Accurate Determination of Comet and Asteroid Orbits Leading to Collision With Earth[3]

CARLOS M. ROITHMAYR
NASA Langley Research Center

LINDA KAY-BUNNELL
Analytical Mechanics Associates Inc.

DANIEL D. MAZANEK
NASA Langley Research Center

RENJITH R. KUMAR AND HANS SEYWALD
Analytical Mechanics Associates Inc.

MATTHEW A. HAUSMAN
Colorado Center for Astrodynamics Research
University of Colorado, Boulder

Introduction

Movements of the celestial bodies in our solar system inspired Isaac Newton to work out his profound laws of gravitation and motion; with one or two notable exceptions, all of those objects move as Newton said they would. But normally harmonious orbital motion is accompanied by the risk of collision, which can be cataclysmic. The Earth's moon is thought to have been produced by such an event, and we recently witnessed magnificent bombardments of Jupiter by several pieces of what was once Comet Shoemaker-Levy 9. Other comets or asteroids may have met the Earth with such violence that dinosaurs and other forms of life became extinct; it is this possibility that causes us to ask how the human species might avoid a similar catastrophe, and the answer requires a thorough understanding of orbital motion.

The two red square flags with black square centers displayed in figure 1 are internationally recognized as a warning of an impending hurricane. Mariners and coastal residents who know the meaning of this symbol and the signs evident in the sky and ocean can act in advance to try to protect lives and property; someone who is unfamiliar with the warning signs or chooses to ignore them is in much greater jeopardy. Although collisions between Earth and large comets or asteroids occur much less frequently than landfall of a hurricane, it is imperative that we learn to identify the harbingers of such collisions by careful examination of an object's path.

An accurate determination of the orbit of a comet or asteroid is necessary in order to know if, when, and where on the Earth's surface a collision will occur. Generally speaking, the longer the warning time, the better the chance of being able to plan and execute action to prevent a collision. The more accurate the determination of an orbit, the less likely such action will be wasted effort or, what is worse, an effort that increases rather than decreases the probability of a collision. Conditions necessary for a collision to occur are discussed, and warning times for long-period comets and near-Earth asteroids are presented.

Orbit determination is the process of using a collection of measurements obtained by observation to calculate a set of orbital elements, six quantities that give (either implicitly or explicitly) the position and

[3]Chapter nomenclature available in chapter notes, p. 217.

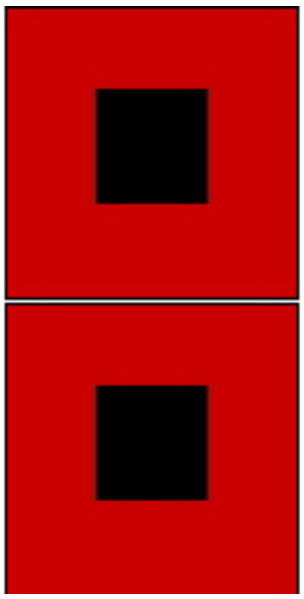

Figure 1. Hurricane Warning Flag.

velocity of an object at a particular instant of time. The classical methods of Laplace, Gauss, and Olbers utilize six angular measurements, obtained two at a time from each of three optical observations, to make a preliminary determination of orbital elements. Additional measurements can be employed together with the iterative method of least squares to improve accuracy of the elements. With all these methods, the quality of the result is affected by errors in the measurements and by the spatial and temporal spacing of observations which are in turn related to the geometry resulting from the orbits of the object and the observatory. The method of least squares makes it possible to take advantage of multiple measurements made from a single observatory or from two or more observatories spread throughout the solar system; the accuracy of orbital elements so obtained is thus a function of the number and spacing of observations, and the number and placement of observatories.

A study of the aforementioned factors and their effects on orbit determination is undertaken in order to identify trends and make a preliminary determination as to the number, placement, and resolution of instruments required to form an effective system for determining orbits of comets and asteroids. In addition, an idea can be obtained of the number and timing of observations that yield the best results, leading to a strategy for using the system to perform observations. In order to keep this stage of the analysis from becoming unnecessarily complicated it is assumed that measurements can be made at will, but of course this cannot always be the case. Orbit determination of long-period comets using a batch filter and sequential filters is examined, and a batch filter is applied to the study of near-Earth asteroids.

Collisions

This analysis is facilitated by the assumption that a collision occurs when the heliocentric distance of a comet or asteroid is identical with that of the Earth; in the case where the object's orbit is not coplanar with the Earth's orbit, the distances must be identical at a point where the object passes through the ecliptic, the plane in which the Earth orbits the Sun. Warning times are studied with the aid of coplanar orbits, and a condition required for a collision in the more general situation of noncoplanar orbits is examined.

Warning Time

The amount of warning before a collision is the interval between the time the object is first discovered (or, to be more precise, the time the orbit becomes known accurately) and the time of collision. Although several orbital periods may elapse during this time, we concern ourselves here with the worst situation in which the comet or asteroid is discovered less than one orbital period before a collision. Unlike large asteroids with relatively short orbital periods, long-period comets (LPCs) do not present themselves for observation over multiple orbits, making it much more difficult to predict a collision decades in advance. Smaller asteroids may escape detection until less than one orbit remains before a collision. Warning time can be obtained through a straightforward application of time-of-flight equations.

Figure 2 shows the Earth E in a circular orbit of radius $r_k = 1$ astronomical unit (au) about the Sun *S*, and a comet (or asteroid) *C* in a coplanar elliptical orbit. The axes s_1 and s_2 lie in the ecliptic, with s_1 in the direction of vernal equinox. It is assumed that discovery of *C* occurs after aphelion and before perihelion, at the point where the red line intersects the orbit of *C*, and E is assumed to be at the wrong place at the wrong time so that E collides with *C* before *C* reaches perihelion, at the point where the blue line intersects the orbits of both objects; the time of flight of *C* between these two points is to be determined. The motion of *S* and *C* is regarded as being governed by two-body orbit mechanics; that is, *S* and *C* are each treated as a particle, or a sphere whose mass is distributed uniformly, and the only forces exerted on *S* and *C* are those of mutual gravitation. The circular orbit of E about *S* also proceeds as two-body motion; however, the gravitational force exerted by E and *C* on each other is left out of account, as are all other perturbing forces acting on E, *C*, and *S*.

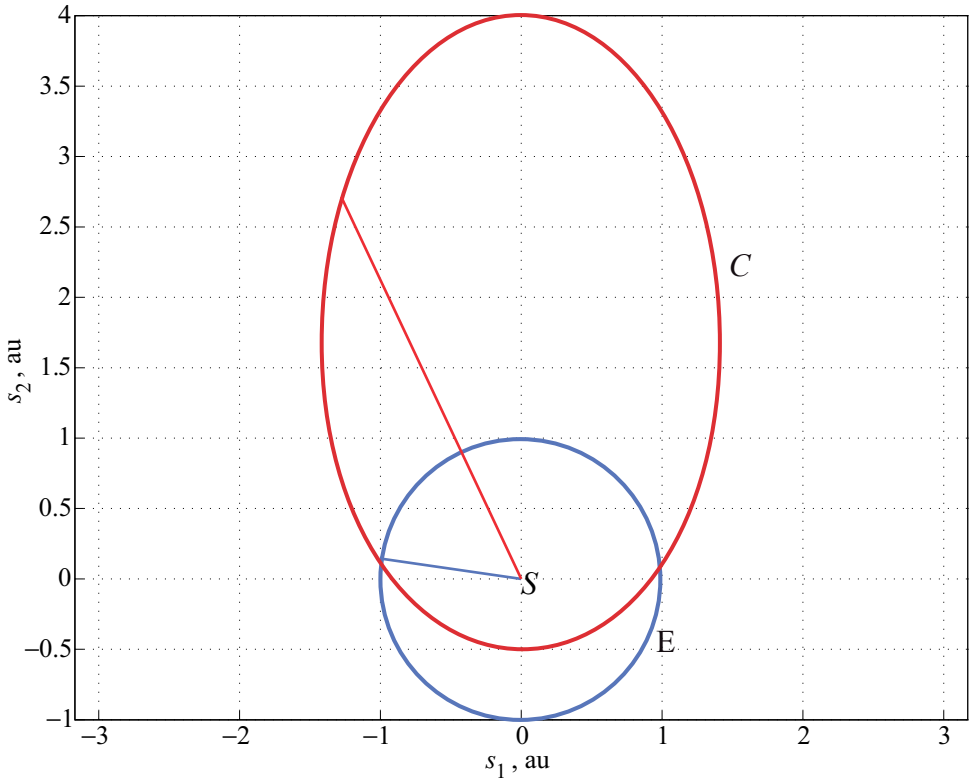

Figure 2. Orbits of *C* and E.

The eccentric anomaly E at any point of an elliptical orbit obeys the relationship in equation (4.2-14) of reference 1:

$$\cos E = \frac{1 - r/a}{e} \tag{1}$$

where a and e are, respectively, the semimajor axis and the eccentricity of the orbit, and r is the distance between the centers of S and C at the point of interest. If E designates the eccentric anomaly at the point of discovery where the heliocentric distance is r, and E_k denotes the eccentric anomaly at the point of collision where the heliocentric distance is r_k = 1 au, then the time of flight for an interval less than one orbital period is obtained via Kepler's equation, as given by equation (4.2-9) of reference 1:

$$t - t_k = \sqrt{\frac{a^3}{\mu}}\left[(E - e\,\sin E) - (E_k - e\,\sin E_k)\right] \tag{2}$$

where μ is the gravitational parameter of the primary, in this case S.

Now, the principal values of the inverse cosine function are $0 \le \cos^{-1} x \le \pi$; E is defined to be 0 at perihelion and π at aphelion. If E and E_k are to have the correct values for the quadrants as illustrated in figure 2, the sign of the right hand member of equation (1) would have to be changed, as would the signs of E, E_k, sin E, and sin E_k in equation (2); however, the absolute value of the time of flight $t - t_k$ would remain unaltered.

Results

Times of flight until collision for LPCs are calculated with the aid of equations (1) and (2). A family of orbits is constructed with perihelia r_p equal to 0.1 and 1 au, and aphelia r_a of 15, 20, 25, …, 100, 200, 300, …, 1000, 2000, …, 50×10^3 au. An aphelion of 15 au corresponds to an orbital period of about 20 years, whereas 50×10^3 au reaches the middle of the Oort cloud and corresponds to a period of 4×10^6 years. (LPCs are often regarded as those with periods greater than 200 years.) The distance r at which the orbit becomes known takes on the values 5, 6, and 7 au; corresponding time of flight, regarded as warning time, is presented in figure 3 as a function of r_a. Warning time does not change appreciably for aphelia in the range $1000 \le r_a \le 50 \times 10^3$ au, and a reduction in r_p by a factor of 10 reduces warning time by approximately the same amount as a reduction of 1 au in r. For the orbits studied, warning times range from 2 years in the case of a 20-year comet with r_p = 1 au, detected at a distance of 7 au from the Sun, to 9.5 months in the case of a comet coming from the Oort cloud with r_p = 0.1 au, detected at 5 au.

Near-Earth asteroids (NEAs) are studied in a similar manner, with r_p equal to 0.2 or 0.9 au, and r_a = 1.4, 1.5, …, 3.0 au; the associated orbital periods range between approximately 9 months, and 2 years and 9 months. The detection distance takes on the values 1.1, 1.2, and 1.3 au, and warning time as a function of r_a is presented in figure 4. Warning times range from approximately 90 days for an asteroid in a 0.9×1.4 au orbit, detected at a heliocentric distance of 1.3 au, to 7 days in the case of an asteroid with a 0.2×3.0 au orbit spotted when it is 1.1 au from the Sun.

Although it has been assumed thus far that the orbits of C and E are coplanar, a collision is possible also when the orbit plane of C is inclined to the ecliptic, so long as an ascending or descending node of C has a heliocentric distance of 1 au. Under these conditions, together with the two-body assumptions set forth in the discussion of warning time, the time of flight is unchanged.

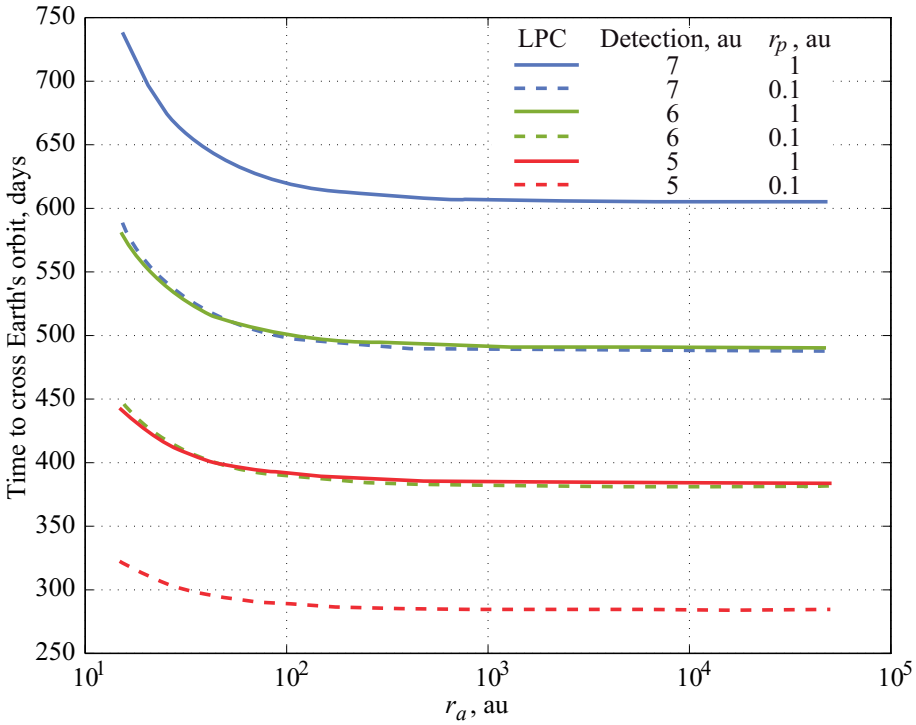

Figure 3. Times of flight to collision, long-period comets.

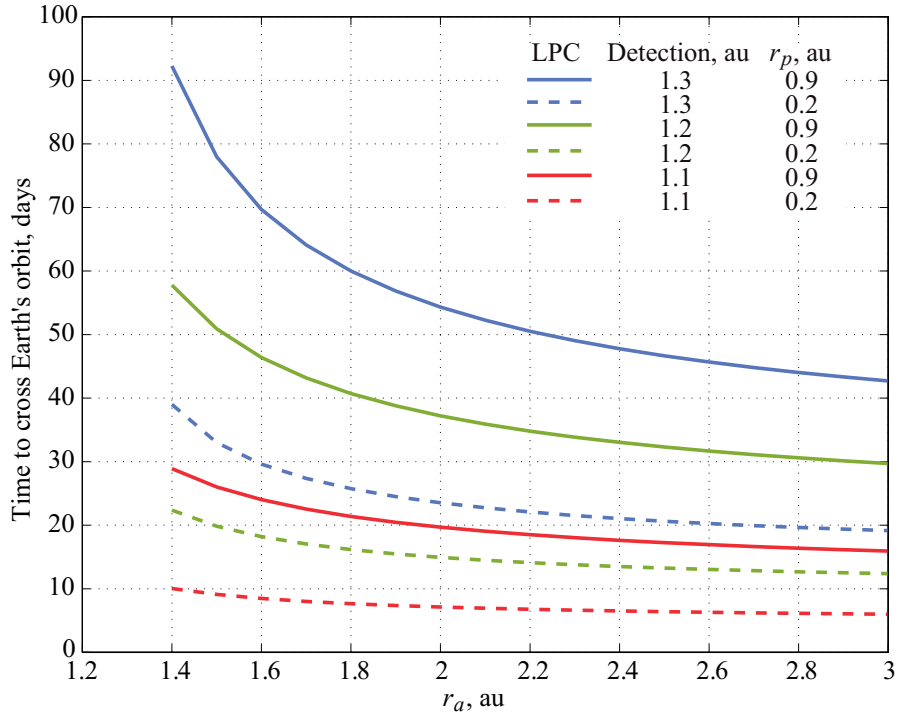

Figure 4. Times of flight to collision, near-Earth asteroids.

Designing a Collision

Forthcoming batch filter analyses rest on the foregoing supposition, namely that a collision with Earth occurs when the object passes through the ecliptic plane and the object's heliocentric distance is equal to r_k, the radius at collision, or 1 au. By design, all of the orbits to be studied meet these conditions; this requires a certain relationship, to be developed presently, between the argument of periapsis on the one hand and, on the other, r_p, r_k, and e.

With the ecliptic chosen as the reference or fundamental plane, an object whose orbit plane is inclined to the ecliptic is, by definition, either at the ascending node or the descending node of the orbit when it passes through the ecliptic. Equations (1.5-4) and (1.5-7) of reference 1 allow us to write

$$r = \frac{r_p(1+e)}{1+e\cos\nu} \tag{3}$$

where r_p is the radius of periapsis, e is the eccentricity of the orbit, and ν is the true anomaly measured in the plane of orbit from the periapsis. At the ascending node $\nu = -\omega$, where ω is known as the argument of periapsis, and at the descending node $\nu = \pi - \omega$. The requirement that $r = r_k$ at one of the nodes therefore can be expressed as

$$r_k = \frac{r_p(1+e)}{1\pm e\cos\omega} \tag{4}$$

where the positive sign preceding cos ω means the condition is imposed at the ascending node, and the negative sign is associated with the descending node. This relationship can be rearranged,

$$\pm\cos\omega = \frac{r_p}{r_k}\left(1+\frac{1}{e}\right) - \frac{1}{e} \tag{5}$$

Now, any member of equation (5) must of course remain between -1 and 1 (inclusive) if it is to be solved for ω,

$$-1 \le \pm\cos\omega \le 1 \tag{6}$$

however, one can simply work with the positive sign because it can be seen that the requirement is the same no matter which sign is used. Substitution from equation (5) into (6) gives

$$-1 \le \frac{r_p}{r_k}\left(1+\frac{1}{e}\right) - \frac{1}{e} \le 1 \tag{7}$$

or

$$\frac{1-e}{1+e} \le \frac{r_p}{r_k} \le 1 \tag{8}$$

and the right hand inequality yields the expected restriction on r_p, namely

$$r_p \le r_k \tag{9}$$

One can deal with the left hand inequality in (8) by appealing to equation (1.7-4) of reference 1 and substituting $(r_a - r_p)/(r_a + r_p)$ for e, which yields a rather sensible result involving the radius at apoapsis,

$$r_k \leq r_a \tag{10}$$

Thus, as long as the inequalities (9) and (10) are satisfied, one is able to solve equation (5) for a value of argument of perihelion that meets the condition necessary for a collision to take place at the ascending or descending node, according to the choice of sign. (Changing the sign of the argument of the function $\cos^{-1}$ yields the supplementary angle: $\cos^{-1}(x) + \cos^{-1}(-x) = \pi$.) The analyst is presented with a second choice of sign because $\cos(\omega) = \cos(-\omega)$; the result of $\cos^{-1}[\cos(\omega)]$ provided by a calculating machine typically lies in the range of principal values of the inverse function, namely $0 \leq \omega \leq \pi$, but a solution on the interval $-\pi \leq \omega \leq 0$ is also correct. Because ω is always measured from the ascending node, $\omega > 0$ always corresponds to a periapsis that is to the north side of the reference plane, and $\omega < 0$ always implies the periapsis is on the south side. Each of the four possible combinations of the two choices is associated with a pre- or postperihelion collision as indicated in table 1.

Table 1. Position of Collision

Node of collision	Perihelial hemisphere	
	North ($\omega > 0$)	South ($\omega < 0$)
Ascending	Preperihelion	Postperihelion
Descending	Postperihelion	Preperihelion

Batch Filter for Long-Period Comets

As mentioned in the introduction, the design of a system for performing observations can be guided by a quantitative study of the way in which orbit determination is influenced by the quality, number, and timing of the observations, as well as the number and location of the observatories. Because the orbital parameters of LPCs are distinct from those of asteroids, the investigation is divided accordingly with comets examined in the present section and asteroids in the following section. It would be highly desirable if a single system could serve effectively in observing the two classes of objects; however, it remains to be seen if this is in fact possible.

In order to gauge the effects of the aforementioned factors on orbit determination, it is necessary to have in hand a way to judge the quality of orbital parameters obtained from observations. To this end the concept referred to as "erroneous predicted miss distance" is introduced.

The study of LPCs begins with an introduction to the method of weighted least squares. The requirement for collision expressed in equation (5) is used to construct a large number of orbital parameter sets, which then yield "true" values of measurements to which errors are introduced intentionally. The measurements thus corrupted account for limits in resolution that can be furnished by an actual telescope and are used to determine or estimate the orbit and obtain an associated erroneous predicted miss distance. The quality of preliminary orbit determination obtained from three observations is examined, followed by a study of the effects of multiple (more than three) observations and observations taken from two or more observatories.

Erroneous Predicted Miss Distance

Orbital elements determined on the basis of observations generally differ from the true orbital elements, that is to say at a particular value of time t, the true position $\mathbf{r}$ and velocity $\mathbf{v}$ differ from the position $\mathbf{r}'$ and velocity $\mathbf{v}'$ resulting from the process of orbit determination. In practice the true orbital elements are unknown; however, they are to be specified in analysis that follows. The two sets of orbital elements can be compared in order to judge the quality of the result of orbit determination; however, it is more convenient to compare a single parameter if at all possible instead of six scalar values associated with position and velocity (or, for that matter, six classical orbital elements). Such a parameter ε, hereafter referred to as an "erroneous predicted miss distance," is now introduced. When ε vanishes, the orbit determined from observations results in a collision at the specified position and at the designated time. (Other metrics can be found in the literature, but seem to require a comparison of two scalars.)

An object C travels along a true orbit, shown in figure 5 with a solid red arc, designed to collide with the Earth E at time t_k, at a point K where the object passes through the ecliptic plane at a heliocentric distance of $r_k = 1$ au. The orbit of C is generally not identical to an orbit determined from measurements, shown with a dashed red arc, and associated with an object C'. An orbit solution is obtained for some epoch t, at which time the true position vector from S to C is $\mathbf{r}(t)$, and the position determined from observation is shown as $\mathbf{r}'(t)$. When the true orbit is propagated to the time of collision t_k, $\mathbf{r}(t_k)$ is of course the position vector from S to K. The orbit of C' can also be propagated to t_k, yielding a position vector $\mathbf{r}'(t_k)$; the magnitude ε of the difference $\mathbf{r}'(t_k) - \mathbf{r}(t_k)$ is defined to be the "erroneous predicted miss distance."

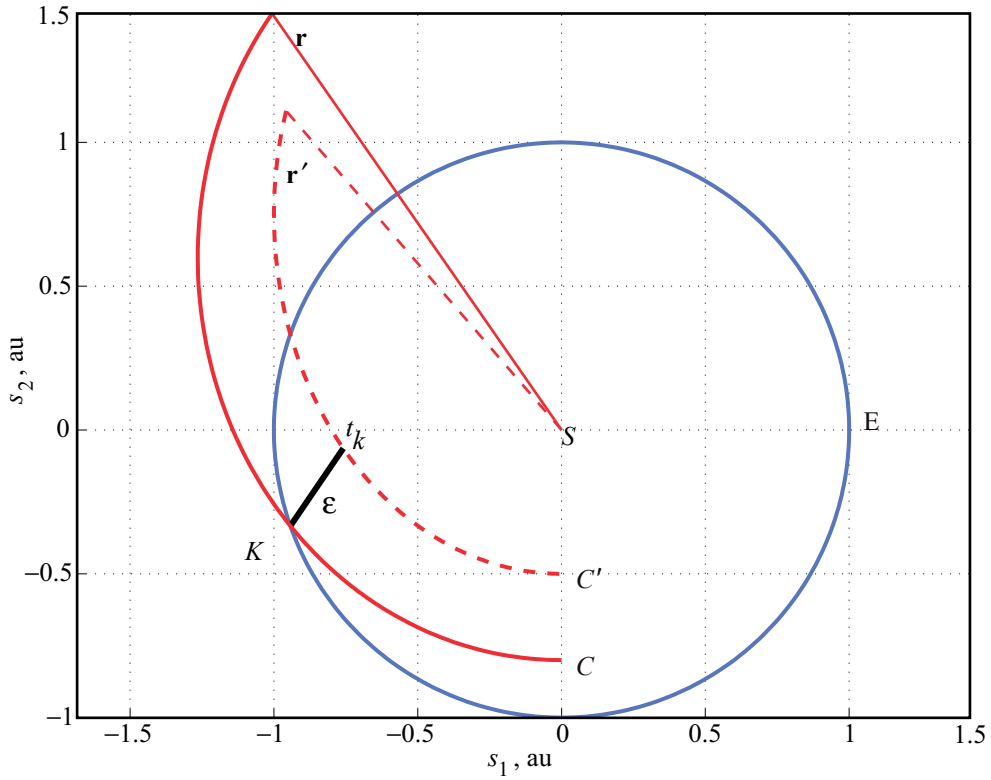

Figure 5. True orbit and orbit determined from observations.

There is one drawback to using ε to judge the quality of an orbit solution: ε may vanish even though $\mathbf{r}'(t) \neq \mathbf{r}(t)$ and $\mathbf{v}'(t) \neq \mathbf{v}(t)$. A solution to the Gauss problem, described in section 5.2 of reference 1, yields the velocities $\mathbf{v}(t_1)$ and $\mathbf{v}(t_2)$ at two specified positions $\mathbf{r}(t_1)$ and $\mathbf{r}(t_2)$, where the time of flight $t_2 - t_1$ between the two positions is also specified, as is the "direction of motion" (i.e., whether the orbit proceeds from $\mathbf{r}(t_1)$ to $\mathbf{r}(t_2)$ via an angular displacement less than or greater than π). Consequently, for any position solution $\mathbf{r}'(t)$ and time of flight $t_k - t$, there exists some velocity $\mathbf{v}'(t)$ that results in C' passing through K at precisely t_k, thus $\varepsilon = 0$ even though the orbit of C' differs from that of C. If by unlucky coincidence an orbit determination algorithm happened to settle on just such a combination of $\mathbf{r}'(t)$ and $\mathbf{v}'(t)$, ε would vanish, giving a false impression that the orbit solution was identical to the true orbit. The effects of such coincidences will be diminished by introducing errors in the measurements in a random fashion, and averaging ε over a large number of measurement sets.

For each orbit examined in the sequel, ε is obtained by the following steps:

1. A set of classical orbital elements is constructed in accordance with equation (5) and used together with orbit propagation software to produce time histories of **r** and **v** for an interval of time over which optical measurements are to be taken.

2. A time of flight to collision $t_k - t$ is determined according to equation (2), whereupon Lagrange coefficients F and G (p. 179, ref. 2) are computed and used to obtain the position of collision $\mathbf{r}(t_k)$.

3. Several values of time t_i (i = 1, 2, 3, …) are selected for obtaining observations. For each such value a position of an observatory is constructed and subtracted from $\mathbf{r}(t_i)$ to yield the position vector from the observatory to the object of interest; this vector is in turn used to calculate a longitude and latitude of the object with respect to a set of inertially fixed heliocentric-ecliptic axes.

4. For each observation, errors are intentionally introduced in the longitude and latitude to reflect the limits in angular resolution of a telescope; the resulting angles serve as measurements.

5. A guess is made as to position and velocity at the time t of one of the observations. The object of the process of orbit determination is to adjust the guess and obtain $\mathbf{r}'(t)$ and $\mathbf{v}'(t)$ that are as consistent as possible with the measurements.

6. The time of flight obtained in step 2 is used together with $\mathbf{r}'(t)$ and $\mathbf{v}'(t)$ to obtain $\mathbf{r}'(t_k)$.

7. The erroneous predicted miss distance ε is the magnitude of $\boldsymbol{\varepsilon} = \mathbf{r}'(t_k) - \mathbf{r}(t_k)$.

Method of Weighted Least Squares

Precision orbit determination can be performed with the linear method of weighted least squares formulated by Carl Friedrich Gauss in 1809, sketched out in references 3 and 4, and explained in detail in reference 5. We have used this method to obtain preliminary determination of orbits using three or four observations as described subsequently, as well as improved orbit determination from many observations and multiple observatories.

The method of weighted least squares can be described briefly as follows: Let $X_1, \ldots, X_6$ be six orbital elements, for example, six scalars associated at a particular instant of time with position $\mathbf{r}(t)$ and velocity

$\mathbf{v}(t)$, to be determined from measurements of longitude $\tilde{\phi}(t_1)$, …,$\tilde{\phi}(t_n)$ and latitude $\tilde{\lambda}(t_1)$, …,$\tilde{\lambda}(t_n)$ obtained at the times t_1, …,t_n of n observations. Using estimates $\hat{X}_1$, …,$\hat{X}_6$ of the orbital parameters, one may compute corresponding values of longitude $\bar{\phi}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_1\right)$, …,$\bar{\phi}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_n\right)$ and latitude $\bar{\lambda}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_1\right)$, …,$\bar{\lambda}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_n\right)$, and form $2n$ residuals by subtracting the computed values from the measurements,

$$y_j\left(\hat{X}_1,\ \ldots,\hat{X}_6\right)=\begin{cases}\tilde{\phi}(t_j)-\bar{\phi}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_j\right) & (j=1,\ \ldots,n)\\ \tilde{\lambda}(t_{j-n})-\bar{\lambda}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_{j-n}\right) & (j=n+1,\ \ldots,2n)\end{cases} \tag{11}$$

The purpose of the method of weighted least squares is to find the values $\hat{X}_1^*$, …,$\hat{X}_6^*$ that minimize the sum of the squares of the weighted residuals,

$$Q=\sum_{j=1}^{2n} w_j y_j^2 \tag{12}$$

where weights w_1, …,w_{2n} are chosen to give more weight or importance to the measurements obtained with a better resolution; $w_j^{-1/2}$ is the assumed accuracy of the jth measurement. Now, for values of X_1, …,X_6 in the neighborhood of $\hat{X}_1$, …,$\hat{X}_6$, the computed values can be expanded in a Taylor series, for example,

$$\bar{\phi}\left(X_1,\ \ldots,X_6,t_j\right)=\bar{\phi}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_j\right)+\sum_{i=1}^{6}\left(X_i-\hat{X}_i\right)\frac{\partial\bar{\phi}}{\partial X_i}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_j\right)+\ \cdots \qquad (j=1,\ \ldots,n) \tag{13}$$

which can be used to write approximate expressions for residuals

$$\begin{aligned} y_j\left(X_1,\ \ldots,X_6\right)&=\tilde{\phi}(t_j)-\bar{\phi}\left(X_1,\ \ldots,X_6,t_j\right)\\ &\approx y_j\left(\hat{X}_1,\ \ldots,\hat{X}_6\right)-\sum_{i=1}^{6}\left(X_i-\hat{X}_i\right)\frac{\partial\bar{\phi}}{\partial X_i}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_j\right) \qquad (j=1,\ \ldots,n)\end{aligned} \tag{14}$$

If one defines arrays

$$\{y\}\triangleq\left[y_1\left(X_1,\ \ldots,X_6\right)\ y_2\left(X_1,\ \ldots,X_6\right)\ \ldots\ y_{2n}\left(X_1,\ \ldots,X_6\right)\right]^{\mathrm{T}} \tag{15}$$

$$\{\hat{y}\}\triangleq\left[y_1\left(\hat{X}_1,\ \ldots,\hat{X}_6\right)\ y_2\left(\hat{X}_1,\ \ldots,\hat{X}_6\right)\ \ldots\ y_{2n}\left(\hat{X}_1,\ \ldots,\hat{X}_6\right)\right]^{\mathrm{T}} \tag{16}$$

$$\{x\}\triangleq\left[\left(X_1-\hat{X}_1\right)\ \left(X_2-\hat{X}_2\right)\ \ldots\ \left(X_6-\hat{X}_6\right)\right]^{\mathrm{T}} \tag{17}$$

and matrices [H] and [W], whose elements are defined as

$$H_{ji} = \begin{cases} \dfrac{\partial \bar{\phi}}{\partial X_i}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_j\right) & (j=1,\ \ldots,n;\ i=1,\ \ldots,6) \\ \dfrac{\partial \bar{\lambda}}{\partial X_i}\left(\hat{X}_1,\ \ldots,\hat{X}_6,t_{j-n}\right) & (j=n+1,\ \ldots,2n;\ i=1,\ \ldots,6) \end{cases} \tag{18}$$

$$W_{ji} = \begin{cases} w_j & (i=j;\ j=1,\ \ldots,2n) \\ 0 & (i\neq j;\ i=1,\ \ldots,2n) \end{cases} \tag{19}$$

then Q can be expressed as

$$Q = \sum_{j=1}^{2n} w_j y_j^2 = \{y\}^{\mathrm{T}}[W]\{y\} \approx \left(\{\hat{y}\}-[H]\{x\}\right)^{\mathrm{T}}[W]\left(\{\hat{y}\}-[H]\{x\}\right)$$

$$= \{\hat{y}\}^{\mathrm{T}}[W]\{\hat{y}\}-\{\hat{y}\}^{\mathrm{T}}[W][H]\{x\}-\{x\}^{\mathrm{T}}[H]^{\mathrm{T}}[W]\{\hat{y}\}+\{x\}^{\mathrm{T}}[H]^{\mathrm{T}}[W][H]\{x\} \tag{20}$$

The value of $\{x\}$ that minimizes Q is obtained by setting $\partial Q/\partial\{x\} = \{0\}$, a 6×1 array of zeros, which yields

$$\{x\} = \left([H]^{\mathrm{T}}[W][H]\right)^{-1}[H]^{\mathrm{T}}[W]\{\hat{y}\} \tag{21}$$

as long as $([H]^{\mathrm{T}}[W][H])$ is nonsingular; this quantity is known as the normal or information matrix and is equal to the inverse of the covariance matrix describing the accuracy of $\{x\}$. When the normal matrix is singular, the orbit cannot be determined uniquely from the measurements and the orbit is said to be unobservable (ref. 5, sec. 4.12). A new estimate of the orbital parameters is formed by adding the adjustment $\{x\}$ to the old estimate $\{\hat{x}\} = [\hat{x}_1\ \ \hat{x}_2\ \ \hat{x}_3\ \ \hat{x}_4\ \ \hat{x}_5\ \ \hat{x}_6]^{\mathrm{T}}$, and the process is repeated, leading to the values of $\hat{x}_1^*,\ \ldots,\hat{x}_6^*$ that minimize Q.

As discussed in reference 5, each row of the $2n \times 6$ mapping matrix [H] can be regarded as the product $[\tilde{H}][\Phi(t_j,t)]$, where $[\tilde{H}]$ is a 1×6 matrix of partial derivatives of a computed value ($\bar{\phi}$ or $\bar{\lambda}$) with respect to $X_1,\ \ldots,X_6$, and $[\Phi(t_j,t)]$ is the 6×6 state transition matrix for the time of the measurement t_j and the time t at which the orbit is to be determined. Both matrices are evaluated with the values $\hat{X}_1,\ \ldots,\hat{X}_6$. Although the rows of [$H$] are often computed with these products, it is also possible to form [H] numerically and this is one option available with the `MATLAB®` function `LSQNONLIN`. With this software, a search direction can be determined with the Levenberg-Marquardt method, a hybrid of the Gauss-Newton approach and the method of steepest descent (ref. 6, pp. 2-16 to 2-21). Parameter options that we set specifically for orbit determination include the minimum change in variables for finite differencing, `DiffMinChange` = 5×10^{-4}, the termination tolerance on the function value, `TolFun` = 1×10^{-10}, and `LargeScale` is set to `off` to select the Levenberg-Marquardt method rather than a large-scale optimization algorithm. The elements of $\{\hat{x}\}$ are scaled so that the first three are in units of au and the last three are in units of km/s. The residuals are rectified so that $-\pi \leq y_j \leq \pi$ and are then multiplied by 1×10^6 to avoid difficulties when the residuals are on the scale of the numerical precision of the computer, and where round off errors can cause a serious loss of accuracy.

Preliminary Orbit Determination

The method of least squares has been used together with the seven-step procedure described in connection with the erroneous predicted miss distance, and MATLAB® software, to determine how the accuracy of preliminary orbit determination based on three observations is affected by the time interval between successive observations, by the distance r at which the object is detected, and by the resolution of the telescope.

Proceeding with the first step, a set of classical orbital elements is specified as $r_a = 10$ au, $r_p = 0.5$ au, $i = 16°$, and $\Omega = 36.5°$, where i is the inclination of the object's orbital plane to the ecliptic, and Ω is the longitude of ascending node measured from the vernal equinox. The argument of perihelion ω is computed according to equation (5) to be 93.017°. Observations are to be performed when the object's heliocentric distance r is in the neighborhood of some specified value, so an initial value of true anomaly $\nu(t_0)$ is obtained by solving equation (3) for ν, chosen ≤ 0 so that the object is traveling toward perihelion. With this set of orbital elements, time histories of $\mathbf{r}$ and $\mathbf{v}$ are recorded over an interval of 98 days.

Continuing with step 3, the observatory is assumed to travel in the ecliptic plane in a circular orbit of radius 1 au; however, *the observatory is not necessarily coincident with the Earth*. At the beginning of the 98-day observation period, the observatory has an initial true longitude $L_0 = 180°$, measured in the ecliptic from the direction of vernal equinox. The times of the three observations are assumed to have equal spacing, $(t_3 - t_2) = (t_2 - t_1)$, ranging from 5 to 49 days. (For example, measurements obtained at the beginning of the 1st, 50th, and 99th days have an interval of 49 days between successive observations, and the total span of the data arc is 98 days.)

Figure 6 illustrates the comet's orbit with a blue curve, and the red portion of the orbit indicates the interval over which observations are made, beginning at $\nu(t_0) = -159.7°$. The green plane represents the ecliptic and the asterisk denotes the Sun. The three black points indicate the positions of the observatory at the beginning, middle, and end of the observation period, where the beginning is marked by the leftmost point.

As described in step 4, each of the six measurements is regarded as the sum of a true angle and an error, which is now assumed to be no greater than the telescope resolution ρ. The MATLAB® function RAND is employed to produce pseudorandom errors uniformly distributed between the limits of the telescope's resolution, $\pm\rho$. In a Monte Carlo approach, an orbit is determined 100 times using a different random number seed in each trial, and an average erroneous predicted miss distance is reported. The assumed distribution of the error is not actually an important consideration when using the method of least squares since this method does not account for such statistics. A uniform distribution is simple to generate and just as valid as any other in this case.

Figure 7 shows, on logarithmic scale of base 10, the average value of ε as a function of the interval between observations, $(t_3 - t_2)$ or $(t_2 - t_1)$, where the time of the second observation t_2 is taken to be the 50th day of the 98-day data arc. The heliocentric distance r at t_2 is varied between 5, 6, and 7 au by using initial values $\nu(t_0) = -155.3°$, $-159.7°$, and $-163.7°$, respectively. Each point on a line represents the average value $\bar{\varepsilon}$ of ε in units of lunar distance from Earth (384 400 km), from a set of 100 measurement errors formed with $\rho = 0.2$ seconds of arc (arcsec). Accuracy of orbit determination becomes better as $r(t_2)$ decreases; however, the most dramatic improvements are obtained by increasing the observation interval to 49 days, at which point all three values of $\bar{\varepsilon}$ are less than 2 lunar distances. Using the laws of logarithms, it is easily shown that a relationship of the form $\bar{\varepsilon} = \bar{\varepsilon}_0 (t_2 - t_1)^m$, where $\bar{\varepsilon}_0$ is a constant,

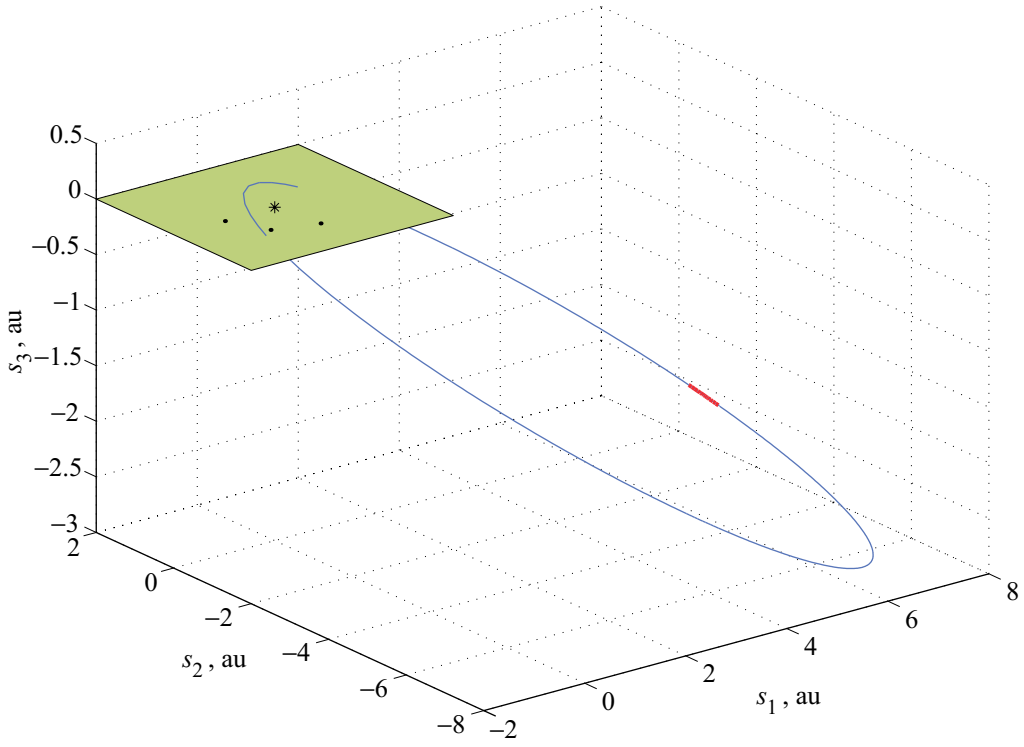

Figure 6. Observation arc.

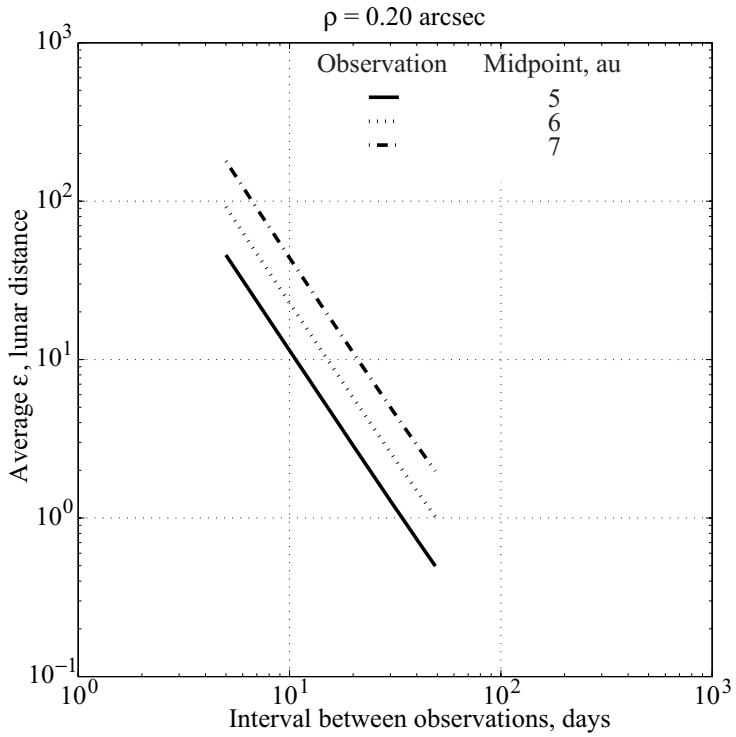

Figure 7. Average ε for various detection distances.

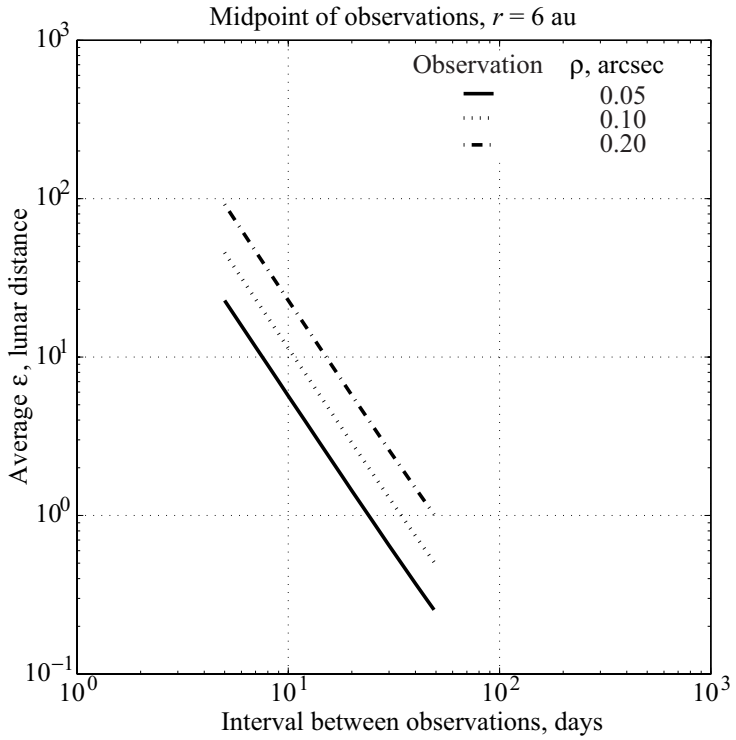

Figure 8. Average ε for various angular resolutions.

becomes $\log_{10}\bar{\varepsilon} = m\,\log_{10}(t_2 - t_1) + \log_{10}\bar{\varepsilon}_0$, which is the equation of a line with slope m. In each line of figure 7, $\bar{\varepsilon}$ decreases by two powers of 10 for every power of 10 increase in $(t_2 - t_1)$; therefore $m = -2$, meaning the accuracy of the preliminary orbit improves with the square of the observation interval.

Figure 8 displays similar information, with $r(t_2)$ held fixed at 6 au and ρ = 0.05, 0.10, and 0.20 arcsec. Telescope resolution is seen to have an effect on preliminary orbit determination accuracy, but the length of the observation interval is again the most significant factor. With an observation interval of 49 days, all values of $\bar{\varepsilon}$ are below 1 lunar distance. Unfortunately, longer observation intervals yield shorter warning times.

In the same vein, an extensive analysis of preliminary determination of LPC orbits is carried out with the following additional details.

In connection with the first step, 1008 sets of classical orbital elements are constructed with r_a = 10, 40, 70, and 100 au, r_p = 0.1, 0.4, 0.7, and 1.0 au, i = 10°, 30°, 50°, …, 170°, and Ω = 0°, 30°, 60°, …, 180°. Time of flight is nearly constant with respect to r_a for r_a > 100 au (see fig. 3); therefore, such orbits are not considered. For each set, the argument of perihelion ω is computed according to equation (5). A 98-day data arc begins with $r(t_0)$ = 6.5 au, and $\nu(t_0)$ is chosen accordingly.

As before, the observatory is assumed to travel in the ecliptic plane in a circular orbit of radius 1 au, and it is worth remembering that *the observatory is not necessarily coincident with Earth*. The effect of observatory position on orbit determination is studied by allowing the initial true longitude L_0 of the observatory to take on four values, each differing by 90°. Thus, a total of 4 × 1008 = 4032 cases are examined. The times of the three observations are equally spaced, with 33 days between subsequent observations.

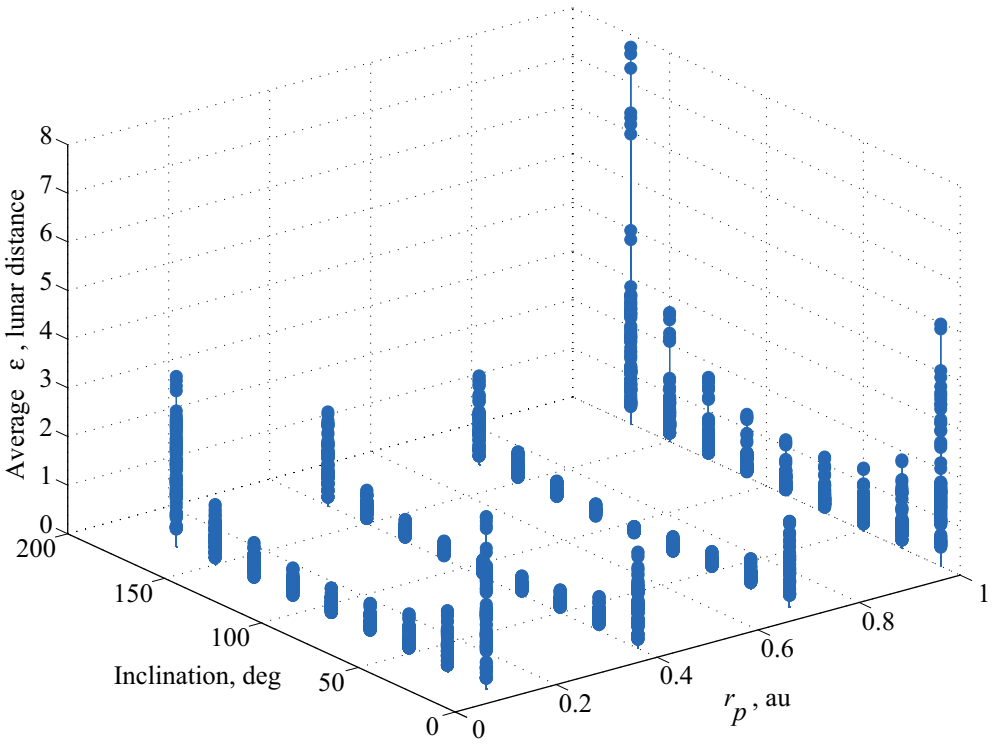

Figure 9. Accuracy of comet orbits, as a function of i and r_p.

For step 4, a telescope resolution of $\rho = 0.1$ arcsec is used to form 100 sets of measurement errors for each orbit to be examined, and the average value of erroneous predicted miss distance is recorded.

The results are displayed in figure 9 as functions of i and r_p; each data point represents $\bar{\varepsilon}$ for one case, and a family of 112 cases (in which r_a, Ω, and L_0 vary) form one stack of points. The best orbit determination is obtained when the orbital planes of the observatory and comet are perpendicular, whereas nearly coplanar orbits yield the poorest determinations. This relationship to inclination stems from the fact that, with three observations, an orbit is unobservable when it is coplanar with the orbit of the observatory. One may form analytic expressions for the elements of the matrix $[\tilde{H}]$ discussed in connection with the method of weighted least squares which are partial derivatives of longitude and latitude. In addition, one may determine which elements of the state transition matrix $[\Phi]$ are zero when two-body mechanics are assumed; a reference orbit coplanar with that of the observatory is used and $[\Phi]$ is written according to equations (9.84)–(9.87) of reference 2. It can then be shown that for three observations, the 6×6 matrix $[H]$ is singular, but for four observations the 6×6 matrix $[H]^{\mathrm{T}}$ $[H]$ is not singular. A study of figure 9 shows that with $r_p = 1$ au, retrograde orbits are harder to determine accurately than prograde orbits. Orbit determination accuracy with $r_p = 0.1$ au is noticeably poorer than with $r_p = 0.4$ or 0.7 au, and even poorer with $r_p = 1$ au. In figure 10 the results are shown as functions of r_a and r_p; each stack contains 252 values of $\bar{\varepsilon}$ associated with varied i, Ω, and L_0. Orbit determination is the most difficult when r_a has the smallest value, 10 au.

Results are not presented as functions of Ω, ω, and $\nu(t_0)$ because there is limited value in doing so. No remarkable relationship appears to exist between $\bar{\varepsilon}$ and Ω. The angle $\nu(t_0)$ simply determines the point of the object's orbit at which observations begin. Finally, argument of perihelion ω is dependent upon r_a and r_p through equation (5).

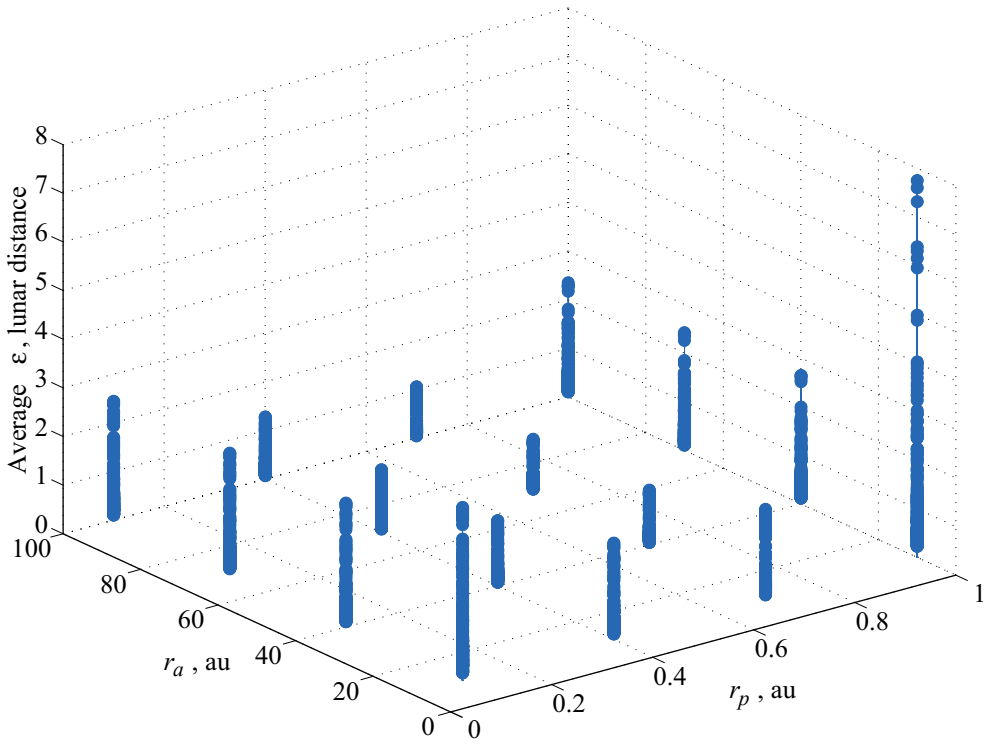

Figure 10. Accuracy of comet orbits, as a function of r_a and r_p.

The 4032 average values of ε are sorted from largest to smallest, and the worst and best cases are identified. The effect of the length of observation interval is then studied for these two cases, with times of the three observations again given equal spacing, ranging from 5 to 49 days between subsequent observations. The worst case, numbered 246, is associated with the orbital parameters r_a = 10 au, r_p = 1 au, i = 170°, Ω = 0°, ω = 0°, and ν(t_0) = −151.68°, and L_0 = 180° for the observatory. Curves for the cases with the same orbital parameters and other values of L_0 are also shown in figure 11. Observatory location affects the accuracy of orbit determination, but not to the same extent as length of the data arc. In all four cases, $\bar{\varepsilon}$ is less than 4 lunar distances with an observation interval of 49 days.

The best case is numbered 3946, with orbital parameters of r_a = 100 au, r_p = 0.7 au, i = 110°, Ω = 120°, ω = 66.69°, and ν(t_0) = −142.86°, and L_0 = 90° for the observatory. Curves for the cases with the same orbital parameters and other values of L_0 are shown in figure 12; long observation intervals once again result in the best preliminary orbit determination. With an observation interval of 49 days, the value of $\bar{\varepsilon}$ associated with each observatory is less than 0.1 lunar distance, or about 6 times the Earth's radius. A slope of m = −2 is evident in figures 11 and 12, confirming the inverse square relationship between $\bar{\varepsilon}$ and observation interval established in figures 7 and 8.

The foregoing results involving preliminary orbit determination corroborate statements made in references 3 and 4, pointing out that the length of the data arc is the single most important factor in determining the accuracy of the orbit solution. The number and precision of the observations, the object's proximity to the observatory when measurements are obtained, and even the use of radar measurements are all secondary to the length of the data arc.

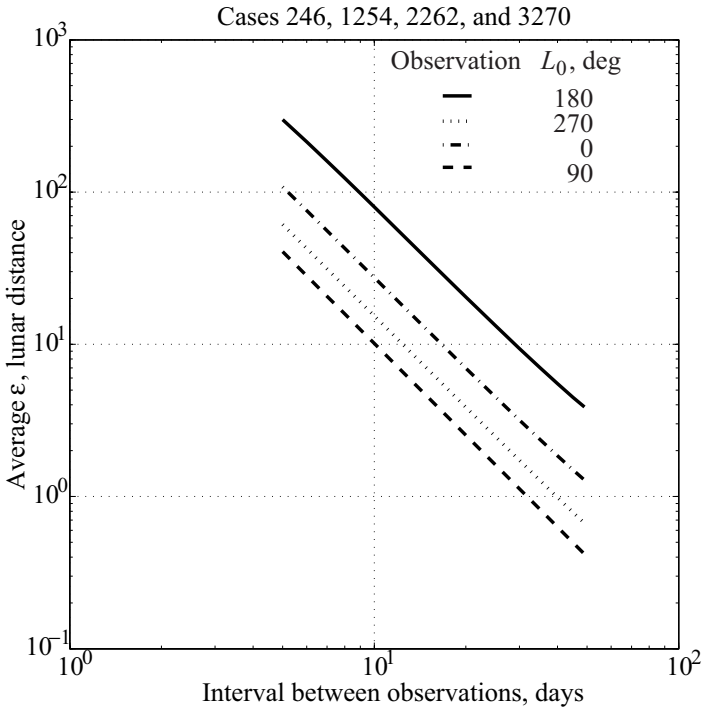

Figure 11. Worst orbit, various observatory longitudes.

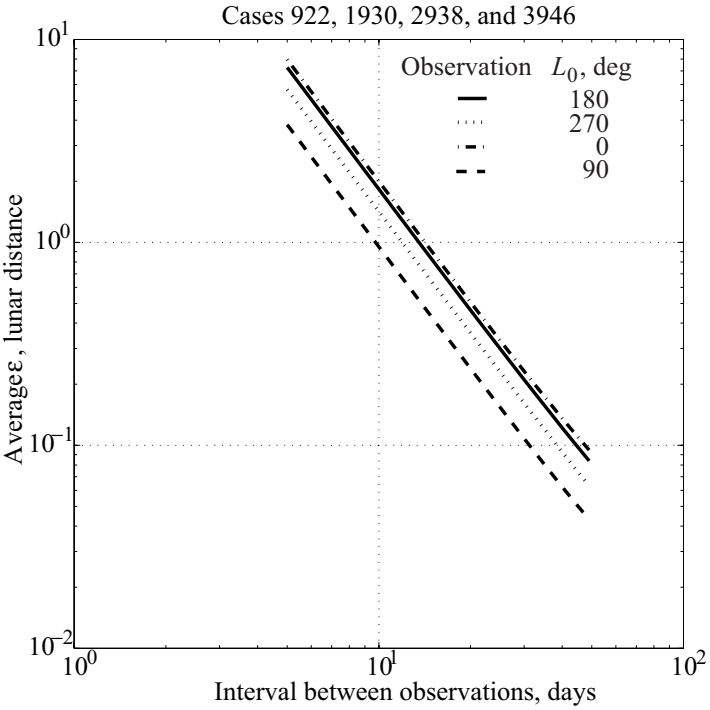

Figure 12. Best orbit, various observatory longitudes.

Multiple Observations

As opposed to the classical methods of Laplace, Gauss, and Olbers, the method of least squares permits the use of a great number of measurements, which can be obtained from one or more observatories. We now proceed to study the effects of multiple observations and observatories on orbit determination of LPCs. A number of observations, ranging from 3 to 99, are taken in equally timed increments over a period of 98 days. We have chosen to obtain an orbit solution for the time of the first observation. Differential correction begins with an estimate $\{\hat{x}\}$ corresponding to an arbitrary day, typically a day after the end of the 98-day observation period.

The forthcoming results, shown in figures 13 through 18, are reported first for a single observatory, and then for multiple observatories.

Single observatory. The orbits identified in the discussion of preliminary orbit determination as resulting in the largest and smallest average value $\bar{\varepsilon}$ of ε, cases 246 and 3946, respectively, are revisited. Figures 13 and 15 show results calculated for various numbers of observations obtained from a single observatory. In each case, a solid curve shows the average erroneous predicted miss distance $\bar{\varepsilon}$ from 100 sets of measurements, each of which is produced with a telescope resolution ρ of 0.1 arcsec.

Results for the worst orbit are contained in figure 13. With 3 observations spaced 49 days apart, $\bar{\varepsilon}$ is nearly 3.8 lunar distances, whereas 99 observations spaced 1 day apart reduce $\bar{\varepsilon}$ to less than 0.22 of a lunar distance, or about 13 Earth radii. In the case of the best orbit, the quality of the result is evidenced by the use of Earth radius (6378 km) as the unit for measuring average erroneous predicted miss distance. The data in figure 15 show that with 3 observations $\bar{\varepsilon}$ is approximately 2.5 Earth radii; 99 observations improve the measure to less than 0.5 Earth radius. In the worst case there is a marked improvement from increasing the number of observations from 3 to 5, followed by a more gradual slope of approximately $m = -1/2$ in going from 10 to 99 observations. In the best case, the slope is also approximately $-1/2$. As discussed previously, this is an indication that the accuracy of the orbit as measured by $\bar{\varepsilon}$ improves as the square root of the number of observations.

As one might expect, better telescope resolution (indicated by a smaller value of ρ) improves orbit determination accuracy. In order to quantify the improvement available with very high resolution measurements (obtained, for example, with interferometry) an analysis is made with 11 observations taken over a 98-day data arc involving a mixture of resolutions; $\rho = 0.01$ arcsec for initial observations, and $\rho = 0.0001$ arcsec for some number of final observations. As described earlier, weights are assigned using the value of ρ with which the corresponding measurements are made, thus giving more weight or importance to the measurements obtained with a better resolution ($w_j = 1/\rho_j^2$). In figures 17 and 18, improvements in $\bar{\varepsilon}$ obtained for the worst orbit (Case 246) and the best orbit (Case 3946), respectively, are shown. The number of final observations made with the better resolution is indicated on the abscissa. A comparison of figures 13 and 17 reveals $\bar{\varepsilon}$ is made better by a factor of 10 when the resolution of 11 observations is improved from 0.1 to 0.01 arcsec. The same conclusion is reached by comparing figures 15 and 18. In each case, a further improvement by a factor of 10 is obtained when the last 4 observations are made with $\rho = 0.0001$ arcsec.

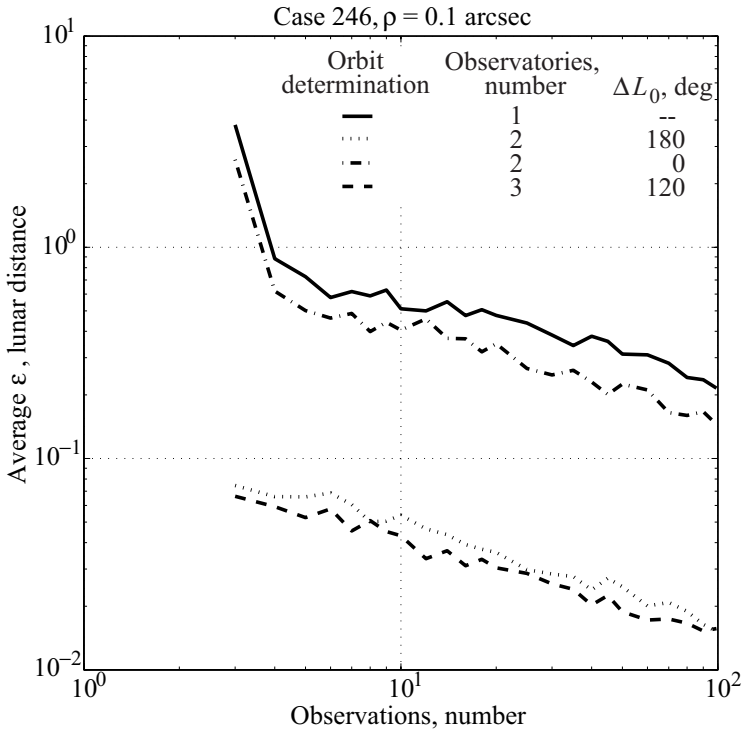

Figure 13. Worst orbit, various observatory configurations.

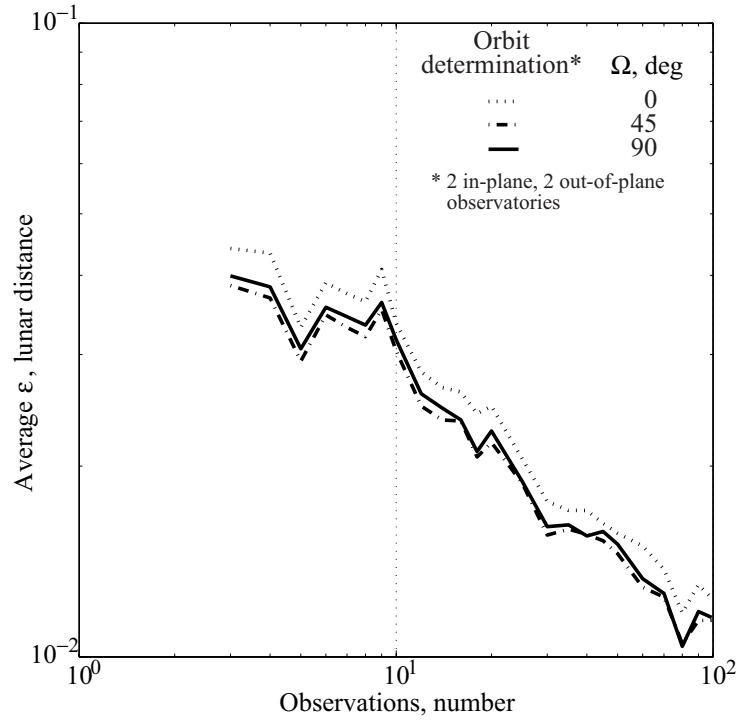

Figure 14. Worst orbit, various observatory configurations (detailed view).

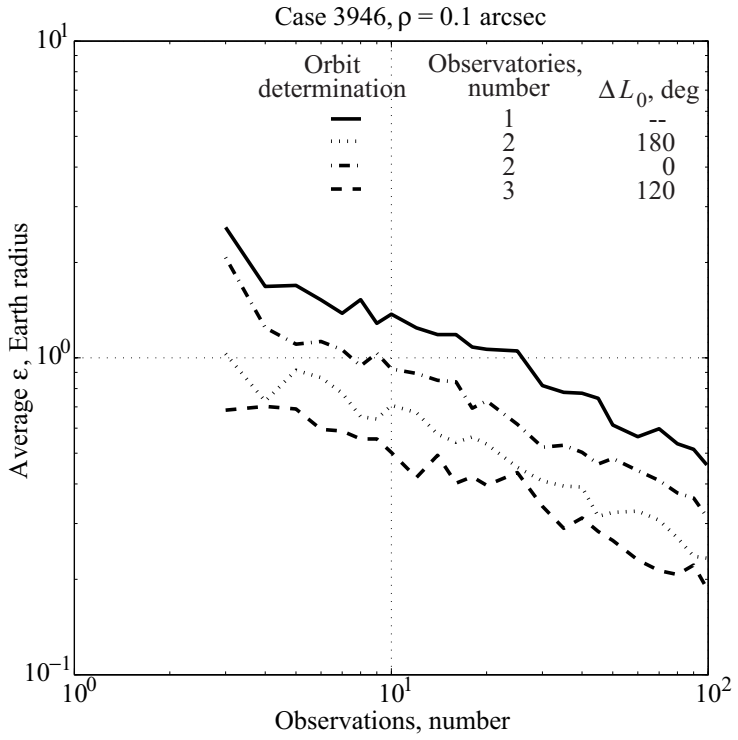

Figure 15. Best orbit, various observatory configurations.

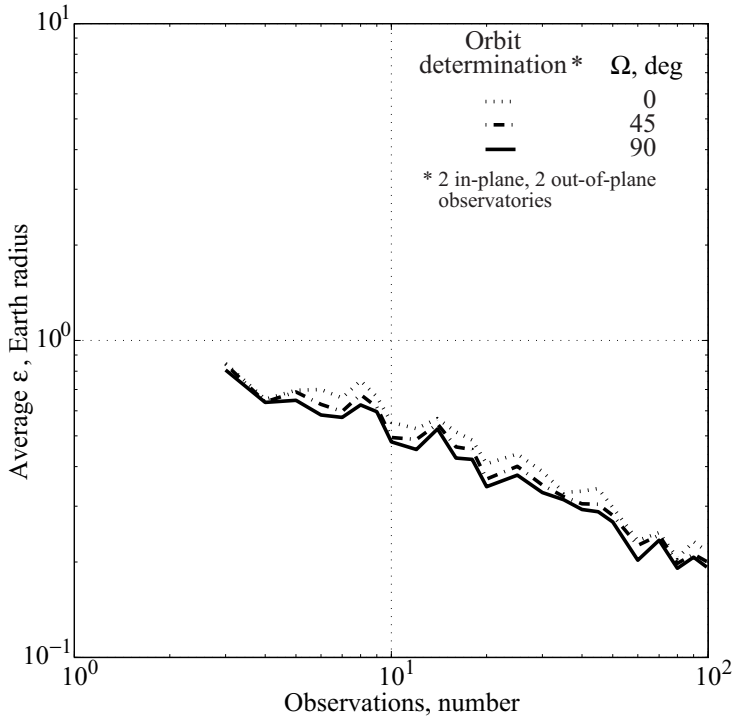

Figure 16. Best orbit, various observatory configurations (detailed view).

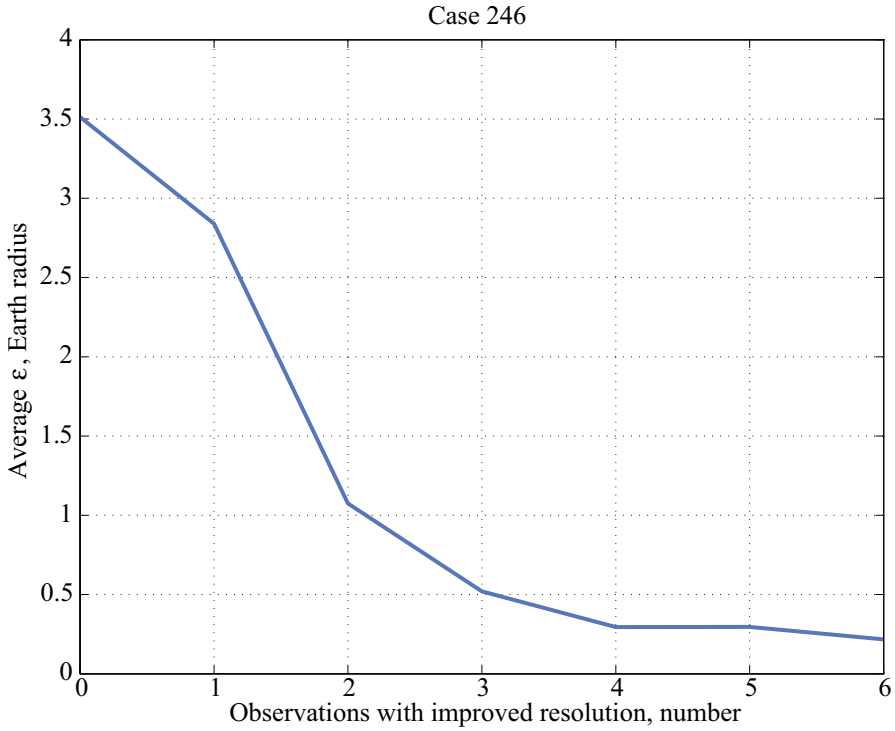

Figure 17. Worst orbit, observations with mixed resolutions.

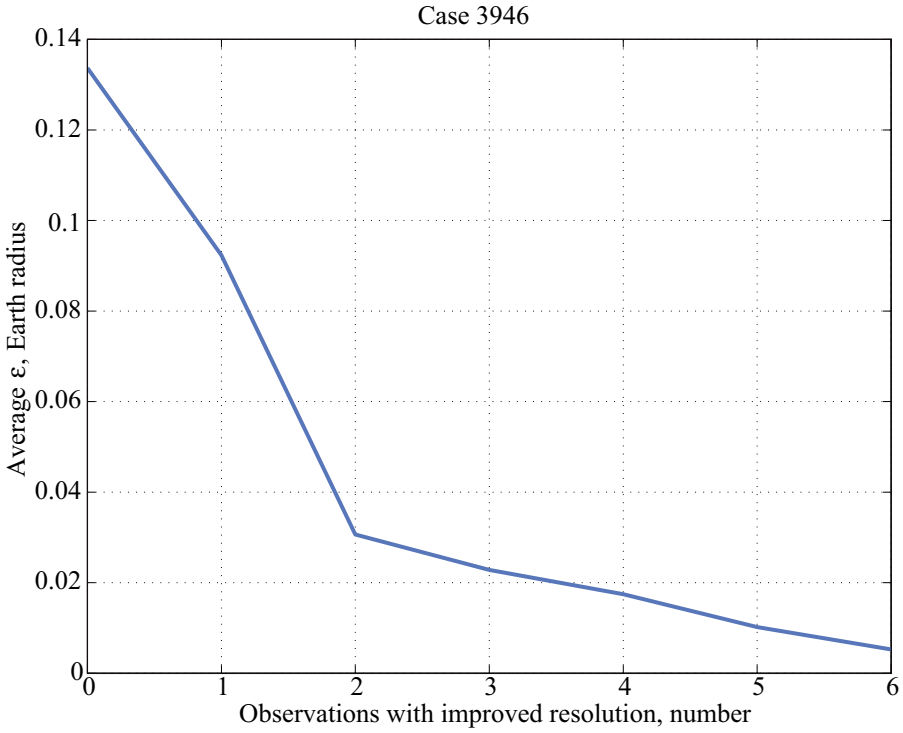

Figure 18. Best orbit, observations with mixed resolutions.

Multiple observatories. As previously mentioned, the method of least squares permits orbit determination to take advantage of measurements not only from additional observations, but also from additional observatories. The benefits of configurations of 2, 3, and 4 observatories are studied with the aid of the worst and best LPC orbits.

The positions of two observatories are constructed such that they have heliocentric circular orbits of radius 1 au in the ecliptic plane, and their true longitudes L_0 are always 180° apart. Three observatories are placed in similar orbits, with their true longitudes phased by 120°. In the case of four observatories, the first two have orbits identical to the configuration of two observatories just described, and the remaining two are in similar coplanar orbits perpendicular to the ecliptic with the ascending node Ω taking on values of 0°, 45°, and 90° in order to determine what effect, if any, Ω has on orbit determination.

In connection with the worst comet orbit, figures 13 and 14 display $\bar{\varepsilon}$ as a function of the number of observations for all observatory configurations described heretofore. It is evident that two or more observatories offer an improvement in $\bar{\varepsilon}$ of nearly a factor of 10 over that from a single observatory. With 99 observations each, two observatories yield $\bar{\varepsilon}$ less than 0.016 lunar distance, slightly less than 1 Earth radius. Three observatories are only marginally better than two observatories. Four observatories are not substantially better than three, and $\bar{\varepsilon}$ is relatively insensitive to the value of Ω for the members of the four-observatory configuration that have orbits normal to the ecliptic. Similarly, figures 15 and 16 show that in the best case substantial improvement in $\bar{\varepsilon}$ is obtained by employing two observatories phased by 180° instead of a single observatory, but the addition of a third or fourth observatory does not appear to be cost effective for this orbit and observatory heliocentric distances of 1 au.

The advantage of two observatories over one may stem from a doubling in the number of observations, or from parallax; in order to gauge the contribution of each, one may form a configuration of two observatories that are coincident with one another, thus eliminating parallax. Results obtained from this configuration are represented with a dash-dot curve in figures 13 and 15. The first of these shows that in the worst case, the improvement obtained from a pair of observatories stems more from parallax than from doubling the number of observations. Figure 15 shows that in the best case, the benefit of two observatories has as much to do with parallax as it does with the availability of twice as many observations.

Of the multiobservatory configurations examined here, two observatories in circular heliocentric orbits of radius 1 au in the ecliptic plane, phased 180° apart, provide the best balance of cost and accuracy in determining the orbits of LPCs. Thirty to forty observations equally spaced over a 98-day interval appear to give results nearly as good as 99 observations taken 1 day apart.

Sequential Filter for Long-Period Comets

The method of weighted least squares, described previously and used to produce all of the results presented earlier, is referred to as a batch processor because it employs observations collected together in a single set; the larger the batch, the more arithmetical operations are required to carry out the matrix multiplications indicated in equation (21). Moreover, when new observations are obtained, they must be added to the previous batch to form an even larger set, and the entire orbit determination procedure must be repeated. Another shortcoming of the method is that there are no provisions for treating the state and observation deviations as random processes and using any associated statistical information to improve the orbit solution.

Remedies for these problems exist in the Kalman filter, a recursive or sequential algorithm in which measurements are used continuously, as they become available, to refine the solution for the orbit. The recursive property of the procedure allows the previous solution to be used as the starting point for the solution to follow. The computational expense of the sequential filter can be less than that of the batch filter because the matrix requiring inversion in the former case can be smaller—when only one measurement is processed at a time, the matrix is reduced to a scalar.

We describe briefly two variations of the Kalman filter, the conventional and the extended algorithms. A method for obtaining the a priori information needed to start the filter is also discussed. Warning time offered by a sequential filter with high resolution optical measurements from a single observatory is compared with that derived from measurements of lower resolution from two observatories. The advantages of measurements of range and range-rate are illustrated. Effects of comet outgassing are considered, followed by a discussion of the construction and use of the body-targeting plane or B-plane in orbit determination analysis. Finally, the calculation of probability and probability of collision are taken up.

The Kalman Filter

The conventional Kalman filter is developed from differential equations for the state, and expressions that relate the measurements to the state, both of which have been linearized about a reference state time history. To reduce the effects of the errors introduced in the process of linearization, the reference state can be replaced after each observation with the improved estimate of the state; this procedure is referred to as the extended Kalman filter. The two varieties of the filter are discussed in turn in what follows.

Conventional Kalman filter. The conventional Kalman filter is described in detail in section 4.7.1 of reference 5 in terms similar to those used earlier in connection with the batch algorithm.

The $n \times 1$ column matrix $\{X\}$ contains n ($n = 6$) orbital elements, dot products of the position $\mathbf{r}(t)$ and velocity $\mathbf{v}(t)$ with three right-handed, mutually perpendicular unit vectors. The state deviation matrix $\{x\}$ is defined as the difference between the true elements $\{X\}$ and the elements $\{X^*\}$ associated with a reference orbit,

$$\{x\} \triangleq \{X\} - \{X^*\} \tag{22}$$

The $m \times 1$ observation deviation, or residual matrix $\{y\}$, is defined as the difference between m measurements $\{Y\}$ and the values $\{Y^*\}$ calculated from the reference orbit,

$$\{y\} \triangleq \{Y\} - \{Y^*\} \tag{23}$$

where the measurements $\{Y\}$ are regarded as the sum of nonlinear functions $\{G(\{X\}, t)\}$ of the state and time, and errors $\{e\}$,

$$\{Y\} = \{G(\{X\}, t)\} + \{e\} \tag{24}$$

Equations for the filter time update are given as equations (4.7.18) and (4.9.50) of reference 5:

$$\{\bar{x}_j\} = [\Phi(t_j, t_{j-1})]\{\hat{x}_{j-1}\} \tag{25}$$

$$[\bar{P}_j] = [\Phi(t_j, t_{j-1})][P_{j-1}][\Phi(t_j, t_{j-1})]^{\mathrm{T}} + [Q_{j-1}] \tag{26}$$

where $\{\hat{x}_{j-1}\}$ and $\{\bar{x}_j\}$ are, respectively, the estimated values of $\{x\}$ following a measurement update at an instant of time t_{j-1}, and prior to a measurement update at t_j. The $n \times n$ state transition matrix $[\Phi(t_j, t_{j-1})]$ that relates the state deviations at times t_{j-1} and t_j is discussed presently. The $n \times n$ estimate error variance-covariance matrices (or simply covariance matrices) of the state deviations $\{\hat{x}_{j-1}\}$ and $\{\bar{x}_j\}$ are denoted by $[P_{j-1}]$ and $[\bar{P}_j]$, respectively. The $n \times n$ matrix Q_{j-1} is the covariance at t_{j-1} of a zero-mean, white sequence known as process noise or state noise.

The $m \times 1$ observation deviation $\{y_j\}$, $m \times n$ observation-state mapping matrix $[\tilde{H}_j]$, and the $n \times m$ Kalman gain $[K_j]$ are assembled in preparation for the measurement update:

$$\{y_j\} = \{Y_j\} - \left\{G\left(\left\{X_j^*\right\}, t_j\right)\right\} \tag{27}$$

$$[\tilde{H}_j] = \frac{\partial\left\{G\left(\left\{X_j^*\right\}, t_j\right)\right\}}{\partial\{X\}} \tag{28}$$

$$[K_j] = [\bar{P}_j][\tilde{H}_j]^{\mathrm{T}}\left([\tilde{H}_j][\bar{P}_j][\tilde{H}_j]^{\mathrm{T}} + [R_j]\right)^{-1} \tag{29}$$

where $\{Y_j\}$ contains the actual measurements at t_j, and where the $m \times m$ matrix $[R_j]$ is the covariance at t_j of $\{e\}$, assumed to have a normal distribution and a mean of zero. Expressions for the elements of $[\tilde{H}_j]$ associated with measurements of longitude and latitude are developed presently, and measurements of range and range-rate are taken up subsequently.

Equations for the filter measurement update are set forth in equations (4.7.16) and (4.7.12) of reference 5,

$$\{\hat{x}_j\} = \{\bar{x}_j\} + [K_j]\left(\{y_j\} - [\tilde{H}_j]\{\bar{x}_j\}\right) \tag{30}$$

$$[P_j] = \left([I] - [K_j][\tilde{H}_j]\right)[\bar{P}_j] \tag{31}$$

where $\{\hat{x}_j\}$ contains the estimated values of the state deviation following a measurement update at time t_j, and $[P_j]$ is the corresponding covariance matrix. After any update performed with measurements obtained at t_j, the best estimate $\{X'_j\}$ of the orbital elements can be formed by adding $\{\hat{x}_j\}$ to the elements $\{X_j^*\}$ associated with the reference orbit at that time, $\{X'_j\} = \{\hat{x}_j\} + \{X_j^*\}$.

In the forthcoming discussion of Gauss's method we take up the matter of obtaining the a priori information needed to start the filtering process, namely the initial values $\{X_0^*\}$ of the elements of the reference orbit, the initial state deviation estimate $\{\hat{x}_0\}$, and a consistent initial covariance matrix $[P_0]$.

The state transition matrix $[\Phi(t_j,t_{j-1})]$ required in equations (25) and (26) for the time update is formed on the assumption that motion of the Sun and the object is governed by two-body gravitational mechanics. The position vector $\mathbf{r}^*_{j-1}$ at time t_{j-1} from the mass center of the Sun to the mass center of the object traveling on the reference orbit, and the inertial time derivative $\mathbf{v}^*_{j-1}$ of that position vector, are related to the relative position $\mathbf{r}^*_j$ and velocity $\mathbf{v}^*_j$ at another time t_j by equations (9.68) of reference 2,

$$\mathbf{r}^*_j = F\,\mathbf{r}^*_{j-1} + G\,\mathbf{v}^*_{j-1} \tag{32}$$

$$\mathbf{v}^*_j = F_t\mathbf{r}^*_{j-1} + G_t\mathbf{v}^*_{j-1} \tag{33}$$

where F, G, F_t, and G_t are known as Lagrangian coefficients, which can be calculated with Battin's universal variables as indicated in equations (9.69) of reference 2. The state transition matrix is partitioned into four 3×3 parts,

$$[\Phi(t_j,t_{j-1})] = \begin{bmatrix} [\tilde{R}(t_j,t_{j-1})] & [R(t_j,t_{j-1})] \\ [\tilde{V}(t_j,t_{j-1})] & [V(t_j,t_{j-1})] \end{bmatrix} \tag{34}$$

where the superscript * is omitted because it is used by Battin to indicate the adjoint of a matrix, but it is to be understood that the partitions are evaluated with $\mathbf{r}^*_{j-1}$, $\mathbf{v}^*_{j-1}$, $\mathbf{r}^*_j$, and $\mathbf{v}^*_j$ of the reference orbit according to equations (9.84)–(9.87) of reference 2. The universal variables U_0, U_1, U_2, and U_3 required to evaluate the Lagrangian coefficients are obtained straightforwardly from the relationships given in problem 4-21 of reference 2. The coefficient C, required to form $[\tilde{V}(t_j,t_{j-1})]$ according to equation (9.87), is expressed in equation (9.74) in terms of U_2, U_4, and U_5. The second of these, U_4, is easily obtained from U_1, U_2, and U_3 with the aid of equation (4.108). The universal variable U_5 can be expressed as in problem 4-30 in terms of U_1 and U_3, except when $\chi = U_1 = 0$, in which case $U_0 = 1$, and $U_2 = U_3 = U_4 = 0$; the variable u also vanishes in view of equation (4.100), which leads to $q = 0$ by way of equation (4.104), and to $U_5(2\chi) = 0$ from equation (4.112). Finally, the first of equations (4.113) yields $U_5(\chi) = U_5(0) = 0$.

In connection with measurements of longitude ϕ and latitude λ, the observation-state mapping matrix $[\tilde{H}_j]$ defined in equation (28) is derived as follows. The position vector from the Sun S to the object C can be written as

$$\mathbf{r} = \mathbf{r}^{SC} = X_1\hat{\mathbf{s}}_1 + X_2\hat{\mathbf{s}}_2 + X_3\hat{\mathbf{s}}_3 \tag{35}$$

where $\hat{\mathbf{s}}_1$, $\hat{\mathbf{s}}_2$, and $\hat{\mathbf{s}}_3$ are a set of right-handed, mutually orthogonal unit vectors fixed in an inertial or Newtonian reference frame N. Unit vectors $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$ lie in the ecliptic plane, with $\hat{\mathbf{s}}_1$ in the direction of vernal equinox, and the direction of $\hat{\mathbf{s}}_3$ is north of the ecliptic. Similarly, the position vector from S to an observatory O can be expressed as

$$\mathbf{r}^{SO} = O_1\hat{\mathbf{s}}_1 + O_2\hat{\mathbf{s}}_2 + O_3\hat{\mathbf{s}}_3 \tag{36}$$

and the position vector $\mathbf{d}$ from O to C is thus

$$\mathbf{d} \triangleq \mathbf{r}^{OC} = \mathbf{r}^{SC} - \mathbf{r}^{SO} = (X_1 - O_1)\hat{\mathbf{s}}_1 + (X_2 - O_2)\hat{\mathbf{s}}_2 + (X_3 - O_3)\hat{\mathbf{s}}_3 \tag{37}$$

A unit vector $\hat{\mathbf{d}}$ that has the same direction as $\mathbf{d}$ can be brought into a general orientation in N by first giving it the same direction as $\hat{\mathbf{s}}_1$ and then subjecting it to a body-two, 3-2-3 rotation sequence with angles of ϕ (longitude), $-\lambda$ (negative of latitude), and 0 (zero). Consequently, $\hat{\mathbf{d}}$ can be expressed as

$$\hat{\mathbf{d}} = \cos\phi\cos\lambda\hat{\mathbf{s}}_1 + \sin\phi\cos\lambda\hat{\mathbf{s}}_2 + \sin\lambda\hat{\mathbf{s}}_3 \tag{38}$$

and relationships for longitude and latitude follow,

$$\tan\phi = \frac{\hat{\mathbf{d}}\cdot\hat{\mathbf{s}}_2}{\hat{\mathbf{d}}\cdot\hat{\mathbf{s}}_1} = \frac{\mathbf{d}\cdot\hat{\mathbf{s}}_2}{\mathbf{d}\cdot\hat{\mathbf{s}}_1} \tag{39}$$

$$\sin\lambda = \hat{\mathbf{d}}\cdot\hat{\mathbf{s}}_3 = \frac{\mathbf{d}\cdot\hat{\mathbf{s}}_3}{d} \tag{40}$$

where d is the magnitude of $\mathbf{d}$, $d = (\mathbf{d}\cdot\mathbf{d})^{1/2}$.

Each observation of an object with an optical telescope yields measurements of ϕ and λ; the corresponding matrix $[\tilde{H}]_j$ is then

$$[\tilde{H}]_j = \begin{Bmatrix} \begin{bmatrix} \frac{\partial\phi}{\partial X_1} & \frac{\partial\phi}{\partial X_2} & \frac{\partial\phi}{\partial X_3} & \frac{\partial\phi}{\partial X_4} & \frac{\partial\phi}{\partial X_5} & \frac{\partial\phi}{\partial X_6} \end{bmatrix} \\ \begin{bmatrix} \frac{\partial\lambda}{\partial X_1} & \frac{\partial\lambda}{\partial X_2} & \frac{\partial\lambda}{\partial X_3} & \frac{\partial\lambda}{\partial X_4} & \frac{\partial\lambda}{\partial X_5} & \frac{\partial\lambda}{\partial X_6} \end{bmatrix} \end{Bmatrix} \tag{41}$$

The partial derivatives of ϕ are obtained by noting that $\frac{\partial}{\partial x}\tan^{-1} u = \frac{1}{1+u^2}\frac{\partial \mathrm{u}}{\partial \mathrm{x}}$; therefore, in view of equation (39),

$$\begin{aligned} \frac{\partial\phi}{\partial X_i} &= \frac{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2}{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2 + (\mathbf{d}\cdot\hat{\mathbf{s}}_2)^2}\frac{\partial}{\partial X_i}\left[\frac{(\mathbf{d}\cdot\hat{\mathbf{s}}_2)}{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)}\right] \\ &= \frac{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2}{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2 + (\mathbf{d}\cdot\hat{\mathbf{s}}_2)^2}\left[\frac{1}{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)}\frac{\partial\mathbf{d}}{\partial X_i}\cdot\hat{\mathbf{s}}_2 - \frac{(\mathbf{d}\cdot\hat{\mathbf{s}}_2)}{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2}\frac{\partial\mathbf{d}}{\partial X_i}\cdot\hat{\mathbf{s}}_1\right] \\ &= \frac{\mathbf{1}}{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2 + (\mathbf{d}\cdot\hat{\mathbf{s}}_2)^2}\left[(\mathbf{d}\cdot\hat{\mathbf{s}}_1)\frac{\partial\mathbf{d}}{\partial X_i}\cdot\hat{\mathbf{s}}_2 - (\mathbf{d}\cdot\hat{\mathbf{s}}_2)\frac{\partial\mathbf{d}}{\partial X_i}\cdot\hat{\mathbf{s}}_1\right] \qquad (i = 1,\ \ldots, 6) \end{aligned} \tag{42}$$

With the aid of equation (37), it can be seen that $\partial\mathbf{d}/\partial X_i = \hat{\mathbf{s}}_i$ for i = 1, 2, 3, and $\partial\mathbf{d}/\partial X_i = \mathbf{0}$ for i = 4, 5, 6; thus,

$$\frac{\partial\phi}{\partial X_1} = \frac{-(\mathbf{d}\cdot\hat{\mathbf{s}}_2)}{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2 + (\mathbf{d}\cdot\hat{\mathbf{s}}_2)^2} \tag{43}$$

$$\frac{\partial\phi}{\partial X_2} = \frac{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)}{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2 + (\mathbf{d}\cdot\hat{\mathbf{s}}_2)^2} \tag{44}$$

$$\frac{\partial\phi}{\partial X_3} = 0 \tag{45}$$

$$\frac{\partial\phi}{\partial X_4} = \frac{\partial\phi}{\partial X_5} = \frac{\partial\phi}{\partial X_6} = 0 \tag{46}$$

The partial derivatives of λ are obtained by noting that $\frac{\partial}{\partial x}\sin^{-1} u = \frac{1}{\sqrt{1-u^2}}\frac{\partial u}{\partial x}$; therefore, in view of equation (40),

$$\begin{aligned}
\frac{\partial\lambda}{\partial X_i} &= \frac{d}{\sqrt{d^2 - (\mathbf{d}\cdot\hat{\mathbf{s}}_3)^2}}\frac{\partial}{\partial X_i}\left[\frac{(\mathbf{d}\cdot\hat{\mathbf{s}}_3)}{d}\right] \\
&= \frac{d}{\sqrt{d^2 - (\mathbf{d}\cdot\hat{\mathbf{s}}_3)^2}}\left[\frac{1}{d}\frac{\partial\mathbf{d}}{\partial X_i}\cdot\hat{\mathbf{s}}_3 - \frac{(\mathbf{d}\cdot\hat{\mathbf{s}}_3)}{2d^3}2\mathbf{d}\cdot\frac{\partial\mathbf{d}}{\partial X_i}\right] \\
&= \frac{1}{\sqrt{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2 + (\mathbf{d}\cdot\hat{\mathbf{s}}_2)^2}}\left[\frac{\partial\mathbf{d}}{\partial X_i}\cdot\hat{\mathbf{s}}_3 - \frac{(\mathbf{d}\cdot\hat{\mathbf{s}}_3)}{d^2}\mathbf{d}\cdot\frac{\partial\mathbf{d}}{\partial X_i}\right] \qquad (i = 1, \ldots, 6)
\end{aligned} \tag{47}$$

from which we obtain

$$\frac{\partial\lambda}{\partial X_1} = \frac{-(\mathbf{d}\cdot\hat{\mathbf{s}}_1)(\mathbf{d}\cdot\hat{\mathbf{s}}_3)}{d^2\sqrt{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2 + (\mathbf{d}\cdot\hat{\mathbf{s}}_2)^2}} \tag{48}$$

$$\frac{\partial\lambda}{\partial X_2} = \frac{-(\mathbf{d}\cdot\hat{\mathbf{s}}_2)(\mathbf{d}\cdot\hat{\mathbf{s}}_3)}{d^2\sqrt{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2 + (\mathbf{d}\cdot\hat{\mathbf{s}}_2)^2}} \tag{49}$$

$$\frac{\partial\lambda}{\partial X_3} = \frac{\sqrt{(\mathbf{d}\cdot\hat{\mathbf{s}}_1)^2 + (\mathbf{d}\cdot\hat{\mathbf{s}}_2)^2}}{d^2} \tag{50}$$

$$\frac{\partial\lambda}{\partial X_4} = \frac{\partial\lambda}{\partial X_5} = \frac{\partial\lambda}{\partial X_6} = 0 \tag{51}$$

Extended Kalman filter. The extended form of the Kalman filter is presented in section 4.7.2 of reference 5; as mentioned previously, it serves to reduce the effects of errors resulting from linearization upon which the conventional filter rests. The error reduction leads to faster convergence of the extended filter in comparison with the conventional form and is accomplished by replacing the reference orbit with the current, best estimate of the true orbit after each observation. As a consequence, the estimate $\{\hat{x}_{j-1}\}$ of the state deviation following a measurement update vanishes, and it becomes unnecessary to perform the time update of the state deviation indicated in equation (25). The time update of the covariance, equation (26), and preparation for the measurement update, equations (27)–(29), proceed as before. It can be seen that, with $\{\bar{x}_j\} = \{0\}$, the measurement update for the state deviation, equation (30), gives way to

$$\{\hat{x}_j\} = [K_j]\{y_j\} \tag{52}$$

The measurement update of the covariance, equation (31), remains the same, and the aforementioned replacement can be expressed as

$$\{X_j^*\} \leftarrow \{X_j^*\} + [K_j]\{y_j\} \tag{53}$$

$$\{\hat{x}_j\} \leftarrow \{0\} \tag{54}$$

A Priori Information Obtained by the Method of Gauss

The venerable method of Gauss for preliminary determination of orbits from optical measurements has been in use for two centuries; the occasion for its creation was the discovery (and subsequent disappearance behind the Sun) in 1801 of the first minor world, Ceres. The asteroid was recovered through the application of the Gaussian method with observations that spanned only 1 month, and Carl Friedrich Gauss became recognized immediately as the premier mathematician in all of Europe.

The method requires only three observations consisting of two angular measurements each, which is in most cases the minimum number of observations needed. (As discussed previously, four observations are required when the orbits of the object and observatory are coplanar.) Gauss's method is used to obtain the a priori information needed to start the filtering sequence, as described in what follows.

A variant of Gauss's method presented (but not labeled as such) in section 5.8 of reference 1 is brought to bear, with the Sun substituted for Earth in the role of primary body, and ecliptic longitude and latitude used in place of right ascension and declination. Battin's expressions for the Lagrange coefficients F and G in terms of universal functions (eqs. (4.84), ref. 2) are used in place of the infinite series suggested in reference 1. The heart of the procedure is an iterative solution of a system of six linear algebraic equations, yielding estimates of the position $\mathbf{r}'(t_2)$ and velocity $\mathbf{v}'(t_2)$ at the time t_2 of the second of three observations. Such a solution is obtained with these steps:

1. Three pairs of measurements of longitude $\tilde{\phi}_i$ and latitude $\tilde{\lambda}_i$ for times t_i (i = 1, 2, 3) are produced according to steps 1, 3, and 4 of the procedure given in the discussion of erroneous predicted miss distance. The intentional errors to be added ($\{e\}_i$ in eq. (24)) are produced with the `MATLAB®` function `RANDN`, yielding a normal distribution with zero mean (in accordance with the Kalman filter assumption) and a standard deviation of 1. Errors are then multiplied by the telescope resolution ρ, but subsequently limited in magnitude; errors less than -2ρ are replaced with -2ρ, and errors greater than 2ρ are replaced with 2ρ.

2. The measurements constructed in step 1 are used to form three unit vectors $\hat{\mathbf{d}}_i$ with the aid of equation (38).

3. An initial guess for the vector $\mathbf{r}'(t_2) = \mathbf{r}^{SO}(t_2) + d\,\hat{\mathbf{d}}_2$ is formed with the known observatory position $\mathbf{r}^{SO}(t_2)$ and a guess for the value of d, such as 5 au.

4. An initial guess of the velocity $\mathbf{v}'(t_2)$ is produced by assuming that the object is in a circular orbit and therefore has a magnitude of $\sqrt{\mu/r_2}$, where r_2 is the magnitude of $\mathbf{r}'(t_2)$. The direction of $\mathbf{v}'(t_2)$ is given by a unit vector normal to $\mathbf{r}'(t_2)$, having the same direction as $(\hat{\mathbf{d}}_1 \times \hat{\mathbf{d}}_3) \times \mathbf{r}'(t_2)$, where the cross product $\hat{\mathbf{d}}_1 \times \hat{\mathbf{d}}_3$ is approximately normal to the orbital plane.

5. Lagrange coefficients F_1 and G_1 that relate $\mathbf{r}'(t_1)$ to $\mathbf{r}'(t_2)$ and $\mathbf{v}'(t_2)$ are constructed from these two vectors and the time difference $t_1 - t_2$. Coefficients F_3 and G_3 corresponding to t_3 are obtained similarly.

6. An iterative solution of equations (5.8-10) in reference 1 is terminated when the magnitudes of all six components of $\mathbf{r}'(t_2)$ and $\mathbf{v}'(t_2)$ change by less than 1×10^{-14} percent from one pass to the next, or after 50 passes, whichever occurs first. It is asserted in reference 1 that this process converges quickly if the time intervals $t_3 - t_2$ and $t_2 - t_1$ are "not too large."

7. Initial elements $\{X_0^*\}$ for a reference orbit are simply dot products of $\mathbf{r}'(t_2)$ and $\mathbf{v}'(t_2)$ with $\hat{\mathbf{s}}_1$, $\hat{\mathbf{s}}_2$, and $\hat{\mathbf{s}}_3$. Initial values of a state deviation $\{\hat{x}_0\}$ are given by dot products of the same unit vectors with the vector differences $\mathbf{r}'(t_2) - \mathbf{r}(t_2)$ and $\mathbf{v}'(t_2) - \mathbf{v}(t_2)$, where $\mathbf{r}(t_2)$ and $\mathbf{v}(t_2)$ are the true position and velocity at t_2.

This procedure is used 100 times with 100 sets of observations to obtain 100 different matrices $\{X_0^*\}$ and $\{\hat{x}_0\}$. Convergence was achieved in all cases. An initial covariance matrix $[P_0]$ describing the uncertainties in $\{\hat{x}_0\}$ is calculated from the 100 state deviation matrices using the `MATLAB®` function `COV`. This a priori information is used to start a Kalman filter sequence. When Monte Carlo simulations are performed, each of the 100 pairs of associated matrices $\{X_0^*\}$ and $\{\hat{x}_0\}$ are used in a separate simulation, but the same matrix $[P_0]$ is used in each simulation.

Single Observatory, and Two Observatories

The results presented in connection with multiple observations indicate that it may be possible to make reliable forecasts with two observatories whose angular resolutions are on the order of 0.1 arcsec, or with a single observatory whose resolution is better by 1 to 3 orders of magnitude. An extended Kalman filter is used to compare the amount of warning time offered by the two alternatives, where warning time is defined (somewhat more rigorously than before) to be the interval between the time a filter yields an erroneous predicted miss distance less than 1 Earth radius, and the designed time of collision. The task is facilitated by introducing a new term, "watch time," defined to be the interval between the time ε is calculated to be less than or equal to 1 lunar distance, and the designed time of collision. A watch is issued prior to a warning for hurricanes and other reasonably predictable natural calamities, and so it would be in the case of a celestial collision.

Single observatory. We assume the single observatory achieves a resolution of 0.0001 arcsec by means of interferometry; two telescopes, each having a resolution of 0.01 arcsec, are placed some distance apart in a circular heliocentric orbit in the ecliptic plane with a radius of 1 au. Observations are obtained at 7-day intervals from only one of the telescopes until $\varepsilon \leq 1$ lunar distance, and the watch time is calculated. Subsequently, both telescopes are used together to produce measurements at 1-day intervals with ρ = 0.0001 arcsec, and the filter continues to process those measurements until $\varepsilon \leq 1$ Earth radius, at which point a warning time is computed.

Two observatories. The configuration of two observatories is the same as that previously described: two telescopes are placed in a circular heliocentric orbit in the ecliptic plane, with a radius of 1 au, and phased 180° apart. For the purpose of making the comparison, one of the telescopes is assumed to be coincident with one of the instruments in the single-observatory configuration. Each telescope has a resolution of 0.01 arcsec. For any one observation, this configuration produces twice as many measurements as the single observatory. Furthermore, two observatories provide parallax that cannot be obtained with the single observatory. Measurements are taken at 7-day intervals until $\varepsilon \leq 1$ lunar distance, at which time the frequency increases to once per day until $\varepsilon \leq 1$ Earth radius.

All measurements contain Gaussian noise and are constructed as described in step 1 of the procedure for obtaining a priori information.

Results. Watch and warning times have been computed for both configurations, using the 1008 hypothetical LPC orbits given previously, and four values of the observatory initial true longitude L_0. The results obtained for $L_0 = 180°$ are presented in figure 19. Computational expense prevents us from performing Monte Carlo analysis and presenting averages; each data point represents a single set of random measurement errors for that orbit.

In almost all cases two observatories, enjoying the advantages of parallax and twice as many measurements, provide a greater watch time than a single observatory taking measurements at the same resolution. The advantage passes to the single observatory after it begins to obtain interferometric measurements; the improvement in resolution by a factor of 100 quickly reduces ε to 1 Earth radius or less, giving a greater warning time in general than two observatories.

Sixteen groups can be identified in the watch and warning times given in figure 19. The first, second, third, and fourth sets of four groups correspond to r_a = 10, 40, 70, and 100 au, respectively. The four groups within each set are associated with r_p = 0.1, 0.4, 0.7, and 1.0 au. The results for other values of L_0 are similar to those shown in figure 19.

When a conventional filter is used in place of the extended algorithm, it ceases to converge on the correct orbit and yields a warning time of zero in approximately 100 of the 1008 orbits, rendering it unsuitable for obtaining watch and warning times. An example of this behavior is exhibited in figure 20. The lack of an update to the reference trajectory in a conventional filter is the source of this defect.

Range and Range-Rate Measurements

The accuracy of orbit determination can be improved greatly with measurements of range and range-rate obtained from radar or lidar instruments. Range is simply the distance d between the observatory and the object, or the magnitude of the vector $\mathbf{d}$ given by equation (37). Range-rate, $\dot{d}$, is the time derivative of d. Measurements are performed by analyzing the returning echo from an electromagnetic pulse that

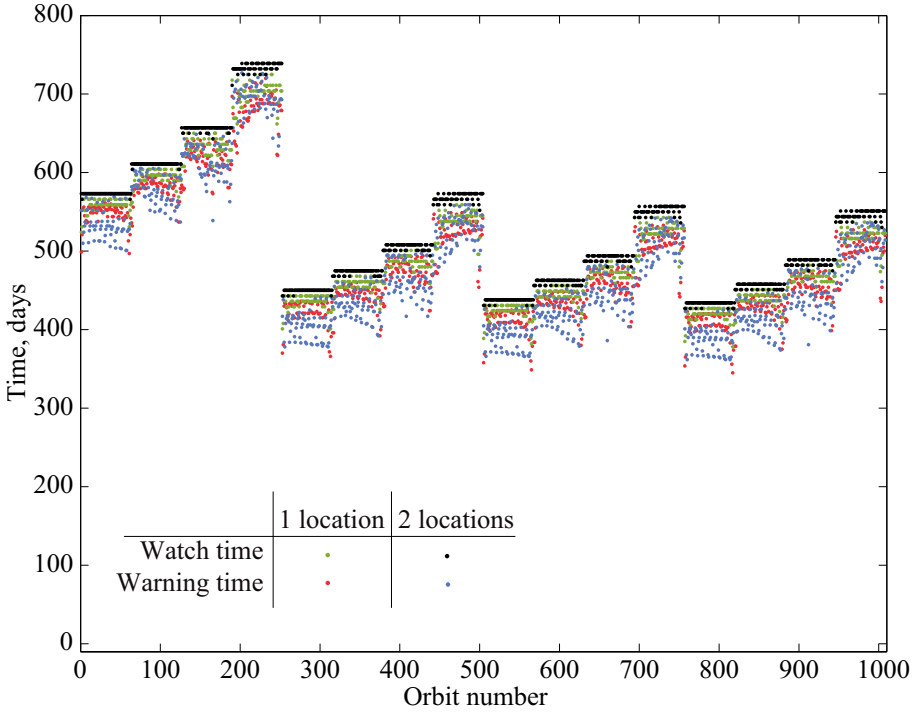

Figure 19. Watch and warning times.

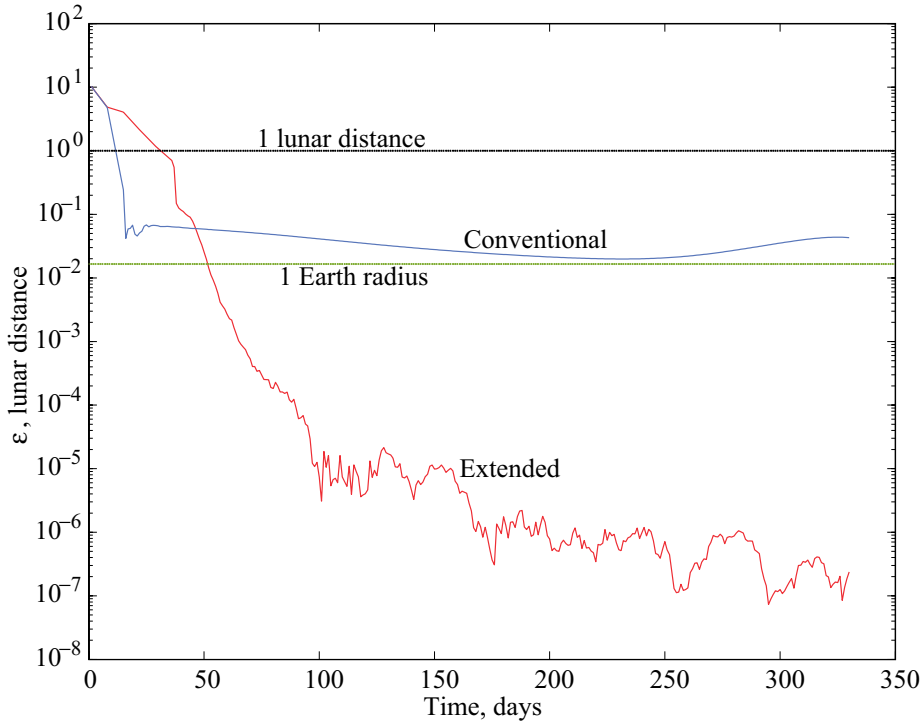

Figure 20. Lack of convergence with a conventional filter.

has been directed toward the object. The distance d is related to the wave's round trip time, and $\dot{d}$ is related to the Doppler shift in the frequency.

As pointed out in the discussion of the Kalman filter, the observation-state mapping matrix $[\tilde{H}_j]$ is dimensioned $m \times n$, the measurement covariance matrix $[R_j]$ is $m \times m$, and the observation deviation matrix $\{y_j\}$ is $m \times 1$, where m is the number of measurements and n is the size of the state, in our case $n = 6$. So far, only angular measurements of longitude and latitude have been considered, corresponding to $m = 2$. The addition of measurements of range and range-rate increases the value of m to 3 and 4, respectively. The covariance matrix $[R_j]$ is diagonal; with $m = 4$, the first two nonzero elements of $[R_j]$ contain the square of the resolution of the optical telescope, the third is the square of the resolution of range measurements, and the fourth is the square of the resolution of range-rate. When observing an asteroid at 20 times the lunar distance, current terrestrial radar systems can achieve resolutions of 10 m and 0.1 mm/s for range and range-rate, respectively.

Expressions for the elements of $[\tilde{H}_j]$ associated with measurements of longitude and latitude are developed in equations (43)–(46) and (48)–(51). The third and fourth rows of $[\tilde{H}_j]$, associated respectively with measurements of range and range-rate, are derived as follows:

$$\frac{\partial d}{\partial X_i} = \frac{\partial}{\partial X_i}(\mathbf{d}\cdot\mathbf{d})^{1/2} = \frac{1}{2}(\mathbf{d}\cdot\mathbf{d})^{-1/2} 2\mathbf{d}\cdot\frac{\partial \mathbf{d}}{\partial X_i}$$

$$= \frac{\mathbf{d}}{d}\cdot\frac{\partial \mathbf{d}}{\partial X_i} \qquad (i = 1,\ldots,6) \tag{55}$$

As mentioned earlier, it is evident from equation (37) that $\partial\mathbf{d}/\partial X_i = \hat{\mathbf{s}}_i$ for i = 1, 2, 3, and $\partial\mathbf{d}/\partial X_i = \mathbf{0}$ for i = 4, 5, 6; therefore,

$$\frac{\partial d}{\partial X_i} = \frac{\mathbf{d}\cdot\hat{\mathbf{s}}_i}{d} \qquad (i = 1,\ 2,\ 3) \tag{56}$$

$$\frac{\partial d}{\partial X_i} = 0 \qquad (i = 4,\ 5,\ 6) \tag{57}$$

Range-rate can be expressed as

$$\dot{d} = \frac{d}{dt}[(\mathbf{d}\cdot\mathbf{d})^{1/2}] = \frac{1}{2}(\mathbf{d}\cdot\mathbf{d})^{-1/2} 2\mathbf{d}\cdot\frac{{}^N d}{dt}\mathbf{d} = \frac{\mathbf{d}}{d}\cdot\frac{{}^N d}{dt}\mathbf{d} \tag{58}$$

where ${}^N d\,\mathbf{d}/dt$ indicates differentiation of $\mathbf{d}$ with respect to time in N, a Newtonian reference frame in which unit vectors $\hat{\mathbf{s}}_i$ (i = 1, 2, 3) are fixed. Strictly speaking, specification of a reference frame is not required because the left hand member of the equation is the derivative of a scalar, but the choice of N leads right away from equation (37) to

$$\frac{{}^N d}{dt}\mathbf{d} = (\dot{X}_1 - \dot{O}_1)\hat{\mathbf{s}}_1 + (\dot{X}_2 - \dot{O}_2)\hat{\mathbf{s}}_2 + (\dot{X}_3 - \dot{O}_3)\hat{\mathbf{s}}_3 \tag{59}$$

Now,

$$\frac{\partial \dot{d}}{\partial X_i} = \frac{1}{d}\frac{\partial \mathbf{d}}{\partial X_i}\cdot\frac{{}^N d}{dt}\mathbf{d} + \frac{\mathbf{d}}{d}\cdot\frac{\partial}{\partial X_i}\left(\frac{{}^N d}{dt}\mathbf{d}\right) - \frac{1}{d^2}\frac{\partial d}{\partial X_i}\mathbf{d}\cdot\frac{{}^N d}{dt}\mathbf{d}$$

$$= \frac{1}{d}\frac{\partial \mathbf{d}}{\partial X_i}\cdot\frac{{}^N d}{dt}\mathbf{d} + \frac{\mathbf{d}}{d}\cdot\frac{\partial}{\partial X_i}\left(\frac{{}^N d}{dt}\mathbf{d}\right) - \frac{1}{d^2}\left(\frac{\mathbf{d}}{d}\cdot\frac{\partial \mathbf{d}}{\partial X_i}\right)\mathbf{d}\cdot\frac{{}^N d}{dt}\mathbf{d} \qquad (i = 1, \ldots, 6) \qquad (60)$$

The last three elements of $\{X\}$ are simply

$$X_i = \dot{X}_{i-3} \qquad (i = 4,\ 5,\ 6) \qquad (61)$$

therefore, in view of equation (59), $\frac{\partial}{\partial X_i}({}^N d\,\mathbf{d}/dt) = \mathbf{0}$ for i = 1, 2, 3, and $\frac{\partial}{\partial X_i}({}^N d\,\mathbf{d}/dt) = \hat{\mathbf{s}}_{i-3}$ for i = 4, 5, 6. Together with the values for $\partial\mathbf{d}/\partial X_i$ used previously, this leads to

$$\frac{\partial \dot{d}}{\partial X_i} = \frac{\hat{\mathbf{s}}_i}{d}\cdot\frac{{}^N d}{dt}\mathbf{d} - \frac{\mathbf{d}\cdot\hat{\mathbf{s}}_i}{d^3}\left(\mathbf{d}\cdot\frac{{}^N d}{dt}\mathbf{d}\right) \qquad (i = 1,\ 2,\ 3) \qquad (62)$$

$$\frac{\partial \dot{d}}{\partial X_i} = \frac{\mathbf{d}\cdot\hat{\mathbf{s}}_{i-3}}{d} \qquad (i = 4,\ 5,\ 6) \qquad (63)$$

The benefits of measurements of range and range-rate are illustrated in figure 21, with curves showing a reduction in average erroneous predicted miss distance $\bar{\varepsilon}$ in units of Earth radius (R_E) as a function of the number of observations, taken once a day. The average is taken from 100 extended Kalman filter simulations, each with a different set of normally distributed measurement errors, and a priori information as discussed earlier. The analysis involves a hypothetical LPC designed to collide with Earth, with orbital elements obtained as described for designing a collision: r_a = 100 au, r_p = 0.7 au, i = 50°, Ω = 60°, ω = 66.68597°, and ν_0 = −137.06483° (corresponding to r = 5 au). The orbital period of this comet is 1839.4 years.

All measurements are obtained at a single observatory. The least accurate determination of the orbit, shown with the solid line, is obtained from measurements that are strictly angular with a resolution of 0.01 arcsec. The slope of the line is approximately −2.3; therefore $\bar{\varepsilon}$ varies inversely with the number of observations to the power 2.3. Greater accuracy is obtained when the angular measurements are supplemented with range measurements having a resolution of 10^3 km, as shown with the dashed curve, and the inclusion of range-rate measurements with a resolution of 1 m/s leads to the most accurate orbit determination, as displayed with the dash-dot curve. The usefulness of range and range-rate measurements is most pronounced over a short data arc of 15 to 30 days; they do not appear to be necessary over an arc of 80 or 90 days.

Modeling Comet Outgassing

A visible comet in the night sky typically appears as a fuzzy object; light is reflected diffusely from a cloud, or coma, consisting of dust evaporated by solar heating from the comet's core, or nucleus. Solar

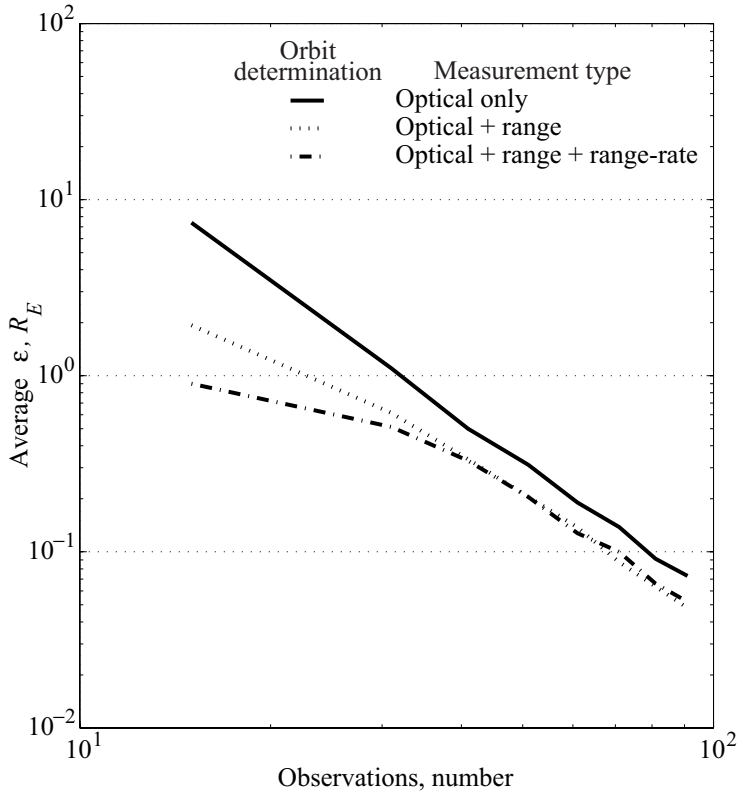

Figure 21. Benefit of range and range-rate measurements.

radiation pressure and the solar wind move the dust and ionized particles away from the nucleus, and a tail is formed. The composition of a comet is not well-known, but the nucleus is generally thought to be a "dirty snowball" made mostly of frozen water and organic and silicate compounds. When the comet is close enough to the Sun that the heat causes these volatile materials to boil off and carry solid particles with them, the comet is said to be "outgassing." Consequently, force is exerted on the comet, and the resulting perturbations to the orbit should be taken into account when attempting to calculate a trajectory precisely enough to determine if a collision with Earth will occur.

The most widely accepted method for modeling force due to outgassing was developed by Marsden, Sekanina, and Yeomans, as described in reference 7, and employed in reference 8, based on the assumption that the comet is an icy conglomerate—an object consisting mostly of frozen water that holds together bits and pieces of rock-like material. The contribution of outgassing to force per unit mass is expressed as

$$\mathbf{f}_g = g(r)[A_1\hat{\mathbf{a}}_1 + A_2\hat{\mathbf{a}}_2 + A_3\hat{\mathbf{a}}_3] \tag{64}$$

where the function $g(r)$ of heliocentric distance r is given by

$$g(r) = \alpha\left(\frac{r}{r_0}\right)^{-m}\left[1+\left(\frac{r}{r_0}\right)^{n}\right]^{-l} \tag{65}$$

The parameters other than r are constants obtained from studies of vaporization rates of comet nucleus material. The normalizing constant α is defined such that $g(1) = 1$, and r_0 is the scale heliocentric distance of high outgassing activity with a value equal to 2.808 au for frozen water. The remaining constants are reported to be $m = 2.15$, $n = 5.093$, and $l = 4.6142$; substituting these values into equation (65) and solving for the normalizing constant yields $\alpha = 0.1113$.

A right-handed, mutually perpendicular set of unit vectors $\hat{\mathbf{a}}_1$, $\hat{\mathbf{a}}_2$, and $\hat{\mathbf{a}}_3$ is defined such that $\hat{\mathbf{a}}_1$ is in the direction of the position vector from S to C, $\hat{\mathbf{a}}_2$ lies in the orbital plane, and $\hat{\mathbf{a}}_3$ is normal to the orbital plane in the direction of the specific inertial angular momentum of C relative to S. Numerical values of the constant coefficients A_1 and A_2 are calculated by studying changes in comet orbital periods, and are listed for more than 20 comets in table I of reference 7. The coefficient A_3 is usually neglected because the force normal to the orbit plane has no detectable effect on the period, therefore values for A_3 are not given. However, a normal force will influence a comet's final Earth-encounter distance, so this component must be included in simulations involving prediction of a collision.

Before undertaking such simulations, it is important to understand the magnitude of the perturbation to the comet's trajectory caused by outgassing. This is accomplished with the aid of dynamical equations formed by adding the perturbing force per unit mass to the right hand member of the equations governing two-body motion,

$$\frac{{}^N d^2}{dt^2}\mathbf{r} = -\frac{\mu}{r^3}\mathbf{r} + g(r)[A_1\hat{\mathbf{a}}_1 + A_2\hat{\mathbf{a}}_2 + A_3\hat{\mathbf{a}}_3] \tag{66}$$

where $\mathbf{r}$ is the position vector from S to C, and ${}^N d^2\mathbf{r}/dt^2$ denotes the second derivative in frame N of $\mathbf{r}$ with respect to time. By comparing numerical solutions obtained through integration of equations (66) with $[g(r) \neq 0]$ and without $[g(r) = 0]$ the contribution of the force of expelled gas, one can quantify its effect on the time history of $\mathbf{r}$. Integration is performed in connection with the hypothetical LPC whose orbital elements are given at the conclusion of the material dealing with range and range-rate measurements. Equations (66) are integrated from t_0 to t_k, at which times $r = 5$ au and 1 au, respectively. The maximum values of the coefficients reported in reference 7 are used: $A_1 = 3.61$ au/10^8 day^2 = 5.4 km/day^2, and $A_2 = 0.3269$ au/10^8 day^2 = 0.49 km/day^2. Because no values for A_3 are given, A_3 is set equal to A_2.

Figure 22 includes time histories of the differences in $\mathbf{r}\cdot\hat{\mathbf{s}}_i$ ($i = 1, 2, 3$) caused by outgassing, as well as the difference in the magnitude of $\mathbf{r}$, which is approximately 4730 km by the time $r = 1$ au. However, the more important result is that differences in position are negligible until about 250 days, when r is approximately 2.2 au; therefore, a two-body trajectory furnishes a good reference orbit until this point, warranting the linearization about such a reference orbit performed in constructing the state transition matrix of equation (34). The state transition matrix should account for outgassing if the filter is to be used with observations of a comet made with r less than, say, 2.5 au, but need not do so for r greater than this number.

The uncertainty that results from outgassing is accounted for with a white state noise of zero mean and time-varying diagonal covariance $[Q(t_j)] \triangleq [Q_j]$, the elements of which are determined by assuming the perturbing force per unit mass is constant during the interval between observations. The uncertainty in

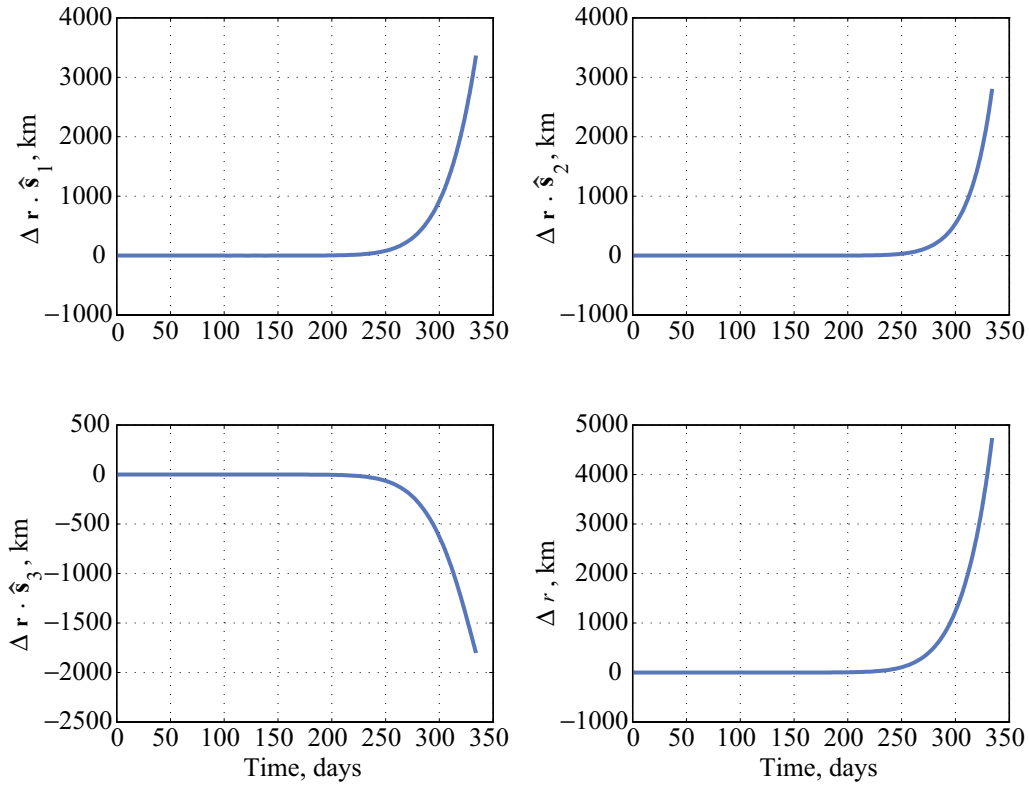

Figure 22. Change in position due to outgassing for a hypothetical comet.

the force per unit mass is reflected in the standard deviations of the coefficients listed in reference 7: $\sigma_{A_1} = 2.096$ km/day^2, and $\sigma_{A_2} = \sigma_{A_3} = 0.085$ km/day^2. The elements of $[Q_j]$ are given by

$$Q_{ii}(t_j) = \left[\frac{\mathbf{f}_g(t_j) \cdot \hat{\mathbf{s}}_i}{2} (t_j - t_{j-1})^2 \right]^2 \qquad (i = 1,\ 2,\ 3) \tag{67}$$

$$Q_{ii}(t_j) = \left\{ \left[\mathbf{f}_g(t_j) \cdot \hat{\mathbf{s}}_{i-3} \right] (t_j - t_{j-1}) \right\}^2 \qquad (i = 4,\ 5,\ 6) \tag{68}$$

where $\mathbf{f}_g(t_j)$ is evaluated at t_j with σ_{A_i} in place of A_i (i = 1, 2, 3) in equation (64). As heliocentric distance r decreases, the elements of $[Q_j]$ grow along with $g(r)$, as do the elements of the Kalman gain matrix $[K_j]$, and the filter places increasing emphasis on the measurements in determining the best estimate of the state deviation.

Once a best estimate $\{X'_j\}$ of the orbital elements is obtained from the sequential filter, an erroneous predicted miss distance is calculated by using the elements of $\{X'_j\}$ as initial values for the differential equations (66) and by performing numerical integration with the `MATLAB®` function `ODE45` from the time t_j of the final observation to the designed time of collision t_k.

As it is not possible to know ahead of time the values of the outgassing coefficients for a newly discovered comet, a Monte Carlo analysis is performed with a set of random values uniformly distributed between the minimum and maximum values of A_1 and A_2 reported in reference 7. Random values for A_3 are obtained from the same distribution used for A_2. The results of this analysis are presented next.

The Body-Targeting Plane

Body-plane (B-plane) targeting is a method commonly used in the preliminary design of interplanetary missions to determine a distance of closest approach of a spacecraft to a celestial body of interest. A miss distance B calculated with this method can furnish a metric for quantifying orbit determination accuracy in connection with hypothetical comets that collide with Earth. The uncertainty in B can be expressed in terms of error ellipses constructed on the B-plane and is actually the more important result when analyzing the accuracy with which an orbit has been determined.

A vector **B**, and a plane containing **B**, are constructed as follows. First, one neglects the gravitational force exerted by the Earth E on the hypothetical object C; therefore, E does not perturb the heliocentric orbit of C, and the velocity $\mathbf{v}^{C/E^*}$ in N of C relative to the mass center E^* of E is regarded as constant in N during the time C is within the Earth's sphere of influence. The magnitude of $\mathbf{v}^{C/E^*}$ is denoted by v_∞. A unit vector $\hat{\mathbf{S}}$ having the same direction as $\mathbf{v}^{C/E^*}$ is shown in figure 23, and is often considered to be parallel to the incoming asymptote of the geocentric hyperbolic trajectory (assuming C is not captured) resulting from gravitational attraction of the Earth that must in fact exist. In the absence of gravitational

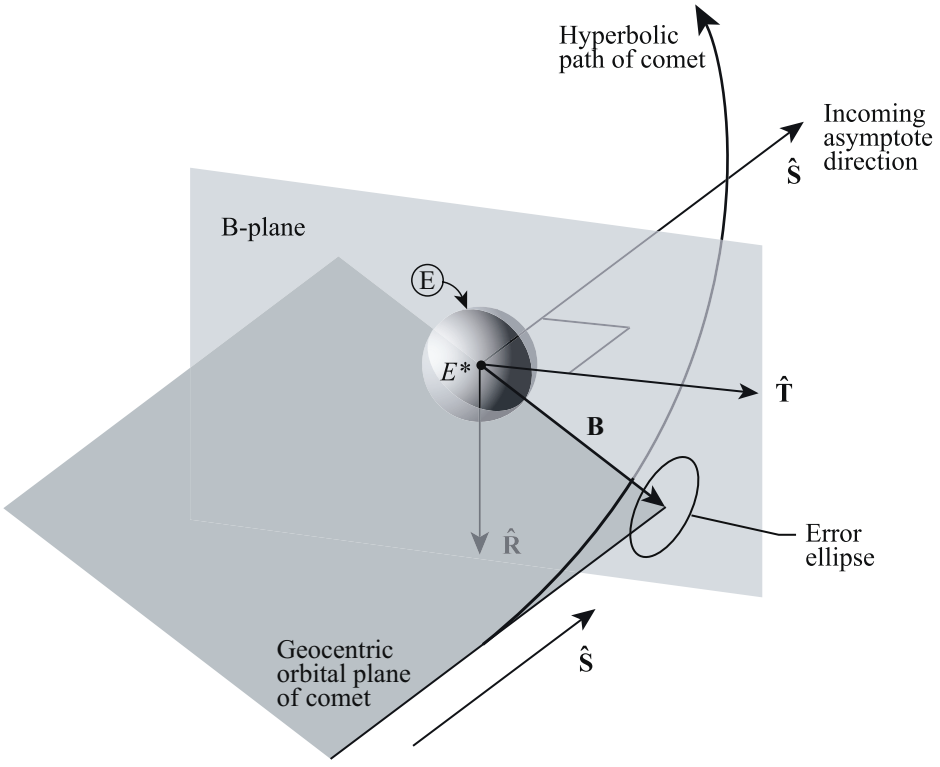

Figure 23. The B-plane.

force exerted by E, *C* continues along the incoming asymptote. The unit vector $\hat{\mathbf{S}}$ can be obtained with the expression

$$\hat{\mathbf{S}} = \frac{\mathbf{v}^{C/E^*}}{\mathrm{v}_\infty} = \frac{\mathbf{v}^{C/S^*}(t_k) - \mathbf{v}^{E^*/S^*}(t_k)}{\mathrm{v}_\infty} \tag{69}$$

where $\mathbf{v}^{C/S^*}(t_k)$ and $\mathbf{v}^{E^*/S^*}(t_k)$ are, respectively, the velocities in *N* of *C* and *E**, relative to the Sun's mass center *S**, at the designed time of collision t_k.

The B-plane is defined such that it contains the point *E** and is perpendicular to $\hat{\mathbf{S}}$; the plane contains two unit vectors $\hat{\mathbf{T}}$ and $\hat{\mathbf{R}}$ that are mutually perpendicular to each other and to $\hat{\mathbf{S}}$. The unit vector $\hat{\mathbf{T}}$ is chosen to be in the ecliptic plane; therefore, it can be obtained from the relationship

$$\hat{\mathbf{T}} = \frac{\hat{\mathbf{S}} \times \hat{\mathbf{s}}_3}{\left[\left(\hat{\mathbf{S}} \times \hat{\mathbf{s}}_3\right) \cdot \left(\hat{\mathbf{S}} \times \hat{\mathbf{s}}_3\right)\right]^{1/2}} \tag{70}$$

where, as one will recall, $\hat{\mathbf{s}}_3$ is perpendicular to the ecliptic plane. Unit vector $\hat{\mathbf{R}}$ is then simply

$$\hat{\mathbf{R}} = \hat{\mathbf{S}} \times \hat{\mathbf{T}} \tag{71}$$

The vector **B** denotes the position from *E** to the point at which the incoming asymptote intersects the B-plane, which is the point of closest approach when one neglects the gravitational attraction of E (or, equivalently, neglects the mass of E). The magnitude *B* of **B** is the corresponding distance of closest approach.

As previously mentioned, the geocentric velocity $\mathbf{v}^{C/E^*} = \mathrm{v}_\infty \hat{\mathbf{S}}$ is assumed to remain constant in *N* for the purpose of constructing **B**; therefore, *C* must necessarily travel on a straight line between the point of closest approach and the position at the designed time of collision, as shown in figure 24. Hence, the specific angular momentum **h** in *N* of *C* relative to *E** is given by

$$\mathbf{h} = \mathbf{B} \times \mathrm{v}_\infty \hat{\mathbf{S}} = \boldsymbol{\varepsilon} \times \mathrm{v}_\infty \hat{\mathbf{S}} \tag{72}$$

where $\boldsymbol{\varepsilon}$ is the position vector from *E** to *C* at the designed time of collision, obtained from the process of orbit determination. From this relationship, an expression for **B** can be derived through premultiplication with $\hat{\mathbf{S}}$ in a cross product,

$$\begin{aligned} \hat{\mathbf{S}} \times \mathbf{h} &= \hat{\mathbf{S}} \times (\boldsymbol{\varepsilon} \times \mathrm{v}_\infty \hat{\mathbf{S}}) = \hat{\mathbf{S}} \times (\mathbf{B} \times \mathrm{v}_\infty \hat{\mathbf{S}}) \\ &= \mathrm{v}_\infty [(\hat{\mathbf{S}} \cdot \hat{\mathbf{S}})\mathbf{B} - (\hat{\mathbf{S}} \cdot \mathbf{B})\hat{\mathbf{S}}] \\ &= \mathrm{v}_\infty \mathbf{B} \end{aligned} \tag{73}$$

because $\hat{\mathbf{S}} \cdot \hat{\mathbf{S}} = 1$, and $\hat{\mathbf{S}} \cdot \mathbf{B} = 0$ by definition. Hence,

$$\mathbf{B} = \frac{\hat{\mathbf{S}} \times \mathbf{h}}{\mathrm{v}_\infty} = \hat{\mathbf{S}} \times (\boldsymbol{\varepsilon} \times \hat{\mathbf{S}}) \tag{74}$$

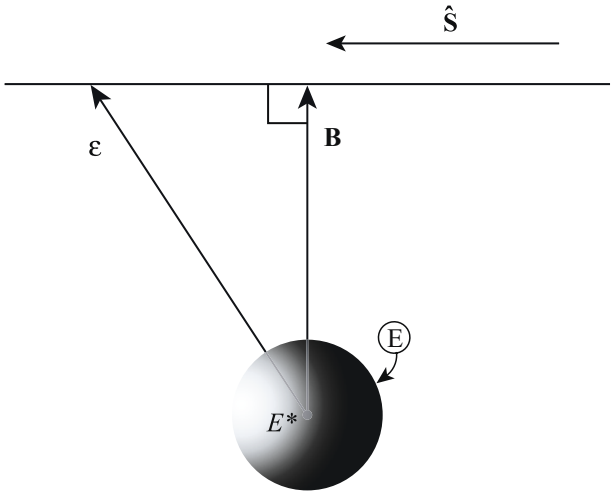

Figure 24. Erroneous predicted miss position, and closest approach.

and, because **B** lies in the plane formed by $\hat{\mathbf{R}}$ and $\hat{\mathbf{T}}$, it can always be expressed as

$$\mathbf{B} = B_R \hat{\mathbf{R}} + B_T \hat{\mathbf{T}} \tag{75}$$

As described previously, the orbit of C is designed to pass through the ecliptic at a heliocentric distance of 1 au, a point that is assumed to be coincident with E^*. A perfect determination of the orbit must necessarily yield $\boldsymbol{\varepsilon} = \mathbf{B} = \mathbf{0}$; however, inevitable random noise present in observations will lead to a prediction that $\boldsymbol{\varepsilon}$ and $\mathbf{B}$ each differ from $\mathbf{0}$. If the magnitude of either of these position vectors is determined to be, say, less than or equal to 1 Earth radius (R_E), one would be justified in predicting a collision unless there is a large uncertainty in the orbit solution as quantified by the covariance of $\boldsymbol{\varepsilon}$ or $\mathbf{B}$, which bears the following relationship to the estimate error covariance matrix $[P_j]$ in equation (31).

A time update according to equation (26) is performed in order to obtain the covariance matrix at the designed time of collision t_k, from the covariance matrix $[P_j]$ after a measurement update at t_j,

$$[\bar{P}_k] = [\Phi(t_k, t_j)][P_j][\Phi(t_k, t_j)]^{\mathrm{T}} + [Q_j] \tag{76}$$

The effects of outgassing can be reflected in $[Q_j]$ as set forth in equations (67) and (68) for short intervals of time. Alternatively, one may divide a large time interval into shorter ones over which $\mathbf{f}_g$ is regarded as constant in N. When 1 year of warning time separates t_k and t_j, dividing the interval into 20 equal parts ensures that $[Q_j]$ is not dependent on the propagation time step and that the constant outgassing force assumption is still valid. The covariance time update for the shorter intervals (each approximately 2 weeks in length) is modified,

$$[\bar{P}_{j+1}] = [\Phi(t_{j+1}, t_j)][P_j][\Phi(t_{j+1}, t_j)]^{\mathrm{T}} + [\bar{Q}_j] \tag{77}$$

to employ the average value of the state noise covariance

$$[\bar{Q}_j] \triangleq \frac{[Q_j] + [Q_{j+1}]}{2} \tag{78}$$

Equation (77) is applied successively until t_{j+1} becomes the time of collision t_k, yielding a 6×6 covariance matrix $[\overline{P}_k]$.

The upper left 3×3 partition of $[\overline{P}_k]$ deals with position and is therefore of the most interest in what follows; it is denoted by the symbol $[\overline{P}_r]$. As indicated in connection with the modeling of comet outgassing, the state transition matrix should be constructed according to reference 8 when outgassing becomes a significant perturbation; in lieu of this, $[\overline{P}_r]$ can be supplemented by adding a covariance matrix assembled from a difference in position associated with uncertainty resulting from outgassing. Equation (66) is numerically integrated from the time of the first observation t_1 to the time of collision t_k, first with $\mathbf{f}_g = \mathbf{0}$, and then three times, with one of σ_{A_i} used in place of A_i (i = 1, 2, 3) in equation (64). The difference in $\mathbf{r}(t_k)$ obtained with σ_{A_i}, and with $\mathbf{f}_g = \mathbf{0}$, is expressed in terms of unit vectors $\hat{\mathbf{s}}_1$, $\hat{\mathbf{s}}_2$, and $\hat{\mathbf{s}}_3$ and presented as a column matrix $\{\delta\}_i$, allowing one to form the 3×3 covariance matrix

$$[\overline{P}_S] \triangleq [\overline{P}_r] + \sum_{i=1}^{3} \{\delta\}_i \{\delta\}_i^{\mathrm{T}} \tag{79}$$

This matrix (like the rest of $[\overline{P}_k]$) is associated with unit vectors $\hat{\mathbf{s}}_1$, $\hat{\mathbf{s}}_2$, and $\hat{\mathbf{s}}_3$, but it is more convenient to work with a covariance that describes the uncertainties in the directions marked by $\hat{\mathbf{S}}$, $\hat{\mathbf{T}}$, and $\hat{\mathbf{R}}$. The transformed covariance, indicated with $[\overline{P}_B]$, is formed as

$$[\overline{P}_B] = [{}^{S}C^{B}]^{\mathrm{T}} [\overline{P}_S] [{}^{S}C^{B}] \tag{80}$$

where the elements of the direction cosine matrix $[{}^{S}C^{B}]$ are defined as ${}^{S}C_{i1}^{B} \triangleq \hat{\mathbf{s}}_i \cdot \hat{\mathbf{S}}$, ${}^{S}C_{i2}^{B} \triangleq \hat{\mathbf{s}}_i \cdot \hat{\mathbf{T}}$, and ${}^{S}C_{i3}^{B} \triangleq \hat{\mathbf{s}}_i \cdot \hat{\mathbf{R}}$ (i = 1, 2, 3).

As discussed in section 4.16 of reference 5, the function

$$\left(\{x_k\} - \{\bar{x}_k\}\right)^{\mathrm{T}} [\overline{P}_B]^{-1} \left(\{x_k\} - \{\bar{x}_k\}\right) = l^2 \tag{81}$$

describes a three-dimensional ellipsoid, where the portion of the state deviation associated with position at t_k is $\{x_k\}$, having covariance $[\overline{P}_B]$ and mean $\{\bar{x}_k\}$, and where l is a constant. The ellipse corresponding to a given value of l, formed by the intersection of the ellipsoid with the B-plane (shown in fig. 23), is obtained from the 2×2 partition of $[\overline{P}_B]$ associated with $\hat{\mathbf{T}}$ and $\hat{\mathbf{R}}$, represented by $[\overline{P}_b]$, which can be diagonalized

$$[\overline{P}_p] = [{}^{B}C^{P}]^{\mathrm{T}} [\overline{P}_b] [{}^{B}C^{P}] \tag{82}$$

where $[{}^{B}C^{P}]$ is a matrix containing the normalized eigenvectors of $[\overline{P}_b]$, with elements defined as ${}^{B}C_{1i}^{P} \triangleq \hat{\mathbf{T}} \cdot \hat{\mathbf{p}}_i$ and ${}^{B}C_{2i}^{P} \triangleq \hat{\mathbf{R}} \cdot \hat{\mathbf{p}}_i$ (i = 1, 2, 3). The diagonal covariance matrix $[\overline{P}_p]$ contains the eigenvalues of $[\overline{P}_b]$ associated with the eigenvectors of unit length $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$, which are parallel to the principal axes of the error ellipse in the B-plane. The orientation of the principal axes of the ellipse with respect to $\hat{\mathbf{T}}$ and $\hat{\mathbf{R}}$ can be obtained as $\theta = \arctan\left({}^{B}C_{21}^{P} / {}^{B}C_{11}^{P}\right)$. The square roots of the two nonzero

elements of $[\bar{P}_p]$ are the semimajor and semiminor axes of the ellipse, denoted as σ_1 and σ_2, respectively. The difference in time that it takes for the comet to travel between the positions marked by **B** and $\boldsymbol{\varepsilon}$ is assumed to be negligible, and the error ellipse therefore describes the uncertainty in the estimate that the comet will pass through the position indicated by **B**. The bivariate Gaussian probability density function can be expressed in invariant form according to equation (2.2-39) of reference 9, or equation (A.19.1) of reference 5,

$$f(x_1,x_2)=\frac{1}{2\pi|P|^{1/2}}e^{-\frac{1}{2}(\{x\}-\{\bar{x}\})^{\mathrm{T}}[P]^{-1}(\{x\}-\{\bar{x}\})} \tag{83}$$

where $[P]$ can be either of the covariance matrices $[\bar{P}_b]$ or $[\bar{P}_p]$, $|P|$ denotes the determinant of $[P]$, and $\{x\}^T=[x_1 \;\; x_2]$ contains the random variables of the state deviation at t_k associated with $\hat{\mathbf{T}}$ and $\hat{\mathbf{R}}$ in the case of $[\bar{P}_b]$, or with $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$ in the case of $[\bar{P}_p]$. The random variables have a mean of $\{\bar{x}\}$, a column matrix containing the two dot products $\mathbf{B}\cdot\hat{\mathbf{T}}=B_T$ and $\mathbf{B}\cdot\hat{\mathbf{R}}=B_R$, or $\mathbf{B}\cdot\hat{\mathbf{p}}_1$ and $\mathbf{B}\cdot\hat{\mathbf{p}}_2$. Contours of constant f are ellipses.

Inspection of results from a Monte Carlo analysis shows the validity of the assumption of linearity inherent in the development of the sequential filter. The object of interest is the hypothetical comet described previously at the conclusion of the section dealing with range and range-rate measurements, without the perturbation induced by outgassing. The extended Kalman filter is employed to produce 100 orbit solutions from 91 optical observations spaced 1 day apart, made from a single observatory whose resolution ρ is 0.01 arcsec. Each solution begins with a preliminary orbit determined by the method of Gauss, and processes a different set of measurements of longitude and latitude, whose random errors are formed according to step 1 of the procedure set forth for obtaining a priori information. The procedure yields 100 values of **B** and associated covariance matrices $[\bar{P}_b]$ (or $[\bar{P}_p]$), each of which corresponds to an ellipse whose center is the position given by **B** and indicated with a point in figure 25. The average values of σ_1, σ_2, θ, B_T, and B_R are given in the second column of table 2. Counterparts to the first three of these parameters may be obtained by forming the covariance of the 100 state deviations $\{x\}$ produced by the filter after the final measurement update, and mapping this single covariance matrix to t_k. Error ellipse parameters obtained in this way are given in the third column of table 2, and the 1σ ellipse centered on the average values of B_T and B_R is shown in figure 25 to be a reasonable representation of the distribution of the individual intersection points.

The time of closest approach t_b can be determined from the designed time of collision t_k by referring to figure 24 and recalling the assumption that the geocentric velocity $\mathbf{v}^{C/E^*}=\mathrm{v}_\infty\hat{\mathbf{S}}$ is assumed to remain constant in N for the purpose of constructing **B**. The time t_k-t_b required for C to travel from the position marked by **B** to the position marked by $\boldsymbol{\varepsilon}$ is given by

$$t_k-t_b=\frac{1}{\mathrm{v}_\infty}(\boldsymbol{\varepsilon}-\mathbf{B})\cdot\hat{\mathbf{S}}=\frac{\boldsymbol{\varepsilon}\cdot\hat{\mathbf{S}}}{\mathrm{v}_\infty} \tag{84}$$

The standard deviation σ_t of this time interval obtained from the 100 orbit solutions is given in the fourth column of table 2.

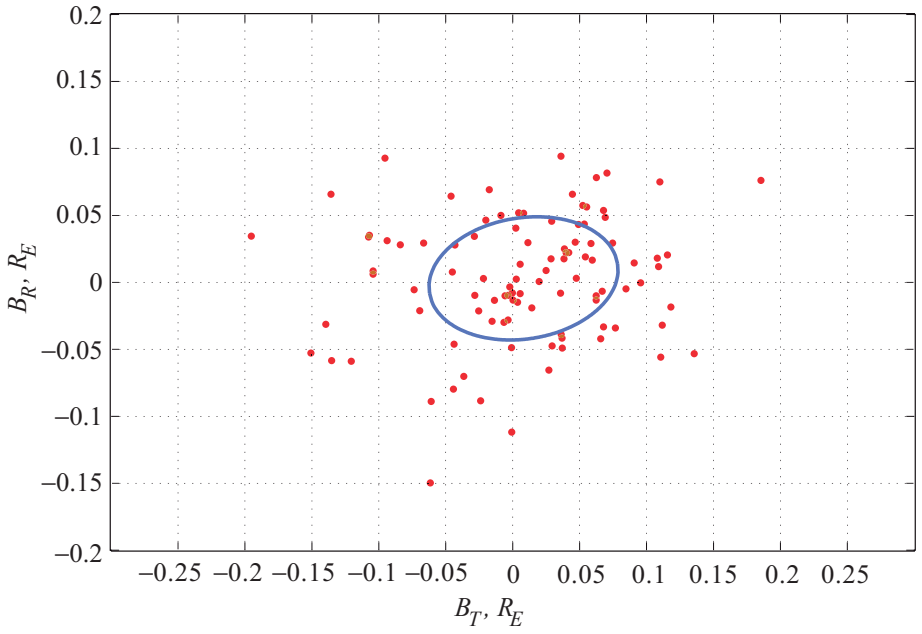

Figure 25. 1σ error ellipse for long-period comet.

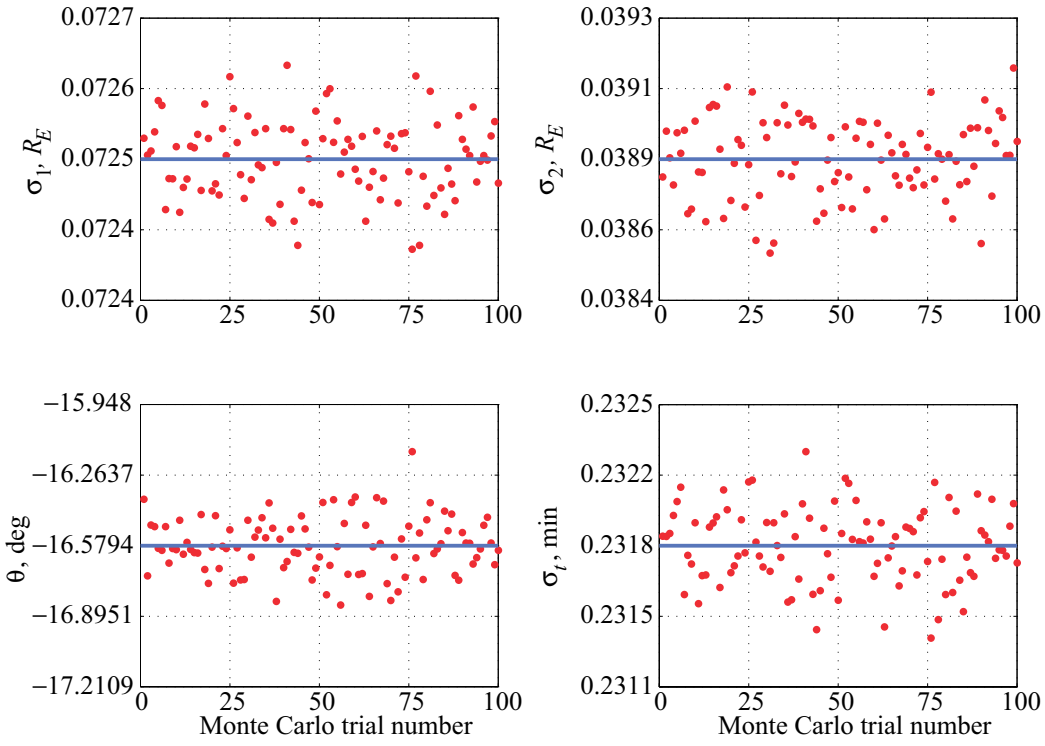

Figure 26. Variation in 1 σ error ellipse for long-period comet.

Alternatively, the accuracy of the estimate of the time at which the comet will arrive at the position marked by **B** is indicated by the extent of the three-dimensional error ellipsoid normal to the B-plane. Letting σ_S denote the square root of the element of $[\overline{P}_B]$ associated with $\hat{\mathbf{S}}$, the uncertainty in time of arrival at the B-plane is given simply by

$$\sigma_t = \sigma_S / \mathrm{v}_\infty \tag{85}$$

Table 2 contains the average values of σ_S and time of arrival standard deviation in the second column and the values obtained from the single covariance matrix in the third column. Values of σ_1, σ_2, σ_S, or σ_t corresponding to 2σ and 3σ ellipses are calculated easily by multiplying the 1σ value in table 2 by 2 or 3.

The similarity of σ_t in the fourth column to the values recorded in the second and third columns shows the errors in time of arrival have the same standard deviation as the travel times between the positions marked by **B** and **ε**. The close agreement in the values of all parameters in the second and third columns of table 2 is one indication the assumption of linearity made in deriving the sequential filter is justified. Additional support is furnished in figure 26: the individual error ellipse parameters for each of the 100 Monte Carlo simulations are virtually the same, with only slight variations distributed on either side of the mean (shown with a solid blue line). The similarity in these error ellipse parameters is evidence of linearity; in other words, the covariance of the state deviation at the designed time of collision is virtually independent of the point at which the orbit intersects the B-plane. Consequently, there is no need to perform additional Monte Carlo analysis.

The foregoing conclusions remain unchanged after including the effect of outgassing, as indicated in table 3. A comparison of tables 2 and 3 reveals that comet outgassing increases the size of the 1σ error ellipse, signaling a less accurate determination of the orbit.

Table 2. 1σ Error Ellipse Parameters for Sample Comet

Parameter	Average	Single	Eq. (84)
σ_1, R_E	0.072535	0.071114	
σ_2, R_E	0.038867	0.045221	
θ, deg	−16.579420	−8.787936	
B_T, R_E	0.008113		
B_R, R_E	0.002959		
σ_S, R_E	0.078602	0.077102	
σ_t, min	0.231822	0.227399	0.227400

Table 3. 1σ Error Ellipse Parameters for Sample Outgassing Comet

Parameter	Average	Single	Eq. (84)
σ_1, R_E	0.287882	0.287184	
σ_2, R_E	0.064515	0.069989	
θ, deg	−31.912291	−31.814645	
B_T, R_E	−0.210096		
B_R, R_E	−0.108374		
σ_S, R_E	0.334448	0.334101	
σ_t, min	0.986446	0.985425	0.774256

Probability Associated With B

The probability that an object passes through the B-plane within a certain distance of the position indicated by **B** is found by integrating the probability density function over the area of a circle. The result can be used in two ways, provided the assumption of linearity is warranted. First, in the unique circumstance that occurs because the orbit of our hypothetical object is known precisely, the probability quantifies the accuracy of the orbit determined by the sequential filter. Second, for the general case in which the true orbit is unknown, a probability of collision between the object and Earth can be obtained.

In what follows, numerical examples are created with the hypothetical comet described in the discussion regarding range and range-rate measurements; again, the filter processes 91 optical observations from a single observatory, taken 1 day apart, and measurement errors are normally distributed, based upon a resolution of $\rho = 0.01$ arcsec. In view of the demonstration in the preceding section, the assumption of linearity is justified, and we dispense with Monte Carlo analysis; furthermore, the eigenvalues and eigenvectors of the covariance matrix $[\bar{P}_b]$ are considered independent of **B** (the mean of the state deviation) obtained from the sequential filter in the orbit determination process. The actual orbital parameters of the hypothetical object are adopted to produce the reference trajectory for the sequential algorithm, and a conventional filter is put into service so that the reference orbit remains unaltered.

The bivariate probability density function given by equation (83) is integrated numerically over the area of a circle whose center is given by the coordinate pair $(\bar{x}_1, \bar{x}_2)$, the elements of which are used to form the matrix $\{\bar{x}\}$. The center of the circle varies over several points in a coarsely spaced grid in the neighborhood of (0, 0) in the B-plane. For each center point, the function $f(x_1, x_2)$ is evaluated at every point of a fine, evenly spaced grid that approximates a circle, and then multiplied by the area enclosed by four neighboring grid points (arranged in a square); the products thus obtained are summed to yield the probability that the object passes through the B-plane, somewhere within the circle.

The radius of the circle is not critical for the purpose of gauging orbit determination accuracy. However, a particular radius is required for calculating probability of collision; the area of the circle must be equal to the effective collision cross section of the planet, which is larger than the physical cross section in order to account for Earth's gravitational attraction that has been ignored up to now. The effective collision radius r_c is given by equation (8.3-30) in reference 1,

$$r_c = \frac{R_E}{\mathrm{v}_\infty} \sqrt{\mathrm{v}_\infty{}^2 + \frac{2\mu_E}{R_E}} \tag{86}$$

where R_E is the physical radius of Earth (regarded as a sphere), and μ_E is the Earth's gravitational parameter.

The probability that the object passes through a circle of radius r_c centered at the Earth's mass center E^*, (0, 0), is found to be 99.9 percent. The probability of passing through circles with other centers is likewise computed, and each concentric ellipse in figure 27 is a locus of points for which the probability is equal to 1 in 10^z ($z = 1, 3, 6$). Because the actual orbit is designed to result in $\mathbf{B} = \mathbf{0}$, it can be said that the higher the probability, the more accurate the orbit determination. With the luxury of knowing the actual orbit corresponds to $\boldsymbol{\varepsilon} = \mathbf{B} = \mathbf{0}$, the information presented in figure 27 gives reason to be 100 times more confident in a filter solution when **B** marks a point on the contour labeled -1 than when it marks a point on the -3 contour.

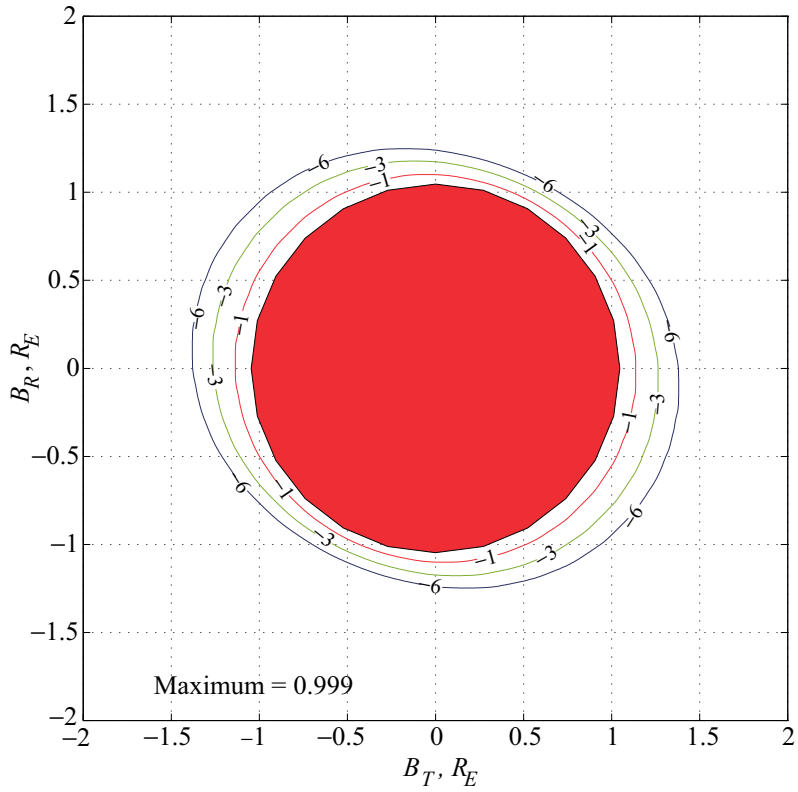

Figure 27. Probability of passing through a circle of radius r_c.

In addition to gauging orbit determination accuracy, figure 27 can be used to indicate probability of collision. In the event the filter solution yields $\mathbf{B} = \mathbf{0}$ (rather unlikely), the probability of collision is 99.9 percent, and the probability is one in a million that the object crosses the B-plane through a circle of radius r_c centered at, say, (1, −1). On the other hand, if the filter solution gives $\mathbf{B} = 1\hat{\mathbf{T}} - 1\hat{\mathbf{R}}$ (a more likely result than $\mathbf{B} = \mathbf{0}$), the probability that the comet passes within a distance r_c of this point is 99.9 percent; furthermore, there is one chance in a million that the comet will instead cross the B-plane at (0, 0) and collide with Earth. This interpretation of figure 27 rests on the assumption of linearity and the concomitant concept that the uncertainty in **B** is described by the same error ellipse, no matter what the direction and magnitude of **B**. In other words, the same error ellipse can be translated or shifted to any point on the B-plane.

The presence of outgassing can be expected to reduce the probability of passing through $\mathbf{B} = \mathbf{0}$; measurements of range and range-rate can be expected to increase it. The orbit solution based on angular measurements with $\rho = 0.01$ arcsec from 91 observations at a single observatory is so accurate that neither of these effects are observed; however, varying the number of observations brings them to light in figure 28. The plot at the top shows the increase in central probability with number of observations. The second and third plots show the respective collision probabilities associated with solutions of $\mathbf{B} = 1\hat{\mathbf{T}}$ and $\mathbf{B} = 1\hat{\mathbf{R}}$. There are six curves in each plot corresponding to three combinations of observation types. The first consists of two angular measurements at each observation, the second includes range measurements at a resolution of 10^3 km, and the third adds measurements of range-rate at a resolution of 1 m/s. For each combination, results with outgassing are compared with those in which outgassing is absent.

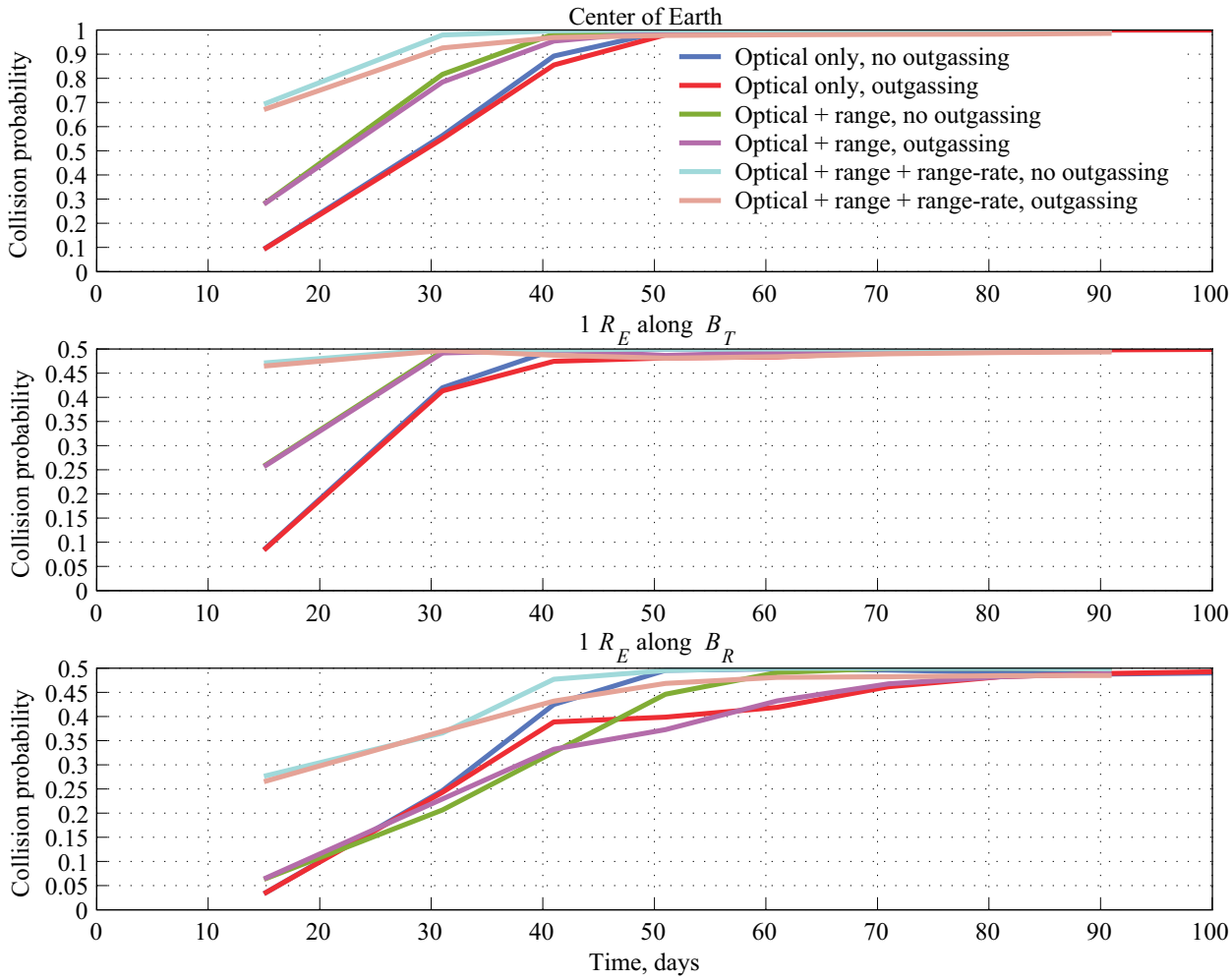

Figure 28. Probability, outgassing, and measurements of range and range-rate.

Batch Filter for Near-Earth Asteroids

Having examined orbit determination of LPCs in some detail, we turn our attention now to potentially dangerous objects whose orbital parameters are rather different, NEAs that are so named when they have perihelial distances less than 1.3 au. These asteroids are further classified by dividing them into groups based upon values of semimajor axis a, aphelial distance r_a, and perihelial distance r_p. Earth-crossing asteroids with $a < 1$ au and $r_a > 0.983$ au are members of the Aten class. Earth-crossing asteroids with $a > 1$ au and $r_p < 1.017$ au belong to the Apollo class. Earth-approaching asteroids with $a > 1$ au and $1.017 < r_p < 1.3$ au are classified as Amors (these asteroid orbits do not actually cross Earth's orbit but come close enough that a perturbation could cause a collision). It is possible that there exists another class of Earth-approaching asteroids for which $r_a < 1$ au, although no asteroids of this type have been detected to date.

The following discussion of analysis is facilitated by the definition of two types of Earth-crossing asteroids that are more general than the foregoing classes. "Interior" asteroids are said to be those with $r_p < 1$ au and $r_a = 1$ au, while "exterior" asteroids have $r_p \leq 1$ au and $r_a > 1$ au. Figures 29 and 30 contain examples of these two types of orbits, where the Earth E is shown in a circular orbit of radius 1 au

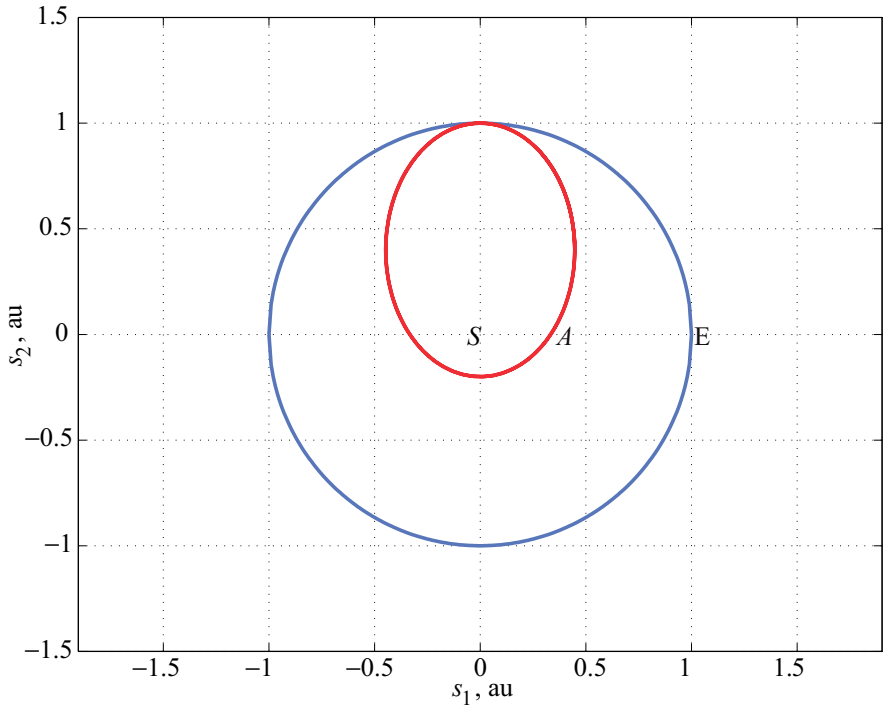

Figure 29. Example of interior asteroid orbit.

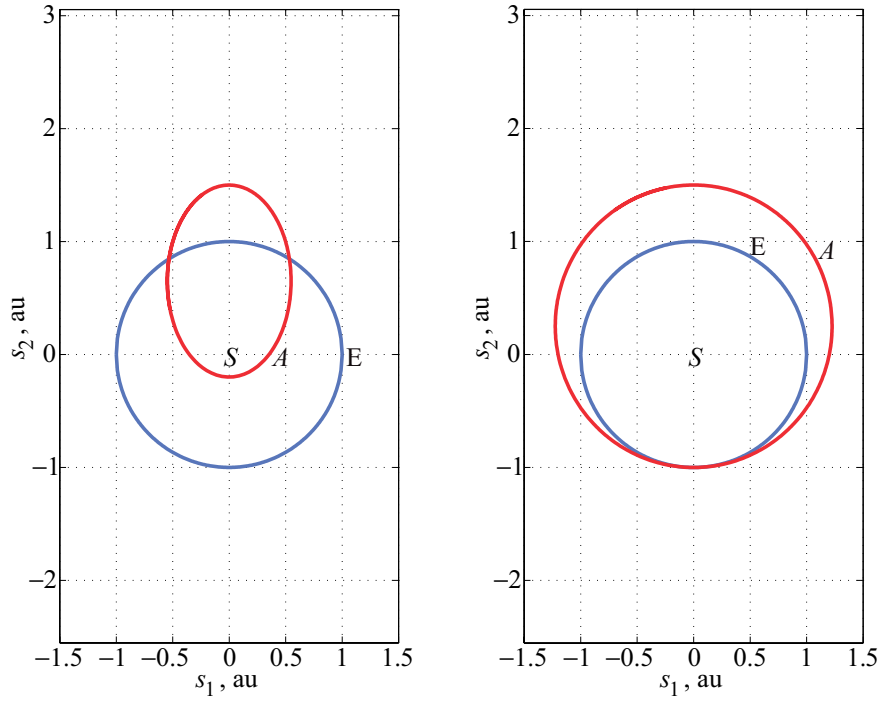

Figure 30. Examples of exterior asteroid orbits.

about the Sun S, and an asteroid A is shown in a coplanar elliptical orbit. The negative signs of eccentric anomalies E and E_k required for preperihelial collision with LPCs (see the discussion of warning time) apply also to exterior asteroids; however, collisions of interior asteroids must always take place after perihelion, therefore the sign changes are unnecessary.

The focus of the present analysis is upon asteroids discovered less than 1 orbital period prior to collision. Preliminary orbit determination for interior and exterior asteroids is performed in much the same way as it is in connection with LPCs, with the following details regarding the seven-step procedure given in connection with erroneous predicted miss distance.

In the first step, the earliest opportunity for observations t_0 is specified as 60 days prior to the time of collision t_k, and the initial value of true anomaly $\nu(t_0)$ is calculated accordingly. The reason for basing $\nu(t_0)$ on time to collision rather than a specified value of r has to do with the observation times chosen in the third step. Four observations permit the determination of orbits inclined 0° or 180° to the ecliptic; we wish to make the fourth and final observation at least 2 weeks prior to collision in order to provide some amount of warning. Hence, the times t_1, t_2, t_3, and t_4 of the observations are taken to be 11, 22, 33, and 44 days after t_0, leaving 16 days between the final observation and collision. A single observatory in a circular heliocentric orbit of radius 1 au in the ecliptic plane is used to make the observations, and the initial true longitude L_0 takes on values of 0°, 90°, 180°, and 270°.

The uniformly distributed measurement errors associated with the fourth step are bounded in magnitude by a telescope resolution of $\rho = 0.1$ arcsec, and once again are constructed with a pseudo-random number algorithm. Average values $\bar{\varepsilon}$ of erroneous predicted miss distances for 100 sets of measurements are shown.

In connection with the fifth step, the position and velocity at t_1 are to be determined by the method of least squares. It so happens that the algorithm is quite sensitive to the initial estimate $\{\hat{x}\}$ of position and velocity; therefore, values corresponding to day 11 are used for exterior asteroids, and to day 12 for interior asteroids. The MATLAB® algorithm failed to converge for a number of measurement sets in a handful of cases; nevertheless, a value of ε is computed from the unconverged estimates. The sensitivity to the initial guess, and the lack of convergence, may be due to changes in an asteroid position that are significantly larger than changes in a comet orbit over the same interval of time.

Orbit determination for interior asteroids is discussed first, and exterior asteroids follow.

Interior Asteroid Orbits

Classical orbital elements of interior asteroids are constructed with $r_a = 1.0$ au, $r_p = 0.2$, 0.4, 0.6, and 0.8 au, $i = -40°$, $-30°$, $-20°$, …, 40°, and $\Omega = 0°$, 45°, 90°, 135°, and 180°. For each of these 180 sets of orbital elements, the argument of perihelion ω is determined to be 180°, according to equation (5). The initial true longitude L_0 of the observatory takes on four values, each differing by 90°, resulting in a total of $4 \times 180 = 720$ cases.

The results from this analysis are presented in figure 31 as a function of asteroid orbital elements i and r_p; each data point represents $\bar{\varepsilon}$ for one case, and a family of 20 cases (in which Ω and L_0 vary) form one stack of points. At the lower values of r_p one can see an inverse relationship of $\bar{\varepsilon}$ to i, which is not unexpected because orbits that are coplanar with the observatory require a minimum of four observations to be determined. Although not as evident at this scale, the relationship exists at the higher values of r_p as well. At $i = 0°$ and $r_p = 0.6$ au one notices two outlying values of $\bar{\varepsilon}$, approximately 0.45 lunar distance. These two cases are distinguished from each other only by the value of L_0 (and of Ω, which is

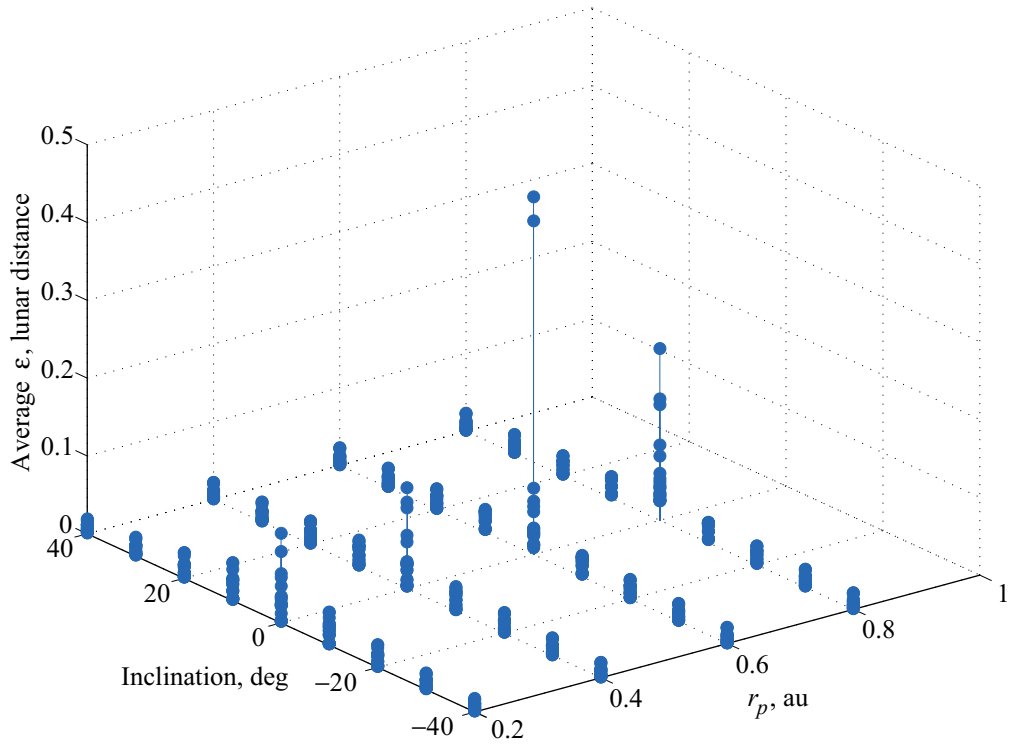

Figure 31. Interior near-Earth asteroids.

undefined for $i = 0$). For a third case in this family the least squares algorithm was unable to converge on a solution for 38 of the 100 measurement sets, and the associated data point is not shown. Of the 719 results displayed, 697 values of $\bar{\varepsilon}$ are less than 0.05 lunar distance, or about 3 Earth radii. When one holds i constant and plots the results in two dimensions, $\bar{\varepsilon}$ is seen to vary proportionally to r_p when $i = \pm 40°$ and inversely to r_p when $i = \pm 10°$. When displayed as functions of orbital elements other than i and r_p, no remarkable trends are visible in the results.

In eight other cases a certain number of measurement sets did not produce convergence. These cases have common values of $r_a = 1$ au, $r_p = 0.2$ au, and $\Omega - L_0 = 45°$. The nine cases that suffered from convergence problems are listed in table 4, together with their orbital parameters, initial true longitude of the observatory, and the number of trials out of 100 that did not converge.

Table 4. Interior Asteroid Orbits With Unconverged Trials

Case	r_a, au	r_p, au	i, deg	Ω, deg	ω, deg	ν_0, deg	L_0, deg	Trials
115	1.0	0.6	0	180	180	123.72	180	38
362	1.0	0.2	–40	45	180	132.52	0	85
367	1.0	0.2	–30	45	180	132.52	0	3
397	1.0	0.2	30	45	180	132.52	0	4
402	1.0	0.2	40	45	180	132.52	0	86
544	1.0	0.2	–40	135	180	132.52	90	1
549	1.0	0.2	–30	135	180	132.52	90	10
579	1.0	0.2	30	135	180	132.52	90	9
584	1.0	0.2	40	135	180	132.52	90	7

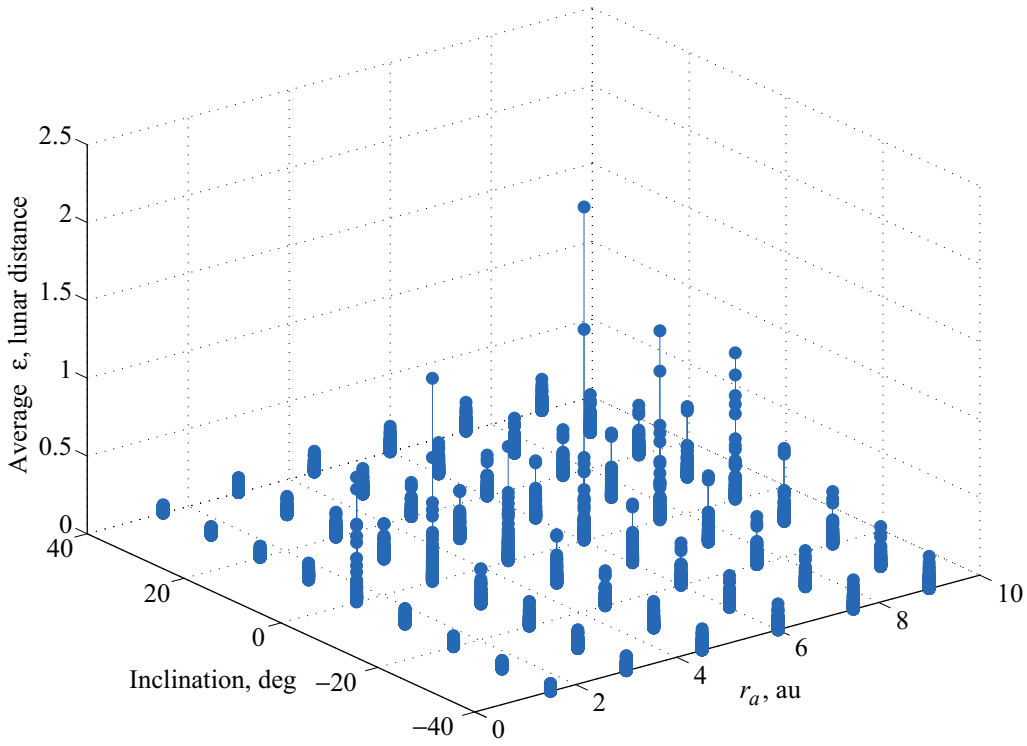

Figure 32. Accuracy of exterior NEA orbits, as a function of i and r_a.

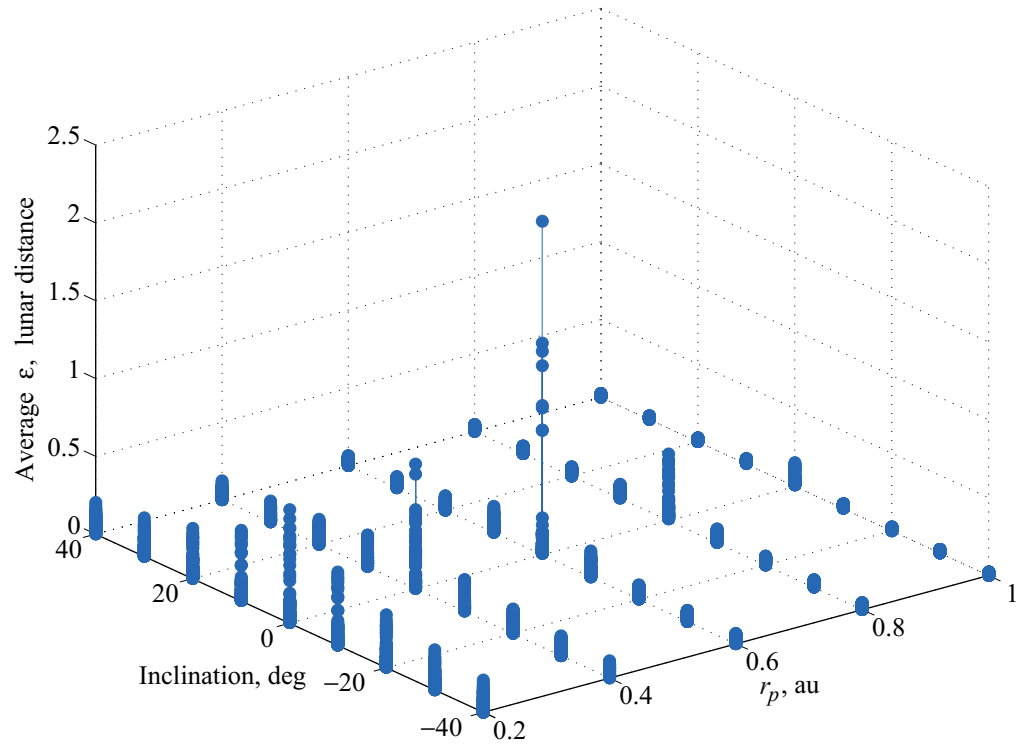

Figure 33. Accuracy of exterior NEA orbits, as a function of i and r_p.

Exterior Asteroid Orbits

For exterior asteroid orbits, 1350 sets of classical orbital elements are constructed with r_a = 1.5, 3.0, 4.5, …, 9.0 au, r_p = 0.2, 0.4, 0.6, …, 1.0 au, i = –40°, –30°, –20°, …, 40°, and Ω = 0°, 45°, 90°, 135°, and 180°. For each set, the argument of perihelion ω is computed according to equation (5). L_0 takes on the same four values as in the interior cases; thus, a total of 4 × 1350 = 5400 cases for exterior asteroid orbits are examined.

The average values $\bar{\varepsilon}$ of erroneous predicted miss distance are shown in figures 32 and 33 as functions of i, and r_a and r_p, respectively. Relatively poor orbit determination is exhibited once again for asteroids traveling in orbits that have low inclination with respect to the observatory, when only four observations are obtained. Accuracy generally decreases for smaller values of r_p and for larger values of r_a.

Of the 5400 cases examined, 5389 yielded $\bar{\varepsilon}$ less than 1 lunar distance. In two of the remaining cases $\bar{\varepsilon}$ was reported to be approximately 4.8 lunar distances; however, the least squares algorithm was unable to converge on a solution for 15 and 38 of the 100 measurement sets, and the associated data points are not shown. The largest value of $\bar{\varepsilon}$ shown in figures 32 and 33 is 2.14 lunar distances, and the smallest is 0.00125 lunar distance, or 0.08 of an Earth radius. In seven other cases involving an object whose orbit is coplanar with that of the observatory (i = 0), a certain number of measurement sets did not produce convergence, as reported in table 5.

Table 5. Exterior Asteroid Orbits With Unconverged Trials

Case	r_a, au	r_p, au	i, deg	Ω, deg	ω, deg	ν_0, deg	L_0, deg	Trials
564	4.5	0.6	0	135	85.59	–120.36	180	15
2991	3.0	0.4	0	0	112.62	–140.51	0	49
3082	3.0	0.8	0	45	62.96	–106.90	0	5
3307	4.5	0.8	0	45	59.10	–102.53	0	32
3532	6.0	0.8	0	45	57.42	–100.64	0	30
3757	7.5	0.8	0	45	56.48	–99.58	0	24
3982	9.0	0.8	0	45	55.88	–98.91	0	19
4479	3.0	1.0	0	135	0	–62.90	0	1
4612	4.5	0.6	0	45	85.59	–120.37	0	38

Although the analysis for NEAs has not been extended to include multiple observatories or improved telescope resolution, it is likely that configurations of two or more observatories, together with multiple observations, will improve preliminary orbit determination for asteroids, as has been demonstrated in connection with LPCs.

Concluding Remarks

An examination has been made of the effects of several factors upon determination of orbits of comets and asteroids on a collision course with Earth, including the time interval between successive observations, the distance at which the object is detected, the resolution of the telescope, the orbital parameters of the object, the number of observations, and, in a rather limited manner, placement in the solar system of one or more observatories. The primary factor is seen to be the length of the data arc. As mentioned at the outset, the analysis is based on a crucial assumption that optical measurements can in fact be obtained from a telescope; the validity of this assumption depends upon several factors that have not yet been

addressed, among them the quality of the instrument, the extent of solar illumination of the object (solar phase angle), the optical reflectivity (albedo) of the object, occultation by the Sun or other bodies, and interference of light from sources such as the stars in the galactic plane, zodiacal dust, and the Sun.

Provided the measurements are available, the following conclusions are made based upon a study of many hypothetical objects, each of which has been given an orbit resulting in a collision with Earth at a specified time and a specified position. As used here, accuracy is expressed in terms of erroneous predicted miss distance, the difference between an object's determined position and the specified position, evaluated at the specified time of collision. A single observatory in a circular heliocentric orbit of radius 1 au in the ecliptic plane, possessing an optical resolution of 0.1 arcsec, provides good preliminary orbit determination accurate to less than 8 lunar distances for long-period comets (LPCs) observed over a 66-day arc at a heliocentric distance of 6.5 au. Multiple observations (approximately 100, spaced 1 day apart) improve the accuracy to less than 14 Earth radii, and a second observatory phased 180° from the first improves the accuracy further still to less than 1 Earth radius. Of the multiobservatory configurations examined here, two observatories of equal resolution in circular heliocentric orbits of radius 1 au in the ecliptic plane, phased 180° apart, provide the best balance of cost and accuracy in orbit determination for LPCs. Thirty to forty observations taken over a 98-day period appear to give results nearly as good as 99 observations taken 1 day apart. For near-Earth asteroids observed over an interval between 60 and 16 days before collision, the accuracy is better than 2 lunar distances for the exterior variety and 0.5 lunar distance for interior asteroids.

The results reported here give a preliminary indication of what is required to predict impending collisions accurately. We must keep in mind that such predictions must be reliable enough to support a decision to expend considerable resources in attempting to alter the orbit of a dangerous body; very accurate knowledge of the orbit is probably essential if orbital modification is to steer the object away from a collision, rather than toward one. Moreover, predictions must have a degree of integrity that virtually eliminates false alarms and unfounded panic. The results indicate that it may be possible to make reliable forecasts with two observatories whose angular resolution is on the order of 0.1 arcsec, or with a single observatory whose resolution is better by 1 to 3 orders of magnitude. Additional study of these two alternatives will have to weigh the advantages in geometry and redundancy of multiple observatories against the expense of putting them in place and maintaining them; a single observatory near Earth could be easier to maintain, but less likely to be in the best position for obtaining measurements.

Outgassing does not perturb a comet's orbit appreciably until it reaches a heliocentric distance of approximately 2.5 au; the perturbation in position is less than the Earth's radius by the time the comet reaches 1 au. Analysis of dangerous LPCs performed with Kalman filters reveals several interesting results. With a mix of weekly and daily optical observations from one or two observatories, a credible warning of collision, in the absence of outgassing, can be expected at least a year in advance if the object is spotted when it is about 6.5 au from the Sun. If a comet is discovered at a heliocentric distance of 5 au, then supplementing daily angular measurements whose resolution is 0.01 arcsec, with range measurements accurate to 1000 km and range-rate measurements good to 1 m/s, is beneficial when dealing with a short data arc of 15 to 30 days, but doing so offers little advantage when the arc is 80 or 90 days long. The assumption of linearity upon which the sequential filters depend is warranted, even in the presence of outgassing, when daily angular measurements with a resolution of 0.01 arcsec are available for 3 months. These sorts of measurements yield extremely accurate orbit determination, even with only 2 months of observations; the addition of range data can reduce the necessary data arc to 1.5 months, and only 1 month is needed when range-rate is included.

Future work should investigate the effects of perturbations that have been ignored here, including the gravitational attraction of Jupiter, Saturn, and other planets, solar radiation pressure, and relativistic effects. The advantages of placing observatories in heliocentric orbits other than those studied here deserves careful consideration. In addition, the consequences of missed observations (due to the reasons stated earlier) should be examined.

References

1. Bate, R. R.; Mueller, D. D.; and White, J. E.: *Fundamentals of Astrodynamics*, Dover Publications, Inc., New York, 1971.
2. Battin, R. H.: *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA, New York, 1987.
3. Chodas, P. W.; and Yeomans, D. K.: Orbit Determination and Estimation of Impact Probability for Near Earth Objects. *Proceedings of the 22nd Annual AAS Rocky Mountain Guidance and Control Conference*, Breckenridge, CO, Feb. 3–7, 1999 (A99-34626 09-12), San Diego, CA, Univelt, Inc. (Advances in the Astronautical Sciences, vol. 101), 1999, pp. 21–40.
4. Chodas, P. W., et al.: Automated Detection of Potentially Hazardous Near-Earth Encounters. AAS 01-461, AAS/AIAA Astrodynamics Specialists Conference, Quebec City, Quebec, Canada, July 30–Aug. 2, 2001.
5. Tapley, B. D.; Schutz, B. E.; and Born, G. H.: *Statistical Orbit Determination*, Academic Press, May 2004.
6. Grace, A.: *Optimization Toolbox User's Guide*, The MathWorks, Inc., Natick, MA, 1994. See also <http://www.mathworks.com/access/helpdesk/help/toolbox/optim/lsqnonlin.shtml>
7. Marsden, B. G.; Sekanina, Z.; and Yeomans, D. K.: Comets and Nongravitational Forces. V., *The Astronomical Journal*, vol. 78, no. 2, March 1973, pp. 211–225.
8. Ekelund, J. E.; and Yeomans, D. K.: A Program for the Accurate Generation of Ephemerides for Halley's Comet. AAS 85-350, *Astrodynamics 1985*: *Proceedings of the AAS/AIAA Astrodynamics Conference*, American Astronautical Society, San Diego, CA, 1986, pp. 795–808.
9. Gelb, A., ed.: *Applied Optimal Estimation*, M.I.T. Press, Cambridge, MA, 1974.