THE USE OF SHORT ARC ORBITAL CONSTRAINTS IN THE ADJUSTMENT OF GEODETIC SATELLITE DATA

Charles Reed Schwarz

NATIONAL AERONAUTICS & SPACE ADMINISTRATION
Washington D.C.

Contract No. NGR36-008-093
OSURF Project No. 2514

Reports of the Department of Geodetic Science

Report No. 118

THE USE OF SHORT ARC ORBITAL CONSTRAINTS IN THE ADJUSTMENT OF GEODETIC SATELLITE DATA

by

Charles Reed Schwarz

Prepared for

National Aeronautics and Space Administration
Washington, D. C.

Contract No. NGR 36-008-093
OSURF Project No. 2514

The Ohio State University
Research Foundation
Columbus, Ohio 43212

January, 1969

Preface

This project is under the direction of Professor Ivan I. Mueller of the Department of Geodetic Science, The Ohio State University. Project Manager is Jerome D. Rosenberg, Geodetic Satellites, Code SAG, NASA Headquarters, Washington, D. C. The contract is administered by the Office of University Affairs, NASA Headquarters, Washington, D. C.

Reports of The Ohio State University Department of Geodetic Science related to NASA Contract No. NGR 36-008-093 and published to date are the following:

The Determination and Distribution of Precise Time, Report No. 70, April, 1966

Proposed Optical Network for the National Geodetic Satellite Program, Report No. 71, May, 1966

Preprocessing Optical Satellite Observations, Report No. 82, April 1967

Least Squares Adjustment of Satellite Observations for Simultaneous Directions or Ranges. Part I: Report No. 86, September, 1967, Part II: Report No. 87, Aug,1968, Part III: Report No. 88, December, 1967

Preprocessing Electronic Satellite Observations, Report No. 100, March, 1968

Comparison of Astrometric and Photogrammetric Plate Reduction Techniques for a Wild BC-4 Camera, Report No. 106, March, 1968

Investigations Into the Utilization of Passive Satellite Observational Data, Report No. 110, June, 1968

Sequential Least Squares Adjustment of Satellite Triangulation and Trilateration in Combination with Terrestrial Data, Report No. 114, in press

Abstract

In this report many different aspects of the orbital constraint adjustment are discussed, comparing it in many ways with the corresponding geometric mode adjustment. The difficulty in modeling the orbit is the biggest drawback to the orbital mode adjustment, and this is described at some length. It is shown that for any given set of data, the orbital constraint adjustment will produce a solution that is as good as or better than the corresponding geometric mode adjustment.

A description of the flow of a particular orbital constraint adjustment computer program is presented. The discussion of this specific program is used to illustrate several general computational and programming considerations.

A series of experiments involving BC-4 photographic plates is described. The path of the satellite appears as a trail of small dot-like images on these plates, and the possibility of imposing an orbital constraint on these trails is discussed. Results are presented of the adjustment of these images in the short arc orbital mode, and the results of several different configurations of geometric mode adjustments are given for the sake of comparison. These results indicate that for short enough arcs the modeling error may be suppressed and the orbital mode adjustment yields an unbiased solution which is almost the same as that obtained by the geometric method. However, the expected strengthening of the solution due to the imposition of the orbital constraint does not seem to be strong enough to show up in these experiments. Moreover, the short arc orbital constraint adjustment appears to be especially prone to a severe accumulation of numerical error, due to the presense of ill-conditioned matrices.

Acknowledgements

The material contained in this report was presented as a thesis in partial fulfillment of the requirements of the degree Master of Science at The Ohio State University. Professor Ivan I. Mueller served as the author's advisor and suggested many of the experiments performed and participated in the analysis of the results. Professor Richard H. Rapp served on the author's reading committee and helped to clarify several sections. Computer time for many of the experiments was furnished by the Ohio State University Computer Center.

Many of the algorithms used in the computer program described in Chapter 4 were taken from a SCATRAN language computer program developed by Professor Richard H. Rapp. The SCATRAN program was also used to check the results of the author's FORTRAN IV program.

The author further wishes to express his appreciation to the U. S. Army Topographic Command which made possible the period of graduate study during which the thesis was written and to Mrs. Jeanne C. Preston who typed this report.

Contents

1. Introduction

The goals of satellite geodesy have at times been classified as either geometric or dynamic. The geometric goals involve the determination of observing station positions, while the dynamic goals involve the description of the earth's gravity field. These divisions correspond to the two main branches of classical geodesy; however, satellite geodesy differs from classical geodesy in that it allows the geodesist to extend his interest to descriptions of the earth that are continental or global in scope.

Investigations whose purpose is to accomplish geometric goals may be further classified into geometric and dynamic methods. The geometric methods treat the satellite much as an unoccupied triangulation or trilateration station. On the other hand, the dynamic methods of geometric satellite geodesy (as differentiated from dynamic satellite geodesy) exploit the fact that the path, or orbit, of the satellite obeys the physical laws of dynamics and can be treated by the methods of celestial mechanics.

Although an artificial satellite obeys the same basic physical laws as the moon and planets, its motion is generally much more complex, expecially in the case of the near earth satellites that are used for geodetic purposes. Whereas analytic theories, valid for fifty or a hundred years, exist for the moon and planets, the motion of an artificial earth satellite may be accurately predicted a few days ahead only with great difficulty.

The dynamic methods are generally recognized as being either long arc or short arc methods, depending on the mathematical description of the orbit they employ. Long arc methods characteristically employ complicated mathematical models that attempt to describe the orbit of the satellite exactly. Short arc methods are characterized by mathematical descriptions of the orbit that are simpler, but are only valid for short periods of time. Just where the dividing line between "long" and "short" falls is pretty much up to the judgment of each author or investigator, but a definition that is considered reasonable by this author is discussed in section 2.3.

Adjustments utilizing the geometric method are said to be of the "geometric mode," while those utilizing the dynamic method are called "orbital mode" or "orbital constraint." Most agencies currently adjusting networks or satellite observing stations utilize the geometric mode. The correctness of the geometric approach is demonstrated by the publication of high-precision global nets by the Smithsonian Astrophysical Observatory and by steady progress toward a global net by the U. S. Coast and Geodetic Survey. Although the adjustment of geodetic satellite data by the use of orbital constraints has been strongly urged by some investigators, orbital mode adjustments, where they exist, are generally used as a tool of analysis rather than as an operational procedure. It appears to this author that the features of the orbital mode adjustment, and its advantages and disadvantages relative to the geometric mode adjustment, have not always been completely understood. For this reason many aspects of the orbital and geometric methods are compared throughout this report, from an analytic point of view in Chapter 2, from the point of view of computational and programming considerations in Chapter 3, and by some numerical results in Chapter 4.

2. Principles and Features of the Orbital Constraint Adjustment

2.1 "Orbital Mode" and "Orbital Constraint" Adjustments

The terms "orbital mode" and "orbital constraint" are used somewhat interchangeably in the literature. It is worthwhile to examine carefully in just what sense these two terms are synonymous, and just what the constraint is that is being imposed. Such an examination serves to make clear under what circumstances the constraint is valid, and helps to define just how long is a "short" arc.

In the "geometric mode" of adjustment, the observation equation relates the position of the observing station to the position of the satellite at the time of the observation. The unknowns are (a) three components of position for each observing station, and (b) three components of position for the satellite for each instant at which an observation is made. Each satellite position is treated as independent of all others. The total number of satellite position unknowns is three times the number of distinct epochs at which the satellite was observed. Thus, geometric mode adjustments characteristically contain an extremely large number of satellite position unknowns. Since each satellite position is treated as an independent set of unknowns, it is necessary that the observations made at any instant be sufficient to determine the satellite position at that instant. Otherwise, the resulting normal equations will be singular and there will be no solution. For instance, observations of range from only two stations or direction from only one station are not usable in the geometric mode.

The orbital constraint adjustment recognizes that the satellite position unknowns are not all independent, for the various positions of each satellite must all lie along some smooth path, or orbit. The constraint, then, is a relation among the satellite position unknowns which arise in the geometrical form of the adjustment. If the constraints are treated as relationships to be rigorously enforced in the adjustment, then the constraint equations may be solved for the various satellite position unknowns and the resulting equations substituted into the observation equations. When this is done, the adjustment is said to be of the "orbital mode."

Since only six independent parameters are required to describe the true orbit, most forms of the orbital constraint equations introduce six new "orbital unknowns." The exceptions to this are the empirical models of the orbit, in which more than six parameters may be used to describe the motion. Since a separate set of three constraint equations may be written for each satellite position, the large number of satellite position unknowns is reduced to a small number, usually six, of orbit unknowns.

The foregoing discussion assumes that the observations depend only on the relative positions of the satellite and the observer. If any rates, for instance range rate, are observed, then the geometrical form will include satellite velocity as well as position unknowns, and the orbital constraint will also require that the various velocity vectors belong to some orbit. For the sake of simplicity the following discussions will assume that only position dependent quantities, such as range, direction, or range difference, have been observed.

The generalization to include satellite velocity unknowns is straightforward, since an additional constraint equation may be written to relate each satellite velocity unknown to the orbit unknowns.

To make the discussion more definite, the relationship between 'orbital mode" and "orbital constraint" may be formalized. Suppose that observations have been made from a network of ground stations to several satellites, or to a single satellite whose path is broken into several orbits. Let t_j denote the jth time epoch, x_i the position vector of the ith ground station, and y_{jk} the position vector of the satellite in the kth orbit at time t_j, both resolved in some suitable coordinate system. Let ℓ_{ijk} be the actual observation from station i to satellite position y_{jk}, and let v_{ijk} represent the "accidental measurement error" component of the observation. Further suppose that the v_{ijk} behave as uncorrelated random variables with zero means and variances $\sigma^2_{ijk} = 1/w_{ijk}$. Each individual observation gives rise to a single equation of the form

$$v_{ijk} + \ell_{ijk} = f(x_i,\ y_{jk},\ t_j). \tag{2.1}$$

The function on the right hand side is expanded in a Taylor series about approximate values of the unknowns, x_i^o and y_{jk}^o, and terms of order higher than the first are neglected in the usual manner.

$$v_{ijk} + \ell_{ijk} = f(x_i^o,\ y_{jk}^o,\ t_j) + \frac{\partial f}{\partial x_i}(x_i - x_i^o) + \frac{\partial f}{\partial y_{jk}}(y_{jk} - y_{jk}^o)$$

or

$$v_{ijk} + \delta\ell_{ijk} = \frac{\partial f}{\partial x_i}\,\delta x_i + \frac{\partial f}{\partial y_{jk}}\,\delta y_{jk} \tag{2.2}$$

where
$$\delta\ell_{ijk} = \ell_{ijk} - f(x_i^o,\ y_{jk}^o,\ t_j),$$

$$\delta x_i = x_i - x_i^o,$$

$$\delta y_{jk} = y_{jk} - y_{jk}^o .$$

The total set of observation equations is written in matrix form as

$$V + L = AX + BY \tag{2.3}$$

or

$$V + L + (A \vdots B) \begin{pmatrix} X \\ \cdots \\ Y \end{pmatrix}$$

where the dotted lines indicate a partitioning. These equations are the basis of the geometric mode adjustment. If the rank of $(A \vdots B)$ is equal to its column dimension, then a least-square solution is given by

$$\begin{pmatrix} X \\ \cdots \\ Y \end{pmatrix} = [(A \vdots B)^T W (A \vdots B)]^{-1} (A \vdots B)^T WL \tag{2.4}$$

where W is a diagonal matrix containing the weights w_{ijk} along its main diagonal.

Let z_k represent the set of parameters that describe the kth orbit. Then the knowledge that the various satellite positions must lie along an orbit is expressed formally by the relation

$$e_{jk} + y_{jk} = g(z_k, t_j). \tag{2.5}$$

Here e_{jk} represents the error in the function $g(z_k, t_j)$. This function is the mathematical model for the orbit. However, we are generally either unable or unwilling to construct a mathematical model which describes reality exactly. Hence the term e_{jk} is called the <u>modeling error</u>. The nature of this term is one of the most important considerations of the whole problem. It generally has several components, each of which requires careful consideration (see section 2.3). Under certain hypotheses, this term may be considered to act as a random variable with mean zero and variance $\sigma^2(e_{jk}) = \epsilon[e_{jk}^2]$, where ϵ denotes the expectancy operator. If the model represented by the function g is constructed carefully enough, the modeling error may be held to negligibly small magnitude, at least for short periods of time.

Since the satellite position y_{jk} contains three components, equation (2.5) is actually a set of three equations. There will be as many of these sets as there are satellite positions. If the satellite velocity were being considered, it would be possible to modify the meaning of e and g and write a relationship of the form

$$e_{jk} + \begin{pmatrix} y \\ \dot{y} \end{pmatrix} = g(z_k, t_j).$$

Equation (2.5) may be linearized by an expansion around approximate values of the orbit parameters z_k^o, yielding

$$e_{jk} + y_{jk} - g(z_k^o, t_j) = \frac{\partial g}{\partial z_k}(z_k - z_k^o)$$

or

$$e_{jk} + \delta y_{jk} = \frac{\partial g}{\partial z_k}\delta z_k. \tag{2.6}$$

The total set of such linearized equations is written in matrix form as

$$E + Y = FZ. \tag{2.7}$$

These are called <u>constraint equations</u>, since they represent constraints among the parameters Y.

There are several alternative methods of incorporating these equations into an adjustment. It is necessary to make one of two hypotheses about the modeling errors E: either (a) the elements of E behave as random variables, so that E may be treated as a vector random variable with zero mean and covariance matrix $\epsilon[EE^T] = \Sigma_E$; or (b) the elements of E, whether random or not, are so small that the elements of the product BE are negligible in comparison with those of V. If neither of these hypotheses can reasonably be made, there is no possibility of obtaining an honest adjustment, and a new model for the orbit must be constructed or the whole experiment must be redesigned.

If hypothesis (a) is used, then equations (2.7) are of exactly the same form as equations (2.3), and may be treated as additional observation equations. Thus the total set of equations becomes

$$\begin{pmatrix} V \\ E \end{pmatrix} + \begin{pmatrix} L \\ O \end{pmatrix} = \begin{pmatrix} A & B & O \\ O & -I & F \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{2.8}$$

and the weight matrix for these equations is

$$\begin{pmatrix} W & O \\ O & \Sigma_E^{-1} \end{pmatrix} .$$

If hypothesis (b) is used, then equations (2.7) may be written simply $Y = FZ$ and are treated as conditions to be rigorously enforced in the adjustment. The least-squares solution of (2.3), subjected to the conditions $Y = FZ$, is found by minimizing the function

$$\varphi = V^T W V + \Lambda^T (Y - FZ) + (Y - FZ)^T \Lambda$$

where Λ is a column matrix of Lagrange multipliers. Setting the variations of φ with respect to X, Y, Z, and Λ each equal to zero produces the four matrix equations

$$\begin{aligned} A^T W V &= 0, \\ B^T W V + \Lambda &= 0, \\ -F^T \Lambda &= 0, \\ Y - FZ &= 0. \end{aligned}$$

Substituting $V = AX + BY - L$ from equation (2.3) and rearranging, these are written in matrix form as

$$\begin{pmatrix} A^T WA & A^T WB & 0 & 0 \\ B^T WA & B^T WB & 0 & I \\ 0 & 0 & 0 & -F^T \\ 0 & I & -F & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ \Lambda \end{pmatrix} = \begin{pmatrix} A^T WL \\ B^T WL \\ 0 \\ 0 \end{pmatrix} . \tag{2.9}$$

The solution in the form of either equations (2.8) or (2.9) is properly called an <u>orbital constraint solution</u>, because the constraints appear explicitly.

Another possibility is to solve equations (2.7) for the satellite position unknowns Y and substitute into equations (2.3), obtaining

$$V + L = AX + B(FZ - E)$$

or

$$(V + BE) + L = (A \quad BF) \begin{pmatrix} X \\ Z \end{pmatrix} \tag{2.10}$$

If hypothesis (a) is used, then this is a reduced set of observation equations, with error term $V + BE$. The minimum variance solution for $\begin{pmatrix} X \\ Z \end{pmatrix}$ requires that the weight matrix be inversely proportional to the covariance matrix of this new error term, which is

$$\epsilon[(V + BE)(V + BE)^T] = W^{-1} + B\Sigma_E B^T .$$

Even though W^{-1} may be a diagonal matrix $W^{-1} + B\Sigma_E B^T$ is a full matrix and its inversion may be a tedious process. Therefore this form of solution is undesirable and is probably never used.

On the other hand, if hypothesis (b) is used, E may be neglected and equation (2.10) may be simplified to

$$V + L = (A \quad BF) \begin{pmatrix} X \\ Z \end{pmatrix} \tag{2.11}$$

The minimum variance solution to this set of equations is then the solution to

$$(A\ BF)^T W (A\ BF) \begin{pmatrix} X \\ Z \end{pmatrix} = (A\ BF)^T WL$$

or

$$\begin{pmatrix} A^T WA & A^T WBF \\ F^T B^T WA & F^T B^T WBF \end{pmatrix} \begin{pmatrix} X \\ Z \end{pmatrix} = \begin{pmatrix} A^T WL \\ F^T B^T WL \end{pmatrix}. \tag{2.12}$$

The solution in this form is properly called an <u>orbital mode adjustment</u>. This term could also be applied to the minimum variance estimate of X and Z in equation (2.10). This solution would be given by equations of the form (2.12), but W would be replaced by $(W^{-1} + B\Sigma_E B^T)^{-1}$.

A few algebraic manipulations will show that equation (2.9) is exactly equivalent to equation (2.12). The second equation in (2.9) is solved for Y, and this is substituted into the other three equations, giving

$$\left.\begin{aligned} & Y = (B^T WB)^{-1}(B^T WL - B^T WAX - \Lambda) \\ & (A^T WA - A^T WB(B^T WB)^{-1} B^T WA)X - A^T WB(B^T WB)^{-1}\Lambda \\ & \qquad = A^T WL - A^T WB(B^T WB)^{-1} B^T WL \\ & -F^T\Lambda = 0 \\ & (B^T WB)^{-1} B^T WAX + (B^T WB)^{-1}\Lambda + FZ = (B^T WB)^{-1} B^T WL \end{aligned}\right\} . \tag{2.9a}$$

The last of these equations is solved for Λ, yielding

$$\Lambda = B^T WL - B^T WAX - B^T WBFZ.$$

This is then substituted into the middle two equations of (2.9a) giving

$$\begin{aligned} A^T WAX + A^T WBFZ &= A^T WL, \\ F^T B^T WL - F^T B^T WAX - F^T B^T WBFZ &= 0, \end{aligned}$$

or

$$\begin{pmatrix} A^T W A & A^T W B F \\ F^T B^T W A & F^T B^T W B F \end{pmatrix} \begin{pmatrix} X \\ Z \end{pmatrix} = \begin{pmatrix} A^T W L \\ F^T B^T W L \end{pmatrix} . \tag{2.9b}$$

But this is exactly the same as equation (2.12). This showns that the orbital constraint solution and the orbital mode solution are indeed algebraically equivalent, although they are different ways of approaching the same problem. Since equation (2.10) is a reduced form of (2.9), it is the much more common practice to form the observation equations in equation (2.11) directly. In effect, this imposes the constraint algebraically rather than numerically.

2.2 Simultaneous and Orbital Modes of Scheduling Observations

Most observing systems observe only one component of position, such as range, or two components, such as direction. If any of these observations are to be usable in a geometric mode adjustment, it is necessary that the observations be made simultaneously from at least two or three stations so that the satellite position will be determined and the adjustment will not be singular. Because of this, the geometric mode of adjustment is often called the "simultaneous mode." Acutally, this is somewhat misleading, since a "simultaneous adjustment" means something quite different. What is really meant is that the observations have been made in a simultaneous mode (or reduced to simultaneity) in order that the adjustment of the observations might be performed in a geometric mode.

In the orbital mode, there is no need for simultaneity of observations. Moreover, there is no need to schedule observations only for times at which the satellite is visible from more than one station. This allows more observations to be made, and also allows very isolated stations to participate in a network. Because of this feature, an orbital mode adjustment has often been envisaged as a method of tying together datums and tying isolated islands to mainland datums (see Figures 2.1 and 2.2).

In practice,use is seldom made of this feature. Most U. S. agencies which are engaged in the tracking of satellites for geometric purposes plan and execute their observations in the simultaneous mode. For instance, one may cite the SECOR program of the U. S. Army Map Service, the PC-1000 program of the U.S. Air Force, and the BC-4 program of the U. S. Coast and Geodetic Survey. As far as is known to this author, only those agencies which have an orbit determination mission in addition to their geodetic mission observe geodetic satellites in a nonsimultaneous mode. For instance, the Baker-Nunn camera program of the Smithsonian Astrophysical Observatory and the U. S. Navy's TRANET program have orbit determination as well as geodetic missions.

2.3 The Equations of Motion

The motion of the satellite through space is governed by Newton's Second Law, which may be stated as the vector differential equation

$$\ddot{X} = F \tag{2.13}$$

where $\ddot{X}$ is the acceleration vector of the satellite and F is the force per unit mass acting on it. Since this is actually a set of three second order differential equations, its solution contains six arbitrary constants of integration. These six constants may take several different forms, depending on

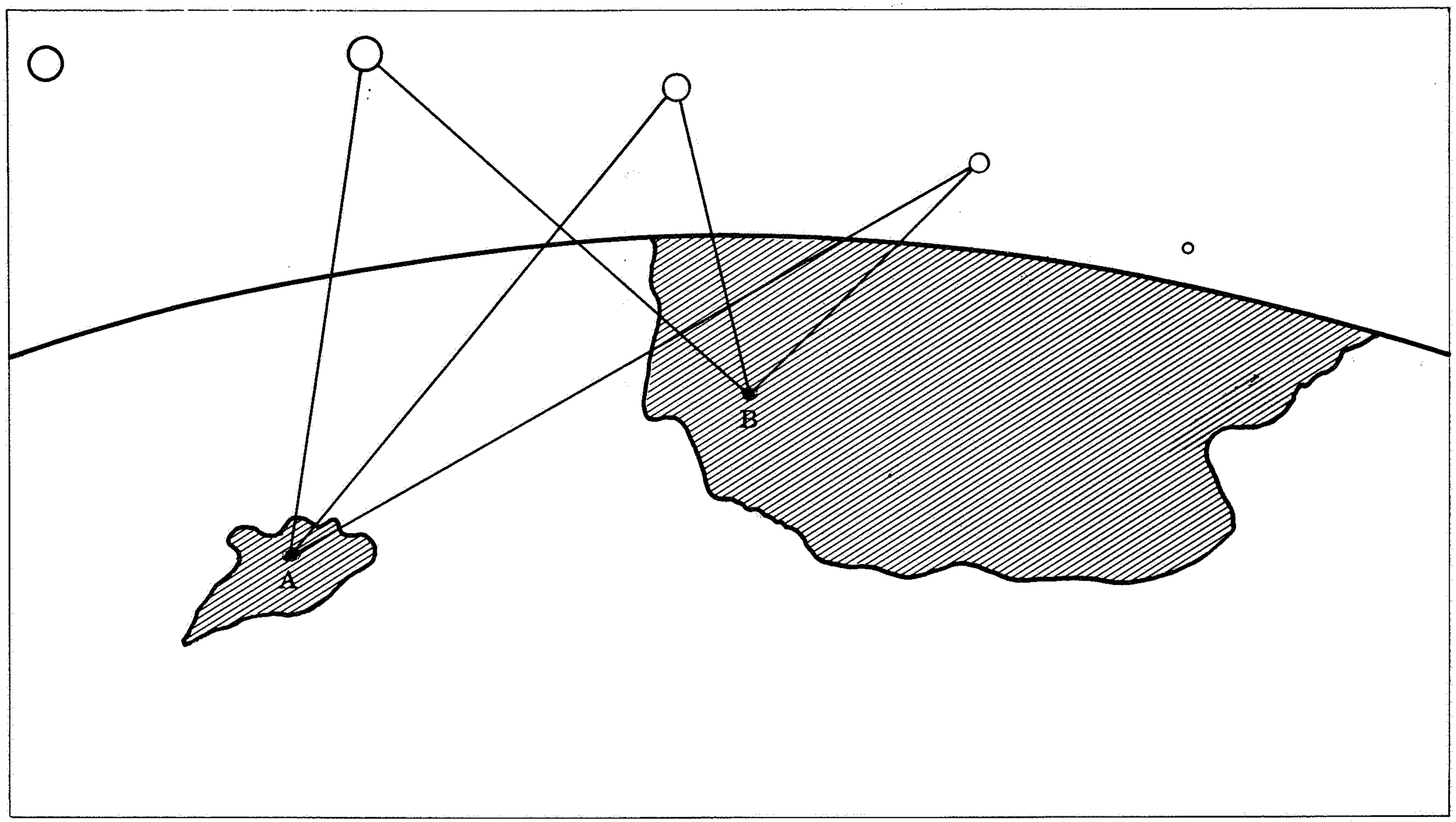


Fig. 2.1 The simultaneous mode of making observations. Each position of the satellite is observed by both stations A and B. Observations may be made only while the satellite is visible from both stations.

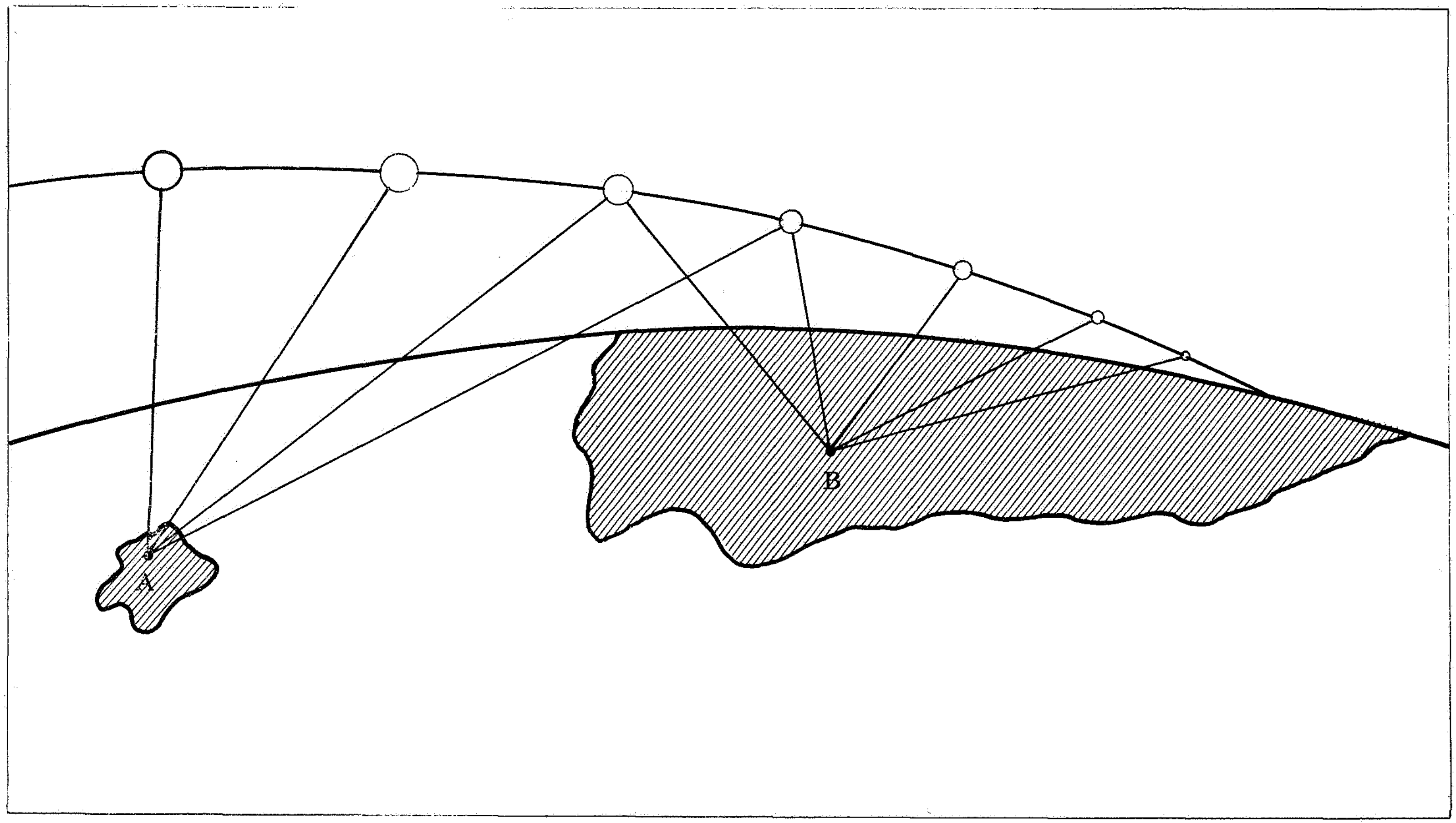


Fig. 2.2 The orbital mode of making observations. All satellite positions visible from one or both stations are usable. The time spans of the observations taken at stations A and B may or may not overlap. Observations of the same satellite position by both stations may or may not be included.

the variables in terms of which this differential equation is solved. However, the important principle is that the orbit is completely characterized by only six numbers, and it is for this reason that most models for the orbit employ six "orbit unknowns."

Formally, the solution to this differential equation may be written

$$\dot{X}(t) = \int_{t_o}^{t} F(t)dt + \dot{X}(t_o)$$

$$X(t) = \int_{t_o}^{t} \dot{X}(t)dt + X(t_o)$$

$$= \int_{t_o}^{t}\int_{t_o}^{t} F(t)dtdt + (t - t_o)\dot{X}(t_o) + X(t_o) \tag{2.14}$$

or

$$X(t) = \int_{t_o}^{t} (t - \tau)F(\tau)d\tau + (t - t_o)\dot{X}(t_o) + X(t_o). \tag{2.15}$$

Here the constants of integration are the position and velocity vectors at the epoch time, $X(t_o)$ and $\dot{X}(t_o)$.

The complexity of the force function F determines the ease with which the indicated integration may be carried out. If it consists solely of a central force field, then the motion is simple two body motion and a closed form solution may be written in terms of the relatively simple equations of the Keplerian ellipse. Because of several factors, principally the oblateness of the earth, the forces acting on a near earth satellite deviate slightly from those of a central force field. Since these deviations are small, the actual motion of an artificial satellite greatly resembles Keplerian motion. The small deviations in the force field are termed "perturbing forces," and the actual motion is called "perturbed two-body motion" or "perturbed Keplerian motion." Since the acceleration in the actual force field differs only slightly from that of a central force field, the actual motion departs only slowly from a Keplerian ellipse. However, the important mathematical difference between motion in the actual force field and motion in a central force field is that no closed form solution exists for the equations of motion (2.13) in the case of perturbed two body motion. Obviously, the solution may always be obtained by a numerical evaluation of the integral in (2.14), although the numerical integration may require a very large number of computations. An approximate expression for the motion may also be obtained by expanding the solution in a series in powers of the small quantities that describe the perturbing forces. These analytic solutions contain both secular and periodic functions of time. They are typically carried to a higher order in the secular terms than in the periodic terms, so that they describe the orbit with moderate accuracy for long periods of time.

Because no closed form solution exists for perturbed two body motion, the geodesist is faced with a choice of a great many approximate solutions which may serve as a model for the orbit. This choice is usually dictated by the accuracy required and the length of time for which an accurate solution is required. Typically, the Keplerian orbit has moderate accuracy for very short periods of time. The numerical integration schemes are very accurate for short periods of time; however, they always include some numerical error, which generally accumulates with time. The analytic solutions are most suitable for describing the orbit over many revolutions. However, since they describe the large motions rather than the small motions of the satellite, they are no more accurate than the simple Keplerian orbit for describing the motion over a small fraction of a revolution.

Several comments may be made concerning the formal equations of motions:

(1) Equation (2.13) is strictly valid only in an inertial coordinate system. As far as can be detected, a nonrotating coordinate system centered at the barycenter of the solar system satisfies this requirement. However, the satellite is carried along by the earth in its orbit, so that it is more natural to write the equations of motion in terms of an earth centered coordinate system. If the equations of motion are written in terms of a barycentric system and then transformed to a nonrotating coordinate system centered at the earth's center of mass, the force function will consist of a dominant force produced on the satellite by the earth, plus a small perturbing force caused by the differences between the forces produced by the sun on the earth and on the satellite. The difference between the forces produced by the moon on the earth and on the satellite may similarly be considered as a small perturbing force.

(2) Equation (2.13) describes a purely Newtonian motion. This is not exactly the true motion in the sense that a geodetic satellite moves with sufficient velocity that relativistic effects should be present. However, the relativistic effects may be treated as small perturbations of the Newtonian motion, specifically as a small contribution to the secular motion of the argument of perigee [Danby, 1962, p. 66]. For orbits lasting several minutes,

or even several weeks, the relativistic effects on the orbit are much too small to be detected, and are never considered for the purpose of geodetic investigations.

(3) The time appearing in equation (2.13) is a purely Newtonian, or uniform, time. This can cause some confusion, since UT1 also enters the observation equation in the adjustment to relate an earth-fixed coordinate system to a nonrotating system.

The force vector on the right-hand side of equation (2.13) theoretically includes all forces acting on the satellite, both gravitational and nongravitational. Obviously it is impossible to model accurately all the forces which could act on the satellite, and some degree of approximation is necessary. The amount of approximation made in modeling the forces determines the ease with which the integral in equation (2.15) may be evaluated. Unfortunately, using an approximate force model also restricts the time interval for which the solution is accurate.

The important forces acting on a near earth satellite are the attractions of the earth, the attractions of the moon and sun, air drag, and solar radiation pressure [Mueller, 1964, p. 173]. The attraction of the earth is usually written as the gradient of the gravitational potential of the earth, where the potential is expressed as a series of solid spherical harmonics, i.e.,

$$F = \nabla V$$

and

$$V = \frac{GM}{r}\left\{1 - \sum_{n=2}^{\infty}\left(\frac{a_e}{r}\right)^n \sum_{m=0}^{n}\left(J_{nm}\cos m\lambda + K_{mn}\sin m\lambda\right)P_{nm}(\cos\varphi)\right\}.$$

Since the earth is nearly spherical, the most influential term is the central force term (J_0 harmonic) $\frac{GM}{r}$. It is nearly 1000 times larger than the J_2 (or oblateness) term, which in turn is about 1000 times larger than any other effect. Lunar and solar forces are even smaller and are important only for satellites higher than those ordinarily used for geodetic purposes. Air drag is important only for very low satellites or those with a large area to mass ratio. Solar radiation pressure is a very small effect which is noticeable only for satellites with a very large area to mass ratio, such as the balloon satellites.

There are many other considerations that affect the accuracy with which the motion may be modeled. To discuss the problem systematically, it is possible to divide the modeling errors into four types, and it is usually necessary to consider all four when discussing a model for the orbit:

(1) Truncation or approximation errors in the force function. These are errors caused by including only the major forces acting on the satellite on the right-hand side of equation (2.13). The force function may be as simple as the central body force (zero-order harmonic) exerted by the earth, or it may include forces arising from zonal and tesseral harmonics in the geopotential, lunar, and possibly solar, gravitational forces, and a model of the air drag exerted by the atmosphere.

(2) Errors in the constants of models that are included in the force function. These would be principally errors in the geopotential coefficients used to describe the gravity field (including the J_0 = GM term), and errors in the model used to describe the variation of density in the atmosphere. Although the expression of the gravity field may be correct in <u>form</u>, the numerical constants

included with the expression may not be absolutely accurate. The determination of these constants is another task of geodesy, but since they must be determined by observations, their numerical value can never be absolutely determined.

(3) Errors in performing the integration indicated in equation (2.15). If any perturbing forces are included in the force function, the integration may be performed only approximately. The method of integration used may be either analytical or numerical. The presence of integration errors means that even if all forces were included in the force model, and even if all numerical constants included in the model were known exactly, the model for the orbit would still contain some error.

(4) Errors due to using a rotating coordinate system, such as the true sidereal system of date, or a nonuniform time, such as Universal Time (see [Kozai, 1960]).

There is also the possibility of modeling the motion of the satellite by an empirical orbit, in which the elements of a Keplerian ellipse are expressed as polynomial (and possiblity trigonometric and hyperbolic) functions of time [Mueller, 1964, p. 213]. These empirical expressions are, strictly speaking, not solutions to the differential equations of motion at all, and the accuracy with which they describe the actual orbit cannot be discussed from exactly the same point of view. Although they resemble the analytic solutions in form, they contain more than six arbitrary constants. A refinement of this approach uses empirical expressions to model the secular variations of the elements, while the analytic expressions of Kaula are used to compute the short period perturbations [Gaposchkin,1966, p. 100; Kaula, 1966, p. 37]. All of these considerations affect the accuracy with which the position of a satellite may be predicted, given a set of initial conditions or constants of integration. The initial conditions usually appear as either six Keplerian elements at a certain epoch time, or else as three components of position and three components of velocity at the epoch time. If the actual orbit and the orbit model start with the same set of initial conditions, the modeling error is necessarily zero at the epoch time. As the time from the epoch increases, the position error of the prediction will typically contain both secular and long period terms. The type of model used determines the amplitude of these terms. In general, the periodic terms produce the largest contribution to the error at first, after which the secular terms predominate.

However, one does not generally know the epoch elements; these are the orbit unknowns in the orbit constraint adjustment. The important question, therefore, is not how well the model may be used for predictions, but how well the model may be made to fit the actual orbit by adjustment of the initial conditions. For short periods of time, it is sufficient to consider the actual orbit to be one generated using a full expression for the gravitational potential of the earth. The effect of further truncation of the force model may then be studied by comparison with such a reference orbit.

Figures 2.3 — 2.5 were obtained by numerically integrating two orbits having the same initial conditions for 1000 seconds. The force function for the first orbit was the geopotential given by the SAO G8 set of coefficients. This set of coefficients was truncated at the seventh degree and seventh order. The

Fig. 2.3 (from [Brown, 1967b]) Least squares adjustment of initial conditions to fit Kozai's zonals through degree 4 to SAO (7, 7) tesserals

Residual in X position coordinate

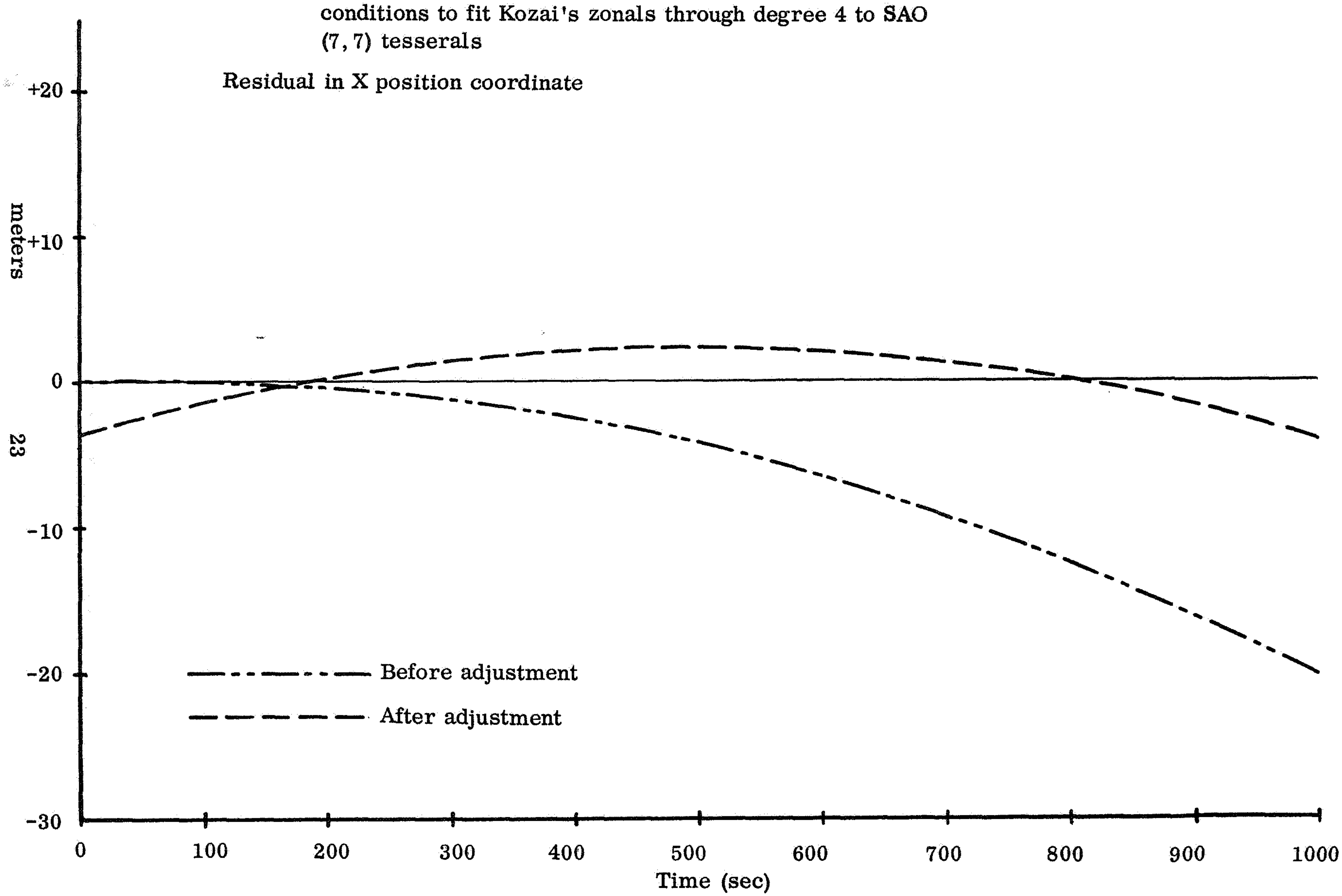


Fig. 2.4 (from [Brown, 1967b]) Least squares adjustment of initial conditions to fit Kozai's zonals through degree 4 to SAO (7, 7) tesserals

Residual in Y position coordinate

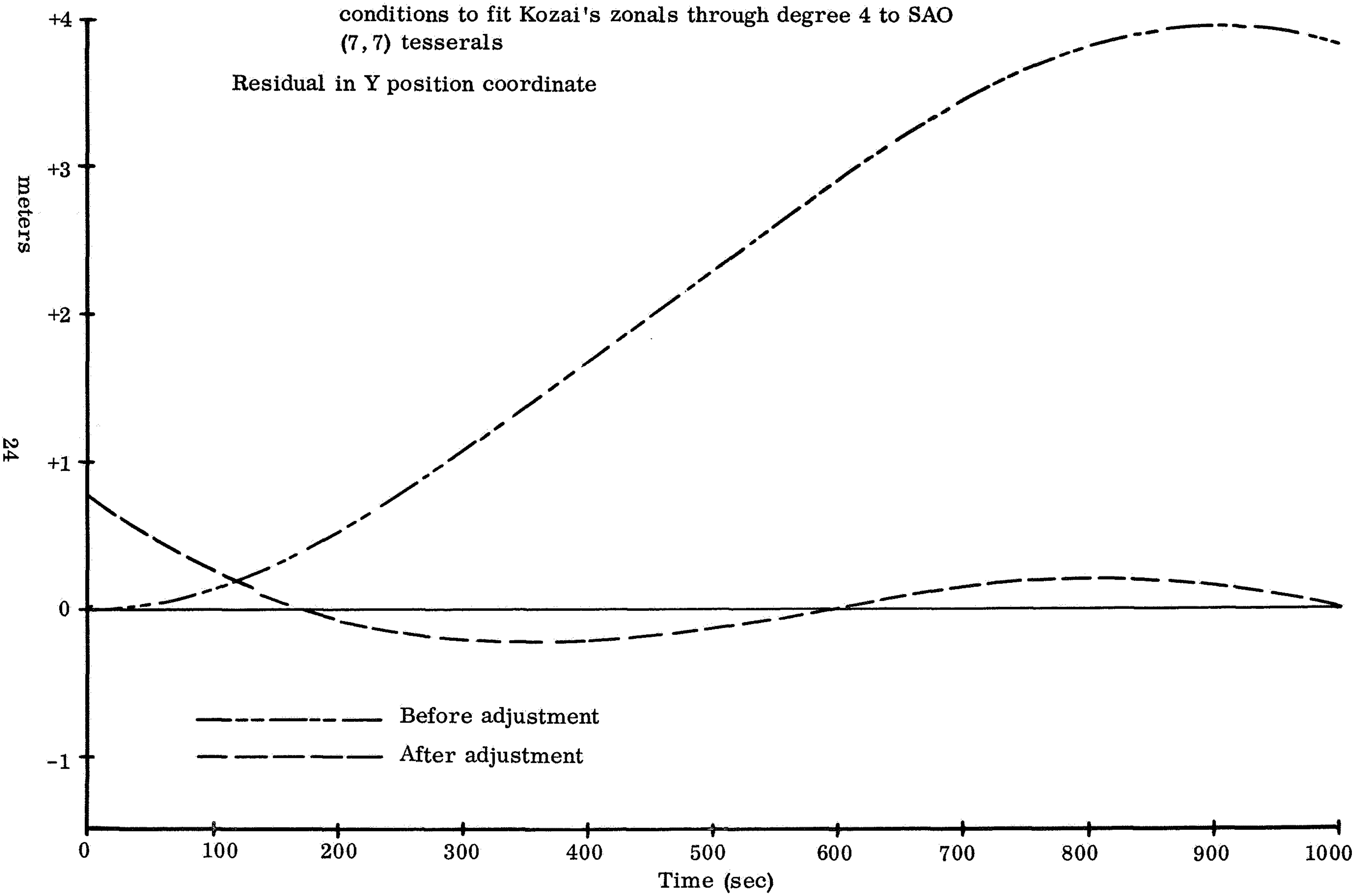
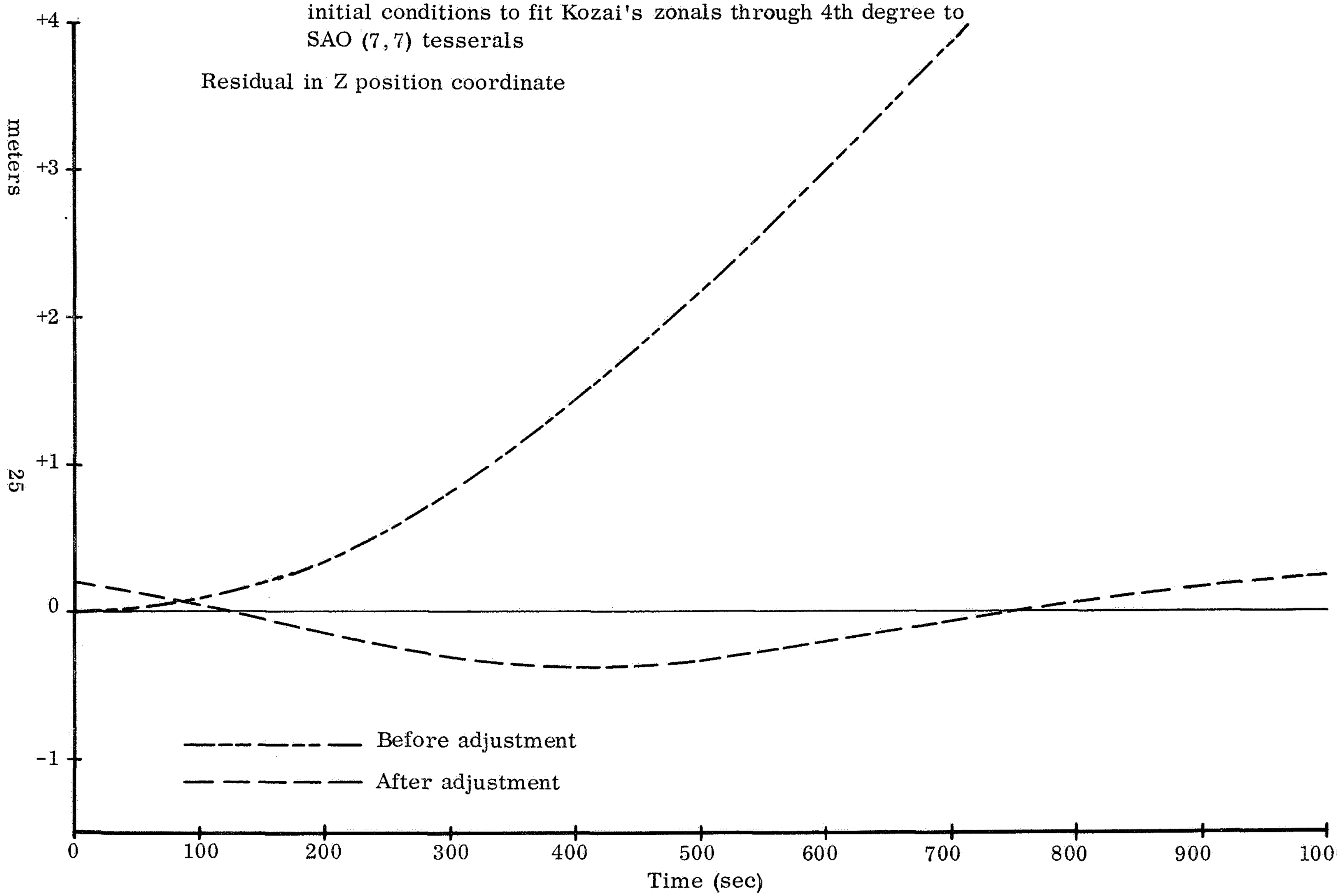


Fig. 2.5 (from [Brown, 1967b]) Least squares adjustment of initial conditions to fit Kozai's zonals through 4th degree to SAO (7, 7) tesserals

Residual in Z position coordinate



second orbit was generated using Kozai's zonal coefficients through degree four. The position differences in X, Y, Z between the two orbits were then used as the observations in a least squares adjustment in which the unknowns were the initial conditions of the second orbit. The residuals plotted are the position differences in X, Y, Z before and after adjustment. These graphs were generated by D. Brown Associates using orbital elements typical of GEOS-A. They illustrate that for sufficiently short arcs, an adjustment of the initial conditions of the orbit can compensate for the systematic error introduced by using a geopotential function limited to zonal harmonics [Brown, 1967].

Figures 2.6 — 2.8 compare a Keplerian orbit to a reference orbit obtained by using only Kozai's zonal coefficients in the same manner. These graphs were generated by this investigator at OSU using orbital elements typical of the PAGEOS satellite. The reference orbit was generated using Hartwell's method of recursive formation of partial derivatives [Hartwell, 1967]. The graphs illustrate that even after the adjustment of initial conditions, the fit of the Keplerian orbit to the reference orbit is not especially good. Since this reference orbit deviates from the actual orbit by less than a meter, the graphs may be interpreted as representing the fit of a Keplerian orbit to the actual orbit.

The differences in all three coordinates were used as observations in these adjustments. In practice all three coordinates of the satellite are rarely observed. If right ascension and declination measurements are used, only two coordinates may be said to be observed, and with range measurements only one coordinate is observed. Figures 2.9 – 2.10 illustrate that a much better

Fig. 2.6 Least squares adjustment of initial conditions to fit Keplerian orbit to Kozai's zonals through degree 4

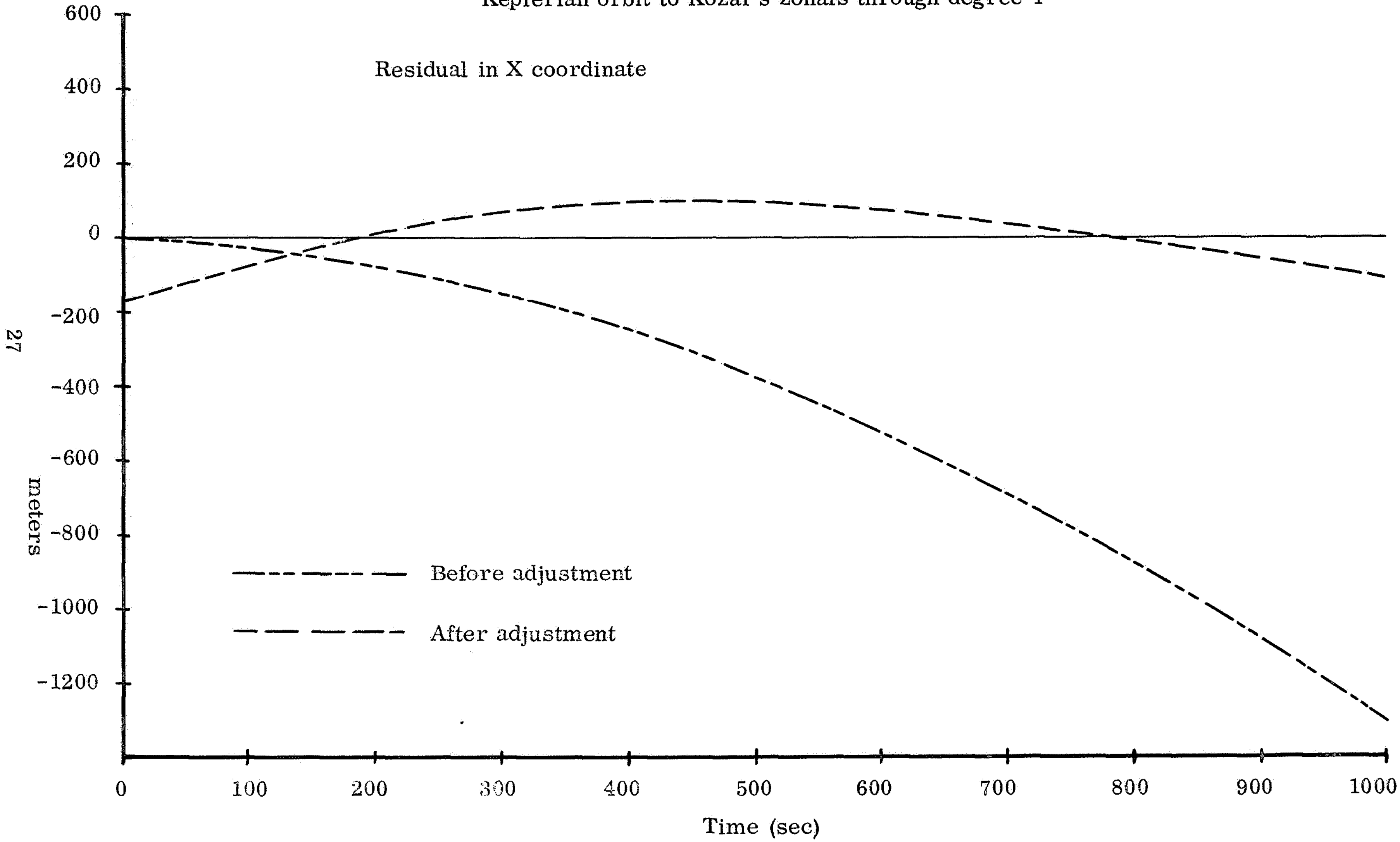


Fig. 2.7 Least squares adjustment of initial conditions to fit Keplerian orbit to Kozai's zonals through degree 4

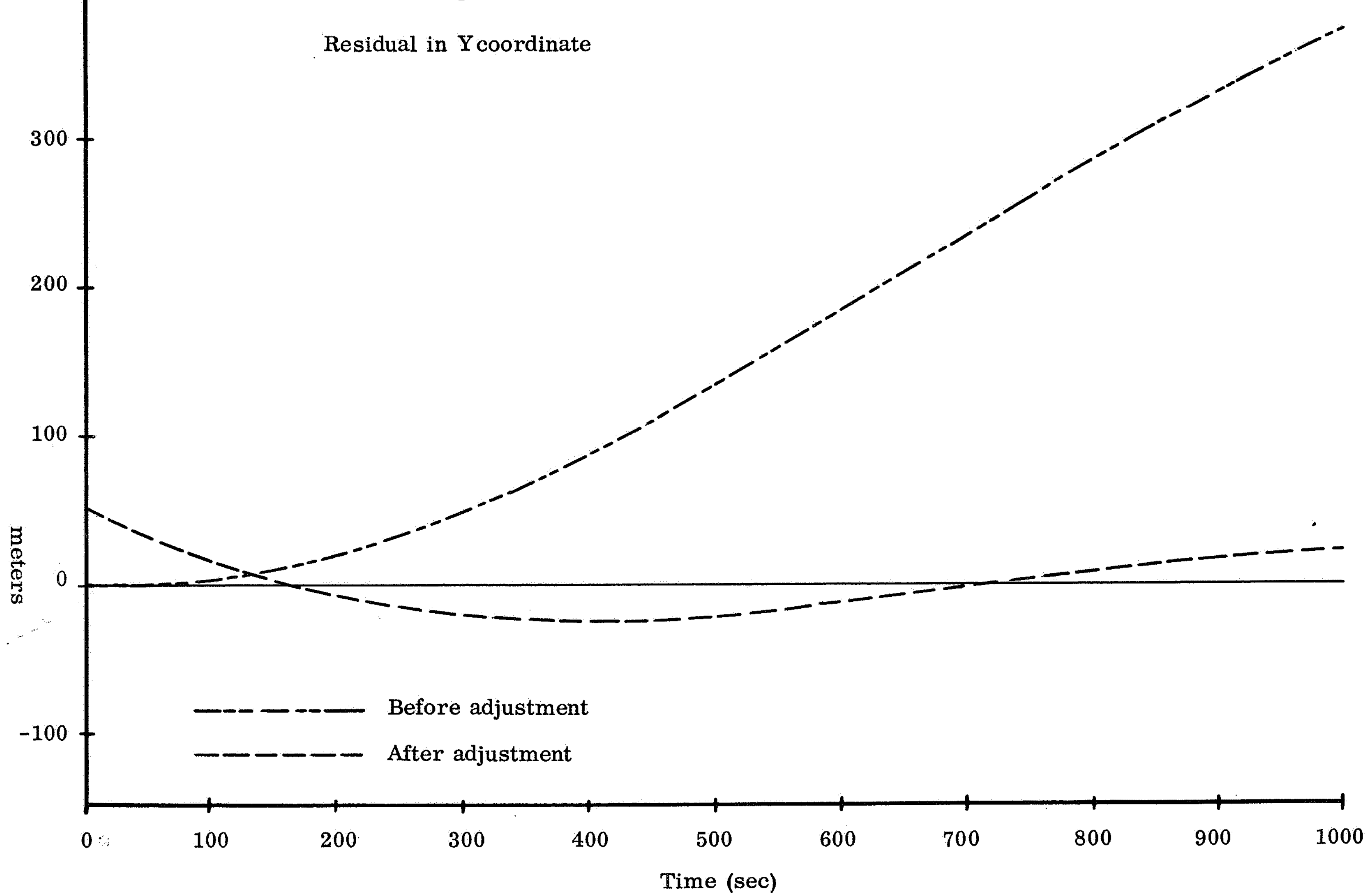


Fig. 2.8 Least squares adjustment of initial conditions to fit Keplerian orbit to Kozai's zonals through degree 4

Residual in Z coordinate

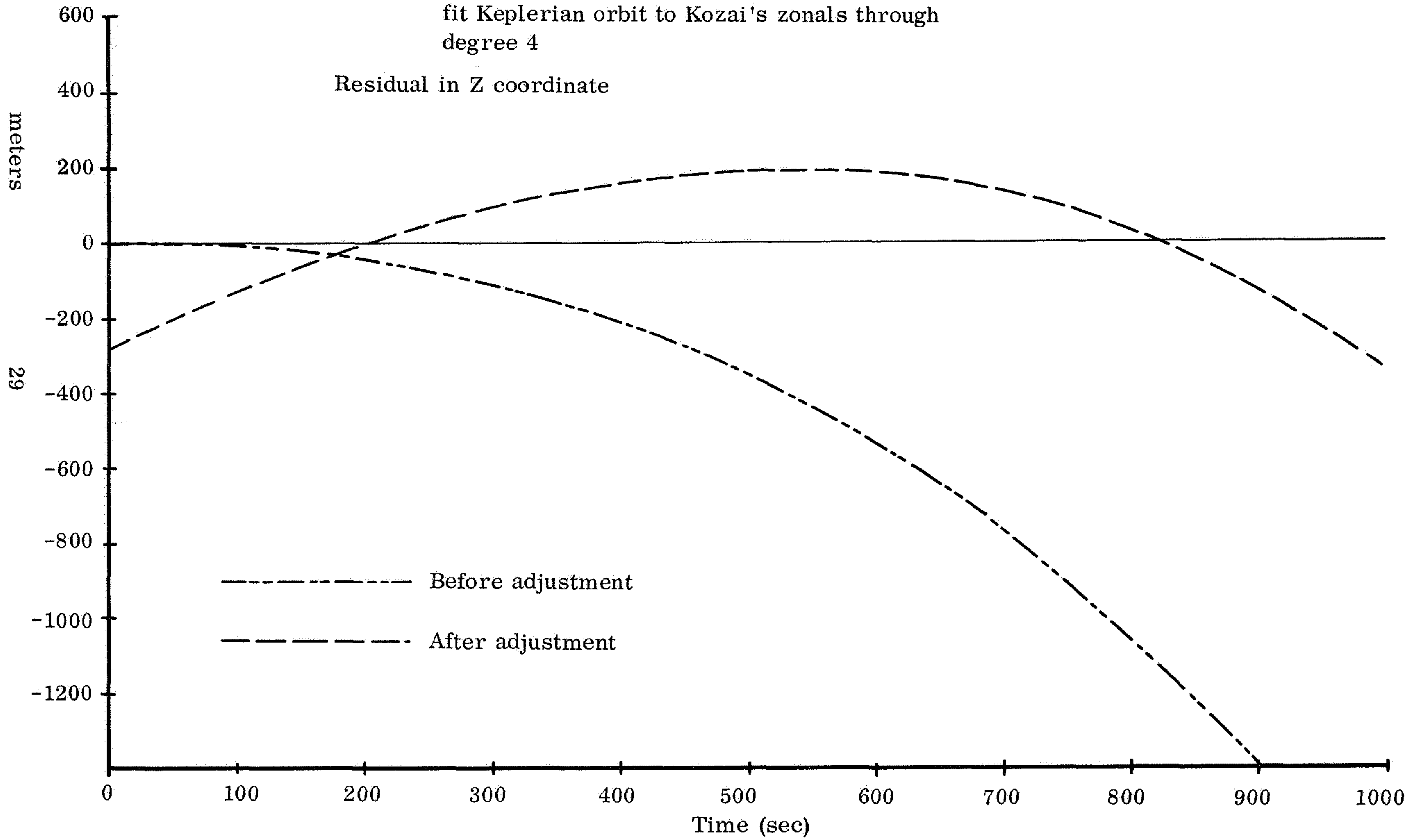


Fig. 2.9 Least squares adjustment of initial conditions to fit a Keplerian orbit to Kozai's zonals through degree 4

X and Y observed

Residual in X coordinate

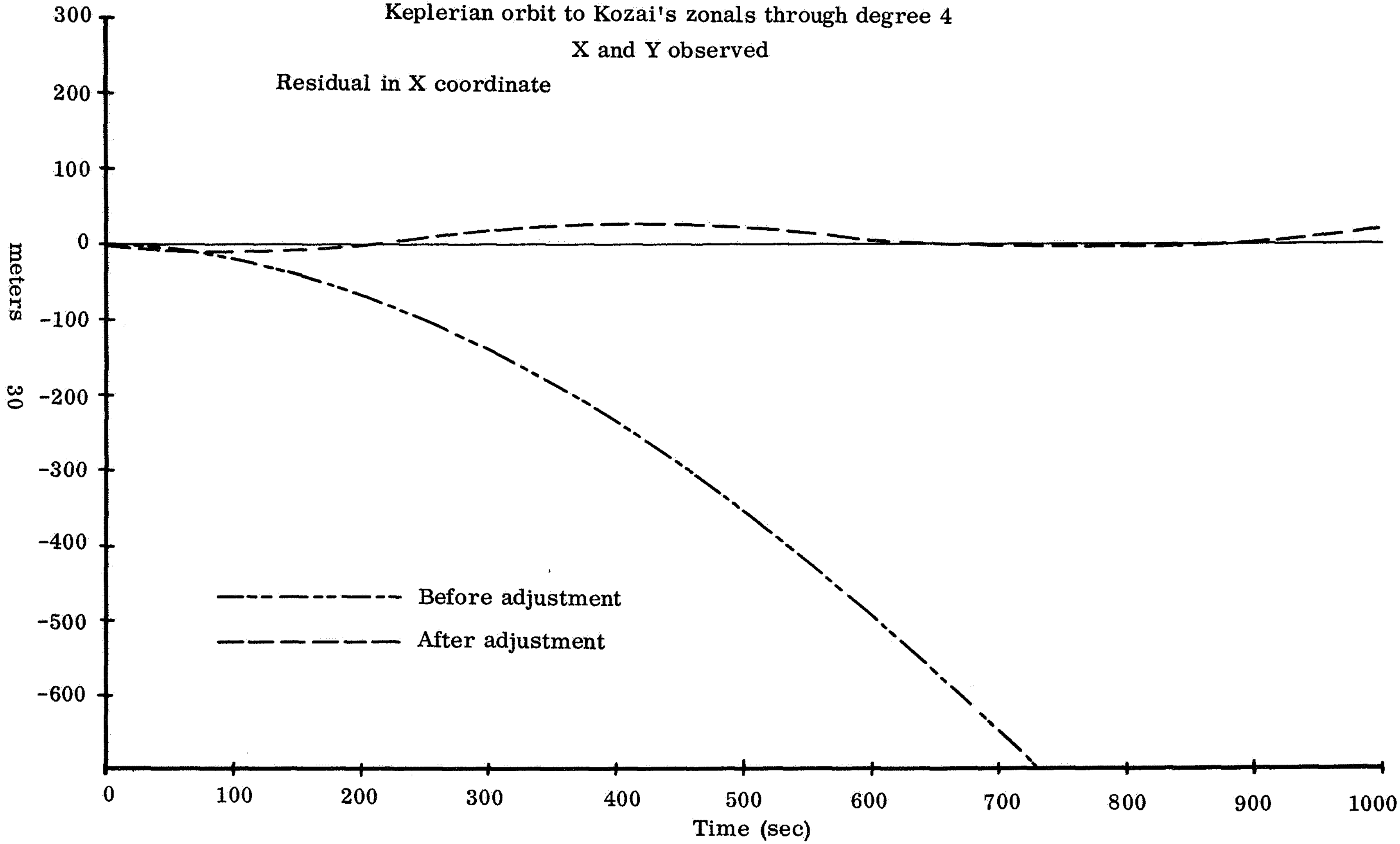
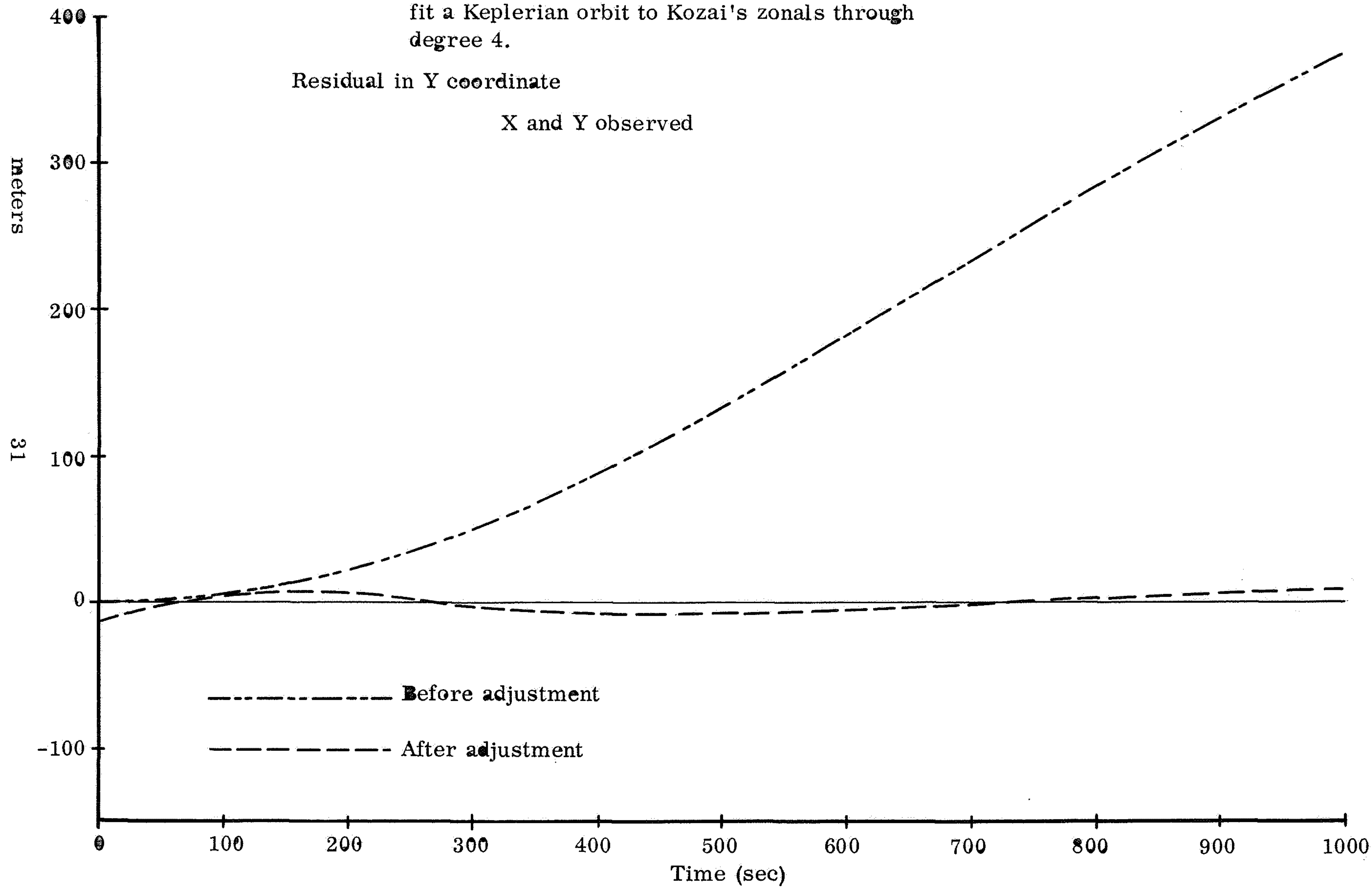


Fig. 2.10 Least squares adjustment of initial conditions to fit a Keplerian orbit to Kozai's zonals through degree 4.

Residual in Y coordinate

X and Y observed



fit may be obtained when only two coordinates are observed. Naturally, the residual after adjustment in the third coordinate is very large in this case. Putting the information contained in all of these graphs together, it is possible to conclude that a Keplerian orbit may be found which represents the actual orbit with an accuracy of a few meters or better in two coordinates (but not three) for a time period of 1000 seconds.

The modeling error term e_{jk} of equation (2.6) designates the error remaining in the modeled orbit after adjustment of the initial conditions. This is seen to be much smaller than the error of using the same model for predictions. The length of time for which this error may be kept small for any given model determines the length of arc for which the orbit constraint is valid. "Small," in this case, should mean at least an order of magnitude smaller than the observational accuracy. The observational accuracy of geodetic tracking equipment is of the order of one or two seconds of arc in direction for optical systems and several meters in range for electronic and laser systems. Therefore, the modeling error should be kept to ten meters or less for optical systems or a meter or less for electronic systems if it is to be neglected. It is now possible to offer a definition of a "short arc." A short arc is the maximum length of arc over which the modeling error of a simple model, such as the Keplerian ellipse or a model using a geopotential with only zonal terms, is not greater than a few meters. For a geodetic satellite this is about 1/8 revolution, or about 15 minutes for a 2 hour satellite such as GEOS-A [Brown, 1968]. This is also the approximate time required for the satellite to pass over a continental network of stations. Thus a "short arc" is pretty much the same as a "pass."

In practice the actual path of the satellite is broken into many short arcs or segments, and each segment is treated as an independent orbit. Thus an adjustment utilizing many independent "orbits" of the same satellite is indistinguishable from one involving many different satellites. If a satellite is tracked by a network of continental extent or less, or if the tracking stations are grouped together so that they may observe the satellite in the simultaneous mode, then the segmenting of the actual satellite path may be done in a natural way, with each orbit corresponding to a period of tracking by a group of stations.

2.4 The Effect of Modeling Errors on the Adjustment

The question of whether the modeling errors behave as random variables is very complicated, since there are so many factors contributing to the modeling error. By the nature of the least-squares solution, the mean of the modeling errors will be near zero. However, if the largest component is the truncation error, there is no reason to suspect that they behave randomly. Indeed, the modeling errors shown in Figures 2.6 – 2.8 are by no means random, but keep the same algebraic sign for long periods of time. It appears that treating the modeling error as random would have a systematic effect on the station determination, at least for a single pass. It has been argued that although the effect of the modeling error may be systematic for a single pass, the effect balances out when a large number of passes are considered. It is by no means clear in any given experiment whether this is true at all, or, if it is, how many passes must be considered in order for the systematic errors to average out to a negligibly small amount. Therefore, the only safe approach is to ensure that the modeling errors, systematic or not, will be kept much smaller than the accidental measurement errors.

There are a few error sources which may reasonably be assumed to act as random variables with zero means. Very short perturbations, whose period is shorter than the interval between successive observations of the satellite, may be treated as random variables that are uncorrelated between observations. Numerical integration error may be treated as a stochastic variable.

Errors in the adopted values of the constants in the model present a slightly different problem, since it would be possible to solve for the most likely values of these constants along with the other parameters of the adjustment. For instance, it is possible to include a finite set of the coefficients of the geopotential as parameters in the orbital mode adjustment. However, observations designed for geodetic purposes will usually give an extremely poor determination of the geopotential coefficients, especially if only short arcs are tracked. Therefore it is usually more desirable not to solve for these coefficients, but to treat them as "unestimated parameters." Since they are determined from observations, the adopted numerical values of the geopotential coefficients have associated expectancies and variances. Although the effects of errors in the adopted values may be systematic for a particular set of coefficients, they may also be treated as statistical quantities with zero expectancies and known variances. If these variances are not negligible in comparison with the observational variance, the statistics of the adjustment may not be reliable. The residuals will reflect both the accidental measurement error and the effects of modeling error, and so may not be good indicators of the quality of the observations. It can be shown that in the presence of unestimated parameters, the estimate of the variance of the observation of unit weight will be too large and the weight coefficient matrix (inverse of normal equation matrix) will be too small [Schwarz, 1967b].

2.5 Orbit Constraint Adjustment and Orbit Determination

There are many strong similarities between the orbit constraint adjustment and the method of orbit determination known in celestial mechanics as "Differential Correction." Differential Correction is a least-squares adjustment of the orbit parameters, usually with the coordinates of the tracking stations held fixed. However, if the coordinates of the tracking stations are not well known or if they are located on different datums whose relationship is not well known, then it may be necessary to include the station coordinates as unknowns in the orbit determination. In this case the system of equations to be solved will be exactly the system that arises in the orbit constraint adjustment , since the same kind of observations and same unknowns enter both problems. Thus the more general form of differential correction (including station unknowns) and the orbital constraint adjustment have identical algebraic structure; however, they differ widely in purpose, emphasis, and experimental design.

In the orbital constraint adjustment, only the station unknowns are of interest. The orbit unknowns are considered to be auxiliary, or "nuisance," parameters. The geodesist is not concerned whether the orbit determined is the correct one or not, as long as it fits the observations reasonably well. Observing networks are designed and observations are generally scheduled so as to optimize the geometry of the station determination, with little or no thought given to how well the orbit will be determined.

In constrast, the station unknowns are the nuisance parameters in the orbit determination problem. The investigator is usually not the least bit concerned about the station determination, but is very concerned with determining the correct orbit, since the orbit determination will very likely serve as the basis for predictions. The observing schedule and tracking network is generally designed to optimize the orbit determination. A strong orbit determination requires that observations be taken from points that are fairly equally spaced all the way around the orbit. Observing networks that have been designed for geodetic purposes usually serve very poorly for orbit determination purposes; although the quality of each individual observation is very high, they are all usually concentrated on a small portion of the orbit, with no observations being taken on a large part of the orbit.

2.6 Comparison of Covariance Matrices Arising from the Geometric and Orbital Mode Adjustments

A question that is often the subject of considerable discussion is whether the geometric mode or the orbital mode adjustment yields the better solution when both use exactly the same data. Intuitively, the orbital mode should give the stronger solution, since it is given more information about the real world (i.e., the fact that the various satellite positions must lie along some orbit). To prove this it is necessary to define exactly what is meant by a better adjustment.

Let there be given a set of observations of a geodetic satellite. Assume that these observations have been performed in the simultaneous mode, so that there are at least 3 observations of each satellite position. It is well known that at best the geometric observations alone will only determine a rigid network of points in space. It is necessary to fix this network to some coordinate system, usually by specifying the coordinates of one station in that coordinate system. This may be done either by completely fixing the point, or by constraining its coordinates to some a priori values with proper weights. In the first instance, the station to be fixed may be eliminated from the adjustment altogether. The solution gives positions and uncertainties of all other stations relative to the one that was fixed. For instance, if the coordinates of one station on a certain datum are held fixed, the coordinate system of the solution may reasonably be said to be the system defined by that datum. On the other hand, it is possible to specify a priori values for the coordinates of some station in a geocentric system, and to constrain the adjustment to these values with a weight matrix that is inversely proportional to the covariance matrix of this knowledge. Since these constraint equations are necessary to prevent singularity of the normal equations, it is not necessary that their weights be at all large. On the contrary, since the position of no station in a geocentric system is well known, these constraints should be fairly weak. The uncertainty of the position of the geocenter will be propagated through to all the other stations by the adjustment, so that the adjusted coordinates of all stations may be said to be geocentric. In this case, no adjusted geocentric station position is less uncertain than the constrained position.

The orbital mode adjustment does not have this problem of fixing the coordinate system. Since the center of mass of the earth lies at one of the foci of the orbital ellipse, the geocenter is determined. In practice this determination is usually quite weak, and it is best to include a priori values of the geocentric coordinates of at least one station [Brown, 1968].

In addition, the geometric mode adjustment will be singular if the scale or orientation of the coordinate system is not specified. For instance, if only range measurements are used the system will lack orientation, and if only direction observations are used the system will lack scale. In the first case three components of direction among the stations must be specified, and in the second the distance between two stations must be specified. Again, this a priori knowledge should be weighted inversely proportional to its covariance matrix.

These a priori constraint equations may be linearized around the a priori values of the station coordinates in the same manner as the observation equations. Let the total set of such linearized constraint equations be denoted $CX = 0$

and let the weight matrix of these equations be W_c. Then the total set of equations for the geometric mode adjustment is

$$\begin{pmatrix} A & B \\ C & O \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} L \\ O \end{pmatrix} + \begin{pmatrix} V \\ V_c \end{pmatrix}, \tag{2.16}$$

with weight matrix

$$\begin{pmatrix} W & O \\ O & W_c \end{pmatrix}$$

where the notation is that used in section 2.1.

The same a priori information about the stations may also be made available to the orbital mode adjustment. The total set of equations for the orbital mode adjustment is then

$$\begin{pmatrix} A & BF \\ C & O \end{pmatrix} \begin{pmatrix} X \\ Z \end{pmatrix} = \begin{pmatrix} L \\ O \end{pmatrix} + \begin{pmatrix} V \\ V_c \end{pmatrix}, \tag{2.17}$$

with the same weight matrix as in (2.16). These two sets of equations are then based on exactly the same information, and may be compared.

Let X_g be the least-squares solution for the station unknowns from the geometric mode adjustment, let X_b be the corresponding solution from the orbital mode adjustment, and let Σ_g, Σ_b be their respective covariance matrices. X_b is then said to be a better estimate of X than X_g if Σ_b is "smaller" than Σ_g. "Smaller" here means that the difference $\Sigma_g - \Sigma_b$ is positive definite, or at least positive semi-definite. This definition is entirely reasonable, since "best" means the minimum variance solution, or the solution whose covariance matrix is "smallest." This is also equivalent to saying that the trace of Σ_b is less then the trace of Σ_g, or that the rms uncertainty of the elements of X_b is less than that of the elements of X_g.

The covariance matrix of a least-square estimate of X is given by

$$\Sigma = \sigma_0^2 Q \tag{2.18}$$

where Q is the partition of the inverse of the normal equations that pertains to X, and σ_0^2 is a constant of proportionality called "the variance of the observation of unit weight." In many adjustments this constant is not sufficiently well known, and must be estimated from the residuals of the adjustment. Nevertheless, (2.18) remains the correct expression for the covariance matrix. Since the system of equations (2.16) and (2.17) express the same observations and have the same weighting matrix, the constant σ_0^2 is the same in both cases. And since σ_0^2 is a positive scalar multiplier, it immediately follows that Σ_b is smaller than Σ_g if and only if Q_b is smaller than Q_g.

As shown in section 2.1, the system of equations (2.17) for the orbital mode adjustment is equivalent to the system (2.16) for the geometric mode adjustment subject to the constraints

$$Y = FZ. \tag{2.19}$$

Since the geometric mode adjustment is assumed to exist, the orbit constraint solution may be formed sequentially; i.e., the system of equations (2.16) may be solved first and then the constraints may be added, yielding a solution which is algebraically equivalent to the least-squares solution of (2.17) [Schwarz, 1967a, p. 37]. The normal equations corresponding to the geometric mode adjustment are

$$\begin{pmatrix} A^TWA + C^TW_cC & A^TWB \\ B^TWA & B^TWB \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} A^TWL \\ B^TWL \end{pmatrix}. \tag{2.20}$$

Let N denote the matrix of coefficients of these normal equations and let M denote its inverse. Further let M be partitioned in the same manner as N, i.e.,

$$N = \begin{pmatrix} A^TWA + C^TW_cC & A^TWB \\ B^TWA & B^TWB \end{pmatrix} = \begin{pmatrix} N_1 & N_2 \\ N_2^T & N_3 \end{pmatrix} ,$$

$$M = N^{-1} = \begin{pmatrix} M_1 & M_2 \\ M_2^T & M_3 \end{pmatrix} .$$

M_1 corresponds to the unknowns X and may be identified with Q_g. Explicitly,

$$M_1 = Q_g = [A^TWA + C^TW_cC - A^TWB(B^TWB)^{-1}B^TWA]^{-1} . \tag{2.21}$$

To add the constraint to the solution of (2.20), the constraint equation (2.19) must be expressed in terms of the same unknowns X and Y. For this purpose let p be the number of satellite position unknowns and r the number of orbit unknowns. Then F is a $p \times r$ matrix. For the orbital mode solution to exist at all, it is necessary that r be less than or equal to p and that F have rank r. Let $s = p - r$ and let F be partitioned into a nonsingular r-square matrix F_1 and an $s \times r$ matrix F_2. Let Y be similarly partitioned into Y_1 and Y_2, i.e.,

$$F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} ; \qquad Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} .$$

Equation (2.19) may now be written as the two equations

$$Y_1 = F_1Z,$$
$$Y_2 = F_2Z.$$

Solving the first of these for Z, and substituting into the second,

$$Z = F_1^{-1} Y_1 ,$$

$$Y_2 = F_2 F_1^{-1} Y_1 ,$$

or

$$\begin{pmatrix} O & -F_2 F_1^{-1} & I \end{pmatrix} \begin{pmatrix} X \\ Y_1 \\ Y_2 \end{pmatrix} = O. \tag{2.22}$$

Let $G = (-F_2 F_1^{-1} \quad I)$ so that (2.22) may be written

$$(O \quad G) \begin{pmatrix} X \\ Y \end{pmatrix} = O. \tag{2.23}$$

Equation (2.23) expresses the orbital constraint in the unknowns X and Y, so that it may be added sequentially to the solution of (2.20). The weight coefficient matrix of the solution to (2.20) is M. Let the weight coefficient matrix after the addition of the constraints be denoted M', and identify Q_b with an upper left partition of M'. Then M' is given by [Schwarz, 1967a, p. 38; Uotila, 1967, p. 65]

$$M' = M - M \begin{pmatrix} O \\ G^T \end{pmatrix} \left[(O \; G) M \begin{pmatrix} O \\ G^T \end{pmatrix} \right]^{-1} (O \quad G) M. \tag{2.24}$$

Carrying through the indicated multiplications

$$M' = M - M \begin{pmatrix} O & O \\ O & G^T (G M_3 \, G^T)^{-1} G \end{pmatrix} M$$

$$= M - \begin{pmatrix} M_2 G^T (G M_3 G^T)^{-1} G M_2^T & \cdots \\ \cdots & \cdots \end{pmatrix} . \tag{2.25}$$

Only the upper left partition of the second matrix on the right is indicated, since it is the only one of immediate interest. Finally, writing only the upper left partition of (2.25),

$$Q_b = M_1' = Q_g - M_2 G^T (GM_3 G^T)^{-1} GM_2^T,$$

or

$$Q_g - Q_b = M_2 G^T (GM_3 G^T)^{-1} GM_2^T. \qquad (2.26)$$

By tracing through the definitions of M_3, G, and M_2, it may be easily seen that GM_3G^T is positive definite and that $Q_g - Q_b$ is positive semi-definite.

This shows that Q_b is "smaller" than Q_g. It means that if the modeling errors are negligible, so that the orbit constraint solution is a proper minimum variance adjustment, then the orbital mode adjustment is "better" than the geometric mode in the sense that it will produce an estimate of the station positions with a smaller covariance matrix.

In practice, it is traditional to estimate the unit variance σ_0^2 from the residuals of the adjustment rather than rely on an a priori value. It may easily happen that the estimated σ_0^2 from the orbital mode adjustment is larger than that from the geometric mode adjustment, and the orbital mode adjustment may thus give a "larger" covariance matrix. However, it is important to note that if the modeling errors are negligible, then both adjustments will yield unbiased estimates of σ_0^2. Let V_g be the residuals from the geometric mode adjustment and V_b those from the orbital mode adjustment, i.e.,

$$V_g = AX_g - L,$$
$$V_b = AX_b - L.$$

Let n be the number of observations and m the number of stations. Then the degrees of freedom are $n - m - p$ in the geometric case and $n - m - r$ in the orbital case.

Further,

$$\epsilon\left[\frac{V_g^T W V_g}{\sigma_0^2}\right] = n - m - p,$$

and

$$\epsilon\left[\frac{V_b^T W V_b}{\sigma_0^2}\right] = n - m - r,$$

so that

$$\epsilon\left[\frac{V_g^T W V_g^T}{n - m - p}\right] = \epsilon\left[\frac{V_b^T W V_b^T}{n - m - p}\right] = \sigma_0^2 .$$

Also,

$$\frac{\epsilon\,[V_g^T W V_g^T]}{\epsilon\,[V_b^T W V_b^T]} = \frac{n - m - r}{n - m - p} \leq 1.$$

This means that the addition of the orbital constraint would cause some of the residuals, and the weighted sum of squares of residuals, to increase from their values in the geometric solution. However, the decrease in the degrees of freedom exactly compensates for this increase, so that the expectancy of the estimated σ_0^2 remains the same.

This development yields a valuable tool for recognizing the presence of modeling errors in the adjustment. The adjustment of a large sample of data in the geometric mode will yield a fairly good idea of what the quality of that kind of data really is. This knowledge may be used to form an a priori value for the unit variance of that kind of data. Then, when more data of the same kind is adjusted in the orbital mode, the estimated unit variance may be tested against its a priori value by the χ^2 test. If the difference is significant for several different sets of data, the cause may very likely be the presence of significant inadequacies in the model for the orbit.

2.7 Summary of the Advantages and Disadvantages of the Orbital Mode Adjustment

In the preceding sections many features of the orbital mode adjustment have been discussed. Many of these may be considered to be advantages or disadvantages when compared to the alternative geometric mode adjustment. The advantages are:

(1) The orbital mode brings more information about the real world into the adjustment. Therefore one would intuitively expect that it would yield a better determination of the station positions. This intuition is correct, if "better" is interpreted to mean possessing a "smaller" covariance matrix.

(2) The orbital mode adjustment contains fewer unknowns, and therefor more degrees of freedom. This means that, apart from other considerations, the statistics of the adjustment will be more reliable. If the total set of normal equations is stored and solved by elimination methods, the fewer unknowns will also mean less storage requirements and fewer arithmetic operations. However, most investigators take advantage of the patterned form of the normal equations and easily reduce the problem to one involving only the station unknowns, so that the number of auxiliary unknowns ceases to be a consideration.

In some cases, such as the measurement of integrated Doppler count, the number of unknowns may grow almost as fast as the number of observations. In such a case it may be necessary to reduce the number of unknowns by imposing an orbital constraint if an overdetermined solution is to exist at all.

(3) The orbital mode adjustment gives a solution in a geocentric rather

than a relative coordinate system. It also gives information about the scale of the system through the adopted value of the constant GM. However, the determination of the geocenter and the scale by short arcs is too weak to be an important advantage.

(4) All observations from a given station, whether or not matched with observations from other stations, are potentially useful. This advantage would be of greater importance if more unmatched data of geodetic quality were available. Also, there is a sufficient abundance of data, almost all of it taken in the simultaneous mode, from almost all stations that have performed observations for geodetic purposes.

The disadvantages of the orbital mode adjustment are:

(1) It is necessary to construct a model for the orbit. The actual motion of a close earth satellite is extremely complex, and any attempt to describe the motion over a long arc becomes very involved. If only a simple model, such as the Keplerian ellipse, is used, then the modeling error is an important consideration, and one must break the actual path into short arcs, whose duration does not exceed 1/8 of the satellite's period.

(2) Some hypothesis about the nature of the modeling error must be made. It may take a great deal of experimentation to determine the nature of the modeling error for a given model, given satellite, given type of data, and given tracking station configuration.

(3) Approximate values of the orbit unknowns must be available. The preliminary determination of the orbit will usually require a separate program and

a good deal of additional data handling. In the case of very short arcs, the preliminary orbit may be difficult to determine and unreliable.

3. Computational and Programming Considerations

There are many different forms which the mathematical structures for the orbital mode adjustment may take. Aside from the considerations discussed in section 2.3, orbit models may differ in the set of parameters used to characterize the orbit and the specific sequence of transformations used to update the satellite position. Furthermore, mathematical structures differ in the specific sequence of transformations used to obtain the computed value of the observed quantity. In this chapter the FORTRAN IV computer program which was used to perform the short-arc experiments described in Chapter 4 will be described and discussed. This program computes differential corrections to station positions and orbit unknowns, based on observations of topocentric right ascension, declination and range. The orbit is modeled by empirical expressions of the type used by the Smithsonian Astrophysical Observatory. The program accepts any mixture and number of range and direction observations, any number of orbits, and up to 150 observing stations.

3.1 Explicit Form of the Orbital Model

An orbit model only slightly more complicated than the Keplerian ellipse is the empirical orbit. This model allows the Keplerian elements to undergo secular and periodic changes with time. These variations are determined strictly empirically; the gravity field of the earth does not appear in the model and there is no integration of equations of motion involved. Empirical orbits have been advocated principally at the Smithsonian Astrophysical Observatory and now form the basis of the SAO Differential Orbit Improvement program [Gaposchkin,1966]. Empirical orbits are also used by the SAO in their Ephemeris 6 predictions, which are published for many different satellites, and by the USCGS in their in-house predictions for the ECHO and PAGEOS satellites.

In an empirical orbit only five, rather than six, elements are used. These are usually taken to be the Keplerian elements ω (argument of perigee), Ω (right ascension of ascending node), i (inclination), e (eccentricity) and M (mean anomaly). Each of these is defined by an equation of the form

$$E_i = E_{i0} + E_{i1}(T - T_0) + E_{i2}(T - T_0)^2 + E_{i3}(T - T_0)^3 + E_{i4}\cos\overline{\omega} + E_{i5}\sin\overline{\omega} \quad (3.1)$$

where E_i is a general symbol for any of the set $\{\omega, \Omega, i, e, M\}$, and T and T_0 are time and reference time respectively; $\overline{\omega}$ is obtained by evaluating only the polynomial part of the expression for ω. Early expressions for empirical orbits were more general than this, allowing polynomial terms of arbitrary degree, trigonometric terms of arbitrary amplitude, period, and phase, and even hyperbolic terms [Mueller, 1964, p. 214]. The form above is somewhat more pleasing than a completely general expression, since it is the form that would be obtained by integrating the Lagrange planetary equations for secular and long period effects.

The coefficients E_{ij} in (3.1) are found by a least squares fit to the data being processed. Thus the elements E_i are "mean" elements only in the sense of "best fitting to this particular set of data."

Since there are six coefficients in each of five expressions, the model allows as many as 30 orbit unknowns. These coefficients characterizing the orbit are usually called the "orbit parameters" to distinguish them from the elements. In practice, fewer than 30 parameters are used. The set $\{\omega_0, \omega_1, \Omega_0, \Omega_1, i_0, e_0, e_1, M_0, M_1, M_2\}$ is normally used to characterize arcs lasting two weeks by the SAO and the USCGS. This set of ten parameters has been found to describe the orbit for several weeks with accuracy sufficient for prediction purposes (several minutes of arc along track and less than a minute of arc across track).

This form of orbit model was used by previous investigators at The Ohio State University because it was thought that it would be more accurate than the Keplerian orbit and almost as simple. The approach of the short arc solution described in [Rapp, 1967] is to constrain the parameters $\{\omega_1\ \Omega_1, e_1, M_2\}$ at values published in the SAO ephemerides, and to solve only for the six parameters $\{\omega_0, \Omega_0, i_0, e_0, M_0, M_1\}$. The mean motion is then computed from

$$n = \frac{dM}{dT} \text{ (polynomial part only)}$$

and the semi-major axis from the modified Third Law of Kepler [Kozai, 1960]

$$a = \left(\frac{GM}{n^2}\right)^{\frac{1}{3}} \left[1 + \frac{1}{3} \frac{J}{p^2} \sqrt{1-e^2}\,(-1 + \frac{3}{2} \sin^2 i)\right]$$

where

$$p = \left(\frac{GM}{n^2}\right)^{\frac{1}{3}} (1-e^2) \text{ and } J = \frac{3}{2}\, a_e^2\, J_2$$

Early experiments by this investigator with data from GEOS-A and PAGEOS indicated that the precision of fit over a short arc was not perceptibly improved by inclusion of the secular variations $\{\omega_1, \Omega_1, e_1, M_2\}$. In most of the experiments described in the next chapter, all parameters except the set $\{\omega_0, \Omega_0, i_0, e_0, M_0, M_1\}$ were set to zero and these six were considered to be unknowns. Used in this manner, the empirical orbit is exactly the same as the Keplerian orbit, except that the modified form of Kepler's third law is used.

The satellite position is computed for any time T by the sequence of calculations diagrammed in Fig. 3.1. The orbit elements, and thus the coordinates α_m, δ_m, R_m, are taken as being in the modified sidereal coordinate system described in [Veis, 1963, p. 12]. This system is almost inertial, and differs from the true sidereal coordinate system of date only by the amount of the precession since 1950.0 and the nutation in right ascension.

3.2 Explicit Form of the Observation Equations for Range and Direction Observations.

Direction observations are assumed to be given in the topocentric sidereal coordinate system, defined by the true equator and equinox of date. This coordinate system was selected because it is the stipulated system for direction observations deposited in the National Space Science Data Center by participants in the National Geodetic Satellite Program. Range observations are, of course, independent of coordinate system, although some coordinate system must be selected for intermediate computations.

The notations used for the various coordinate systems are:

$u, v, w,$	—	average terrestrial coordinate system.
$u', v', w',$	—	instantaneous terrestrial coordinate system, differing from the average system only by polar motion.
ξ, η, ζ or x, y, z	—	geocentric sidereal coordinates, related to u', v', w', by an R_3 rotation through the Greenwich Apparent Sidereal Time.
x', y', z'	—	topocentric sidereal coordinates.
$\alpha, \delta, R; \alpha', \delta', R'$	—	spherical coordinates, referred to the geocentric and topocentric sidereal coordinate systems respectively.
α_m, δ_m, R_m	—	spherical coordinates in the modified sidereal (orbital) system.

The transformations between these systems are all found in [Veis, 1963].

The form of the linearized observation equation is

$$V + \delta L = A\delta x + B\delta b \tag{3.2}$$

where $\delta L = (\delta\alpha', \delta\delta', \delta R')^T$ is the vector of "observed minus computed" discrepancies in α', δ', R'; $dx = (\delta u, \delta v, \delta w)^T$ is the vector of corrections to the approximate terrestrial station coordinates; and δb is the vector of corrections to the approximate orbital parameters. The vector b of orbit unknowns has a variable length in the program. It may consist of any subset of as many as 21 of the 30 orbit parameters, and may be different for each orbit, although for short arc applications it was almost always specified to consist of the six parameters $b = \{\omega_0, \Omega_0, i_0, e_0, M_0, M_1\}$. The matrices A and B are

$$A = \frac{\partial(\alpha', \delta', R')}{\partial(u, v, w)}; \qquad B = \frac{\partial(\alpha', \delta', R')}{\partial b}.$$

If right ascension and declination observations are given, only the first two equations of (3.2) are generated; if range is observed, only the last equation is formed. Equation (3.2) has the same form and meaning (although different notation) as equation (2.11) in Chapter 2.

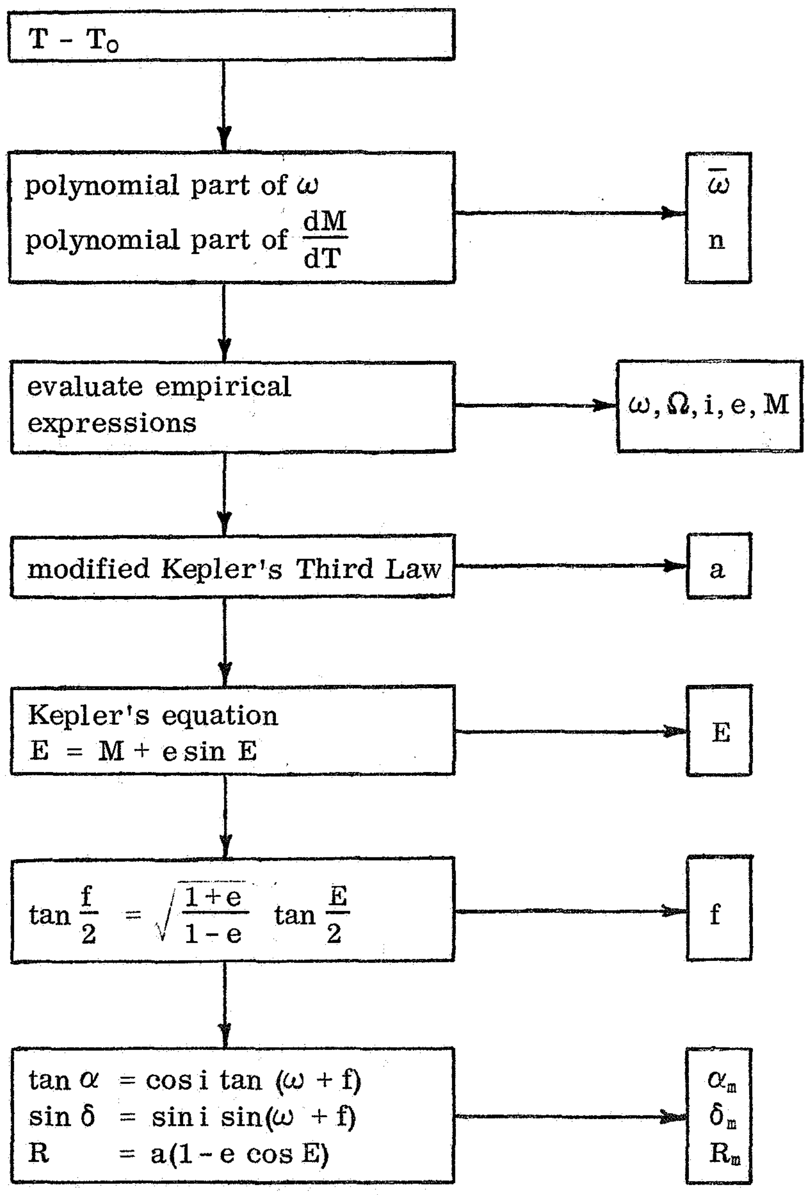


Fig. 3.1 Computation of satellite position from empirical expressions

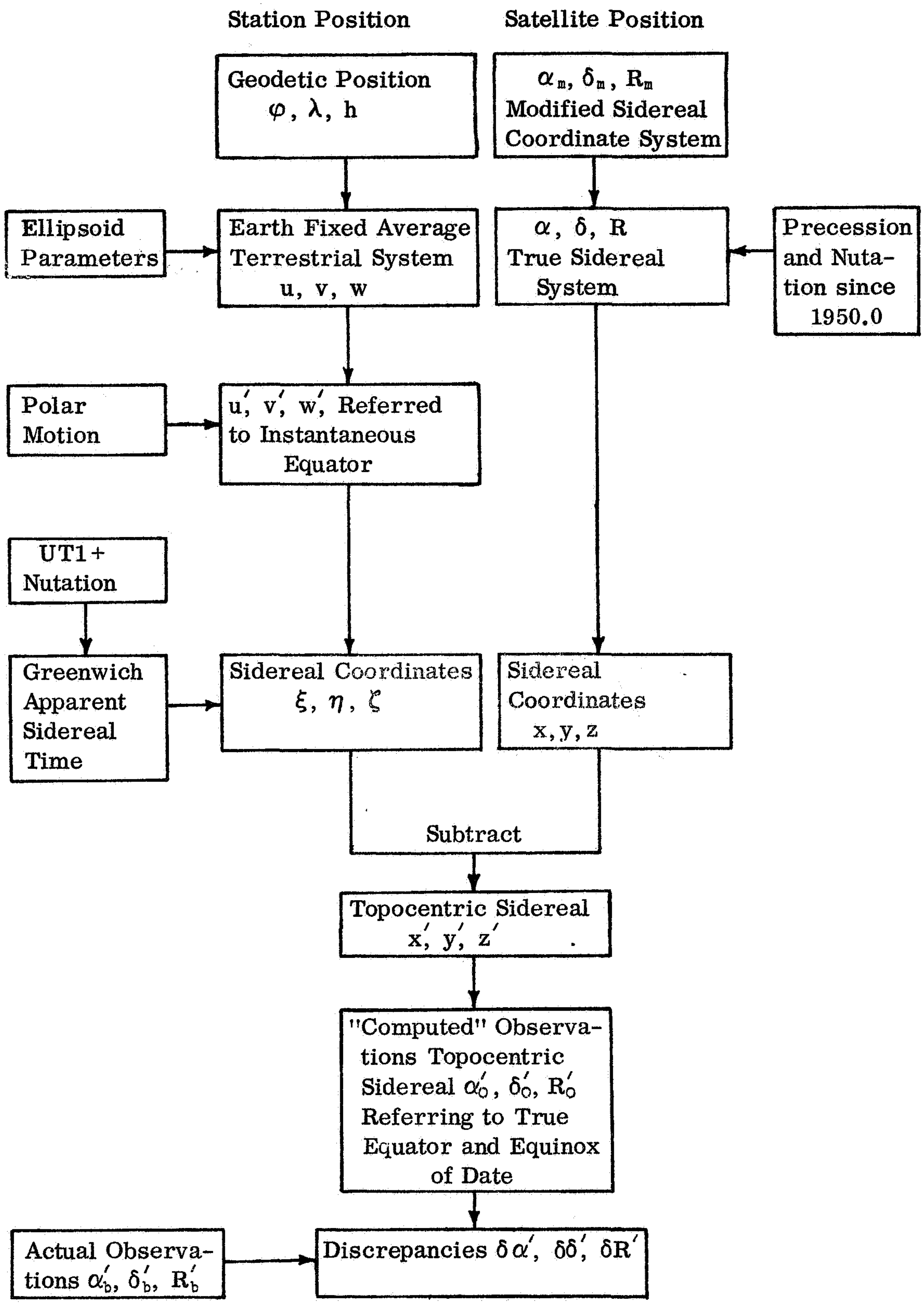


Fig. 3.2 Sequence of computations leading to "observed minus computed" discrepancies

The sequence of observations leading to the "observed minus computed" discrepancies is outlined in Figure 3.2. The A and B matrices are computed as products of matrices

$$A = -Q \times P; \qquad\qquad B = Q \times S \times U \times O. \tag{3.3}$$

The component matrices of these products are:

$$Q = \frac{\partial(\alpha', \delta', R')}{\partial(u', v', w')} ,$$

$$S = \frac{\partial(u', v', w')}{\partial(\alpha, \delta, R)} ,$$

$$U = \frac{\partial(\alpha, \delta, R)}{\partial(\omega, \Omega, i, e, M, a)} ,$$

$$O = \frac{\partial(\omega, \Omega, i, e, M, a)}{\partial b} ,$$

$$P = \frac{\partial(u', v', w')}{\partial(u, v, w)} .$$

The individual components of these matrices are given in the Appendix.

The matrix O, and thus B, has a variable number of columns, depending on the number of unknowns associated with the orbit in question. The computer program allows these matrices to have up to 21 columns.

The advantage of the arrangement of matrices in equation (3.3) is that the matrices P and the product $SUO = S \times U \times O$ may be formed by the same block of coding for both direction and range measurements. The program then branches, forming only the first two rows of Q for direction observations and only the last row in the case of a range observation. The resulting partition of

Q is then multiplied into P and SUO. A further advantage is that P and SUO depend only on the time and the orbit, not on the station. The program requires that observations be grouped by orbit, but arrangement within orbit is arbitrary. However, if observations are grouped by epoch within the orbit, it is only necessary to compute the P and SUO matrices once for each epoch. This results in considerable savings of time in some cases, since when observations are taken in the simultaneous mode, two or more stations will observe at each epoch. The Q matrix depends on the particular station involved and must be computed anew for each observing station.

3.3 Accumulation of Normal Equations and Elimination of Orbit Unknowns.

As in the geometric mode adjustment, the orbital mode adjustment leads to a very highly patterned form of normal equations. In order to discuss this pattern, it is necessary to devise a system of subscripts. Let

> i be an index denoting the ith station,
>
> k be an index denoting the kth orbit,
>
> j_k be an index denoting the jth time epoch within the kth orbit.

Also let,

> s be the total number of stations ($1 \leq i \leq s$),
>
> p be the total number of orbits ($1 \leq k \leq p$),
>
> e_k be the total number of time epochs in the kth orbit ($1 \leq j_k \leq e_k$).

By attaching subscripts to equation (3.2), the observation equation for the observation(s) made on the kth orbit at the j_kth epoch from the ith station is written

$$A_{ijk}\ \delta X_i + B_{ijk}\ \delta b_k = \delta L_{ijk}\ .$$

The weight matrix for this observation is W_{ijk}, where

$$W_{ijk} = \begin{pmatrix} \text{var}\ (\alpha') & \text{cov}\ (\alpha',\ \delta') \\ \text{cov}\ (\alpha'\ \delta') & \text{var}\ (\delta') \end{pmatrix}^{-1}$$

for a right ascension and declination pair, and $W_{ijk} = (\text{var}\ (R'))^{-1}$ for a range observation.

The pattern of the observation equations is displayed in Figure 3.3. The pattern of the resulting normal equations is shown in Figure 3.4, where the notation popularized by Duane Brown is used. The partitions of the normal equations are given by

$$\left.\begin{aligned}
\dot{N}_i &= \sum_{k=1}^{p} \sum_{jk=1}^{e_k} A^T_{ijk}\ W_{ijk}\ A_{ijk} \\
\overline{N}_{ik} &= \sum_{jk=1}^{e_k} A^T_{ijk}\ W_{ijk}\ B_{ijk} \\
\ddot{N}_k &= \sum_{jk=1}^{e_k} \sum_{i=1}^{s} B^T_{ijk}\ W_{ijk}\ B_{ijk} \\
\dot{K}_i &= \sum_{k=1}^{p} \sum_{jk=1}^{e_k} A^T_{ijk}\ W_{ijk}\ L_{ijk} \\
\ddot{K}_k &= \sum_{jk=1}^{e_k} \sum_{i=1}^{s} B^T_{ijk}\ W_{ijk}\ L_{ijk}
\end{aligned}\right\} . \tag{3.4}$$

In adjustments performed on computers it is almost always convenient to skip the formation of the full observation and normal equation matrices, proceeding directly to the formation of the nonzero partitions of the normal equations. In this case all of the nonzero partitions are indicated in (3.4). As each observation is processed, its contributions to the sums indicated in

(3.4) are computed and added in. Space is reserved in the computer for the maximum of 150 $\dot{N}_i$ and $\dot{K}_i$ blocks, but only for one each of the $\ddot{N}_k$, $\overline{N}_k$, and $\ddot{K}_k$ blocks. Since all of the observations for the kth orbit are grouped together, this space is used to accumulate the $\ddot{N}_k$, $\overline{N}_k$, and $\ddot{K}_k$ blocks. These blocks are then written on tape and the space released to process the next orbit. This allows the processing of an unlimited number of orbits. Each individual observation equation is saved on another tape to be used in the subsequent residual analysis.

The largest block for which space must be reserved is the $\overline{N}_k$ block, whose size is 450 × 21. It would be possible to conserve even more computer space by reserving room for only one (3 × 21) $\overline{N}_{ik}$ subblock. However, this would require that the input be ordered so that all the observations from each station are grouped together within each orbit. This in turn would mean that input decks that were ordered for processing in the geometric mode would have to be reordered for processing in the orbital mode. Since it was desired to compare geometric and orbital mode solutions for several sets of data, and since enough space was available in the computer, it was decided to reserve space for the larger $\overline{N}_k$ block.

There is also the consideration that many of the $\overline{N}_{ik}$ subblocks within each $\overline{N}_k$ may be zero, and thus the program may have to process a large number of zero blocks. Each $\overline{N}_{ik}$ is nonzero only if the ith station made at least one observation on the kth orbit. It was felt that in most applications enough blocks would be nonzero so that the effort saved by not processing zero

δX_1	$\delta X_2 \ldots\ldots$	δX_s	δb_1	$\delta b_2 \ldots\ldots$	δb_p	
A_{111}			B_{111}			L_{111}
	A_{211}		B_{211}			L_{211}
			$\vdots$			$\vdots$
		A_{s11}	B_{s11}			L_{s11}
A_{121}			B_{121}			L_{121}
			$\vdots$			$\vdots$
		A_{s21}	B_{s21}			L_{s21}
			$\vdots$			$\vdots$
A_{1e1}			B_{1e1}			L_{1e1}
			$\vdots$			$\vdots$
		A_{se1}	B_{se1}			L_{se1}
A_{112}				B_{112}		L_{112}
				$\vdots$		$\vdots$
		A_{s12}		B_{s12}		L_{s12}
		$\vdots$		$\vdots$		$\vdots$
		A_{se2}		B_{se2}		L_{se2}
						$\vdots$
A_{11p}					B_{11p}	L_{11p}
					$\vdots$	
		A_{sep}			B_{sep}	L_{sep}

Fig. 3.3 Form of observation equations

$$
\begin{array}{cccc|cccc|c}
\delta X_1 & \cdots & \cdots & \delta X_s & \delta b_1 & \delta b_2 & \cdots & \delta b_p & \\
\underline{\dot{N}_1} & O & \cdots & O & \bar{N}_{11} & \cdots & \cdots & \bar{N}_{1p} & \dot{K}_1 \\
 & \underline{\dot{N}_2} & \cdots & O & \bar{N}_{21} & \cdots & \cdots & \bar{N}_{2p} & \dot{K}_2 \\
 & & \ddots & \vdots & \vdots & & & \vdots & \vdots \\
 & & & \underline{\dot{N}_s} & \bar{N}_{s1} & \cdots & \cdots & \bar{N}_{sp} & \dot{K}_s \\
\hline
 & & & & \underline{\ddot{N}_1} & O & & O & \ddot{K}_1 \\
 & & & & & \underline{\ddot{N}_2} & & O & \ddot{K}_2 \\
 & & & & & & \ddots & \vdots & \vdots \\
 & & & & & & & \ddot{N}_p & \ddot{K}_p
\end{array}
$$

or

$$
\begin{array}{c|cccc|c}
\delta x & \delta b_1 & \delta b_2 & \cdots & \delta b_p & \\
\dot{N} & \bar{N}_1 & \bar{N}_2 & \cdots & \bar{N}_p & \dot{K} \\
\hline
 & \underline{\ddot{N}_1} & O & \cdots & O & \ddot{K}_1 \\
 & & \underline{\ddot{N}_2} & \cdots & O & \ddot{K}_2 \\
 & & & \ddots & \vdots & \vdots \\
 & & & & \ddot{N}_p & \ddot{K}_p
\end{array}
$$

or

$$
\left(\begin{array}{c|c} \dot{N} & \bar{N} \\ \hline \bar{N} & \ddot{N} \end{array}\right)
\left(\begin{array}{c} \delta X \\ \hline \delta b \end{array}\right)
=
\left(\begin{array}{c} \dot{K} \\ \hline \ddot{K} \end{array}\right)
$$

Fig. 3.4 Form of normal equations

blocks would be more than matched by the extra bookkeeping effort that would have been made necessary.

Once all of the blocks have been formed, it is possible to eliminate the orbit unknowns and form a reduced set of normal equations which involve only the station unknowns. These are given by

$$(\dot{N} - \overline{N}\ \ddot{N}^{-1}\ \overline{N}^{T})\ \delta X = \dot{K} - \overline{N}\ \ddot{N}^{-1}\ \ddot{K} \tag{3.5}$$

where

$$\overline{N}\ \ddot{N}^{-1}\ \overline{N}^{T} = \sum_{k=1}^{p} \overline{N}_k \ddot{N}_k^{-1} \overline{N}_k^{T},$$

$$\overline{N}\ \ddot{N}^{-1}\ \ddot{K} = \sum_{k=1}^{p} \overline{N}_k \ddot{N}_k^{-1} \ddot{K}_k .$$

The dimension of the reduced normal equations is equal to the total number of station unknowns, and these equations may be formed in one pass over the tape containing the $\ddot{N}_k$, $\overline{N}_k$, and $\ddot{K}_k$ blocks only if the whole matrix of reduced normal equations will fit into the machine. Although the different blocks of normal equations will be formed for up to 150 stations, the program proceeds to a solution only if the networks being adjusted contain 30 or fewer stations.

After this step, the contributions to the normal equations generated by a priori constraints on station positions and constraints among stations may be added to the reduced normal equations.

In the geometric mode, an off-diagonal block of the reduced normal equations will be zero only if the two stations to which it corresponds never observed any satellite position together. Because of the restrictions imposed by the requirement for intervisibility, it is quite likely that there will be many zero blocks. In the orbital mode, the same rule holds, but "any satellite position" is replaced by "any orbit." This means that there will be a great many more nonzero blocks than in the corresponding geometric mode adjustment. In fact, for networks of continental extent or smaller, there will very likely be no zero blocks in the reduced normal equations.

Since it is not likely to have any special properties, equation (3.5) is solved for δX by direct inversion of the matrix of reduced normal equations. The solution for the orbit unknowns is then obtained spearately for each orbit from

$$\delta b_k = \ddot{N}_k^{-1} (\ddot{K}_k - \overline{N}_k^T \delta X).$$

Finally, the tape containing the observation equations is read, the residuals for each equation are computed, and the customary statistical analyses are performed.

3.4 Inclusion of Constraints Among Station Unknowns

For the purpose of this discussion, let the $\dot{N}$ matrix be partitioned into 3×3 subblocks $\dot{N}_{1j}$, where the notation $\dot{N}_{11}$ corresponds to $\dot{N}_1$ in the previous section. Although the constraints among stations are most conveniently discussed in terms of additive contributions to the $\dot{N}$ matrix, these contributions are more conveniently added to the reduced normal equations, both for numerical reasons and because these equations give rise to nonzero off-diagonal blocks in the $\dot{N}$ matrix.

Four types of constraints are processed by the orbit constraint program being described.

(1) A priori weighted constraints on station positions. These generate observation equations of the form $I\delta X_i = O$ and contributions to $\dot{N}$ of the form

$$\dot{W}_i \delta X_i = O$$

where $\dot{W}_i$ is the inverse of the covariance matrix describing the accuracy with which the approximate (a priori) coordinates of the ith station represent that station's terrestrial (geocentric) coordinates.

(2) Absolute constraints on station positions, or fixing of station positions. A station is easily fixed by removing the three rows and columns corresponding to that station from the matrix of reduced normal equations. This must be done after all other constraints involving that station have been processed.

(3) Chord distance between stations. Often highly accurate baselines make it desirable to constrain the distance between stations. This is especially true of stations connected to the super-accurate baselines currently being established by the USCGS as part of the U. S. World Geodetic Satellite Program. The geodetic coordinates may be easily converted to a chord distance d_{ij}^b between the ith and jth station. The single nonlinear equation generated by this observation is

$$[(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2]^{\frac{1}{2}} = d_{ij}^b + v_{ij} .$$

In linearized form this is

$$(T_{ij} \;\; - \;\; T_{ij}) \begin{pmatrix} \delta X_i \\ \delta X_j \end{pmatrix} = d_{ij}^b - d_{ij} + v_{ij}$$

where

$$T_{ij} = \left(\frac{x_i - x_j}{d^o_{ij}} \quad \frac{y_i - y_j}{d^o_{ij}} \quad \frac{z_i - z_j}{d^o_{ij}} \right)$$

and

$$d^o_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2]^{\frac{1}{2}},$$

evaluated at the approximate values of the station coordinates. The contributions to the normal equations generated by this observation are

$$\begin{array}{ll} \text{to } \dot{N}_{ii} \text{ and } \dot{N}_{jj}, & T^T_{ij}\, w_{ij}\, T_{ij}\,; \\ \text{to } \dot{N}_{ij} \text{ and } \dot{N}_{ji}, & -T^T_{ij}\, w_{ij}\, T_{ij}\,; \\ \text{to } \dot{K}_i, & T^T_{ij}\, w_{ij}\, (d^b_{ij} - d^o_{ij}); \\ \text{to } \dot{K}_j, & -T_{ij}\, w_{ij}\, (d^b_{ij} - d^o_{ij}); \end{array}$$

where $w_{ij} = \text{var}\,(d_{ij})^{-1}$.

(4) Relative position between stations. If two stations are fairly close together, it may be desirable to preserve their relative surveyed position in rectangular coordinates. This requires that the two stations be connected by both horizontal and vertical control. The constraint is expressed by the three equations

$$X_i - X_j = \Delta X^b_{ij} + V_{ij}.$$

The contributions to the normal equations are

$$\begin{array}{ll} \text{to } \dot{N}_{ii} \text{ and } \dot{N}_{jj}, & W_{ij}\,; \\ \text{to } \dot{N}_{ij} \text{ and } \dot{N}_{ji}, & -W_{ij}\,; \\ \text{to } \dot{K}_i\,, & W_{ij}\,(\Delta X^b_{ij} - \Delta X^o_{ij}); \\ \text{to } \dot{K}_j\,, & -W_{ij}(\Delta X^b_i - \Delta X^o_{ij})\,; \end{array}$$

where $\Delta X^{o}_{ij} = X_i - X_j$ evaluated at the approximate values of the coordinates, and W_{ij} is the inverse of the covariance matrix describing the accuracy of the relative position in rectangular coordinates.

3.5 The Need for Approximate Orbital Elements

One very troublesome aspect of the orbital mode adjustment is the need for approximate orbital elements. Not only must these elements be fairly good approximations if the solution is to converge at all, but they should be quite accurate approximations if it is to converge in a reasonable number of iterations.

It is not at all obvious to the user of a short arc orbital constraint program how the approximate orbit elements are to be obtained. There do not appear to be any simple graphical or desk calculator methods which will yield an accurate set of Keplerian elements from a set of direction or range observations, so that it appears to be necessary to go to the computer just to generate a set of approximate elements.

This problem was first encountered when the orbital constraint program described in this chapter was used to adjust the BC-4 data described in Chapter 4. A first attempt at obtaining approximate orbital elements entailed the use of the computer program described in [Wintz, 1961]. This program, which is available in the library of the Department of Geodetic Science, computes a set of Keplerian elements using three direction observations from a single station by the Gauss method. (Wintz estimates that the effort required to perform this task on a desk calculator is about five man-days — a completely unreasonable effort in view of the present accessibility of computers [Wintz, 1961, p. 16].) It quickly became apparent that this approach would not suffice for the short arcs described by the balloon satellites as they passed through the 33° field of view of the BC-4 camera. Such disconcerting results as negative eccentricities, negative mean motions, and semi-major axes much smaller than an earth radius, were obtained. Moreover, when the same orbit was determined separately by the observations from two different stations, the two sets of elements obtained bore no resemblance to each other. The cause of the difficulty was assumed to be that the three to five minute time span of the arcs did not afford a sufficient geometrical separation of the observations. This experience was a preview of many more difficulties to be encountered in the determination of orbits from the BC-4 data.

Since not even an approximate orbit could be extracted from the data, the orbit elements that had been used by the Coast and Geodetic Survey for their predictions were obtained. A differential correction orbit determination program was written. This was actually a simplified version of the orbital mode adjustment program, with the station unknowns removed. The USCGS elements were used as a first approximation, and a second approximation was obtained from the data by differential orbit correction. These second approximations were used as input to the orbital constraint adjustment, and were usually sufficiently accurate that a satisfactory adjustment was obtained in one or two iterations.

The differential orbit correction process usually required four or five, and sometimes as many as ten, iterations to compute corrections of two or three degrees of arc to the angular elements. If the first approximations were not good to within a few degrees of arc, the differential orbit correction program would often compute a hyperbolic orbit at some iteration. At this point Kepler's equation, which is valid in its usual form only for elliptic orbits, would fail to converge and the whole process would break down. This indicated both that the orbit was not well determined and that the observations were highly nonlinear functions of the Keplerian elements. In later experiments with the orbital mode adjustment program, it was noticed that this program could easily pick up corrections of one mile to station positions in a single iteration, but that corrections of one minute of arc to the angular orbit elements necessitated several iterations.

Because of the difficulty of obtaining approximate orbit elements, consideration was given to the possibility of redesigning the adjustment so that the approximate values of the orbit elements would not appear. In the geometric mode adjustment, it is possible to design the problem so that the approximate satellite positions do not appear at all. All geometric mode adjustments that utilize the concept of intersecting planes, such as that described by [Aardoom, Girnius, and Veis, 1966], have this property. Krakiwsky demonstrates a geometric adjustment where the approximate satellite positions algebraically cancel out and disappear in the formation of the normal equations [Krakiwsky and Pope, 1967, p. 42]. It appears that the approximate values of nuisance parameters can always be algebraically eliminated from the normal equations of an adjustment if the mathematical model may be written linearly in those parameters. There does not appear to be any natural and reasonably simple way of doing this for the orbital mode adjustment. Although it may be possible to write a model for the short arc orbit constraint adjustment that is linear in the Keplerian elements (and non-linear in the observations), such a model would have to embody some orbit determination process, such as the Gauss method, and this would introduce intolerable complications.

4. The Use of Orbital Constraints in the Case of Satellite Trails on BC-4 Photographic Plates

The purpose of the experiments described in this chapter was to determine the feasibility of adjusting satellite directions determined from BC-4 photographic plates in the short arc orbital mode, and to compare the solution obtained in the orbital mode with the corresponding geometric mode solution. These experiments were part of a continuing series of investigations conducted at OSU into the various methods of utilizing passive satellite observational data, especially the data deposited in the National Space Science Data Center. Results of previous investigations are reported in [Hotter, 1967; Hornbarger, 1968; and Veach, 1968].

4.1 The Chopped Satellite Trail

The BC-4 is the only camera in widespread use designed specifically for photographing passive satellites. It consists of a modified Wild RC-5 aerial camera mounted on a T-4 theodolite base. Originally a 300 mm Astrotar lens was used, but this is now being replaced by a 450 mm Cosmotar lens especially designed for satellite photography. The field of view is approximately $33^\circ \times 33^\circ$ for the 300 mm lens and $20^\circ \times 20^\circ$ for the 450 mm lens.

The BC-4 has no sidereal or great circle drives, but is always operated in a stationary mode. Star images are exposed before and after the pass of the satellite with a capping shutter in front of the lens. As the satellite passes through the field of view, the capping shutter is held open and the satellite trail is chopped by system of rotating disk shutters. The exposure time and chopping interval are selected according to the expected speed of the satellite through the field of view, and the f-stop selected is based on the expected magnitude of the satellite. An exposure time of 1/60 sec and a chopping interval of 1/5 sec are typical of plates of ECHO II. This yields a trail of 400-600, sometimes more, satellite images across the plate. For the PAGEOS satellite, the exposure time is generally 1/30 or 1/15 sec, and the chopping interval is generally 4/5 sec. This yields a trail of about 200-400 images across the plate. The capping shutter is programmed to code the trail by omitting certain images. This facilitates identification of the images by time of exposure. A great deal of effort is put into the timing of the images, with the result that the accuracy of the time associated with each satellite image is no worse than 150 microseconds [Taylor, 1963].

The BC-4 camera is providing the data for the establishment of the U.S. World Geometric Satellite Net. This program is directed by the U.S. Coast and Geodetic Survey of the Environmental Science Service Administration, which schedules observations and reduces and analyzes the data. Observations are scheduled only in the simultaneous mode. Successful simultaneous photography by two or more stations is termed an event. Since the passage of the satellite through the field of view takes two to five minutes, "simultaneous" in this context means that the time spans of the two plates overlap. In practice, observations are scheduled and the cameras are aimed so that the overlap is large and the satellite passes near the center of each plate at nearly the same time. The opening of the shutter is controlled by the station clocks. Since each clock defines its own time system, and because of the finite time required for the light from the satellite to reach the stations, the images are not exactly simultaneous. The star images are measured by the USCGS and are used to perform a sophisticated stellar camera calibration, which is summarized in [Hotter, 1967]. The accuracy of the computed direction of the optical axis is thought to be about $0\rlap{.}''5$.

Each satellite image is also measured. For the data submitted to the National Space Science Data Center, each satellite image is converted to a right ascension and declination by use of the plate constants. The satellite images are treated as unknown stars in this conversion. The right ascensions and declinations of the satellite images include atmospheric refraction and stellar aberration, so that the coordinates deposited in the NSSDC are "apparent" in the sense of the apparent place of a star. The time associated with each image is the UT1 of the instant the image was received (mid-opening of the shutter) modified by the addition of 44 ms to refer to the old adopted longitude of the U. S. Naval Observatory. In order to obtain the geometrical direction of the center of the satellite, it is necessary to correct the listed coordinates for phase angle, parallactic refraction, and parallactic aberration (light-time correction) [Veach, 1968, p. 91].

4.2 Satellite Images and Curve Fitting

The sheer abundance of data available from even a single BC-4 plate creates several interesting problems, since each user must decide how best to use such a huge amount of data. The method employed by the USCGS involves the fitting of polynomials to the satellite trail, but not all investigators agree that this is the best approach. In particular, it may be possible to use only a selected set of images and obtain the same information by processing the data in the orbital mode.

4.21 The Use of Every Single Image. The most obvious approach is to treat each image as a separate observation. However, there are several good reasons why this is unsatisfactory. The main problem is that there is just too much data to be processed for the amount of geometrical information that may be obtained. Although they may provide statistical information, adjacent images are so close together in direction that they provide essentially the same geometrical information. Thus it would appear that the same amount of geometrical information could be obtained from only a selected set of images. Although using the several hundred images on each plate will not necessarily harm an adjustment, it is usually not advisable to generate superfluous observation equations, since the processing of a greater number of equations on the computer will usually result in a greater accummulation of round-off error. Furthermore, the generation of several hundred observation equations that give little useful information would be a needless waste of computer time. A second consideration is that certain components of error, such as emulsion creep or anomalous refraction, may be nearly the same in adjacent images. This could mean that the errors in adjacent images are significantly correlated, and that an adjustment that fails to take account of this time-wise (or serial) correlation may be biased. A third consideration is that the images on the two or more plates that constitute an event are not simultaneous. Therefore some sort of curve fitting procedure, if only as an interpolation tool, will have to be employed anyway.

4.22 The Use of a Few Selected Images. If only a few, such as 10 or 20 images, are selected from the plate, and if these are spaced fairly well apart, then the correlation between adjacent images arising from such effects as emulsion creep and image motion may be neglected. However, if a single plate reduction is done for the whole plate, the directions computed for different images will still be correlated. The plate constants are determined from measurements of the star images, and thus have statistical uncertainties. Since the computed right ascensions and declinations of the different satellite images all involve the same set of plate constants, these will all be correlated; i.e., although the plate coordinates (x, y) of the different images are not correlated, the right ascensions and declinations are. This consideration also applies to photographs of sequences of flashing lights, such as the ANNA or GEOS optical beacons. Since this consideration applies to other photographic systems, such as the MOTS and PC-1000, it will be examined in some detail.

Let the error in the right ascension of a single image be written

$$\delta\alpha_1 = \delta\alpha_{1m} + \delta\alpha_{1s} + \delta\alpha_{1p}$$

where $\delta\alpha_{1m}$ is the component of the error arising from accidental measurement error on the comparator, $\delta\alpha_{1s}$ is the error from shimmer, image motion, and emulsion creep, and $\delta\alpha_{1p}$ is the error arising from any errors in the plate constants; i. e.,

$$\delta\alpha_{1m} = \frac{\partial\alpha_1}{\partial x_1}\,\delta x_1 + \frac{\partial\alpha_1}{\partial y_1}\,\delta y_1,$$

etc. There are, of course, several other small sources of error, but consideration of them is not necessary for this analysis. Since the dominating source of error in the plate constants is the errors in the catalogued positions of the stars, it may reasonably be assumed that these components of error are independent. Then the variance of this right ascension may be written

$$\sigma^2\{\alpha_1\} = \epsilon[\delta\alpha_1^2] = \epsilon[\delta\alpha_{1m}^2] + \epsilon[\delta\alpha_{1s}^2] + \epsilon[\delta\alpha_{1p}^2].$$

Similarly, the variance of the right ascension of another image from the same plate is

$$\sigma^2\{\alpha_2\} = \epsilon[\delta\alpha_{2m}^2] + \epsilon[\delta\alpha_{2s}^2] + \epsilon[\delta\alpha_{2p}^2].$$

The covariance between α_1 and α_2 is then given by

$$\operatorname{cov}\{\alpha_1\alpha_2\} = \epsilon[\delta\alpha_1\delta\alpha_2]. \tag{4.1}$$

This expression will contain the expectancies of six cross products. The measurement errors may reasonably be assumed to be independent, and, if the images are sufficiently well spaced, the errors arising from shimmer and image motion may also be assumed to be independent. However, the errors in the two images arising from the plate constants will be very nearly the same, and will both be very nearly the error in the right ascension of the estimated direction of the optical axis. Thus the only cross product which does not vanish in (4.1) will be

$$\epsilon[\delta\alpha_{1p}\,\delta\alpha_{2p}] \approx \epsilon[\delta\alpha_p^2].$$

The problem presented by covariances between different images is that they lead to large, nondiagonal, weight matrices. Covariance between the right ascension and declination components of one image presents little problem, since it leads only to a weight matrix with 2 $\times$ 2 blocks on the main diagonal. However, covariance between different images produces blocks whose dimension is twice the number of images taken from each plate. For example, let n be the number of images used from a plate and let the observation equations be arranged in the order $\{\alpha_1, \delta_1, \alpha_2 \ldots \delta_n\}$. For simplicity, assume that right ascensions and declinations have the same variance σ^2 and that right ascensions are uncorrelated with declinations. Also let

$$r = \frac{\epsilon[\delta\alpha_p^2]}{\sigma^2}$$

be the correlation between any pair of right ascensions or declinations. Then the covariance matrix for the observations on a single plate is

$$\begin{array}{cc} & \begin{array}{cccccc} \alpha_1 & \delta_1 & \alpha_2 & \delta_2 & & \alpha_n \quad \delta_n \end{array} \\ \sigma^2 & \begin{pmatrix} \underline{1} & 0 & r & 0 & \ldots & r & 0 \\ & \underline{1} & 0 & r & \ldots & 0 & r \\ & & \underline{1} & 0 & \ldots & r & 0 \\ & & & & & \vdots & \vdots \\ & & & & & \underline{1} & 0 \\ & & & & & & 1 \end{pmatrix} \end{array} \qquad (4.2)$$

The total weight matrix for the adjustment of all observations on all plates is made up of blocks such as this along the main diagonal. Since this matrix is highly patterned, its inverse may be formed analytically with little trouble. However, most computer programs for the adjustment of satellite directions are not set up to deal with correlations of this type, since a rigorous treatment of these correlations would necessitate storing the observation equations arising from all the images on a plate in the machine at the same time, and this would result in a somewhat cumbersome program. The orbital constraint program described in the previous chapter will take account only of correlation between the right ascension and declination of a single image, and this is also true of the geometric mode program used for the comparisons described in section 4.3. Therefore, it is reasonable to consider what will happen if these correlations between different images are neglected and a diagonal weight matrix is used.

First, the solution will still be unbiased even if the wrong weight matrix is used [Hamilton, 1964, p. 146]. However, the solution will not have the property of minimum variance, since this requires that the weight matrix be inversely proportional to the covariance matrix of the observations. Let the weight coefficient matrix of the solution obtained using a diagonal weighting matrix be denoted Q_D, and let that of the minimum variance solution be denoted Q_{MV}. Then

$$Q_{MV} \le Q_D$$

since Q_{MV} is minimum by definition. However, it is also possible to find an upper bound for Q_D in terms of Q_{MV} [Magnus and McGuire, 1962, p. 469]. This is given by

$$Q_D \leq \frac{1}{4} (\lambda_{max} + \lambda_{min}) \left(\frac{1}{\lambda_{max}} + \frac{1}{\lambda_{min}}\right) Q_{MV}$$

where λ_{max} and λ_{min} are the maximum and minimum eigenvalues of the true correlation matrix of the observations. The correlation matrix in (4.2) is highly patterned, and its eigenvalues turn out to be a function only of the correlation coefficient r and the number n of images used from the plate. The applicable values are

$$\lambda_{max} = 1 + (n-1)r \ ,$$

$$\lambda_{min} = 1 - r \ .$$

The factor $\frac{1}{4} (\lambda_{max} + \lambda_{min}) \left(\frac{1}{\lambda_{max}} + \frac{1}{\lambda_{min}}\right)$ is tabulated in Table 4.1 for $r = 0.06$ and $r = 0.1$. These correlation coefficients were used because they represent the correlation between individual images on BC-4 plates for the 300 mm and 450 mm lenses, respectively. The interpretation of this table is that the solution obtained by neglecting correlations between different images is no worse than this factor times the covariance matrix of the minimum variance solution. Thus it is possible to make a quantitive judgment as to whether or not the worsening of the solution caused by neglect of these correlations is tolerable.

Table 4.1
Worsening Factors for the Effect of the Neglect of Correlations Between Observations

Number of Images per Plate	Worsening Factor	
	r = 0.06	r = 0.1
1	1.0	1.0
2	1.0036	1.010
3	1.007	1.021
4	1.013	1.034
5	1.019	1.049
10	1.062	1.146
15	1.117	1.260
20	1.179	1.383
25	1.245	1.511

4.23 Curve Fitting and the Orbital Constraint. In the method used by ESSA to process the BC-4 plates, there is no problem of correlation between images, since only one image is considered and one pair of observation equations per plate is generated. This method entails the use of two polynomials to describe the trail of the satellite across the plate, and a single "fictitious image" is computed from these polynomials. Since the polynomial curve fitting smoothes out measurement error, shimmer, and other errors, this fictitious image is of higher quality than any of the single images. The two polynomials describe the plate coordinates x and y as functions of time, where the x and y axes lie in the photographic plate and are oriented in the direction of, and perpendicular to, the approximate direction of the satellite trail. (The ESSA computations never involve the right ascension and declination of each image.) The satellite trail on the plate appears to be a straight line to the unaided eye; however, careful analysis shows that higher degree polynomials fit significantly better [Veach, 1968, p. 73]. Normally fifth degree polynomials are used, but seventh degree polynomials have also been used, especially for the ECHO plates which generally have a greater number of images than PAGEOS plates. The coefficients of the polynomials are determined by least-squares curve fitting to the corrected x and y plate coordinates. The fictitious image is computed by selecting an epoch, usually specified as the UT1 of the time the light left the satellite, correcting this epoch for light travel time, transforming it into the time system in which the curve fitting was done (e.g., the station clock system), and evaluating the x and y polynomials for that epoch. The light-time correction and the transformation between time systems is discussed in detail and an example of the computations is given in [Veach, 1963, p. 100].

The USCGS procedure of using two polynomials has the effect of collapsing all of the images into a single image, and thus may be viewed as an information compression technique. Thus an important question is whether any significant information is lost in the compression process. A companion question is whether the information contained in the single fictitious image could be obtained by using some subset of the images on the plate. Another question, discussed by [Veach, 1968] and by [Hornbarger, 1968], is whether the same information could be obtained by dividing the plate into small areas and doing an astrometric plate reduction (which is much cheaper than the photogrammetric reduction), fitting polynomials, and obtaining a fictitious image for each area.

The technique of fitting a curve to the satellite trail has a certain intuitive justification; since the path of the satellite in space is a smooth curve, its trail on the photographic plate should be constrained to be a smooth curve also. The problem is that the curve fit uses too many parameters to describe the constraint. If fifth degree polynomials are used, a total of twelve parameters are used to describe the trail on each plate. If two plates are involved in the event, a total of 24 independent parameters are used to describe the motion of the satellite, and if three plates are involved, 36 independent parameters are estimated. However, the path of the satellite should be equally well described by the six parameters of a Keplerian or other simple orbit model for the two to five minutes spanned by the plate. It follows that the excess over six of curve fitting parameters are superfluous, and the problem is said to be "overparameterized."

The danger of overparameterization is that the observations may fit the constraint <u>too well</u>. Brown claims that there are several possible sources of error that are periodic across the plate with periods of the order of one or a few minutes [Brown, 1967, p. 91]. The curve fit will conform to these low frequency components of error, rather than constrain the images to the true projection of the satellite path onto the plate. Brown concludes that a low frequency error whose amplitude is one micron could easily be accommodated by the curve fit, and that this error could occur in and dominate the fictitious image. It appears then that the images on the satellite trail should be constrained, but the constraint should be expressed by short arc methods rather than by polynomial curve fitting.

4.3 Experiments Performed with BC-4 Plates

In order to answer some of the questions raised in the previous section, a selection of BC-4 data obtained from the NSSDC was processed with the short arc orbital constraint program described in the previous chapter. In order to afford a comparison with the geometrical approach, much of the same data was also processed by the geometric mode adjustment program described by [Krakiwsky and Pope, 1967].

4.31 Description of the Experiments. The data was obtained from a tape furnished to the NSSDC by ESSA in the spring of 1968. A copy of this tape was obtained from the NSSDC by OSU. This tape contains right ascensions and declinations for each image on 298 plates, all of which are being used by ESSA in the establishment of the World Geometric Satellite Net. Almost all of these plates are photographs of the PAGEOS satellite. With an average of 250-400 images per plate, the tape contains in the neighborhood of 100,000 images. A set of 12 plates, taken from three stations, was extracted from this tape. The data set was later expanded to include 45 plates and four stations.

The stations involved were the following:

1 Thule, Greenland

2 Beltsville, Maryland

3 Moses Lake, Washington

38 Revilla Gigedo, Mexico

The first set of data, involving stations 2, 3, and 38, consisted of the following:

Event	Date	Plates	Stations
2497	Aug 19, 1966	A404, E328	3, 2
3539	Dec 17, 1966	A491, E393	3, 2
4212	Mar 5, 1967	A567, U411	3, 38
4236	Mar 8, 1967	A571, E422, U416	3, 2, 38
4267	Mar 12, 1967	A574, E450, U423	3, 2, 38

The station, event, and plate numbers are those assigned by the USCGS.

The BC-4 cameras apparently never attempt to take more than one plate on a given pass of the satellite (as do the MOTS and PC-1000 cameras when photographing flashes from the flashing light satellites). In fact, observations are only rarely scheduled for two successive passes of the satellite. Thus it is necessary to define an "orbit," in relation to BC-4 photography, to be the trail of the satellite imaged on the two or more plates of an event, and "orbit" and "event" are used interchangeably.

The preprocessing of the 12 plates began with the determination of a preliminary orbit for each of the five events. The difficulties encountered at this stage are discussed in section 3.5. The preliminary (second approximation) orbits computed by differential orbit correction were then used to compute the range to the satellite for the time of each image. These ranges were used to correct all images on each of the plates for parallactic refraction, phase angle, and parallactic aberration, according to the formulae described in [Veach, 1968, p. 91]. A standard temperature and atmospheric pressure were used for the parallactic refraction correction.

Approximately 20 epochs falling within the time span of each event were selected; i. e., these 20 epochs were approximately equally spaced between the time the latest starting trail started and the time the earliest ending trail ended. A quasi-simultaneous image was constructed for each of these epochs by second order interpolation. These images were designated "individual images." The epochs were always selected so that they would fall close to the time of an actual image on the plate. Since some images were missing (due to coding of the trail, partial cloud cover, or other reasons), it was not always possible to ensure that these images were equally spaced in time. The interval between individual images varied somewhat, due to the varying interval of simultaneous tracking for the events. The average interval between individual images was about 15 sec. These 20 individual images per plate were adjusted in both the geometric and orbital modes. In addition, subsets of the data consisting of 10 individual images per plate and three individual images per plate were adjusted in the geometric mode.

Several investigations were conducted by Veach to determine the possibility of fitting the right ascensions and declinations from the NSSDC tape to time polynomials. Using the analysis of variance techniques described in [Veach, 1968, p. 73], he concluded that the satellite trail could also be described by fifth degree polynomials for the right ascension and declination. A pair of polynomials was fit to the satellite trail for all plates and a single fictitious image was formed for each plate. These fictitious images were then adjusted in the geometric mode. (Since all the data in the event (pass) was collapsed to a single epoch, there could be no corresponding orbital mode adjustment.)

Another experiment utilized three fictitious images per plate. The satellite trail was divided into three segments and an independent pair of polynomials was fit to each segment. Third degree polynomials were found by Veach to be adequate for these shorter trails. A fictitious image was taken from each segment, and these were adjusted in the geometric mode. A corresponding orbital mode adjustment was not made, since this would have had the effect of imposing a constraint on the satellite trail twice.

In all experiments, in both the geometric and orbital modes, the coordinates of Beltsville on the North American Datum were held fixed and the NAD distance from Beltsville to Moses Lake was constrained. Plate U416 of event 4236 gave difficulties in both the geometric and orbital modes, and was deleted from all experiments reported here.

The results of the 12 plate series of experiments are shown in Table 4.2. The experiments with the geometric mode were

(1) 3 individual images per plate,

(2) 10 individual images per plate,

(3) 20 individual images per plate,

Table 4.2

Results of Twelve Plate Series of Experiments

Experiment	Moses Lake X	Moses Lake Y	Moses Lake Z	Revilla Gigedo X	Revilla Gigedo Y	Revilla Gigedo Z	Degrees of freedom	A priori standard deviation of a single observation	A posteriori standard deviation of a single observation
	-2127800 m +	-3786000 m +	4655800 m +	-2160900 m +	-5642800 m +	2035100 m +			
Geometric Mode									
1. 3 individual images	-1.7 (11.5) [9.3]	-25.5 (54.4) [43.8]	57.9 (42.2) [34.0]	-11.2 (29.0) [23.4]	0.4 (59.3) [47.8]	141.1 (78.9) [63.6]	16	1".7	2".11
2. 10 individual images	-6.7 (6.1) [5.1]	-44.7 (28.9) [24.0]	63.7 (22.3) [18.5]	-45.9 (15.4) [12.8]	-50.4 (31.5) [26.4]	110.5 (41.7) [34.6]	65	1.7	2.05
3. 20 individual images	-3.1 (3.7) [3.4]	-16.8 (17.3) [15.7]	37.1 (13.5) [12.4]	-32.2 (9.6) [8.7]	-20.9 (19.2) [17.5]	62.2 (25.4) [23.1]	148	1.7	1.87
4. 3 fictitious images	-0.3 (3.3) [5.5]	-11.2 (15.5) [25.8]	42.2 (12.1) [20.2]	-44.7 (8.3) [13.8]	-44.1 (17.0) [28.3]	89.3 (22.6) [37.7]	16	1.0	0.61

Table 4.2 (continued)

Experiment	Moses Lake X	Moses Lake Y	Moses Lake Z	Revilla Gigedo X	Revilla Gigedo Y	Revilla Gigedo Z	Degrees of freedom	A priori standard deviation of a single observation	A posteriori standard deviation of a single observation
	-2127800 m +	-3786000 m +	4655800 m +	-2160900 m +	-5642800 m +	2035100 m +			
5. 1 fictitious image	-2.5 (2.0) [6.7]	-16.6 (9.6) [32.0]	39.8 (7.4) [24.7]	-40.6 (5.0) [16.7]	-65.4 (10.4) [34.7]	118.7 (13.8) [46.0]	2	0".7	0".21
Orbital Mode									
6. Every 20th image uncorrected	2.2 (4.5) [5.0]	11.9 (19.0) [21.1]	17.8 (14.0) [15.5]	-12.0 (11.8) [13.1]	-32.6 (21.6) [24.1]	77.0 (26.9) [29.9]	290	2.5	1.5
7. 20 individual images	-3.4 (3.1) [3.3]	-21.7 (13.9) [14.6]	43.1 (11.8) [12.3]	-33.6 (10.4) [10.9]	-26.9 (17.6) [18.5]	70.0 (23.5) [24.7]	451	1.7	1.62

Beltsville is held fixed on North American Datum. For each experiment the first row gives the solution coordinates. The second row in parentheses gives the standard deviation based on the estimated unit variance from the adjustment. The third row gives standard deviations based on a priori observational standard deviations.

(4) 3 fictitious images per plate, obtained by breaking each trail into three segments and fitting third degree polynomials,

(5) 1 fictitious image per plate, using a fifth degree polynomial curve fit for the entire trail, analogous to the procedure of the USCGS.

The experiments using the short arc program were:

(6) every 20th image, taken directly from the NSSDC tape and uncorrected for parallactic refraction, phase angle, or light-time,

(7) 20 images per plate.

The results of the first three experiments were about as expected, except that the differences between the three estimates of the standard deviation of an individual image are too high. The fifth experiment, using one fictitious image per plate, gave unbelievably low uncertainties. This can only be explained by the fact that this adjustment had only two degrees of freedom.

Exactly the same set of data was used for experiment 3 in the geometric mode and experiment 7 in the orbital mode. The results from these two experiments show more similarity than any other pair of experiments. Since the two adjustments used the same set of data, they must be compared on the basis of a single standard deviation of a single direction, regardless of the estimated unit variances from the respective adjustments. On this basis, the similarity between these two adjustments becomes even more striking. Neither adjustment is clearly stronger than the other. The strengthening of the solution, which was expected because of the analysis presented in section 2.6, does not appear in these results. The estimated standard deviation obtained from the orbital mode adjustment indicates that modeling error is suppressed well below the level of observational error and that the observations fit the orbit constraint about as well as was expected. The similarity between the two solutions indicates that the geometric solution is not greatly changed by the imposition of the orbital constraint. This is especially interesting when compared to the large change in the solution between 20 individual images and 10 individual images per plate.

In order to confirm the results of the 12 plate experiments, a larger set of data, consisting of 45 plates taken from four stations, was selected. These plates were

Event	Date	Plates	Stations
2492	Aug 18, 1966	B328, A404	2, 3
3173	Nov 18, 1966	B686, A478	1, 3
3185	Nov 19, 1966	B688, A479	1, 3
3539	Dec 17, 1966	E393, A491	2, 3
3538	Dec 17, 1966	B748, E392	1, 2
3560	Dec 19, 1966	B753, E394	1, 2
3561	Dec 19, 1966	B754, E395	1, 2
3795	Jan 14, 1967	B799, A513	1, 3
3837	Jan 20, 1967	B825, A521	1, 3
3935	Jan 31, 1967	B835, E416	1, 2
4061	Feb 14, 1967	B870, A544	1, 3
4182	Mar 1, 1967	E436, A557, U406	2, 3, 38
4196	Mar 3, 1967	A562, U408	3, 38
4212	Mar 5, 1967	A567, U411	3, 38
4236	Mar 8, 1967	E422, A571	2, 3
4244	Mar 9, 1967	E444, U417	2, 38
4251	Mar 10, 1967	B890, E446, U419	1, 2, 38
4259	Mar 11, 1967	E448, U421	2, 38
4267	Mar 12, 1967	E450, A574, U423	2, 3, 38
4276	Mar 13, 1967	A576, U425	3, 38
4292	Mar 15, 1967	B893, U427	1, 38

Since the orbit constraint is written in terms of geocentric coordinates, it was felt that the adjustments should be performed in a coordinate system that is more nearly geocentric than that defined by the North American Datum. The SAO C-5 Datum was selected for this purpose, and the C-5 coordinates of Beltsville were held fixed in all experiments in the 45 plate series. Thus the adjusted station coordinates from these adjustments may be said to be in the C-5 system. The datum shifts used for the transformation from NAD to the C-5 Datum were

$$\Delta x = 21\text{m},$$
$$\Delta y = -143\text{m},$$
$$\Delta z = -173\text{m}.$$

The spatial chord distance between Beltsville and Moses Lake, as determined from their NAD coordinates, was again constrained.

The same preprocessing procedure was used for the 45 plate set as had been used for the 12 plate set. The same set of adjustments was performed, with the exceptions of the three individual and three fictitious images per plate. The results of these experiments are tabulated in Table 4.3.

All of the experiments in the 45 plate series yielded acceptable estimates of the unit variance, so that most of the results are directly comparable. Many of the observations made about the 12 plate experiments are corroborated by the results of the larger set. The corresponding geometric and orbital mode adjustments no longer show such striking similarities since the data used was not exactly the same, the main difference being that observations from the ends of the longer trails were used in the

Table 4.3
Results of the 45 Plate Series of Experiments

Experiment	Moses Lake			Revilla Gigedo			Thule			Degrees of freedom	A posteriori stan. dev. of a single obs.
	X	Y	Z	X	Y	Z	X	Y	Z		
	-2127800 m +	-3785800 m +	4656000 m +	-2160900 m +	-5642700 m +	2035300 m +	546500 m +	-1390000 m +	6180200 m +		
Geometric mode											
1. 20 individual images	-31.1 (2.9)	-105.8 (11.1)	32.2 (9.3)	-62.6 (5.4)	-21.1 (9.0)	60.4 (12.1)	87.9 (9.9)	-11.3 (11.5)	35.3 (15.6)	448	1."9
2. 10 individual images	-30.7 (4.1)	-85.3 (16.0)	35.0 (13.7)	-63.0 (7.7)	-12.1 (12.6)	45.7 (17.6)	82.2 (14.4)	-26.3 (16.7)	29.1 (23.0)	223	2.0
3. 1 fictitious image	-26.7 (4.6)	-88.2 (19.0)	26.1 (14.9)	-56.5 (9.7)	-57.6 (14.1)	78.7 (19.3)	67.1 (14.4)	-2.3 (16.6)	47.0 (23.2)	13	0.7
Orbital Mode											
4. 20 individual images	-20.6 (3.6)	-85.3 (12.5)	51.7 (8.2)	-55.4 (5.8)	-16.3 (10.1)	77.3 (11.1)	88.9 (8.8)	-0.2 (10.8)	41.2 (13.6)	1568	1.7
5. 10 individual images	-32.8 (5.2)	-129.2 (19.6)	61.7 (13.0)	-65.1 (8.6)	-43.1 (15.6)	91.9 (16.7)	105.3 (13.4)	-26.8 (16.8)	110.1 (21.6)	742	1.8
6. 20 uncorrected images	-22.5 (4.1)	-79.8 (15.1)	31.5 (9.7)	-77.6 (6.4)	-32.4 (11.4)	79.8 (12.7)	91.8 (9.2)	-18.5 (14.8)	13.5 (14.6)	1192	1.7

Beltsville is held fixed on the SAO C-5 datum. The first row of each entry gives the adjusted coordinate, and the second row gives the standard deviation.

orbital mode adjustment. The single fictitious image appears to contain about as much information as 10 individual images. The 20 individual image adjustments in both modes appear to be stronger than either the adjustments of 10 individual images or the one fictitious image, and the improvement is about a factor of $\sqrt{2}$. This indicates that information is indeed lost in the curve fitting procedure: more information is available on the plate than is carried by the single fictitious image. It also appears that if only ten images, rather than several hundred, were measured the same amount of information could be obtained from the plate. If 20 images are measured, even more information may be obtained.

Although 20 individual images may contain more information than the single fictitious image or the 10 individual images, the "worsening" of the solution caused by neglecting the correlation between images on the same plate is greater when 20 individual images are used (see Table 4.1). This may indicate that this information cannot be properly utilized unless one is willing to consider correlation between images in the adjustment procedure. On the other hand the problem of correlation between images is not so serious if the plate is reduced astrometrically. Veach discusses the practicability of making a separate astrometric determination of the plate constants for each image, using only stars in the vicinity of that image. If this approach is used, correlation will be present only to the extent that there is overlap between the sets of stars used to determine the astrometric plate constants for different images [Veach, 1968, p. 25].

The adjustments with the 20 uncorrected images, taken directly from the NSSDC tape, were performed for both the 12 plate and the 45 plate data sets in order to determine the significance of the corrections for parallactic refraction, phase angle, and parallactic aberration. Surprisingly, smaller residuals were obtained than with most of the adjustments involving corrected images. In view of the spreads of the adjusted coordinates provided by the other adjustments, the solutions provided by the uncorrected observations appear to be quite reasonable in most coordinates. This is only an indication that the effect of the systematic error in the observations is less than the effect of sample variation, and is certainly no justification for allowing known systematic errors to remain in the observations.

It appears clear from the a posteriori standard deviations of the adjustments that the quality of a single image from a BC-4 plate is about $1\overset{''}{.}7$. A careful analysis of the residuals from the orbital mode adjustments indicated that two statistical populations were present: the standard deviation of a single image from a plate taken with a 300mm lens is about $2\overset{''}{.}0$, and that for a 450mm lens is about $1\overset{''}{.}4$. These results agree well with technological analyses of the error components [Taylor, 1963, p. 8; Veach, 1968, p. 19]. The standard deviation of a fictitious image taken from a fifth degree curve fit to the entire satellite trail appears to be about $0\overset{''}{.}7$, which agrees well with the results obtained by ESSA [Schmid, 1968, p. 6]. The difference in the quality of the images formed by the two lenses indicates that the dominant contribution to the standard deviation of a single image is the plate measurement error. If the contribution from the uncertainty of the direction of the optical axis is about 0".5, then the correlation between different individual images from the same trail, according to equation (4.1), is only about 0.06 for the 300mm lens. According to Table 4.1, neglecting this small amount of correlation and treating the individual images as independent observations will cause worsening of the solution by factors of 1.062 for 10 images per plate, and 1.179 for 20 images per plate, which are certainly tolerable.

4.32 Behavior of Short Arc Orbit Constraint Adjustment. Throughout all of the experiments with the short arc adjustment program, it was noticed that the solution produced by this program was somewhat unstable. When iteration of a solution was required, quite large changes in the elements of the solution vector would produce only very small improvement in the quadratic form of the residuals. Furthermore, the deletion of a small amount of data sometimes resulted in large changes in the solution vector. These symptoms indicated that the problem was ill-conditioned. There is ample physical reason to suspect that the short arc orbital mode adjustment is an ill-conditioned problem. Since the orbital mode adjustment process is also an orbit determination process, orbits are determined from observations on only a short portion of the path. For the Keplerian orbit model, this means that the size, shape, position, and orientation of an ellipse must be determined from observations of a short portion of the ellipse. This is a difficult task to perform accurately even by graphical methods; in a least squares adjustment, this physical (or geometrical) ill-conditioning is reflected as a numerical ill-conditioning of the normal equations.

The length of the satellite trails on the BC-4 plates in the 12 plate set varied between 2.8 and 5.4 minutes of time. Since the period of the PAGEOS satellite is about three hours, these trails cover between 5° and 10° of mean anomaly.

Since the ill-conditioning is associated with the determination of the orbit parameters rather than the station positions, it was expected that the ill-conditioning would be found in the partitions of the normal equations pertaining to the orbit parameters. These are the $\ddot{N}_k$ matrices, which must be inverted when the reduced normals are formed according to the expression $\dot{N} - \sum_k \overline{N}_k \ddot{N}_k^{-1} \overline{N}_k^T$. Since difficulties were also encountered in the differential orbit correction process, in which the normal equations consist solely of the $\ddot{N}_k$ matrices, it was suspected that the inversion of these matrices was the cause of numerical difficulties.

Since it appeared that there was a practical lower limit as well as an upper limit to what constitutes a short arc, the condition of these matrices for the five events (orbits) of the 12 plate set was investigated. The M, N and P condition numbers described in [Fadeev and Fadeeva, 1963, p. 120] were computed for each of the $\ddot{N}_k$ matrices and several different combinations of data. These numbers are defined by

$$M - \text{number} = \frac{1}{n} M(A) M(A^{-1}),$$

$$N - \text{number} = \frac{1}{n} N(A) N(A^{-1}),$$

$$P - \text{number} = \frac{\max |\lambda_i|}{\min |\lambda_i|}$$

where

$$M(A) = n \max |a_{ij}| ,$$

$$N(A) = [\mathrm{Tr}\,(A^T A)]^{\frac{1}{2}};$$

n is the dimension of the arbitrary matrix A and the λ_i are its eigenvalues. The P - number is probably the best guide to the condition of a matrix, since it generally lies between the other two. Although ill-conditioning is not quantitively defined, Fadeev and Fadeeva give the impression that matrices with condition numbers over 20,000 may be troublesome.

In addition, the RMS error in inversion was computed. This quantity was defined as the RMS of the elements of the matrix $AA^{-1} - I$. Obviously when this matrix is computed its elements consist solely of round-off error, so that their RMS depends on the particular algorithm used for inversion and the precision carried, as well as the condition of the matrix. For this reason, this quantity gives an indication of the numerical error to be expected when an ill-conditioned matrix is inverted by a particular algorithm on a particular machine. In these investigations, the single precision floating print arithmetic of the IBM 7094 was used, which carries a 27 binary bit mantissa (about seven to eight decimal digits). For the inversion the library matrix inversion package of the OSUSYS processor was used. This subroutine is a standard Gauss-Jordan elimination process, and is substantially the same as the library subroutine of the IBJOB processor. This combination matched the inversion process embodied in the orbital constraint adjustment program.

Normal equations were formed for the six sets of data described below:

(a) Observations were taken only from the interval of time overlap of the plates of the event, i.e. only from the time interval for which quasi-simulatneous individual images could be formed.

(b) Individual images were added from the ends of the longer trail in the event.

(c) Artificial observations were constructed to extend the longer trail to 10 minutes.

(d) Both trails were artificially extended to 10 minutes.

(e) One trail was extended to 18 minutes.

(f) Both trails were extended to 18 minutes.

The dimension of the $\ddot{N}_k$ matrices was six. The matrix of reduced normal equations, whose dimension was nine, was also analyzed to see what effect the ill-conditioning would have on this matrix. The $\ddot{N}_k$ matrices are identified by event numbers in Table 4.4, and the six sets of data are identified by the letters a-f.

The condition numbers in Table 4.4 show that the $\ddot{N}_k$ matrices are very poorly conditioned indeed. With a few exceptions, the condition of these matrices improves as the length of the trails is extended. However, the improvement is not very much, and is certainly not enough to eliminate the numerical problems. The RMS error in inversion varies in about the same manner as the condition numbers (again with few exceptions).

The problem presented by ill-conditioned matrices is not that they cannot be inverted, since they usually can be, but that they are especially prone to the accumulation of numerical round-off error during the inversion process. One author, investigating the numerical accuracy of different

Table 4.4
Condition Numbers of $\ddot{N}_k$

Matrix	Minutes of data (1 min ≈ 2°)		Condition numbers units = 10^6			RMS error in inversion
	Longest trail	Shortest trail	M	N	P	
2497 a	4.3	4.3	108	10	64	0.0578
b	5.1	4.3	70	6.8	37	0.0330
c	10.0	4.3	107	10	63	0.0463
d	10.0	10.0	63	6.7	35	0.0181
e	18.0	10.0	44	5	32	0.0195
f	18.0	18.0	48	6	33	0.0127
3539 a,b	2.8	2.8	451	44	132	0.356
c	10.0	2.8	226	24	133	0.0880
d	10.0	10.0	53	5	29	0.0233
e	18.0	10.0	19822	1896	264	8.958
f	18.0	18.0	139	13	53	0.0467
4212 a	2.9	2.9	727	32	176	0.0324
b	5.4	2.9	324	14	75	0.0261
c	10.0	2.9	131	6	35	0.00538
d	10.0	10.0	86	4	23	0.00643
e	18.0	10.0	77	3.5	20	0.00406
f	18.0	18.0	63	2.9	17	0.00263
4236 a	3.4	3.4	1448	62	351	0.0660
b	5.0	3.4	1120	48	263	0.0155
c	10.0	3.4	500	21	117	0.0251
d	10.0	10.0	185	8	46	0.0059
e	18.0	10.0	123	5	31	0.0064
f	18.0	18.0	80	3.6	21	0.0046
4267 a	2.5	2.5	913	39	46	0.0109
b	5.0	2.5	501	22	118	0.0344
c	10.0	2.5	185	9	50	0.0100
d	10.0	10.0	83	3.8	22	0.00219
e	18.0	10.0	60	2.7	16	0.00215
f	18.0	18.0	47	2.1	12	0.00187

Table 4.5 Condition Numbers of Reduced Normal Equations				
	Condition numbers units = 10^3			RMS error in inversion
	M	N	P	
a	125	4.7	23	$.37 \times 10^{-6}$
b	104	4.1	22	$.30 \times 10^{-6}$
c	55	1.9	10	$.19 \times 10^{-6}$
d	26	0.8	4.4	$.07 \times 10^{-6}$
e	19	0.7	3.6	$.07 \times 10^{-6}$
f	9	0.3	1.4	$.03 \times 10^{-6}$

inversion algorithms, has noted that "for some ill-conditioned matrices accumulation of error may extend to every significant digit in the final results of the calculations with little or no indication that the output is erroneous [Longley, 1967]. Examining the RMS error in inversion, it appears that for satellite trails of the length found on BC-4 plates, and for the precision and algorithms used by the short arc adjustment program, only about one significant decimal digit may be expected in the inverse of the $\ddot{N}_k$ matrices. The possibility that the numerical error may be even worse at times was illustrated by the fact that negative numbers were found on the main diagonal of the computed inverse of these positive definite matrices in a few instances. Orbits for which this happened were always deleted from the data set.

The condition of the matrix of reduced normal equations is not especially bad, and its inversion certainly does not appear to present any problems. Howevery, this fact does not eliminate the problem of numerical error. Experience indicates that the magnitude of the elements of $\overline{N}_k \ddot{N}_k^{-1} \overline{N}_k^T$ is about the same as, or only slightly less than, that of the elements of $\dot{N}$. Thus the elements of the reduced normal equations may contain only one or two significant digits, and the inverse cannot then be any better. Of even more concern is the possibility that the constant column of the reduced normal equations, computed according to the expression $\dot{K} - \sum_k \overline{N}_k \ddot{N}_k^{-1} \ddot{K}_k$, may suffer a similar severe loss of significant digits.

The possibility that the condition of the $\ddot{N}_k$ matrices could be improved by writing the observation and normal equations in terms of a different set of orbit unknowns was also investigated. The programs were rewritten so that the orbit unknowns were the rectangular position and velocity vectors at epoch, X_0, $\dot{X}_0$. No significant improvement in the $\ddot{N}_k$ matrices was found when these unknowns were used, nor was there any significant change in the rate of convergence of the solution.

There are several ways of coping with the problem of numerical error. The precision of the computations may be extended, or an algorithm which is less effected by ill-conditioning may be selected. It has been noted that the class of orthogonalization methods of matrix inversion is generally much superior to the class of elimination methods in this regard [Longley, 1967]. On the other hand, one could only accept arcs which span at least 30° of mean anomaly (1/12 revolution), although this would impose severe restrictions on what constitutes a short arc.

5. Conclusion

Throughout this report the orbital mode adjustment has been compared to the corresponding geometric mode of adjustment from many viewpoints. In many respects, the orbital mode adjustment is the more difficult adjustment to program, run, and analyze. It requires the selection of a model for the orbit and careful analysis to ascertain that the model is valid for the type of observations, type of orbit, and length of pass to be processed. It requires the development of fairly complicated preprocessing program to produce a good set of approximate orbit elements. It is plagued by problems of ill-conditioning, and the results it produces require careful skilled analysis to assure that they are valid. The seemingly most important advantage of the orbital mode adjustment is that it must produce a stronger solution from the same set of data; however, the amount of this strengthening was not found to be significant, at least not for direction observations typical of the BC-4 cameras, when the same set of observations were adjusted in both modes.

There are a few instances in which the use of the orbital mode may allow the use of a significant amount of data that is otherwise unusable. For example, the optical beacon of the GEOS-A satellite was programmed to produce several sequences of flashes as the satellite passed over the continental United States. These flashes were observed principally by MOTS and PC-1000 cameras. All of these flashes may be tied together by the orbit when the data is adjusted in the short arc mode. For many flash sequences, only a single station was able to obtain successful photography, the photography that was scheduled to be taken from other stations being lost due to cloud cover or other reasons. In the short arc mode these single (unmatched) plates could be used and would strengthen the solution. This fact would be of more importance were there not also a sufficient number of passes on which successful MOTS and PC-1000 photography was obtained on all flash sequences.

By comparison with the geometric mode, the orbital mode is at least seen to be a valid form of adjustment. The fact that the two adjustments gave almost identical solutions when the same set of data was used in both indicates that it is possible to suppress the modeling error below the level of the observational error, and to obtain a valid solution in the orbital mode in spite of the ill-conditioning problem. In all of the experiments performed with the BC-4 data, the orbital model was the simple Keplerian orbit. This indicates that the Keplerian orbit suffices for direction observations with an accuracy of $1''.5$-$2''.0$ (about as good as can be obtained for a single image with present optical equipment); however, it may not be good enough for range observations. Since a typical slant range to the PAGEOS satellite is 5000 km, the linear uncertainty of a single BC-4 image is 30-40 meters. Range observations, such as those produced by the SECOR equipment, may be an order of magnitude better than this. In general, it appears that the most appropriate model for short arcs is a numerical integration of a force function derived from a truncated expression of the geopotential. This model has the advantage that short period and secular effects of the J_2 harmonic are completely accounted for, so that the model very closely matches the actual orbit, at least for short periods of time. Since it is only necessary to integrate the orbit through a short period of time, and since many satellite positions are needed, the machine time necessary to perform the numerical integration is not much greater than that for a Keplerian orbit. Models utilizing numerical integration of a force function derived from a truncated geopotential are used by D. Brown Associates in their NEO-EMBET series of programs and by Cubic Corporation in programs developed for the processing of SECOR data.

One aspect of the adjustment of geodetic satellite data that has not been discussed in this report is the possibility of error modeling, where certain systematic errors in the observations are modeled and the unknown parameters in these models are solved for along with all other unknowns in the adjustment. Error modeling is probably never necessary for optical direction observations, since most errors affecting these observations are well known and can be removed. On the other hand, systematic effects such as frequency drift, zero-set error, etc., may be quite significant in electronic measurements. If error models are utilized, orbital constraints may become quite necessary, since the elimination of the satellite position unknowns allows the addition of new unknowns, such as the coefficients of the error models.

Although the orbital mode adjustment possesses many advantages over the geometrical mode, few of these advantages are often realized to a significant degree in practice. However, since it is always capable of producing a solution that is at least as good as the geometrical mode solution, the orbital mode adjustment does offer a valuable tool for comparison and analysis, and in some cases may be the clearly superior method of adjustment of geodetic satellite data.

Appendix
Explicit Form of Matrices Arising in the Observation Equations

$$Q = \frac{\partial(\alpha', \delta', R')}{\partial(u', v', w')} = \begin{pmatrix} \frac{1}{R'\cos\delta'} & 0 & 0 \\ 0 & \frac{1}{R'} & 0 \\ 0 & 0 & 1 \end{pmatrix} R_3(90^\circ)\, R_2(90^\circ - \delta')\, R_3(\alpha')\, R_3(-\theta)$$

$$= \begin{pmatrix} \frac{-\sin t'}{R'\cos\delta'} & \frac{\cos t'}{R'\cos\delta'} & 0 \\ \frac{-\cos t'\sin\delta'}{R'} & \frac{-\sin t'\sin\delta'}{R'} & \frac{\cos\delta'}{R'} \\ \cos t'\cos\delta' & \sin t'\cos\delta' & \sin\delta' \end{pmatrix}$$

where $t' = \alpha' - \theta$ and θ is the Greenwich Apparent Sidereal Time

$$S = \frac{\partial(u', v', w')}{\partial(\alpha, \delta, R)} = R_3(\theta)\, R_3(-\alpha)\, R_2(\delta - 90^\circ)\, R_3(-90^\circ) \begin{pmatrix} R\cos\delta & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -R\cos\delta\sin t & -R\sin\delta\cos t & \cos\delta\cos t \\ R\cos\delta\cos t & -R\sin\delta\sin t & \cos\delta\sin t \\ 0 & R\cos\delta & \sin\delta \end{pmatrix}$$

where $t = \alpha - \theta$

$$P = R_2(x_p)\, R_1(y_p) \approx \begin{pmatrix} 1 & 0 & -x_p \\ 0 & 1 & y_p \\ x_p & -y_p & 1 \end{pmatrix}$$

where x_p, y_p are the components of polar motion.

$$U(1,1) = \frac{\partial\alpha}{\partial\omega} = \cos i \sec^2\delta$$

$$U(1,2) = \frac{\partial\alpha}{\partial\Omega} = 1$$

$$U(1,3) = \frac{\partial\alpha}{\partial i} = -\cos(\alpha-\Omega)\sin(\alpha-\Omega)\tan i$$

$$U(1,4) = \frac{\partial\alpha}{\partial e} = \frac{\cos i \sin f}{\cos^2\delta}\left(\frac{1}{1-e^2} + \frac{1}{1-e\cos E}\right)$$

$$U(1,5) = \frac{\partial\alpha}{\partial M} = \frac{\cos i}{\cos^2\delta}\quad\frac{\sqrt{1-e^2}}{(1-e\cos E)^2}$$

$$U(1,6) = \frac{\partial\alpha}{\partial a} = 0$$

$$U(2,1) = \frac{\partial\delta}{\partial\omega} = \sin i\cos(\alpha-\Omega)$$

$$U(2,2) = \frac{\partial\delta}{\partial\Omega} = 0$$

$$U(2,3) = \frac{\partial\delta}{\partial i} = \sin(\alpha-\Omega)$$

$$U(2,4) = \frac{\partial\delta}{\partial e} = \sin i\cos(\alpha-\Omega)\sin f\left(\frac{1}{1-e^2} + \frac{1}{1-e\cos E}\right)$$

$$U(2,5) = \frac{\partial\delta}{\partial M} = \sin i\cos(\alpha-\Omega)\frac{\sqrt{1-e^2}}{(1-e\cos E)^2}$$

$$U(2,6) = \frac{\partial\delta}{\partial a} = 0$$

$$U(3,1) = \frac{\partial R}{\partial\omega} = 0$$

$$U(3,2) = \frac{\partial R}{\partial\Omega} = 0$$

$$U(3,3) = \frac{\partial R}{\partial i} = 0$$

$$U(3,4) = \frac{\partial R}{\partial e} = \frac{a(e-\cos E)}{1-e\cos E}$$

$$U(3,5) = \frac{\partial R}{\partial M} = \frac{a e\sin E}{1-e\cos E}$$

$$U(3,6) = \frac{\partial R}{\partial a} = (1 - e \cos E)$$

where E is the eccentric anomaly and f is the true anomaly.

$$\bar{\omega} = \sum_{j=0}^{3} \omega \;(T - T_0)^j$$

$$D = 1 - \omega_4 \sin \bar{\omega} + \omega_5 \cos \bar{\omega}$$

$$C_j = (T - T_0)^j \qquad j = 1, 4$$

$$C_5 = \cos \bar{\omega}$$

$$C_6 = \sin \bar{\omega}$$

$$\frac{\partial \omega}{\partial \omega_j} = \frac{C_j}{D} \qquad j = 1\text{–}4$$

$$\frac{\partial \omega}{\partial \omega_j} = C_j \qquad j = 5, 6$$

$$\frac{\partial \omega}{\partial \Omega_j} = \frac{\partial \omega}{\partial i_j} = \frac{\partial \omega}{\partial e_j} = \frac{\partial \omega}{\partial M_j} = 0 \qquad j = 1\text{–}6$$

$$\left.\begin{array}{lll}
\dfrac{\partial E_i}{\partial E_{ij}} = C_j & & j = 1\text{–}6 \\[2ex]
\dfrac{\partial E_i}{\partial \omega_j} = (-E_{i4} \sin\bar{\omega} + E_{i5} \cos\bar{\omega}) C_j & & j = 1\text{–}4 \\[2ex]
\dfrac{\partial E_i}{\partial \omega_j} = 0 & & j = 5, 6 \\[2ex]
\dfrac{\partial E_i}{\partial i_j} = \dfrac{\partial E_i}{\partial e_j} = \dfrac{\partial E_i}{\partial M_j} = 0 & & j = 1\text{–}6
\end{array}\right\} \quad E_i = \{\Omega, i, e, M\}$$

$$\frac{\partial n}{\partial M_j} = j\, C_{j-1} \qquad j = 2\text{–}4$$

$$\frac{\partial a}{\partial M_j} = -\frac{2a}{3n}\frac{\partial n}{\partial M_j} \qquad j = 2\text{-}4$$

$$\frac{\partial a}{\partial M_j} = 0 \qquad j = 1,\ 5\text{-}6$$

$$\frac{\partial a}{\partial \omega_j} = \frac{\partial a}{\partial \Omega_j} = \frac{\partial a}{\partial i_j} = \frac{\partial a}{\partial e_j} = 0 \qquad j = 1\text{-}6$$

References

Aardoom, L., A. Girnius, and G. Veis (1966). "Geometric Methods," in "Geodetic Parameters for a 1966 Smithsonian Standard Earth," (C. A. Lundquist and G. Veis, editors), Special Report No. 200, Vol. I. Smithsonian Astrophysical Observatory, Cambridge, Mass.

Brown, Duane C. (1967a). Review of Current Geodetic Satellite Programs and Recommendations for Future Programs. D. Brown Associates, Inc., Eau Gallie, Florida.

Brown, Duane C. (1967b). Quarterly Progress Report for GEOS-A Short Arc Orbital Data Reduction and Analysis. D. Brown Associates, Inc., Eau Gallie, Florida.

Brown, Duane C. (1968). Lectures to the Summer Institute in Geodetic Science. The Ohio State University, Columbus, Ohio, June 17-21.

Danby, John M. A. (1962). Fundamentals of Celestial Mechanics. Macmillan Co., New York.

Faddeev, D. K. and V. N. Faddeeva (1963). Computational Methods of Linear Algebra. Translated by Robert C. Williams. W. H. Freeman, San Francisco.

Gaposchkin, E. M. (1966). "Orbit Determination," in "Geodetic Parameters for a 1966 Smithsonian Standard Earth," (C. A. Lundquist and G. Veis, editors), Special Report No. 200. Smithsonian Astrophysical Observatory, Cambridge, Mass.

Hamilton, Walter C. (1964). Statistics in Physical Science. Ronald Press, New York.

Hartwell, J. Graham (1966). "The Recursive Formulation of a Taylor Series for Orbit Integration," in Geodetic Data Analysis for GEOS-A, An Experimental Design. D. Brown Associates, Inc., Eau Gallie, Florida.

Hornbarger, Daniel H. (1968). "Comparison of Astrometric and Photogrammetric Plate Reduction Techniques for a Wild BC-4 Camera." Reports of the Department of Geodetic Science No. 106. The Ohio State University, Columbus.

Hotter, Frank D. (1967). "Preprocessing Optical Satellite Observations." Reports of the Department of Geodetic Science No. 82. The Ohio State University, Columbus.

Kaula, William M. (1966). Theory of Satellite Geodesy. Blaisdell Publishing Company, Waltham, Mass.

Kozai, Y. (1960). "Effect of Precession and Nutation on the Orbital Elements of a Close Earth Satellite." Astronomical Journal, 65, 621-623.

Krakiwsky, Edward J., and Allen J. Pope (1967). "Least Squares Adjustment of Satellite Observations for Simultaneous Directions or Ranges." Part 1 of 3: Formulation of Equations. Reports of the Department of Geodetic Science No. 86. The Ohio State University, Columbus.

Longley, James W. (1967). "An Appraisal of Least Squares Programs from the Point of View of the User." Journal of the American Statistical Association, 62, 819-841. Washington, D. C.

Magnuss, T. A., and J. B. McGuire (1962). "Comparison of Least Squares and Minimum Variance Estimates of Regression Parameters." Annals of Mathematical Statistics, 33, 2.

Mueller, Ivan I. (1964). Introduction to Satellite Geodesy. Frederick Ungar Publishing Company, New York.

Rapp, R. (1967). Unpublished, untitled report.

Schmid, Hellmut H. (1968). "Application of Photogrammetry to Three-Dimensional Geodesy (Photogrammetric Satellite Triangulation)." Invited paper presented to Commission III, XIth Congress of the International Society of Photogrammetry, Lausanne, Switzerland, 8-20 July.

Schwarz, Charles R. (1967a). "Optimal Filter and Least Squares Estimation." Army Map Service Geodetic Memorandum No. 1610. Washington, D. C.

Schwarz, Charles R. (1967b). "Effects of Unestimated Parameters in Least Squares Adjustments." Presented to the Symposium on Computational Photogrammetry of the Potomac Region of the American Society of Photogrammetry, Dec. 4-8.

Taylor, Eugene A. and B. Frank Lampton (1963). "A Report on the Camera Calibration Phase of the USCGS Satellite Geodesy Program." U. S. Coast and Geodetic Survey, Washington, D. C.

Uotila, Urho A. (1967). "Introduction to Adjustment Computations with Matrices." Department of Geodetic Science, The Ohio State University. Columbus.

Veach, James P. (1968). "Investigations into the Utilization of Passive Satellite Observational Data." Reports of the Department of Geodetic Science No. 110. The Ohio State University. Columbus.

Veis, G. (1963). "Precise Aspects of Terrestrial and Celestial Reference Frames." Special Report No. 123. Smithsonian Astrophysical Observatory, Cambridge, Mass.

Wintz, E.K. (1961). "Geodetic Field Use of an Optically Visible Earth Satellite." M.S. Thesis. Department of Geodetic Science, The Ohio State University, Columbus.

FOR THE DEPARTMENT OF GEODETIC SCIENCE

Project Supervisor Ivan I. Mueller Date 1.15.1969

FOR THE OHIO STATE UNIVERSITY RESEARCH FOUNDATION

Executive Director Robert C. Stephenson Date 1/21/69