NASA Technical Memorandum 100716

On Principles, Methods and Recent Advances in Studies Towards a GPS-Based Control System for Geodesy and Geodynamics

Demitris Delikaraoglou
NASA/Goddard Space Flight Center
Greenbelt, Maryland

National Aeronautics and
Space Administration

Goddard Space Flight Center
Greenbelt, Maryland 20771

1989

LIST OF CONTENTS

LIST OF FIGURES

PREFACE

In the past decade or so, there has been considerable interest and progress in the development and utilization of space techniques for precise measurements of geodetic baselines, earth orientation, and various geodynamic studies, especially for measuring large-scale distortions within plates and determining the rates of interplate motion. These methods rely heavily on extra-terrestrial reference sources such as the distant quasars or other compact extragalactic objects used in *Very Long Baseline Interferometry* (VLBI), or the moon used for *Lunar Laser Ranging* (LLR) and the low-earth satellites such as LAGEOS and STARLETTE used in *Satellite Laser Ranging* (SLR). Notably, VLBI and SLR have achieved significant superiority over other conventional approaches for measuring vector baselines very precisely. As currently applied, VLBI and SLR have reached a level of maturity that to date can be used to measure routinely baseline vectors with lengths up to intercontinental distances with repeatabilities of 0.01 parts per million or better in both length and orientation.

In both techniques a number of fixed stations are used to determine the variations of the earth's angular orientation in space, to measure plate tectonic motions through monitoring of the locations of the fixed stations with respect to each other, and to contribute to the maintenance of a reference frame with respect to which the motion of additional points of interest can be determined by means of mobile VLBI and SLR equipment. However, in addition to the high cost of instrumentation and operation, such mobile systems can still be somewhat limited in their ability to occupy sites which are not easily accessible, thus limiting their use for many regional geodetic and geodynamic applications where more measurements of this type are needed, at more frequent intervals in time and space. Operational costs are particularly high for SLR due to the system's susceptibility to weather. Typically, 5 to 30 days for fixed (and up to 60 days for mobile) site occupations are required if the length of intersite baselines up to intercontinental distances were to be determined to a precision of 3-5 cm. By contrast, for the most basic IRIS network of fixed VLBI sites (i.e. the POLARIS and Wettzell observatories) the time typically required to achieve sub-decimeter accuracies in the determination of baselines of similar lengths corresponds to observation intervals of the order of 24 hours. Mobile VLBI systems are less sensitive to adverse weather, but involve considerable operations since the smaller-diameter antennas yield less-sensitive interferometers than normally achievable with the larger fixed antennas which, in turn, impose severe limitations in the observing schedules (often restricting observations to the stronger sources) and tend to distort the experimental geometry and observing strategies. This decrease in sensitivity can, in principle, be compensated for by high-gain antennas, low-noise receivers and multi-observing sessions but not without the expense of all the attendant complexities and increased cost of operations.

Although these technologies are becoming an increasingly important tool for geodynamic studies, the future role of mobile VLBI and SLR may well be fulfilled by using alternative techniques such as those utilizing the signals from the *Global Positioning System* (GPS) which, already without the full implementation of the system, offers a favorable combination of cost and accuracy and has consistently demonstrated the capability to provide high-precision densification control in the regional and local areas of the VLBI and SLR networks. Although GPS itself is still technically in its testing phase and is not expected to become fully operational until the early 1990's with the

placement of 18 operational and 3 active spare satellites in equally-spaced orbits (three satellites in each of six orbital planes), it has already proved its usefulness in measuring relative and absolute positions time and again. Numerous studies, tests, comparisons and actual full-scale projects have shown that currently accuracies of a few parts in 10^6 over distances up to a few hundred kilometers and observing time periods ranging from of 3-5 hours to as little as a few minutes are being achieved routinely with standard receiver equipment. Accuracies of a few parts in 10^7 over longer distances are becoming increasingly likely with extended modelling of the dominant error sources, whereas with carefully designed experiments, these errors may be reduced with special efforts to improve the satellite ephemerides to a few parts in 10^8. This level of performance requires, in addition to the special orbit refinement efforts, high-performance GPS receivers and reliable atmospheric calibrations (especially of the wet troposphere, for instance, by water-vapour radiometers).

The high-performance, low-cost versatility and potential of GPS-based geodetic systems have sparked an intense interest within several government agencies in the U.S., Canada, Europe and Australia which have initiated several system studies and identified promising geodetic applications, including the expected modes of operation. Performing simultaneous observations at widely separated permanent control sites together with mobile GPS receivers would conceivably allow centimeter-level accuracies for baselines up to thousands of kilometers in length. Even back-packable GPS satellite observing systems would thus afford the ability to measure baselines between VLBI or SLR fixed sites and arbitrary locations accessible to the GPS systems. This combination will allow several more points to be tied to the VLBI and SLR networks with comparable accuracy for detailed regional and local monitoring that is rarely available today. This premise has led the U.S. National Aeronautics and Space Administration (NASA) and the Jet Propulsion Laboratory (JPL) to explore the feasibility of the fiducial concept approach as it is commonly referred to (e.g. Davidson et al, 1985) which, as envisaged, will enable the joint determination of GPS satellite orbits and geodetic baselines in a manner much more suitable for the kind of measurements needed for continuous tectonic monitoring of areas of geophysical and geodynamic interest. Similar plans currently underway in Canada (e.g. Delikaraoglou et al, 1986) include the implementation of an Active Control System (ACS) based on the concept of a sparse network of control stations where continuous GPS tracking is carried out. This GPS network would be combined with a number of widely spaced, fixed VLBI antennas to ensure consistent scale and orientation of the terrestrial networks positioned with GPS with respect to the ACS system.

There are several factors that must be considered in the assessment of the potential applications and the need for a strategy towards the implementation of a fully operational fiducial or ACS network. There seem to be three main tasks in the way to the full implementation of such a system: setting up automated GPS stations (e.g. the Canadian ACS units are designed to be microcomputer-controlled, with suitable communication interfaces and ability to monitor the station operations and accomplish other station-keeping functions automatically); financing the whole operation; and analyzing the data to generate and distribute the satellite orbits and organize an efficient system and means of transmitting ACS data to the various users. All that can be stated about the first two is that they are quite considerable. The last one is also on the formidable side; of the three, this is the only one which this work is concerned with. The main approach has been based on three main issues, that firstly, to obtain accuracies at 0.01 ppm or better, existing models and software for the processing of GPS differential observations should be improved considerably through analysis of existing and future GPS observations; furthermore, investigations in improving the accuracy of

the orbital information for the GPS satellites should continue through modelling refinements of the orbital dynamics of this particular type of satellite; and finally, consideration of the problem of the optimal GPS network design strategies should be undertaken.

In assembling the material for this report we have reviewed in Section 1 the need for both VLBI and SLR vis-à-vis GPS and have outlined the capabilities and limitations of each technique and how their complementary applications can be of benefit to geodetic and geodynamic operations. To establish the accuracy requirements using GPS for a relative positioning accuracy at the level of 0.01 ppm or better, the effects of various biases and errors are briefly examined in Section 2. Ways of minimizing these effects in general and within the context of this study in particular are also examined. Of these, the orbit errors have been recognized since the early days of GPS as constituting a severe problem. Things have improved considerably lately, although these errors still remain mostly a nuisance that can be mitigated with various sophisticated models, but still limit the usefulness of the data in many applications. For this reason, the development of the models pertinent to the problem of orbit estimation for the GPS satellites receives particular attention in Section 3, where some detailed background and the models used in this study are presented. In doing so, we have followed the standard formalism based on celestial mechanics which can give further insight into the way empirical methods work and perhaps into how to improve them, in the hope that this rather detailed introduction to the otherwise basic concepts can be of some value to those who wish to understand the nature of the orbit errors better and how these are shaped for the particular type of GPS satellite. The basic principles of the coordinate and time frames inherent in the reduction of GPS data are reviewed in Section 4. Current activities in establishing GPS-based automated geodetic control systems are reviewed in Section 5. Strategies for the simultaneous determination of GPS satellite orbits and geodetic baselines are also examined in Section 5, followed by actual results of data reductions carried out to test these strategies and the general methodology described in this report. Finally, a summary of the goals achieved with this study and directions for future work are given in Section 6.

Acknowledgements

This research was carried out while the author held a U.S. National Research Council National Academy of Sciences Research Associateship at NASA Goddard Space Flight Center. This support and the additional funding for travel is gratefully acknowledged. Dr. R. H. Manka and Mr. D. T. McHugh at the Associateship Program Office made sure with utmost thoroughness that my visit to GSFC was uneventful, pleasant and scientifically productive. During this time, I have accumulated reasons for being thankful, particularly to Dr. David Smith, Head, Geodynamics Branch, NASA/GSFC for acting as advisor during my stay at NASA and for making possible my visit to the Geodynamics Branch, and also to all the colleagues and personnel of the Branch for assisting me in many ways.

Dr. Oscar L. Colombo of EG&G/WASC Inc. gave my ideas the benefit of the doubt and listened to them patiently, while also sharing his time to familiarize me with the intricacies of his solar radiation model, a key element in the estimation of the GPS orbits and a necessary one before embarking on some extensive and expensive calculations of my own. Assistance from, and discussions with, Drs. Erricos Pavlis and Nikita P. Zelensky also of EG&G/WASC Inc. are much appreciated and acknowledged. Thanks are also due to Dr. Yehuda Bock of MIT for the provision of the GPS data from the POLARIS sites. Last, but not least, thanks are also extended to the Canadian Geodetic Survey for granting me leave of absence and for the provision of the Summer '86 GPS data.

SECTION 1

THE NEED FOR VLBI/SLR VIS-À-VIS GPS

1.1 Requirements for Geodetic and Geodynamic Control

Geodesy and geodynamics derive much of their observational techniques from precise positioning measurements made by the ever-increasing use as reference sources of artificial satellites or extra-terrestrial objects such as quasars. The modern measurement types include: very long baseline interferometry (VLBI); lunar and satellite laser ranging (LLR and SLR respectively); and "ranging" to the Global Positioning System (GPS) of satellites. In all these techniques, the observed objects are used to approximate a non-rotating reference frame either directly, as in the case of the slow moving quasars, or from the dynamical theories of their motion in the case of the moon and the artificial satellites. Notably, VLBI and SLR have the capability of providing three-dimensional coordinates of relative positions of reference points with an accuracy of 5 to 10 cm or better, practically for any baseline length. Repeated observations at that level of accuracy can indicate variations in the positions due to crustal dynamic processes of the earth on both regional and global scales, and if carried out at frequent time intervals, variations in the earth's rotation.

Fixed VLBI or SLR stations alone can be used to study such motions if they are located in the geophysically active regions, and to contribute to the maintenance of a reference framework with respect to which the motion of additional points of interest can be determined by means of mobile units. In fact, mobile systems are being used extensively within the current activities of NASA's Crustal Dynamics Project (CDP) to measure baselines between the fixed sites and locations accessible to the mobile systems, thus resulting in much stronger networks and making it possible to directly assess deformations associated with tectonic activity in regions within plate boundaries and their peripheries. Such observations are extremely valuable and well suited to the capabilities of these techniques. Today, however, there is an increasing need for more geodynamic support requiring accuracies of the order of 0.1 to 0.01 ppm over distances from 100 to 10000 km (Hothem and Williams, 1985) and at more frequent intervals both in a spatial and temporal sense. Such measurements could provide the necessary information for the adequate characterization of plate velocities and the complex patterns of crustal strain so necessary for the full understanding of the tectonic processes occuring at the interplate boundaries. VLBI and SLR networks in this context are expensive propositions, however, since rapid repeat measurements are difficult to carry out. In the near future, it may be possible to monitor geodetic networks with spaceborne laser systems that range to corner retroreflectors on the earth's surface (Cohen et al., 1986). The feasibility of such systems, however, has not yet been demonstrated and the state of current technology suggests that their deployment may be a decade or more away.

This role may well be fulfilled by using alternative techniques such as those utilizing the signals from GPS which, already without the full deployment of the system, has the same uncertainty level for baselines less than about 1000 km in length as

mobile VLBI or SLR. Given the small, relatively inexpensive and highly portable receivers, along with the shorter observing requirements and hence, the smaller operating costs of GPS campaigns, it appears that GPS may offer a definite advantage for this type of application. A similar impact is expected to be realized in the areas of the conventional geodetic networks. In their educated opinion, Vanicek et al. (1984) and Delikaraoglou and Steeves (1985) predicted that the shift to relative positioning over thousands of kilometers instead of a few tens of kilometers or less, will eventually give way to a simple reference framework of accurately positioned sites that will act not only as foothold points for the relative positioning techniques, but also provide the link to specific coordinate systems. In fact, just such a reference framework is currently being tested in both the U.S., as part of the validation of JPL's proposed Fiducial Network System concept (e.g. Thornton et al, 1986), and in Canada (Delikaraoglou et al, 1986) as part of the development of a GPS-based Active Control System (ACS)-see also Section 5.

In the remainder of this section we review briefly the VLBI and SLR techniques vis-à-vis GPS mainly for the purpose of gaining a perspective in their respective application areas and for placing the role of GPS in context. We shall not dwell in detail into the theory of these techniques as it can be found adequately described, for example, in IEEE (1985), AGU (1985) and Wells et al. (1986). Rather, we shall attempt to outline the relative merits and disadvantages of each and explain how their complimentary application can be beneficial.

1.2 Capabilities and Limitations of Available Techniques

1.2.1 Very Long Baseline Interferometry (VLBI)

The technique of very long baseline interferometry (VLBI) has been used extensively since the late sixties to provide, among other applications, measurements of the earth orientation, global and regional crustal motion, deep-space spacecraft navigation, and three-dimensional vector baselines up to intercontinental distances (Clark et al., 1985; Davidson and Trask, 1985; Herring, 1986). Using the principles of wave interference and bandwidth synthesis, signals from distant quasars or other extragalactic sources received in the antennas of two or more radio telescopes are amplified and translated to a lower frequency band under control of a hydrogen maser frequency standard; in turn they are digitized, time-tagged and recorded on wide bandwidth magnetic tapes. Tape recordings from different sites are then played back and cross-correlated to obtain estimates of *group delay* (i.e. the difference in arrival times of the quasar signal wavefronts at the radio telescopes) and *phase-delay* of the sample cross-correlation function and its time rate of change or *delay rate*. Group delays are measured in two widely separated frequency bands in order to derive a new observable which is free of ionospheric (or more generally, plasma) delay. For geodetic VLBI these frequencies are in the microwave portion of the electromagnetic spectrum; e.g. in the S-band (≈ 2.2 GHz) and X-band (≈ 8.4 GHz) used in the Mark-III system. These observables depend among other factors on the direction of the source(s) and the vector separation and orientation of the antennas of the observing sites. Hence, it is possible to estimate these and other quantities by measuring the delays and their rates of change for many different sources and many baselines observing simultaneously through a multi-parameter least-squares fitting procedure in which geophysical, astrometric and station coordinates are adjustable and/or recoverable parameters.

To date, VLBI can provide the highest possible geodetic control needed for the precise monitoring of crustal deformations. This is mainly due not only to its high accuracy but also to its excellent repeatability. The latter is mainly ensured because of the very geometric nature of VLBI and the coordinate system defined by the observed radio sources whose positions as viewed from an earth observer define essentially a quasi-inertial reference system (which is as close as we can get to an inertial system once the small undetected proper motions and the earth's own inherent acceleration in space are neglected). In the analysis of VLBI observations, this system is eventually related to a geocentric, earth-fixed coordinate system with the orientation of the earth being defined by the pole position, UT1, and by the standard formulas for sidereal time, precession, nutation and earth tides, etc. In this context there seems to be a generally accepted consensus that VLBI is the only technique that would allow the unambiguous establishment of the future conventional terrestrial reference frame through its realization by a global network of VLBI stations and the global tectonic motions of those positions. Satellite techniques like SLR and GPS do not have direct access to such a quasi-inertial reference frame. Besides the lack of direct connections of the coordinate system actually employed and the inertial frame, these techniques are sensitive by contrast to gravitational forces hence, insofar as this knowledge is incomplete, errors in computing the force field in which the satellites move could and do lead to incorrect satellite positions which, in turn, affect the geodetic coordinates of the ground stations. Furthermore, such techniques basically observe range or range differences and contain no information about the rotational or translational degrees of freedom of the satellite networks, thus resulting in additional estimability problems (Grafarend and Müller, 1985; Delikaraoglou,1985) which can be overcome only after carefully chosen geometric or dynamic constraints are imposed.

The use of VLBI is not limited to large fixed radio telescopes. Since the early 1980's, mobile systems have been developed by JPL for the NASA Crustal Dynamics Project and have been operated in conjunction, for example, with several fixed large antennas in the Western United States and Canada to establish relative motions near the tectonic plate boundaries in California and Alaska. The small (from 3.0 m to 9.0 m diameter) antennas used with such systems are a major advantage for geodynamic applications, particularly in allowing the establishment of control within the global networks of the fixed sites and more frequent measurements (both in space and time) to adequately determine the deformations within the plates and their peripheries. Like their fixed counterparts, mobile VLBI systems have essentially all-weather capability but involve considerably more complex operations and poorer system consistency, mainly because the smaller size antennas yield less sensitive interferometers than normally achievable with the larger fixed antennas. A direct consequence of this is the reduction of quasar signal-to-noise ratio which, in turn, imposes severe limitations in the observing schedules (often restricting observations to the stronger sources) and tends to distort the experimental geometry and observing strategies. This decrease in sensitivity can of course in principle be compensated for by higher-gain antennas, lower-noise receivers and multi-observing sessions, and indirectly by differencing the vectors between individual mobile units and a common fixed site, but not without the expense of all the attendant complexities and increased costs of operations. Naturally, there are other error sources inherent in VLBI which impart temporal variations on the baseline components. Many of these sources of error are in fact shared by systems like GPS as will be discussed in Section 2. Given the complexity of the VLBI operations, costs in maintaining VLBI networks and correlator facilities are obviously high, and rapid repeat measurements are difficult. This is, however, a small price to pay in light of the usefulness of the technique in providing a stable reference system for other geodetic and

geodynamic work including the determination of accurate orbits for systems like GPS, as will be seen in some detail in Section 5.

1.2.2 Satellite Laser Ranging (SLR)

Satellite Laser Ranging (SLR) systems have been in use since the mid-1960's. By tracking earth-orbiting satellites from a global network of laser tracking stations, information on the forces acting on the satellite can be derived from the analysis of measurements of the round-trip travel time (and hence the path length) of a photon pulse from a laser site to the corner-cube reflectors mounted on the surface of a satellite. Overall laser tracking system capabilities have increased in accuracy by a factor of 3 every 5 years over the last 20 years as noted by Pearlman (1982). To date, the distance from a tracking site to the satellite can be determined to a precision of a few centimeters using current technology and techniques. Most SLR systems now use short pulse length (100-200 ps) neodymium YAG lasers with receiving sensitivities ranging from a single photon to more than 100 photons and best single-shot range accuracy at the level of 2 cm.

Currently there are some 25 fixed SLR systems worldwide. In addition, NASA has developed and deployed eight trailer-based mobile Laser systems (MOBLAS) and four highly compact transportable-class (TLRS) systems with a further two transportable sytems (MTLRS-1 and -2) having been built in Europe. The TLRS systems were developed mainly in response to a need of obtaining SLR data at remote sites and are rapidly being relocated in support of local and regional tectonic studies. Much like the mobile VLBI systems, the compact size of the TLRS systems was achieved at the expense of greatly reduced operational signal levels (about two orders of magnitude lower) as a result of the smaller telescope apertures (typically 25 cm vs. 75 cm) and lower laser output energies (5mJ vs. 100 mJ) relative to the larger MOBLAS systems (Degnan, 1985).

Of the retroreflector-equipped satellites the most fruitful in terms of geodetic results, have been the dedicated, fully reflecting laser satellites LAGEOS and STARLETTE, and only recently the Japanese Experimental Geodetic Payload satellite, AJISAI. Both the geodetic and geodynamic inferences drawn from SLR data depend mainly on dynamic techniques which require accurate descriptions of the earth's gravity field and the precise definition of the coordinate frame of the tracking network locations, including the description of polar motion and specification or determination of fundamental constants (e.g., GM, speed of light, etc). In fact, for current laser tracking systems, errors in the gravity field models are one of the dominant errors in the recovery of site locations and certain orbital and other geophysical parameters. Several other error sources are also present that limit the accuracy of the present systems. These include uncertainties in the relative timing errors, atmospheric refraction, epoch time errors, system delay calibration errors, etc. There are also other factors that are detrimental to SLR activities. Although susceptibility to adverse weather is a major drawback, system availability and scheduling are also often cited as limiting the efficiency factor at about 30% for the current systems. As a result, operational costs are particularly high for SLR. Typically, 5 to 30 days of occupation for fixed sites (and up to 60 days for mobile systems) are required to determine the length of baseline vectors up to intercontinental distances to a precision level of 3-5 cm.

In spite of these disadvantages, SLR systems have provided very precise measurements for establishing the orientation of the earth within a quasi-inertial coordinate frame (Smith et al., 1987). This is ensured in part, by the methodology used which interrelates the laser observations through a theory of motion describing the satellite's orbital behaviour in time and subsequently determines the positions of the tracking stations in an appropriate terrestrial coordinate system. The latter, being on the earth's surface, rotates with it while the satellite orbit plane is, to a first approximation, independent of the rotation of the earth. Hence, the satellite motion (described in the inertial frame), and the stations (located in the earth-fixed coordinate frame) are linked through the earth's angular velocity vector which is oriented by the polar motion and UT1 solutions. In the case of SLR, particularly for the LAGEOS satellite, the data are sufficiently abundant to actually permit a complete solution of station positions and earth orientation parameters simultaneously with the satellite orbits-typically at a level of precision of 2 mas for the x,y axis of the earth's rotation axis and 0.2 msec for UT1 for data in 5-day intervals (Tapley et al., 1985).

1.2.3 Global Positioning System (GPS)

The Global Positioning System (GPS) of satellites is currently being developed by the U.S. Department of Defense (DoD) as the next navigation system of the future. A detailed description of GPS may be found in ION (1980, 1984), NOAA (1985), ARL (1986) and Wells et al (1986). In summary, the system is designed to provide all-weather, real-time precise position and time information with a higher degree of accuracy than ever before available. When fully operational (in the early 1990's as now postulated following the delays of the launch schedule caused by the Space Shuttle Challenger accident), the satellite segment of the system will consist of 18 operational satellites and three active spares placed in six orbital planes inclined at 55° to the equator (the present prototype satellites are at inclinations of 63° instead) at a height of about 26,000 km, having orbital periods approximately 12-hours, thus ensuring that any place on the earth will be able to "see" a minimum of four and up to eight satellites at any instant in time.

The GPS satellites carry an ensemble of atomic clocks onboard which are carefully monitored by the U.S. Naval Office (USNO) and their time scales are maintained trackable with respect to UTC-USNO to within ± 100 nsec, which ensures high satellite oscillator stability and precise timing information. Each satellite transmits at two L-band frequencies, L_1 (1575.42 MHz) and L_2 (1227.60 MHz) which allows for the elimination of the first-order ionospheric effects. Three different spread spectrum codes are modulated onto the L-band signals, including a data (D-code) modulation consisting of a satellite broadcast ephemeris message (at the rate of 50 bits/sec) and two pseudo-random noise (PRN) codes known as the coarse/acquisition (C/A) and precise (P) codes. The latter are modulated onto the L-bands at frequencies of 1.023 MHz and 10.23 MHz respectively, with the respective codes repeating every 1 msec for the C/A code and approximately every 38 weeks for the P code (although in practice the P code for each satellite is reset every week). It is the current DoD policy that the future availability of these codes, when the operational system is in place, would eventually rely on the so-called *Precise Positioning Service (PPS)* to be based on the P code on two frequencies and the *Standard Positioning Service (SPS)* to be based on the C/A code only on the L_1 frequency. In order for the SPS to be freely available, it will

necessarily be poorer in accuracy, while the PPS will be more precise but will be encrypted and thus limited in access to civil users (Baker, 1986). Furthermore, it is entirely conceivable that DoD may severely degrade the relative accuracy of the SPS in the interest of national defense.

The present (1987) status of the system is that 11 (prototype) satellites are currently in orbit, with 7 satellites being fully operational allowing 4 to 5 hours per day of four-satellite coverage in the United States and Canada, Europe, a large portion of South America, Australia, the Near East and the polar regions. Even though the system is still technically in its testing phase, it has already demonstrated its capability for determining baselines with an accuracy of 0.1 ppm over distances of a few hundred kilometers and, most importantly, with much shorter required observing periods for such accuracies than with other techniques. This, together with the all-weather capability and the true portability of the currently available GPS equipment combined with the relatively low cost of operations, clearly illustrate the apparent advantages of GPS. This should not, however, mitigate the few disadvantages inherent in the system as well, particularly the dependence of the user on the policy of the DoD. Possible arbitrary, sudden system shutdowns and/or failures and lack of precise orbit ephemerides are extreme examples of this policy which are obviously nonexistent in systems like VLBI and SLR. The possibilities of selective availability and degraded accuracy should have some effect on static relative positioning (especially on establishing appropriate modelling strategies) with the major factors affecting the accuracy of relative coordinate determinations being:

- the accuracy of the orbit representation, and
- our capability to adequately model atmospheric refraction errors, particularly those of the wet troposphere and especially over long baselines,

which would be very critical in the context of the high accuracies needed for geophysical and geodynamic applications. In spite of these drawbacks, however, GPS remains an extremely attractive measuring technique likely to fulfill the future role of mobile VLBI and SLR to provide high-precision densification control in the local and regional areas of the VLBI and SLR networks. This premise has led the U.S. National Geodetic Survey (NGS), NASA, JPL, and others to explore this feasibility through several campaigns that have been initiated to use GPS for the study of crustal motion (e.g. Strange, 1985; Thornton et al., 1986). In Section 5 we shall report on some of the most recent results of GPS relative positioning in terms of the levels of accuracy achieved by employing the methodology described in this report for minimizing the various error biases, particularly those of the satellite orbits.

SECTION 2

THE ACCURACY OF GPS RELATIVE POSITIONING

Since GPS has been designed primarily as a military navigation system to provide all-weather, real-time navigation, its principal mode of operation is real-time point positioning by ranging using the codes of either the available PPS or SPS signals. This is the way GPS is designed to be used for navigation, although even in this mode it could provide geodetic accuracies if used on the same station over an extended period of time (Wells et al., 1986). The type of observable, however, which can be used to perform precise relative positioning independent of the coded signals is that of the carrier beat phase. Carrier beat phase measurements with GPS are similar to the continuously integrated Doppler count measurements obtained from the Doppler TRANSIT satellite system. That is, a carrier phase measurement is essentially the phase of the signal which remains when the incoming Doppler-shifted satellite carrier signal is beat with the nominally-constant reference frequency generated in the receiver. Clearly, if the transmitter remains at a fixed distance from the receiver , the so-measured phase of the incoming signal remains constant. Viewed from this perspective, it is evident that the phase of the signal changes with the change of the range between the satellite and receiver: the phase changes by 1 cycle wavelength of the carrier frequency (19 cm on the L_1 and 25 cm on the L_2 GPS frequencies respectively). This type of measurement leads potentially to the most precise information about the satellite-to-receiver ranges one can obtain from GPS. However, the problem of utilizing this potential is one of ambiguity: it is quite difficult to accurately locate the cycle of the carrier whose phase is being measured. With currently available instrumentation, phase changes measured to an accuracy of a fraction of a cycle are possible. Hence, carrier beat phase measurements are sensitive to sub-decimeter range changes. However, the success of achieving this level of positioning accuracy from this type of data hinges on the capability to resolve this cycle indeterminacy (or ambiguity) including the recovery of phase cycle slips often encountered during tracking.

Relative or differential GPS positioning requires that two or more receivers are observing simultaneously the same satellites. Given the altitude of the satellites in the system, this implies a limit of about 4000-5000 kilometers on the length of baselines that can be measured and also that the data on both stations must be correlated for a solution. One can envision various levels of such differential operations. At the lowest level, it is necessary only that the operating receivers track GPS at the same time, although they need not track the same satellites. In this case, the effect of the GPS timing and propagation errors would be correlated. One level higher, the receivers would track the same satellites, although not necessarily in the same sequence. In this case, the effects of ephemeris errors and specific satellite timing biases would be correlated. At the highest level, the receivers would be forced to track the same signal wavefront from the same satellite simultaneously, in which case the correlation between the effects of satellite and propagation error sources would be at a maximum. This latter mode of operation, which is very similar to the conventional VLBI, is most often used for geodetic and geodynamic applications, since it leads to one or two orders of magnitude improvement in the results.

Differential observations can be processed on an individual baseline-by-baseline basis or all together as a network. The data processing can furthermore be done either

with the direct "raw" phase observable (i.e., in an undifferenced mode) or with some linear combination of the measured phases from the different stations. The linear combinations are usually obtained from the raw phases which are differenced further between-stations, between-satellites, and between-epochs and in combinations thereof that remove, or greatly reduce specific types of errors. A *single difference* is the difference in ranges (or phases) from two receivers observing simultaneously the same satellite. These observables cancel any satellite-specific biases such as satellite clock offsets and to a great extent (depending on baseline lengths which are short compared to the altitude of the satellites) satellite ephemeris and atmospheric refraction errors. *Double differences* are the differences between two single differences of two satellites. These observables tend to further cancel any receiver-specific biases such as receiver clock offsets (which are common contributors to both single differences), but they are mathematically correlated with others in the same stations or to the same satellites. One differencing step further, *triple differences* are obtained by differencing between double differences over two consecutive time epochs. This combination also cancels any residual satellite and receiver clock errors and any receiver-satellite pair-specific biases such as cycle ambiguities. Like the double-difference observables, triple differences are also mathematically correlated.

Theoretically, subject to including appropriate covariance matrices reflecting these correlations, the data reduction results with any of these observables should be identical. Conversely, neglecting these correlations could lead to significantly different results (Ashkenazi and Yau, 1986). Undifferenced carrier phases may also be processed directly by including extra nuisance parameters in the model to account for the various biases (Goad, 1985; Lindlohr and Wells, 1985). Furthermore, it has been shown that the undifferenced and differenced techniques are equivalent, if the same biases that are cancelled in the differenced observables are included as nuisance parameters in the undifferenced model (ibid; Bock et al., 1986; Schaffrin and Grafarend, 1986). In practice, however, these methods will give different results, particularly over the longer baselines (> 200-300 km) which are of interest to geodynamics, mainly because significant atmospheric variations reduce the correlations among the observations, thus offsetting the advantage of using such observations to reduce the existing errors in the first place. In fact, this is the main reason why the use of double differences weighted according to baseline length is argued (e.g. by Bock et al., 1986) to be the preferred method of processing since the residual atmospheric effects in these observables are approximately proportional to baseline length.

As the discussion has alluded to already, there are two major types of requirements in GPS operations that differential techniques are capable of supporting. Firstly, differential GPS must be capable of improving the performance of the system beyond the limits of its non-differential expected accuracy envelope. This is brought about merely by utilizing the differential observables just mentioned. Secondly, differential GPS must merely be capable of detecting when GPS is outside its expected accuracy range, whether this is due to intentional or unintentional degradation, signal propagation, timing, orbital or geometric causes, all of which gain particular significance in precise operations for geodynamics. This requires particular attention in understanding the "cause and effect" relationship between the various biases and errors that are inherent in the GPS measurements.

In the remainder of this section, we shall attempt to identify and describe, mostly in non-mathematical terms, these error sources with respect to the accuracy of their modelling and to describe current efforts to deal with their effect in general, and in the present study in particular.

2.1 Biases and Errors

The accuracy of positions obtained by GPS is dependent on two general influences: the errors affecting the measurements themselves and the geometric strength of the relative satellite-station configuration being observed. The geometrical aspects relating to the latter influence have been discussed, e.g. by Vanicek et al. (1984). Hence, nothing more will be said about them here since they are of no direct consequence to the present study. Later we shall only discuss the relation between orbital errors and relative position accuracies as it can be established from a purely geometric point of view.

The measurement errors generally fall into three categories: those associated with the satellites, with the signal propagation, and with the receivers. Errors from each of these sources have complex spectral properties and there are correlations between some of them, both in the temporal and spatial sense. These will be discussed immediately below.

2.1.1 Satellite Orbital Biases

Satellite-associated biases consist of errors in the orbital ephemerides (i.e., the satellite is not where the GPS ephemeris used in the data reduction tells us it is) and errors in the satellite clocks (i.e., the satellite time scale is not perfectly synchronized with the various other time scales usually involved in the data reduction-see also Section 4). These errors are uncorrelated between satellites; they affect both code and carrier beat phase measurements equally; and they depend on the number and the location of the tracking stations providing data for orbit determination and monitoring of the satellite clocks (i.e., the GPS control stations), the force models used and the satellite configuration geometry itself (Swift, 1985; Fliegel et al., 1985). Satellite ephemeris errors are the most difficult to deal with. Clocks would be improved, or their undesirable effects would be eliminated or greatly removed by the differencing of simultaneous observables from two or more receivers and satellites. Ephemeris errors on the other hand, would require better understanding and estimation of the forces acting on the satellites; a process which is hampered however, by insufficient knowledge since these forces cannot measured directly from the ground tracking sites. The main approach to solving the problems created by the imperfect knowledge of the physical phenomena responsible for these effects has been usually directed towards developing more "exact" models to begin with. This is the standard approach followed by many to date. In this study, while following the same example in making the same choice, we also have opted for the alternative approach (cf. Section 3): to use parametric models of various complexities for the ephemeris parameters themselves (i.e., parameters that can be adjusted as part of the orbit estimation process itself).

The broadcast ephemeris is known to be in error, typically by 20 to 80 m, which clearly makes it unsuitable for any precise applications, particularly over long baselines. It is also now known that deficiencies exist in the solar radiation force model presently employed in the estimation and prediction of the orbits of the prototype satellites by the Naval Surface Weapons Center (NSWC), where the so-called precise ephemeris is also being computed, and that a new improved model is to be required for

the operational satellites (Fliegel et al., 1985). Although the accuracy of the NSWC precise ephemeris is not exactly known, in an earlier investigation, Swift's (1985) evaluation indicated that it may be in error by as much as 6 m, 43 m and 25 m in the radial, along-track, and across-track directions respectively. As soon as the operational configuration of the satellites is in place, it is expected that the operational ephemeris accuracy will be of the order of 1.5 m in the radial direction (Van Dierendonck et al., 1980), although this is still thought to be quite optimistic given the current experiences. Rather, actual errors of about 5 m are thought to be more likely (Wells et al., 1986).

The relation between orbit errors and resulting relative positioning accuracies can be established from a purely geometrical point of view. The exact linear relation between the basic differential range observable $\Delta\rho^j$ and the baseline vector **B** between two sites can be expressed as shown in Fig. 2.1 (cf. also Vanicek et al., 1984) by

$$\begin{aligned} \mathbf{u}^j \mathbf{B} &= - \mathbf{u}^j \mathbf{e}^j_1 \Delta\rho^j \\ &\approx - [1 - \tfrac{1}{4} \omega^4] \Delta\rho^j \end{aligned} \qquad (2.1)$$

where

$$\mathbf{u}^j = \tfrac{1}{2} (\mathbf{e}^j_1 + \mathbf{e}^j_2) \qquad (2.2)$$

is the average of the unit vectors $\mathbf{e}^j_1$ and $\mathbf{e}^j_2$ in the direction of the satellite.

Assuming that an error in the satellite position introduces an error into the baseline vector in such a way that the observable is unaffected, one can obtain a first-order approximation of the error equation that relates these errors by replacing $\mathbf{e}^j$ and **B** by $\mathbf{e}^j+\mathbf{de}^j$ and **B+dB**. This results in the error equation

$$(\mathbf{de}^j_1 + \mathbf{de}^j_2)\, \mathbf{B} + (\mathbf{e}^j_1 + \mathbf{e}^j_2)\, \mathbf{dB} = 0 \qquad (2.3)$$

which defines a constraint for the effect of the orbital errors on a baseline determination and makes a unique solution for **dB** possible if the condition that the magnitude of **dB** be a minimum is imposed along with (2.3). The form of this error equation indicates that the errors introduced into the baseline vector depend on the magnitude and the orientation of the orbit error vector relative to the baseline vector **B**. Buffett (1985) has shown that the solution of the error equation that maximizes the effect of the orbit errors is obtained when **dB** is parallel to $\mathbf{e}^j_1$ and $\mathbf{e}^j_2$ which corresponds to a more convenient rule of thumb that has also been adopted by several investigators (e.g. Vanicek et al., 1985), i.e.

$$(dB/B) = (d\rho/\rho) \qquad (2.4)$$

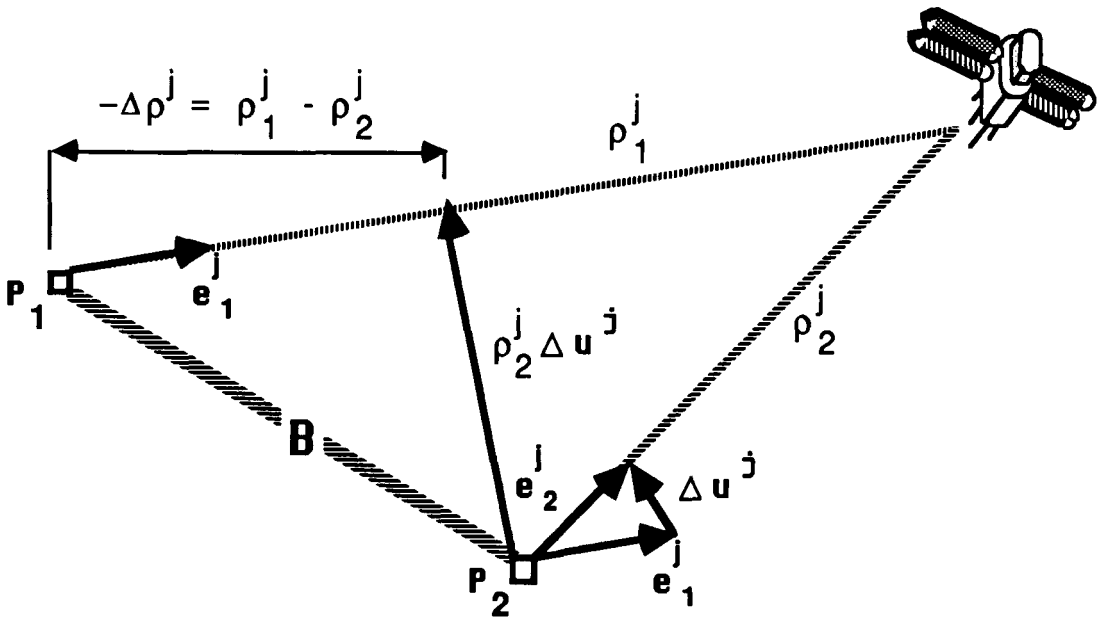

Fig. 2.1 - Differential range geometry.

where $d\rho$ is the orbit error expressed in terms of a topocentric position error and ρ is the average topocentric distance (≈20,000 km) to the satellite. The implication of the rule of thumb for relative positioning is that although for short baselines (< 200 km) sensitivity to the orbit errors is reduced by a factor of 100 more, for centimeter geodesy over distances greater than a few hundred kilometers, orbit errors still limit system performance. This implies that for regional and global geodynamic applications which require 0.01 ppm accuracies (i.e., 1 cm over 100 to 1000 km), the satellite orbit errors should not exceed 2 m (in equivalent range error) for centimeter accuracies over 100 km and 20 cm for the same accuracy over 1000 km. Consequently, the need to improve the GPS orbits has been recognized early on to be a necessary part of improving the current GPS capabilities.

2.1.2 Propagation Biases

Propagation errors consist of refraction errors due to the influence of the ionosphere and the troposphere, multipath, imaging and phase-center variations due to the mutual interference of two or more signals emitted from the same source and

reflected along different paths by surrounding reflective surfaces in the vicinity of the GPS antennas.

2.1.2.1 Ionospheric delay errors

GPS code measurements are subject to ionospheric group delay, and carrier beat phase measurements to ionospheric phase delay. At the GPS frequencies, the plasma of the ionosphere (50-1000 km above the earth surface) ionized by ultraviolet radiation from the sun can result in range delays varying from 50 m up to 150 m (at maximum sunspot activity, midday, satellite near horizon) to less than a meter (sunspot minimum, night, satellite at zenith). Since ionospheric refraction is frequency-dependent, measurements on the L_1 and L_2 frequencies can be compared to estimate the first-order effect represented by an expression of the form (Spilker, 1980)

$$a/f^2 + b/f^3 + c/f^4 + \ldots \qquad (2.5)$$

where f is the carrier frequency and a,b,c,... are parameters which are functions of the total electron content along the signal path. For the carrier beat phase measurements, for instance, this correction to the L_1 phase observations takes the form

$$d\phi_1 = [\phi_1 - \phi_2(f_1/f_2)]\, f_2^2/(f_2^2 - f_1^2) \qquad (2.6)$$

where ϕ_1, ϕ_2 are the phase observations in the L_1 and L_2 frequencies. This first-order correction removes about 99% of the range delay effect, leaving some 8 cm at the most as a result of neglecting the higher order effects b/f^3 and c/f^4 (ibid). However, there is an undesirable by-product of this correction: noise on each of the L_1 and L_2 phase observables will result in increased measurement noise in the corrected observation. Assuming for the L_2 a noise amplification of 1.35 times that for the L_1 observable, a simple propagation of errors using (2.6) reveals that the noise level of the dual-frequency estimate is 3.3 times that of the uncorrected L_1 observable which illustrates that there is a point at which making the dual-frequency correction becomes counterproductive. In differential positioning the corresponding effects at the ends of a baseline are obviously of concern. Clearly, for short baselines (<50 km) where there is very high correlation of the effect, ionospheric errors can be greatly reduced simply by single differencing or even be ignored altogether. This has been verified experimentally by several investigators (e.g. Kleusberg, 1986). Of course for double-difference observables, there is a further amplification of the noise level due to the linear combination for both the dual-frequency correction and the double-differencing operation. Thus, the use of dual-frequency combinations is generally not advisable for such short baselines. For longer baselines of 100 to 1000 km, there is a different situation since signals between satellites spaced far apart (for optimal geometry) will generally pass through different and uncorrelated parts of the ionosphere; thus one is obliged to remove the ionospheric effect using the dual-frequency correction. In this

situation the horizontal gradients of the ionosphere will be the determining factor of the residual effect. Usually, the largest gradients will be in the north-south direction at mid-afternoon and in the east-west direction at sunrise, and will reach values between about 4 m/1000 km during the maximum solar activity cycle and 1 m/1000 km during solar minimum (Clynch and Coco, 1986).

A simple way to account for these residual effects is to consider that the total ionospheric effect may be expressed as a function of the total electron content N_e of the ionospheric layer as

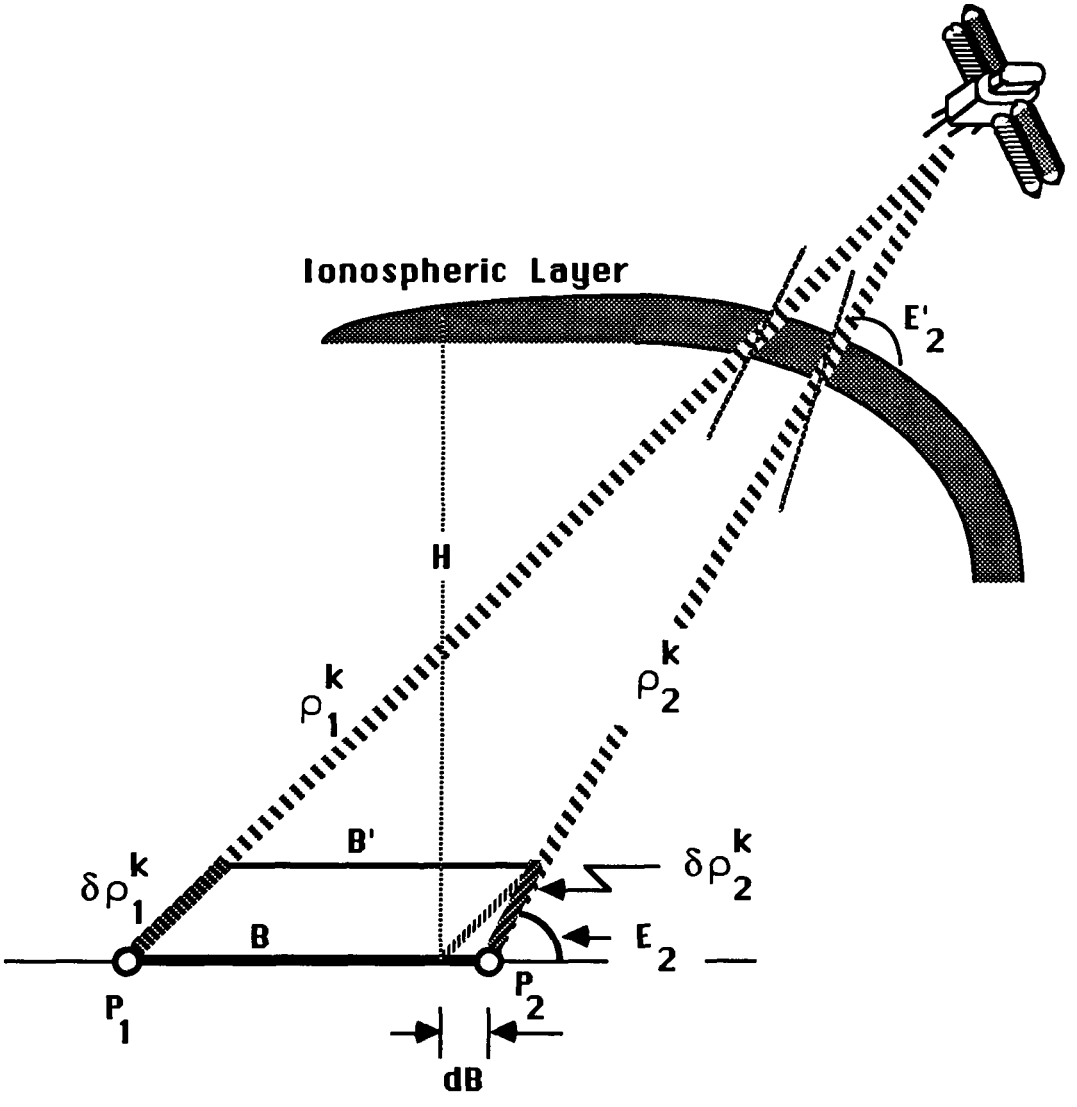

Fig. 2.2 - Influence of ionospheric refraction on GPS observations.

$$\delta\rho_{ion} = (k\, N_e / f_L^2) = (k\, N_v / f_L^2 \sin E') + (k\, \delta N_v / f_L^2 \sin E') \qquad (2.7a)$$

$$= IER + (IRE_v / \sin E') = IER + IRE_s \qquad (2.7b)$$

where k is a known constant (k=40.5284), f_L is the carrier frequency in hertz, N_v is the estimate of the vertical electron content and δN_v denotes an error in N_v. IER is the first-order effect accounted for by the two-frequency correction (2.6), while IRE_s expresses the residual ionospheric error in terms of the vertical electron density error IRE_v. The simple mapping function

$$IRE_s = IRE_v / \sin E' \qquad (2.8a)$$

is a result of assuming a thin ionospheric layer at a mean ionospheric height (usually set at H = 400 km; i.e. 50 km above the layer of maximum density in the mid-latitude regions) onto which the entire electron content is assumed to be condensed; E' is the elevation angle of the slant range at the crossing point of the ionospheric layer. In more rigorous terms, the ionospheric crossing point should be described by two parameters: the station latitude and the local solar time (given that the ionosphere is strongly dependent on the hour angle of the sun; i.e. day vs. night observations). In simple terms, however, it can be determined as a function of the satellite elevation and the geocentric station vector R as

$$\cos E' = [R / (R+H)] \cos E \quad . \qquad (2.9a)$$

The residual ionospheric error IRE_s can then be computed as

$$IRE_s = RREN \times (IER)^2 \qquad (2.8b)$$

where RREN is a normalizing factor approximated as

$$RREN = a + b\, e^{-E'/20.7} \qquad (2.10)$$

where a, b are coefficients to be determined in the course of the adjustment of the data along with other nuisance parameters.

Unmodelled ionospheric effects in GPS carrier beat phase observations lead to an increased noise level by a factor of 2-3 in the adjustment residuals (cf. Kleusberg, 1986) and to a scale effect on the so-determined baselines (cf. Beutler et al., 1987). As a consequence, a baseline B' derived from single-frequency GPS observations could be shorter with respect to the true baseline B as shown in Fig. 2.2 by an amount (dB)

proportional to the difference of the ionospheric error $(d\rho_{ion})_{12} = \delta\rho_2^k - \delta\rho_1^k$ at the two ends of the baseline; i.e.

$$dB \cos dA = (d\rho_{ion})_{12} / \cos E \qquad (2.11)$$

where dA and E are the difference in the corresponding azimuths and the average elevation angle to the satellite respectively. The combined effect of all observations can be simply determined by the geometrical configuration and the exhibiting balance of all azimuths and elevations over all satellites observed within a given session; i.e. as a rule of thumb,

$$\Delta B/B = \sum_i [(d\rho_{ion})_{12} / B \cos dA \cos E]_i \qquad (2.12)$$

expressed in mm/km. This amounts to about 0.25 ppm per 1 m of unmodelled (or improperly modelled) ionospheric error. For numerical computations, Georgiadou and Kleusberg (1988) have shown that $(d\rho_{ion})_{12}$ can be conveniently expressed as

$$(d\rho_{ion})_{12} = dV\ B\ [R/(R+H)^2]\ \sin E\ \cos E\ (\cos z')^{-3} \times$$
$$\times \{1 + R/r\ \sin E + (R/r\ \sin E)^2\} \qquad (2.13)$$

where r is the geocentric distance of the satellite, z' is the zenith distance corresponding to the given elevation angle E and computed from (2.9a) as

$$\sin z' = [R/(R+H)] \cos E\ ; \qquad (2.9b)$$

the vertical ionospheric delay dV is a function of the electron content

$$dV = 40.3\ N_V / f_L^2 \qquad (2.14)$$

which itself is the most influencing factor given that its variability is the strongest both spatially and temporally as well as regionally. Layer height plays an important role, particularly in elevations of about 15^o. For a vertical delay of 1 m and a baseline of 1 km, $(d\rho_{ion})_{12}$ as computed from (2.13) may vary in the fashion illustrated in Fig. 2.3 showing the dependence of the effect on the assumed height of the ionospheric layer at three altitudes: H=300 km (low), H=400 km (normally assumed) and H=500 km.

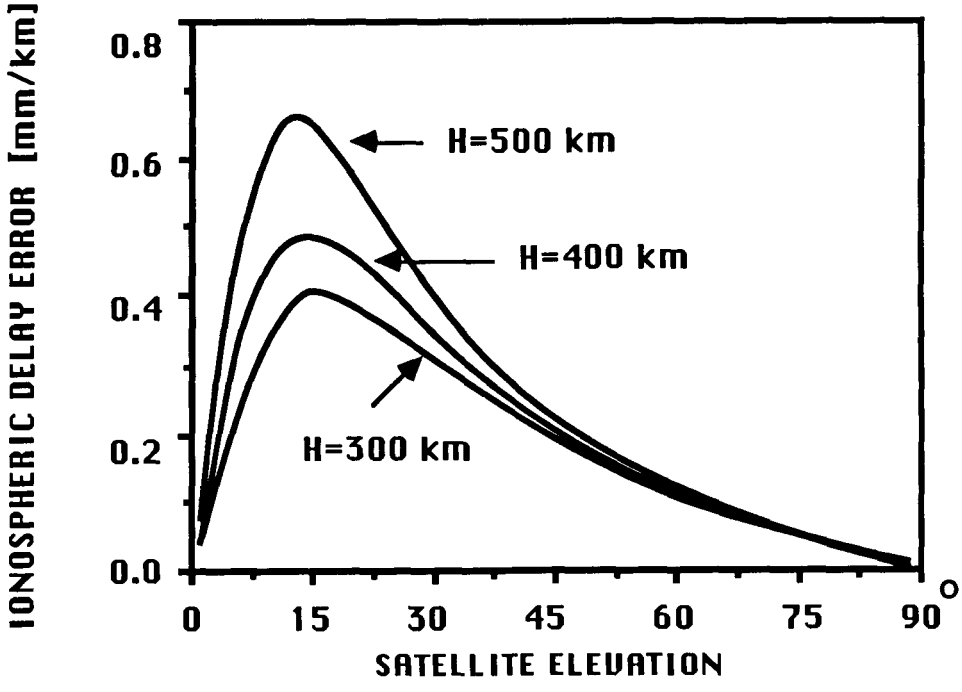

Fig. 2.3 - Error in differential ionospheric path delay
per 1 m of unmodelled (or improperly modelled) vertical ionospheric delay.

At sites where single-frequency GPS receivers might be operating, the vertical ionospheric delay, dV, along the paths of the observed satellites can be determined from the carrier phase observations of one or more dual-frequency receivers located in the vicinity of the GPS survey. This is a practical application where the usefulness of an Active Control System of GPS sites can be indeed very significant. In fact, some experimental analyses of data collected in Canada over short baselines (< 30 km) have already been conducted (ibid) to test this premise. Although positive indications were drawn from the results of these tests, further tests are necessary over long baselines before any definite conclusions are drawn on this issue.

2.1.2.2 Tropospheric delay errors

GPS signals, like VLBI, are experiencing a delay due to the neutral atmosphere (which includes the troposphere and other regions up to some 80 km altitude) which is essentially independent of the frequency of the carrier. The tropospheric contribution of the slant range to a satellite is usually expressed as

$$(\delta\rho)_{trop} = (\delta\rho)_{dry}\ R(E)_{dry} + (\delta\rho)_{wet}\ R(E)_{wet} \qquad (2.15)$$

where $(\delta\rho)_{dry}$ and $(\delta\rho)_{wet}$ are the contributions of the dry and wet troposphere at local zenith and $R(E)_{dry}$ and $R(E)_{wet}$ are corresponding elevation angle mapping functions. The modelling of the effect of the dry troposphere is easier and fairly accurate to handle as demonstrated by SLR and VLBI data analyses. On the other hand, the wet portion of the troposphere is much more difficult to model mainly due to uncertainties in determining the representative humidity along the signal path which is not necessarily well correlated with surface conditions. This error can be reduced by using *water vapour radiometers* (WVR) which, unfortunately, are expensive and may not operate under all weather conditions or offer direct external calibrations of the low elevation angle observations necessary in continental scale GPS experiments due to the low elevation limit of 20° for current instruments.

The best models available to date can typically reduce the combined effect by 92-95% or leaving up to 2-5 cm on the average, depending on the amount of atmospheric information available. In more extreme cases, e.g., at typical humid coastal sites, Ware et al. (1987) have found, for instance, 12-cm rms errors in the vertical component of GPS baselines when surface measurements were used in a tropospheric model similar to

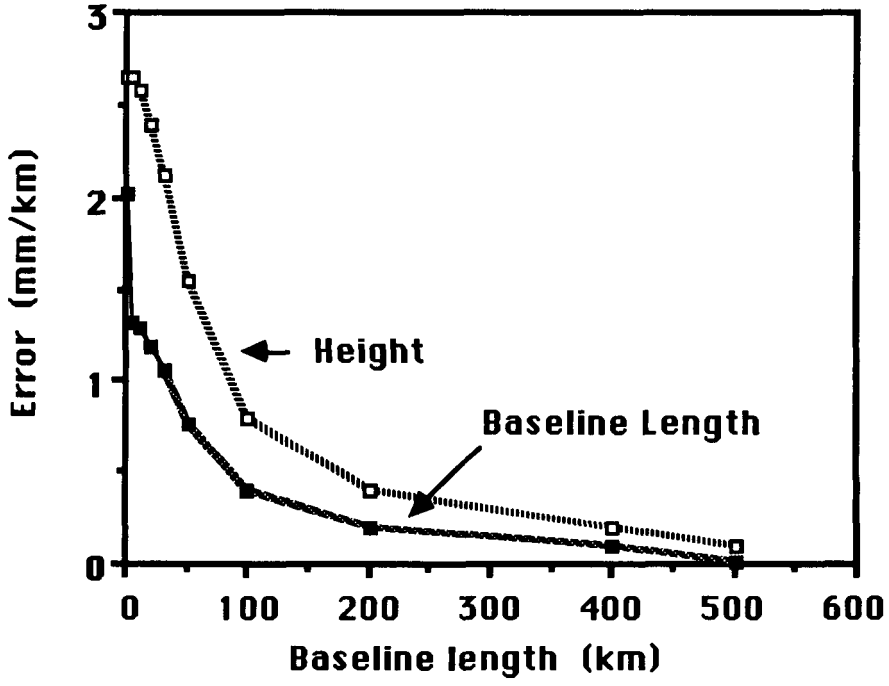

Fig. 2.4 - Effect of wet tropospheric refraction modelling errors on GPS results translated from VLBI experiences based on observation periods of 24 hours (after Kouba, 1987).

(2.15) to estimate path delay. In most GPS data reductions, zenith tropospheric delay parameters are statistically estimated. This technique essentially allows accounting for a spatially and temporally average troposphere for each station and for the time over which the zenith scale parameters are estimated. The wet tropospheric variations around these averages are then the dominant tropospheric errors which, if unaccounted for, in turn map into errors in the baseline length and height differences between observing sites. For relative positioning GPS operations over baselines less than 200-300 km other errors such as orbital and ionospheric errors have often obscured the influence of the wet tropospheric mis-modelling on the determined baselines. Translating the VLBI experience into the GPS situation, Kouba (1987) has suggested a simplified description of the effect of the wet troposphere on the determined baseline lengths and height differences between GPS sites using a model of the form

$$\sigma^2_i = s^2 \ [1-\exp(-B^2/d^2)] \ / \ B^2 \qquad (2.16)$$

where σ_i is in mm/km, the index i stands for baseline length or height difference, and B is the baseline length in km. Equation (2.16) assumes that the effect is governed by a second-order Gauss-Markov process described by a discrete variance s^2 (s = 40 mm for length, s = 80 mm for height differences which correspond to long baseline uncorrelated tropospheric conditions based on VLBI experience) and a correlation distance d usually set at 30 km which (although somewhat arbitrary) reflects the common portion of the wet atmosphere seen from the two ends of the baseline. The underlying property of this process is that the baseline correlation diminishes with increasing distance and approaches unity as the station separation approaches that of a zero baseline. Numerical values from the effect computed from Eq. (2.16) are plotted in Figure 2.4 showing that short baselines are subject to a systematic error of about 1 ppm for baseline length and nearly 3 ppm for height differences. For distances over 50 km the effect decreases proportionally to the inverse of baseline length, vanishing at about 500 km which corresponds to the VLBI results. At distances greater than 500 km the wet tropospheric scale and height effects are assumed to be random under the assumptions of the underlying Gauss-Markov process upon which these values are based. Treuhaft and Lanyi (1985) have shown that if a more detailed model of atmospheric structure is used, water-vapour fluctuation-induced errors are caused mainly by temporal fluctuations due to spatial patterns which are moved over a site by the wind and hence, vary with the observing schedule. In Treuhaft's and Lanyi's educated opinion such errors can be reduced at their optimal level if the length of time for the zenith solutions is less than approximately 4 hours, which coincidentally is typical for the GPS observing sessions. To date, to the author's best knowledge, not enough experiments have been conducted to assess the character of the noise contribution of the WVR instruments vis-à-vis that of the GPS data and to resolve the issue of the usefulness of WVR's for individual GPS experiments. In that light, the practical benefits of using a calculated tropospheric covariance according to (2.16) or similar models for the parameter estimation scheme should be studied further.

2.2 Multipath and Imaging

Multipath effects result from the mutual interference of two or more signals emitted from a satellite that travel along different paths due to surrounding reflective surfaces (Fig. 2.5), thus arriving out of phase in the antenna causing interfering signals which may not be decodable or understandable by the receiver. Antenna imaging is a secondary phenomenon, similar to multipath, caused by nearby conducting objects coupling electrically with the antenna, thus essentially defining an "image" of it with the

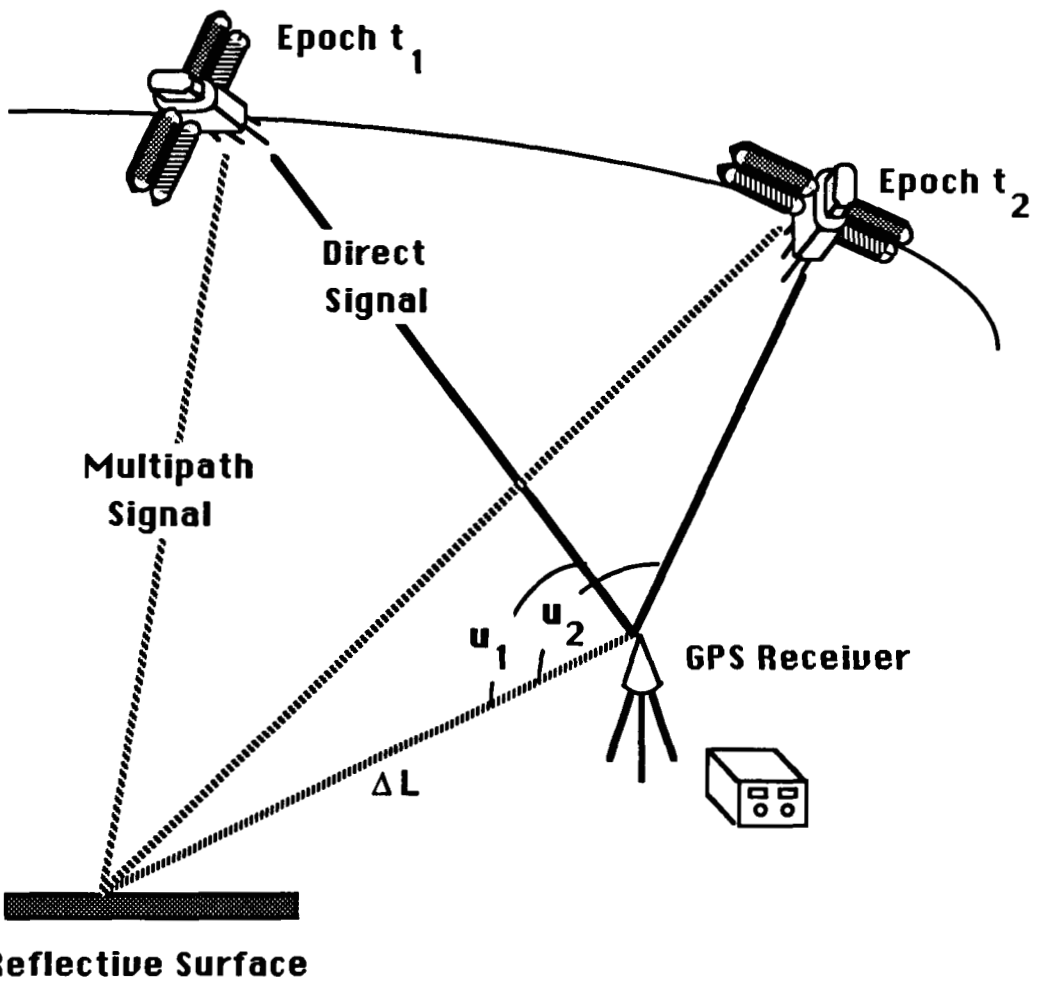

Fig. 2.5 - Geometry of multipath effect on GPS signals.

resulting amplitude and phase characteristics (particularly of the phase center of the antenna) being significantly different from those of the isolated antenna.

Based among other factors on the geometry of the satellites and the antenna characteristics, multipath effects on carrier phase observations exhibit themselves as a very irregular function of time. The phase difference between two "out-of-step" arriving signals is given directly as (Tranquilla, 1987)

$$\Phi = 2\pi \, \Delta L / \lambda \qquad (2.17)$$

in rad/m, where ΔL is the path length difference. Assuming that the antenna characteristics remain unchanged, the interference patterns of direct and reflected multipath waves will exhibit themselves in the data time series as we move through the observation plane. For instance, for a series of carrier phase GPS observations on a single satellite (cf. Fig. 2.5) corresponding to an observation angle sweep of 5° (or a time interval of 5-10 minutes) say, at $u_1=45°$ and $u_2=50°$, the multipath-induced phase change is computed from Eq. (2.17) approximately as $0.064 \times 2\pi \, \Delta L/\lambda$ rad/m which in turn, after setting it equal to one period implies $\Delta L/\lambda=15.5$ or approximately $\Delta L=3$ m at the L_1 frequency; clearly a very significant effect. In practical terms, this corresponds to a single reflective surface located at a distance of 3 m. A reflective surface at a closer distance would produce low-frequency variations, whereas a surface at a greater distance would produce high-frequency variations. This has been confirmed by experimental observations (e.g Bletzacker, 1985; Evans, 1986; Tranquilla et al., 1987) which have shown that multipath effects can induce periodic effects ranging from sub-minute to 2-3 min periods (caused, for example, by diffused forward scattering multipath such as from overhead wires at low elevation angles); 5-10 min periods varying with satellite elevation (caused, for example, by specular reflections from reflective surfaces in the vicinity of the antenna); or up to 25-60 min periods (usually associated, for example, with reflections from fresh or saltwater surfaces).

Although it has been suggested that the only means to minimize these undesirable effects is through antenna beam shaping (e.g., the use of polarized antennas, large ground planes and sheets of absorbing material), it may be possible to reduce the effect of multipath either by averaging over long enough time periods, so that the relative phase of the direct and reflected signals changes by at least one full cycle or alternatively by accommodating spectral analysis techniques during preprocessing to detect and suppress any periodic variations due to multipath found in the data series. Such an approach has been tested in the software developed in the course of this study and seems to work well when dealing with this problem at the stage of cycle-detection in the preprocessing of carrier phase data. More will be said about this below.

2.3 Cycle Slips

As pointed out in the beginning of this section, carrier beat phase measurements lead potentially to the most precise information about the slant ranges to a satellite one can obtain from GPS. The problem of utilizing this potential however is twofold:

(a) Associated with the observations is an ambiguity in the number of whole cycles between the satellite and the receiver at the time of initial signal acquisition (see also Section 4.3 and Fig. 4.2). Although with presently available instrumentation the fractional part of the carrier phase may be measured with an internal precision of 1 to 3 mm, and with an overall accuracy (including the influence of the atmosphere) at the 1- to 4-cm level, it is quite difficult to locate accurately the exact cycle of the carrier whose phase is being measured, to the effect that by an incorrect determination of the integer cycle ambiguity, an equivalent range error of a multiple k of the wavelength of the carrier may result; i.e., $k \times 19$ cm for the L_1 frequency (or $k \times 25$ cm for the L_2 frequency).

The estimation of this indeterminacy is usually carried out in the stage of the adjustment of the available observations by introducing an additional unknown bias to be estimated along with other nuisance parameters. The remaining problem in this approach stems from the difficulty in pinpointing the exact integer value (as the number of cycles can only be an integer) of the so-determined real-valued ambiguities. So far, many different approaches are used to resolve this ambiguity problem (see Langley et al., 1984; Bender and Larden, 1985; Bock et al., 1986), but as yet there is no efficient one for adjusting real and integer numbers simultaneously, not to mention the complications arising from the aliasing of the cycle ambiguities with the effects of multipath, atmospheric delays and orbital errors, particularly for data collected over long baselines. In this latter case, one can only allow these ambiguity parameters to assume real values.

(b) The ambiguity problem is further compounded by the frequent occurrence of cycle slips. This is a commonplace phenomenon observed in GPS data when a satellite signal is obstructed in any way and can no longer be tracked. When signal lock is resumed, the fractional part of the measured carrier beat phase would still be the same as if tracking had been maintained all along. The integer number of cycles, however, exhibits a discontinuity, or cycle slip, which effectively adds a new ambiguity to the data. Such occurrences can be quite frequent, including many instances of loss of lock not recognized by the receiver (e.g., by the quality and tracking mode parameters of its measurements report), such as is shown, for instance, in Fig. 2.6 for the data of a typical observing session with TI-4100 receivers during the March '85 NASA/JPL experiment.

Therefore, a vital question related to the preprocessing of GPS data is how well is one able to determine the limits of reconstruction of the number of whole cycles when phase lock is lost and regained; and hence, how well can one recover from such cycle slips by using various models for the actual change in range due to satellite motion and the phase error introduced by imperfect clocks. This initial assessment of the data is particularly important as it eliminates any damaged data which would otherwise alias the results of the orbit determination and baselines.

2.3.1 Detection and Removal

There several possible approaches to dealing with the cycle slip problem. An often-used approach (Hilla, 1986) is to perform a positioning adjustment in which the station positions are held fixed, and to edit the data by inspecting the corresponding

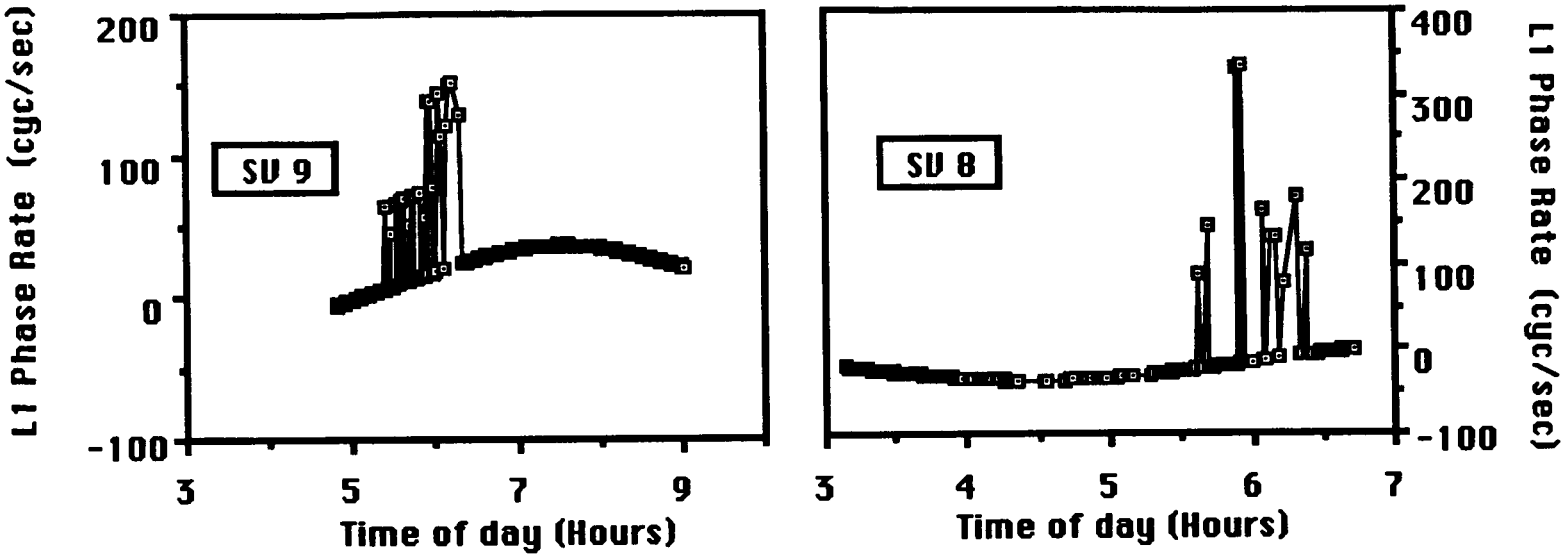

Fig. 2.6 - Cycle slips and outlier observations can be readily seen in the rate of change-of-phase observable

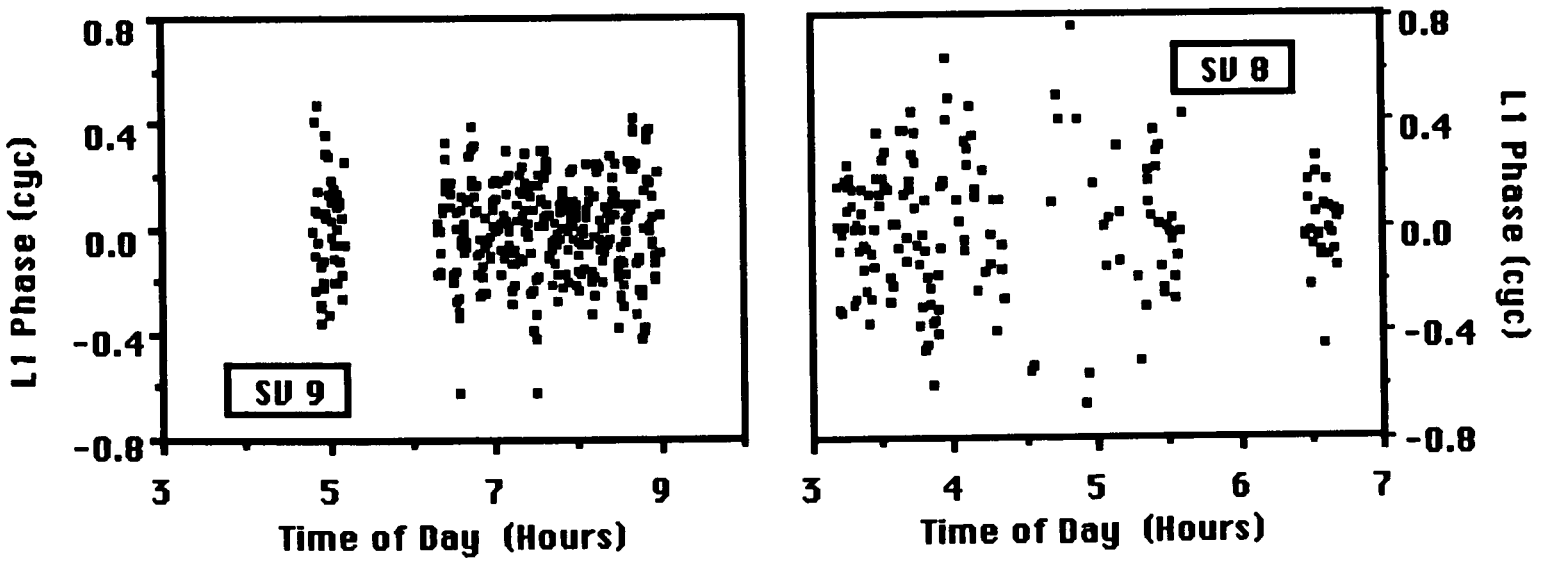

Fig. 2.7 - Cycle slip removal residuals following simple polynomial fits to the phase observable

residual series for abrupt discontinuities. This has been shown to work, but it can only recover large cycle slips (less than a few tens of cycles). A similar approach is to use between-receivers, -satellites, and -time triple-difference observables (which are not affected by receiver or satellite clock errors) to determine station locations and subsequently, once convergence of station coordinates has been achieved, to search through the triple-difference residuals to identify large discontinuities in the double-difference observables which form the specific triple differences (Remondi, 1985). This can be tedious and time consuming. A common approach is to fit polynomials to various types of phase data (e.g. raw phase, phase with orbital motion removed, between-receivers phase, between-receivers and -satellites phase, etc.), and then use these polynomials to project the phase into the future. Using this approach Clynch and Coco (1986) have found, for instance, that at least a third-order polynomial is needed to model raw phase at the 1-3 cycle level for about 100 sec, whereas a second-order polynomial models at the same level about 180 sec of data from which orbit dynamics have been removed. These results varied as expected when orbit-removed data sets for various clock types were examined. Using observations spread over a 20-minute period, double-differenced carrier phases were predicted to within one half a cycle for up to 3 minutes using crystal oscillators, 6.5 minutes using rubidium clocks, and 9 minutes using hydrogen maser clocks.

In our analysis we used a similar approach based on a piecewise polynomial fitting procedure applied on data series of raw phase or between-stations single-difference measurements. Essentially, one may consider a time series of observed carrier phase measurements as composed of constituents that are of interest (e.g., satellite dynamic motion and satellite-to-receiver distance changes), and those noise or nuisance constituents (among them the cycle slips) which obscure the constituents one would like to study. The latter can be either *periodic*, rendering the series *coloured*, or some other type such as *constant shifts* or *trends* rendering the series *non-stationary*, or both. For example, from the discussion in this section it is easily understood that:

- the occurrence of cycle slips observed in GPS data induces a datum shift in the observed time series which remains constant as long as carrier phase lock is maintained. That is, when signal lock is resumed, the fractional part of the measured carrier phase would still be the same as if tracking had been maintained all along. The integer number of cycles, however, exhibits a discontinuity or cycle slip which appears in the data as an abrupt jump in the carrier phase history (cf. Fig. 2.6).

Touching for a moment on the aforementioned other aliasing effects on these cycle slips:

- carrier phase measurements like all GPS measurements being intimately connected to precise timing, are also prone to the deterministic drifts of the satellite and receiver clocks involved in the measuring process. For the satellite rubidium and cesium atomic frequency standards, the latter are usually assumed to be linear or quadratic in time. Similar behaviour is often assumed valid for describing the long-term stability of the high-quality quartz crystal clocks of currently available GPS receivers;

- a similar type of drift, which has also been identified in experimental observations, can be due to temperature variations inside the receiver (ibid);

- multipath and imaging delays, which as already mentioned are usually noted when operating in a highly reflective environment, can also corrupt measured carrier phases at any elevation angle inducing periodic systematic noise in the observed time series.

All these non-stationarity-inducing constituents can be considered as systematic noise in the preprocessing stage, as long as we know their general form; that is, for example, the time at which cycle slips occur and the kind of trend (linear, quadratic, etc.). The main point here is that an ideal preprocessing algorithm attempting to recover cycle slips in the GPS data should take into account the interrelationship of the hidden couplings exerted by these effects. Knowing as much about the constituents of the observed data series of carrier phase measurements, in the preprocessing stage, it is possible to use a piecewise continuous function of degree n

$$F(t) = \sum_{j=1}^{n} a_j \Phi_j(t_m) \equiv \phi^k(t_m) \qquad (2.18)$$

to approximate m carrier phases ϕ^k (or single differences) observed by a station (or two stations respectively if single differences are used) to satellite k where a_j are coefficients to be determined and

$$\Phi_j \equiv [\, \Phi_0, \Phi_1, ..., \Phi_n, \cos \omega_i t, \sin\omega_i t \,] \qquad (2.19)$$

are base functions providing for datum shift biases, linear trends, quadratic (or other type) trends and periodic components, ω_i, of the kind described above.

In practice, if this approach is to be applied solely for cycle detection and recovery purposes, only the coefficients pertaining to the datum shift base functions (Φ_0) are of interest. If there are q cycle slips, present in the data series of a particular session, one ends up dividing the total session in sub-blocks I_q free of cycle slips, so that the size of cycle slips between blocks is determined by a series of a continuous function of the form (2.18) with identical coefficients except for the zero-order term; i.e.,

$$F(t) = \sum_{i=1}^{q} a_{oi} + \sum_{j=1}^{n} a_j \Phi_j(t_m) \qquad (2.20)$$

so that an additional bias term is added to the original function for each sub-block I_q.

Although seemingly straightforward, the foregoing approach tacitly assumes that we know exactly the time at which the cycle slips occur in the data, that is to say, the division of the observation session into cycle slip free sub-blocks is known a priori. Such breaks, however, are often not obvious at the time of observation and can only be detected by a post-mission close examination of the data. In an effort to make this detection as automatic as possible, we have further implemented an algorithm which takes advantage of the short-term characteristics of the observed carrier phases to detect the occurrence of cycle slips. This is based on the fact that the change in GPS range

is exceptionally smooth (because of the extreme distance-3 to 4 earth radii-to the GPS satellites) and hence, highly predictable. Clock and propagation effects do not normally destroy this smoothness, except perhaps in the case of sudden ionospheric disturbances and highly dynamic receiver motion.

Consider a sequence of carrier phase measurements ϕ_j, j=1,2,...,N. Let P_1 be the value at time t of the unique polynomial of degree zero (i.e. a constant) passing through the point (t_1,ϕ_1); so $P_1 = \phi_1(t_1)$. Likewise define P_2, P_3, ..., P_N. Also let P_{12} be the value of the unique polynomial of first degree passing through both (t_1,ϕ_1) and (t_2,ϕ_2). Likewise define P_{23}, P_{34}, ..., $P_{(N-1)N}$. Similarly, one can define the higher order polynomials, up to $P_{123..N}$, which is the value of the unique interpolating polynomial passing through all N points. This process can be schematically visualized as a tableau of various P's with "ancestors" on the left leading to a single "dependent" at the extreme right; e.g., for N=4

$$
\begin{array}{lllll}
t_1: & \phi_1 = P_1 & & & \\
 & & P_{12} & & \\
t_2: & \phi_2 = P_2 & & P_{123} & \\
 & & P_{23} & & P_{1234} \\
t_3: & \phi_3 = P_3 & & P_{234} & \\
 & & P_{34} & & \\
t_4: & \phi_4 = P_4 & & &
\end{array}
\qquad (2.21)
$$

Each unique polynomial P can be constructed by using the recursive relationship between a "dependent" and its two "ancestors"

$$
P_{i(i+1)...(i+j)} = \frac{(t-t_{i+j})\, P_{i(i+1)...(i+j-1)} + (t_i-t)\, P_{(i+1)(i+2)...(i+j)}}{(t_i-t_{i+j})} \qquad (2.22)
$$

which holds because the two ancestors already agree at points $t_{i+1}...\ t_{i+j+1}$. A further improvement on this recurrence relationship is achieved by keeping track of the small differences

$$
C_{j,i} = P_{i...(i+j)} - P_{i...(i+j-1)} \qquad (2.23a)
$$

$$
D_{j,i} = P_{i...(i+j)} - P_{(i+1)...(i+j)} \quad , \quad j=1,2, \ ..., \ N-1. \qquad (2.23b)
$$

These differences satisfy the relation

$$C_{j+1,i} - D_{j+1,i} = C_{j,i+1} - D_{j,i} \qquad (2.24)$$

which is useful in proving that the recursive relationship in (2.22) can be converted to one involving only the small differences **C** and **D**, so that

$$D_{j+1,i} = \frac{(t_{i+j+1}-t)(C_{j,i+1}-D_{j,i})}{(t_i-t_{i+j+1})} \qquad (2.25a)$$

$$C_{j+1,i} = \frac{(t_i-t)\ (C_{j,i+1}-D_{j,i})}{(t_i-t_{i+j+1})} . \qquad (2.25b)$$

In practice, at each level j, after each column in the tableau (2.21) is completed based on N given carrier phase values, if a predicted carrier phase value, $\hat{\phi}(t)$, at epoch t is sought, one can compute such a value by forking up or down the tableau which is equivalent to an interpolating or extrapolating value from the given ϕ_i values plus a set of **C** and **D** corrections which form the "straightmost" route through the tableau so that the partial approximations are centered insofar as possible on the target epoch t. The use of **C** and **D** corrections effectively makes the interpolation or extrapolation one order higher. The algorithm is easily implemented to follow the L_1 and L_2 signals and project the phase into the future. If the difference between the observed and projected values exceeds a pre-defined threshold, the algorithm interprets it as the occurrence of a cycle slip, so that an additional term is added in (2.20). The level of tolerance of these projections is obviously different for various types of clocks and the sampling rates of the data. From tests with TI-4100 data it has been found that 5 consecutive data points at 30-sec intervals were usually needed usually to model the raw phase at the 1-3 cycle level for about 150 sec, which compares well with similar findings by Clynch and Coco (1986), who used algebraic polynomials and data from which orbit dynamics had been removed. This is considered adequate for the purpose of identifying the time of occurrence of suspected cycle slips in the data.

The foregoing process was applied to several data sets prior to the analysis leading to the results to be reported in Section 5. In each case a system of equations (2.20) was solved simultaneously in a least-squares procedure similar to the least-squares spectral analysis described by Wells et al (1985) to estimate the coefficients a_{oi} and a_j. For the base functions Φ_j, Chebychev polynomials were found to perform better than algebraic polynomials simply because the Chebychev approximation is very nearly the same polynomial, which among other polynomials of the same degree, yields the smallest maximum deviation from the true function $\phi(t)$. The degree of the polynomials required for cycle slip detection depends mainly on the length of the data series and the sampling of the data, as well as the smoothness (or noisiness) of the data to be approximated. Experience, for instance, with several data sets from the Spring '85 campaign has demonstrated that a degree 6 or 7 was typically required in order to produce piecewise fits of the raw phase over a data span of 1 hour with an rms error at

the 1-cycle level or better, whereas a degree of 3 or 4 was adequate when the satellite dynamics were removed prior to the fit. These results varied as expected when data sets collected with different clock types were examined. Fig. 2.7 is a typical example of the fits obtained from the single-difference data from the Owens-Mojave baseline on day 90 when both stations were operating with cesium clocks. In this case, no periodic terms were necessarily included in the fitting process to recover the cycle slips present in the original data series shown in Fig. 2.6.

2.4 Clock Errors

In GPS, like VLBI, the tracking data being recorded are based on one-way signal travel time and hence, inherent to the measurements are the errors due to the receiver clock and those due to the clocks of the transmitting satellites. Although the rubidium and cesium frequency standards used onboard the GPS satellites are considered more than adequate for most applications of GPS, the requirements of geodynamic applications demand that biases in both the satellite and receiver clocks be properly modelled or otherwise be accounted for. As explained at the beginning of this section, a common way to eliminate or greatly reduce these biases is through double differencing or equivalently by including separate clock bias parameters for each satellite and receiver clock at each measurement epoch (which is equivalent to a "white noise" process). The differencing techniques are usually applied to a set of "quasi-simultaneous" measurements from different stations and satellites made over such a short time interval that the errors due to the clocks can be considered constant. In reality, however, one cannot disregard the fact that the estimation problem at hand (be it orbit determination or baseline estimation, or both) involves measurements over longer intervals of time, which in turn is forcing us to make some assumptions about the behaviour of the clock errors as a function of time. For example, in order to increase the degrees of freedom in the adjustment of the observations, a bias, a drift, and a drift rate are sometimes used (i.e., a quadratic clock model is assumed). This might be a satisfactory model to use for the satellite clocks and it is easy to handle, but is not generally accurate for describing the behaviour of the receiver clocks (except when a short observing period is involved or a high-precision frequency standard like a hydrogen maser is used). For the most precise applications, it is generally necessary to try to eliminate these errors altogether through direct modelling as opposed to allowing them to be absorbed by other parameters.

One possible way of doing so, is to use a continuous time-state representation from which it is possible to derive (albeit at the cost of some added computational effort) the corresponding discrete time-state representation at the observation times. From experience gained so far with GPS operations, it is generally accepted that the frequency standards used in the currently available GPS geodetic receivers are quite stable; however, their behaviour is not deterministic, but rather undergoing stochastic variations in both the time and the frequency domains. From empirical verification of experimental work (e.g., Heroux, 1986), it has been determined by examining residuals after fitting a quadratic polynomial to GPS interstation clock errors that the residuals from such fits have random walk characteristics which are actual changes in the frequency behaviour of the clocks; hence, such variations can be interpreted as a random walk in frequency (which gets integrated into time), and an independent random walk in time. Stated differently, the variances of these two random walks can always be associated with the general clock characteristics such as long-term frequency stability

and short-term time stability. This makes a continuous model more suitable than any discrete time model since GPS measurements in most receivers are sampled at intervals from 6 sec to 5 min during each observing session, while the random walk behaviour is essentially continuous. The practical implication of this sampling process is that at the observation epochs there is a correlation between the random input to time and the random input to frequency caused by the continuous integration of the random walk in frequency. This is often disregarded for the sake of more simplified models.

The form of the continuous process model underlined by such characteristics can be described as (e.g., Gelb, 1974; Jones and Tryon, 1987)

$$dX(t) = F\,X(t)\,dt + D\,dt + dw(t) \tag{2.26}$$

with

$$X(t) = \begin{vmatrix} x(t) \\ w(t) \end{vmatrix}, \quad F = \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} \tag{2.27}$$

and

$$dW(t) = \begin{vmatrix} \varepsilon\ (t) \\ \eta\ (t) \end{vmatrix}, \quad D = \begin{vmatrix} 0 \\ d \end{vmatrix} \tag{2.28}$$

where x(t) is the clock error state, w(t) denotes the frequency random walk component, d denotes the deterministic frequency drift and $\varepsilon(t)$ and $\eta(t)$ are independent white noise processes with spectral densities σ_ε^2 and σ_η^2 respectively. For a discrete time interval δt, the prediction of the state vector $X(t+\delta t)$ is

$$\begin{aligned} X(t+\delta t) &= \Phi(\delta t)\,X(t) + \int_0^{\delta t} \begin{vmatrix} 1 & \delta t - \tau \\ 0 & 1 \end{vmatrix} d\tau \\ &= \Phi(\delta t)\,X(t) + D(\delta t) \end{aligned} \tag{2.29}$$

where

$$\Phi(\delta t) \equiv I + F\,\delta t = \begin{vmatrix} 1 & \delta t \\ 0 & 1 \end{vmatrix} \tag{2.30}$$

and

$$D(\delta t) = d \begin{vmatrix} \delta t^2/2 \\ \delta t \end{vmatrix} \tag{2.31}$$

with the covariance matrix of the error introduced by the random input to this prediction given as

$$Q(\delta t) = \int_0^{\delta t} \begin{vmatrix} 1 & \delta t - \tau \\ 0 & 1 \end{vmatrix} \begin{vmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_\eta^2 \end{vmatrix} \begin{vmatrix} 1 & 0 \\ \delta t - \tau & 1 \end{vmatrix} d\tau \qquad (2.32a)$$

$$= \begin{vmatrix} \delta t\, \sigma_\varepsilon^2 + \tfrac{1}{3} \delta t^3 \sigma_\eta^2 & \tfrac{1}{2} \delta t^2 \sigma_\eta^2 \\ \tfrac{1}{2} \delta t^2 \sigma_\eta^2 & \delta t\, \sigma_\eta^2 \end{vmatrix} \qquad (2.32b)$$

which illustrates how, although the white noise processes involved are independent, prediction over even a finite time interval introduces a correlation between time and frequency through Q.

In a GPS data reduction scheme there are usually several clocks involved. The foregoing formalism allows the state equation for N clocks to be obtained by extending the two-variable state vector of each clock into a hypervector of 2N elements, and the corresponding transition matrix Φ to a 2N $\times$ 2N matrix consisting of 2×2 blocks (of the form (2.30)) on the main diagonal and zero off-diagonal blocks. Similarly, the covariance matrix Q of the random input error is extended to a 2N $\times$ 2N block-diagonal matrix with the 2×2 blocks on the diagonal being of the form (2.32a or 2.32b) where σ_ε^2 and σ_η^2 generally may be chosen to be different for each clock. Going one step further, during the adjustment of the observations, usually one clock is used as the reference clock and the differences between each clock in the observing session and the reference clock are inferred from the single-difference observations. This differencing produces an observation vector y(t)

$$y(t) = A(t)\ X(t) + v(t) \qquad (2.33)$$

where A(t) is the design matrix of the measurement partials with respect to the clock state vector and is of the form

$$A(t) = \begin{vmatrix} 1 & 0 & -1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & \ldots & 0 & 0 \\ . & . & . & . & . & . & \ldots & . & . \\ 1 & 0 & 0 & 0 & 0 & 0 & \ldots & -1 & 0 \end{vmatrix} \qquad (2.34)$$

and has dimensions (N-1) $\times$ 2N, and v(t) represents a random vector of observational errors assumed to be normally distributed with zero mean and an associated covariance matrix R independent of the input random noise dW and assumed to be diagonal with constant elements equal to the variances of the observational errors (usually known from the electronics of the receiver).

SECTION 3

ESTIMATION OF GPS ORBITAL PARAMETERS

3.1 General Considerations

In general, every observation ℓ to a GPS satellite (be it pseudorange, integrated Doppler or carrier beat phase) is a function of the satellite's position at the observation time; i.e.,

$$\ell = \ell(\mathbf{z}, \mathbf{p}, \mathbf{x}) + \text{noise} \qquad (3.1)$$

where $\mathbf{x}$ is a vector of parameters of interest such as station coordinates, time delays, etc., $\mathbf{p}$ is a vector of parameters which describe the forces acting on the satellite, and $\mathbf{z}$ is the vector of orbital parameters which describe the motion of the satellite around the earth in a unique way. However, the choice and the number of such parameters are by no means unique.

According to Newton's second law, the satellite motion is defined as a particular solution of a system of second-order differential equations:

$$\mathbf{r}'' = \mathcal{F}(t ; \mathbf{r}, \mathbf{r}', \mathbf{p}) \qquad (3.2)$$

where $\mathbf{r} = \mathbf{r}(t)$ is the position vector of the satellite in an inertial Cartesian frame, $\mathbf{r}'$ is the satellite's velocity vector, and $\mathbf{r}''$ is the acceleration vector. The parameters $\mathbf{p}$ describe, for instance, the gravitational attractions of the earth, sun and the moon, solar radiation pressure, etc. Their choice and their number depend mainly on the lengths of the satellite arcs considered and as such are not unique. Equation (3.2) does not define the satellite's orbit uniquely either. Fundamentally, the procedure for determining the orbit of the satellite can be considered as the process of assuming the satellite to be under the influence of a known force field and then using observations such as ℓ to determine which solution to the equations of motion one should "choose". This is because, even if the forces acting on the satellite were known, an infinite number of solutions to the differential equations in (3.2) exist until boundary conditions are imposed-such as values for the six components of the initial position and velocity, six Keplerian elements, or similar elements at some chosen epoch.

Let $\mathbf{z}_o$ be a vector of such a set of initial conditions for the satellite state and $\mathbf{p}_o$ and $\mathbf{x}_o$ be vectors of initial estimates for the parameters $\mathbf{p}$, and $\mathbf{x}$ respectively. Generally, the orbit estimation model consists of two sets of functions $\mathcal{F}_1$ and $\mathcal{F}_2$

$$\mathcal{F}_1(\mathbf{r},\mathbf{p},\mathbf{x}\ ;\ \ell) = 0 \qquad (3.3a)$$

$$\mathcal{F}_2(\mathbf{r},\mathbf{z_o}) = 0 \qquad (3.3b)$$

where $\mathcal{F}_1$ represents the purely geometric relationship between satellite positions and observables, whereas $\mathcal{F}_2$ represents the orbital motion model relating the satellite positions at any epoch and the initial state vector. If the initial estimates used are reasonably "good"; i.e., the components of the force model used to define the equations of motion are sufficient to yield a computed orbit close to the "true" orbit, their corrections would be proportionately small so that it may be possible to obtain them by solving a first-order approximation of the non-linear estimation model (3.1) based on linearized observation equations of the form

$$\delta\ell = (\partial\ell/\partial z)_{z=z_o}\ \delta z + (\partial\ell/\partial p)_{p=p_o}\ \delta p + (\partial\ell/\partial x)_{x=x_o}\ \delta x \qquad (3.4a)$$

$$= D_z\ell\ \delta z + D_p\ell\ \delta p + D_x\ell\ \delta x \qquad (3.4b)$$

with D_z, D_p, D_x denoting the partial derivatives with respect to $\mathbf{x}$, $\mathbf{p}$ and $\mathbf{z}$ respectively. In the sequel the following notation will be used extensively to denote:

D_yx ... the partial derivative of a scalar x (or vector $\mathbf{x}$) with respect to a scalar y (or vector $\mathbf{y}$) ;

d_yx ... the total derivative of a scalar x (or vector $\mathbf{x}$) with respect to a scalar y (or vector $\mathbf{y}$).

Evidently, any variations of the initial state vector errors δz_o would give rise to corresponding variations in the satellite state errors δz related by

$$\delta z = D_{z_o}z\ \delta_{z_o} + D_pz\ \delta p \qquad (3.5)$$

where the components of $D_{z_o}z$ and D_pz are solutions of a special form of the linearized equations of motion known as the variational equations and represent the sensitivities of the orbit to perturbations in the parameters $\mathbf{z_o}$ and $\mathbf{p_o}$. Consequently, Eq. (3.4) takes the form

$$\delta\ell = D_z\ell\ D_{z_o}z\ \delta z_o + D_z\ell\ D_pz\ \delta p + D_p\ell\ \delta p + D_x\ell\ \delta x\ . \qquad (3.4c)$$

As will be discussed in some detail in Section 5, there can be variations in the way these parameters are treated in an adjustment process which, in turn, define different orbit solution strategies or modes of analysis, each with different capabilities to reduce particular systematic errors and improve the orbit and baseline determinations. A completely general parameter estimation process should be able to solve for the best estimates of any combination of the orbital parameters

$$z_{oi1}, z_{oi2}, z_{oi3}, z_{oi4}, z_{oi5}, z_{oi6}, \quad i=1,2,\ldots,n_s$$

$$p_1, p_2, \ldots, p_n,$$

where n_s is the number of satellite arcs being observed, and n is the number of dynamic parameters of the force model. Such complete generality is not always necessary however. If for instance, only relatively short arcs (up to one revolution or 12 hours duration for the case of the GPS satellites) were to be processed, it would allow modelling of $\mathcal{F}$ in equation (3.2) using relatively few parameters, p_n, and would solve or update the set of state vector parameters z_{oi}. Moreover, one would assume that some or all of these parameters (e.g., low-degree potential coefficients, lunar and solar gravitations, etc.) are partially known a priori, a fact that can be easily implemented through appropriate variance/covariance information in the adjustment process. Often, even this set may contain more "free" parameters than required. If the observations originate from a relatively small area and if the number of stations observing simultaneously is small, it will probably not make sense to solve for all z_{oi} elements either; instead, one would perhaps only be able to solve for one element (e.g., responsible for a possible along-track error) per satellite arc with any degree of certainty.

3.2 Equations of Motion

In the satellite equations of motion (3.2) **r** and **r'** can be given in any convenient system. In this section the equations of motion will be derived in a geocentric, inertial, Cartesian coordinate system as defined in detail in Section 4. Neglecting relativistic effects, the Newtonian equations of motion in the inertial frame are given as

$$\mathbf{r}''(t) = \nabla V(\mathbf{r},t) + \mathbf{a}(\mathbf{r},\mathbf{r}',t) \qquad (3.6)$$

where $\nabla \equiv D_{\mathbf{r}}$ is the gradient operator, **a** is the vector of all disturbing accelerations acting on the satellite and ∇V is the gradient of the earth's gravitational potential referenced in the inertial frame; hence, its explicit functional relationship with time so that

$$V(\mathbf{r},t) = V(\mathbf{r}) + V_1(\mathbf{r},t) \qquad (3.7a)$$

where $V(\mathbf{r})$ is the gravitational potential referring to an earth-fixed frame moving relative to the true inertial frame. Expressing $V(\mathbf{r},t)$ in (3.6) by the way of Eq. (3.7a) results in a convenient decomposition of the gravitational potential in a radial part

$$V(\mathbf{r}) = (\mu/r) \quad \rightarrow \quad \nabla V(\mathbf{r}) = -(\mu \mathbf{r}/r^3) \tag{3.7b}$$

where μ is the product of the gravitational constant G and the mass of the earth M, and a small disturbing part $V_1(\mathbf{r},t)$ for which

$$\nabla V_1(\mathbf{r},t) \ll \nabla V(\mathbf{r}) \tag{3.8}$$

so that a regrouping of the terms $\mathbf{a}(\mathbf{r},\mathbf{r}',t)$ and $\nabla V_1(\mathbf{r},t)$ as

$$\mathbf{p}(\mathbf{r},\mathbf{r}',t) = \mathbf{a}(\mathbf{r},\mathbf{r}',t) + \nabla V_1(\mathbf{r},t) \tag{3.9}$$

leads to a homogeneous and a perturbed decomposition part of Eq. (3.6); i.e.,

$$\mathbf{r}''(t) = \nabla V(\mathbf{r}) + \mathbf{p}(\mathbf{r},\mathbf{r}',t) \tag{3.10}$$

which essentially describes the central problem of dynamic satellite geodesy.

3.2.1 The Homogeneous or Two-Body Problem

The solution of (3.10) where $\mathbf{p}(\mathbf{r},\mathbf{r}',t) \equiv 0$ is known as the homogeneous or the two-body problem and leads to the well-known result that the trajectory of the motion is a conic where the particular form (i.e., ellipse, parabola or hyperbola) depends on the initial conditions. Since the GPS satellites are in elliptical or near-circular orbits, the geometric interpretation of the solution is illustrated in Fig. 3.1. The osculating ellipse is the two-body orbit the satellite follows on the orbital plane with the center of mass of the earth as one of its focii. Since the gravitational force field is radially symmetric in this case (i.e. $\mathbf{r}$ and $\mathbf{r}'$ are collinear), the motion is confined to the plane which is tangent to the true orbit at any instant of time t. The stationarity of the orbital plane, the perigee, and the size and shape of the orbit together with the constant period of the motion lead to a realization that the Keplerian motion is fully describable in terms of six parameters, of which one is a function of time. The choice of these parameters is not unique. A particular set of parameters, the familiar Keplerian elements, are particularly common for the study of the satellite perturbations. They are defined as shown in Fig. 3.1. The *right ascension* (Ω) and *inclination* (i) define the orientation of

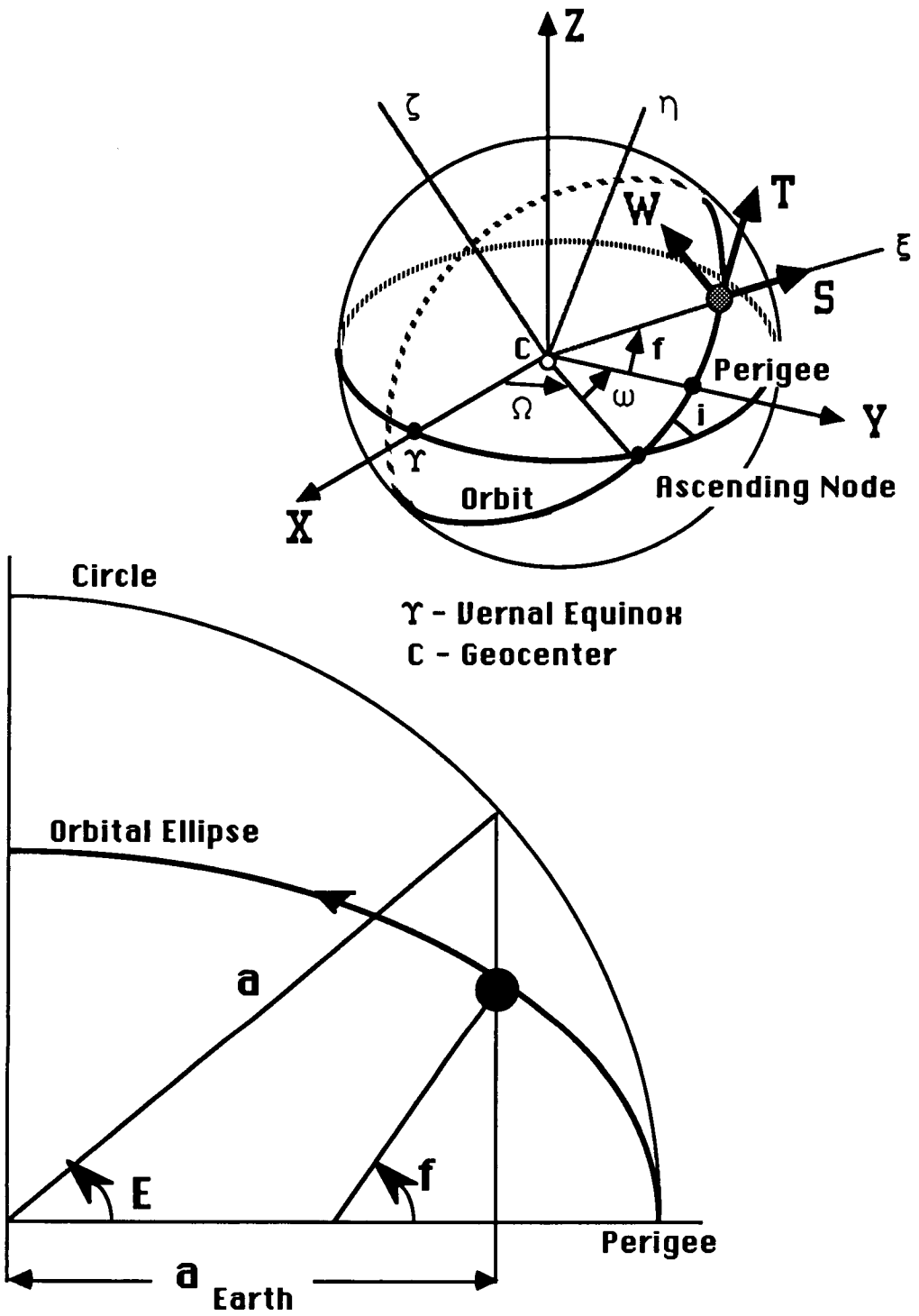

Fig. 3.1 - (a) above, the satellite orbit as projected onto a unit sphere.
(b) below, the geometry of the orbital ellipse.

the orbital plane in space; the *argument of perigee* (ω) defines the location of the perigee in the orbit; the *semi-major axis* (a) and *eccentricity* (e) define the size and shape of the orbit respectively, whereas one of the anomalies is normally used to define the position of the satellite in the orbital ellipse; i.e., the *true anomaly* (f), or the *eccentric anomaly* (E) which relates to the true anomaly by

$$\tan(f/2) = [(1+e)/(1-e)]^{1/2} \tan(E/2) \qquad (3.11)$$

and the *mean anomaly* M, which would be characterized as the angular distance from perigee if the satellite were moving at a constant velocity equal to the *mean motion* n

$$n = (\mu\, a^{-3})^{1/2} = d_t M(t). \qquad (3.12)$$

That is,

$$M = n\,(t-T_P),\ T_P \text{ ... time of perigee passage} \qquad (3.13a)$$

usually given explicitly through Kepler's equation

$$M = E - e \sin E \qquad (3.13b)$$

which can be solved via a series of expansion or iteration methods.

These six elements, which may be considered as the six integration constants of the homogeneous equations of motion, are uniquely related to x,y,z and x',y',z' at any instant of time and can be used instead of them, the advantage of doing so being mainly that they do not vary nearly as much along the orbit (with the exception of the anomalies) as the Cartesian coordinates and velocities do. The transformation from the Keplerian representation to the position and velocity vectors is given as (Kaula, 1966)

$$\mathbf{r} \equiv \mathbf{r}(x,y,z) \equiv \mathbf{r}(a,e,i,\omega,\Omega,M) \qquad (3.14a)$$

$$\equiv R_3(-\Omega)\, R_1(-i)\, R_3(-\omega)\, \mathbf{q} \qquad (3.14b)$$

$$\mathbf{r}' \equiv \mathbf{r}'(x,y,z) \equiv \mathbf{r}'(a,e,i,\omega,\Omega,M) \qquad (3.15a)$$

$$\equiv R_3(-\Omega)\, R_1(-i)\, R_3(-\omega)\, \mathbf{q}' \qquad (3.15b)$$

where $R_i(\theta)$ are the usual rotation matrices for a rotation angle θ , and

$$\mathbf{q} \equiv \mathbf{q}[q_1, q_2, q_3]^T \tag{3.16a}$$

and

$$\mathbf{q}' \equiv \mathbf{q}'[q'_1, q'_2, q'_3]^T \tag{3.16b}$$

with

$$q_1 = a\,(\cos E - e) \tag{3.17a}$$

$$q_2 = a\,(1-e^2)^{1/2}\,\sin E \tag{3.17b}$$

$$q_3 = 0 \tag{3.17c}$$

and

$$q'_1 = -\,n\,a\,\sin E\,/\,(1-e\cos E) \tag{3.18a}$$

$$q'_2 = n\,a\,(1-e^2)^{1/2}\,\cos E\,/\,(1-e\cos E) \tag{3.18b}$$

$$q'_3 = 0. \tag{3.18c}$$

3.2.2 The Inhomogeneous or Perturbed Problem

In order to obtain a solution of the perturbed problem of the equations of motion, it is necessary to recast the non-linear vector second-order differential equations (3.2) to a non-linear system of three scalar differential equations of second order whereby the six integration constants are now considered as functions of time; i.e.,

$$\mathbf{r} \equiv \mathbf{r}(x,y,z) \equiv \mathbf{r}(a(t),e(t),i(t),\omega(t),\Omega(t),M(t)) \tag{3.19a}$$

$$\mathbf{r}' \equiv \mathbf{r}'(x,y,z) \equiv \mathbf{r}'(a(t),e(t),i(t),\omega(t),\Omega(t),M(t)) \tag{3.19b}$$

This is the familiar *method of variation of arbitrary constants* which conveniently bypasses the problem posed in (3.2) that a direct solution in terms of Cartesian coordinates is intractable. Simply put, since the effect of small perturbations acting on the satellite will cause the satellite orbit to deviate slowly from an unperturbed Keplerian orbit, **r** and **r'** will also still correspond uniquely to an orbital ellipse that instantaneously corresponds to the perturbed motion, hence the problem reduces to solving for the time variations of the Keplerian elements instead.

Lagrange obtained this transformation of the equations of motion in an ingenious but rather complex way described in detail in Brouwer and Clemence (1961) and Kaula (1966). The end product is a system of six first-order differential equations

$$\sum_{k=1}^{6} [\, s_j, s_k]\, d_t s_k = D_{s_j} R \quad , \quad j=1,2,...,6 \tag{3.20}$$

where $d_t s_k$ denotes the derivative of the s_k element with respect to time; R is the perturbing force function whose form will depend on the particular type of perturbing source (e.g., earth oblateness, lunar-solar attractions, etc.); s_k or s_j represent the Keplerian elements a,e,i,ω,Ω,M, and $[s_j,s_k]$ are the Lagrange brackets defined as

$$[s_j,s_k] = \sum_{k=1}^{6} [\, D_{s_j} r \; D_{s_k} r' - D_{s_j} r' \; D_{s_k} r \,]. \tag{3.21}$$

A desirable property of the Lagrange brackets that facilitates their evaluation is their invariance with time , so that once evaluated at any convenient point in the orbit, such as, for instance, at the perigee where E=0, equation (3.20) can be inverted to yield the *Lagrange planetary equations* of the time derivatives of the Keplerian elements; i.e.,

$$d_t a = (2/na)\, D_M R \tag{3.22a}$$

$$d_t e = [(1-e^2)/(na^2 e)]\, D_M R - [(1-e^2)^{1/2}/(na^2 e)]\, D_\omega R \tag{3.22b}$$

$$d_t \omega = [-\, \cot i/(na^2(1-e^2)^{1/2})]\, D_i R + [(1-e)^{1/2}/(na^2 e)]\, D_e R \tag{3.22c}$$

$$d_t i = [\cot i/(na^2(1-e^2)^{1/2})]\, D_\omega R - [1/(na^2(1-e)^{1/2})\sin i]\, D_\Omega R \tag{3.22d}$$

$$d_t \Omega = [1/(na^2(1-e)^{1/2}\sin i)]\, D_i R \tag{3.22e}$$

$$d_t M = n-[(1-e^2)/(na^2 e)]\, D_e R - [2/(na)]\, D_a R \tag{3.22f}$$

which are of the general form

$$d_t s_k = s(a,e,i,\omega,\Omega,M;\, \mathbf{p}, \mathbf{z_o}) \tag{3.23}$$

where $\mathbf{z_o}$ is the vector of the initial conditions at t_o

$$z_o = \{a(t_o), e(t_o), i(t_o), \omega(t_o), \Omega(t_o), M(t_o)\}. \qquad (3.24)$$

3.3 Likely Causes Shaping Orbital Errors

3.3.1 Effects of Errors in the Initial Conditions

As with all geodetic satellites the orbit determination problem for the GPS satellites in theory involves a complex process of integration of the equations of motion, but in practice ultimately reduces to an orbit improvement problem using observations which are themselves functions of many satellite positions, station positions and other "nuisance" parameters like satellite and receiver timing delays, atmospheric propagation delays and so on. As already mentioned, the formulation of the differential equations of motion as just outlined leads to an infinite number of solutions even if all disturbing forces are known; that is, until boundary conditions are imposed-such as the values for the initial position and velocity. The GPS observations are then used to determine as accurately as possible these initial conditions. Consequently, errors in the GPS orbits can arise from errors in the forces acting on the satellites and errors in the computed boundary conditions. In this context, the geopotential and a few other forces like the solar radiation are usually considered as the sole source of error in the satellite forces, and tracking station locations as the main source of error in obtaining erroneous boundary conditions. In principle, errors in the tracking station locations can be investigated separately from errors in the satellite forces, although in practice it is frequently difficult to completely separate orbit errors into those directly related to the satellite motion and those related to the station positions. This is because both are largely interrelated: to determine the orbits one needs to know both the geopotential and the positions of the tracking stations; and to obtain the position of other stations with observations, it is necessary to know the orbit and, thus, the forcing field.

To understand the effect of an incorrect initial state on a computed orbit, it is necessary to look back at the perturbed equations of motion in (3.10) written as

$$\mathbf{r}'' + \frac{\mu}{r^3}\mathbf{r} = \mathbf{A} \qquad (3.25)$$

where **A** denotes the disturbing acceleration caused by the perturbing forces acting on the satellite. When a first-order perturbation is applied on (3.25), we obtain

$$\delta\mathbf{r}'' + \frac{\mu}{r^3}\delta\mathbf{r} - \frac{3\mu}{r^4}\delta\mathbf{r}\,\mathbf{r} = \delta\mathbf{A} \qquad (3.26)$$

where $\delta\mathbf{A}$ is considered to represent the perturbing acceleration resulting from either an incorrect set of geopotential coefficients, the initial state, or the force model. The net

effect on the computed orbit becomes quite clear if the resulting satellite position error vector δr is expressed in the orthogonal coordinate system {S,T,W} of the radial, along-track and out-of-plane components shown in Fig. 3.1(a) that moves with the satellite in the inertial frame and has its origin on the satellite as defined by the reference orbit. If $\{ \delta\xi, \delta\eta, \delta\zeta \}$ are the corresponding components of the error vector δr in this frame, then

$$\delta r = \delta\xi\, \hat{\xi} + \delta\eta\, \hat{\eta} + \delta\zeta\, \hat{\zeta} \qquad (3.27)$$

where $\hat{\xi}$, $\hat{\eta}$ and $\hat{\zeta}$ are the unit vectors in the radial, along-track and out-of-plane directions, which notably are time dependent with

$$d_t\hat{\xi} \approx d_t M\, \hat{\xi} \qquad (3.28a)$$

$$d_t\hat{\eta} \approx -d_t M\, \hat{\eta} \qquad (3.28b)$$

$$d_t\hat{\zeta} \approx 0\,. \qquad (3.28c)$$

In this frame

$$\delta A = D_r\delta R\, \hat{\xi} + (1/r)\, D_u\delta R\, \hat{\eta} + (r \sin u)^{-1}\, D_i\delta R\, \hat{\zeta} \qquad (3.29)$$

where δR represents errors in the force function, $u=f+\omega$ is the argument of latitude, and i is the satellite inclination. Substitution of (3.27) into (3.26) yields in view of (3.28)

$$d^2_t\delta\xi - \left[(d_t M)^2 + \frac{2\mu}{r^3} \right] \delta\xi - 2\, d_t M\, d_t\delta\eta - d^2_t M\, \delta\eta$$

$$= D_r\, \delta R \equiv \delta S \qquad (3.30a)$$

$$d^2_t\delta\eta + \left[\frac{\mu}{r^3} - (d_t M)^2 \right] \delta\eta + 2\, d_t M\, d_t\delta\xi - d^2_t M\, \delta\xi$$

$$= (1/r)\, D_u\, \delta R \equiv \delta T \qquad (3.30b)$$

$$d^2_t\delta\zeta + \frac{\mu}{r^3}\delta\zeta = (r \sin u)^{-1} D_i \delta R \equiv \delta W \qquad (3.30c)$$

where δS, δT, and δW are the transformed errors in the acceleration components defined in the {S,T,W} frame. Neglecting the small eccentricities of the GPS satellites, equations (3.30) can be further simplified as

$$d^2_t\delta\xi - 3\, n_o^2\, \delta\xi - 2\, n_o\, d_t\delta\eta = \delta S \qquad (3.31a)$$

$$d^2_t\delta\eta + 2\, n_o\, \delta\xi = \delta T \qquad (3.31b)$$

$$d^2_t\delta\zeta + n_o^2\, \delta\zeta = \delta W \qquad (3.31c)$$

known as Hill's equations (cf. also Colombo, 1985) which interestingly reveal that across-track errors are independent of the radial and along-track ones.

The general solution of these equations can always be considered as being composed of *a particular solution* of the inhomogeneous equations, that is, when the equations (3.31) include terms explicitly dependent on the gravitational potential or on other external forces such as solar radiation, and *a general solution* of the homogeneous part of the equations (i.e. $\delta R \equiv 0$) which reflects how orbital errors $\Delta a(t_o)$, $\Delta e(t_o)$, $\Delta i(t_o)$, $\Delta\omega(t_o)$, $\Delta\Omega(t_o)$, $\Delta M(t_o)$ in the initial state at t_o will propagate as functions of time. The *complete solution* is obviously their sums.

The form of the induced errors in the orbit from an incorrect initial state can be obtained by solving the linearized homogeneous equations of motion with the errors of the initial state as initial conditions, i.e.

$$\delta a(t) = \Delta a(t_o) \qquad (3.32a)$$

$$\delta e(t) = \Delta e(t_o) \qquad (3.32b)$$

$$\delta i(t) = \Delta i(t_o) \qquad (3.32c)$$

$$\delta\omega(t) = \Delta\omega(t_o) \qquad (3.32d)$$

$$\delta\Omega(t) = \Delta\Omega(t_o) \qquad (3.32e)$$

$$\delta M(t) = -(3/2)\ (n_o/a)\ \Delta a(t_o)\ (t-t_o) + \Delta M(t_o) \qquad (3.32f)$$

or equivalently, in the radial, along-track and across-track system

$$\delta\xi(t) = d_t\delta\xi(t_o)\ (\sin n_o t/n_o)$$

$$- [(2/n_o)\ d_t\delta\eta(t_o) + 3\ \delta\xi(t_o)]\ \cos n_o t$$

$$+ [(2/n_o)\ d_t\delta\eta(t_o) + 4\ \delta\xi(t_o)] \qquad (3.33a)$$

$$\delta\eta(t) = (2/n_o)\ d_t\delta\xi(t_o)\ \cos n_o t$$

$$+ [(4/n_o)\ d_t\delta\eta(t_o) + 6\ \delta\xi(t_o)]\ \sin n_o t$$

$$+ [\delta\eta(t_o) - (2/n_o)\ d_t\delta\xi(t_o)]$$

$$- [\ 3\ d_t\delta\eta(t_o) + 6\ n_o\ \delta\xi(t_o)]\ t \qquad (3.33b)$$

$$\delta\zeta(t) = \delta\zeta(t_o)\ \cos n_o t + d_t\delta\zeta(t_o)\ (\sin n_o t/n_o) \qquad (3.33c)$$

and consequently

$$d_t\delta\xi(t) = d_t\delta\xi(t_o)\ \cos n_o t + [2\ d_t\delta\eta(t_o)$$

$$+ 3\ n_o\delta\xi(t_o)]\ \sin n_o t \qquad (3.34a)$$

$$d_t\delta\eta(t) = -\ 2\ d_t\delta\xi(t_o)\ \sin n_o t$$

$$+ [(4\ d_t\delta\eta(t_o) + 6\ n_o\ \delta\xi(t_o)]\ \cos n_o t$$

$$- [\ 3\ d_t\delta\eta(t_o) + 6\ n_o\ \delta\xi(t_o)] \qquad (3.34b)$$

$$d_t\delta\zeta(t) = -\ n_o\ \delta\zeta(t_o)\ \sin n_o t + d_t\delta\zeta(t_o)\ \cos n_o t\ . \qquad (3.34c)$$

As to be expected, if there is an error in the period of the satellite motion, the satellite along-track error grows linearly with time, whereas the satellite altitude/ radius exhibits an error which will not average to zero and hence, will likely translate directly into the determination of the station positions in a systematic way. These terms, although they arise mathematically from a solution of the homogeneous perturbed equations of motion, should not be construed as trivial additions to the perturbed satellite motion from the physical point of view. By contrast they illustrate the very resonant character of the linearized equations of motion. Like the orbits of the altimetry satellites (also usually having eccentricities approaching zero) for which Colombo (1984a) showed that the same expressions also hold, the GPS orbits can behave like an undamped dynamic system whose unforced response produces oscillations in the satellite

motion at the system's natural frequencies: 0 and n_o. When driven by a disturbing force that has such a periodic component at the natural frequencies, the forced response of the system will include similar oscillatory terms which grow larger as their frequencies approach 0 or n_o (cf. Fig. 3.2).

3.3.2 Effect of Unmodelled or Improperly Modelled Forces

The particular solution of (3.31) reflects the contribution of the forces acting on the satellite. Although $\delta\mathbf{A}$ may well be thought of as being due to errors in the geopotential resulting from errors δC_{nm} and δS_{nm} in the harmonic coefficients, these are not expected to be significant for (the very well known to date) low degree and order terms which are usually required in the case of the GPS satellites for modelling the earth's gravitational force during the orbit integration. Rather it is likely that $\delta\mathbf{A}$ would be due mainly to unmodelled or improperly modelled forces, particularly of the non-gravitational nature like the solar radiation pressure. In his more recent work Colombo (1985) related his earlier developments on this aspect from the altimetry and the satellite-to-satellite tracking satellites to the GPS case suggesting that the same type of forced response may occur under the influence of non-gravitational forces, such as:

- the solar radiation effects whereby the solar radiation pressure resulting from the electromagnetic interaction between radiation from the sun and the surface of the spacecraft, may repeat itself at each revolution if the orientation of the spacecraft with respect to the sun is maintained (which happens anyway with the GPS satellites). Any departure from perfect periodicity will be due to the imperfections of the force model, the slow changes in the relative geometry of the earth, the sun, and the satellite, which would reasonably produce some resonant orbit error close to the fundamental frequency n_o;

- the variable part of the earth's gravitational field arising from tidal and other deformations of the solid earth and the oceans. This likewise depends on the relative geometric arrangement of the spacecraft, the sun or the moon and the earth, and so may cause resonance effects for similar reasons; and

- the constant force along the spacecraft's y axis caused by the lack of perfect alignment of the solar panels (Fliegel et al, 1985) known as the y-bias effect. This is causing periodic changes (given the way the satellite orientation is maintained in space) almost at a rate of once per revolution which is also likely to excite resonance.

If errors in the initial state and any unmodelled (or incorrectly modelled) forces are large enough, their resonant effects will have a simple mathematical approximation expressing the complete solution of (3.31) in the form (Colombo, 1984b; 1985)

$$\Delta u_i(t) = \sum_{k=0}^{K} A_{k,i}\, t^k \cos n_o t + \sum_{k=0}^{K} B_{k,i}\, t^k \sin n_o t + \sum_{j=0}^{J} C_{j,i}\, t^j \qquad (3.35)$$

where i=1,2,3 and $\Delta u_i(t) \equiv \{\Delta\xi, \Delta\eta, \Delta\zeta\}$. In fact, Colombo (ibid) has verified with simulations that with this parameterization, some terms will effectively be very close to zero, and others will be related in a simple way, thus leaving at the most, 18 parameters or about 6 per component defining

$$\Delta\xi(t) = A_\xi \cos n_o t + B_\xi \sin n_o t + C_\xi\, t \cos n_o t + D_\xi\, t \sin n_o t + E_\xi \qquad (3.36a)$$

$$\Delta\eta(t) = A_\eta \cos n_o t + B_\eta \sin n_o t + C_\eta\, t \cos n_o t + D_\eta\, t \sin n_o t + E_\eta + F_\eta\, t + G_\eta\, t^2 \qquad (3.36b)$$

$$\Delta\zeta(t) = A_\zeta \cos n_o t + B_\zeta \sin n_o t + C_\zeta\, t \cos n_o t + D_\zeta\, t \sin n_o t + E_\zeta + F_\zeta\, t \qquad (3.36c)$$

from which similar expressions for the velocity components can be easily derived by a further simple differentiation. Numerical results with actual data using this formulation in the present work will be reported in Section 5.

Having outlined the adverse effects that may arise on the GPS orbits by the likely errors in the initial state and the force models, it is in order that the mechanisms which shape them now be considered.

3.4 Perturbations to GPS Orbits

The GPS satellites are designed to provide precise three-dimensional position, velocity and time information anywhere in the vicinity of the earth. Perturbations due to:

- the gravitational attraction of the earth;
- the gravitational perturbations of the moon, the sun and the planets (i.e. the third-body effects);
- solar radiation pressure;
- atmospheric drag;
- magnetic forces; and
- the variable part of the earth's gravitational potential arising from tidal and other deformations of the solid earth and the oceans,

act to drive the satellites away from their operational orbits. Because of the high demand for accuracy, the Ground Control takes care that the satellite subtracks are maintained over the lifetimes of the satellites. This orbit maintenance is taking place in

the form of small satellite manoeuvres to periodically correct the orbit and maintain the repetitive nature of the satellite ground tracks.

The major perturbations on the GPS satellites act on the spacecraft in a variety of ways and timescales, not all of which are important or relevant to the determination of the orbits usually used for the determination of geodetic baselines; these are usually of up to 1 week in duration. Typically, the disturbing forces just outlined may be considered according to the characteristic amount of time over which they act; i.e. generally,

- **secular perturbations** which are constant over time (e.g. gravitational perturbations due to the earth's oblateness);

- **short-period perturbations** superimposed on a precessing Keplerian ellipse which act on periods of a few hours; i.e., the orbital period of the satellites (e.g. gravitational perturbations for the low-degree harmonics of the geopotential);

- **long-period perturbations** which act over periods of a few days to several weeks (e.g. the third-body effects); and

- **resonant perturbations** arising from the repetitive ground-track condition as a result of particular combinations of the tesseral and sectorial harmonics with the given secular rates in ω, Ω, and M; and the forced response effects induced under the repetitive influences of such causes as those outlined in Section 3.3.2.

As it turns out, the orbital elements of the GPS satellites undergoing variations of interest are ω, Ω, M and a. Of these, the three angular ones undergo secular, long-period and resonant perturbations. The semi-major axis, a, does not vary secularly or in long periodic terms, but is influenced almost solely by resonant perturbations which change its value by a few meters per day at the most, but in turn, is the primary cause of a nodal longitude acceleration analogous to the one experienced by geosynchronous satellites.

3.4.1 Disturbing Function of the Gravitational Potential

3.4.1.1 Linear perturbations

The integration of the Lagrange planetary equations as given in (3.22) requires that the partial derivatives of the disturbing function R be evaluated with respect to the Keplerian elements. The development of the disturbing function due to the geopotential begins with the standard representation of the earth's gravitational potential in spherical harmonics C_{nm}, S_{nm} (up to a maximum degree N)

$$V = V_{nm}(r,\phi',\lambda)$$

$$= (GM/r)\,\{1 + \sum_{n=2}^{N} \sum_{m=0}^{n} (a_E/r)^n\, P_{nm}(\cos\phi') \left[C_{nm} \cos m\lambda + S_{nm} \sin m\lambda\right]\} \qquad (3.37)$$

in terms of the the mean earth radius a_E, the geocentric radial distance r, the geocentric latitude ϕ' and longitude λ of the satellite and the associated Legendre functions P_{nm} (Heiskanen and Moritz, 1967). For the treatment that follows, equation (3.37) has to be expressed in terms of the Keplerian elements and θ, the Greenwich sidereal angle that accounts for the earth's rotation, in the form (Kaula, 1966)

$$R(a,e,i,\omega,\Omega,M,t) = \sum_{n} \sum_{m} V_{nm}(a,e,i,\omega,\Omega,M,t,\theta) \qquad (3.38)$$

where

$$V_{nm}(a,e,i,\omega,\Omega,M,t,\theta) = (GM/a^{n+1}) \sum_{p=2}^{n} F_{nmp}(i) \sum_{q=0}^{\infty} G_{npq}(e) \times S_{nmpq}(\omega,\Omega,M,\theta(t)) \qquad (3.39)$$

with

$$S_{nmpq} = \begin{cases} C_{nm}, & \text{n-m even} \\ -S_{nm}, & \text{n-m odd} \end{cases} \cos[(n-2p)\omega + (n-2p+q)M + m(\Omega - \theta(t))]$$

$$+ \begin{cases} S_{nm}, & \text{n-m even} \\ C_{nm}, & \text{n-m odd} \end{cases} \sin[(n-2p)\omega + (n-2p+q)M + m(\Omega - \theta(t))] \qquad (3.40)$$

The functions $F_{nmp}(i)$ and $G_{npq}(e)$ are known as the inclination and eccentricity functions respectively, which involve the evaluation of binomial coefficients and factorials. Ordinarily these functions (and their derivatives) can be very troublesome and computationally expensive to evaluate especially for high degrees, n, of the harmonic expansion of the potential. For the near-circular GPS orbits two important properties of $G_{npq}(e)$ somewhat alleviate this problem: that they are approximately proportional to $e^{|q|}$, so that they converge rapidly for small eccentricities and with increasing q; and that

$$\lim_{e \to 0} G_{npq}(e) = \begin{cases} 1 & \text{if } q = 0 \\ 0 & \text{otherwise} \end{cases} \qquad (3.41)$$

This is a tremendous simplification since the inner summation of (3.39) is restricted to a single term; namely, q=0. Furthermore for the GPS satellites, because they are in high stable orbits at about 20,000 km altitude they are much less affected by the short wavelength features of the geopotential. Thus, only a subset of the spherical harmonic coefficients (usually up to degree and order 8) is used in the integration process which simplifies the computation of the eccentricity and inclination functions with any of the many analytical expressions found, for example, in Allan (1965), Kaula (1966), the recurrence relations in Kostelecky (1985) or the Fast Fourier Transform relations proposed by Goad (1987). Here the notation in the disturbing function in (3.38) by the way of (3.39) and (3.40) has been chosen to correspond closely with that of Kaula (1966). Thus $G_{npq}(e)$ is identical to the expressions given in Kaula, whereas the $F_{nmp}(i)$ can be evaluated from the expression given by Allan (1965) as

$$F_{nmp}(i) = \frac{(n+m)!}{2^n p!\,(n-p)!}\, i^{n-m} \sum_k (-1)^k \binom{2n-2p}{k} \binom{2p}{n-m-k} c^{3n-m-2p-2k}\, s^{m-n+2p+2k} \qquad (3.42)$$

($c \equiv \cos \tfrac{1}{2} i$ and $s \equiv \sin \tfrac{1}{2} i$) which is somewhat more simple.

The representation in (3.39) of the potential admits a rather elegant form for solving the Lagrange planetary equations when the angular rates of a, e and i are constant, while ω, Ω and M are linear functions of time. Careful inspection of (3.38) indicates that the single largest effect in R by the gravitational field would be produced by the earth's oblateness; i.e., the second zonal harmonic term C_{20}, causing most of the secular variations in the mean Keplerian orbit. Consequently, by introducing

$$R = (\mu/a)\ C_{20}\ (a_E/a)\ F_{201}(i)\ G_{210}(i) \qquad (3.43)$$

for the disturbing function into the Lagrange planetary equations, the secular variations of the Keplerian elements would be written as

$$d_t a = 0 \qquad (3.44a)$$

$$d_t e = 0 \tag{3.44b}$$

$$d_t i = 0 \tag{3.44c}$$

$$d_t \omega = \tfrac{3}{4}\, n\, C_{20}\, [(a_E/a)^2 / (1-e^2)^2]\, (1-5\cos^2 i) \tag{3.44d}$$

$$d_t \Omega = 3n\, C_{20}\, [(a_E/a)^2 / 2(1-e^2)^2] \cos i \tag{3.44e}$$

$$d_t M = n\, \{1 - 3C_{20}\, [(a_E/a)^2 / 4(1-e^2)^{3/2}]\, (3\cos^2 i - 1)\}. \tag{3.44f}$$

To first order the change in $s_k \equiv \{a,e,i,\omega,\Omega,M\}$ over a complete orbit is

$$\Delta_1 s_k = \oint d_t s_k \, dt \ , \tag{3.45}$$

with the integration taking place along the reference orbit where a, e and i are assumed constant during the time of integration and the radial distance of the satellite is the same as for an unperturbed orbit, and

$$\omega = \omega_o + d_t\omega\ (t-t_o) \tag{3.46a}$$

$$\Omega = \Omega_o + d_t\Omega\ (t-t_o) \tag{3.46b}$$

$$M = M_o + d_t M\ (t-t_o) \tag{3.46c}$$

$$\theta = (\Omega_o - \theta_o) + (d_t\Omega - \dot{\omega}_E)\ (t-t_o) \tag{3.46d}$$

where t_o is the time of the initial conditions and $\dot{\omega}_E$ is the rotation rate of the earth. The final result is:

$$\Delta_1 a_{nmpq} = \frac{2\sqrt{(\mu/a)}(a_E/a)^n\, F_{nmp}(i)\, G_{npq}(e)(n-2p+q)S_{nmpq}(\Theta)}{[(n-2p)d_t\omega + (n-2p+q)d_t M + m(d_t\Omega - \dot{\omega}_E)]} \tag{3.47a}$$

$$\Delta_1 e_{nmpq} = \frac{\sqrt{(\mu/a^3)}\,(a_E/a)^n F_{nmp}(i)\, G_{npq}(e) \times \times (1-e^2)^{1/2}\,[(1-e^2)^{1/2}(n-2p+q)-(n-2p)]S_{nmpq}(\Theta)}{e[(n-2p)d_t\omega + (n-2p+q)d_tM + m(d_t\Omega-\omega_E)]} \quad (3.47b)$$

$$\Delta_1 i_{nmpq} = \frac{\sqrt{(\mu/a^3)}(a_E/a)^n F_{nmp}(i) \times \times G_{npq}(e)[(n-2p)\cos i - m]S_{nmpq}(\Theta)}{(1-e^2)^{1/2}\sin i[(n-2p)d_t\omega + (n-2p+q)d_tM + m(d_t\Omega-\omega_E)]} \quad (3.47c)$$

$$\Delta_1 \omega_{nmpq} = \frac{\sqrt{(\mu/a^3)}\,(a_E/a)^n [(1-e^2)^{1/2}\, e^{-1}F_{nmp}(i)\, d_e G_{npq}(e) - -\cot i\,(1-e^2)^{-1/2}\, d_i F_{nmp}(i)\, G_{npq}(e)]\, S_{nmpq}(\Theta)}{[(n-2p)d_t\omega + (n-2p+q)d_tM + m(d_t\Omega-\omega_E)]} \quad (3.47d)$$

$$\Delta_1 \Omega_{nmpq} = \frac{\sqrt{(\mu/a^3)}(a_E/a)^n\, d_i F_{nmp}(i)\, G_{npq}(e) \int_{\Theta} S_{nmpq}(\Theta)}{(1-e^2)^{1/2}\sin i\,[(n-2p)d_t\omega + (n-2p+q)d_tM + m(d_t\Omega-\omega_E)]} \quad (3.47e)$$

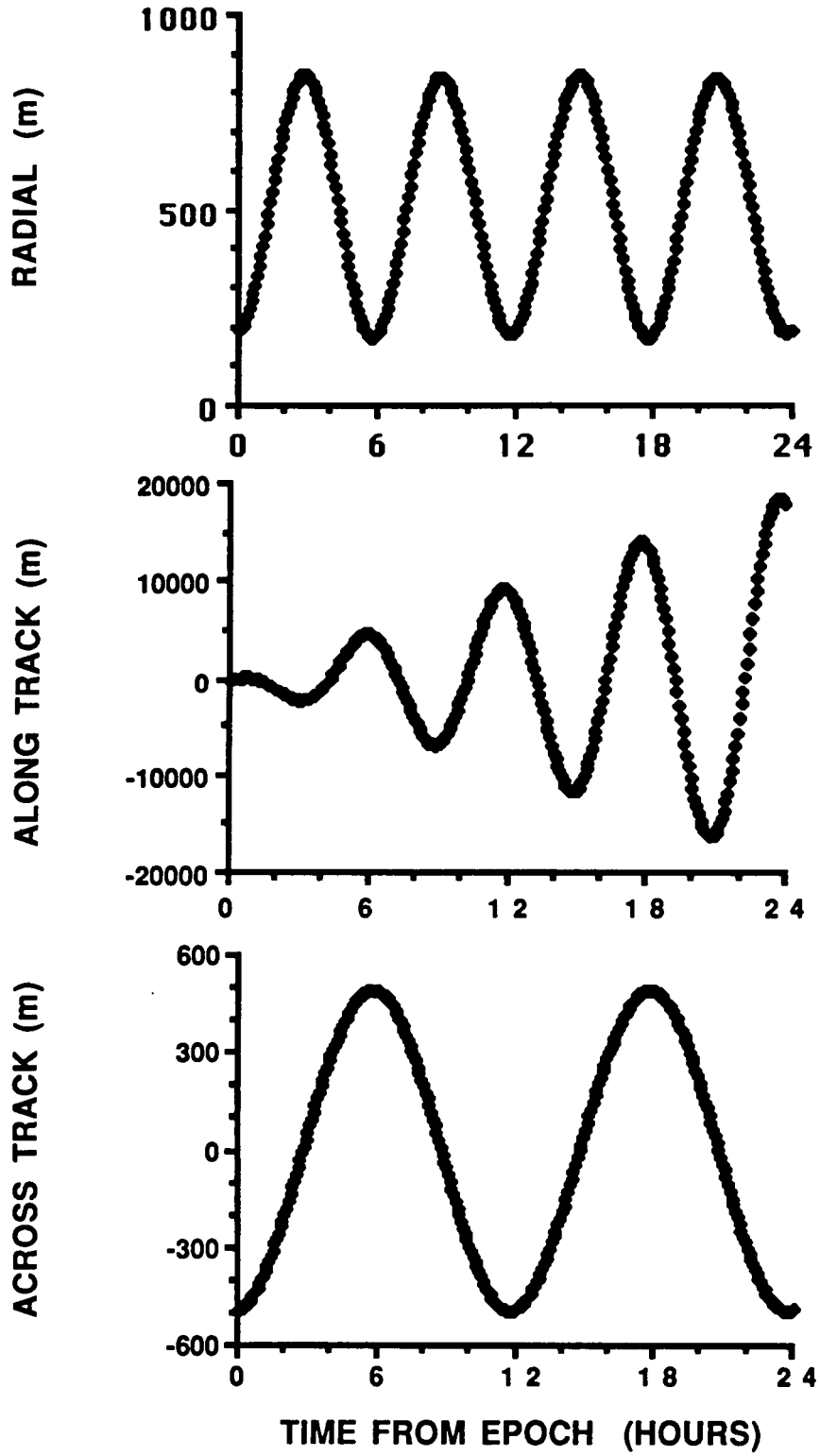

Fig. 3.2 - Perturbation effects on the GPS satellite orbits due to the second zonal harmonic of the geopotential.

$$\Delta_1 M_{nmpq} = \frac{\sqrt{(\mu/a^3)}\ (a_E/a)^n\ F_{nmp}(i)\ [-(1-e^2)\ e^{-1}\ d_e G_{npq}(e)\ -\ -(2n+1)\ G_{npq}(e)]\ \int_{\Theta} S_{nmpq}(\Theta)}{[(n-2p)d_t\omega\ +\ (n-2p+q)d_t M\ +\ m(d_t\Omega - \omega_E)]} \qquad (3.47f)$$

where $\Theta \equiv [(n-2p)\omega + (n-2p+q)M + m(\Omega-\omega_E)]$, and $d_i F_{nmp}(i)$ and $d_e G_{npq}(e)$ are the derivatives of the eccentricity and inclination functions with regard to i and e respectively.

The typical perturbations (3.47) are short-period effects impressed upon a precessing Keplerian orbit. The combined secular and periodic perturbations are quite significant even for short orbital arcs. These vary from about 800 m in i to about 5000 m in Ω. When transformed into the {S,T,W} coordinate system attached to the satellite, which is closer to giving a direct measure of the satellite position error, these effects are as shown in Fig. 3.2 for a 2-day arc. The perturbations due to the higher harmonics of the gravitational potential are found to be quite significant, but are attenuated quite rapidly because of the high altitude of the GPS orbits. Fig. 3.3 shows, again resolved into radial, along- and across-track components, the growth of the effect of the higher-order terms up to degree and order 4 over the 2-day arc. The resulting combined satellite errors exhibit a similar time dependence to the errors described by Eq. (3.34). This illustrates that over orbital arcs it is indeed possible to "soak up" most of the satellite position error due to this geopotential error by appropriate adjustment of the satellite's initial orbit parameters.

3.4.1.2 Higher order (non-linear) perturbations

From the equations already derived for the mean rates of change over one revolution, one can obtain the effect of the first- and second-order gravitational perturbations by integrating $\Delta_1 s_k$ with respect to the given element over its cycle period. Approximating the second-order solution by a Taylor's series expansion

$$d_t s_k = d_t s_k(s_o) + D_{s_i} s_k\ \Delta_1 s_i + \ldots$$

$$= d_t(\Delta_1 s_k) + d_t(\Delta_2 s_k) \qquad (3.48)$$

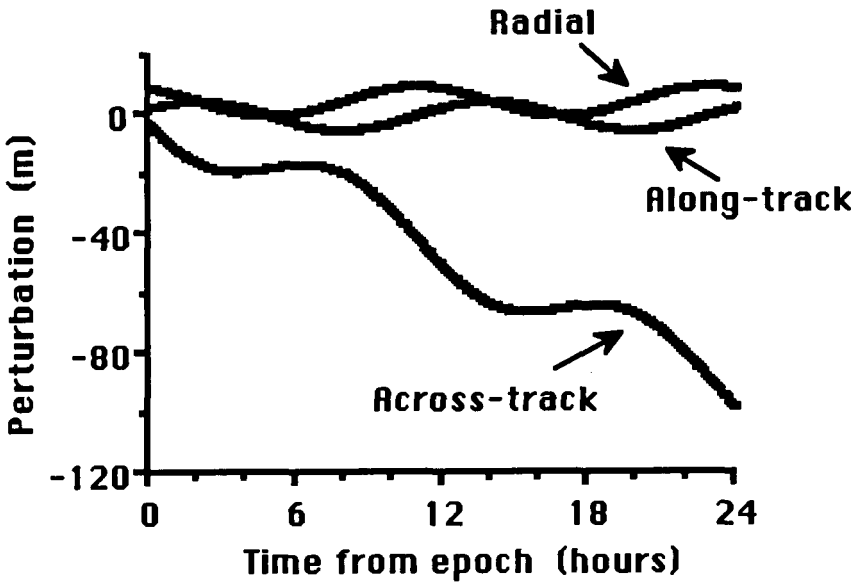

Fig. 3.3 - Perturbation effects on GPS orbits due to higher degree and order terms of the geopotential excluding C_{20}

in which, upon subtracting out the first approximation $d_t(\Delta_1 s_k)$, yields

$$d_t(\Delta_2 s_k) = D_{s_i s_k}\ \Delta_1 s_i \tag{3.49}$$

from which the second-order perturbations are computed as

$$\Delta_2 s_k = \oint_t [\sum_i\ D_{s_i s_k}\ \Delta_1 s_i]\ dt\,. \tag{3.50}$$

Normally, finding a second-order solution by evaluating (3.50) would be computationally inconvenient since the evaluation of $\Delta_2 s_k$ would require a considerable amount of differentiations and multiplications. Fortunately, again because of the near-circular GPS orbits, matters are greatly simplified. This is because, as it turns out from a careful inspection of the terms involved in (3.50), only the changes in the mean anomaly (arising from the perturbations in the semi-major axis), the node and the perigee are significant-in addition to the movement of the earth; furthermore, the first-order perturbations which should be included in (3.50) require the effects of only the terms R_{2010} of the disturbing potential for each $\Delta_1 s_j$ which entails considerable simplification.

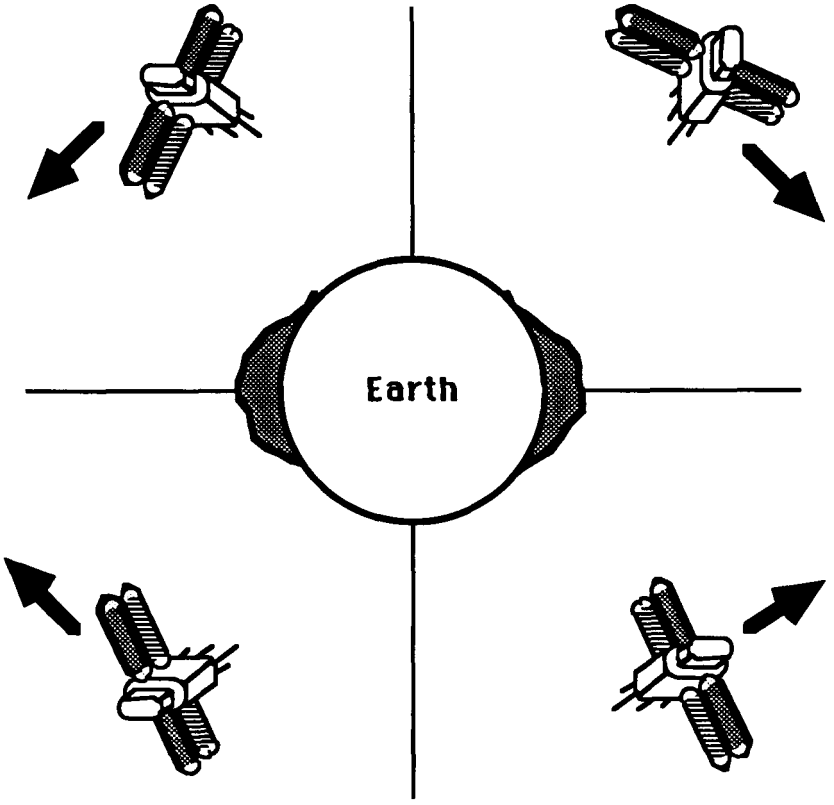

Fig. 3.4 - Schematic representation of the earth's ellipticity causing resonant effects on the GPS orbits.

3.4.2 Resonant Perturbations

The possible existence of resonance is well known in geodetic orbit theory. Resonance occurs for orbits whose periods have some simple ratio to the length of a day. For such orbits the satellite is influenced by the same disturbing gravitational feature at repeated intervals and resonance is thereby induced in the satellite motion. The GPS satellites experience a type of resonance analogous to the type experienced by geosynchronous satellites arising from the tesseral and sectorial harmonics, particularly C_{22} and S_{22}. This is illustrated in Fig. 3.4. In a reference system rotating with the earth, it is clear from symmetry that when the satellite is off either principal axis of the earth's equatorial ellipse, the bulge nearer the satellite has a stronger attraction on the satellite due to its shorter distance from it. If, for instance, the bulge is "ahead" of the satellite, it tends to exert a pull along its orbit thus transferring angular momentum to the satellite which, in turn, should result in an increase in the period and potential energy, and a decrease in the mean linear and angular velocities. Because of the conservation of angular momentum, the increase in the latter results in an increase in the semi-major axis of the orbit. A similar type of resonance (but one harder to visualize in such simple terms) exists for the GPS satellites arising from

particular combinations of tesseral and sectorial harmonics. Because of the high altitude of the orbits the GPS satellites tend to have low angular velocities, the satellites slowly drift in time by some 4 minutes a day which appears as an eastward shift of the satellite subtracks by approximately 1^{o}. Because the commensurability of the GPS satellite orbital periods with the earth's rotation rate is not exact, the longitude of the ascending node has a secular rate upon which resonant effects are impressed.

Generally, the corresponding variational equation for this nodal regression can be derived from the equation of the earth-referred longitude of the node crossing given generally for any satellite by

$$\lambda_s = (1/s_o)(\omega + M) + \Omega - \theta \qquad (3.51)$$

where s_o is the ratio of two integers in the relationship $s = s_o + \Delta s$ where s is the node-to-node satellite revolutions per earth rotation and $\Delta s<1$. For the GPS satellites obviously $s_o = 2$, and $\Delta s \approx 0.012$. Note λ_s is given as the sum of angles lying in two different planes and should not be confused with the conventional geodetic longitude measured along the equator. From (3.51)

$$\dot{\lambda}_s = \tfrac{1}{2}[(\dot{\omega} + \dot{M}) + \dot{\Omega} - \omega_E] \qquad (3.52)$$

or

$$d_t\lambda_s = \tfrac{1}{2}[(d_t\omega + d_tM)] + \dot{d_t\Omega} - \omega_E$$

$$= (n/2)\ [1-(2/n^2a)]D_aR - \tfrac{1}{2}[(1-e^2)/(n^2a\ e)]\ D_eR +$$

$$+ \{(1-\tfrac{1}{2}\cos i)/[n^2a(1-e^2)^{1/2}]\}\ D_iR - \omega_E \qquad (3.53)$$

in view of (3.22).

The form of the disturbing function for the resonant effects can be derived from equation (3.38) by considering the limits on the summations to be restricted only to terms in which

$$\tfrac{1}{2}[(d_t\omega + d_tM)] + \dot{d_t\Omega} - \omega_E \approx 0\ ; \qquad (3.54)$$

i.e., restricting the indices n, m, p and q to those satisfing the conditions:

(i) $n - 2p = m/2 \Rightarrow [n-(m/2)]$ even ,

$$p = (2n - m) / 4 \qquad (3.55a)$$

(ii) $$q = 0 \qquad (3.55b)$$

due to the small eccentricity of the GPS orbits, so that such terms as V_{3210}, V_{4410}, V_{5220} etc. will give rise to terms with an integral multiple of $[\frac{1}{2}(\dot{\omega} + \dot{M}) + \dot{\Omega} - \omega_E]$. Because of the presence of the $(a_E/r)^n$ term in the disturbing function (3.37), the dominant effect will rise from the harmonics for which n is small. A further simplification is possible by making the observation that

$$\lambda_{nm} = (1/m) \arctan (S_{nm}/C_{nm}) \qquad (3.56)$$

so that the disturbing function for the resonant perturbations becomes

$$R = R_{nm} = (\mu/a) \sum_{n,m} (a_E/a)^n \sqrt{(C_{nm} + S_{nm})} \times$$

$$\times F_{nm,n-(m/2)}(e) \cos[m(\omega + M + \Omega - \theta - \lambda_{nm})]. \qquad (3.57)$$

Because of the presence of the mean anomaly M in the disturbing function, there will be a major perturbation of the semi-major axis from the first of the equations (3.22), i.e.

$$d_t a = - (2/na) \sum_{n,m} (a_e/a)^n \sqrt{(C_{nm} + S_{nm})} \times$$

$$\times F_{nm,n-(m/2)}(e) \; m \sin[m(\omega + M + \Omega - \theta - \lambda_{nm})]. \qquad (3.58)$$

During integration of the Lagrange planetary equations,

$$d_t\psi = m \, d_t\lambda_S - q \, d_t\omega \qquad (3.59)$$

(denoting $\psi = m(\lambda_S - \lambda_{nm}) - q\,\omega$) is explicitly assumed constant by considering that the only variations are the secular rates d_tM, $d_t\omega$ and $d_t\Omega$ due to oblateness and the

rotation rate of the earth. The resultant perturbations of the Keplerian elements are sinusoidal with periods that are submultiples of the orbital period, earth rotation, and any combination of these. Consideration of Eq. (3.55) however, reveals that $d_t\psi$ must also be constant for any term V_{nmpq} that satisfies Eq. (3.55). This implies that other resonant terms must be added to those already defined. These correspond to those with periods P_c (usually termed "circulation" periods) related to the orbital period P through the resonant parameter R_c

$$R_c = P_c / P = 1 + (s_o / \Delta s) \qquad (3.60)$$

$$= 1 + (2 / \Delta s)$$

recalling that $s_o = 2$ for the GPS satellites. Generally, resonance can be observed physically for all satellites making more than 10 revolutions per day, whereas for those satellites for which P < 10 revolutions per day, it is possible to observe resonance only when the orbit period is nearly commensurate with the earth's rate of rotation (i.e. Δs is small); put differently, when the resonant parameter R_c is less than 10, resonant effects can be ignored. For the GPS satellites given that $\Delta s \approx 0.012$, R_c reaches values in the hundreds ($R_c \approx 161$), so that the resonant perturbations begin to appear as forced along-track oscillations given as

$$\Delta \approx a\,(\Delta M + \Delta\omega + \Delta\Omega\ \cos i) \qquad (3.61)$$

impressed on the motion of the secularly precessing ellipse. ΔM, $\Delta\omega$ and $\Delta\Omega$ in (3.61) are evaluated with Eq. (3.47) simplified by the assumptions (3.41) directly applicable to near-circular orbits, and fixing m to $m=s_o$ for the fundamental resonance or $m = k\ s_o$ for other resonances.

In addition, resonant perturbations would occur when the divisor in (3.47) approaches zero, i.e.

$$(n-2p)d_t\omega + (n-2p+q)d_tM + m\,(d_t\Omega - \omega_E) \approx 0\ . \qquad (3.62)$$

A particular case relevant to the GPS satellites falling into this class occurs near the critical inclination $i=63.4^o$ for which

$$\dot{\omega} \approx 0\ .$$

Currently the prototype Block I satellites have inclinations in the vicinity of the critical inclination so perigee motion is very slow. Block II satellites are planned to be near 55° inclination, so that the perigee advance should be at a faster rate.

3.4.3 Lunisolar (Third Body) Perturbations

The gravitational perturbations caused by a third body on the satellite's orbit are treated by defining a function R which is the third-body potential. This takes the form

$$R_d = (GM\ m_d/r_d)\ \{\ [1 - (2r/r_d)\ \cos\psi + (r^2/r_d)]^{-1/2}\ (r/r_d)\ \cos\psi\} \qquad (3.63)$$

where m_d and r_d is the mass and the geocentric distance of the third body, ψ is the angle between the geocentric vectors **r** to the satellite, and $\mathbf{r_d}$ to the third body respectively. For the case of the GPS satellites, only the sun and moon effects are usually treated in this context.

Generally for short satellite GPS arcs (<1/3 revolution or ≈ 4 hours duration), one can safely assume that the third body is stationary during that period. In this way, treating the earth, the satellite and the third body as point masses, the gravitational perturbations resulting from the moon and the sun can be developed analytically in a manner very similar to the case of the gravitational perturbations. The form of the disturbing function in this case takes the form (Kaula, 1962 ; Buffett, 1985)

$$R = (G\ m_d/r_d) \sum_{n=2}^{n_{max}} (a/r_d)^n \sum_{m=0}^{n} \sum_{p=0}^{n} \sum_{q=-\infty}^{\infty} F_{nmp}(e) \times$$

$$\times H_{nqp}(e)\ S^*_{nmpq}(\omega,\Omega,M) \qquad (3.64)$$

where

$$S^*_{nmpq} = \begin{cases} A_{nm}, & \text{n-m even} \\ -B_{nm}, & \text{n-m odd} \end{cases} \cos[(n-2p)\omega_d + (n-2p+q)\ M_d + m\ \Omega_d]$$

$$+ \begin{cases} B_{nm}, & \text{n-m even} \\ A_{nm}, & \text{n-m odd} \end{cases} \sin[(n-2p)\omega_d + (n-2p+q)\ M_d + m\ \Omega_d]\ . \qquad (3.65)$$

Here $H_{npq}(e)$ are the second inclination functions (in place of $G_{npq}(e)$ for the earth gravitational case), and A_{nm} and B_{nm} are spherical harmonics given as

$$A_{nm} = [(n-m)!/(n+m)!]\ \varepsilon_m\ P_{nm}(\sin\delta_d)\ \cos m\alpha_d \qquad (3.66a)$$

$$B_{nm} = [(n-m)!/(n+m)!]\ \varepsilon_m\ P_{nm}(\sin\delta_d)\ \sin m\alpha_d \qquad (3.66b)$$

$$\varepsilon_m = \begin{cases} 1, & m=0 \\ 2, & m\neq 0 \end{cases} \qquad (3.66c)$$

with α_d, δ_d and $\{a_d,e_d,i_d,\omega_d,\Omega_d,M_d\}$ being the declination and right ascension and the orbit elements of the third body. The general expressions for the $H_{nqp}(e)$ functions can be found, for example, in Buffett (1985); they need not be repeated here.

For a more complete treatment of the third-body disturbing function, this study has followed Buffett's (1985) suggestion incorporating the motion of of the third body following an earlier development by Giacaglia (1973) whereby the spherical harmonics A_{nm} and B_{nm} are rotated in the ecliptic reference frame resulting in Keplerian elements of the third body which are referred to in the ecliptic frame and hence can be more precisely approximated using the mean elements of the IAU 1980 Nutation Theory (Wahr, 1981; Seidelmann, 1982). In this way, the difficulty in approximating the equatorial elements of the sun and moon as simple functions in time is eliminated. The final result is a very complex-looking disturbing function

$$R = (G\, m_d/r_d) \sum_{n=2}^{n_{max}} (a/r_d)^n \sum_{m=0}^{n} \sum_{s=-n}^{n} \sum_{p=0}^{n} \sum_{t=0}^{n} \sum_{q=-\infty}^{\infty} \sum_{u=-\infty}^{\infty} F_{nmp}(i)$$

$$\times F_{nst}(i_d)\ H_{npq}(e)\ G_{ntu}(e_d)[(n-s)!/(n+m)!]\ \varepsilon_m$$

$$\times U^*_{nms}\ S_{nmsptqu}(\omega,\Omega,M) \qquad (3.67)$$

where

$$U^*_{nms} = U_{nms} \quad , \quad s \geq 0 \qquad (3.68a)$$

$$U^*_{nms} = (-1)^{n-s}\ [(n-s)!/(n+s)!]\ U_{nms} \quad , \quad s < 0 \qquad (3.68b)$$

$$U_{nms} = \sum_{r=r_1}^{r_2} (-1)^{n-m-r} \binom{n+m}{m+s+r} \binom{n-m}{r} \cos^{2r+m+s}(\varepsilon/2) \times \times \sin^{2n-2r-m-s}(\varepsilon/2) \qquad (3.68c)$$

with

$$r_1 = \max \{0;-m-s\}$$

$$r_2 = \min \{n-s;n-m\},$$

and

$$S_{nmsptqu} = \begin{cases} \tfrac{1}{2} \{ \cos [\{(n-2p)\omega+(n-2p+q)M+m\Omega\} + \{(n-2t)\omega_d+(n-2t+u)M_d+s(\Omega_d+\pi/2)\}] + \\ \qquad \cos [\{(n-2p)\omega+(n-2p+q)M+m\Omega\} - \{(n-2t)\omega_d+(n-2t+u)M_d+s(\Omega_d+\pi/2)\}] \} & \text{... s even} \\ \tfrac{1}{2} \{ \sin [\{(n-2p)\omega+(n-2p+q)M+m\Omega\} + \{(n-2t)\omega_d+(n-2t+u)M_d+s(\Omega_d+\pi/2)\}] + \\ \qquad \sin [\{(n-2p)\omega+(n-2p+q)M+m\Omega\} - \{(n-2t)\omega_d+(n-2t+u)M_d+s(\Omega_d+\pi/2)\}] \} & \text{... s odd} \end{cases} \qquad (3.69)$$

for the zonal harmonics, and

$$S_{nmsptqu} = \begin{cases} \cos [\{(n-2p)\omega+(n-2p+q)M+m(\Omega+\pi/2)\} - \{(n-2t)\omega_d+(n-2t+u)M_d+s(\Omega_d+\pi/2)\}] & \text{... (m-s) even} \\ (-1)^{n-s} \sin[\{(n-2p)\omega+(n-2p+q)M+m(\Omega+\pi/2)\} - \{(n-2t)\omega_d+(n-2t+u)M_d+s(\Omega_d+\pi/2)\}] & \text{... (m-s) odd} \end{cases} \qquad (3.70)$$

for the non-zonal harmonics.

The magnitude of the lunar perturbations even for short GPS orbital arcs can be of the order of 50-500 m. By comparison, given that the value of m_S / r_S^3 for the sun is ≈ 0.46, the perturbations of the solar third-body effects are about half of those of the moon. The gravitational attractions of the planets are even much smaller than those of the sun, so that their effects on GPS satellites are usually considered negligible. The indirect effect on GPS satellite motion caused by tidal deformations of the solid earth and the oceans due to the sun and the moon, is insignificant for short orbital arcs. For longer arcs (> 2 days duration), perturbing accelerations due to the ocean tide of the order of 10^{-9} m/s^2 can excite resonant effects resulting in orbital perturbations of over 2 meters. These can be accounted for by way of empirical accelerations described in Section 3.3.2.

3.6 Solar Radiation Pressure

Radiation from the sun carries with it not only energy but also momentum. When this radiation is absorbed by some object like a satellite, the momentum transferred from the radiation creates a repulsive pressure upon the impact of the light protons on the satellite. Generally, the pressure normal to the incident direction is

$$P_S = (I/c) \qquad (3.71)$$

where I, the intensity (solar flux) of the solar radiation at the mean earth-distance from the sun (i.e., one astronomical unit), is approximately 1360 Joule/m^2 sec and c is the speed of light in free space, so that $P_S \approx 4.6 \times 10^{-6}$ N/m^2. The resulting force on the spacecraft's surface is proportional to the effective area A of the surface normal to the incident radiation

$$F_r = C_R A \qquad (3.72)$$

where C_R is the surface reflectivity constant. The acceleration acting on a spacecraft of mass m at a distance r_{ss} ($\mathbf{r}_{ss} = \mathbf{r} - \mathbf{R}_{sun}$) from the sun is

$$\mathbf{r}''_{sp} = \frac{F_r\, r^2_{sun}}{m\, r^3_{ss}}\, \mathbf{r}_{ss} \qquad (3.73a)$$

$$= P_S \frac{C_R A}{m} \frac{\mathbf{r}_{ss}}{r_{ss}^3} r^2_{sun} \qquad (3.73b)$$

where r_{sun} (=149 598 500 km) denotes an astronomic unit and F_r is directed opposite to the satellite-sun line.

The GPS satellites have irregular, complex surface geometries and are constructed of materials with some surface portions reflecting diffusely and other portions reflecting specularly. In addition, they experience thermal changes in the surface area of the satellite that is illuminated by the sun as the satellite orbits the earth. Hence, in reality, at any moment the satellite will experience a net force that is not necessarily along the satellite-sun line plus a torque. A practical implication arising from such a complex satellite shape is that it is rather inappropriate to assign a single C_R value (especially for long arcs) and A for the GPS satellites. Clearly, a solar radiation model which was to take into account the separate contributions of the satellite's main body, the solar panels, the antenna array and the engine assembly would require a very complex model like the ROCK-IV model (Flieger et al, 1985). For short arcs a simpler "flat-plate model" with a single C_R estimated from observations would be quite adequate.

A practical solution for the effects of the solar radiation pressure requires several simplifying assumptions. The discontinuity of the force due to the shadowing effects of the earth as the satellite enters and exits the shadow zone (Fig. 3.5) is taken into account by the inclusion of a shadow function factor ν so that

$$r''_{sp} = \nu \; P_s \frac{C_R A}{m} \frac{r_{ss}}{r^3_{ss}} r^2_{sun} \; . \qquad (3.73c)$$

The development of an analytical theory in the presence of a shadow function is quite complicated for the reason that points of discontinuity are functions of the orbital elements and of the sun's position. Kaula (1966) has treated the problem in an iterative manner using a simple case of a cylindrical shadow model. Lala (1971) developed an elegant semi-analytical model designed for differential orbit improvement programs utilizing high-accuracy observations, and as such, is particularly suitable for the case of the GPS satellites as our implementations have shown. The shadow function used is of the form

$$\nu = \tfrac{1}{2}[\, 1 + \sin\chi \; (1-\cos^2\chi)^{-\frac{1}{2}} \,] \quad , \; \chi = \lambda - \Phi \qquad (3.74)$$

where λ is the satellite angular distance from the center of the earth's shadow and Φ is the angular distance between the shadow center and the shadow boundary intersection with the satellite orbit, with all angles being measured at the earth's center; consequently, when the satellite is in the sunlight portion of the orbit

$$\chi \in [0,\pi] \rightarrow \nu = 1 \qquad (3.75a)$$

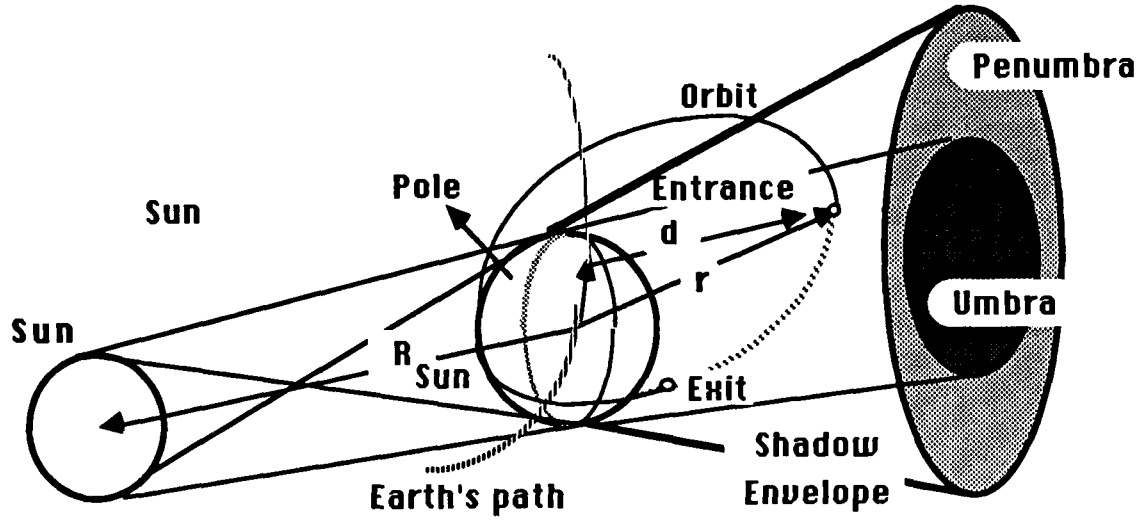

Fig. 3.5 - Solar radiation geometry

and when the satellite is in the shadow portion of the orbit

$$\chi \in [-\pi/2, 0] \rightarrow \nu = 0. \qquad (3.75b)$$

Unlike the gravitational field case where the usual method of obtaining the change in the Keplerian elements of the orbit with time was to first find the disturbing function and then use Lagrange's planetary equations, this approach cannot be used in the present problem since the force cannot, in general, be written as the gradient of the potential. Hence no disturbing function exists. Instead the force must be dealt with directly. Once the force is found, it can be used in the Gaussian form of Lagrange's equations in the manner indicated below.

The solar radiation pressure accelerations can be expressed in terms of orthogonal components (cf. Fig. 3.1a) $\mathbf{s} \equiv [S,T,W]^T$ to be used in Gauss' form of Lagrange's planetary equations. These components are derived from the Cartesian components $\mathbf{r}''_{sp}$ of the satellite acceleration vector via a rotation:

$$\mathbf{s} = [S,T,W]^T = R_{sr}\, \mathbf{r}''_{sp} \qquad (3.76)$$

where

Table 3.1 - Solar radiation pressure parameters for use in Gaussian form of Lagrange's planetary equations (after Lala, 1971).

q	a	e	i	ω	Ω	M
2FA	$2a+F\sqrt{(1-e^2)}$	$F\sqrt{(1-e^2)}$	$F\sqrt{(1-e^2)}$	$F\sqrt{(1-e^2)}/e$	$F\sqrt{(1-e^2)}/e$ $+\sin i$	$F\sqrt{(1-e^2)}/e$
I	1	2 + Me	2 + Me	2 + Me	1 + Me	2 + Me
XC_o	$e\,A_2$	A_2	0	0	$-2\,A_1$	$2\,A_1$
XC_1	A_2	$e\,A_2$	B_1	B_2	$e\,B_1$ $-e\,\mathrm{ctg}i\,B_2$	$-3e\,A_1$
XC_2	0	$(1-e^2)\,A_2$	$-e\,B_1$	$-e\,B_2$	$(1-e^2)\,A_1$ $+e^2\,\mathrm{ctg}i\,B_2$	$-(1-3e^2)A_1$
XS_o	$-A_1$	$-e\,A_1$	$-B_2$	B_1	$-e\,\mathrm{ctg}i\,B_1$	$-2e\,A_2$
XS_1	0	$-(1-e^2)\,A_1$	$e^2\,B_2$	$-e\,B_1$	A_2 $+e^2\,\mathrm{ctg}i\,B_1$	$-(1-2e^2)A_2$

$F = (a^2/\mu)\,r''_{sp}$

$A_1 = r_{11}$ element of R_{rs}

$A_2 = r_{21}$ element of R_{rs}

$B_1 = W\cos\omega$

$B_2 = W\sin\omega$

(1) M_e = max power of the eccentricity retained in the series

(2) $XC_i = -e\,XC_{i-1}$, $i>2$

$XS_i = -e\,XS_{i-1}$, $i>1$

$$R_{sr} = R_3(\omega(t))\ R_1(i(t))\ R_3(\Omega(t))\ R_1(-\varepsilon)\ R_3(-\lambda_{sun}(t))$$

$$= \{r_{ij}\} \quad , \quad i,j=1,2,3 \qquad (3.77)$$

with λ_{sun} being the ecliptic longitude of the sun.

For short satellite arcs, it makes sense to assume that the orientation of the orbit with respect to the sun is constant. In this way, the elements of the rotation matrix R_{sr} can be treated as constant which in turn allows the components of **s** to be treated as functions only of the true anomaly of the satellite. Introducing (3.77) into (3.76) followed by the insertion of the latter into the Gaussian form of the Lagrange planetary equations leads to

$$D_t z = 2n\ FA \{ \sum_{i=0}^{I} XC_i \cos^i f + \sum_{i=0}^{I-1} XS_i \cos^i f \sin f \} \ , z \equiv [a,e,i,\omega,\Omega,M] \qquad (3.78)$$

where n is the mean satellite motion and the parameters FA, XC_i and XS_i are given in Table 3.1 for each Keplerian element. The trigonometric terms $\cos^i f$ are usually replaced by equivalent expressions in terms of the mean anomaly M using the Hansen coefficients $X_n^{m,k}$; that is

$$\cos^i f = X_o^{o,i} + \sum_{u=1}^{\infty} [\, X_u^{o,i} + X_{-u}^{o,i} \,] \cos uM \ . \qquad (3.79)$$

Under the foregoing assumption that the orientation of the orbit is constant with respect to the sun within the time interval of integration, the solution of the equations of motion would comprise only short-period and secular terms. Over longer arcs, the coefficients FA, XC_i and XS_i could not be considered to be constant with the integration, resulting instead in long-period perturbations. This effect would be best observed in the semi-major axis which does not undergo a long-period motion from any other source. This perturbation is (cf. Murphy and Felsentreger, 1966)

$$\Delta a_{sr} = -\ 2(a^2/\mu)\ P_s\ (A/m)\ \left[S \cos E + T \sqrt{(1-e^2)}\ \sin E \right]_{E_2}^{E_1} \qquad (3.80)$$

where E_1 and E_2 are the eccentric anomalies of the satellite at the exit from and the entrance into the earth's shadow. For longer arcs, the assumption of constant orbit

orientation with respect to the sun can still be tolerated if the formulation in (3.36) is adopted to account for any unmodelled effects.

For GPS satellites the acceleration of a satellite due to the solar radiation pressure is of the order of 10^{-7} m/sec^2; for short satellite arcs the perturbation effects on the orbit can be of the order of 2-10 m, whereas they may reach values from less than 100 m to over 800 m after 2 days.

The indirect solar radiation arriving at the satellite after being reflected from the earth's surface; i.e. *the albedo effect*, may be considerable especially on the orbits of satellites with high surface-to-mass ratios like the GPS spacecraft. The effect of the albedo pressure on the GPS satellite orbits is considered to be of the order of 1-2% of the direct solar radiation effect (King et al., 1985), which incidentally is of the same order of magnitude as the effect of the earth and ocean tides, and hence, can be neglected for most applications or considered in later developments as required.

SECTION 4

TIME, COORDINATE FRAMES, AND OBSERVABLES

For the processing of GPS data we usually require results in three dimensions although the GPS system is inherently four-dimensional. In doing so, there are a number of coordinate and time frames which are inherent in these computations and hence, deserve particular attention. To begin with, there are various coordinate systems which are naturally involved in the determination of the satellite orbits. The satellites are moving with respect to an inertial reference frame and hence, a fundamental requirement in the processing of GPS data is the definition of an inertial coordinate system for the formulation and the solution of the equations of motion. Other coordinate systems are required in order to develop the perturbing forces acting on the satellites. The station coordinates on the other hand are usually expressed in an earth-centered, earth-fixed (ECEF) reference frame which is moving with respect to the inertial reference frame because of the earth's motion. The measurements made by the GPS receivers are essentially timing differences of the time of transmission (in the satellite time scale) and the time of arrival (in the receiver time scale) of a particular electromagnetic signal transmitted by the satellite and therefore are affected by the timing errors in the clocks involved. Such measurements necessarily involve detailed and careful consideration of a number of time scales. In particular, time enters into GPS data processing in three different ways: through the time tags of the measurements, through the time tags of the satellite ephemerides, and through the measurements themselves. These are intimately connected with the definition of the observables and for that reason, this section deals with the brief description of the celestial and terrestrial reference systems used in the orbit computations. The time systems inherent in the GPS system and the GPS observables will also be discussed.

4.1 Time and GPS

In order to fully appreciate the role of time in GPS data processing, it is necessary to define the various time systems involved and their attendant time scales. Some of these definitions are standard. Others are particular to the GPS system and the GPS data processing. These will assist in the subsequent development of the observation equations and the formulation of strategies relating to the elimination of GPS clocks and bias elimination errors.

In general, there are three time systems that are used in GPS computations: dynamical time, atomic time, and sidereal time. Like any orbital body the motions of the GPS satellites can be expressed most readily as a function of a *dynamical time system* as defined by the theories of Newtonian or relativistic mechanics. This is the fundamental uniform time system that defines the independent variable in the satellite's equations of motion. When we generate a set of four-dimensional coordinates of the GPS satellites-"ephemerides"-we implicitly use dynamical time. In this system, the motions of the satellites and the perturbing forces acting on them can be used themselves to measure

dynamical time. For all practical purposes, the dynamical time is equivalent to an *Atomic Time System* (Moritz and Mueller, 1986). The latter is more easily realized by the atomic clocks, e.g. the international network of clocks coordinated by the Bureau International de l'Heure-BIH, that maintain the International Atomic Time scale (TAI). The atomic time scale defined by such clocks is not itself a true or perfect time scale but is the most accurate measure of "true" time and is used to define the origin and the rate of other time scales such as (1) the *Terrestrial Dynamical Time* , TDT (which represents the dynamical time for motion of bodies within the earth's gravitational field) being related to TAI by

$$TDT = TAI + 32.184 \text{ sec} ; \qquad (4.1)$$

(2) the *Universal Coordinated Time*, UTC which runs at the same rate as TAI but is incremented by 1-second jumps (leap seconds) when necessary to account for the slowing down of the earth's rotation with respect to the sun and is being transformed to TAI by adding a specific integral number of seconds (23 at present), and (3) the *GPS System Time.* The GPS System Time used by the Control Segment is derived from UTC as maintained by the atomic clocks of the U.S. Naval Observatory. GPS time was set to UTC on 6 January, 1980 but because the leap seconds that are introduced in the UTC time are not accounted for in the GPS time, there presently exists an integer-second difference between the two time systems. Currently this difference is 4 sec, i.e.

$$\text{GPS Time} = \text{UTC} + 4 \text{ seconds} \qquad (4.2)$$

which implies a constant offset of 19 sec between GPS time and TAI, i.e.

$$\text{GPS Time} = \text{TAI} - 19 \text{ seconds.} \qquad (4.3)$$

In addition to the atomic time scales which provide a uniform measure of time, *Sidereal Time* is required to relate the inertial and terrestrial systems. *Universal Time*, which is the most common form of sidereal time, is a rotational time scale affected by the irregularities of the earth's rotation. By correcting observed universal time for the effects of polar motion, according to the definition of UT1, this time reflects the actual angular orientation of the earth with respect to the vernal equinox at that instant. UT1 is derived from observations made by BIH and is published as corrections to UTC, i.e.

$$UT1 = UTC + DUT1. \qquad (4.4)$$

Since UTC is related to GPS System Time, the transformation between UT1 and GPS time is accomplished by a simple shift in origin.

The relationship between UT1 and the Greenwich sidereal time is defined as

$$\text{GMST at } 0^h \text{ UT1} = 24110^s.54841 + 8640184^s.812866\ T_u + 0^s.093104\ T_u^2 - 6^s.2 \times 10^{-6}\ T_u^3 \qquad (4.5)$$

where GMST is the mean sidereal time defined by the Greenwich Hour Angle of a fictitious sun S_u and the mean equator and equinox of date, and T_u is the number of Julian centuries elapsed since JD 2451545.0 (36525 days of Universal Time elapsed since J2000=Jan. 1.5, 2000).

Individual atomic clocks maintain their own time scale often referred to as *clock time*. This is true also for the clocks of the GPS satellites and receivers, each of which maintain a scale (*satellite clock time* vs. *receiver clock time*) which is a representation of the "true" time (be it GPS time or UTC), but in practice it has some error with respect to it. The GPS satellites carry both rubidium and cesium frequency standards. The space environment is kind to the GPS clocks, but the latter are physically left to drift off the standard GPS time system. However, their performance is continuously monitored by the GPS Control Segment and the amount by which they have drifted is accurately known. For example, synchronization between the GPS satellite clocks is kept to within about 20 nsec by the broadcast clock corrections, and the GPS time scale is synchronized to UTC to within 100 nsec. Future experiments like the Advanced Clock/Range Experiment (ACRE) are designed to test compact H-Maser frequency standard capability for future GPS satellites. This, in addition to laser ranging capability at the cm-level (for this experiment Block I GPS satellite #12 is to be retrofitted with laser retroreflectors for 1992 launch (Buisson, 1987)) will allow simultaneous comparison of H-Maser, cesium and rubidium clocks in space and the capability to separate orbit position and velocity errors from satellite clock errors.

As with any measurement system based on the motion of satellites such as SLR and LLR, NNSS/TRANSIT, as well as GPS, another time scale implicit in the measurements is one that relates to the time tags of the satellite ephemerides. This is often referred to as the *Satellite Ephemeris Time* or SET (not to be confused with the formal Ephemeris Time (ET) system defined by the orbital motion of the earth about the sun). Just as individual clocks can have differences with respect to a time scale such as GPS time or UTC, the motion of the satellite can be considered as defining a unique time scale which in general will represent the satellite ephemeris time but will have some errors with respect to it. Given, for instance, a set of satellite positions (say, from range observations) and an ephemeris that relates the satellite positions to SET, one can define a time scale close to the SET but different from it due to errors in the ephemeris and the satellite positions themselves. This realization of SET may be referred to as *Satellite Orbital Time* (SOT). SOT can be considered as an approximate measure of dynamical time and should not be confused with satellite clock time.

The time system inherent in the GPS computations involves all the time scale relationships just outlined. The physical GPS observable, be it a pseudorange or a carrier beat phase, is essentially the difference between the reference phase of the satellite oscillator at the time of signal emission and the reference phase of the receiver oscillator at the time of signal reception. In rigorous terms this clock difference is composed of the accumulated sum of five components: the difference between station clock time and proper time (ideally UTC) at the station, the difference between proper time and coordinate time at the station (i.e. UTC-TDT), the difference between the coordinate time at signal emission and signal reception, the difference between coordinate and proper time at the satellite (TDT - UTC + Δt_{rel}; Δt_{rel} = relativistic clock correction),

and finally the difference between proper time and clock time at the satellite. In commonly used GPS terminology, these differences can be readily interpreted (disregarding for the moment the influence of atmospheric effects and instrumental delays): the first and fifth differences are nothing else but the familiar station and satellite clock errors respectively; the second and fourth differences are general relativistic time transformations; the third difference is the solution to the light-time equation connecting the signal emission and reception event; i.e., the geometric range between satellite and receiver defined by the epochs of these events. If during the modelling of the observations these relationships are rigorously taken into account, then the GPS computation time may be equivalent to any of the above time scales. For most applications certain time offsets between these time scales may be assumed negligible and ignored altogether, hence defining by default the computational time by the time scale of one reference clock (at the satellite or receiver level), by the average clock times of several receiver clocks or even (as is often done routinely in GPS data reductions) by the time offsets obtained from the GPS pseudorange navigation solutions.

These concepts will hopefully became more apparent by seeing how these time scales are used in the development of the GPS carrier phase observation equations.

4.2 Coordinate Frames

Although the observables obtained from GPS receivers are instrumented differently from one receiver type to another, it can be shown that they are all, be it pseudorange, carrier phase or Doppler, functions of the instantaneous geometric ranges between the satellites and ground receivers and their time derivatives. This geometric range is by far the larger portion of the observed range, and also the main cause of complexity in building the observational model because of the numerous transformations which are necessary to relate the inertial reference frame used for the formulation and solution of the equations of motion and for relating the satellite position to the earth-fixed, earth-cantered reference frame which the station positions are referenced to.

The coordinate system adopted for the solution of the equations of motion in the present work is based on the fundamental astronomic system defined by the geocentric, equatorial frame with the equator and equinox of J2000.0 and the 1976 IAU conventions and the 1980 Nutation Series (Leiske et al., 1977; Seidelmann, 1982).

For the numerical solution, in order to evaluate the accelerations due to the earth's gravitational potential, the inertial coordinates of the satellite are rotated into the average terrestrial system (i.e. the system in which the spherical harmonic coefficients are referred to) using a series of rotations

$$r_{AT} = S\,N\,P\,r_I\,. \qquad (4.6)$$

where S, N and P denote transformations to be explained below for polar motion and UT1, nutation and precession respectively. When the gravitational accelerations are evaluated in the AT system, they are transformed back into the inertial system using the transpose of the complete transformation matrix S N P.

In (4.6) the orientation of the AT system with respect to the true system of date is determined via S which represents the transformation

$$S = R_2(-x)\ R_1(-y)\ R_3(GAST) \qquad (4.7)$$

where x,y are the coordinates of the polar motion (readily obtained from published BIH values) and the Greenwich apparent sidereal time is related to GMST (Eq. (4.5)) by

$$GAST = GMST + \delta\psi \cos(\bar{\varepsilon} + \delta\varepsilon) \qquad (4.8)$$

where $\bar{\varepsilon}$ is the obliquity of the mean equator of date given by

$$\bar{\varepsilon} = 23^{o}\ 26'\ 21.''448 - 46.''8150\ T - 5.''9 \times 10^{-4}\ T^2 + 1.''813 \times T^3 \qquad (4.9)$$

and $\delta\varepsilon$ is the nutation in obliquity. $\delta\varepsilon$ along with $\delta\psi$, the nutation in longitude, and ε ($\equiv \bar{\varepsilon} + \delta\varepsilon$), the obliquity of the ecliptic, relate the mean and true systems of date, Fig. 4.1a. Following Wahr's (1981) nutation theory for an elastic, elliptical and oceanless earth, $\delta\varepsilon$ and $\delta\psi$ can be expressed as a series expansion of linear combinations of five fundamental arguments a_i, i.e.,

$$\delta\varepsilon = \sum_{i=1}^{106} A_i(T) \cos\left[\sum_{j=1}^{5} k_{ij}\, a_i\right] \qquad (4.10a)$$

$$\delta\psi = \sum_{i=1}^{106} B_i(T) \sin\left[\sum_{j=1}^{5} k_{ij}\, a_i\right] \qquad (4.10b)$$

where $A_i(T)$, $B_i(T)$, k_{ij} are known coefficients, and the a_i terms represent:

- the mean anomaly of the moon

$$a_1 = \ell = 485866.''733 + (1325^{r} + 715922.''633)\ T + 31.''310\ T^2 + 0.''064\ T^3\ ; \qquad (4.11a)$$

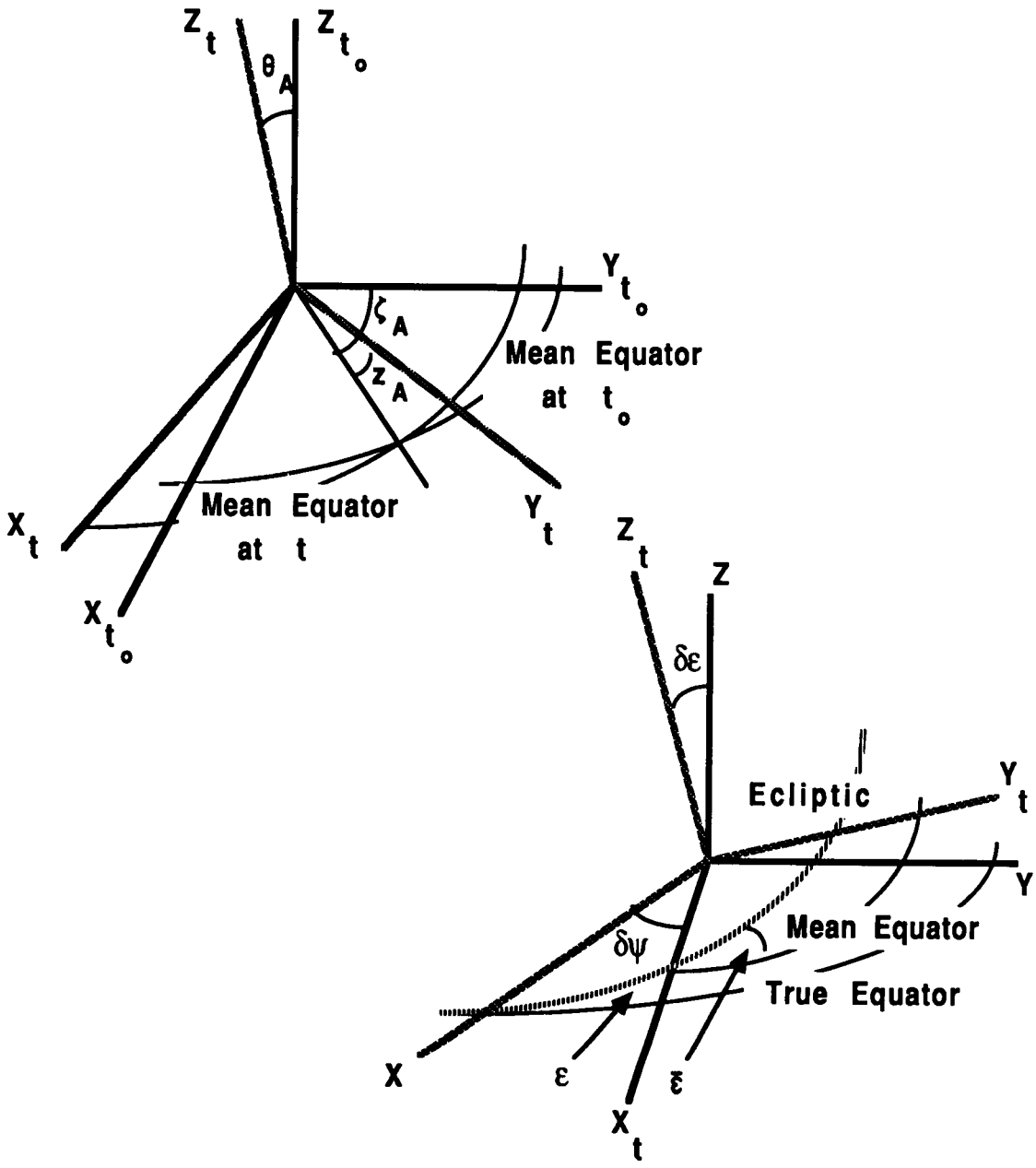

Fig. 4.1 - (a) above, Precession angles.
(b) below, Nutation angles.

- the mean anomaly of the sun

$$a_2 = \ell' = 1287099".804 + (99^r + 1292581".224)\ T$$
$$- 0".577\ T^2 + 0".012\ T^3\ ; \qquad (4.11b)$$

- the mean argument of latitude of the moon (F≡L-Ω)

$$a_3 = F = 335778".877 + (1342^r + 295263".137)\ T$$
$$+ 13".257\ T^2 + 0".011\ T^3\ ; \qquad (4.11c)$$

-the mean elongation of the moon from the sun (D≡L-L', and L and L' are the astronomic longitude of the moon and the sun respectively)

$$a_4 = D = 1072261".307 + (1236^r + 1105601".328)\ T$$
$$- 6".891\ T^2 + 0".019\ T^3\ ; \qquad (4.11d)$$

- the mean longitude of the ascending lunar node measured in the ecliptic

$$a_5 = \Omega = 450160".280 + (5^r + 482890".539)\ T$$
$$- 7".455\ T^2 + 0".008\ T^3\ ; \qquad (4.11e)$$

where $1^r \equiv 360^o \equiv 1296000"$. An additional set of terms, although not as yet implemented in the present software, can be added to the nutation variations (4.10) to account for (a) the out-of-plane nutations (not originaly included in the IAU 1980 Nutation series) which are identical to (4.10) except for the replacement of sin ↔ cos

$$\delta\varepsilon = \sum_{i=1}^{106} C_i(T) \sin \left[\sum_{j=1}^{5} k_{ij}\ a_i \right] \qquad (4.12a)$$

$$\delta\psi = \sum_{i=1}^{106} D_i(T) \cos \left[\sum_{j=1}^{5} k_{ij}\ a_i \right] \qquad (4.12b)$$

and (b) for the free-core nutations with nominal period of 460 days (Sovers and Border, 1987).

Applying these rotations defines the transformation between the mean system of date and the true system of date through the nutation matrix N

$$N = R_1(-\varepsilon)\, R_3(-\delta\psi)\, R_1(\bar{\varepsilon})\ . \qquad (4.13)$$

The transformation from the mean equatorial system of date to the equatorial system of the reference epoch (e.g. J2000.0) is represented by the three equatorial arguments ζ_A, Z_A and θ_A (Fig. 4.1b)

$$\zeta_A = 2306".2181\ T + 0".30188\ T^2 + 0."017998\ T^3 \qquad (4.14a)$$

$$z_A = 2306".2181\ T + 1".09468\ T^2 + 0."018203\ T^3 \qquad (4.14b)$$

$$\theta_A = 2004".3109\ T - 0".42665\ T^2 - 0."041833\ T^3 \qquad (4.14c)$$

defining the precession matrix P as

$$P = R_3(-90^o - z_A)\, R_1(\theta_A)\, R_1(90^o - \zeta_A) \qquad (4.15)$$

which completes the model for the orientation of the earth in our adopted J2000.0 system of reference.

4.3 Observables

GPS observations can always be thought of as some type of biased range. Code pseudorange or carrier beat phase observation equations contain various bias terms modelling the effects of the imperfections of the time scales used to generate and time tag these observables, atmospheric delays, etc. In the case of the carrier phase the corresponding observation equation contains an additional linear bias term-the initial phase ambiguity. A comprehensive summary of the simplified form of these equations and the various linear combinations has been given by Wells et al. (1986). A more detailed account of the mathematical development of these equations can be found, for instance, in Remondi (1984) and Wei (1986) just to mention a few. In this overview, only the carrier beat phase observation equation will be given for completeness, since this was the observable dealt with mostly in this study.

It is well known that the phase Φ of an ideal oscillator relates uniquely to the true time τ (which provides the theoretical frame for deriving the observation equations) by

$$\Phi(\tau) = \Phi(\tau_o) + \int_{\tau_o}^{\tau} f \, d\tau \qquad (4.16a)$$

where f is the time-dependent frequency of the oscillator. For an oscillator of high stability, over a short interval $(\tau-\tau_o)$, (4.16a) can be written as

$$\Phi(\tau) = \Phi(\tau_o) + f_o \, (\tau-\tau_o) \qquad (4.16b)$$

where f_o is the nominal frequency of the oscillator. In practice, the carrier beat phase measurement at some time epoch $T(\tau)$, measured at the receiver time scale, is based on the phase alignment of the receiver-generated signal with the incoming carrier signal without the knowledge of which cycle would represent perfect cycle synchronization (see Fig. 4.2). Hence, the total phase ϕ consists of an integer number of cycles $N(\phi;T_o,T)$ counted since the initial lock-on time T_o, a fractional part $Fr(\phi)$ and an unknown integer number N of cycles at the initial epoch T_o, so that

$$\phi = N(\phi;T_o,T) + Fr(\phi) + N(T_o) \qquad (4.17a)$$

$$= \phi_{measured} + N(T_o) \, . \qquad (4.17b)$$

Ideally, if the receiver maintains continuous lock on the signal, N is a unique number for each satellite-receiver pair. Otherwise, every time the receiver loses lock, a cycle slip occurs which effectively introduces a new ambiguity term (see also Section 2). Generally the frequency of the receiver oscillator is not constant with respect to "true time" and thus the clock time will have a non-linear relationship with respect to it. For the nominal L_1 (=1575.42 MHz) frequency a 0.6 nsec error in clock time will create approximately 1 cycle of phase error. Since phase is physically invariant, the received phase of the receiver-generated carrier signal at reception time T is the same as the phase of the satellite-generated carrier signal at transmission time t, and therefore

$$\phi = \phi(t) - \phi(T) \qquad (4.18)$$

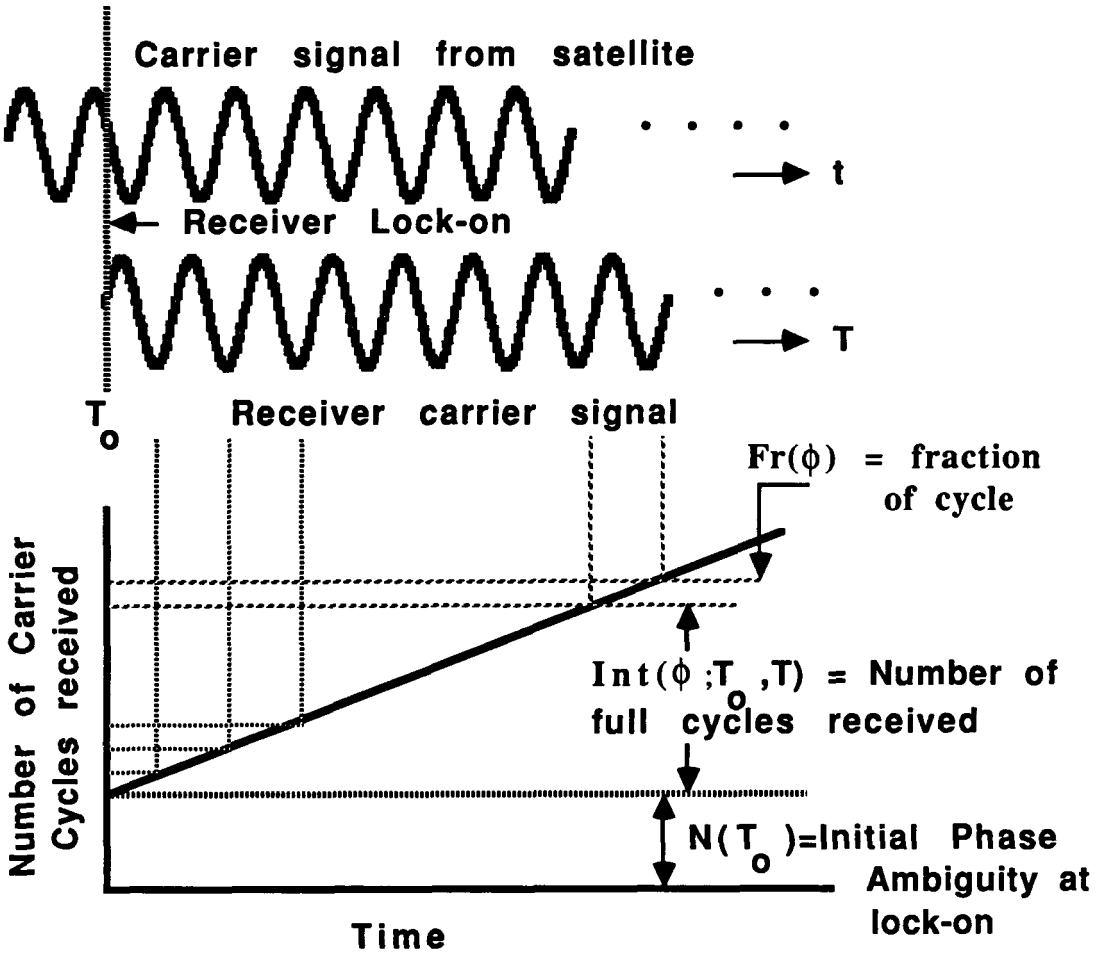

Fig. 4.2 - Schematic representation of carrier beat phase observable generation.

or after considering Eq. (4.16b) and the relationship of receipt and transmit times

$$t + dt + (\rho/c) = T + dT + \delta_{atm} , \qquad (4.19)$$

the equation for the instantaneous carrier beat phase observable for one satellite, one receiver and one epoch is readily obtained as

$$\phi = - (f/c)\ \rho(\tau) - f\ [dt(\tau)-dT(\tau)] - \delta_{atm} + N \qquad (4.20)$$

where any satellite and station frequency deviations from the nominal value have been absorbed in the terms dt and dT respectively; c is the speed of light in the vacuum; and δ_{atm} denotes the delay caused by the travel of the signal through the earth's atmosphere.

This is caused by a change in phase velocity but, as the effect is small compared to the geometric range, it can be modelled as a time delay, phase delay or simply as an increase in range (cf. Section 2); ρ denotes the geometric range to the satellite determined by the position vectors of the satellite and the receiver in the chosen coordinate frame. In (4.20) the time-dependent terms have been purposely expressed in terms of the "true time". This does not usually present any problems for the clock terms, if these are eliminated by differencing. The geometric range ρ, however, is largely dependent on using a correct time scale relating the time of transmission to the satellite ephemeris time from which the satellite position vector and hence (together with a priori information on the station's position) the range is calculated. This calculation is usually done iteratively, with the accuracy of the calculation being dependent on the difference of the time of reception used (i.e. measurement time tag) and the time of reception expressed in terms of the ephemeris time. The effect of time tag error on the computation of the geometric range is readily seen in the Taylor series approximation

$$\rho(\tau) = \rho(T+\delta\tau) = \rho(T) + \delta\tau\,\dot{\rho}(T) + [\,(\delta\tau)^2\,\ddot{\rho}(T)/2] \qquad (4.21)$$

where $\dot{\rho}$ and $\ddot{\rho}$ are the rate of change (Doppler) and range acceleration (negative of the Doppler slope). For the height of the GPS satellites, the rate of change reaches a maximum of about 800 m/sec, whereas the maximum value of the second derivative of the range is about 175 mm/sec^2, or less than 1 Hz/sec. Clearly, if the tag error is less than one μsec both of these terms can be ignored, otherwise additional terms will be required in the adjustment to account for this type of error: ideally one extra parameter for every receiver and every epoch. However, this obviously would lead to a prohibitively large number of parameters and hence is seldom done; instead, the technique of differencing between-receivers, -satellites and -time, and combinations thereof is often preferred (albeit at the expense of some loss of geometry, and redundancy of the data as well as the introduction of correlations).

The combination of Eq. (4.16) through (4.21) leads to the complete form of the observation equation of a carrier beat phase measurement from station 1 to satellite j

$$\phi^j_1 = -\,(f/c)\,[\rho^j(T_1) + \delta\tau_1\,\dot{\rho}^j(T_1)] - f\,[dt^j(t^j_1) - dT_1(T_1)] - \delta^j_{1,atm} + N^j_1\,. \qquad (4.22)$$

The single-difference phase observable, defined as the difference of simultaneous measurements made by two stations, 1 and 2, to the same satellite j (with simultaneity customarily defined at the receiver time scale level), is readily obtained from (4.22) as

$$\phi^j_{12} = \phi^j_2 - \phi^j_1 \qquad (4.23a)$$

$$= \{- (f/c)\ [\rho^j(T_2) + \delta\tau_2\ \dot{\rho}^j(T_2)]$$

$$- f\ [dt^j(t^j_2) - dT_2(T_2)] - \delta^j_{2,atm} + N^j_2\ \}$$

$$- \{- (f/c)\ [\rho^j(T_1) + \delta\tau_1\ \dot{\rho}^j(T_1)]$$

$$- f\ [dt^j(t^j_1) - dT_1(T_1)] - \delta^j_{1,atm} + N^j_1\ \} \qquad (4.23b)$$

$$= - (f/c)\ \{[\rho^j(T_2) - \rho^j(T_1)] + [\delta\tau_2\ \dot{\rho}^j(T_2) - \delta\tau_1\ \dot{\rho}^j(T_1)]\}$$

$$- f\ [dt^j(t^j_2) - dt^j(t^j_1)] + f\ [dT_2(T_2) - dT_1(T_1)]$$

$$- [\delta^j_{2,atm} - \delta^j_{1,atm}] + [N^j_2 - N^j_1]. \qquad (4.23c)$$

The satellite clock errors still appear in (4.23c) because the two observed signals were transmitted at different times, so that the term containing these clock errors describes the change in satellite clock error between transmit times. In theory, for the cesium clocks of the satellites this term may be considered as insignificant, although with the postulated future dithering of the signals this may well not be the case. In practice, the magnitude of this term depends on two factors: (a) the difference in transit times, and hence the length of the baseline between the two stations and their relative geometry with respect to the satellite; and (b) the level of synchronization between the two station time scales.

Similarly, the equation of a between-receivers and -satellites, or double-difference carrier beat phase observable can be readily obtained from (4.23) for two stations (1 and 2) and two satellites (j and k) as

$$\phi^{jk}_{12} = \phi^k_{12} - \phi^j_{12} \qquad (4.24a)$$

$$= - (f/c)\ \{[\rho^k(T_2) - \rho^k(T_1)] + [\delta\tau_2\ \dot{\rho}^k(T_2) - \delta\tau_1\ \dot{\rho}^k(T_1)]\}$$

$$- f\ [dt^k(t^k_2) - dt^k(t^k_1)] + f\ [dT_2(T_2) - dT_1(T_1)]$$

$$- [\delta^k_{2,atm} - \delta^k_{1,atm}] + [N^k_2 - N^k_1]$$

$$- \{ - (f/c)\ \{[\rho^j(T_2) - \rho^j(T_1)] + [\delta\tau_2\ \dot{\rho}^j(T_2) - \delta\tau_1\ \dot{\rho}^j(T_1)]\}$$

$$- f\,[dt^j(t^j_2) - dt^j(t^j_1)] + f\,[dT_2(T_2) - dT_1(T_1)]$$

$$- [\delta^j_{2,atm} - \delta^j_{1,atm}] + [N^j_2 - N^j_1]\,\} \qquad (4.24b)$$

$$= - (f/c)\,\{[\rho^k(T_2) - \rho^k(T_1)] - [\rho^j(T_2) - \rho^j(T_1)]\}$$

$$- (f/c)\,\{[\delta\tau_2\,\dot{\rho}^k(T_2) - \delta\tau_1\,\dot{\rho}^k(T_1)] - [\delta\tau_2\,\dot{\rho}^j(T_2) - \delta\tau_1\,\dot{\rho}^j(T_1)]\}$$

$$- f\,\{[dt^k(t^k_2) - dt^k(t^k_1)] - [dt^j(t^j_2) - dt^j(t^j_1)]\}$$

$$- \{[\delta^k_{2,atm} - \delta^k_{1,atm}] - [\delta^j_{2,atm} - \delta^j_{1,atm}]\}$$

$$+ [N^k_2 - N^k_1] - [N^j_2 - N^j_1]\,. \qquad (4.24c)$$

In this equation all receiver clock errors have been removed, whereas the term containing the satellite clock errors has been reduced by a factor $c/\rho \approx 10^5$ compared to that of the single-difference equation, thus effectively making this observable also insensitive to the satellite clock errors.

Double differences are mathematically correlated because of the differencing process. This correlation depends on how the double differences are formulated from the corresponding single differences. For instance, if they are always formed between two successive satellites, the correlation coefficient between two so-formed double differences will be -0.5. Preferable to this scheme, is to account for this correlation (as it was done in the software of the present study) by considering the measurements at the same time epoch and to construct an elementary transformation

$$\mathbf{C}_{DD} = \mathbf{D}\,\mathbf{C}_{SD}\,\mathbf{D}^T \qquad (4.25)$$

where $\mathbf{C}_{SD}$ is the covariance matrix of the individual single differences at that epoch assumed uncorrelated and with variance σ^2; D is the double-difference operator transforming the single-difference observation space into the double-difference observation space, i.e.

$$D = \begin{vmatrix} -1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & -1 & 0 & 0 & \dots & 0 & 0 \\ . & . & . & . & \dots & . & . \\ 1 & 0 & 0 & 0 & \dots & -1 & 0 \\ 1 & 0 & 0 & 0 & \dots & 0 & -1 \end{vmatrix} \qquad (4.26)$$

leading to the covariance matrix associated with these double-difference observations

$$C_{DD} = \sigma^2 \begin{vmatrix} 2 & 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & 2 & 1 & 1 & \dots & 1 & 1 \\ . & . & . & . & \dots & . & . \\ 1 & 1 & 1 & 1 & \dots & 2 & 1 \\ 1 & 1 & 1 & 1 & \dots & 1 & 2 \end{vmatrix} \qquad (4.27)$$

or equivalently to the weight matrix

$$P_{DD} = \frac{\sigma^{-2}}{n+1} \begin{vmatrix} n-1 & -1 & -1 & \dots & -1 & -1 \\ -1 & n-1 & -1 & \dots & -1 & -1 \\ . & . & . & \dots & . & . \\ -1 & -1 & -1 & \dots & n-1 & \\ -1 & -1 & -1 & \dots & -1 & n \end{vmatrix} \qquad (4.28)$$

which is of dimension n × n where n+1 is the number of satellites observed at that particular observing epoch. The weight matrix for an entire session has a block-diagonal form with each individual sub-block being of the form (4.28) and the size of different sub-blocks being generally variable depending on the number of satellites observed at that particular epoch corresponding to each sub-block.

SECTION 5

STRATEGIES FOR ORBIT DETERMINATION AND BASELINE ESTIMATION

The objectives of future GPS-based geodetic systems have often been summarized as being twofold:

- to provide accurate positions for the thousands of horizontal and vertical control points that comprise the current national geodetic databases in support of the uniform topographic mapping efforts and the growing need for better geodetic control in multipurpose cadastre systems; this would require both stations for remote (100's of kilometers) positioning with decimeter accuracy, and stations for local (10's of kilometers) positioning to a 2-3 cm accuracy level;

- to measure vector baselines over distances from 100 to 4000 km both in length and orientation with an accuracy better than 0.01 ppm and to apply this capability to the study of the tectonic and crustal deformations.

To fulfill these objectives, and in order to fully utilize available time in the field without an increase in the workforce, it is desirable to have highly compact stations which can be easily transportable, set up and operated by a minimum number of personnel (e.g. 1-3 persons) and by implementing automated techniques, be able to carry out such measurements over short observation periods (e.g. less than 4 hours including calibration assemblies). Any conceivable strategy to meet these needs must be three-tiered: the highest level being the orbital positions of the GPS satellites, the intermediate level being a set of continuously operating GPS receivers at several kilometers' spacing, and the lowest level, the traditional monumented control. There are currently two independent efforts leading to the development and setting up of such a general system :

- the development by the Jet Propulsion Laboratory for NASA (and independently by the U.S. Geological Survey (USGS), the National Geodetic Survey (NGS) and the University of Texas) of a general system based on the Fiducial Network concept; and

- the development of an Active Control System (ACS) currently being implemented by the Canadian Geodetic Survey and the Geological Survey of Canada and in some slight variation in scope by the Texas Department of Highways and Public Transportation (Merrell, 1986).

5.1.1 The Fiducial Network System

The Fiducial Network System (FNS) development, like the ACS (to be briefly described below), leads to a design that includes three major components: mobile GPS receivers, a network of well known fiducial stations and a central processing facility. The GPS data acquired simultaneously at the mobile and fiducial sites are brought to a central processing facility where the differential observables are reduced, calibrated and analyzed to jointly determine the orbits and the geodetic baselines in the framework of the few fiducial sites whose coordinates are accurately known from VLBI or SLR (Davidson et al., 1985). The need for cm-level baseline accuracies over distances of several thousand kilometers imposes a requirement for a regional network of fiducial sites whose relative positions are known to approximately to 1-3 cm (Dixon, Golombek and Thornton, 1985) which can only be provided by VLBI or SLR and improved (by some factor of 3) tropospheric calibration accuracies. The extent of these and other calibration problems is still the subject of investigations by JPL and several other groups. Obviously, for some applications this requirement may be relaxed somewhat to a few cm considering that the sensitivity of the results to fiducial baseline accuracy is nearly proportional to the length of the baseline to be measured.

The same cm-level requirement in baseline accuracy from a few hours of observations translates into a requirement of a few cm of measurement noise over a few seconds including system noise, uncalibrated atmospheric effects (hence, simultaneous acquisition of the L_1 and L_2 carriers to assist in the ionospheric calibrations, and preferably the L_1 and L_2 code to assist the ambiguity resolution, is a critical support requirement), multipath, and other instrumental errors. Similarly, the requirement of simultaneous tracking of all satellites in view dictates the use of omnidirectional antennas or appropriately designed phase array antennas.

The central processing facility for FNS should be suitably equipped to receive and process all data taken by the fiducial and mobile site receivers, determine precise orbits and post-fit GPS ephemerides self-consistent over extented periods at least at the level of 0.1 ppm or better.

To satisfy the minimal regional geodetic requirements, at least three fiducial sites are required in order to provide substantial support for the determination of orthogonal (i.e. east-west, north-south) baselines. Ideally, the fiducial sites should be sufficiently close to the areas to be surveyed so that adequate simultaneous viewing between mobile and fiducial sites is assured, whereas the mobile sites can lie inside or outside the fiducial network (as long as they are within the effective area of the FNS sites) without suffering noticeable degradation in recoverable baseline accuracies. Theoretically, this range of mobile stations from FNS sites can be of the order of thousands of kilometers due to the high GPS orbits, however, in practical terms the optimal effective range will be somewhat limited by the correlation distance of atmospheric conditions affecting refraction, i.e. typically 200-300 km.

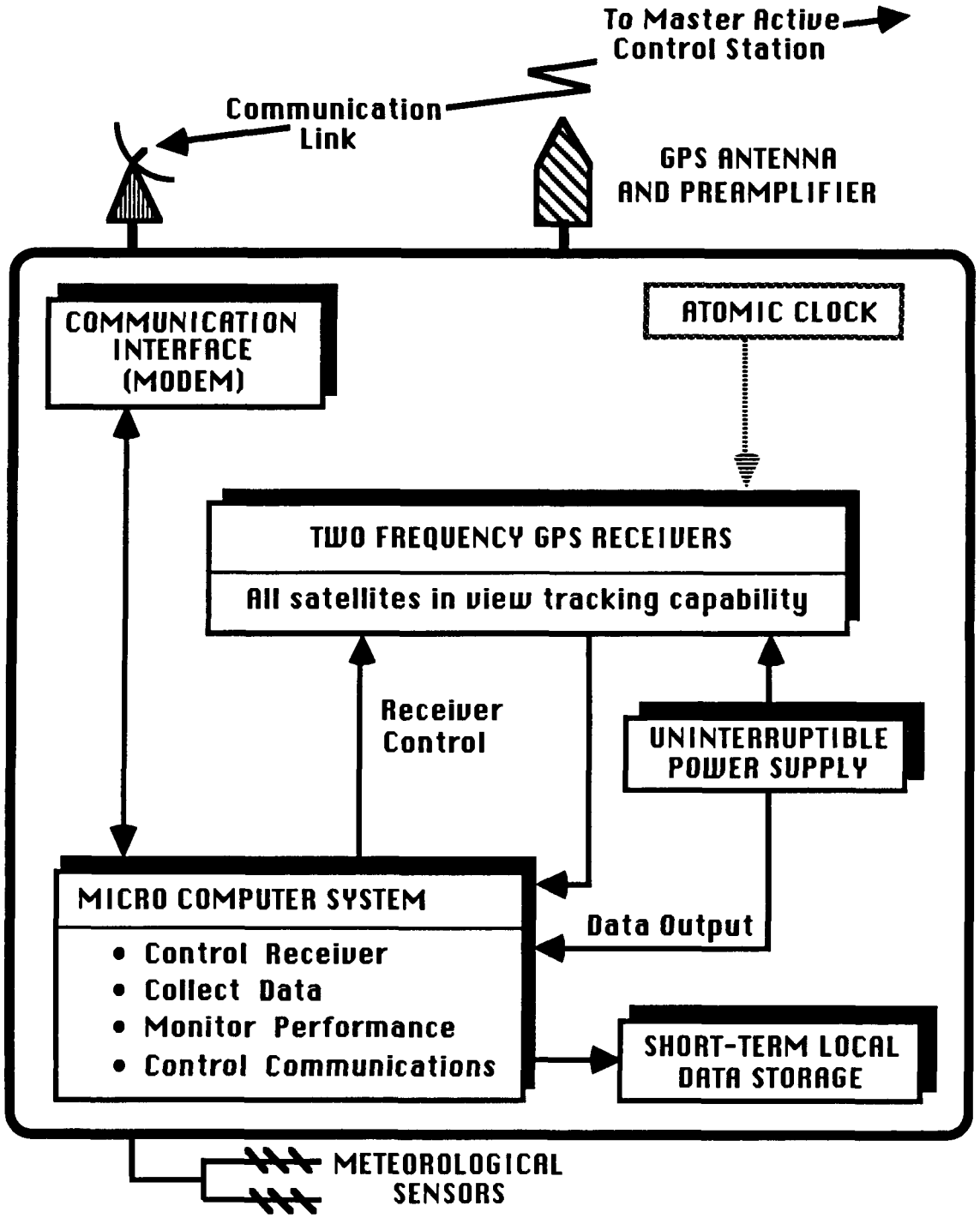

Fig. 5.1 - Schematic representation of the Canadian GPS-based Active Control System station operation.

5.1.2. The Active Control System

The Active Control System (ACS) under development by the Canadian Geodetic Survey (CGS) is based on the concept of a sparse network of control stations where continuous GPS tracking is carried out. A number of these Active Control Points (ACP) are to be permanently collocated at a number of widely spaced, fixed VLBI antennas thus ensuring the highest possible stability and long-term continuity of the ACS framework and consistent scale and orientation of the terrestrial networks positioned by GPS. With continuous tracking capabilities, the ACS will perform four major functions: (a) the data will be used along with similar data collected at stations like the FNS sites in other countries for the computation of precise orbits for distribution to various GPS users; (b) user control network connections based on ACP and user data sets combined. Because the coordinates of the ACS stations will be precisely determined (due to the continuous collection of data in these stations), they will provide precise control without the requirement for GPS users to occupy control stations; (c) broadcasting differential corrections for navigation purposes; and (d) as a means of verifying the integrity of the GPS system. Currently, the VLBI station at Yellowknife is the first such prototype ACP station in Canada in operation, which is also contributing data to the NGS/UoTexas Fiducial Network System between Austin, Texas, Westford, Mass. and Mojave, Ca. At least two new stations are planned to be brought "on-line" as soon as early next year as part of the ongoing developments. For the final ACS system a complement of as many as 12-14 stations are envisaged (Delikaraoglou et al., 1986).

The ACS system is being designed with 4 major features: automatic tracking stations, a tracking system design capable of collecting both code and carrier beat phase data, the capability to generate precise station position estimates with data taken over short periods of time, and the capability of the system to "self-check" its performance.

In Figure 5.1 the major elements of the system are depicted in a functional form. An Active Control Point (ACP) unit would consist mainly of GPS receiving equipment (capable of tracking all visible satellites at the site) and a microcomputer housed in either existing facilities (e.g. provincial Mapping Agencies) or in unattended, environmentally-controlled shelters. Automatic weather sensors and likely a high-precision frequency standard will also be part of the ACP units. Currently, in the prototype stage of the development, TI-4100 receivers are used to evaluate the various functions of the ACP operations. However, the choice of GPS receivers to be eventually deployed in these sites is not planned to be restricted to one particular type.

The planned operations of the ACPs require that all the functions of the system are microcomputer-controlled on the site with most functions such as satellite receiver control, tracking schedules, data acquisition tasks, data management, etc. be accomplished by software residing in the controlling microcomputer without interruption of the receiver tracking functions. Likewise, other housekeeping functions at the ACP site such as tracking status, error diagnostics, station operational status, and recording of weather data from the meteorological sensors will also be carried out by the ACP controller.

The data collected at each ACP site will be appropriately edited, verified and suitably compressed before it is stored locally on disk or transferred via the chosen communication link to the Master Active Control Station (MACS). MACS is essentially the central control node for the ACS network and the control center of the ACS's continuous operations and quality assurance. It is also the facility where computation of

GPS precise orbits would take place. Effective real-time data distribution on a regional and continental basis will depend on cost-effective mobile satellite communication services being available. This is expected early in the 1990's with the use of the new satellite communications system known as the Mobile Satellite (MSAT) which is currently entering its final stage of development and is expected to be fully deployed about the same time as GPS becomes fully operational and the full ACS network is in place (Delikaraoglou, 1986).

5.2 Modes of Analysis for Orbit Determination and Baseline Estimation

The problem of accurately modelling the forces acting on the GPS satellites has been discussed to some extent in Section 3 where it was indicated that forces which contribute to the satellite motion comprise:

- gravitational attractions of the earth;
- gravitational attractions of the sun, moon, and the planets (i.e. the third-body effects);
- the variable part of the earth's gravitational fields arising from the tidal and other deformations of the solid earth and the oceans;
- solar radiation pressure (both direct and albedo effects including the so called Y-bias effect (Fliegel et al., 1985));
- atmospheric drag forces; and
- magnetic forces and other smaller effects.

The central problem in any orbit determination scheme is to accommodate a relatively sophisticated model for the above force fields with the degree of sophistication being, loosely stated, a function of the length of the arcs under consideration. Whereas in theory, as already seen in Section 3, the solution of this determination problem involves a rather complex integration process of the satellite's equations of motion, in practice, it ultimately reduces to an orbit improvement problem using observations which are themselves functions of many satellite positions, ground positions and other nuisance parameters such as satellite and receiver clock errors, atmospheric delay parameters, etc. In the case of the GPS satellites and the carrier beat phase measurements, this usually involves a simultaneous adjustment of:

- satellite epoch states;
- station and satellite clock parameters;
- selected station locations;
- zenith tropospheric delays;
- satellite solar radiation pressure coefficient(s);
- empirical accelerations; and
- carrier phase ambiguities.

Over the past 2 years high-precision orbit determination results for the GPS satellites have been reported (e.g. Abbot et al., 1986; Beutler et al., 1986; Williams, 1986) from which the quality of baseline results have improved by nearly two orders of magnitude over that period; i.e., from several parts in 10^6 in 1984 to between 2 parts in 10^7 and 5 parts in 10^8 today for the longer baselines. In achieving these results, two methods of orbital analysis have been used almost exclusively:

- the *"Fiducial Network"* mode based on the approach just outlined in which the GPS orbits and determined baselines are defined in the framework of a few fiducial stations whose coordinates are accurately known from VLBI or SLR; and

- the *"Free Network"* mode approach (e.g. Beutler, 1985; Delikaraoglou, 1985) in which a combination of station and satellite coordinates are allowed to vary simultaneously in an adjustment by way of imposing relatively strong a priori constraints on the orbits and letting the terrestrial network adjust free.

Fundamental to each of these analysis techniques are also considerations regarding the *length of the satellite arcs* vis-à-vis the *network configuration*. Single-pass arcs (about 1/3 satellite revolution or 4-5 hours duration) require less elaborate force models (cf. Buffett, 1985) and hence, fewer parameters to solve for, whereas longer multi-day arcs (1 day to 1 week duration or more) are preferable in order to define a more stable reference frame, but they require additional care to account for any unmodelled (or improperly modelled) contributions of the forces acting on the satellites. The latter are particularly important for long arcs since, if unaccounted for, they can adversely affect the results in a systematic way thus becoming (along with data noise) the limiting error source. Tracking networks of different scales and geometries allow us to differentiate between orbital and other distance-dependent error sources in GPS such as unmodelled atmospheric delays. To date for instance, the U.S. Department of Defense has relied on its global network of the GPS monitor stations and the use of the pseudorange and integrated Doppler observables to produce predicted (broadcast) orbits at the tens-of-meters accuracy level. However, for the more stringent orbital requirements of precise geodetic and geodynamic studies with GPS, the optimal network configuration for the determination of GPS orbits and station coordinates is not well established yet. Ideally, dense regional and local networks in the area of interest at the scales (i) 100-500 km, (ii) 500-2000 km, and (iii) 2000-4000 km would provide the test ground for isolating possible existence of systematic errors specific to the space/terrestrial conditions at the time of the measurements and thus would allow for a comprehensive study of various orbital analysis techniques, particularly the ones that use increasingly shorter satellite arcs. Results of various comparison tests carried out as part of this study to assess each strategy's practical capability to reduce particular systematic errors and improve the orbit and baseline determinations will be reported later on in this section. Before doing so, however, we will discuss another outstanding item tacitly assumed as part of each of these modes of analysis: the effect of the a priori covariance information on the so-determined orbits and baselines.

5.4 A priori Information in the GPS Satellite Networks

In any adjustment of differential GPS observations the various unknown parameters can conveniently be grouped into three categories:

- station parameters or baseline components;
- satellite epoch state vectors;
- other dynamical satellite parameters like solar radiation pressure, instrumental-type parameters like clock errors and other non-geometric parameters like atmospheric errors.

In practice, provided that enough differential observations of adequate geometry of adequate geometry have been collected, a least-squares adjustment can be carried out in which the primary interest is in the coordinates of the terrestrial points and hence, a solution is often sought from a partitioned system of normal equations of the form

$$\begin{bmatrix} N_{ff} + P_f & N_{fg} & N_{fs} & N_{fb} \\ N_{fm}^T & N_{mm} + P_f & N_{ms} & N_{mb} \\ N_{fs}^T & N_{ms}^T & N_{ss} + P_f & N_{sb} \\ N_{fb}^T & N_{mb}^T & N_{sb}^T & N_{bb} + P_f \end{bmatrix} \times \begin{bmatrix} X_f \\ X_m \\ X_s \\ X_b \end{bmatrix} = \begin{bmatrix} A_f^T P L \\ A_m^T P L \\ A_s^T P L \\ A_b^T P L \end{bmatrix} \qquad (5.1)$$

where

$$A = [\, A_c \mid A_b \,] \quad \ldots\ldots \text{ first partitioning} \qquad (5.2a)$$

$$= [\, A_g \; A_s \mid A_b \,] \quad \ldots\ldots \text{ second partitioning} \qquad (5.2b)$$

$$= [\, A_f \; A_m \mid A_s \mid A_b \,] \quad \ldots\ldots \text{ third partitioning} \qquad (5.2c)$$

is the partitioning of the design matrix A of the network;

$$X = [\, X_c \; X_b \,]^T \quad \ldots\ldots \text{ first partitioning} \qquad (5.3a)$$

$$= [\, X_g \; X_s \;|\; X_b \,]^T \qquad \text{..... second partitioning} \qquad (5.3b)$$

$$= [\, X_f \; X_m \;|\; X_s \;|\; X_b \,]^T \qquad \text{..... third partitioning} \qquad (5.3c)$$

is the corresponding partitioning of the unknown vector into a sub-vector X_c of coordinate unknowns (both ground X_g and satellite X_s) and a sub-vector X_b of nuisance parameters (non-geometric, instrumental, etc); X_g is further partitioned in a set of coordinate unknowns X_f corresponding to fixed or constrained (i.e. fiducial) stations and X_m corresponds to mobile or new stations to be surveyed; N_{ij} are, respectively, the partitioned parts of the normal equations pertaining to the partition of X; $P_i \equiv C^{-1}(X_i)$ are similarly weight matrices corresponding to the a priori information (i.e. $C^{-1}(X_{o,i})$ the a priori covariance matrices) available on the parameters $X_{o,i}$ assuming that there is no correlation between the different sets of parameters.

The correspondence of this partitioning scheme to the aforementioned two main modes of analysis is now readily apparent. The "fiducial network" mode corresponds to the case whereby $0 < P_s < \infty$, $P_f \rightarrow \infty$ and $0 < P_m < \infty$. Respectively, the "free network" mode case corresponds to $0 < P_s < \infty$, $0 < P_f < \infty$ and $0 < P_m < \infty$ whereby furthermore, the first two sets of conditions of a priori information are necessary since they generally reduce the datum defect inherent in the free network adjustment (Delikaraoglou, 1985). Generally there are a number of practical consequences arising from the a priori information on either the satellite parameters or the station coordinates of some of the stations and the effect which they may have on the variances of the coordinates of the rest of the determined stations and the baseline components between such stations. These can be readily understood by examination of the more general form of the normal equations (5.1).

It can be shown that starting from the second partitioning of the parameter vector X, and after successive elimination of X_b and then X_s and some lengthy derivations, the post-fit covariance matrix of the station coordinates X_g including a priori information is given by

$$\mathbf{C}(X_g) =$$

$$= \left| \begin{array}{c:c} (N_{ff} + P_f) - N_{fs}(N_{ss} + P_s)^{-1} N_{fs}^T & N_{fm} - N_{fs}(N_{ss} + P_s)^{-1} N_{ms}^T \\ \hdashline N_{fm}^T - N_{ms}(N_{ss} + P_s)^{-1} N_{fs}^T & N_{mm} - N_{ms}(N_{ss} + P_s)^{-1} N_{ms}^T \end{array} \right|$$

$$= \begin{vmatrix} B_{11} & B_{12} \\ B_{21}^T & B_{22} \end{vmatrix}^{-1} = \begin{vmatrix} Q_{11} & Q_{12} \\ Q_{21}^T & Q_{22} \end{vmatrix} \qquad (5.4)$$

where the submatrices Q_{ij} are given by

$$Q_{11} = [\, B_{11} - B_{12}\,(B_{22})^{-1}\,(B_{12})^T\,]^{-1}$$

$$= [\, [(\, N_{ff} + P_f\,) - N_{fs}\,(\, N_{ss} + P_s\,)^{-1}\,(N_{fs})^T\,]$$

$$- \{(\, N_{fm} - N_{fs}\,(\, N_{ss} + P_s\,)^{-1}\,(N_{ms})^T\,\}$$

$$\times [\, N_{mm} - N_{ms}\,(\, N_{ss} + P_s\,)^{-1}\,(N_{ms})^T\,]^{-1}$$

$$\times \{(N_{fm})^T - N_{ms}\,(\, N_{ss} + P_s\,)^{-1}\,(N_{fs})^T\,\}\,]^{-1} \qquad (5.5a)$$

$$Q_{22} = [\, (B_{22})^{-1} - (B_{22})^{-1}(B_{12})^T\, Q_{11}\, B_{12}(B_{22})^{-1}\,]$$

$$= [\, B_{22} - (B_{12})^T\,(B_{11})^{-1}\,B_{12}\,]^{-1}$$

$$= [\, [(\, N_{mm}\,) - N_{ms}\,(\, N_{ss} + P_s\,)^{-1}\,(N_{ms})^T\,]$$

$$- \{(\, N_{fm})^T - N_{ms}\,(\, N_{ss} + P_s\,)^{-1}\,(N_{fs})^T\,\}$$

$$\times [\, (N_{ff} + P_f\,) - N_{fs}\,(\, N_{ss} + P_s\,)^{-1}\,(N_{fs})^T\,]^{-1}$$

$$\times \{\, N_{fm} - N_{fs}\,(\, N_{ss} + P_s\,)^{-1}\,(N_{ms})^T\,\}\,]^{-1} \qquad (5.5b)$$

$$Q_{12} = -\, Q_{11}\, B_{12}\,(B_{22})^{-1}$$

$$= [\, B_{22} - (B_{12})^T\,(B_{11})^{-1}\,B_{12}\,]^{-1}$$

$$= -\, Q_{11}\,[\, N_{fm} - N_{fs}\,(\, N_{ss} + P_s\,)^{-1}\,(N_{ms})^T\,]$$

$$\times [\, (N_{mm} - N_{ms}\,(\, N_{ss} + P_s\,)^{-1}\,(N_{ms})^T\,]^{-1} \qquad (5.5c)$$

$$Q_{21} = (Q_{12}\,)^T\ . \qquad (5.5d)$$

The covariance matrix for the baseline components between stations in the f- and m-sets respectively are readily obtained by a linear transformation:

$$C(B_{f,m}) = \begin{vmatrix} -I & I \end{vmatrix} \begin{vmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{vmatrix} \begin{vmatrix} -I \\ I \end{vmatrix} \qquad (5.6a)$$

$$= Q_{11} + Q_{22} - Q_{12} - (Q_{12})^T \qquad (5.6b)$$

$$= Q_{11} + Q_{22} + Q_{11}\, B_{12}\, (B_{22})^{-1} + (B_{22})^{-1}\, (B_{12})^T\, Q_{11}. \qquad (5.6c)$$

The effect of a priori information can now be readily seen for the following cases:

(a) the satellite parameters are constrained and there is no knowledge on the coordinates of the f-set of stations ($P_s \rightarrow \infty$, $P_f \rightarrow 0$), e.g. in a case of reducing GPS data at completely unknown stations with orbits assumed completely known and held fixed in the adjustment. Under these assumptions equations (5.4) through (5.6) reduce to

$$C(X_f) = [\, N_{ff} - N_{fm}(N_{mm})^{-1}(N_{fm})^T \,]^{-1} \qquad (5.7a)$$

$$C(X_m) = [\, N_{mm} - (N_{fm})^T(N_{ff})^{-1}(N_{fm}) \,]^{-1} \qquad (5.7b)$$

$$C(B_{f,m}) = C(X_f) + C(X_m) + C(X_f)\, (N_{fm})^T\, (N_{mm})^{-1} + (N_{mm})^{-1}\, (N_{fm})^T\, C(X_f); \qquad (5.7c)$$

(b) the satellite parameters are partially known and the coordinates of the f-set of stations constrained ($0 < P_s < \infty$, $P_f \rightarrow \infty$), e.g. in the case of reducing GPS data from fiducial stations for the determination of orbits along with the coordinates of unknown stations (m-set of stations)

$$C(X_f) = 0 \qquad (5.8a)$$

$$C(X_m) = [\, N_{mm} - N_{ms}(N_{ss} + P_s)^{-1}(N_{ms})^T \,]^{-1} \qquad (5.8b)$$

$$C(B_{f,m}) = [\, N_{mm} - N_{ms}(N_{ss} + P_s)^{-1}(N_{ms})^T \,]^{-1} = C(X_m) \; ; \qquad (5.8c)$$

(c) satellite parameters and the coordinates of the f-set of stations constrained ($P_s \rightarrow \infty$, $P_f \rightarrow \infty$), e.g. the case of reducing GPS data between fiducial and unknown stations and holding the orbits (as determined from the fiducial stations, for instance, in a previous run) fixed during the determination of the coordinates of the unknown stations (*m-set of stations*)

$$C(X_f) = 0 \qquad (5.9a)$$

$$C(X_m) = (N_{mm})^{-1} \qquad (5.9b)$$

$$C(B_{f,m}) = (N_{mm})^{-1} = C(X_m) \; ; \qquad (5.9c)$$

(d) satellite parameters and the coordinates of the f-set of stations partially known ($0 < P_s < \infty$, $0 < P_f < \infty$), e.g. the case of reducing GPS data in the free network mode by imposing relatively strong constraints on the orbits and some stations (i.e. the f-set of stations) and letting the rest of the network (i.e. the m-set of stations) free to adjust

$C(X_f)$ = equation (5.5a)

$C(X_m)$ = equation (5.5b)

$C(B_{f,m})$ = equation (5.6).

A careful examination of these covariance expressions leads to some useful insight as to how the results may vary under different circumstances of usage of *a priori* information. For instance, from (a) and (c) it is readily seen that when the satellite orbits are assumed known, the uncertainty of the coordinates of new surveyed stations increases as the uncertainty of the fiducial stations increases. A similar increase is also to be expected in the new stations when the coordinates of the fiducial stations are constrained and the uncertainty in the satellite orbits increases; i.e., case (c) vis-à-vis case (b) and a further increase is to be expected when both the uncertainty in the satellite orbits and the coordinates of the f-set of stations increases; i.e., case (b) vis-à-vis case (d). The last observation indicates a significant weakness, and as it turns out from a practical viewpoint, a problematic area of the free network mode of analysis vis-à-vis the fiducial network mode: unless appropriate, realistic orbital constraints are chosen for the former, variable results in the solution are to be expected. This fact alone suggests that the free network mode should be an unlikely candidate to use for precise geodynamic applications where the utmost certainty in the results is a must.

For the baseline components, the examination of the covariance expressions (5.8c) and (5.9c) indicates that given that the coordinates of fiducial stations are often assumed known, increasing the uncertainty in the satellite orbits leads to an increased uncertainty in the components of the baselines connecting fiducial and non-fiducial stations; also if the satellite orbits are assumed known, increasing the uncertainty of the

fiducial stations also increases the uncertainty in the components of the baselines between fiducial and non-fiducial stations (cf. Eq. (5.7c) vs. Eq. (5.9c)).

5.4 Experimental Results

The methodology described thus far has been implemented in software and tested with various GPS data sets. In this brief overview, we shall report on representative tests relating to different aspects of orbit determination and the results obtained with sample data available from two such GPS campaigns designed to deploy a large number and distribution of receivers (Fig. 5.2). The first campaign took place in the Spring of 1985 and involved TI-4100 receivers (Henson et al., 1985) deployed at some ten sites in Southern California, JPL Series-X receivers (Crow et al., 1984) operating at Mojave and Owens Valley in California, and prototype Macrometer-II receivers (Ladd and Councelman, 1985) operating at the three POLARIS sites at Haystack, MA, Richmond, FL and Ft. Davis, TX. Six days of data (days 90-95) from this campaign

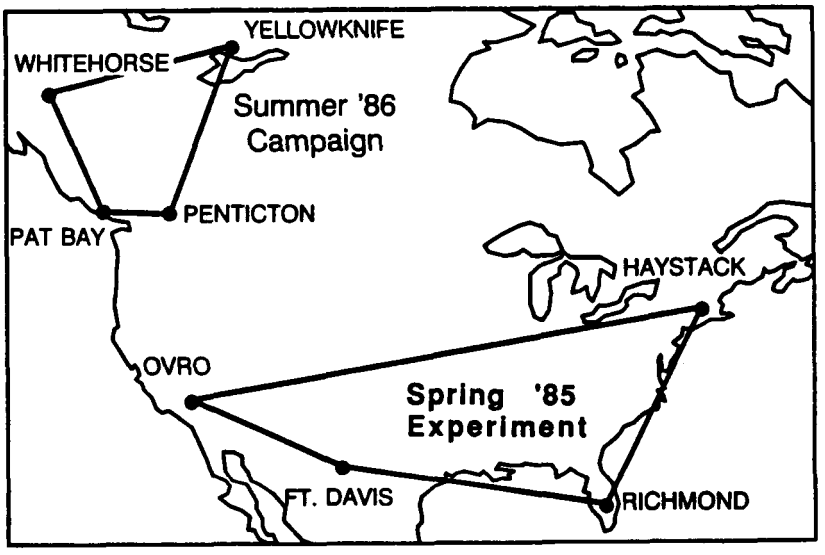

Fig. 5.2 - Site locations for the Spring '85 and Summer '86 GPS Experiments

were used in our tests. The second campaign took place in the Summer of 1986 for the main purpose of establishing first epoch measurements in Iceland, but it also involved observations from GPS receivers providing simultaneous coverage of the northern hemisphere in Canada, USA, Sweden and Denmark. For the purpose of this report, results with only a 2-day portion of the data (days 193-194) collected in Northwestern Canada during this campaign will be reported.

Most of the analyses reported herein were carried out using a modified version of NASA's GEODYN precision orbit determination program and software developed for the preprocessing of the raw GPS data (including formatting, editing, calibration and cycle-slip detection). GEODYN provided integrated trajectories in which the earth's gravity field model was expressed by spherical harmonics (up to degree and order 8 for the case of the GPS satellites in accordance with the discussion in Section 3); point mass attractions from the sun and moon and a direct solar radiation pressure model (i.e., a constant acceleration along the sun-satellite vector) were included, plus the model of empirical accelerations developed by Colombo (1986), i.e. Eq.(3.36).

The observations used to compute the orbits for the Spring '85 Experiment were from Macrometer data collected at the three POLARIS sites. A copy of these observations was transferred already "cleaned" from cycle slips with the compliments of the MIT GPS group. Data from the TI-4100 receivers at Owens Valley, Mojave, Ft. Davis and Hat Creek were pre-processed at GSFC mostly with software developed in the course of this study. Similarly, TI-4100 data from the Summer '86 GPS campaign were transferred from CGS and pre-processed at GSFC for editing, phase connection, data compression and atmospheric calibrations.

The observables used in our analyses were basically "between-stations" single differences. Dual-frequency L_1 and L_2 observations were used throughout. Since this type of observable is affected by receiver clock errors, the clock behaviour was modelled as "white noise"; i.e., one common clock bias was solved for per measurement epoch for stations with no external frequency standards; this is effectively equivalent to double differencing but it requires many more bias parameters-an actual double differencing scheme is currently implemented in GEODYN which hopefully will eliminate the constraints imposed by this equivalent "white noise" model in the future.

A number of tests were carried out to investigate the differences between the fiducial concept and the free network approach in both the orbit determination and relative positioning with GPS vis-à-vis the impact on these determinations of the length of the satellite arcs and the variations in the tracking network configurations. For this purpose identical data sets were processed in each of these modes and the results were compared.

Spring '85 Experiment tests

For the fiducial network strategy, typically, the coordinates of the stations at Haystack, Richmond and Ft. Davis were held fixed at their VLBI values, whereas the coordinates at Owens Valley were adjusted along with the satellite orbits, clock biases, zenith tropospheric delays, etc. This configuration provided a well-distributed set of points to test the orbit quality and baseline repeatabilities over distances of 1000-4000 km. Descriptions of various results obtained from the analysis of this data set were also reported by Pavlis et al (1987) and Colombo and Zelensky (1987).

For the free network approach, Haystack was held fixed and the coordinates of all the other stations were estimated along with the satellite orbits. As already mentioned, this approach requires that some a priori (variance/covariance) information about the orbits be introduced into the normal equations of the adjustment process. In the results to be reported below, a priori rms errors for the state vectors were chosen at the rather

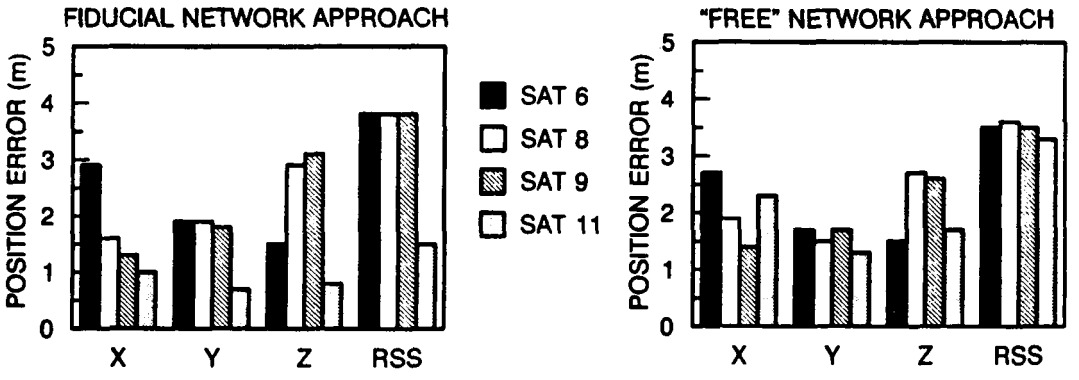

Fig. 5.3 - Orbit precision with 6-day multi-arc solutions using TI-4100 data from the March '85 Experiment

extreme values of 20 m in position and 0.2 m/sec in velocity. Several tests in the large domain 2-20 m and 0.2-2 m/sec were also carried out with variable results depending on the imposed constraints. As already pointed out, this is currently one of the most problematic areas of this approach, that is to say, to come up with realistic estimates of the imposed constraints on the orbits. This also verified the practical problem hinted in the examination of the covariance expressions in the previous section and confirmed the adverse effect in the solution results under different circumstances of a priori information.

Fig. 5.3 shows plots of formal orbit errors from a 6-day solution for the GPS satellites 6, 8, 9 and 11 indicating the level of precision of the computed orbits. GPS 11 was a well-tracked satellite (i.e. with good temporal coverage) during this period, whereas the other satellites were more sparsely tracked each day. As the figure indicates, in the fiducial network approach orbit precision of better than 1 m in all position components is achievable from multi-day arcs for a well-tracked satellite, and 2-3 m for sparsely tracked ones. By comparison, for the free network approach, the formal errors are consistently higher for GPS 11 in spite of its good tracking coverage, whereas they are quite comparable for the other three satellites (somewhat worse results were obtained in another test whereby Ft. Davis was held fixed instead of Haystack in order to test the sensitivity of this approach to the choice of the fixed station).

Since in general, formal errors would tend to underestimate orbit accuracy, a comparison between solutions of two overlapping arcs of different length were made in which the recovery of the position and orientation of Owens Valley as determined with respect to the three POLARIS sites was compared. The discrepancy vector obtained in each case was resolved for an easier comparison into a coordinate system whose axes are aligned

- along the true baseline (in this case the VLBI baseline) from the fixed to the adjusted station;
- along the axis in the direction of increasing azimuth for the baseline; and
- along the axis in the direction of increasing elevation for the baseline (which is almost identical to the local geodetic height component).

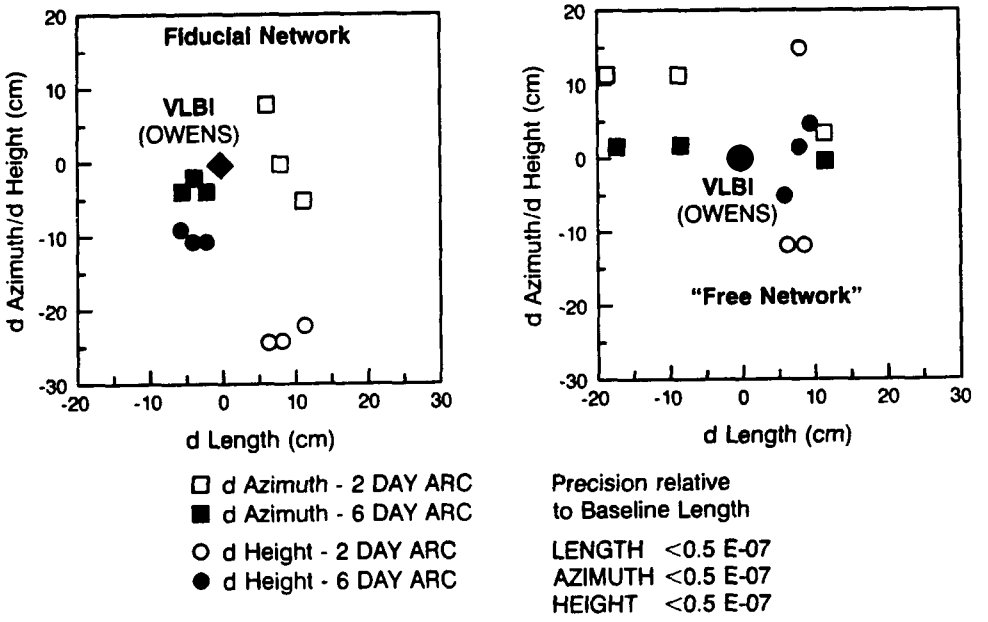

Fig. 5.4 - GPS vs. VLBI at Owens (non-fiducial station) illustrating the effect of variation of the arc length

Fig. 5.4 shows the results of this comparison for a 2- (days 90-91) vs. a 6-day arc (days 90-95) in each of the fiducial and free network modes respectively. In the fiducial network case, the azimuth discrepancies of the three baselines Haystack/Owens, Richmond/Owens, and Ft. Davis/Owens, as well as the discrepancies in their lengths (3929, 3741 and 1508 km respectively) are all within 10 cm for the 2-day arc, whereas the elevation discrepancies are all at the 25-cm level with respect to the VLBI values at Owens. The corresponding azimuth discrepancies in the free network case are also within 10 cm, whereas the baseline length and elevation discrepancies show a reverse tendency compared to the previous case. That is, elevation discrepancies were reduced to the 15-cm level, whereas baseline discrepancies up to 20 cm were noted. As may have been expected, there was a very noticeable improvement in all cases when the 6-day arc results were compared. The fiducial network results were for all baselines within 5 cm in azimuth and length, and 10 cm in elevation which corresponds to better than 1 part in 10^8 relative to the length of these baselines. Considering that the accuracy of the GPS orbits and relative positioning accuracies are related by the rule of thumb

$$(dr/r) = (dB/B)$$

where B is the baseline length, r is the average range to the satellite (about 26,000 km), dB is the baseline error and dr is the orbit error, a 10-cm length discrepancy (as noted in the fiducial network case) over the 3000-km average length of these baselines would infer an orbit error of 0.8 m. Similarly a discrepancy of 20 cm (as noted in the free network case) would correspond to 1.7 m orbit error. These values may give a more representative measure of the orbit accuracy than would formal errors and would

indicate what one could expect from multi-day arcs for well-tracked satellites in each case.

In order to assess further the quality of the so-computed GPS orbits, a set of daily solutions was carried out to test baseline repeatability. Using TI-4100 data, the station locations at Mojave and Hat Creek with respect to Owens Valley were independently computed each day. The orbits used in these solutions were those computed from the 6-day arcs using Macrometer data from the three POLARIS sites and Owens Valley and they were held fixed in these tests, so these comparisons provided another test not only of the accuracy of the orbits but also of their robustness, and of the models used to provide the integrated trajectories. The noted differences from the VLBI values are shown in Fig. 5.5a and 5.5b. Four daily solutions and their respective means and associated standard deviations are shown for the east, north and vertical directions and the corresponding length of each baseline is also shown. In the case of Mojave, the location of the SLR monument is also indicated as a further measure of comparison. In the case of the Owens-Mojave (245 km) baseline the standard errors of the mean in all components except the height were better than 5 cm which corresponds to a precision better than 2 parts in 10^7.

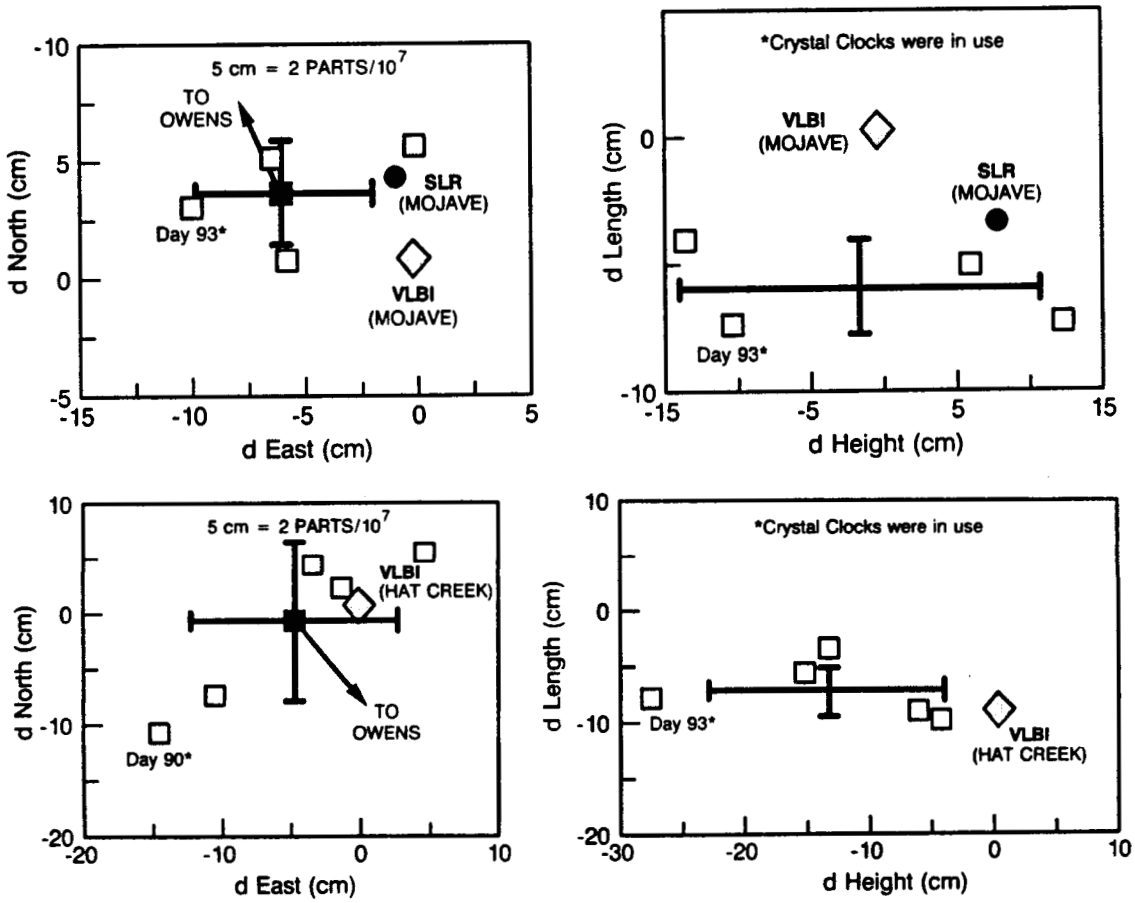

Fig. 5.5 - Day-to-day repeatability (a) above, Owens-Mojave (245 km) (b) below, Owens-Hat Creek (484 km); 6-day orbits (days 90-95) kept fixed in daily solutions.

The higher scatter observed in the respective height differences corresponds to a precision of 5 parts in 10^7 which is probably due to the troposphere. It would appear that for the shorter baselines, tropospheric delay fluctuations may be affecting repeatability more than is the case with the continental baselines where orbital inaccuracies may be the more dominant source of error. From the North/East plot, it is interesting to note that there seems to be a great deal of correlation between the noted differences and the direction to Owens which may be indicative of the geometric weakness of the daily observability windows for this baseline. Similar results were obtained for the daily solutions carried out for the Owens-Hat Creek (484 km) baseline. Again all components, including the height, showed a scatter indicating a precision at the level of 2 parts in 10^7. A similar correlation of the discrepancies in the north and east components and the direction to Owens (note that Hat Creek, Owens and Mojave are nearly along the same line) is also noticeable.

Summer '86 Northwestern Canada Campaign tests

The tests performed thus far were with a valuable body of data available from sites with well-known coordinates from VLBI and from several days of observations. There are situations however, when none of these conditions are fulfilled. In such a situation the free network approach can be proven extremely useful. This was the situation, for instance, with the 2 days of data collected at the regional network in Northwestern Canada in the Summer of 1986. Although Yellowknife, Whitehorse and Penticton are VLBI sites in Canada, the connections of the GPS markers to their VLBI counterparts were not available which rendered their use as fiducial locations useless for the purposes of our tests. It was possible however to conduct further tests using the free network approach. Here we just give an outline-without presenting complete results (which are to be reported in a subsequent detailed report)-of the evaluation of this data set using orbits computed with initial conditions derived from the broadcast ephemerides.

On days 193-194 (July 11-12) four TI-4100 receivers were operated by CGS at the above sites and the Pacific Geoscience Center (Pat Bay, Sidney, B.C.). Satellites 3, 6, 8, 9, 11, 12 and 13 were observed on both days for 3-5 hours each night. In addition to these satellite passes these satellites were observed 12 hours earlier, but this data was not incorporated in these tests. In summary, our processing consisted of the following steps:

- We used the broadcast ephemerides to obtain the initial conditions needed in the CGS software (in which similar models to the ones used in GEODYN have been implemented) to integrate the orbits;
- In these integrated trajectories a force model similar to the one used in the previous tests was used, except for the solar radiation model which included only one radiation pressure coefficient per day which was considered adequate for the short arcs in these tests;
- The L_1 and L_2 observations were combined to form the ionospheric-free linear combination;
- Clock errors were similarly treated as "white noise"; and

- The observations were processed individually for each day using the following options:

(1) the coordinates of Yellowknife were constrained to their VLBI values with a priori variances of 1 m^2 per coordinate;
(2) the coordinates of Whitehorse, Penticton and Pat Bay were treated as unknowns; and

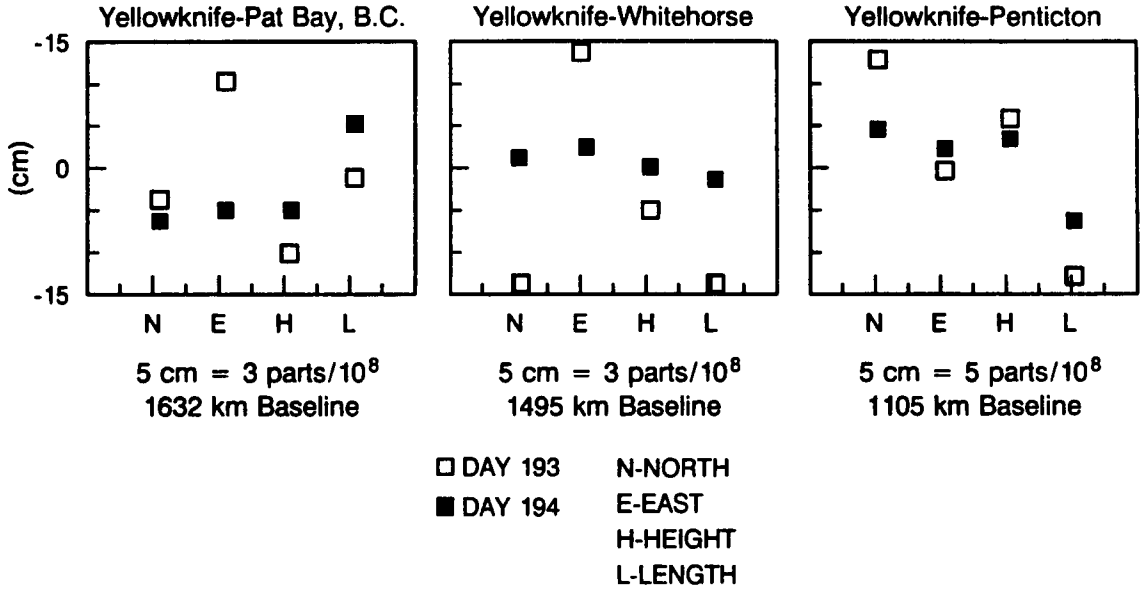

Fig. 5.6 - Baseline repeatability using the "free" network approach with single pass, short-arcs for satellites 6, 8, 9, 11.

(3) satellite state vectors for each satellite arc were computed using constraints of 20 m in the Keplerian elements of each arc expect for the argument of the perigee value which were held fixed.

It should be noted in this connection that in contrast to the multi-day arcs, the geometry is very limited in such arcs of a few hours' duration. In this particular network-satellite configuration, for instance, there was considerable lack of data from the northwest direction which certainly compromised viewing geometry. Fig. 5.6 shows the baseline repeatability results using these single-pass arcs. The rms scatter of the two daily solutions of the components of the three baseline vectors (Yellowknife-Penticton/1105 km, Yellowknife-Whitehorse/1495 km, and Yellowknife-Pat Bay/ 1632 km) about their weighted means were about 5 cm on the average and less than 15 cm in the extreme cases noted in the baseline between Yellowknife and Penticton. Nevertheless these values are of the order of 5 parts in 10^8 which would suggest an orbit error of 1.3 for these short arc orbits.

SECTION 6

SUMMARY AND DIRECTIONS FOR FUTURE WORK

We have demonstrated, albeit with a limited data set, that dual-frequency GPS observations and interferometric-type analysis techniques make possible the modelling of the GPS orbits for several days with an accuracy of a few meters. The use of VLBI or SLR sites as fiducial stations together with refinements in the orbit determination procedures can and have greatly reduced the systematic errors in the GPS satellite orbits used to compute the positions of non-fiducial locations. In general, repeatability and comparison with VLBI of so-determined locations are of the order of 2 parts in 10^7 to 5 parts in 10^8 for baseline lengths < 2000 km. The free network approach offers a powerful alternative technique capable of achieving nearly equivalent results for less than 2-day or shorter arcs (with somewhat worse results for longer arcs) particularly under conditions where one has no access to precise orbits or a network of VLBI/SLR fiducial stations in the vicinity of the survey. However, the requirement of having to define appropriate constraints needed to be imposed in the estimation process in this latter case is a drawback of this approach worth bearing in mind.

Multi-day arcs (3-6 days) generally provide better orbits and baselines than single-day passes or up to 2-day arcs. In achieving these levels of repeatability and accuracy several refinements of the orbit error models are necessary. To that extent, Colombo's (1986) general force model based on the likely resonant character of the GPS orbit errors seems capable of removing the main unmodelled (or incorrectly modelled) forces like solar radiation pressure, Y-bias, etc. thus confirming in practice the theoretical considerations that lead to its development.

Considering the results of this analysis and similar results reported by several other groups, it seems that achieving high-quality orbits for the GPS satellites is, to date, a lesser problem than other error sources. From the results obtained and the experience gained so far it is clear that of the limiting factors to further orbit and baseline accuracy improvement, fiducial station coordinate errors, better handling of propagation medium errors (particularly of the wet troposphere), and further modelling of the orbital dynamics for the GPS satellites are areas to concentrate on for future work. All these errors can be better understood and reduced via tracking networks of various scales and geometries which would allow us to differentiate between the various error sources. The obvious question now is what levels of repeatability can be achieved with GPS by repeated occupations of the same sites at frequent intervals; the only reason for not having an answer yet is that not enough repeat experiments (other than the three occupations of some California sites during the Spring and Fall '85 and again during the Summer '86 campaigns) have been carried out so far. To date, several national and international campaigns (many of them NASA sponsored) have taken place and more are being planned for the future. Yet GPS work seems set to continue at many centers, in a largely decentralized way unless the data collected during these campaigns are shared for the purpose of carrying out such studies that hopefully will offer a unique opportunity to study these aspects further.

In retrospect, this report which is mainly a synthesis of problems, assumptions, methods and recent advances in the studies towards the establishment of a GPS-based system for geodesy and geodynamics marks only one phase in the continuing effort for the development of such a system. At the outset of this project, the physical phenomena relevant to the GPS carrier phase observable which does not require access to the (likely to be restricted in the future) coded information in the satellite signals, and general orbit force models were reviewed. Models for these phenomena were described for implementation in software, a large part of which was developed and tested in-house at GSFC. This activity, however, is by no means complete. Several enhancements and improvements are still possible especially if these models (and the software modules now available at GSFC) are to be eventually implemented in packages like NASA's GEODYN II. Some of these include the full implementation of the double-difference module; the 6-parameter solar radiation general force model that does not require specific physical information on the satellites (e.g., panel shape, dimensions, material, etc.) and yet as demonstrated here is as good or better than those models that rely on the restricted information; adjustment strategies for simultaneous orbit and station position estimation; representation of refraction and clock parameters with time-varying polynomials or continuous stochastic processes of given correlation functions, and ability to write out residuals and selected partials for further processing outside GEODYN. Finally, it should be mentioned that although this study and the results reported here have relied on the use of the carrier beat phase observable, there might be good reason to experiment also with the synergistic mix of the carrier phase with the P-code pseudorange observable. The activity at JPL and recent results presented at the October 1987 Crustal Dynamics Meeting at NASA's Goddard Space Flight Center have given positive indications that further improvements in precision can be achieved by the combination of the two. In any event, the sole fact remains that GPS results already agree impressively well with collocated VLBI results. To some, including the author, it now seems reasonable to expect within the next few years that more evidence will show GPS to be as powerful and reliable a tool as mobile VLBI and SLR are today, but largely more economical. Promises of pocket-sized receivers and wallet-sized prices in the not-too-distant future may well prove that GPS is indeed out of this world!

References

Abbot, R.I., Y. Bock, C.C. Councelman and R.W. King (1986) - "GPS orbit determination", Proc. Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

AGU (1985) - "LAGEOS - Scientific results", reprinted from Journal of Geophysical Research, Vol. 90 (B11), Sept.

Allan, R.R. (1965) - "On the motion of nearly synchronous satellites", Royal Aircraft Establishment, Farnborough, Hants.

ARL (1986) - Proc. Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX, Vol. I & II, Applied Research Lab., Univ. of Texas.

Ashkenazi, V. and J. Yau (1986) - "Significance of discrepancies in the processing of GPS data with different algorithms", Bull. Geodesique 60, pp. 229-239.

Baker, P. (1986) - "GPS Policy", Proc. Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

Bender, P.L. and D.R. Larden (1985) - "GPS carrier phase ambiguity resolution over long baselines", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Beutler, G., W. Gurtner, I. Bauersima and R. Langley (1985) - "Modelling and estimating the orbits of the GPS satellites", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Beutler, G., W. Gurtner, M. Rothacher, T. Schildknect and I. Bauersima (1986) - "Determination of GPS orbits using double difference carrier phase observations from regional networks", Proc. Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

Beutler, G., I. Bauersima, W. Gurtner, M. Rothacher and T. Schildknect (1987) - "Atmospheric refraction and other important biases in GPS carrier phase observations", presented at the XIX General Assembly of the IUGG, Vancouver, Canada, Aug. 9-22.

Bletzacker, F.R. (1985) - "Reduction of multipath contamination in a geodetic GPS receiver", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Bock, Y., S.A. Gourevitch, C.C. Councelman III, R.W. King and R.I. Abbot (1986) - "Interferometric analysis of GPS phase observations", Manuscripta Geodetica.
Brouwer, D. and G.M. Clemence (1961) - "Methods of Celestial Mechanics", Academic Press, New York.

Buffett, B.A. (1985) - "A short-arc orbit determination for the Global Positioning System" Univ. of Calgary Report 20013.

Buisson, J. (1987) - "The Advanced Clock/Range Experiment", presented at the NASA Crustal Dynamics Meeting, GSFC, Oct. 21-23.

Clark, T.A., B.E. Corey, J.L. Davis and G. Elgered (1985) - "Precision geodesy using the Mark-III Very Long Baseline Interferometer System", in IEEE Transactions on Geoscience and Remote Sensing, Special issue on satellite geodynamics, Vol. GE-23 (4), July.

Clynch, J.R. and D.S. Coco (1986) - "Error characteristics of high quality geodetic GPS measurements: clocks, orbits and propagation effects", Proc. Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

Cohen, S.C., J.J. Degnan, J.L. Bufton, J.B. Garvin and J.B. Abshire (1986) - "The Geoscience Laser Altimetry/Ranging System (GLARS)", NASA TM 87803, Sept.

Colombo, O.L. (1984a) - "Altimetry, Orbits and Tides", NASA TM 86180, GSFC, Greenbelt, MD.

Colombo, O.L. (1984b) - "The Global Mapping of gravity with two satellites", Netherlands Geodetic Commission, Publ. of Geodesy, Vol. 7 (3).

Colombo, O.L. (1986) - "Ephemeris errors of GPS satellites", Bulletin Geodesique 60, pp. 64-88.

Colombo, O.L. and N. Zelensky (1987) - "Precise determinations of GPS orbits and of station positions", Spring American Geophysical Union Meeting, Baltimore, May.

Crow, R.B., F.R. Blerzacker, R.J. Najarian, G.H. Purcel, J.I. Statman and B.J. Thomas (1984) - "SERIES-X Final engineering report", Publ. D-1476 Jet Propulsion Laboratory, Pasadena.

Davidson, J.M., C.L. Thorton, C.J. Vegos, L.E. Young and T.P. Yunck (1985) - "March 1985 demonstration of the fiducial network concept for GPS Geodesy: a preliminary demonstration", First Inter'l Symp. on Precise Positioning with GPS Rockville, MD.

Davidson, J.M. and D.W. Trask (1985) - "Utilization of Mobile VLBI for geodetic measurements", in IEEE Transactions on Geoscience and Remote Sensing, Special issue on satellite geodynamics, Vol. GE-23 (4), July.

Degnan, J.J. (1985) - Satellite Laser Ranging: current status and future prospects", in IEEE Transactions on Geoscience and Remote Sensing, Special issue on satellite geodynamics, Vol. GE-23 (4), July.

Delikaraoglou, D. (1985) - "Estimability analyses of the free networks of differential range observations to GPS satellites", in 'Optimization of geodetic networks', E. Grafarend and F. Sanso (Ed), Springer - Verlag, ISBN 0-387-15739-5, New York and Berlin.

Delikaraoglou, D. and R.R. Steeves (1985) - "The impact of VLBI and GPS on geodesy in Canada", First Inter'l Symp. on Precise Positioning with GPS Rockville, MD.

Delikaraoglou, D., R.R. Steeves and N. Beck (1986) - "Development of a Canadian Active Control System using GPS", Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

Delikaraoglou, D. (1986) - "Report on a Canadian Active Control System using GPS", Proc. of the 42nd Annual Meeting of the Institute of Navigation, Seattle, WA.

Dixon, T.H., M.P. Golombek and C.L. Thornton (1985) - "Constraints on Pacific plate kinematics and dynamics with Global Positioning System measurements", JPL Geodesy and Geodynamics preprint No. 118.

Evans, A.G. (1986) - "Comparison of GPS pseudorange and biased Doppler range measurements to demonstrate signal multipath effects", Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

Fliegel, H.F., W.A. Feess,W.C. Layton and N.W. Rhodus (1985) - "The GPS radiation force model", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Gelb, A. (1974) - "Applied Optimal Estimation", MIT Press, Cambridge, Mass.

Georgiadou, Y and A. Kleusberg (1988) - "On the effect of ionospheric delay on geodetic relative GPS positioning", Manuscripta Geodetica 13, pp. 1-8.

Giacalia, G.E.O. (1973) - "Lunar perturbations on artificial satellites of the earth", Smithsonian Astronomical Observatory, Spec. Rep. 352.

Goad, C.C. (1985) - "Precise relative position determination using GPS carrier phase measurements in a non-differenced mode", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Goad, C.C. (1985) - "An efficient algorithm for the evaluation of inclination and eccentricity functions", Manuscripta Geodetica, 12, 11-15.

Grafarend, E. and V. Müller (1985) - "The critical configuration of satellite networks especially of laser and Doppler type, for planar configurations of terrestrial points", Manuscripta Geodetica 10, pp. 131-152.

Heiskanen, W.A. and H. Moritz (1967) - "Physical Geodesy", W.H. Freeman, San Fransisco.

Henson, D.J., E.A. Collier and K.R. Schneider (1985) - "Geodetic Applications of the Texas Instruments TI-4100 GPS Navigator", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Heroux, P. (1986) - "Evaluation of a sequential method for cycle slip detection using carrier phase time derivatives", Geodetic Survey of Canada, Ottawa, Ontario.

Herring, T.A. (1986) - "Very Long Baseline Interferometry and its contribution to Geodynamics", in "Space Geodesy and Geodynamics", A.J. Anderson and A. Cazenave (Ed.), Academic Press, London, New York.

Hilla, S.A. (1986) - "Processing cycle slips in non-differenced phase data from the Macrometer Model V-1000™ receiver", Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

Hottem, L.D. and G.E. Williams (1985) - "Factors to be considered in the development of specifications for geodetic surveys using related positioning GPS techniques", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Jones, R.H. and Tryone, P.V. (1987) - "Continuous time series models for unequally spaced data applied to modelling atomic clocks", SIAM, Vol. 8 (1).

IEEE (1985) - "Transactions on Geoscience and Remote Sensing", Special issue on satellite geodynamics, Vol. GE-23 (4), July.

ION (1980) - "Global Positioning System: Papers published in Navigation, Vol. I.

ION (1984) - "Global Positioning System: Papers published in Navigation, Vol. II.

Kaula, W.M. (1962) - "Development of the lunar and solar disturbing functions for a close earth satellite", Astron. Journal 67, 300.

Kaula, W.M. (1966) - "Theory of satellite geodesy", Blaisdell Publ., Waltham, Mass.

King, R.W., E.G. Masters, C. Rizos, A. Stolz and J. Collins (1985) - "Surveying with GPS", Monograph 9, School of Surveying, The University of New South Wales, Kensington, N.S.W., Australia.

Kleusberg, A. (1986) - "Ionospheric refraction in geodetic relative GPS positioning", Manuscripta Geodetica.

Kostelecky, J. (1985) - "Recurrence relations for the normalized inclination functions", Bull. Astro. Inst. Czechosl. 36, pp. 242-246.

Kouba, J. (1987) - "GPS capabilities and limitations for geodynamics", EOS, Transactions of the AGU, Vol. 68 (16), p. 286 (abstract only).

Ladd, J.W. and C.C. Councelman (1985) - "The Macrometer II™ Dual-band Interferometric Surveyor", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Lala, P. (1971) - "Semi-analytical theory of solar radiation pressure perturbations of satellite orbits during short time intervals", Bull. Astro. Czech. 19, pp. 233-239.

Langley, R.B., G. Beutler, D. Delikaraoglou, B. Nickerson, R. Santerre, P. Vanicek and D.E. Wells (1984) - "Studies in the application of the Global Positioning System to differential positioning", Tech. Rep. 108, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B.

Lieske, J.H., T. Lederle, W. Fricke and B. Morando (1977) - Expressions for the precession quantities based upon the IAU (1976) System of Astronomic Constants", Astron. Astrophysics, 58, 1-16.

Lindlohr, W. and D.E. Wells (1985) - "GPS design using undifferenced carrier beat phase observations", Manuscripta Geodetica, Vol. 10, pp. 255-295.

Merrell, R.L. (1986) - "Application of GPS for transportation related engineering surveys", Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

Moritz, H. and Mueller, I.I. (1986) - "Earth rotation: theory and observation", Uncar/ Continuum, New York.

Murphy, J. and T. Felsentreger (1966) - "Solar Radiation Pressure", in "The National Geodetic Program", part I, NASA SP-365, pp. 354-357.

NOAA (1985) - "Positioning with GPS", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Pavlis, E.C., D. Delikaraoglou and O.L. Colombo (1987) - "Geodetic results of the analysis of GPS data from the Spring '85 Experiment", Spring American Geophysical Union Meeting, Baltimore, May.

Pearlman, M.R. (1982) - "Some current issues in Satellite Laser Ranging", presented at the 4th Inter'l workshop on Laser Instrumentation, Univ. Texas at Austin.

Remondi, B. (1984) - "Using the Global Positioning System (GPS) phase observable for relative geodesy: modelling, processing and results", Ph.D. dissertation, Center for Space Research Rep. CSR-82-2, The University of Texas at Austin, TX.

Remondi, B. (1985) - "Modelling the GPS carrier phase for geodetic applications", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Schaffrin, B. and E. Grafarend (1986) - "Difference theorems for bias elimination with GPS observations", Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

Seidelmann, P.K. (1982) - "1980 IAU Theory of Nutation: Final report of the IAU Working Group on Nutation", Celestial Mechanics, 27, pp. 79-106.

Smith, D.E., R. Kolenkiewicz, P. Dunn, M.H. Torrence, E.E. Pavlis, J.W. Robbins, R.J. Williamson, S.M. Klosko, L.Carpenter and S.K. Fricke (1987) - "Global Laser Solution: SL7.1", presented at the NASA Crustal Dynamics Meeting, GSFC, Oct. 21-23.

Sovers, O.J. and J.S. Border - "Observation model and parameter partials for the JPL geodetic GPS modelling software GPSOMC", JPL Publ. 87-21.

Spilker, J.J. (1980) - "Signal structure and performance characteristics", in "Global Positioning System": papers published in Navigation, Vol. I, The Institute of Navigation, Washington, D.C.

Strange, W. (1985) - "Cm-level three dimensional crustal motion surveys with GPS", EOS, Transactions of the AGU, Vol. 66 (46), p.847 (abstract only).

Swift, E.R. (1985) - NSWC's GPS orbit/clock determination system", First Inter'l Symp. on Precise Positioning with GPS, Rockville, MD.

Tapley, B.D., B.E. Schutz and R.J. Eanes (1985) - "Station coordinates, baselines and earth orientation from LAGEOS laser ranging : 1976-1984", in "LAGEOS - Scientific results", AGU special issue reprinted from Journal of Geophysical Research, Vol. 90 (B11), Sept.

Thornton, C.L., W.G. Melbourne and T.H. Dixon (1986) - "NASA GPS-based geodetic program in Mexico and the Caribbean", Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

Tranquilla, J.M. (1987a) - "Some comments on the occurrence of multipath and imaging problems in electronic navigation and positioning applications", Department of Electrical Engineering, University of New Brunswick, Fredericton, N.B.

Tranquilla, J.M., J. Carr, R. Basett, B. Colpitts and S. Best (1987) - "Field experiments in NAVSTAR signal multipath, imaging and phase center errors", Final Report to Department of Fisheries and Oceans, Ottawa.

Treuhaft, R.N. and G.E. Lanyi (1985) - "The effect of the dynamic wet troposphere on VLBI measurements", TDA Progress Report 42-84, Jet Propulsion Laboratory, Pasadena.

Van Dierendonck, A.J., S.S. Russel, E.R. Kopitzke and M. Birnbaum (1980) - "The GPS Navigation Message", in "Global Positioning System": papers published in Navigation, Vol. I, The Institute of Navigation, Washington, D.C.

Vanicek, P., R.B. Langley, D.E.Wells and D. Delikaraoglou (1984) - "Geometrical aspects of differential positioning", Bullettin Geodesique, 55, pp. 37-52.

Vanicek, P., G. Beutler, A. Kleusberg, R.B. Langley, R. Santerre and D.E. Wells (1985) - "DIPOP: Differential positioning program package for the Global Positioning System", TR No. 15, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B.

Wahr, J. (1981) - "The forced nutations of an elliptical, rotating, elastic and oceanic earth", Celestial Mechanics, 64, pp. 704-727.

Ware, R.H., K.J. Hurst and C. Rocken (1987) - "Global change and GPS", EOS, Transactions of the AGU, Vol. 68 (16), p. 286 (abstract only).

Wei, Z. (1986) - "Positioning with NAVSTAR, the Global Positioning System", Rep. No. 370, Department of Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio.

Wells, D.E., N. Beck, D. Delikaraoglou, A. Kleusberg, E.J. Krakiwsky, G. Lachapelle, R.B. Langley, M. Nakiboglu, K.P. Schwarz, J.M. Tranquilla and P. Vanicek (1986) - "Guide to GPS Positioning", Canadian Institute of Surveying and Mapping, P.O. Box 5378 STN F, Ottawa, Ontario, Canada.

Williams, B.G. (1986) - "GPS satellite orbit determination results from the March 1985 field test", Fourth Inter'l Geodetic Symp. on Satellite Positioning, Austin, TX.

NASA
National Aeronautics and Space Administration

Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM-100716		

4. Title and Subtitle	5. Report Date
On Principles, Methods and Recent Advances in Studies Towards a GPS-Based Control System for Geodesy and Geodynamics	February 1989
	6. Performing Organization Code: 621.0
7. Author(s): Demitris Delikaraoglou	8. Performing Organization Report No.: 89B00058
	10. Work Unit No.
9. Performing Organization Name and Address: Geodynamics Branch, Goddard Space Flight Center, Greenbelt, Maryland 20771	11. Contract or Grant No.
	13. Type of Report and Period Covered: Technical Memorandum
12. Sponsoring Agency Name and Address: National Aeronautics and Space Administration, Washington, D.C. 20546-0001	14. Sponsoring Agency Code

15. Supplementary Notes

The author held a U. S. National Research Council National Academy of Sciences Research Associateship at the NASA Goddard Space Flight Center while on leave of absence from the Canadian Geodetic Survey, Surveys and Mapping Branch, Ottawa, Ontario, Canada K1A 0E9.

16. Abstract

Although Very Long Baseline Interferometry (VLBI) and Satellite Laser Ranging (SLR) are becoming increasingly important tools for geodynamic studies, their future role may well be fulfilled by using alternative techniques such as those utilizing the signals from the Global Positioning System (GPS). GPS, without the full implementation of the system, already offers a favorable combination of cost and accuracy and has consistently demonstrated the capability to provide high precision densification control in the regional and local areas of the VLBI and SLR networks.

This report reviews VLBI and SLR vis-à-vis GPS and outlines the capabilities and limitations of each technique and how their complementary application can be of benefit to geodetic and geodynamic operations. It demonstrates, albeit with a limited data set, that dual-frequency GPS observations and interferometric type analysis techniques make possible the modelling of the GPS orbits for several days with an accuracy of a few meters. The use of VLBI or SLR sites as fiducial stations together with refinements in the orbit determination procedures can greatly reduce the systematic errors in the GPS satellite orbits used to compute the positions of non-fiducial locations. In general, repeatability and comparison with VLBI of the GPS determined locations are of the order of between 2 parts in 10^7 and 5 parts in 10^8 for baseline lengths < 2000 km.

This report is mainly a synthesis of problems, assumptions, methods and recent advances in the studies towards the establishment of a GPS-based system for geodesy and geodynamics and is one phase in the continuing effort for the development of such a system. To some, including the author, it seems reasonable to expect within the next few years that more evidence will show GPS to be as a powerful and reliable a tool as mobile VLBI and SLR are today, but largely more economical.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Geodynamics, Global Positioning System, GPS, Satellite Laser Ranging, SLR, Geodesy, Very Long Baseline Interferometry, VLBI	Unclassified - Unlimited Subject Category 46

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified		