

Chapter 1

Orbit Determination Concepts

1.1 INTRODUCTION

The treatment presented here will cover the fundamentals of satellite orbit determination and its evolution over the past four decades. By satellite orbit determination we mean the process by which we obtain knowledge of the satellite's motion relative to the center of mass of the Earth in a specified coordinate system. Orbit determination for celestial bodies has been a general concern of astronomers and mathematicians since the beginning of civilization and indeed has attracted some of the best analytical minds to develop the basis for much of the fundamental mathematics in use today.

The classical orbit determination problem is characterized by the assumption that the bodies move under the influence of a central (or point mass) force. In this treatise, we focus on the problem of determining the orbits of noncelestial satellites. That is, we focus on the orbits of objects placed into orbit by humans. These objects differ from most natural objects in that, due to their size, mass, and orbit characteristics, the nongravitational forces are of significant importance. Further, most satellites orbit near to the surface and for objects close to a central body, the gravitational forces depart from a central force in a significant way.

By the state of a dynamical system, we mean the set of parameters required to predict the future motion of the system. For the satellite orbit determination problem the minimal set of parameters will be the position and velocity vectors at some given epoch. In subsequent discussions, this minimal set will be expanded to include dynamic and measurement model parameters, which may be needed to improve the prediction accuracy. This general state vector at a time, t , will be denoted as $\mathbf{X}(t)$. The general orbit determination problem can then be posed as follows.

If at some initial time, t_0 , the state \mathbf{X}_0 of a vehicle following a ballistic trajectory is given and if the differential equations that govern the motion of the vehicle are known, then the equations of motion can be integrated to determine the state of the vehicle at any time. However, during an actual flight, the initial state is never known exactly. Moreover, certain physical constants as well as the mathematical specification of the forces required to define the differential equations of motion are known only approximately. Such errors will cause the actual motion to deviate from the predicted motion. Consequently, in order to determine the position of the spacecraft at some time $t > t_0$, it must be tracked or observed from tracking stations whose positions are known precisely. With the observations of the spacecraft motion, a better estimate of the trajectory can be determined. The term “better estimate” is used since the observations will be subject to both random and systematic errors and, consequently, the estimate of the trajectory will not be exact.

The observational data will usually consist of such measurements as range, range-rate, azimuth, elevation, or other observable quantities. That is, the state variables (position, velocity, unknown model parameters, etc.) will not be observed, but rather the observable will usually be some nonlinear function of the state variables.

The problem of determining the best estimate of the state of a spacecraft, whose initial state is unknown, from observations influenced by random and systematic errors, using a mathematical model that is not exact, is referred to as the problem of state estimation. In this presentation, such a procedure will be referred to as the process of *orbit determination*. The word “best” is used to imply that the estimate is optimal in some as yet undefined statistical sense.

When an estimate of the trajectory has been made, the subsequent motion and values for the observations can be predicted. In the orbit determination procedure, the process of predicting the state of a vehicle is referred to as “generating an ephemeris.” An ephemeris for a space vehicle is a table of its position and velocity components as a function of time. The predicted values will differ from the true values due to the following effects:

1. Inaccuracies in the estimated state vector (i.e., position and velocity vector) caused by errors in the orbit determination process, such as:
 - a. Approximations involved in the method of orbit improvement and in the mathematical model,
 - b. Errors in the observations,
 - c. Errors in the computational procedure used in the solution process.
2. Errors in the numerical integration procedure caused by errors in the dynamical model and computer truncation and roundoff errors.

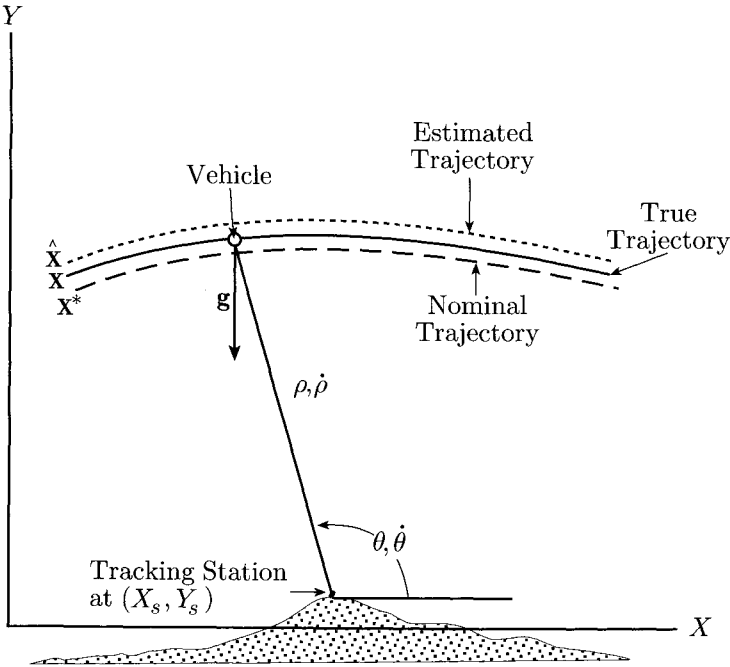


Figure 1.2.1: Uniform gravity field trajectory.

Consequently, the process of observation and estimation must be repeated continually as the vehicle's motion evolves. Furthermore, the orbit determination procedure may be used to obtain better estimates of the location of tracking stations, adjust the station clocks, calibrate radar biases, obtain a better estimate of geophysical constants, and so on. In fact, any observable quantity directly affecting either the motion of the vehicle or the observation-state relationship can be determined by appropriately using the observational data.

To formulate the orbit determination problem, we will use elements of probability and statistics as well as matrix theory. Appendixes A and B contain introductory material on these topics.

1.2 UNIFORM GRAVITY FIELD MODEL

To illustrate some of the basic ideas involved in the orbit determination process, consider the flight of a vehicle moving under the influence of a uniform gravitational field, as shown in Fig. 1.2.1. In the preliminary design of the trajectory, an initial state \mathbf{X}_0^* (i.e., the initial position and the initial velocity vectors), is

selected in such a way that the vehicle will follow a desired nominal (or design) trajectory. In practice, however, the design conditions are never realized. Hence, the *true initial state*, \mathbf{X}_0 , will differ from the *nominal initial state*, \mathbf{X}_0^* , and consequently, the true trajectory followed by the vehicle will differ from the nominal trajectory. An indication of these two trajectories is given in Fig. 1.2.1. Here the true trajectory is denoted as \mathbf{X} , the nominal trajectory as \mathbf{X}^* , and the best estimate of the true trajectory is indicated by $\hat{\mathbf{X}}$. To determine an estimate, $\hat{\mathbf{X}}$, of the true trajectory of the vehicle, tracking information must be used.

Figure 1.2.1 shows a tracking station at a point whose coordinates are (X_s, Y_s) . Assume that the range (linear distance along the line of sight), ρ , and the elevation, θ , of the line of sight to the satellite can be measured. The tracking information, or observations, ρ and θ , at any time depend on the true state of the satellite and the position of the tracking station. Since the true position of the satellite is unknown, the calculated or modeled values of the observations will depend on the nominal position of the satellite and the location of the tracking station at the time the observation is made. The difference between observed and calculated observations at the observation time provides the information used to obtain an improved estimate of the satellite's motion. Other observations such as range rate, $\dot{\rho}$, and elevation rate, $\dot{\theta}$, which depend on both the velocity and position of the satellite and tracking station, could also be used to determine $\hat{\mathbf{X}}$. Note that range rate is the line of sight velocity that is, the projection of the spacecraft's velocity vector on to the range vector

$$\dot{\rho} = \frac{\dot{\boldsymbol{\rho}} \cdot \boldsymbol{\rho}}{\rho}.$$

1.2.1 FORMULATION OF THE PROBLEM

For the uniform gravity field model, the differential equations of motion can be expressed as

$$\begin{aligned}\ddot{\mathbf{X}}(t) &= 0 \\ \ddot{\mathbf{Y}}(t) &= -g,\end{aligned}\tag{1.2.1}$$

where g is the gravitational acceleration and is assumed to be constant. Integration of Eq. (1.2.1) leads to

$$\begin{aligned}X(t) &= X_0 + \dot{X}_0 t \\ Y(t) &= Y_0 + \dot{Y}_0 t - g \frac{t^2}{2}, \\ \dot{X}(t) &= \dot{X}_0 \\ \dot{Y}(t) &= \dot{Y}_0 - gt\end{aligned}\tag{1.2.2}$$

where t represents the time, and the reference time, t_0 , is chosen to be zero. The subscript 0 indicates values of the quantities at t_0 . Now, providing the values of

X_0 , Y_0 , \dot{X}_0 , \dot{Y}_0 , and g are given, Eq. (1.2.2) can be used to predict the position and velocity of the vehicle at any time. However, as previously indicated, the values of these quantities are never known exactly in practice, and it is the task of the orbit determination procedure to estimate their values.

1.2.2 THE EQUATION OF THE ORBIT

The parameter t can be eliminated in the last two of Eq. (1.2.2) to obtain the curve followed by the vehicle in the (X, Y) plane; that is,

$$t = \frac{X - X_0}{\dot{X}_0}. \quad (1.2.3)$$

Hence,

$$Y(t) = Y_0 + \frac{\dot{Y}_0}{\dot{X}_0}(X - X_0) - \frac{1}{2} \frac{g}{\dot{X}_0^2}(X - X_0)^2. \quad (1.2.4)$$

Equation (1.2.4) can be recognized readily as the equation of a parabola. This equation is analogous to the conic curve obtained as the solution to motion in an inverse square force field (see Chapter 2). It should be noted that Eqs. (1.2.3) and (1.2.4) still involve the values of the four initial conditions. The relation indicates that if the vehicle was at X' , then it must also have been at Y' obtained from Eq. (1.2.4). It is also obvious that any real value of X is possible, and it has associated with it a unique Y . However, the converse is not true. For each possible value of Y , there are usually two values of X possible. The time epoch, t' , corresponding to the point (X', Y') can be determined from the relation $t' = (X' - X_0)/\dot{X}_0$. This relation is analogous to Kepler's equation for the motion in an inverse square force field. Since the epoch of time must be associated with the determination of the state of the vehicle if an ephemeris is to be generated and since the time is an integral part of all observations, the solution in the form of Eqs. (1.2.3) and (1.2.4) is usually not convenient for most orbit determination procedures.

In the subsequent discussions, either the differential equations, Eq. (1.2.1), or their solution in the form of Eq. (1.2.2) will be used.

1.2.3 THE ROLE OF THE OBSERVATION

As an example of the orbit determination procedure, consider the situation in which the state is observed at some time epoch, t_j . Then if X_j , Y_j , \dot{X}_j and \dot{Y}_j are given at t_j , Eq. (1.2.2) can be used to form four equations in terms of four unknowns. This system of equations can be used to determine the unknown

components of the initial state. For example, from Eq. (1.2.2) it follows that

$$\begin{bmatrix} X_j \\ Y_j + gt_j^2/2 \\ \dot{X}_j \\ \dot{Y}_j + gt_j \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_j & 0 \\ 0 & 1 & 0 & t_j \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ Y_0 \\ \dot{X}_0 \\ \dot{Y}_0 \end{bmatrix}.$$

Then, the initial state can be determined as follows:

$$\begin{bmatrix} X_0 \\ Y_0 \\ \dot{X}_0 \\ \dot{Y}_0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_j & 0 \\ 0 & 1 & 0 & t_j \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} X_j \\ Y_j + gt_j^2/2 \\ \dot{X}_j \\ \dot{Y}_j + gt_j \end{bmatrix}. \quad (1.2.5)$$

Providing that the matrix inverse in Eq. (1.2.5) exists, the true initial state $\mathbf{X}_0^T = [X_0, Y_0, \dot{X}_0, \dot{Y}_0]$ can be determined.

Unfortunately, in an actual trajectory or orbit determination process, the individual components of the state generally cannot be observed directly. Rather, the observations consist of nonlinear functions of the state; for example, range, elevation, range rate, and so on. In this case, the nonlinear observation-state relationship is

$$\begin{aligned} \rho &= \sqrt{(X - X_s)^2 + (Y - Y_s)^2} \\ \tan \theta &= (Y - Y_s)/(X - X_s) \\ \dot{\rho} &= \frac{1}{\rho} [(X - X_s)(\dot{X} - \dot{X}_s) + (Y - Y_s)(\dot{Y} - \dot{Y}_s)] \\ \dot{\theta} &= \frac{1}{\rho^2} [(X - X_s)(\dot{Y} - \dot{Y}_s) - (\dot{X} - \dot{X}_s)(Y - Y_s)], \end{aligned} \quad (1.2.6)$$

where X_s, Y_s, \dot{X}_s , and \dot{Y}_s are the position and velocity components, respectively, of the tracking station. Generally the station velocity components \dot{X}_s and \dot{Y}_s will be zero relative to an Earth-fixed frame unless we are accounting for tectonic plate motion. Now, if $\rho, \theta, \dot{\rho}$ and $\dot{\theta}$ are given at t_j , then Eq. (1.2.2) can be substituted into Eq. (1.2.6) to obtain

$$\begin{aligned} \rho_j &= \sqrt{(X_0 - X_s + \dot{X}_0 t_j)^2 + (Y_0 - Y_s + \dot{Y}_0 t_j - gt_j^2/2)^2} \\ \theta_j &= \tan^{-1}[(Y_0 - Y_s + \dot{Y}_0 t_j - g \frac{t_j^2}{2}) / (X_0 - X_s + \dot{X}_0 t_j)] \\ \dot{\rho}_j &= \frac{1}{\rho_j} [(X_0 - X_s + \dot{X}_0 t_j)(\dot{X}_0 - \dot{X}_s) \\ &\quad + (Y_0 - Y_s + \dot{Y}_0 t_j - g \frac{t_j^2}{2})(\dot{Y}_0 - gt_j - \dot{Y}_s)] \end{aligned} \quad (1.2.7)$$

$$\begin{aligned}\dot{\theta}_j = & \frac{1}{\rho_j^2} [(X_0 - X_s + \dot{X}_0 t_j)(\dot{Y}_0 - g t_j - \dot{Y}_s) \\ & - (\dot{X}_0 - \dot{X}_s)(Y_0 - Y_s + \dot{Y}_0 t_j - g \frac{t_j^2}{2})].\end{aligned}$$

Symbolically Eq. (1.2.7) represents four nonlinear algebraic equations in terms of the unknown components of the initial state. If \mathbf{Y}^T represents the vector whose components are the observations $[\rho_j, \theta_j, \dot{\rho}_j, \dot{\theta}_j]$, then Eq. (1.2.7) can be expressed as

$$J(\mathbf{X}_0) \equiv \mathbf{Y} - G(\mathbf{X}_0, t) = 0, \quad (1.2.8)$$

where $G(\mathbf{X}_0, t)$ is a 4×1 vector of nonlinear functions consisting of the right-hand side of Eq. (1.2.7). Eq. (1.2.8) can be solved iteratively by the Newton–Raphson iteration procedure. If \mathbf{X}_0^n represents the n th approximation to the solution, then

$$J(\mathbf{X}_0^{n+1}) \cong J(\mathbf{X}_0^n) + \left[\frac{\partial J}{\partial \mathbf{X}_0^n} \right] [\mathbf{X}_0^{n+1} - \mathbf{X}_0^n],$$

according to a 2-term Taylor series approximation. Now, setting $J(\mathbf{X}_0^{n+1}) = 0$ to comply with Eq. (1.2.8) and solving for \mathbf{X}_0^{n+1} leads to

$$\mathbf{X}_0^{n+1} = \mathbf{X}_0^n - \left[\frac{\partial J}{\partial \mathbf{X}_0^n} \right]^{-1} J(\mathbf{X}_0^n), \quad (1.2.9)$$

where

$$\frac{\partial J}{\partial \mathbf{X}_0^n} = - \frac{\partial G(\mathbf{X}_0^n, t)}{\partial \mathbf{X}_0^n}.$$

The process can be repeated until $\|\mathbf{X}_0^{n+1} - \mathbf{X}_0^n\| \leq \epsilon$, where ϵ is a small positive number.

Several facts are obvious from this example:

1. The linear dependence of the state at time t on the initial conditions, as obtained in Eq. (1.2.2), usually does not occur.
2. In general, the equations that define the solution to the orbit determination problem are nonlinear, and the solution must be obtained by a numerical iteration procedure. Note that this iteration procedure requires an initial guess for \mathbf{X}_0 . This initial guess must be close enough to the true solution for convergence to occur.
3. Since the equations are nonlinear, multiple solutions may exist; only one will correspond to the minimizing or correct solution. Generally the initial guess for \mathbf{X}_0 will be close to the truth so convergence to an incorrect solution is not a problem.

This situation represents a special case of the usual orbit determination problem; that is, the case where there are as many independent observations at one epoch of time as there are unknown components of the state vector. Usually, the number of observations available at any time will be smaller than the number of components in the state vector. For example, the radar may have the capability of measuring only range, ρ , and range rate, $\dot{\rho}$, or range and elevation, θ , at any one point in time. The observation set (ρ, θ) is not adequate to determine the state $(X_0, Y_0, \dot{X}_0, \dot{Y}_0)$, since using Eq. (1.2.6) will lead to a set of two nonlinear equations in terms of four unknowns. However, if observations are available at any two time points t_1 and t_2 , then the set $(\rho_1, \theta_1, \rho_2, \theta_2)$ can be used to determine the trajectory. From Eq. (1.2.7), the following equations can be obtained:

$$\begin{aligned}
 \rho_1 &= \sqrt{(X_0 - X_s + \dot{X}_0 t_1)^2 + (Y_0 - Y_s + \dot{Y}_0 t_1 - g t_1^2/2)^2} \\
 \theta_1 &= \tan^{-1}[(Y_0 - Y_s + \dot{Y}_0 t_1 - g t_1^2/2)/(X_0 - X_s + \dot{X}_0 t_1)] \\
 \rho_2 &= \sqrt{(X_0 - X_s + \dot{X}_0 t_2)^2 + (Y_0 - Y_s + \dot{Y}_0 t_2 - g t_2^2/2)^2} \\
 \theta_2 &= \tan^{-1}[(Y_0 - Y_s + \dot{Y}_0 t_2 - g t_2^2/2)/(X_0 - X_s + \dot{X}_0 t_2)].
 \end{aligned} \tag{1.2.10}$$

Equation (1.2.10) represents four nonlinear equations in terms of four unknowns from which the unknowns can be determined as indicated by Eq. (1.2.9).

The approaches just outlined have two fundamental assumptions that cannot be realized in practice. These assumptions are:

1. Perfect knowledge of the differential equations that describe the motion.
2. Perfect observations.

The assumption of perfect knowledge of the differential equations implies that all the forces are modeled perfectly and that the values for the constant parameters used in the force model are exact. Usually, the true values for all of the parameters are not defined with complete numerical accuracy. But what is even more relevant, simplifying assumptions usually are made to reduce the computational requirements. Under such assumptions, certain unimportant or “negligible” components of the forces acting on the vehicle are omitted. As a consequence, the mathematical model used to describe the dynamical process will differ from the true process.

In general, the observations will be corrupted by random and systematic observation errors that are inherent in the measurement process. Hence, the observed value of the range, ρ , or elevation, θ , will differ from the true value because of these errors. In addition, the true and observed values for the range and the elevation will differ because of imperfect knowledge of the tracking station location, (X_s, Y_s) .

Because of these factors, four observations will not be sufficient to determine the planar motion trajectory exactly. In the usual operational situation, many observations are made, and the trajectory is selected to give the best agreement with the observations obtained. In general the mechanism for obtaining the best estimate is to linearize the problem by expanding the equations of motion and the observation-state relationship about a reference trajectory. Deviations from the reference trajectory (e.g., position and velocity) are then determined to yield the “best” agreement with the observations. As we will see, the “best” agreement generally is based on the least squares criterion.

1.2.4 LINEARIZATION PROCEDURE

To illustrate the linearization procedure, return to the flat Earth example discussed previously. Assuming that there are errors in each of the initial position and velocity values and in g , the state vector is $X^T = [X, Y, \dot{X}, \dot{Y}, g]$. It is a straightforward procedure to replace X by $X^* + \delta X$, Y by $Y^* + \delta Y$, \dot{X} by $\dot{X}^* + \delta \dot{X}$, \dot{Y} by $\dot{Y}^* + \delta \dot{Y}$, and g by $g^* + \delta g$ in Eq. (1.2.2) to obtain the equation for perturbation, or deviation, from the reference trajectory. The $(\)^*$ values represent assumed or specified values. If the equations for the reference trajectory, designated by the $(\)^*$ values, are subtracted from the result obtained by this substitution, the equations for the state deviations are obtained as

$$\begin{aligned}\delta X &= \delta X_0 + \delta \dot{X}_0 t \\ \delta Y &= \delta Y_0 + \delta \dot{Y}_0 t - \delta g \frac{t^2}{2} \\ \delta \dot{X} &= \delta \dot{X}_0 \\ \delta \dot{Y} &= \delta \dot{Y}_0 - \delta g t \\ \delta g &= \delta g.\end{aligned}\tag{1.2.11}$$

Equation (1.2.11) can be expressed in matrix form in the following manner:

$$\begin{bmatrix} \delta X \\ \delta Y \\ \delta \dot{X} \\ \delta \dot{Y} \\ \delta g \end{bmatrix} = \begin{bmatrix} 1 & 0 & t & 0 & 0 \\ 0 & 1 & 0 & t & -t^2/2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta X_0 \\ \delta Y_0 \\ \delta \dot{X}_0 \\ \delta \dot{Y}_0 \\ \delta g \end{bmatrix}.\tag{1.2.12}$$

Equation (1.2.12) can be used to predict the deviation of the vehicle from the reference trajectory, $\mathbf{X}^*(t)$, at any time $t > 0$. Note that the deviations from the true state, $\delta \mathbf{X}(t) \equiv \mathbf{x}(t)$, will be caused by deviations in the original state or deviations in the gravitational acceleration, δg . However, the quantities $\delta X_0, \delta Y_0, \dots$, are not known, and it is the problem of the orbit determination procedure to estimate their values.

In this example, both Eqs. (1.2.2) and (1.2.12) are linear. In the case of central force field motion, the equations analogous to Eq. (1.2.2) are not linear, however, and the orbit determination process is simplified considerably by the linearization procedure. Care must be taken so that deviations from the reference conditions remain in the linear range.

As mentioned earlier, the observations ρ and θ , which are taken as the satellite follows the true trajectory, are nonlinear functions of X , Y , X_s , and Y_s . Furthermore, they contain random and systematic errors represented by ϵ_ρ and ϵ_θ ; hence,

$$\begin{aligned}\rho &= \sqrt{(X - X_s)^2 + (Y - Y_s)^2} + \epsilon_\rho \\ \theta &= \tan^{-1} \left(\frac{Y - Y_s}{X - X_s} \right) + \epsilon_\theta.\end{aligned}\quad (1.2.13)$$

Our objective is to linearize the observations with respect to the reference trajectory. This can be accomplished by expanding Eq. (1.2.13) in a Taylor series about the reference or nominal trajectory as follows:

$$\begin{aligned}\rho &\cong \rho^* + \left[\frac{\partial \rho}{\partial X} \right]^* (X - X^*) + \left[\frac{\partial \rho}{\partial Y} \right]^* (Y - Y^*) + \epsilon_\rho \\ \theta &\cong \theta^* + \left[\frac{\partial \theta}{\partial X} \right]^* (X - X^*) + \left[\frac{\partial \theta}{\partial Y} \right]^* (Y - Y^*) + \epsilon_\theta.\end{aligned}\quad (1.2.14)$$

Note that the partials with respect to \dot{X} , \dot{Y} , and g are zero since they do not appear explicitly in Eq. (1.2.13). By defining

$$\begin{aligned}\delta \rho &= \rho - \rho^* \\ \delta \theta &= \theta - \theta^* \\ \delta X &= X - X^* \\ \delta Y &= Y - Y^*,\end{aligned}\quad (1.2.15)$$

we can write Eq. (1.2.14) as

$$\begin{aligned}\delta \rho &= \left[\frac{\partial \rho}{\partial X} \right]^* \delta X + \left[\frac{\partial \rho}{\partial Y} \right]^* \delta Y + \epsilon_\rho \\ \delta \theta &= \left[\frac{\partial \theta}{\partial X} \right]^* \delta X + \left[\frac{\partial \theta}{\partial Y} \right]^* \delta Y + \epsilon_\theta.\end{aligned}\quad (1.2.16)$$

The symbol $[]^*$ indicates that the value in brackets is evaluated along the nominal trajectory. In Eq. (1.2.16), terms of order higher than the first in the state deviation values have been neglected assuming that these deviations are small. This requires that the reference trajectory and the true trajectory be close at all times in the

interval of interest. Now, if the following definitions are used,

$$\begin{aligned}
 \mathbf{y}^T &\equiv [\delta\rho \ \delta\theta] \\
 \tilde{H} &\equiv \begin{bmatrix} \left[\frac{\partial\rho}{\partial X}\right]^* & \left[\frac{\partial\rho}{\partial Y}\right]^* & 0 & 0 & 0 \\ \left[\frac{\partial\theta}{\partial X}\right]^* & \left[\frac{\partial\theta}{\partial Y}\right]^* & 0 & 0 & 0 \end{bmatrix} \\
 \mathbf{x}^T &\equiv [\delta X \ \delta Y \ \delta\dot{X} \ \delta\dot{Y} \ \delta g] \\
 \boldsymbol{\epsilon}^T &\equiv [\epsilon_\rho \ \epsilon_\theta],
 \end{aligned} \tag{1.2.17}$$

then Eq. (1.2.16) can be expressed as the matrix equation,

$$\mathbf{y} = \tilde{H}\mathbf{x} + \boldsymbol{\epsilon}. \tag{1.2.18}$$

In Eq. (1.2.18), \mathbf{y} is called the *observation deviation vector*, \mathbf{x} is the *state deviation vector*, \tilde{H} is a mapping matrix that relates the observation deviation vector to the state deviation vector, and $\boldsymbol{\epsilon}$ is a random vector that represents the noise or error in the observations. It is common practice to refer to \mathbf{x} and \mathbf{y} as the *state and observation vectors*, respectively, even though they really represent deviations from nominal values.

1.2.5 STATE TRANSITION MATRIX

The matrix multiplying the initial state vector in Eq. (1.2.12) is referred to as the *state transition matrix*, $\Phi(t, t_0)$. For the state deviation vector, \mathbf{x} , defined in Eq. (1.2.17), this matrix is given by

$$\Phi(t, t_0) = \begin{bmatrix} 1 & 0 & t & 0 & 0 \\ 0 & 1 & 0 & t & -t^2/2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{1.2.19}$$

The state transition matrix maps deviations in the state vector from one time to another. In this case, deviations in the state are mapped from t_0 to t . Note that we have arbitrarily set $t_0 = 0$. For this linear example, the mapping is exact, but in the general orbit determination case, the state equations will be nonlinear, and the state transition matrix is the linear term in a Taylor series expansion of the state vector at time t about the state vector at a reference time t_0 . Hence, the state deviation vector and observation deviation vector at any time can be written in

terms of \mathbf{x}_0 , as follows:

$$\begin{aligned}\mathbf{x}(t) &= \Phi(t, t_0)\mathbf{x}_0 \\ \mathbf{y}(t) &= \tilde{H}(t)\mathbf{x}(t) + \epsilon(t).\end{aligned}\tag{1.2.20}$$

$\mathbf{y}(t)$ can be written in terms of \mathbf{x}_0 as

$$\mathbf{y}(t) = H(t)\mathbf{x}_0 + \epsilon(t),\tag{1.2.21}$$

where

$$H(t) = \tilde{H}(t)\Phi(t, t_0).\tag{1.2.22}$$

The problem that remains now is to determine the best estimate of \mathbf{x}_0 given the linearized system represented by Eqs. (1.2.20) and (1.2.21). Our problem can now be summarized as follows. Given an arbitrary epoch, t_k , and the state propagation equation and observation-state relationship

$$\begin{aligned}\mathbf{x}(t) &= \Phi(t, t_k)\mathbf{x}_k \\ \mathbf{y} &= H\mathbf{x}_k + \epsilon,\end{aligned}\tag{1.2.23}$$

find the “best” estimate of \mathbf{x}_k . In Eq. (1.2.23), \mathbf{y} is an $m \times 1$ vector, \mathbf{x} is an $n \times 1$ vector, ϵ is an $m \times 1$ vector, and H is an $m \times n$ mapping matrix, where n is the number of state variables and m is the total number of observations. If m is sufficiently large, the essential condition $m \geq n$ is satisfied.

In contrast to the algebraic solutions described earlier in this chapter, the system represented by the second of Eq. (1.2.23) is always underdetermined. That is, there are m -knowns (e.g., the observation deviations, \mathbf{y}) and $m + n$ unknowns (e.g., the m -observation errors, ϵ , and the n -unknown components of the state deviation vector, \mathbf{x}_k). In Chapter 4, several approaches for resolving this problem are discussed. The most straightforward is based on the method of least squares as proposed by Gauss (1809). In this approach, the best estimate for the unknown state vector, \mathbf{x}_k , is selected as the value $\hat{\mathbf{x}}_k$, which minimizes the sum of the squares of the calculated values of the observations errors. That is, if \mathbf{x}_k^0 is any value of \mathbf{x}_k , then $\epsilon^0 = \mathbf{y} - H\mathbf{x}_k^0$ will be the m -calculated values of the observation residuals corresponding to the value \mathbf{x}_k^0 . Then, the best estimate of \mathbf{x}_k will be the value that leads to a minimal value of the performance index, $J(\mathbf{x}_k^0)$, where

$$J(\mathbf{x}_k^0) = 1/2(\epsilon^{0T}\epsilon^0) = 1/2(\mathbf{y} - H\mathbf{x}_k^0)^T(\mathbf{y} - H\mathbf{x}_k^0).\tag{1.2.24}$$

For a minimum of this quantity, it is necessary and sufficient that

$$\left. \frac{\partial J}{\partial \mathbf{x}_k^0} \right|_{\hat{\mathbf{x}}_k} = 0; \quad \delta \mathbf{x}_k^{0T} \left. \frac{\partial^2 J}{\partial \mathbf{x}_k^0 \partial \mathbf{x}_k^0} \right|_{\hat{\mathbf{x}}_k} \delta \mathbf{x}_k^0 > 0.\tag{1.2.25}$$

From the first of the conditions given by Eq. (1.2.25), it follows that

$$(\mathbf{y} - H\hat{\mathbf{x}}_k)^T H = 0$$

or by rearranging that

$$H^T H \hat{\mathbf{x}}_k = H^T \mathbf{y}. \quad (1.2.26)$$

If the $n \times n$ matrix $(H^T H)$ has an inverse, the solution can be expressed as

$$\hat{\mathbf{x}}_k = (H^T H)^{-1} H^T \mathbf{y}.$$

Chapter 4 expands and clarifies the procedure and techniques for determining, and assessing the accuracy of, the estimate $\hat{\mathbf{x}}_k$. Further details are given on computational methods for determining $\hat{\mathbf{x}}_k$ in Chapter 5.

1.3 BACKGROUND AND OVERVIEW

As noted in this introductory chapter, orbit determination methodology can be separated into two general classes, for example, classical (or deterministic) orbit determination and modern (or statistical based) orbit determination. In the classical approach, observational errors are not considered and the problem is reduced to making the necessary number of measurements to determine the orbit. There is an equivalence between the number of observations and the number of unknowns. The modern orbit determination problem recognizes the influence of observation errors and, to minimize the effects of the observational error, more observations are used than the number of parameters to be estimated. Modern orbit determination methodology dates back to the beginning of the 19th century, when Legendre (1806), in attempting to determine the orbits of the planets, recognized the importance of observational errors and proposed the least squares solution as a method for minimizing the effect of such errors. Gauss (1809) presented a probabilistic basis for the method of least squares, which was closely akin to the maximum likelihood method proposed by Fisher (1912). Gauss (1809) stated that he had used the method since 1795 and historians (Cajori, 1919) generally agree that Gauss originated the method. The notation of Gauss is still in use. Laplace (1812) added refinements to the computational approach and to the statistical foundations. Markov (1898), in his work on the mathematical foundations of probability theory, clarified many of the concepts associated with the method of least squares. Helmert (1900) extended the application to astronomy and geodesy. In the 20th century, Kolmogorov (1941) and Wiener (1949) linked the problem to modern systems theory and set the stage for the sequential filter formulation for processing tracking data. In 1959, Swerling gave the first conceptual approach to the sequential orbit determination problem. Kalman (1960) presented a mathematically rigorous approach for the sequential processing of observations of a linear dynamic system. In a later work, Kalman and Bucy (1961)

gave an approach for the extended Kalman filter, which is the sequential approach for the general nonlinear orbit determination problem. Following the publication of Kalman's paper in 1960, there was an extensive series of publications related to this topic during the next three decades. Many of these investigations will be referenced in the subsequent chapters. A summary of the historical development of the least squares approach is given by Sorenson (1970).

The main thrust of this book will be aimed at developing computational data processing techniques for estimating the state of orbiting satellites. As noted in the previous discussion, the problem formulation will lead to a nonlinear relation between the observations and the epoch state of the satellite. Therefore, in modern system theory terms, we will be concerned with estimating the state of a nonlinear dynamic system, using observations of nonlinear functions of the system state, where the observations are corrupted by random and systematic observational errors. Although the main focus for the physical models will be the orbit determination problem, the problem will be cast in the state space notation defined by modern system theory. As such, the computational approaches presented in the treatment will be applicable to a wide range of problems, including missile guidance, attitude control, economic system analysis, chemical process control, and a wide variety of manufacturing processes. Once the defining equations from these processes have been reduced to the state space formulation, the computational techniques developed here will be applicable.

1.4 SUMMARY

Following the introduction of the basic concepts of nonlinear parameter estimation in Chapter 1, Chapter 2 provides a detailed discussion of the characteristics of the orbit for Earth-orbiting satellites. This is the primary dynamical system of interest in this treatment. Chapter 3 describes the various observations used in the Earth-orbiting satellite orbit determination problem. Chapter 4 describes the properties of the Weighted Least Squares and Minimum Variance Estimation algorithms as they are expressed in a batch estimation form and describes the transformation to the sequential estimation algorithm. The relation to the Kalman filter is noted. In Chapter 5, the computation algorithms for solving the linear system based on both the Cholesky Decomposition and the Orthogonal Transformation approach are presented. In addition, square root sequential estimation algorithms are given. Chapter 6 describes the error analysis approaches that evolve from the Consider covariance analysis. The algorithms developed here account for the effects of uncertainties on unestimated parameters that appear in either the dynamic or measurement equations.

As noted previously, a primary emphasis of this treatment will be devoted to developing computationally robust and efficient algorithms for processing the

observations of Earth-orbiting satellites.

1.5 REFERENCES

- Cajori, F., *A History of Mathematics*, MacMillan Co., New York, 1919.
- Fisher, R. A., "On an absolute criteria for fitting frequency curves," *Mess. Math.*, Vol. 41, pp. 155–160, 1912.
- Gauss, K. F., *Theoria Motus Corporum Coelestium*, 1809 (Translated into English: Davis, C. H., *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections*, Dover, New York, 1963).
- Helmert, F. R., "Zur Bestimmung kleiner Flächenstücke des Geoids aus Lothabweichungen mit Rücksicht auf Lothkrümmung", *Sitzungsberichte Preuss. Akad. Wiss.*, Berlin, Germany, 1900.
- Kalman, R. E., "A New Approach to Linear Filtering and Prediction Theory," *J. Basic Eng.*, Vol. 82, Series E, No. 1, pp. 35–45, March 1960.
- Kalman, R. E. and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," *J. Basic Eng.*, Vol. 83, Series D, No. 1, pp. 95–108, March 1961.
- Kolmogorov, A. N., "Interpolation and Extrapolation of Stationary Random Sequences," *Bulletin of the Academy of Sciences of the USSR Math. Series*, Vol. 5, pp. 3–14, 1941.
- Laplace, P. S., *Théorie Analytique de Probabilités*, Paris, 1812 (The 1814 edition included an introduction, *Essai Philosophique sur les Probabilités*, which has been translated into English: Dale, A. I., *Philosophical Essay on Probabilities*, Springer-Verlag, New York, 1995).
- Legendre, A. M., *Nouvelles méthodes pour la détermination des orbites des comètes*, Paris, 1806.
- Markov, A. A., "The law of large numbers and the method of Least Squares," (1898), *Izbr. Trudi., Izd. Akod. Nauk*, USSR, pp. 233–251, 1951.
- Sorenson, H. W., "Least squares estimation: from Gauss to Kalman," *IEEE Spectrum*, Vol. 7, No. 7, pp. 63–68, July 1970.
- Swerling, P., "First order error propagation in a stagewise differential smoothing procedure for satellite observations," *J. Astronaut. Sci.*, Vol. 6, pp. 46–52, 1959.
- Wiener, N., *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*, John Wiley & Sons, Inc., New York, 1949.

1.6 EXERCISES

1. Write a computer program that computes $\rho(t_i)$ for a uniform gravity field using Eq. (1.2.10). A set of initial conditions, X_0 , \dot{X}_0 , Y_0 , \dot{Y}_0 , g , X_s , Y_s , and observations, ρ , follow. With the exception of the station coordinates, the initial conditions have been perturbed so that they will not produce exactly the observations given. Use the Newton iteration scheme of Eq. (1.2.9) to recover the exact initial conditions for these quantities; that is, the values used to generate the observations. Assume that X_s and Y_s are known exactly. Hence, they are not solved for.

Unitless Initial Conditions

$$\begin{aligned} X_0 &= 1.5 \\ Y_0 &= 10.0 \\ \dot{X}_0 &= 2.2 \\ \dot{Y}_0 &= 0.5 \\ g &= 0.3 \\ X_s &= Y_s = 1.0 \end{aligned}$$

Time	Range Observation, ρ
0	7.0
1	8.00390597
2	8.94427191
3	9.801147892
4	10.630145813

Answer

$$\begin{aligned} X_0 &= 1.0 \\ Y_0 &= 8.0 \\ \dot{X}_0 &= 2.0 \\ \dot{Y}_0 &= 1.0 \\ g &= 0.5 \end{aligned}$$

2. In addition to the five state variables of Exercise 1, could X_s and Y_s be solved for given seven independent observations of ρ ? Why or why not? Hint: Notice the relationship between X_0 and X_s and between Y_0 and Y_s in Eq. (1.2.7) for ρ_j . See also Section 4.12 on observability.