ASTROMETRIC PARAMETER ESTIMATION SUITABLE FOR SIMULATIONS

MARC A. MURISON
Astronomical Applications Department, U.S. Naval Observatory, 3450 Massachusetts Ave NW, Washington DC 20392
murison@aa.usno.navy.mil

November 7, 2000

ABSTRACT

I show that one can accurately solve for the standard errors of the five classical astrometric parameters via a simple linear least squares approach using scan-direction timing observations.

Key words: FAME, parameter estimation, least squares, simulations D:\FAME\ObservationDensity\ParameterEstimation.lwp

1. INTRODUCTION

In FAME simulations, it is useful to determine the mission-averaged errors of the classical astrometric parameters position, proper motion, and parallax. The two FAME observables are a timing in the scan direction and a position measurement in the cross-scan direction. Ideally, one would use both observables to calculate the astrometric parameters and errors for each observed star, using least squares as a parameter estimator. The cost function of the estimation algorithm would simply be the distance on the sky of the observation point (as derived from the observables) from the actual location of the observed star. However, in general for FAME the scan-direction timing information is orders of magnitude more precise than the cross-scan information. Since the observations for any given star are distributed widely in scan angle (the local direction of the scan motion), the cost function just defined is dominated by the large cross-scan position variance. This renders the parameter estimation process a meaningless exercise in manipulating numerical noise.

The alternative is to use only the scan-direction timings when the cross-scan single-measurement errors are large. An observation then consists of an “observation line”, which is the limiting case of an infinitely squashed error ellipse with infinite semimajor axis. The location of the observation line perpendicular to the line (i.e., the scan direction) is precisely measured, but, effectively, the observation can fall anywhere on the line (cross-scan direction). An obvious measure to use as the estimation cost function is then the perpendicular distance from the observation line to the star position.

When parallax is taken into account, the perpendicular distance is a nonlinear function of the astrometric parameters. Since the parameters are all small (position is measured relative to a nearby fiducial point, e.g. the catalog position of the star), one can linearize the perpendicular distance in the parameters. This allows a straightforward use of linear least squares to determine the parameters and, more importantly from a simulation perspective, the parameter errors, from a set of timing observations. For FAME, typical errors are of order 100 μas or μas/yr, so the second-order cost function terms neglected by the linearization are a factor of order 10^4 smaller than the first-order term. The other error inherent in this approach is the neglect of cross-scan information in the observations. For the majority of the FAME program stars, this error is insignificant. Hence, in general this approach is formally quite precise. On the other hand, for the FAME grid (or fiducial) stars, where the precision of the cross-scan information is within an order of magnitude of that of the scan-direction information, a method that takes advantage of that cross-scan information will likely yield somewhat better results than the method presented in this paper.

This paper develops the perpendicular distance approach. Section 2 examines the geometry of the problem. The goal of this section is to derive a useable cost function for the least squares estimator. The least squares implementation is explicitly shown in section 3.

2. PERPENDICULAR DISTANCE FROM THE OBSERVATION LINE

2.1. Equations of the Observation Line and the Perpendicular Line

Define a local coordinate frame $[\Delta\lambda, \Delta\beta]$ whose origin is located at some reference position $[\lambda_{ref}, \beta_{ref}]$ on the sky in ecliptic coordinates, and assume the scale of $\Delta\lambda$ and $\Delta\beta$ are such that the local frame is suitably Euclidean (Figure 1). The reference position could be, e.g., a star's catalog position, or in a simulation it could be the center of a grid cell. Suppose a star is at some epoch t_0 located at the position $[\Delta\lambda_0, \Delta\beta_0]$ in the local frame. After a time $t - t_0$ the star's proper motion causes it to move to $[\Delta\lambda_0, \Delta\beta_0] + (t - t_0)[\mu_\lambda, \mu_\beta]$, where $\vec{\mu} = [\mu_\lambda, \mu_\beta]$ is the proper motion. Superimposed on the proper motion is the star's annual parallax, which is an ellipse, say of semimajor axis a. The center of the parallax ellipse moves according to the star's proper motion. Hence, at time t the geometry of a star's position is as illustrated in Figure 1.

The instantaneous direction of motion of the spacecraft field of view as it scans the sky defines the *scan angle*, designated q in Figure 1. The scan angle is the angle of the FOV motion with respect to a local ecliptic meridian through the star. It can be shown (Murison, 2000) that the scan angle is a function of position on the sky relative to the Sun and of the precession cone angle (the angle between the spacecraft spin axis and the precession axis, which is in the nominal direction of the Sun), given by

$$\sin q = Q \tag{1}$$

and

$$\cos q = \frac{[\sin^2(\lambda-\lambda_\odot)-\cos^2(\lambda-\lambda_\odot)\sin^2\beta]\cos\psi}{\sin(\lambda-\lambda_\odot)[1-\cos^2(\lambda-\lambda_\odot)\cos^2\beta]} + \frac{[\cos^2(\lambda-\lambda_\odot)-\sin^2\psi]\cos(\lambda-\lambda_\odot)\sin\beta}{Q\sin(\lambda-\lambda_\odot)[1-\cos^2(\lambda-\lambda_\odot)\cos^2\beta]} \quad (2)$$

where Q is the pair of quadratic solutions

$$Q = \frac{\cos(\lambda-\lambda_\odot)\cos\psi\sin\beta}{1-\cos^2(\lambda-\lambda_\odot)\cos^2\beta} \pm \frac{|\sin(\lambda-\lambda_\odot)|\sqrt{\sin^2\psi-\cos^2(\lambda-\lambda_\odot)\cos^2\beta}}{1-\cos^2(\lambda-\lambda_\odot)\cos^2\beta} \quad (3)$$

ψ is the precession cone angle, and $\lambda_\odot$ is the ecliptic longitude of the Sun. These equations are very useful in simulations. Furthermore, λ and β may be expressed as explicit functions of the Euler angles which tie the external reference frame to the spacecraft body frame (Murison, 2000). The Euler angles in turn may be derived from simple analytical models or full numerical integrations of the rigid body equations of motion, depending on the desired completeness of the model of the spacecraft scanning motion.

If, in eqs. (1)-(3), we use $\lambda = \lambda_{ref}$ and $\beta = \beta_{ref}$ instead of $\lambda = \lambda_\star$ and $\beta = \beta_\star$, we introduce an error in the determination of q by these equations. To first order, the errors in $\sin q$ and $\cos q$ are proportional to $\lambda_\star - \lambda_{ref}$ and $\beta_\star - \beta_{ref}$. As will be seen later, $\sin q$ and $\cos q$ appear in the astrometric parameter estimation only when multiplied by small parameters, so the error introduced in the calculation of q by using the local coordinate origin rather than the true position of the star as propagated through the analysis is second-order and therefore ignorable. Hence, for purposes of astrometric parameter estimation using simulated observations, q as determined from eqs. (1)-(3) can be treated as a calculated and therefore known quantity, at least to first order, by setting $\lambda = \lambda_{ref}$ and $\beta = \beta_{ref}$.

Suppose an observation is made at time t. For FAME, the precision of observations in the cross-scan direction for most stars — ~2-3 arcsec — is quite bad due to the cross-scan CCD binning required by the observing scheme, while the precision in the scan direction is over three orders of magnitude better at 580 microarcseconds (μas). Thus, for most stars, there is effectively no cross-scan information contained in the individual observations. For the ~10^3 grid (or fiducial) stars, no CCD binning is performed, so the full cross-scan precision of ~2 milliarcseconds (mas) is available. This is still an order of magnitude larger than the scan-direction precision. Hence, in terms of information contained in the observables that contributes to the precision of position measurements, the cross-scan component of the observations is inconsequential for the majority of stars and, at best, relatively unimportant for the small subset of grid stars. Therefore, we will consider only the perpendicular distance y of the star from the *observation line*, as illustrated in Figure 1. The observation line is defined as the line passing through a given observation and perpendicular to the scan direction at the time of the observation. In Figure 1, the scan-direction measurement is ΔS, and the cross-scan measurement is ΔC. (The observations are with respect to the local coordinate frame origin.)

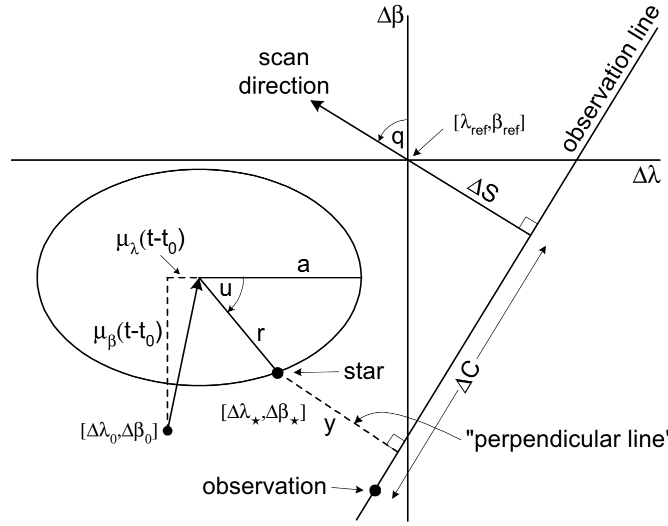

Figure 1 — Geometry of an observation.

The perpendicular distance y of the an observed star from the observation line will be a function of time and of the astrometric parameters $[a, \Delta\lambda_0, \Delta\beta_0, \mu_\lambda, \mu_\beta]$, where a is the parallax ellipse semimajor axis (henceforth referred to as the parallax), $[\Delta\lambda_0, \Delta\beta_0]$ is the position at epoch t_0, and $[\mu_\lambda, \mu_\beta]$ is the proper motion. The equation of the observation line is

$$\Delta\beta = \Delta\lambda \tan q + b \quad (4)$$

where b is the intercept in the local frame. From an observation we have a point on the observation line, given by (cf. Figure 1)

$$\begin{bmatrix} \Delta\lambda \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} -\Delta C\cos q + \Delta S\sin q \\ -\Delta C\sin q - \Delta S\cos q \end{bmatrix} \quad (5)$$

Plugging (5) into (4), we solve for b and obtain the equation for the observation line,

$$\Delta\beta = \Delta\lambda \tan q - \frac{\Delta S}{\cos q} \quad (6)$$

Notice that the dependence on the cross-scan measurement, ΔC, has dropped out.

Similarly, the equation of the perpendicular line is, from the geometry of Figure 1,

$$\Delta\beta = -\Delta\lambda \cot q + b \quad (7)$$

for some intercept b. Now, in the local frame the star position is

$$\begin{bmatrix} \Delta\lambda_\star \\ \Delta\beta_\star \end{bmatrix} = \begin{bmatrix} \Delta\lambda_0 + (t-t_0)\mu_\lambda + r\cos u \\ \Delta\beta_0 + (t-t_0)\mu_\beta - r\sin u \end{bmatrix} \quad (8)$$

where the parallax ellipse parameters r and u will be determined later. Plugging (8) into (7), we have the equation for the perpendicular line,

$$\Delta\beta = \Delta\beta_\star - (\Delta\lambda - \Delta\lambda_\star)\cot q \quad (9)$$

2.2. The Perpendicular Distance

The perpendicular distance $|y|$ follows from the intersection point of the two lines. Setting eq. (6) equal to eq. (9), we find that the location of the intersection point is

$$\begin{bmatrix} \Delta\lambda_i \\ \Delta\beta_i \end{bmatrix} = \begin{bmatrix} \Delta\lambda_\star + (\Delta S + \Delta\beta_\star \cos q - \Delta\lambda_\star \sin q)\sin q \\ \Delta\beta_\star - (\Delta S + \Delta\beta_\star \cos q - \Delta\lambda_\star \sin q)\cos q \end{bmatrix} \quad (10)$$

Now, the perpendicular distance is

$$y = \sqrt{(\Delta\lambda_\star - \Delta\lambda_i)^2 + (\Delta\beta_\star - \Delta\beta_i)^2} \quad (11)$$

Hence, using (10), we have the simple result

$$|y| = |\Delta S + \Delta\beta_\star \cos q - \Delta\lambda_\star \sin q| \quad (12)$$

where $\Delta\lambda_\star$ and $\Delta\beta_\star$ are given by (8).

2.3. Parallax Ellipse Geometry

As the Earth orbits the Sun, the finite distance of a star will result in the apparent motion of that star with respect to very distant stars. This motion is an ellipse whose semimajor axis is inversely proportional to the star's distance from the Sun. A star in the direction of one of the ecliptic poles will execute a circle. For ecliptic latitudes closer and closer to the ecliptic plane, the eccentricity of the parallactic ellipse increases, until the ellipse of a star on the ecliptic has degenerated into a line segment. Hence, the eccentricity is a function of ecliptic latitude. Consider the pole case which is a circle: $x^2 + y^2 = a^2$. As viewed from the star, the Earth's path is this circle. Now tilt the solar system by an angle $\frac{\pi}{2} - \beta$, which is equivalent to considering a star with ecliptic latitude β. The transformation of coordinates is $\xi = x$ and $\eta = y \sin\beta$. Hence, the equation of the circle as projected onto the inclined plane is

$$\frac{\xi^2}{a^2} + \frac{\eta^2}{a^2 \sin^2\beta} = 1 \quad (13)$$

which is an ellipse of eccentricity

$$e = \sqrt{1 - \left(\frac{b}{a}\right)^2} = |\cos\beta| = \cos\beta \quad (14)$$

Changing to polar coordinates, $\xi = r\cos u$ and $\eta = r \sin u$, we obtain from (13) the equation of the ellipse in polar form,

$$r = \frac{a|\sin\beta_\star|}{\sqrt{1 - \cos^2\beta_\star \sin^2 u}} \quad (15)$$

where $\beta_\star$ is the ecliptic latitude of a star, a is the parallax (i.e., the ellipse semimajor axis), and u is the polar angle.

The ellipse polar angle is related nonlinearly (by projection) to the true anomaly v of the Earth's (assumed circular) orbit around the Sun, and hence to the longitude of the Sun. The ellipse polar angle u is the projection of the mirror of v. We need to relate the linear motion of the Sun to the nonlinear (with time) angle u. Consider the circle and the angle E shown in Figure 2. If the circle corresponds to the parallax ellipse at the ecliptic pole (so that it reflects the Earth's circular orbit), then the angle E is related linearly to the ecliptic longitudes of the star and the Sun and may be written

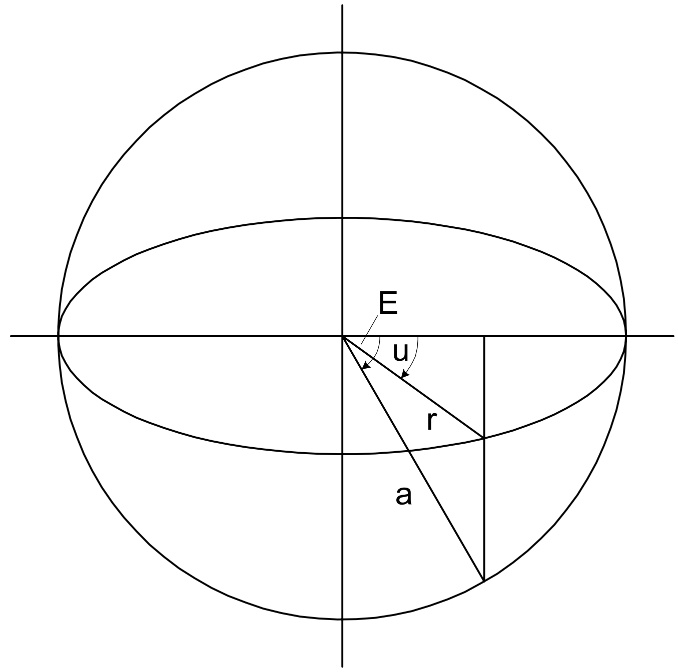

Figure 2 — Projection of the tilted parallax circle into an ellipse. The projection of the angle E is the ellipse angle u. The angle E is related to the solar longitude by $E = \frac{\pi}{2} - (\lambda_\star - \lambda_\odot)$.

$$E = \frac{\pi}{2} - (\lambda_\star - \lambda_\odot) \quad (16)$$

where $\lambda_\odot$ is the ecliptic longitude of the Sun. As the latitude of the star moves away from the pole, the circle becomes an ellipse, and the angle E projects onto the angle u, as illustrated. Hence, we have the relation

$$a\cos E = r\cos u \quad (17)$$

between E and u.[1] Plugging (15) into (17), we find

$$\left.\begin{aligned} \cos^2 u &= \frac{\cos^2 E}{1 - \cos^2\beta_\star \sin^2 E} \\ \sin^2 u &= \frac{\sin^2\beta_\star \sin^2 E}{1 - \cos^2\beta_\star \sin^2 E} \end{aligned}\right\} \quad (18)$$

By (18) and (16), (15) becomes the simple relation

$$r = a\sqrt{1 - \cos^2\beta_\star \cos^2(\lambda_\star - \lambda_\odot)} \quad (19)$$

This is the equation of the position of a star on the parallactic ellipse as a function of time and ecliptic latitude.

From the geometry of Figure 1, we can write the ecliptic coordinates of the star in terms of the astrometric parameters:

$$\begin{aligned} \lambda_\star &= \lambda_{ref} + \Delta\lambda_\star = \lambda_{ref} + \Delta\lambda_0 + (t - t_0)\mu_\lambda + r\cos u \\ \beta_\star &= \beta_{ref} + \Delta\beta_\star = \beta_{ref} + \Delta\beta_0 + (t - t_0)\mu_\beta - r\sin u \end{aligned} \quad (20)$$

2.4. First-Order Approximation of the Perpendicular Distance

Up to this point, the expressions describe the observation geometry exactly — no geometrical approximations have been made. Unfortunately, eq. (12) is nonlinear in the unknown parameters $\vec{p} \equiv [a, \Delta\lambda_0, \Delta\beta_0, \mu_\lambda, \mu_\beta]$, which would require the use

[1] The angle E is called the *eccentric anomaly* in the two-body problem of celestial mechanics. It is similarly used there to render a polar form of the equation of the two-body elliptical orbit into an analytically simpler shape. This then leads to Kepler's equation relating the linear mean motion to the nonlinear true anomaly.

of a nonlinear algorithm for parameter estimation. However, we can take advantage of the fact that the unknown parameters are all small (i.e., $p_i p_k \ll p_j$). Hence, we may easily calculate a first-order approximation to eq. (12) that is linear in the parameters, which then allows us to use much faster and simpler linear least squares for parameter estimation. Substitute (20) into (19) and, to first order, we find

$$r \simeq a\sqrt{1-\cos^2\beta_{ref}\cos^2(\lambda_{ref}-\lambda_\odot)} \tag{21}$$

which is conveniently linear in the unknown parameter a and independent of all the other unknown parameters. Using eqs. (18)-(20) in (12) and expanding again to first order, we obtain the useful result

$$\begin{aligned}|y| \;\simeq\; \big|\,\Delta S & \\ &- a\,[\sin(\lambda_{ref}-\lambda_\odot)\sin q \\ &\quad + \sin\beta_{ref}\cos(\lambda_{ref}-\lambda_\odot)\cos q] \\ &+ [\Delta\beta_0 + (t-t_0)\mu_\beta]\cos q \\ &- [\Delta\lambda_0 + (t-t_0)\mu_\lambda]\sin q\,\big|\end{aligned} \tag{22}$$

Eq. (22) gives the perpendicular distance as a function of time (both explicitly with the proper motion components and implicitly via the longitude of the Sun, $\lambda_\odot$), the high-precision timing observable ΔS, and the derived parameter q.

3. LINEAR LEAST SQUARES SOLUTION

Let the least squares cost function be the perpendicular distance as approximated by (22). Then

$$\begin{aligned}\chi^2 = \sum_{i=1}^{n} \; & w_i\{\Delta S_i + a\,R_i \\ & + [\Delta\beta_0 + (t_i-t_0)\mu_\beta]\cos q_i \\ & - [\Delta\lambda_0 + (t_i-t_0)\mu_\lambda]\sin q_i\}^2\end{aligned} \tag{23}$$

where

$$\begin{aligned}R_i \;\equiv\; & \sin[\lambda_{ref}-\lambda_\odot(t_i)]\sin q_i \\ & + \sin\beta_{ref}\cos[\lambda_{ref}-\lambda_\odot(t_i)]\cos q_i\end{aligned} \tag{24}$$

the $w_i = \frac{1}{\sigma_i^2}$ are the individual observation weights, and the q_i are calculated from eqs. (1)-(3). The normal equations follow from determining the stationary points of χ^2 in the parameter space,

$$\frac{\partial\chi^2}{\partial\vec{p}} = 0 \tag{25}$$

where $\vec{p} \equiv [a, \Delta\lambda_0, \Delta\beta_0, \mu_\lambda, \mu_\beta]$ are the unknown parameters.

Note that eq. (23) is linear in the unknown parameters. We may therefore employ linear least squares for our parameter estimation. Taking the derivatives of (23), we find that eq. (25) can be written

$$\boldsymbol{M}\cdot\vec{p} = \vec{b} \tag{26}$$

where

$$\boldsymbol{M} = \begin{bmatrix} \sum_{i=1}^n w_i R_i^2 & \sum_{i=1}^n w_i R_i \sin q_i & -\sum_{i=1}^n w_i R_i \cos q_i & \sum_{i=1}^n w_i (t_i-t_0) R_i \sin q_i & -\sum_{i=1}^n w_i (t_i-t_0) R_i \cos q_i \\ \sum_{i=1}^n w_i R_i \sin q_i & \sum_{i=1}^n w_i \sin^2 q_i & -\sum_{i=1}^n w_i \sin q_i \cos q_i & \sum_{i=1}^n w_i (t_i-t_0)\sin^2 q_i & -\sum_{i=1}^n w_i (t_i-t_0)\sin q_i\cos q_i \\ -\sum_{i=1}^n w_i R_i \cos q_i & -\sum_{i=1}^n w_i \sin q_i \cos q_i & \sum_{i=1}^n w_i \cos^2 q_i & -\sum_{i=1}^n w_i (t_i-t_0)\sin q_i\cos q_i & \sum_{i=1}^n w_i (t_i-t_0)\cos^2 q_i \\ \sum_{i=1}^n w_i (t_i-t_0) R_i \sin q_i & \sum_{i=1}^n w_i (t_i-t_0)\sin^2 q_i & -\sum_{i=1}^n w_i (t_i-t_0)\sin q_i\cos q_i & \sum_{i=1}^n w_i (t_i-t_0)^2\sin^2 q_i & -\sum_{i=1}^n w_i (t_i-t_0)^2\sin q_i\cos q_i \\ -\sum_{i=1}^n w_i (t_i-t_0) R_i \cos q_i & -\sum_{i=1}^n w_i (t_i-t_0)\sin q_i\cos q_i & \sum_{i=1}^n w_i (t_i-t_0)\cos^2 q_i & -\sum_{i=1}^n w_i (t_i-t_0)^2\sin q_i\cos q_i & \sum_{i=1}^n w_i (t_i-t_0)^2\cos^2 q_i \end{bmatrix} \tag{27}$$

and

$$\vec{b} = \begin{bmatrix} -\sum_{i=1}^n w_i\,\Delta S_i R_i \\ \sum_{i=1}^n w_i\,\Delta S_i \sin q_i \\ -\sum_{i=1}^n w_i\,\Delta S_i \cos q_i \\ \sum_{i=1}^n w_i\,(t_i-t_0)\,\Delta S_i \sin q_i \\ -\sum_{i=1}^n w_i\,(t_i-t_0)\,\Delta S_i \cos q_i \end{bmatrix} \tag{28}$$

The solution for the unknown parameters is then

$$\vec{p} = \boldsymbol{C}\cdot\vec{b} \tag{29}$$

where $\boldsymbol{C} \equiv \boldsymbol{M}^{-1}$ is the *covariance matrix*. A straightforward application of propagation of errors (e.g., Press et al., 1992) yields the formal parameter errors from the diagonal components of the covariance matrix:

$$\sigma_k = \sqrt{\boldsymbol{C}_{kk}}\,\sigma_{\Delta S} \tag{30}$$

for parameter p_k, and where $\sigma_{\Delta S}$ is the standard error of the timing observations ΔS. The formal parameter cross-correlations are contained in the off-diagonal components of the covariance matrix,

$$\sigma_{jk} = \sqrt{\boldsymbol{C}_{jk}}\,\sigma_{\Delta S}, \qquad j \neq k \tag{31}$$

The standard error $\sigma_{\Delta S}$ is assumed for simplicity to be the same for all observations (i.e., $\sigma_i \equiv \sigma_{\Delta S}$), so that it is a common factor that can be pulled out of $\boldsymbol{C}$ in eqs. (30) and (31). This need not, of course, be the case in actual practice.

It is worth noting that the parameter standard errors σ_k are only functions of time t, ecliptic coordinates $(\lambda_{ref}, \beta_{ref})$, and precession cone angle ψ. They are not functions of the parameters $\vec{p}$ themselves. Therefore, we don't need to know $\vec{p}$ in order to determine $\sigma_{\vec{p}}$, which is a useful general property of linear least squares parameter estimation.

This approach has been implemented in certain FAME simulations (Murison, 2000, and successive papers) and found to work quite well.

The author thanks George Kaplan and Rob Olling for helpful discussions in the early stages of this work.

REFERENCES

Murison, M.A. (2000). "The Observation Density and Scan Angle Distributions of a Spin-Stabilized, Precessing, Scanning Satellite Instrument: I. Geometry", Astronomical Applications Department Technical Note AA2000-06, http://aa.usno.navy.mil/AA/TechNotes/. Also available in the FAME technical memorandum series, FTM2000-16.

Press et al. (1992). *Numerical Recipes*, 2nd edition, Cambridge University Press.