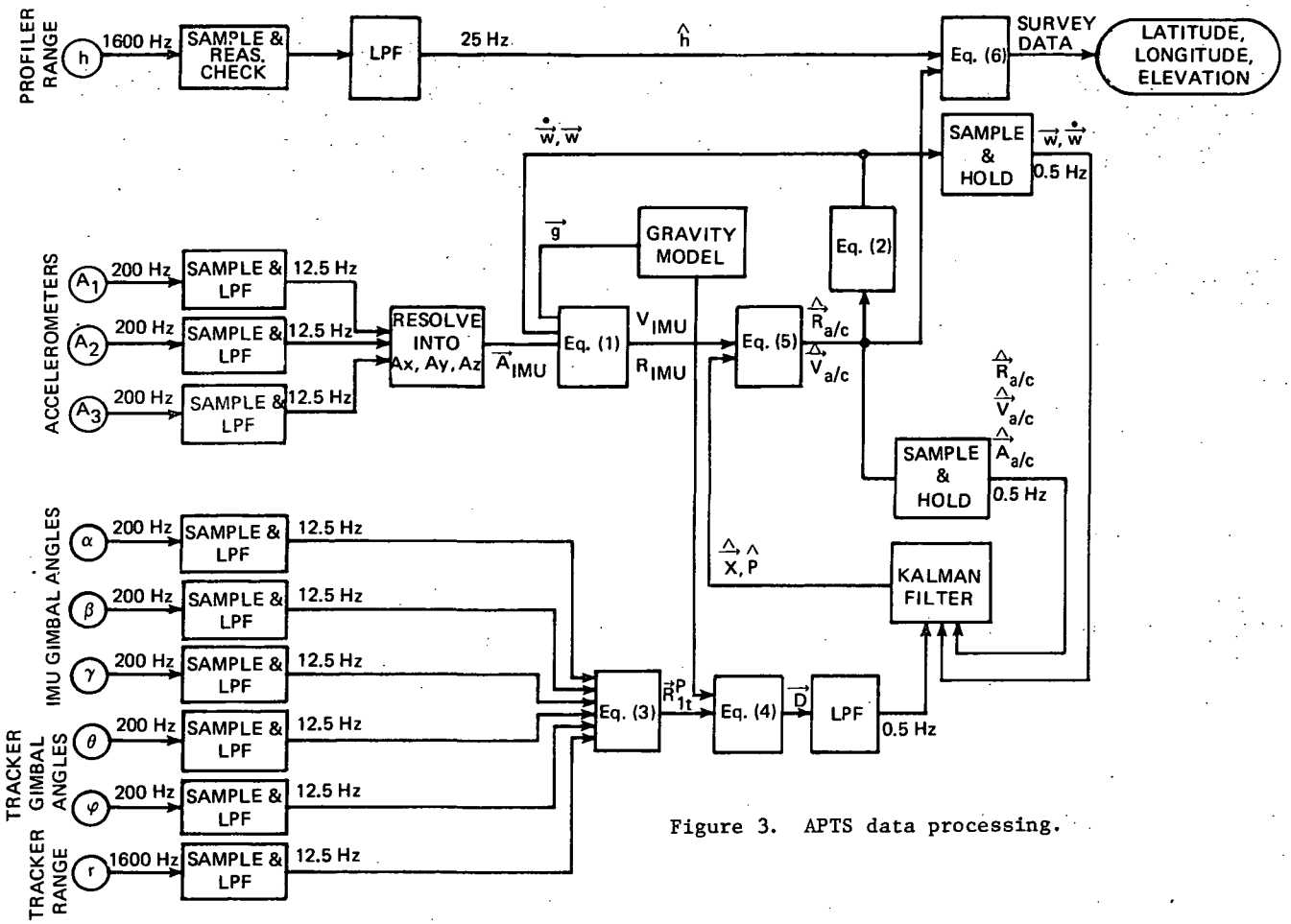



Figure 3. APTS data processing.

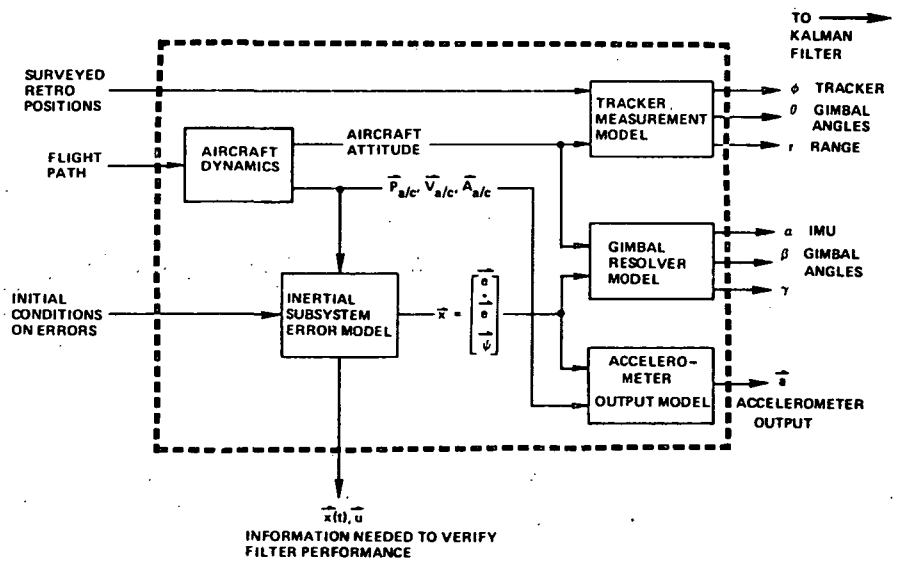



Figure 4. Truth model simulation.

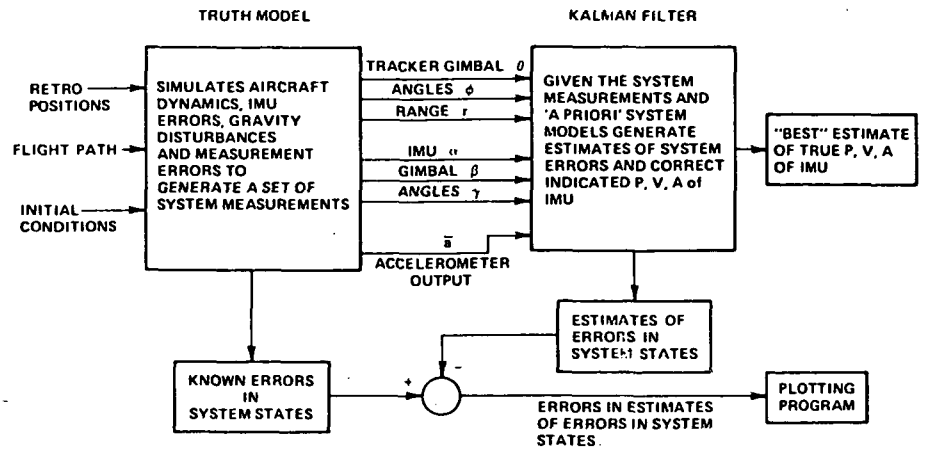



Figure 5. Interaction of truth model and Kalman filter.

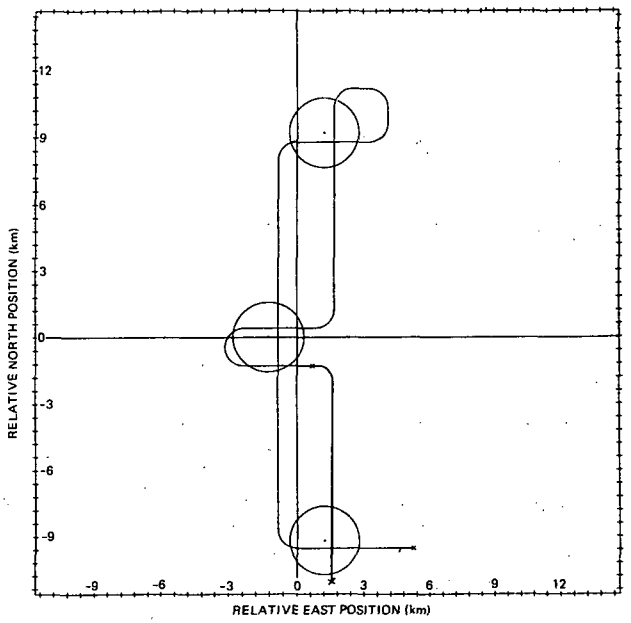



Figure 6. Flight trajectory and retroreflector geometry for computer simulation of APT navigation performance.

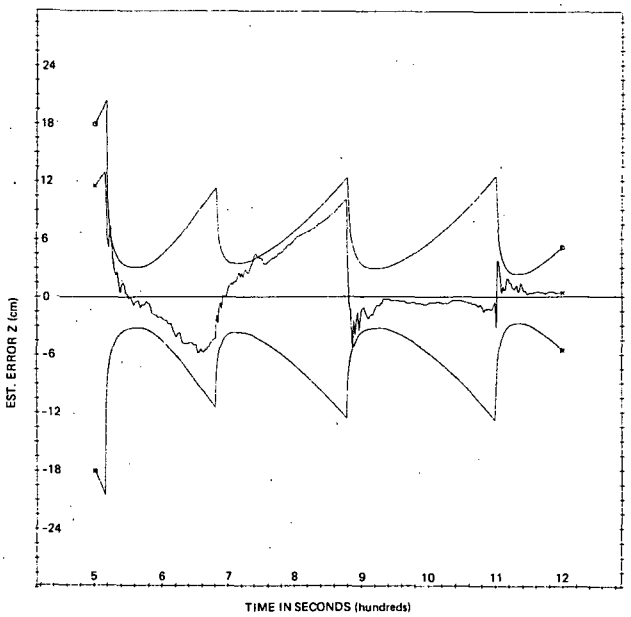



Figure 7. Vertical position error and covariance envelope from navigation simulation.


6. Broxmeyer, C. 1964. Inertial navigation systems. New York: McGraw-Hill.
7. Youmans, D.G. 1977. Flight testing of an airborne laser terrain profiler. Cambridge, MA: The Charles Stark Draper Laboratory, Inc. report R-1106.
8. Garofalo, F.J. and Lemos, L.K. 1978. Aerial profiling of terrain system truth model documentation. Cambridge, MA: The Charles Stark Draper Laboratory, Inc., report R-1182.
9. Soltz, J.A. 1978. APT in-flight navigation equations. Cambridge, MA: The Charles Stark Draper Laboratory, Inc., report R-1192.
10. Hursh, J.W., Mamon, G., and Soltz, J.A. 1977. "Aerial profiling of terrain." In Proceedings: 1st international symposium on inertial technology for surveying & geodesy. pp. 121-30. Ottawa, Canada: International Association of Geodesy, and the Canadian Institute of Surveying.

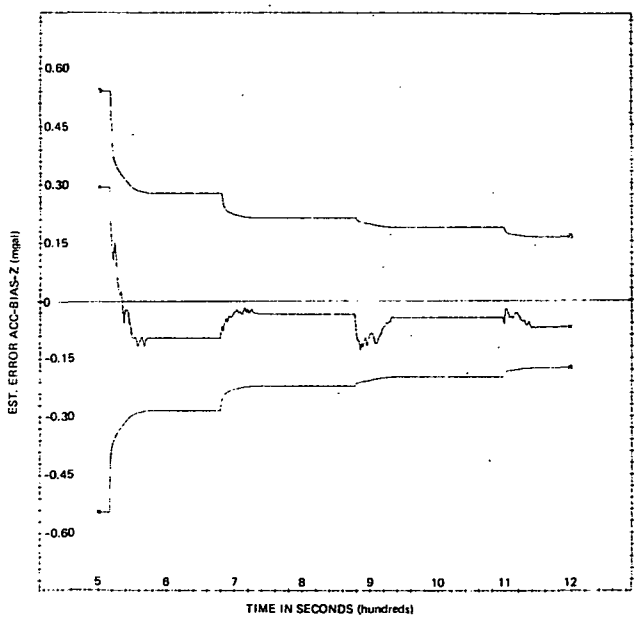



Figure 8. Vertical acceleration error and covariance envelope from navigation simulation.

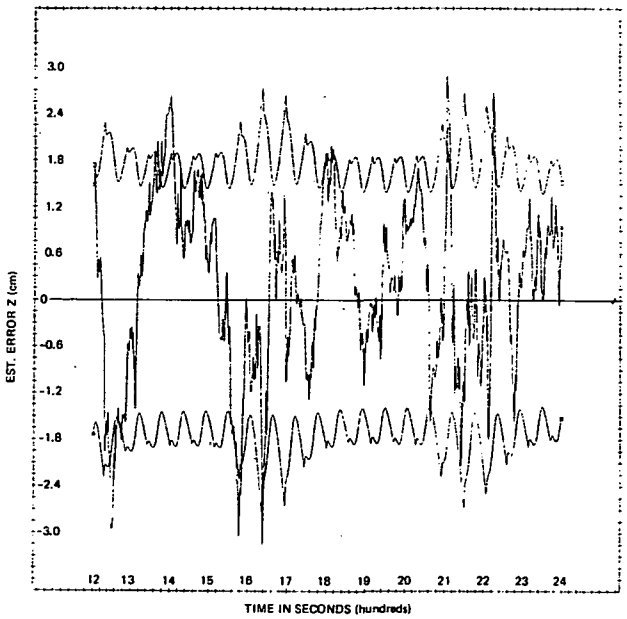



Figure 10. Vertical position error and covariance envelope from crustal-measurement simulation.

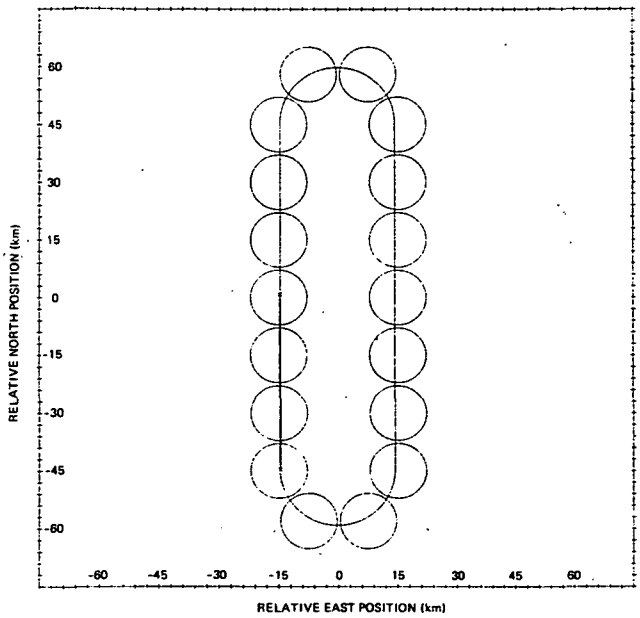



Figure 9. Flight trajectory and retroreflector geometry for simulation of APTS measurement of crustal movements.

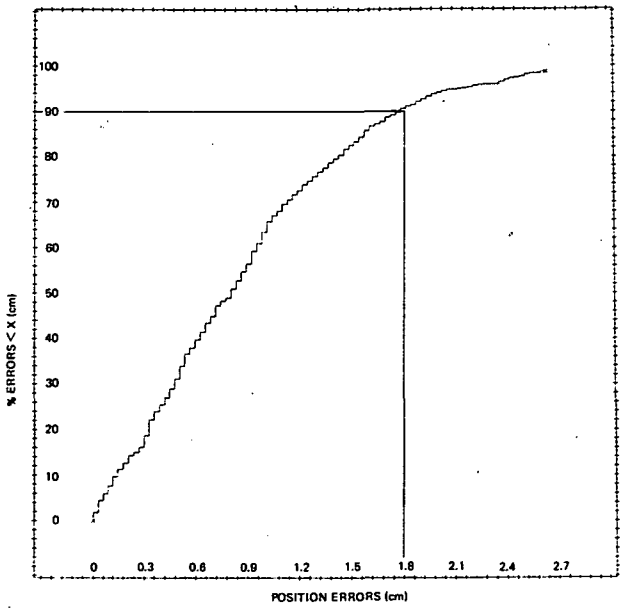



Figure 11. Integrated histogram of vertical position errors from crustal movement simulation.

# Models for Extracting Vertical Crustal Movements from Leveling Data

Sandford H. Holdahl
National Geodetic Survey, National Ocean Survey/NOAA
Rockville, Maryland 20852

Abstract. Various adjustment strategies are now being used in North America to obtain vertical crustal movements from repeated leveling. The more successful models utilize polynomials or multiquadric analysis to describe elevation change with a velocity surface. Other features permit determination of nonlinear motions, motions associated with earthquakes or episodes, and vertical motions of blocks where boundaries are prespecified. The preferred models for estimating crustal motions permit the use of detached segments of releveling to govern the shape of a velocity surface and allow for input from nonleveling sources such as tide gages and paired lake gages. Some models for extracting vertical crustal movements from releveling data are also excellent for adjusting leveling networks, and permit mixing old and new data in areas exhibiting vertical motion. The new adjustment techniques are more general than older static models and will undoubtedly be used routinely in the future as the constitution of level networks becomes mainly relevelings.

## Introduction

In the United States, most leveling surveys have been performed to support individual engineering or mapping projects. Until recent years the timing and arrangement of the surveys were rarely influenced by the geodesist's need to detect verical crustal movement. The most prominent factors influencing the network development were the availability of cooperative funding from local government and the desire to eliminate what were regarded as weaknesses in the network. Generally, the development of most national networks geographically follows the development of a nation, and "ideal" plans for quickly establishing a network of strategically spaced lines are rarely implemented.

Because of the manner in which most national networks evolve and are maintained, the geodesist is challenged to find ways of detecting crustal movements from scattered releveling over an original network which is also not time-homogeneous. The detection of vertical crustal movements is important to the earth scientist, but for the geodesist it is also necessary to model such movements when adjusting networks of leveled height differences of different dates. The geodesist then wants to find models which bring consistency between the observations, the detected movements or velocities, and the heights being published.

Various methods for determining and predicting vertical crustal movements have been used in North America. They are described in the following pages. Each of the methods works well in particular circumstances, two of the methods are general enough to be used frequently in height computations. Figure 1 through 4 schematically illustrate characteristics of leveling networks. Various line types indicate that observations were obtained in different years.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

## Methods and Models

### Method 1

Occasionally the distribution of original and repeated levelings is almost ideal for a small area (see Figure 1). Two levelings covering the study area, each accomplished within a short time period and adequately separated from each other in time, may be adjusted independently. After the adjustments, movements are calculated by comparing the two sets of adjusted heights. To make the comparison, a movement is assumed to be known for one of the common points; usually the movement at that point is taken to be zero and the computed movements are considered relative. If one of the points is a tide gage, absolute movement at that point can be inferred from the tide gage record. Absolute movements at other points can be calculated by adding in a constant when making the comparison. Velocities are obtained by dividing the movements by the time elapsed between epochs.

Method 1 is worth mentioning because it does not involve complicated mathematical models and thus avoids the need to develop special computer programs. However, the network in Figure 1 is handled equally well by the more general methods described next.

### Method 2

Often this method is more applicable for crustal movement determination than Method 1 because data requirements are less restrictive. The original and repeat observations existing in an area will generally not be separated by a constant time interval. By forming velocity difference observations from repeat levelings, the data are effectively made homogeneous. Since velocity observations are independent of date, velocity misclosures should theoretically equal zero if leveling is perfect and the assumption of constant movement is correct.

The velocity difference, $\Delta v$, between points connected by releveling is computed according to:

$$\Delta v = \frac{\Delta h_2 - \Delta h_1}{\Delta t} \qquad (1)$$

where $\Delta h_1$ and $\Delta h_2$ are the old and new observed height differences respectively, and $\Delta t$ is the time elapsed between levelings.

The variance of the velocity difference is computed using equation (2),

$$m^2_{\Delta v} = (m_1^2 + m_2^2)S / \Delta t^2 \qquad (2)$$

FIGURE 1

**Complete Releveling - Two Leveling Epochs**

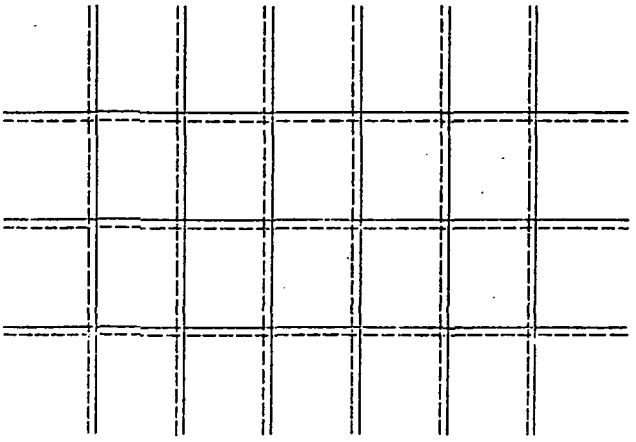


FIGURE 2

**Scattered Relevelings**

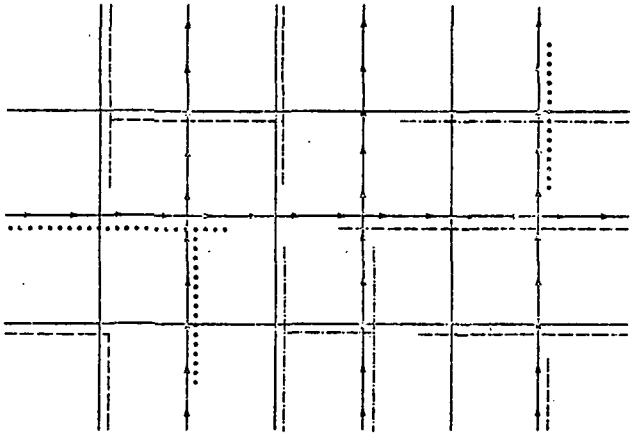


FIGURE 3

**Multiple Relevelings, Single Levelings**

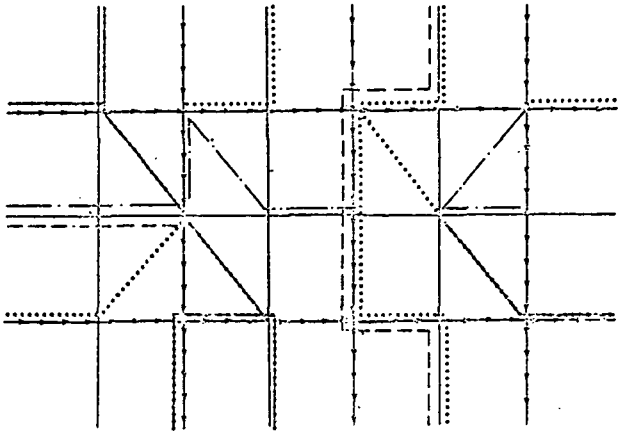
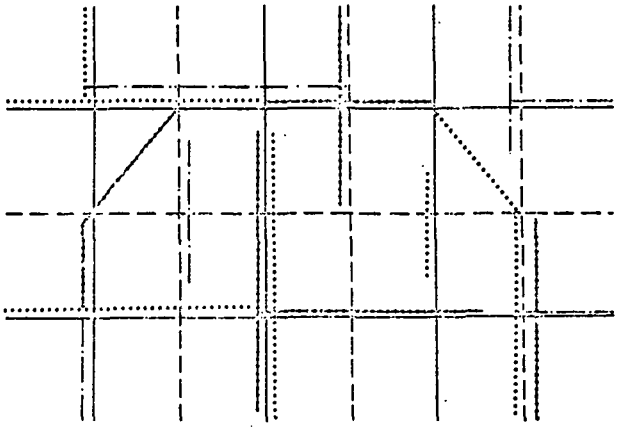


FIGURE 4

**Multiple Relevelings, Single Levelings, and Detached Relevelings**



where $m_1^2$ and $m_2^2$ are the a priori unit variances of the old and new levelings, and S is the distance between link terminals.

The velocity differences forming a network may be disconnected but not by long distances. Constraints requiring velocities in the same small locality to be identical can hold the segments together. These constraints are reasonable only if the geographical separation of observations is not great or the variation in velocity occurs very gradually over the unconnected area. Weighted constraints that allow some motion between neighboring junctions may also be used to hold disconnected segments together, but selecting a weight may be like guesswork.

Method 3

Method 3, which fits a velocity surface through a field of velocity differences, is described in detail by Vanicek and Cristodulides [1974]. The primary advantage of surface fitting is its usefulness on networks of scattered relevelings of the type shown in Figure 2.

The velocity surface V can be expressed by a generalized two-dimensional polynomial:

$$V(x,y) = \sum_{i=1}^{k} c_i f_i(x,y) \qquad (3)$$

where $f_i$ are arbitrarily chosen, linearly independent functions of the coordinates x and y, and $c_i$ are the coefficients which best fit the observations. A velocity difference can be written as:

$$\Delta V(x,y) = \sum_{i=1}^{k} c_i \Delta f_i(x,y) \qquad (4)$$

where $\Delta f_i(x,y) = f_i(x_B,x_B) - f_i(x_A,x_A)$, for a pair of connected bench marks, A and B. Because of simplicity, the two dimensional-algebraic functions $x^iy^j$, i, j = 0, 1, 2, . . ., m are frequently used for the $f_i$.

In the above type of adjustment, the unknowns are the coefficients $c_i$, and the observations are velocity differences computed from levelings using equation 1. The origin of the coordinate system is usually taken to be a point near the center of the study area. Once the coefficients are known, velocities for desired points are calculated using equation 3.

Although method 3 works well on the network shown in Figure 2, it would not fully utilize the measurements shown in Figure 4. To do so would require a nondiagonal weight matrix or preprocessing to obtain weighted mean velocity differences when there is more than one releveling. Neither does the model utilize the information found in circuit misclosures. The main advantage of Method 3 is that it minimizes the number of unknowns in the solution.

Method 4

This technique uses simple polynomials to describe height variations at selected bench marks in the study area. If implemented in its most basic form, the observations are differences of elevation rather than velocity differences. It is assumed that elevation differences are connected as shown in Figure 3.

At the onset, we pick a starting or reference time, $t_o$. Then, for example, the height of a bench mark A at time $t_i$ is written as follows:

$$h_{a,i} = h_{a,o} + a_1(t_i - t_o) + a(t_i - t_o)^2 + \ldots \quad (5)$$

where $h_{a,o}$ is the elevation of bench mark A at the reference time. The observation equation for Method 4 is:

$$r_{b-a,i} = h_{a,i} - h_{b,i} - \Delta h_{b-a,i} \quad (6)$$

where $\Delta h_{b-a,i}$ is the observed difference of elevation between bench marks A and B at time $t_i$.

Ordinarily the data redundancy will not permit the use of polynomials higher than degree 3; there is not much advantage in a higher degree even if permitted by the data. When the degree of the polynomial (equation 5) is 2, then $a_1$ is the velocity of elevation change at time $t_o$ and $a_2$ the acceleration. At a time other than $t_o$ the instantaneous velocity at bench mark A is calculated according to

$$V_{a,i} = a_1 + 2a_2(t_i - t_o) \quad (7)$$

When the degree of the polynomial (equation 5) is of degree 1, $a_1$ is a constant velocity.

The above method has some very nice advantages:

(1) If there are three or more relevelings over the same segments, all can be put into the adjustment without resorting to a nondiagonal weight matrix. Single levelings, in appropriate locations, can also add strength to the solution.
(2) Velocity and velocity difference observations are easily introduced to the adjustment. Velocities which have been inferred from tide gage records are entered as weighted parameters; velocity difference observations, computed from pairs of lake level gage records may be entered as differences between the $\Delta t$ coefficients corresponding to the two points on the sides of the lake.
(3) The solution produces a homogeneous set of heights which correspond to a selected point in time, $t_o$.
(4) Each point polynomial may have its own degree, the degree being limited only by common sense and the number of excess observation of different date contacting the point and connecting it to the network.

Regarding item (4), it is occasionally difficult to decide how many unknowns can be solved for at each point. Unusual configurations of observations may cause one to guess incorrectly. Therefore, for large complex networks, it is helpful to have a preprocessing program or subroutine to determine solvability.

Method 5

This is a combination of Methods 3 and 4. For any bench mark A in the study area, we can give the following expression for its height at time $t_i$

$$h_{a,i} = h_{a,o} + V(x_a,y_a)(t_i - t_o) \quad (8)$$

where, for example,

$$V(x_a,y_a) = c_o + c_1 x_a + c_2 y_a + c_3 x_a y_a + c_4 x_a^2 + \ldots \quad (9)$$

The unknowns in the adjustment are the height at each point corresponding to time $t_o$, and the coefficients $c_k$, k = 1,2,3,...m which define the velocity surface. If u is the number of unknown junction heights, then the total number of unknowns is u + m. The observation equation is then written as in equation (6). Note that the constant term of equation (9) drops out. The constant term is the absolute velocity of height at the origin of the network. If known, this can be conveniently specified and its uncertainty propagated into computed velocities.

Method 5 has the important advantages of Methods 3 and 4. Height differences are adjusted rather than velocity differences; therefore, no processsing is required to convert leveling observations to velocity differences. Method 5 is preferable to Method 3 when the number of unknowns in the adjustment does not tax the computer.

The choice of whether to use Method 4 or Method 5 will depend on the configuration of relevelings and the extent of the geodesist's foreknowledge of the movement pattern in the

study area. In general, Method 4 is more sensitive, but the relevelings must be interconnected; otherwise, each independent subnetwork must have its own initialization in height and velocity.

The solutions produced by Methods 4 and 5 are conceptually different. Method 5 gives a solution that requires all bench marks in a relatively small locality to take on the same velocity, because velocity is a function of position. Method 4 does not naturally provide for this local consistency, but it can be forced by the addition of appropriate constraints between the velocity unknowns of points in the same locality.

Generally, the weighted sum of squared residuals (V'PV) from Method 4 will be less than the corresponding sum from Method 5. Method 5 produces a smoothly fluctuating velocity surface, whereas Method 4 accommodates the observations with any number of bumps and dips having whatever amplitudes are required to minimize corrections to the observations. It is probably misleading to argue which method is best from this point of view because both produce results of high value and the comparison of the two solutions may be of most interest. The larger separations of the two solutions can be regarded as local velocity anomalies. These should be examined closely as they may be indications of local accumulations of systematic errors in the leveling data.

Polynomial expressions for velocity surfaces may produce problems with computer graphics because the fitted surfaces quickly taken on extreme values outside the data area. This may ruin the scale of three-dimensional plots, produce error messages, or use excessive computer time for contouring. To avoid these problems, multiquadric (MQ) analysis has been used as an alternative to polynomials in crustal movement investigations [Holdahl and Hardy, 1977].

Method 6

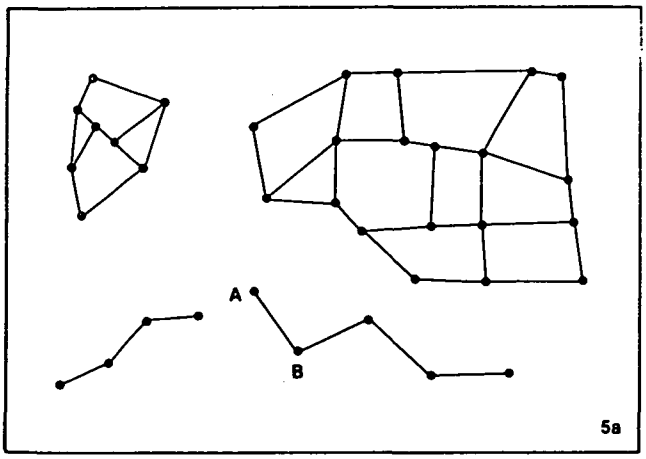
This is essentially the same as Method 5 except that we employ MQ analysis, and replace equation (9) with

$$V(x_a,y_a) = \sum_{j=1}^{k} c_j Q\left[x_a,y_a,x_j,y_j,D\right] \quad (10)$$

The $C_j$ are undetermined coefficients; Q is a quadric kernel function; the $x_j,y_j$ are the positions of nodal points; and D is a geometric parameter which may or may not be needed depending on the quadric form. Nodal points should be located where there are solvable point velocities or tilt information. If the hyperboloid is selected as the quadric form, we then have the following expression for velocity of elevation change:

$$V(x_a,y_a) = \sum_{j=1}^{k} c_j \left[(x_a - x_j)^2 + (y_a - y_j)^2 + D^2\right]^{\frac{1}{2}}. \quad (11)$$

Substituting (11) into (8) gives us a model for leveling adjustments, which advantageously produces automated graphic displays of the velocity surface and velocity error surface without destruction of scale by an extreme value calculated at an uncontrolled edge of a rectangular study area.



5a

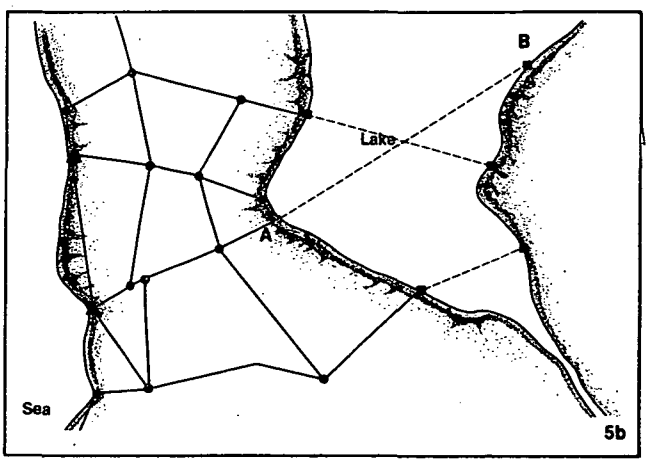
In Method 6, nodal points should be placed at each location that has a solvable point velocity, as in Method 4. Additional nodal points can be added wherever detached tilt information is located.

Scattered or detached relevelings (see Figure 5a) cannot contribute to the determination of the absolute position of a velocity surface unless they are individually initialized by a measured or assumed velocity. However, detached relevelings can be used to aid in determining the local shape of a velocity surface. This was mentioned previously as being the original motivation for developing Method 3. Methods 5 and 6 also permit the use of detached relevelings. In Method 5, the velocity difference observation equation resulting from a detached releveling between points A and B would be written as

$$r(x,y) = \sum_{i=1}^{nx} \sum_{j=1}^{ny} c_{ij} \left(x_b^i,y_b^i - x_a^i,y_a^j\right) - \Delta V_{b-a}. \quad (12)$$

In Method 6,

$$r(x,y) = \sum_{j=1}^{k} c_j \left[Q(x_b,y_b,x_j,y_j,D) - Q(x_a,y_a,x_j,y_j,D)\right] - \Delta V_{b-a} \quad (13)$$



5b

where $\Delta V_{b-a}$ is obtained from the observed height differences, as in equation (1). These same equations can be used to incorporate relative velocities between pairs of water level gages on a lake (see Figure 5b). In the latter case, $\Delta V_{b-a}$ would be the slope of a line fitted through a plot of differences between readings from gages at A and B obtained over a period of years.

For two or more relevelings over a detached segment, equations (12) and (13) are not appropriate. In Models 5 and6, it is better to use the usual height difference observation equations, and add another observation (fictitious or otherwise) that specifies a height at $t_o$ for one of the points on the detached segment. This eliminates concern for correlation of velocity difference observations. In setting up crustal movement studies, it has been convenient to have only one fictitious fixed height, which may be called "mean sea level." It has a height of 0.0 meter at the reference time, $t_o$, and other heights at $t_o$ are provided to the adjustment as fictitious height difference measurements from it with an appropriate weight. These fictitious height difference measurements are, in fact, contraints. No special coding is required if they are treated as measurements.

Method 7

Vanicek revised his original model (Method 3) to consider episodic and nonlinear movements. In the expanded model the "observations" are relative movements obtained by comparing repeated measurements of height differences, whereas in Method 3 they are velocity differences.

The mathematical formulation for the movement surface, u is

$$u(x,y,t) = \sum_{k=1}^{n_t} c_{ok}T_k(t) + \sum_{i=0}^{n_x}\sum_{j=0}^{n_y}\sum_{k=1}^{n_t} c_{ijk}x^i y^j T_k(t) \quad i+j\neq 0$$

$$= \underline{T}(t)\underline{c}_o + X(x,y,t)\underline{c} \qquad (14)$$

where

$$T_k(t) = t^k \qquad k=1,n_p$$

$$T_k(t) = \begin{cases} 0 & t<b_k \\ (t-b_k)/(e_k-b_k) & b_k\leq t\leq e_k \\ 1 & t>e_k \end{cases} \qquad k=n_p+1,n_p+n_e$$

and $\underline{c}$ is the n-vector of unknown coefficients $c_{ijk}$ for the previously chosen values $n_x, n_y$ of maximum power in x and y.

In these formulas x and y are local horizontal Cartesian coordinates calculated from latitude and longitude through the following transformation equations:

$$x = (\phi-\phi_o)R \; ; \; y = (\lambda-\lambda_o)R\cos\phi_o \qquad (15)$$

where $(\phi_o,\lambda_o)$ is the centroid of all bench marks and R is the mean radius of the Earth. Time, t, is reckoned from a stipulated date, $t_o$, for which $u(x,y,t_o)$ is everywhere equal to zero. In addition, $b_k, e_k$, for $k=n_p+1, n_p+n_e$, are the beginning and ending dates of $n_e$ movement episodes so that $n_t=n_p+n_e$. We note that the episodic movements are treated as linear within the duration of the episode.

Observation equations for m releveled segments can now be written:

$$\underline{\Delta h}(x_1,y_1,x_2,y_2,t_2) - \underline{\Delta h}(x_1,y_1,x_2,y_2,t_1)$$

$$= \underline{d}(x_1,y_1,x_2,y_2,t_1,t_2)$$

$$= \sum_{i=0}^{n_x}\sum_{j=0}^{n_y}\sum_{k=1}^{n_t} c_{ijk}(x_2^i y_2^j - x_1^i y_1^j)\left[T_k(t_2)-T_k(t_1)\right]$$

$$+ \underline{r}(x_1,y_1,x_2,y_2,t_1,t_2) \qquad i+j\neq 0$$

$$= \underline{B}(x_1,y_1,x_2,y_2,t_1,t_2)c+r \qquad (16)$$

where $\underline{d}$ denotes the m-vector of the differences of leveled height differences, $\Delta h$, and $\underline{r}$ is the residual vector. If

$$m > n = (n_x\cdot n_y + n_x + n_y)n_t \qquad (17)$$

we can find the solution, $\underline{c}$, through the method of least squares. The normal equations are solved, in the computer program, through orthogonalization.

The shift coefficients, $\underline{c}_o$, cannot be determined from the releveled segments alone. Movement u* (x,y,t) of at least one, but generally $n_g$ tide gages (x,y) must be determined from sea level records at $n_d$ dates to allow for evaluating the shift coefficients. The following $n_g\cdot n_d$ observation equations can be then formulated:

$$u^*(x_i,y_i,t_j) = \underline{T}(t_j)\underline{c}_o + \underline{X}(x_i,y_i,t_j)\underline{c} + \underline{r}^*, \qquad i=1,n_g;\; j=1,n_d \qquad (18)$$

If $n_g\cdot n_d > n_t$, then the equations (18) may be solved for $\underline{c}_o$, again using the method of least squares.

For each tide gage the uplift u* must be determined so as to satisfy the following condition:

$$u^*(x,y,t_o) = 0. \qquad (19)$$

As with Method 3, the advantage is primarily the reduced number of unknowns. However, the same disadvantages remain: (1) observations may be correlated and (2) information contained in circuit misclosures is not utilized.

The advantage of being able to estimate episodic and nonlinear movements is very attractive, but these same features can easily be incorporated in Methods 5 and 6 without concern for correlation of observations. Equation (8), for expressing height of point A at time $t_i$, can be modified to include terms corresponding to the elevation change associated with earthquakes that occurred between the times when levelings were accomplished within the study area:

$$h_{a,i} = h_{a,o} + V(x_a,y_a)(t_i-t_o) + \sum_{j=1}^{ne} u_j(x_a,y_a,x_j,y_j,t_j,d_j) \quad (20)$$

where ne is the number of events or episodes, $t_j$ the time of an earthquake, $d_j$ the depth of the earthquake in units identical to x and y, and $(x_j,y_j)$ the location of an event. A logical choice for the function u, suggested by R. Snay, is one where episodic elevation change decays with distance from the event:

$$u(x,y,x_j,y_j,t_j,d_j) = 0, \quad \text{if } t_i < t_j$$

$$= \sum_{j=1}^{ne} a_j\left[(x-x_j)^2+(y-y_j)^2+d_j^2\right]^{-\frac{1}{2}}, \quad \text{if } t_i > t_j. \quad (21)$$

The coefficients, $a_j$, are to be solved for in the adjustment. A nice advantage is that only one unknown is introduced for each earthquake or episode. If Figure 5c, three events are illustrated, and the contribution of those events can be evaluated at any time or location following the adjustment. If episodic motions are not modeled in the adjustment, observations which are suspected of having been affected by earthquakes must be removed prior to adjustment. Removal of observations should be the last alternative, and is difficult to justify except when the leveling is suspect or when insufficient relevelings exist to permit solving for episodic motion.

Another form of flexibility involves modeling of vertical block motions characterized by discontinuities of movement at fault lines, Methods 5 and 6 can be modified to accomplish this, As in Figure 5d, we can divide a study area into three blocks, P, Q, R, and express, for example, the height of a point A, at time, $t_i$, on block P, as follows:

$$h_{a,i} = h_{a,o} + V_P(x_a,y_a)(t_i-t_o) \quad (22)$$

where $V_P$ describes the velocity surface of block P. The height difference between points A and B, where B is on block Q, is given by

$$\Delta h_{b-a,i} = h_{b,o} - h_{a,o} + [V_Q(x_b,y_b) - V_P(x_a,y_a)](t_i-t_o) \quad (23)$$

Essentially, every point is located on one of the blocks, and each block has its own velocity surface. Equation (22) can be supplemented with terms which permit episodic or nonlinear vertical motions within selected blocks.

By permitting the model to solve for block motion, episodic motions, and accelerations we greatly increase its flexibility. But there is the concern that almost any kind of blunder or systematic error may be modeled as crustal movement. With an inflexible model the opposite is true, i.e., unusual movements will be forced to occur at constant rates and be partially absorbed by large corrections to the observations. The ideal adjustment model, then, is one that is very flexible and provides the geodesist with the possibility of describing vertical movement of any type; and the best strategy for the adjustment is to use only as much of that flexibility as is prudent after considering the seismicity, geology, and engineering activity in the area.

## Solvability

Networks of relevelings can become complicated in the sense that casual observation of the net may not reveal which unknowns are solvable. An algorithm has been developed by Allen Pope of the National Geodetic Survey, that uses the observation equations to resolve such questions.

If we use Method 4, the solvability algorithm will resolve exactly which heights and point velocities are solvable. When Methods 5, 6, and 7 are used, the solvability algorithm will show that all coefficients of the velocity surface are unsolvable if one attempts to solve for too many. It has been very helpful to use the solvability algorithm as an analysis tool by first formulating the leveling adjustment using the observation equations according to Method 4, i.e., pretending to solve for the reference-time heights and velocities at all junctions; and secondly formulating the problem the way it would actually be adjusted, using Method 6. When solvability fails using Method 6, the user can identify the cause by reviewing the output of solvability as applied to Method 4. It tells which points have solvable velocities. The number of coefficients that may be used to describe the velocity surface is equal to the number of solvable point velocities, plus the number of paired junctions which do not have solvable velocities but have relative velocity information in between. These latter pieces of floating tilt information must be counted by looking at the network diagram.

## Systematic Errors

Certain leveling errors are time-dependent. Therefore, an attempt should be made to eliminate them prior to adjustment. Without elimination, their influence is modeled as vertical movement. Where short time intervals or slow velocities are involved, the error due to uncorrected systematic leveling errors may be larger than the real movements.

One of the errors most damaging to leveling is caused by refraction. During normal daylight working hours the line of sight is bent upwards. The uphill sight will bend more than the downhill sight because the density of air changes most rapidly near the ground, the hotter air being nearest the ground. The amount of bending is proportional to the square of the sight length, the leveled height difference, and the vertical temperature change, $\Delta t$, between heights of 50 and 250 cm. $\Delta t$ is dependent on time of day, season local turbidity of the atmosphere, cloud cover,

the direction of leveling, and the slope of the leveling path.

Few countries have applied the refraction correction that was developed by T. J. Kukkamaki in 1937. The measurement of $\Delta t$, for input to the formula, has been considered an awkward task for a leveling team. The correction itself has long been considered too small, probably because most of the documented experience with measuring $\Delta t$ comes from northern countries where the sun exerts a lesser influence because of its lower declination. In the lower latitudes of the United States large temperature gradients have been observed 5 to 10 times as great as the average values observed in England. Further, experiments and microclimate theory support the idea that refraction error is less on the north face of a mountain than on the south face. This produces a north-south accumulation of error when leveling on undulating terrain [Holdahl 1978]. It was thought that only the leveling in countries with large mountains suffered significantly from refraction error. This is incorrect because terrain that merely undulates in the north and south direction can yield a large accumulation of error if leveling extends for several hundred kilometers.

Refraction error affects computations of crustal movements in several ways.

(1) Two levelings accomplished in distinctly different seasons (i.e., seasons with different declinations of the sun) will usually yield an apparent relative elevation change.

(2) Single levelings, that are not corrected for refraction, but are permitted as observations in some adjustment models, will contribute to circuit misclosures in a way that cannot be distinguished from a contribution to the same circuit misclosure caused by a real crustal movement.

(3) A releveling accomplished with a maximum sight length specification, which significantly differs from the specification used in the original survey, will normally yield an apparent crustal movement if all other conditions are equal.

Because it has rarely been measured, it is necessary to estimate $\Delta t$ if a refraction correction is to be applied to old leveling data. A method based on historic measurements of solar radiation is being developed to accomplish this [Holdahl 1978]. This method is untested at this time. Until some method is shown to be corrective, it is doubtful that high reliability can be associated with conclusions derived from large networks of releveling in areas of undulating terrain and high levels of solar radiation.

Another leveling error, which might be termed systematic, results from neglect of gravity anomalies. Some of the above-described methods for calculating crustal motions are not vulnerable to this error (Methods 1,2,3, and 7). Methods 4,5, and 6 would, however, be adversely influenced by gravity anomalies if all the following conditions exist:

(1) the height differences were not provided in geopotential units;

(2) the gravity anomaly is large and height rapidly changes over the study area;

(3) single levelings, i.e., segments or links which have not been releveled, are used in the network adjustment.

Misclosures of leveled circuits are theoretically equal to zero only when the height differences are computed in geopotential units, thus taking into consideration the variation of true gravity along the level lines. For small study areas, where variation in the gravity anomaly is small or relevelings exist over the entire net, the use of geopotential units is unimportant if only the velocities of elevation change are desired. The heights, rather than the velocities, are most sensitive to gravity anomalies. The gravity effect tends to cancel in the velocity determination when releveling is complete. Networks with a significant percentage of unreleveled segments should usually be adjusted using geopotential units because the cancellation may be incomplete or nonexistent.

Smaller types of known systematic errors exist and their influence is dealt with by corrections to the observed height differences. These include the

(1) Astronomic correction: accounts for deviation of the vertical due to positions of the Sun and Moon

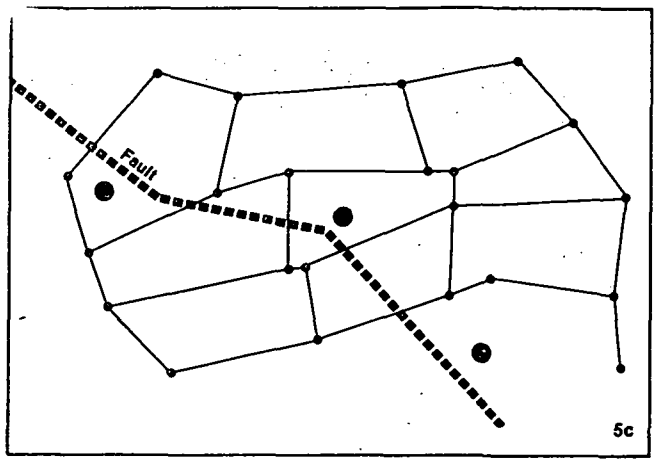
(2) Rod correction: accounts for minor scale error due to deviation of calibrated rod length from nominal rod length,

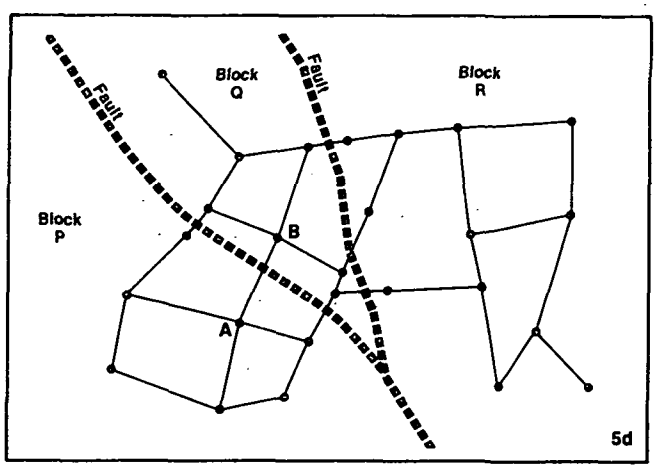
(3) Rod temperature correction: accounts for contraction and expansion of the rod length due to heat,

(4) Collimation correction: accounts for nonparallelism of the telescope and an equipotential surface passing through its center,

(5) Orthometric correction: corrects for convergence of normal equipotential surfaces. Satisfactory only when gravity anomalies are near zero.

Unfortunately, there seem to be some systematic errors which are unknown or poorly understood. This is evident from sea slope determinations that have been accomplished at different times in the last several decades on the California coast. These determinations show a spread that is too wide to be attributed to random leveling errors. The most recent sea slope determination agrees well with estimates obtained by oceanographic techniques but the disagreement of present and former determinations is still a mystery. The geodesist should use height-velocity adjustment models to filter data and calculate motions, always being alert for measurements that may be contaminated by systematic error accumulation. At the same time the user must be aware that any model has "built-in" assumptions and constraints which restrict the ways in which motion may be resolved, while nature produces an endless variety of ways to exhibit motion.




References

Hardy, R. L., Research results in the application of multiquadric equations to surveying and mapping problems, Surveying and Mapping, December, Vol. XXXV, No. 4, pp. 321-332, 1975.

Holdahl, S., Studies of precise leveling at California fault sites, NOAA Technical Report NOA NGS , 1970.

Holdahl, S., Recent elevation change in Southern California, NOAA Technical Memorandum, NOS NGS-7 19 pp, Rockville, 1977.

Holdahl, S., Removal of refraction errors in geodetic leveling, IAU Symposium No. 89, Refractional Influences in Geodesy, Uppsala, Sweden, 1978.

Holdahl, S., and R. L. Hardy, Solvability and multiquadric analysis as applied to investigations of vertical crustal movements, Proceedings of the 1977 International Symposium on Recent Crustal Movements, Palo Alto, California, 1977.

Holdahl, S., and N. Morrison, Regional investigations of vertical crustal movements in the U.S. using precise relevelings and mareograph data, Tectonophysics, 23 (4), 373-390, 1974.

Vanicek, P., and D. Cristodulidis, A method for the evaluation of vertical crustal movement from scattered geodetic relevelings, Canadian Journal of Earth Science, 11, 605-610, 1974.

Vanicek, P., M. R. Elliot, and R. O. Castle, Four-dimensional modeling of recent vertical crustal movements in the area of the Southern California uplift, Proceedings of the 1977 International Symposium on Recent Crustal Movements, Palo Alto, California, 1977.

Vanicek, P. and A. C. Hamilton, Further analysis of vertical crustal movement observations in the Lac. St. Jean Area, Quebec, Canadian Journal of Earth Science, 9, 1139-1147, 1972.

# Detection of Accelerated Crustal Movements Based on Terrestrial Techniques

Robert O. Castle
U.S. Geological Survey
Menlo Park, California 94025

Summary. Reported accelerations in the vertical displacement field disclosed through repeated levelings are unassailable where the magnitudes of the measured displacements that define these accelerations overwhelm any conceivable survey error and the nature of the measured movement is fully consistent with its predicted occurrence. Specific examples include those accelerations associated with seismic slip events and massive withdrawals of underground fluids. Aseismic accelerations based on progressively smaller vertical displacements are increasingly equivocally defined; nonetheless, the reality of sharply defined aseismic accelerations associated with modest vertical displacements is independently confirmed by both continuous sea-level measurements [Wyss, 1977] and lake-level records [Wilson and Wood, 1978]. Although replication of observed elevations during the periods that both preceded and followed aseismic uplift in southern California [Savage and Prescott, 1979] is excellent evidence of the existence of major vertical accelerations, unambiguous data sets of this sort rarely occur in the geodetic record. Moreover, because the detection of vertical accelerations commonly depends on the results of a single leveling, validation of these accelerations may require detailed assessment of the accuracy of the critical survey. Where the interval between levelings is significantly greater than the inferred period of the acceleration, characterizations of accelerations based solely on the results of repeated level surveys becomes significantly less meaningful. However, four-dimensional modeling techniques that depend only on the existence of continuous and/or discontinuous releveled segments [Vaníček et al., 1979], rather than continuous line or network relevelings, may permit the detection and relatively unambiguous representation of otherwise unrecognizable crustal accelerations.

Continuing efforts directed toward the recognition of accelerations in the vertical displacement field should be based ideally on the results of repeated level surveys coupled with those additional measurements, such as continuous sea-level records, repeated gravity surveys and so forth, that provide temporal constraints on any measured vertical movements. In addition, repeated levelings designed to detect vertical accelerations should be tailored to the need. If, for example, the purpose is the best possible characterization of the vertical movement history athwart an active fault, the most useful program probably will continue to consist of frequently repeated surveys along the same line referred to a control point well removed from the deformational field associated with continuing fault movement. On the other hand, where a generalized representation of regionally developed accelerations is desired, four-dimensional surface fitting that utilizes segmented relevelings randomly distributed in both space and time probably is the most cost-effective approach for meeting this goal. While surface-fitting techniques tend to subdue short wavelength features, they are especially well suited to the depiction of those accelerations accompanying artificially induced subsidence and broadly defined tectonic deformation such as that recognized in southern California.

## References

Savage, J.C., and W.H. Prescott, Geodimeter measurements of strain during the southern California uplift, Jour. Geophys. Research [in press], 1979.

Vaníček, Petr, M.R. Elliott, and R.O. Castle, Four-dimensional modeling of recent vertical crustal movements in the area of the southern California uplift, Tectonophysics [in press], 1979.

Wilson, M.E., and S.H. Wood, Salton Sea water-level records (1952-1977) and the southern California uplift (abs.), Trans. Amer. Geophys. Union, 59, 242, 1978.

Wyss, Max, The appearance rate of premonitory uplift, Bull. Seismological Society America, 67, 1091-1098, 1977.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

Page Intentionally Left Blank

# Gravity Field, Geoid and Ocean Surface by Space Techniques

R. J. Anderle
Naval Surface Weapons Center
Dahlgren, Virginia 22448

Abstract. Knowledge of the earth's gravity field continued to increase during the last four years. Altimetry data from the GEOS-3 satellite has provided the geoid over most of the ocean to an accuracy of about one meter. Increasing amounts of laser data has permitted the solution for 566 terms in the gravity field with which orbits of the GEOS-3 satellite have been computed to an accuracy of about one to two meters. The combination of satellite tracking data, altimetry and gravimetry has yielded a solution for 1360 terms in the earth's gravity field. A number of problems remain to be solved to increase the accuracy of the gravity field determination. New satellite systems would provide gravity data in unsurveyed areas and correction for topographic features of the ocean and improved computational procedures together with a more extensive laser network will considerably improve the accuracy of the results.

Major improvements in our knowledge of the earth's gravity field have been obtained since the GEOPS 4 Conference on The Geoid and Ocean Surface in August 1973 (Rapp, 1974). A dramatic improvement was obtained by use of the GEOS-3 satellite altimetry data, which has defined the geoid over most of the ocean areas to about a one meter accuracy (Brace, 1977, Hadgigeorge in press, Kahn et. al. 1977, Marsh et. al. 1978, Marsh et. al. in press, Rapp 1978, Rapp in press, Yionoulis et. al., in press). In each solution for the geoid based upon the altimetry data, errors in the satellite position were larger than uncertainties in the altimeter measurements. In order to reduce the effects of the orbit errors, bias parameters for each satellite pass were determined which minimized the discrepancy between geoid determinations at intersections of north-westerly satellite subtracks with south-westerly subtracks. Some analysts further reduced the dependence of the geoid determination on the accuracy of the satellite orbit by fitting the altimetric data to a low order reference geoid. Anderle (1977) raised the question of biases in the computed geoid arising from systematic errors in the computed satellite orbits, but Rapp (private communication) found that the different approaches to the computation of the geoid agreed to about one meter. The computed ocean geoids are also subject to effects of ocean tides and geostrophic effects. However, a new model of the principal ocean tide constituent for one degree areas agrees with tide measurements to 5 cm (Schwiderski, in press), and models for six additional constituents have been computed by Estes (1977).

Solutions for the earth's potential continued to include an ever increasing number of coefficients as additional satellite observations were acquired at higher levels of precision and as computer programs were extended. High quality laser data played an important role in the solutions by Gaposchkin (in press) and Lerch (1978). Table 1 lists the more recent solutions obtained at various agencies. Some solutions were based solely on observations of satellite motion while others, indicated by "general" under the column headed optimization, included gravimetric, astro-geodetic, and/or altimetric data. The solutions allowed the computation of the GEOS-3 satellite altitude to an accuracy of 1.5 m (Douglas and Anderle, 1977). The latest solution is believed to represent the ocean geoid to an accuracy of one or two meters (Lerch, et.al., 1978).

Another new source of data on the earth's gravity field acquired recently was from satellite-to-satellite tracking. Data were acquired between two low orbiting satellites in the Apollo-Soyuz experiment and also between high and low satellites, ATS-F and GEOS-3. Agreement of gravity anomalies for two degree and five degree squares computed from high-low satellite to satellite tracking data with anomalies computed from this data was around five to ten milligals (Hajela, 1977, Vonbun 1977, Vonbun 1978).

Highly precise values for effects of gravity coefficients of specific order have been obtained by analysis of resonance effects on satellite motion (King-Hele, et. al., 1978, Klokocnik, 1978, Reigber in press) and evaluated by Wagner (in press). Such results are useful in evaluation of general solutions based on less sensitive data.

A number of outstanding problems need to be addressed in the course of further refinement of our knowledge of the gravity field:

1. drag and solar radiation effects on satellite motion,
2. other small effects,
3. correlation of coefficients,
4. combination of heterogeneous data,
5. computational problems, and
6. instrument limitations.

The following comments expand briefly on these problems:

1. Changes in the atmospheric drag due to solar radiation and changes in magnetic flux affect the motion of satellites in ways which can be misinterpreted as effects of the earth's gravity field. The effects of direct solar radiation and earth's albedo on the more complicated satellite satellite configuration can similarly be misinterpreted.

2. Other smaller effects which must be accurately modeled or determined include

   a. crustal motion,

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics. October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

b. polar motion and earth's rotation,

c. solid earth, ocean and atmospheric tides,

d. ionospheric effects on electronic measurements,

e. tropospheric effects on measurements, and

f. ocean topography on geoids computed from altimetric data.

3. The separation of gravity coefficients computed from observed satellite motion is made difficult by the limited number of satellites with different orbital motion for which accurate observations are available because satellites with different orbital inclinations or orbital periods are sensitive to different orders and degrees of gravity coefficients. Yet, apart from resonance effects, satellite motion is insensitive to high order terms in the earth's gravity field. Separation of the coefficients is possible through the use of altimetric data over the oceans and gravimetric data. But such data are not available in many regions, and the precision of such data is insufficient to compute accurate satellite orbits.

4. Neglect of systematic instrument and environmental effects, truncation of the gravity field representation, and other model errors have generally yielded standard errors of solutions for gravity coefficients which are overly optimistic. As a result, solutions based on combinations of various types of data have usually required the use of arbitrary weights for the different data classes employed in order to obtain a reasonable contribution to the solution from each set of data.

5. The computation of the gravity field from observed satellite motion is a costly undertaking. The high cost is one factor which limits the number of coefficients used to define the gravity field in such computations, and the frequent use of an inconsistent number of gravity parameters for different data which are then combined in a single solution.

6. Instrument limitations for lasers include weather and the cost of installations. S-band radars require a transponder on the satellite, have limited availability due to their heavy workload, and are subject to ionospheric refraction errors. Doppler receivers require a transmitter on the satellite and are of lower precision than lasers on S-band radars.

Solutions to many of these problems will be addressed by the panel members. Gaposchkin and Smith will discuss primarily advanced computational techniques; Fishell and Reigber will discuss advanced measurement techniques, and Whitehead will address the topographic effects on the ocean geoid.

References

Anderle, R. J. , L. K. Beuglass and M. G. Tanenbaum, Accuracy of computed orbits of GEOS-3 satellite, Naval Surface Weapons Center Technical Report, TR-3470, May 1976.

GRAVITY FIELD DETERMINATIONS

| AGENCY | DESIGNATION | NUMBER OF COEFFICIENTS | OPTIMIZATION | REFERENCE |
|---|---|---|---|---|
| DMA | DOD WGS-72 | 472 | General | Seppelin, 1972 |
| NSWC | NWL 10-E | 401 | NAVSAT | --------- |
| NSWC | NWL 1G | 396 | GEOS-3 | Anderle, 1976 |
| NOAA | P = 4 | 252* | General | Chovitz, in press |
| GRGS/SFB | GRIM 2 | 950 | General | Balmino, 1977 |
| SAO | SE VI.3 | 604 | General | Gaposchkin, in press |
| GSFC | GEM 9 | 566 | Satellite | Lerch, 1977 |
| GSFC | GEM 10B | 1360 | General | Lerch, 1978 |

*104 density squares

TABLE 1

Anderle, R. J. and Rose L. Hoskin, Correlated errors in altimetry geoids, Geophysical Research Letters 4 (10), 421-423, October 1977.

Anderle, R. J., The Ocean geoid deduced from GEOS-3 satlllite radar altimetry data, Second International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics, National Technical University, Athens, in press.

Arnold, D. A., E. M. Gaposchkin, Y. Kozai, J. Latimer, C. G. Lehr, C. A. Lundquist, G. Menders, M. R. Pearlman, J. M. Thorp, C. R. H. Tsiang, G. Veis, F. L. Whipple, M. R. Williamson and J. Wohn, Smithsonian Astrophysical Observatory, National Geodetic Satellite Program, pp. 795-948, National Aeronautics and Space Administration, Washington, D.C., 1977.

Balmino, G., B. Moynot and Ch. Reigber, GRIM 2 1976 gravimetric geoid - evaluation, EOS, Transactions, American Geophysical Union, 58 (5), May 1977.

Berbert, J. H., J. Brownd, T. Felsentreger, D. Harris, T. S. Johnson, M. A. Khan, F. Lerch, J. Marsh, J. Murphy, B. Putney, J. Reece, J. Richardson, M. Sandson, P. Schmidt, D. Smith, S. Vincent, and C. Wagner, NASA Goddard Space Flight Center, National Geodetic Satellite Program 295-484, National Aeronautics and Space Administration, SP-365, Washington, D.C., 1977.

Blaha, Georges, Refinements in the combined adjustment of satellite altimetry and gravity anomaly data, Air Force Geophysical Laboratory Report, TR-77-0164, July 1977.

Brace, Kenneth L., Preliminary ocean-area geoid from GEOS-3 radar altimetry, presented at NASA GEOS-3 Investigations Final Meeting, New Orleans, November 1977.

Chovitz, B. and Koch, Global gravity field augmented by altimetry data, Journal of Geophysical Research, in press.

Douglas, M. S. and R. J. Anderle, Accuracy of computed orbits of GEOS-3 satellite, GEOS-III Principal Investigators Meeting, New Orleans, La., 18-19 November 1977.

Estes, R. H., A computer software system for the generation of global ocean tides including self-gravitation and crustal loading effects, Business and Technological Systems, Inc., Report TR-77-41, February 1977.

Fang, B. T., Satellite-to-satellite tracking orbit determination, American Institute of Aeronautics and Astronautics, reprint 78-48, January 1978.

Gaposchkin, E. M., Gravity field determination from laser observations, Phil. Trans. R. Soc. Lond. A. 284, 515-527, 1977.

Gaposchkin, G. M., Glogal gravity field, Second International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics, National Technical University, Athens, in press.

Gerber, Mark A., Gravity gradiometry, Astronautics and Aeronautics, 16 (5), 18-26 May 1978.

Hadgigeorge, G. and G. Blaha, Gravity field determination from combination of altimetric and gravimetric data, Second International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics, National Technical University, Athens, in press.

Hajela, D. P., Recovery of 5° mean gravity anomalies in local areas from ATS-6/GEOS-3 satellite-to-satellite range rate observations, Air Force Geophysical Laboratory Technical Report 0272, September 1977.

Jordan, Stanley K., Fourier physical geodesy, Air Force Geophysical Laboratory Technical Report 78-0056, March 1978.

Kahn, W. D., J. W. Siry, R. D. Brown and W. T. Wells, Ocean gravity and geoid determination, Goddard Space Flight Center preprint X921-77-259, October 1977.

King-Hele, D. G., Doreen M. C. Walker, and R. H. Gooding, Evaluation of 14th order harmonics in the geopotential, Royal Aircraft Establishment Technical Report 78015, February 1978.

Klokocnik, J. and J. Kostelecky, Evaluation of the 14th order harmonics from resonant inclination variations, Bulletin Astronomical Inst. Czech., 29, 10-14, 1978.

Klokocnik, J., Some characteristics of the earth gravitational field from orbital resonances of the intercosmos satellites, Second International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics, National Technical University, Athens, in press.

Lerch, F. J., S. M. Klosko and R. E. Laubscher, and C. A. Wagner, Gravity model improvement using GEOS-3 (GEM 9 and 10), Goddard Space Flight Center Document X-921-77-246, September 1977.

Lerch, F. J., C. A. Wagner, S. M. Klosko, R. P. Belott, R. E. Laubscher and W. A. Taylor, Gravity model improvement using GEOS-3 altimetry (GEM 10A and 10B), EOS, Transactions, American Geophysical Union,, 59 (4), April 1978.

Marsh, James G., Marie-Jeanne Munteanu, Thomas V. Martin, John J. McCarthy and Phyllis S. Chovitz, Estimation of the mean sea surface in the North Atlantic using GEOS-3 altimetry data, EOS, Transactions, American Geophysical Union, 59 (4), April 1978.

Marsh, James G., Thomas V. Martin, John J. McCarthy and Phyllis J. Chovitz, Estimation of mean sea surfaces in the North Atlantic the Pacific and the Indian Ocean using GEOS-3 altimeter data, Second International Symposium on the Uses of Artificial Satellites for Geodesy and Geodynamics, National Technical University, Athens, in press.

Mather, R. S., The Analysis of GEOS-3 altimeter data in the Tasman and Coral Seas, National Aeronautics and Space Administration Technical Memorandum 78032, November 1977.

Rapp, Richard H., The Geoid: Definition and determination, EOS, Transactions, American Geophysical Union, 55 (3), 118-126, March 1974.

Rapp, Richard H., Gravity anomaly and geoid undulation results in local areas from GEOS-3 altimeter data, EOS, Transactions, American Geophysical Union, 59 (4), April 1978.

Rapp, Richard H., Altimeter derived gravity anomalies, Second International Symposium on the Uses of Artificial Satellites for Geodesy and Geodynamics, National Technical University, Athens, in press.

Reigber, Ch. and R. Rummell, 12th order harmonics from resonant perturbations of satellites using non-homogeneous weight functions, Second International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics, National Technical University, Athens, in press.

Seppelin, Thomas, Department of Defense World Geodetic System 1972, The Canadian Surveyor 28 (5), 496-605, December 1974.

Schwiderski, Ernst, Global ocean tides. Part I: a detailed hydrodynamical interpolation model, Naval Surface Weapons Center TR-3866, in press.

Vonbun, F. O., Probing the earth's gravity field by means of satellite-to-satellite tracking, Phil. Trans. R. Soc. Lond. 284, 475-483, 1977.

Vonbun, F. O. and W. D. Kahn, Gravity anomalies determined from tracking the Apollo-Soyuz, EOS, Transactions, American Geophysical Union, 59 (4) April 1978.

Wagner, C., Accuracy of resonance gravity coefficients, Planetary and Space Science, in press.

Wagner, C., F. Lerch, J. Brownd, and J. Richardson, Improvement in the geopotential derived from satellite and surface data (GEM 7 and 8), Journal of Geophysical Research 82 (5), February 1977.

Yionoulis, S. M., A. Eisner, V. L. Pisacane, H. D. Black, and L. L. Pryor, GEOS-3 ocean geoid investigation, NASA Contractor Report 141440, May 1978.

# Recent Advances in Analytical Satellite Theory

E. M. Gaposchkin
Smithsonian Astrophysical Observatory
Cambridge, Massachusetts 02138

Abstract. Recent work on analytical satellite-perturbation theory has involved the completion of a revision to 4th order for zonal harmonics, the addition of a treatment for ocean tides, an extension of the treatment for the noninertial reference system, and the completion of a theory for direct solar-radiation pressure and earth-albedo pressure. Combined with a theory for tesseral-harmonics, lunisolar, and body-tide perturbations, these formulations provide a comprehensive orbit-computation program. Detailed comparisons with numerical integration and observations are presented to assess the accuracy of each theoretical development.

## Introduction

Attempts to find analytical descriptions of satellite motion predate the age of artificial earth satellites. The considerable body of theory in existence at that time became the foundation on which to build solutions to specific problems arising from the desire to calculate trajectories of artificial satellites. The celebrated volume 64 of the Astronomical Journal can be considered the beginning of the field of celestial mechanics for satellite geodesy. Three papers in particular appeared in that issue [Brouwer, 1959; Garfinkel, 1959; Kozai, 1959]; of these authors, Garfinkel and Kozai are contributing to the field today. Those articles all address essentially the same problem — perturbations due to $J_2$, $J_3$, and $J_4$ — and are significant in two respects: That problem is still receiving attention, and the methods employed then are still in use. One part of this paper is devoted to what is called the main problem of satellite theory; some of the present results will be reviewed. The method used then by Kozai to integrate the Lagrange planetary equations is almost commonplace now. The device of canonical transformations employed by Brouwer and Garfinkel is now used almost exclusively when higher order solutions are developed for specific problems. The von Ziepel method of finding a canonical transformation led to a significant generalization by Hori [1966, 1973] in the method of Lie Series, a method that has become the sine qua non of modern methods. Just because the beginning work was similar to the present, however, does not mean that no progress has been made. On the contrary, enormous strides have been taken and considerable work probably remains.

The main motivation for developing elaborate analytical descriptions for satellite motion is to aid in understanding the forces causing the motion. A second practical reason is the potential economy available for certain applications. As our understanding of the driving forces increases and as the observational accuracy improves, the requirements for accuracy become correspondingly stringent. An accuracy goal — say 1 cm — is easily set, but how to verify that an accuracy has been achieved is not so clear.

In 1967, I presented a paper similar to this one [Gaposchkin, 1968], describing a philosophy of how to develop, combine, and verify a complete satellite theory and outlining the status at that time. Basically, the method consisted of two steps — a comparison with numerical integration to verify that the mathematical problem had been properly solved, followed by a comparison with observations to verify that the mathematical problem was an adequate description of the physical problem. I am not so optimistic now. First, preserving the accuracy of numerical-integration methods for long time periods poses considerable problems [Velez, 1975; Balmino, 1975]. In this context, a long period is measured in terms of the number of revolutions of the fastest body, normally the satellite; integrations for more than 1000 revolutions of anything are difficult and time consuming. Although the theory can be used to test the integration, rather than the other way around, even the use of numerical integration to test short-period perturbations has proved difficult. Second, verifying the theory by means of data analysis presents significant problems. Many cases occur in which several forces have similar qualitative orbital effects that tend to cancel. A good example is shown by lunar and solar perturbations, where the direct effects, the effects due to tidal deformation, and the effects due to precession and nutation all have the same origin and produce perturbations with the same spectral character; in some cases, these effects add, and in others, they cancel. A third aspect of this verification process is the necessity to know certain physical quantities. While this is not a problem in the comparison with numerical integration, it is critical in the analysis of observations. In reality, then, the verification must be combined with the determination of physical quantities. Finally, although individual components of the satellite theory can be verified, in practical terms, the theory must all fit together and several interactions should be taken into account. Indeed, there are significant difficulties, but the situation is far from hopeless, and in the following, I describe where we are today in the theory, the verification, and the data analysis.

## Selection of Variables

The first thing to select is a set of variables to be used for the analysis. The most popular set is the Kepler elements ($\omega$ = argument of perigee, $\Omega$ = right ascension of the ascending node, I = inclination, e = eccentricity, M = mean anomaly, and a = semimajor axis). In practice, a is obtained from the mean motion n according to Kepler's third law $n^2a^3$ = constant, and n then becomes the sixth variable. This is done for the practical reason that n is the more easily and more accurately determined quantity, and a becomes a derived quantity. This set of variables has the conceptual

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

advantage that each has a simple physical meaning. Kepler elements have the one drawback that for zero eccentricity or inclination, they become degenerate. This degeneracy is also a problem for small eccentricities and inclinations, in that these variables become highly correlated in any adjustment procedure that attempts to determine them by a statistical process using observations. Finally, this degeneracy presents analytical difficulties; some series expansions become much longer than they would if another set of variables were chosen. Nevertheless, Kepler elements continue to be the most widely used both for analytical work and for reporting results.

For the reason cited above, some fundamental analysis is now done with another set of variables; a list of candidate variables was given by Gaposchkin [1973]. For the following, the inclination degeneracy has not been addressed, but the more important and relevant degeneracy in eccentricity has been overcome by using the nonsingular variables

$$\xi = e \cos \omega, \quad \eta = e \sin \omega, \quad M + \omega, \quad \Omega,$$

$$L^2 = a, \quad H = G \cos I \quad ,$$

where

$$G^2 = L^2(1 - e^2) = L^2(1 - \xi^2 - \eta^2) \quad .$$

In that set of variables,* the Kepler elements are recognizable, along with the Delaunay variables $\ell = M$, $g = \omega$, $h = \Omega$, L, G, and H. The Delaunay variables are used to derive the long-period and secular perturbations and are then formally combined into the set of nonsingular variables for numerical evaluation. It can be shown that these nonsingular variables satisfy the d'Alembert characteristic with respect to eccentricity, although not with respect to sin I. The d'Alembert characteristic is that the lowest powers of e and sin I of the coefficient $\begin{matrix}\sin \\ \cos\end{matrix}(kM + q\omega)$ in the trigonometric series are $\alpha$ and $\beta$, respectively, where

$$\alpha = |k - q| \pmod 2 \quad ,$$

$$\beta = |q| \pmod 2 \quad .$$

Therefore, any expression satisfying the d'Alembert characteristic is well behaved as e (or sin I) goes to zero. If it is necessary to compute perturbations for small inclinations, then the variable $M + \omega + \Omega$ satisfies the d'Alembert characteristic and is suitable.

For the main problem, we have formally obtained expressions for the nonsingular variables analytically. However, if the computer word has sufficient accuracy, then the nonsingular variables can be calculated numerically, and the well-conditioned properties of these variables can be realized. Therefore, it is entirely adequate, if necessary, to derive perturbations in Kepler elements and numerically combine then into nonsingular variables for calculating an ephemeris. Of course, a unified treatment in nonsingular variables would be preferable, although it is not always the most convenient solution. We can then develop perturbations in the most convenient set of variables for the particular problem and then unify the variables at the calculation stage.

Mean elements are a key to the construction of an analytical theory. In the framework of perturbation theory, the mean elements are the zero-order reference for the development. They become the constants of integration and therefore play a similar role to the initial conditions in solving differential equations. Each perturbation theory has implicit in it a definition of mean elements, and generally the relation between mean elements and the initial conditions does not receive any attention.

In the present situation, several perturbation theories are employed in the same computation, with the mean elements being empirically obtained from observations. This situation is rigorously correct when a single perturbation theory is employed, provided suitable partial derivatives are available. In the general case, the mean elements must have the same formal definition, to the accuracy of the theory.

A second aspect of mean elements concerns their constancy. If they were truly constants of the motion, if we knew all the numerical constants entering the theory, and if our observations were without error, then the mean elements for a satellite would be the same at different epochs. Any variations in them would have to be due to errors in the theory, errors in some numerical constants, or errors in the data. Assuming that we can control errors in the theory and the data, then the variations in the mean elements can be used to get information on the numerical constants (i.e., the physical parameters) entering the orbit theory. In fact, this has been the basis of much of the geodetic and geophysical information obtained from satellite data.

## Analytical Methods

Basically, three methods are used to develop satellite theory. To begin with, we cannot hope to find exact closed-form solutions to the equations of motion and must seek approximate solutions by some perturbation method. The simplest method is to recast the equations of motion in our chosen set of variables, for example, Kepler elements. This results in a rigorously equivalent set of six coupled first-order differential equations, called the Lagrange planetary equations. For some perturbations (for example, for tesseral harmonics), we can expand these variables around a reference orbit, say a precessing Keplerian ellipse, by Fourier series. The equations can then be treated as a forced harmonic oscillator with constant coefficients, and this approximation to the equations of motion can be integrated term by term. The approximation can be further improved by using this first-order solution as the reference orbit, expanding it in Fourier series, and so on. Beyond the second order, however, the method is usually replaced by the more general one of canonical transformation.

* Expressions are often written in this expanded set of variables, but formally, this is only a notational convenience.

The theory of canonical transformation goes back to the last century, when it was developed to solve mechanics problems. It uses more fundamental properties of dynamical systems, which are beyond the scope of this discussion. Suffice it to say that it is rigorously equivalent and can be either easier or more difficult than the use of Lagrange planetary equations. Originally, canonical-transformation theory was thought to require the use of canonical variables, but the generalization by Hori [1966] proved that it can be employed for any variable provided the basic equations can be solved. In the theory of canonical transformation, the key is to find a solution to a single partial-differential equation. Although in general this is difficult, a method has been developed for the satellite main problem that will automatically find a suitable approximation to the equation. Therefore, this method is applicable to obtaining a solution to any order and can be automated, to some extent, on a computer. Such an approach is in general use now for higher order solutions.

When the first two methods are inadequate for some reason, a third one must be used. This is called a semianalytical solution, in that part of the solution (integration) can be accomplished analytically, while the remaining part must necessarily be done numerically. Recourse to this method is required when closed-form expressions for the force function are complicated or impossible to find. Examples are lunisolar perturbations when the analytical description of the moon's motion to suitable accuracy would be prohibitive and radiation-pressure perturbations when the shadow function must be obtained numerically. This method integrates over the short-period perturbation analytically and then integrates the averaged force function numerically to obtain the perturbations. In this case, the numerical integration can take relatively large time steps and is therefore economical. It also conveniently separates long-period and short-period perturbations.

## Current Status

A third-order solution to the main problem — i.e., for the motion of a satellite in the geopotential containing only $J_2$, $J_3$, and $J_4$ — has been obtained by Kinoshita [1977]. Third-order periodic perturbations with fourth-order secular perturbations are derived by the method due to Hori [1966]. All quantities are expanded into power series in the eccentricity, but the solution is closed with respect to inclination. A comparison with results obtained by numerical integration of the equations of motion indicates that the solution can predict the position of a close-earth satellite with an accuracy of better than 1 cm over a period of 1 month. For this check, a special-purpose Taylor-type integrator is adopted, in which the positions and velocities are expanded into a power series of time and the coefficients of the series are determined by recurrence formulas [Rabe, 1961; Deprit and Zahar, 1966].

Periodic perturbations due to tesseral harmonics are a first-order linear theory based on integration of the Lagrange planetary equations as developed by Kaula [1966]. The theory also includes the interaction with $J_2$ and second-order interactions with the mean motion through Kepler's third law. Although the theory is essentially that of Kaula, the details of the calculation have been considerably revised with the inclination function as described by Gaposchkin [1973] or Kinoshita [1977] and the eccentricity function calculated as Hansen coefficients.

The lunisolar perturbations in satellite motion are obtained by a semianalytical method [Kozai, 1973]. The disturbing function is expressed by the orbital elements of the satellite and the geocentric polar coordinates of the moon and the sun. These coordinates are obtained by using the larger terms in Brown's theory [United States Naval Observatory, 1954]: 26 terms in longitude, 14 in latitude, and 12 in the parallax. The secular and long-period perturbations are derived by numerical integration, and the short-period perturbations, analytically. Perturbations due to the solid body tide can be included in the same way.

The orbital elements of a close-earth satellite have perturbations caused by the motion of the equatorial plane of the earth due to precession and nutation. Kozai and Kinoshita [1973] derived exact differential equations for the perturbations of satellite orbital elements due to the motion of the earth's equatorial plane and solved them to second order in precession. This theory, in fact, defines the reference system used for satellite motion, in which the inclination and the argument of perigee are referred to the equator of date and the longitude of the ascending node is measured from a fixed point along a fixed plane and then along the equator of date.

The perturbations of a spherical satellite due to direct solar radiation are computed according to a semianalytical algorithm due to Aksnes [1976], which is based on expressions derived by Kozai [1961]. Through some simple modifications, the algorithm also holds when $e = 0$ and $i = 0$. The perturbations are obtained by summing over the sunlit segment of the satellite's orbit during each revolution or partial revolution. The end points of the segment are evaluated numerically once per revolution. Testing of the algorithm is done by means of numerical integration of the equations of motion and through comparisons with observations of the balloon satellite 1963 30D during a 200-day interval.

The perturbations due to solar radiation diffusely reflected from the earth have been treated by Lautman [1977a], who used a semianalytical method based on the assumptions that the satellite is spherically symmetric and that solar radiation is reflected from the earth according to Lambert's law with uniform albedo. Expressions for the radiation-pressure force are developed into series in true anomaly v. The perturbations within a given revolution are obtained analytically by integrating with respect to v, while holding all slowly varying quantities constant. The long-term perturbations are then obtained by summing the net perturbations at the end of each revolution. This theory has been extended [Lautman, 1977b] to account for the increasing reflectivity of the earth toward the poles; the earth's albedo is assumed to have a latitude dependence given by $a = a_0 + a_2 \sin^2 \phi$.

The short-period perturbations due to air drag are computed according to a computation due to Sehnal and Mills [1966]. The density function of the earth's atmosphere includes the effect of the atmospheric bulge described by Jacchia. The method of solution is, in essence, numerical construction of the disturbing function. The secular acceleration for geodetic satellites is more accurately given by analysis of the data. The short-period perturbation is usually less than 1 cm per revolution, and this development is not currently used.

The analysis of ocean tides is done along the following lines. Recall that in calculating lunisolar perturbations, the body-tide potential is easily included by introducing the Love number $k_\ell$. An alternative way to describe the lunisolar gravitational potential is essentially given by Doodson [1921]. In this case, the potential at the earth's surface can be written

$$\frac{U}{g} = \mathrm{Re} \sum_{\ell ms} F_{\ell m} D_{\ell ms} (i)^{\ell+m} (-1)^{[(\ell+m+1)/2]} \times \bar{P}_{\ell m}(\phi)\, e^{i(\sigma_s(t)+m\lambda)} ,$$

where Re {x} is the real part of x, $i^2 = -1$, [x] is the integer part of x, $F_{\ell m}$ are numerical factors, $D_{\ell ms}$ are the s coefficients as determined by Doodson by Fourier analysis, $\bar{P}_{\ell m}$ are fully normalized associated Legendre functions, and $\sigma_s(t)$ is the time variation given in terms of the six chosen variables.

Cartwright and Edden [1973] have provided values of $D_{\ell ms}$ based on modern values of the solar and lunar ephemerides: $F_{20} = -12.020364$ cm, $F_{21} = F_{12} = 13.879920$ cm. Expression (1) is the potential at the surface of the earth, and we can continue this potential analytically to satellite altitudes as

$$\frac{U}{g} = \mathrm{Re} \sum_{\ell ms} \left(\frac{r}{a_e}\right)^{\ell} F_{\ell m} D_{\ell ms} (i)^{\ell+m} (-1)^{[(\ell+m+1)/2]} \times \bar{P}_{\ell m}(\phi)\, e^{i(\sigma_s(t)+m\lambda)}$$

and then develop satellite perturbations due to the sun and the moon. The body tide can be defined in terms of the complex Love number

$$k_\ell^* = k_\ell \left(1 - i \frac{\mu}{2Q}\right)$$

[Munk and MacDonald, 1960, p. 153]. Considering the deformation, $U_{tide}$ can be analytically continued to satellite altitudes as

$$U_{tide} = g\, \mathrm{Re} \sum_{\ell ms} \left(\frac{a_e}{r}\right)^{\ell+1} k_\ell^* F_{\ell m} D_{\ell ms} \times (i)^{\ell+m} (-1)^{[(\ell+m+1)/2]} \bar{P}_{\ell m}(\phi) \times e^{i(\sigma_s(t)+m\lambda)} .$$

Similarly, each component of an ocean tide of height $\zeta_s$ (for driving function $D_{\ell ms}$ for argument $\sigma_s$) can be expressed in spherical harmonics $\bar{\mathscr{C}}_{\ell ms}$, where $\bar{\mathscr{C}}_{\ell ms}$ is the fully normalized complex representation of the ocean tide. If the ocean tide is expressed as a surface layer, then the external potential at satellite heights can be written

$$U_{ocean} = 4\pi G \rho_\omega a_e\, \mathrm{Re} \sum_{\ell ms} \frac{1 + k'_\ell}{2\ell + 1} \left(\frac{a_e}{r}\right)^{\ell+1} \times (i)^{\ell+m} (-1)^{[(\ell+m+1)/2]} \bar{\mathscr{C}}_{\ell ms} \times \bar{P}_{\ell m}(\phi)\, e^{i(\sigma_s(t)+m\lambda)} . \quad (1)$$

As noted by many [Gaposchkin, 1973; Lambeck et al., 1974; Felsentreger et al., 1976; Goad and Douglas, 1978], the ocean and body tides enter the potential in exactly the same form, and by satellite analysis, we can sense only the linear combinations,

$$g F_{\ell m} D_{\ell ms} k_\ell^{(s)} + 4\pi G \rho_\omega a_e \frac{1 + k_\ell^{\prime(s)}}{2\ell + 1} \bar{\mathscr{C}}_{\ell ms} ,$$

where $k_\ell$ and $k'_\ell$ are assumed to depend on the frequency of the argument $\sigma_s$. Thus, we are obliged to assume one tide to determine the other.

For frequencies far from resonance ($23^h53^m0^s.04$), $k_\ell$ and presumably $k'_\ell$ can be taken from seismic models [Longman, 1962, 1963; Farrell, 1972]. Both Jeffreys and Vicente [1957a,b] and Molodensky [1961] pointed out that nearly diurnal earth tides should be amplified because of the existence of a resonance between the elastic mantle and the liquid core. In Table 1, the variation in $k_2$ predicted by Molodensky's Model II is given.

In any event, the expressions for the potentials [eqs. (2) through (4)] can easily be used to obtain satellite perturbations, as the arguments are given as linear functions of time, and the Lagrange planetary equations can be integrated directly as a forced harmonic oscillator.

## Accuracy Assessment

The various perturbation theories described above are all included in a general-purpose orbit-determination program that accepts observations of direction, range difference, and range rate. Each individual theory has been tested, and now we wish to study the accuracy of the combined theory. For this purpose, we used a numerical integration program to calculate simulated (errorless) data. The general-purpose integration program, developed by Krogh [1973], is an Adams-type integrator that has the option of variable or fixed step size once the integration has been started. The variable-step-size option uses a desired accuracy as input. The program performs the integration in coordinates $(x, y, z, \dot{x}, \dot{y}, \dot{z})$ in single precision on a CDC 6000 computer that has 14-decimal-digit accuracy. The force package allows use of an arbitrary gravity field represented in spherical harmonics, moon and sun positions (we use the same routines as the analytical theory), radiation pressure, and a drag model based on the same physical assumptions as the analytical developments described earlier.

TABLE 1. Values of $k_2$ for Tidal Terms Near Resonance (Based on Molodensky, 1961)

| Doodson | Darwin | $D_{\ell ms}$ | $\beta$ | $k_2$ | $\sigma$ (°/hr) |
|---|---|---|---|---|---|
| 255.555 | M2 | 0.90809 | 0.84 | 0.3015 | 28.984104 |
| 185.555 | $00_1$ | -0.01624 | -4.54 | 0.3025 | 16.139102 |
| 167.555 | $\phi_1$ | -0.00755 | -124.06 | 0.3277 | 15.123206 |
| 166.554 | $\psi_1$ | -0.00422 | -728.10 | 0.4551 | 15.082135 |
| 165.565 | | -0.07186 | 206.19 | 0.2580 | 15.043275 |
| 165.555 | K1 | -0.53011 | 192.30 | 0.2609 | 15.041069 |
| 165.545 | | 0.01051 | 180.17 | 0.2635 | 15.038862 |
| 163.555 | P1 | 0.17543 | 55.49 | 0.2898 | 14.958931 |
| 145.555 | O1 | 0.37694 | 6.79 | 0.3001 | 13.943036 |
| 056.554 | SA | 0.01156 | 1.70 | 0.3015 | 0.041067 |

$$\beta = \frac{41.87}{0.2136 - 100[(\sigma + w)/\sigma]} + 1.9 \quad ,$$

$$k_2 = 0.3015 + 0.2109 \times 10^{-3} \frac{\sigma(\sigma + 2w)}{w^2} \beta \quad ,$$

where $\sigma = \theta - w$, $\sigma$ being the external driving frequency and $\theta$ being the earth-fixed driving frequency.

We originally planned to use simulated data to verify the short-period perturbations, largely because we believed that the accuracy of a general-purpose numerical integration could not be relied on, but secondarily to conserve computer time. For the short-period perturbations, simulated range observations were computed for subsets of the forces. The simulated observations were used in the general orbit-determination program, and the mean elements were computed by least squares, thus avoiding the problem of explicitly relating the initial conditions of the numerical integration and the mean elements. This testing was done for satellites in orbits similar to those of Geos 1 and Geos 3. Our main interest was the difference in eccentricity between the two satellites, as one of our concerns was the validity of solutions for small eccentricity. The results of this testing are given in Table 2.

During the first phase of the study, the variable step size was exercised in the numerical integration; the accuracy sought was $10^{-4}$ cm/sec. Because of the excellent agreement of the theory for the main problem ($J_2$, $J_3$, $J_4$), the integration ephemeris was thought to be of sufficient accuracy. However, when working with a combined tesseral- and zonal-harmonics field, the accuracy never was better than 19 cm root mean square. The tesseral-harmonics theory alone (with $J_2 = 0$) gave perfect agreement.

Since their amplitudes are approximately 1 m, the $J_2$ interaction terms could be suspect, but they have been carefully checked by me and by H. Kinoshita. Furthermore, if the interaction terms need revision, we would expect that adding more harmonics would increase the error. In fact, increasing the field from $\bar{C}_{22}$, $\bar{S}_{22}$ to the complete field through $\bar{C}_{44}$, $\bar{S}_{44}$ (i.e., eight times the number of coefficients) only decreased the goodness of fit, from 19 to 22 cm.

At that point, the accuracy of the integrator was questioned, and some simple tests were made with fixed-step-size runs. In all cases, the trajectories differed by more than 1 m. For example, the difference between the variable step size and the fixed 0.25-sec step size can be written rather well as

TABLE 2. Test of Orbit Determination

| Force Function | Interval (days) | Geos 1 $\sigma$ (m) | Geos 3 $\sigma$ (m) |
|---|---|---|---|
| $J_2$ | 1 | 0.01 | 0.01 |
| $J_2$, $J_3$, $J_4$ | 1 | 0.01 | 0.01 |
| $J_2$, $J_3$, $J_4$, $J_5$ | 1 | 0.04 | |
| $J_2$, $J_3$, $J_4$ | 12 | | 0.06 |
| $J_2$, $\bar{C}_{22}$, $\bar{S}_{22}$ | 1 | 0.19 | |
| $\bar{C}_{22}$, $\bar{S}_{22}$ | 1 | 0.01 | |
| $J_2$, $J_3$, $J_4$, $\bar{C}_{22} - \bar{S}_{44}$ | 1 | 0.22* | |
| $J_2$, $J_3$, $J_4$, $\bar{C}_{22} - \bar{S}_{44}$ | 6 | 0.36 | |
| $J_2$ + sun + moon + body tide | 1† | 0.12 | |
| Sun + moon + body tide | 1† | 0.12 | |
| $J_2 - \bar{S}_{44}$ | 1 | 0.25‡ | |
| $J_2 - \bar{S}_{44}$ | 1 | 0.24** | |

* The integrator used a variable step size.

† The dynamical effects of the moving equator are not included in either the numerical integrator or the analytical theory.

‡ The integrator used a fixed step size of 0.5 sec.

** The integrator used a fixed step size of 0.25 sec.

$$\epsilon = 250\ t^3\ \text{cm} \quad ,$$

where t is in days. Two trajectories that differ by more than 2.5 m can be fit to an accuracy of 22 cm. Clearly, the process of fitting mean elements was able to adjust or compensate for the difference between the two trajectories. This difference is due to errors in the numerical-integration algorithm, and at this point, it is not known what is the true trajectory. The nature of the integration error is that much of it can be absorbed in the mean elements of the analytical theory. In fact, this should not be surprising. An along-track acceleration can be modeled by a change in the mean motion. In any event, what can be said is that short-period perturbations can be modeled with an accuracy of 22 cm. The limiting factor in this assessment may be either the theoretical expressions or the trajectory; further investigation is warranted.

A similar situation obtains in the case of lunisolar perturbations, where the principal effects are of much longer period than 1 day and cannot be evaluated with such a short interval. The short-period terms have an amplitude of about 1 m for Geos-type satellites. From the tests outlined in Table 2, these terms are computed to 12 cm, or about 10%. We could expect to have more than two digits even from the simplest theory. The lunisolar perturbations and the tesseral harmonics share the common factor that the disturbing function explicity contains the time. This is not true for the zonal harmonics, where the disturbing function depends only on position. This factor may limit the accuracy of either the numerical integration or the analytical solutions.

## Analysis of the Data

Analysis of tracking data is performed primarily to obtain geophysical information; a secondary consideration is verification of the models. (In the latter, I do not include the determination of numerical parameters, which fall under the primary objective.) In the results given here, both considerations are important. The first set of data will concern the Lageos satellite, which has exhibited some small and unexpected orbital changes. The second set of data is on the Geos 1 and Geos 2 satellites, which provide useful information about ocean tides and core-mantle resonance. Since in both cases the results cannot be unambiguously checked, they are open to interpretation and are potentially subject to errors or oversights in the very complex software packages used in analyzing tracking data.

The Lageos satellite was designed to be a stable platform, with a well-defined orbit. Its orbital characteristics, given in Table 3, were carefully chosen in several respects: The orbit is sufficiently high that the effects of the anomalous gravity field are reduced, minimizing the uncertainty in ephemeris calculation. The mean motion minimizes any resonances with the gravity field. The very small area-to-mass ratio reduces the size of nongravitational perturbations (solar pressure, albedo pressure, and atmospheric drag). The symmetrical cross section allows

TABLE 3. Orbital Characteristics of Lageos

| | |
|---|---|
| a = | $12.270 \times 10^6$ m |
| e = | 0.0046 |
| I = | 109°.86 |
| n = | 6.3866 rev/day |
| A/m = | 0.0069 $cm^2/g$ |

greatest ease in attempting to model these forces. Most important is that Lageos is equipped with cube-corner reflectors, which enables precision laser ranging to be done. Orbits for the first 586 days of the satellite's lifetime have been computed, from which it can be immediately seen that indeed the overall objectives have been met and the orbit is known very well.

Figure 1 is a plot of the semimajor axis of Lageos. These are independently determined values, each based on 8 days of tracking data. The perturbations described above have all been included, and the remarkable 50-cm decrease in the semimajor axis is an unmodeled effect. A candidate source for this effect is the use of an inappropriate value of A/m in calculating radiation-pressure and albedo perturbations. Over this interval, the radiation pressure contributed a 20-cm decrease in the semimajor axis, and the albedo, an additional 10 cm. To compensate for the full 50-cm decrease, however, A/m would have to be increased by a factor of 2.7, far outside the plausable uncertainty in A/m; in addition, increases in the specular diffuse coefficient, which ranges from 1 to 1.44, and in the solar constant would be necessary. A second candidate is drag from the neutral atmosphere. The equivalent drag is enormous, amounting to that occurring at altitudes of 2000 km, but adding significant drag will reduce the good agreement in other orbital elements. In fact, most atmospheric models do not attempt to model drag above 2000 km, and our knowledge of atmospheric drag at 6000-km altitudes is extremely limited. Some sort of charge buildup and interaction with the magnetosphere is also possible, and a model of that should be explored. However, we would expect to see some change in $\dot{a}$ corresponding to changing magnetospheric conditions. Nonisotropic radiation of heat is another possibility. Again, we would expect to see a change as the spacecraft spins down. A final possibility is the Poynting Robertson effect [Robertson, 1937]. This effect, which is viewed in celestial mechanics as an aberration, is due to the conservation of momentum of photons reradiated from the satellite both along and opposite the motion. The effect is always along track and has the same effect as drag; it can be calculated based on the incoming flux or in terms of the temperature at which the photons are reradiated. This distinction is important if radiation from the earth contributes significantly to the temperature of the satellite.

Figure 2 presents the variation in inclination for Lageos over the same period. Owing to the essentially equatorial distribution of laser tracking stations, the satellite is not observed at maximum latitude and the inclination is not so well determined as possible. With improved

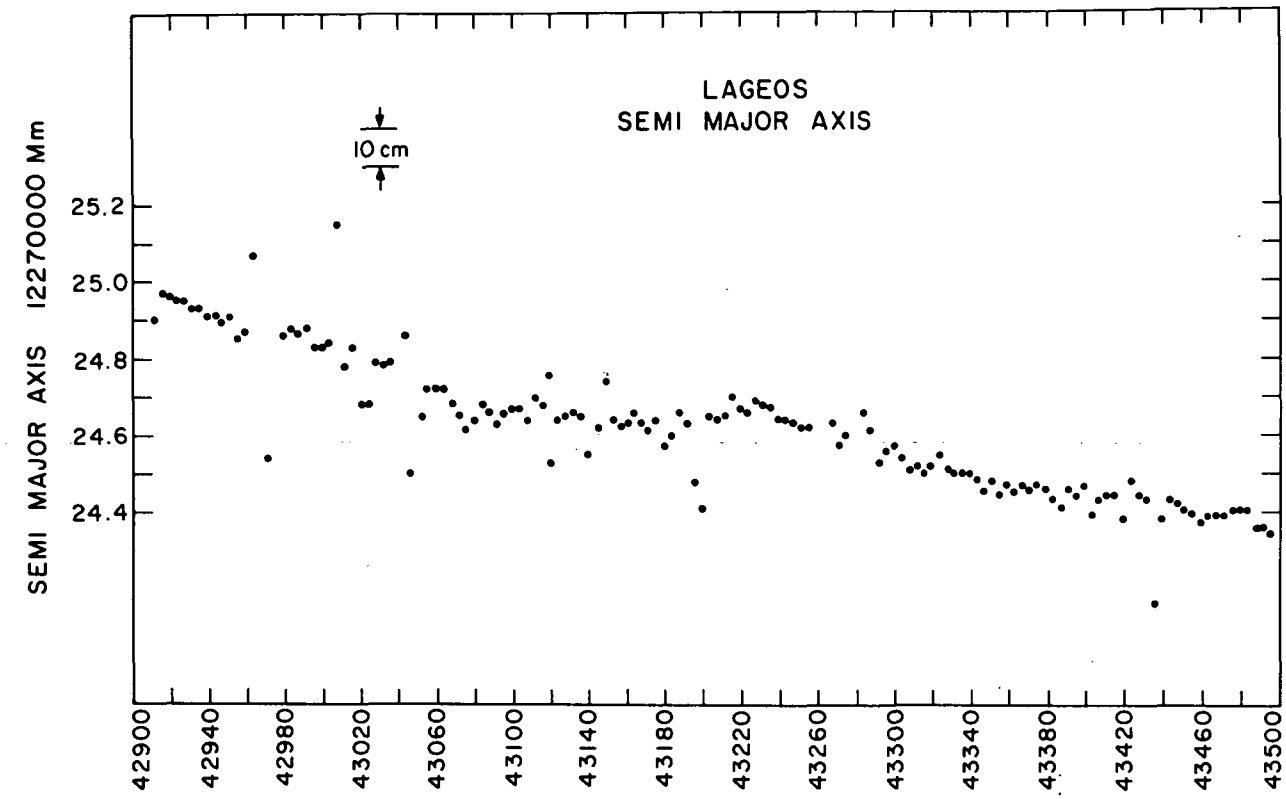



Fig. 1. Semimajor axis of Lageos.

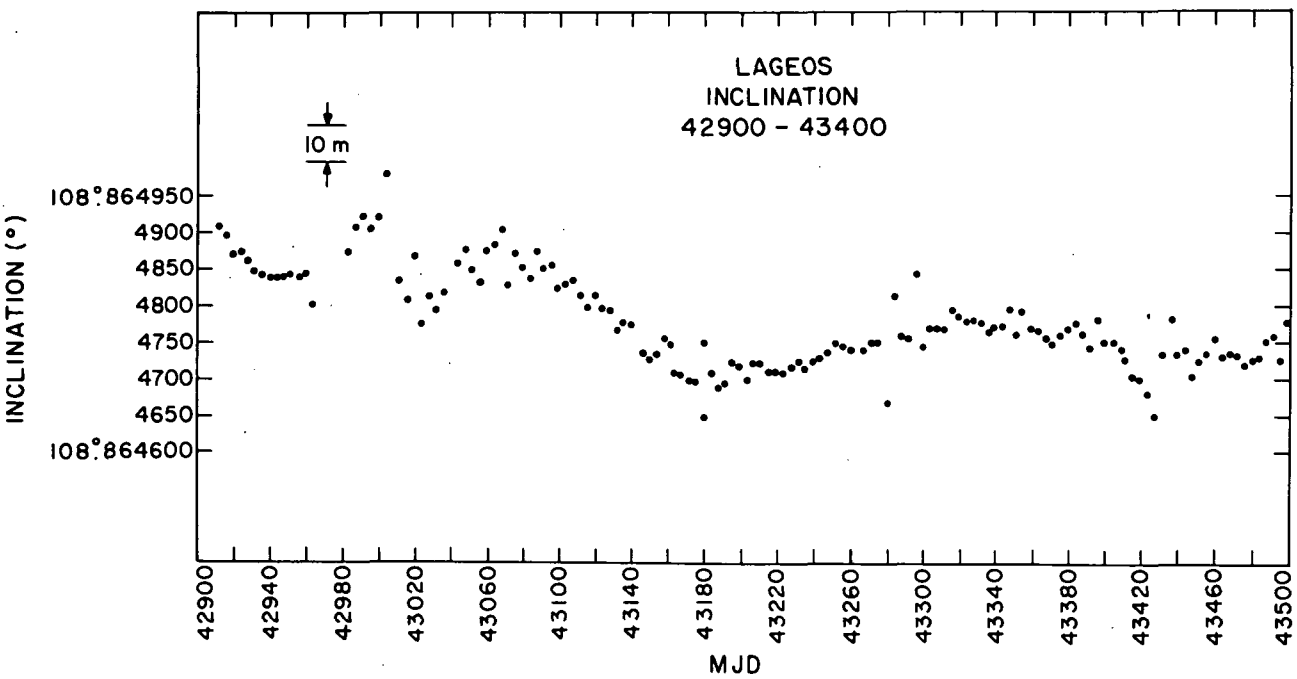



Fig. 2. Inclination of Lageos.

models, some coming from this work, the inclination and other orbital elements will be computed with greater accuracy. However, even now, some useful information can be derived on the tides; this is discussed in the following.

The study of tides is one of the fruitful applications of long-period perturbations. As Lambeck et al. [1974] showed, ocean tides as well as body tides give rise to sensible perturbations in close-earth satellites. Currently, not enough information exists about the tides to make a definitive analysis, and preliminary results are given here to illustrate some of the difficulties. Though sensible, the perturbations are small and

TABLE 4. Data Analyzed

| | Geos 1 | Geos 2 |
|---|---|---|
| Inclination | 59°.383 | 105°.804 |
| $\dot{\omega}$ | 0°.652/day | -1°.620/day |
| $\dot{\Omega}$ | -2°.246/day | 1°.400/day |
| First time | 41501 (MJD) | 41501 (MJD) |
| Last time | 42239 (MJD) | 42278 (MJD) |
| Interval | 738 days | 777 days |

a long time series is needed even to hope to get good numerical values. Other physical effects, at or near the tidal perturbation frequencies, also come into play. For example, some effects basically cannot be separated; also, there are some very fundamental gaps in our understanding.

Two series of orbits have been computed on the Geos 1 and Geos 2 satellites, as summarized in Table 4. In analyzing the mean elements, a gravity field and a Love number $k_2$ are adopted, among other constants. We took $k_2 = 0.29$. The inclination of these series has been analyzed and interpreted as a perturbation due to ocean tides.

Individual tidal constituents can be isolated by the frequency. In performing harmonic analysis, we should also determine terms that should exist owing to other effects. For example, odd zonal harmonics of the gravity field will not be known without error, and hence a term in sin $\omega$ will arise in the data. Such a term could obviously be used to improve knowledge of the odd zonal harmonics. Also, we suspect that any analytical theory will have an error term that is called mixed secular, i.e., t cos $\omega$t, and such a term should be introduced.

The harmonic analysis was performed for the tidal constituents K2S + K2M, S2, K1S + K1M, and P1. The M2 and O1 tides could also be studied, but their periods are much shorter, 10 to 15 days. Table 5 gives the results of applying equation (1) to the observed variations in inclination. Also listed in Table 5 are the periods of the satellite perturbations, values for some constituents from terrestrial observations [Lambeck, 1975], and results of similar analysis by Felsentreger et al. [1976]. The fully normalized coefficients in complex form are given, together with the equivalent amplitude and phase, as is conventionally published for ease of comparison. In general, the complex tidal component is determined to about 1%.

The first thing to note is that for Geos 2, the rates of perigee $\dot{\omega}$ and node $\dot{\Omega}$ are almost equal and opposite. Therefore, with a relatively short time series, the terms in sin/cos $\Omega$ arising from K1S + K1M cannot easily be separated from the sin $\omega$ term arising from the uncertainty in the odd zonal harmonic. The first values in Table 5 for that constituent represent an attempt to obtain both quantities. The second line corresponds to not including a sin $\omega$ term in the harmonic analysis; this result is more plausible and agrees with Felsentreger et al.

A similar circumstance occurs with K2S + K2M, in which case, we adopt values not including a cos $2\omega$ term, which is equivalent to assuming no error in the even zonal harmonics.

The next point concerns the adopted value for the body tide, expressed in terms of the Love number $k_2$. Table 6 provides the change in the ocean tide's complex amplitude (real part only) corresponding to the change in Love number $k_2$ from 0.29 (as adopted) to 0.302, the value obtained from Farrell [1972]. These changes are comparable to the derived amplitudes and would have a significant effect on any interpretation made of the ocean-tide values. Furthermore, if we adopt Molodensky's theoretical model for the change of Love number $k_2$ with frequency, then the amplitudes of the derived ocean tide would be modified by the values given in the last column of Table 6.

Another aspect of the ocean tide is that each tidal constituent $\sigma_s$ gives rise to a complete spherical-harmonics description of the ocean tide and a set of harmonics $\mathscr{C}_{\ell ms}$. The satellite orbit

TABLE 5. Ocean Tides Determined from Satellite Data

| Tide | | Geos 1 | | | | Geos 2 | | | | Lageos | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Darwin | Doodson | Period (days) | $\mathscr{C}_{\ell m}$ (cm) | $C^+_{\ell m}$ (cm) | $\epsilon^+_{\ell m}$ (deg) | Period (days) | $\mathscr{C}_{\ell m}$ (cm) | $C^+_{\ell m}$ (cm) | $\epsilon^+_{\ell m}$ (deg) | Period (days) | $\mathscr{C}_{\ell m}$ (cm) | $C^+_{\ell m}$ (cm) | $\epsilon^+_{\ell m}$ (deg) |
| K2S + K2M | 275555 | 80 | 1.711 + 0.931i | 1.26 | 29 | 128 | 0.791 - 2.235i | 1.53 | 289 | | | | |
| | | | | | | | 0.598 + 0.536i | 0.52 | 42 | | | | |
| S2 | 273555 | 56 | 2.660 - 0.226i | 1.72 | 355 | 433 | 2.817 + 2.606i | 2.48 | 319 | 280 | -0.321 + 2.588i | 1.683 | 97 |
| K1S + K1M | 165555 | 160 | 2.130 + 3.676i | 5.49 | 60 | 257 | 13.260 - 7.389i | | | | | | |
| | | | | | | | 5.30 - 2.44i | 7.54 | 335 | | | | |
| P1 | 163555 | 85 | -4.779 + 1.376i | 6.42 | 164 | 631 | 1.707 + 1.128i | 2.63 | 33 | 221 | -0.349 + 2.644i | 3.43 | 98 |

| | | Lambeck [1975] | | Felsentreger et al. [1976] Geos 1 | | Geos 2 | |
|---|---|---|---|---|---|---|---|
| | | $C^+_{\ell m}$ (cm) | $\epsilon^+_{\ell m}$ (deg) | $C^+_{\ell m}$ (cm) | $\epsilon^+_{\ell m}$ (deg) | $C^+_{\ell m}$ (cm) | $\epsilon^+_{\ell m}$ (deg) |
| S2 | 273555 | 3.59* | 327 | 1.7 | 350 | 1.0 | 62 |
| K1S + K1M | 165555 | 3.3 | 318 | 8.8 | 345 | 5.7 | 334 |
| P1 | 163555 | 0.7 | 140 | 5.0 | 178 | 4.9 | 127 |

* Includes atmospheric tide.

TABLE 6. Equivalent Ocean Tide

| Tide | $D_{\ell m}$ | $\Delta C_{\ell m}$ (cm) | $k_2$ | $\Delta C_{\ell m}$ (cm) |
|---|---|---|---|---|
| M2 | 0.90809 | 1.908 | 0.3015 | 1.8288 |
| K2 | 0.11498 | 0.242 | 0.3015 | 0.2316 |
| S2 | 0.42248 | .888 | 0.3015 | 0.8508 |
| N1 | -.53011 | -1.114 | .2609 | 2.7014 |
| P1 | .17543 | .369 | .2898 | -0.0061 |
| O1 | .37694 | .792 | .3001 | 0.6667 |

$$\Delta k_\ell = 0.302 - 0.29 = 0.012$$

responds to a linear combination of these $\bar{\mathscr{C}}_{\ell ms}$, where here we have determined the combined effect, i.e., a lumped coefficient. To illustrate this, consider K2S + K2M. From Geos 1, we have the value $\bar{\mathscr{C}}_{22} = 1.711 + 0.931i$ and from Geos 2, $\bar{\mathscr{C}}_{22} = 0.598 + 0.536i$. If we assume that the only other tidal term is $\bar{\mathscr{C}}_{42}$, we can obtain the two linear equations in two unknowns. It turns out that the factor multiplying $\bar{\mathscr{C}}_{42}$ for Geos 1 is negligible, so we immediately have

$$\bar{\mathscr{C}}_{22} = 1.711 + 0.931i \quad ,$$

$$\bar{\mathscr{C}}_{42} = 7.44 + 2.64i \quad ,$$

or in terms of conventional harmonics,

$$C^+_{22} = 1.26 \text{ cm}, \quad \epsilon^+_{22} = 28.^\circ56 \quad ,$$

$$C^+_{42} = 1.76 \text{ cm}, \quad \epsilon^+_{42} = 19.^\circ5 \quad .$$

These values of $C^+_{22}$ would have to be modified if $k_2$ takes a value different from 0.29. However, $C^+_{42}$ is unchanged.

## Conclusions

The accuracy of analytical perturbation theory for short-period perturbations seems certain only to about 20 cm. It may be better than that, and additional, rather straight-forward analysis should clarify the situation. In any case, the numerical verification of analytical theory is not so simple, and the unification of a number of different developments based on different principals still requires some study.

The long-term analysis of mean elements is providing some new information about the ocean and body tides, although many questions remain to be resolved. The relation between ocean and body tides and the possible frequency dependence of $k_\ell$ and the loaded Love number $k'_\ell$ must be resolved before definitive information on ocean tides can be obtained. A number of simple tests can be made, on which the currently derived values fail. They should agree with a numerical model, and the relative amplitudes of tides close in frequency should correspond to the relative amplitude of the driving force as represented by $D_{\ell ms}$.

Finally, some new information should come from analysis of Lageos data. The unexplained secular decrease in semimajor axis will not hinder the planned purposes of Lageos. In all other respects, the satellite is being used as it was intended.

Acknowledgment. This work was supported in part by Contract F19628-78-C-0216 from the U.S. Air Force.

## References

Aksnes, K., Short-period and long-period perturbations of a spherical satellite due to direct solar radiation, Celest. Mech., 13, 89-104, 1976.

Balmino, G., Numerical methods of orbital dynamics, in Satellite Dynamics, ed. by G. E. O. Giacaglia, Springer Verlag, Berlin, pp. 50-97, 1975.

Brouwer, D., Solution of the problem of artificial satellite theory without drag, Astron. J., 64, 378-397, 1959.

Cartwright, D. E., and A. C. Edden, Corrected table of tidal harmonics, Geophys. J., 33, 253-264, 1973.

Deprit, A., and R. Zahar, Numerical integration of an orbit and its concomitant variations by recurrent power series, Zeit. Angew. Math. Phys., 17, 426-430, 1966.

Doodson, A. T., The harmonic development of the tide-generating potential, Proc. Roy. Soc. A, 100, 305-329, 1921.

Farrell, W. E., Deformation of the earth by surface loads, Rev. Geophys. Space Phys., 10, 761-797, 1972.

Felsentreger, T. L., J. G. Marsh, and R. W. Agreen, Analysis of the solid earth and ocean tidal perturbation on the orbits of the Geos-1 and Geos-2 satellites, J. Geophys. Res., 81, 2557-2562, 1976.

Gaposchkin, E. M., Satellite orbit analysis at Smithsonian Astrophysical Observatory, in Space Research VIII, ed. by A. P. Mitra, L. G. Jacchia, and W. S. Newman, North-Holland Publ. Co., Amsterdam, pp. 76-80, 1968.

Gaposchkin, E. M., Satellite dynamics, in 1973 Smithsonian Standard Earth (III), ed. by E. M. Gaposchkin, Smithsonian Astrophys. Obs. Spec. Rep. No. 353, pp. 85-192, 1973.

Garfinkel, B., The orbit of a satellite of an oblate planet, Astron. J., 64, 353-367, 1959.

Goad, C. C., and B. C. Douglas, Lunar tidal acceleration obtained from satellite derived ocean tide parameters, J. Geophys. Res., 83, 2306-2310, 1978.

Hori, G., Theory of general perturbations with unspecified canonical variables, Publ. Astron. Soc. Japan, 18, 286-296, 1966.

Hori, G., Theory of general perturbation, in Recent Advances in Dynamical Astronomy, ed. by B. D. Tapley and V. Szebehely, D. Reidel Publ. Co., Dordrecht-Holland, pp. 231-249, 1973.

Jeffreys, H., and R. O. Vicente, The theory of nutation and the variation of latitude, Mon. Not. Roy. Astron. Soc., 117, 142-161, 1957a.

Jeffreys, H., and R. O. Vicente, The theory of nutation and the variation of latitude. The Roche model core. Mon. Not. Roy. Astron. Soc., 117, 162-173, 1957b.

Kaula, W. M., Theory of Satellite Geodesy, Blaisdell Publ. Co., Waltham, Mass., 124 pp., 1966.

Kinoshita, H., Third-order solution of an artificial satellite theory, Smithsonian Astrophys. Obs. Spec. Rep. No. 379, 104 pp., 1977.

Kozai, Y., The motion of a close earth satellite, Astron. J., 64, 367-377, 1959.

Kozai, Y., Effects of solar radiation pressure on the motion of an artificial satellite, Smithsonian Astrophys. Obs. Spec. Rep. No. 56, pp. 25-33, 1961.

Kozai, Y., A new method to compute lunisolar perturbations in satellite motions, Smithsonian Astrophys. Obs. Spec. Rep. No. 349, 27 pp., 1973.

Kozai, Y., and H. Kinoshita, Effects of motion of the equatorial plane on the orbital elements of an earth satellite, Celest. Mech., 7, 356-366, 1973.

Krogh, F. T., On testing a subroutine for the numerical integration of ordinary differential equations, J. Assoc. Comp. Mach., 20, pp. 545-562, 1973.

Lambeck, K., A. Cazenave, and G. Balmino, Solid earth and ocean tides estimated from satellite orbit analyses, Rev. Geophys. Space Phys., 12, 421-434, 1974.

Lambeck, K., Effects of tidal dissipation in the oceans on the moon's orbit and the earth's rotation, J. Geophys. Res., 80, 2917-2925, 1975.

Lautman, D. A., Perturbations of a close-earth satellite due to sunlight diffusely reflected from the earth, Celest. Mech., 15, 387-420, 1977a.

Lautman, D. A., Perturbations of a close-earth satellite due to sunlight reflected from the earth. II. Variable albedo, Celest. Mech., 16, 3-25, 1977b.

Longman, I. M., A Green's function for determining the deformation of the earth under surface mass loads. 1. Theory, J. Geophys. Res., 67, 845-850, 1962.

Longman, I. M., A Green's function for determining the deformation of the earth under surface mass loads. 2. Computations and numerical results, J. Geophys. Res., 68, 485-496, 1963.

Molodensky, M. S., The theory of nutation and diurnal earth tides, in Quatrième Symposium International sur les Marées Terrestres, Comm. Obs. Roy. de Belgique No. 188, Ser. Geophys. No. 58, pp. 25-56, 1961.

Munk, W. H., and G. J. F. MacDonald, The Rotation of the Earth: A Geophysical Discussion, Cambridge Univ. Press, Cambridge, 323 pp., 1960.

Sehnal, L., and S. B. Mills, The short-period drag perturbations of the orbits of artificial satellites, Smithsonian Astrophys. Obs. Spec. Rep. No. 223, 30 pp., 1966.

Rabe, E., Determination and survey of periodic Trojan orbits in the restricted problem of three bodies, Astron. J., 66, 500-513, 1961.

Robertson, H. P., Dynamical effects of radiation in the solar system, Mon. Not. Roy. Astron. Soc., 97, 423-438, 1937.

United States Naval Observatory, Improved Lunar Ephemeris, 1952-1959, Nautical Almanac Office, 422 pp., 1954.

Velez, C. E., Stabilization and real world satellite problem, in Satellite Dynamics, ed. by G. E. O. Giacaglia, Springer Verlag, Berlin, pp. 136-153, 1975.

# Recent Advances in Computational Techniques

David E. Smith
Goddard Space Flight Center
Greenbelt, Maryland 20771

Abstract. The determination of very precise orbits and geodynamic parameters from laser tracking data requires the continual development and improvement of the software systems and computational techniques. Computational accuracies at the few centimeter level are presently required to match the performance of the present day laser ranging systems and altimeters and in the next few years the accuracies are expected to increase further. In this paper and the major error sources in orbit determination are briefly discussed and the present and future modeling activities needed to meet the accuracy requirements of the next few years are described.

## Introduction

The precise computation of the motion of earth satellites has become a critical component of several space techniques for studying the earth. Perhaps the two most relevant at the present time are the computation of orbits for the determination of crustal motions and the computation of orbits for altimeter satellites. These and other applications require orbital accuracies in the centimeter range and challenge our present-day capability to accurately model many of the forces that perturb satellites and also our computational techniques. Laser ranging systems operated by NASA Goddard Space Flight Center are now at about the 5 cm accuracy level and systems capable of 2 cm are presently under development (Silverberg, this volume). However, our present ability to determine the orbits of satellites has not yet achieved the same level of accuracy and historically orbital accuracy has lagged behind the observational accuracy by several years because the improved observations are needed to improve upon the models for the spacecraft behavior. This situation is particularly true for the gravity field, which for most satellites and applications is still the largest source of orbital error.

Figure 1 shows the developmetn of orbit determination for arc lengths of three to five revolutions over the last several years in comparison with the improvement in the quality of laser range measurements. Figure 1 represents the situation for a typical low altitude spacecraft; such as Beacon Explorer C or GEOS-3 which have altitudes of about 1000 km or less. For most higher altitude satellites, such as Lageos at 6000 km, the orbital fit will be closer to the data quality. Another important consideration in interpreting Figure 1 is that the orbit fit curve is largely based on the experience in the San Andreas fault experiment (Smith, et al., 1979) in which two or three laser tracking stations within one or two thousand kilometers of each other were used. If these stations had been on different continents the fit might have been larger.

For these short orbital arcs of a few revolutions (3 to 8 hours) the largest source of error in the orbital computations is usually the gravity field but as the arc length increases to days, weeks, and months other sources of error usually begin to dominate. Figure 2 illustrates the way the orbit error typically grows as a function of arc length for the major perturbing forces. The vertical axis is arbitrary in scale and is only representative of the relative magnitudes. The exact variation of each of the curves will depend on the satellite, its orbit and the sophistication of the modeling of the perturbing force used in the orbit computation. For example, numerical errors in the computations are usually negligible if an analytical theory is being used for the spacecraft motion but can become extremely important if a numerical integration system is used because the error accumulates as the arc length increases.

At GSFC the major software system for the computation of precision orbits is the Geodyn program (Putney, 1977). This program system has the ability to determine orbits from a variety of tracking data types and is capable of estimating various geophysical parameters such as polar motion and earth rotation, tidal parameters, geopotential coefficients, as well as parameters associated with the perturbing forces, such as drag coefficients, and measurement errors, such as range biases.

In the following section the status of our modeling of the different perturbing forces at GSFC will be briefly described together with our plans for future improvements.

## Orbital Perturbing Forces

### Gravity

One of the areas that has shown greatest improvement during the last decade has been our knowledge of the gravity field of the earth. The inclusion of high precision range measurements from laser tracking systems and Unified S-Band doppler data and, more recently, altimeter data has enabled our definition of the gravity field to extend out to degree and order 36 with specialized altimeter and surface gravity solutions extending out to degree and order 180. These new fields (Goddard Earth Models - GEM) developed at GSFC have permitted improvements of at least an order of magnitude in the determination of short orbital arcs of satellites over the past decade. Figure 3 shows a comparison of the abilities of three gravity models, GEM's 1, 7 and 9, to fit five consecutive passes of laser data from a single tracking station. These five passes, obtained at GSFC in 1974 on the Beacon Explorer C satellite, when analyzed by the GEM 1 gravity field developed in 1970-71 (Smith, et al., 1973) could only be satisfied at the 2 meter level even though the data was of 10 cm quality.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

The same data analyzed a few years later with ghe GEM 7 model (Wagner, et al., 1977) could be satisfied to about the 50 centimeter level and more recently the GEM 9 (Lerch, et al., 1978) model fits to 12 cm. The improvement from GEM 1 to GEM 9 has been brought about largely by the inclusion in the later models of large quantities of laser tracking data on several satellites, particularly GEOS-3, but not the Beacon Explorer C data shown in Figure 3. The slight curvature of the GEM 9 results in Figure 3 show that some gravitational signature still exists in the data and that some improvement still remains to be made although this may well be the most difficult.

For longer arcs the gravity error increases to about the 50 cm to 1 meter level after one week with the GEM 9 field for low altitude satellites such as Lageos, reaching 50 cm after about one month.

The present plans at GSFC are to continue to improve our knowledge of the gravity field so that the locations of mobile and fixed laser stations can be determined to the few centimeter level for the precise measurement of crustal motions; and also for the precise analysis of the GEOS-3 and Seasat altimeter data.

Air Drag

As the orbital arc length increases the second most important perturbing force (after gravity) for low altitude satellites is usually the effect of air drag. The general form of the perturbation is

$$\text{accel.} = -\frac{1}{2} C_D \frac{A}{M} \rho v^2$$

where $C_D$ is the drag coefficient, A is the spacecraft cross-sectional area, M is the spacecraft mass, $\rho$ is the air density and $v$ is the spacecraft velocity. At the present time the density models used in computing the drag acceleration are based on the work of Jacchia and include variations in solar activity, diurnal terms, geomagnetic effects and semi-annual and seasonal latitude variations. In order to improve the responsiveness of the model to unmodeled changes in density we have introduced a time dependent parameter ($\dot{\rho}$) that enables us to account for systemated changes during the orbital arc. In addition, we are introducing a capability to estimate the drag coefficients ($C_D$) for specific periods during the orbital arc. Thus it will be possible to vary the drag coefficient from one day to the next and thereby modify the drag acceleration without any change to the density model. We believe this may accomodate density changes that last for short periods of time that are not represented in our models.

Another capability that exists in the GSFC Geodyn program is to accurately model the cross-sectional area of the spacecraft. This facility was introduced for the Beacon Explorer C, GEOS-3 and Seasat spacecraft because of their irregularity in shape and the need for very precise orbit calculations on these satellites. In all these cases drag was a major influence in the orbit computations and inclusion of a variable cross-sectional area could improve upon the computational accuracy. The technique is incorporated by computing externally to the main program the cross-sectional area as a function of angle of incidence and including this information in tabular form in the orbit program. At each integration step the appropriate area is deduced from the table and used in the drag calculations.

Solar Radiation Pressure

The form of the perturbation by solar radiation pressure is

$$\text{accel.} = -C_R \frac{A}{M} \frac{(\text{solar constant})}{C}$$

where $C_R$ is a constant that accounts for the type of reflection, specular diffuse, etc., that is taking place at the satellite, $\frac{A}{M}$ is the area (A) to mass (M) ratio and C is the velocity of light. The model used in the Geodyn program includes a solar flux varying with solar distance, and approximations for absorption and refraction at the terminator. The incident area is variable in the same manner as for air drag and takes into account spacecraft attitude, shape, shadowing and varying reflective properties over the spacecraft. Although these computations are reasonably precise we believe that errors are occurring at the umbra/penumbra/full sunlight boundaries where the numerical integration procedure jumps over one or even two boundaries in one step. The effect of this error is estimated to be a slight offset in the boundary location.

Albedo Radiation Pressure

At the present time we do not have an albedo model in our orbit computation program but one is under consideration. The basic form of the perturbation is

$$\text{Accel.} = -C_R \frac{A}{M} \frac{(\text{Albedo})}{C} \left(\frac{R}{D}\right)^2$$

and is similar to that for solar radiation pressure except for the (R/D) term which shows that the acceleration follows the inverse square law (R is the earth radius, D is the radial distance of the satellite). The difficulty with evaluating this perturbing force is that the albedo is variable in both space and time and needs to be evaluated at every integration step for the entire surface of the earth observable from the spacecraft. This procedure is computationally very time consuming. Simplification of the albedo into only day/night effects, for example, will probably underestimate the effect and provide deceiving results. Our present considerations are directed towards the computation of the long-term effects of albedo by digitally representing seasonal albedo maps of the world derived from satellite meteorological data.

Earth and Ocean Tides

Our present modeling of earth tides in the Geodyn program is a single second degree spherical harmonic with one amplitude and phase. This model is used to compute the gravitational effects of earth tides on the satellite orbit and to compute the body tides on the locations of the

tracking stations. Our ocean tide models only account for the displacement of the ocean surface (Hendershott model globally, Mofjeld in Northwest Atlantic) for the analysis of satellite altimeter data and do not include the gravitational effect on the satellite. However, considerable accomodation of the ocean tide effect on the orbit can be achieved by suitable modification of the solid earth tidal amplitude and phase. For example, for Beacon Explorer C we were able to model approximately 90 to 95% of the combined long period earth and ocean tidal perturbation of the satellite with a Love number ($k_2$) of 0.245, and a phase lag of 3.2 degrees (Smith, et al., 1972) used in the solid earth tide model. This accomodation of the oceans within the earth model ignores any frequency dependent terms in either the earth or ocean tides. Consequently, we are planning to allow for a frequency dependence of both the amplitude and phase of the solid-earth which we expect to permit complete accomodation of the ocean tides. In addition, we expect to incorporate a spherical harmonic representation of the major ocean tides ($M_2$, $S_2$, $K_1$, $K_2$, $N_2$, $O_1$, $P_1$, etc), derived from the numerical integration of Laplace's tidal equations, into our program system so that we can use these models for both altimetry and orbital analysis and also be able to use altimetry and orbital data to improve on the coefficients in the tidal expansions.

Numerical Problems

Numerical integration systems introduce errors of rounding and truncation into the orbit calculations that can become significant for long orbital arcs. In the Geodyn program system a typical step size within the integrator will be about sixty seconds but as the size of the gravity models has increased we are finding that this figure needs to be reduced to perhaps forty seconds in order to properly account for the high frequency terms; and the CPU time increases accordingly. Generally, with step sizes of the order of sixty seconds integration error can be kept to the order of a meter after about 30 days. This error is also predominantly along track and is an acceleration similar to drag. Thus, if drag is being adjusted in the orbit determination process then the integration error will be largely absorbed into the adjustment. In this case integration error is not a major contributor to orbit error.

Another aspect of numerical problems in orbit determination is the core size required to determine large spherical harmonic gravity fields and associated station coordinates. Some of our large gravity field and station coordinate solutions have nearly 2,000 unknowns and require several million bytes on an IBM 360/95 type computer. This storage is not always available and if the computations are attempted in smaller core the CPU and/or IO time increases. Nearly all our computations in satellite orbit, geodesy and geodynamics are requiring greater accuracy today than a few years ago and this means an increase in both CPU time and core storage.

Because of this need for additional precision and complexity in the computations we are giving consideration to the advantages and disadvantages of parallel processor computers and dedicated minicomputers for some of this work.

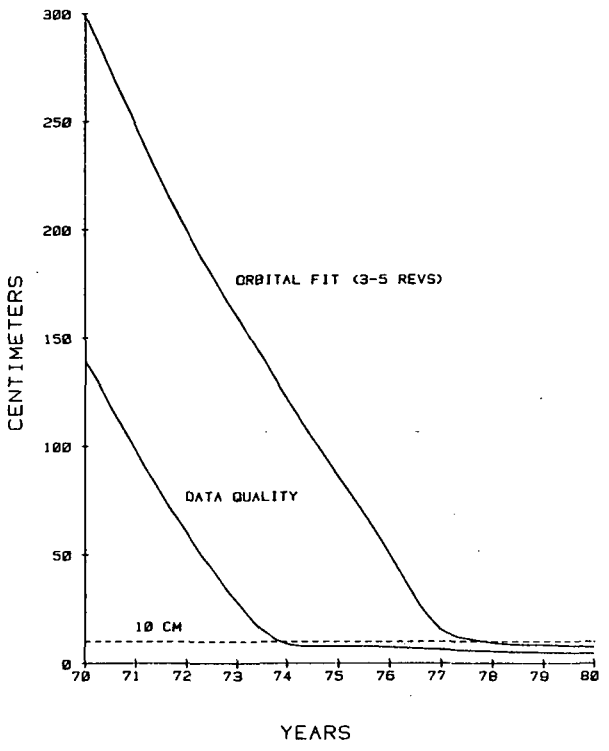



Fig. 1. Development of orbit determination capability.

Conclusions

The application of satellite geodetic techniques to problems in solid earth and ocean dynamics is requiring ever increasing accuracy in the computation of satellite orbits. This need presents considerable difficulty in the modeling of the many perturbing forces that influence the spacecraft motion. With increasing complexity and accuracy is the need for faster and larger computing facilities.

In this paper I have briefly described the status of the major GSFC orbit and geodynamic parameter estimation program (Geodyn) and the degree of complexity that we are finding necessary to meet orbital and geophysical accuracies. The computing of precision orbits at the centimeter level is difficult to obtain and even more difficult to maintain for any length of time and may have to be limited to satellites, such as Lageos, that are carefully configured to minimize the perturbing forces.

References

Lerch, F.J., S.M. Klosko, R.E. Laubscher, and C.A. Wagner, Gravity model improvement using GEOS-3 (GEM 9 & 10), Goddard Space Flight Center, Report X-921-77-246, 1977.

Putney, B., General theory for dynamic satellite geodesy in National Geodetic Satellite Program, NASA SP-365, 319-334, 1977.

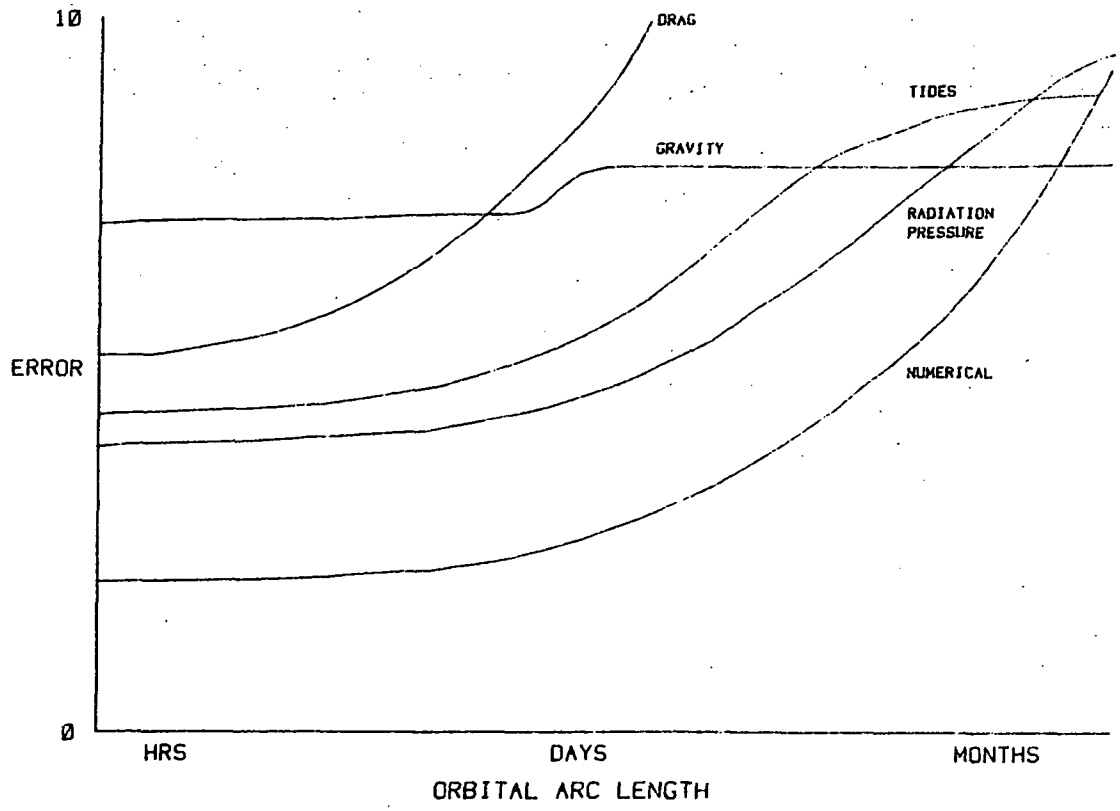



Fig. 2. Relative orbit error for low altitude spacecraft.

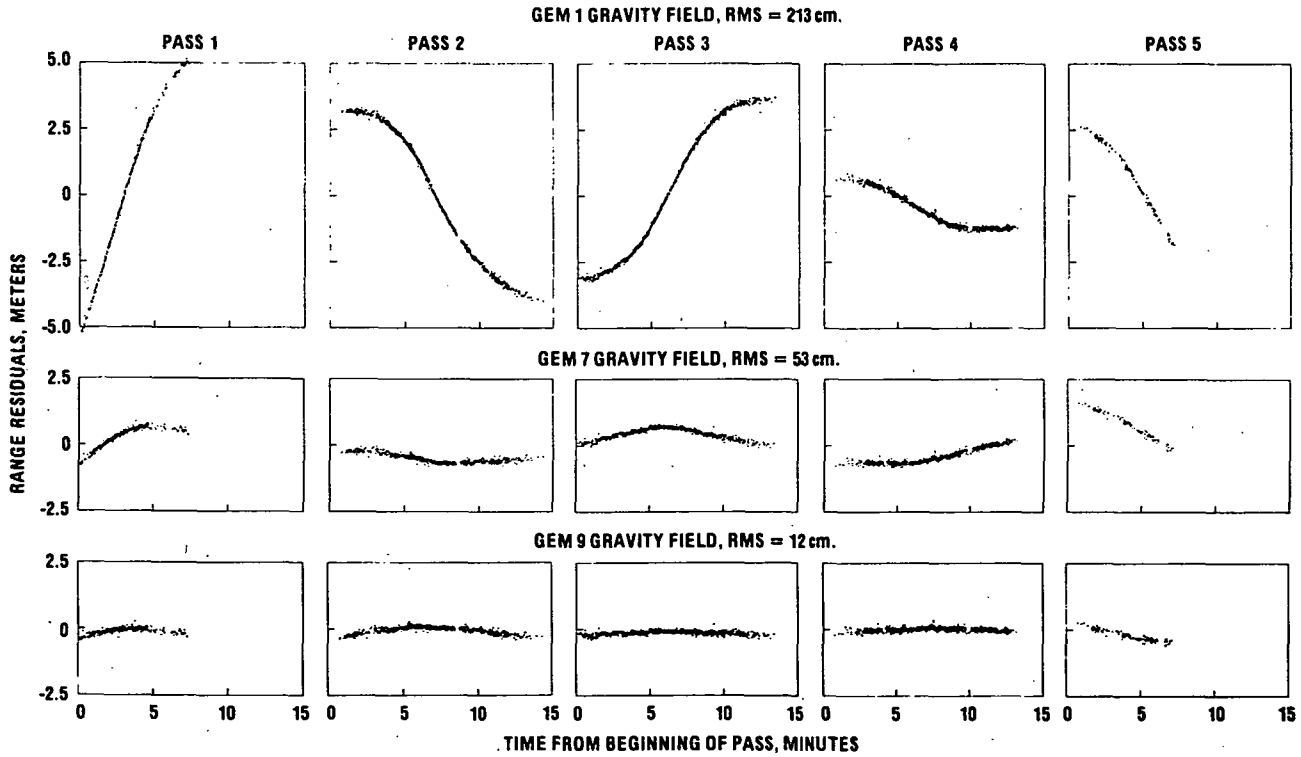



Fig. 3. Laser residuals from five pass orbit of Beacon Explorer C in 1974.

Smith, D.E., R. Kolenkiewicz and P.J. Dunn, A determination of the earth tidal amplitude and phase from the orbital perturbations of the Beacon Explorer C spacecraft, NATURE, 244, 498, 1973.

Smith, D.E., R. Kolenkiewicz, P. J. Dunn and M.H. Torrence, The measurement of fault motion by satellite laser ranging, Tectonophysics, 1979 (in press).

Wagner, C., F. Lerch, J. Brownd, and J. Richardson, Improvement in the geopotential derived from satellite and surface data (GEM 7 and 8), J. Geophys. Res., 82, 901, 1977.

Page Intentionally Left Blank

# A Drag-Free Lo-Lo Satellite System for Improved Gravity Field Measurements

Robert E. Fischell and Vincent L. Pisacane
Applied Physics Laboratory, Johns Hopkins University
Laurel, Maryland 20810

Abstract. Since the force of gravity is inversely proportional to the square of the distance between two masses, it is obvious that the effect of small gravitational anomalies on the motion of an orbiting satellite will increase as the satellite altitude decreases. However, at very low altitudes, the effect of atmospheric drag results in drastically reduced orbit lifetimes and considerable uncertainty in satellite motions. The concept suggested herein employs a DISturbance COmpensation System (DISCOS) on each of a pair of satellites at very low altitudes to provide refined measurements of the earth's gravitational field. The DISCOS maintains the satellites in orbit and essentially eliminates motion uncertainties due mostly to drag and to a lesser extent from solar radiation pressure. By a closed-loop measurement of the relative range-rate between the two low satellites, one can determine the earth's gravitational field with a considerably greater accuracy than could be obtained by tracking a single satellite.

## Introduction

Since the advent of artificial satellites of the earth, the science of geodesy has advanced with remarkable rapidity. Without employing orbiting satellites, our knowledge of the earth's gravity field could not be determined to nearly the level to which it is now known. This improved knowledge of geodesy is accomplished essentially by studying the long period motions of orbiting satellites, particularly satellites at quite moderate altitudes. The lower the altitude of the satellite, the more pronounced the gravitational effect of the earth's gravity field, particularly the higher order harmonics. Although it is obviously desirable to go to still lower altitudes, the uncertainty of the along track force caused by air drag results in a severe degradation of the accuracy to which the satellite motion can be measured.

To determine the earth's gravity field, with greater precision it is necessary to make increasingly more accurate measurements of the satellite's motion. A basic limitation in the use of a single satellite to determine the earth's gravity field is that it is very difficult to measure the minute difference in velocity of the low orbiting satellite due to small gravitational effects in the presence of the overwhelming 7 kms/sec of satellite orbit velocity.

Suggested herein is the concept of providing a drag-free capability to each one of a pair of satellites at very low altitudes. The principal advantages of such a system are that 1) A drag free system allows the satellite to go to extremely low altitudes where they are most sensitive to gravitational anomalies and meaningful velocity measurements can be accomplished because the satellite is free of the effects of atmospheric drag; 2) Precise measurements of the relative velocities of two satellites in identical orbits is vastly simpler than the measurement of the absolute velocity of a single satellite. This is because it is less difficult to directly measure a small quantity rather than to make a measurement of a small quantity in the presence of a very large quantity.

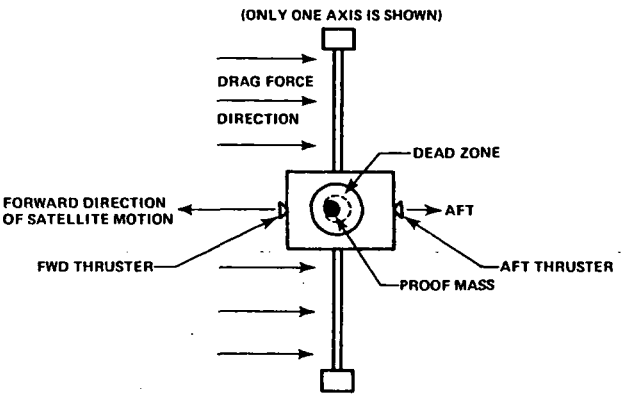



**Fig. 1 Illustration of DISCOS as used on the Triad Satellite**

## DISCOS Concept and Performance

A DISturbance COmpensation System (DISCOS) has been shown to perform satisfactorily in orbit on the TRIAD satellite. An illustration of a single axis of the three-axis DISCOS system that was used on TRIAD is shown in Figure 1. The TRIAD satellite was extended in the vertical direction to achieve gravity-gradient stabilization. As such, it presented a considerable cross-sectional area on which the atmosphere acted to create drag. At the center of the satellite was a dense spherical mass called the proof mass which was contained within a spherical cavity that shielded the proof mass from the effect of external disturbance forces such as solar radiation pressure and air drag. Contained within the spherical cavity at the center of the satellite was an imaginary spherical cavity known as the dead zone. When the effect of drag was sufficient to push the satellite backward until the dead zone touched the surface of the proof mass, the aft thruster fired thereby causing the satellite to move forward relative to the proof mass. By this means, the proof mass was allowed to fly in an orbit determined purely by gravitational forces, and the satellite, because of the DISCOS system, was constrained to follow the motion of the proof mass. In Figure 2 is shown a particular 12 minute period of data collected from the TRIAD satellite. In this figure is shown the along-track position on the proof mass expressed in millimeters as a function of time from the start of a pass of the satellite. During the time of this pass, the principal disturbance force on the satellite was drag, so that at approximately one minute of time, the integrated

Proc. of the 9th GEOP Conference. *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280. The Ohio State Univ., Columbus, Ohio 43210.

drag force was sufficient to cause the proof mass to touch the imaginary dead zone. This is a consequence of 1.0 mm of motion of the proof mass from the geometric center of the cavity. At that point, which is shown in Figure 2 as the first of three "Aft Thruster Fires" indications, the impulse of the aft thruster causes a velocity of the satellite relative to the proof mass such that proof mass has an apparent motion back towards the geometric center of the cavity. However, due to the fact that the drag forces are continuing at this time, the proof mass never reaches the center of the cavity. Four minutes after the first of the thruster firings, the proof mass again touches the dead zone. The parabolic shape of the displacement curve of the proof mass as a function of time is an indication of an essentially constant drag force during this four minute period. From the times of 5 minutes to approximately 11 minutes, the satellite is again pushed away from the proof mass but this time the different shape of the essentially parabolic curve indicates that the drag force at that time was smaller. Finally we see at 11 minutes the third of the firings of the aft thruster. This data is typical of what occurs in a DISCOS system as the firing of the thrusters are used to keep the proof mass centered in the satellite's spherical cavity.

## Application of DISCOS to Satellites at Very Low Altitude

The TRIAD satellite containing the first three-axis DISCOS flew at an altitude of approximately 700 km and used three pounds of cold gas (Freon) propellant to achieve 18 months of drag free operation. If one is to maintain a spacecraft at a much lower altitude or if one is to maintain drag free operation for a longer period of time more total impulse from the propellant is required. This can be accomplished either by providing a greater quantity of propellant and/or by providing a much higher specific impulse as compared with cold gas. In Figure 3 is shown the concept of a satellite which could be one of a pair used in a drag free lo-lo system for obtaining more refined measurements of the earth's gravity field. The satellite shown in Figure 3 has a DISCOS system at its center mass and two very large tanks containing propellant (e.g., hydrazine) located symmetrically on each side. It is estimated that for a satellite of this size, 2000 kilograms of hydrazine could be readily provided. Figure 3 also shows the concept of simple angular momentum flywheel that could be used to stabilize the satellite in roll while aerodynamic forces could be used to stabilize the satellite in pitch and yaw. A radar altimeter is shown which might provide improved data on ocean surface topography which of course could contribute materially to our knowledge of the earth's gravity field and ocean dynamics. Furthermore, this concept shows a vector and scalar magnetometers which are indicative of another class of measurement which is best accomplished with satellites at very low altitudes.

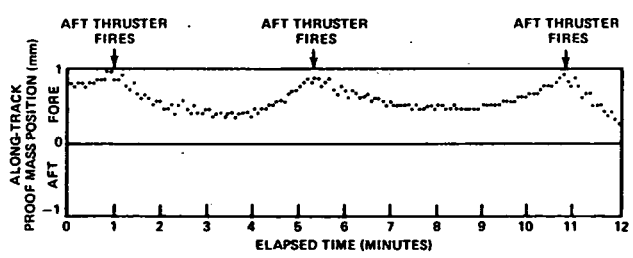



Fig. 2 DISCOS proof mass motion.

In Figure 4 is shown the propellant usage for such a satellite in the altitude region between 100 and 250 kms. The propellant usage in Figure 4 is expressed in kilograms/month for a satellite using hydrazine as the DISCOS system propellant and having a frontal area of 1 square meter. Two curves are shown in Figure 4; one for a solar minimum and the other for a solar maximum. If a satellite such as that shown in Figure 3 can contain 2000 kilograms of hydrazine, it will provide approximately 2 months of drag-free operation at a 125 km and approximately 20 months operation (at a solar minimum) at an altitude of 150 km.

A consideration which makes the satellite concept as shown in Figure 3 reasonable from a cost standpoint is the advent of the Shuttle as a launch vehicle. The Shuttle is particularly well suited to launch satellites that are extremely heavy but that require very low altitude for their operation. If longer times in orbit are required, the Shuttle astronauts could rendezvous with the pair of drag-free satellites and refuel these spacecraft in-orbit. Although there would be times when it is desirable to fly the satellite at altitudes as low as 125 to 150 kilometers, most of the time a considerable improvement in the knowledge of the earth's gravity field would result from flying the satellites at altitudes in the region of 250 kilometers. At such an altitude, even at the solar maximum, 2000 kilograms of hydrazine would provide approximately 200 months of in-orbit operation without refueling. This mission duration time would be satisfactory for significantly improving the accuracy to which we know of the earth's gravitational field.

Another effect that must be considered in allowing spacecraft to fly at extremely low altitudes is the aerodynamic heating that results from friction of the high speed spacecraft with the earth's atmosphere. In Figure 5 is shown the temperature at the stagnation point which would be obtained at the leading surface of a satellite that was flying at altitudes between 100 and 250 kilometers. The assumptions here are that the front surface has an infra-red emmisivity of 0.8 and there is no energy input from the sun. From these curves (which show solar maximum and solar minimum), it can be seen that at an altitude of 125 kilometers, the skin temperature at the stagnation point under the conditions described above is below 300°C. Such a temperature could readily be withstood even by aluminum. When one rises to an altitude of 150 kilometers, Figure 5 shows that the temperature of the front surface is not very much different from room temperature and therefore would not compromise the design of a spacecraft. In essence, Figure 5 shows us that there is no significant temperature barrier to operating a single satellite or a pair of satellites at extremely low altitudes; i.e., above 125 km.

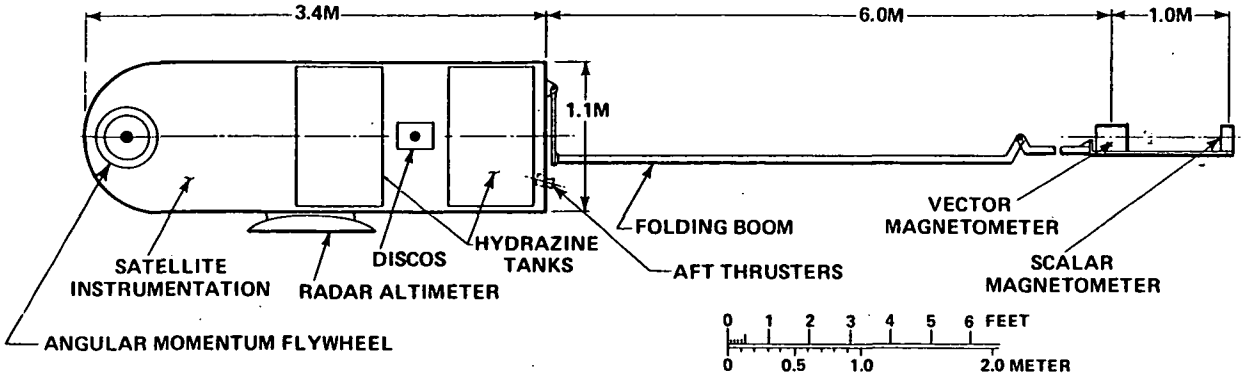



Fig. 3 Geomapsat conceptual design.

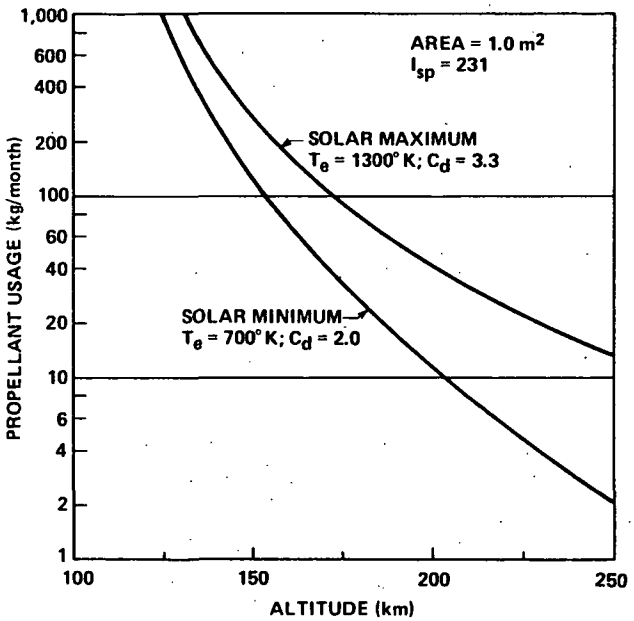



Fig. 4 Propellant usage for satellites at very low altitude.

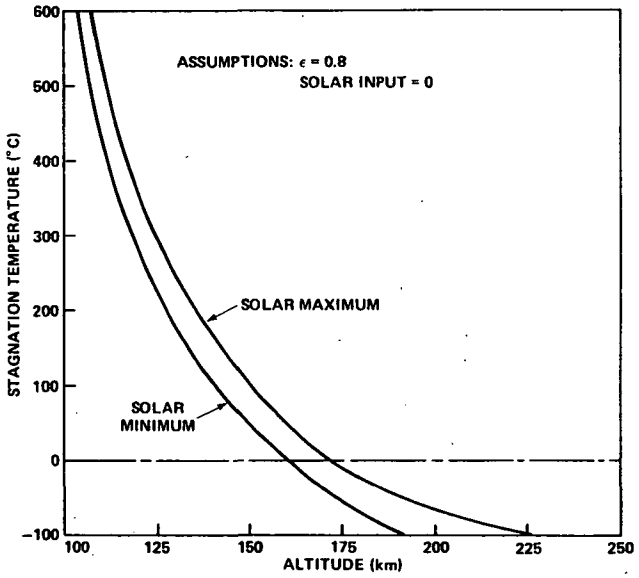



Fig. 5 Effect of altitude on stagnation temperature.

## Analysis of Lo-Lo System of Satellites

In this section, an approximate expression is developed for the relationship between the gravitational environment and the relative range-rate between the satellites of a lo-lo system. Through judicious approximations an analytic formulation can be obtained. This is useful to enhance conceptual understanding and to provide a basis for evaluating and interpreting the more detailed numerical simulation which follow.

Assume that the two satellites in Figure 6 are in identical circular orbits but separated in phase by the angle $\alpha$ and have mean anomalies $M_i$ where i = 1 or 2 which identifies each one of the two spacecraft. Let T be the kinetic energy per unit mass and U the gravitational energy per unit mass. If the rotation of the gravitational field is neglected it is possible to introduce the principle of conservation of energy

$$T + U = K \qquad (1)$$

where K is a constant. Let $L_i$, $H_i$, $C_i$ be the spatial perturbations of the spacecraft from the cir-

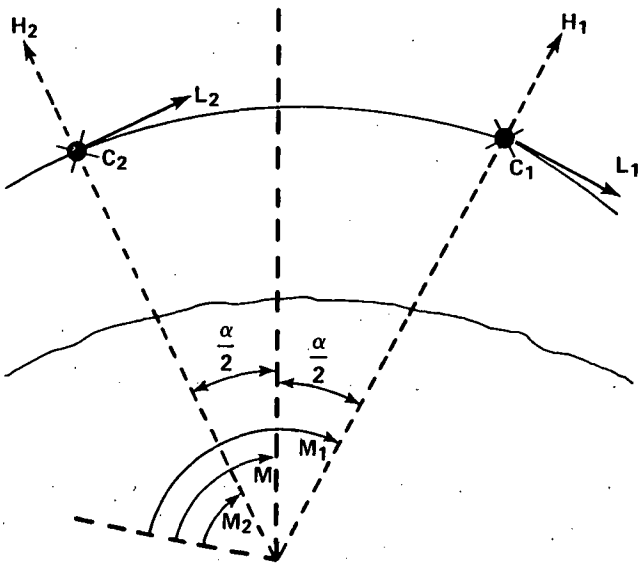



Fig. 6 Geometry of lo-lo system of satellites and definition of $L_i$, $C_i$, $H_i$ coordinate systems, i = 1 to 2.

cular reference orbit consistent with the reference systems in the figure. Then

$$T = \tfrac{1}{2}\left[(v_o + \dot{L}_i)^2 + \dot{C}_i^2 + \dot{H}_i^2\right] \tag{2}$$

where $v_o$ is the nominal circular velocity. Let

$$U = V - \frac{k}{a} \tag{3}$$

where V is the perturbing potential per unit mass, k is the gravitational constant and a the radius to the satellite. Substituting Eqs. (2) and (3), Eq. (1) gives

$$\tfrac{1}{2} v_o^2 + v_o \dot{L}_i + V_i - \frac{k}{r} + \left[\tfrac{1}{2}\left(\dot{L}_i^2 + \dot{H}_i^2 + \dot{C}_i^2\right)\right] = K \tag{4}$$

The bracketed term is of second order since $\dot{L}_i$, $\dot{H}_i$, $\dot{C}_i \ll v_o$ and can be neglected.

For closed circular motion

$$K = -\frac{k}{2a} \tag{5}$$

and

$$v_o = \left(\frac{k}{a}\right)^{\frac{1}{2}} \tag{6}$$

Substituting Eqs. (5) and (6) into Eq. (4) give

$$\dot{L}_i = -v_o^{-1} V(M_1) . \tag{7}$$

The relative range-rate between the two spacecraft in the same plane is

$$\dot{\rho} = (\dot{H}_1 + \dot{H}_2) \sin\frac{\alpha}{2} + (\dot{L}_1 - \dot{L}_2) \cos\frac{\alpha}{2} . \tag{8}$$

For small separation distances, i.e., $\alpha$ small, it is appropriate the range-rate be

$$\dot{\rho} \approx \dot{L}_1 - \dot{L}_2 \tag{9}$$

which using Eq. (7) becomes

$$\dot{\rho} = -v_o^{-1} (V(M_1) - V(M_2)). \tag{10}$$

Under the assumptions stated, this equation demonstrates that the relative range-rate is a measure of the difference in the perturbing potential. This result has been previously derived by Wolff [1969] and Comfort [1973]. Now, let the perturbing potential be expressed by

$$V = \sum_{\ell=2}^{\infty} V_\ell . \tag{11}$$

For a near circular orbit Kaula [1966] gives

$$V_\ell = \frac{k}{R}\left(\frac{R}{a}\right)^{\ell+1} \sum_{n=o}^{p} \sum_{p=o}^{\ell} F_{\ell mp}\, S_{\ell mp}(\omega, M, \Omega, \theta) \tag{12}$$

where $F_{\ell mp}$ is soley a function of the inclination and

$$S_{\ell mp} = \begin{bmatrix} C_{\ell m} \\ -S_{\ell m} \end{bmatrix}^{(\ell-m)\text{even}}_{(\ell-m)\text{odd}} \operatorname{Cos}\left[(\ell-2p)(M+\omega) + m(\Omega-\theta)\right] \qquad \begin{bmatrix} S_{\ell m} \\ C_{\ell m} \end{bmatrix}^{(\ell-m)\text{even}}_{(\ell-m)\text{odd}} \operatorname{Sin}\left[(\ell-2p)(M+\omega) + m(\Omega-\theta)\right] \tag{13}$$

where R is a scaling distance, $\omega$ the argument of perigee, $\Omega$ the right ascension of the ascending note, $\theta$ the Greenwich sidereal time and $\ell$ and m the order and degree respectively. For a polar satellite $(\Omega-\theta)$ is constant since the rotation of the earth has been neglected here. Consequently $V_\ell$ is a harmonic function with arguments $(\ell-2p)(M+\omega)$.

Partitioning the potential in components $V_\ell$ permits Eq. (10) to be rewritten as

$$\dot{\rho} = \sum_{\ell=2}^{\infty} \dot{\rho}_\ell \tag{14}$$

where

$$\dot{\rho}_\ell = -v_o^{-1}\left[V_\ell(M_1) - V_\ell(M_2)\right] \tag{15}$$

For the lo-lo configuration

$$M_1 = M + \frac{\alpha}{2} \quad , \quad M_2 = M - \frac{\alpha}{2} \tag{16}$$

so that

$$S_{\ell mp}(M_1) - S_{\ell mp}(M_2) = 2 \sin\frac{(\ell-2p)\alpha}{2}\, \bar{S}_{\ell mp}(M,\omega,\Omega,\theta) \tag{17}$$

where $\bar{S}_{\ell mp}$ is the derivative of $S_{\ell mp}$ with respect to its argument and $\omega = \omega_1 = \omega_2$, $\Omega = \Omega_1 = \Omega_2$. Substituting Eqs. (12) and (17) into Eq. (15) gives

$$\dot{\rho}_\ell = -v_o^{-1}\left(\frac{k}{R}\right)\left(\frac{R}{a}\right)^{\ell+1} \sum_{m=o}^{p} \sum_{p=o}^{\ell}\left[2 \sin\frac{(\ell-2p)\alpha}{2}\right] F_{\ell mp}\, \bar{S}_{\ell mp} \tag{18}$$

The bracketed term represents a function dependent on the separation of the two spacecraft that modulates the amplitude of the range-rate. This term demonstrates the tuning capability of the lo-lo configuration. For a particular value of $\ell$ say $\ell_o$, Eq. (18) shows that frequencies $\ell_o$, $\ell_o$-2, $\ell_o$-4, $\cdots$ times $(M+\omega)$ exist where contributions to the range-rate with frequency $\ell_o(M+\omega)$ results from coefficients $C_{\ell m}$ and $S_{\ell m}$ such that

$$\ell = \ell_o + 2p \qquad p = 0,1,2,\cdots \tag{19}$$

In this context, Eq. (19) shows that by selecting a separation

$$\alpha = \frac{\pi(2n+1)}{\ell_o} \qquad n = 0,1,2,\cdots \tag{20}$$

the contributions to the range-rate from coefficients $C_{\ell m}$ and $S_{\ell m}$

$$\ell = \ell_o + 2p \qquad p = 0,1,2,\cdots \tag{21}$$

will be doubled. Conversely, all coefficients such that

$$\ell = 2m(\ell_o + 2p) \qquad n = 0,1,2,\cdots \tag{22}$$

will contribute nothing to the signal. This tuning capability is illustrated in Figure 7. Another way to conceptualize tuning is to consider the separation to be such that the variations in the gravitational field are either negatively or positively correlated. Enhancement in the signal, will occur with the former and destruction with the latter.

An estimate of the amplitude of the range-rate for a specific gravity anomaly can be obtained as follows. The gravity anomaly $\Delta g$ at the spacecraft can be written in terms of the disturbing potential V as

$$\Delta g = -\left.\frac{\partial V}{\partial r}\right|_{r=a} - 2\frac{V}{a} \tag{23}$$

Substituting Eqs. (11) and (12) for V gives

$$\Delta g = \frac{k}{R^2}\sum_{\ell=2}^{\infty}(\ell-1)\left(\frac{R}{a}\right)^{\ell+2}\sum_{m=o}^{\ell}\sum_{p=o}^{\ell} F_{\ell mp}\, S_{\ell mp} \tag{24}$$

Let $\Delta g_\ell$ be defined by

$$\Delta g = \sum_{\ell=2}^{\infty} \Delta g_\ell \tag{25}$$

where

$$\Delta g_\ell = (\ell-1)\frac{k}{R^2}\left(\frac{R}{a}\right)^{\ell+2}\sum_{m=o}^{\ell}\sum_{p=o}^{\ell} F_{\ell mp}\, S_{\ell mp} \tag{26}$$

Comparing Eqs. (12) and (26), gives

$$V_\ell = \frac{a}{\ell-1}\Delta g_\ell \tag{27}$$

Since the interest is in recovering a gravity anomaly whose magnitude is specified at the surface, let

$$\overline{\Delta g}_\ell = \left.\Delta g_\ell\right|_{a=R} \tag{28}$$

From Eq. (26), it follows that

$$\Delta g_\ell = \left(\frac{R}{a}\right)^{\ell+2}\overline{\Delta g}_\ell \tag{29}$$

so that Eq. (27) can be rewritten as

$$V_\ell = \frac{R}{\ell-1}\left(\frac{R}{a}\right)^{\ell+1}\overline{\Delta g}_\ell \tag{30}$$

Substituting this into Eq. (15) gives

$$\dot\rho_\ell = -v_o^{-1}\frac{R}{\ell-1}\left(\frac{R}{a}\right)^{\ell+1}\left[\overline{\Delta g}_\ell(M_1+\omega) - \overline{\Delta g}_\ell(M_2+\omega)\right] \tag{31}$$

as the expression for range-rate. The maximum amplitude for $\dot\rho_\ell$ will occur at the spacing given by Eq. (20) where

$$\left.\dot\rho_\ell\right|_{max} = -v_o^{-1}\frac{2R}{\ell-1}\left(\frac{R}{a}\right)^{\ell+1}\overline{\Delta g}_\ell\left(M+\omega+\frac{\pi}{2}\right). \tag{32}$$

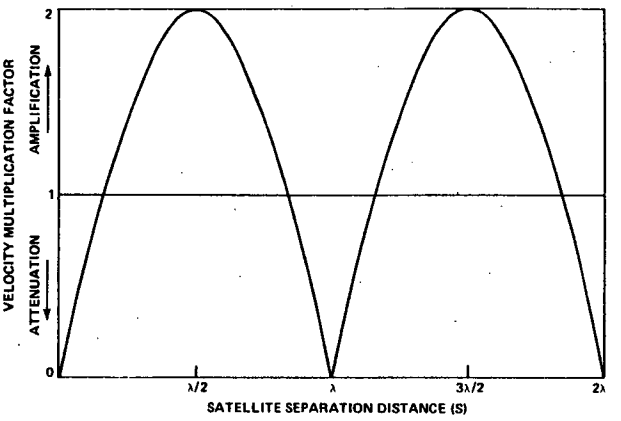



**Fig. 7 Effect of separation distance on velocity amplification.**

This equation has been used as the basis to determine the accuracy in the range-rate measurement that is necessary to recover a specified gravity anomaly. Results are presented in Figure 8 where the curves represent the amplitude and wavelength of the gravity anomalies that produce a one way range-rate of the specified instrument accuracy. At the longer wavelengths a small angle approximation for optimum separation becomes less valid. A recent study by Goldfinger [1978] have shown that for harmonics of order two even at the optimal spacing the error is less than a factor of 2.

Corruption of the signal by errors in the knowledge of the initial conditions has been discussed by Pisacane [1978]. There, a computer simulation study show that the effects of the initial condition errors can be effectively eliminated by high pass filtering.

## Velocity Measurement Concept

The relative velocity measuring instrumentation concept is indicated in Figure 9. The basic method can be described by thinking of the second satellite as a pure transponder whereby the tone $(Nf_1)$ transmitted by the first satellite is simply sent back with a doppler offset resulting from the relative motion between the satellites. The same doppler effect would also be experienced on the return trip resulting in a received frequency of $Nf_1(1 - \dot\rho/c)^2$ at the first satellite; where $\dot\rho$

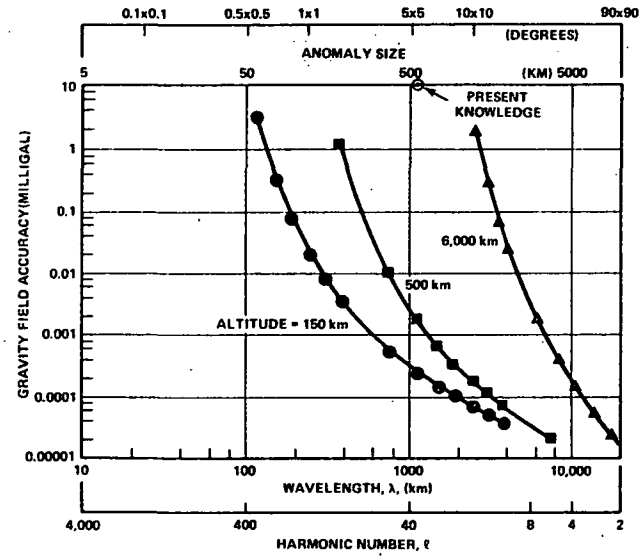



Fig. 8 Anticipated accuracy for recovery of the gravity field based on relative velocity measurement accuracy of $10^{-4}$ mm/sec.

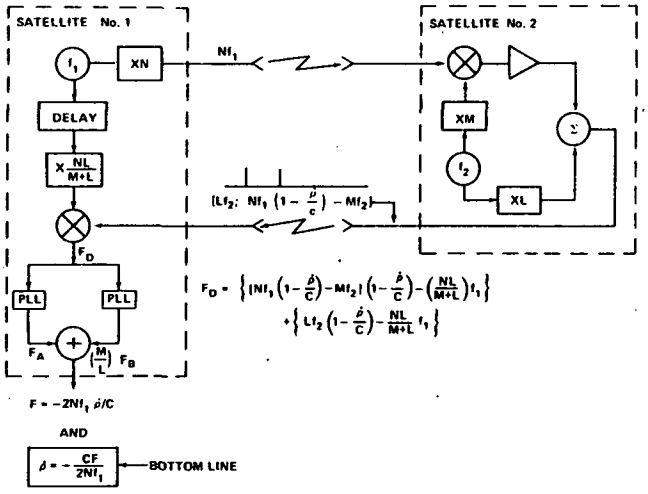



Fig. 9 System for determining relative velocity of a lo-lo satellite pair.

| | | |
|---|---|---|
| FREQUENCY | = | 100 GHz |
| ANTENNA BEAM WIDTH | = | 2° x 10° |
| ANTENNA SIZE | = | 18mm x 90mm (0.7" x 3.5") |
| NOISE FIGURE | = | 10 dB |
| BANDWIDTH | = | 1.0 Hz |
| TRANSMITTED POWER | = | 2.0 mW |
| RANGE | = | 112 km |
| SIGNAL-TO-NOISE RATIO | = | 45 dB |
| INTEGRATION TIME | = | 5.0 SECONDS |
| 1σ PHASE JITTER | = | 0.06 DEGREES |
| 1σ VELOCITY NOISE | = | $10^{-4}$ mm/sec |

Fig. 10 An example of lo-lo closed-loop link parameters.

represents an average velocity over the measurement interval. If the return signal is then differenced with the original transmission frequency, the velocity component, approximately $2Nf_1 \dot{\rho}/c$, will be separated for measurement. The actual implementation incorporates a frequency translation in the second satellite to avoid oscillation difficulties. However, the implementation provides for removal of both oscillator offset errors thereby providing the same final result.

If both oscillators were properly characterized as having constant but unknown frequencies, i.e., having only bias errors, the technique just discussed would fully remove all oscillator errors. When time varying errors are included, their effect can in concept be removed by matching the delays into the differencing circuits. The delay shown in the figure between the $f_1$ oscillator and the return mixer is intended to carry out this function. Removal of time varying effects of the $f_2$ oscillator requires that the two delays between the $f_2$ oscillator and the summing circuit in the second satellite be matched. Assuming that complete oscillator errors are removed by the implementation, the remaining error sources are due to phase variations in uncommon circuits, propagation effects and thermal noise limitations in the link.

Assuming all other effects can be made negligible, the signal-to-noise performance shown in Figure 10, can be realized with quite reasonable antenna dimensions and transmission powers. While the system is not fundamentally limited to this precision, a high degree of circuit phase stability will be required to achieve the indicated performance. Prior to further evaluation of the instrumentation limitations, the $10^{-4}$ mm/sec noise value is being considered as the practical measurement limit.

## Conclusions

The advent of the shuttle launch system makes it possible to launch a pair of satellite containing several thousand kilograms of propellant into a low altitude orbit.

These large quantities of propellant can be used in a DISCOS system to maintain a satellite in orbit and to free the satellite from the otherwise overwhelming disturbance effects of atmospheric drag.

A pair of such satellites at very low altitudes can greatly refine our knowledge of the earth's gravitational field.

Acknowledgments. The authors want to thank Thomas Thompson and Andrew Goldfinger of the Applied Physics Laboratory for their assistance in helping in the preparation of this report.

## References

The Staff of the Space Development Department, The Johns Hopkins University Applied Physics Laboratory, and The Staff of the Guidance and Control Laboratory, Stanford University, "A Satellite Freed of All but Gravitational Forces: TRIAD I, J. of Spacecraft and Rockets, Vol. 88, No. 9, pp. 637-644, 1974.

Wolff, M., "Direct measurements of the Earth's gravitational potential using a satellite pair," Journal of Geophysical Research, vol. 74, no. 22, 15 Oct. 1969.
Comfort, G., "Direct mapping of gravity anomalies by using Doppler Tracking between a satellite pair," Journal of Geophysical Research, vol. 78, no. 29, 10 Oct. 1973.
Kaula, W. M., Theory of Satellite Geodesy, Blaidsdell, Waltham, Mass., 1966.
Goldfinger, A., Private communication, November 1978.
Pisacane, V.L., "Filtering of Satellite-to-Satellite Observations: Low-Low Configuration," presented at the American Geophysical Union 1977 Spring Annual Meeting, Washington, D.C., May 30-June 3, 1977; (Acta Astronautica, in press).

Page Intentionally Left Blank

# Improvements of the Gravity Field from Satellite Techniques As Proposed to the European Space Agency

Ch. Reigber
Lehrstuhl für Astronomische und Physikalische Geodäsie
Technische Universität München, Federal Republic of Germany

Abstract. The paper gives a summary of the European Earth Sciences Space Programme and the requirements for earth gravity field mapping resulting from this programme. Three satellite experiments for gravity field improvement proposed to the European Space Agency in the last years are shortly characterized. One of these experiments, the low-low-SST-SLALOM experiment, based on laser interferometry for a "two target-one Spacelab telescope" configuration, is discussed in more detail. Reasons for the low-low concept selection are given and some mission aspects and a possible system concept for a compact ranging, acquisition and tracking system are presented.

## Introduction

Improvements of our knowledge of the earth's gravity field include improvements in the determination of the size of spatial features of the spherical harmonic description of the field as well as the increase in spatial resolution. So far we are on a continuous way of gravity field improvements by space methods since results for the first four zonal harmonics were presented by O'Keefe et al. [1959].

Up to about 1976 besides a slow increase in resolution especially the accuracy of the resolved harmonic components was improved by satellite orbit perturbation analysis combined with surface gravity data analysis. But we were still in the large scale regime which is classified by wavelengths $\lambda$ larger than about 1200 km and which represents signatures of anomaly sources in the deeper and upper parts of the earth's mantle.

Beside using these large scale global gravity models in satellite geodesy for example for satellite orbit determination and global geoid representation Kaula [1972], Marsh and Marsh [1976] and Lambeck [1976] started to correlate patterns of the global free air anomaly field with geological provinces, convection and density inhomogeneites, respectively.

Our picture of at least one equipotential surface of the earth's gravity potential - the geoid - cleared up drastically in ocean areas in the last two years when a number of results from the GEOS 3 altimeter analysis were published [e.g. Anderle 1978, Marsh et al. 1978, Rapp 1977].

A quick glance on the preliminary DOD GEOS 3 geoid [National Research Council, 1978] already demonstrates the high resolution achieved in ocean areas and the strong visible correlation of the geoid with geological structures like oceanic trenches and island arcs. This example shows that satellite geodesy has already obtained geoid features in the medium wavelength ($200 < \lambda < 1200$ km) and short wavelength ($\lambda < 200$ km) region in some parts of the globe and starts to bridge to small scale geodesy, geophysics and geology.

In order to obtain precise medium and short wavelength gravity information over the entire globe - which is essential for our understanding of physical characteristics of material in the asthenosphere, lithosphere and surface topography and of dynamical and thermodynamical processes in the earth's interior - continuous theoretical, technological and financial efforts have to be undertaken in the 80's.

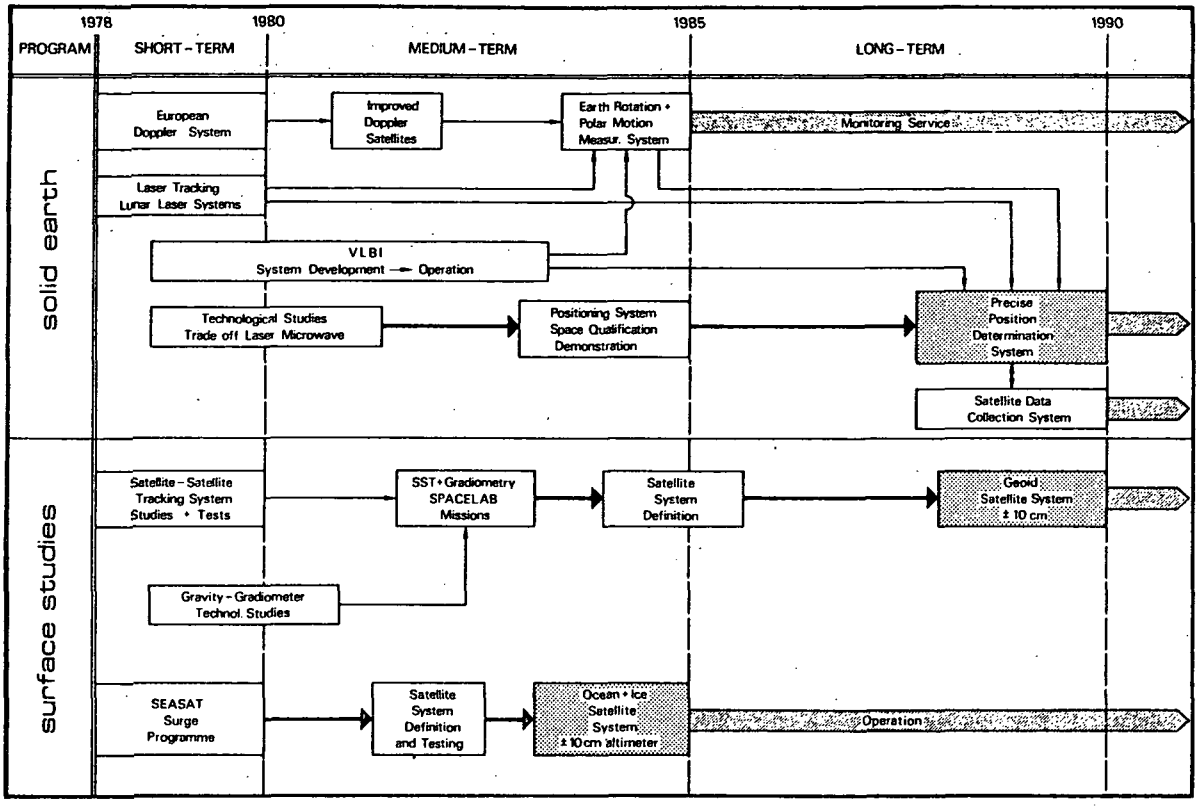
The European Earth Sciences community after having contributed in the past to gravity field mapping - through the developement of methods, ground tracking and satellite instrumentation, surface and tracking data collection and analysis of data - has demonstrated its willingness of contributing to this aim through the Earth Sciences Programme proposed to the European Space Agency during the European Workshop on Space Oceanography, Navigation and Geodynamics (SONG) at Schloss Elmau in early 1978 [ESA, 1978 a].

## Objectives for gravity field mapping in the European Earth Sciences Programme

During this workshop two parallel Earth Sciences Space Programmes for the next 10-15 years were identified along with three supporting ground based programmes which could meet the European possibilities and requirements. The two space programmes are (Figure 1):

- the SOLID EARTH PROGRAMME which is intended to prepare a possible operational system for earthquakes prediction. It implies the developement of a very precise (1-5 cm) position determination system supported by an Earth Rotation and Polar Motion monitoring service and the deployment of a reasonable number of automatic geo-

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280. The Ohio State Univ., Columbus, Ohio 43210.




SONG-Workshop Space Programme Proposal

Fig. 1. Modified version of the Steering Comittee's scheme of the SONG-Workshop Space Programme Proposal [Kovalevsky et al., 1978].

physical stations on ground;

- the SURFACE STUDIES PROGRAMME which is aimed at a global study of the oceans and the ice coverage of the earth for better understanding of the physics of the hydrosphere and its relation with the atmosphere. It implies the launch of a "Geoid Satellite" for highly precise geoid mapping subsequent to possible missions of prototype instrumentation launched by the two forthcoming European Space Transportation Systems SPACELAB and ARIANE.

The reasons for fine structure gravity field mapping in the context of these two programmes are:

Solid earth programme

- Determination of accurate satellite orbits resulting in improved station position, polar motion and earth rotation results and vice versa.
- Investigation and modelling of mechanisms and processes which form and/or move lithospheric plates for developing earthquake prediction models on a regional or global scale.

Surface programme

- Precise geoid determination with spatial resolution down to 200 km as global or regional static reference surface for investigation of general ocean circulation, current systems and tides.
- Recovery of detailed regional structures of the gravity field on continents and continental margins for resource exploitation and lithospheric structure description.

Space Experiments proposed to ESA

Already before the formulation of the Earth Sciences Space Programme three space experiments were proposed to the European Space Agency which - with appropriate mission and system characteristics - could meet most of the gravity requirements in the Solid Earth and Surface studies programmes.

These experiments are:

- The DUMB BELL gravity gradient sensor put forward by G. Colombo to the Agency [c.f. European Space Agency 1976, Colombo et al. 1976].
  The Dumb Bell system is a space borne gradiometer consisting of two spacecrafts which are connected by a long (10-20 km) wire and was proposed to be launched in a low perigee

($q \approx 300$ km) near polar orbit by a conventional launch vehicle. Because of the extremly long arm the system would be much more sensitive to local gravity features than usual space borne gradiometer systems [e.g. Forward, 1973] if system noise could be kept small.

- The TWIN PROBE experiment submitted to ESA by Bertotti and Querzola [1977]. In its proposed form the experiment constitutes a low orbit satellite-to-satellite tracking experiment between Spacelab and a specific arrangement of target satellites, in order to get rid of nongravitational forces. The method consists of two equal pairs of target satellites, each pair composed of two dense and equally shaped satellites but different in mass. With the masses M', M" and the positions $\bar{X}_i'$, $\bar{X}_i''$ of the ith-pair it is possible to derive on the basis of the principle of equivalence the position of an ideal point [Bertotti and Colombo, 1972] $\bar{X}_i = (M'\bar{X}_i' - M''\bar{X}_i'')/(M' - M'')$ which does not feel surface forces in case the surface forces at positions $\bar{X}_i'$, $\bar{X}_i''$ can be assumed to be equal. Applying this method to both pairs of twin probes one obtains in principle the pure gravitational orbits $\bar{X}_1(t)$, $\bar{X}_2(t)$ of the two ideal points where the orbits or the relative motion between the two ideal points have to be reconstruncted from measurements between Spacelab and the four target satellites.
- The SLALOM (Satellite Laser Low Orbit Mission) experiment proposed by Reigber [1978] and Balmino [1978] for regional medium wavelength gravity mapping in the context of a preliminary call for experiment proposals for the early phase of Spacelab utilisation in Europe. This low orbit SST-experiment is a follow-on project of the former DIABOLO-Experiment [Balmino et al., 1976] and Laser-SST-Experiment [Reigber et al., 1976] put forward in connection with the call for proposals for First Spacelab Payload Experiments.

The DUMB BELL experiment - because of Prof. Colombo's association with the Smithsonian Astrophysical Observatory - cannot be considered as an original European experiment.

Since the TWIN PROBE experiment has many overlaps with the SLALOM experiment and because for the latter a mission and system definition study is just under way, [ESA, 1978 b]* some mission and system aspects of only the SLALOM system will be discussed in more detail in the sequel.

## The SLALOM experiment

The SLALOM-experiment is considered to be a "low-low" satellite-to-satellite tracking (ll-SST) experiment. In the usual terminology this characterizes the situation where from a low ($h < 1000$ km) orbiting observing system a spacecraft in a slightly different low orbit and equipped with transponders, corner cubes etc. is tracked. The configuration envisaged for SLALOM is shown in Figure 4. In contrast to the ll-mode we have the "high-low" (hl) mode where the tracking spacecraft is in a high (usually geostationary) orbit. The usual type of intersatellite tracking data is range rates.

Both modes have already been proved practically with good results for long wavelength gravity signal detection in the context of the "high-low" ATS6/GEOS3 and ATS6/APOLLO-SOYUZ SST experiments [Hajela 1977, Marsh et al., 1977, Vonbun et al.,1977] and with almost no result for the gravity signal detection part of the "low-low" APOLLO/SOYUZ Doppler tracking experiment [Weiffenbach et al., 1976].

Before discussing some of the SLALOM mission and system aspects it seems reasonable to explain the reasons for proposing the "low-low" concept solution (c.f. also Rummel et al., 1978). As shown in the last chapter the gravity requirements in the European Earth Sciences programme are mostly related to the medium wavelength domain of the field structure. This information could in principle be obtained from ll-SST as well as from hl-SST, if the same quantity is observed, the measurement accuracy is the same in both cases, the low orbit has the same mean altitude in both configurations and the intersatellite distance in the "low-low" case is not much smaller than the characteristic wavelength of the medium scale region.

This is because the measurement itself has the same sensitivity to medium and short wavelength features of the field in the "high-low" case as in the "low-low" case. The only difference between these two modes is that in the ll-mode the long and medium - wavelength contribution of the spectrum to the observed signal becomes smaller and smaller with decreasing intersatellite di-

* ESA Contract No 3483/78/F/DK(SC) with Messerschmitt-Bölkow-Blohm (MBB) Space Division (G. Barthel, T. Halldorson, D. Meissner), Lehrstuhl für Astron. und Physik. Geodäsie and Sonderforschungsbereich 78 (SFB 78) (Ch. Reigber, R. Rummel), Groupe de Recherches de Geodesie Spatiale (GRGS) (G. Balmino, L. Castel)

TABLE 1. R.M.S. Velocity difference in mm/sec between degrees $l_1$ and $l_2$ at altitude h = 250 km for different intersatellite distances

| Inter-satellite Distance | l=2-18 | 19-36 | 37-72 | 73-180 | 181-2000 |
|---|---|---|---|---|---|
| | long | | medium | | short |
| 35 500 km | 31.548 | 1.567 | 0.379 | 0.041 | $0.190.10^{-3}$ |
| 300 km | 5.636 | 1.028 | 0.325 | 0.040 | $0.190.10^{-3}$ |
| 200 km | 3.945 | 0.806 | 0.277 | 0.038 | $0.189.10^{-3}$ |
| 100 km | 2.076 | 0.481 | 0.185 | 0.029 | $0.179.10^{-3}$ |
| 50 km | 1.066 | 0.264 | 0.109 | 0.019 | $0.145.10^{-3}$ |
| 10 km | 0.218 | 0.057 | 0.025 | 0.005 | $0.480.10^{-4}$ |
| 1 km | 0.022 | 0.006 | 0.003 | $0.5.10^{-3}$ | $0.546.10^{-5}$ |

stance so that the low frequency contribution is more and more damped. These conclusions can be drawn from Table 1 and Figures 2 and 3.

The velocity difference variances in these graphs were derived from the expression [ESA, 1978 c]

$$\sigma^2(|\dot{\bar{X}}_{12}|)_l = \frac{1}{|\dot{\bar{X}}_m^2|}\left(1-\left(\frac{r_P}{r_Q}\right)^{l+1}\right)^2 \sigma^2(T)_l \quad (1)$$

where

$\dot{\bar{X}}_m = \left(\frac{GM}{r_P}\right)^{1/2}$ ... mean velocity of satellite

$\sigma^2(T)_l = \left(\frac{R^2}{r_P^2}\right)^{l+1} \frac{R^2}{(l-1)^2} c_l$ ... degree variances of disturbing potential [c.f. Rummel 1975]

$c_l = \frac{425.28\ (l-1)}{(l-2)(l+24)}\ \text{mgal}^2$ ... gravity anomaly degree variance model [Tscherning and Rapp 1975]

and the random measurement noise in the velocity differences is modelled by, [ESA, 1978 c],

$$\epsilon^2(\dot{\rho}_{12}) = m_o^2\, e^{-c\psi} \quad , \quad (2)$$

with

$m_o$ ... noise level

c ... inverse relative correlation length

$\psi$ ... spherical distance.

From this error model, which approximates white noise, the degree-order variance is obtained as

$$\epsilon^2(\dot{\rho}_{12})_{lm} = \frac{1}{2}\, m_o^2\, \frac{P^2}{t_c^2} \quad , \quad (3)$$

with P the orbital period and $t_c$ the correlation length in time units. Defining the maximal resolution which can be achieved for a definite $m_o$ by a signal to noise ratio of 1 : 1 one can derive the additional conclusions:

- for a full medium wavelength description of gravity the altitude of the lower orbit has to be less than 300 km
- a measurement noise level of not larger than ± 10 µm/sec is a definite requirement
- there is no gain by closer intersatellite distances if the measurement accuracy is not increased simultaneously.

Because of the high resolution of ± 10 µm/sec for the range rate which is not achievable with doppler measurements at microwaves we think laser velocity measurements by interferometrical methods are the only way out. This leads - because of laser energy requirements, pointing requirements etc. - to the feasibility of only a "low-low" experiment with an intersatellite distance of not larger than about 350 km.

The Shuttle/Spacelab system with its common user facilities, subsatellite ejection and crew intervention possibilities is thought to be well suited as a platform for such a sophisticated and probably heavy instrumentation.

## Mission aspects

The main objective of the SLALOM ex-

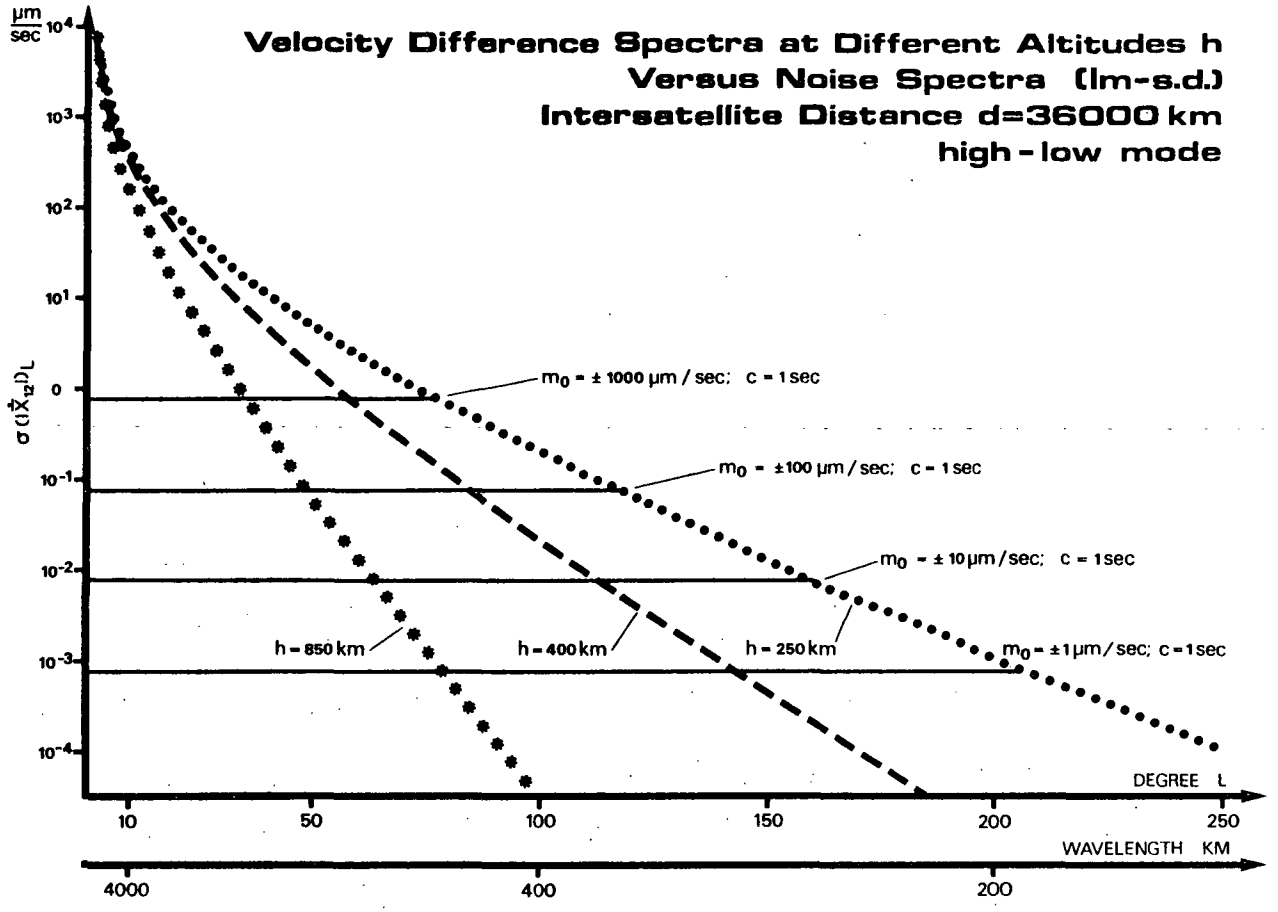



Fig. 2.

periment is to perform with a laser ranging instrumentation during a 7 - 10 days Shuttle/Spacelab mission period over a specific geographical area (e.g. The East Mediterranean region) a cycle of range and range rate measurements. These measurement profiles will in a post flight analysis be used for gravity parameter recovery over this limited area or part of it. Since the resolution of gravity information will strongly depend on the cross-track spacing of the observed profiles above the area of interest the drift of the orbit has to be selected in such a way that an optimal number of ground tracks in the measurement area is obtained without retracing tracks within the mission period. This means to ask for a near resonant Shuttle orbit with a drift period nearly as long as the mission duration. A drift period of 10 days can be reached for a mean semimajor axis $\bar{a}$ = 6603 km, mean inclination $\bar{i}$ = 50 deg. and mean eccentricity $\bar{e}$ = 0.002. The orbit drift rate would be -2.25 deg./rev. This is the optimum drift rate that can be obtained for a Shuttle mission duration of 10 days.

As explained in the next chapter the frequency translated Michelson interferometer principle will be applied for range rate measurements using a highly stabilized continuous wave (CW) gas laser. Spacelab will serve as platform for the laser telescope and subsystems and a subsatellite - equipped with corner cubes and released from Spacelab - as passive target.

The quantity observed in this flight configuration is the rate of change of the distance between the ranging instrumentation zero point and the target reflection point. This raw measurement reflects not only the instantaneous motions of the objects due to differential gravitational and nongravitational forces accelerations but also the movements of the whole Shuttle/Spacelab system due to internal disturbances, the motions of both Spacecrafts around their center of mass (C.M.) and the relative motion of the ranging instrumentation zero point with respect to the C.M. of Shuttle/Spacelab system. All but the gravity field induced effects have to be eliminated or modelled before the data can be used for gravity parameter determination.

This is a very difficult task even

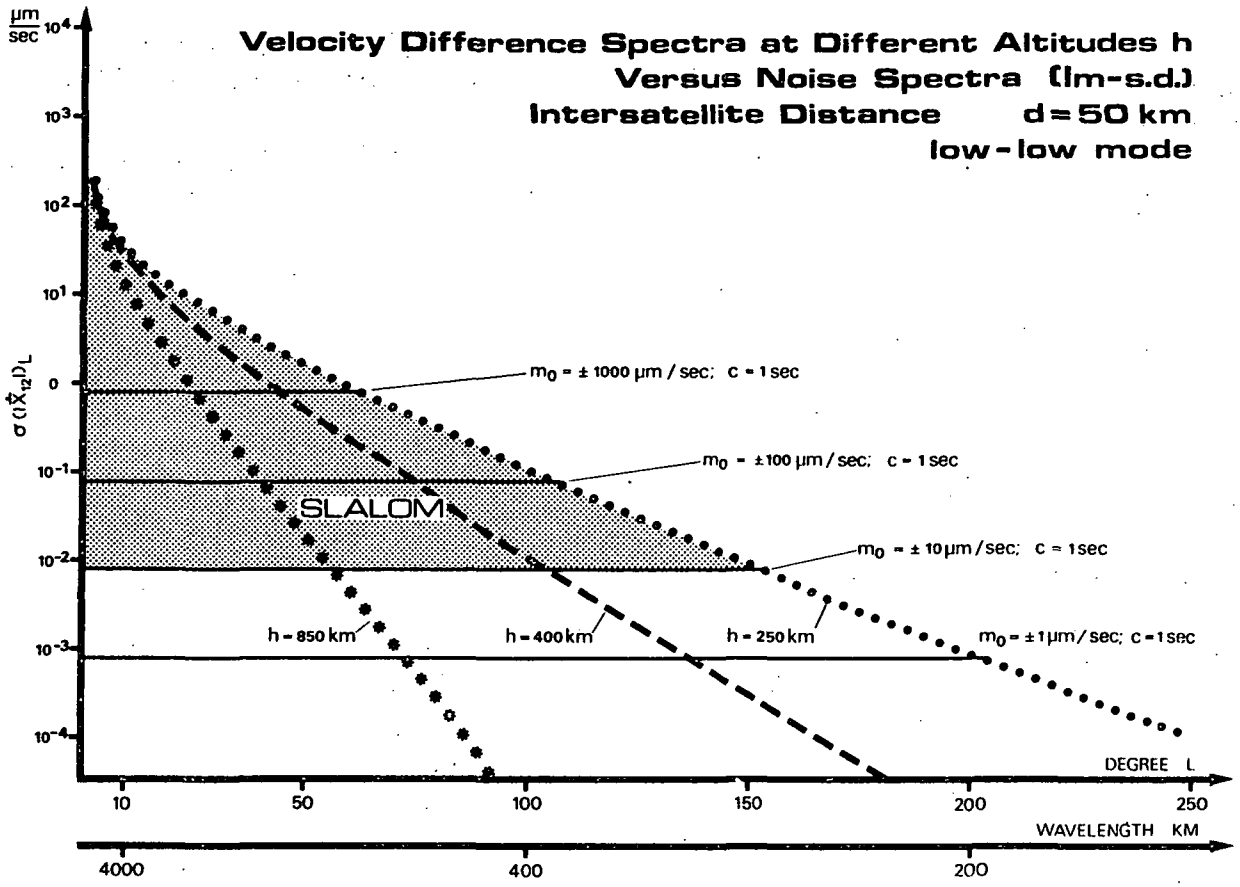



Fig. 3.

when introducing sophisticated hardware for attitude control, C.M. determination and elimination of surface forces effects.

Out of four experiment configurations which were identified in the SLALOM System/Mission definition study [ESA, 1978 b] a concept based on two passive target satellites of the same cross section-to-mass ratio and one laser telescope on the Spacelab pallet was found to be the most appropriate solution for removing all problems mainly caused by critical orbiter operations, crew motions and C.M. shifts. Additionally, because of the same area-to-mass ratio differential surface forces effects will be minimized. This configuration together with the basic equations is illustrated in Fig. 4.

In this configuration the quantities simultaneously observed by one instrument on board of the Shuttle are range and range rates to the two targets and the angular distance $\psi$ between $T_1$ and $T_2$. The finally analyzed observation is the range rate $\dot{\rho}_{12}$ between $T_1$, $T_2$

$$\dot{\rho}_{12} = \frac{1}{\rho_{12}}(\rho_1\dot{\rho}_1+\rho_2\dot{\rho}_2 - \cos\psi(\dot{\rho}_1\rho_2+\dot{\rho}_2\rho_1) + \rho_1\rho_2 \sin\psi\,\dot{\psi}) \quad (4)$$

with

$$\cos\psi = \bar{e}_1\cdot\bar{e}_2;\ \sin\psi\,\dot{\psi} = \bar{e}_1\cdot\dot{\bar{e}}_2+\dot{\bar{e}}_1\cdot\bar{e}_2 .$$

From the variance expression of this quantity one obtains as configuration constraints and measurement accuracy requirements if the range rate accuracy of $\pm$ 10 μm/sec should be propagated into $\dot{\rho}_{12}$

- one target as close to the Shuttle as possible during the experiment phase ($\rho_2 < 20$ km)
- intersatellite angular distance $\psi$ small; if possible even smaller than the $2^o \times 2^o$ field of view (FOV) of the telescope ($\psi < 2^o$)
- high accuracy for the ranges ($\sigma(\rho) \approx \pm 0.1$ m), precise angular distance ($\sigma(\psi) \approx \pm 10''$) and very precise rate of change of angular distance ($\sigma(\dot{\psi}) \approx \pm 0''.01$).

The two configuration conditions can be satisfied by deploying from the Shuttle the two target satellites with a definite velocity change $\Delta\bar{v}$, exactly controlled in amount and direction or by a controlled Shuttle deceleration.

With an area-to-mass (A/M) ratio of the target satellites larger than the Shuttle A/M (2.045E-3 in X(POP) ± Z nadir drift mode) by 1-30% and ejection in along-track direction with a $\Delta v$ between 1-40 cm/sec target orbits could be reached which would be trackable from the Shuttle within a 350 km distance during the whole mission. The usual form of the relative distance in the $(\rho,t)$-plane is a parabola. An example that would fit to the SLALOM mission requirements is shown in Figure 5.

For this example within the experiment period of about 6 days no safety problems would occur and all before mentioned requirements are fulfilled with $\rho_1 < 350$ km, $\rho_2 < 20$ km, $\psi < 1^o$ .

With the mentioned accuracies for the range $\rho$ and the angular quantities $\psi$ and $\dot{\psi}$, the maximum standard deviation of the range rate, $\sigma_{max}(\dot{\rho}_{12})$, would be ± 13 µm/sec.

Assuming that the measurements are not influenced by optical disturbances the final range rate $\dot{\rho}_{12}$ is solely affected by the difference in the instantaneous state of the two target satellites and the motion due to differential gravitational and surface forces accelerations. All variations induced by non earth gravity influences in the raw data $\dot{\rho}_{12}$ have to be eliminated.

This is apart from the differential air drag effect easily be done because the effects are either small or can be modelled precisely. At SLALOM altitude the differential drag effect will probably not be zero but will be small because of the same cross section-to-mass ratio of both target satellites.

Such small drag effects are caused by small scale density changes in the upper atmosphere and have to be expected especially during periods of increasing geomagnetic activity. It is difficult to give some realistic figures of mass density changes for horizontal distances of 50 to 250 km at an altitude of 225 km.

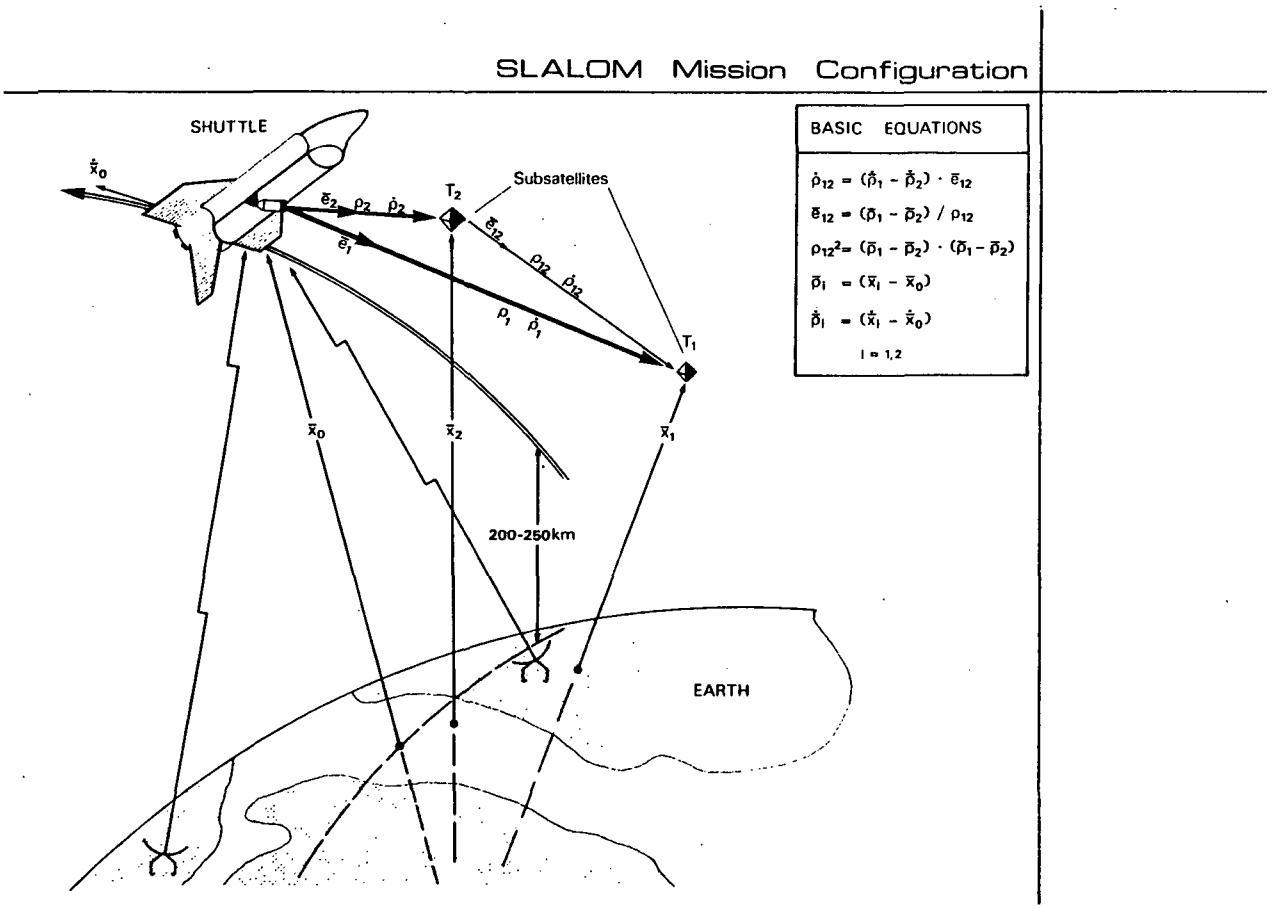



Fig. 4. SLALOM baseline experiment configuration

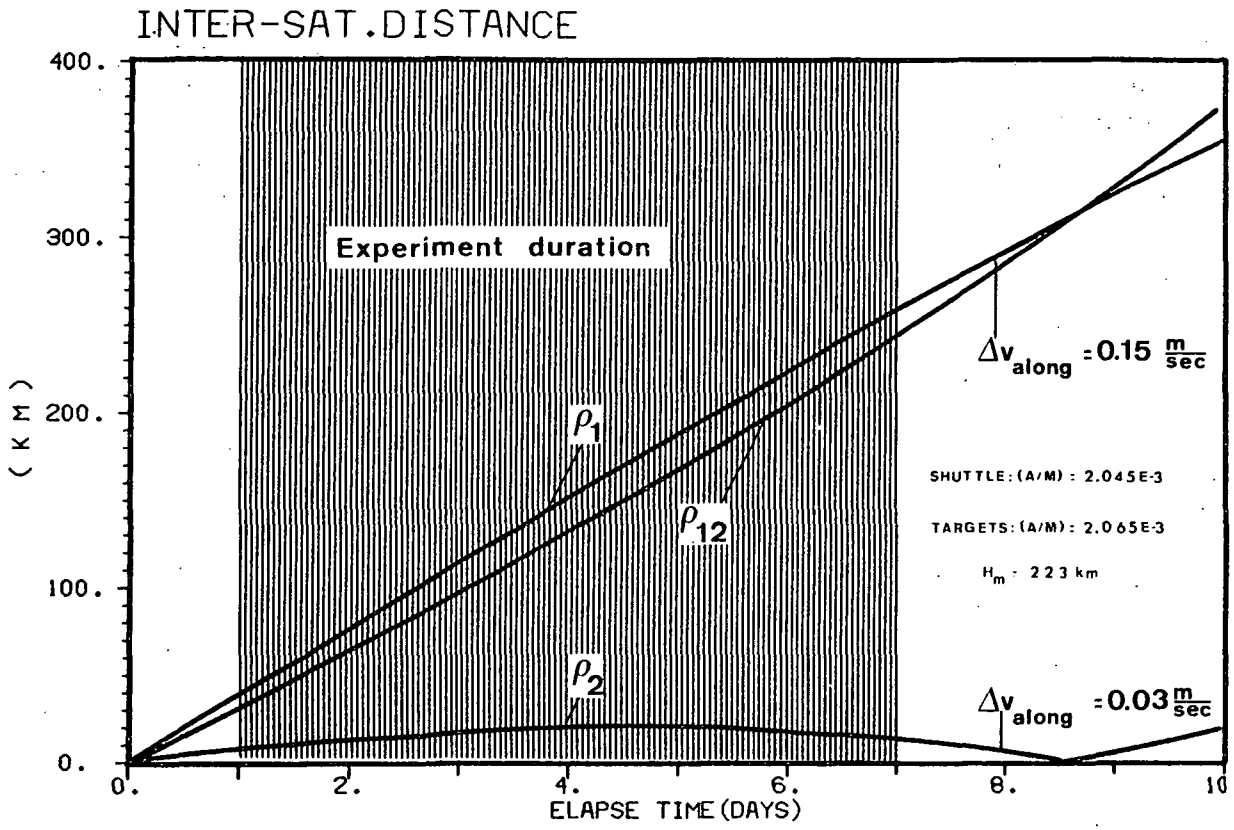



Fig. 5. Intersatellite distances (c.f. Figure 4) obtained from orbit integration using CIRA 65 standard atmosphere.

The only information we have is from in situ measurements of the neutral atmospheric composition obtained by gas analyzers aboard of satellites [Trinks and von Zahn, 1975]. These data indicate the presence of small scale density variations with peak-to-peak amplitudes of 5% up to 30% under disturbed conditions and with amplitudes of 1% - 10% during quiet-time conditions [Prölss and von Zahn, 1975, Prölss and Fricke,1975]. The wavelengths of these fluctuations range from 100 to some hundred kilometers but can probably even be shorter.

Taking as a reasonable number for the relative difference of the atmospheric density at the two positions $\bar{x}_1$, $\bar{x}_2$ with horizontal distance of 200 km a value of $\Delta R/R = 0.1$ will result at $h = 225$ km in a differential drag acceleration $\Delta F_D$ of

$$\Delta F_D = F_D \cdot \Delta R/R \approx 1.10^{-6} \text{ m/sec}^2 \doteq 0.1 \text{ mgal}.$$

As it is seen from Figure 6 - which shows gravity induced acceleration difference spectra (derived by differentiating eq. (1)) and acceleration noise spectra at $h = 250$ km - this unmodelled drag acceleration would be small enough to allow the recovery of medium wavelength structures of the gravity field.

On the other hand unmodelled residual accelerations of this order in satellite hight would result in 1-2 mgal errors of recovered gravity information on the earth surface because of the strong amplification of the medium frequency components of the noise spectrum in the downward continuation process by a factor of about 20 [c.f. Rummel,1975].

Taking the "two target - one telescope" concept as baseline experiment configuration the mission operations will start with alignment of the optical system, switching on laser to standby mode, possible orbiter manoeuvres for subsatellite launch and then the sequential deployment of the two target satellites along track with a positive impulse and a well defined spin. At very near distances initial acquisition could be performed by the orbiter KU-band radar system. As soon as an initial orbit is computed, acquisition and tracking can be shifted to the laser instrumentation. When the angular distance between the two target satellites is small enough so that the satellites remain within the field of view of the telescope (approximately after one day) tracking of both targets will be performed simultaneously.

Before passing the selected area of

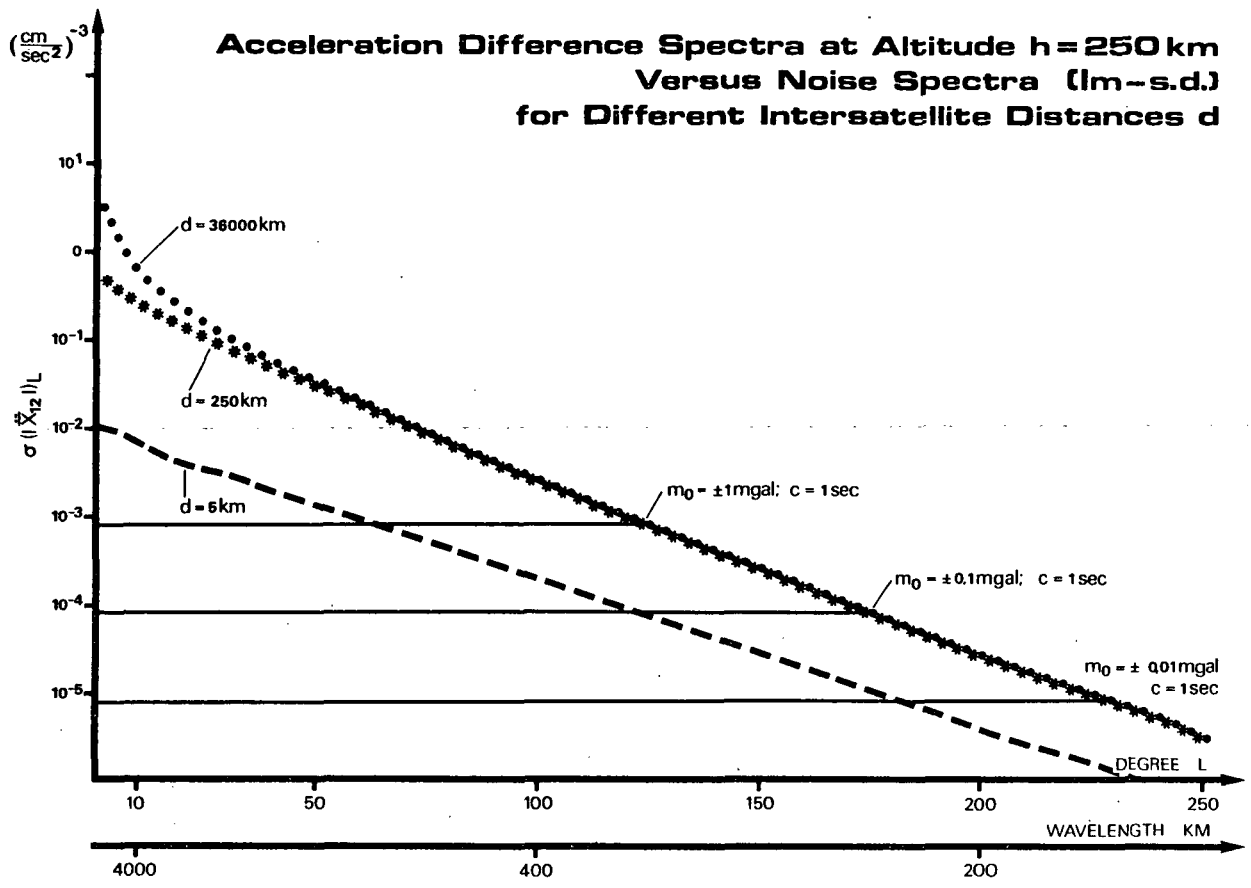



Fig. 6.

investigation the laser telescope will be reorientated to the predicted target position by means of a pointing platform. Reacquisition will be performed with the aid of the laser instrumentation and the instrument will then be switched to tracking mode. Then range, range rate and angular position measurements are obtained with a high repitition rate over a period of about 5 - 6 minutes for both target satellites. These data will be stored for later precise double or triple differential orbit correction and gravity signal recovery.

## System Concept

The SLALOM system consists of two targets - a long range and short range satellite - and a laser ranging instrumentation mounted on a support structure which will be directly attached on the orbiter bay fittings.

The main functions of the ranging system are range measurements, range rate measurements and acquisition and tracking. All these functions have to be performed simultaneously for two target satellites with high accuracy.

In a laser ranging instrumentation study performed by MBB and conducted by ESTEC possible solutions for the different functions were identified [c.f. ESTEC, 1978].

In case of ranging an instrumentation based on pulse transit time or phase shift methods is considered to give comparable results. For pulse transit time measurements a Nd:YAG laser or laser diodes are considered as effective transmitter candidates whereas for phase shift methods continuous wave solid state and gas laser like laser diodes, HeNe- ($\lambda$=3.37 $\mu$m) Argon- ($\lambda$=0.51 $\mu$m) and $CO_2$- ($\lambda$=10.6 $\mu$m) gas lasers could be used as emitters.

Out of the three methods which are mainly used for velocity measurements with lasers - range increment measurement method, doppler shift measurement of microwave modulated on a CW laser, phase shift (or doppler shift) measurement by interferometrical methods - only the interferometrical methods are capable of reaching the required range rate accuracy of $\pm$ 10 $\mu$m/sec for SLALOM. Possible lasers for emission are $CO_2$-, HeNe-, and Argon gas lasers. With the

classical Michelson Doppler interferometer [Watrasiewicz and Rudd, 1976] no possibility exists for distinguishing between positive and negative range rates as they appear in the SLALOM experiment. This difficulty can be avoided by changing to a frequency translated Michelson interferometer which has the capability of bidirectional counting. This is achieved by frequency modulating the local oscillator with a single constant frequency. The problem of low signal-to-noise ratio in case of large frequency - receiving - bandwidth (which is necessary in case of SLALOM because of large range rate changes) can be overcome by frequency off-setting the local oscillator. A basic requirement for a high doppler shift frequency resolution is a high short time laser stability. For a 10 μm/sec resolution the stability of the laser in the detector's integration time has to be better than $10^{-11}$ for a $CO_2$ laser at $\lambda$ = 10.6 μm and about 20 times better for a HeNe laser with $\lambda$ = 0.6 μm.

The SLALOM ranging instrumentation functions for target acquisition and tracking are: Illumination of the satellites for optical acquisition, directional sensing of the targets and beam deflection for scanning over the field of view of the telescope (acquisition) and over a partial area of it around the most probable position of the target (tracking method).

For sake of simplicity the laser illumination should be performed with one of the already existing lasers for range or range rate measurements. This solution would guarantee coaxial alignment with the instrument's pointing direction. For optical sensing of the targets the Instrumentation Study [ESTEC, 1978] proposes the use of image dissector tubes - as used for example in the IPS star tracker - which are synchronized with the motion of beam deflectors based on rotating mirrors or piezoelectric drives for fine pointing.

In Figure 7 a schematic diagram for a possible SLALOM instrumentation with a two target ranging and tracking capability is shown. A Ritchey Chretien type telescope with an aperture of 0.2m is used for transmission and likewise for reception. Within this telescope all optical subunits will be integrated leading to minimized optical distortions because of identical pointing axes and same thermal conditions.

Finally something remains to be said

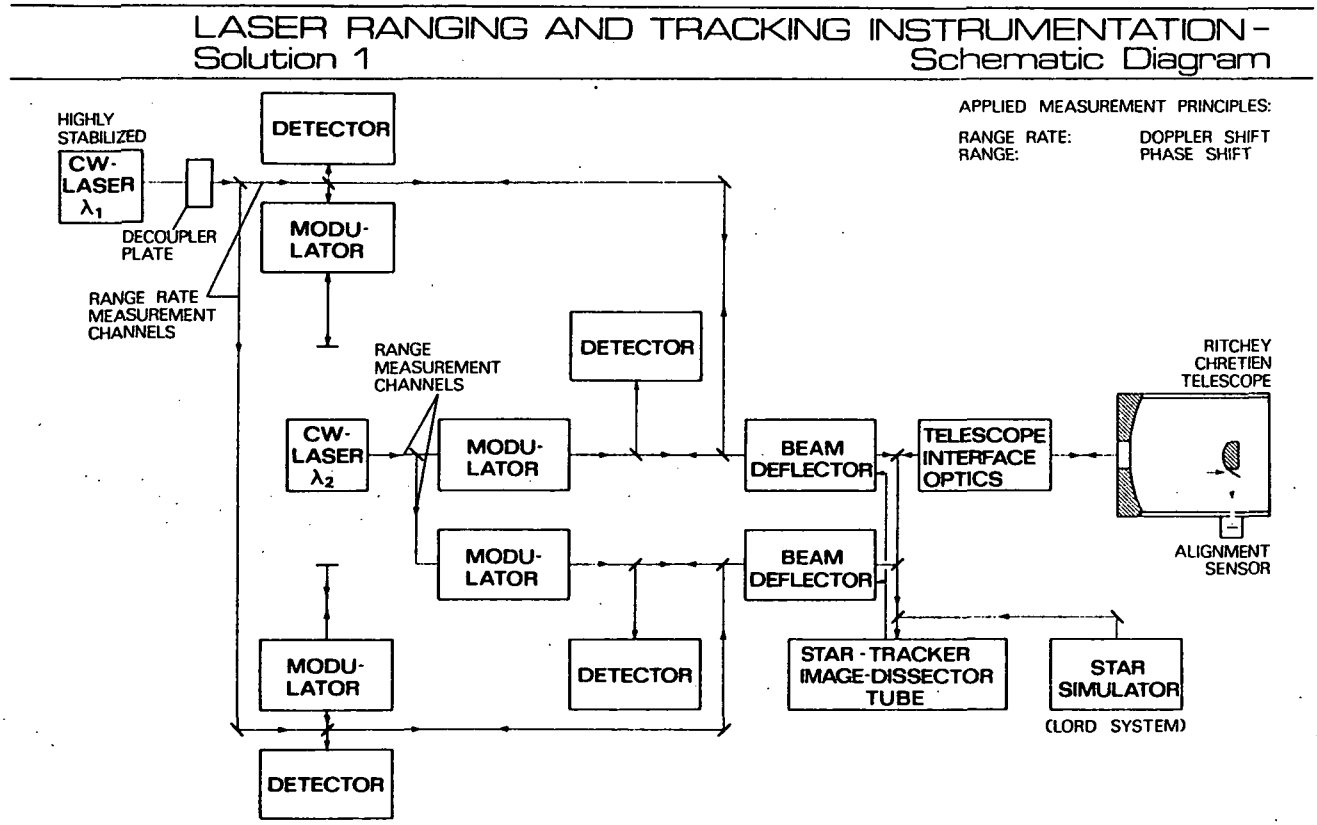



Fig. 7. Possible laser ranging and tracking instrumentation for SLALOM experiment (Schematic diagram from ESTEC, 1978: Laser Ranging Instrumentation-Interim Report).

about the subsatellites lay-out. For this one has to consider that

- the area to mass ratio for both subsatellites has to be equal and within some percent similar to the Shuttle cross section-to-mass ratio for the experiment flight attitude (e.g. $2.045 \times 10^{-3}$ $m^2/kg$ for (-Z/LV; Y/VV)-mode)
- the reflecting cross section has to be large because of the maximum range of 300 km
- doppler shift due to rotation should be minimal.

These conditions are best met for a subsatellite covered with corner cubes and which is of octahedron type. With a wall assembly made of glas ceramics plus for instance lead inlay and a 50 cm length of each octahedron axis the satellite mass would be about 53 kg. This would result in the required A/M ratio of $2.045 \times 10^{-3}$.

## Conclusion

Although the early Shuttle/Spacelab missions are to a certain extent not ideal with respect to earth surface coverage and although still a number of problem areas exist for the SLALOM experiment - e.g. laser stability requirements, angular distance rate accuracy, high simultaneous pointing requirements, blackouts and doppler shifts due to target rotation, slow convergence in the double differential orbit improvement process and strong amplification of unmodelled contributions of the range rate signal in the gravity recovery process - we believe that the SLALOM experiment will provide a compact gravity sensing system for medium and fine structure resolution in the near future. It may also have its implications on the developement of spaceborne laser systems for other geodetic applications.

The handicap of the short Spacelab mission duration is on the other hand balanced out by a number of advantageous features of the Shuttle/Spacelab system for a SLALOM like experiment and the possibility of repeating the experiment in other missions. For the future it is even possible to think in terms of longer Spacelab missions or a retrievable detached laser platform.

Acknowledgement. This study was supported by ESA Contr.No. 3483/78. The author wishes to thank Mr. Barthel and Mr. Halldorson (MBB) for providing details on system aspects and Dr. Rummel (SFB 78) for numerous valuable discussions and many helpful suggestions.

## References

Anderle, R.J., The Ocean Geoid Deduced from GEOS-3 Satellite Radar Altimetry Data, presented at Second International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics, Athens 1978.

Balmino, G., Orszag, A., Lefebvre, M., Dumoulin, J., and Pieplu, J.L., Diabolo-Laser, FSLP Proposal AO5/55/EO/F to ESA, 1976.

Balmino, G., and Lefebvre, M., SLALOM Proposal to CNES-Headquarters, 1978

Bertotti, B., and Querzola, B., Measurement of the gravitational field of the earth by the twin probe method, proposal to ESA, 1977.

Bertotti, B., and Colombo, G., Precision measurement of the sun's gravitational field by means of the twin probe method, Astroph. and Space Sci. 17, 223, 1972.

Colombo, G., Arnold, D.A., Binsack, J.H., Gay, R.H., Grossi, M.P., Lautman,D.A., and Orringer, O., DUMB BELL Gravity Gradient Sensor. A New Application of Orbiting Long Tethers, SAO Reports on Geoastronomy 2, 1976.

European Space Agency, Space Oceanography, Navigation and Geodynamics, Proceedings of a European Workshop held at Schloss Elmau, Germany, Jan. 1978, ESA SP-137, 1978 a.

European Space Agency, Slalom Mission/ System Definition Study, Midterm Report, 1978 b.

European Space Agency, Slalom Mission/ System Definition Study, Final Report (in preparation), 1978 c.

European Space Agency, DUMB BELL Mission for Gravity Field and magnetospheric Studies, Report on the Mission Definition Study, DP/PS(76) 22, 1976.

European Space Research and Technology Center, Laser Ranging Instrumentation, Interim Report, 1978

Forward, R.L., Review of Artificial Satellite Gravity Gradiometer Techniques for Geodesy, Hughes Aircraft Company Research Laboratories, Research Report 489, 1973.

Hajela, D.P., Recovery of 5° Mean Gravity Anomalies in Local Areas from ATS-6/GEOS-3 Satellite-to-Satellite Range Rate Observations, Air Force Geophysical Laboratory Technical Report 0272, 1977.

Kaula, W.M., Global gravity and Tectonics. In: E.C. Robertson, J.F. Hays and L. Knopoff (editors), The Nature of the Solid Earth, Mc Graw-Hill, 1972.

Kovalevsky, J., Melchior, P., Sigl, R., and Veis, G., Report of the Steering Committee. In: ESA SP-137, Proceedings of the Space Oceanography Navigation and Geodynamics Workshop, 1978.

Lambeck, K., Lateral density anomalies in the upper mantle. J. Geophys.Res. 81, 1976.
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 Technical Memorandum 79553, 1977.
Marsh, B.D., and Marsh, J.G., On Global Gravity Anomalies and Two-Scale Mantle Convection, J. Geophys. Res. 81,29, 1976.
Marsh, J.G., Martin, T.V., Mc Carthy, J.J., and Chovitz, P.J., Estimation of Mean Sea Surfaces in the North Atlantic the Pacific and the Indian Ocean Using GEOS-3 Altimeter Data, presented at second International Symposium on the Use of Artificial Satellites for Geodesy and Geodynamics, Athens 1978.
National Research Council, Prospects in Geodesy, Report of Committee on Geodesy, National Academy of Sciences, Washington 1978.
O'Keefe, J.A., Eckels, A., and Squires, R.K., The Gravitational Field of Earth, Astro. J., 64, 1959.
Prölss, G.W., and Fricke, K.H., Neutral Composition Changes During a Period of Increasing Magnetic Activity, Planet. Space Sci. 24, 1976.
Prölss, G.W., and von Zahn, U., Large and small scale changes in the disturbed upper atmosphere, J. Atmos. Sci.38, 1976.
Rapp, R.H.,Mean Gravity Anomalies in Sea Surface Heights Derived from GEOS-3 Altimeter Data, Department of Geodetic Science, 268, The Ohio State University, Columbus, 1977.
Reigber, Ch., Sigl, R., Schneider, M., and Ilk, K.H., Laser SST-Experiment FSLP Proposal AO5/83/EO/D to ESA, 1976.
Reigber, Ch., SLALOM-Experiment, Proposal for Early Phase of Spacelab Utilisation to Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt, BPT, 1978.
Rummel, R., Reigber, Ch., and Ilk, K.H., The Use of Satellite-to-Satellite Tracking for Gravity Parameter Recovery: In ESA SP-137, Proceedings of the Space Oceanography Navigation and Geodynamics Workshop, 1978.
Rummel, R., Downward Continuation of Gravity Information from Satellite-to-Satellite Tracking or Satellite Gradiometry in Local Areas, Department of Geodetic Science 221, The Ohio State University, 1975.
Tscherning, C., and Rapp, R.H., Closed Covariance Expressions for Gravity Anomalies, Geoid Undulations and Deflections of the Vertical Implied by Anomaly Degree Variance Models, ibid. No 208, 1974.
Trinks, H., and von Zahn, U., The ESRO 4 gas analyzer, Rev.Sci.Instrum. 46, 1975.
Vonbun, F.O., Kahn, W.D., Wells, W.T., and Conrad, T.D.,Gravity Anomalies Determined from Tracking the Apollo-Soyuz, NASA Technical Memorandum 78031.
Watrasiewicz, B.M., and Rudd, M.J., Laser Doppler Measurements, Butterworths & Co (Publishers)Ltd., 1976.
Weiffenbach, G.C., Grossi, M.D., and Shores, P.W., Apollo-Soyuz Doppler-Tracking Experiment MA-089, Final Report, Smithsonian Astrophysical Observatory, 1976.

# Problems in Determining Sea Surface Topography

John A. Whitehead, Jr.
Woods Hole Oceanographic Institution
Woods Hole, Massachusetts 02543

Abstract. Anticipated problems for determining ocean dynamics signals from sea surface topography are discussed. The needs for repeated tracks are listed if oceanic tides or ocean turbulence are to be determined.

People want to observe sea surface topography for many reasons and I would like to discuss the ocean dynamics aspects, in contrast to the geodynamics aspects which I feel have been presented by others.

Geodesists and geophysists are interested in things that are shown in Figure 1, which I will call oceanographic noise. This is presented as a function of typical wavelength. Approximate amplitudes are also given. In the middle we see the footprint size of SEASAT represented as a footprint. Features smaller than the footprint lie outside our interests. A big signal of amplitude 40 meters or so comes from trenches. The various global highs and lows are not particularly related to surface tectonics in any clearly understood way that I know of. They are up to 100 meters in amplitude. Ridge systems give a very small signal but they do apparently affect the second derivatives very strongly.

I'm not really concerned with the geophysical signal here. What I'm interested in telling you about is shown in Figure 2, entitled Geodesists and Geophysists noise.

In this figure we present the deviation of the ocean surface from the geoid as a function of length scale. The span of such wavelengths is represented by a horizontal line and I caution you that this is meant to be very crude. Equally crude estimates of amplitudes are included. Starting on the left we see capillary waves, spray, foam, and seaweed in the millimeter to centimeter range. Next come wind waves in the $10^{-1}$ to $10^{3}$ meter range and as we know, these are very variable both spatially and temporally. Amplitudes vary from a tenth of a meter to ten meters. We believe there is a fairly good sized minimum (we're not really sure) in the wavelength range from about a kilometer to roughly 50 km. In this range lies the footprint size of SEASAT, represented as a footprint. Above 50 km or so we get a strong contribution from oceanic eddies. These are low frequency events with periods of weeks or longer and with amplitudes from 20 cm to one meter. There are stationary counterparts to these in various frontal zones near the Gulf Stream, the current that goes around Antarctica, and so forth. From 1000 to 5000 km lie tides, which are really shallow water gravity waves, influenced by rotation of the earth, and with easily identified, sharp frequencies of various sorts. Equatorial currents and the non-frontal aspects of the big ocean currents appear to be at most a few thousand km. Little is known of any structures of 10,000 km, but there may be some long basin modes (periods up to about 30 hours).

First, let us note that the wind-driven waves contribute the most height and lie below the footprint size. This causes me as much worry as anything, although I've been assured that empirical corrections can be made down to the 10 cm level of SEASAT. One additional feature is that

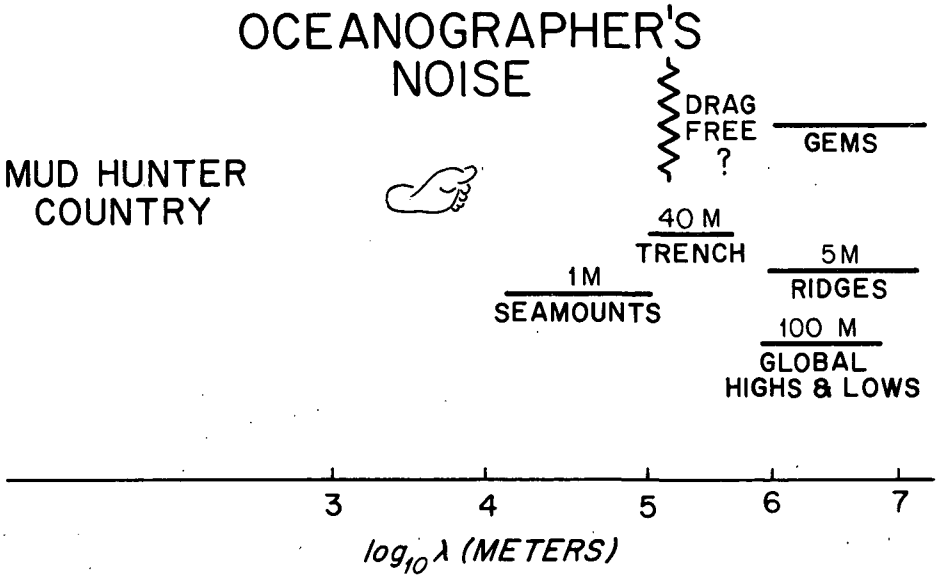



Fig. 1. Oceanographer's noise.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

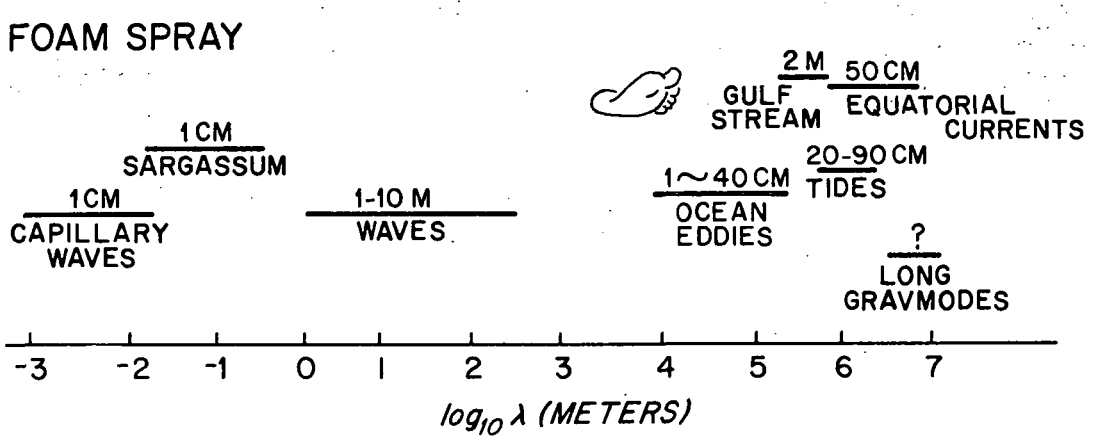



Fig. 2. Geodesists' and Geophysicists' noise.

patches of waves (under storms) are from 100 km to 1000 km in size generally, and thus overalp ocean eddies.

Ocean eddies, of the size 100 km or so, have been observed by standard oceanographic methods for the past ten years. Figure 3 (from MODE Atlas (1977)) shows estimated pressure at 150 meters depth from float and hydrographic data, in units of equivalent centimeters of head. The region encompasses a 300 × 300 km region and the eddies are of order 100 km. The amplitude peak to peak is about 25 centimeters. To see how this signal stands up against the geoidal deflection, Figure 4 shows a sketch of the GEOS altimeter output for the region. There is change in depth of the geoid of about 2 to 3 meters. It is obvious from this that the oceanic turbulence is going to be a small relative signal indeed and we would really need to take continuous tracks for intervals of well over a month in order to get the signal.

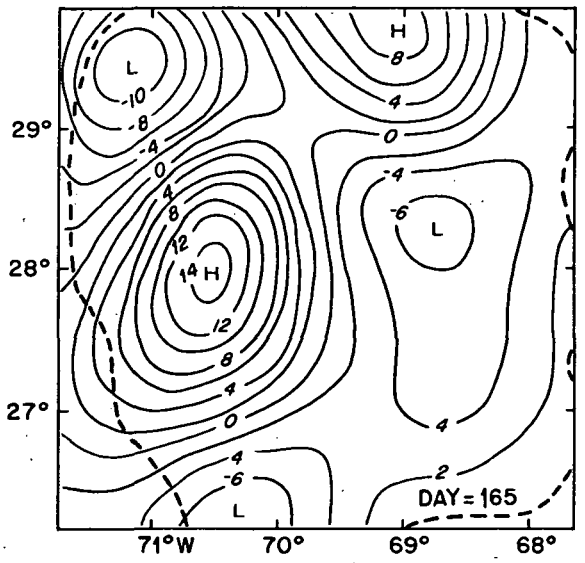



Fig. 3. Pressure in units of equivalent centimeters of head, as deduced from float and density data. Contour interval is 1 cm. Source: MODE Atlas, p. 131.

Not all eddies have such a small amplitude. Figure 5 shows some typical sections of isotherms in relatively quiet regions of the ocean. The undulations in the isotherms are due to such turbulence as we saw on the previous slide with estimated surface deflections of 10 cm. However, Figure 6 shows a section further north, where there is evidence of very much stronger eddy activity. These may get up to 50 or 60 centimeters surface deflection.

One should be able to take advantage of the time dependence of these eddies to filter out some of their contributions to "geodesists noise." Steady counterparts to these eddies such as the edge of the Gulf Stream or the Antarctic circumpolar front won't be so easily isolated except that, fortunately, they tend to wiggle about. There will be no way to determine the gravity field to 10 centimeters in the 100 km wavelength region in my opinion short of extensive ship surveys, and handling this "wiggle of streams" problem is essential if those currents are to be resolved.

Tides are another feature whose time dependence may aid in their analysis, especially since

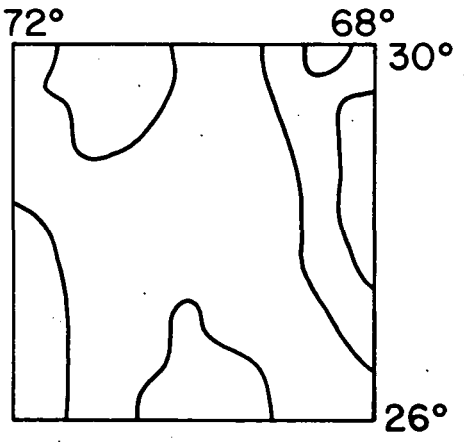



Fig. 4. Geoidal variation as seen from the GEOS-3 satellite of the MODE area. Contour interval 1 meter.

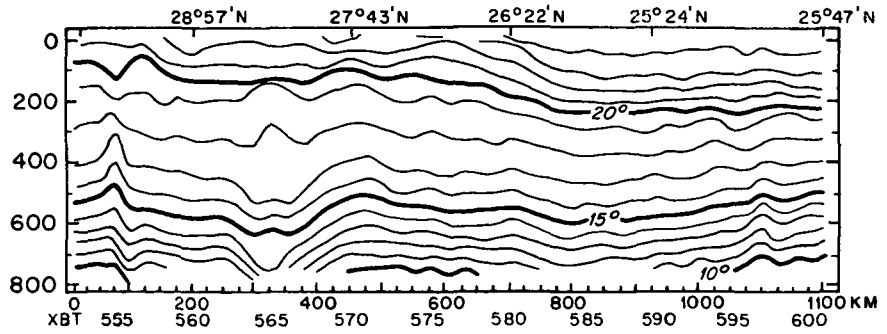



Fig. 5. Section of isotherms of Atlantic Ocean at approximately 30° north.

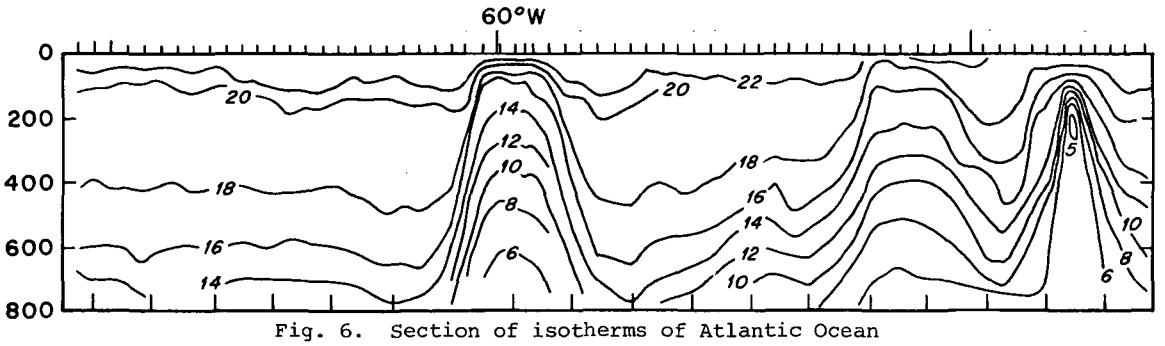



Fig. 6. Section of isotherms of Atlantic Ocean at approximately 50° north. There is evidence of much stronger eddy activity.

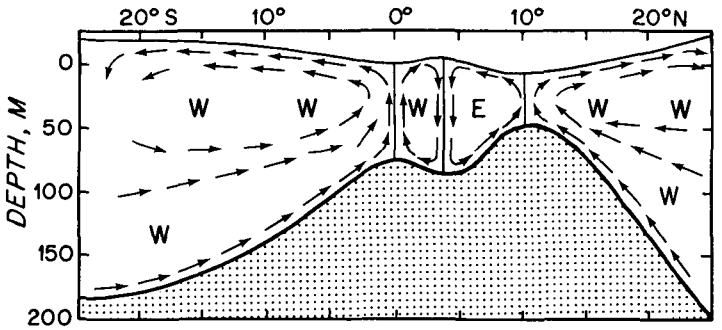



Fig. 7. Hypothesized north-south section of the equatorial ocean. The surface deflection has never been measured.

their frequencies are so sharp. It would appear that there are some interesting spatial-temporal biasing problems connected with the tides since the basic coarse grid for SEASAT takes three days to get established. The problem of observing the tides is old and fascinating, but according to the most recent survey (Hendershott (1977)) deep ocean tides all over the world are not known.

There has been a lot of debate over whether it is necessary to subtract out a geoid which is accurate to ten centimeters. The present planning is for SEASAT to cover the entire earth oceans twice in the first year, hence the geoid plus the steady oceanic currents could be reasonably well determined. My view is that, although it would be desirable to know the orbit of SEASAT to an accuracy of 10 cm, it does not appear to be crucial, since the orbit is extremely smooth for short wavelengths. Although the orbit may bump around at long wavelengths (over 1000 km) we have good spherical harmonic gravity models such as the GEM series to determine this orbit correction. By an appropriate least squares fitting, slow changes of the orbit, which I understand will be less than 10 cm per orbit, will contribute to only the lowest orders of the apparent gravity field, and can therefore be easily subtracted.

Lastly I would like to tell of possible long wavelength features such as long basin-filling sloshing modes with periods up to 30 hours or so. We know little about them and they should be relatively easy to pick out with the assistance of a good gravity model. They might be less than ten centimeters in amplitude, and if so the basic problem will be the usual one of separating a weak signal from the noise.

In terms of stationary long modes, Figure 7 shows an oceanographer's view of the tropics from the famous Sverdrup, Johnson, and Fleming (1942) textbook. The anticipated surface deflection is schematically sketched in and spectral compon-

ents of the geoid should be known to sufficient accuracy to aid in observing such a surface. To assist in such an endeavor, however, one could look for a banded structure in an east-west direction. This is absent in the gravitational field as far as we know.

There are many more ocean surface features which haven't been mentioned, many possessing their own particular challenges. Some examples are the sea level adjustment on shelf areas, level changes across straits such as Gibraltar, sea level changes between Pacific and Atlantic, and the response of an oceanic surface during an earthquake event.

## References


Hendershott, Myrl C., in The Sea, Vol. 6, E. D. Goldberg, I. N. McCave, J. J. O'Brien, and J. A. Steele, Eds., John Wiley and Sons, New York, 1048 p, 1977.

MODE Atlas, Valery Lee and Carl Wunsch, Eds. Atlas of the Mid-Ocean Dynamics Experiment. POLYMODE Office 54-1417, Massachusetts Institute of Technology, Cambridge, MA 02139, 274 p, 1977.

Sverdrup, H. V., Martin W. Johnson, and Richard H. Fleming, The Oceans. Prentice Hall, Inc. Englewood Cliffs, NJ, 1087 p, 1942.

# World Gravity Standards

Urho A. Uotila
Department of Geodetic Science, The Ohio State University
Columbus, Ohio 43210

In order to use gravity anomalies in geodetic computations and geophysical interpretations, the observed gravity values from which anomalies are derived should be referred to one consistent world wide system. We are very fortunate that we have one unique reference system available. The International Gravity Standardization Net 1971 was adapted by the International Union of Geodesy and Geophysics at Moscow in 1971 [Morelli, et al, 1974]. The network was a result of extensive cooperation by many organizations and individuals around the world. The final results were produced by a small subgroup of Special Study Group No. 5 of the International Association of Geodesy. This network contains more than 1800 stations around the world. The data used in the adjustment included more than 25,000 gravimeter, pendulum and absolute measurements. It was claimed that "standard errors for IGSN 71 gravity values are less than ±0.1 mgal." It certainly was a tremendous improvement over the old "Potsdam system."

At the XVIth General Assembly of the International Association of Geodesy in 1975 a resolution was passed, and Working Group No. 2: "World Gravity Standards" was established to maintain the IGSN as the international gravity reference standard and to provide advice and assistance to the International Gravity Bureau in problems related to gravity standards. The Earth Physics Branch of Energy, Mines and Resources in Canada was requested to set up a technical service for the maintenance of IGSN. This service involves the use of existing EPB facilities to maintain a data bank of observations related to IGSN, and station descriptions. New data will be transmitted to EPB through the International Gravity Bureau in Paris, who will also act as the central distribution agency for revised IGSN information.

As you might be well aware, there have been many new absolute measurements of gravity since the adoption of IGSN 71. The most extensive comparison of the new absolute measurements and the IGSN 71 values was reported by Cannizzo, Gerotti and Marson [1978]. They reported results of 25 absolute gravity measurements carried out at 17 stations in Europe using new Italian transportable apparatus in 1976 and 1977. They concluded that the accuracy of the new absolute measurements is 10 $\mu$gal. Figure 1 shows the gravity differences between gravity values of IGSN 71 and new absolute measurements [Cannizzo, et al, 1978]. Most of the gravity differences are less than 0.1 mgal, which was the claimed accuracy for IGSN 71 gravity values. It seems that there is no linear scale difference but some local systematic differences.

During 1977 Marson and Alasia [1978] measured absolute gravity at six stations in the U.S.A. The preliminary analyses of the results show a good agreement with IGSN 71 values. There have been also other absolute measurements such as Hammond, et al [1978] and Arnautov, et al [1977]. All comparisons have shown reasonable agreements between new absolute values and the IGSN 71 values.

The effect of new absolute gravity measurements in the IGSN 71 values have been studied [Uotila, 1978]. From these studies it has been concluded that it is not useful to make a global readjustment of the IGSN 71 in the near future. The IGSN 71 serves as a good standard for relating gravity measurements to the absolute system in world wide bases as far as computations of mean anomalies and production of gravity anomaly maps are concerned. The Working Group No. 2 has further recommended that the IGSN 71 values should not be fixed in new local, national or continental adjustments, but entered properly weighted using their variance-covariance matrix.

There has been some discussion to produce gravity station networks, where the accuracy of the gravity values would be about 10 $\mu$gal. If we wish to have that level accuracy of the gravity values in a network, we must have much more accurate measurements of gravity differences between the stations than currently are available or have many more measurements of gravity with a good absolute apparatus. In order to improve the accuracy in measurements of gravity differences, we must establish good calibration lines with more accurate gravity values than currently are available. We have to observe or model also gravity variations caused by local and global environmental factors, such as tides, water level, etc. There is much to be done before we can have a large net of gravity stations with 10 $\mu$gal accuracy.

## References

Arnautov, G.P., J.D. Boulanger, E.N. Kalish, Yu.E. Stus and V.G. Tarasyuk, Measurements of Absolute Gravity Value at Base Gravity Stations in (Liodovo), Moscow, Tallin and Tbilisi, Ac. Sc. USSR, Soviet Geophys. Com., 1977.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

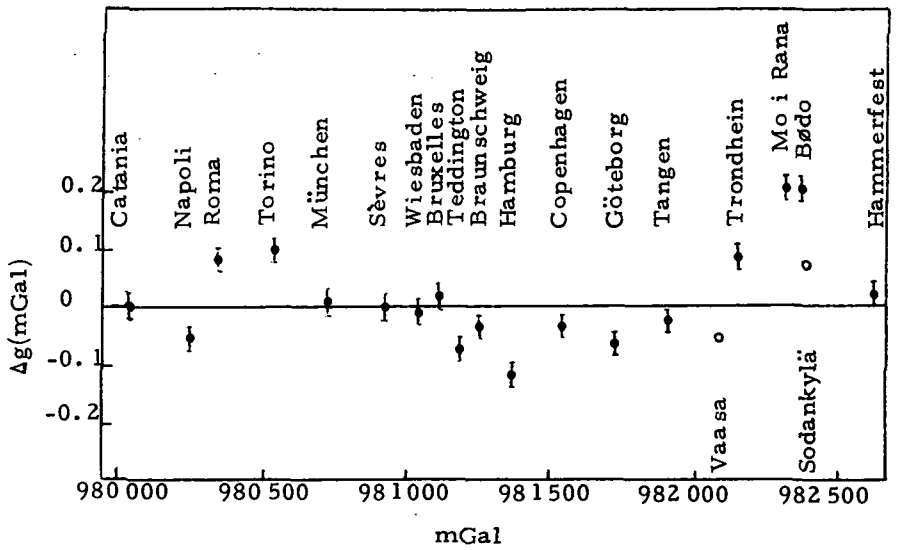



Fig. 1 - Gravity difference ($g_{IGSN\ 71} - g_{abs}$) with indication of the standard deviation: • with reference to IGSN 71, ∘ with reference to Finnish net.


Cannizzo, L., G. Cerutti and I. Marson, Absolute Gravity Measurements in Europe, Il Nuovo Cimento, Vol. 1C, N. 1, Gennaio-Febbraio, 1978.

Hammond, J.A. and R.L. Iliff, The AFGL Absolute Gravity Program, Proceedings of the 9th GEOP Research Conference, Application of Geodesy to Geodynamics, Editor, I.I. Mueller, Department of Geodetic Science Report No. 280, The Ohio State University, Columbus, Ohio, 1978.

Marson, I. and F. Alasia, Absolute Gravity Measurements in the United States of America Air Force Geophysics Laboratory, AFGL-TR-78-0126, Hanscom AFB, Mass., 1978.

Morelli, C., C. Gantar, T. Honkasalo, R.K. McConnell, J.G. Tanner, B. Szabo, U. Uotila and C.T. Whalen, The International Gravity Standardization Net 1971 (IGSN 71), International Association of Geodesy, Special Publication No. 4, Paris, 1974.

Uotila, Urho A., Studies in Gravimetric Geodesy, Department of Geodetic Science Report No. 281, The Ohio State University, Columbus, Ohio, 1978.

# The Measurement of Surface Gravity

N79 21494

J. C. Harrison
Cooperative Institute for Research in Environmental Sciences
University of Colorado/NOAA, Boulder, Colorado 80309
and

L.J.B. LaCoste
LaCoste and Romberg, Inc.
6606 North Lamar Blvd., Austin, Texas 78752

LaCoste and Romberg G and D gravity meters are normally employed when attempting high precision measurement of gravity differences on land, and we therefore discuss the capabilities and limitations of these instruments. Their design differs only trivially (mainly in the reset mechanism) from that described in the 1945 patent (LaCoste, 1945) and shown in figure 1. A negative length spring 4 with wire added to bring it to the zero length condition supports the beam 3. The beam pivots about the line joining the points of attachment of the springs 5 to the support rods 6 and theory (LaCoste, 1935) shows that for equilibrium of the beam in a horizontal position the distance, $\ell$ , of the upper support 35 of the zero length spring above this pivot line is proportional to $g$ . The meter is read by moving the support 35 vertically to bring the beam into this position. The change $\delta\ell$ in $\ell$ required to do this as the meter is read first in one place and then in another is proportional to the gravity difference $\delta g$ by the relation $\delta\ell/\ell = \delta g/g$ . The meters are built with $\ell$ = 2.5 cm so that the (worldwide) 7 gal range of the G meter requires moving the support $1.75 \times 10^{-2}$ cm , and 1 microgal accuracy means positioning the support to within 0.25 A.

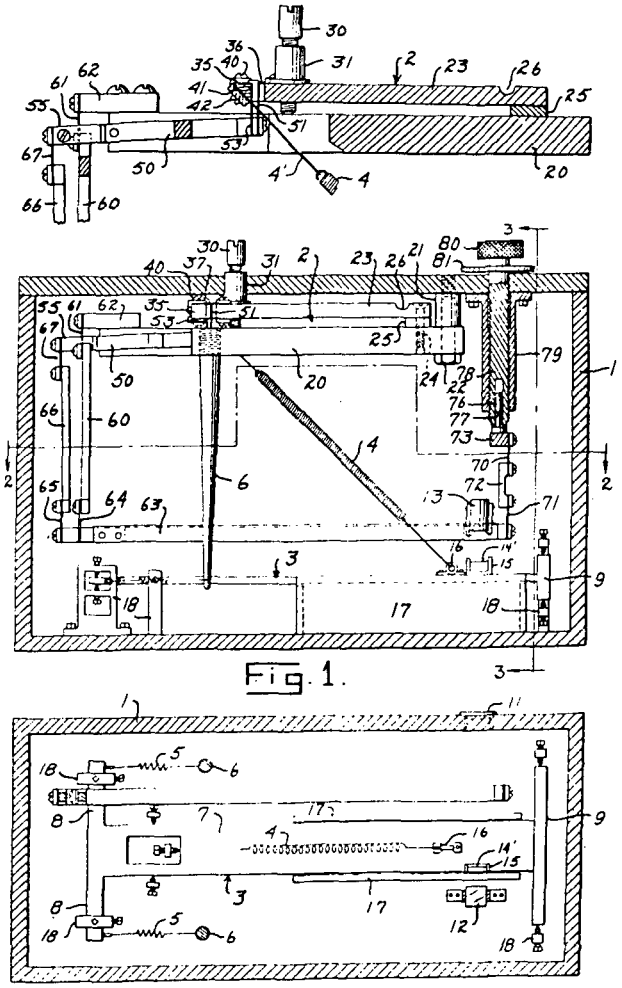


Fig. 1. Construction of LaCoste-Romberg land gravity meter.

These movements are generated with a measuring screw and double reduction lever system. In an ideal system the vertical motion of the support is proportional to the rotation of the measuring screw (which is driven through a gearbox) but in practice there are departures from this ideal due to periodic errors in the screw, eccentricity in the screw resulting in wobble, and non-linearity of the lever system. The screw problems result in errors with periodicities of once and twice per turn of the screw, or about 70 and 35 mgal with the G meter. There is some variation between meters, but 35 μgal (about $5 \times 10^{-6}$ of full scale) is a typical amplitude for this error in the G meter, and it is the most important source of error. The reduction factor of the lever system varies smoothly over its range, leading to departures of perhaps a part in 1000 of full scale from perfect linearity. These departures are determined by weighing a 200 mgal rider at various parts of the meter's range, and a calibration curve based on these weighings is supplied with the meter. The main improvement in accuracy of the D meter over the G is that the periodic screw errors have been reduced by increasing the reduction ratio of the lever system to reduce the range of the meter to 200 mgal. Thus the screw errors (still about 5 ppm of full scale) have been reduced to about 1 μgal. They can be determined in the G meter by weighing a small rider at many points on the meter's range but this is laborious. It is advisable to run the measuring screw back and forth some before starting a day's readings in order to spread out the lubricant. The reduction factor of the lever system should be stable and there is no reason to suppose that the calibration factor will change in a larger ratio than the meter drift is of total gravity.

When the state of strain in a metal is changed it does not attain its final strain immediately following the change in stress, but the last .02% or so of response takes place slowly over a matter of hours. The high magnification of astatic gravity meters means that the mass moves a long way for a small change in gravity, with a consequently large change of tension in the spring. The change in moving a gravity meter beam from reading line to stop corresponds to many hundreds of milligals equivalent spring

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics*, October 2-5, 1978. Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

tension, even though this motion may be produced by only a few mgal change in gravity. Hysteresis on the hundreds of mgals tension will be seen in the meter reading. Great care is taken during manufacture to ensure that the spring length is exactly the same when the beam is clamped, as when it is at the reading line. The meter is clamped for an hour, then read; then left at the reading line for an hour and read again, and so on. Adjustments continue until a single smooth tidal curve can be put through the two sets of data. However if the arrestment gets out of adjustment the meter will drift rapidly after unclamping. Astatic gravity meters should not be used in the deflection mode for measuring earth tides as the changes in spring length result in appreciable hysteresis.

The spring material has the property that there is an inflection in the variation of its rigidity with temperature. This allows the meter to be thermostated at the temperature of zero temperature coefficient and renders the meter rather insensitive to changes in external temperature, although large fluctuations will induce transients. An outer crudely thermostated box could be added by the user, if the meter is to be used in applications where large fluctuations in external temperature are inevitable.

A large hollow container is added to the rear of the beam to bring its center of volume above the pivot axis and thus compensate for changes in atmospheric pressure. The coefficient is reduced to below 10 μgal/cm of mercury and the meter sealed. While the seal is not perfect, changes in internal pressure are slowed down to the point where their effects will be removed as "drift".

The spring material is magnetic and the spring is carefully demagnetized when a meter is assembled. Sensitivity to horizontal magnetic field is detected by reading the meter in various orientations. Demagnetization continues until the variation in reading with azimuth is reduced below about 10 μgal. Sensitivity to vertical field is tested by bringing a large permanent magnet up to the meter - typically changes of the order of 100% of the earth's field produce effects of the order of 100 μgal. The meter is then doubly shielded and this reduces the effects below the detectable level. However, magnetic effects are always a danger; meters should never be exposed to large magnetic fields and gravity stations should not be located in places with abnormal magnetic fields. Periodic checks should be made by reading the meter in various azimuths, to ensure that the meter is still compensated magnetically.

Current accuracy of the G meter is about 30 μgal rms unless precautions are taken to avoid periodic screw errors when considerable improvement can be expected. Accuracies of 3-6 μgal are reported for a single measurement with the D meter, with 1 μgal possible by repetition or network adjustment.

The introduction of portable free-fall absolute gravity apparatus has been an important innovation. J.A. Hammond and R.L. Iliff discuss the absolute gravity program of the Air Force Geophysical Laboratory in a companion paper.

There has been little change over the last 10 years either in the names of the commonly used sea gravimeters or in their principles of operation. The changes have been mainly refinement and improvement of design, and have resulted in substantial upgrading of performance, so that the state of the art accuracy is now about 1 mgal as against 5 mgal ten years ago. The meters are mounted on gyrostabilized platforms to keep the sensitive axes vertical. The platform is slaved to stay aligned with the gyros (either a single two-axis gyro or two single-axis gyros). The orientation of the gyros can be slowly changed by the outputs of horizontal accelerometers mounted on the platform, and this feedback is arranged so that the means of these outputs are nulled. The detailed behaviour of the platform depends on exactly how the accelerometer outputs are processed to provide the feedback to the gyros (LaCoste, 1967; Talwani, 1970) and it can be made to behave like a simple pendulum with a chosen damping and period. The period must be long enough that the platform is unaffected by the horizontal wave accelerations and a 4-6 minute period, 0.707 critical damping, combination functions excellently for this purpose, although we shall later discuss applications where these periods are made much larger.

The Askania sea gravity meter is now manufactured by the Bodenseewerk. Two models are available: the Kss 5 is a refurbished Gss 2 sensor (Graf, 1958; Graf and Schulze, 1961; Schulze, 1962) in which the main improvement is more accurate location of, and increased tension in, the constraining filaments, mounted on an Anschütz platform. The meter is designed to operate with horizontal accelerations up to 50 gal and vertical up to 100 gal; accuracies of 1 mgal in "rough" sea and 2.5 mgal in "very rough" sea are claimed. A new sensor, the Gss 3 (Figure 2), consists of a tubular mass constrained by 5 filaments and 2 springs to move in a straight line. Most of the weight is supported by a spring. A feedback loop using a capacitative displacement sensor and an electromagnetic thruster keeps the mass stationary relative to its supporting structure, and the current in this thruster provides the measure of gravity. The straight line motion eliminates cross-coupling effects and an accuracy of better

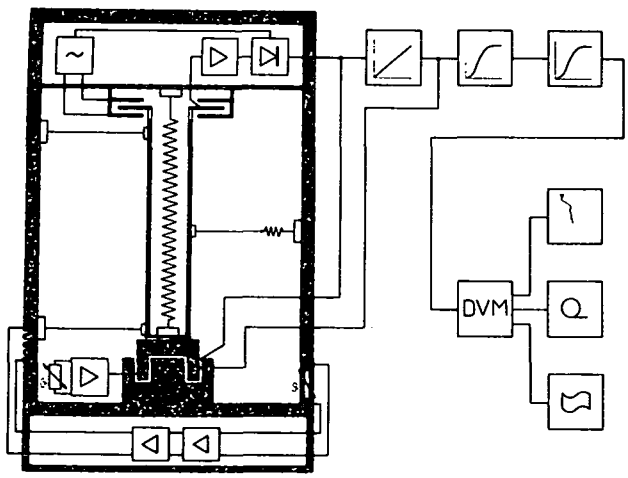



Fig. 2. The Gss-3 sea gravity meter sensor

than 1 mgal up to 200 gal wave acceleration is specified. This sensor is operated with a newly developed KT 30 stable platform.

The LaCoste and Romberg meter is still substantially as described by LaCoste (1967) and the main improvement in performance results from the use of cross-correlation techniques in correcting for cross-coupling errors (LaCoste, 1973a). The correction for inherent cross-coupling itself is, in fact, trivial as beam position and horizontal acceleration are both sensed anyway, and a correction proportional to their product is easily computed. More serious errors of cross-coupling type arise in these (and presumably other) sensors because of unwanted motions in degrees of freedom which have been supposedly supressed, but which in fact occur because materials are not infinitely rigid. This problem is especially severe with the LaCoste-Romberg instrument owing to the small clearances on the air dampers. As long as these errors are small they will be linearly proportional to certain products of accelerations and velocities in the x, y and z directions, the exact combinations which are important depending on the construction of the sensor and the source of the errors A number of such corrections are computed and applied along with the correction for inherent cross-coupling. The constants of proportionality may change with age or shipment of the meter but can be determined empirically by cross-correlation of the short period gravity variations with the acceleration and velocity products known to be important, using the reasonable assumption that gravity and wave accelerations are uncorrelated. This technique allows one both to correct data already obtained and to correct the compensation of the meter, provided the necessary information was recorded during the survey. As an example we show (Figure 3) data obtained with 3 LaCoste-Romberg meters in the North Sea by the (British) Institute of Geological Sciences (M. Tully, personal communication) in up to Force 7 sea conditions. Meters S-40 and S-75 had been in use for some time and had been correctly compensated on the basis of earlier cross-correlation analyses. Meter S-84 was a new system being used at sea for the first time, and the raw data differs from that obtained with S-40 by up to 7 mgal. Correction on the basis of cross-correlation analysis brings the two into perfect agreement. S-75 read systematically about 2 mgal higher, probably as a result of vibration. The mean of 143 cross ties with S-75 was 0.84 mgal, and with S-40 was 1.04 mgal.

However, the real solution is to build a more rigid meter and LaCoste and Romberg have been experimenting for some time with a straight line meter similar in some respects to the Gss 3 but employing an infinite period-zero length spring suspension (LaCoste, 1973b).

Bell Aerospace have brought out the BGM-3 system to replace their BGM-2. The gravity sensor, a pendulous mass with capacitative position sensing and electromagnetic feedback, has been significantly improved by a simplified design and use of improved materials and manufacturing techniques. Accuracies of 1 mgal are claimed in up to 100 gals wave acceleration. The meter can

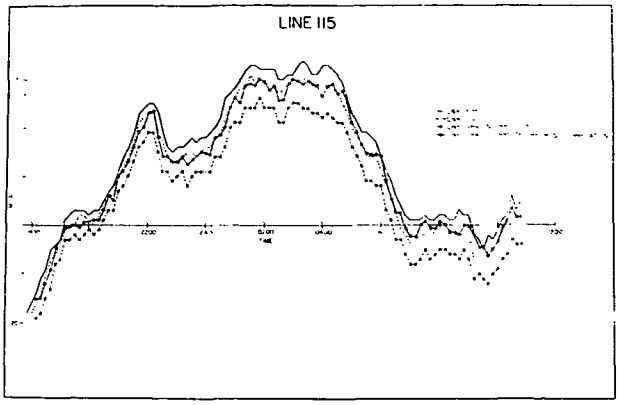



Fig. 3. North Sea profile obtained by the Institute of Geological Sciences (M. Tully, personal communication) showing effectiveness of cross-correlation correction.

accept inputs from a shipboard inertial navigator and will then produce free air anomalies in real time.

Woods Hole Oceanographic Institution (C. Bowin, personal communication) have continued to use a vibrating string accelerometer (Bowin et al., 1969, 1972) on a Sperry Mark 19 Mod 3c gyrocompass mount and use a Hewlett-Packard 2114 computer for instrument control and data recording. A more portable system has also been developed using a Aeroflex ART-57 table. RMS differences of crossings in the open ocean have varied between 1 and 3 mgal.

With sea gravity meter accuracy at the 1 mgal level the Eötvös correction becomes a major source of error in the gravity anomalies. Satellite navigation systems have improved navigation at sea very considerably and one can probably count on getting good fixes every two hours or so. However, if sea or current conditions are variable, or the ship does any manoevering between fixes, the velocity uncertainties may introduce considerable noise into the anomaly profiles. Determination of velocity to 1/10 knot over a 6 minute interval requires fixes accurate to about .01 mile or 20 m. Electronic navigational aids may provide this accuracy in the survey area but more probably they will not. In this latter case it is worth considering the use of inertial navigation in conjunction with the other methods. The ship may carry an inertial navigator. There are many implementations of inertial navigation and the performance depends on this implementation and the quality of components used. All ideal (meaning that no errors in indicated position are produced by the navigator's history of acceleration over the earth's surface) systems however, behave as undamped pendulums with 84 minute period. Initial errors, component imperfections or deflections of the vertical excite free oscillations of the navigator. An example of such oscillations is shown in Figure 4, which is a plot of longitude indicated by a Honeywell SPN system using electrostatic gyros minus LORAN longitude on a U.S. Navy Oceanographic Office aircraft gravity test in 1976 (J. Ford, personal communication). The 84 minute Schuler oscilla-

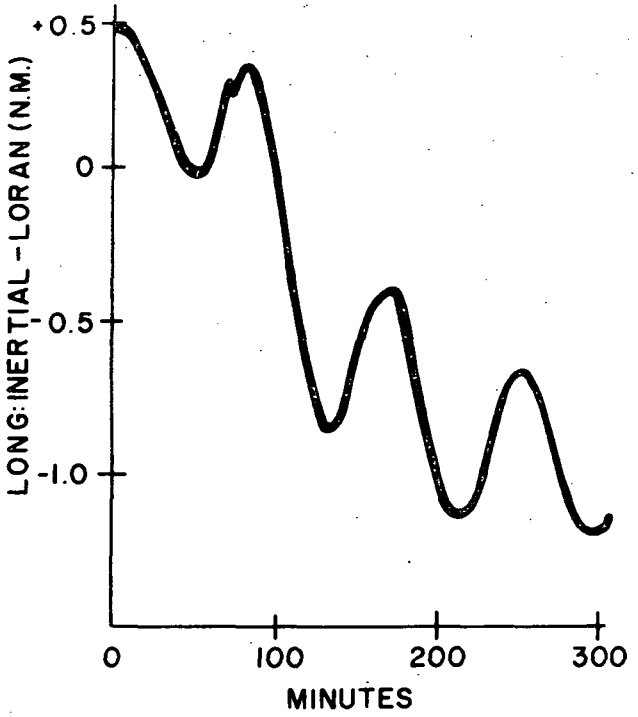



Fig. 4. Difference between inertially determined and LORAN longitudes during Naval Oceanographic Office airborne tests (J. Ford, personal communication) showing Schuler oscillations of the inertial navigator.

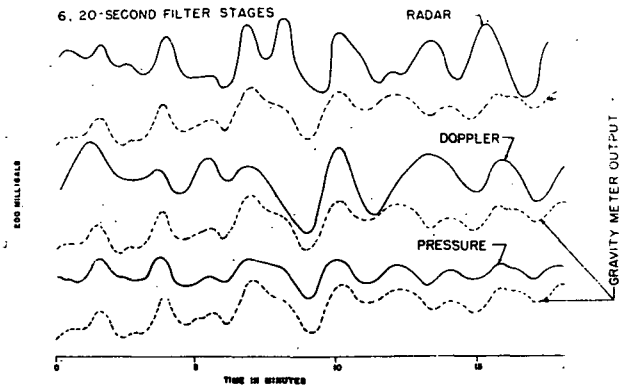



Fig. 5. Vertical acceleration of plane from radar, Doppler radar and barometric altimeters compard with observed gravity.

tions are apparent and velocity errors of up to about 2 knots occur. Such oscillations can of course be removed by real time Kalman or post facto filtering using the electronic positioning information.

The gyrostabilized platform used in sea gravity meters is almost an inertial navigator in itself and LaCoste has done some interesting experiments to ascertain whether it could be used in this capacity. He added a third gyro to provide azimuth stabilization and has shown how the accelerometer and gyro outputs may be combined to pro-

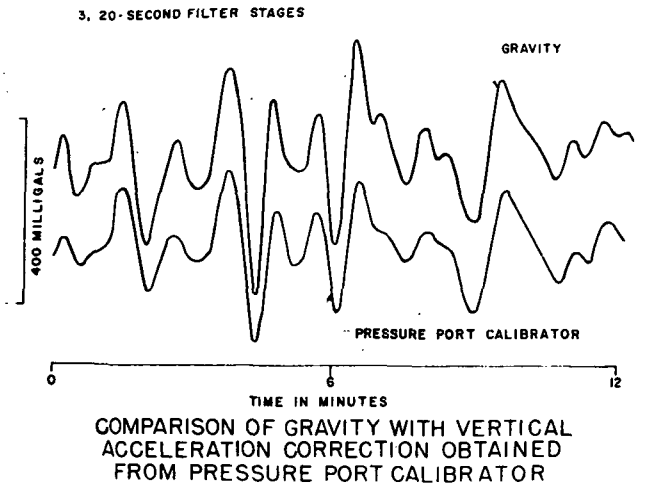



Fig. 6. Gravity compared with acceleration from barometric altimeter.

vide a Schuler-type equation for the Eötvös correction even if the platform itself is not Schuler tuned. (Valliant & LaCoste, 1976) The danger with using a Schuler tuned platform is that any malfunction may initiate oscillations which, being undamped, can spoil results for an indefinite time. However, LaCoste-Romberg platforms have been used satisfactorily with periods up to and including 84 minutes; the 84 minute period is not necessary for the theory to be valid but the gyros behave better in long period servo loops as less external precession is required. Gravity profiles with Eötvös effects corrected with the platform navigator are significantly smoother than when Lorac fixes were used (LaCoste, 1977). Adding the third gyro is very much cheaper than an entire inertial navigator, and while the performance of the platform is not comparable with that of a good commercial inertial navigator over long periods of time, it is able improve the determination of Eötvös corrections in ships and aircraft when used to interpolate within a framework of satellite or electronic fixes.

Probably the most interesting developments in recent years have come in gravity measurements from aircraft. An aircraft normally gives a much smoother ride than a ship, so the gravity meter errors are correspondingly less. The problems are in correcting for the Eötvös effect and vertical accelerations of the aircraft over short time intervals that the gravity information is not smeared out by the high speed of the vehicle to the point of losing important detail. A low-speed aircraft is therefore advantageous and the most impressive results to date have been obtained by Bill Gumert of the Geoscience Division of Carson Helicopters using a Sikorsky S-61 helicopter at speeds of 70 knots. He used a Del Norte trisponder positioning system which (with modification) gives positions within a 3 m circle of error out to 160 km range. This is adequate for 1 mgal accuracy in Eötvös correction over 1 minute intervals. The comments made above concerning the value of inertial navigators for interpolating between electronic fixes are equally

applicable to aircraft work. Gumert finds that the LaCoste-Romberg inertial system gives results agreeing to 2 mgal with those derived from the trisponder positions.

A simple harmonic height variation of ±5 cm with a period of 1 minute produces accelerations of 50 mgal amplitude, 35 mgal rms value. These accelerations decrease for longer period motions as the inverse square of the period; for example to 2 mgal maximum, 1.6 mgal rms at 5 minutes. The correction for these height variations is thus critical and the quality of the altimetry controls the detail which can be obtained. Many types of altimeters are available. Figure 5 compares vertical accelerations derived from data taken with a radar, a Doppler radar and a barometric altimeter, with the gravity meter record during 15 minutes of the previously mentioned Oceanographic Office airplane tests in 1976 (data filtered with 6 stages of 20 sec R-C smoothing). The correspondence between observed gravity and the pressure altimetry is impressive but the scale bar is 200 mgal and there are differences of the order of 5 mgal between the curves. Figure 6 shows less filtered data from the same test shorter periods are present and the acceleration amplitudes are bigger. Bill Gumert (personal communication) in his helicopter work has used a precise pressure sensor (Rosemont 1201 F), a laser altimeter (Geodolite 3A) and a radar altimeter (Honeywell AN/APN-194). The pressure altimeter has a repeatability of ± 7.5 cm over a 50 m altitude range but is not an absolute instrument. The laser altimeter range but has some sensitivity to the color of the target, while the radar altimeter gives absolute altitudes to about 1m. In use, the pressure and radar altimeters are intercompared frequently over flat areas of known elevation to provide an absolute calibration for the former. The rms difference at 244 line crossings during a survey in New York State was 2.3 mgal and eleven of these were misties of over 5 mgal at the beginning of lines where the instruments probably had not had time to stabilize. Comparison with ground data is about as good as the internal consistency of the airborne data (2-3 mgal). Current accuracies are about 2 mgal after smoothing to remove features of wavelength shorter than 3-4 miles. This is a most important achievement holding much promise for mdium scale structural investigations which, by keeping the topography at arms length, eases the computation of terrain corrections very considerably and, in poorly surveyed areas, eliminates the necessity of leveling to determine station heights. Comparable accuracies can probably be obtained with fixed wing aircraft with the horizontal distance scale expanded in the ratio of the speeds - for example 10 miles wavelength with a 350 knot aircraft - although there may be more high frequency altitude variation in the faster moving airplane which could increase the required averaging time slightly.

References


Bowin, C., C.G. Wing and T.C. Aldrich, Tests of the M.I.T. vibrating string gravimeter, 1967. Jour. Geophys. Res., 74, 3278-3280, 1969.

Bowin, C., T.C. Aldrich and R.A. Folinsbee, VSA gravity meter system: tests and recent developments, Jour. Geophys. Res., 77, 2018-2033, 1973.

Graf, A., Das Seegravimeter, Z. für Instrumkde, 60, 151-162, 1958.

Graf, A. and R. Schulze, Improvements on the sea gravimeter Gss 2, Jour. Geophys. Res., 66, 1813-1821, 1961.

LaCoste, L.J.B., A Simplification in the Conditions for the Zero-length-Spring Seismograph, Bull. Seis. Soc. Amer., 25, 176-179, 1935.

LaCoste, L.J.B. and A. Romberg, U.S. Patent 2,377,889, 1945.

LaCoste, L.J.B., Measurement of gravity at sea and in the air, Rev. Geophys. 5, 477-526, 1967.

LaCoste, L.J.B., Cross-correlation method for evaluating and correcting shipboard gravity data, Geophysics, 38, 701-738, 1973a.

LaCoste, L.J.B., U.S. Patent 3,717,036, 1973b.

LaCoste, L.J.B., 1974 test of LaCoste and Romberg inertial navigation system, Geophysics, 42, 594-601, 1977.

Schulze, R., Automation of the sea gravimeter Gss 2, Jour. Geophys. Res., 67, 3397-3401, 1962.

Talwani, M., Chap. 8 Gravity (espec. ϕ 3) in The Sea, Vol. IVa, Interscience- John Wiley, New York, 1970.

Valliant, H.D. and L.J.B. LaCoste, Theory and evaluation of the LaCoste and Romberg three-axis inertial platform for marine gravimetry, Geophysics, 41, 459-467, 1976.

Wing, C.G., M.I.T. vibrating string surface-ship gravimeter, Jour. Geophys. Res., 74, 5882-5894, 1969.

Page Intentionally Left Blank

# The AFGL Absolute Gravity Program

James A. Hammond and Robert L. Iliff
Air Force Geophysics Laboratory, Terrestrial Sciences Division
Hanscom AFB, Massachusetts 01731

Abstract. A brief discussion of the AFGL's program in absolute gravity is presented. Support of outside work and in-house studies relating to gravity instrumentation are discussed. A description of the current transportable system is included and the latest results are presented. These results show good agreement with measurements at the AFGL site by an Italian system and with previous measurements by Hammond and Faller. The accuracy obtained by the transportable apparatus is better than 0.1 $\mu m/sec^2$ (10 $\mu gal$) and agreement with previous measurements is within the combined uncertainties of the measurements. The instrument will be used extensively for field measurements in 1979.

The Air Force Geophysics Laboratory's program in absolute gravity can be divided into three main areas: support of outside research into measurement techniques and of comparative measurements by other absolute instruments; the study of the physics of the measurement techniques and the development of new instrumentation; and measurements in the laboratory and at selected field sites with the AFGL transportable system. The predominant focal point of this paper is the work on current measurements, but in the interest of completeness we will also briefly discuss the first two areas.

The outside work supported by AFGL includes that of Dr. James Faller and Mr. Robert Rinker of the Joint Institute for Laboratory Astrophysics in Boulder, Colorado (JILA/NBS). AFGL has supported the development of a novel system for the isolation of a reference reflector in an interferometer type of absolute gravity instrument. This system, which uses an electro-mechanical feedback system to synthesize a very long period vertical mass-spring support, is being designed and built into a package which should be capable of directly supporting the reference reflector on a gravity instrument.

AFGL supported the visit to the U.S.A. of the transportable system developed by the Istituto di Metrologia "G. Colonetti" (IMGC) of Torino, Italy with the cooperation of the Bureau International de Poids et Mesures (BIPM). This work was supported by a grant to the IMGC administered through the European Office of Aerospace Research and Development as well as by in-house support from AFGL and the Defense Mapping Agency Geodetic Survey Squadron (DMA/GSS). The work involved transporting the equipment and two people to six sites (Hanscom AFB, MA; Denver, CO; Holloman AFB, NM; San Francisco, CA; Bismarck, ND; Miami, FL). The system had a mass of about 1500 kg when packaged for air transport and the entire operation required six weeks to complete (with a final remeasurement at Hanscom AFB, seven measurements were made). The uncertainty obtained was about .1 $\mu m/sec^2$ (10 $\mu gal$) at most sites. [Marson and Alasia, 1978]

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

In the area of studying the physics of measurement techniques we are planning to do accurate measurements of the effect of air resistance on the free fall type of measurement. The current data analysis method allows very small effects to be seen in the deviation of the fall of the reflector from a purely uniform acceleration.

Another concern is that the laser wavelength standard is reproducible and stable. Periodic measurement of the laser used in our system will be done to assure our wavelength standard is not perturbed by time or the effects of transporting the equipment.

AFGL is looking at new developments in electronics and other areas to solve some of the current problems with this kind of instrumentation. In particular, several techniques for making the system simple to operate are employed in the current AFGL system. The system is completely automated and data are analyzed and corrected for gravity tides in real time. Optical and mechanical alignment are simplified over previous systems and self-checks on timing accuracy can be performed independent of a gravity measurement. A new timing and data analysis system is being obtained that should have increased reliability over the current system.

Measurements are currently being made with a system that incorporates the mechanical parts from the first generation instrument [Hammond, 1970], and uses a control system and support base (with optics) built at AFGL. The timing and data analysis equipment were integrated by JILA. Figure 1 shows schematically the absolute gravity system. The laser length standard is a Lamb-dip stabilized He-Ne laser which is periodically compared with an Iodine stabilized laser in our laboratory. The oscillator is a Rubidium frequency standard and the timing electronics allow very precise ($\pm$ 125 psec) measurement of a large number of time values during the fall of the reflector. The reference system in the current arrangement is simply a retroreflector mounted to the base with no seismic isolation.

Figure 2 is a photo of the system as it looked at the time of this symposium. The vacuum chamber has been reduced in height by 45 cm so that the free fall path is now about 60 cm. A smaller vacuum pump is used (30 l/sec pumping speed) and the pump magnetic field is reduced considerably from the earlier system. An "old fashioned", simple free fall technique is used because several apparently inherent problems obtained with a "chamber-in-a-chamber" system, resulting in our setting that chamber aside, at least temporarily. This system has a total mass of about 700 kg when packed for air transport and it is contained in nine or ten boxes which can be handled by one or two people.

The first field measurements were made in June of 1978, approximately six months after the decision to convert the old vacuum chamber for use with the new system. At the time of that field trip we were using a computation technique that used 150

time measurements from three different positions in the free fall path. This technique produced good statistics and fairly good repeatability at the AFGL site, but a systematic effect was known to exist prior to the June 1978 field trip.

In spite of this systematic effect, (which has since been eliminated) several important things were demonstrated by these field measurements at Denver, CO, Holloman AFB, NM, and San Francisco, CA:

1. Portability of the system, short operating time (three sites in ten days)
2. 10 μgal accuracy is possible at most sites
3. Sites must be chosen carefully.

The third result expresses a difficulty that proved to be very serious in San Francisco. The site chosen was the actual IGSN-71 site in a museum in Golden Gate Park. Preliminary measurements with a short period seismometer did not indicate such a serious noise problem, but it was so bad that we were only able to get a standard error of .15 μm/sec$^2$ (15 μgal) with about 2400 drops. The IMGC system didn't experience as much trouble because they use a seismometer for an inertial support of the reference retroreflector.

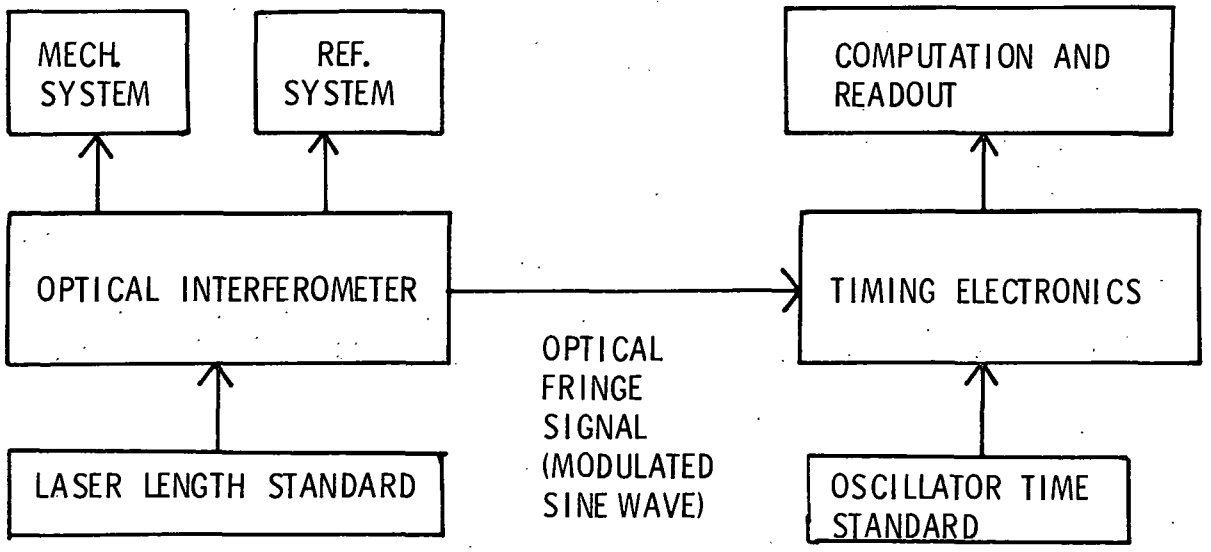
In August of 1978 a least-squares program was adopted for doing the data analysis. This is a program in which the positions ($X_i$) and the times ($t_i$) are fit to a constant acceleration formula:

$$X_i = 1/2\ g\ t_i^2 + v_o\ t_i + X_o$$

The results are: g, $v_o$, $S_o$ (the last two are of no interest, except to fix the positions of the measurement), and a table of residuals for each drop. The residuals for each position in the path can be averaged and then plotted as in Figure 3. These residuals represent the deviation of the relative path difference between the reference reflector and the freely falling reflector from what it would be if the reference reflector were not accelerating at all, and the free falling reflector were accelerating uniformly at g. Thus, a vibration of the reference reflector at a constant frequency appears as a vibration whose frequency in space decreases as the falling reflector moves to the bottom. In Figure 3, then, the ordinate gives the magnitude of these average residuals in Angstroms and the abscissa gives the position of the object in the vacuum chamber measured from the start of the measurement (approximately 8 cm from the zero velocity position).

Even with the vibration shown in Figure 3 the g value obtained showed a much reduced bias when compared with the IMGC measurement and with the older measurement [Hammond and Faller, 1970]. The repeatability and the standard error of the g value were as good as before. A rough calculation showed that even this small vibration, initiated by the release of the falling body, could produce a systematic effect of the order of .60 μm/sec$^2$. In fact, the g value then obtained agreed with the IMGC to within the uncertainties of the measurements.

We decided to get rid of the vibration even though it didn't cause a large bias with the least squares analysis technique. To do this the chamber was isolated from the reference reflector and the rest of the optics by placing it on a separate vibration isolation system. The results of averaging 150 drops worth of data are shown in Figure 4. The solid line is the result of subjecting synthetic data, with a 3 μm/sec$^2$/m gradient included, to the same least squares analysis. Thus most of the systematic appearance, if not all, is caused by the vertical gradient. If one attempts to fit the




ABSOLUTE GRAVITY INSTRUMENT SCHEMATIC

Fig. 1. Schematic diagram of a system for measuring absolute gravity.

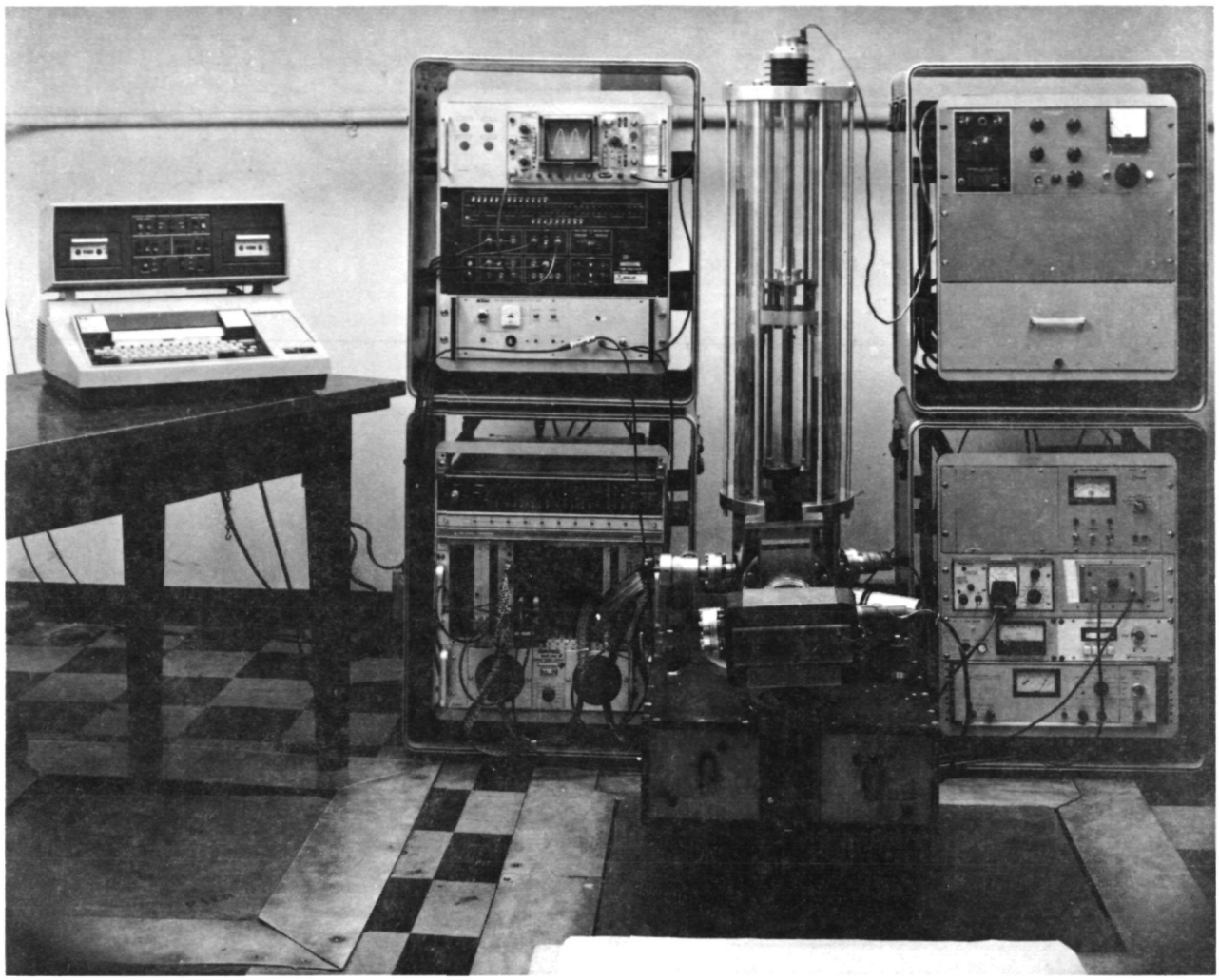


Fig. 2. The AFGL Transportable Absolute Gravity Instrument.

gradient as an additional parameter, the least squares fit becomes poorly determined and the g values have a higher scatter.

The most recent value obtained at AFGL on pier 1, Haskell Observatory is:

| | | |
|---|---|---|
| Measured Value | 9803783.21 | ± 0.03 μm/sec$^2$ |
| Gradient Correction | + 3.77 | ± 0.03 μm/sec$^2$ |
| Velocity of Light Correction | - .25 | ± 0.01 μm/sec$^2$ |
| Wavelength Uncertainty | | ± 0.03 μm/sec$^2$ |
| Estimated Uncertainty for Atmospheric Pressure and Other Possible Systematic Effects | | ± 0.05 μm/sec$^2$ |
| Value at Floor Level | 9803786.73 | ± 0.07 μm/sec$^2$ |

This can be compared with the IMGC 1978 measurements and the 1970 Hammond-Faller value at this site.

| | | |
|---|---|---|
| Marson and Alasia (1978) | 9803786.59 | ± 0.10 μm/sec$^2$ |
| Hammond and Faller (1971) | 9803786.71 | ± 0.42 μm/sec$^2$ |

There is a discrepancy at close to a significant level between the AFGL and IMGC values. Effects such as water table level and air pressure could contribute to the difference. Pressure changes could result in as much as 0.10 μm/sec$^2$ full swing. Those who use these data will have to do the best they can (weighted means) to incorporate all possible measurements at any site. Future operators of these types of instruments must make it a rule to make some notation about the barometric pressure. Discrepancies in gravity tide effects could also contribute to these differences. It is possible to perform corrections for tides which fully include the effects of ocean loading.

In summary, the AFGL program in absolute gravity, which contributed so significantly to the IGSN-71 with absolute gravity measurements at eight sites, will continue to produce significant results and advances in the future. There are several important ways for absolute measurements of gravity to help in solving problems in geodynamics. The necessity for accurate calibration lines for relative instruments, as well as for the absolute values for networks and at sites of special significance, should be obvious enough that looking at instrument acquisition as a long term goal would appear more important than making patchwork measurements with prototype instruments. The agencies

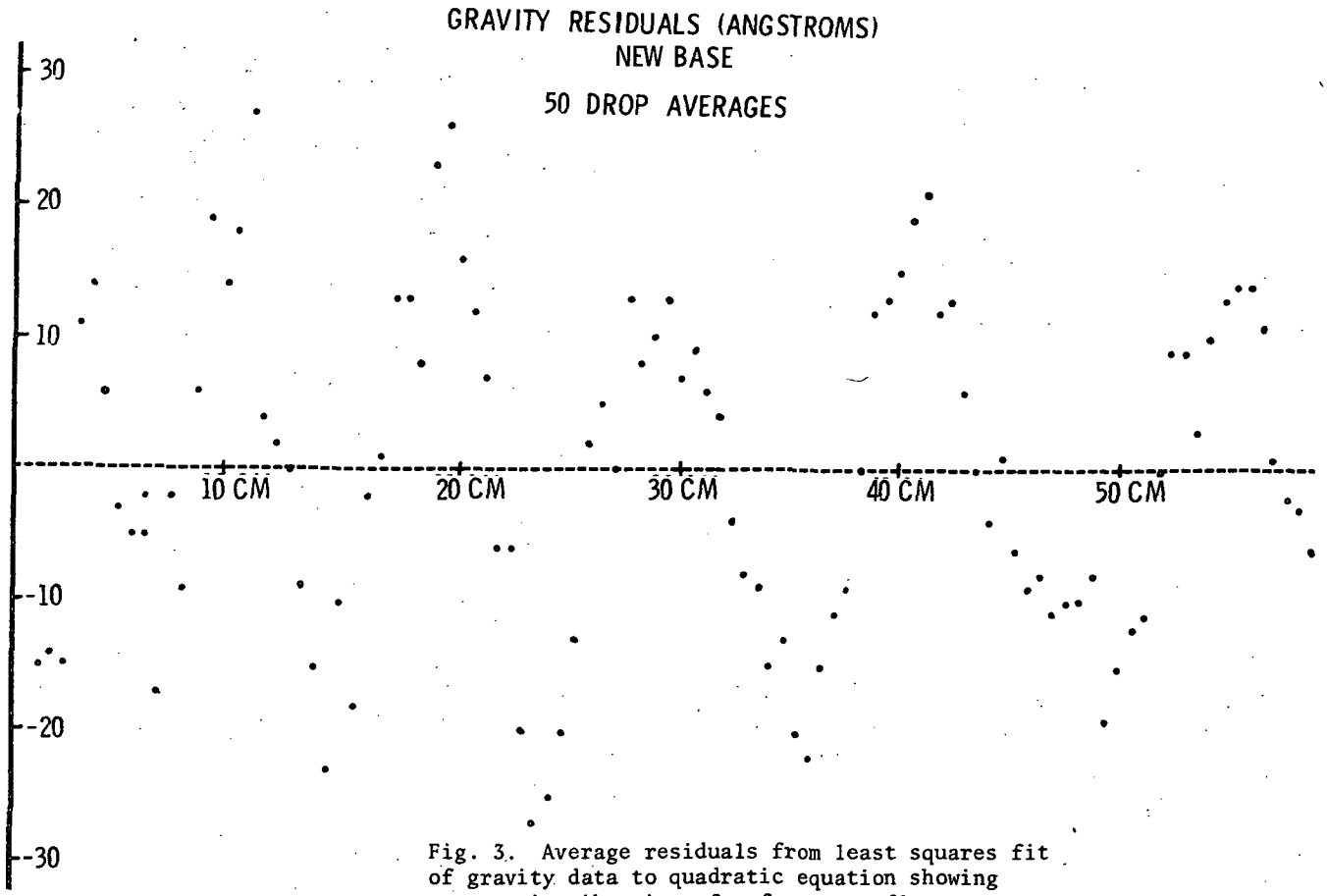



Fig. 3. Average residuals from least squares fit of gravity data to quadratic equation showing systematic vibration of reference reflector.

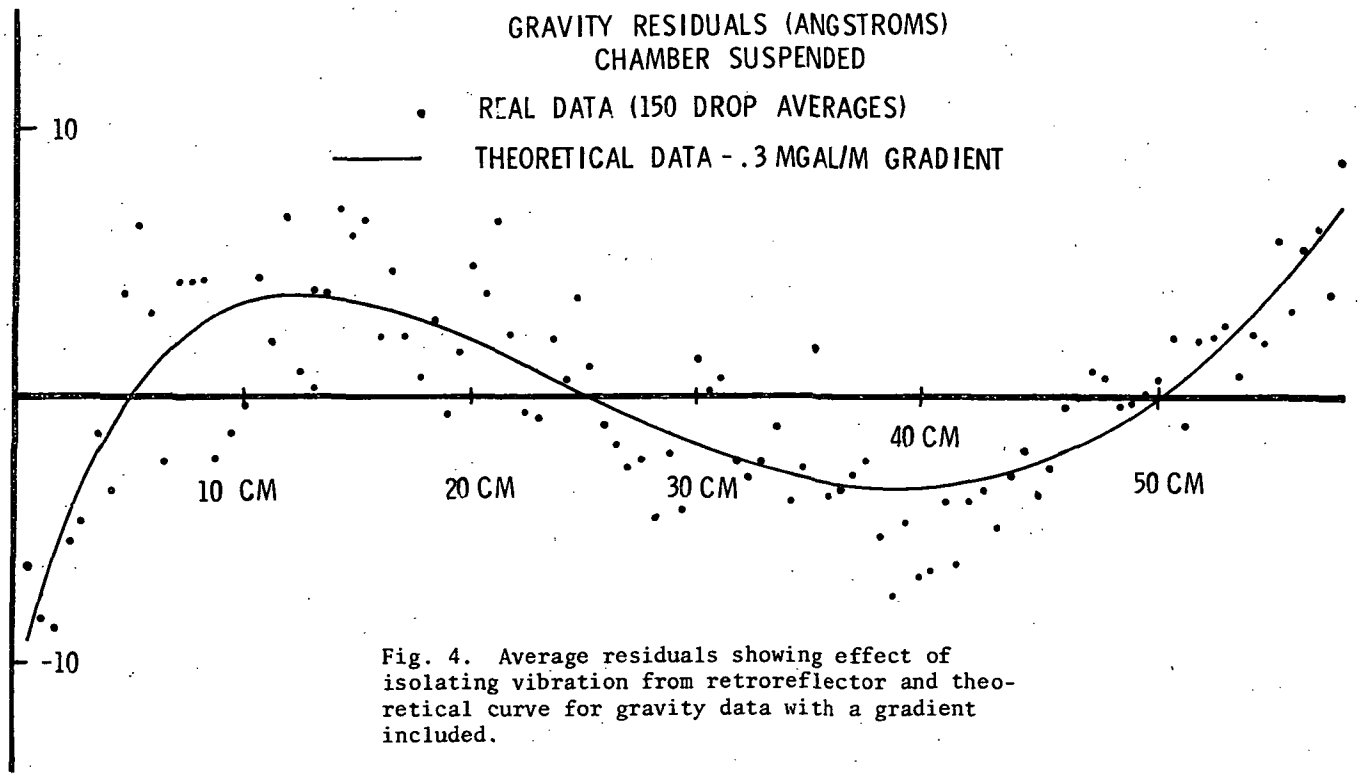



Fig. 4. Average residuals showing effect of isolating vibration from retroreflector and theoretical curve for gravity data with a gradient included.

interested should be looking at more than just the numbers describing the results. They should look at the systems with an eye to seeing them become operational instruments rather than specialized systems which only a Ph.D. physicist can operate.

## References


Hammond, J., A laser-interferometer system for the absolute determination of the acceleration of gravity, Report, AFCRL-70-0163, Air Force Cambridge Research Laboratories, Bedford, Mass., March 3, 1970.

Hammond, J., and J. Faller, Results of absolute gravity determinations at a number of different sites, J. Geophys. Res., 76, 7850, 1971.

Marson, I., and F. Alasia, Absolute gravity measurements in the United States of America, Report, AFGL-TR-78-0126, May 1978.

Page Intentionally Left Blank

# Present Status of Marine Gravity

A. B. Watts
Lamont-Doherty Geological Observatory of Columbia University
Palisades, New York 10964

## Introduction

The technique of measuring gravity at sea has been greatly improved by the development of spring-type surface-ship gravimeters which can be operated in a wide variety of sea conditions (LaCoste, 1967; Graf and Schulze, 1961). Since about 1961 surface-ship gravity measurements have been obtained on a routine basis in each of the world's major ocean basins. By combining surface-ship gravity measurements with measurements obtained earlier on board submarines (Vening Meinesz, 1948; Worzel, 1965) it is now possible to construct gravity anomaly maps of large regions of the world's oceans. These maps form an important basis for geodetic, geological and geophysical studies.

Since the development of the concept of plate tectonics there have been two approaches to the interpretation of marine gravity data that have proved particularly useful. The first, based on the pioneering studies of Vening Meinesz (1941) and Gunn (1947), uses relatively short-wavelength (wavelength $\lambda \lesssim 400$ km) gravity anomalies in oceanic regions to provide information on the long-term ($>10^6$ years) mechanical properties of the oceanic lithosphere (for example, Walcott, 1970; Watts and Cochran, 1974; Watts et al., 1975). The second, uses relatively long-wavelength ($\lambda \gtrsim 400$ km) gravity anomalies in oceanic regions to provide information on the forces which operate on the plates and which may ultimately drive them (for example, Anderson et al., 1973; Sclater et al., 1975; Watts, 1976; McKenzie, 1977).

The purpose of this paper is to present a brief review of some of the most recent developments in marine gravity. The extent of marine gravity data coverage is illustrated in a compilation map of the main free-air gravity anomaly maps of the world's oceans which have been published since 1974. A brief discussion of some of the main results in the interpretation of marine gravity is given and some comments made on recent determinations of the gravity field in oceanic regions using satellite radar altimeters.

## Gravity Measurements

During the past few years there have been increased efforts to obtain gravity measurements in oceanic regions, particularly aboard U.S., U.S.S.R., and Japanese research vessels (Table 1). Gravity measurements have now been obtained along more than 2 million nautical miles of ship's tracks. Although the accuracy of gravity measurements obtained on individual ship's cruises depends on the types of navigation and instrumentation used, the standard error of these measurements (Table 1) based on studies of discrepancies at intersecting ship's tracks is estimated to be in the range of 5 to 10 mgal.

Figure 1 summarizes the regions of the world's oceans where gravity anomaly maps have been constructed. This figure only includes those maps with an areal extent of $4 \times 10^4$ km$^2$ or greater. These maps are contoured either at 10 mgal or 25 mgal intervals and include a compilation of all available surface-ship, submarine and land gravity measurements. A significant proportion of the data used in these maps is now available from the NGDC*.

## Interpretation

Studies have now been carried out which have used marine gravity data to determine information on the deformation (or flexure) of the oceanic lithosphere caused by surface loads such as sediments (Gunn, 1944; Walcott, 1972; Cochran, 1973) and seamounts (Gunn, 1943; Walcott, 1970; Watts and Cochran, 1974; Watts et al., 1975). An important parameter in these studies is the effective flexural rigidity which is determined mainly by the effective elastic thickness of the oceanic lithosphere. By comparing observed gravity anomalies with calculated anomalies based on simple elastic or viscoelastic models it has been possible to estimate the effective elastic thickness and how it may vary with crustal age. The main results of these studies, summarized in Watts (1978), is that surface loads formed at or near mid-ocean ridge crests are associated with relatively small values of the effective elastic thickness while surface loads formed on relatively old lithospheric plates are associated with relatively large values.

Figure 2 is a plot of "isostatic response function" for the East Pacific rise crest and Hawaiian-Emperor seamount chain in the Pacific ocean. This figure shows that the range of wavelengths which provide information on isostasy at the East Pacific rise is 30 to 300 km while that for the Hawaiian-Emperor seamount chain is 200 to 800 km. The importance of these functions, however, (for example, Lewis and Dorman, 1970; McNutt and Parker, 1978) is that they can be easily compared to different models of isostasy. Figure 2 shows that the observed "isostatic response function" for the East Pacific rise crest and Hawaiian-Emperor seamount can be explained by a simple flexure model with values of the effective elastic thickness of the oceanic lithosphere in the range 2 to 6 km and 20 to 30 km respectively.

Recent studies have shown (Detrick and Watts, in preparation; Watts, Bodine, and Ribe, in preparation) that the "isostatic response functions" in Figure 2 can explain gravity data over a wide variety of other geological features. In particular, the ridge

*NGDC. National Geophysical Data Center, Boulder, Colorado.

Proc. of the 9th GEOP Conference. *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280. The Ohio State Univ., Columbus, Ohio 43210.

TABLE 1

Principal Marine gravity Operations Over the World's Oceans 1973 - 1978

| Institution | Country | Principal Ships | Gravimeter | Stable Platform | Coverage |
|---|---|---|---|---|---|
| Bedford Institute of Oceanography | Canada | Hudson | Graf-Askania Gss-2 | Anschütz | Pacific, Atlantic |
| | | Baffin | Graf-Askania Gss-2 | Anschütz | Atlantic |
| Centre National Pour l'exploration des Oceans (CNEXO) | France | Jean Charcot | Graf-Askania Gss-2 | Anschütz | Atlantic |
| German Hydrographic Institute | Germany | Komet | Graf-Askania Gss-2 | Anschütz | Atlantic |
| | | Meteor | Graf-Askania Gss-2 (also Gss-3) | Anschütz | Atlantic |
| Institute of Oceanology, Moscow | USSR | Akademik Kurchatov | Graf-Askania Gss-2 | Russian built gyrostabilized | Atlantic |
| | | Vityaz | Graf-Askania Gss-2 | | Pacific |
| Lamont-Doherty Geological Observatory | US | Vema | Graf-Askania Gss-2 | Alidade | Pacific, Indian, Atlantic |
| | | Robert D. Conrad | Graf-Askania Gss-2 | Anschütz, Aeroflex | Pacific, Indian, Atlantic |
| National Oceanic and Atmospheric Administration | US | Surveyor | LaCoste-Romberg S-51 | Gyrostabilized | Pacific |
| | | Researcher | Lacoste-Romberg S-52 | Gyrostabilized | Atlantic |
| Ocean Research Institute, University of Tokyo | Japan | Hakuhō-Maru | TSSG (VSA) | Gyrostabilized | Pacific |
| | | Hakuhō-Maru* | LaCoste-Romberg S-32 | Gyrostabilized | Pacific |
| Woods Hole Oceanographic Institution | US | Chain | VSA | Sperry MK19 | Pacific, Indian |
| | | Atlantis | VSA | Sperry MK19 | Atlantic |

*Joint United States-Japan cooperative program.

crest function adequately explains gravity data over the Walvis and Ninetyeast aseismic ridges while the Hawaiian-Emperor seamount chain function adequately explains data over some Mid-Pacific seamounts and the Louisville ridge. These results are in general agreement with the observation that a number of geological features on the ocean floor originated either at or near a mid-ocean ridge crest (ridge crest and fracture zone topography, Walvis and Ninetyeast ridges) or as a relatively young load on an old lithospheric plate (Hawaiian-Emperor seamount chain, Mid-Pacific seamounts, Louisville ridge).

Although these studies have used marine gravity data to provide information on the mechanical behavior of the oceanic lithosphere, they provide little information on the forces which may be operative on the plates. The main problem is that the mechanical and thermal properties of the oceanic lithosphere serve to obscure the gravity effect of deeper processes in the Earth such as mantle convection.

A useful approach to this problem has been to examine the relationship between long-wavelength gravity anomalies and deviations in expected depth of the sea-floor (or residual depth anomalies) for broad regions of the world's oceans (Anderson et al., 1973; Sclater et al., 1975; Watts, 1976). These studies show that a good correlation between gravity and residual depth anomalies exists, at least for the North Atlantic and Central Pacific Oceans. A correlation between long-wavelength gravity and residual depth anomalies makes a good argument for convection. Recently, however, Cochran and Talwani, (1978) concluded from a global data set that there was, in general, a poor visual correlation between long-wavelength gravity and depth anomalies in the world's oceans. In addition, Detrick and Crough (1977) have proposed the residual depth anomaly in the Central Pacific ocean formed by lithospheric thinning over an underlying "hot spot". Future studies should therefore attempt to establish a relationship between gravity and residual depth anomalies as a function of wavelength since this information appears to be the most likely to constrain models of mantle convection (McKenzie, 1977).

## GEOS-3 Satellite Altimeter Data

With the advent of satellite altimetry it is now possible to determine the shape of the marine geoid with a great deal of accuracy (Leitao et al., 1975; Leitao and McGoogan, 1975). In the absence of noise, gravity anomalies derived from GEOS-3 altimeter data, for example, would be equivalent to gravity anomalies measured on surface-ships. In the presence of noise, however, surface-ship gravity measurements provide the best means to determine the short-wavelength gravity field in the oceans while GEOS-3 altimeter data provide the best means to determine the long wavelengths.

In a recent study Rapp (in press) has recovered 1 x 1° average gravity anomalies from GEOS-3 altimeter data and compared them with averages determined from surface ship and land measurements. The RMS difference between predicted and terrestrial 1 x 1° gravity anomaly averages was ±16 mgal for the Philippine sea region and ±8 mgal for the East Coast, U.S.

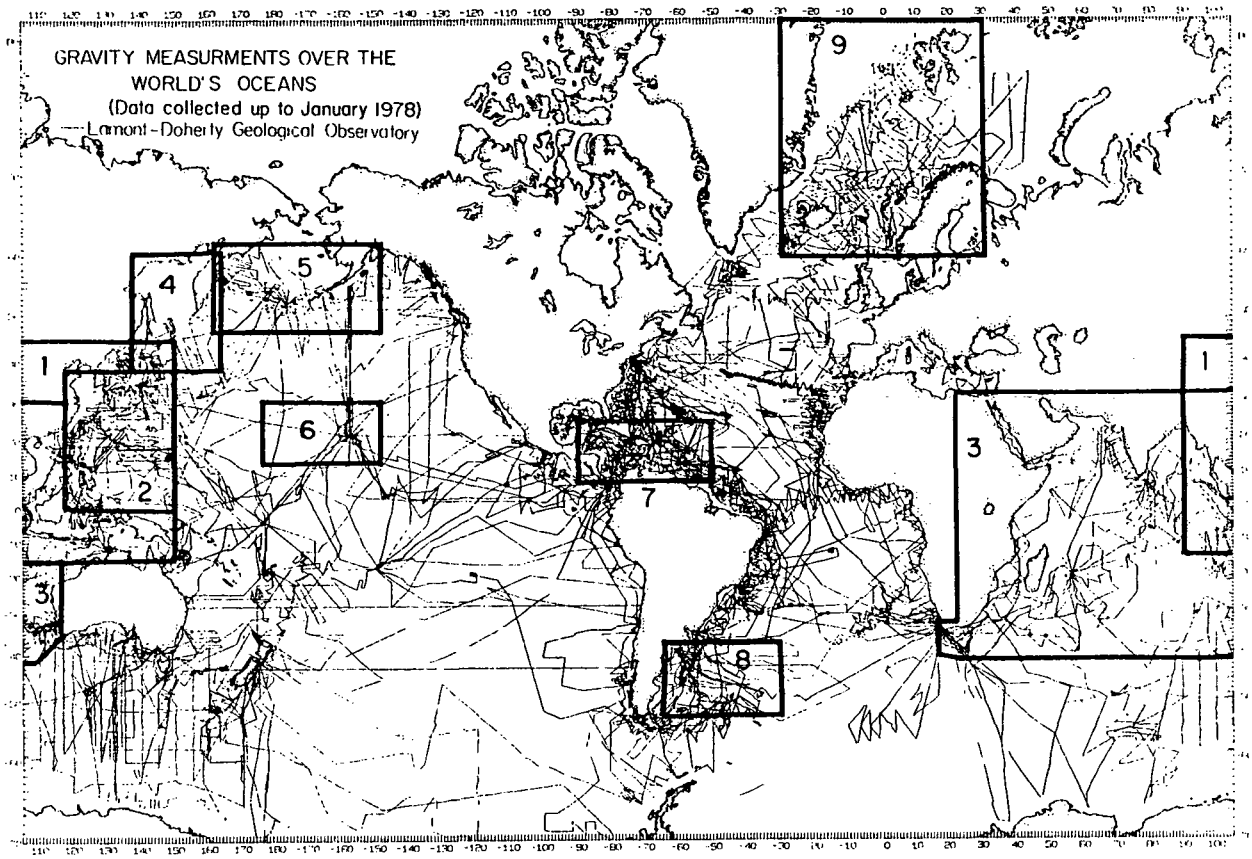



Figure 1. Location map showing coverage of gravity data of Lamont-Doherty research vessels R/V Robert D. Conrad, Vema and Eltanin (Table 1) up to October 1978. The heavy lines outline those recently published free-air gravity anomaly maps of the world's oceans with an area greater than $4 \times 10^4$ km$^2$. The sources of the maps are as follows: 1. Watts, Bowin, Bodine (1978); 2. Watts (1976); 3. Talwani and Kahle (1975); 4. Watts, Bodine, Kogan (in press); 5. Watts (1975); 6. Watts and Talwani (1974); 7. Bowin (1976); 8. Rabinowitz (1977); 9. Grønlie and Talwani (1978).

region. Thus, in these regions, which include a variety of different geological features, gravity anomalies can apparently be recovered from GEOS-3 altimeter data with a resolution of about 200 km and a standard error of about ±12 mgal.

The overall usefulness of GEOS-3 altimeter data for lithospheric studies can be evaluated by comparing these estimates of resolution with those which are required to define isostasy at the East Pacific rise crest and Hawaiian-Emperor seamount chain (Fig. 2). The "isostatic response functions" in Figure 2 explain surface-ship gravity data in the region of the Hawaiian-Emperor seamount chain and East Pacific rise crest with an average standard error of ±12 mgal and ±4 mgal respectively. These errors can be attributed to features of the gravity field of these regions which are not related to isostasy. Thus in order to provide information on the state of isostasy of the Hawaiian-Emperor seamount chain, a resolution of at least 200 km (Fig. 2) with a standard error of better than ±12 mgal is required, while at the East Pacific rise crest a resolution of at least 30 km (Fig. 2) with a standard error of better than ±4 mgal is required.

These considerations suggest GEOS-3 altimeter data may provide useful information on the state of isostasy of relatively young loads on old lithospheric plates (Hawaiian-Emperor seamount chain) but appears unlikely to provide useful information on isostasy of features formed on young oceanic crust near mid-ocean ridge crests.

The most promising use of GEOS-3 altimeter data appears to be in the improved definition of the long-wavelength gravity field. Of particular interest is the information which may be present in GEOS-3 altimeter data on deep processes in the Earth such as those which may be associated with mantle convection.

Acknowledgments. The marine gravity program at Lamont-Doherty Geological Observatory is presently supported by the Office of Naval Research contract N00014-75-C-0210

and National Science Foundation grants NSF OCE-77-07941, NSF OCE 77-10647, and NSF OCE 77-25993.

## References


Anderson, R.N., D.P. McKenzie, and J.G. Sclater, Gravity, bathymetry and convection in the Earth, Earth & Planet. Sci. Lett., 18, 391-407, 1973.

Bowin, C., Free-air gravity anomaly map of the Caribbean, Geol. Soc. Amer. Mem. 169, 1976.

Cochran, J.R., Gravity and magnetic investigations in the Guiana Basin, Western Equatorial Atlantic, Geol. Soc. Amer. Bull. 84, 3244-3268, 1973.

Cochran, J.R., and M. Talwani, Free-air gravity anomalies in the world's oceans and their relationship to residual elevation, Geophys. J. R. astr. Soc. 50, 495-552, 1977.

Detrick, R.S., and S.T. Crough, Island subsidence hot spots and lithospheric thinning, J. Geophys. Res. 83, 1236-1244, 1978.

Graf, A, and R. Schulze, Improvements on the Sea Gravimeter Gss2, J. Geophys. Res., 66, 1813-1821, 1961.

Gronlie, G., and M. Talwani, Geophysical Atlas of The Norwegian Greenland Sea, VEMA Research Series IV, Lamont-Doherty Geol. Obs., Palisades, N.Y., 1978.

Gunn, R., A quantitative study of isobaric equilibrium and gravity anomalies in the Hawaiian Islands, Franklin Inst. Jour., 236, 373-390, 1943.

Gunn, R., A quantitative study of the lithosphere and gravity anomalies along the Atlantic coast, Franklin Inst. Jour., 237, 139-154, 1944.

Gunn, R., Quantitative aspects of juxtaposed ocean deeps, mountain chains and volcanic ranges, Geophysics, 12, 238-255, 1947.

LaCoste, L.J.B., Measurement of gravity at sea and in the air, Rev. of Geophys., 5, 447-526, 1967.

Leitao, C.D., and McGoogan, J.T., SKYLAB radar altimeter: short-wavelength pertubations detected in ocean surface profiles, Science, 186, 1208-1209, 1975.

Leitao, C.D., C.L. Purdy, and R.L. Brooks, Wallops GEOS-C altimeter processing report, NASA TM X - 69357, 68p., 1975.

Lewis, B.T.R., and L.M. Dorman, Experimental isostasy, 2. An isostatic model for the USA derived from gravity and topographic data, J. Geophys. Res., 75, 3367-3386, 1970.

McKenzie, D., Surface deformation, gravity anomalies and convection, Geophys. J. R. astr. Soc. 48, 211-238, 1977.

McNutt, M.K., and R.L. Parker, Isostasy in Australia and the evolution of the compensation mechanism, Science, 199, 733-775, 1978.

Rapp, R.H., Geos-3 data processing for the recovery of Geoid undulations and gravity anomalies, J. Geophys. Res., submitted to special GEOS-3 volume.

Rabinowitz, P., Free-air gravity anomalies - Argentine Continental Margin, Amer. Assoc. Petrol. Geol., Tulsa, Oklahoma, 1977.

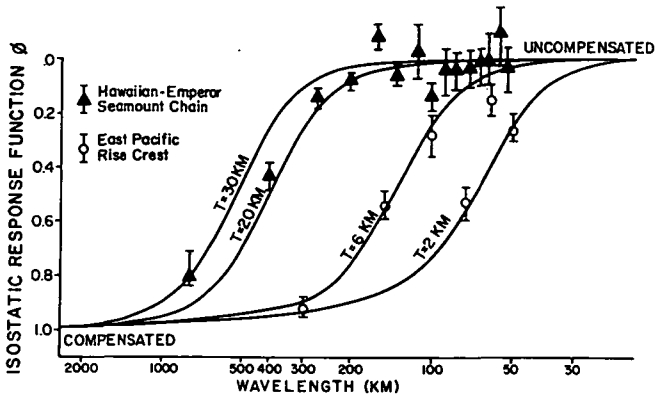




Figure 2. Plot of "isostatic response function" against wavelength for the Hawaiian-Emperor seamount chain and East Pacific rise crest in the Pacific ocean. This figure shows the range in wavelengths for which isostasy is effective at these features. The data in this plot is based on two recent studies by Watts (1978) and Cochran and Watts (in preparation).


Sclater, J.G., L.A. Lawver, and B. Parsons, Comparison of long wavelength residual elevation and free-air gravity anomalies in the North Atlantic and possible implications for the thickness of the lithospheric plate, J. Geophys. Res., 80, 1031-1052, 1975.

Talwani, M., and H.-G. Kahle, Free-air gravity anomaly maps of the Indian Ocean, in Geological and Geophysical Atlas of the Indian Ocean, (ed.): G. Udintsev, Moscow, 81-104, 1975.

Vening Meinesz, F.A., Gravity expeditions at Sea, 1923-1938, Vol. IV; Neth. Geod. Comm., Delft, 233p., 1948.

Vening Meinesz, F.A., Gravity over the Hawaiian Archipelago and over the Madiera Area; Conclusions about the Earth's crust, Proc. Kon. Ned. Akad. Wetensia, 44p., 1941.

Walcott, R.I., Flexure of the lithosphere at Hawaii, Tectonophysics, 9, 435-446, 1970.

Walcott, R.I., Gravity, flexure and the growth of sedimentary basins at a continental edge, Geol. Soc. Amer. Bull., 83, 1845, 1972.

Watts, A.B., Gravity field of the northwest Pacific ocean basin and its margin: Aleutian Island arc-trench system, Geol. Soc. Amer. Map and Chart Series, MC-10, 1975.

Watts, A.B., Gravity field of the northwest Pacific ocean basin and its margin: Philippine Sea, Geol. Soc. Amer. Map and Chart Series, MC-12, 1976.

Watts, A.B., Gravity and bathymetry in the Central Pacific Ocean, J. Geophys. Res., 81, 1533-1553, 1976.

Watts, A.B., An analysis of isostasy in the world's oceans. Part 1 - Hawaiian-Emperor seamount chain, in press, J. Geophys. Res., 1978.

Watts, A.B., J.H. Bodine, and M.G. Kogan, Gravity field of the northwest Pacific ocean basin and its margin: Kuril Island arc-trench system,

Geol. Soc. Amer. Map and Chart Series MC-27, in press.
Watts, A.B., C.O. Bowin, and J.H. Bodine, Free-air gravity field, in A geophysical Atlas of the East and Southeast Asian Seas, Geol. Soc. Amer. Map and Chart Series MC-25, 1978.
Watts, A.B., and J.R. Cochran, Gravity anomalies and flexure of the lithosphere along the Hawaiian-Emperor Seamount Chain, Geophys. J. R. astr. Soc., 38, 119-141, 1974.
Watts, A.B., J.R. Cochran, and G. Selzer, Gravity anomalies and flexure of the lithoslphere; A three-dimensional study of the Great Meteor Seamount, Northeast Atlantic, J. Geophys. Res., 80, 1391-1398, 1975.
Watts, A.B., and M. Talwani, Gravity field of the northwest Pacific ocean basin and its margin: Hawaii and vicinity, Geol. Soc. Amer. Map and Chart Series, MC-9, 1975.
Worzel, J.L., Pendulum gravity measurements at sea 1936-1959, J. Wiley & Sons, N.Y., London, Sydney, 422p. 1965.

Page Intentionally Left Blank

# Geoid Anomalies and the Near-Surface Dipole Distribution of Mass

D. L. Turcotte
Department of Geological Sciences, Cornell University
Ithaca, New York 14853

and

J. R. Ockendon
Computing Laboratory, Oxford University
Oxford, England OX13PR

Abstract. Although geoid or surface gravity anomalies cannot be uniquely related to an interior distribution of mass, they can be related to a surface mass distribution. However, over horizontal distances greater than about 100 km, the condition of isostatic equilibrium above the asthenosphere is a good approximation and the total mass per unit column is zero. Thus the surface distribution of mass is also zero. For this case we show that the surface gravitational potential anomaly can be uniquely related to a surface dipole distribution of mass. Variations in the thickness of the crust and lithosphere can be expected to produce undulations in the geoid.

## Introduction

The gravitational potential and acceleration can in general be obtained by integrating over any specified distribution of mass. In many cases, however, the detailed distribution of mass in the crust and mantle may be unknown. In these cases unique relationships between the gravity and geoid anomalies and surface distreibutions of density may be of considerable use. One example of such a relationship is the Bouguer formula for the gravity anomaly $\Delta g$,

$$\Delta g = 2\pi G \sigma(x,y) \tag{1}$$

where G is the gravitational constant and the surface density distribution is

$$\sigma(x,y) = \int_0^h \Delta\rho\ (x,y,z)\,dz \tag{2}$$

where $\Delta\rho$ is the density anomaly. The Bouguer formula is valid if the horizontal scale of the density variation is large compared with the vertical scale h and h<<a where a is the radius of the earth.

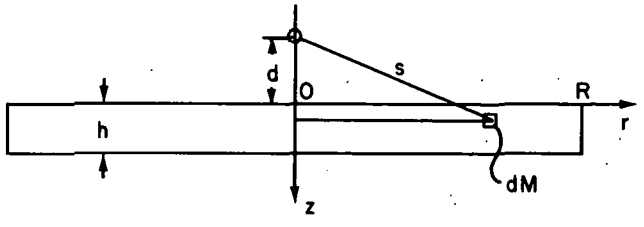



Fig. 1. Illustration of the circular disk formulation.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

Using the technique of matched asymptotic expansions, Ockendon and Turcotte (1977) have derived a power series expansion for the gravitational acceleration and potential caused by slowly varying density changes. They find that if the near surface density distribution is in isostatic equilibrium then the gravitational potential anomaly $\Delta U$ is given by

$$\Delta U = -2\pi G\delta\ (x,y) \tag{3}$$

where the surface dipole density distribution is

$$\delta(x,y) = \int_0^h z\ \Delta\rho(x,y,z)\,dz \tag{4}$$

The conditions for the validity of this relation are the same as for the Bouguer formula with the additional isostatic requirement that $\sigma = 0$, i.e., that the gravity anomaly given by the Bouger formula (1) is zero.

It is the purpose of this paper to give two elementary planar derivations of (3) and to test its validity for near surface density variations on the earth.

## Disk Approximations

We first consider a circular disk of thickness h and radius R as shown in Figure 1. The density of the disk is a function of the vertical coordinate z, $\rho(z)$, but not of r. The contribution to the gravitational acceleration of each element of mass in the disk is

$$d\vec{g} = -\frac{G\vec{s}\,dM}{|s|^3} \tag{5}$$

Integrating over the volume of the disk to obtain the vertical components of the gravitational acceleration on the axis at a distance d above yields

$$g_z = 2\pi G \int_0^h \int_0^R \frac{(d+z)\,r\rho(z)\,dr\,dz}{[r^2 + (d+z)^2]^{1/2}} \tag{6}$$

First integrating with respect to r and then taking the limit $R \to \infty$ gives

$$\lim_{R\to\infty} g_z = 2\pi G \int_0^h \rho(z)\,dz \tag{7}$$

which is the Bouguer formula previously given in (1).

The gravitational potential due to each element of mass is

$$dU = \frac{G\,dM}{|s|} \tag{8}$$

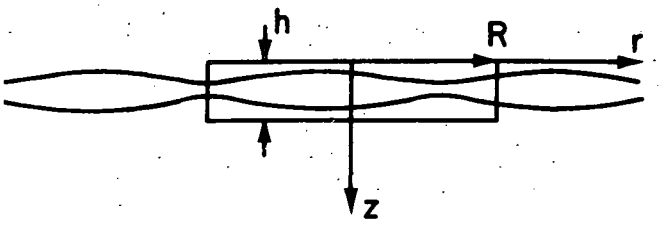



Fig. 2. Illustration of the application of Gauss' theorem to a thin layer of mass anomalies.

Integrating over the volume of the disk to obtain the potential at a height d on the axis of the disk yields

$$U = 2\pi G \int_0^h \int_0^R \frac{r\rho(z)\, dzdr}{[r^2 + (d+z)^2]^{1/2}} \tag{9}$$

Integration with respect to r and expanding for large R gives

$$U = 2\pi G \left[R\int_0^h \rho(z)dz - \int_0^h (d+z)\,\rho(z)\,dz + \frac{1}{2R}\int_0^h (d+z)^2\rho(z)\,dz + 0\,(R^{-3})\right] \tag{10}$$

First applying the condition of isostasy, i.e.,

$$\int_0^h \rho(z)\, dz = 0$$

and then taking the limit $R \to \infty$ yields

$$U = -2\pi G\int_0^h z\rho(z)dz \tag{11}$$

which is the formula previously given in (3).

## Mass-Layer Approximations

For mass anomalies confined to thin layers it is useful to integrate (5) over the cylindrical volume illustrated in Figure 2. Gauss' theorem may then be used to convert one of the volume integrals to a surface integral with the result

$$\iint \vec{g}\cdot\vec{ds} = -4\pi G\iiint \rho\, dv \tag{12}$$

In the limit $h \to 0$ these integrals can be evaluated to yield (Officer, 1974, pp. 262-269)

$$g_z^+ - g_z^- = 4\pi G\sigma \tag{13}$$

However by symmetry

$$g_z^+ = -g_z^- \tag{14}$$

so that

$$g_z^- = 2\pi G\int_0^h \rho\, dz \tag{15}$$

which is the Bouguer formula (1).

A surface dipole mass distribution is obtained by taking two equal surface mass distributions of opposite sign, $\sigma_1 < 0$ on $z = 0$ and $\sigma_2 = -\sigma_1$ on $z = h$ (Fig. 3a); the limits $\sigma_2 \to +\infty$, $\sigma_1 \to -\infty$ and $h \to 0$ are taken such that

$$\int_0^h z\rho dz = \sigma_2 h = \delta$$

is finite. It follows from (13) and (14) that

$$\begin{aligned} g_z &= 0 \quad z > h,\ z < 0 \\ g_z &= 4\pi G\sigma_2, \quad 0 < z < h \end{aligned} \tag{16}$$

Using the relationship between the gravitational field and potential, $g_z = \partial U/\partial z$ the difference in the potential across the dipole layer is

$$U^+ - U^- = 4\pi\, G\sigma_2\ h = 4\pi G\delta \tag{17}$$

We choose our origin for U such that

$$U^+ = -U^- \tag{18}$$

so that the distribution of U illustrated in Figure 3c is obtained and

$$U^- = -2\pi G\delta = -2\pi G\int_0^h z\rho dz \tag{19}$$

which is the same as (3).

In order to establish the quantitative validity of (3) we consider the gravity and potential fields due to spherical harmonic distributions of mass on spherical surfaces. The gravitational field just outside a spherical surface due to a surface mass distribution $\sigma_n S_n$ (where $S_n$ is the spherical surface harmonic of order n) on the surface is given by (Jeffreys, 1976, p. 234)

$$g = 4\pi G\left(\frac{n+1}{2n+1}\right)\sigma_n S_n \tag{20}$$

The wavelengths of the surface mass distribution can be related to the order of the harmonic n by

$$\lambda = \frac{2\pi a}{n} \tag{21}$$

where a is the radius of the sphere (of the earth). For short wavelength distributions we take the limit $n \to \infty$ in (20) with the result

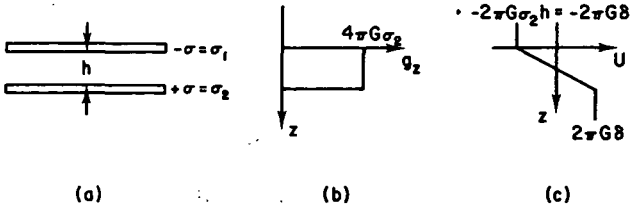



Fig. 3. Gravitational acceleration (b) and potential (c) associated with a dipole mass distribution (a).

$$\lim_{n\to\infty} g = 2\pi G \sigma_n S_n \qquad (22)$$

which is the Bouguer formula.

In order to represent a dipole distribution of mass we consider a spherical harmonic distribution of surface mass on a sphere of radius r = a-h with amplitude $K_n\sigma_n S_n$ in addition to the distribution of surface mass $\sigma_n S_n$ on the sphere r = a. The resulting gravitational potential just outside the outer sphere is given by (Jeffreys, 1976, p. 237)

$$U = \frac{-4\pi G}{(2n+1)} \sigma_n S_n a \left[1 + K_n \left(1 - \frac{h}{a}\right)^{n+2}\right] \qquad (23)$$

The condition of isostasy requires an equal mass defect on the inner sphere to that on the outer sphere. Allowing for the difference in area we require (Jeffreys, 1976, p. 237)

$$K_n = -\frac{a^2}{(a-h)^2} \qquad (24)$$

Substitution of (24) into (23) yields

$$U = \frac{-4\pi G}{(2n+1)} \sigma_n S_n a \left[1 - \left(1 - \frac{h}{a}\right)^{n}\right] \qquad (25)$$

Taking the limit h/a → 0 in (25) gives

$$\lim_{h/a\to 0} U = \frac{-4\pi n G \sigma_n S_n h}{2n+1} \qquad (26)$$

Next taking the limit n → ∞ we find

$$\lim_{n\to\infty} U = -2\pi G \sigma_n S_n h \qquad (27)$$

and noting that $\sigma_n S_n h$ is the surface dipole distribution of mass this is the same as (3).

By using (25) we can compare the results for finite depths of compensation h with the limiting solution given in (27). This is done in Figure 4. The ratio of the surface potential U from (25) to the value $2\pi G\sigma_n S_n h$ is given as a function of 1/n for h = 25, 50, 100, 200, and 400 km. The corresponding wavelengths from (21) are also included. We see that the approximation breaks down as expected when the wavelength is of the same order as the depth of compensation, i.e., as λ/h → 1. For depths of compensation of 50 km or less the error in using (27) is 10% or less over a wide range of wavelengths. It should be emphasized that the case of two mass layers is an extreme case of compensation at depth. The realistic case of distributed mass with depth will lead to lower errors than those given in Figure 4 if the mass differences are limited to a depth h.

## Discussion

The gravitational field and potential outside a closed surface can be uniquely related to a surface distribution of mass. The Bouguer formula (1) relates the local gravity anomaly to the magnitude of the surface mass distribution σ. However, over horizontal distances greater than about 100 km on the earth's surface, the condition of isostasy is a good approximation and requires that the surface mass distribution of σ be zero.

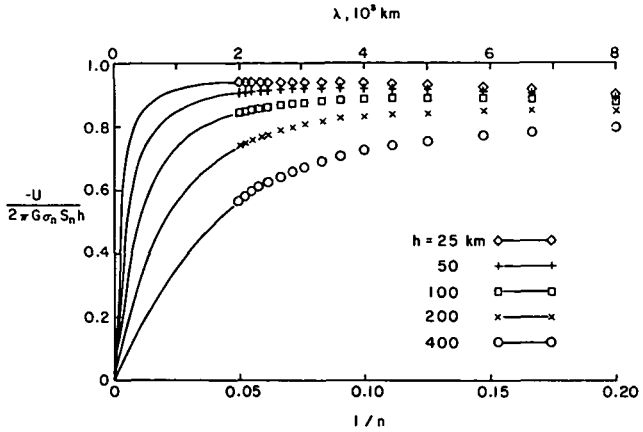



Fig. 4. The ratio of the surface potential U on the earth from (25) to the value for a surface dipole layer (27) as a function of 1/n for various values of the depth of compensation h. Also included are the values of the wavelength λ corresponding to the value of 1/n from (21).

If the surface mass distribution is zero the gravitational field and potential outside a closed surface can be uniquely related to a surface dipole distribution of mass. In this case (3) relates the local gravitational potential anomaly to the magnitude of the surface dipole layer σ. The measured distribution of surface potential anomalies can be directly used to obtain a surface distribution of the density dipole strength. This surface mass dipole distribution can be directly related to the density distribution in the crust and lithosphere, although there will also be other, deeper contributions to the external gravitational field and potential. We have shown that the local association of the potential anomaly with the dipole density distribution is a good approximation for the depths of compensation associated with the crust or lithosphere.

The gravitational potential anomaly is directly proportional to the geoid anomaly. The geoid anomaly is measured directly by radar altimetry from the GEOS-3 satellite. Haxby and Turcotte (1978) have shown that several measurements of geoid anomalies can be related to density variations in the crust and lithosphere using (3).

Acknowledgements. This research was supported by the National Aeronautics and Space Administration, NASA Grant NSG 5060. The authors would like to acknowledge stimulating discussions with A. Kuckes and W. Kaula. This is contribution 608 of the Department of Geological Sciences, Cornell University.

## References

Haxby, W.F., and D.L. Turcotte, On isostatic geoid anomalies, J. Geophys Res., in press, 1978.

Jeffreys, H., The Earth, Cambridge University Press, Cambridge, 1976.

Ockendon, J.R., and D.L. Turcotte, On the gravitational potential and field anomalies due to thin mass layers, Geophys. J.R. astr. Soc., 48, 479-492, 1977.

Officer, C.B., Introduction to Theoretical Geophysics, Springer-Verlag, New York, 1974.

# Interpretation of Altimeter Data

Micheline C. Roufosse
Harvard-Smithsonian Center for Astrophysics
Cambridge, Massachusetts 02138

Abstract. Two methods are used to interpret the whole range of signals contained in the Geos 3 altimeter data. They each address a different class of events and thus complement each other in their ability to provide information on the state of convection in the earth's mantle. The long wavelength section of the spectrum yields information on the depth of the convection cells and the viscosity variations inside those cells through a study of the variations of the admittance as a function of wavelength. The short wavelength section of the spectrum provides information on the time evolution of the lithosphere, considered as a thin elastic plate, by studying its response to loads at several points in its evolution. The variation of the flexural rigidity with age is obtained from that study.

## Introduction

The geoid heights derived from the Geos 3 experiment contain signals covering a whole spectrum of wavelengths that are related to various processes. It is thus useful to separate the whole spectrum into two large classes: short wavelength that covers signals smaller than 500 to 1000 km width and large wavelength that includes signals of larger width. These two ranges of wavelengths must be handled separately; their different origins require different methods of interpretation. The overall problem addressed by both methods is to characterize convection in the earth's mantle by use of two complementary approaches.

## Long Wavelength Study

The central idea of the long wave study was developed by McKenzie and his coworkers in a series of numerical studies on convective flow [McKenzie, Roberts, and Weiss, 1974; McKenzie and Weiss, 1975; McKenzie, 1977]. Several two-dimensional models with different parameters were studied in order to determine how they affect the state of convection in the earth's mantle and how they relate to the observed quantities accessible to geophysicists. The result of this analysis is that the most relevant quantity to study is the behavior of the admittance as a function of wavelength; the admittance is defined as the ratio, in wavenumber space, of the gravity to the bathymetry.

The study of a wide range of numerical models shows that the admittance is insensitive to the Rayleigh number and the degree of internal heating but is strongly affected by viscosity variations and the depth of the convecting layer or the deformability of the lower boundary. So far, very little information has been available on the variation of the admittance with wavelength in oceanic regions as a result of inadequate gravity data coverage. Data collected over oceanic regions during the Geos 3 mission have solved that problem; and currently, adequate geoid height data are available over most oceanic regions. It is thus possible to derive the gravity field directly on the geoid by combining radar altimeter data, range-rate residuals, and surface ship observations. Each of the three sets of data must be written into a coherent network that reduces crossing errors to a minimum. Deriving the gravity field is then a linear inversion problem, which can be treated by any standard method. The Backus-Gilbert method, however, offers the advantage that both the gravity field and an estimate of its error as a function of latitude and longitude are obtained directly. Both bathymetry and gravity must then be Fourier-transformed into wavenumber space and divided by one another to give the admittance as a function of wavelength. This study will put more definite constraints on the lower boundary and viscosity variations characterizing convection in the mantle.

## Short Wavelength Study

The short wavelength signals in the geoid heights yield information on the time behavior of the lithosphere. Following Crough [1975], we can consider the lithosphere as a thin plate whose thickness increases with increasing time up to a certain age, of the order of 80 m.y., and then continues to increase at a progressively lower rate until it reaches equilibrium thickness. Since the thickness of a plate influences its mechanical properties, it is possible to study the time evolution of the lithosphere by observing how it deforms when loaded by seamounts placed at several points along its evolutionary path. To examine the mechanical properties of the lithosphere, we assumed the thin-plate model developed by McKenzie and Bowin [1976]. In this model, the lithosphere consists of a thin elastic plate overlying a fluid medium; the plate is being loaded by bathymetric features such as seamounts, island chains, and ridges and is subsequently deformed. The magnitude and wavelength of the deformed area depend mostly on the flexural rigidity, which is proportional to the cube of the lithospheric thickness. By studying the correlation function between the geoid height and the bathymetry, we can determine the flexural rigidity of the area under investigation. This can be done in one of two ways: the first is to Fourier-transform the geoid height and the bathymetry into wavenumber space, divide the geoid height by the bathymetry, and obtain the response function as a function of wavelength; the flexural rigidity can then be deduced from the characteristics of this function. The second method is to calculate a theoretical filter Z(k) in wavenumber space by using the thin-plate model [McKenzie and Bowin, 1976] and varying the values for the

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

flexural rigidity:

$$Z(k) = \frac{3(\rho_c - \rho_w)}{2r\rho_e\gamma} \frac{(1 - e^{-wkt}) e^{-wkd}}{[1 + (wk)^4] wk} \quad (1)$$

where

$$\gamma = \left[\frac{(\rho_m - \rho_c)}{F}\right]^{1/4} \quad (2)$$

$$w = \frac{2\pi}{n\Delta\gamma} \quad (3)$$

In these expressions, $\rho_c$, $\rho_w$, $\rho_m$, and $\rho_e$, are, respectively, the crustal, water, mantle, and mean-earth densities, r is the earth's equatorial radius, t is the crustal thickness, d is the water depth, g is the average gravity, F is the flexural rigidity, n is the number of points in the filter, and $\Delta$ is the spacing between consecutive points of the filter. The filter derived in equation (1) is then Fourier-transformed into direct space and convolved with the bathymetry, resulting in a theoretical geoid height. The value for the flexural rigidity that gives the best agreement between predicted and observed geoid heights is the one that will be selected for each area studied.

In practice, the method chosen will depend on the type of data available. The first method is more adequate when comparing gravity and bathymetry data from surface ships because both sets of data give equispaced points and can thus be easily Fourier-transformed. The second method, however, is preferable when dealing with Geos 3 altimeter data because it is not dependent on having both bathymetry and geoid-height data in a Fourier-transformable format. The Geos 3 data are easily transformed, but the bathymetry data must be reconstructed, as rigorously as possible, along the subsatellite position by using bathymetric contour charts; this operation generally results in poor accuracy and irregular point spacing. The second method, the two-dimensional approach, is thus the one we have used to study the evolution of the lithosphere. The regions

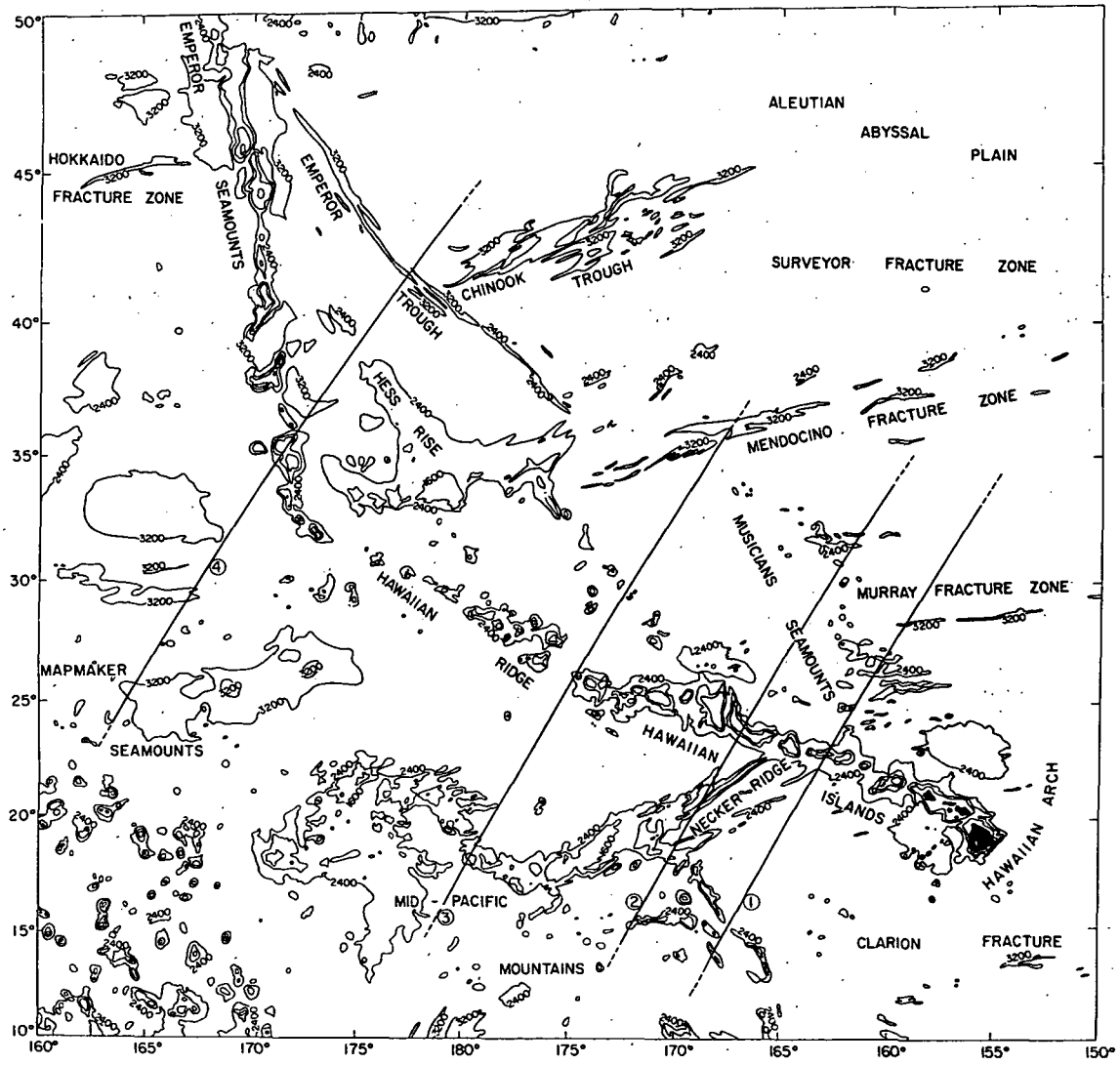



Fig. 1. Geos 3 passes studied in the Hawaiian-Emperor Seamounts region.



C-4

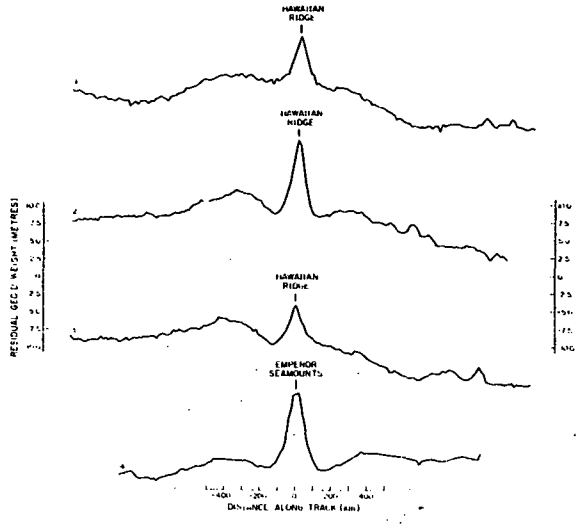



Fig. 2. Observed geoid height profiles in the Hawaiian-Emperor Seamounts region, represented with respect to a reference geoid of degree and order 12.

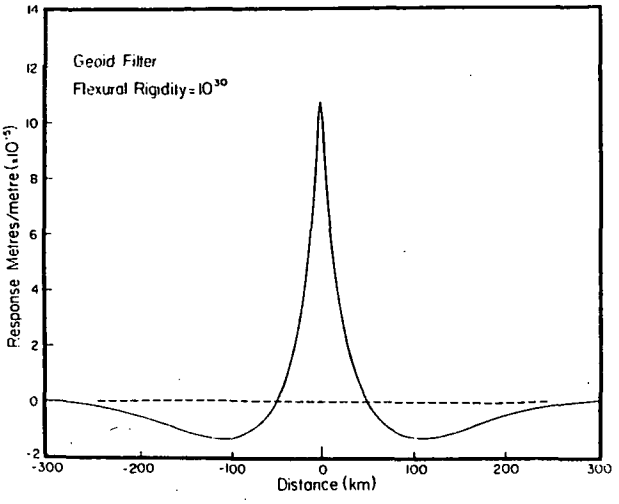



Fig. 3. Theoretical geoid filter calculated with a flexural rigidity of $10^{30}$ dyne cm.

studied so far are the Hawaiian-Emperor Seamount chain, the Marshall-Gilbert Island chain, and the Crozet Islands; these areas will be part of a much larger network.

The Hawaiian-Emperor Seamounts were selected first because they have been previously studied by other methods from other sets of data [Watts and Cochran, 1974; Walcott, 1976]. This area therefore constitutes an ideal testing ground for the two-dimensional technique. The altimeter passes from Geos 3 selected in that area are superposed on a map of the region in Figure 1, and their profiles are shown in Figure 2. The profiles are represented with respect to a reference geoid of degree and order 12 calculated from Standard Earth IV spherical-harmonics coefficients; they all show the features typical of the region — a sharp peak centered on the island chain flanked by a shallow depression and superposed on an asymmetrical bulge.

We then calculated theoretical filters using values for flexural rigidity ranging from $10^{29}$ to $10^{31}$ dyne-cm; an example, with a flexural rigidity of $10^{30}$ dyne-cm, is represented in Figure 3. After convolving the filters with the reconstructed bathymetry, we got the results shown in Figure 4. The top profile in the figure is the observed geoid represented with respect to a reference geoid of degree and order 16, which was chosen in order to remove the unwanted long-wavelength features; the middle profile is the best-fitting predicted geoid, obtained with a value of $10^{30}$ dyne-cm for the flexural rigidity; and the bottom profile is the bathymetry reconstructed along the subsatellite positions from the bathymetric charts designed at Scripps Institute of Oceanography by Chase, Menard, and Mammerickx [1970]. Our values for all passes are in close agreement with those determined by other methods [Watts and Cochran, 1974; Suyenaga, 1977].

In the framework of the time evolution of the lithosphere, two important observations were made in that region:

1. Predictions for the mid-Pacific Mountains are best when a lower value for flexural rigidity is used. This is in quantitative agreement with a study by Larson [1976] on the evolution of the Western Central Pacific Ocean. Interpreting the magnetic anomalies, he characterized the mid-Pacific Mountains as a slowly spreading center. A reduced plate thickness would thus be expected, as observed in the present work.

2. The flexural rigidity associated with the Emperor Seamounts is somewhat smaller then the easternmost active volcanoes of the Hawaiian chain, as shown in Figure 5; it is on the order of $8 \times 10^{29}$ dyne-cm. The smaller value can be explained by taking into account the age of the seamounts along the chain [Clague and Jarrard, 1973]. In a recent study, Watts [1978] observed the correlation between gravity and bathymetry data obtained from surface ships over several sections of the Pacific Ocean: the East Pacific Rise, the Hawaiian-Emperor Seamounts, and the Kuril Rise. He deduced that the relevant factor related to flexural rigidity is the age of the lithosphere at the time of loading. Although the lithosphere in the case of the Emperor Seamounts is older than it is at the head of the Hawaiian chain, the load there is proportionally older, and therefore the lithosphere at the time of loading was younger, thus requiring a smaller flexural-rigidity value. So far, only one Geos 3 track is available for interpretation in that region; the difference observed is within the range of permissible uncertainties, but we cannot draw any definitive conclusions until more tracks along the seamount chain have been studied.

The two other areas investigated — the Marshall Gilbert Island and the Crozet Island — exhibited quite similar behavior to that found above. In the Geos 3 track shown crossing the Gilbert Islands chain in Figure 6, the top profile represents the bathymetry reconstructed from the Chase et al. chart, and the bottom profile is the

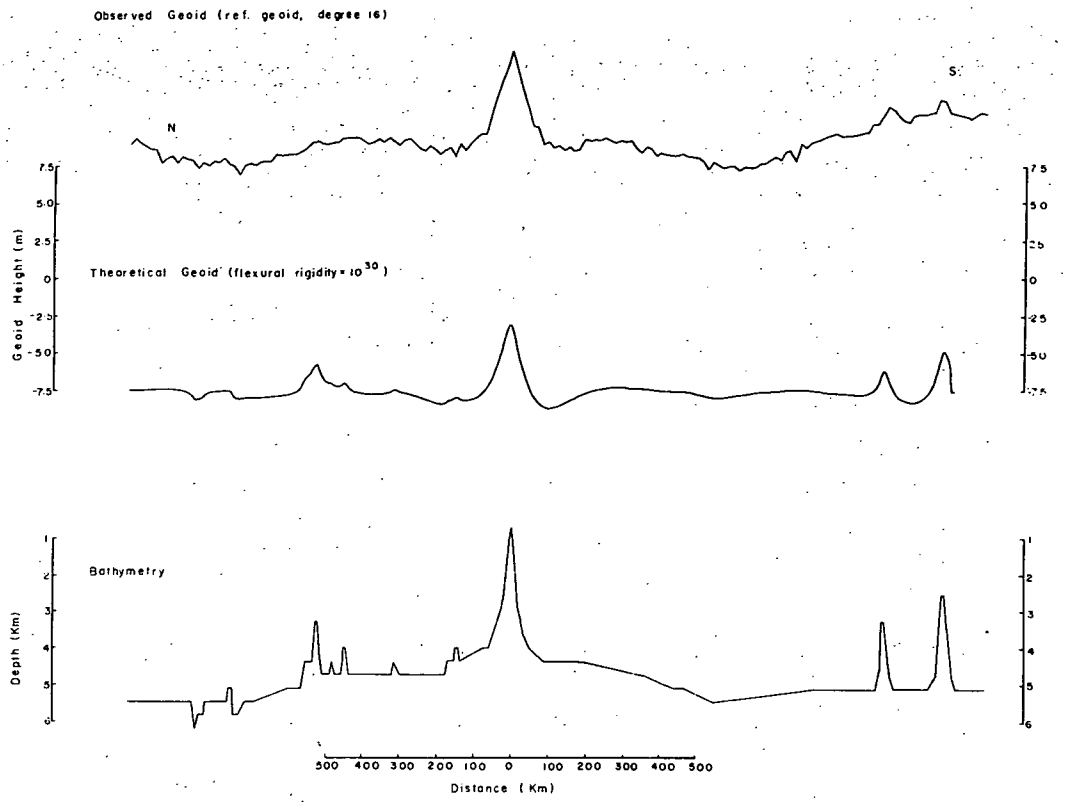



Fig. 4. The top profile represents the observed geoid height with respect to a reference geoid of degree and order 16 in the Hawaiian region; the intermediate profile is the predicted geoid, calculated with a filter of flexural rigidity of $10^{30}$ dyne-cm, and convolved with the bathymetry represented on the bottom profile.

observed geoid height with respect to a 16th-order reference geoid. The Gilbert Islands are the sharp feature seen at 1200 km. Figure 7 shows a series of predicted geoids in the same area computed with several values for flexural rigidity. Both the width and the height of the signal vary considerably, and the best-fitting value in this case is between 0.5 and 0.75 x $10^{30}$ dyne-cm. The lithosphere is quite old in that area, of the order of 120 m.y., and therefore a large lithospheric thickness is expected, which is inconsistent with the small value found for the flexural rigidity. To reconcile the present observation with Watt's model, it could be speculated that the Gilbert Islands constitute old loads.

The study of the Crozet Plateau was done in collaboration with Dr. Anny Cazenave, from Groupe de Recherche et de Geodesie Spatiale and Centre National d'Etudes Spatiales, Toulouse, who provided the relevant Geos 3 profiles. In a recent work, Cazenave and Lambeck (in preparation) analyzed the geoid anomalies in that region using the three-dimensional approach developed by Watts, Cochran, and Selzer [1975] for their study of the Great Meteor Seamount. Cazenave and Lambeck found that flexural-rigidity values ranging from 0.75 to 1 x $10^{30}$ dyne-cm gave an excellent fit between observed and predicted geoids. When we applied the two-dimensional approach described above to the Crozet Islands, we obtained a flexural-rigidity value similar to theirs; this can be seen by comparing the observed geoid plotted in Figure 8 and the predicted geoids shown in Figure 9.

In the Crozet Plateau region, the age of the load is unknown, and the age of the lithosphere, according to Schlich [1975], is Upper Cretaceous. A comparison of Cazenave and Lambeck's method with ours suggests that the two-dimensional approach is ideal for studying linear features such as island or seamount chains, while it offers less precision for dealing with individual features. In the case of the Crozet Plateau, only those tracks crossing the maximum altitude of the plateau gave a correct value for the flexural rigidity; all others resulted in larger values, owing to the fact that they reproduced only the lower bathymetric points, whereas, in reality, the actual observed geoid is influenced by nearby masses. Therefore, the feature being studied will dictate whether we use the two- or the three-dimensional approach.

In the future, we intend to collect and study as many features as possible with various ages for the load and various ages of the lithosphere in order to deduce a relationship between flexural rigidity and age of the lithosphere. That study will first be carried systematically in the Pacific Ocean and then will be extended to all oceanic regions. This could be used as a method to derive the age of unknown loads.

Acknowledgment. This work was supported in part by contract SR 33852 with the Massachusetts

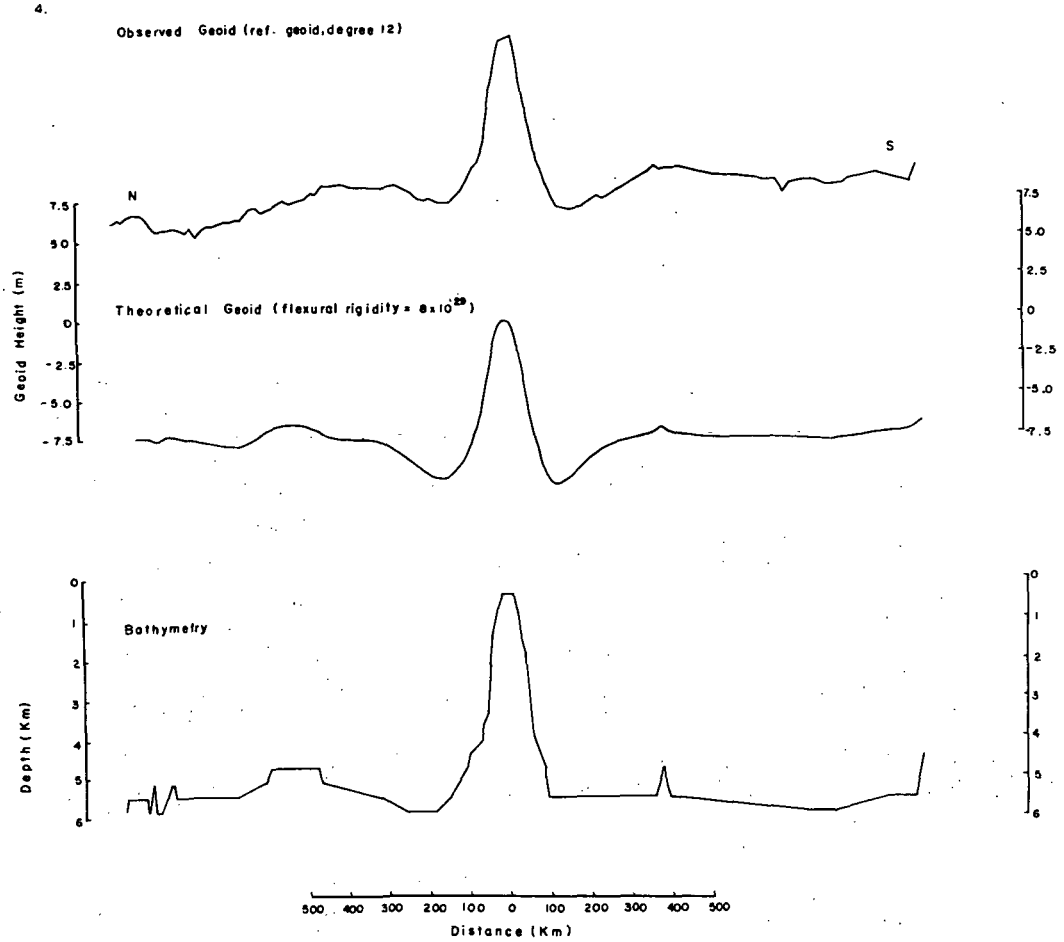



Fig. 5. The top profile represents the observed geoid with respect to a reference geoid of degree and order 12 in the Emperor Seasmount region; the intermediate profile is the predicted geoid calculated with a filter of flexural rigidity 8 x $10^{29}$ dyne-cm and convolved with the bathymetry represented on the bottom profile.

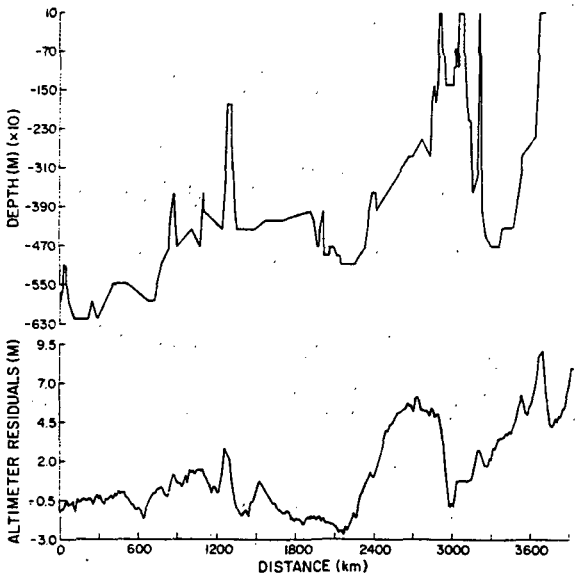



Fig. 6. The top profile represents the bathymetry in the Gilbert Islands region; the bottom profile represents the observed geoid with respect to a reference geoid of degree and order 16.

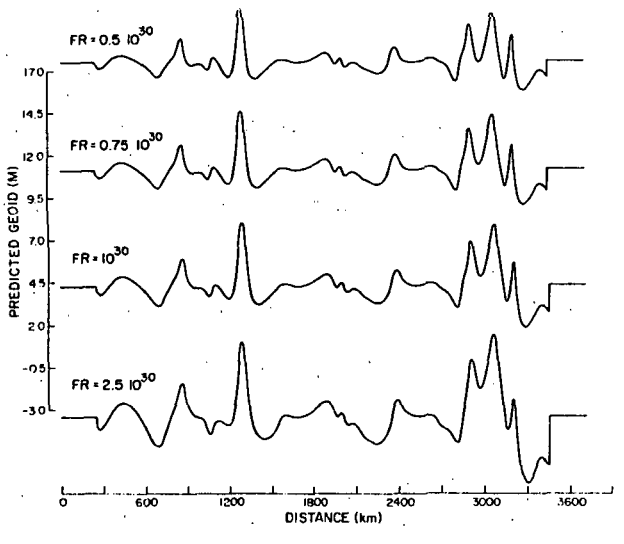



Fig. 7. The four profiles represent the predicted geoid in the Gilbert Islands region calculated with filters of different flexural rigidities: 0.5, 0.75, 1., and 2.5 (x $10^{30}$) dyne-cm, respectively.

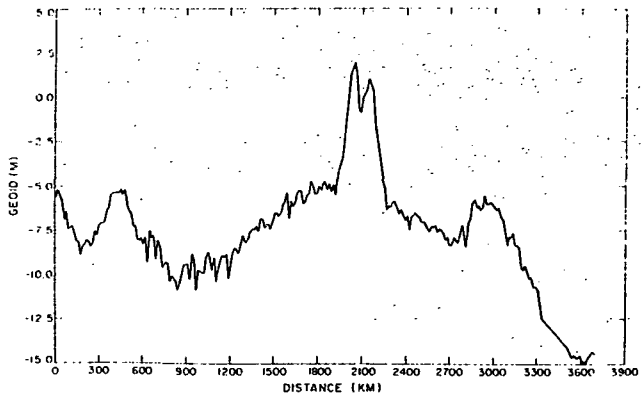



Fig. 8. Observed geoid with respect to a reference geoid of degree and order 16 in the Crozet region.

Institute of Technology and in part by contract F19628-78-C-0003 from the Air Force Geophysical Laboratory.

References

Chase, T.E., H.W. Menard, and J. Mammerickx, Bathymetry of the North Pacific, Scripps Institution of Oceanography and Institute of Marine Resources, 1970.

Clague, D.A., and R.D. Jarrard, Tertiary Pacific plate motion deduced from the Hawaiian-Emperor chain, Geol. Soc. Am. Bull., 84, 1135-1154, 1973.

Crough, S.T., Thermal model of oceanic lithosphere, Nature, 256, 388-390, 1975.

Larson, R.L., Late Jurassic and early Cretaceous evolution of the western central Pacific Ocean, Journ. Geomagn. Geochem., 28, 219-236, 1976.

McKenzie, D.P., Surface deformation, gravity anomalies and convection, Geophys. Journ. Roy. Astron. Soc., 48, 211-238, 1977.

McKenzie, D.P., J.M. Roberts, and N.O. Weiss, Convection in the earth's mantle: towards a numerical simulation, Journ. Fluid Mech., 62, 465-538, 1974.

McKenzie, D.P., and N.O. Weiss, Speculations on the thermal and tectonic history of the earth, Geophys. Journ. Roy. Astron. Soc., 42, 131-174, 1975.

McKenzie, D.P., and C. Bowin, The relationship between bathymetry and gravity in the Atlantic Ocean, Journ. Geophys. Res., 81, 1903-1915, 1976.

Schlich, R., Structure et age de l'Ocean Indien Occidental, Mem. Hors-Sene No. 6, Soc. Geol. France, Paris, 103, 1975.

Suyenaga, W., Earth deformation in response to surface loading, EOS, Trans. AGU, 58, 1231, 1977.

Walcott, R.I., Lithospheric flexure, analysis of gravity anomalies and the propagation of seamount chains. In "The Geophysics of the Pacific Ocean Basin and its Margin," ed. by G.H. Sutton, M.H. Manghnani, and R. Moberly, AGU Geophys. Mono. 19, Washington, D.C., pp. 431-438, 1976.

Watts, A.B., An analysis of isostasy in the world's oceans: Part 1 - Hawaiian-Emperor Seamount chain. Journ. Geophys. Res. (in press).

Watts, A.B., and J.R. Cochran, Gravity anomalies and flexure of the lithosphere along the Hawaiian-Emperor Seamount chain, Geophys. Journ. Roy. Astron. Soc., 38, 119-141, 1974.

Watts, A.B., J.R. Cochran, and G. Selzer, Gravity anomalies and flexure of the lithosphere: a three-dimensional study of the Great Meteor Seamount, Northeast Atlantic, Journ. Geophys. Res, 80, 1391-1398, 1975.

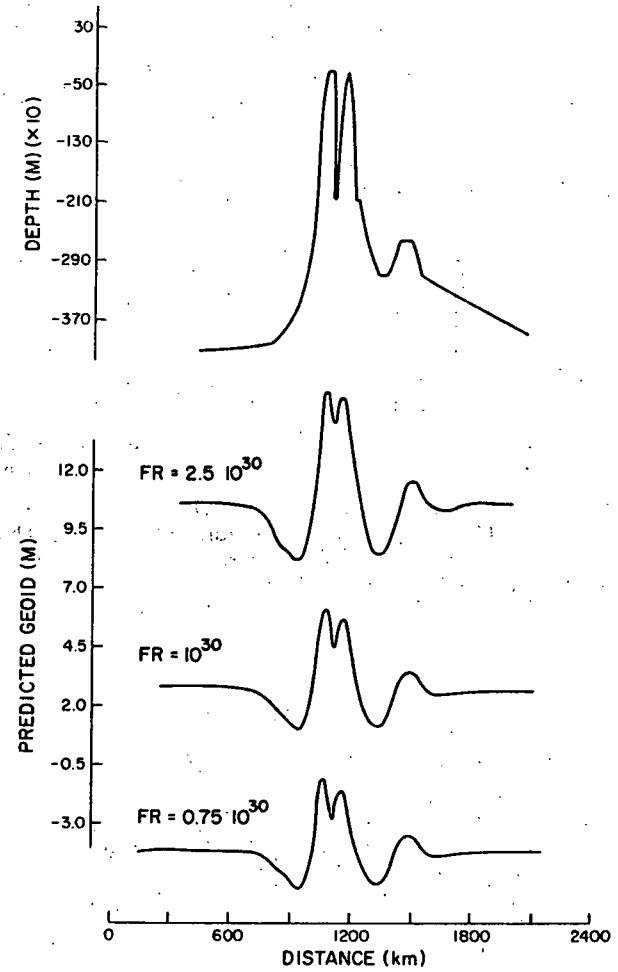



Fig. 9. The top profile is the bathymetry in the Crozet region; the three bottom profiles represent the predicted geoid calculated with filters of different flexural rigidities: 2.5, 1., and 0.75 (x $10^{30}$) dyne-cm, respectively.

# Comparison of Surface and Satellite Gravity Data

Richard H. Rapp
Department of Geodetic Science, The Ohio State University
Columbus, Ohio 43210

Abstract. Satellite derived potential coefficients (GEM 9) are compared to terrestrial gravity data by degree in terms of coefficient differences and in terms of mean anomaly differences. We found the root mean square undulation difference (to degree 20) was ±9 m and the anomaly difference was ±7 mgals with GEM 9 commission errors of ±1.7 m and ±3.8 mgals. The standard deviations of the GEM 9 implied undulations increased from ±4 cm at degree 2 to ± 53 cm at degree 20. The corresponding values implied by a recent (June 1978) terrestrial 5° field were ±2.53 m and ±0.38 m (at degree 20).

Comparisons of 5° equal area and 1° x 1° blocks showed discrepancies of ±11 and ±25 mgals respectively when using the GEM 9 coefficients to degree 20. Comparisons between Geos-3 altimeter derived anomalies and 1° x 1° terrestrial data showed that ±6-8 mgals is a reasonable accuracy estimate for the altimeter derived anomalies. Limited comparisons have also been made with anomalies derived from satellite to satellite tracking data indicating an accuracy of about ±6 mgals for the recovery of 5° equal area blocks.

## Introduction

The determination of the gravity field of the earth has been one of the classic goals of geodesy. The uses of the gravity field in geodesy originally related to geoid undulation and deflection of the vertical computation. Later applications arose in trajectory and orbit computations. Now we see needs for the global gravity field for better understanding the process in the earth's interior.

Initially gravity measurements were made with pendulums and then gravimeters which made accurate relative measurements. Even with rapid progress in equipment and techniques there are gaps in the earth's terrestrial gravity coverage.

The use of satellites to determine potential coefficients improved the situation with regard to the long wavelength behavior of the gravity field. Specifically, gravity anomalies can be derived from these potential coefficients. In early computations, a comparison of anomalies derived from potential coefficients with terrestrial mean anomalies was made to evaluate various potential coefficient sets derived from the analysis of satellite orbits (Kaula, 1966). Such procedures not only gave some indications of which potential coefficient solutions might be more reliable, but they also gave some confidence that there was some agreement between the satellite derived anomalies and the terrestrial anomalies.

These anomaly comparisons have continued for a number of different purposes (Lambeck, 1971, Rapp, 1972, 1975). This paper is an attempt to look at the current situation in several different ways.

## The Terrestrial Data

For our comparisons we will be using a recently (June, 1978) updated set of 1° x 1° mean gravity anomalies. This updating started from a set of 35011 anomalies supplied by the Defense Mapping Agency Aerospace Center in St. Louis. We updated this set of anomalies by adding, replacing or deleting 11933 values. These values were obtained from various sources such as recently published maps or data sent by various organizations for our use. The final data set contained 39405 1° x 1° anomalies some of which had been estimated thru geophysical correlation techniques. The location of these anomalies is shown in Figure 1.

In the update that we performed we would often find anomaly estimates from two sources that were widely different. As an example I show in Table 1 1° x 1° anomaly estimates for three blocks from different sources.

Clearly the differences are not small. Nor are such discrepancies unusual. However the number (on the order of 100-200) of such discrepancies are small on a percentage basis. Thus there are a number of areas where we have a poorly defined 1° x 1° mean anomalies.

In summary we will work in our comparisons with 39405 1° x 1° mean anomalies where the root mean square standard deviation is ±16 mgals but where some standard deviations may be as large as ±81 mgals.

Table 1. Location of Larger Discrepancies Between Terrestrial Data Sources

| $\varphi^\circ$ | $\lambda^\circ$ | Source A | Source B | Difference |
|---|---|---|---|---|
| 27 | 85 | -15±10 mgals | -147±9 mgals | -132 |
| 62 | 216 | -44±15 mgals | 106±19 mgals | 150 |
| -5 | 134 | -55±11 mgals | 72±22 mgals | 127 |

## Anomaly Computations from Potential Coefficients

The usual procedure to compute gravity anomalies from fully normalized potential coefficients ($\bar{C}_{\ell m}, \bar{S}_{\ell m}$) is (Rapp, 1977a)

$$\Delta g = \frac{kM}{r^2}\sum_{\ell=2}^{\infty}(\ell-1)\left(\frac{R_B}{r}\right)^{\ell}\sum_{m=0}^{\ell}(\bar{C}_{nm}\cos m\lambda + \bar{S}_{\ell m}\sin m\lambda)\,\bar{P}_{\ell m}(\sin\bar{\varphi}) \qquad (1)$$

where $R_B$ is the radius of the Bjerhammar sphere which is somewhat imbedded within the earth and r is the geocentric distance to the point in question. In

Proc. of the 9th GEOP Conference. *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

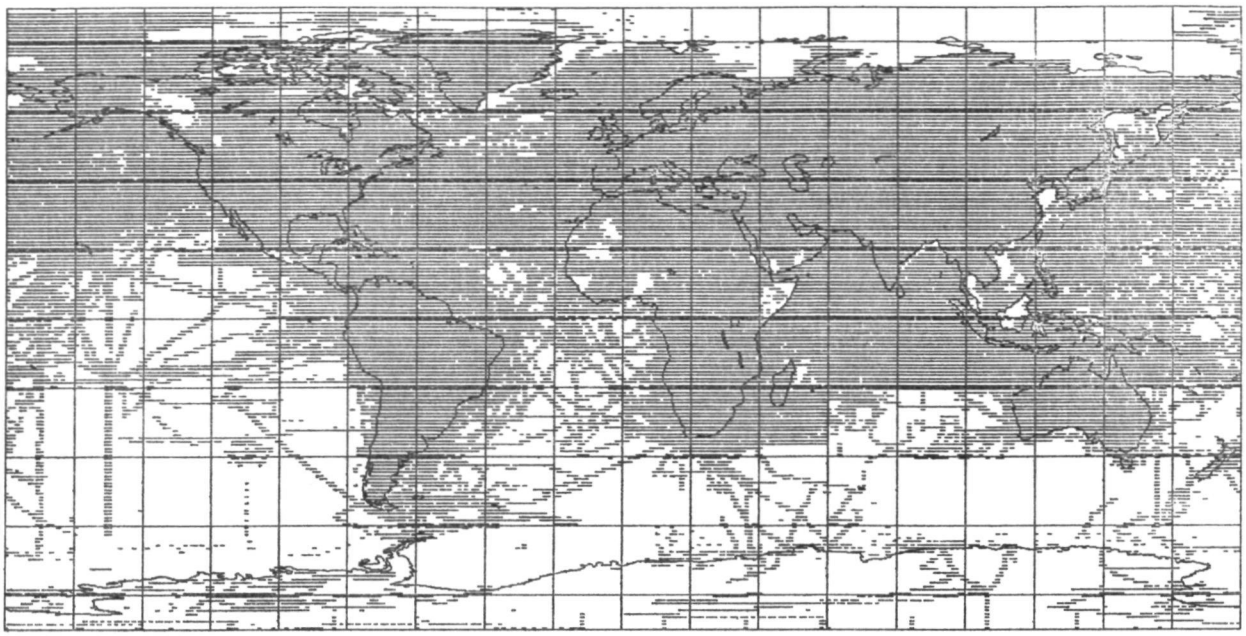


Figure 1. Location of the 39405 1° x 1° Anomalies of the June 1978 Data Tape.

practice $R_e$ is usually taken to be that equatorial radius (or scale factor) used in the determination of the potential coefficients. Equation (1) will yield the anomalies with respect to an ellipsoid of defined flattening implying a set of reference $\bar{C}_{\ell,0}$ coefficients ($\ell$ even) which is subtracted from the original $\bar{C}_{\ell,0}$ values. Such anomalies are given with respect to an ellipsoid with no atmospheric mass. Most terrestrial anomalies have been computed with respect to a gravity formula in which the mass of the atmosphere is included. For consistency purposes 0.87 mgals should be subtracted from the anomaly obtained from (1).

Another problem continues to appear that is related to the convergence of the equation (1) at the surface of the earth. Various solutions to this problem have been discussed. Recently Moritz (1978) has argued that although the series diverges at the surface of the earth, a practical convergence can be expected when using a finite set of coefficients. Arnold (1978) recently claims to have proven that the spherical harmonic expansion does converge on the surface. Sjöberg (1977) has given some simple examples demonstrating divergence. To avoid the question I suggested (Rapp, 1977b) that anomalies could be evaluated using (1) on a sphere enclosing all the masses of the earth. Then these anomalies could be downward continued (by collocation, for example) to the terrestrial surface. Numerical tests indicated better agreement (±1 mgal for 5° anomalies) with the terrestrial data when this approach was used than when a direct evaluation on the surface was used. Clearly more study is needed in this area. For this paper all anomaly evaluations have been carried out at the surface of the earth ignoring the convergence problem.

## Potential Coefficients from Terrestrial Anomalies

We should note here that potential coefficients can also be determined from a global estimate of the terrestrial gravity field (Rapp, 1977a). Using the 1° x 1° data set previously discussed we computed a set of 1654 5° equal area anomalies using procedures described in Rapp (1978). These anomalies were used to generate potential coefficients to degree 20 using equation (6) of Rapp (1977a). These coefficients will be compared to satellite derived coefficients in a later section.

## Comparison Quantities

We can compare the satellite and terrestrial data in several ways. The most obvious is the computation of anomalies from potential coefficients using (1) and the comparison with terrestrial data in various size blocks. The comparison can be made by computing the root mean square difference between the satellite and terrestrial data. This difference will be caused by three factors: 1) errors in the terrestrial data; 2) errors in the potential coefficients; and 3) errors caused by the neglect of higher degree terms in the spherical harmonic expansion. Kaula (1966) has described methods to separate these terms.

We can also compare the potential coefficients from the satellite and terrestrial results. This comparison is instructive to consider, by degree, the differences in terms of anomalies and undulations. The mean square undulation difference would be given by:

$$\delta N_\ell^2 = R^2 \sum_{m=0}^{\ell} (\Delta \bar{C}_{\ell m}^2 + \Delta \bar{S}_{\ell m}^2) \qquad (2)$$

where R is a mean earth radius. The anomaly dif-

Table 2. Comparison of Terrestrial Anomalies to Anomalies Implied by GEM 9

| | Block Size | |
|---|---|---|
| $\ell$ (max) | 5° | 1° |
| 12 | ±10 mgals | - |
| 20 | ±11 mgals | ±25 mgals |

ference by degree would be given by:

$$\delta g_\ell^2 = \gamma^2 (\ell - 1)^2 \sum_{m=0}^{\ell} (\Delta \bar{C}_{\ell m}^2 + \Delta \bar{S}_{\ell m}^2) \qquad (3)$$

where $\gamma$ is a mean value of gravity over the earth. The total difference between the two sets of coefficients could be expressed as:

$$\delta N^2 = \sum_{\ell=2}^{\ell_{max}} \delta N_\ell^2 \qquad (4)$$

$$\delta g^2 = \sum_{\ell=2}^{\ell_{max}} \delta g_\ell^2 \qquad (5)$$

Similar equations can be written for the accuracy of the various quanties given the standard deviations of the potential coefficients.

Results Using the GEM 9 Potential Coefficients

To implement the comparisons described in the previous section we will use the GEM 9 (Lerch,et als., 1977) potential coefficients. This coefficient set is complete to degree 20 with some higher order terms. It is based solely on satellite data.

Table 2 shows the root mean square difference between the anomalies computed from the potential coefficients and the terrestrial data. The 5° comparisons were made using 1062 blocks whose standard deviations were less than ±6 mgals. The 1° x 1° comparisons were made using 16579 blocks whose standard deviations were less than ±16 mgals.

In Figures 2 (for geoid undulations) and 3 (for anomalies) information is given, by degree, for the following quantities:

1. Root mean square value implied by the GEM 9 coefficients;
2. Root mean square difference between the GEM 9 coefficients and the coefficients implied by the 5° block terrestrial data;
3. The standard deviations computed from the accuracies of the terrestrial coefficients and the GEM 9 coefficients.

From Figure 1 we see that the GEM 9 undulation has a standard deviation of ±46 cm at degree 12 while the terrestrial standard deviation is ±50 cm and the differences at that degree is ±80 cm. At the higher degrees the differences and the standard deviations approach the magnitude of the undulation at that degree . Similar comments can be made for the anomaly information in Figure 3.

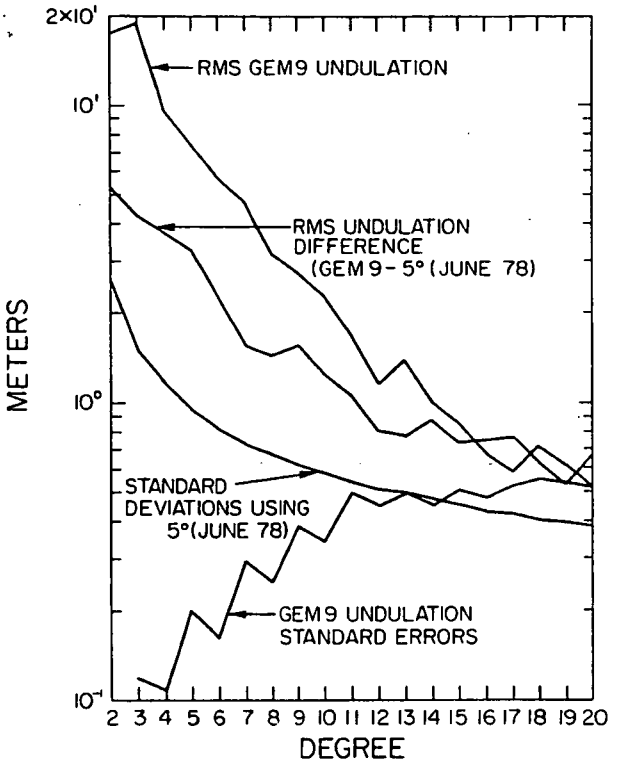



Figure 2. GEM 9 and 5° Terrestrial Implied Potential Coefficient Comparisons in Terms of Geoid Undulations.

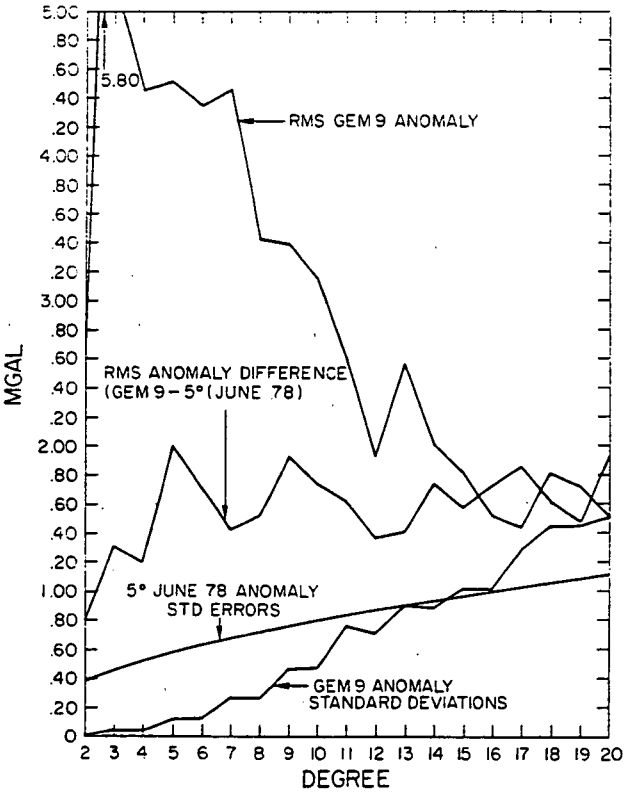



Figure 3. GEM 9 and 5° Terrestrial Implied Potential Coefficients in Terms of Gravity Anomalies.

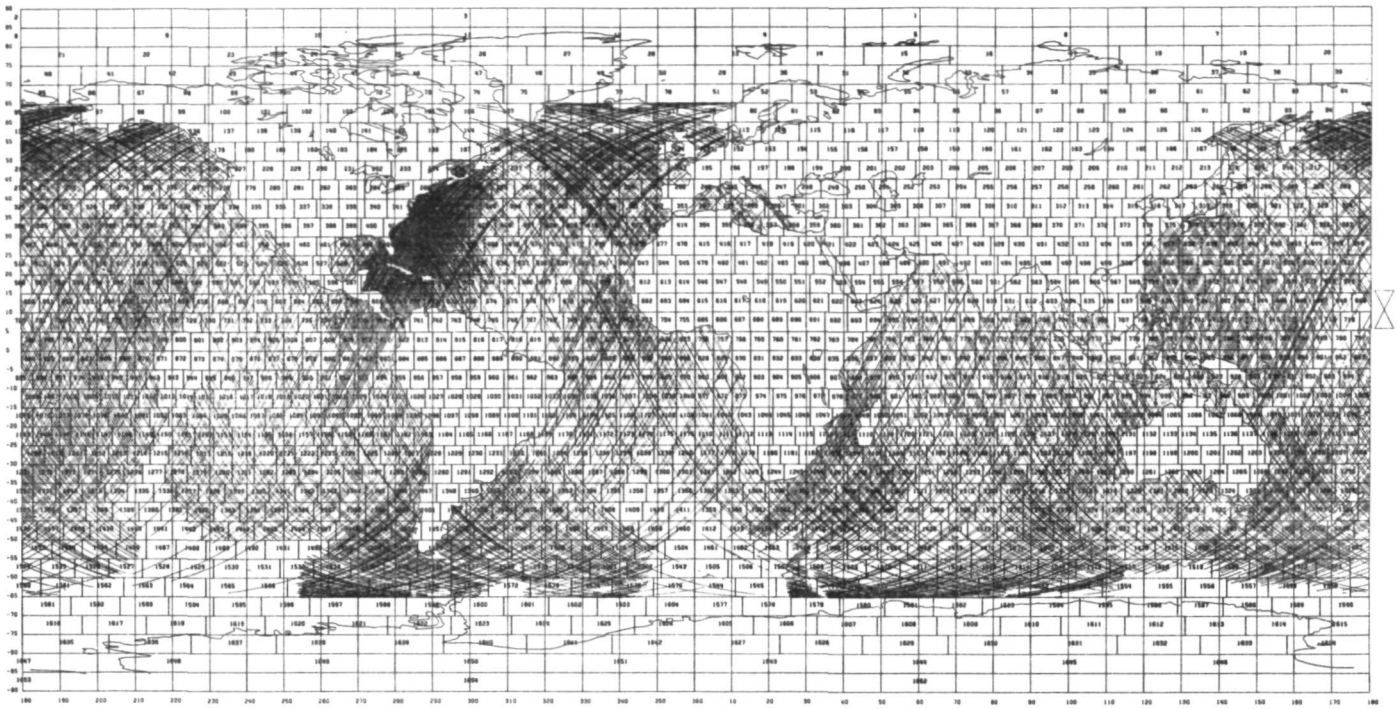


Figure 4. Location of Edited Geos-3 Altimeter Data and 5° Equal Area Blocks.

Table 3 shows comparisons between the two potential coefficient sets when all coefficients between degree 2 and 20 are considered.

## Geos-3 Altimeter Results and Comparisons

The Geos-3 altimeter data has greatly improved our knowledge of gravity (and undulations) at sea in 5° and 1° x 1° mean anomalies. The location of edited Geos-3 data available at The Ohio State University is shown in Figure 4 along with the location of the 5° equal area blocks. Anomalies and undulations have been computed from this data using the procedures described in Rapp (1977d, 1979). From this data we have now computed 29478 1° x 1° anomalies and undulations. Of these there are 27465 1° x 1° anomalies that have standard deviations ≤ ±15 mgals. The location of these anomalies is shown in Figure 5.

The 1° x 1° anomalies derived from the Geos-3 satellite altimeter data have been compared to the terrestrial data in two data sets. The first set lies off the East Coast of the United States in an area where the altimeter data is dense and the terrestrial data is of above average reliablity. The second set comprises the whole altimeter derived anomalies compared to the available terrestrial data subject to the following accuracy limitation (which also applies to the first data set): Comparisons between the two anomalies are only made if the terrestrial anomaly standard deviation is ≤ 25 mgals and the altimeter derived anomaly standard deviation is ≤ 15 mgals. The results are given in Table 4. By including the comparison with the GEM 9 anomalies (using the potential coefficients to $\ell = 20$) we can see the improvement the altimeter results have given over the GEM 9 anomaly field. The accuracy estimates (of about ±8 mgals) for the altimeter derived anomalies apears consistent with the RMS anomaly differences.

In some cases we have found very large discrepancies between the altimeter derived anomalies and the terrestrial data. Specifically we found 17 differences greater than 100 mgals and 203 differences greater than 50 mgals. I give in Table 5, 5 blocks where the differences are large. A number of these cases occur in areas where the anomaly field is changing quite rapidly and only one ship track is available thru a block.

Table 3. Root Mean Square Difference and Commission Errors of the GEM 9 and 5° Terrestrial Implied Potential Coefficients

| | Undulation | Anomaly |
|---|---|---|
| RMS Difference | ±9.1 m | ±7.0 mgals |
| GEM 9 Commission Errors | ±1.7 m | ±3.8 mgals |
| 5° Terrestrial Comm. Errors | ±3.9 m | ±3.5 mgals |

Table 4. Comparison of Altimeter Derived Anomalies and Terrestrial Anomalies in 1° x 1° Blocks

| | Set 1 | Set 2 |
|---|---|---|
| RMS Diff. (GEM 9 - Terr.) | ±32 mgals | ±25 mgals |
| RMS Diff. (Alt. - Terr.) | ±11 mgals | ±15 mgals |
| RMS Terr. Std. Dev. | ±11 mgals | ±15 mgals |
| RMS Alt. Std. Dev. | ± 7 mgals | ± 8 mgals |
| Number of Comparisons | 659 | 16579 |

Table 5. Information Related to Large 1° x 1° Anomaly Differences Between Terrestrial and Altimeter Derived Values

| $\varphi^\circ$ | $\lambda^\circ$ | $\Delta g_{ALT}$ | $\Delta g_{TERR}$ | Difference |
|---|---|---|---|---|
| 7 | 153 | 8± 7 mgals | -258±17 mgals | 266 mgals |
| 47 | 153 | -15±10 mgals | -202±23 mgals | 187 mgals |
| 56 | 162 | -52± 9 mgals | -188±21 mgals | 136 mgals |
| 46 | 171 | -46± 9 mgals | 85±13 mgals | -131 mgals |
| 37 | 213 | -14± 8 mgals | 86±18 mgals | -102 mgals |

It seems clear that there is generally good agreement between the terrestrial and altimeter 1° x 1° anomalies consistent with an accuracy estimate of ±8 mgals for the altimeter data. The large discrepancies discussed above indicate areas where more detailed information is needed on the anomaly field.

Computations have also been made in computing 5° equal area mean anomalies and undulations. The predicted accuracy of the 5° anomalies is on the order of ±3 mgals which is consistent with comparisons made with terrestrial data.

## Satellite to Satellite Tracking Results

A recent data type for the recovery of gravity anomalies is that of satellite to satellite tracking. Such data is currently available only in limited areas and only experimental types of results have been obtained.

One experiment has involved the tracking of the Apollo spacecraft by the ATS-6 satellite. Since the Apollo vehicle was at an altitude of only about 230 km the range rate signal could be strongly perturbed by local anomalies. The analysis of this data and a description of the experiment is found in Vonbun et al. (1977). Using limited data they were able to recover some 5° x 5° mean free-air anomalies to an accuracy of about ±7 mgals based on a comparison with ground truth.

Another experiment involving the ATS-6 satellite has used Geos-3. A description of this experiment and some data analysis may be found in Marsh et al. (1977) The analysis of some of this data for various size mean anomaly blocks has recently been described by Hajela (1978). In this study 5° equal area anomalies were recovered to an accuracy of about ±6 mgals.

At this point we do not have sufficient data to significantly improve our surface gravity field from satellite to satellite tracking data. However test results on anomaly recovery are sufficiently promising that more such data should be sought.

## Conclusions

This paper has been a brief survey of ways in which satellite derived gravity data compares with surface data. These comparisons have been performed using block means (such as 5° equal area and 1° x 1°) and in terms of potential coefficients. We found that the differences between the terrestrial data and GEM 9 potential coefficients was ±11 mgals for 5° blocks and ±25 mgals for 1° x 1° blocks. Much of this difference is caused by the fact that the GEM 9 set is complete to degree 20 only.

Comparisons with potential coefficients derived from the terrestrial 5° data, showed poor agreement at the lower degree. Overall there was a ±9.1 m undulation difference and a ±7.0 mgal anomaly difference. The standard deviation, by degree of the undulation or anomaly difference, was found to approach the actual magnitude of the quantity near degree 18. The standard deviation at a given degree, for the undulation,

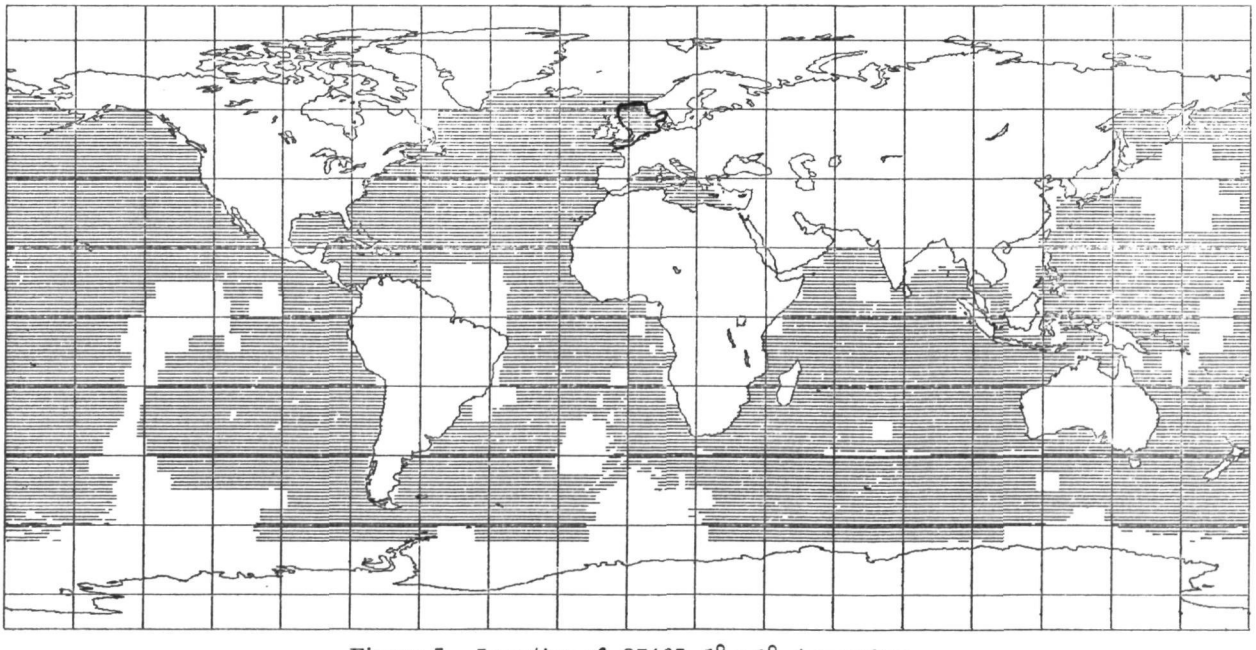


Figure 5. Location of 27465 1° x 1° Anomalies Computed from Geos-3 Altimeter Data.

was smaller for the GEM 9 coefficients than for the terrestrial derived coefficients up to degree 13. Figures 2 and 3 summarized the differences found.

Comparisons of the Geos-3 altimeter derived 1° anomalies and the surface data indicated the predicted standard deviations of about ±7 mgals were reasonable. However, some large discrepancies exist between the altimeter derived and terrestrial anomalies ranging up to 266 mgals. The use of the altimeter anomalies may be a way to detect bad 1° x 1° anomaly estimates.

Finally satellite to satellite tracking results were briefly discussed noting the achievement of encouraging results with additional data needed.

Acknowledgments. The work described in this paper was supported in part by funds supplied thru NASA grants NGR 36-008-161, and NSG 5275 which are administered thru the Goddard Space Flight Center.

## References

Arnold, D., The Spherical - Harmonics Expansion of the Gravitational Potential of the Earth in the External Space and Its Convergency, Gerlands Beitr. Geophysik, Leipzig, 87, 2, 81-90, 1978.

Hajela, D.P., Improved Procedures for the Recovery of 5° Mean Gravity Anomalies from ATS-6/Geos-3 Satellite to Satellite Range-Rate Observations using Least Squares Collocation, Department of Geodetic Science Report No. 276, The Ohio State University, Columbus, 1978.

Kaula, W., Tests and Combinations of Satellite Determinations of the Gravity Field with Gravimetry, Journal of Geophysical Research, 71, 5303-5314, 1966.

Lambeck, K., Comparison of Surface Gravity Data with Satellite Data, Bulletin Geodesique, No. 100, 203-219, 1971.

Lerch, F., S. Brownd, S. Klosko, Gravity Model Improvement Using Geos-3 (GEM 9 and GEM 10), Abs., EOS, Vol. 58, No. 6, 1977.

Marsh, J., B. Marsh, T. Conrad, W. Wells and R. Williamson, 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 T.M. 79553, Goddard Space Flight Center, Greenbelt, Maryland, December, 1977.

Moritz, H., On the Convergence of the Spherical Harmonic Expansion for the Geopotential at the Earth's Surface, Bullettino Di Geodesia E Scienze Affini, Nos. 2, 3, 4, 1978.

Rapp, R.H., Comment on: "Comparison of Surface Gravity Data with Satellite Data" by K. Lambeck, Bulletin Geodesique, No. 105, 343-347, 1972.

Rapp, R.H., Comparison of the Potential Coefficient Models of the Standard Earth (II and III) and the GEM 5 and GEM 6, Bulletin Geodesique, No. 117, 279-287, 1975.

Rapp, R.H., Determinations of Potential Coefficients to Degree 52 from 5° Mean Gravity Anomalies, Bulletin Geodesique, Vol. 51, No. 4,301-323, 1977a.

Rapp, R.H., Geos-3 Data Processing for the Recovery of Geoid Undulations and Gravity Anomalies, Journal of Geophysical Research, 1979 (in press).

Rapp, R.H., Mean Gravity Anomalies and Sea Surface Heights Derived from Geos-3 Altimeter Data, Department of Geodetic Science Report No. 268, The Ohio State University, Columbus, December 1977d.

Rapp, R.H., Results of the Application of Least-Squares Collocation to Selected Geodetic Problems, in Approximation Methods in Geodesy, ed. by Moritz and Sunkel, Herbert Wichmann Verlag Karlsruhe, 1978.

Rapp, R.H., The Use of Gravity Anomalies on a Bounding Sphere to Improve Potential Coefficient Determination, Department of Geodetic Science Report No. 254, The Ohio State University, Columbus, 1977b.

Sjöberg, L., On the Errors of Spherical Harmonic Developments of Gravity at the Surface of the Earth, Department of Geodetic Science Report No. 257, The Ohio State University, Columbus, 1977.

Vonbun, F., W. Kahn, W. Wells, T. Conrad, Gravity Anomalies Determined from Tracking the Apollo-Soyuz, NASA Technical Memorandum 78031, December 1977.

# Accuracy of Vertical Deflection Determination by Present-Day Inertial Instrumentation

K. P. Schwarz
Department of Surveying Engineering, University of New Brunswick
Fredericton, N.B., Canada E3B 5A3

Abstract. The conventional procedures to determine deflections of the vertical in mountainous terrain require time-consuming astronomical or gravimetric methods and their application is therefore restricted to a small number of stations. The interpolation of vertical deflections between such stations can be performed by the Inertial Surveying System currently used for position determination. The principle of such a procedure is outlined and the existing implementations are discussed.

An analysis of results obtained in the Canadian Rocky Mountains indicates that the observation of deflection differences along the same line can be repeated with a precision of about 0".5 but that there are systematic discrepancies between the forward and the backward running of the same line. A comparison with the available astronomically determined deflections also shows systematic differences of 2" to 3". These errors are most likely due to the 'overshooting' of the Kalman procedure at gradient changes. It appears that the software can be altered in such a way that deflection differences between stations, not more than half an hour of travel time apart, can be determined by the inertial system with an accuracy of better than ± 1".

## 1. Introduction

Two methods have conventionally been used to determine the deflections of the vertical $\xi$ and $\eta$ which define the difference in direction between the ellipsoidal normal and the actual gravity vector. The first approach uses integral formulas to determine the deflection components from gravity anomalies $\Delta g$. In Vening-Meinesz' integral

$$\begin{Bmatrix}\xi\\ \eta\end{Bmatrix} = \frac{1}{4\pi G}\iint_{\sigma} \Delta g \frac{dS}{d\psi}\begin{Bmatrix}\cos\alpha\\ \sin\alpha\end{Bmatrix} d\sigma \qquad (1.1)$$

$\xi$ and $\eta$ are in principal computed at the surface of the geoid. Here G denotes a mean value of gravity for the whole earth, $S(\psi)$ is Stokes' function, $\psi$ is the spherical distance, $\alpha$ is the azimuth, and $\sigma$ indicates integration over the earth. In Molodenski's approach $\xi$ and $\eta$ are determined at the earth's surface. The second method uses astronomically determined latitude and longitude $(\Phi, \Lambda)$ and geodetic latitude and longitude $(\phi,\lambda)$ to obtain deflection components at the observation point by the simple relations

$$\begin{aligned}\xi &= \Phi - \phi\\ \eta &= (\Lambda - \lambda)\cos\phi\end{aligned} \qquad (1.2)$$

Usually, the deflection coverage of larger areas is rather sparse because of the time-consuming data acquisition procedures. This is especially true for mountainous terrain where a dense coverage would be required to adequately represent the slope changes of the equipotential surfaces. The amount of work required for this is prohibitive in most cases. Methods to interpolate deflections between stations where the gravity vectors are known are therefore of great interest. Two ways to approach this problem have evolved in recent years. They could be called computational and observational interpolation. In the first approach all information about the anomalous gravity field in a certain area is combined to predict deflection values at the specific point. Methods differ as to the way in which the different data groups are combined and represented. But all have in common that they employ heterogeneous data and thus avoid the limitations which are often encountered when using one type of observations only. The actual resolution of these methods depends to a large extent upon the amount, the accuracy, and the distribution of the data. With a scarce coverage as for instance in mountainous areas it is impossible to recover any details. While this approach is characterized by the optimal use of the available information, the second approach relies on an instrument which is capable of measuring changes of the direction of the gravity vector with reference to an initial point. In this case a detailed mapping of the deflection changes along the path of the instrument is possible. Thus, a relative geoid can be computed which is then oriented by the absolute deflection values obtained by other means. Inertial systemsare capable of performing such an observational interpolation and from their error characteristics an application in mountainous terrain seems to be especially promising.

The discussion will concentrate on deflection interpolation with such instruments. This limited application should not obstruct the view for one of the main advantages of these systems: the capability to obtain position and gravity field information at the same time. It seems that because of historic subdivisions in geodesy the full potential of inertial systems is not yet utilized.

The following sections will be somewhat biased towards the Litton 'Inertial Surveying System'. This does not indicate a preference but has been dictated by the fact that the only data available to the author had been taken with this system.

## 2. Movement of an Inertial System in a Local Gravitational Field

An inertial measuring unit consists basically of three mutually orthogonal accelerometers and of an assembly of gyroscopes establishing a reference frame with known orientation to the accelerometer triad. Usually, the accelerometers will

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280. The Ohio State Univ., Columbus, Ohio 43210.

be aligned along the output axes of the gyros. The output of the accelerometer triad are three components of specific force

$$f_i = \ddot{r}_i + g_i \qquad (2.1)$$

where $\ddot{r}_i$ are the inertially referenced accelerations expressed as the second derivatives of a radius vector with respect to time and $g_i$ are the components of the gravitational acceleration at the system location due to all bodies in the universe. For surveys on the surface of the earth the origin of the inertial system is usually translated to the mass center of the earth thereby making the variations of the gravitational effect of all extra-terrestrial bodies smaller than $2.10^{-7}$. Thus, for relative accuracies of about $2.10^{-7}$ only the effect of the earth's gravitational field has to be considered. Since the measuring accuracy of available inertial systems is of the order of $10^{-5}$ an earth centered origin will be assumed hereafter.

The accelerometer triad can be related to the inertial triad by connecting the two radius vectors by a rotation matrix C

$$r^I = Cr^A \qquad (2.2)$$

where the superscripts I and A refer to the inertial and to the accelerometer frame respectively. Differentiating twice with respect to time we obtain

$$\ddot{r}^I = C\ddot{r}^A + 2C\dot{r}^A + C\ddot{r}^A \qquad (2.3)$$

We now can distinguish three special cases. If C is independent of time,only the first term on the right-hand side remains and equation (2.3) expresses the rotation between two inertial frames. Such a system can be instrumented by mounting the accelerometers on a gimballed platform and keeping its orientation fixed in inertial space. These systems are called space stabilized. Honeywell's Geo-spin is a system developed along these lines for geodetic purposes. In the second case the only time dependency allowed in C will be the rotation of the earth. Such a system will again make use of a gimballed platform which now will be constantly torqued in such a way that it stays orthogonal to a reference ellipsoid. These systems are called local-level and can directly be related to the geodetic ($\phi$, $\lambda$, h) - coordinates; Litton's 'Inertial Surveying System' and Ferranti's system work with this concept. If finally an arbitrary time dependency is allowed in C, equation (2.3) represents a strapdown system. In this case the inertial instruments are mounted along axes attached to the vehicle and the orientation changes arbitrarily with respect to inertial space. So far, systems of this kind have not been developed for geodetic applications.

The principle of inertial geodesy can best be seen from equation (2.1). If the gravity vector $g_i$ is known we can obtain position by integrating twice

$$\ddot{r}_i = f_i - g_i \qquad (2.1a)$$

Usually, only an approximation to $g_i$ is available, either in form of the normal gravity vector $\gamma_i$ or in form of a higher order approximation from one of the satellite solutions. In that case the differences between the reference field and the actual field can be determined by measurement using the normal case as a first approximation. Thus, position and gravity field determination become intertwined in an iterative procedure. This concept will be used in the sequel for a local-level system.

Another approach which shows clearly the interdependence of geometry and physics starts from the holonomity problem. The transformation of locally imperfect differentials into locally perfect differentials for frames used in geodesy has been discussed in detail by Grafarend (1975).

The specific force equation for a local-level system is obtained from equation (2.1) and (2.3)

$$f_i = C\ddot{r}_i^A + 2C\dot{r}_i^A + C\ddot{r}_i^A + g_i \qquad (2.4)$$

The first three members on the right-hand side are usually expressed in terms of vehicle velocity, earth rotation rate, and ellipsoidal radii (see e.g. Britting, 1971). The important point is that an ellipsoidal surface is used for all computations and that small correction terms are applied to account for the deviations between model and reality. This is done by splitting the gravity vector g into a normal and an anomalous part

$$g_i = \gamma_i + \delta g_i \qquad (2.5)$$

or

$$\text{grad } W = \text{grad } U + \text{grad } T \qquad (2.6)$$

where W is the gravity potential, U is the normal ellipsoidal gravity potential, and T is the anomalous gravitational potential. Similarly, $g_i$ is the gravity vector, $\gamma_i$ the normal gravity vector, and $\delta g_i$ the gravity disturbance vector. The vectors are now expressed in spherical coordinates with geocentric latitude $\bar{\phi}$, longitude $\lambda$, and radius vector r. Using the usual series expansion of the normal potential U (see e.g. Heiskanen and Moritz (1967), p. 230), we obtain

$$U = \frac{kM}{r}\{1 - \sum_{n=1}^{\infty} J_{2n}\left(\frac{a}{r}\right)^{2n} P_{2n}(\sin\bar{\phi})\} + \psi(r,\bar{\phi}) \qquad (2.7)$$

where

$$J_{2n} = (-1)^{n+1}\frac{3e^{2n}}{(2n+1)(2n+3)}\left(1-n+5n\frac{C-A}{ME^2}\right),$$

kM is the gravitational constant times the mass of the earth, $J_{2n}$ are the even harmonic coefficients of the expansion, a is the semi-major axis of the ellipsoid, $P_{2n}(\sin\phi)$ are Legendre polynomials, E is the linear eccentricity $E = (a^2 - b^2)^{1/2}$, e is the first eccentricity $e=E/a$, C and A are the earth's moments of inertia around its axis of rotation and around an axis in the equatorial plane respectively, and $\psi(r, \phi)$ is the centrifugal potential. We then have

$$\gamma = \text{grad } U = \frac{1}{r}\frac{\partial U}{\partial \bar{\phi}} i_{\bar{\phi}} + \frac{1}{r \cos\bar{\phi}}\frac{\partial U}{\partial \lambda} i_{\lambda} + \frac{\partial U}{\partial r} i_r \quad (2.8)$$

where

$$\frac{\partial U}{\partial \lambda} = 0$$

because of rotational symmetry. The other two partial derivatives are

$$\frac{1}{r}\frac{\partial U}{\partial \bar{\phi}} = \gamma_{\bar{\phi}} = \frac{kM}{r^2}\sum_{n=1}^{\infty} J_{2n}\left(\frac{a}{r}\right)^{2n} P'_{2n}(\sin\bar{\phi}) + \frac{\partial \psi}{\partial \bar{\phi}} \quad (2.9)$$

where

$$P'_{2n}(\sin\bar{\phi}) = -\cos\bar{\phi}\sum_{k=1}^{n}(4n - 4k + 3) \cdot P_{2n-2k+1}(\sin\bar{\phi})$$

and

$$\frac{\partial U}{\partial r} = \gamma_r = -\frac{kM}{r^2}\{1 - \sum_{n=1}^{\infty}(1 + 2n) J_{2n}\left(\frac{a}{r}\right)^{2n} \cdot P_{2n}(\sin\bar{\phi}) + \frac{\partial \psi}{\partial r} \quad (2.10)$$

Since the series (2.9) and (2.10) converge very fast approximations of the form

$$\gamma_{\bar{\phi}} = \gamma_e (a_1 + a_2 \sin^2\bar{\phi}) \sin\bar{\phi}\cos\bar{\phi} \quad (2.11)$$

and

$$\gamma_r = \gamma_e (1 + b_1 \sin^2\bar{\phi} + b_2 \sin^4\bar{\phi}) \quad (2.12)$$

can be used where $\gamma_e$ refers to normal gravity at the equator. The coefficients $a_1$, $a_2$, $b_1$, $b_2$ depend on the reference system chosen. The relative accuracy of these formulas is about $10^{-8}$. With the same accuracy normal gravity along the ellipsoidal normal $\gamma_n$ can be obtained by using

$$\gamma_n = \gamma_{\bar{\phi}} \sin\varepsilon - \gamma_r \cos\varepsilon \quad (2.13)$$

where $\varepsilon = 0.50\, e^2 \sin 2\phi$.

The absence of odd degree terms in formula (2.7) is necessary in order to maintain the same ellipsoidal reference surface for all computations. An inclusion of the $J_3$ - term as for instance in (Britting, 1971) is inconsistent with the use of the ellipsoid as a computational surface. If higher order approximations are used for the gravity field, formulas for the appropriate surfaces must be developed.

Conceptually, the anomalous part of the gravity field can be treated in exactly the same way as the normal part. Using the expansion of the anomalous potential T into spherical harmonics

$$T = \frac{kM}{r}\{1 - \sum_{n=2}^{\infty}\left(\frac{a}{r}\right)^n J'_n P_n(\sin\phi) - \sum_{n=2}^{\infty}\sum_{m=1}^{n}\left(\frac{a}{r}\right)^n (J_{nm}\cos m\lambda + K_{nm}\sin m\lambda) \cdot P_{nm}(\sin\phi)\} \quad (2.14)$$

where $J'_n$ are the zonal coefficients minus the normal part and $J_{nm}$, $K_{nm}$ are the tesseral harmonic coefficients, we can again differentiate with respect to $\phi$, $\lambda$, and r. There are, however, two difficulties with this approach. First, only truncated series (2.14) are available from satellite observations which will not give the required local details. Second, the evaluation of such series will be too laborious for real time computations. For the following discussion we will therefore assume that only the normal part of the gravity field as represented by equations (2.7) to (2.13) is known.

Equation (2.5) shows that the deviations from the normal model are given by the gravity disturbance vector $\delta g_i$ which has the components

$$\begin{bmatrix} \delta g_{\phi} \\ \delta g_{\lambda} \\ \delta g_r \end{bmatrix} = \begin{bmatrix} \frac{1}{r}\frac{\partial T}{\partial \phi} \\ \frac{1}{r\cos\phi}\frac{\partial T}{\partial \lambda} \\ \frac{\partial T}{\partial r} \end{bmatrix} \quad (2.15)$$

Using spherical approximations the right-hand side can be expressed in terms of $\xi$, $\eta$, and $\Delta g$

$$\begin{bmatrix} \delta g_{\phi} \\ \delta g_{\lambda} \\ \delta g_r \end{bmatrix} = \begin{bmatrix} -\gamma_o \xi \\ -\gamma_o \eta \\ \Delta g + \frac{2G}{R} N \end{bmatrix} \quad (2.16)$$

where $\gamma_o$ denotes normal gravity at the ellipsoid, G and R are mean values of gravity and earth radius respectively, and N is the geoidal undulation at the point.

Changes of these quantities from one station to the next can be determined by using two properties of an inertial system: the capability to align to the local vertical and the faculty to keep an orientation fixed in space. The first property allows determination of the direction of the local gravity vector each time the system stops. The second property makes it possible to transport an orthogonal frame established at an initial point to other points on the earth's surface and to use it as a reference. Thus, the actual changes of the gravity vector can be compared to the changes of the normal gravity vector i.e. changes of the gravity disturbance vector (2.16) can be determined. If the gravity vector is known at the initial point it can be determined in all subsequent points. Strictly speaking, an iterative process would be required, expressing the fact that position and gravity field determination cannot be separated. In practice, the iterative corrections will often be negligible because of the small distances between stations.

It should be well understood that the changes of the gravity disturbance vector are not continuously recorded but can only be determined at discrete points where the vehicle stops. For continuous recording gradiometers must be added to the inertial system. However, it is of interest for the following discussion to relate the gravity disturbance vector to changes in the gravity field and in vehicle motion. A derivation of the relevant formulas is given in (Moritz 1975). Using the notations

$$T_i = \frac{\partial T}{\partial X_i} \qquad T_{ij} = \frac{\partial^2 T}{\partial X_i \, \partial X_j}$$

we can write

$$T_i(t) = T_i^0 + \int_{t_o}^{t} T_{ij}(s) \{u_i^o - \int_{t_o}^{s} \bar{f}_j(r)\, dr + \int_{t_o}^{s} T_j(r)\, dr\}\, ds \,, \tag{2.17}$$

where

$$\bar{f}_j = f_j - \gamma_j \,,$$

u denotes the velocity, and the superscript zero indicates an initial value. The last term on the right of formula (2.17) describes the interaction between gravitation and inertia. Since its effect will be very small in the applications considered here, it will be neglected hereafter. This approximation does not affect the following argument. Using the relations (2.14) and (2.15) results in

$$\begin{aligned}
\xi &\approx \xi_o + \frac{1}{\gamma_o r} \int_{t_o}^{t} (T_{\phi\phi} v_\phi + T_{\phi\lambda} v_\lambda + T_{\phi r} v_r)\, ds \\
\eta &\approx \eta_o + \frac{1}{\gamma_o r \cos\phi} \int_{t_o}^{t} (T_{\lambda\phi} v_\phi + T_{\lambda\lambda} v_\lambda + T_{\lambda r} v_r)\, ds \\
\Delta g &= \Delta g_o - \int_{t_o}^{t} (T_{r\phi} v_\phi + T_{r\lambda} v_\lambda + T_{rr} v_r)\, ds + \\
&\quad + \frac{2G}{R}(N_o - N) \,,
\end{aligned} \tag{2.18}$$

where

$$v_i = \int_{t_o}^{t} \bar{f}_i(s)\, ds$$

because of $u_i^o = 0$. For local applications the term with $(N^i - N_o)$ can be neglected. Except for small corrections the $v_i$ represent the velocity components and with

$$v_r << (v_\lambda, v_\phi)$$

in many cases the deflections of the vertical can be expressed by the approximation

$$\begin{aligned}
\xi &= \xi_o + \frac{1}{\gamma_o r} \int_{t_o}^{t} (T_{\phi\phi} v_\phi + T_{\phi\lambda} v_\lambda)\, ds \\
\eta &= \eta_o + \frac{1}{\gamma_o r \cos\phi} \int_{t_o}^{t} (T_{\lambda\phi} v_\phi + T_{\lambda\lambda} v_\lambda)\, ds \,.
\end{aligned} \tag{2.19}$$

This formula shows that changes in $\xi$ and $\eta$ are dependent on the ratio $v_\phi / v_\lambda$, i.e. on the instrument heading. This is especially apparent for an L-shaped traverse which first runs east - west and then south - north. We obtain

$$\Delta\xi_{E-W} = \frac{1}{\gamma_o r} \int_{t_o}^{t} T_{\phi\lambda} v_\lambda\, ds$$

$$\Delta\xi_{S-N} = \frac{1}{\gamma_o r} \int_{t_o}^{t} T_{\phi\phi} v_\phi\, ds$$

and

$$\Delta\eta_{E-W} = \frac{1}{\gamma_o r \cos\phi} \int_{t_o}^{t} T_{\lambda\lambda} v_\lambda\, ds$$

$$\Delta\eta_{S-N} = \frac{1}{\gamma_o r \cos\phi} \int_{t_o}^{t} T_{\lambda\phi} v_\phi\, ds.$$

## 3. Implementation of the Measuring Principle in the Inertial Surveying System

Fig. 3.1 ilustrates the principle of determining changes in the gravity disturbance vector by an inertial measuring unit. At an initial point $P_1$ the system is aligned to the local gravity vector $g_1$ by a levelling procedure which drives the two 'horizontal' velocity outputs to zero and by establishing astronomical north via gryocompassing. Basically, a local astronomical $(\Phi,\Lambda)$ - system is established. The small angle $\theta_i$ between $g_i$ and $\gamma_i$ is called the total deflection of the vertical. The initial frame is transported to $P_2$ making corrections for the rotation rate of the earth by continuously torquing the platform. Similarly, compensation of changes of the normal gravity vector are included in the specific force equation. At $P_2$ the changes of the gravity disturbance vector will cause a small misalignment of the platform with respect to the local vertical and the resulting velocity readings in the 'horizontal' accelerometers can be resolved into the components $\Delta\xi$ and $\Delta\eta$.

At this point two different procedures are possible. The first one is used in the 'Rapid Geodetic Survey System' (RGSS) and is illustrated in fig. 3.2. In this case the velocity readings are recorded but the initial frame is left unchanged, i.e. only the above mentioned torques and normal gravity corrections are applied. The reference surface for the computations is then an ellipsoid which is slightly tilted against the global reference ellipsoid because the alignment has been made with respect to the local vertical. If $\phi$, $\lambda$, h, $\xi$, and $\eta$ are known in the initial point this tilt can theoretically be removed. Since this system measures differences in ellip-

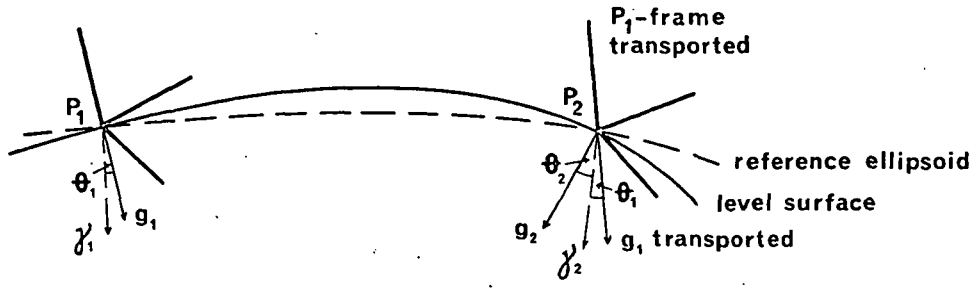



Fig. 3.1 Principle of measuring deflection changes

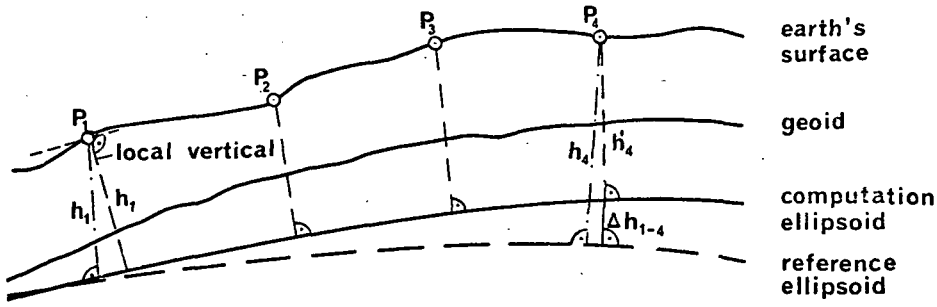



Fig. 3.2 Principle of 'Rapid Geodetic Survey System'

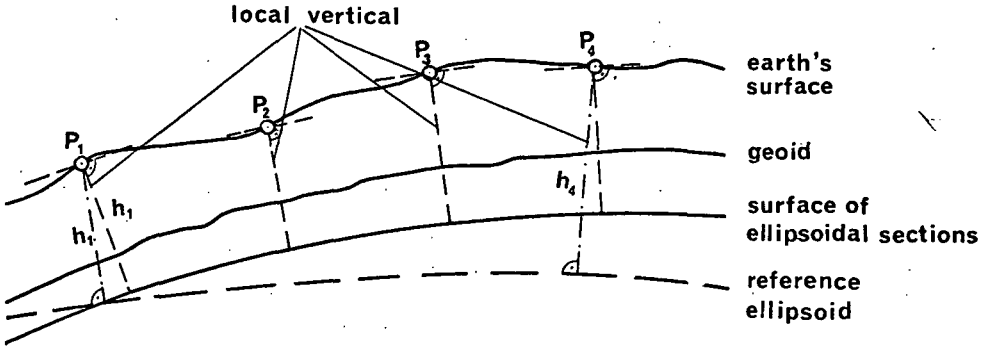



Fig. 3.3 Principle of 'Inertial Positioning System'

soidal height h it is also possible to use the deviation between measured and known height difference $\Delta h_{1-4}$ to retilt the ellipsoid. By using the deflection information provided at the zero updates the changes of the geoidal undulation between $P_1$ and $P_4$ can be computed.

The second procedure is used in the 'Inertial Positioning System' (IPS) and is illustrated in fig. 3.3. In this case the frame is realigned to the local vertical at each zero update. This means that the geoid is approximated by a sequence of ellipsoidal sections. The height differences determined from this surface will approximate levelled height differences. The interpretation of the computed latitude and longitude differences is somewhat problematic.

Theoretically, the anholonomity problem creeps in at this point. Practically, a piecewise mapping onto the ellipsoid will give results which are acceptable within the limits of present measuring accuracy. With improved systems this procedure should, however, be avoided.

So far, measuring errors have not been considered. They will disturb the simple relations discussed above. Certain error sources produce accelerations which are very similar to those generated by changes in the gravity field. A separation can be achieved by an adequate measuring process. Changes in the gravity field are position dependent, at least at the level of accuracy considered here, while most instrumental errors are time dependent. A well designed survey can help to separate the two disturbances. Reoccupation of stations after certain time intervals and checks at stations with a known gravity disturbance vector will provide a control of the instrumental errors.

The present 'Inertial Surveying System' controls the different error sources by a hierarchy of biases. They are either added to the specific force equations or used to modify the torquing commands. In this concept the gravity disturbances are considered as one of several sources of noise. Optimal filtering techniques are used to eliminate this noise. Thus, deflection changes are absorbed

into bias changes. The separation from instrumental errors, especially gyro drift, is done under the assumption that the correlation functions are known. Two sets of biases are important for deflection determination: the alignment biases and the Kalman biases.

The first group, consisting of three gyro biases and one accelerometer bias, is determined during the levelling and gyrocompassing procedure at the initial point and remains constant for one mission. It fixes the tilt of the computation surface against the global reference ellipsoid and also introduces a scale factor in the height computation. Since a number of different effects are lumped into the gyro biases the resulting tilt cannot be considered as representing the gravity disturbances at the initial point. This will have a second-order effect on the computation of deflection differences but will in general be negligible for local applications.

The second group of biases, the Kalman biases, are determined at each zero update. In this case, the value of each bias $b_i$ is recomputed using the new data x (velocities) according to a priori knowledge contained in the gain matrix K. The formula

$$b_{i+1} = b_i + K_i \, (x - Bb_i) \qquad (3.1)$$

expresses this relation. The matrix B gives a functional relationship between x and b. Two sets of Kalman biases are important for the determination of deflection changes. The sum of tilt corrections for each axis and the accelerometer biases. The procedure illustrated by fig. 3.3 combines tilt correction and accelerometer bias to obtain deflection components. No tilt corrections are made in the procedure described by fig. 3.2 and the deflections can be derived from the accelerometer biases only.

It will be shown in the next section that the use of Kalman estimation, well suited for error control, does not always give reliable results for the determination of deflection components.

## 4. Analysis of Results

The data used in this analysis have been provided by the Geodetic Survey of Canada. They were taken during a campaign in the Okanagan Valley of the Canadian Rocky Mountains in 1975. Fig. 4.1 shows the survey line which is a paved road between Curve and Bottom and unpaved between Bottom and Kobau. All stations marked by a triangle have astronomically determined deflections of the vertical. The height profile is shown in fig. 4.2. It should be noted that a rather extreme terrain has been selected with a number of sharp curves in the second part of the line and a height difference of about 1600 m.

The surveys were made by car during a period of about three weeks in May and June and usually a forward and a backward running were made with one alignment. The method used is that described by Fig. 3.3. Not all legs of the traverse were observed with the same frequency; the minimum number of double runs was 5 the maximum number 10, with an average of about 8 runs.

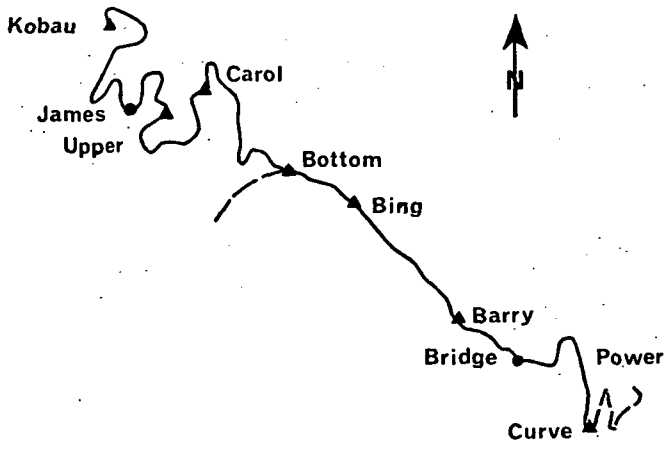



Fig. 4.1 Map of the survey route

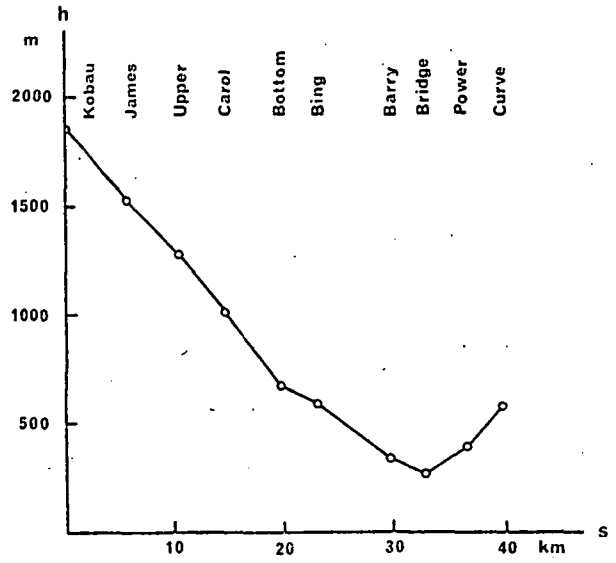



Fig. 4.2 Topographic profile of the survey route

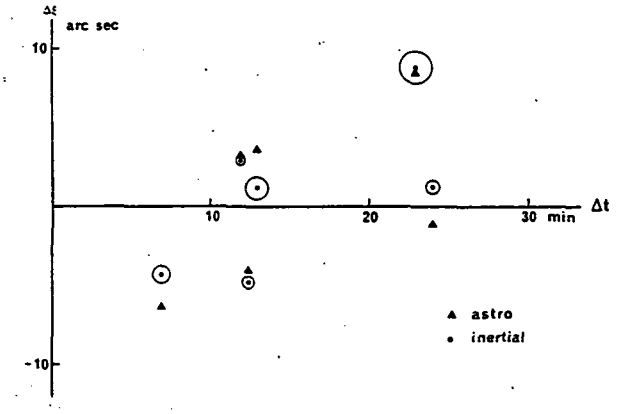



Fig. 4.3 Accuracy of system derived $\Delta\xi$ - values

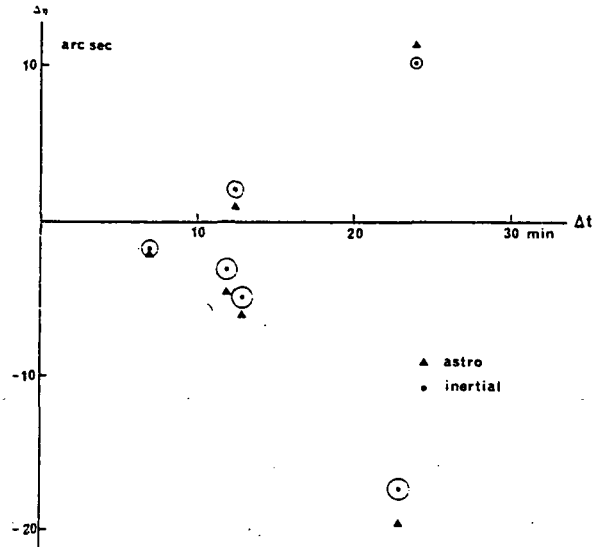



Fig. 4.4 Accuracy of system derived Δη - values

Since only the differences Δξ and Δη between stations can be determined, the most obvious approach is to compute these differences and to estimate standard deviations for each difference. Since sample sizes were rather small in some cases, the hypothesis was tested if all variances could be considered as representing the same population. The standard deviations of Δξ and Δη were almost identical with $\sigma_{\Delta\xi} = \pm$ 0".54 and $\sigma_{\Delta\eta} = \pm$ 0".56, so that the standard error for a deflection difference could be estimated from a sample size of about 100. The result was

$$\sigma_I = \pm 0".55$$

for the mean of a forward and a backward running. The standard deviations of the individual differences were compared to $\sigma_I$ and all except one passed the F-test on a 5% - level. The one rejected standard error was too small. These results show that deflection differences can be determined with a high precision. This means that the repeatability of the results is very good.

As to the accuracy fig. 4.3 and 4.4 should be consulted. They show the deflection differences as functions of the travel time Δt between stations. Mean values of the system determined differences are marked by a dot, while the differences of the astronomical deflections are represented by a triangle. The individual standard errors (1σ) are indicated by a circle. No standard deviations were available for the astronomically determined differences but judging from the observation method they should in general be below 0".5. We will therefore use

$$\sigma_A = \pm 0".5$$

as standard deviation of the astronomically determined differences.

The figures show very clearly that the deviations between astronomically determined and system derived differences is much larger than could be expected from the standard deviations. The standard error $\sigma_{A-I}$ (astronomical - inertial) is

$$\sigma_{A-I} = \pm 1".73 .$$

Considering the size of the deflection differences it can be concluded that the inertial system recovers deflection changes with a good accuracy. Considering the size of $\sigma_I$ and $\sigma_A$ it must be concluded that there are systematic differences between the two data groups. Although $\sigma_I$ and $\sigma_A$ are almost equal there is some reason to believe that the differences derived from the inertial system are systematically wrong. One indication is given by the large differences between forward and backward runnings.

If we compute the mean of the differences between stations using only forward runs in one case and only backward runs in the other we obtain the results summarized in table 4.1.

The standard deviations $\sigma_F$ and $\sigma_B$ belong to the means of the forward and the backward runs respectively. The standard deviations $\sigma_{F-B}$ characterize the deviations between the individual forward and backward runs. The sample size is about 50 in each case. Using an F-test at a 5% - level it must be concluded that there are systematic deviations between the forward and backward runnings.

Part of these differences can be explained by the 'slowness' of the Kalman estimation to adapt to a new situation. If we look at formula (3.1)

$$b_{i+1} = b_i + K_i\ (x - Bb_i) \qquad (3.1)$$

the new estimate $b_{i+1}$ is composed of the old estimate $b_i$ and a portion representing the influence of the new data. This influence is weighted by the gain matrix K which is dependent on the a priori correlation function and previous estimates. Thus, the old estimate $b_i$ may to a large extent determine the value of $b_{i+1}$, i.e. the estimation is somewhat slow to follow changes in the value of b. The situation is illustrated in fig. 4.5 The full line represents a deflection profile, the dashed line its estimation by the Kalman procedure when coming from the left side. There is a kind of 'overshooting' due to the influence of the old estimate which makes the difference $P_2P_3$ too small. When coming from the right side $P_2P_3$ will be determined correctly but in this case $P_1P_2$ will be wrong. This explains the differences between forward and backward runnings. The following example will demonstrate that this effect is

| Deflection Component | $\sigma_{F-B}$ | $\sigma_F$ | $\sigma_B$ |
|---|---|---|---|
| Δξ | ± 1.43 | ± 0.64 | ± 0.96 |
| Δη | ± 3.86 | ± 1.12 | ± 1.03 |

Table 4.1 Comparison of standard deviations

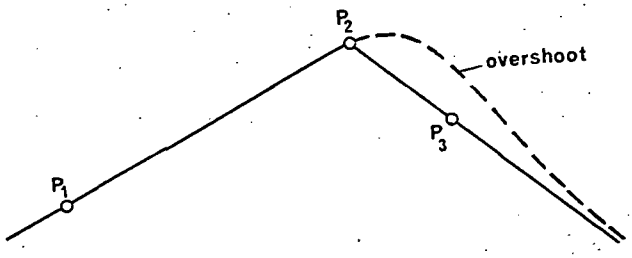



Fig. 4.5 'Overshooting' of Kalman filter

present in the data and leads to observable systematical errors.

Fig. 4.6 shows the $\xi$-profile between the stations Barry and Carol and outlines the Kalman estimation of the difference Bing-Bottom. Coming from Barry 'overshooting' at Bing will produce a difference which is too small. Coming from Carol 'undershooting' will also give too small a difference. The actual values determined as means of 10 measurements are

$$\Delta\xi_{BI-BO} = -4\overset{''}{.}08 \pm 0\overset{''}{.}07$$

$$\Delta\xi_{BO-BI} = -4.42 \pm 0\overset{''}{.}32$$

The 'correct' difference from astronomical observations is $\Delta\xi = -6\overset{''}{.}53$. It should be noted that the line between Bing and Bottom is rather straight and has a length of only 6.6 km. Thus, there is no other obvious explanation for errors of this size.

The interpretation of the results from the curved part of the line is more difficult. As has been shown in section 2 changes in the gradient of $\xi$ and $\eta$ are likely to occur with each change in platform heading. Thus, for lines having several sharp curves between stations the unwanted effects of the Kalman procedure may either accumulate or cancel. This is exactly the pattern which evolves for the winding part of the line. Some system derived deflection differences agree very well with those obtained from astronomical observations, others deviate by 2" to 3". It appears that these deviations are of a size which can be expected from the slow adaption of the Kalman procedure. It is difficult to say, however, if this is the main effect or if changes in thermal and magnetic gradients as functions of platform heading also play a major role in changing the drift rates and by this the deflection estimates. There is one

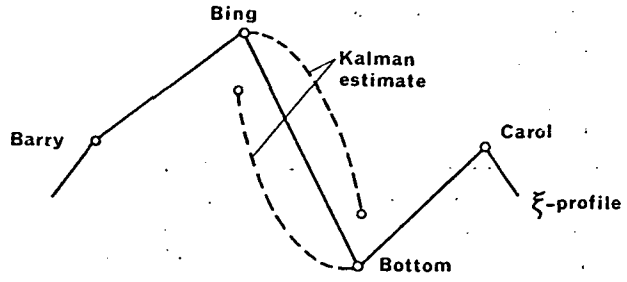



Fig. 4.6 Explanation of systematic $\Delta\xi$ - discrepancy between Bing - Bottom

observation which would indirectly confirm a highly nonlinear drift for the data analysed here. The value of $\sigma_I = \pm 0\overset{''}{.}55$ has been computed by making the usual linear drift removal between Curve and Kobau. When not removing any drift from the data the corresponding value drops to $\sigma_I' = 0\overset{''}{.}50$. This shows that the assumption of linear drift is a wrong model on a line like this because it does not improve results. Since drift is definitely present it must be suspected that it is highly nonlinear. This nonlinearity can probably be correlated to changes in platform heading.

In order to control these different effects it will be necessary to refine the mathematical model and to change the software accordingly. The gravity disturbance vector should be modelled as a position dependent quantity rather than a time dependent bias term of stochastic nature. An adequate model of the changes in gyro drift due to platform heading can probably be obtained by stationary experiments. Otherwise, a change of the measuring procedure would be necessary in order to determine drift changes during field operations.

The accuracy of deflection determination with present- day inertial instrumentation can be summarized in two numbers. Using the system as it is systematic differences of 2" to 3" must be expected even over relatively short distances. With changes in the software a standard error of $\pm 1$" or better can be expected between stations not more than half an hour of travel time apart. It should be noted that these results have been obtained in mountainous terrain and that no restrictions with respect to the course of the survey have been imposed. Results obtained by Fishel and Roof (1977) seem to confirm the above findings.

Acknowledgements. The data analysed in this paper have been made available by the Geodetic Survey of Canada. The cooperative support of several persons in this agency is gratefully acknowledged. The paper is a contribution to the Research Agreement No. 132728 between the Department of Energy, Mines, and Resources and the University of New Brunswick.

References

Britting, K.R. (1971), Inertial Navigation Systems Analysis. Wiley-Interscience, New York.

Farrel, F.L. (1976), Integrated Aircraft Navigation. Academic Press, New York.

Fishel, N. and E. Roof (1977), Results of Tests Using an Inertial Rapid Geodetic Survey System (RGSS). Proceedings of the American Congress on Surveying and Mapping, pp. 100 - 153.

Grafarend, E. (1975), Three Dimensional Geodesy I - The Holonomity Problem. Zeitschrift für Vermessungswesen, 100, pp. 269-281.

Heiskanen, W.A. and H. Moritz (1967), Physical Geodesy. W.H. Freeman and Co., San Francisco.

Lyon, F. (1977), Optimized Method for the Derivation of the Deflections of the Vertical from RGSS Data. Proceedings of the "First International Symposium on Inertial Technology for Surveying and Geodesy", Ottawa, pp. 417-428.

Luetzow, von B. (1977), A Review of Past, Present and Future USAETL Inertial/Gradiometric Geodesy Activities and Programs. Proceedings of the "First International Symposium on Inertial Technology for Surveying and Geodesy", Ottawa, pp. 60-72.

Moritz, H. (1975), Combination of Aerial Gravimetry and Gradiometry. Report No. 223 of the Department of Geodetic Science, The Ohio State University, Columbus.

Schwarz, K.P. (1977), Airborne Inertial Systems for Gravity Determination in Ocean Areas. Proceedings of the "First International Symposium on Inertial Technology for Surveying and Geodesy", pp. 351-360.

Todd, M. (1977), The Development of the Inertial Rapid Geodetic Survey System at USAETL. Proceedings of the "First International Symposium on Inertial Technology for Surveying and Geodesy", Ottawa, pp. 113-120.

Page Intentionally Left Blank

N79 21501

# Implications of Cavity, Topographic and Geologic Influences on Tilt and Strain Observations

J. C. Harrison
Department of Geological Sciences and CIRES, University of Colorado
Boulder, Colorado 80309

The existence of cavity effects was pointed out by King and Bilham in 1973, about a year after our first GEOP conference on earth and ocean tides. The principle involved is illustrated by figure 1 which shows the cross-section of a circular tunnel with tiltmeters installed at A and B. In the lower part of this figure the tunnel is shown deformed by horizontal compressive stress and two effects are obvious. Firstly, the tiltmeters suffer a very local strain-induced tilt which is not really what we mean by "tilt" at all and prompts the question "what is tilt anyway"; and secondly, the strain across the tunnel is three times as great as it would have been had the tunnel not been present. Clearly one cannot install a tiltmeter or strainmeter in an underground cavity and expect to measure a tilt or strain representative of the surrounding rock without first understanding the effect of the cavity. Changes in strain are likely to be associated with tilting in all cases of interest, be one concerned with earth-tides, build-up of tectonic strain in an earthquake region, or teleseismic and near-field tilts and strains following earthquakes. This discussion of cavity, topographic and geologic effects is equally valid for all these applications.

The concept of scale is very important. With earthtides and many other problems we are dealing with large scale deformations and considering macroscopically simple earth models, such as, for example, radially symmetric spherical earths. However, we are observing these deformations on a microscopically complicated earth, in cavities, often close to irregular topography and in the presence of local variations in elastic properties due to geological inhomogeneities. As none of these irregularities introduces net forces or couples, the influence of each is restricted to its immediate vicinity (St. Venant's principle) and its effect on the overall deformation of the earth is small. Unfortunately practical reasons often force us to measure in just those sites where the perturbations are large, and the traditional sites for tilt and strain observations - disused mines and tunnels - are often particularly badly affected. It is, however, usually the large scale "homogeneous" deformation which is of interest. The large scale strain obeys the large scale (plane or spherical) free surface boundary conditions and therefore has only 3 independent components. The homogeneous tilt is the tilt of the free horizontal surface and the boundary conditions ensure that this is the same as the tilt of a vertical line element; however, the tilt of an inclined line element is affected by strain even in the homogeneous case.

A measurement of strain is a measurement of change in distance per unit distance ($\delta l/l$) between two points on the earth's surface or on the interior of a cavity inside the earth. As long as we are in the linear regime of elastic behaviour and small deformations, this strain will be a linear combinations of the three components of the homogeneous strain. An ellipsoidal or infinitely long cylindrical cavity strains uniformly and we can define the entire strain tensor for the cavity; an irregular cavity deforms in an irregular manner and the best we can do is to define linear strain between two specified points. An observed tilt is similarly the sum of the homogeneous tilt plus strain-coupled tilts from each component of homogeneous strain.

Interpretation of an observed tilt or strain thus requires that the appropriate coupling coefficients be determined and this is normally done by numerical calculations using finite element modeling. (Levine and Harrison, 1976; Berger and Beaumont, 1976; Emter et al., 1977). This modeling is greatly facilitated if there is a clear separation of scales, so that the geology is on a large scale relative to the topography and the topography on a large scale relative to the cavity. Then the effects may be computed separately and the total estimated by a series of matrix multiplications. Otherwise all effects must be included in the one finite element model and the calculation may become very unwieldy.

Ellipsoidal and cylindrical cavities were treated analytically by Harrison (1976). Strain is measured correctly along the axis of an infinitely long tunnel; for a tunnel of finite length a circular cross-section is optimal and for such a tunnel a length/diameter ratio greater than 20:1 results in less than 1% strain error. Narrow cavities are extremely compressible in the direction of their short dimension; a fine example of this effect comes from the Schiltach Observatory in the Black Forest where a x58 strain magnification has been observed across a narrow cleft, in good agreement with finite element calculations which predict x53 (Emter et al., 1977). The floor of an infinitely long tunnel shows no tilt coupling in the direction of the tunnel axis, and the sides of a vertical borehole tilt as if the hole were not present.

Finite element calculations have given insight into the behaviour of more complex cavities. The walls of a tunnel of square cross-section bend outwards as a result of cross-tunnel tension, while the floor remains flat (figure 2). Cracks (figure 3) or geometrical irregularities (figure 4) however induce local tilts and the cracks, behaving as narrow cavities, exhibit very large strains.

The cavity effects have very important implications as tilt and strain have usually been measured in underground cavities to avoid meteor-

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

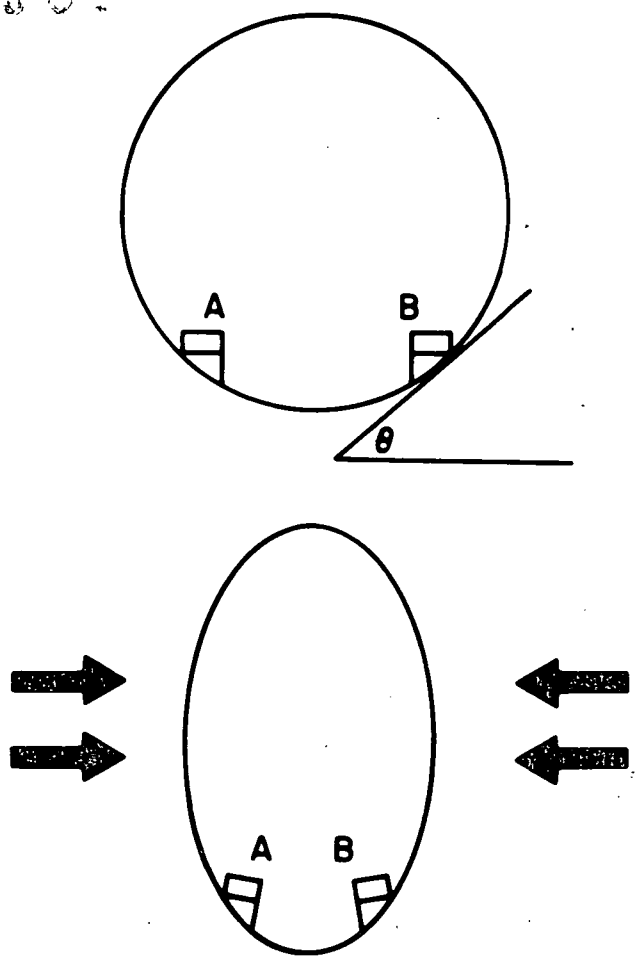



Fig. 1. A circular tunnel is deformed by horizontal strain. Tiltmeters at A and B record strain-coupled tilts and the horizontal strain across the tunnel is three times as large as they would have been had the cavity not been present.

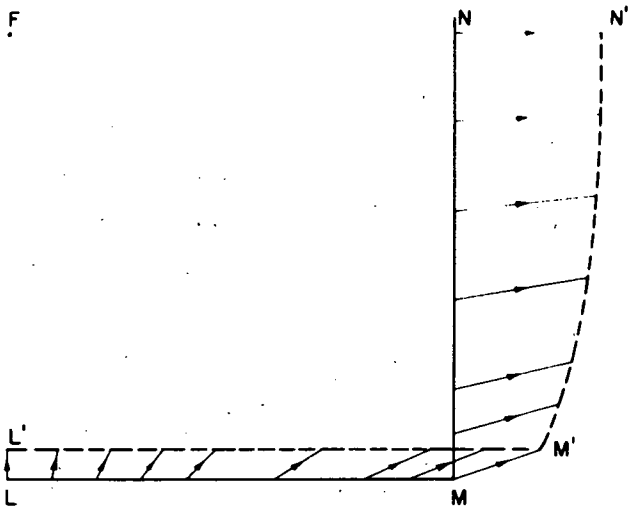



Fig. 2. Deformation of a square tunnel by horizontal tension. Only the bottom right corner of the tunnel is shown; the solid line represents the undeformed tunnel and the dashed line the deformed corner.

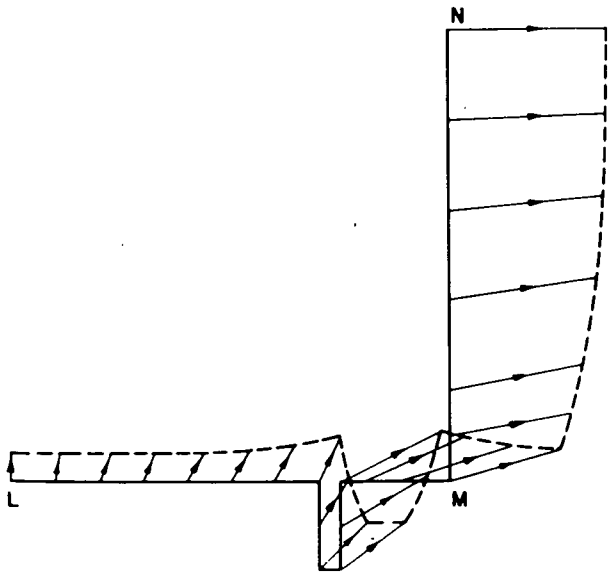



Fig. 3. Deformation of the tunnel of Figure 2 showing the effect of a crack.

ological perturbations. These cavities were excavated with explosives resulting in a fractured aureole surrounding the cavity in addition to cracks and joints which may be present naturally. Tilts and strains must be measured over baselines at least as long as the thickness of this fractured aureole. Strain measurements have been made over long baselines, typically 30m or more, along the axes of tunnels. They are not much affected by the cavity effects that one knows about and corrections are normally small. However, these observations may be affected by fractures and material inhomogeneities of which one is not aware, and differences in tidal strains measured with end to end strainmeters along tunnels in Great Britain suggest some such effect (Evans et al., in press).

For tilt the situation is entirely different. Tilt measurements have traditionally been made with short base (< 1m) instruments in geometrically complex situations. Finite element modeling would be very difficult and is, in practice, impossible because the fracture pattern in the vicinity of the tiltmeter is unknown. Thus all earth-tide tilt observations made with short base instruments are essentially worthless, and the observational techniques must be completely revised. One exception is when the tilt is primarily a loading tilt from a nearby ocean; in this case the load induced tilt is much larger than both the load induced and body-tide strain, and the strain-coupled tilt is correspondly less important.

Topographic effects are also more important than anyone had expected. Finite element calculations for a hill and valley situation with 45° slopes, figure 5, shows that the local strain in some places is actually of opposite sign to the homogeneous while in others, it is 3 times as large; up to 3.7 time if one considers strains measured across the valley. Even a valley with slopes of 1 in 10 produces a 36% modification of strain and tilt-strain coupling coefficients of the same order of magnitude.

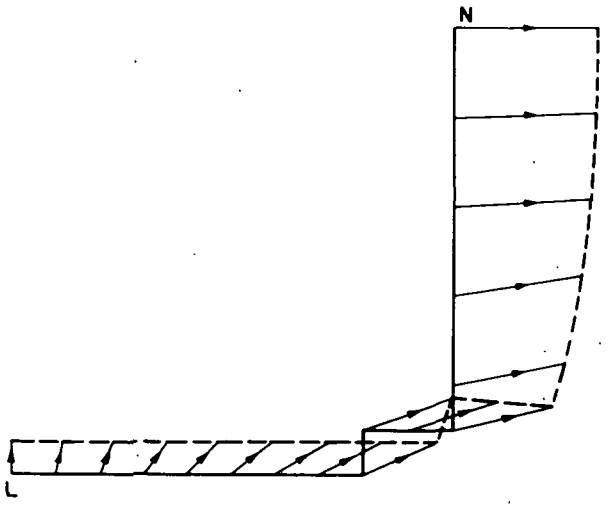



Fig. 4. Deformation of the tunnel of Figure 2 showing the effect of a geometrical irregularity (ledge).

As might be expected from the preceding discussion, material inhomogeneities ("geologic effects") also produce strain perturbations and tilt-strain coupling. Three examples are shown in Figures 6, 7 and 8; the first, the effects of a sediment-granite contact crudely modeling the Front Range-plains boundary near Boulder, Colorado; the second computations by Beaumont and Berger (1974) showing the effects of an assumed change in elastic constants due to dilatancy on the tilt and strain earth tides; and the third (unpublished) showing computations by the author of the effects of an assumed partially moulten zone beneath Yellowstone. These effects are large - modifications of the homogeneous tilt and strain tides are of the order of many tens percent, up to 100%, they are localized to the vicinity of the of the inhomogeneity; and the tilt anomalies appear to be larger but more restricted in area than those in strain.

A general agreement between tilt and strain tide observations and the predictions of these cavity, topographic and geologic influences is now well established. In general it is not possible to correct tilt observations made with short base tiltmeters with useful accuracy, because of the difficulty of correcting for the very small scale geometric and material inhomogeneities which have important influences on such measurements. All such measurements, that is all but a very few of body tide tilt observations made to date, are therefore worthless considering the accuracies required to contribute useful information about the earth. The strain situation is generally better because strain measurements have been made with long (30m - 100m) instruments along the axes of tunnels, which happens to be the correct technique from the point of view of cavity effects. Corrections for the cavity are therefore small and probably realistic. Levine and Harrison (1975) and Berger and Beaumont (1976) have corrected earth tide strain observations for the topographic and geologic effects. These corrections generally improve agreement between theory and observation and this is particularly

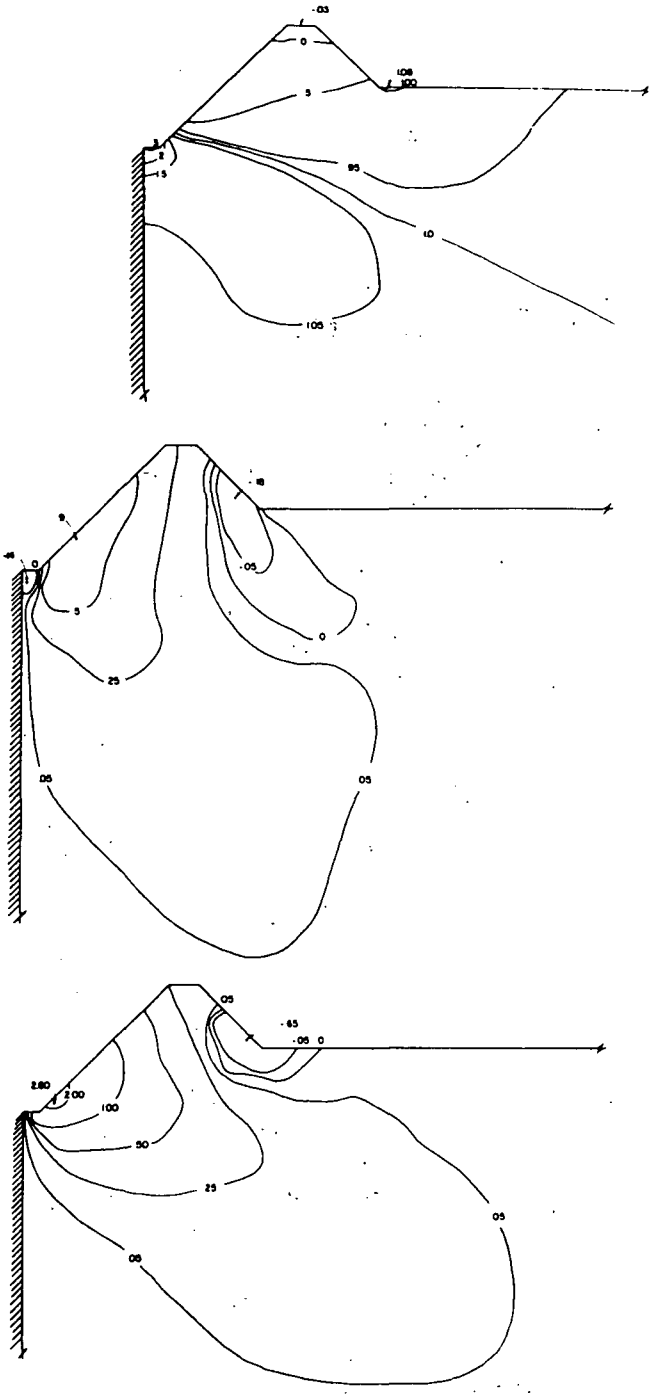


Fig. 5. Effect of topography with 45° slopes in modifying tilt and strain. The topography is symmetrical about the lachured line on the left. (top) horizontal strain-strain coupling factor (middle) tilt-strain coupline factor for a horizontal element (bottom) tilt-strain coupling factor for a vertical element.

obvious when they are large (20-30%). Never-theless only about half of the observations are in agreement with theory within the observational accuracy; this may be because the ocean loads are not accurately known but there is also the possibility that the modeling calculations are

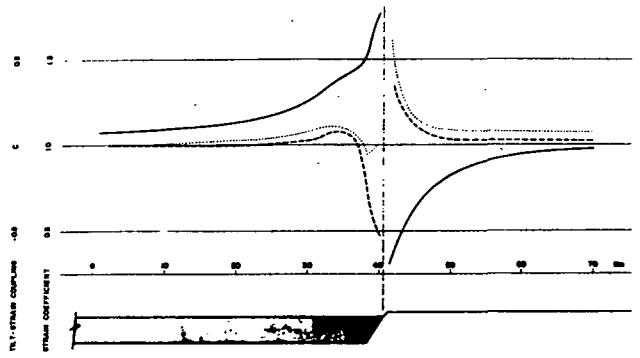
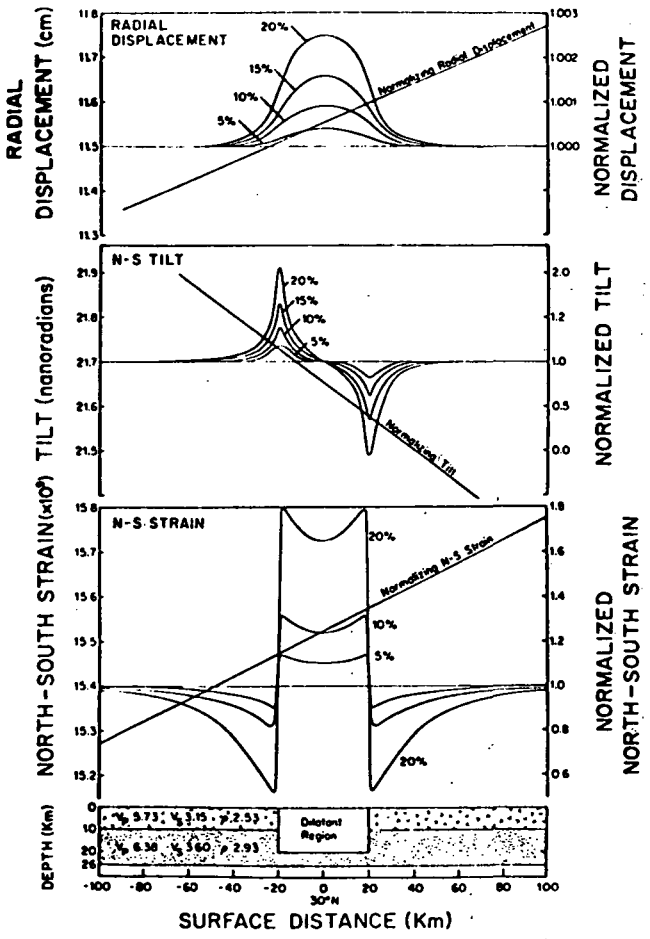


Fig. 6. Strain-strain (solid line), tilt-strain (horizontal element-dashed line) and tilt-strain (vertical element-dotted line) coupling factors for an idealized cross-section near Boulder, Colorado. Basement (unshaded) properties are Vp = 5.8 km/s, Vs = 1.80 km/s.




not as accurate as the finite element computations becuase unknown but significant faults, joints, elastic inhomogeneities, etc. have not been included.

In summary tilt and strain tide observations are importantly (pathologically at the 100%, typically at the few 10s% level) affected by cavities, topography and geological inhomogeneities; gravity observations are practically unaffected. It is important that tilt and strain be observed with long base instruments because small irregularities, too small to model realistically, can have important local effects. On the other hand we do not really know what we mean by "long" because we do not know on what, if any scale, rock behaves as a continuous homogeneous medium; earth tide observations can help us answer this question, and different areas probably behave in different ways as a result of differing rock types and tectonic histories. The traditional earth tide observatory, and abandoned mine or tunnel is a very poor place to measure body tides because of complicated cavities, topography and geology. Instead the ideal site for observing the body tide is in flat terrain with horizontally layered, mechanically homogeneous geology. Strain will be measured with long surface or trench mounted laser strain meters and tilt with long, surface or trench mounted liquid levels, or with borehole tiltmeters. Horizontal geological discontinuities can produce large perturbations of the tilt and strain tides and these perturbations, using the known homogeneous tidal strains and tilts, can be used to explore local structure in favorable cases and, through possible time

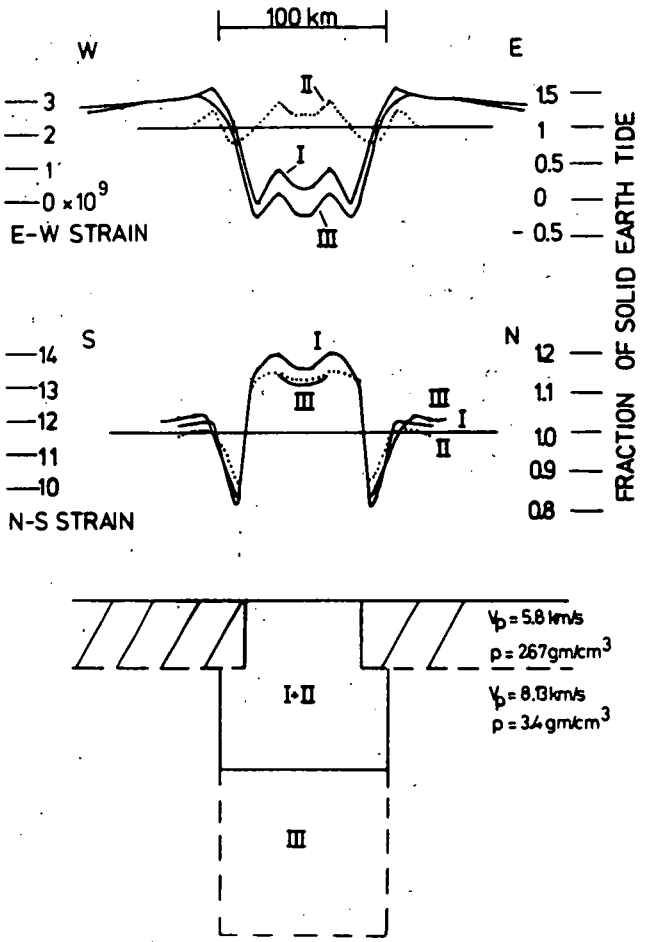



Fig. 7. The effect of a dilatant zone in modifying earth tides. The normalized functions represent the fractional change in radial displacement, tilt and strain as a function of position when the P-velocity in the dilatant zone is reduced by 5, 10, 15, and 20 percent (from Beaumont and Berger, 1974).

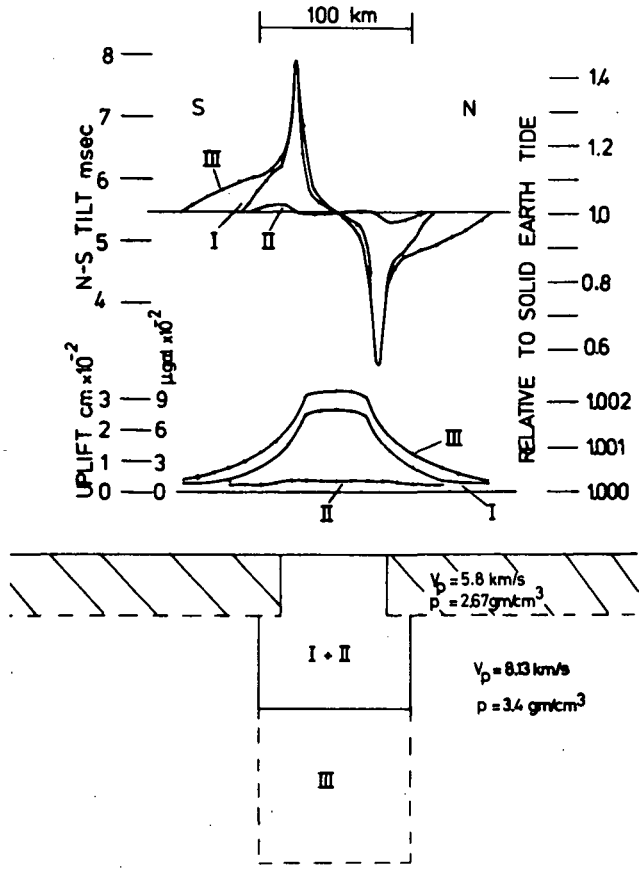



Fig. 8. Modification of tides due to partially moulten zone beneath Yellowstone National Park on the assuption of a 10% reduction in Vp. Models I and II are for the body shown extending to 100 km with, in I material properties corresponding to flat inclusions and in II, properties corresponding to round inclusions. Model III has the material properties of I but extends to 200 km depth. (a) radial displacement and N-S tilt and (b) E-W and N-S strain.

variations of tidal admittances, in earthquake prediction. Ocean load uncertainties currently preclude the use of tidal observations for determining whole earth structure or body tide energy dissipation, but it does seem possible that they can be utilized for exploration on a smaller scale - at least this is a possibility that deserves further investigation.

## References


Beaumont, C. and J. Berger, Earthquake prediction: modification of the earth tide tilts and strains by dilatancy, Geophys. J. Roy. Astron. Soc., 39, 111-121, 1974.

Berger, J. and C. Beaumont, An analysis of tidal strain observations from the United States of America. II. The inhomogeneous tide. Bull. Seis. Soc. Amer., 66, 1821-1846, 1976.

Emter, D., A. Jensch and H. Kiessel, Finite element estimates of elastic effects on tidal tilt and strain with special respect to results from the Schiltach Observatory, in press, proc. 8th Int. Earth Tide Symposium, Bonn, 1977.

Evans, R., J. Beavan, R. Bilham and G. King, A survey of strain tides in Great Britain, in press.

Harrison, J.C., Cavity and topographic effects in tilt and strain measurement, Jour. Geophys. Res., 81, 319-328, 1976.

Levine, J. and J.C. Harrison, Earth tide strain measurements in the Poorman Mine near Boulder, Colorado, Jour. Geophys. Res., 81, 2543-2555, 1976.

Page Intentionally Left Blank

# Open Ocean Tide Modelling



M. E. Parke
NRC Associate, NOAA/PMEL, Seattle, Washington 98105

Abstract. The period since the first GEOP conference in 1972 has seen marked changes in global tidal modelling, with many new models produced in the past two years. Two trends have been evident. The first centers on the incorporation of terms for ocean loading and gravitational self attraction into Laplace's Tidal Equations (LTE). The second centers on a better understanding of the problem of near resonant modelling and the need for realistic maps of tidal elevation for use by geodesists and geophysicists. These trends are described. Although new models still show significant differences, especially in the South Atlantic, there are significant similarities in many of the world's oceans. This allows suggestions to be made for future locations for bottom pressure guage measurements. Where available, estimates of M2 tidal dissipation from the new models are significantly lower than estimates from previous models. The new estimates are consistent with recent estimates of the rate of decelleration of the lunar longitude.

## Introduction

Over the past seven years since the first GEOP conference in 1972 there have been two marked trends in tidal modelling. The first centers on the incorporation of the terms for ocean loading and gravitational self attraction into Laplace's Tidal Equations (LTE). The second centers on a better understanding of the problem involved with near resonant modelling and the need for realistic maps of tidal elevation for use by geophysicists and geodesists. It is the purpose of this article to describe these two trends, and to look at what improvements are likely in the future.

Because of the scope of the article, and since many researchers have calculated models only for M2, only M2 models after 1972 will be presented. Calculations for other constituents will be mentioned in passing. For a good summary of modelling prior to 1977 and a detailed discussion of modelling techniques see Hendershott [1977]. Another excellent summary can be found in Cartwright [1977].

## Ocean Loading and Self Gravitation

The terms for ocean loading and self attraction were incorporated into LTE by Hendershott [1972]. The effect that these terms have on LTE may be summarized as follows. LTE may be written after Hendershott [1977] as a single elliptic elevation equation in negative mercator coordinates as

$$\mathcal{L}(\zeta_o) + \varepsilon^2 \mathrm{sech}^2\tau\, \zeta_o = \mathcal{L}(\Gamma/g-\delta) + \overline{F} \qquad (1)$$

where

$$\mathcal{L} = QH\nabla^2 + [(QH)_\phi - (i/s)(QH\tanh\tau)_\tau]\partial\phi + [(QH)_\tau + (i/s)(QH\tanh\tau)_\phi]\partial\tau \qquad (2)$$

$$\overline{F} = (1/\rho D_o)[(QF^\phi)_\phi + (QF^\tau)_\tau] - (i/\rho D_o s)[(Q\tanh\tau F^\phi)_\tau - (Q\tanh\tau F^\tau)_\phi] \qquad (3)$$

and

$$\varepsilon^2 = 4\Omega^2 a/gD_o \qquad Q = 1/(s^2-\tanh^2\tau)$$
$$D = D_o H(\phi,\tau) \qquad s = \sigma/2\Omega \qquad (4)$$

where $\tau,\phi$ are the mercator latitude and longitude, $\zeta_o$ is the observed ocean tide, $F^\tau, F^\phi$ are the meridional and zonal components of dissipative stress, $\sigma$ is the tidal forcing frequency, $\Omega$ is the earth's angular rate of rotation, $\Gamma$ is the total tide generating potential, $\delta$ is the geocentric solid earth tide, D is the local depth of the ocean, and a is the radius of the earth. The exact form of dissipative stress for each model will depend on the choice of $F^\tau$ and $F^\phi$.

For a rigid earth, $\Gamma = U$ (the astronomical potential) and $\delta = 0$, i.e.

$$\Gamma/g - \delta = U/g \qquad (5)$$

In the presence of solid earth deformation and tidal loading $\Gamma/g - \delta$ may be expressed as follows. In the usual Love number notation (Munk and MacDonald, 1960, p. 24,29,30), the $n^{th}$ spherical harmonic component of the potential is

$$\Gamma_n = (1+k_n)U_n + (1+k_n')g\alpha_n\zeta_{on} \qquad (6)$$

while the $n^{th}$ spherical harmonic component of the solid earth tide is

$$\delta_n = h_n U_n/g + h_n'\alpha_n\zeta_{on}. \qquad (7)$$

Here the $n^{th}$ spherical harmonic of the observed ocean tide is

$$\zeta_{on} = \sum_m C_{nm} Y_n^m(\sin\theta) \qquad (8)$$

where $Y_n^m$ is the spherical harmonic normalized after Backus [1958], i.e. $|Y_n^m| = 1$, $U_n$ is the $n^{th}$ spherical harmonic component of the astronomical potential (in practice only $U_2$ need be considered), $\theta$ is the latitude, and

$$C_{nm} = \iint \zeta_o(\phi',\theta') Y_n^m(\sin\theta')\cos\theta'\,d\phi'\,d\theta'. \qquad (9)$$

Thus if one defines

$$\overline{\zeta} = \sum_n (1+k_n-h_n)U_n/g \qquad (10)$$

then

$$\Gamma/g - \delta = \overline{\zeta} + \sum_n (1+k_n'-h_n')\alpha_n \iint \zeta_o Y_n^{m*}\cos\theta'\,d\phi'\,d\theta' \qquad (11)$$

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280. The Ohio State Univ., Columbus, Ohio 43210.

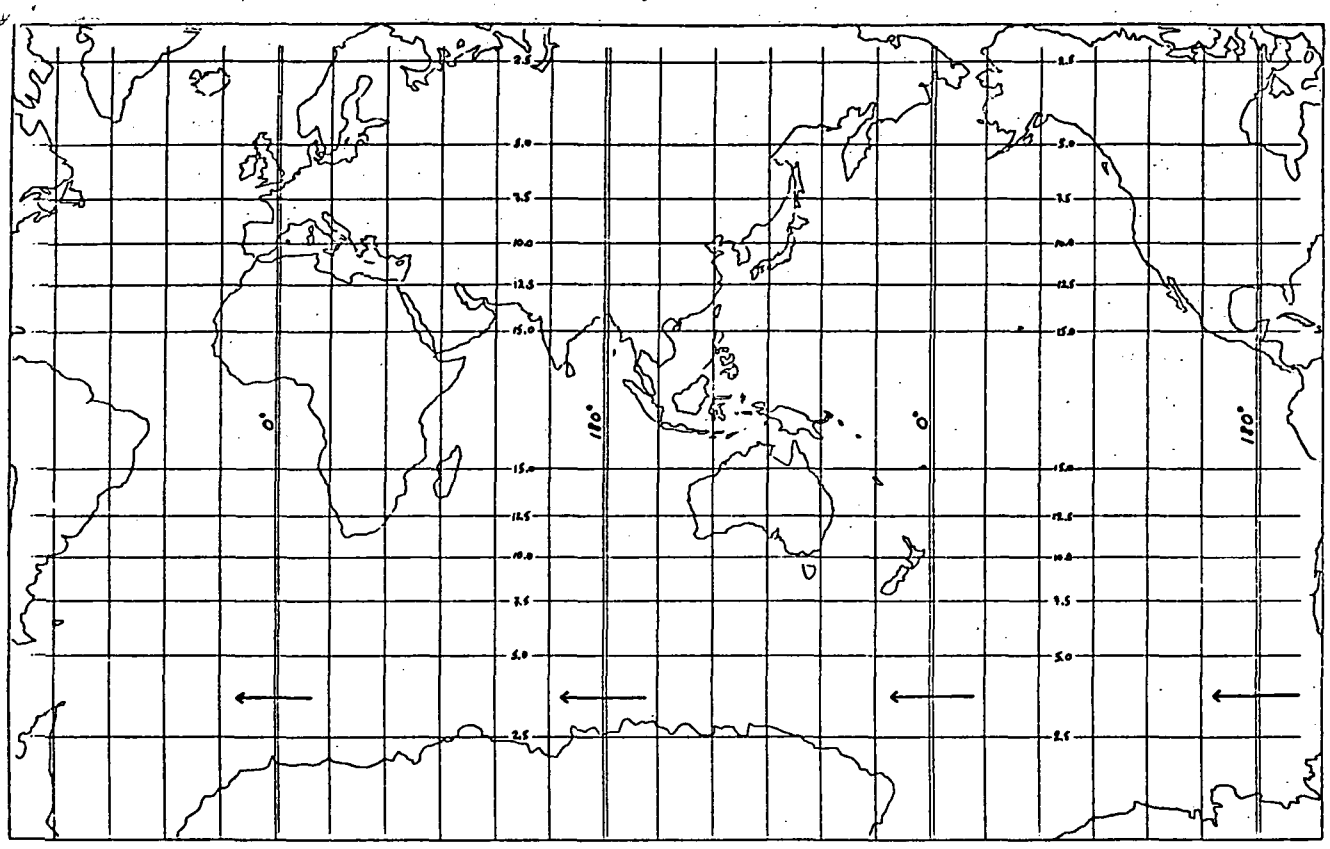


Fig. 1. The M2 gravitational potential divided by g seen by an observer fixed to the surface of the earth in the absence of ocean tidal effects, i.e. $(1+k_2-h_2)U_2$. (amplitudes in cm, phases in degrees)

or by placing the summation inside the integral,

$$\Gamma/g - \delta = \overline{\zeta} + \iint \zeta_o(\phi',\theta')G(\phi',\theta'|\phi,\theta)\cos\theta' d\phi' d\theta' \quad (12)$$

where $G(\phi',\theta'|\phi,\theta)$ is a Green's function evaluated by Farrell [1972]. In negative mercator coordinates

$$\Gamma/g - \delta = \overline{\zeta} + \iint \zeta_o(\phi',\tau')G(\phi',\tau'|\phi,\tau)\mathrm{sech}^2\tau' d\phi' d\tau' \quad (13)$$

The presence of this global integral makes solution difficult. Hendershott [1972] proposed the iterative sequence

$$\mathcal{L}(\zeta_o^1) + \varepsilon^2\mathrm{sech}^2\tau\zeta_o^1 = \mathcal{L}(\textstyle\sum_n(1+k_n-h_n)U_n/g) + \overline{F} \quad (14)$$

$$\mathcal{L}(\zeta_o^i) + \varepsilon^2\mathrm{sech}^2\tau\zeta_o^i = \mathcal{L}(\textstyle\sum_n(1+k_n-h_n)U_n/g) + \overline{F} + \iint \zeta_o^{i-1}G(\phi',\tau'|\phi,\tau)\mathrm{sech}^2\tau' d\phi' d\tau' \quad (15)$$

and found it divergent for M2 in the absence of interior dissipation, i.e. $\overline{F} = 0$. Gordeev, et. al [1977] showed that this procedure will converge in the presence of interior dissipation. Their calculations based on the Longman [1963] Green's function are consistent with the importance of these terms. Parke [1978a] showed dramatically that the loading and self gravitation terms are important for M2 by comparing the surface potential seen by an observer fixed to the surface of the earth with and without ocean effects, i.e.

$$\Gamma/g = (1+k_2-h_2)U_2/g \quad (16)$$

vs

$$\Gamma/g = (1+k_2-h_2)U_2/g + \textstyle\sum_n(1+k_n'-h_n')\alpha_n\zeta_{on} \quad (17)$$

where the ocean terms were calculated based on the Parke-Hendershott model discussed in the next section. These charts are shown in figures 1 and 2. Note the considerable distortion in the North Atlantic caused by the ocean effects. Historically researchers have had trouble with local models of the North Atlantic using the equilibrium potential. Perhaps this is the reason why.

Over the past several years the importance of the terms for ocean loading and self attraction has gradually become accepted, with models sans ocean loading effects being published as late as 1977. Starting in 1972, the first iteration of Hendershott is a model sans loading effects and is presented in figure 3. The calculation is based on a six degree Mercator grid with specified elevation boundary conditions and implicit dissipation, i.e. $\overline{F} = 0$. Energy is allowed to freely flow through the boundaries to be dissipated in the shallow seas and shelves.

Zahel [1973] calculated a model for K1 using a four degree spherical mesh graded toward the

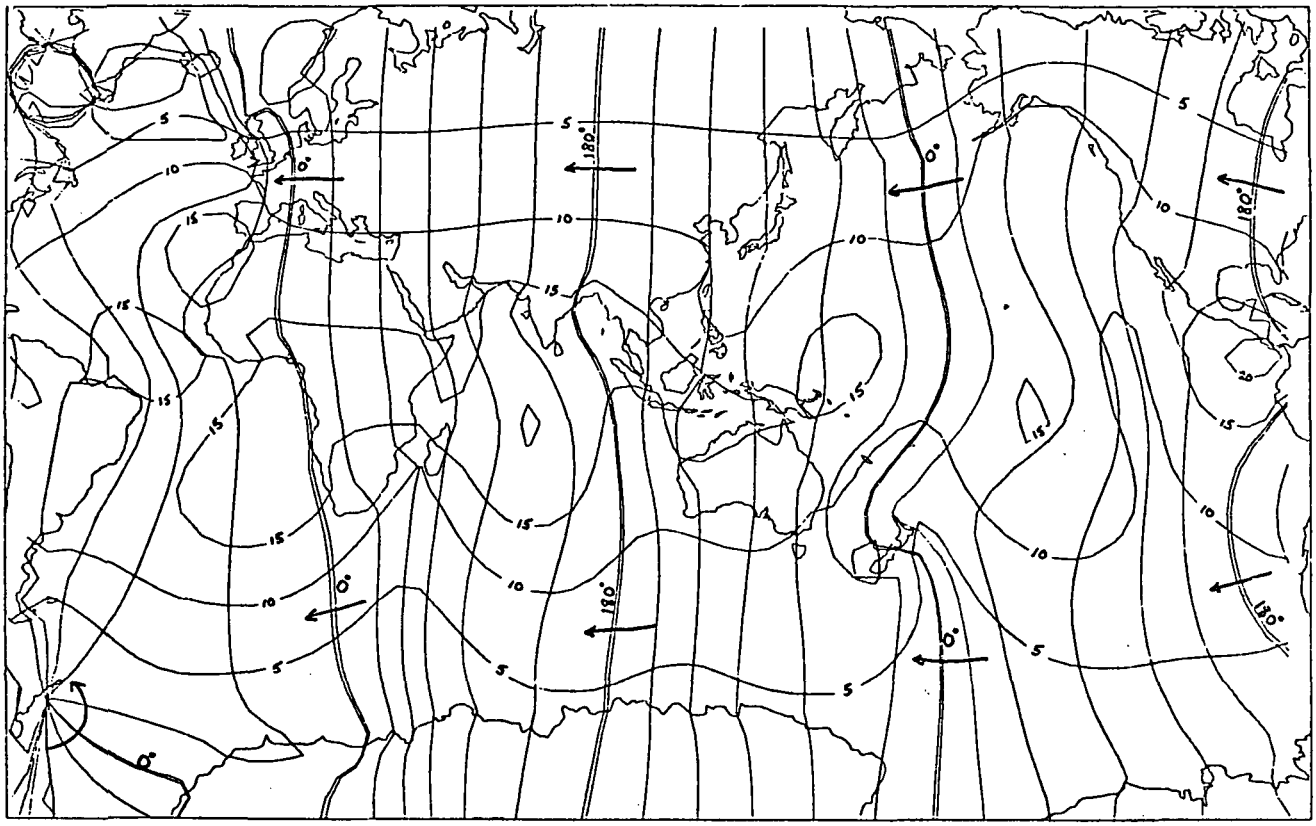


Fig. 2. The M2 gravitational potential divided by g seen by an observer fixed to the surface of the earth when the effects of ocean loading and gravitational self attraction from the Parke-Hendershott [1978] tidal model are added, i.e. $(1+k_2-h_2)U_2 + \sum_n (1+k_n'-h_n')\alpha_n \zeta_{on}$. (amplitudes in cm, phases in degrees)

poles, reflecting boundary conditions, and dissipation of the form

$$F/\rho D = -r((u^2+v^2)^{1/2}/D)\underline{u} + AV^2\underline{u} \qquad (18)$$

where r is a bottom drag coefficient and A a lateral eddy viscosity with r = .003 and A = $10^7$ $cm^2$/sec. Zahel presented a similar calculation for M2 on a one degree mesh at the IUGG meeting in Grenoble in 1975, and subsequently published [Zahel,1977]. Figure 4 is a copy of the cotidal chart presented in 1975. Zahel has recently completed a model including loading effects (Trevor Baker, personnal communication).

Estes [1977] repeated Zahel's procedure for M2 using a two degree mesh and allowing for deformation of the solid earth by the astronomical forcing. Resulting amplitudes were smaller as would be expected. This calculation was extended to S2, N2,K2,K1,O1, and P1. The M2 model was used as a starting point for the iterative procedure of Hendershott described above. Convergence was found after 16 iterations. This model is shown in figure 5.

Parke [1978b] solved the modified LTE on a six degree Mercator mesh using specified elevation boundary conditions and no interior dissipation. This was done by using a set of test functions similar to the iterates of Hendershott as a basis set for a least squares fit to the complete equations. The test functions were generated by the same method as the iterates, except that periodically the best least squares fit was subtracted from the equations and an iterative sequence started on the residuals. This was done to aid numerical stability. Solutions for M2 and S2 were found to be unrealistically resonant, while for K1 the iterative procedure of Hendershott was found to be convergent and the solution showed every sign of being far from resonance.

Accad and Pekeris [1978] calculated models for M2 and S2 using a two degree mercator mesh with implicit dissipation determined at the coasts by a modified Proudman boundary condition. Instead of treating the edge of the coastal shelves as a step function, they treat it as a linear ramp. For a step function, the assumption that the energy contained in the upper layer (of depth h') is dissipated rather than reflected leads to the usual Proudman condition

$$hu_n = (gh')^{1/2}\zeta_o \qquad (19)$$

where h is the depth offshore from the shelf and $u_n$ is the normal velocity. When the edge of the shelf is considered as a ramp, this condition becomes

$$hu_n = (gh)^{1/2}(1-R)/(1+R)\zeta_o \qquad (20)$$

where R is a complex reflection coefficient depending on S (the width of the ramp), h, and h'. Accad and Pekeris used the values S = 100km and h' = 10m, while h was taken to be the observed

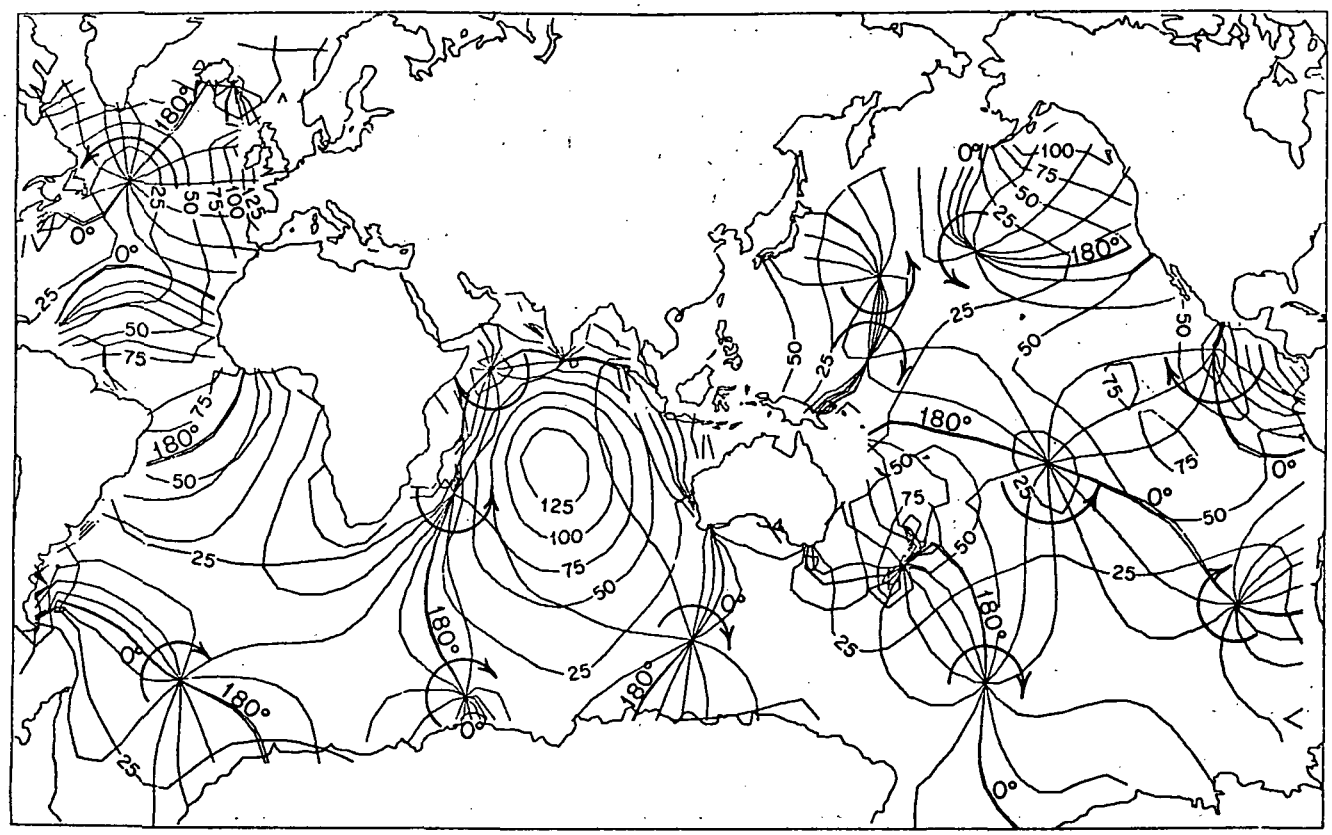



Fig. 3. The Hendershott [1972] M2 tidal model. (amplitudes in cm, phases in degrees)

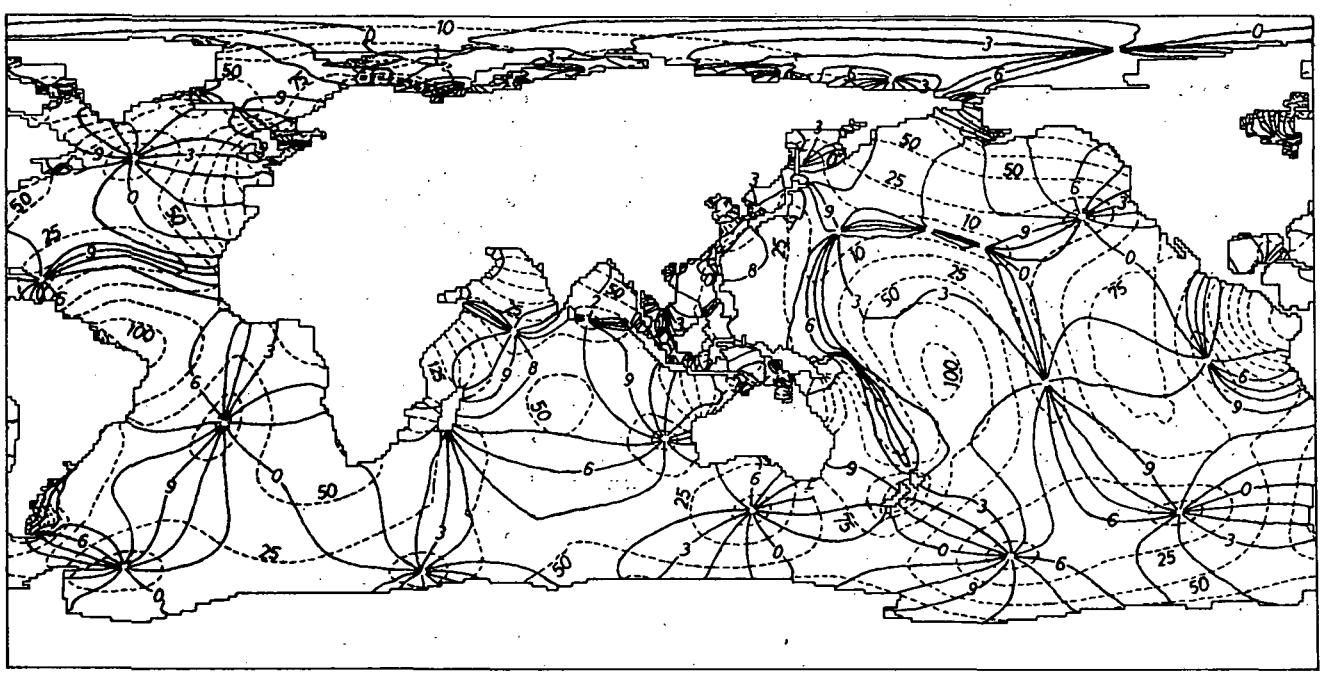



Fig. 4. The Zahel [1977] M2 tidal model. (amplitudes in cm, phases in lunar hours)

value of the ocean depth closest to the grid point on the coast. Ocean loading terms were approximated by

$$\sum_n (1+k'_n-h'_n)\alpha_n \zeta_{on} \simeq 0.085\zeta_o \simeq \sum_n 0.085\zeta_{on} \qquad (21)$$

or

$$(1+k'_n-h'_n)\alpha_n \simeq 0.085 \qquad (22)$$

for all n. The M2 results are given in figure 6.

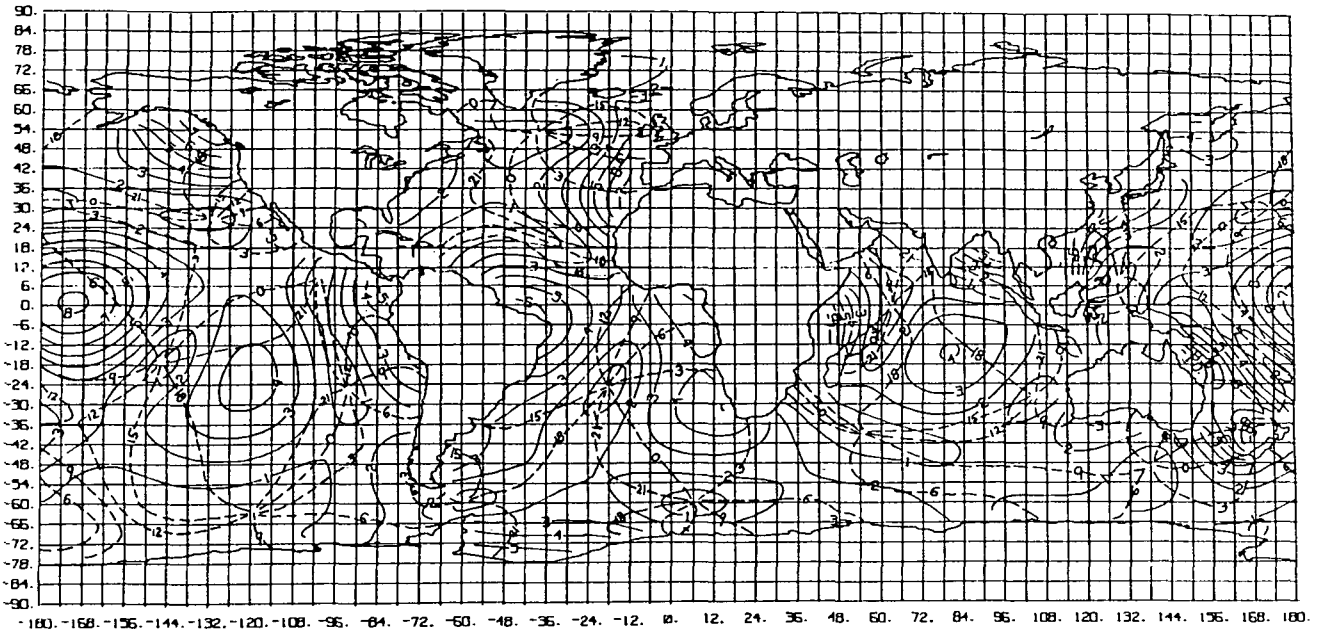


Fig. 5. The Estes [1977] M2 tidal model with the effects of ocean loading and self attraction included. (amplitudes in decimeters, phases in lunar half hours)

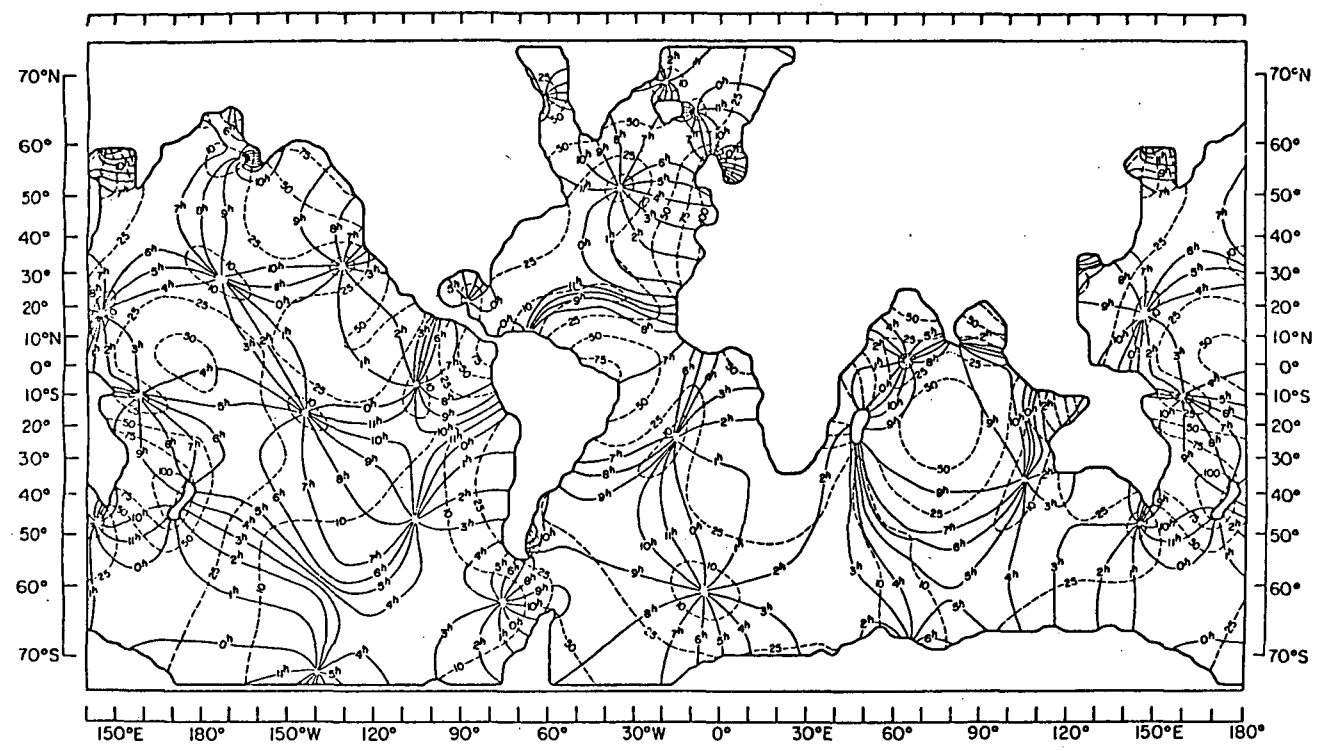



Fig. 6. The Accad and Pekeris [1978] M2 tidal model. (ampl. in cm, phases in lunar hours)

## Near Resonant Modelling

Over the past few years there has been increasing geophysical interest in finding a realistic representation of the open ocean tide. Historical models of the semi-diurnal tides (primarily M2) show significant differences, especially in the Pacific and South Atlantic. Considering the closeness to resonance of the problem, though, the level of agreement is actually quite remarkable. The fact that there are significant differences, however, means that to produce realistic maps of the ocean tidal elevation the problem of resonance must be handled. This problem arises because small errors in how one's model represents the real ocean basins cause small errors in the frequencies at which the model basin resonates. Near resonance this causes a significant error in the assigned amplitudes for these modes.

As an example of a near resonant one mode sys-

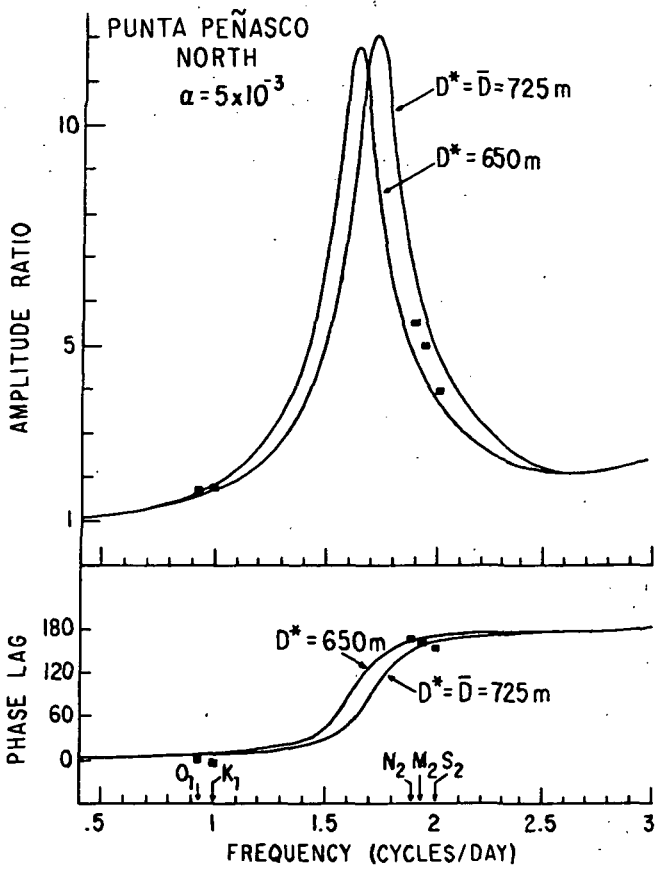



Fig. 7. A comparison of the tidal response at Punta Peñasco North to the tide at the mouth of the Gulf of California for two models with different mean depth, after Stock [1976].

tem, figure 7 shows the model response of the Gulf of California at two different mean depths after Stock [1976]. Far from resonance, e.g. O1 and K1, the difference between the responses is small. Near resonance, however, e.g. for M2, S2, and N2, the difference is quite significant. In a one mode system such as the Gulf of California, this effect can be adjusted with a single parameter such as the mean depth. The global tide, however, appears to be more complex. Four possible approaches are:

(1) If one knew the near resonant global normal modes, one could adjust one's model accordingly.

(2) One can adjust model parameters. Schwiderski [1978] follows this approach with a one degree graded spherical mesh with linear eddy dissipation with eddy viscosity A given by

$$A = aH(\lambda,\phi)L/2(1+\mu\cos\phi) \qquad (23)$$

and a linear bottom friction term with coefficient

$$B = bL^2\mu\cos\phi \qquad (24)$$

where a and b are trial and error parameters, L is the equatorial mesh size, $\mu$ is the mesh grading parameter, and $H(\lambda,\phi)$ is the local depth of the ocean. Note that A is directly proportional to H, and that B has no depth dependence. The linear bottom friction term is adjusted within set limits at each boundary point and island station to force as close as possible agreement between the observed and calculated tides at these locations. Loading terms are approximated after Pekeris by

$$\sum_n (1+k'_n-h'_n)\alpha_n\zeta_{on} \simeq 0.1\zeta_o \qquad (25)$$

where the factor 0.1 is attributed to Pekeris. A cotidal chart copied from a plot provided by Ron Estes is shown in figure 8. One consequence of forcing agreement at island stations is the creation of small scale local distortions. See for example the Southwest Pacific.

(3) Parke and Hendershott [1978] took the set of test functions used by Parke [1978b] to solve the modified LTE in the absence of interior dissipation and fit them in the least squares sense to interior (island) data. This was done for M2, S2, and K1 and represents a first order correction to the problem. The resulting models are on a six degree Mercator mesh with specified elevation boundary conditions and dissipation only in shallow seas and on shelves. Encouraging is the fact that the M2 calculation was shown to be stable over a ±5% variation in the mean depth. All three representations conserve mass. A cotidal chart for M2 is given in figure 9.

(4) Estes (personnal communication) proposes combining models of the tide with other data such as altimetry and gravity measurements, with the model value at each point considered simply as another datum.

It is interesting and provocative that the Schwiderski and Parke-Hendershott models show many qualitative and quantitative similarities throughout much of the worlds oceans. Starting with the Pacific, both show a convergence of phases between Japan and New Guinea, amphidromes off California, Latin America, Chile, and in the Central Pacific. West of the California amphidrome Schwiderski shows small amplitudes and a convergence of phases, while Parke-Hendershott shows a double amphidrome. It should be noted that the absolute difference between these two cases is small. The amphidrome that Parke-Hendershott show southeast of New Zealand has moved much farther to the east and south in the Schwiderski model. Both show two anti-amphidromes in the Pacific; Parke-Hendershott at approximately 4°N 174°E and 2°S 126°W while Schwiderski places them at approximately 6°S 176°E and 4°S 129°W. The dominant feature of the Indian Ocean for all models is a central anti-amphidrome. Parke-Hendershott place this at approximately 18°S 78°E while Schwiderski places it at approximately 20°S 78°E. In the South Atlantic Parke-Hendershott show a region of low amplitudes extending eastward from South America while Schwiderski shows a double amphidrome. Both show an anti-amphidrome next to the tip of Africa with Parke-Hendershott placing it at approximately 32°S 5°E and Schwiderski at 30°S 10°E.

## Discussion

Table 1 summarizes the global tide models calculated since 1972. Recent estimates of model dissipation by Parke-Hendershott [1978] of 2.22 x $10^{19}$ ergs/sec and by Accad and Pekeris of 2.55 x

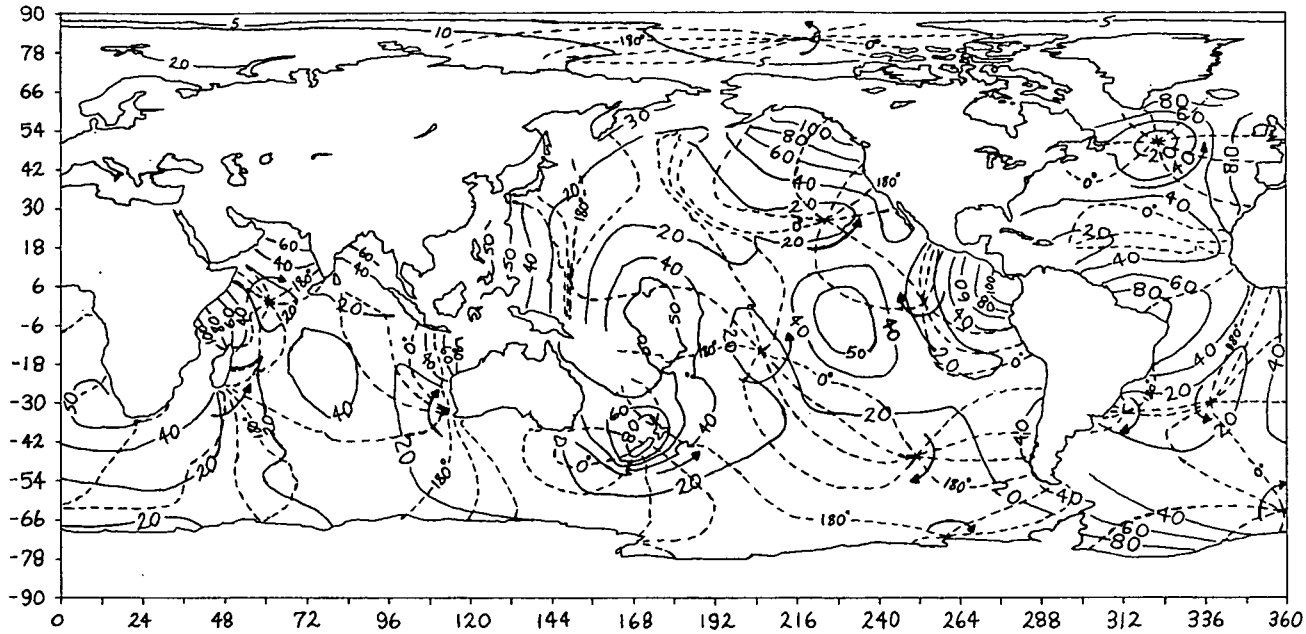



Fig. 8. The Schwiderski [1978] M2 tidal model. Many shallow seas and shelves such as the Patagonian shelf show spacial structure much too fine to be resolved at these scales and so these regions have been left blank. (amplitudes in cm, Phases in degrees)

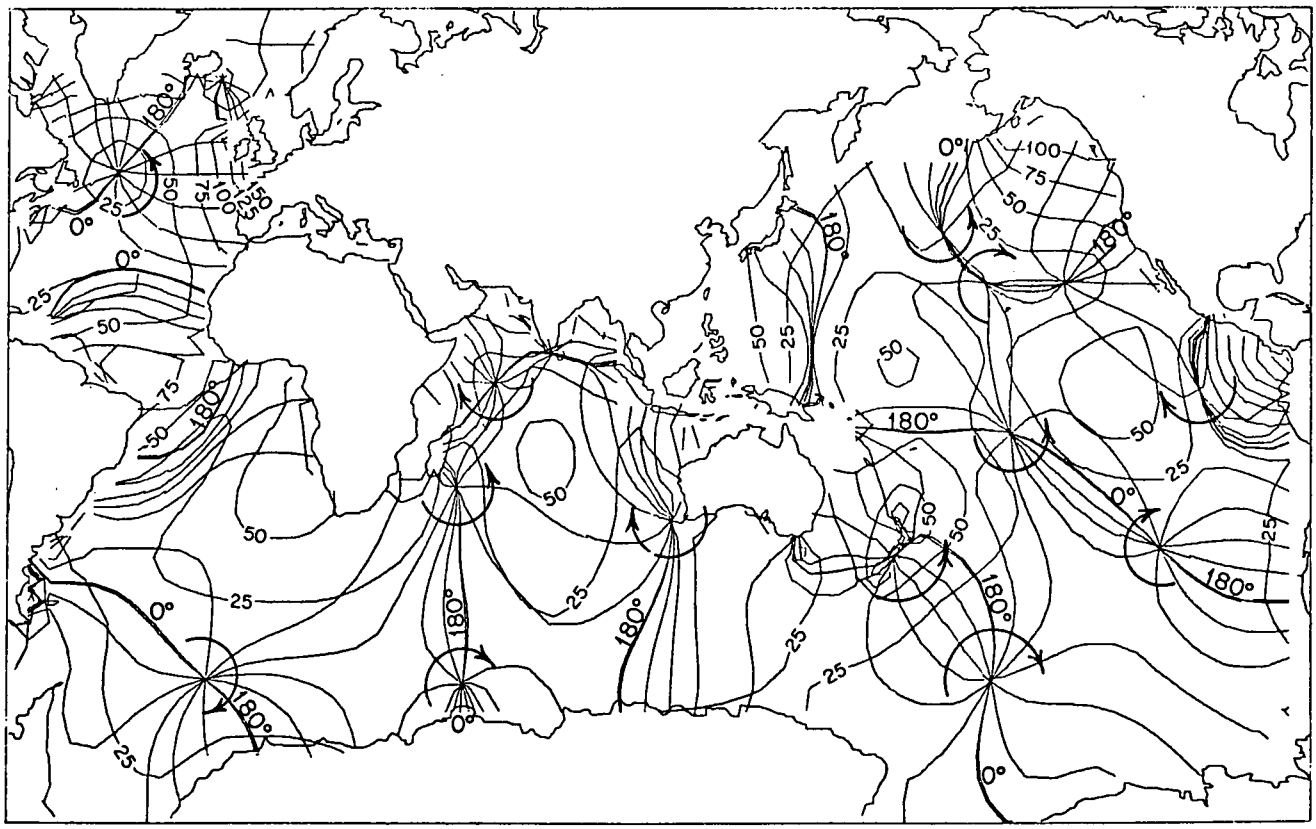


Fig. 9. The Parke-Hendershott [1978] M2 tidal model. (amplitudes in cm, phases in degrees)

$10^{19}$ ergs/sec are significantly lower than previous model estimates. Recent astronomical estimates include Muller [1977] with 3.3 ± .2 x $10^{19}$ ergs/sec and Goad and Douglas [1978] with 3.3 ± .4 x $10^{19}$ ergs/sec. It should be noted that the astronomical estimates are for all lunar dissipations not just M2. Estimates of the non-M2 ocean contribution to lunar dissipation vary from 0.3 - 1.0 x $10^{19}$ ergs/sec. This leaves a range for the M2 ocean tidal dissipation based on the Muller estimate of 2.1 - 3.2 x $10^{19}$ ergs/sec. Miller [1966] estimates empirically a M2 dissipation in the shallow seas and shelves of 0.9 - 2.5 x $10^{19}$ ergs/sec. Combining these two esti-

TABLE 1: Summary of Global Ocean Tidal Models since 1972

| Model | Constituent(s) | Mesh | Boundary Conditions | Form of Dissipation | Dissipation ($\times 10^{19}$ erg $s^{-1}$) | Loading Terms |
|---|---|---|---|---|---|---|
| Hendershott [1972] | M2 | 6° mercator | specified elevation | implicit, in shallow seas and shelves only | 3.08 | yielding to astronomical force only |
| Zahel [1973] | K1 | 4° spherical | reflecting | bottom stress in shallow water | - | none |
| Zahel [1973] | M2 | 1° spherical | reflecting | bottom stress in shallow water | 3.77 | none |
| Estes [1977] | M2,S2,N2,K2, K1,O1,P1 | 2° spherical | reflecting | bottom stress in shallow water | - | yielding to astronomical force only |
| Estes [1977] | M2 | 3° spherical | reflecting | bottom stress in shallow water | - | complete potential using Green's function |
| Schwiderski [1978] | M2 | 1° spherical | reflecting | bottom friction $bL^2\mu\cos\phi$ eddy dissipation $a/2\ L\ H(\lambda,\phi)(1+\mu\cos\phi)$ | - | estimate loading terms with $0.1\zeta_o$ |
| Parke-Hendershott [1978] | M2,S2,K1 | 6° mercator | specified elevation | implicit, in shallow seas and shelves only | M2 = 2.22<br>S2 = 0.208<br>K1 = 0.221 | complete potential using Green's function |
| Accad and Pekeris [1978] | M2,S2 | 2° mercator | modified Proudman condition using ramp shaped shelf edge | implicit, in shallow seas and shelved only | M2 = 2.55<br>S2 = 0.526 | estimate loading terms with $0.085\zeta_o$ |

mates, one would expect M2 ocean tidal dissipation to be in the range 2.1 - 2.5 x $10^{19}$ ergs/sec which is in excellent agreement with the new model estimates.

In the last section, it was noted that there exists a marked similarity between the models of Parke-Hendershott and Schwiderski. There are also a number of qualitative similarities with the loading solutions of Estes and Accad and Perkeris. All four models show two anti-amphidromes in the Pacific and one in the Indian Ocean, although Estes' anti-amphidrome in the eastern Pacific is much further south than that for the other three. Estes also shows an anti-amphidrome just west of the tip of Africa similar to that of Parke-Hendershott and Schwiderski. All four show amphidromes in the North Atlantic, in the South Atlantic near Antartica, off the coast of California, and in the central Pacific.

Anti-amphidromes represent ideal places for bottom pressure guage measurements, since they are locations of large ampitude and spacial stability. Because of the strong similarities in the location and number of anti-amphidromes in the above M2 models, a strong recommendation can be made as to where measurements should be taken. Thus the following list of locations is recommended; near 1°S 175°E, near 3°S 128°W, near 19°S 78°E, and near 31°S 8°E. Also since there are relatively large variations from model to model in the South Atlantic between South America and Africa, 30°S 30°W is recommended.

Although new models for other constituents are not shown here, similar recommendations can be made for S2. Accad and Pekeris' S2 model shows anti-amphidromic points at 2°N 176°E, 10°N 142°W, and 18°S 74°E while Parke-Hendershott show anti-amphidromic points at 6°N 172°E, 8°N 142°W, and 22°S 75°E. Therefore ideal locations for S2 bottom pressure guage measurements would be at 4°N 174°E, 9°N 142°W, and 20°S 74°E. Not surprisingly, two of the three locations are almost identical with those for M2.

For the future, there are three potential directions for improvement in the ability to represent the open ocean tide. All models refer back to known observations either directly through specified elevation boundary conditions, indirectly through adjustment of a friction parameter(s), or simply as a measure of model quality. Unfortunately, coastal observations are on the wrong side of the shelves, and quite often in bays or estuaries. Island observations are often inside lagoons. The growing collection of bottom pressure guage measurements will help. Cartwright and Zetler (personnel communication with Zetler) are collecting present measurements for publication by the International Association for the Physical Sciences of the Ocean (IAPSO). Shelf models of such difficult regions as the Patagonian Shelf will also help. Satellite altimetry measurements will ultimately provide spacial (if not as accurate) coverage of the global tide. In special regions where there are high tidal amplitudes and short length scales, so that the spacial structure of the tide can be separated from orbit errors, existing satellite data can quite possibly be utilized.

Acknowledgements. This paper was developed as part of a research associateship with the National Research Council. The author would like to thank Ron Estes for forwarding a cotidal chart for Schwiderski's model and for forwarding a copy of Accad and Pekeris' preprint.

References


Accad, Y., and C. L. Pekeris, Solution of the tidal equations for the M2 and S2 tides in the world oceans from a knowledge of the tidal potential alone, preprint, Department of Applied Mathematics, the Weizmann Institute, Rehovot, Israel, 1978.

Backus, G., A class of self-sustaining dissipative sperical dynamos, Ann. Phys. (N.Y.),4, 372-447, 1958.

Cartwright, D. E., Oceanic tides, Rep. Prog. Phys., 40, 665-708, 1977.

Estes, R. H., A computer software system for the generation of global ocean tides including self-gravitation and crustal loading effects, X-920-77-82, Goddard Space Flight Center, Greenbelt, Md., 60 pages, 1977.

Farrell, W. E., Deformation of the earth by surface loads, Rev. Geophys. Space Phys., 10, 761-797, 1972.

Goad, C. C., and B. C. Douglas, Lunar tidal acceleration obtained from satellite-derived ocean tide parameters, Jour. Geophys. Res., 83, 2306-2310, 1978.

Gordeev, R., B. Kagan and E. Polyakov, The effects of loading and self attraction on global ocean tides, the model and the results of a numerical experiment, J. Phys. Oceanogr., 7, 161-170, 1977.

Hendershott, M. C., The effects of solid earth deformation on global ocean tides, Geophys. J. Roy. Astron. Soc., 29, 389-402, 1972.

Hendershott, M. C., Numerical models of ocean tides in The Sea, Chapter 6, Goldberg, McCave, O'Brien, Steele eds., Wiley-Interscience, New York, 1977.

Longman, I., A Green's function for determining the deformation of the earth under surface mass loads, Jour. Geophys. Res., 68, 485-496, 1963.

Miller, G., The flux of tidal energy out of the deep oceans, Jour. Geophys. Res., 71 (4), 2485-2489, 1966.

Muller, P., Determination of the rate of change of G and the tidal acceleration of earth and moon from ancient and modern astronomical data, SP43-36, The Jet Propulsion Laboratory, Pasadena, Ca., 24 pages, 1976.

Munk, W.H., and G. MacDonald, The Rotation of the Earth, Cambridge University Press, London, 1960.

Parke, M. E., and M. C. Hendershott, M2,S2,K1 models of the global ocean tide on an elastic earth, Marine Geodesy, in press, 1978.

Parke, M. E., Global modelling of tides on an elastic earth, Proc. Int. Long Wave Symposium, Ottawa, Canada, 1978a.

Parke, M. E., Global numerical models of the open ocean tides M2,S2,K1 on an elastic earth, PhD Thesis, Univ. of California at San Diego, 1978b.

Schwiderski, E. W., Global ocean tides, part I: a detailed hydrodynamical interpolation model, NSWC/DL TR-3866, Naval Surface Weapons Center, Dahlgren, Virginia, 1978.

Stock, G., Modelling of tides and tidal dissipation in the Gulf of California, PhD Thesis, Univ. of California at San Diego, 1976.

Zahel, W., The diurnal K1 tide in the world ocean - a numerical investigation, Pure and Applied Geophysics, 109, viii, 1819-1825,1973.

Zahel, W., Proc. IRIA Int. Colloq. on Numerical Methods of Science and Technical Computation, Berlin: Springer-Verlag, 1977.

Page Intentionally Left Blank

# What Can Earth Tide Measurements Tell Us About Ocean Tides or Earth Structure?

T. F. Baker
Institute of Oceanographic Sciences, Bidston Observatory
Birkenhead, Merseyside, England

Abstract. A brief review is given of the current experimental problems in Earth tides using comparisons of tidal gravity and tilt measurements in Europe with loading calculations as examples. This review shows the limitations of present day instrumentation and installation techniques and some of the ways in which they can be improved.

With these limitations in mind, we go on to a general discussion of many of the geophysical and oceanographic investigations that are possible with Earth tide measurements. In particular, we concentrate on the percentage accuracies required in the measurements in order to obtain new information about the Earth or Oceans.

## 1. Introduction

In recent years there has been a very significant advance in the quality of Earth tide instrumentation. Various gravimeters, tiltmeters and strainmeters are now available which, for any tidal constituent between one cycle per day and four cycles per day, give a signal/noise ratio which is comparable with, and in some cases significantly better than, that obtained with most ocean tide instrumentation. However, due to certain experimental difficulties, which are discussed in Section 2, there has not been the same degree of advance in the interpretation of the recorded signal.

On the theoretical side there have recently been some very important advances. The solution of the loading Green's function problem for a spherical, radially stratified, gravitating, elastic Earth model by Farrell [1972] and the introduction of the finite element method to model more complex local and regional structures, [Beaumont and Berger 1974, Harrison 1976 and Berger and Beaumont 1977] have now made it possible to attempt a realistic interpretation of the recorded signal.

Despite all the above advances in solving the experimental and theoretical problems of Earth tides, progress in using the Earth tide signal to obtain useful geophysical and oceanographic information has been slow and we are still at a very preliminary stage. Progress has been limited by the efforts required to solve the experimental problems discussed in the next section. Clearly, it is important to briefly review these experimental limitations before going on, in the later sections, to discuss the feasibility of some of the geophysical and oceanographic objectives.

An examination of the literature reveals essentially four stages in the development of the study of Earth tides, each associated with a different level of interpretation of the recorded signal. At the first stage there is the publication of a list of experimental tidal parameters with no interpretation other than a rough comparison with the parameters expected for an oceanless, elastic Earth and the noting of any large unexpected anomalies. The second stage is a general comparison of experimental results (usually $M_2$ and/or $O_1$) with a body tide and load tide calculation for a given standard Earth model and a single ocean tide model. The third stage is a comparison of the measurements with a range of possible given seismic Earth models and/or a range of possible given ocean tide models and thereby choosing the model(s) which give the best fit to the data. The ultimate objective is the fourth stage, the actual inversion of the Earth tide measurements in order to obtain improved models of the Earth structure or of the ocean tides. Clearly, we are particularly interested in the results from stages three and four, but unfortunately the majority of published papers are in the first two categories.

Before going on to review some of the recent developments, it is important to mention two more secondary objectives of tidal investigations. Firstly, there is an increasing demand for body tide and load tide parameters as corrections to geodetic measurements as the accuracy of these measurements increases. Corrections are required for satellite altimetry, laser ranging, VLBI, first order geodetic levelling and microgravimetric surveys. Generally corrections at the 1 centimetre and 1 microgal level are required and the above experimental and theoretical developments are now beginning to make this possible. Secondly, the improvement in instrumentation has led to increasing interest in the difficult problem of recording and interpreting long period tilt and strain signals as precursors to earthquakes. Here the resolution of a tidal signal of roughly the predicted amplitude gives some of the necessary assurance that the instrument is correctly coupled to the Earth and producing meaningful long period signals.

## 2. The Experimental Problems

Figures 1 and 2 illustrate the current experimental problems of Earth tides in a convenient and concise way. The $M_2$ observed load values are calculated by subtracting the theoretical body tide from the observations (assuming a tilt diminishing factor of 0.69 and a gravimetric factor of 1.160 with zero phase lag). The tilt observations are taken from Melchior [1976], Ostrovsky [1976] and Lecolazet, Steinmetz and Wittlinger [1970]. The gravity observations are from Melchior, Kuo and Ducarme [1976] and Baker [1977]. The contours give the computed $M_2$ tidal loading amplitude and phases for Europe. The Farrell [1972] Green's function for a Gutenberg-Bullen A Earth model has been used and convolved with the Hendershott $M_2$ world ocean tide model. It is important to

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

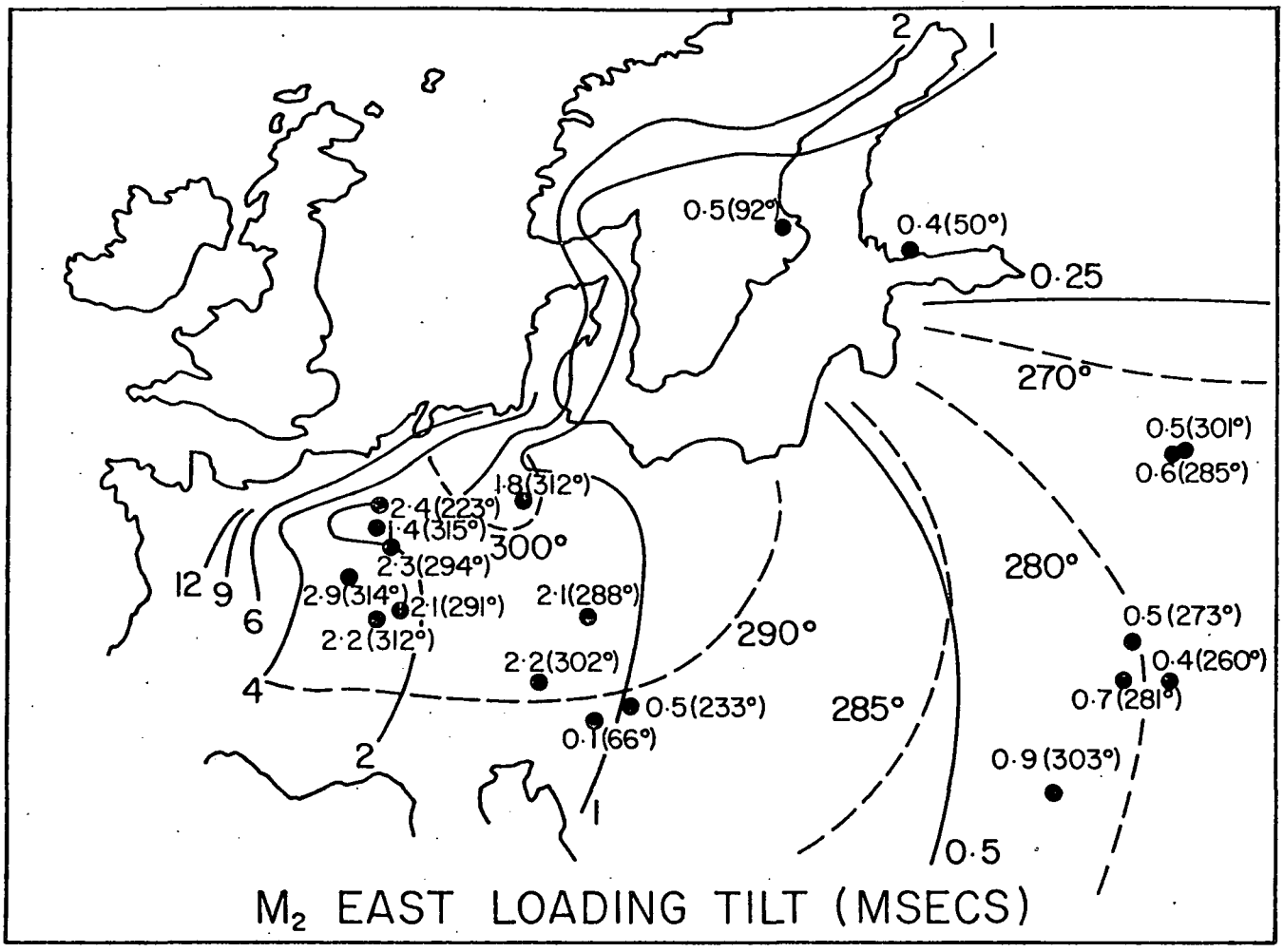



Fig. 1. The observed and calculated $M_2$ east loading tilt in Europe. The contours show the calculated load distribution using the models described in the text. Amplitudes are in milliseconds of arc and the phase lags are with respect to the tidal potential in the Greenwich meridian.

accurately model the large amplitude tides in the adjacent shelf seas and therefore, for the seas around the British Isles, the Hendershott model has been replaced by a more detailed hydrodynamical model of the $M_2$ tide [Flather 1976]. It should be noted that the Arctic Ocean and the small tides in the Baltic and Mediterranean have not been included. (For the stations immediately adjacent to the Adriatic and Mediterranean Seas the effect of neglecting these loads can be estimated from the calculations of Chiaruttini [1976]. For the two stations nearest the Adriatic we should subtract about 0.15 microgals from the amplitudes and add about 7° to the phases of our calculated loads.) Despite these limitations the calculations give a reasonable approximation to the load signal and in particular give an indication of the spatial variability of the signal. (It should be noted that for the tidal gravity map the uncertain contribution from distant oceans gives an uncertainty in amplitude and phase equivalent to perhaps ±0.5 microgals which is fairly uniform over the geographical area).

Generally there is an overall agreement between theory and experiment. A detailed inspection shows however that, for both the tilt and gravity, there is a large variability over short distances which is inconsistent with the expected spatial variation of the loading signal. It is this scatter of the observed data that is limiting the progress in interpretation of the Earth tide measurements in different areas of the World. The error limits as calculated from the residual spectra are usually less than 0.1 msec in tilt and less than 0.1 microgals in gravity. Therefore, systematic perturbations or systematic experimental errors are present in the data.

In the case of the east-west tilt results, strain induced tilt (strain-tilt coupling) perturbations are present due to the cavity, the topography and the local geology (see for example, Harrison 1976). The east-west rather than the north-south tilt results have been plotted for two reasons. Firstly, the load tilt is of the order of 2-3 times larger in the east-west azimuth in central Europe and secondly the strain induced tilt perturbations are usually less in this azimuth due to the small east-west $M_2$ body strain in these latitudes.

The typical scatter of the results in West and

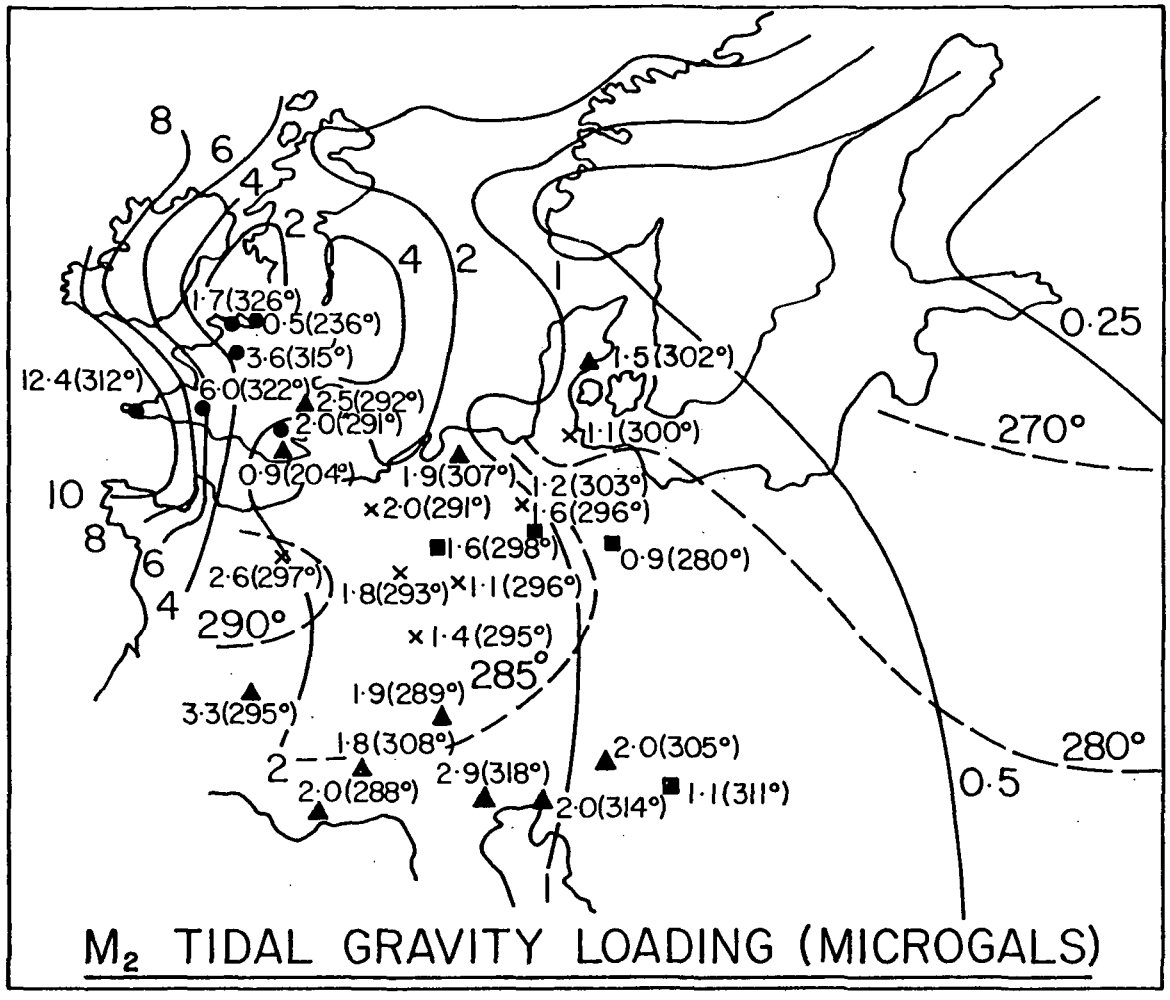



Fig. 2. The observed and calculated $M_2$ gravity loading in Europe. The contours show the calculated load distribution using the models described in the text. Amplitudes are in microgals and the phase lags are with respect to the tidal potential in the Greenwich meridian. ● LaCoste-Romberg tidal nulled gravimeters; ▲ Geodynamics TRG-1 gravimeters; ■ Askania gravimeters; × mean of various gravimeters.

Central Europe is of the order of ±0.5 msec (considering both in-phase and quadrature components) which is in fact only ±6% of the total $M_2$ east-west signal (8-9 milliseconds). The extreme case is an anomaly of 16% of the total $M_2$ signal. These results are all in mines and tunnels.

The typical scatter of results in the Soviet Union about the mean value is only ±0.2 msec (±3% of the total $M_2$ signal). The Ostrovsky tiltmeters are installed at the bottom of specially constructed shafts 10-20 metres deep and 1-1.5 metres in diameter, in either bed-rock or in unconsolidated sediments. The small scatter is probably due to the symmetrical cavity. (It should be noted, however, that in the north-south direction the percentage scatter of the observations is 2-3 times larger). These results, together with those of Beaumont [1978] using metal vaults installed on the bed-rock at a depth of 5-6 metres, suggest that with near surface types of installation it may be possible to reduce unknown local strain induced tilt perturbations to less than 5% of the signal. However, these installation techniques reduce the strain induced tilt perturbations at the expense of increased noise, particularly in the diurnal band. Future measurements with long baselength tiltmeters and strainmeters in trenches may give better results over a wider range of frequencies. The measurements of Berger and Wyatt [1973] with an 800 metre surface laser strainmeter and of Michelson and Gale [1919] with a 150 metre trench tiltmeter show that longer baselengths should give the desired improvement.

Boreholes offer another possible solution to the tilt perturbation problem. The comparison of the signals from 3 adjacent boreholes by Zschau [1976] is very encouraging since for $M_2$ agreement was obtained to ±1% for the east-west azimuth. However, as yet, there are too few published results with which to make a complete assessment of boreholes.

These new developments, as well as reducing the cavity perturbation, allow a more flexible choice of site for geophysical investigations and, where possible, for avoiding the complex topography and geology often associated with mines.

The scatter of the gravity results in figure 2 cannot be explained by local perturbations such

as the cavity, since the effects on gravity are negligible. The typical scatter of the results in Central Europe is ±0.3 to ±0.4 microgals which is of the order of ±1% of the total $M_2$ signal (30-45 microgals in these latitudes). These errors, however represent about ±20% of the $M_2$ load signal. This scatter is due to the uncertainty in the amplitude and phase calibrations of different instruments. Clearly we have an uncertainty of the order of ±1% in amplitude calibration and an uncertainty in the instrumental phase lags of the order of 1°. There are differences in the manufacturers calibrations and also in the calibration procedures of each research group. Also taking the mean value of several types of instrument at a single station leads to further confusion. Clearly we must aim for all instruments in an area to be calibrated relative to each other to an accuracy of better than ±0.2% in amplitude and ±0.1° in phase, so that the loading signal can then be defined to better than 5% accuracy. This can only be achieved by inter-calibrating gravimeters at regular intervals at fundamental base reference stations in each geographical area and carefully monitoring any changes in instrumental parameters during profile measurements.

In this context the measurements of Torge and Wenzel [1977] provide an interesting example. They have carefully compared most of the major types of modern gravimeter at Hannover for several months and also examined the instrumental phase lags and damping factors as functions of frequency. At the frequency of $M_2$ they have found instrumental lags of up to 2° and damping factor corrections of up to 2.5%. They have also found a wide range of noise levels for the different gravimeters as determined by the residual variance after tidal analysis.

This brief review of the experimental problems of tilt (strain) and gravity measurements shows that many problems still remain, but significant work is now in progress towards their solution. The importance of these experimental limitations will become clear in the later sections.

## 3. The Body Tide and Earth Structure

It is now well established using a range of possible seismic models that the global body tide Love numbers are only uncertain through a range of 1 or 2%. In view of the above experimental problems and the uncertainties in the ocean tide, it is therefore reasonable to assume in most work that the global Love numbers are known parameters.

The phase lag of the body tide is less well established. It is usually inferred indirectly from free oscillation Q values [Lagus and Anderson 1968]. It has often been stated that, in principle, upper bounds on the phase lag could be determined from very accurate measurements in areas where the load is small, such as in the centre of continents. However, since the expected phase lag is very small, there are severe problems due to both the experimental difficulties discussed previously and the uncertain load contributions from distant oceans.

As an example, consider the Russian tilt results in figure 1. In that area the load tilt is roughly in quadrature with the body tide. Therefore, the mean of the observed amplitudes can be brought into better agreement with the load calculations by introducing a phase lag in the $M_2$ body tide tilt of between 1° and 2°. This phase lag is unreasonably large, therefore a more likely explanation is either a small uncertainty in the load calculations or a small phase lag (1-2°) in the Ostrovsky tiltmeters.

A problem that has received a lot of interest in the last few years is the measurement of the core resonance in the diurnal band (near $\psi_1$) arising from the interaction of the liquid outer core with the elliptical rotating mantle. The effects of the resonance have been established for several years now by observing a decrease of the $K_1$ gravimetric factor relative to $O_1$ of the order of 1% and the increase of the $K_1$ tilt diminishing factor relative to $O_1$ of the order of 7% (see for example Lecolazet and Melchior 1977). Lecolazet and Melchior report a variation in the amplitude of $\psi_1$ from gravimeter data greater than the errors determined from the variance of the diurnal residuals. They suggest a possible variation of the resonance frequency. Recently, the resonance has been investigated using 2 years data from a laser strainmeter [ Levine 1978 ] and $1\frac{1}{2}$ years data from the superconducting gravimeter [ Warburton and Goodkind 1978 ].

Although the presence of the resonance is now well established, investigations of the detailed structure (frequency and Q value) clearly require very long series of high quality observations with careful examination of ocean loading and meteorological perturbations.

## 4. Earth Structure from Tidal Loading Measurements

Due to tidal loading from the adjacent sea area the tidal tilt signal within 50 kilometres of the coastline is often observed to be several times larger than the theoretical body tide tilt. If one, or more, tidal constituents in the sea area are known to a few percent accuracy from oceanographic measurements, then it may be possible to determine some parameters concerning the crustal or upper mantle structure from the tilt observations. Similarly, the tidal strain signal is larger adjacent to a sea area, although usually by a smaller factor with respect to the body tide than is the case for the tilt signal. The main difficulty has been in finding suitable sea or ocean areas where it can be assumed that the main marine tidal constituents are known to a reasonable accuracy.

Beaumont and Lambert [1972] have calculated the tilt loading Green's functions for several laterally homogeneous crustal and upper mantle models. They examine the ways in which tilt loading studies differ from or complement, seismic studies. There are three major differences: the loading signal depends upon the Earth's elastic properties averaged over an area rather than averaged along the path of the seismic wave; the loading signal is essentially independent of density; and the frequency of the tidal loading signal is several orders of magnitude lower than for the seismic signals. Beaumont and Lambert conclude that the loading tilt essentially gives information on the Young's modulus, and they suggest the

inversion of load tilt data together with seismic and gravity data in order to give improved estimates of the density and elastic parameters in layered models of the crust and upper mantle.

Continuing this work Beaumont [1978] and Beaumont and Boutilier [1978], using measurements from 5 tilt stations, have confirmed, using different seismic models, that the most appropriate model for the crust beneath Nova Scotia is that given by a seismic refraction profile along the coast of Nova Scotia. The work shows the importance of using tilt differences to eliminate the uncertain tides from distances greater than 500 kilometres. The differential tilt signal then depends upon the crustal structure and the nearby shelf tide, which for $M_2$ is defined to an accuracy of 2% in amplitude and 2° in phase using ocean bottom tide gauges. From the comparisons of the observations with the model results they also conclude that, except for an anomaly in the borehole results at Bedford, Massachusetts, the strain induced tilt perturbations are reasonably small.

Baker [1977], using tilt observations in a mine at Llanwrst, North Wales, finds that a crustal model from seismic investigations in the Irish Sea with a three layered, 30 kilometre thick crust, fits the observations better than the standard Gutenberg Bullen Earth model with a 38 kilometre crustal thickness. Measurements at different points within the mine, however, show that even in the tunnel azimuth there are differences in the observed tilts of the order of 10% due to strain induced tilt pertubrations which make conclusive geophysical interpretation difficult.

Using the tilt Green's function for the above Irish Sea layered model, numerical experiments have been performed in order to investigate the sensitivity of the tilt Green's function to variations in the individual model parameters (P and S wave velocities, density, Young's modulus and Poisson's ratio) of the main crustal layer (depth 4 kilometres to 24 kilometres). A change in the P wave velocity of 5% gives a maximum change in the Green's function of the order of 5% but the Green's function is affected over distances up to 2 or 3 times the depth of the crustal layer. The range of distances involved illustrates the non-uniqueness of the inverse problem.

Other tests have involved the variation of the depth of the Moho, introduction of a crustal low velocity layer, the removal of the uncertain lower crustal layer and variations of the upper mantle parameters. In all these tests it has been assumed that the main crustal parameters have already been determined by explosion seismology, but the parameters are allowed to vary through their possible range of uncertainty in order to investigate how the inversion of tilt loading data can give additional information on the structure. Although it is found that the parameter variations give changes in the Green's function of up to 10% or more, the convolution of the Green's function over a sea area reduces the percentage change in the total tilt signal, due to the load contribution from distances where the Green's function is unchanged. Differential tilt measurements at a suitable separation will, of course, give a larger percentage change than a single measurement.

The convolutions for Llanrwst (16 kilometres from the Irish Sea) show that, although there is a difference of 5% between the predicted tilts with the Gutenberg Bullen model and the local seismic model, reasonable parameter variations of the seismic model only change the predicted tilt by typically 2-4%. Such small changes are difficult to detect for two reasons. Firstly, the uncertainties in the $M_2$ shelf tide are often of this order. Secondly, as we found in Section 2 it is difficult to guarantee that there are no unknown strain induced tilt perturbations at the 2-5% level which can bias the geophysical conclusions. A large number of tilt measurements in an area with careful topographic and local geological finite element model corrections would be required.

Where there is no previous seismic information available, or where there are large uncertainties in the structure, then tilt measurements can be used to determine the main crustal parameters. From similar tests carried out on the strain load Green's function, it is found to be of the order of 3 times more sensitive to the crustal parameters than the tilt Green's functions. It is therefore suggested that strainmeters (preferably in trenches) may provide a better means of investigating the laterally homogeneous layered Earth structure adjacent to a known tidal load (unfortunately, however, strain is susceptible to coupling perturbations over a wider spatial range than is the case for tilt; Harrison [1976]).

The load tilt signal may also provide useful ocean tide information in the first few hundred kilometres away from the coastline, if the principal crustal and upper mantle parameters are already known from seismology (see Section 6).

In all the above work we have assumed that the Earth responds elastically to the load. Zschau [1977] has pointed out that in some areas the anelastic response may be important and thus there may be a phase lag in the response. For the Irish Sea area, the anelastic response appears to be small. For the Llanrwst tunnel azimuth, 16 months Askania tiltmeter observations give an $M_2$ phase lag with respect to the elastic load calculations of $0.1^\circ \pm 0.2^\circ$. From the preliminary analysis of three months observations from the Askania in a borehole in the Lake District [Baker, Edge, Jeffries 1977] an $M_2$ phase lag of $-0.7^\circ \pm 0.3^\circ$ with respect to the elastic load calculations is obtained. (It should be noted, however, that the uncertainty in the Irish Sea $M_2$ tide is probably of the order of $\pm 1^\circ$.) The largest phase lags are predicted at ocean ridges where Zschau gives a tilt phase lag Green's function of nearly $5^\circ$ in magnitude. If the phase lag after convolution is sufficiently large, it may be possible to detect the anelastic response using differential tilt measurements on islands on the mid-Atlantic ridge now that the tides are becoming more precisely defined from an array of temporary ocean bottom tide gauges on the mid-Atlantic ridge [Cartwright, 1977].

## 5. Lateral Heterogeneities in Earth Structure

In the above discussions it has been assumed that laterally homogeneous layered Earth models

are appropriate in a given area. Beaumont and Berger [1974] have shown that the tidal tilt and strain signals differ significantly from their normal values (i.e. the values on a laterally homogeneous Earth) for measurements in the vicinity of a vertical interface or discontinuity between two different elastic media. For example, they find an increase of tidal tilt of 50% over the normal tilt amplitude for a 15% contrast in P wave velocity across an interface. The effects can be detected for a distance away from the discontinuity which depends upon the depth of the discontinuity. Clearly, since the effects on the tilt and strain signal are significantly larger than those found for the changes in the layered models discussed in Section 4, then the possibilities of obtaining useful information on Earth structure are significantly larger. Experiments to investigate such discontinuities have recently commenced (e.g. Baker, Edge, Jeffries 1977, Grosse-Brauckmann, Herbst and Rosenbach 1977, Zschau and Gerstenecker 1977). One important fundamental question remains to be answered however. Since vertical discontinuities in elastic parameters are normally associated with past or present tectonic activity, the associated areas are usually very complex in terms of their geology (and topography). This may lead to significant unknown coupling perturbations over a wide range of scale lengths. Is it possible to separate the desired signal from these perturbations?

Another interesting possibility arising from a lateral contrast of elastic structure is discussed by Beaumont and Berger [1974]. This is the time variation in the tidal amplitudes due to changes in elastic parameters in a fault zone prior to an earthquake. Such precursory effects should be very easy to detect since local coupling perturbations of the regional signal are not important and all that is required is a stable calibration. Latynina and Rizayeva [1975] have investigated a possible 6% change in the observed $M_2$ strain before an earthquake in the Soviet Union.

## 6. Ocean Tides from Tidal Loading Measurements

The gravity loading Green's function for a layered Earth model is relatively insensitive to reasonable variations in the elastic parameters if the measurements are at a distance greater than 10 kilometres from the coast so as to avoid the effects of the various near surface elastic layers. Finite element calculations by Zurn, Beaumont and Slichter [1976] for a subducting lithospheric plate also show that the vertical body tide displacement only differs from its normal value by 0.8% over the edge of the plate. Calculations by Beaumont [1978] show that the gravity loading signal can be modified by at most 10% due to the lateral change of crustal thickness at a continental margin.

For most gravity measurements it is safe to assume that the Earth structure is well known and to use the tidal gravity loading signal (after subtracting an assumed body tide - Section 3) to examine the ocean tide distribution. Again the normal approach is to check the agreement or disagreement between the tidal gravity observations and one or more ocean tide models (see for example Warburton, Beaumont and Goodkind 1975, Beaumont 1978, Beaumont and Boutilier 1978, Baker 1977). Beaumont [1978] using both tilt and gravity observations finds that an $M_2$ ocean tide intermediate between Tiron et. al. and Zahel's models is required for the North-West Atlantic. This conclusion is later verified with loading calculations using a new $M_2$ map based on several ocean bottom tide gauges and a hydrodynamical model for the Bay of Fundy - Gulf of Maine [Beaumont and Boutilier 1978]. Baker [1977] finds that the Hendershott $M_2$ model for the North East Atlantic gives a much better agreement with the tidal gravity measurements in the British Isles than the Pekeris and Accad or Zahel $M_2$ models.

Unfortunately ocean tidal maps rarely have any estimates of uncertainty attached. Differences between the maps for the major constituents of the order of 50% in amplitude and 60° in phase (or more) are found in some areas of the world. It is however important to distinguish between two types of ocean tide model. Some models [Hendershott, Bogdanov and Magarik and Tiron et. al.] constrain the tides to fit coastal tide gauges, whilst others simply allow a no flow condition across the coastlines. It is therefore perhaps not surprising that the former often give better agreement with Earth tide measurements since in the near loading area they are constrained to give a first approximation to the ocean tide.

In the introduction we discussed four stages in the development of the study of Earth tides. The third stage was the comparison of the observed signal with a range of possible Earth and/or ocean models. This is the stage of development of the work discussed in the previous sections and also for most of the published work on tidal gravity.

Jachens and Kuo [1976], Kuo and Jachens [1977] and Kuo, Jachens and Lee [1977] have, however, progressed one stage further by actually inverting their tidal gravity data in order to obtain improved models of the ocean tide distribution. In the last few years they have carried out an intensive programme of measurement in North America and Europe using Geodynamics TRG-1 tidal gravimeters.

Jachens and Kuo [1976], using a 'trial and error' method, obtained a new $O_1$ cotidal chart for the North Atlantic that is in better agreement with their tidal gravity observations than the starting model of Tiron et. al. This trial and error inversion is later confirmed by using a Lagrangian multiplier inversion method. Kuo and Jachens [1977] and Kuo, Jachens and Lee [1977] then go on to use the methods of linear programming in order to find the coefficients of a fourth order two dimensional polynomial which is used to correct an initial ocean tidal chart. Using 17 tidal gravity stations and 62 coastal and island tide gauges they have calculated new $M_2$ and $O_1$ tidal maps for the North East Pacific. Similarly using 25 tidal gravity stations and 90 tide gauges they obtain a new $M_2$ chart for the North Atlantic. No indication is given concerning the errors or the uniqueness (resolution) of the resulting solutions. Instead the new tidal maps are tested using a few ocean bottom tide gauge measurements that were not included in the original inversion. For $M_2$ in the North East Pacific the Tiron model has large phase

errors (50°-80°) with respect to the ocean bottom measurements, which are reduced to less than 5° in the inversion solution. For $O_1$ in the NE Pacific errors of the order of 20% and 20° in the Tiron map with respect to the bottom tide gauges are reduced to less than 3% and 1°. For $M_2$ in the N.Atlantic errors of up to 25% and 15° in the Tiron map are reduced to the order of 5% and 5°.

Beaumont and Boutilier [1978] find that the N.Atlantic $O_1$ chart of Jachens and Kuo is inconsistent with their tiltmeter measurements and they suggest that the gravimeters from which the chart is derived are subject to instrumental phase lags and calibration errors.

Baker [1977] has used measurements from 8 tidal gravity stations in the British Isles to test the feasibility of the inverse ocean tide problem (see figure 2). Two LaCoste-Romberg tidal nulled gravimeters were used together with the results from a Geodynamics gravimeter [Melchior, Kuo and Ducarme 1976]. In order to ensure that their relative amplitude and phase calibrations were known, measurements were made with all 3 instruments at Bidston. For the $M_2$ load convolutions the Hendershott model was used together with a more detailed $M_2$ shelf tide numerical model for the seas around the British Isles [Flather 1976]. Comparisons with coastal tide gauges show that the typical errors in the shelf model are 10% in $M_2$ amplitude and 10° in $M_2$ phase. The $M_2$ tide is therefore relatively well known and this allows a useful test of the capabilities of tidal gravity.

It is found that the non-uniqueness of the inverse problem can be reduced by considering pairs of gravity stations. If the stations are suitably situated then a large (2 microgals) tidal gravity difference arises from the adjacent sea area whilst the contributions from other sea areas and distant oceans are considerably reduced. Using pairs of stations average correction factors were found for the Celtic Sea, Irish Sea and east English Channel $M_2$ tides. Similar results were also obtained from a least squares adjustment of the tides in the 3 sea areas. Adjustments to the tides in the other sea areas by reasonable amounts that are consistent with the oceanographic uncertainties does not affect the solution. Within the errors determined from the residuals of the least squares adjustment, the results are consistent with the known shelf model uncertainties determined from comparisons of the model with coastal tide gauges [Flather 1976]. In particular the gravity measurements require an adjustment to the amplitude of the east English Channel $M_2$ tide of 15% which is consistent with the tide gauge comparisons.

These results show that tidal gravity measurements can be inverted in order to obtain better tidal maps in ocean areas where the tides are less well known. Two important points are illustrated by the work. Firstly, the choice of sites is of importance in reducing the non-uniqueness of the inverse problem. Gravity stations should be chosen such that a reasonably large differential tidal gravity between pairs of stations is obtained from the ocean area of interest. Other ocean areas contributing significantly to the differential signal must be constrained either by oceanographic information or another gravity measurement. With the typical noise level of modern gravimeters (of the order of 0.05 microgals internal error in the semi-diurnal band from a few months observations), a differential gravity signal of at least 1 microgal is required in order to obtain useful ocean tide information.

The second, and perhaps more important point, is the problem of relative amplitude and phase calibrations. Since the tidal gravity load signal is only a small fraction of the body tide signal (in most cases less than 10%) then the problem is ill-conditioned since a small calibration error represents a large error in the observed load tide. This problem was discussed in Section 2 with respect to figure 2, where it was concluded that all instruments in an area should be inter-calibrated at a reference station to better than ±0.2% in amplitude and ±0.1° in phase. It should be emphasised that for most problems only relative calibration of the gravimeters is important. Absolute calibrations to these accuracies are still very difficult to achieve. A small uniform error in the amplitude or phases of the observations in a gravity profile has a similar effect on the tidal gravity inversion as the effect of a small error in the assumed body tide or of an error in the calculated contributions from distant oceans.

It should be noted that the problems should be less severe for observing the diurnal loads near the Equator or the semi-diurnal and diurnal load tides near the Poles where the body tide contributions are much smaller.

It is also important to note that the required accuracy for tidal gravity implies that great attention must be paid to the tidal analysis procedures [Baker 1978 a, Yaramanci 1977 and 1978]. With different analysis methods applied to tidal gravity data, it is found that differences in amplitude of 0.5% for the major constituents can arise from incorrect use of these methods.

Even though the modern gravimeter has extremely high total signal/noise ratio the useful signal/noise ratio is far less. For this reason the limitation to gravity differences of greater than 1 microgal was suggested above. This however implies that tidal gravity can make very little contribution to defining the ocean tide for the smaller constituents or in areas where the loading from the major constituents is very small. For the small constituents tidal tilt has advantages over gravity, since within 100 km of the coastline the load tilt is comparable with, or very much larger than, the body tilt. The useful ocean tide information obtained would probably be limited, however, to the first few 100 kms away from the coastline. Of course, strain induced tilt perturbations must be reasonably small ($\lesssim$5% of the amplitude of the constituent) and the overall crustal and upper mantle parameters (in particular the depth of the Moho and average crustal seismic velocities) must be known from seismic information (see Section 4).

Inversion of tidal gravity and tilt measurements can make a contribution to the problem of mapping the ocean tides provided attention is given to all the above experimental problems. The main contribution will probably be near 'anti-amphidromes' where the amplitude and phase of the ocean tides are such that the gravity loading signal is particularly large. The spatial averages given by loading measurements give complementary information to

oceanographic measurements. Programmes to measure ocean tides with ocean bottom pressure sensors have now commenced, but due to financial and other constraints the progress is slow [Cartwright 1977]. The additional information provided by loading measurements and eventually the developing satellite technology will help in giving the necessary global coverage.

Finally, a mention should be made of some other oceanographic problems on which loading measurements may give some information. Warburton and Goodkind [1978] have observed unexplained time variations in the amplitudes and phases of the major waves on their super-conducting gravimeter. They suggest possible variations in the ocean tides. Seasonal modulations of $M_2$ have been observed in short analyses of tilt at Llanrwst arising from the loading of the shallow water non-linear waves in the $M_2$ group [Baker 1978 b, Yaramanci 1978]. Also for measurements adjacent to a tidal loading area, the residuals and residual spectra after removing the tides may contain some interesting oceanographic information. Although there are problems with measurements in mines regarding the systematic perturbations of the signal, clearly the relatively high signal/noise ratio has some advantages for these types of problems where small signals are of interest. Measurements in mines and tunnels can also be used where the effects of perturbations of the absolute signal magnitude are of less importance. For example, the tilt loading signal can be used to examine the shape of the response function of a sea area on the assumption that the perturbing load strain has the same response function as the load tilt. Measurement of the average response function for a sea area obtained from load measurements would be relatively free of any non-linear effects that are local to an individual estuary or tide gauge.

## 7. Concluding Remarks

It has been shown that significant results are being obtained from the comparison of Earth tide measurements with a range of possible seismic and ocean tide models. However, in order to progress to the inversion of the measurements so as to obtain new models of the Earth's structure or of the ocean tides, various experimental errors must be reduced still further. Progress is already being made in this direction and some of the remaining problems on which future work is necessary have been emphasised.

## References


Baker, T.F., Earth tides, crustal structure and Ocean tides, Proceedings of 8th International Symposium on Earth tides, Bonn, 1977.

Baker, T.F., A review of the objectives of tidal analysis, Bul. Inf. Mar. Terr., 78, 4571-4578, 1978a.

Baker, T.F., Non-equilibrium influences on the tidal signal, Bul. Inf. Mar. Terr., 78, 4596-4610, 1978b.

Baker, T.F., R.J. Edge and G. Jeffries, Notes on a new experimental tilt programme in the English Lake District, Proceedings of 8th International Symposium on Earth Tides, Bonn, 1977.

Beaumont, C., Tidal loading: crustal structure of Nova Scotia and the $M_2$ tide in the north west Atlantic from tilt and gravity observations, Geophy. J.R. astr. Soc., 53, 27-53, 1978.

Beaumont, C., and J. Berger, Earthquake Prediction: modification of Earth tide tilts and strains by dilatancy, Geophys. J.R. astr. Soc., 39, 111-121, 1974.

Beaumont, C., and R. Boutilier, Tidal loading in Nova Scotia: results from improved ocean tide models, Canadian J. Earth Sci., 15, 981-993, 1978.

Beaumont, C. and A. Lambert, Crustal structure from surface load tilts, using a finite element model, Geophys. J.R. astr. Soc., 29, 203-226, 1972.

Berger, J., and C. Beaumont, An analysis of tidal strain observations from the U.S.A. II. The inhomogeneous tide, Bull. Seism. Soc. Am., 66, 1821-1846, 1976.

Berger, J., and F. Wyatt, Some observations of Earth strain tides in California, Phil. Trans. R. Soc. Lond., A274, 267-277, 1973.

Cartwright, D.E., Oceanic tides, Rep. Prog. Phys., 40, 665-708, 1977.

Chiaruttini, C., Tidal loading on the Italian peninsula, Geophys. J.R. astr. Soc., 46, 773-793, 1976.

Farrell, W.E., Deformation of the Earth by surface loads, Rev. Geophys. Space Phys., 10, 761-797, 1972.

Flather, R.A., A tidal model of the north-west European continental shelf, Mem. Soc. R. Sci. Liège, 10, 141-164, 1976.

Grosse-Brauckmann, W., K. Herbst, O. Rosenbach, A tilt profile across the Harz Mountains, Proceedings of 8th International Symposium on Earth tides, Bonn, 1977.

Harrison, J.C., Cavity and topographic effects in tilt and strain measurement, J. Geophys. Res., 81, 319-328, 1976.

Jachens, R.C., and J.T. Kuo, The $O_1$ tide in the North Atlantic Ocean as derived from land-based tidal gravity measurements, Proceedings of 7th International Symposium on Earth tides, 165-175, 1976.

Kuo, J.T., and R.C. Jachens, Indirect mapping of ocean tides by solving the inverse problem for tidal gravity observations, Ann. Geophys., 33 (1/2), 73-82, 1977.

Kuo, J.T., R.C. Jachens and S.S. Lee, The north eastern Pacific $O_1$ and the North Atlantic $M_2$ ocean tides as derived from inversion, Proceedings of 8th International Symposium on Earth tides, Bonn, 1977.

Lagus, P.L. and D.L. Anderson, Tidal dissipation in the Earth and planets, Phys. Earth Planet. Int., 1, 505-510, 1968.

Latynina, L.A. and S.D. Rizayeva, On the changes in the tidal deformations before earthquakes, Izvestiya phys. solid Earth, 11 (9), 607-9, 1975.

Lecolazet, R., and P. Melchior, Experimental determination of the dynamical effects of the liquid core of the Earth, Ann. Geophys., 33, (1/2), 11-22, 1977.

Lecolazet, R., L. Steinmetz, and G. Wittlinger, First results in a campaign to measure tidal tilt in eastern France, Proceedings of 6th International Symposium on Earth tides, Comm. Obs. R. Belg., 9, 23-29, 1970.

Levine, J., Strain-tide spectroscopy, Geophys. J.R. astr. Soc., 54, 27-41, 1978.

Melchior, P., Clinometric stations in Europe, Proceedings of the 7th International Symposium on Earth tides, 41-69, 1976.

Melchior, P., J.T. Kuo and B. Ducarme, Earth tide gravity maps for Western Europe, Phys. Earth planet. Int., 13, 184-196, 1976.

Michelson, A.A., and H.G. Gale, The rigidity of the Earth, Astrophysical Journal, 50, 330-345, 1919.

Ostrovsky, A.Y., Results of observations of tidal tilts of the Earth's surface on the territory of the USSR for the period 1957-1972, Proceedings of 7th International Symposium on Earth tides, 121-126, 1976.

Torge, W. and H.G. Wenzel, Comparison of Earth tide observations with nine different gravimeters at Hannover, Proceedings of 8th International Symposium on Earth tides, Bonn, 1977.

Warburton, R.J., C. Beaumont and J.M. Goodkind, The effect of ocean tide loading on tides of the solid Earth observed with the superconducting gravimeter, Geophys. J.R. astr. Soc., 43, 707-720, 1975.

Warburton, R.J. and J.M. Goodkind, Detailed gravity-tide spectrum between one and four cycles per day, Geophys. J.R. astr. Soc., 52, 117-136, 1978.

Yaramanci, U., The differences between various tidal analysis methods, Proceedings of 8th International Symposium on Earth tides, Bonn, 1977.

Yaramanci, U., A unified approach to signal analysis in Earth tides, Ph.D. thesis University of Liverpool, 1978.

Zschau, J., Tidal sea load tilt of the crust and its application to the study of crustal and upper mantle structure, Geophys. J.R. astr. soc., 44, 577-593, 1976.

Zschau, J., Phase shifts of tidal sea load deformations of the Earth's surface due to low viscosity layers in the interior, Proceedings of the 8th International Symposium on Earth tides, Bonn, 1977.

Zschau, J. and C. Gerstenecker, Non local anisotropy of Earth tidal tilt in the vicinity of a deep going active thrust, Proceedings of 8th International Symposium on Earth tides, Bonn, 1977.

Zurn, W., C. Beaumont, L.B. Slichter, Gravity tides and Ocean loading in southern Alaska, J. Geophys. Res., 81, 4923-4932, 1976.

Page Intentionally Left Blank

# High Precision Tide Spectroscopy

John M. Goodkind
Department of Physics, University of California, San Diego
La Jolla, California 92093

Abstract. Diurnal and long period earth tides have been measured to high accuracy and precision with the superconducting gravimeter. The results provide new evidence on the geophysical questions which have been attacked through earth tide measurements in the past. In addition they raise new questions of potential interest. Slow fluctuations in gravity of order 10 μgal over periods of 3 to 5 months have also been observed and are discussed.

## Introduction

The low noise and drift of the superconducting gravimeter have enabled us to investigate a number of phenomena at new levels of sensitivity. Most of the results described below were obtained from the first 18 months of data at Piñon Flat, California starting in September 1973, and is reported in detail in the literature[Warburton, et al, 1975, 1976, 1977, 1978]. However, we have also begun an analysis of an additional 3 1/2 years of data for which preliminary results on long period tides are discussed here. In the coming year we will begin analysis of records from a total of 6 instruments, four of which are presently in operation.

The geophysical objectives of earth tide measurements have been to measure properties of the interior of the earth by determining such things as the frequency of the core resonance, the phase shift on earth tides, or a frequency dependence of the elastic constants of the earth. In order to meet those objectives, the perturbing effects of ocean tides, the atmosphere, and other environmental influences on gravity must be understood and accounted for in the data. The measurements described here have made new and essential contributions to those objectives. However, I also wish to point out how the new levels of sensitivity can uncover new problems which may ultimately be of more interest than the original objectives.

## Ocean and Atmospheric Loading

Our first investigation of the ocean loading problem is described in Warburton, et al [1975]. An accurate calibration is of crucial importance for this work since, in order to measure the ocean load part of the tide, the theoretical solid earth tide must be subtracted from the measured tide. Our instrument was calibrated, using the direct attraction of a mercury filled sphere, to an accuracy of 0.2%. A computation of expected tidal amplitudes using the best available ocean tide models for $O_1$ and $M_2$ agreed with the observations within the calibration error. The agreement is not proof that the models used were correct or unique but it does demonstrate the level at which ocean models can be tested. They will be tested even more critically when superconducting gravimeters are located at numerous sites around the world.

The atmosphere is the next largest perturbing influence on gravity tides but a quantitative investigation of its effects was not possible without the superconducting gravimeter. The dependence of the pressure-gravity admittance on frequency at tidal and non-tidal frequencies was investigated in Warburton, et al [1977]. The incoherent fluctuations due to weather and local thermal effects show an admittance which increases with decreasing frequency, reaching a maximum value of about 0.3 μgal/mbar at frequencies lower than about 5 cycles/day. At integral multiples of 1 cycle/day the thermally generated atmospheric tides are globally coherent. At up to 4 cycles/day they can be well represented by spherical harmonics as shown by Chapman and Lindzen [1970]. With this representation and the load Love numbers computed by Farrell [1973] it is possible to compute the expected admittances at the first four harmonics. The admittances computed in this way, ignoring the oceans, agree well with the observed values. However, one would expect the oceans to have a substantial effect so that it appears as if the oceans do not respond coherently on a large scale to the atmospheric tides. Additional measurements are required to confirm this effect.

## Anomalous Tides

To some extent the influence of oceans and atmosphere can be eliminated by examining only differences in gravimetric factors and phases between tides which differ in frequency by only one or two cycles per year. If the frequencies are at least a few cycles per year away from 1, 2, 3, or 4 cycles per solar day, then the atmosphere will have no systematic influence. The oceans should not exhibit any resonances sharp enough to make substantial differences on tides separated by only one cycle/year. Therefore, anomalies such as this could be evidence for resonances in the interior of the earth [Warburton, et al 1978] or for a universal preferred reference frame [Warburton, et al 1976], or some other unanticipated effects. The core resonance effect is clearly observable by observation of the relative amplitudes of the $P_1$ and $K_1$ tides. However, precise determination of the frequency and determination of the Q of the resonance depend on measurement of the $\Psi_1$ tide which is very small. In our results this tide appears to be strongly affected by an anomalous ocean tide [Warburton, et al 1978] (which is itself a consequence of the wobble) so that results which are now being obtained at Boulder, Colorado, where the ocean loading is smaller, will be important for understanding this core effect. This inland station will also allow us to determine if the apparent anomalies at $\rho_1$, $M_1$, and $J_1$ [Warburton, et al 1978] were caused by the ocean, and perhaps to set more stringent upper limits on a preferred frame effect [Warburton, et al 1976].

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

Observation of the tides in the 3 and 4 cycles/day bands is especially useful for testing for non-linearities in either the instrument or in the tides themselves [Warburton, et al 1978]. The instrument, when used in electrostatic feedback, had small quadratic and cubic terms in its response function which were measured by this means. The non-linear response of the ocean at $M_4$ led to an anomalous $M_4$ gravity tide at La Jolla, California which was substantially smaller 100 kilometers inland at Piñon Flat. This information along with that currently being obtained at 2 other inland stations in southern California could be used to determine the spatial distribution of the non-linear $M_4$ ocean tide along the coast of California. A peculiar feature of the 3 cycles/day band is that the measured gravimetric factor is close to the theoretical value for the solid earth. This seems to indicate that the ocean loading effect in this band is much smaller than in the 1, 2, and 4 cycles/day bands.

Another surprising feature of the data was some small temporal fluctuations of the tidal amplitudes [Warburton, et al 1978]. In the case of $S_2$ the fluctuations were primarily in the phase rather than the amplitude and the phase fluctuations correlated with the fluctuations in phase of the atmospheric $S_2$ tide. At other frequencies no such obvious identification has, as yet, been possible. Attempts to correlate the fluctuations with fluctuations in the ocean tides at La Jolla and Los Angeles have not been conclusive.

## Long Period Tides

The long period tides are well above the noise level of the instrument even though Piñon Flat is close to the node at 35° latitude, but they yield surprisingly low gravimetric factors. For the fortnightly tide it is between 0.7 and 0.8, and for the monthly tide between 0.85 and 0.92. These two values for each are obtained by making different assumptions about drift as discussed below. (For the 3 year record, which has been analyzed, at present the correlations between least squares fit amplitudes of the various long period tides are finite so that it is not possible to compute simple statistical error bars.) If the low gravimetric factors continue to appear in subsequent data the most likely interpretation will be that the monthly and fortnightly tides on the ocean are far from equilibrium. If this is entirely a consequence of short wavelength responses of the oceans then the data from Boulder will be influenced less by the ocean and show larger gravimetric factors.

The longest period phenomenon which we have investigated is the Chandler wobble with period 436 days. The wobble of the pole results in a change in latitude which in turn changes the centrifugal force at a fixed position on the earth. This results in an apparent periodic variation in gravity of amplitude 2.7 μgal at Piñon Flat. A least squares fit to the data, either the IPMS data on polar motion, or a sine wave at 436 day period, leads to the same result. In both cases the resulting least squares fit amplitude is ∿6 μgal. If all other terms which are being fit simultaneously are held fixed and the frequency of the Chandler wobble term is swept, the resulting amplitude for the wobble term is independent of frequency between 0.76 and 0.84 cycles/year. This indicates that for this record we are measuring some broad band phenomenon other than the Chandler wobble.

## Secular Variations of Gravity

The greatest potential contribution of this instrument to geodesy is its capability for measuring the non-periodic, long term changes in gravity which could result from tectonic uplift or subsidence. The data which is discussed here was obtained with an early version of the instrument without some additional stabilizing coils which are now in use. During the first 18 months there was a significant linear drift with slope 139 μgal/year obtained by least squares fit. At the end of this period the instrument was shut down for about two weeks. When it was started up again it was "annealed" by raising the temperature and the magnetic field above the final operating values. For the subsequent 18 months the slope was 4.8 μgal/year. Figure 1 shows the residual signal from this three year record after removal of all tides, atmospheric effects, and the linear drifts. The peak-to-peak variations are approximately 10 μgal and take place over periods of 3 to 5 months. Thus, an apparent linear drift of 4.8 μgal/year over 18 months is probably a consequence of these 3 to 5 month, non-monotonic, variations.

There are theoretical and experimental reasons for expecting the drift of the instrument to be a logarithmic function of the time. [Prothero, et al 1968] However, attempts to fit a logarithmic term

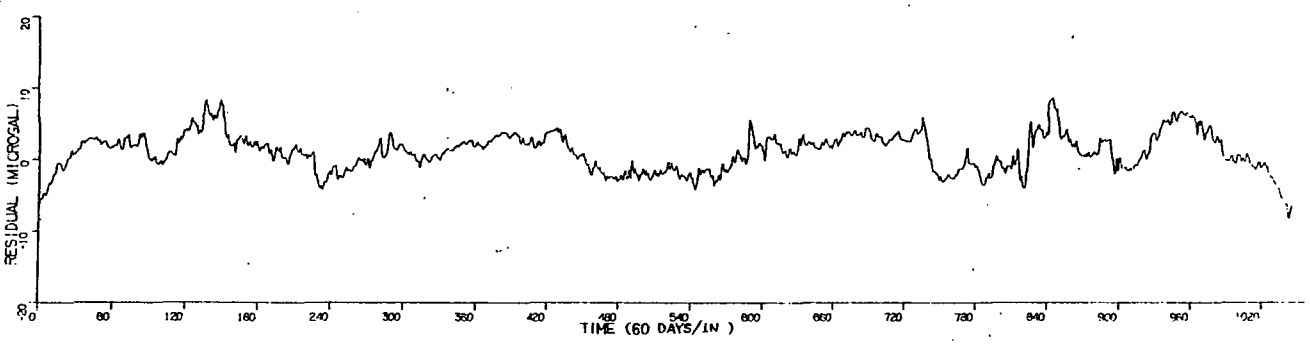



Fig. 1. Gravity residual for a 3 year record.

in place of the linear term left a large linear drift on the first 18 months of data. Including both linear and logarithmic terms in the fit resulted in a lower variance of the residual, but does not substantially alter the linear drift term. Thus the drift appears to be well accounted for by a linear term. For this record, with its two different linear drift rates, the amplitude of the long period tides and the Chandler wobble depends on whether the logarithmic term is included. The last two years of data has not yet been reduced, but since it was obtained with an instrument which had stabilizing coils, there should be no ambiguities due to instrument drift.

Conclusive evidence about instrument drift must come from two instruments run side-by-side. For approximately three months the first instrument with stabilizing coils was run alongside the instrument which obtained 3 years of data at Piñon Flat. The residual signals from each,after subtraction of tides, revealed a correlated change in drift rate during that time. [Goodkind 1979] However, the mount for the new instrument in those tests was tilting so as to generate an artificial signal. We hope to make a more valid side-by-side test of drift rate during the coming year.

## Future Prospects

At present 4 instruments are in operation at Piñon Flat, Lytle Creek and Otai Mountain in California, and at Boulder, Colorado. Two more will be placed in the field in October and November at Greenbelt, Maryland and The Geysers, California. Thus we will have records from six different locations to compare during the coming year. This will allow us to improve the statistics on the Chandler wobble by stacking the records. It will also allow us to investigate regional variations in all of the phenomena discussed above. The influence of the oceans should be revealed by differences between records in Boulder and California which do not appear between the 3 records in Southern California.

The addition of our own barometric pressure gauges, recorded simultaneously with the gravity data, will allow better removal of pressure effects from the data. We are also installing our computer controlled data systems which can communicate two ways over the telephone with our laboratory minicomputer. This will ensure records with fewer interruptions and very efficient reduction of the data. Thus with the data obtained in the coming year we hope to begin to get answers to some of the questions raised above and also to discover if the slow variations in gravity which we have observed are a consequence of geophysical pressures of interest to geodesy and geodynamics.

Acknowledgments. This study was supported by the following agencies and grants: NASA NGR 05-009-246, NSF EAR 75-21621, USDI 14-08-0001-G-297, USDI 14-08-0001-G-374, and NBS 5-9013.

## References

Chapman, S., and R. S. Lindzen, Atmospheric Tides, Gordon & Breach, New York, 1970.

Farrell, W. E., Phil. Trans. R. Soc. London A, 275, 253, 1973.

Goodkind, J. M., Tectonophysics, 52, 1979.

Prothero, W. A., and J. M. Goodkind, Rev. Sci. Inst. 39, 1257, 1968.

Warburton, R. J., C. Beaumont, and J. M. Goodkind, Geophys. J. R. Astr., 43, 707, 1975.

Warburton, R. J. and J. M. Goodkind, Astrophys. J., 208, 881, 1976.

Warburton, R. J. and J. M. Goodkind, Geophys. J. R. Astr., 48, 281, 1977.

Warburton, R. J. and J. M. Goodkind, Geophys. J. R. Astr., 52, 117, 1978.

Page Intentionally Left Blank

# Linear and Nonlinear Interactions Between the Earth Tide and a Tectonically Stressed Earth

Christopher Beaumont
Oceanography Department, Dalhousie University
Halifax, Nova Scotia, Canada B3H 3J1

Abstract. The earth tide provides a spatially and temporally predictable force that deforms the Earth and can be measured as changes in gravity, tilt, and strain at the Earth's surface. All things being equal, tidal constituent amplitudes and phases will not change with time. However, in the vicinity of earthquake focal regions conditions may not be equal. Crustal rocks stressed to more than ~0.6 of their failure strength exhibit material properties over and above that of linear elasticity. These effects, including dilatancy, are known from laboratory measurements but have not been proven in situ.

Interactions between the earth tide and crustal rocks that are under high tectonic stress are discussed in terms of simple phenomenological models. In particular, the difference between a nonlinear elastic model of dilatancy and a dilatancy model that exhibits hysteresis is noted. It is concluded that the small changes in stress produced by the earth tide act as a 'probe' of the properties of crustal rocks. Observations of earth tide tilts and strains in such high stress zones may, therefore, provide keys to the constitutive properties and the tectonic stress rate tensor of these zones.

## Introduction

Research concerning earth tides has concentrated almost exclusively on the elastic behaviour of the Earth's crust and mantle, and the boundaries between elastic materials that give rise to the so called geologic, topographic and cavity effects. There are, however, zones of high tectonic stress within the Earth's crust that may have a more complex rheology which can be investigated with the earth tide. Small changes in stress associated with the tide act as a 'probe' both of the state of tectonic stress and of the constitutive properties of these zones. This concept, which has yet to be observed in in situ experiments, is based on results from laboratory measurements of the continuum properties of intact rock samples subject to deviatoric stress in excess of ~0.6 of their failure strength. These ideas are discussed in terms of simple phenomenological models. A rigorous treatment, including other possible rheologies, is not warranted until observations exhibiting anomalous tidal variations have been recorded.

Beaumont and Berger (1974) rather naively suggested that temporal variations in the Earth's admittance to the earth tide should accompany $V_p/V_s$ seismic velocity anomalies, if the seismic velocity anomalies result from changes in the elastic properties of the crust. Their theoretical predictions suggested significant tidal admittance anomalies in the vicinity of 'elastic dilatant' crustal inclusions. These results have been criticized from the standpoint that the dilatant behaviour of rocks as observed in the laboratory (see, for example, Brace et al, 1966) is not merely a change in linear elastic properties. Consequently, an elastic inclusion model for tidal response represents an oversimplification. This paper provides an opportunity to extend Beaumont and Berger's results to more realistic rheologies.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*. Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

## Tidal Interactions with a Nonlinearly Elastic Crust

In this paper the normal definition of dilatancy, a volumetric change induced by a deviatoric stress, is used. Infinitesimal linear elasticity theory is represented by the tensor equation

$$\varepsilon_{ij} = AI_1 \delta_{ij} + G\sigma_{ij} \qquad i,j = 1,2,3 ,$$

where $\varepsilon_{ij}$ and $\sigma_{ij}$ are the strain and stress tensors, $\delta_{ij}$ is the Kronecker delta, A and G are constants and $I_1$ is the first invariant of the stress tensor = $\sigma_{ii}$ (summation convention over repeated indices implied). This equation is separable into isotropic and deviatoric parts,

$$\varepsilon_{ii} = 3AI_1 + G\sigma_{ii}$$

and

$$\varepsilon_{ij} = G\sigma_{ij} , \quad i \neq j ,$$

which demonstrates that deviatoric stress, $\sigma^D_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_{ii}$ , can induce only deviatoric strain and no volumetric change. Therefore, linearly elastic materials undergoing infinitesimal strain do not exhibit dilatancy.

Stuart and Dieterich (1974) considered a model of nonlinear elasticity complete to second order in stress (Reiner, 1945),

$$\varepsilon_{ij} = \phi_0 \delta_{ij} + \phi_1 \sigma_{ij} + \phi_2 \sigma_{ij} , \qquad i,j = 1,2,3 ,$$

where $\phi_k = \phi_k(I_1, I_2, I_3) \qquad k = 0,1,2 ,$

and $I_1$, $I_2$ and $I_3$ are the scalar invariants of the stress tensor. Such a model requires seven instead of two elastic constants, that is $\alpha$, B, C, H and M in addition to A and G.

$$\varepsilon_{ij} = (\alpha + AI_1 + BI_1^2 + CI_2)\delta_{ij} + (G + HI_1)\sigma_{ij} + M\sigma_{ik}\sigma_{kj} ,$$

$$i,j,k = 1,2,3 .$$

These constants determine the quadratic behaviour of nonlinear elasticity. Isotropic and deviatoric stress-strain relationships are cross-coupled, therefore, deviatoric stress induces

volumetric changes. In addition, the behaviour of such nonlinear elastic materials is anisotropic. This anisotropy, as emphasized by Stuart and Dieterich, is not intrinsic. It is a physical anisotropy that is caused by the stresses, but not the stress-strain relationship itself.

A possible nonlinear stress-strain relationship (figure 1) illustrates the interaction among a nonlinear elastic material, the tectonic stress, $\sigma$, and the earth tide stress, $\sigma_E$. This simplified approach ignores the fact that the $\sigma$ versus $\varepsilon$ or $\theta$ curves are functions of the stress invariants. No matter what the state of tectonic stress, strain in any direction will follow a curve of the type shown in figure 1, either concave or convex. The tidal stress (~0.01 bar), being much smaller than the range of tectonic stress (~5 kbar) or earthquake stress drops (~100 bar), will sample an essentially linear elastic behaviour proportional to the elastic tangent moduli of the nonlinear tectonic stress-strain curve. Therefore, the tidal admittance remains linearly elastic, but almost certainly anisotropic, and changes with increasing tectonic stress. Figure 1 suggests a decrease in tidal strain with increasing tectonic stress, but the exact form of the change will be more complex, some components of the strain tensor may increase while others decrease. Such changes can in theory be predicted in detail from the constitutive law and a knowledge of the tectonic and tidal stress tensors. At present, mere identification of temporal variations in the tidal admittance would be sufficient to demonstrate both a rheology more complex than linear elasticity and a changing tectonic stress. Heaton (1975) has outlined some of these ideas and has also pointed out that anisotropy will in general produce phase changes in the tidal admittance in addition to the amplitude variations.

Model of Dilatancy

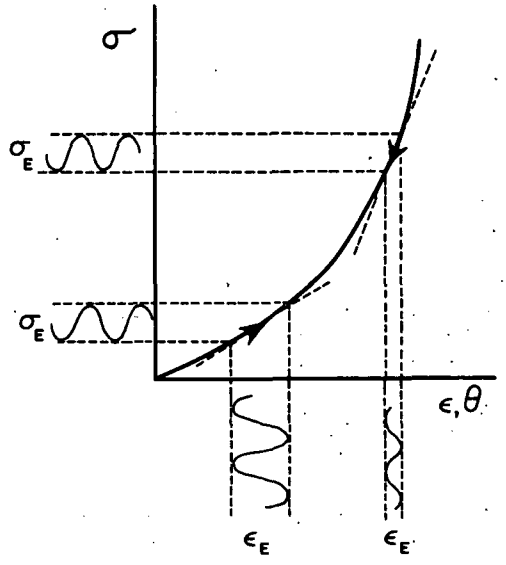



Fig. 1. A typical stress ($\sigma$) versus strain ($\varepsilon$) or dilatation ($\theta$) graph for nonlinearly elastic crustal rocks. Superimposed tidal stress ($\sigma_E$) induces a tidal strain ($\varepsilon_E$) proportional to the tangent to the tectonic stress-strain curve.

Tidal Interactions with a Crust that Exhibits Stress Hysteresis

Intact rock samples, when subject to cyclical stress that approaches the failure strength of the rock, exhibit dilatancy in the manner predicted by the Stuart and Dieterich model. That is, the stress-strain relationship is no longer linear for stresses in excess of ~0.6 of the failure strength and deviatoric stress induces volumetric strain. However, such samples exhibit stress hysteresis in addition to a nonlinear behaviour. Strain is a double valued function of stress, the value depending on whether stress is increasing or decreasing. Results from Scholz and Kranz (1974) (figure 2) on intact Westerly granite illustrate this behaviour. Repeated stress cycles (1, 18, 19) reduce the minimum deviatoric stress required for the onset of dilatancy to ~300 bars, but in each stress cycle a broadly similar hysteresis loop is fillowed. A widely accepted explanation is that the energy loss, measured by the area of the hysteresis loop, is the work done against friction in opening and closing pre-existing microcracks within the rock sample (figure 2). An idealized representation of the behaviour of the $\theta$ (dilatational), $\phi$ and z components of strain in the cylindrical rock sample (figure 2, lower

Dilatancy of Intact Westerly Granite (Scholz and Kranz, 1974)

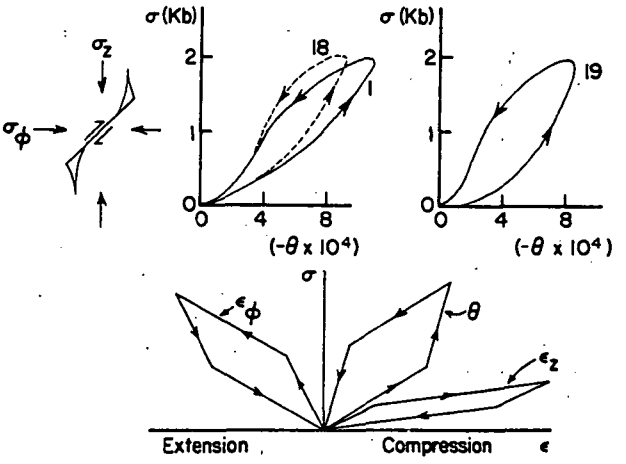



Fig. 2. Dilatancy of intact Westerly granite (from Scholz and Kranz, 1974). The panels (left to right, top to bottom) illustrate: 1) the opening of a microcrack under applied stress; 2) and 3) the changes in dilatation ($\theta$) during stress cycles 1, 18 and 19; 4) idealized stress-strain relationships for the $\varepsilon_\phi$, $\theta$, and $\varepsilon_z$ components of strain.

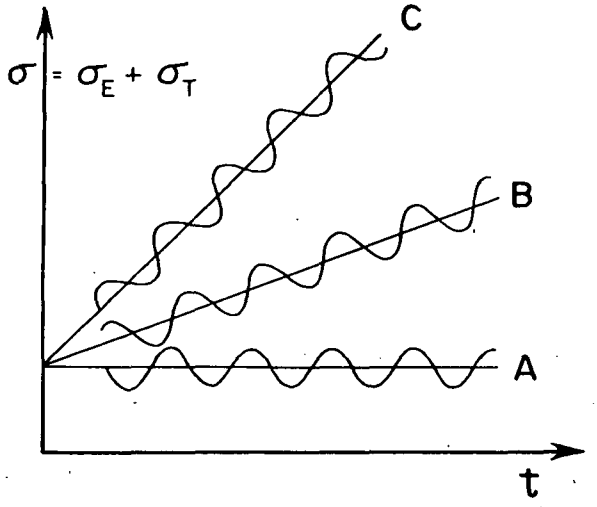



Fig. 3. The variation of superimposed tidal and tectonic stress ($\sigma$) with time (t). This simplified graph illustrates a linearly increasing tectonic stress and a tidal stress comprised of a single sinusoid. The significance of A, B and C is discussed in the text.

panel) demonstrates that all strain components are double valued. This model assumes that all cracks start to open and close at the same state of stress. It provides a first approximation to the true behaviour of the rock sample and can be represented mathematically in terms of an elastic-plastic material with internal back stresses. This description is discussed in detail in a later section.

The main concern is the interaction between the earth tide and crustal zones that exhibit such a dilatant behaviour with hysteresis. It will be seen that the interaction is different from that predicted for a nonlinearly elastic zone. Consider three simple cases (figure 3) in which the total stress is represented by a superposition of a linearly changing tectonic stress and a simple sinusoidally varying tidal stress. In A, the tectonic stress rate is either zero or very slowly varying ($d\sigma_T/dt \ll (d\sigma_E/dt)_{mean}$). In B, the tectonic stress rate is approximately equal to the mean tidal stress rate ($d\sigma_T/dt \sim (d\sigma_E/dt)_{mean}$). In C, the tectonic stress rate is sufficiently high that there is no decrease in total stress ($d\sigma_T/dt > (d\sigma_E/dt)_{max}$).

The interaction of these three possible stress-time regimes with the idealized hysteresis curve (figure 4) suggests that the character of the tidal response (strain or tilt) is peculiar to each of the three tectonic stress rate regimes. Consequently, observations of tidal strain can, in principle, be used to measure tectonic stress rates.

In figure 4A the tectonic stress rate is much less than the mean tidal stress rate. For states of total stress, $\sigma$, that are below that required for dilatancy (line 1-2) a superimposed tidal stress ($\sigma_E$) produces a normal amplitude elastic

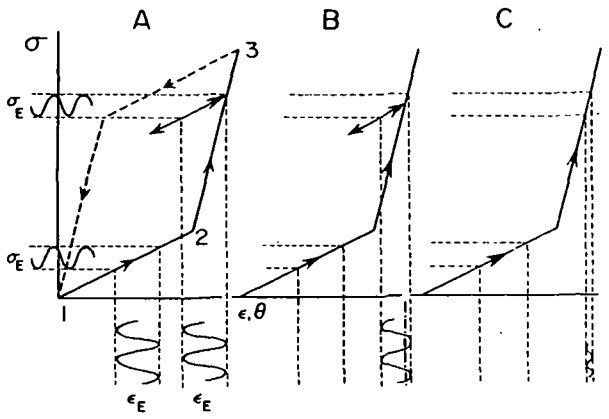



Fig. 4. Interaction between the earth tide ($\sigma_E$, $\epsilon_E$) and an idealized stress hysteresis loop for the three tectonic stress rates, A, B, and C from figure 3.

tidal strain ($\epsilon_E$). Moreover, the same response occurs even when the total stress is sufficiently high to produce a dilatant response (line 2-3). This result is a direct consequence of the general character of hysteresis curves. That is, any small stress decrease will be accompanied by an elastic strain recovery. The dilatant strain is not recovered until all the elastic strain has been recovered. Consequently, for small tectonic stress rates the crust retains its normal elastic admittance to the earth tide stress.

In figure 4C the tectonic stress rate is sufficiently high that there is never a decrease in total stress. In the pre-dilatancy region the response is identical to that of A. There will be no tidal anomaly. However, once the tectonic stress has moved into the dilatant region the strain is forced to change in proportion to the slope of the 2-3 line because there is no stress recovery. The tidal strain undergoes an anomalous decrease, or increase (see figure 2) but remains linearly related to the tidal potential.

The most interesting interaction is that shown in figure 4B. In the dilatant region stress reduction is accompanied by an elastic strain recovery proportional to the slope of line 1-2. The response remains elastic until the stress increases to the former stress maximum. Stress increases beyond the previous maximum induce strains proportional to the slope of line 2-3. The overall response (strain or tilt) is nonlinear. The magnitude of the nonlinearity depends on the relative slopes of lines 1-2 and 2-3, and on the relative size of the tectonic and mean tidal stress rates, the maximum nonlinearity occurring when these are equal. The main consequence of a nonlinear response will be the appearance of numerous additional lines in the tidal spectrum at sum and difference frequencies of the tidal constituents. Their detectability depends on the signal to noise ratio at the nonlinear interaction frequencies.

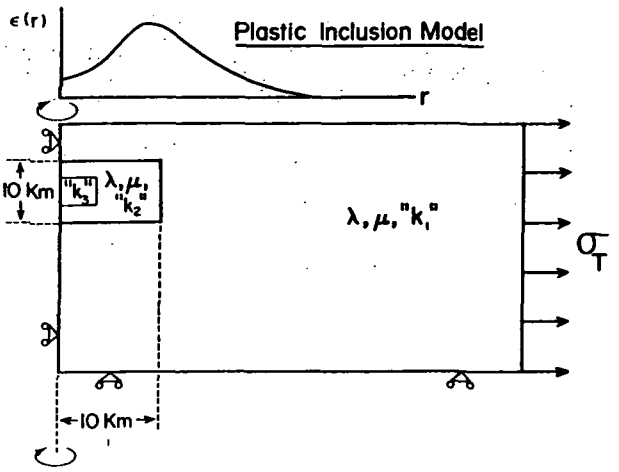



Fig. 5. Elastic-plastic model of a weak crustal inclusion. The model has uniform elastic properties ($\lambda$ and $\mu$) but the yield strength, characterized by $k_1$, $k_2$, and $k_3$, varies with position. Details of the precise yield criteria are discussed in the text.

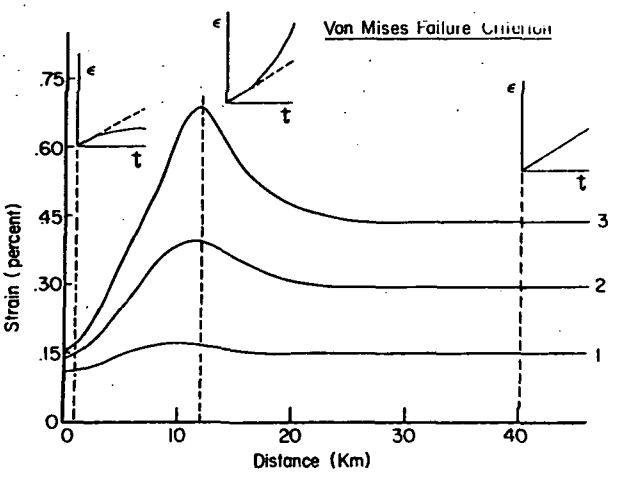



Fig. 6. Surface radial strain anomaly ($\varepsilon(r)$) as a function of three equal increases (1, 2, 3) of tectonic stress ($\sigma_T$) for the elastic-plastic model with inclusions that have a Von Mises yield criterion. See text for details.

## Plastic Models of Crustal Inclusions that Exhibit Hysteresis

It is anticipated that only confined volumes (crustal inclusions) will experience high stress concentrations at any given time. Only these inclusions will have a rheology that deviates from the normal elastic response. Alternatively, inclusions that have repeatedly experienced failure will almost certainly appear 'weak' in the sense that they will exhibit a departure from linear elasticity at lower stresses than the surrounding intact crust. An inclusion model (figure 5) similar to that employed by Beaumont and Berger (1974) but with an elastic-plastic rheology is used to explore the latter alternative. The model, a part of an axisymmetric half space with a 10 km radius, 10 km deep disc shaped inclusion buried at a depth of 4 km, is intended to be very simple. The half space has uniform elastic properties ($\lambda$,$\mu$) but the inclusion is characterized by yield strengths $k_2$ and $k_3$ that are less than that, $k_1$, for the surrounding crust. The result of particular interest is the character of the surficial strain, $\varepsilon(r)$, or tilt as a function of tectonic stress, $\sigma_T$. The results (figures 6 and 7) were obtained for two different yield criteria and associated plastic flow laws by finite element modelling (Bathe, Wilson and Iding, 1974, and Bathe, Ozdemir and Wilson, 1974).

The radial strain anomaly as a function of increasing stress (figure 6) is for a model with a Von Mises failure criterion (Prager and Hodge, 1951),

$$F(k,\sigma) = J_1 - k \ ,$$

where $J_1$ is the first invariant of the deviatoric stress tensor and $k = \sigma_y^2/3$, where $\sigma_y$ is the yield stress in simple tension. Yielding occurs when $F \geq 0$. The curves 1, 2 and 3 illustrate the strain anomaly for three equal stress increments. The strain as a function of position is constant for stress states $F < 0$. However, once yielding has occurred (curve 1), the buried inclusion appears 'weaker' and a strain anomaly develops in response to increasing stress by virtue of the contrast in properties between the inclusion and the surrounding crust. The inset figures illustrate the time variation of surface radial strain at selected points under the assumption that stress increases linearly with time. These figures may be interpreted in a similar manner to figure 4 because tectonic stress is proportional to time. The transition from an elastic (dashed line) to an anomalous response is now smooth, unlike that of figure 4, because the surrounding elastic crust 'filters' the plastic response of the inclusion. Strain recovery (not shown) on stress reduction will follow a hysteresis curve very much like that observed by Scholz and Kranz (1974) (figure 2, smooth curves).

In fact, there is an almost perfect analogy between the behaviour of the elastic-plastic inclusion model and that of a rock sample in the laboratory. For the rock sample, yielding (slip on pre-existing micro-cracks) occurs over a range of stress because each micro-crack has its own yield strength. Similarly, the crust is inhomogeneous and includes many inclusions like that of figure 5, each of which has its own yield strength. The elastic-plastic inclusion model is equivalent to a rock sample with one or more micro-cracks concentrated in one region. The analogy between a rock sample and the crust may be made complete by choosing a plastic yield criterion which has the same form as that for slip across micro-cracks. A suitable criterion is the Mohr-Coulomb criterion, which when generalized to three dimensions becomes the Drucker-Prager yield criterion,

$$F(I_1,J_2,k') = \alpha I_1 + J_2^{1/2} - k'$$

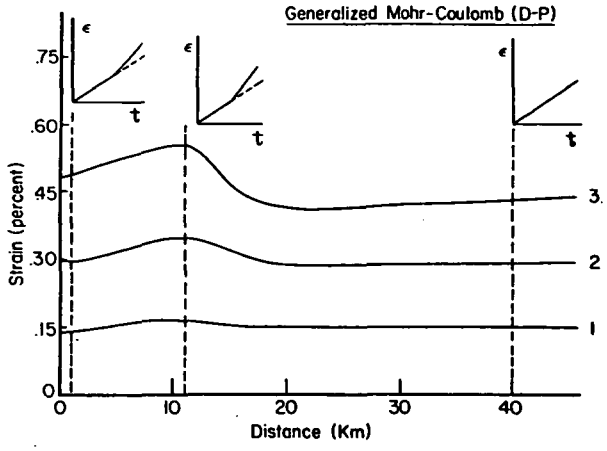



Fig. 7. Surface radial strain anomaly ($\varepsilon(r)$) as a function of three equal increases (1, 2, 3) of tectonic stress ($\sigma_T$) for the elastic-plastic model with inclusions that have a Drucker-Prager (generalized Mohr-Coulomb) yield criterion. See text for details.

(Drucker and Prager, 1952), where $J_2$ is the second invariant of the deviatoric stress tensor and $\alpha$ and $k'$ are material properties related to the cohesion c and angle of friction $\psi$. $\alpha = 2\sin\psi/\sqrt{3}(3-\sin\psi)$, $k' = 6c\cos\psi/\sqrt{3}(3-\sin\psi)$. It is also interesting that the associated flow law is dilatant.

$$\dot{\theta}^P = 3\alpha\lambda \ ,$$

where $\dot{\theta}^P$ is the rate of plastic dilatation and $\lambda$ is a constant proportional to the rate of plastic working, $\sigma_{ij}\dot{\varepsilon}^P_{ij}$, or defined in terms of the plastic strain rate and the yield function.

$$\dot{\varepsilon}^P_{ij} = \lambda \partial F/\partial\sigma_{ij}$$

(Drucker and Prager, 1952).

The results for the Drucker-Prager model (figure 7) suggest that the surface tectonic strain anomalies will not be as large as those for the Von Mises model. The difference is partly a result of the two yield crieteria, though values for the constants $\alpha$, $k'$ and $k$ were chosen so that yielding occurs at the same tectonic stress for both models. A more important factor is the effect of dilatancy. The major difference between the strain for the two models at distances $\lesssim$10 km is due to strain induced by tectonic uplift of the zone over the dilatant inclusion. As the value of $\alpha$ is reduced toward zero, the results of the Drucker-Prager model trend smoothly to those of the Von Mises model.

When the tectonic stress is reduced, the strain behaviour of the inclusion models is very similar to that of rock samples. The plastic inclusions do not possess internal 'back stresses'; therefore, the plastic strain would be irrecoverable if the inclusions were not embedded in an elastic matrix. 'Back stresses' in the elastic matrix ensure that the plastic strain is recovered. That there is irrecoverable plastic work done during this process ensures that the stress-strain relation will exhibit hysteresis over a stress cycle. The situation in a stressed rock sample is very similar. The micro-cracks have no intrinsic 'back stresses.' It is the stress in the elastic matrix that ensures that dilatancy is recovered.

The only weakness of the plastic flow laws that have been used is that the rheology is elastic-perfectly plastic; that is, there is no strain hardening. Such a model suggests that once slip across a micro-crack has been initiated it will proceed at constant stress.

We are now in a position to predict the form of tidal interactions with the elastic-plastic inclusion models. The $\varepsilon$ versus t inset graphs of figures 6 and 7 are the equivalents of the $\varepsilon$ versus $\sigma$ graphs of figure 4. The interpretation is exactly the same with the addition that the tidal anomalies will vary with position on the surface of the model. At large distances from the inclusion the response remains normally elastic; no anomalous tidal admittance is predicted. In the neighbourhood of the inclusion the character of the tidal admittance will depend on the relative tidal and tectonic stress rates.

## Discussion and Conclusions

The intent of this paper has been a discussion of phenomenological models of linear and nonlinear variations in tidal admittance that result from stress dependent properties of crustal rocks. The proposed constitutive relations are those observed for intact laboratory samples at deviatoric stress levels in excess of ~0.6 of their failure strength. If crustal rocks exhibit the same properties in situ, it follows that nonlinear variations in tidal admittance of the kind predicted would indicate: 1) that tectonic stresses are sufficiently large to induce nonlinear behaviour, 2) that crustal dilatancy with hysteresis is occurring, and 3) that tectonic stress rates are comparable to the tidal stress rate. The absence of a nonlinear tidal admittance is not such a useful result for it merely indicates that one or more of the above conditions has not been met.

The most interesting condition concerns the need for comparable tectonic and tidal stress rates. This condition does not arise for constitutive models of the type proposed by Stuart and Dieterich (1974), nor is a nonlinear tidal admittance predicted. Unfortunately, tidal stress rates are not optimal. Mean tectonic stress rates are of the order 1-20 bars/year ($10^{-1}$ - 2 MPa/year), if it is assumed that earthquake stress drops are a measure of the tectonic stress accumulation between repeat earthquakes in the same area. The assumptions involved in this estimate are discussed by Dieterich (1978). Mean tidal stress rates for mid-latitudes are larger, from 50-100 bars/year (5-10 MPa/year), where the mean tidal stress rate is taken to mean $|\dot{\sigma}_E|$. Consequently, nonlinear interactions are predicted only under what might be termed an accelerated tectonic stress rate. Earthquake precursors suggest that accelerated tectonic stress rates are probable before earthquakes.

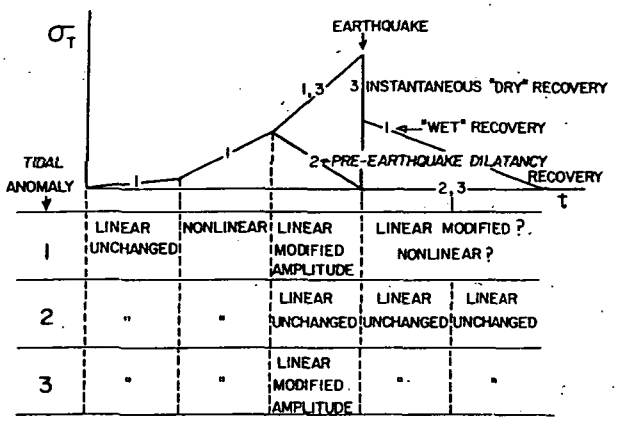



Fig. 8. Possible variations in tidal admittance during an earthquake cycle ($\sigma_T$ v t) for three postulated earthquake models. 1 is the dilatancy-diffusion model in which post-earthquake dilatancy recovery is controlled by the diffusion of fluids. 2 is a model in which dilatancy is recovered before the earthquake occurs. 3 is a model that postulates instantaneous dilatancy recovery at the time of the earthquake. The details of the variation of tidal admittance will depend on the relative tidal and tectonic stress rates, and the tectonic stress required for dilatancy to occur.

The very appearance of a stress induced precursor and its disappearance before an earthquake indicates that there is stress redistribution during the period for which a precursor exists. If precursors are a manifestation of stress induced dilatancy that is recovered as a broad zone of cracks coalesce to form a fault zone, there must be significant stress redistribution during the process. The same conclusion is reached if the diffusion of fluids is also involved. Possible variations in tidal admittance for postulated pre-, co-, and post-seismic processes are summarized in figure 8.

A comment on tidal triggering of earthquakes is relevant at this point. Heaton (1975), among others, has noted that the probability that an earthquake will be tidally triggered is greatest when $|\dot{\sigma}_E| >> |\dot{\sigma}_T|$. This result assumes that the earthquake occurs through a Mohr-Coulomb type failure at that phase of the tide when shear stress on the fault plane is maximized and normal stress is minimized. This is exactly the condition for which a nonlinear tidal admittance is least likely. Conversely, tidal triggering of earthquakes is least likely when $|\dot{\sigma}_E| << |\dot{\sigma}_T|$, but this is the most favoured condition for an anomalous tidal admittance. The conclusion, which can be tested, is that earthquakes that appear to be tidally triggered are unlikely to have anomalous tidal precursors in the period immediately before the earthquake.

The overall conclusions may be summarized in the following manner. The Earth's admittance to the earth tide may be sensitive to tectonic stress if in situ crustal rocks exhibit the same stress dependent properties as those observed for intact laboratory samples. The earth tide tilt and strain in a region of high and variable tectonic stress will exhibit linear variations if the Earth is nonlinearly elastic. This result is similar to that predicted by Beaumont and Berger (1974) with the additional effect that the tidal admittance will be anisotropic. A more interesting result is predicted if crustal rocks exhibit stress hysteresis. At high tectonic stress the admittance of the Earth to the earth tide will be a function of the tectonic stress rate. In particular, a nonlinear admittance is predicted when the tidal and tectonic stress rates are approximately equal.

Acknowledgments

I would like to thank Ross Boutilier for assistance with computer programming of the elastic-plastic finite element models.

References

Bathe, K-J., H. Ozdemir, and E. L. Wilson, Static and dynamic geometric and material nonlinear analysis, report UC SESM 74-4 of the Structural Engineering Laboratory, University of California, Berkeley, California, 1974.

Bathe, K-J., E. L. Wilson, and R. H. Iding, 'NONSAP', a structural analysis program for static and dynamic response of nonlinear systems, report UC SESM 74-3 of the Structural Engineering Laboratory, University of California, Berkeley, California, 1974.

Beaumont, C., and J. Berger, Earthquake prediction: modification of the earth tide tilts and strains by dilatancy, Geophys. J. R. astr. Soc., 39, 111-121, 1974.

Brace, W. F., B. W. Paulding, Jr., and C. H. Scholz, Dilatancy in the fracture of crystalline rocks, J.Geophys.Res., 71, 3939-3953, 1966.

Dieterich, J. H., Preseismic fault slip and earthquake prediction, J. Geophys. Res., 83, 3940-3947, 1978.

Drucker, D. C., and W. Prager, Soil mechanics and plastic analysis or limit design, Q. appl. Math., 10, 157-165, 1952.

Heaton, T. H., Tidal triggering of earthquakes, Geophys. J. R. astr. Soc., 43, 307-326, 1975.

Prager, W., and P. G. Hodge, Theory of perfectly plastic solids, Chapman and Hall, New York, 1951.

Reiner, M., A mathematical theory of dilatancy, Am. J. Math., 67, 350-362, 1945.

Scholz, C. H., and R. Kranz, Notes on dilatancy recovery, J. Geophys. Res., 79, 2132-2135, 1974.

Stuart, W. D., and J. H. Dieterich, Continuum theory of rock dilatancy, Proc. Third Cong. Int. Soc. Rock Mech., IIA, 530-534, 1974.

# The Inverse Problem: Ocean Tides Derived from Earth Tide Observations

John T. Kuo
Lamont-Doherty Geological Observatory, Columbia University
Palisades, New York 10964

Summary. Evidence has been accumulated to the point that open ocean tides can indeed be mapped by solving the inverse problem using land-based and island-based tidal gravity observations, supplemented by shore and island ocean tide-gauge observations and a few deep ocean-bottom observations. Past efforts have been towards a better understanding of open ocean tides both through numerical integration of Laplace's tidal equations and through direct observations of tides in the open oceans. The co-amplitude and co-phase tidal charts principally of the tidal constituent $M_2$ calculated by numerical integration of Laplace's tidal equations generally fail to agree among themselves, and with the tidal observations on mid-ocean islands. The problem of open ocean tides remains open.

During the last few years, we have indirectly mapped both the $M_2$ and $O_1$ ocean tides in the northeastern Pacific Ocean, based on the inverse scheme of "Linear Programming." A comparison between the inversion results of the $M_2$ and $O_1$ ocean tides and the three ocean-bottom observations made by Scripps, which were not included in the inversion scheme, gives an excellent agreement and assures that the proposed technique of indirect mapping of open ocean tides by means of tidal gravity observations can contribute significantly to attack the classic problem of open ocean tides.

Ever since we proposed the possibility of indirectly mapping ocean tides by means of land- and island-based tidal gravity measurements (Kuo et. al., 1970), skepticism has been raised principally by physical oceanographers, concerning its actual feasibility, typically such as the recent one by Zetler (1978):

> "It has been demonstrated that land-based earth tides observed on a gravimeter near an ocean are modified by the ocean tidal loading. It is more controversial, however, whether it is possible to map the open ocean tides by solving the inverse problem using land-based gravity measurements, shore constraints, and a few ocean-bottom stations. Certainly it can be done with an infinite of stations of perfect precision; the number and precision of land and ocean measurements necessary to achieve required accuracy have not yet been determined."

Evidence has been accumulated to the point that open ocean tides can indeed be mapped by solving the inverse problem using land- and island-based tidal gravity measurements, coupled with shore and island ocean tidal measurements and a few ocean-bottom measurements. There is definitely no need to have an infinite of stations of perfect precision. The degree of precision of tidal gravity measurements, from our experience, must be 1% or better, which is perfectly achievable.

Much effort has been directed toward a better understanding of open ocean tides, both through numerical integration of Laplace's tidal equations, (Pekeris and Accad, 1969; Zahel, 1970; Hendershott and Munk, 1970; and others) and through direct measurements of tides in the deep oceans (Nowroozi et al., 1969; Filloux, 1971, Munk et al., 1970; and others). The co-amplitude and co-phase tidal charts principally of the tidal constituent $M_2$ calculated by numerical integration of Laplace's tidal equations are apparently quite sensitive to the boundaries of the ocean basins and the law of friction, and generally fail to give close agreement with the tidal observations on mid-ocean islands. Moreover, the agreement among various tidal charts is still poor as shown in Figures 1 and 2, comparing the tidal charts for example, the $M_2$ and $O_1$ in the northeastern Pacific Ocean.

During the last few years, we have indirectly mapped both the $M_2$ and $O_1$ ocean tides in the northeastern Pacific Ocean, based on the inversion scheme of "Linear Programming." A comparison between the inversion results of the $M_2$ and $O_1$ ocean tides and of the three ocean-bottom observations made by Munk et al. (1970) and Irish et al. (1971), which were not included in the inversion scheme, gives an excellent agreement.

The basic data used in the inversion are principally from the tidal gravity observations made on North America and on the islands in the northeastern Pacific Ocean and from the tide gauge observations on the coasts and islands. A total of 17 tidal gravity stations and 62 coastal and island tide gauge stations was used for the northeastern Pacific Ocean. The distributions of the tidal gravity stations and the tide gauge stations of the contiguous continents and islands of the northeastern Pacific used in the inversion are shown in Figure 3.

The $M_2$ and $O_1$ worldwide tide maps of Tiron et.al. (1967), designated as Tiron $M_2$ Map and Tiron $O_1$ Map, respectively, as shown in Figures 4 and 5, were adopted as the starting models and were digitized at 2° by 2° spacings. As a matter of fact, Kuo and Jachens (1977) have shown that a starting model is not of crucial importance in the inversion. For oceanic regions within 2.5° of the tidal gravity observational stations, these maps were modified to conform with coastal observations. The predicted tidal gravity effects arising solely from ocean tides were calculated by numerical convolution of Tiron $M_2$ and $O_1$ Maps with linear combination of a mass loading Green's function for an oceanic crust model (Farrell, 1972) and, in addition, of a Newtonian attraction Green's function for a density coating layer on the surface of a sphere, which is not accounted for in the formulation of the loading Green's function by Farrell (1972).

A fourth order two-dimensional polynomial was selected as the highest order polynomial surface

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

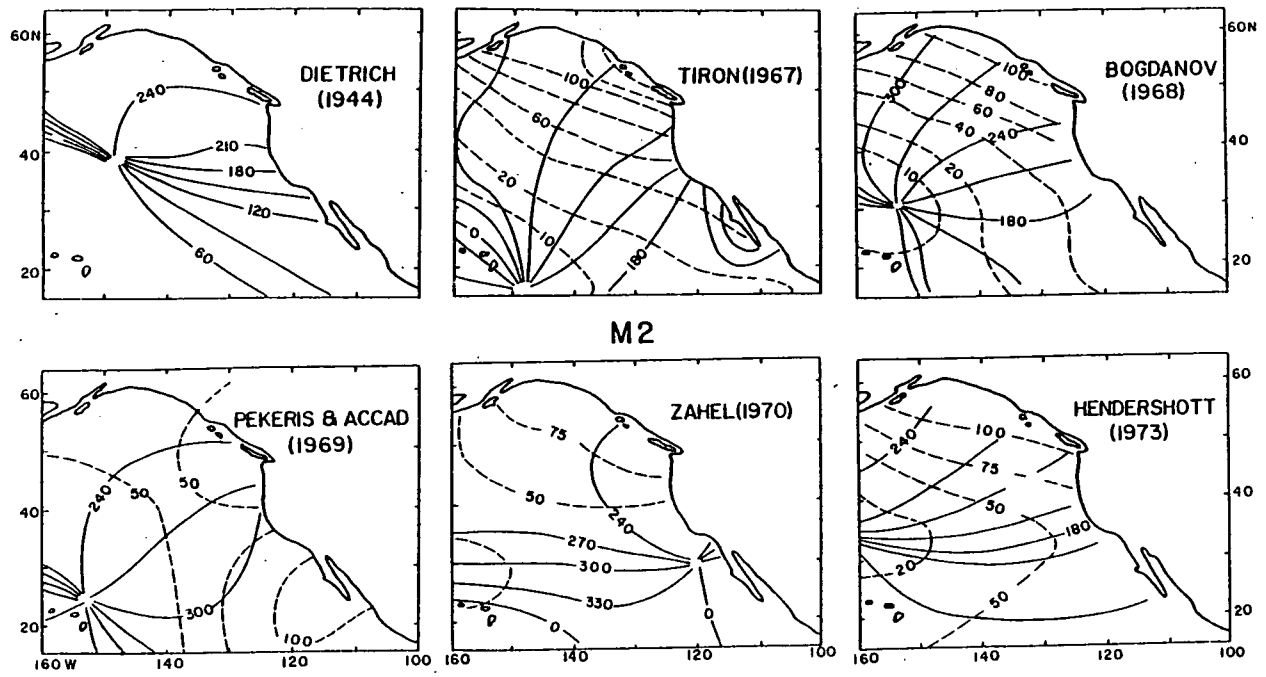



Figure 1. The $M_2$ ocean tidal constituent as derived by empirical (Dietrich, 1944) and Laplace tidal equations by various authors.

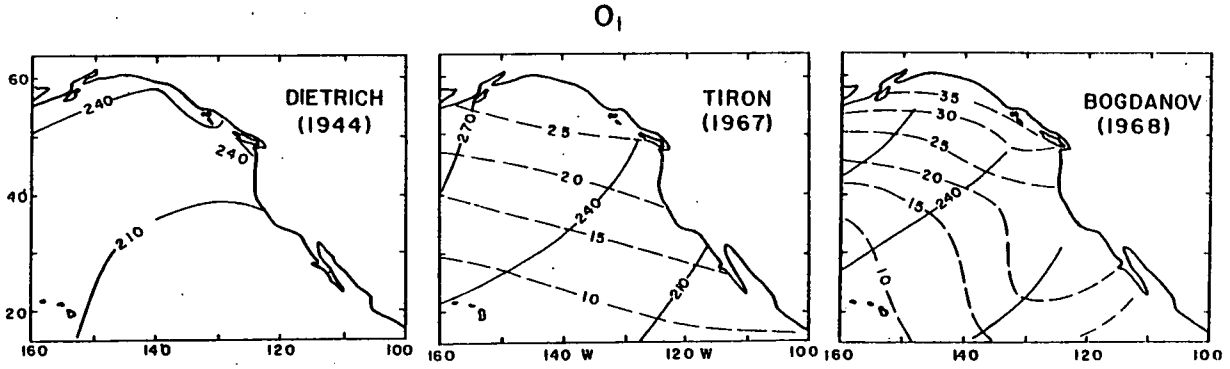



Figure 2. The $O_1$ ocean tidal constituent as derived by empirical (Dietrich, 1977) and Laplace tidal equations by various authors.

correction applied to Tiron $M_2$ and $O_1$ Maps. The limited distribution of the available tidal gravity data does not warrant a surface of order higher than fourth.

The inversion scheme is based on "Linear Programming." The reason of choosing this Linear Programming Inversion is very simple. Linear programming is concerned with the optimum operation of interdependent variables, that is the minimization of a linear objective function, whose variables satisfy a system of linear inequalities (Danzig, 1977). Unlike the other numerical solutions of minimization problems by iterative procedures, the solution to a linear programming problem, if exists, is unique and gives a true minimum for the entire system, i.e. a global minimum. Thus, linear programming procedure can be ideally used to seek the optimal inverse solution subject to constraints imposed by the tidal gravity and ocean tide gauge observations.

Figures 6 and 7 give the new $M_2$ and $O_1$ tide maps for the northeastern Pacific Ocean, bounded on the west by 160°W on the south by 15°N, and on the north and east by North America, resulting from the application of the linear programming inversion. The most prominent feature of the new $M_2$ map is the existence of a single amphidromic system in the region, centered at approximately 26°N latitude, 137°W longitude and rotates counterclockwise. The amplitude contours display a low amplitude trough of less than 20 cm. The amplitude increases smoothly north of the trough attaining magnitudes

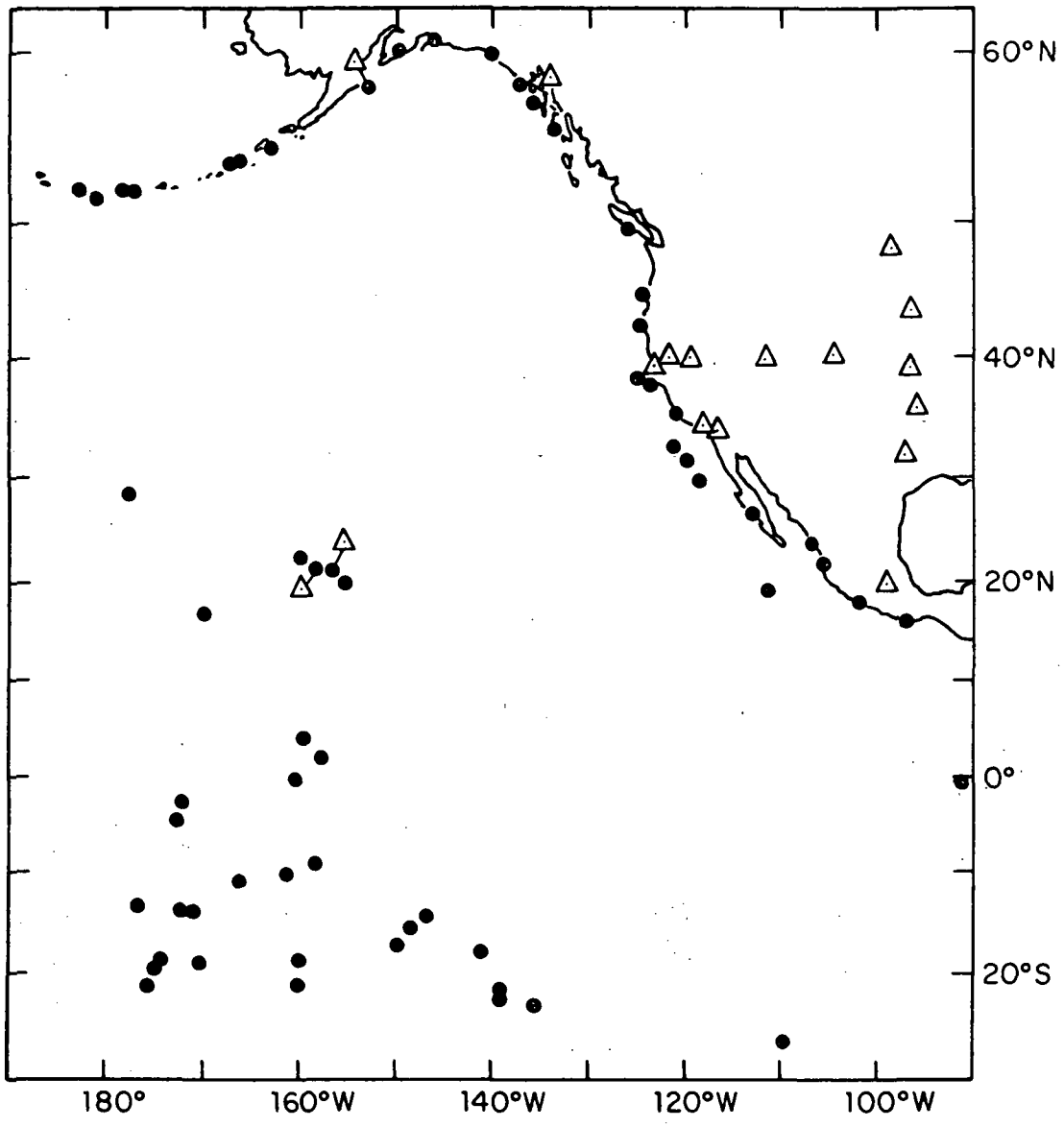



Figure 3. Distribution of tidal gravity, tide gauge and ocean-bottom tide stations in the northeastern Pacific Ocean.

greater than 120 cm near the head of the Gulf of Alaska. The prominent feature of the new $O_1$ map is the modification of both the coamplitude and cophase lines from that of the Tiron $O_1$ Map. The new $O_1$ map shows a delay of the cophase lines, compared with the cophase lines of the Tiron $O_1$ Map. Except that the coamplitude line of 25 cm remains nearly the same as Tiron $O_1$ Map, all the other coamplitude lines of 10, 15, and 20 cm are curved somewhat to conform with the west coast of North America.

The goodness of the inversion results must be critically tested to insure the validity of the inversion procedure. It can be best accomplished by comparing the inversion results with observations The observations from the three ocean-bottom sites, Kathy, Filloux, and Josie II, shown as solid circles in Figure 8 were not included as the basic data in the inversion procedure, and are far removed from possible local perturbation introduced by islands or submarine topographic features. The comparison between the values obtained by the various tidal charts, the inversion procedure and these corresponding observed values is given in Tables I and II.

The agreement between the inversion results and the observations at all three sites is better than 6 cm in amplitude and better than 5° in phase for $M_2$. Although there are no available observed data of $O_1$ for the station Josie II, the agreement for $O_1$ at the stations, Kathy and Filloux is better than 1 cm in amplitude and 1° in phase. The inversion results appear to be quite insensitive to uncertainties in knowledge of worldwide ocean tidal distributions and to possible biases in tidal gravimeter calibration. The agreement at Kathy and Filloux for both $M_2$ and $O_1$ is embarrassingly good, probably somewhat fortuitous, since the starting

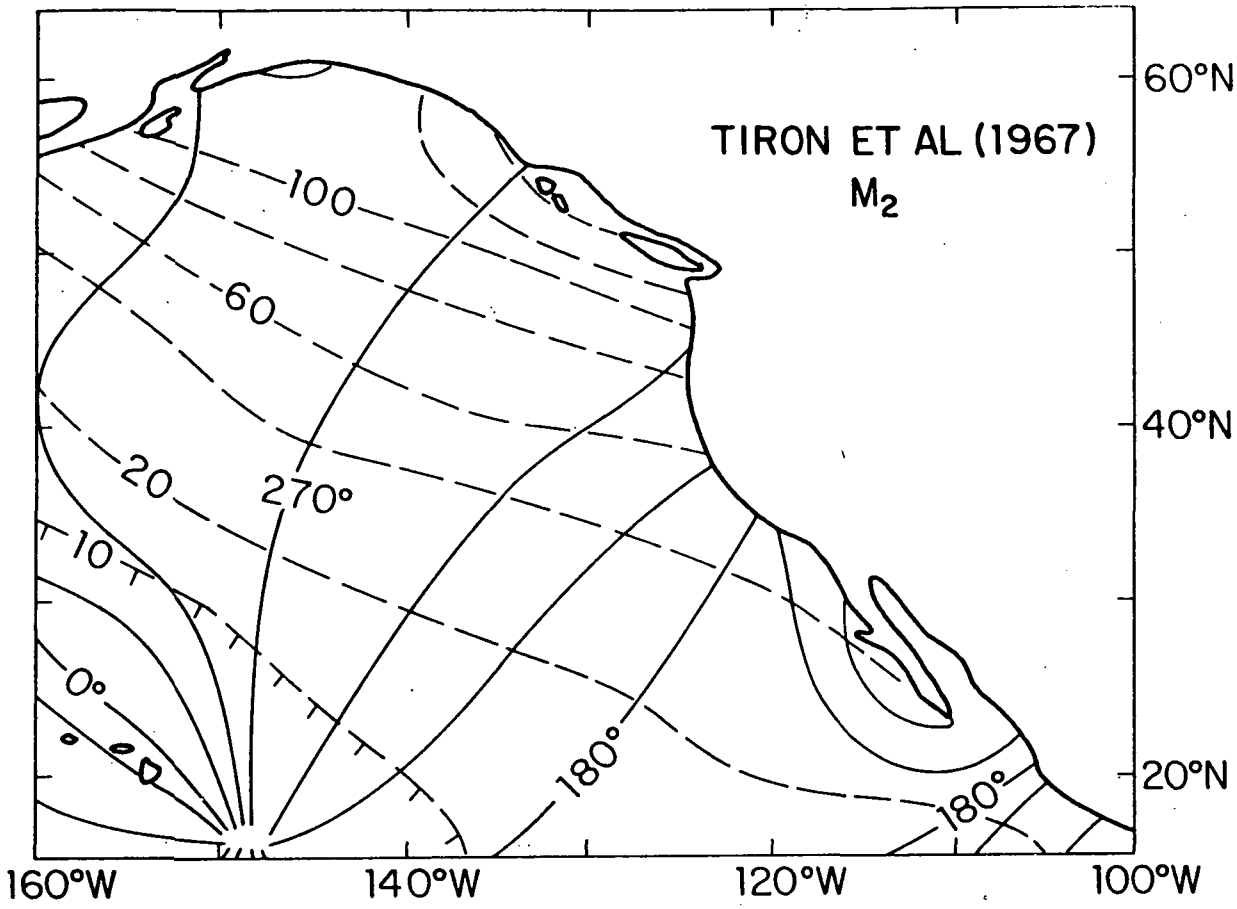



Figure 4. Tiron et al., (1967), Original $M_2$ Tidal Map.

TABLE I

COMPARISONS OF THE INVERSED RESULTS OF $M_2$ AND $O_1$ WITH THE THREE OCEAN-BOTTOM TIDAL OBS. (MUNK ET.AL., 1970) AND AVAILABLE TIDAL CHARTS AT THE STATIONS, KATHY, FILLOUX AND JOSIE II.

| STATION | | OBSERVED (O) A(cm)/G(deg) | AVAILABLE COTIDAL CHARTS A/G D | T | B | P&A | Z | H* | INVERSION A/G |
|---|---|---|---|---|---|---|---|---|---|
| Kathy 124°25.8'W 27°45.0'N | $M_2$ | 28.6/128.0 | -/115 | 31/180 | 41/172 | 75/305 | 35/305 | 60/160 | 28.3/132.9 |
| | $O_1$ | 17.5/199.0 | -/190 | 15/219 | 21/208 | - | - | - | 18.0/199.0 |
| Filloux 129°01.1'W 24°46.9'N | $M_2$ | 18.8/107.0 | -/92 | 20/186 | 35/172 | 52/300 | 32/320 | 55/157 | 17.8/104.5 |
| | $O_1$ | 15.6/201.3 | -/190 | 12/220 | 22/210 | - | - | - | 15.5/200.0 |
| Josie II 144°59.7'W 34°00.3'N | $M_2$ | 26.6/267.0 | -/75 | 27/261 | 15/232 | 25/262 | 35/270 | 30/197 | 20.0/270.0 |
| | $O_1$ | Not determined | -/210 | 15/245 | 25/235 | - | - | - | 17.0/224.0 |

*D = Dietrich, 1977; T = Tiron et.al., 1967; B = Bogdanov et al., 1968

P&A = Pekeris & Accad, 1969; Z = Zahel, 1970; H = Hendershott, 1973.

O = Observed (Munk et al., 1970)

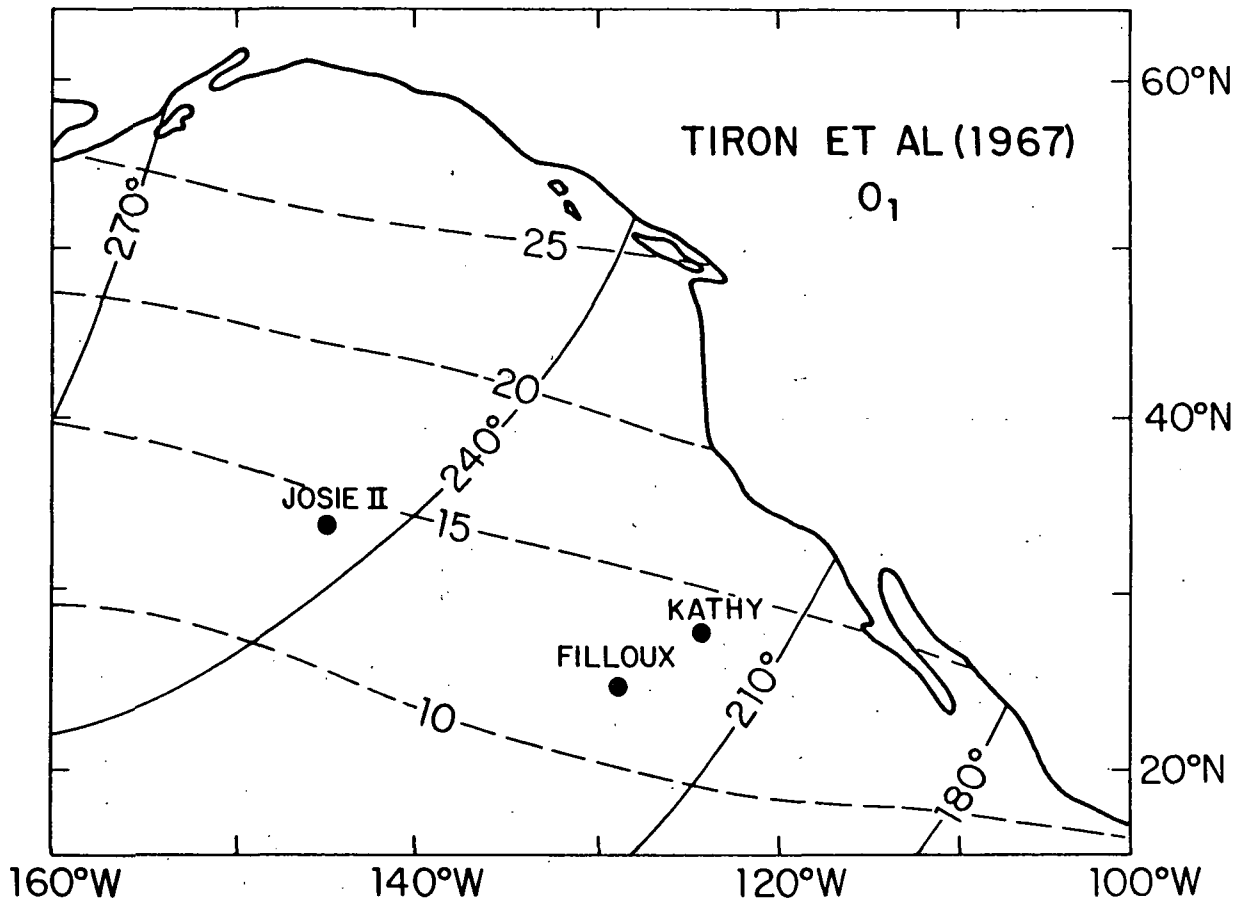



Figure 5. Tiron et al., (1967), Original $O_1$ Tidal Map.

TABLE II

DIFFERENCES

| STATION | | D-O | T-O | B-O | P&A - O | Z-O | H-O | INV-O |
|---|---|---|---|---|---|---|---|---|
| Kathy | $M_2$ | -/-13° | +2.4/+52 | +12.4/+44 | +46.4/+177 | +6.4/+177 | +31.4/+32 | -0.3/+4.9 |
| | $O_1$ | -/-9 | -2.4/+20 | +3.5/+9 | - | - | - | +0.5/0 |
| Filloux | $M_2$ | -/-15 | +1.2/+79 | +16.2/+65 | +33.2/+193 | +13.2/+213 | +36.2/+50 | -1.0/-2.5 |
| | $O_1$ | -/-11.3 | -3.6/+18.7 | 6.4/+8.7 | - | - | - | -0.1/-1.3 |
| Josie II | $M_2$ | -/-192 | 0.4/-6 | -11.6/-35 | -1.6/-5 | +8.4/+3 | +3.4/-70 | -6.6/+3.0 |
| | $O_1$ | Not Observed | - | - | - | - | - | - |

model could not be read reliably to one degree and one cm.

The crucial difference between the present procedure and those of past workers is that land-based and island-based tidal gravity observations were used as the bases for interpolating between widely spaced direct ocean tidal observations rather than the traditional empirical procedures based on the numerical integration Laplace tidal equations. For comparison, Figure 9 shows a map

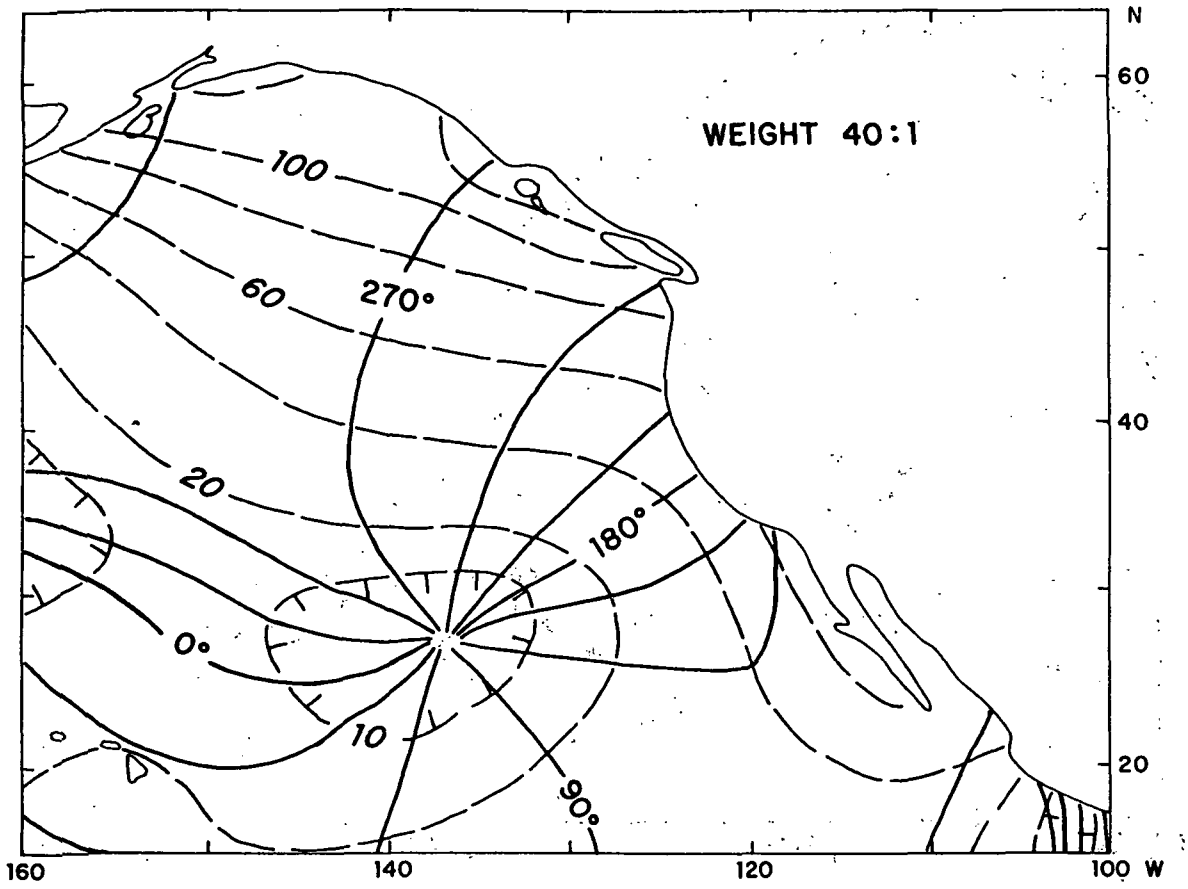



Figure 6. $M_2$ obtained from inversion for the northeastern Pacific Ocean.

derived by Munk et al. (1970) based on ocean-bottom tidal measurements and empirical inference.

The inversion results of the $M_2$ and $O_1$ tides in the northeastern Pacific Ocean assure that earth tidal gravity can contribute significantly to attack one of the remaining classic geodynamic problems, - the open ocean tides. The worldwide open ocean tides can be mapped by means of solving the inverse problem of ocean tides for a series of tidal gravity observations, complemented by coastal, shelf, and island tidal gauge observations and a limited number of ocean-bottom tidal observations.

Acknowledgements. The author gratefully acknowledges the support of the Office of Naval Research under the contract T0-0605, and of the National Science Foundation under the grant NSF-EAR-76-24383.

References

Bogdanov, K.T. and V. Magarik, Numerical solutions to the problem distribution of semidiurnal tides $M_2$ and $S_2$ in the world ocean (transl.) Dokl. Akad Nauk SSSR, 172, 1315-1317, 1967.

Danzig, G.D., Linear Programming, Past and Future, Science Technology and the Modern Navy, Thirtieth Anniversary, 1946-76, Ed. Edward I. Salkovitz, Office of Naval Research, Dept. of the Navy, 84-95, 1976.

Dietrich, G., "Die Schwingungssysteme der halb-und eintagigen Tiden in den Ozeanen," Veroffentl. Inst. Meereskende, Univ. Berlin, A 41, 1-68, 1944.

Farrell, W.E., "Deformation of the earth by surface loads, Rev. Geophysics 10 (3), 761-797, 1972.

Filloux, J.H., "Deep sea tide observations from the northeast Pacific," Deep Sea Res., 18, 275-284, 1971.

Hendershott, M.C. and W.H. Munk, "Tides," Ann. Rev. Fluid Mech., 2, 205-224, 1970.

Hendershott, M.C., "The effects of solid-earth deformation on global ocean tides," Geophy. J., Roy. Aston. Soc., 29, 389-402, 1972.

Irish, J., W.H. Munk and F. Snodgrass, $M_2$ amphidrome in the northeast Pacific, Geophy. Fluid Dyn., 2, 355-360, 1971.

Kuo, J.T., R.C. Jachens, M.Ewing and G. White, "Transcontinental tidal gravity profile across the United States," Science, 168, 969-971, 1970.

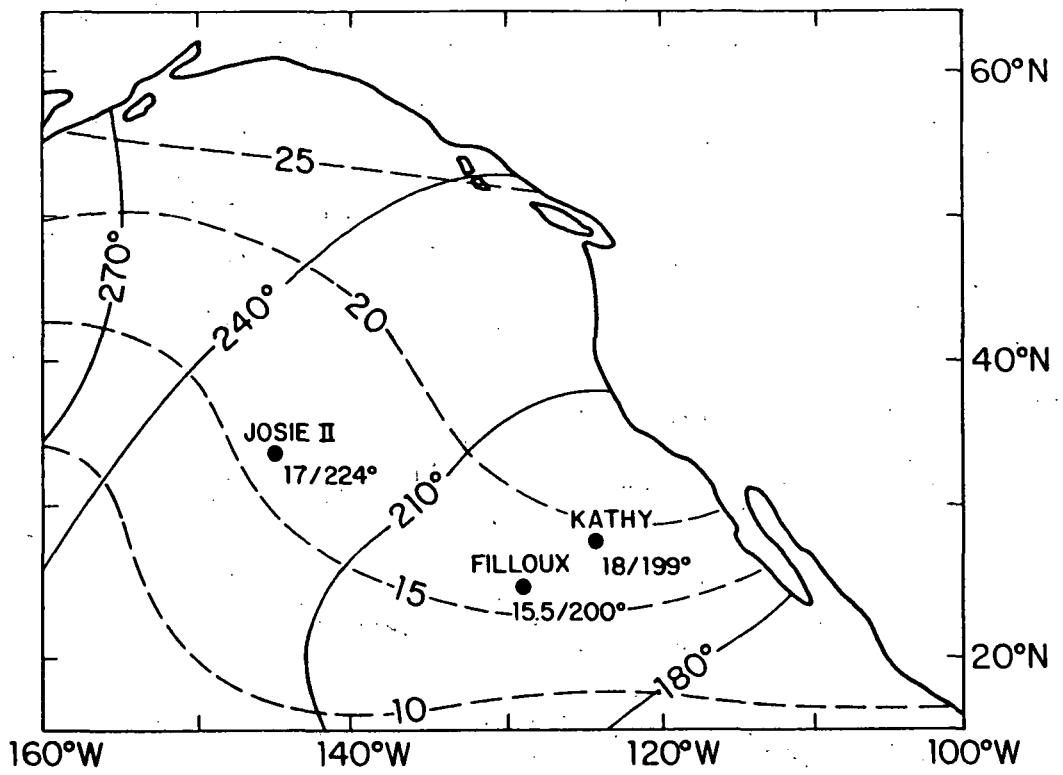



Figure 7. $O_1$ obtained from inversion for the north-eastern Pacific Ocean.

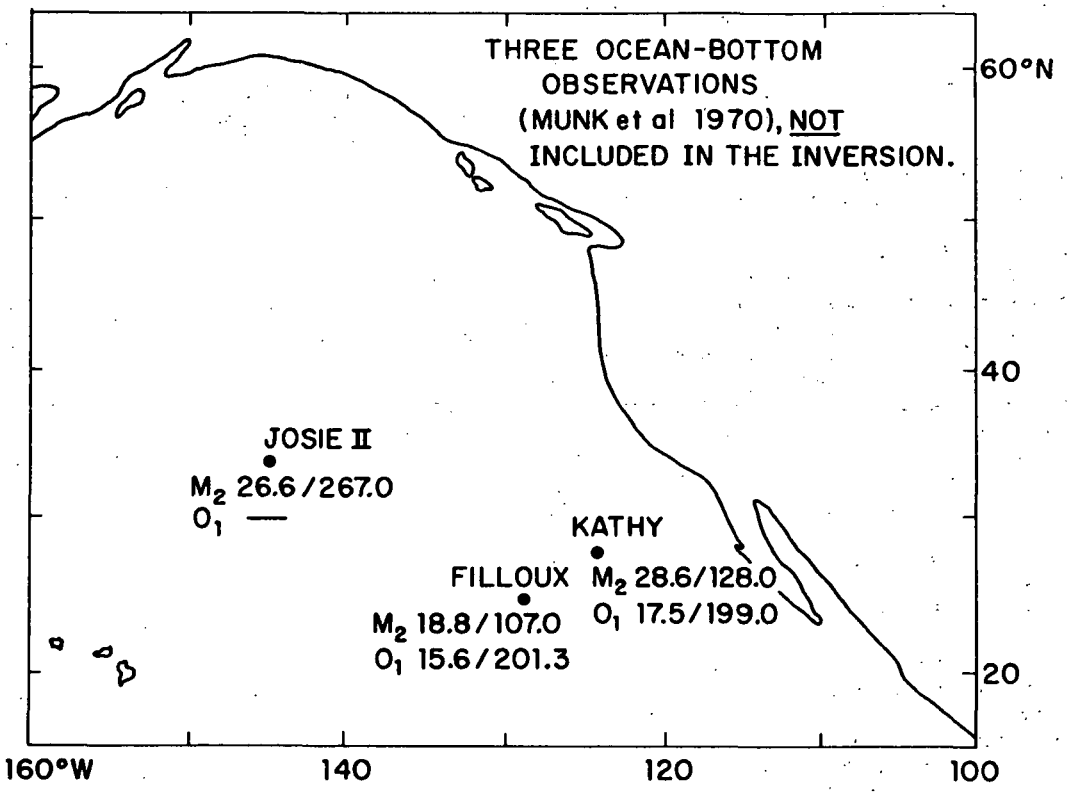



Figure 8. Three ocean-bottom observations (Munk et.al. 1970), not included in the inversion.

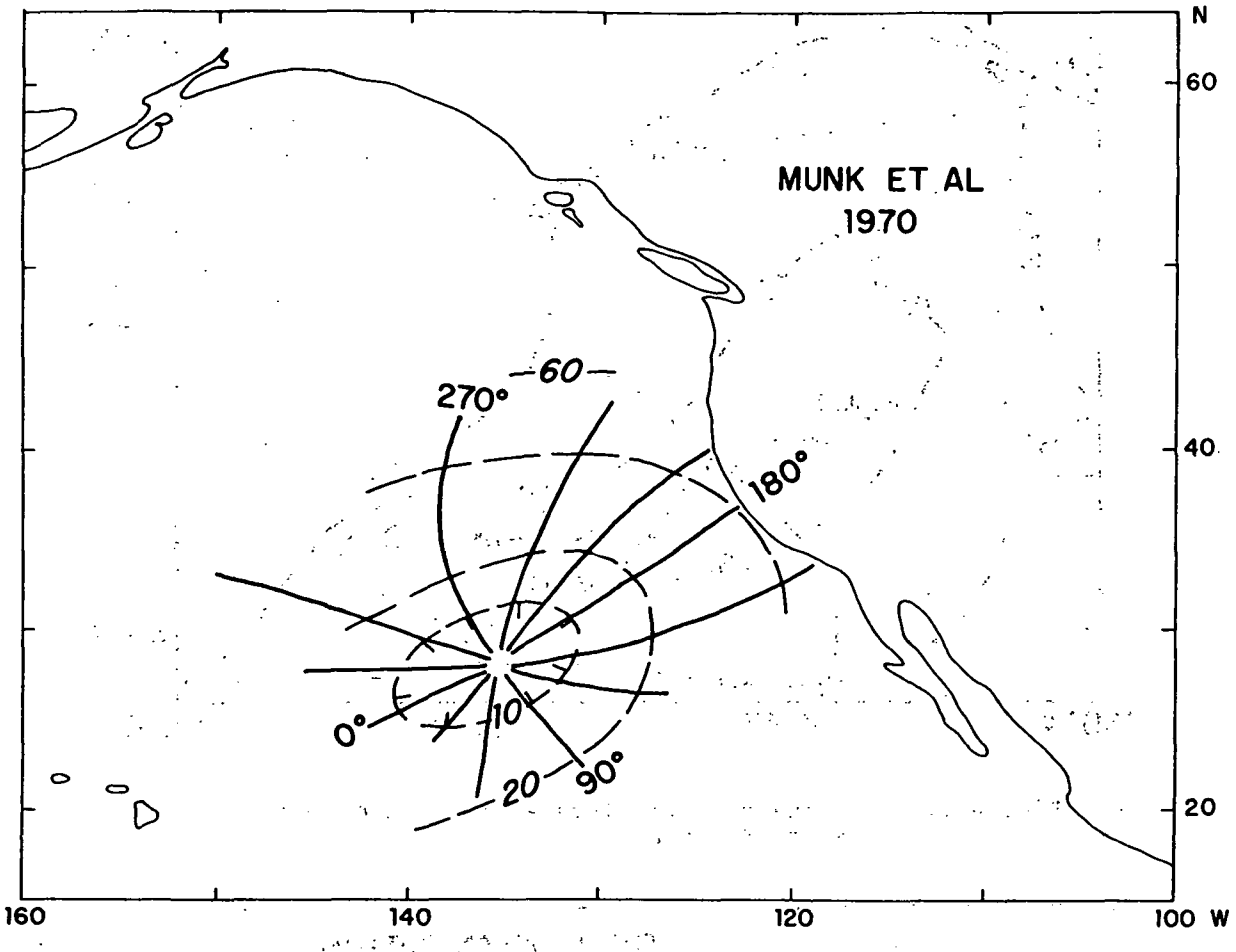



Figure 9. The $M_2$ ocean tide as derived by Munk et al. (1970) based on ocean-bottom measurements.

Kuo, J.T. and R.C. Jachens, "Indirect Mapping of ocean tides by solving the inverse problem for the tidal gravity observations," Annales de Geophysique, 33(1/2), 73-83, 1977.

Munk, W.H., F. Snodgrass and M. Wimbush, "Tide off-shore transition from California coastal to deep sea waters," Geophy. Fluid Dyn., 1, 161-235, 1970.

Nowroozi, A.A. and J.T. Kuo and M. Ewing, "Solid earth and oceanic tides recirded on the ocean floor off the coast of northern California," J. Geophy. Res., 74, 605-614, 1969.

Pekeris, C.L., and Y. Accad, "Solution of Laplace's equations for the $M_2$ tide in the world oceans," Phil. Trans. Roy. Soc. London, A. 265, 413-435, 1969.

Tiron, K.D., Y.N. Sergeev, and A.N. Michurin, Maps of tides in the Pacific, Atlantic and Indian Oceans, Vest Leningrad University 24, 123-135, 1967.

Zahel, W., Die reproduktion Gezeitenbedingter Bewengungsvorgange in Weltozean mittles des hydrodynamisch - numerischen Verfahrens, Mitt, Inst. Meereskunde der Univ. Hamburg, 17, 1970.

Zetler, B.D., Tide prediction, from Geophysical Predictions, Studies in Geophysics, National Research Council, National Academy of Science, 166-177, 1978.

# Tidal Friction in the Solid Earth

Jochen Zschau
Institute of Geophysics, Kiel University, Federal Republic of Germany

Abstract. The earth's imperfectly elastic response to body and loading tidal forces is discussed using complex Love numbers and complex mass loading coefficients. Exact analytical expressions have been derived relating the energy dissipation within an inhomogeneous, compressible solid earth to the surface values of these complex characteristic numbers, thus relating the global dissipation function Q to the phase shifts in the potential, gravity, tilt, strain and displacement tides.

Integration of a global ocean tidal model shows that energy dissipated in the solid earth due to ocean loading is at least 10 % of that dissipated in the body tide; however both body and loading tide together do not account for more than a few percent of the astronomically observed dissipation.

The commonly used relation $Q^{-1} = -\tan\varphi$ where $\varphi$ is the observed phase lag only applies when selfgravitation and hydrostatic prestress are ignored. It, therefore, is not applicable to the earth and in fact there is no unique relation between the global Q and the tidal phase shifts, this relation being very dependent on the distribution of Q with depth. Determinations of the global Q from satellite observations may be in error by 70 %, and calculations on the basis of seismic Q-models predict phase shifts in the gravity tide of only a few thousands of a degree in place of the currently predicted tenths.

Unlike in the body tide case, dissipation in the loading tide is sensitive to properties of the asthenosphere, and phase shifts in the $M_2$ loading tides in displacement, gravity and tilt may be as high as several degrees for loads near ocean ridges and subduction zones.

## Rate of Tidal Dissipation in the Solid Earth

Dissipation of body and loading tidal energy within the solid earth may be determined from its complex Love numbers and complex mass loading coefficients, respectively [Zschau, 1979 a, b] . Such the expressions for the dissipated energy turn out to be fairly simple,

$$\Delta E_n = - \frac{2\times n+1}{4\ GR}\ K_n \iint_S \psi_n^2\ dS \qquad (1)$$

for the body tide, and

$$\Delta E_n = \frac{2\times n+1}{4\ GR}\ (H_n^* - K_n^*) \iint_S \psi_n^{*2}\ dS \qquad (2)$$

for the loading tide, where $\Delta E_n$ is the energy dissipated during one cycle of harmonic loading, $K_n$ and $H_n^*$, $K_n^*$ are the imaginary parts of the surface Love numbers and mass load coefficients, respectively, $\psi_n$ is the amplitude of the body force potential, $\psi_n^*$ is the amplitude of the surface load potential, R is the earth's radius and G is the gravitational constant. n describes the degree of the expansion into spherical harmonics. The integration is taken over the surface of the earth. The expressions above have been obtained without approximating the real earth by an incompressible and homogeneous one as was necessary in former calculation, for instance by Munk & Mac Donald [1960] .

Using these formulas, and assuming the mantle Q structure LMS as obtained from the observation of the earth's free oscillation [see Smith, 1972] , a body tide solid dissipation rate of

$$3.19 \times 10^{17}\ \text{erg/s},$$

i.e. about 1 % of the astronomically observed dissipation rate, has been obtained. The corresponding computation for the loading tide dissipation rate within the earth's crust and mantle, carried out on the basis of a global $M_2$ ocean tide model [Hendershott, 1972] , yields a minimum value of roughly 10 % of the body tide solid dissipation rate, i.e.

$$3.21 \times 10^{16}\ \text{erg/s}.$$

This value has been determined from the low degree harmonics of the ocean tide distribution up to n=25, and, therefore, does not represent the high amplitude tides in the shelf areas. The latter may contribute significantly to the total dissipation, because the dissipated energy is proportional to the square of the marine tidal amplitude. On the other hand, the elastic strain energy stored in the mantle as well as the energy dissipated in the mantle decreases with increasing degree n of the spherical harmonic loading for $n > 5$ as may be seen from Fig. 1. This suggests that solid earth dissipation in the shelf areas does not change the total dissipation rate drastically. Anyway, both body tide - and loading tide dissipation together most probably do not account for more than a few percent of the astronomically observed dissipation rate.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

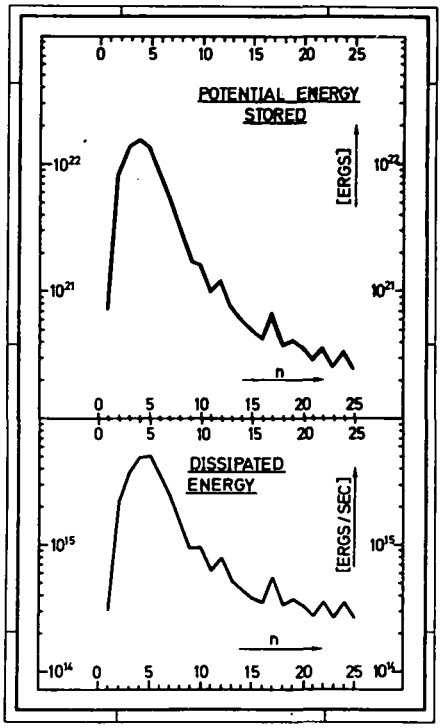



Fig. 1. Loading tide shear strain energy stored and dissipated within the solid earth [from Zschau, 1979b]. n is the degree of the spherical harmonic expansion of Hendershott's global $M_2$ ocean tide model [Hendershott, 1972]. Results are valid for the Gutenberg Earth with the free oscillation Q structure LMS for the mantle and the surface wave Q structure for the crust.

## Body Tide Phase Shifts and the Earth's Dissipation Function Q

In Zschau [1979b] the earth's global dissipation function Q has been related to the phase shifts in the potential, gravity, tilt, strain and displacement body tides. The following expressions valid for constant Q distribution within the earth have been obtained:

Tangential displacement: $\tan\varphi = L/l = -\,0.897\ Q^{-1}$

Potential variation due to tidal deformation: $\tan\varphi = K/k = -\,0.612\ Q^{-1}$

Radial displacement: $\tan\varphi = H/h = -\,0.555\ Q^{-1}$

Surface areal strain: $\tan\varphi = \frac{H-3L}{h-3l} = -\,0.318\ Q^{-1}$

Tilt: $\tan\varphi = \frac{K-H}{1+k-h} = +\,0.217\ Q^{-1}$

gravity variation: $\tan\varphi = \frac{H-\frac{3}{2}K}{1+h-\frac{3}{2}k} = -\,0.0508\ Q^{-1}$

(h,H), (k,K), and (l,L) complex Love numbers; the index n has been omitted here.

| | | CONVENTIONAL | CORRECT | |
|---|---|---|---|---|
| | OBSERVED DELAY ANGLES | $1/\tan\alpha_{h-\frac{3}{2}k}$ | $Q_{gl}$ (CONSTANT Q) | $Q_{gl}$ (LOW Q ZONE) |
| RECORDING GRAVIMETERS | 0.1° | 79 | 28 | 5.5 |
| | 0.2° | 39 | 14 | < 3 |
| SATELLITE OBSERVATIONS | | $1/\tan\alpha_k$ | | |
| | 0.5° | 60 | 36 | 34 |

TABLE 1. Conventional and correct global Q values ($Q_{gl}$) corresponding to tidal delay angles of gravity and satellite observations.

The above equations show that the commonly used relation

$$Q^{-1} = -\tan\varphi \qquad (3)$$

where $\varphi$ is the observed phase shift is not applicable in the earth tide case. For the tidal gravity variation this had already been pointed out by Slichter [1960] who instead of (3) introduced the formula

$$Q^{-1} = -\frac{\delta}{\delta - 1}\tan\varphi \qquad (4)$$

where $\delta$ is the gravimetric $\delta$-factor, and $\varphi$ is the gravity phase shift. With $\delta = 1.16$ this yields

$$\tan\varphi = -\,0.138\ Q^{-1} \qquad (5)$$

which is still too large by more than a factor of 2 as is obvious from the comparison of this formula with the corresponding one given above. One may show that Slichter's formula is equivalent to

$$Q^{-1} = \frac{H-\frac{3}{2}K}{h-\frac{3}{2}k} \qquad (6)$$

which does not correspond to the basic definition of $Q^{-1}$ as the strain energy dissipated during one cycle of loading over $2\pi$ the peak energy stored in the system. The general expression for the earth's global dissipation function Q in terms of its complex Love numbers is, however,

$$Q^{-1} = \frac{Kh - H\times(1+k)}{h\times(1+k-h)} \quad \text{[Zschau, 1979b]} \qquad (7)$$

This equation is valid for an incompressible body with homogeneous density. None of the above phase shifts is equivalent to this expression, hence in general

$$Q^{-1} \neq -\tan\varphi \qquad (8)$$

As shown in Zschau [1979b] the reason for this is that in the case of tidal deformations, selfgravitation and hydrostatic prestress cannot be ignored.

Besides the fact that (3) is not applicable in the tidal case, there, furthermore, is no unique relation between the global Q and the tidal phase shifts, this relation being very dependent on the distri-

bution of Q with depth. For instance, a low Q zone in the upper mantle such as given in model LMS [see Smith, 1972] may alter the ratio between observed tidal phase shifts and the global (average) Q up to a factor of 5.

Table 1 gives some examples for the errors involved when not taking account of the above aspects: Let the phase delay of the body tide gravity be $0.1^{\circ}$ with respect to the external forces. Proceeding in the conventional manner, i.e. using Slichter's formula with $\delta = 1.16$ gives the wrong global body tide $Q = 79$. In the case of constant Q within the mantle, we find the correct value to be $Q = 28$. If we assume a low Q asthenosphere, i.e. let the real Q distribution within the earth differ by only a constant factor from the free oscillation Q model LMS, we find that the global Q has to be chosen as low as 5.5 to correspond to the gravity phase delay of $0.1^{\circ}$. This is less than 7 % of the value $Q = 79$, obtained by the conventional method. For the same reason it turns out that the body tide gravity phase delay due to friction within the solid earth amounts more likely to a few thousandth of a degree than to a few tenth of a degree as expected so far. We, therefore, suggest that the average O1 gravity phase shift of $-0.2^{\circ}$ as observed for Europe may not be attributed to imperfect elasticity in the mantle like it is proposed by Melchior et al. [1976], but rather to the indirect effect of the $O_1$ tide in the oceans. There seems to be no chance at all at the moment to get information on the mantle Q from body tide gravity investigations.

The delay angle of the potential bulge due to tidal deformation of the earth has been determined from the orbits of artificial satellites to be $0.5^{\circ}$ [Lambeck et al., 1974] . Lambeck et al. relate this delay angle to a mantle Q of 60 which one obtains by using formula (3). The correct Q corresponding to this delay angle is 36, if the Gutenberg earth and constant Q values within the mantle are adopted. For the LMS equivalent model, we have calculated the global Q which corresponds to the delay angle of $0.5^{\circ}$ to be 34. This shows that the delay angle of the tidal potential bulge is less sensitive to the geometry of the Q distribution within the earth than the phase delay of the tidal gravity at the deformed surface as measured by a gravimeter. The Q values of 34 and 36 are much lower than the lowest limit of the upper mantle seismic Q.

The usage of the exact theoretical relationship between the bulge of the tidal potential and the body's global Q may also be important for other planets as for instance for Mars. From observations of the secular acceleration of the Mars satellite Phobos, Smith and Born [1976] deduced the phase angle of the potential bulge due to the body tide of Mars. They related this phase angle to the global Q of Mars by the simple formula (3), and found a global Q between 50 and 150. However, as one cannot neglect selfgravitation and the hydrostatic prestress for Mars, equation (3) is not applicable, and, therefore, the Q between 50 and 150 is probably too high for Mars, provided that the observed phase angle of the tidal bulge is true.

Similar considerations as above may also be important for the determination of the lunar global Q from observations of its physical librations [see Yoder, 1978] .

## The Effect of Imperfect Mantle Elasticity on Loading Tides

Unlike in the body tide case, also high degree harmonics of the load are important in the loading tide case. It is found that for the free oscillation Q model LMS, the loading tides of harmonic degrees slightly less than 100 are strongly effected by the low Q asthenosphere, the global loading tide Qs being up to nearly 7 times smaller than the global body tide Q (see Fig. 2). Correspondingly, the loading tide phase shifts due to imperfections in the elasticity of the mantle are by more than one order of magnitude higher than those of the body tide. Near ocean ridges and near subduction zones the $M_2$ loading tide phase shifts may even be as high as a several degrees for the displacements as well as for gravity and tilt, provided the Maxwell constitutive law is valid. From the computation of phase shift Green's functions it is obvious that these maximum phase shifts occur at about 80 to 100 km distance from the load, this distance depending on the depth of the assumed low viscous asthenosphere (see Fig. 3). Our numerical results suggest that loading tide investigations could become an effective tool for studying the upper mantle viscosity in regions where viscosities lower than $10^{19}$ Poise may be expected.

Acknowledgement: A portion of this study was completed while the author was a visiting fellow at the Cooperative Institute for Research in Environmental Sciences, Boulder, Colorado.

## References

Hendershott, M.C., The effects of solid earth deformation on global ocean tides, Geophys. J.R. astr. Soc. 29, 389, 1972.

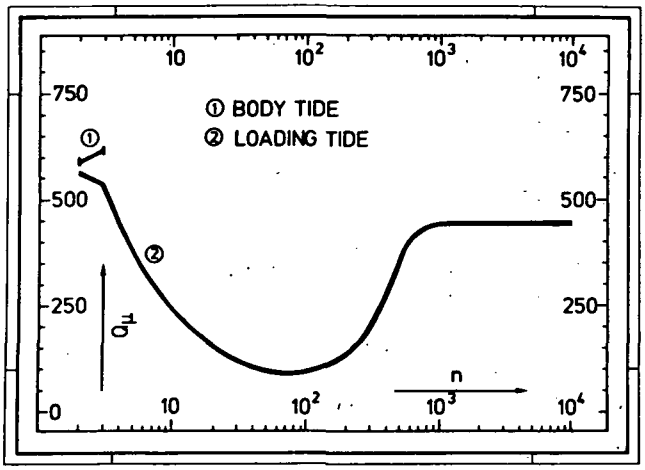



Fig. 2. Body tide Q and loading tide Q for different degrees n of the spherical harmonic expansions of the tides [ from Zschau, 1979b ]. Results are valid for the Gutenberg Earth with the LMS Q structure from free oscillation data for the mantle, and the MM 8 Q structure from surface wave data for the crust. The low Q values for n slightly lower than 100 in the loading tide case, are due to the low Q asthenosphere in model LMS. The high Q values for higher n and for lower n are due to the high Q values in the crust and in the lower mantle, respectively.

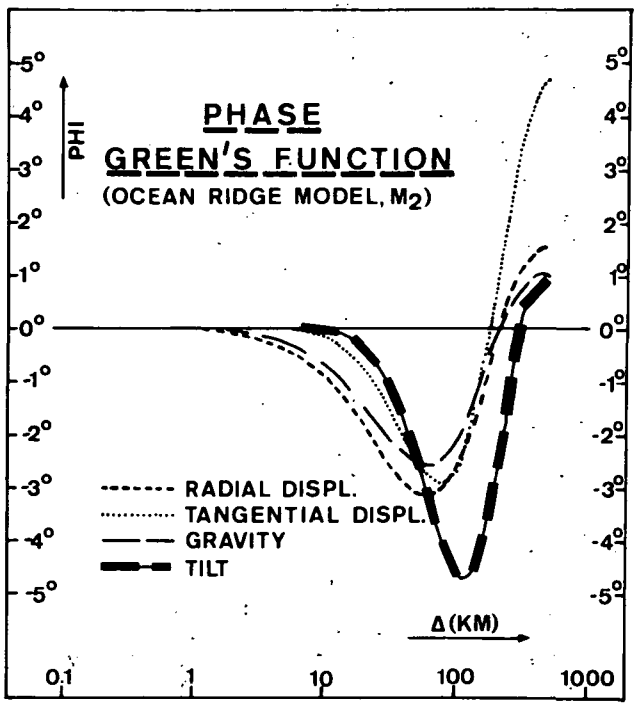



Fig. 3. $M_2$ phase Green's functions for an ocean ridge model [ from Zschau, 1979a,b ]. They represent the phase shifts between the viscoelastic earth's responses to a varying point load, and the responses of the corresponding elastic earth. The high phase shifts at distances of about 100 km from the load are due to the low viscous asthenosphere. Δ : distance from the point load, PHI: phase shift.


Lambeck, K., Cazenave and G. Balmino, solid earth and ocean tides estimated from satellite orbit analysis, Rev. Geophys. Space Phys. 12, p. 421, 1974.

Melchior, P., J.T. Kuo, and B. Ducarme, earth tide gravity maps for Western Europe, Phys. Earth Planet, Interiors, 13, p. 184, 1976.

Munk, W.H., and G.J.F. MacDonald, The rotation of the earth, Cambridge University Press, New York, 1960.

Slichter, L.B., in P. Melchior: Compte Rendu des Réunions de la Commission Permanente des marées terrestres à l'Assemblée Générale d'Helsinki, Marées Terr. Bull. Inf. 21, p. 369, 1960.

Smith, S.W., The anelasticity of the mantle, in: The upper mantle, edited by A.R. Ritsema, Tectonophysics 13, p. 601, 1972.

Smith, J.C., and G.H. Born, secular acceleration of Phobos and Q of Mars, Icarus 27, p. 51, 1976.

Joder, C.F., Effects of the Spin-Spin interaction and the inelastic tidal deformation on the Lunar physical librations, 1979, in print (Preprint from JPL, Pasadena, California 91103).

Zschau, J., Phase shifts of tidal sea load deformations of the earth's surface due to low viscosity layers in the interior, Proceed. 8th Int. Symp. on Earth Tides held at Bonn 1977, in print (1979 a)

Zschau, J., Tidal friction in the solid Earth: Loading tides versus body tides, in: P. Brosche and J. Sündermann, Tidal friction and the earth's rotation, Springer 1979b, in print.

# Body Tides on an Elliptical Rotating Earth

John M. Wahr
CIRES, University of Colorado
Boulder, Colorado 80309

Abstract. The complete tidal response of an elliptical, rotating, elastic Earth is found to contain small displacements which do not fit into the conventional Love number framework. Corresponding observable tidal quantities (gravity, tilt, strain, Eulerian potential, etc.) are modified by the addition of small (≲1%) latitude dependent terms.

## Introduction

Traditional body tide calculations assume a spherically symmetric non-rotating Earth, with surface motion at a given frequency described by a small set of 'Love numbers', principally h, k and ℓ. Simple linear combinations of these numbers are found to compare with experiment.

The spherically symmetric case has been extended by others (Jeffreys and Vicente, 1957a, 1957b; Molodensky, 1961; Shen and Mansinha, 1976; Sasao *et al.* 1978) to include effects of non-sphericity on the diurnal tides. Particular attention has been focused on the dynamical effects of a rotating fluid core with an accompanying elliptical core-mantle boundary. The mantle has been repeatedly modelled as a spherically symmetric, non-rotating elastic shell, and the fluid core assumed to be spherically stratified. In each of these expanded treatments, the computed tidal response consists of the conventional body tide, represented at the free surface by the Love numbers h, k and ℓ, together with a toroidal component describing the Earth's nutational behavior. In this way the diurnal Love numbers are modified for the most important effects of rotation and ellipticity. The nutations, traditionally computed for a rigid Earth (Woolard, 1953; Kihoshita, 1977) are also improved by including spherically stratified elastic behavior throughout the Earth.

## Qualitative Effects of Rotation and Ellipticity

The work described below extends these results by also including the effects of rotation in the inner core and mantle, together with elliptically stratified, realistic material properties throughout the Earth. Numerical investigations suggest a computed surface solution accuracy of at least three parts in a thousand, the order of the Earth's ellipticity. All important tidal bands are considered.

It is found that for the complete rotating, elliptical problem, surface motion is too complex to be represented by the familiar Love numbers: h, k and ℓ. Even the inclusion of nutational motion into the diurnal solution and length of day changes into the long period solution is not sufficient. Other small (≲1%) unmodelled motions remain. These new terms, which will not be detailed here, can be handled by defining more Love numbers. The utility of this procedure, however, is not always maximal. It is, of course, physical observables which are the ultimate computational goal. For an elliptical, rotating mantle the relations between an expanded Love number set and these observables can be disturbingly complex. Consequently, it is often more useful to present results for the observational quantities, directly.

As an example, consider the tidal gravity signal as measured by a standard gravimeter. For a second degree semi-diurnal tide, it has the form

$$\text{Gravity} = -\frac{2}{r_o}\, g\, V_{n=2}(\omega)\left[\delta_2(\omega)\, Y_{n=2}^{m=2}(\theta,\phi) + \delta_4(\omega)\, Y_4^2(\theta,\phi)\right] e^{i\omega t} \qquad (1)$$

where ω is the perturbing frequency, $Y_n^m$ a surface spherical harmonic, $r_o$ the Earth's mean radius (6371 km), g the computed gravity acceleration at the Earth's equator assuming a spherical density distribution ($g = 979.8259$ cm/sec$^2$), $V_n(\omega)$ the appropriate term in the luni-solar potential as observed at the Earth's equator, and $\delta_2(\omega)$ and $\delta_4(\omega)$ are gravimetric factors which reflect the dynamical behavior of the Earth at the frequency, ω. (For a spherical Earth $\delta_4 = 0$).

It is convenient to think of eq. (1) as representing a latitude dependent gravimetric factor:

$$\delta = \delta_2 + \delta_4 \frac{Y_4^2(\theta,\phi)}{Y_2^2(\theta,\phi)} \qquad (2)$$

Similar latitude dependent phenomenon are present in all tidal bands and for all observable quantities (i.e.), surface gravity, tilt, strain, latitude and longitude variations, tidal perturbations in inertial space gravity, inertial space station displacements, etc. They are found at about the one per cent level or less. For example, computations of the $M_2$ semi-diurnal gravimetric factors shown in eq. (1) give, for Earth model PEM-C (Dziewonski, *et al.*, 1975):

| | | |
|---|---|---|
| spherical case: | $\delta_2 = 1.159$ | $\delta_4 = 0$ |
| rotating, elliptical case: | $\delta_2 = 1.160$ | $\delta_4 = -.005$ |

Acknowledgements. This work is supported under NASA grant NSG 7319.

## References

Dziewonski, A. M., A. L. Hales, and E. R. Lapwood, Parametrically simple Earth models consistent with geophysical data, Physics of the Earth and Planetary Interiors, 10, 12-48, 1975.

Jeffreys, H. and R. O. Vicente, The theory of nutation and the variation of latitude, Mon. Not. R. Astr. Soc., 117, 142-161, 1957a.

Jeffreys, H. and R. O. Vicente, The theory of nutation and the variation of latitude: the Roche Model Core, Mon. Not. R. Astr. Soc., 117, 162-173, 1957b.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280. The Ohio State Univ., Columbus, Ohio 43210.

Kinoshita, H., Theory of the rotation of the rigid Earth, Celestial Mech., 15, 277-326, 1977.

Molodensky, M. S., The theory of nutation and diurnal Earth tides, Comm. Obs. R. Belgique, 288, 25-56, 1961.

Sasao, T., S. Okubu and M. Saito, A simple theory on dynamical effects of stratified fluid core upon nutational motion of the earth, to be published in the Proceedings of IAU Symposium No. 78 "Nutation and the Earth's Rotation", Kiev, May 1977), in press. 1978.

Shen, P-Y. and L. Mansinha, Oscillation, nutation and wobble of an elliptical rotating Earth with liquid outer core, Geophys. J. R. Astr. Soc., 46, 467-496, 1976.

Woolard, E. W., Theory of the rotation of the earth around its center of mass. Astronomical Papers for the American Ephemeris and Nautical Almanac (Gov't. Printing Office, Washington, DC, 1953), Vol. XV,Pt.1, 1953.

# Observation of the Nearly Diurnal Resonance of the Earth Using a Laser Strainmeter

Judah Levine *
Time and Frequency Division, National Bureau of Standards
Boulder, Colorado 80302

Abstract. We have used two years of strain tide data to study the response of the Earth to the diurnal and semidiurnal tidal excitations. Our results show that there is significant structure in the response of the earth to tidal excitations near one cycle/sidereal day. This structure agrees with the resonance behavior predicted from the calculations of the forced elastic-gravitational response of an elliptical, rotating earth with a liquid outer core. The data can also be used to test for possible preferred frames and spatial anisotropies. We find that upper bounds on the parameterized post-Newtonian (PPN) parameters which characterize these effects are $\alpha_2 \leq 0.007$ and $\zeta_w \leq 0.005$.

## Introduction

We have analyzed approximately two years of strain-tide data obtained using the 30-meter laser strainmeter we have previously described (Levine and Hall, 1972).

## Data Acquisition

The data were obtained using a 30-meter long laser strainmeter located in the Poorman Mine, an unworked gold mine located approximately 8 km west of Boulder, Colorado at latitude 40.03°N, and longitude 254.67°E.

The heart of the strainmeter is an evacuated 30-meter Fabry-Perot interferometer located along the length of the tunnel. The axis of the interferometer is 7° west of North.

The interferometer is illuminated by a 3.39-μm helium-neon laser. A servo loop piezoelectrically tunes the laser to keep its wavelength coincident with one of the transmission maxima of the long interferometer. The frequency of the laser, f, is therefore related to the length of the interferometer, L, by

$$f = \frac{nc}{2L}$$

where n is an integer and c is the velocity of light. Thus

$$\frac{\Delta f}{f} = -\frac{\Delta L}{L} \quad .$$

A second 3.39-μm laser is stabilized using saturated absorption in methane. The beat frequency between the two lasers is extracted for further processing. Then

$$\frac{\Delta f_{beat}}{f} = \frac{\Delta L}{L}$$

or

$$\Delta f_{beat} = 8.85 \times 10^{13} \frac{\Delta L}{L}$$

so that a measurement of the fluctuations in the beat frequency provides a direct measurement of the fractional change in the length of the long path. The relationship between beat frequency and strain has no adjustable constants or calibration factors.

The beat frequency is digitally recorded 10 times/hour along with other information including local barometric pressure, etc. The digital data are bandpass filtered using a symmetric convolution filter and then decimated to one sample every two hours for comparison with theory.

## Data Analysis

We have used all of the components published by Cartwright and Edden (1973) in our analysis. For each component with frequency $f_k$ we construct a time series of the form

$$a_k \cos(2\pi f_k t + \phi_k + \alpha_k) Y_n^m(\theta,\varphi) T(n,m,\theta,\varphi,\theta_s)$$

where $a_k$ is the amplitude given by Cartwright and Edden, $\phi_k$ is the phase of the component, $Y_n^m$ is the spherical harmonic computed at the station co-latitude $\theta$ and East Longitude $\varphi$. The quantity T is a function converting potential to strain along the axis of the strainmeter $\theta_s$ (Levine and Harrison, 1976). The quantity $\alpha_k$ is -90° if (n+m) is odd and is zero otherwise (Cartwright and Tayler, 1971).

The theoretical series were further modified by a function to correct for local topography, local crustal inhomogeneities, cavity effects and ocean loads (Levine and Harrison, 1976). We assumed initially that these effects do not vary rapidly with frequency, so that all of the diurnal components have the same correction as $O_1$, and that all of the semidiurnal components have the same correction as $M_2$.

The terms are then grouped by frequency. Each frequency group contains all of the terms (regardless of parentage) which differ by less than one cycle/year from each other.

This process produces 48 time series. We fit these series to our data by the least-squares method using an adjustable amplitude and an adjustable phase for each cycle/year group.

The results of this process are shown in Fig. 1 for the diurnal amplitude, Fig. 2 for the diurnal phase, and Fig. 3 for the semidiurnal amplitude and phase. The error bars represent one standard deviation and are obtained from estimates of the

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978*, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

*Fellow, Joint Institute for Laboratory Astrophysics of the National Bureau of Standards and the University of Colorado.

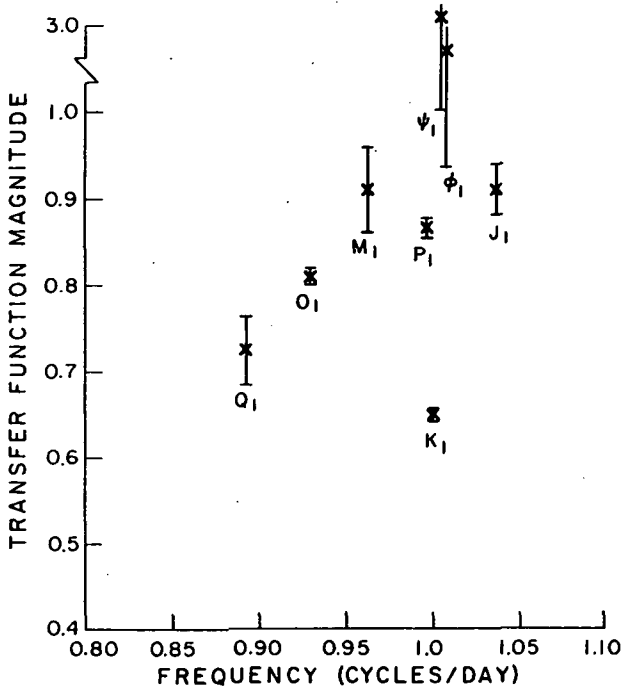



Fig. 1. Normalized transfer function for the diurnal tides. For clarity only the major components are plotted. The error bars are one standard deviation.

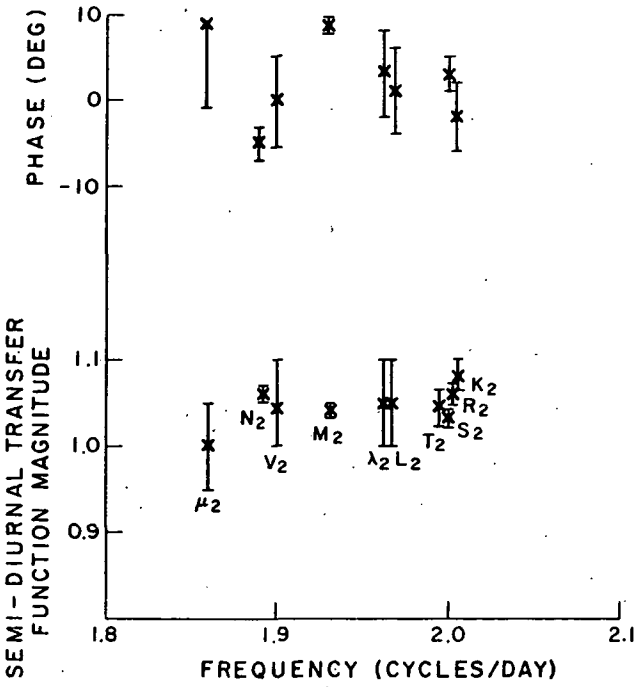



Fig. 3. Normalized transfer function amplitude (lower curve) and phase (upper curve) for the semidiurnal tides. For clarity only the major components are plotted. The error bars are one standard deviation.

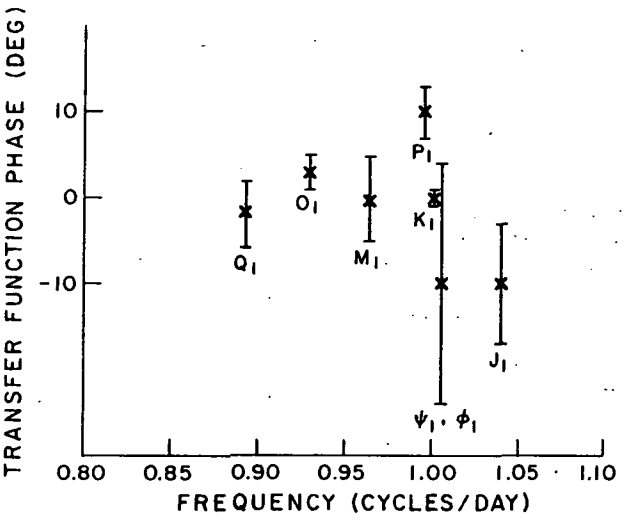



Fig. 2. Normalized transfer function phase for the diurnal tides. For clarity only the major components are plotted. The error bars are one standard deviation.

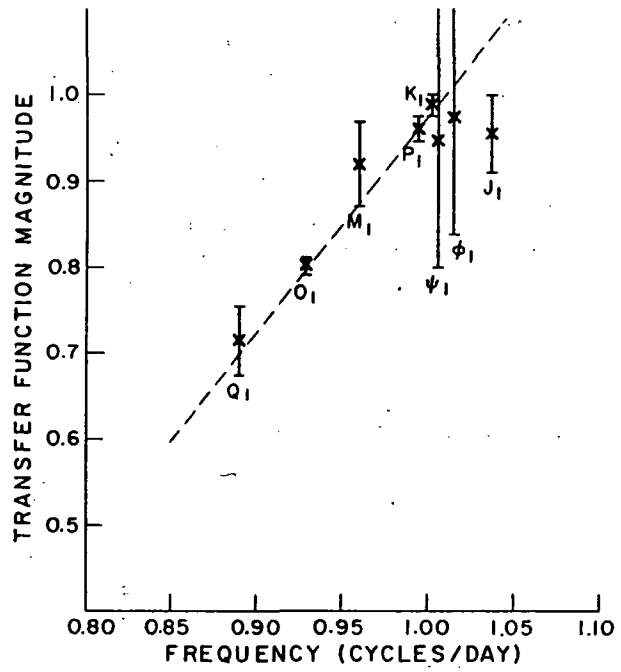



Fig. 4. Normalized transfer function amplitude for the diurnal tides when the frequency-dependent Love numbers are inserted into the fitting function. The dashed line is obtained by fitting a straight line to the transfer function amplitudes weighted by their respective uncertainties.

noise.

The amplitude of the diurnal transfer function shows a statistically significant structure. There appears to be a significant decrease in the transfer function for the lower frequency diurnal components. We attribute this to a slow change in the contribution of the ocean load to our observed data.

More significantly there is a dip in the transfer function near one cycle/day (we have not plotted the $S_1$ amplitude at one cycle/day since it is heavily contaminated by thermoelastic processes). This dip is consistent with the effects of the nearly diurnal resonance associated with the liquid core. The consistency can be shown more clearly by inserting the resonance directly into the fitting function. This may be done using the frequency dependent Love numbers published by Molodensky (1961) and Shen and Mansinha (1976). These Love numbers produce changes in the transfer function at all of the diurnal frequencies but, except near one cycle/day, the new transfer function lies within one standard deviation of the old one. The modified transfer function is shown in Fig. 4. As can be seen from the figure the resonance models account for the dip in the transfer function near one cycle/day. Unfortunately, the data cannot be used to infer the fine structure near the resonance. Furthermore, the insertions of realistic estimates for the energy loss in the nearly diurnal band may affect the shapes predicted by all of the models somewhat. It is unlikely we will be able to test any of these effects using our data. The strongest test for the resonance obtainable from our data comes from the components $P_1$ and $K_1$ which are really on the tail of the resonance function. Our measurements at $\phi_1$ and $\psi_1$ (where the resonance has the largest effect) have error bars whose size probably does not permit them to be used in a quantitative comparison between theory and experiment or in attempts to differentiate among the various calculations of the effect of the resonance.

These results may be compared with the analysis of Warburton and Goodkind (1977) using data from a cryogenic gravimeter, with the gravimeter data analysis of Abours and Lecolazet (1978) and with the analyses of Lecolazet and Melchior (1977). Their results are generally in agreement with ours, confirming the general shape of the resonance. None of the analyses is able to make a quantitative comparison with theory because of the relatively poor signal-to-noise ratio in the measurements of the crucial components $\phi_1$ and $\psi_1$.

The semidiurnal transfer function amplitude shows far less structure, and the agreement between experiment and theory is quite good. There is no evidence of anomaly at $S_2$ (two cycles/day) confirming that the anomaly at $S_1$ is almost certainly of thermoelastic origin.

We may place upper limits on the various PPN parameters by calculating the magnitudes of the anomalous tidal components in terms of the PPN parameters $\alpha_2$ and $\zeta_w$. In this way we conclude $\alpha_2 \leq 0.007$ and $\zeta \leq 0.005$.

## Conclusions

These results confirm the general correctness of the various published earth models. From the point of view of the current discussion they are also significant in that they illustrate the sort of measurements that can be performed with laser strainmeters. These measurements show that we can make meaningful measurements of diurnal strain changes at the $10^{-10}$ level, and that almost continuous operation for several years is possible.

From the geodetic point of view, the limitation on the utility of laser strainmeters arises from their sensitivity to local effects, especially to spurious motions near the piers.

Spurious motions of the piers at the level of millimeters will even play a significant role in geodetic measurements made over much longer baselines. Motions of this magnitude represent fractional changes of parts in $10^8$ even over 50 km baselines, so that such effects will make significant contributions to the error budget of any geodetic instrument now in operation or under construction.

It is important to compare the data obtained with laser strainmeters with measurements obtained using electromagnetic distance measuring equipment operating over parallel baselines. If this comparison shows that the two provide a consistent picture of the regional strain field, then laser strainmeters may prove useful in measuring regional strain at fixed observatories. In this service they can provide significantly higher sensitivity than any other technique.

## References


Abours, S. and R. Lecolazet, New results about the dynamical effects of the liquid outer core as observed in Strasbourg, Proc. 8th Int. Conf. on Earth Tides, 1978 (in press).

Berger, J. and J. Levine, The spectrum of the earth noise from $10^{-8}$ to $10^{+2}$ Hz, J. Geophys. Res., 79, 1210-14, 1974.

Cartwright, D.E. and A.C. Edden, Corrected tables of tidal harmonics, Geophys. J. Roy. Astr. Soc., 33, 253-2640, 1973.

Cartwright, D.E. and R.J. Tayler, New computations of the tide-generating potential, Geophys. J. Roy. Astr. Soc., 23, 45-74, 1971.

Lecolazet, R. and P. Melchior, Experimental determination of the dynamical effects of the liquid core of the earth, Ann. Geophys., 33, 11-22, 1977.

Levine, J. and J.L. Hall, Design and operation of a methane absorption stabilized laser strainmeter, J. Geophys. Res., 77, 2595-2609, 1972.

Levine, J. and J.C. Harrison, Earth tide strain measurements in the Poorman Mine near Boulder, Colorado, J. Geophys. Res., 81, 2543-55, 1976.

Molodensky, M.S., The theory of nutations and diurnal earth tides, Comm. Obs. R. Belgique, 188, 25-56, 1961.

Shen, Po-Yu and L. Mansinha, Oscillation, nutation and wobble of an elliptical rotating earth with liquid outer core, Geophys. J. Roy. Astr. Soc., 46, 467-496, 1976.

Warburton, R.J. and J.M. Goodkind, Detailed gravity tide spectrum between 1 and 4 cycles per day, Geophys. J. Roy. Astr. Soc., 52, 117-130, 1977.

# The Influence of Earth Tides on Earth's Coordinates

R. O. Vicente
Department of Applied Mathematics, Faculty of Sciences
Lisbon, Portugal

The importance of Earth tides on Earth coordinates has been considered only recently for several reasons: 1) the precision that we are obtaining nowadays for Earth's coordinates shows that the effects of Earth tides appear on the values obtained for the coordinates; 2) the possibility of determining, by observations, the values of Earth tides; 3) the consideration of theoretical models that can compute the values of Earth tides.

We are going to examine briefly some of these reasons. First of all, it should be pointed out that we have to be careful about the definitions of what we mean by Earth coordinates.

When it was only possible to obtain a precision of a few meters for the values of the Earth coordinates, some of the systems of reference adopted in the dynamics of the Earth rotation could not be distinguished within that precision. But when we are aiming at precisions of a few centimeters, it is very important to define carefully the systems of reference employed for our coordinates on the Earth.

We can consider, for instance, astronomical and geodetic coordinates of a point at the Earth's surface.

The astronomical coordinates are referred to the astronomical local vertical, and therefore, to the instantaneous axis of rotation of the Earth, briefly called the axis of rotation. The intersection of this axis with the Earth's surface defines the geographic poles. The complement of the acute angle between the astronomical vertical and the axis of rotation of the Earth is called the astronomical latitude. For the definition of astronomical longitude we need a reference meridian passing by the geographic poles [Woolard and Clemence, 1966].

Considering the case of the real Earth, there are known irregularities in the direction of the astronomical vertical from place to place over the Earth, and, therefore, the astronomical meridians, parallels of latitude and the equator are irregular curves of double curvature, but they do not depart very far from plane curves.

The instruments (visual zenith telescopes (VZT), photographic zenith telescopes (PZT) and astrolabes), employed in the classical techniques of determining polar motion, refer their observations to the astronomical vertical.

Depending on the adopted pole and reference meridian, we can have different definitions of the coordinates of a point on the Earth's surface. The danger, nowadays, is that different people speak about different poles and, therefore, the coordinates they are speaking about are not the same.

The simplest Earth model considers the Earth as a solid rigid body. Even this simple model shows the very important distinction between free and forced motions, originating periodic displacements of the axes employed in the dynamics of the Earth rotation. They have been called nutations because the displacements are periodic.

The equations of motion can be represented by the vector equation

$$\frac{d\vec{H}}{dt} = \vec{G}$$

showing that the time derivative of the angular momentum vector $\vec{H}$ around the centre of mass is equal to the vector moment $\vec{G}$ of the external forces, mainly due to the Sun and Moon.

When there are no external forces acting on the Earth $\vec{G}=0$, therefore, $\vec{H}$ is constant and fixed in space. We have the so-called free nutation because this type of motion exists even without external forces. The important consequence is the fact that any free nutation changes the Earth coordinates showing, for instance, variations of latitude, but it does not affect star places on the celestial sphere. The name wobble has, unfortunately, been recently employed to designate the free nutations.

When there are external forces acting on the Earth, $\vec{G}\neq 0$, and the position of $\vec{H}$ is not fixed in space. We obtain the forced nutations that are so important in any problem dealing with the transformation of reference systems from one epoch to another because they affect star places.

One aspect of the actions of the Sun and the Moon on the Earth concerns the tidal attractions of the luni-solar forces, which deform the Earth, giving rise to the tides of the solid Earth.

Different components of the Earth tide produce the nutations that are important in astronomy, and the forced nutations correspond to diurnal tides, so there is a close connection between nutations and tides.

It is fortunate that all the studies concerned with the tides of the solid Earth employ the same standard, that is, the bodily tide numbers $\underline{h}$, $\underline{k}$ and $\underline{\ell}$ are defined considering a statical theory applied to an elastic solid Earth with spherical symmetry, and the disturbing tidal potential is a spherical harmonic of the second degree [Jeffreys, 1976].

The value of $\underline{k}$ is related to the period of the free Eulerian nutation, and the value of $\Lambda=1+k-\ell$ is influenced by disturbances that affect the position of the astronomical vertical of a place on the Earth, that is, there are variations in the geographic coordinates of a place with periods depending on the periods of the tidal forces.

The actions of the Sun and Moon deflect the position of the astronomical vertical, and the maximum deviation is of the order of 0".05; they are the most regular of the variations in direction of the astronomical vertical. There are remaining variations due to geophysical causes, not yet very well known, and the tidal variations are very much modified by local geophysical factors.

In practice, the determination of the values of the bodily tide numbers is more difficult because there are local corrections.

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

TABLE 1. Diurnal Waves

| | Jeffreys-Vicente 1957 | | Molodensky 1961 | Shen-Mansinha 1976 | Observations |
|---|---|---|---|---|---|
| | Central Particle | Roche | 2 | $\beta_o$ | |
| | | | $\delta=1+h-\frac{3}{2}k$ | | |
| $K_1$ | 1.183 | 1.185 | 1.151 | 1.1380 | 1.1406 A.M.<br>1.1507 W.M. |
| $P_1$ | 1.209 | 1.172 | 1.161 | 1.1539 | 1.1699 A.M.<br>1.1664 W.M. |
| $O_1$ | 1.221 | 1.211 | 1.166 | 1.1596 | 1.1522 A.M.<br>1.1676 W.M. |
| | | | $\gamma=1+k-h$ | | |
| $K_1$ | 0.714 | 0.693 | 0.721 | 0.7340 | 0.7422 A.M.<br>0.7501 W.M. |
| $P_1$ | 0.676 | 0.696 | 0.695 | 0.7019 | 0.7068 A.M.<br>0.7167 W.M. |
| $O_1$ | 0.658 | 0.658 | 0.686 | 0.6895 | 0.6785 A.M.<br>0.6752 W.M. |
| | | | $\Lambda=1+k-\ell$ | | |
| $K_1$ | 1.120 | 1.162 | | 1.1689 | |
| $P_1$ | 1.149 | 1.180 | | 1.2033 | |
| $O_1$ | 1.160 | 1.183 | | 1.2055 | |

A.M. Arithmetic Mean
W.M. Weighted Mean

It is convenient to have the possibility of determining the values of $\underline{k}$ and $\underline{\ell}$ by processes that do not depend on any geophysical assumptions. This fact shows the importance of having good determinations of the period of the free Eulerian nutation, called the Chandler period, that permit the determination of the value of $\underline{k}$, and the advantage of careful analysis of the observations of variation of latitude that give the possibility of computing $(1+k-\ell)$. Some of the modern techniques that can observe polar motion have an important role to play in these studies.

The observations of Earth tides are interpreted in amplitude ratio form (observed to theoretical amplitude), and we can consider, for instance, the following linear combinations for the bodily tide numbers:

for the horizontal component $=1-h+k$
for the vertical component $=1+h-\frac{3}{2}k$

The computed values of Earth tides show that the maximum effect will correspond to a radial deformation of about 30 cm. We can see the importance of Earth tides for the present day goal of achieving precisions of a few centimeters on the Earth's coordinates.

It has been demonstrated that the application of a statical theory for the semi-diurnal, fortnightly and semi-annual tides does not introduce appreciable errors in theoretical models [Jeffreys, 1949]. But for diurnal tides we have to consider a dynamical theory and the importance of the liquid core of the Earth has been proved. The main waves are:

| Doodson's code number | Tidal component | Nutation component |
|---|---|---|
| 165.555 | $K_1$ | Precession |
| 163.555 | $P_1$ | Semiannual |
| 145.555 | $O_1$ | Fortnightly |

The observations of Earth tides show a very irregular distribution over the Earth. The majority of stations are concentrated in Europe and North America. Local effects have been detected, for instance, due to the proximity of oceans and big rivers.

It is fortunate that there is an "International Centre for Earth Tides" where all the observations are collected and the computations performed, therefore, we have a consistent treatment of the data which is very important.

Some results of theoretical and observed values for $\delta$, $\gamma$ and $\Lambda$ are indicated in Table 1.

The observations correspond to 21 stations with very irregular distribution on the Earth [Melchior, 1971]. We can see that the consideration of arithmetic mean or weighted mean even alters some values. This is a good example about the difficulties encountered in the combination of observations which become important for the precision we are trying to achieve.

The consideration of the fourth decimal place has probably no physical meaning at the present time.

The theoretical models indicated correspond to different approximations that were made. The values computed by Jeffreys-Vicente [1957] and Molodensky [1961] were based on models of the internal struc-

ture of the Earth that appeared about 25 years ago, taking account of the liquid core.

Jeffreys-Vicente employed two models (central particle and Roche) which were considered as representing extreme cases of the possible behaviour of the core. This way of looking at the problem corresponds to setting up an upper and a lower bound to the behaviour of the core, considering that the knowledge about the structure of the core was not very detailed at the time. More recent models, based on a better knowledge of the structure of the core, have confirmed the general trends found by the central particle and Roche models.

Molodensky model 2 considers an inner core. It should be pointed out that the existence of the inner core introduces difficulties, and it has been shown that the partial differential equations to be satisfied are hyperbolic and the boundary conditions have to be considered carefully [Jeffreys and Vicente, 1966, p. 24].

Shen-Mansinha [1976] employed a model $\beta_o$ corresponding to an elliptical rotating Earth with a liquid outer core. The consideration of the other models, designated by $\beta=-0.2$ and $\beta=+0.2$, does not alter the conclusions.

The theoretical values obtained for the vertical component $\delta$, written on Table 1, show that model $\beta_o$ presents slightly smaller values for all waves in comparison with the other models.

Considering the wave $O_1$, we see that the vertical component $\delta$ is more affected by different types of models than the horizontal component $\gamma$.

For the horizontal component $\gamma$, the model $\beta_o$ shows slightly greater values than Molodensky model 2, being more in agreement with the observations.

A few theoretical values of $\Lambda$ are computed and there is no comparison with the observations because it is very difficult, at the present time, to obtain reliable observational values for these waves.

The models show general agreement with the observations because we cannot rely too much on the geophysical meaning of the third and fourth decimal places. The differences among the theoretical models are within five per cent.

Another important aspect is the frequency dependence of the bodily tide numbers which was revealed by the earlier investigations [Jeffreys and Vicente, 1957]. This result means that the behaviour of the Earth is very much conditioned, for the diurnal tides, by the period of the waves considered, as it is shown by the values written on Table 1.

The existence of damping and some recent determinations of Q values, depending on the period of the motion considered, show some of the difficulties encountered in the researches about Earth tides.

The global values of the bodily tide numbers have been employed to allow for the influence of Earth tides on the coordinates of a point on the Earth, but the improvement in the precision of the determination of coordinates would benefit from local observations of Earth tides by the appropriate techniques.

The waves with greater amplitudes appear in the following sequence: $K_1$ (due to the Moon and Sun), $O_1$ (due to the Moon) and $P_1$ (due to the Sun and Moon); and the frequencies are well defined by the theory of the diurnal tides. The great advantage of the bodily tides is the fact of their well-known periodicities.

The analyses of the observations, made by some of the modern techniques, should reveal these periodicities if they correspond to an adequate interval of time. It is nevertheless necessary that the observational programme is well planned and does not suffer from discontinuities. The modern techniques, employing artificial satellites, the Moon and radio interferometry, offer the advantage of providing global values.

We are attempting to obtain precisions of the order of a few centimeters and that raises the question if we are nowadays approaching the noise level of the system. In this case, we mean by noise level of the system, the possible range of values for the coordinates of a point on the Earth which are affected by numerous phenomena, some of them not very well known, and corresponding to the more general case of an Earth model considered as a collection of particles subject to so many forces.

## References


Jeffreys, H., Dynamic effects of a liquid core, Monthly Not. R. Astr. Soc., 109, 670-687, 1949.

Jeffreys, H., The Earth, Camb. Univ. Press, 1976.

Jeffreys, H. and R. O. Vicente, The theory of nutation and the variation of latitude, Monthly Not. R. Astr. Soc., 117, 142-173, 1957.

Jeffreys, H. and R. O. Vicente, Comparison of forms of the elastic equations for the earth, Mém. Cl. Sci. Acad. Roy. Belgique, 2e. sér., 37, fasc. 3, 1966.

Melchior, P., Precession-nutations and tidal potential, Celestial Mech., 4, 190-212, 1971.

Molodensky, M. S., Communications Obs. Roy. Belgique, No. 188, 1961.

Shen, P. Y. and L. Mansinha, Oscillation, nutation and wobble of an elliptical rotating earth with liquid outer core, Geophys. J. R. Astr. Soc., 46, 467-496, 1976.

Woolard, E. W. and G. M. Clemence, Spherical Astronomy, Academic Press, 1966.

Page Intentionally Left Blank

# How Do Earth Tides Affect Astronomers?

Tetsuo Sasao[1]
Cooperative Institute for Research in Environmental Sciences
University of Colorado/NOAA
Boulder, Colorado 80309

Earth tides affect astronomical observations of the Earth's rotation in two ways. They deflect verticals and change the polar moment of inertia of the Earth, thus causing periodic variations in its rotation rate (Table 1). Since astronomers observe stars usually around midnight at almost the same local sidereal time $\alpha = L_\odot + 180°$ in every station, where $L_\odot$ is the mean longitude of the Sun, the diurnal and semidiurnal tides, as well as nutations, produce apparently common variations with long "aliased" periods in the daily mean latitude and time values of different stations. Therefore, detection of polar motions is little contaminated by the Earth-tide effects, because they are largely absorbed in "common" z, $\tau$ and n-terms in the conventional observation equation:

$$\text{latitude:}\quad \Delta\phi = x\cos\lambda + y\sin\lambda + z$$

$$\text{time:}\quad \omega(\text{UTO-UTC}) = \tan\phi(x\sin\lambda - y\cos\lambda + \tau) + n$$

where x, y are "pole coordinates", $\phi$, $\lambda$ are latitude and east longitude of a station and $\omega$ is the mean rotation rate of the Earth. Nutation observations, on the contrary, are disturbed rather seriously by the diurnal tides with the same arguments and, through the aliasing, by semidiurnal and long period tides.

A major concern of astronomers in analysing the nutations is to detect evidence for the fluid-core resonance. Observed values contained in the squares of Figure 1 seemingly follow a theoretical resonance curve based on a realistic Earth model. Effects of the diurnal tides in time and latitude expected in $\phi = 39°08'$ (latitude of 5 ILS stations) are shown by the arrows in Figure 1 under the assumption $\Lambda = 1 + k - \ell = 1.2$. It is evident that $O_1$ and $P_1$ tides may deviate the observations of the fortnightly and semiannual nutations noticeably. It does not seem appropriate, however, to correct astronomical data for the diurnal-tide effects by simply assuming $\Lambda = 1.2$, because ocean-tide effects must hardly be negligible. The problem may become important in more detailed studies of the internal constitution of the Earth by means of nutation observations.

Consequences of the oscillations of the verticals due to $M_2$-tide can be observed in the astronomical latitude and time data and are regarded as one of the means for determination of the $\Lambda$-factor. In order to decide, for example, that $\Lambda = 1.2$ but not 1.1, however, one needs accuracy higher than 0.001 arcseconds or 3 cm. It does not seem easy, in general, to attain such accuracy by conventional astronomical observations with the typical error of single observation reaching 0.2 arcseconds or 6 m. As a matter of fact, 52 determinations of the $\Lambda$-factor so far obtained

TABLE 1. Effects of Earth Tide on Astronomical Observations

| Irregularities in the Earth's Rotation | | | |
|---|---|---|---|
| | Polar motion | Nutation | Changes in L.O.D. |
| Latitude | $x(t)\cos\lambda + y(t)\sin\lambda$ | $\sin[L(t) - \alpha]$ | ——— |
| Time | $\tan\phi[x(t)\sin\lambda - y(t)\cos\lambda]$ | $\tan\phi\ \cos[L(t) - \alpha]$ | $n(t)$ |

| Earth Tide Effects | | | | |
|---|---|---|---|---|
| | Long period tide | | Diurnal tide Vert. defl. | Semidiurnal tide vert. defl. |
| | Vertical deflection | Changes in C | | |
| Latitude | $\Lambda\sin2\phi\cos L(t)$ | ——— | $\Lambda\cos2\phi\sin[L(t) - \alpha]$ | $\Lambda\sin2\phi\cos[L(t) - 2\alpha]$ |
| Time | ——— | $k\ \sin L(t)$ | $\Lambda\tan\phi\cos[L(t) - \alpha]$ | $\Lambda\sin[L(t) - 2\alpha]$ |

$\Lambda = 1 + k - \ell$

$L(t)$: Argument of periodic disturbance due to the Sun and Moon

$\alpha$ : Local sidereal time

Proc. of the 9th GEOP Conference, *An International Symposium on the Applications of Geodesy to Geodynamics*, October 2-5, 1978, Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

[1]Permanent address: International Latitude Observatory of Mizusawa, Mizusawa, Iwate, Japan)

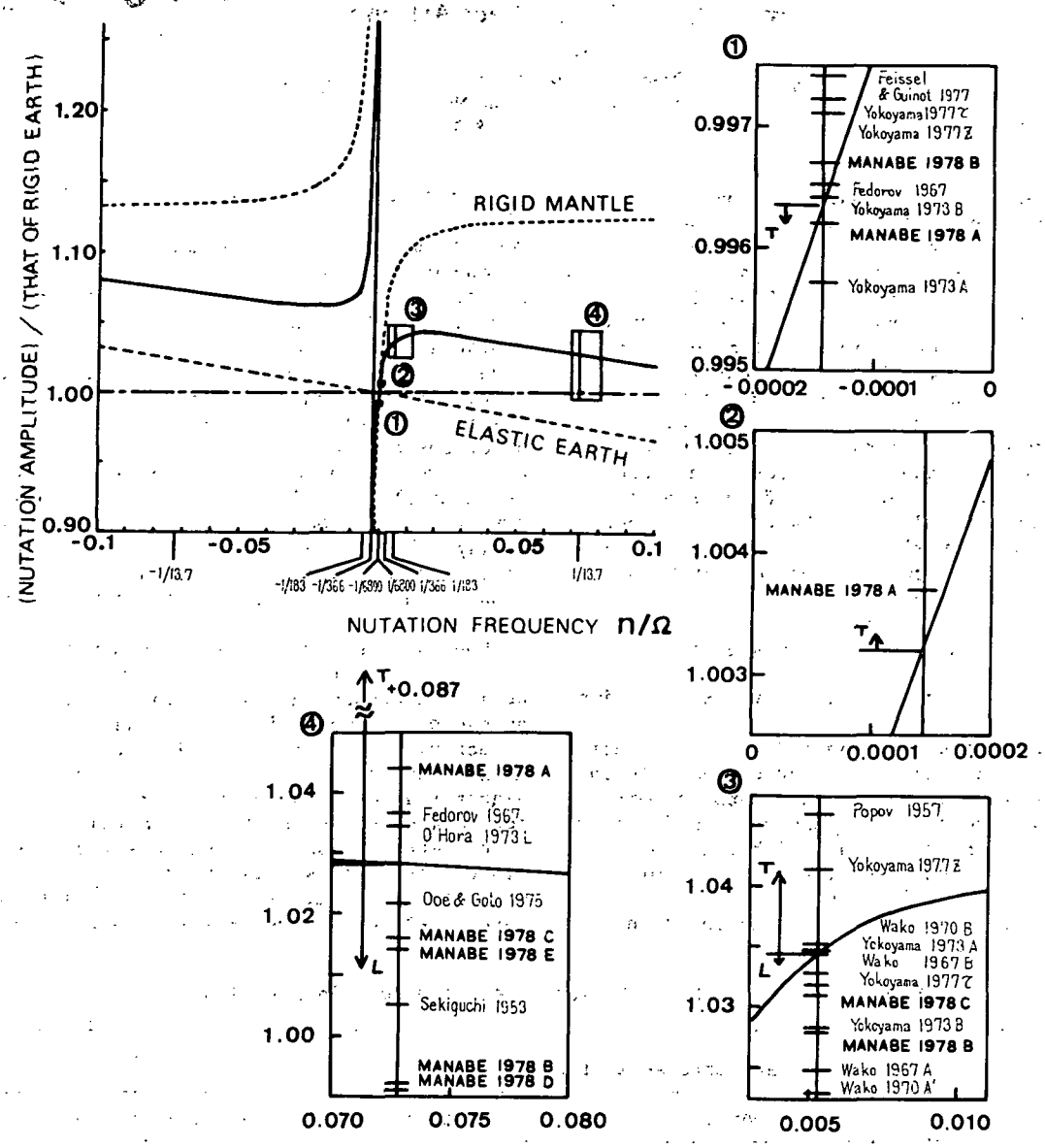



Fig. 1. Effects of diurnal Eaeth tides on nutation observations. Arrows T and L show the effects of diurnal tides on main nutation components in time and latitude observations, respectively, calculated for $\phi = 39°08'$ with an assumption $1 + k - \ell = 1.2$.

are scattered from 0.1 to 2.3, as shown in Figure 2. Although some authors have pointed out the possible secular change in the Λ-factor, recent analysis of past ILS data done by Manabe, Sakai and Sasao has revealed no firm evidence for the change (Figure 3). Different values of the Λ-factor detected in different stations are sometimes attributed to the difference between oceanic and continental stations. However, deviation of the 18.6-year principal nutation from Woolard's value derived by Manabe et al. from latitude data of 5 ILS stations shows even larger station-to-station differences than those of the $M_2$-tide (Figure 4), though the principal nutation must almost be unaffected by ocean-tide effects. It thus seems necessary to examine carefully the scale value problem and other possible sources of instrumental errors before considering any inverse problems for the ocean-tide effects on the basis of the astronomical $M_2$-tide data.

Changes in the rotation rate of the Earth due to zonal deformation caused by $M_m$ and $M_f$ tides are analyzed so as to determine Love number k. It should be noted that one needs 0.001 arcseconds accuracy again in order to decide whether $k = 0.30$ or 0.27. Nevertheless, the reported difference between k-numbers derived from $M_m$ and $M_f$ waves appears significant (Table 2). It might be interesting to note here a possibility that the $M_f$-wave is disturbed by the aliased "uncorrected" part of the fortnightly nutation arising from the fluid-core effect, which does not seem to have been taken into consideration fully.

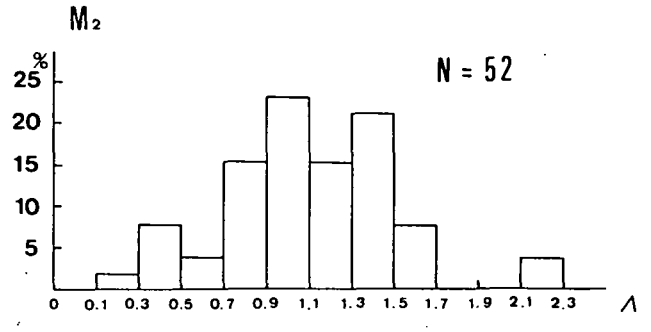



Fig. 2. Histogram of astronomically inferred values of $\Lambda = 1 + k - \ell$ factor.

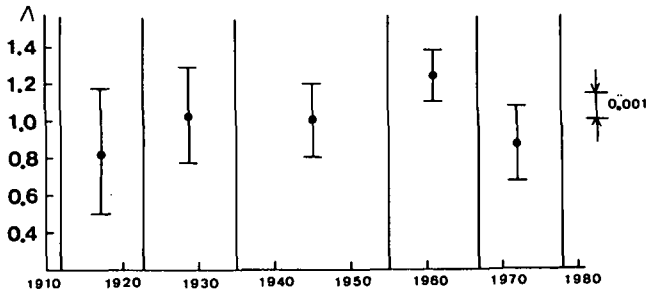



Fig. 3. Values of $\Lambda = 1 + k - \ell$ factor estimated for different periods of ILS observations.

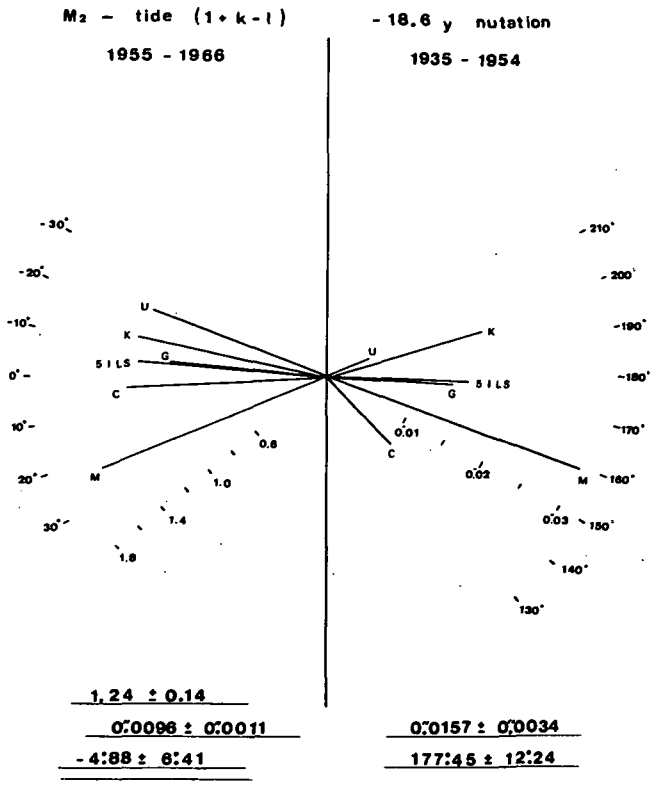



Fig. 4. Amplitude and phase diagram showing station-to-station differences of the estimated values of the $1 + k - \ell$ factor and retrograde component of "uncorrected" 18.6-year principal nutation. M, K, C, G, U correspond to 5 ILS stations: Mizusawa (Japan), Kitab (USSR), Carloforte (Italy), Gaithersburg (USA) and Ukiah (USA).

TABLE 2. Estimation of Love Number k from Changes in the Earth's Rotation Rate Due to $M_f$ and $M_m$ Tide

| Theoretical Values (in msec) | |
|---|---|
| $M_m$ | $-0.77(k/0.29)\sin(L_{(} - r')$ |
| $M_f$ | $-0.71(k/0.29)\sin 2L_{(}$ |
| ... | $-0.30(k/0.29)\sin(2L_{(} - \Omega)$ |

| Observed k | | |
|---|---|---|
| Authors | $M_f$ | $M_m$ |
| Guinot (1970) | $0.331 \pm 0.061$ | $0.265 \pm 0.068$ |
| Pilinik (1970) | $0.300 \pm 0.005$ | |
| | $k(M_f) - k(M_m) = 0.030 \pm 0.006$ | |
| Guinot (1974) | $0.334 \pm 0.005$ | $0.295 \pm 0.011$ |
| Djurovic (1975) | 0.343 | 0.301 |

# Geodynamic Problems



W. M. Kaula
Department of Earth and Space Sciences
University of California, Los Angeles, California 90024

Introduction. This article is a transcription of a talk given at the end of the conference, with the intentions of summarizing some leading themes thereof and supplementing the report of the conference with my impressions of other geodynamic problems to whose solution geodesy could contribute. The emphasis of the article is on our understanding of the solid earth; suggestions of what measurements should be undertaken are based on estimates of instrumental feasibilities in Kaula et al. [1978], a committee report which was rather widely reviewed.

The organization of the conference report is by observational technique. A problem-oriented approach should take a different cut: the most obvious is a combination of spatial and temporal spectra, or length and time scales. Figure 1 is an attempt at such a cut; the discussion is succeeding sections is based thereon: (1) earth evolution & mantle convection; (2) lithosphere-asthenosphere-surface load interaction; (3) glacier-ocean-solid earth interaction; (4) solid earth interactions with the sun, moon, core, oceans, and atmosphere; (5) zones of strain accumulation; and (6) earthquakes.

## 1. Earth Evolution and Mantle Convection

The greatest events in earth history--in the sense of the amount of mass and energy involved--were formation (probably by planetesimal infall), possibly a major impact great enough to knock off the protolunar material, and separation of the core. From isotopic constraints and dynamical plausibility, all these events happened more than $4.4 \times 10^9$ years ago. Since then, the most important process has been solid state convection in the mantle. The major long term questions are: the energy sources for this convection and the geodynamo: the division between primordial and radiogenic, etc.; the manner of crustal separation and ocean & atmosphere outgassing; the degree of inhomogeneity, thermal as well as compositional, of the lower and upper mantle; the extent to which there has been material as well as heat transfer among parts of the mantle; and the nature of fluctuations about a general trend of declining activity. Although past history is significantly constrained by petrological and isotopic data as well as by thermomechanical reasoning, our ideas about mantle convection must be largely based on its present state and recent history: in particular, the last $\sim$200 My, for which there exist seafloor spreading and other evidences to infer the major surface motions. This restriction raises the question of whether current conditions are atypical. For example, present heat flow could be misleading if the mantle convective system can undergo fluctuations in heat transfer of a factor of two or more, as do some numerical experiments in convection. Certainly the present is atypical in the extent to which the land is broken up into several continents.

Major constraints on mantle convection are the properties of the lithosphere, which is part of the flow system. The plate velocities, oceanic topography, and plate margin locations--particularly the subducted slabs--are all significant boundary conditions. Because of the strong temperature dependence of the rheology, the lithosphere is a remarkably thick boundary layer, and screens much of the properties of mantle convection from observation. Geodetic data contribute to inferring these convective properties in two ways: (1) the variations in the gravity field indicate the amount of density inhomogeneity; and (2) the rates of uplift and sinking in response to the glacier-ocean mass transfer, discussed below, indicate the effective viscosity of the mantle. Both of these properties relate to stress, the distribution of which determines the flow.

The magnitudes of density anomalies $\Delta\rho$ inferable from gravity anomalies $\Delta g$ depends on their characteristic vertical length scale L, beyond which they are in effect compensated. This length scale $L_D$ is probably much less than the mantle thickness M. As is well known, the visible density anomalies constituted by the topography must be largely compensated in the uppermost 100 km. If this is true within the lithosphere, then it must be all the more so in the more plastic interior. Hence we can write, roughly,

$$\Delta g \propto 2\pi G \Delta\rho L_D \qquad (1)$$

Because the rest of the earth attracts the density anomaly stresses are set up, which will also be proportionate to the length scale $L_D$:

$$\sigma \propto g\Delta\rho L_D \propto g\Delta g/G2\pi \qquad (2)$$

Viscosity is related to stress through the strain rate, which can be inferred from the plate velocities v:

$$\sigma \propto \nu\dot{\varepsilon} \propto \nu v/L_V \qquad (3)$$

If we assume the two length scales $L_V$ and $L_D$ to be the same and take $\Delta g$ to be the typical value for $1^o \times 1^o$ square, say 25 mgal, use $10^{22}$ for $\nu$ and 5 cm/yr for v, then L is $\sim$300 km and $\sigma$ is $\sim$60 bars.

The foregoing estimate is, at best, an order-of-magnitude, and is likely to be deceptive when applied to a regime which must be organized at some level, as is convection. But the order-of-magnitude of 100-500 km for the length scale is what arises from other estimates, such as those based on Rayleigh number considerations. Hence it is desirable that the gravity be known to a resolution not worse than 100 km to infer mantle

Proc. of the 9th GEOP Conference. *An International Symposium on the Applications of Geodesy to Geodynamics, October 2-5, 1978.* Dept. of Geodetic Science Rept. No. 280, The Ohio State Univ., Columbus, Ohio 43210.

convective patterns therefrom.

To infer whether there are density inhomogeneities at depth require consideration of, firstly, the spectral distribution of gravity and secondly the correlation of gravity with surface topography. The leading feature of the gravity spectrum is the "$10^{-5}/\ell^2$" law: a fairly sharp drop off in the rms magnitude of the normalized potential coefficients $\bar{C}_{\ell m}$, $\bar{S}_{\ell m}$ with spherical harmonic degree $\ell$. The principal properties of the cross-correlation between gravity and topography are (1) negligible correlation for the long wavelengths $\ell \lesssim 5$ and (2) moderately positive correlation for the long wavelengths $\ell \gtrsim 6$ [Kaula, 1977]. Studies which assume that the density spectrum $\Delta\rho_\ell$ is "white": i.e., a comparable amount of variability in each degree conclude that the steep drop off in the gravity spectrum requires inhomogeneities at depth [Kaula, 1977; Lambeck, 1976]. The equally plausible assumption that the length scale L in eqs. 1-3 also has a spectrum $L_\ell$ such that $L_\ell$ varies inversely with $\ell$ (i.e., directly with horizontal wavelength) suggest that $\Delta\rho_\ell$ may vary directly with $\ell$, because of the stresses entailed; requiring density anomalies at depth all the more. Analyses based on numerical models of convection get a flatter spectrum than $10^{-5}/\ell^2$, with even a reversal in slope at the longest wavelengths when the bottom boundary conditions are free [McKenzie, 1977].

For the present, it seems prudent to assume that there are density anomalies throughout the mantle comparable in magnitude to those near the surface. Hence any study of gravity should employ harmonic analysis, to separate shallow and deep effects.

## 2. Lithosphere-Asthenosphere-Surface Load Interaction

Moving to shorter wavelengths and more recent phenomena makes it fruitful to explain associated topographic and gravitational features by specific physical models, as discussed by Watts, Roufosse, and Turcotte in these proceedings. The common theme of all these studies is to treat the lithosphere as an elastic layer over a fluid half space. The leading property then becomes the flexural rigidity [Walcott, 1970]:

$$D = \frac{\mu(\mu+\lambda)T^3}{3(2\mu+\lambda)} \approx \frac{2\mu T^3}{9} \qquad (4)$$

in which $\lambda,\mu$ are the elastic moduli and T is the effective thickness of the lithosphere. In general, the flexural rigidity is less than what is calculated from seismic values for the elastic moduli $\lambda,\mu$ and the Rayleigh-wave lid thickness for T, as should be expected. Typical results are $\sim 10^{30}$ dyne-cm (compared to $\sim 10^{32}$ dyne-cm from seismic data), and an inverse correlation with both lithospheric age (since creation at the ocean rise) and duration of loading.

More dynamic problems utilizing gravity & topography are those associated with subduction zones, where the plate velocity and thus the asthenospheric viscosity may be of significance. Thus, for example, the topographic and gravitational high oceanward from the trench have been explained as due to the elastic bending of the lithosphere [Watts & Talwani, 1974]. Probably the most elaborate models leading to prediction not only of gravity anomalies and topography but also to rate-of-topographic rise are the thermomechanical finite element calculations of Himalaya and Zagros underthrusting by Bird [Bird et al., 1975; Bird, 1978].

A major plate tectonic problem to which it is hoped future geodetic data will apply is the variation of plate velocities about their values inferred from remanent magnetic striations. The plate velocities appear to be quite steady for periods of about $10^7$ years, after which there is an adjustment of the velocity pattern, apparently due to change in physical circumstances at plate margins [Larson & Pitman, 1972]. This dependence on margin conditions is consistent with plate tectonic models which attempt to account for the velocity patterns as arising from plausible combinations of forces on the margins and on the lithosphere-asthenosphere interface, dependent on the extent to which the plate is oceanic or continental. These studies conclude that the plates move mainly as a consequence of push at the rise and pull at the trenches; the asthenosphere acts as a minor drag on the lithospheric plates [Forsyth & Uyeda, 1975; Richardson et al., 1976]. Since single events, like the 1960 Chile earthquake, may account for several decades' motion over a major segment of plate margin, it is plausible that some of this jerkiness is transmitted to the overall plate motion. This post-seismic stress propagation may be geodetically detectable over 1000's of kilometers.

Geodetic techniques may also detect variations about steady rates in vertical motion arising from tectonic causes both close to and remote from plate margins. Estimates of vertical motion rates from either geologic observations or thermomechanical modeling suggest that vertical motion rates greater than $\sim$1 mm/yr should be exceptional. While the inherent sluggishness of the solid earth makes faster rates implausible and, at present, inexplicable, our understanding is imperfect enough that such rates cannot be ruled out a priori. It is therefore important to determine relative uplift rates more reliably between sites for which the geologic setting is well known.

## 3. Glacier-Ocean-Solid Earth Interaction

The dominant effect on the rate of uplift or sinking, and on the height relative to sea level, of the solid earth in the time scale of $\sim$300 to 30,000 years is the $3\times10^{22}$ gram mass transfer from continental glaciers to the ocean surface which mostly took place 18,000 to 8,000 years ago. Because the distribution of the glacial unloading is relatively well known [Paterson, 1972; Peltier & Andrews, 1976] and of the ocean loading very precisely known, measurements of the response of the solid earth are the most significant data on its rheology.

The principal data type pertaining to this regime is carbon-14 dating of ancient shorelines, which gives points in the historical record over the last few 1000 years. The principal interpretations are that the characteristic decay

time $\tau(\lambda)$ varies directly with wavelength $\lambda$ for $\lambda \gtrsim 1000$ km and inversely for $\lambda \lesssim 1000$ km. The former appears consistent with a mantle of $10^{22}$ poise; the latter, with an elastic lithosphere $\sim$100 km thick. Current vertical rates are 10 mm/yr near the center of the formerly glaciated area, but $\sim$1 mm/yr away from it.

Since most of the adjustment to the glacier-ocean load transfer appears to be completed, geodetic leveling measurements are to some extent supplementary. However, they are a valuable supplement, because the $C^{14}$ shoreline data give rather incomplete coverage of the peripheral bulge area which is critical to inferring elastic and non-linear rheological effects.

The post-glacial response is one of the two or three principal evidences that the mantle indeed flows. However, there are some cautions as to the application of the inferred viscosity to the longer term convective problem [Kaula, 1979]. Foremost among these is the possibility that the post-glacial effect is transient. While it may seem strange to call an effect on a $\sim 10^3$ year timescale 'transient', it is true that steady state rheology has been attained in the laboratory only at strains appreciably greater than the $10^{-4}$ to $10^{-3}$ characteristic of post-glacial load transfer [Goetze & Brace, 1972; Weertman, 1978]. The present level-of-detail of the data warrants at most linear viscoelastic models [Clark et al., 1978]. To go beyond these models and infer non-linear viscosity or transient effects may require an amount of detail which only geodetic data can supply.

## 4. Solid Earth Interations with the Sun, Moon, Core, Oceans, and Atmosphere

Although tidal friction can be said to apply to the entire history of the earth, most of these exogenic effects lie well within a $\lesssim$300 year timescale in their characteristic periods. The common thread of the exogenic disturbances is that they are all quite small in their effects at the earth's surface, and hence are observable geodetically only if they affect things which can be measured very precisely, such as pole position or tide height.

Because rotation and tides are precisely measurable and mathematically modelable, they have an intellectual appeal which may result in their receiving more attention that may be warranted by the criterion of illuminating causes. As discussed by Lambeck in this volume, matters which are still very much in doubt are: the relative magnitudes of atmospheric, earthquake, and aseismic creep effects on the free polar wobble; the amount of non-tidal rotational acceleration; the manner of dissipation of both polar wobble and tides; and the nature of core-mantle coupling. While for all these problems an increase in accuracy of measurement of their rotational & tidal effects should be of some help, it also can be said for all of them that enhanced insight will require either or both of appreciable improvement in non-geodetic modeling or significant additional non-geodetic observations.

Thus to infer whether the atmosphere makes a major contribution to the excitation of the long term variations in LOD and free polar wobble already observable requires a better understanding of both seasonal and decade-scale changes: such questions as the year-to-year variations in atmospheric mass distribution, water and energy transfers between the ocean & atmosphere, and the influence of the solid earth on the atmosphere, most likely through volcanism [Lambeck & Cazenave, 1976, 1977; Wilson & Haubrich, 1976]. All of these phenomena have broad spectra, and hence require observations capable of measuring changes on a much finer scale than global. Furthermore, some of the most feasible observations--such as satellite photography of cloud motions--require significant supplemental observations and modeling to be applied. The recent revisions of the magnitude-moment relationships for great earthquakes [Kanamori, 1977] revive seismicity as a possible source of wobble excitation, but there has not yet been a convincing modeling of the effect of any one earthquake on the pole path.

The non-dissipatory part of the response of the earth to tidal and rotational effects depends on bulk properties, and hence is more amenable to mathematical attack. Significant advances have been made in modeling the effects of the ocean [Dahlen, 1976] and the core [Smith, 1977]. On the other hand, the dissipatory part of the response is probably associated more with the liquid-solid interfaces: ocean-crust and core-mantle. Consequently, the dissipation is more difficult to model. The mechanisms of both the polar wobble damping and tidal friction are still not understood. However, some constraints can be placed on both these processes, even if details of how the dissipation occurs are unknown. Thus the tidal friction at present must be anomalously high, because the present 1/Q extrapolated into the past brings the moon close to the earth much too recently. But this high tidal dissipation is not surprising: by paleontological and other indications, the present continental configuration is remarkably broken up, so that a greater-than-usual portion of tidal energy is transferred from the second to the higher harmonics of the tide, eventually to be dissipated, willy-nilly. An interesting theme to pursue might be the extent to which the need to maintain a small tidal 1/Q constrains the tectonic style in the past. Some geological discussions suggest that there is a real dearth of subduction zone associations from Proterozoic, 0.6 to $2.5 \times 10^9$ years ago.

In summary, measurements of pole position and LOD appear to be in the Hubble situation: "We do not know <u>what</u> the future will want, but we do know they will want it accurate". To which we can add "continuous": as new techniques are phased in, there should be some years' overlap with the older measurements.

## 5. Zones of Strain Accumulation

The somewhat arbitrary distinction from topic #2, lithospheric-asthenospheric interaction, is that we are concerned here about endogenic motions in the solid earth at rates sufficient both to be measured and to test tectonophysical models. These criteria narrow attention to plate margins

(using a broad definition of the term) and very recent times. The equally arbitrary distinction from topic #6, earthquakes, is that we are concerned here about the entire process of strain change in a region, rather than just the phenomena associated with particular abrupt releases of strain. The possibility of gradual, non-earthquake release of strain makes the title 'strain accumulation' somewhat incomplete in its implications.

As discussed by Savage in this volume, the rate-of-strain as measured near active plate margins in California, Japan, and New Zealand appears to be on the order of $0.3 \times 10^{-6}$/yr. This strain accumulation is distributed across a zone $\sim$200 km wide. Such a magnitude is roughly consistent with the prediction from plate tectonics,

$$\dot{\varepsilon} \approx \frac{v}{W} \qquad (5)$$

where $\dot{\varepsilon}$ is strain rate, v is the relative velocity of the plates and W is the width of the strain accumulation zone. However, this generalization is based on relatively few surveys of differing temporal intervals in regions of differing tectonic context and seismicity. Factors aside from relative plate velocities which may affect the pattern of rate-of-strain include: (1) the regional fault configuration; (2) the depth to an 'asthenosphere' or other place which can undergo plastic deformation; and (3) the recent seismic history of the region. Thus, for example, the San Andreas fault appears to have at least three distinctly different regional behaviors. In northern California, the strain evolution appears to be still strongly influenced by the 1906 earthquake: since then, the seismicity has been much lower than in the previous $\sim$50 years, and the strain accumulation appears to be much more about the Hayward & Calaveras faults [Thatcher, 1975ab]. In the Hollister region, there are sporadic motions at $\sim$ monthly intervals associated with very small earthquakes, suggestive of slip on a shallow plastic zone [Huggett et al., 1977, Johnston et al., 1977]. In southern California, where the San Andreas bends appreciably and there are subsidiary faults, in six years of precise measurements there has been a steady north-south contraction of $0.3 \times 10^{-6}$/year [Savage et al., 1978].

The San Andreas, being a transcurrent fault, is relatively simple among zones of strain accumulation. However, it shows significant variations on a length scale of $\sim$300 km and on a temporal scale of $\sim$100 years for major events, as indicated by the sedimentary record in the south [Sieh, 1978] as well as by the historical record in the north [Thatcher, 1975ab]. A consequence of these irregularities is that the strain accumulation pattern in a region may be quite different from that suggested by simple plate tectonic considerations, as is the case now in southern California. Hence, while more geodetic measurements are manifestly desirable, it can be expected that a purely empirical approach will not solve the problem. There must be a development of models as well as elaboration of the data. Other improvements which should be forthcoming are the combination of horizontal and vertical data in analysis and better integration of geodetic measurements in the geological context.

## 6. Earthquakes

As emphasized in the preceding section, earthquakes are not isolated phenomena. However, they do have a distinctive character, and there is nothing like a $\geq$7.5 magnitude earthquake to test a model of a strain accumulation zone. From a purely scientific point-of-view, it is a pity they do not happen more often. We are concerned here, firstly, about when, where, & why earthquakes occur in a strain accumulation region and, secondly, for a seismic event the pattern of precursory, coseismic, and subsequent activity and its causes.

It has been evident for sometime that earthquakes are most likely to occur at gaps in strain accumulation zones: places where earthquakes have not recently occurred and hence where presumably the greatest stresses have accumulated [e.g., Kelleher et al., 1973]. However, the stress redistributed by earthquakes in adjacent regions can vary significantly; some stress can apparently be released aseismically, or, in complex fault zones, by motion on adjacent faults; and tectonic configurations vary appreciably in their ability to withstand stress. Hence in an earthquake-prone region the recurrence interval can vary over almost an order-of-magnitude. For example, at Pallett Creek, in the south-central San Andreas fault, the interval has varied from 50 to 300 years in the last 1400 years [Sieh, 1978].

So far, geodetic measurements have contributed only slightly to the depiction of motions precursory to a sizeable earthquake. The 1964 Niigata earthquake (7.5) was preceded by rises on the order of 10 cm during the previous 60 years, with an acceleration in the last 5 years [Kisslinger, 1974]. The 1971 San Fernando shock (6.4) was located about 30 km from the center of an uplift of 20 cm over 10 years, with some migration of this center toward the shock a couple of years before the event [Castle et al., 1975]. The 1973 Point Mugu quake (6.0) was preceded by uplifting and then downwarping of about 4 cm during the previous 13 years [Castle et al., 1977]. However, these changes are not remarkably bigger than those inferred from leveling data in relatively non-seismic regions [Brown, 1976], and more recent experience with the Palmdale bulge has indicated that leveling must be treated with care and used in combination of other data [Kumar & Strange, 1978]. This data should include both gravity measurements and horizontal geodetic survey. However, the former can be appreciably affected by non-tectonic processes such as ground water movement [Lambert & Beaumont, 1977], while the latter have so far been only sketchily related to a shock [Thatcher, 1974].

A much better geodetic description of strain accumulation prior to a major earthquake can be expected in the next decade or two. The San Andreas has been anomalously quiet since the intersification of instrumentation programs. More

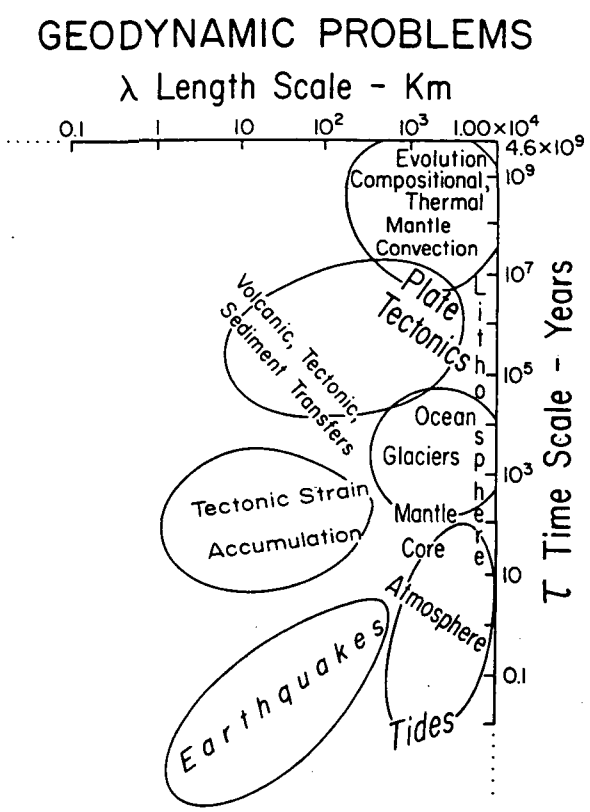



Fig. 1. Length scale versus time scale breakdown of geodynamic problems.

careful attention to leveling observation, computation, and adjustment should also payoff. Point measurements--strain, tilt, and gravity--appear to have greater inherent difficulties of sorting out regional from local effects [Johnston et al., 1977; McHugh & Johnston, 1977; Lambert & Beaumont, 1977]. For a while at least, seismicity itself will remain the leading precursory data type: see, e.g., Shimazaki [1978] and Ishida & Kanamori [1978] for studies on two different scales.

Postseismic geodetic measurements [e.g., Brown et al., 1977] are important because they indicate some aspects of the stress redistribution precursory to subsequent creep and seismicity in the zone, and because they are affected by the same tectonic context as the earthquake itself. In fact, so far earthquake models have been more successful in explaining post-seismic adjustment than the precursory buildup and the earthquake itself [e.g., Nur & Mavko, 1974; Anderson, 1975; Rundle & Jackson, 1977; Savage & Prescott, 1978; Walcott, 1978]. The situation appears to be analogous to that discussed in connection with rotation: phenomena dependent on bulk properties, such as viscoelastic post-seismic adjustment, are easier to model than those dependent on interface properties, such as earthquake occurrence. Geodetic measurements also constrain the bulk response more than the interface effects, for which more stress- and energy-sensitive indicators (primarily seismic data, but also heat flow, radon, etc.) will continue to be needed, as well as experimental rheology [Dieterich, 1978].

## Conclusions

It is many years since Love [1911] wrote his book of similar title, but progress seems to be accelerating in recent years in both measurement and modeling of geodynamic phenomena. Rather than attempt to recapitulate the main themes, I close with a personal wish list, in priority:

1. Intensify geodetic measurements in strain accumulation zones within ∿100 km of the best-studied faults. The data, as well as modeling considerations, indicate that this is where most of the action occurs, and that it is quite observable with current techniques at major plate boundaries [Savage & Prescott, 1978; Walcott, 1978]. This is not to say that a comprehensive insight does not require widespread measurements entailing space techniques; the globe is one big coupled system [Anderson, 1975]. Rather, given limited resources, the most likely payoff, scientifically as well as practically, is close to the big faults: scientifically, because meaningful models must deal with stress, which in turn entails strain & strain rate, taking detailed measurements to describe. The rocks know each other mainly through their nearest neighbors.

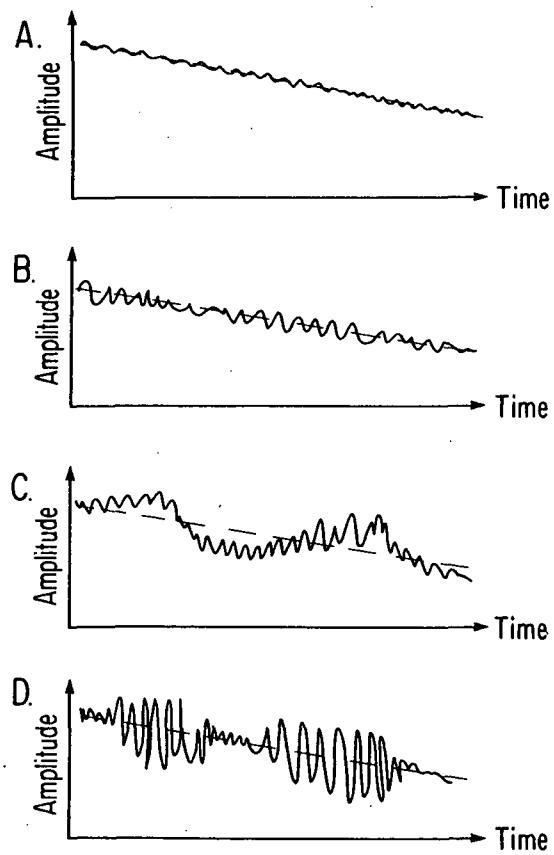



Fig. 2. Schematic representation of the possible forms of short term oscillations about long term geodynamic trends.

2. Develop better devices for geodetic measurements over distances less than ∿30 km. 'Better' means not only more accurate, but more rapid & responsive. Some that look promising are:

a. Three-wavelength ranging system good to within $5 \times 10^{-8}$, such as that described by Levine in this volume;

b. mini-interferometers, such as those described by Shapiro in this volume;

c. three-wavelength angle measuring devices, to obtain level differences more rapidly.

Since important tectonic phenomena are inherently inaccessible to direct observation and must be inferred from surface measurements, there is no distinct point of diminishing returns in the accuracy of measurements, while we are still far from saturation in spacing and frequency.

3. Tidy-up the vertical networks. There have been too many suspicious leveling results, such as the now-here, now-gone discrepancy from tidal bench marks between San Diego and San Francisco [Douglas, 1978]. Leveling has the inherent accuracy to tell us a lot which should be realized. While it is indeed possible that short term variations may have appreciably higher rates than the long term trends (see Figure 2), there are limits to what is physically plausible, particularly over distances more than some 10's of kilometers.

4. Measure gravity globally to 100 km resolution. As discussed in section 2, such detail is needed to see through the lithosphere to infer the underlying mantle convective action. The SEASAT altimeter will obtain it for most of the oceans, but satellite-to-satellite tracking systems, such as described by Fischell in this volume, appear to be the way over the land.

5. Develop portable 3 μgal (30nm/sec) gravity meters. Despite the difficulties with local effects, a comparably accurate gravity meter seems to be a necessary concommittant to leveling.

The earth is rather big and has been around a long time. While geodesy has gained a global range spatially, it is still limited temporally to the last finger-snap of geologic time. However, there are important, as well as entertaining, processes on the range of time scales within geodetic reach. The task of geodesy might be said to measure the short-term fuzz on the long-term trends: for each phenomenon, to determine which sort of pattern--A, B, C, or D in Figure 2--prevails.

Acknowledgment. This work was supported by NASA grant NSG-5263.

References

Anderson, D.L., Accelerated plate tectonics, Science, 187, 1077-1080, 1975.

Bird, G.P., Initiation of intracontinental subduction in the Himalaya, J. Geophys. Res., 83, 4975-4987, 1978.

Bird, G.P., M.N. Toksöz, and N.H. Sleep, Thermal and mechanical models of continent-continent convergence zones, J. Geophys. Res., 80, 4405-4416, 1975.

Brown, L.D., and J.E. Oliver, Vertical crustal movements from leveling data and their relation to geologic structure in the eastern United States, Revs. Geophys. Sp. Phys., 14, 13-35, 1976.

Brown, L.D., R.E.Reilinger, S.R. Holdahl, and E.I. Balazs, Postseismic crustal uplift near Anchorage, Alaska, J. Geophys. Res., 82, 3369-3378, 1977.

Castle, R.O., J.P. Church, M.R. Elliott, and N.L. Morrisson, Vertical crustal movements preceding and accompanying the San Fernando earthquake of February 7, 1971: a summary, Tectonophysics, 29, 127-140, 1975.

Castle, R.O., J.P. Church, M.R. Elliott, and J.C. Savage, Preseismic and coseismic elevation changes in the epicentral region of the Point Mugu earthquake of February 21, 1973, Bull. Seism. Soc. Amer., 67, 219-232, 1977.

Clark, J.A., W.E. Farrell, and W.R. Peltier, Global changes in post-glacial sea level: a numerical calculation, Quaternary Res., 9, 265-287, 1978.

Cochran, J.R. and M. Talwani, Free-air anomalies in the world's oceans and their relationship to residual elevation, Geophys. J. R.A.S., 50, 495-552, 1977.

Dahlen, F.A., The passive influence of the oceans upon the rotation of the earth, Geophys. J. R.A.S., 46, 363-406, 1976.

Dieterich, J.H., Preseismic fault slip and earthquake prediction, J. Geophys. Res., 83, 3940-3948, 1978.

Douglas, B., Geodetic and steric leveling: together at last?, EOS Trans. AGU, 59, 1052-1053, 1978.

Forsyth, D., and S. Uyeda, On the relative importance of the driving forces of plate motion, Geophys. J. R.A.S., 43, 163-200, 1975.

Goetze, C., and W.F. Brace, Laboratory observations of high-temperature rheology of rocks, Tectonophysics, 13, 583-600, 1972.

Huggett, C.R., L.E. Slater, and G.R. Huggett, Fault slip episodes near Hollister, California; initial results using a multiwavelength distance measuring instrument, J. Geophys. Res., 82, 3361-3368, 1977.

Ishida, M., and H. Kanamori, The foreshock activity of the 1971 San Fernando earthquake, California, Bull. Seism. Soc. Amer., 68, 1265-1280, 1978.

Johnston, M.J.S., A.C. Jones, and W. Daul, Continuous strain measurements during and preceding episodic creep on the San Andreas fault, J. Geophys. Res., 82, 5683-5691, 1977.

Kanamori, H., The energy release in great earthquakes, J. Geophys. Res., 82, 2981-2987, 1977.

Kaula, W.M., Geophysical inferences from statistical analyses of the gravity field, Proc. Symp. "Geod. Sci. in a Changing World", Rep., 250, Dept. Geod. Sci., Ohio State Univ., 119-141, 1977.

Kaula, W.M., Problems in understanding vertical movements and earth rheology, in Mörner, N., ed., Earth Rheology, Isostasy, and Eustasy, John Wiley & Sons, Chichester, in press, 1979.

Kaula, W.M., and 11 others, Geodesy: Trends & Prospects, Nat'l. Acad. Sci., Washington, 86 pp., 1978.
Kelleher, J., L. Sykes, and S. Oliver, Possible criteria for predicting earthquake locations and their application to major plate boundaries of the Pacific and Caribbean, J. Geophys. Res., 78, 2547-2585, 1973.
Kumar, M., and W.E. Strange, Regional elevation and gravity changes in the Palmdale bulge area, EOS Trans. AGU, 59, 1051, 1978.
Lambeck, K., Lateral density anomalies in the upper mantle, J. Geophys. Res., 81, 6333-6340, 1976.
Lambeck, K., and A. Cazenave, Long term variations in the length of day and climatic changes, Geophys. J. R.A.S., 46, 555-573, 1976.
Lambeck, K., and A. Cazenave, The earth's variable rate of rotation: a discussion of some meteorological and oceanic causes and consequences, Phil. Trans. Roy. Soc. Lond. A, 284, 495-506, 1977.
Lambert, A., and C. Beaumont, Nano-variations in gravity due to seasonal groundwater movements: implications for the gravitational detection of tectonic movements, J. Geophys. Res., 82, 297-306, 1977.
Larson, R.L., and W.C. Pitman, World-wide correlation of Mesozoic magnetic anomalies and its implications, Geol. Soc. Amer. Bull., 83, 3645-3662, 1972.
Love, A.E.H., Some Problems of Geodynamics, Cambridge Univ. Press, 180 pp., 1911.
McHugh, S., and M.J.S. Johnston, An analysis of coseismic tilt changes from an array in central California, J. Geophys. Res., 82, 5692-5698, 1977.
McKenzie, D., Surface deformation, gravity anomalies, and convection, Geophys. J. R.A.S., 48, 211-238, 1971.
Nur, A., and G. Mavko, Post-seismic viscoelastic rebound, Science, 183, 204-206, 1974.
Paterson, W.S.B., Laurentide ice sheet: estimated volumes during late Wisconsin, Revs. Geophys. Sp. Phys., 10, 885-917, 1972.
Peltier, W.R., and J.T. Andrews, Glacial-isostatic adjustment - I, the forward problem, Geophys. J. R.A.S., 46, 605-646, 1976.
Richardson, R.M., S.C. Solomon, and N.H. Sleep, Intraplate stress as an indicator of plate tectonic driving forces, J. Geophys. Res., 81, 847-856, 1976.
Rundle, J.B., and D.D. Jackson, A kinematic viscoelastic model of the San Francisco earthquake of 1906, Geophys. J. R.A.S., 50, 441-458, 1977.
Savage, J.C., and W.H. Prescott, Asthenospheric readjustment and the earthquake cycle, J. Geophys. Res., 83, 3369-3376, 1978.
Savage, J.C., W.H. Prescott, M. Lisowski, and N. King, Strain in southern California: measured uniaxial north-south regional contraction, Science, 202, 883-885, 1978.
Shimazaki, K., Correlation between intraplate seismicity and interplate earthquakes in Tohoku, Northeast Japan, Bull. Seism. Soc. Amer., 68, 181-192, 1978.
Sieh, K.E., Prehistoric large earthquakes produced by slip on the San Andreas fault at Pallett Creek, California, J. Geophys. Res., 83, 3907-3939, 1978.
Smith, M.L., Wobble and nutation of the earth, Geophys. J. R.A.S., 50, 103-140, 1977.
Thatcher, W., Episodic strain accumulation in southern California, Science, 194, 691-695, 1974.
Thatcher, W., Strain accumulation and release mechanism of the 1906 San Francisco earthquake, J. Geophys. Res., 80, 4862-4872, 1975a.
Thatcher, W., Strain accumulation on the northern San Andreas fault zone since 1906, J. Geophys. Res., 80, 4873-4880, 1975b.
Walcott, R.I., Flexural rigidity, thickness, and viscosity of the lithosphere, J. Geophys. Res., 75, 3941-3954, 1970.
Walcott, R.I., Geodetic strains and large earthquakes in the axial tectonic belt of North Island, New Zealand, J. Geophys. Res., 83, 4419-4429, 1978.
Watts, A.B., and M. Talwani, Gravity anomalies seaward of deep-sea trenches and their tectonic implications, Geophys. J. R.A.S., 36, 57-90, 1974.
Weertman, J., Creep laws for the mantle of the earth, Phil. Trans. Roy. Soc. Lond. A, 288, 9-26, 1978.
Wilson, C.R., and R.A. Haubrich, Meteorological excitation of the earth's wobble, Geophys. J. R.A.S., 46, 707-743, 1976.

Page Intentionally Left Blank

Applications of Geodesy to Geodynamics
9th GEOP Conference
Oct. 2-5, 1978

Participants

Leendert Aardoom
Working Group for Satellite Geodesy
Delft Univ. of Technology
P.O. Box 581
Apeldoorn, Netherlands

John N. Alt
Woodward-Clyde Consultants
Three Embarcadero Center
Suite 700
San Francisco, CA 94111

Richard J. Anderle
Naval Surface Weapons Center
Box 3436 College Station
Fredericksburg, VA 22401

Allen Joel Anderson
University of Uppsala
Dept. of Solid Earth Physics
Box 556
S-75122 Uppsala, Sweden

Edward G. Anderson
Univ. of New Brunswick
Dept. of Surveying Engineering
Fredericton, NB, Canada
E3B 5A3

L. F. Baank
Dept. of Development
Maystraat 39
Paramaribo, Surinam

T. F. Baker
Institute of Oceanographic Sciences
Bidston Observatory
Birkenhead, Merseyside, England

Chris Beaumont
Dept. of Oceanography
Dalhousie University
Halifax, Nova Scotia
Canada B3H 4J1

John Beavan
Lamont-Doherty Geol. Obs.
Palisades, N Y 10964

Larry D. Beers
Defense Mapping Agency
U.S. Naval Observatory, Bldg. 56
Washington, D.C. 20305

Peter L. Bender
NBS/JILA
National Bureau of Standards
Boulder, CO 80302

Paul B. Beruff
DMHH-TC-GSS
T.E. Warren AFB
Cheyenne, WY 82001

William Bishop
National Geodetic Survey
Code C133
11400 Rockville Pike
Rockville, MD 20852

J. A. R. Blais
Geodetic Survey of Canada
615 Booth St.
Ottawa, Canada K1A 0E9

A. G. Bobba
Canada Centre for Inland Waters
Burlington, Ontario
Canada

John D. Bossler
Commerce, NOAA, NOS, NGS
6001 Executive Blvd.
Rockville, MD 20852

Robert F. Brammer
TASC
6 Jacob Way
Reading, MA 01867

Tom Brocher
Princeton Univ.
Geology Dept.
Princeton, N J 08540

Duane Brown
Geodetic Services Inc.
P.O. Box 3668
Indialantic, FL 32903

Gerry H. Cabaniss
AF Geophysics Lab
AFGL/LWH
Hanscom AFB, MA 01731

Odile Calame
CERGA
Ave Copernic
06130 Grasse, France

Michele Caputo
Ist. di Fisica, Univ. d. Studi, Roma
Piazzale della Scienze, 5
Rome, Italy

William E. Carter
National Geodetic Survey
C-133
Rockville, MD 20852

Robert O. Castle
Geological Survey
345 Middlefield Rd.
Menlo Park, CA 94025

Anny Cazenave
Groupe De Recherche de
Geodesie Spatiale
GRGS/CNES
18 Ave. E. Belin
31055 Toulouse, Cedex, France

William H. Chapman
U.S. Geological Survey
National Center, Stop 521
Reston, VA 22092

Michael A. Chinnery
Lincoln Lab M.I.T.
42 Carleton St.
Cambridge, MA 02142

Tom Crough
Princeton University
Geology Dept.
Princeton, N J 08540

I. Csikos
Royal Netherlands Meteorol. Inst.
3732 GK. De Bilt
The Netherlands

B. Louis Decker
Defense Mapping Agency Aerospace
Center
Second & Arsenal Streets
St. Louis Air Force Station
MO 63118

G. de Jong
Delft University of Technology
Thysseweg 11
Delft, The Netherlands

J. De May
CIRES
Univ. of Colorado
Boulder, CO 80309

Steve Dickman
Dept. of Geod. Sciences
State Univ. of New York
Binghamton, N Y 13901

Inez J. Dimitrijevich
DMAAC, 2nd and Arsenal Streets
St. Louis AFS, MO 63118

Vojislav Dimitrijevich
DMAAC (retired)
4400 Lindell Blvd.
St. Louis, MO 63108

Hermann Drewes
Inst. für Theoretische Geodäsie
Universität Hannover
Nienburger Str. 6
D-3000 Hannover 1, W. Germany

Ward L. Ebert
The Johns Hopkins Univ.
Applied Physics Lab.
Johns Hopkins Road
Laurel, MD 20810

Harry W. Emrick
Colorado School of Mines
Office of Continuing Ed.
Golden, CO 80401

W.E. Farrell
Systems, Science & Software
P.O. Box 1620
La Jolla, CA 92038

Martine Feissel
Bureau International de l'Heure
61 Avenue de l'Observatoire
75014 Paris, France

Charles J. Finley
NASA, Code ERG-2
Washington, D.C. 20546

R.E. Fischell
Applied Physics Lab.
Johns Hopkins Univ.
Johns Hopkins Road
Laurel, MD 20810

Thomas L. Fischetti
NASA
600 Independence Avenue
Washington, D.C. 20546

Evan Fishbein
Pennsylvania State Univ.
Dept. of Geosciences
446 Deike Bldg.
University Park, PA 16802

Edward A. Flinn
NASA Headquarters
Code ERG
Washington, D.C. 20546

Angel Fonseca
Ministerio Energia y Minas and
UGGI
Torre Norte PISO 19 C-SB
Caracas 101, Venezuela

E.M. Gaposchkin
Smithsonian Astrophysical Obs.
60 Garden St.
Cambridge, MA 02138

Robert R.M. Georgevic
Univ. of Petroleum and Minerals
2325 Woodlyn Road
Pasadena, CA 91104

John G. Gergen
NGS
500 Azalea Drive
Rockville, MD 70850

John M. Goodkind
Univ. of California, San Diego
Dept. of Physics, B-019
La Jolla, CA 92093

Erik W. Grafarend
Astronomical & Physical Geodesy
Univ. FAF at Munich
Werner Heisenberg Weg 39
D-8014 Neubiberg
West Germany

Larry Greischar
Univ. of Wisconsin
1215 W. Dayton St.
Madison, WI 53706

Bernard Guinot
Bureau International de l'Heure
Observatoire de Paris
61, Avenue de l'Observatoire
75014 Paris, France

T.L. Gunther
Computer Sciences Corp.
8605 Seven Lucks
Bethesda, MD 20034

D.P. Hajela
Va. Polytechnic Inst. & State Univ.
219 C Norris Hall, Civil Eng. Dept.
Blacksburg, VA 24060

Lars Ake Haller
The National Land Survey of Sweden
S-801 12 Gävle
Sweden

James A. Hammond
Air Force Geophysics Lab.
Hanscom AFB, MA 01731

J.C. Harrison
CIRES
University of Colorado
Boulder, CO 80309

Rick Heestand
Princeton Univ.
Geology Dept.
Princeton, N J 08540

Heinz Henneberg
Univ. of Zulia
Apartado 6
Maracaibo, Venezuela

Jack E. Henrich
IBM
18100 Frederick Pike
Gaithersburg, MD 20760

S. Hieber
E.S.A.
8-10, rue Mario Nikis
75738 Paris, France

Sandford R. Holdahl
National Ocean Survey
11400 Rockville Pike, Rm. 426 H
Rockville, MD 20852

Larry Hothem
NOAA, NOS
National Geodetic Survey
ATTN: C133
Rockville, MD 20852

Phineas J. Icenbice, Jr.
JMR Instruments, Inc.
20621 Plummer Street
Chatsworth, CA 91311

Robert C. Jachens
U.S. Geological Survey
345 Middlefield Rd.
Menlo Park, CA 94025

William M. Kaula
U.C.L.A.
Dept. of Earth & Space Sciences
Los Angeles, CA 90024

Robert W. King
Mass. Institute of Technology
Dept. of Earth & Planetary Sciences
54-627
Cambridge, MA 02139

W.J. Klepczynski
U.S. Naval Observatory
Washington, D.C. 20390

Louis H. Knipling, Jr.
Naval Post Grad. Sch. Code 6405
Monterey, CA 93940

Horst Kremers
Technical University
Postfach 31 0801
D-1000 Berlin (West) 31
West Germany

Muneendra Kumar
National Geodetic Survey
Rm. 312, Rockwal Bldg.
11400 Rockville Pike
Rockville, MD 20852

J.T. Kuo
Columbia University
Henry Krumb School of Mines
New York, N Y 10027

K. Lambeck
Australian National Univ.
R.S.E.S.
P.O. Box 4
Canberra 2600, Australia

Anthony Lambert
Earth Physics Branch
1 Observatory Crescent
Ottawa, Ontario K1A 0Y3
Canada

Alfred Leick
University of Maine
103 Boardman Hall
Orono, ME 04773

G.W. Lennon
Flinders University
School of Earth Sciences
Bedford Park, South Australia 5042

Judah Levine
National Bureau of Standards
Division 542
Boulder, CO 80302

James Lewkowicz
Boston College, Physics
Chestnut Hill, MA 02167

John Linton
York Univ., Dept. of Physics
4700 Kweele St.
Downsview, Ontario
M3J 1P3, Canada

Charles A. Lundquist
NASA/Marshall Space Flt. Ctr.
Code ES01
Huntsville, AL 35812

Peter F. MacDoran
Tracking & Orbit Determination Sec.
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91103

L. Mansinha
Dept. of Geophysics
Univ. of Western Ontario
London, Ontario
Canada N6A 5B7

Charles F. Martin
Defense Mapping Agency
U.S. Naval Observatory, Bldg. 56
Washington, D.C. 20305

Marco Mattina
University of Bologna
Viale Risorgimento 2
40136 Bologna, Italy

William H. McCain
US Navoceano Code 3541
NSTL Station
Bay St. Louis, MS 39522

Dennis D. McCarthy
U.S. Naval Observatory
Washington, D.C. 20390

C. Meertens
CIRES
University of Colorado
Boulder, CO 80309

W.G. Melbourne
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91103

Jason Morgan
Princeton University
Geology Dept. Guyot Hall
Princeton, N J 08540

Jesus Moron
Universidad del Zulia
Calle 45 No. 156-45
Maracaibo, Venezuela

A. George Mourad
Battelle Columbus Laboratories
505 King Avenue
Columbus, OH 43201

Marie-Jeanne Munteanu
NASA
Goddard Space Flt. Ctr. Code G21
Greenbelt, MD 20771

Edo Nyland
Institute of Earth and Planetary Physics
Univ. of Alberta
Edmonton, Alberta
Canada T6G 2J1

Jeffrey J. Olson
University of California, San Diego
Dept. of Physics, B-019
La Jolla, CA 92093

Hyman Orlin
National Academy of Sciences
2101 Constitution Ave., N.W.
Washington, D.C. 20418

Haim Papo
Technion, IsraelInst. of Tech.
Dept. of Civil Engineering
Technion City, Haifa, Israel

Michael E. Parke
Pacific Marine Environment Lab.
3711 15th Avenue, N.E.
Seattle, WA 98105

Richard P. Peat
Defense Mapping Agency Topographic Ctr.
6500 Brookes Lane
Washington, D.C. 20315

Robert K. Perry
Naval Research Laboratory
Code 8106
Washington, D.C. 20375

J. Peters
Earth Physics Branch
Energy, Mines & Resources
1 Observatory Crescent
Ottawa, Ontario, Canada

Greg Petrie
Battelle Northwest Labs.
P.O. Box 999
Richland, WA 99352

Jacub Rais
National Mapping (Bakosurtanal)
Jl. Wahidin I, No. 11
Jakarta, Indonesia

W.J. Rapatz
Institute of Ocean Sciences
P.O. Box 6000
9860 West Saanich Road, Sidney, B.C.
Canada, V8L 4B2

C. Reigber
TU Muenchen
Technische Universit.
Arcisstr. 27
8 Muenchen 2, West Germany

Robert Reilinger
Cornell University
210 KimballHall
Ithaca, N Y 14853

N.A. Renzetti
JPL-Caltech
4800 Oak Grove Drive
Pasadena, CA 91103

Douglas S. Robertson
National Geodetic Survey
NOS, NOAA, C133
Rockville, MD 20852

M.G. Rochester
Physics Dept.
Memorial Univ. of Newfoundland
St. John's, Newfoundland
Canada A1B 3X7

Thomas P. Rooney
Air Force Geophysics Laboratory
Hanscom AFB, MA 02173

Micheline Roufosse
Smithsonian Astrophysical Obs.
60 Garden Street
Cambridge, MA 02138

Tetsuo Sasao
International Latitude Observatory
Mizusawa, Iwate-Ken
Japan

Koichi Sato
International Polar Motion Service
Internation Latitude Observatory
Mizusawa, Iwate-Ken
Japan

James C. Savage
U.S. Geological Survey
345 Middlefield Road
Menlo Park, CA 94025

Klaus Schüller
Univ. Bonn, Institut für
Theoretische Geodaesie
53 Bonn 1, Nussallee 17
West Germany

B.E. Schutz
University of Texas
Dept. of Aerospace Engr.
and Engr. Mech.
Austin, TX 78712

K.P. Schwarz
Dept. of Surveying Engineering
University of New Brunswick
Fredericton, N.B. Canada E3B 5A3

Jiro Segawa
Ocean Research Institute
University of Tokyo
1-15-1, Minamidai, Nakano-ku,
Tokyo, 164, Japan

Irwin I. Shapiro
Mass. Institute of Technology
M.I.T. 54-620
Cambridge, MA 02139

James A. Shearer
Air Force Geophysics Lab
26 Paul Revere Road
Acton, MA 01720

Peter J. Shelus
McDonald Observatory
Astronomy Dept.
Univ. of Texas at Austin
Austin, TX 78712

Eric C. Silverberg
University of Texas at Austin
Dept. of Astronomy
Austin, TX 78712

David E. Smith
NASA Goddard Space Flight Ctr.
Code 921
Greenbelt, MD 20906

D.E. Smylie
York University, Dept. of Physics
4700 Kweele Street
Downsview, Ontario
M3J 1P3, Canada

Richard A. Snay
Nat'l. Oceanic & Atmospheric Adm.
C1314
Rockville, MD 20852

Tomas Soler
NGS, NOS, NOAA C 1x6
Rockville, MD 20852

Sean C. Solomon
Mass. Institute of Technology
M.I.T. 54-522
Cambridge, MA 02139

J. Arnold Soltz
Charles Stark Draper Lab, Inc.
Rm. 3115, MS No. 20
555 Technology Square
Cambridge, MA 02139

F.N. Spiess
Marine Physical Laboratory
Univ. of California, San Diego
San Diego, CA 92152

Hans Sünkel
Inst. of Physical Geodesy
Technical Univ. at Graz
Steyrergasse 17
A-8010 Graz, Austria

Byron D. Tapley
University of Texas at Austin
Aerospace Engineering/
Engr. Mech.
WRW 402
Austin, TX 78712

Wolfgang Torge
Inst. Theor. Geod., T.U. Hannover
Nienburger Str. 6
D3000 Hannover
Germany (Fed. Rep.)

D.L. Turcotte
Cornell University
Dept. of Geological Sciences
Kimball Hall
Ithaca, N Y 14853

Petr Vanicek
Dept. of Surveying Engineering
Univ. of New Brunswick
Box 4400
Fredericton, N.B. Canada E3B 5A3

Johan H. Vas
Higher Geodesy Surinam
Maystraat 39
Paramaribo, Surinam
South America

Rob Van der Voo
Univ. of Michigan
Dept. of Geol. and Mineralogy
Ann Arbor, MI 48109

George Veis
National Technical Univ.
9, K. Zographou Str.
Athens 147
Greece

N.J. Vlaar
State Univ. Utrecht
V.M.L., Lucas Bolwerk
6, Utrecht, Netherlands

R.O. Vicente
University of Lisbon
R. Mestre Aviz, 30, R/C
Lisboa 3, Portugal

Joseph J. Von Schwind
Naval Postgrad. School
Code 68 VS,
Dept. of Oceanography
Monterey, CA 93940

J. Wahr
CIRES
Univ. of Colorado
Boulder, CO 80309

Richard J. Warburton
University of California, San Diego
Dept. of Physics, B-019
La Jolla, CA 92093

A.B. Watts
Lamont-Doherty Geological
Observatory
Palisades, N Y 10964

Donald O. West
Woodward-Clyde Consultants
Suite 700, No. 3 Embarcadero
Center
San Francisco, CA 94111

Charles T. Whalen
National Geodetic Survey
NOAA-NOS-NGS
Rockville, MD 20852

James H. Whitcomb
Seismological Lab.
Caltech 252-21
Pasadena, CA 91125

John A. Whitehead
Woods Hole Oceanographic Inst.
Woods Hole, MA 02543

Charles A. Whitten
NOAA, NOS (retired)
9606 Sutherland Road
Silver Spring, MD 20901

John Wickham
Univ. of Oklahoma
Dept. of Geol. & Geoph.
Norman, OK 73069

Sydney O. Wigen
Fed. Govt. Canada
Institute of Ocean Sciences
Box 6000, Sidney B.C.
V8L 4B2 Canada

G.A. Wilkins
Royal Greenwich Observatory
Herstmonceux Castle
Hailsham, E. Sussex
BN27 1RP, England

Peter Wilson
Institut f. Angewandte Geodaesie
Weinbergstr. 9
623 Frankfurt Sindlinger
West Germany

Gernot M.R. Winkler
Time Service Division
U.S. Naval Observatory
Washington, D.C. 20390

Susanna Zerbini
University of Bologna
Instituto di Geofisica
Viale Risorgimento 2
40136 Bologna, Italy

J. Zschau
Institut für Geophysik
Neue Universitat
2300 Kiel
West Germany

Participants from The Ohio State University
Department of Geodetic Science
1958 Neil Avenue
Columbus, Ohio 43210

Muzaffer Adiguzel
Ali Badei Ali
Brent A. Archinal
Anthony Ashiofu
Farid Askari
Edward M. Baker
Yehuda Bock
R. Ben Buckner
Robert Burtch
Roberta Carroll
Charles W. Challstrom
Michael A. Chapman
Thomas Croxell
Milud El-Gammudi
Kamil Eren
Tomas Garzon-Rojas
Carol A. Greco
Joshua Greenfield
John Hannah
Billy Hamed Isah
Masao Ishihara
Huseyin Baki Iz
Robert P. Jacober, Jr.
Christopher Jekeli
Kostas Katsambalos
Lenny A. Krieg
Swatchai Kriengkraipet
Joseph Loon
Alphonse Lwangasi
Susan May
Surindar P. Mehta
Dennis G. Milbert
Jagdish Mitter
Ivan I. Mueller
Dawood S. Muhammad
Leslie E. Ott
Despina Pavlis
Erricos Pavlis
Richard H. Rapp
Gregory A. Robinson
Halil Alev Seker
Asok K. Sen
Lakshmi P. Sharma
Helen E. Sharpe
Anwar Siala
Randal E. Simpson
Stephen D. Taylor
Thomas O. Tindall
Urho A. Uotila
Boudewijn H.W. van Gelder
Susan Von Gruenigen
Richard Ward
Indrajith D. Wijayaratne
Edward J. Wright
David B. Zilkoski

## Author Index