

Spin-Axis Attitude Estimation

Sergei Tanygin¹ and Malcolm D. Shuster²

Du spinnst doch!
Du hast doch'n Knall!
Du spinnst doch!
Du spinnst doch total!³

Die Prinzen (1993)

Abstract

Spin-axis attitude estimation is examined in a manner analogous to the study of three-axis attitude estimation. Measurement modeling issues are given careful consideration, as are those of representation, frame, and constraint. Three approaches to spin-axis attitude estimation are presented and compared numerically. A thorough covariance analysis of all algorithms is performed.

Introduction

Spin-Axis attitude estimation⁴ receives much less attention today than it did in the past. If we examine the now venerable *Spacecraft Attitude Determination and Control* [1], edited by James R. Wertz in 1978, which was an excellent summary of

¹Senior Astrodynamics Specialist, Analytical Graphics, Inc., 220 Valley Creek Blvd., Exton, Pennsylvania 19341. E-mail: stanygin@agi.com.

²Director of Research, Acme Spacecraft Company, 13017 Wisteria Drive, Box 328, Germantown, Maryland 20874. E-mail: mdshuster@comcast.net, website: <http://home.comcast.net/~mdshuster>.

³You're like really so spinning (*lit. crazy*)! / You're like really so whacked! / You're like really so spinning! / You're really spinning to the max! (based on an ancient ballad, modern free translation into Valspeak by the second author). Before becoming a leading European rock band, the "princes" were all members of the choir of the Thomaskirche in Leipzig, whose most illustrious choirmaster and organist was Johann Sebastian Bach.

⁴We use the terms *estimation* to refer to the mathematical process of calculating an estimate from data and *determination* to refer to the entire process of determining the estimate including sensor performance, telemetry encoding and transmission, software, etc. Thus, for us, *attitude estimation* is a subprocess of *attitude determination*. Likewise, the adjustment of sensor data to correct for known sensor and environmental effects we call *compensation*, and the estimation of the compensation parameters we call *calibration* or simply *bias determination* when the parameters are limited to biases. Although we write *spin-axis attitude*, we really mean *single-axis attitude*. The spacecraft may be stabilized about three axes.

practice for near-Earth spacecraft during the 1970s and has been the *vademecum* of countless newcomers to spacecraft attitude estimation, we will find among its nearly 900 pages hundreds of pages devoted to spin-axis attitude estimation but fewer than twenty pages devoted to three-axis attitude estimation, and this includes the ten pages devoted to the attitude representations. This great disparity of emphasis, of course, reflected the relative emphasis on spin-stabilized and three-axis-stabilized spacecraft in that decade, at least at NASA Goddard Space Flight Center. Today, the emphasis is reversed, and one sees papers at conferences on spin-axis attitude estimation with far less frequency than one sees papers on three-axis attitude estimation.

In 1983, when reference [2] was published, the trend toward more three-axis-stabilized spacecraft and three-axis attitude determination systems was already well underway. The principal contribution of reference [2] to spin-axis attitude estimation was to bring to it some of the tools of three-axis attitude estimation, in particular, a greater reliance on vector methods rather than on spherical trigonometry. The latter, of course, can never be eliminated completely, at least not if one includes Earth horizon scanners in the attitude sensor suite.

The two estimation domains, spin-axis and three-axis attitude estimation, can never be made completely similar because of basic differences between the two. For example, the typical measurements for three-axis attitude estimation are directions, i.e., unit vectors, while those for spin-axis attitude estimation are usually angles (or, equivalently, cosines). At an even more basic level, rotations form a group, while directions in three dimensions certainly do not. Nonetheless, much of the inner workings of the algorithms could be made very similar, and this approach led to more efficient algorithms for spin-axis attitude estimation than had hitherto been available. These were applied first to the AMPTE Mission [3], launched in 1984. This is not to say that earlier work on spin-axis attitude estimation was suddenly swept aside, nothing of the sort. That work [1, 4–8] had been very fruitful and remains very much in use today. But the final computation step was certainly much improved by the new approach, as was the quantitative study of the contribution of various error processes. Like the 1983 paper, the present work concentrates entirely on the problem of spin-axis attitude estimation alone, and does not address the associated problem of calibration and compensation, which have been important topics from the beginning. As in basic studies of three-axis attitude estimation, we assume, effectively, that compensation of the sensor data has already been carried out.

There are more similarities connecting spin-axis attitude estimation and three-axis attitude estimation than there are differences separating them. The 1983 paper [2] on the former already contains material similar to that in a 2002 paper on constraint in (three-axis) attitude estimation [9], which, in turn, has informed the present article. The time is ripe to reexamine spin-axis attitude estimation from the broader perspective and deeper understanding gained from the past two decades of three-axis attitude estimation, whose level of sophistication and attention to statistics [9, 10] has grown considerably during that period, while the methods of spin-axis attitude estimation have been largely static. What will arise from this reexamination, we hope, will be a greater appreciation of the nature of spin-axis attitude estimation within the general framework of spacecraft attitude estimation and greater assurance in constructing new spin-axis attitude estimation algorithms.

The previous work [2] was devoted almost entirely to single-frame estimation. The present work, which began as Part II of a sequence beginning with reference [2],

published 24 years earlier, is devoted entirely to batch spin-axis attitude estimation, which occupied only one page of reference [2]. Thus, the present work and reference [2] are complementary. On the other hand, the single-frame algorithms are only a special case of the batch algorithms, so that reference [2] is almost entirely absorbed in the present work. Not absorbed are the review of research published before 1983, the explicit two- and three-measurement algorithms, and the derivation of the nadir-angle measurement (which can be found also in reference [1]). Reference [2] remains important background for the present work, which, however, is largely self-contained. For the most part, the notation of this work follows that of reference [11].

Other Work

The significant literature before 1983 was reviewed in reference [2]. The methodology of reference [2], as we have said, was applied to the AMPTE spacecraft, launched in 1984 [3] and, together with reference [1], has been the starting point for many later studies. Van der Ha [12] has carried out a careful covariance analysis of spin-axis attitude estimation taking into account also the effect of uncertainty in the biases. Emara-Shabaik [13] has carried out similar calculations of the spin-axis attitude and biases. Fraiture [14] has developed a sampling method for spin-axis and bias estimation which greatly simplifies the estimation of the latter. His publication is noteworthy also for supplying an informative historical perspective on the development of spin-axis attitude estimation methodology at ESA. Van der Ha [15] has offered a deterministic approach, the equal-chord method, to the problem of spin-axis attitude estimation. His earlier work [12] has also received a much more detailed treatment recently [16]. Recently, he presented a methodology for estimating spin-axis attitude from Sun-aspect data alone [17].

The present work, like reference [2], has considered only the estimation problem. Our focus has been largely on the geometrical aspects of spin-axis attitude estimation. With the improvement of Earth-horizon scanners over the past two decades, the contribution of biases to spin-axis attitude errors has been much reduced, except for systematic errors in the direction of the spin-axis with respect to the spacecraft body frame (dynamic imbalance) and radiance corrections. Radiative corrections have been studied for several decades. An extensive study has been carried out by Phenneger et al. [18]. Van der Ha et al. have given more recent values for all corrections in their study of the CONTOUR spacecraft [19].

Spin-Axis Attitude Measurements and Models⁵

Z-Axis Magnetometer

The simplest spin-axis attitude measurement is the magnetic-field component along the spacecraft spin axis, which we take to be the spacecraft body z-axis. Thus,

$$z_B = \frac{B_z^m}{|\mathbf{B}|} = \hat{\mathbf{n}} \cdot \hat{\mathbf{B}} + v_B \quad (1)$$

and we assume

$$v_B \sim \mathcal{N}(0, \sigma_B^2) \quad (2)$$

⁵For the most part, we simply summarize in this section the needed material from reference [2].

where $|\mathbf{B}|$ is the magnitude of the magnetic field, obtainable from a field model given the position of the spacecraft with respect to an Earth-fixed coordinate system, v_B is the sensor noise, assumed to be Gaussian and zero-mean with angle-equivalent standard deviation σ_B of about 0.5 deg for magnetic latitudes smaller than 70 deg in magnitude. The spin-axis attitude is represented by the unit vector $\hat{\mathbf{n}}$, the direction of the spacecraft spin axis. The superscript m will generally denote a measured quantity.

Single-Axis Sun Sensor

A single-axis Sun sensor, sometimes called a Sun aspect sensor, simply measures the angle between the Sun direction and the spin-axis, and consists essentially of a “protractor,” which measures the crossing time t_S and the Sun angle β when the Sun line is in the “protractor” plane. Thus,

$$z_S = \cos \beta^m = \hat{\mathbf{n}} \cdot \hat{\mathbf{S}} + v_S \quad (3)$$

with

$$v_S \sim \mathcal{N}(0, \sigma_S^2) \quad (4)$$

and typically in noise equivalent angle $\sigma_S \approx 0.5$ deg for digital spinning Sun-aspect sensors. For an analogue (V-slit) Sun aspect sensor, the accuracies can be higher.

Earth Horizon Scanner

The Earth horizon scanner begins by measuring a dihedral angle, the Earth width, which is given by

$$\Omega^m = \omega(t_{\text{LOS}}^m - t_{\text{AOS}}^m) \quad (5)$$

where t_{AOS} and t_{LOS} are the respective times at which the Earth is first detected (arrival of signal) and at which the Earth signal is lost by the scanner (loss of signal),⁶ ω is the spacecraft spin rate, and Ω is the Earth width. The spin rate ω is also a measured quantity determined, for example, from successive Sun crossings. We shall assume, however, that it is essentially constant and determined from very many Sun crossings, so that its measurement error can be neglected compared to those in t_{AOS} and t_{LOS} . This dihedral angle is converted to an arc length, the nadir angle η , according to references [1] and [2]

$$\cos \eta \equiv \hat{\mathbf{n}} \cdot \hat{\mathbf{E}} = A^{-1}[\cos \rho \cos \gamma \pm \sin \gamma \cos(\Omega/2)(A - \cos^2 \rho)^{1/2}] \quad (6a)$$

with

$$A = \cos^2 \gamma + \sin^2 \gamma \cos^2(\Omega/2) \quad (6b)$$

Here, ρ is the angular radius of the Earth as seen from the spacecraft, and γ is the scanner half-cone angle [2]. The sign ambiguity in equation (6a) must be removed by reference to other measurements, such as the Earth-Sun dihedral angle below.⁷ For example, only one of the two possible values of $\cos \eta$ is consistent, in general, with both equations (6) and equation (14) below. In those special cases where both

⁶These are sometimes denoted by *S/E* (space-to-Earth) and *E/S* (Earth-to-space), respectively.

⁷In single-frame spin-axis attitude estimation. In batch estimation the ambiguity is eliminated by the data in other frames. Van der Ha et al. prefer to use a differential-corrector method [19].

signs are possible in equation (6a), the two solutions turn out to be equivalent. The Earth measurement z_E is then

$$z_E = \cos \eta^m = \hat{\mathbf{n}} \cdot \hat{\mathbf{E}} + v_E \quad (7)$$

Because of equations (6), the effective Earth measurement noise has a complicated dependence on sensor parameters and altitude. We have from simple sensitivity analysis⁸

$$v_E \approx \frac{\omega}{2} \frac{\sin \eta}{\cot \gamma - \cot(\Omega/2) \cot \eta} (\sin(\Omega/2)) (\Delta t_{\text{LOS}} - \Delta t_{\text{AOS}}) \quad (8)$$

and we assume

$$\Delta t_{\text{AOS}} \sim \mathcal{N}(0, \sigma_{\text{AOS}}^2) \quad \text{and} \quad \Delta t_{\text{LOS}} \sim \mathcal{N}(0, \sigma_{\text{LOS}}^2) \quad (9\text{ab})$$

The error levels σ_{AOS} and σ_{LOS} are sensitive to the bandwidth and the triggering circuitry of the Earth horizon scanner. In low-Earth orbit, the standard deviation of v_E is typically in the angle-equivalent range from 0.1 to 0.5 deg.⁹

Earth-Sun Dihedral Angle

We define the Earth-Sun dihedral angle ψ_{ES} to be the dihedral angle from the plane of the Sun direction and the spacecraft spin axis to the plane of the nadir vector and the spacecraft spin axis, measured in the same sense as the spacecraft spin. Thus,

$$\psi_{\text{ES}} = \omega(t_S - (t_{\text{LOS}} + t_{\text{AOS}})/2) \quad (10)$$

and we assume, as usual,

$$\Delta t_S \sim \mathcal{N}(0, \sigma_{t_S}^2) \quad (11)$$

From the definition, it follows that [2]

$$\sin \psi_{\text{ES}} = \frac{\hat{\mathbf{n}} \cdot (\hat{\mathbf{S}} \times \hat{\mathbf{E}})}{|\hat{\mathbf{n}} \times \hat{\mathbf{S}}| |\hat{\mathbf{n}} \times \hat{\mathbf{E}}|} \quad (12)$$

The dependence of ψ_{ES} on $\hat{\mathbf{n}}$ is complicated by the denominators. However,

$$|\hat{\mathbf{n}} \times \hat{\mathbf{S}}| = \sqrt{1 - z_S^2} \quad \text{and} \quad |\hat{\mathbf{n}} \times \hat{\mathbf{E}}| = \sqrt{1 - z_E^2} \quad (13\text{ab})$$

so that

$$\begin{aligned} z_{\text{ES}} &\equiv \sqrt{1 - z_S^2} \sqrt{1 - z_E^2} \sin \omega(t_S - (t_{\text{LOS}} + t_{\text{AOS}})/2) \\ &= \hat{\mathbf{n}} \cdot (\hat{\mathbf{S}} \times \hat{\mathbf{E}}) + v_{\text{ES}} \end{aligned} \quad (14)$$

The Earth-Sun dihedral angle measurement is now linear in $\hat{\mathbf{n}}$ (and now, effectively, an arc-length measurement), but the effective measurement noise will now contain

⁸Note an error in the denominator of equation (52) of reference [2], which should have been identical to equation (8) above. The authors are grateful to Professor Jozef C. van der Ha for pointing out the error in reference [2], which had reappeared in a preliminary version of the present work. Van der Ha also has a derivation of equations (6) in reference [15] which is different in method and form from that in reference [2], but obviously mathematically equivalent. His example in reference [12] parallels our equations (6).

⁹These error levels are the case for Earth-horizon scanners which are sensitive only to a very narrow band of the infra-red spectrum and, as a result, are not sensitive to ‘‘cold-cloud’’ effects. In Earth-horizon scanners of the 1980s and before, in which the sensitive infra-red band was a much larger 14–16 microns, cold-cloud effects could introduce errors of 2.0 deg. Earth radiance corrections have also been reduced (or, at least, much simplified) by the inclusion of an Earth albedo sensor as part of the Earth-horizon sensor.

terms arising from v_S , v_E , Δt_S , Δt_{AOS} and Δt_{LOS} . That price is fair, however, since there is much more to be gained from having a simpler sensitivity matrix than from having a simpler covariance matrix, which, in practice, we will tend to approximate rather crudely in any event. From equation (14),

$$z_{ES} = \sin \beta^m \sin \eta^m \sin \psi_{ES}^m \quad (15)$$

from which

$$\begin{aligned} v_{ES} &= z_{ES}^{\text{true}} (\cot \beta \Delta \beta + \cot \eta \Delta \eta + \cot \psi_{ES} \Delta \psi_{ES}) \\ &= (\hat{\mathbf{n}} \cdot (\hat{\mathbf{S}} \times \hat{\mathbf{E}}))^{\text{true}} \left[-\frac{\cos \beta}{\sin^2 \beta} v_S - \frac{\cos \eta}{\sin^2 \eta} v_E + \cot \psi_{ES} \Delta \psi_{ES} \right] \\ &\sim \mathcal{N}(0, \sigma_{ES}^2) \end{aligned} \quad (16)$$

The error in the Earth-Sun dihedral $\Delta \psi_{ES}$ is a simple function of the errors in the crossing times.

The computation of σ_{ES}^2 is straightforward but unenlightening. Note that v_E and $\Delta \psi_{ES}$ will be correlated unless $\sigma_{AOS}^2 = \sigma_{LOS}^2$. In a similar way we can define z_{BE} and z_{BS} . Note, however, that only two of these three dihedral angles can be independent measurements.

If we write our concatenated measurement for frame k as, for example,

$$\begin{aligned} \mathbf{Z}_k &= [z_{B,k} \quad z_{S,k} \quad z_{E,k} \quad z_{ES,k}]^T \\ &= [\hat{\mathbf{B}}_k \quad \hat{\mathbf{S}}_k \quad \hat{\mathbf{E}}_k \quad (\hat{\mathbf{S}}_k \times \hat{\mathbf{E}}_k)]^T \hat{\mathbf{n}} + [v_{B,k} \quad v_{S,k} \quad v_{E,k} \quad v_{ES,k}]^T \\ &\equiv H_k \hat{\mathbf{n}} + \mathbf{v}_k \end{aligned} \quad (17)$$

where T denotes the matrix transpose, then the 4×4 measurement covariance matrix will be correlated as

$$R_k = \begin{bmatrix} \times & 0 & 0 & 0 \\ 0 & \times & 0 & \times \\ 0 & 0 & \times & \times \\ 0 & \times & \times & \times \end{bmatrix} \quad (18)$$

where “ \times ” shows the location of possible nonzero entries. Again, the computation of R_k from the information provided above is straightforward but unenlightening. There are, of course, many choices for the effective measurements depending on the sensor suite and the choice of dihedral angles.¹⁰

The Cost Function¹¹

In the previous section, we have created an effective measurement vector which is linear in the spin-axis vector and satisfies

$$\mathbf{Z}_k = H_k \hat{\mathbf{n}} + \mathbf{v}_k, \quad k = 1, \dots, N \quad (19)$$

with measurement noise

$$\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, R_k), \quad k = 1, \dots, N \quad (20)$$

¹⁰While our model for the sensitivity matrix H_k is unassailable, that for the measurement variances is very crude. The equations for the measurement noise terms point the way to a more realistic model. In actual mission practice, however, one has seldom gone beyond the assumption of constant measurement variances.

¹¹It is here that the new material of this work begins.

The measurement noise is generally assumed also to be white. From these forms we may construct our estimators.

The maximum-likelihood estimate [10] for such a set of measurements is simply

$$\hat{\mathbf{n}}^* \equiv \arg \min_{|\hat{\mathbf{n}}|=1} J(\hat{\mathbf{n}}) \quad (21)$$

that is, the value of the unit vector $\hat{\mathbf{n}}$ for which $J(\hat{\mathbf{n}})$ is a minimum, and $J(\hat{\mathbf{n}})$ is the data-dependent part of the negative-log-likelihood function [10, 20], namely,

$$J(\hat{\mathbf{n}}) = \frac{1}{2} \sum_{k=1}^N (\mathbf{Z}_k - H_k \hat{\mathbf{n}})^T R_k^{-1} (\mathbf{Z}_k - H_k \hat{\mathbf{n}}) \quad (22)$$

The minimization of $J(\hat{\mathbf{n}})$ directly in terms of $\hat{\mathbf{n}}$ is complicated by the norm constraint

$$\hat{\mathbf{n}}^T \hat{\mathbf{n}} = 1 \quad (23)$$

the treatment of which is the main feature of this paper.

$J(\hat{\mathbf{n}})$ is a quadratic function of $\hat{\mathbf{n}}$. Therefore, it may be written as

$$J(\hat{\mathbf{n}}) = \mathbf{J} + \mathbf{G}^T \hat{\mathbf{n}} + \frac{1}{2} \hat{\mathbf{n}}^T \mathbf{F} \hat{\mathbf{n}} \quad (24)$$

with

$$\mathbf{J} \equiv \frac{1}{2} \sum_{k=1}^N \mathbf{Z}_k^T R_k^{-1} \mathbf{Z}_k \quad (25a)$$

$$\mathbf{G} \equiv - \sum_{k=1}^N H_k^T R_k^{-1} \mathbf{Z}_k \quad (25b)$$

$$\mathbf{F} \equiv \sum_{k=1}^N H_k^T R_k^{-1} H_k \quad (25c)$$

The scalar \mathbf{J} is just $J(\mathbf{0})$, that is, $J(\hat{\mathbf{n}})$ evaluated at zero argument, a physically impossible value for the spin-axis vector because of the norm constraint. The column array \mathbf{G} is the gradient of $J(\hat{\mathbf{n}})$ at zero argument, taking the three components of $\hat{\mathbf{n}}$ as independent. Lastly, the symmetric matrix \mathbf{F} is the Hessian matrix of $J(\hat{\mathbf{n}})$, again taking the three components of $\hat{\mathbf{n}}$ as independent. These three quantities are clearly unphysical individually, but also very useful, since, once computed, one need never deal with the individual measurements again.

The minimization of the cost function of equation (22) is complicated by the presence of the constraint, equation (23). There are two approaches to obtaining the minimizing value of $\hat{\mathbf{n}}$. First, one can minimize $J(\hat{\mathbf{n}})$ explicitly over the three-dimensional $\hat{\mathbf{n}}$ by using Lagrange's method of multipliers or, secondly, one can replace the three-dimensional argument of J by a two-dimensional variable, so that the constraint is satisfied automatically. We consider both approaches in the present work.

We will assume for the first part of this work that \mathbf{F} is nonsingular. Singular cases will result when the vectors implicit in the individual cosine measurements (limited to a subset of $\hat{\mathbf{B}}_k, \hat{\mathbf{E}}_k, \hat{\mathbf{S}}_k, k = 1, \dots, N$, generally without measurements of dihedral angles) are coplanar. Such a situation can also occur by accident due to delicate cancellations in the construction of \mathbf{F} , but the likelihood of that occurring is too minute to merit serious consideration. We treat the singular cases explicitly near the end of this article, in which we present additional algorithms applicable specifically to the singular cases.

Method 1: The Lagrange-Multiplier Method

Using Lagrange's method of multipliers, which was also employed in the non-Euclidean single-frame context of reference [2], we write

$$J'(\hat{\mathbf{n}}) = J(\hat{\mathbf{n}}) + \frac{1}{2}\lambda(\hat{\mathbf{n}}^T\hat{\mathbf{n}} - 1) \quad (26)$$

with λ a yet unspecified constant, and minimize $J'(\hat{\mathbf{n}})$ without constraint. This leads directly to

$$\frac{\partial J'}{\partial \hat{\mathbf{n}}}(\hat{\mathbf{n}}^*) = \mathbf{G} + (\mathbf{F} + \lambda I_{3 \times 3})\hat{\mathbf{n}}^* = \mathbf{0} \quad (27)$$

or

$$\hat{\mathbf{n}}^* = -(\mathbf{F} + \lambda I_{3 \times 3})^{-1}\mathbf{G} \quad (28)$$

and the satisfaction of equation (23) now requires

$$\hat{\mathbf{n}}^* \cdot \hat{\mathbf{n}}^* = \mathbf{G}^T(\mathbf{F} + \lambda I_{3 \times 3})^{-2}\mathbf{G} = 1 \quad (29)$$

which may be solved for λ by the Newton-Raphson method, namely,

$$\lambda_o = 0 \quad (30a)$$

$$D_{i-1} = (\mathbf{F} + \lambda_{i-1} I_{3 \times 3})^{-1} \quad (30b)$$

$$\mathbf{n}_{i-1} = -D_{i-1}\mathbf{G} \quad (30c)$$

$$\lambda_i = \lambda_{i-1} - \frac{1}{2} \frac{(1 - \mathbf{n}_{i-1}^T \mathbf{n}_{i-1})}{\mathbf{n}_{i-1}^T D_{i-1} \mathbf{n}_{i-1}} \quad (30d)$$

and

$$\lambda = \lim_{i \rightarrow \infty} \lambda_i, \quad \hat{\mathbf{n}}^* = \lim_{i \rightarrow \infty} \mathbf{n}_i \quad (31ab)$$

This approach, first applied in a somewhat different form and context in reference [2], is also to be found *mutatis mutandis* in Part II of reference [9] for three-axis attitude estimation.

Because equation (29) is nonlinear, it may have many solutions for λ . In fact, it can have as many as six solutions, but only one is physical. The nature of the Lagrange multiplier is examined in Appendix A.

The cost function $J(\mathbf{n})$ as a function on \mathcal{R}^3 is differentiable everywhere and bounded below by zero. Therefore, it has a unique minimum provided \mathbf{F} is positive definite. In the case of vanishing measurement noise ($\mathbf{G} = \mathbf{G}^{\text{true}}$) this minimum must be the true value of the spin-axis attitude¹²

$$\hat{\mathbf{n}}^{\text{true}} = -\mathbf{F}^{-1}\mathbf{G}^{\text{true}} \quad (32)$$

which must also satisfy the unit-norm constraint, equation (23), and, therefore,¹³

$$\lambda^{\text{true}} = 0 \quad (33)$$

¹²Note that we do not allow the case $\mathbf{G}^{\text{true}} = \mathbf{0}$, since this would result in an unphysical solution and is therefore impossible.

¹³By λ^{true} we mean the value for noise-free measurements, i.e., at the *true* value of the measurements.

Since S^2 , the unit sphere in \mathcal{R}^3 , is a compact subspace of \mathcal{R}^3 , it follows in the general case that $J(\hat{\mathbf{n}})$ must also attain its respective minimum on S^2 , and therefore, $\mathbf{F} + \lambda I_{3 \times 3}$ cannot be singular. Since this quantity is positive definite for vanishing measurement noise and it must be a continuous function of the measurement noise (that is, of $\mathbf{G} - \mathbf{G}^{\text{true}}$) it cannot become singular for some value of \mathbf{G} infinitesimally different from \mathbf{G}^{true} . If it could, then clearly the value of \mathbf{G} would be unphysical.¹⁴ Therefore, we must have that $\mathbf{F} + \lambda I_{3 \times 3}$ is always positive-definite for \mathbf{F} positive-definite.

Covariance Matrix for the Lagrange-Multiplier Method

To calculate the covariance matrix of the Lagrange-method estimate we note from equation (28) that the error in $\hat{\mathbf{n}}^*$ is given to first order in the measurement noise by

$$\begin{aligned} \Delta \hat{\mathbf{n}}^* &= -(\mathbf{F} + \lambda^{\text{true}} I_{3 \times 3})^{-1} \Delta \mathbf{G} \\ &\quad + (\Delta \lambda) (\mathbf{F} + \lambda^{\text{true}} I_{3 \times 3})^{-2} \mathbf{G}^{\text{true}} \end{aligned} \quad (34)$$

with

$$\Delta \mathbf{G} = - \sum_{k=1}^N H_k^T R_k^{-1} \mathbf{v}_k \quad (35)$$

and we have noted that

$$\Delta(\mathbf{F} + \lambda I_{3 \times 3})^{-1} = -(\Delta \lambda) (\mathbf{F} + \lambda^{\text{true}} I_{3 \times 3})^{-2} \quad (36)$$

to first order in $\Delta \lambda$. Note also that

$$E\{\Delta \mathbf{G}\} = \mathbf{0} \quad \text{and} \quad \text{Covar}\{\Delta \mathbf{G}\} = \mathbf{F} \quad (37\text{ab})$$

where $E\{\cdot\}$ denotes the expectation, $\text{Covar}\{\cdot\}$ the covariance matrix, and $\text{Var}\{\cdot\}$ (below) the variance.

It follows to first order that

$$\Delta \hat{\mathbf{n}}^* = -\mathbf{F}^{-1} \Delta \mathbf{G} + (\Delta \lambda) \mathbf{F}^{-1} \hat{\mathbf{n}}^{\text{true}} \quad (38)$$

From equation (29) we obtain likewise to first order

$$\hat{\mathbf{n}}^{\text{true}T} \mathbf{F}^{-1} \Delta \mathbf{G} - (\hat{\mathbf{n}}^{\text{true}T} \mathbf{F}^{-1} \hat{\mathbf{n}}^{\text{true}}) \Delta \lambda = 0 \quad (39)$$

from which, trivially,

$$\Delta \lambda = (\hat{\mathbf{n}}^{\text{true}T} \mathbf{F}^{-1} \hat{\mathbf{n}}^{\text{true}})^{-1} \hat{\mathbf{n}}^{\text{true}T} \mathbf{F}^{-1} \Delta \mathbf{G} \quad (40)$$

and, therefore,

$$E\{\Delta \lambda\} = 0 \quad \text{and} \quad \text{Var}\{\Delta \lambda\} = (\hat{\mathbf{n}}^{\text{true}T} \mathbf{F}^{-1} \hat{\mathbf{n}}^{\text{true}})^{-1} \quad (41\text{ab})$$

and also

$$E\{(\Delta \lambda) \Delta \mathbf{G}\} = (\hat{\mathbf{n}}^{\text{true}T} \mathbf{F}^{-1} \hat{\mathbf{n}}^{\text{true}})^{-1} \hat{\mathbf{n}}^{\text{true}} \quad (42)$$

We see from equation (41b) that although \mathbf{F} is of order N/σ^2 , where N is the number of measurements, and σ is the typical measurement error level, λ^{rms} is only of order \sqrt{N}/σ^2 . Thus, while large in absolute terms, λ^{rms} is generally minute in comparison with the scale of \mathbf{F} .

¹⁴For example, $\mathbf{G} = \mathbf{0}$ is unphysical for \mathbf{F} positive-definite.

The calculation of the spin-axis attitude covariance matrix is now straightforward with the result

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} = \Lambda \mathbf{F}^{-1} \Lambda^T \quad (43)$$

where

$$\Lambda = I - (\hat{\mathbf{n}}^{\text{true}T} \mathbf{F}^{-1} \hat{\mathbf{n}}^{\text{true}})^{-1} \mathbf{F}^{-1} \hat{\mathbf{n}}^{\text{true}} \hat{\mathbf{n}}^{\text{true}T} \quad (44)$$

It is easily verified to within errors of third order in the measurement noise that¹⁵

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}\hat{\mathbf{n}}^{\text{true}} = \mathbf{0} \quad (45)$$

Method 2: Unconstrained Brute-Force Method

Although the spin-axis attitude estimate, as defined by equation (21), must have unit norm, we can define, nonetheless, an unconstrained estimate \mathbf{n}_{uc}^* as the value which minimizes $J(\mathbf{n})$ without constraint. This leads immediately to

$$\frac{\partial J}{\partial \mathbf{n}}(\mathbf{n}_{\text{uc}}^*) = \mathbf{G} + \mathbf{F}\mathbf{n}_{\text{uc}}^* = \mathbf{0} \quad (46)$$

or, if \mathbf{F} is invertible,¹⁶

$$\mathbf{n}_{\text{uc}}^* = -\mathbf{F}^{-1} \mathbf{G} \quad (47)$$

If there were no measurement noise, then \mathbf{n}_{uc}^* as given by equation (47) would have unit norm and be the exact spin-axis vector, as noted in equation (32). Thus, \mathbf{n}_{uc}^* must be close to the desired $\hat{\mathbf{n}}^*$. We can write, certainly,

$$\mathbf{n}_{\text{uc}}^* = \hat{\mathbf{n}} + \Delta \mathbf{n}_{\text{uc}}^* \quad (48)$$

We do not attach physical significance to \mathbf{n}_{uc}^* , because, not necessarily having unit norm, it cannot be a direction. However, because the measurement noise terms are Gaussian, independent, and zero-mean, so approximately is the “estimate error” $\Delta \mathbf{n}_{\text{uc}}^*$. In fact, we must have that the error covariance matrix associated with \mathbf{n}_{uc}^* satisfies

$$P_{\mathbf{nn}}^{\text{uc}} \equiv E\{\Delta \mathbf{n}_{\text{uc}}^* \Delta \mathbf{n}_{\text{uc}}^{*\text{T}}\} = \mathbf{F}^{-1} \equiv \mathbf{P} \quad (49)$$

provided that \mathbf{F} is invertible. The unconstrained estimate \mathbf{n}_{uc}^* will be a sufficient statistic [10] for $\hat{\mathbf{n}}$, so that the entire content of equation (22) can be recapitulated as

$$J_{\text{eff}}(\hat{\mathbf{n}}) = \frac{1}{2} (\mathbf{n}_{\text{uc}}^* - \hat{\mathbf{n}})^T (P_{\mathbf{nn}}^{\text{uc}})^{-1} (\mathbf{n}_{\text{uc}}^* - \hat{\mathbf{n}}) \quad (50)$$

$J_{\text{eff}}(\hat{\mathbf{n}})$ can differ from the expression of equation (22) only by a constant. In fact, by a rearrangement of the terms of equation (22) we can write

$$J(\hat{\mathbf{n}}) = J_{\text{eff}}(\hat{\mathbf{n}}) + \mathbf{J} - \frac{1}{2} \mathbf{n}_{\text{uc}}^{*\text{T}} \mathbf{F} \mathbf{n}_{\text{uc}}^* \quad (51)$$

Thus, \mathbf{J} , \mathbf{n}_{uc}^* and $P_{\mathbf{nn}}^{\text{uc}}$ represent the data equally well as \mathbf{J} , \mathbf{G} and \mathbf{F} , provided \mathbf{F} is invertible.

¹⁵This is seen most easily by noting that $\hat{\mathbf{n}}^{*\text{T}} \hat{\mathbf{n}}^* = 1$. Hence, $\hat{\mathbf{n}}^{\text{true}T} \Delta \hat{\mathbf{n}}^* = O(|\Delta \hat{\mathbf{n}}^*|^2)$, and $P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}\hat{\mathbf{n}}^{\text{true}} = O(|\Delta \hat{\mathbf{n}}^*|^3)$.

¹⁶Note that we cannot permit $\mathbf{G} = \mathbf{0}$, because that would lead to $\mathbf{n}_{\text{uc}}^* = \mathbf{0}$ and an indeterminate $\hat{\mathbf{n}}_{\text{uc}}^*$. Clearly, \mathbf{F} would need to be singular in that case.

The correctly constrained spin-axis attitude estimate can now be written

$$\begin{aligned}\hat{\mathbf{n}}^* &= ((P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{uc}})^{-1} + \lambda I_{3 \times 3})^{-1} (P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{uc}})^{-1} \mathbf{n}_{\text{uc}}^* = (I_{3 \times 3} + \lambda P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{uc}})^{-1} \mathbf{n}_{\text{uc}}^* \\ &= (\mathbf{F} + \lambda I_{3 \times 3})^{-1} \mathbf{F} \mathbf{n}_{\text{uc}}^* = (I_{3 \times 3} + \lambda \mathbf{F}^{-1})^{-1} \mathbf{n}_{\text{uc}}^*\end{aligned}\quad (52)$$

which are all equivalent, clearly, to equation (28). The Lagrange multiplier, likewise, is a solution of

$$\hat{\mathbf{n}}_{\text{uc}}^{*\text{T}} (I_{3 \times 3} + \lambda P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{uc}})^{-2} \mathbf{n}_{\text{uc}}^* = 1 \quad (53)$$

Thus, the unconstrained “estimate” of the spin-axis vector becomes simply an intermediate step in the estimation of the constrained spin-axis vector. This is the case for all of the nonsingular algorithms developed in this work.

Because the unconstrained estimate of the spin-axis vector is expected to be close to the correct estimate of the spin-axis vector, one can obtain a first-order approximation, simply by unitizing it. Thus, we write the brute-force unconstrained estimate

$$\hat{\mathbf{n}}_{\text{uc}}^* \equiv \frac{\mathbf{n}_{\text{uc}}^*}{|\mathbf{n}_{\text{uc}}^*|} \quad (54)$$

To calculate the covariance matrix for this unconstrained estimate note

$$\Delta \hat{\mathbf{n}}_{\text{uc}}^* = (I_{3 \times 3} - \hat{\mathbf{n}}^{\text{true}} \hat{\mathbf{n}}^{\text{true}\text{T}}) \Delta \mathbf{n}_{\text{uc}}^* \equiv \Lambda_{\text{uc}} \Delta \mathbf{n}_{\text{uc}}^* \quad (55)$$

from which we can write the corresponding covariance matrix

$$P_{\hat{\mathbf{n}}_{\text{uc}}^* \hat{\mathbf{n}}_{\text{uc}}^*} = \Lambda_{\text{uc}} \mathbf{F}^{-1} \Lambda_{\text{uc}}^{\text{T}} \quad (56)$$

in analogy to equation (43). If \mathbf{F} shows strong correlations, then Λ and Λ_{uc} will be very different, and the unconstrained estimate will be a poor approximation. Clearly, we must have $P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} \leq P_{\hat{\mathbf{n}}_{\text{uc}}^* \hat{\mathbf{n}}_{\text{uc}}^*}$. In any event, the computation of a correctly constrained estimate is so simple, that the unconstrained approximation has little practical value as a final result.

Method 3: The Incremental Vector Method

Since $\hat{\mathbf{n}}$ is a direction, it is characterized by only two independent parameters, and, therefore, there can be only two independent *incremental* parameters. Likewise, since $\hat{\mathbf{n}}_i$ is a unit vector, then to $O(|\epsilon_i|^2)$ the increments in three-dimensional space must be along the two directions perpendicular to $\hat{\mathbf{n}}_i$. Thus, if $\{\hat{\mathbf{n}}_i, \hat{\mathbf{a}}_i, \hat{\mathbf{b}}_i\}$ constitutes a (proper) orthonormal triad of vectors, we may write

$$\begin{aligned}\hat{\mathbf{n}} &= \hat{\mathbf{n}}_{i-1} + \epsilon_{a,i} \hat{\mathbf{a}}_{i-1} + \epsilon_{b,i} \hat{\mathbf{b}}_{i-1} + O(\epsilon_{a,i}^2, \epsilon_{b,i}^2) \\ &= \hat{\mathbf{n}}_{i-1} + C_i \epsilon_i + O(|\epsilon_i|^2)\end{aligned}\quad (57)$$

where ϵ_i , a 2×1 array, and the 3×2 array C_i are given by

$$\epsilon_i = [\epsilon_{a,i} \quad \epsilon_{b,i}]^{\text{T}} \quad \text{and} \quad C_i \equiv [\hat{\mathbf{a}}_{i-1} \quad \hat{\mathbf{b}}_{i-1}] \quad (58\text{ab})$$

We initialize our iterative algorithm with¹⁷

$$\hat{\mathbf{n}}_o \equiv \hat{\mathbf{n}}_{\text{uc}}^* \quad (59)$$

¹⁷Again, we cannot allow \mathbf{F} to be singular or \mathbf{G} to vanish.

with $\hat{\mathbf{n}}_{\text{uc}}^*$ given by equation (54). Thus, we write

$$\begin{aligned} J(\hat{\mathbf{n}}) &= J(\hat{\mathbf{n}}_{i-1} + C_i \boldsymbol{\epsilon}_i) \\ &= \mathbf{J} + \mathbf{G}^T(\hat{\mathbf{n}}_{i-1} + C_i \boldsymbol{\epsilon}_i) + \frac{1}{2}(\hat{\mathbf{n}}_{i-1} + C_i \boldsymbol{\epsilon}_i)^T \mathbf{F}(\hat{\mathbf{n}}_{i-1} + C_i \boldsymbol{\epsilon}_i) \end{aligned} \quad (60)$$

The estimation of $\boldsymbol{\epsilon}_i$ is straightforward and leads to

$$\boldsymbol{\epsilon}_i^* = -(C_i^T \mathbf{F} C_i)^{-1} C_i^T (\mathbf{G} + \mathbf{F} \hat{\mathbf{n}}_{i-1}) \quad (61)$$

and

$$\hat{\mathbf{n}}_i = \hat{\mathbf{n}}_{i-1} + C_i \boldsymbol{\epsilon}_i^* + O(|\boldsymbol{\epsilon}_i|^2) \quad (62)$$

In order to maintain the unit norm of $\hat{\mathbf{n}}_i$ at each iteration, we should mechanize the calculation of $\hat{\mathbf{n}}_i$ as

$$\mathbf{n}_i = \hat{\mathbf{n}}_{i-1} + C_i \boldsymbol{\epsilon}_i^* \quad \text{and} \quad \hat{\mathbf{n}}_i = \mathbf{n}_i / |\mathbf{n}_i| \quad (63\text{ab})$$

As $i \rightarrow \infty$, we must have¹⁸ $\boldsymbol{\epsilon}_i^* \rightarrow \mathbf{0}$ and $\hat{\mathbf{n}}_i \rightarrow \hat{\mathbf{n}}^*$. To see that equations (63) are consistent with equation (28), we remark that it cannot be true that $\mathbf{G} + \mathbf{F} \hat{\mathbf{n}}_{i-1} \rightarrow \mathbf{0}$ because of the norm constraint. Therefore, taking the limit $i \rightarrow \infty$ (assuming the $\hat{\mathbf{a}}_i$ and $\hat{\mathbf{b}}_i$ have been defined so that the limit C_∞ exists), it must be true that $\mathbf{G} + \mathbf{F} \hat{\mathbf{n}}^*$ can have components only in the null space of C_∞ , that is, in the subspace spanned by $\hat{\mathbf{n}}^*$. Therefore, for some value of λ it must be true on taking the limit $i \rightarrow \infty$ that

$$\mathbf{G} + \mathbf{F} \hat{\mathbf{n}}^* = -\lambda \hat{\mathbf{n}}^* \quad (64)$$

which is equivalent to equation (27).

The companion matrix for $\boldsymbol{\epsilon}_i$, the i -th increment, is

$$(F_{\boldsymbol{\epsilon}\boldsymbol{\epsilon}})_i = E \left\{ \frac{\partial^2 J}{\partial \boldsymbol{\epsilon}_i \partial \boldsymbol{\epsilon}_i^T} \right\} = C_i^T \mathbf{F} C_i \quad (65)$$

and for the converged result

$$(F_{\boldsymbol{\epsilon}\boldsymbol{\epsilon}})_\infty = C_\infty^T \mathbf{F} C_\infty \quad (66)$$

The quantities $(F_{\boldsymbol{\epsilon}\boldsymbol{\epsilon}})_i$ in equation (65) and $(F_{\boldsymbol{\epsilon}\boldsymbol{\epsilon}})_\infty$ in equation (66) are not the Fisher information matrix (although they may be very near to it in value), because in theory the latter is always evaluated at $\hat{\mathbf{n}}^{\text{true}}$ and not at $\hat{\mathbf{n}}_i$ or $\hat{\mathbf{n}}^*$.¹⁹ We must not confuse the converged tangent plane spanned by the basis $\{\hat{\mathbf{a}}_\infty, \hat{\mathbf{b}}_\infty\}$ with the tangent plane for the definition of the spin-axis attitude errors. The individual bases $\{\hat{\mathbf{a}}_i, \hat{\mathbf{b}}_i\}$ span only the plane of the i -th spin-axis attitude increment. For the computation of the Fisher information matrix we must employ the plane spanned by $\{\hat{\mathbf{a}}^{\text{true}}, \hat{\mathbf{b}}^{\text{true}}\}$, which is perpendicular to $\hat{\mathbf{n}}^{\text{true}}$. This distinction is often disregarded and can lead to imprecision in our statements about the Fisher information matrix and the covariance matrix. In reference [9], which addressed a similar distinction in three-axis attitude estimation (*sic*), the distinction is reinforced by denoting the increment in the i -th Newton-Raphson iteration by $\boldsymbol{\epsilon}_i$ and the estimation error by $\tilde{\boldsymbol{\epsilon}}$. We shall follow that practice in this work.

¹⁸When obvious from the context, we will not denote the dimension of zero arrays explicitly.

¹⁹Of course, in mission operations, where one must compute confidence intervals for the estimate of a variable, and the true value of that variable is not available, one has no choice but to evaluate the covariance matrix at the estimate.

It follows that the estimate-error covariance matrix for $\tilde{\boldsymbol{\epsilon}}$ is

$$P_{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}} = F_{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}}^{-1} \quad (67)$$

with

$$F_{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}} = C^T F C \quad (68)$$

provided that the Fisher information matrix is invertible (i.e., the spin-axis attitude is observable). Here, C is defined similarly to C_i in equation (58) but in terms of $\hat{\mathbf{a}}^{\text{true}}$ and $\hat{\mathbf{b}}^{\text{true}}$. Note again that both $P_{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}}$ and $F_{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}}$ are symmetric 2×2 matrices. For convenience we write C rather than C^{true} .

$P_{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}}$ and $F_{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}}$ are matrix representations in the two-dimensional tangent space to $\hat{\mathbf{n}}^{\text{true}}$ with basis $\{\hat{\mathbf{a}}^{\text{true}}, \hat{\mathbf{b}}^{\text{true}}\}$. If we wish to expand this basis to three-dimensional space by appending the basis vector $\hat{\mathbf{n}}^{\text{true}}$, then the resulting 3×3 estimate-error covariance matrix becomes

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\{\hat{\mathbf{a}}^{\text{true}}, \hat{\mathbf{b}}^{\text{true}}, \hat{\mathbf{n}}^{\text{true}}\}} = \begin{bmatrix} P_{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}} & \mathbf{0}_{2 \times 1} \\ \mathbf{0}_{1 \times 2} & 0 \end{bmatrix} \quad (69)$$

which is explicitly singular. The superscript denotes the basis. For an arbitrary basis of the three-dimensional space, this becomes more generally

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} = C P_{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}} C^T \quad (70)$$

The last identity follows from

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} \equiv E\{\Delta\hat{\mathbf{n}}^* \Delta\hat{\mathbf{n}}^{*\text{T}}\} = C E\{\Delta\tilde{\boldsymbol{\epsilon}} \Delta\tilde{\boldsymbol{\epsilon}}^{\text{T}}\} C^T = C P_{\tilde{\boldsymbol{\epsilon}}\tilde{\boldsymbol{\epsilon}}} C^T \quad (71)$$

Thus, by explicit construction it follows that

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} \hat{\mathbf{n}}^{\text{true}} = \mathbf{0} + O(|\Delta\hat{\mathbf{n}}|^3) \quad (72)$$

the direct result of equation (23), from which

$$\hat{\mathbf{n}}^{\text{true}\text{T}} \Delta\hat{\mathbf{n}}^* = O(|\Delta\hat{\mathbf{n}}|^2) \quad (73)$$

We have already seen the consequences of this in equation (45).

Note, finally, the relationships

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} = C(C^T F C)^{-1} C^T \quad (74a)$$

$$P_{\mathbf{m}\mathbf{m}} = F^{-1} \quad (74b)$$

and consequently

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} = C(C^T(P_{\mathbf{m}\mathbf{m}}^{\text{uc}})^{-1} C)^{-1} C^T \quad (74c)$$

It is not necessarily true, as we see from equation (74c), that $P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}$ can be obtained from a simple bilinear operation on $P_{\mathbf{m}\mathbf{m}}$. Despite the redundancy, there is an economy and universality to the covariance matrix of equation (74a) which is referred to a single space-fixed basis, which are lacking in the matrix of equation (69) which is referred to one of a continuum of body-fixed bases.

Method 4: The Incremental-Angle Method

Except for the treatment of the measurement of the nadir angle and the Earth-Sun dihedral angle, the treatment up to now has been purely vectorial. However, the spherical angles of the spin-axis attitude provide an equally valid parameterization,

which merits our attention. This method is essentially the one used at the beginning of the space age for least-square spin-axis attitude estimation, which is prominent in reference [1]. The processing of the data by vectorial methods simply makes it more efficient.

We write the spin-axis vector as a function of the two spherical angles

$$\hat{\mathbf{n}}(\theta_1, \theta_2) = \begin{bmatrix} \sin \theta_1 \cos \theta_2 \\ \sin \theta_1 \sin \theta_2 \\ \cos \theta_1 \end{bmatrix} \quad (75)$$

and define the cost function of the spherical angles in the usual way as

$$J(\boldsymbol{\theta}) = J(\hat{\mathbf{n}}(\boldsymbol{\theta})) \quad (76)$$

with $\boldsymbol{\theta} = [\theta_1, \theta_2]^T$. Then the gradient vector and Hessian matrix are given by

$$\frac{\partial J}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta}) = M^T(\boldsymbol{\theta}) \frac{\partial J}{\partial \hat{\mathbf{n}}}(\boldsymbol{\theta}) = M^T(\boldsymbol{\theta})(\mathbf{G} + \mathbf{F}\mathbf{n}(\boldsymbol{\theta})) \quad (77)$$

$$\begin{aligned} \frac{\partial^2 J}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}(\boldsymbol{\theta}) &= M^T(\boldsymbol{\theta}) \frac{\partial^2 J}{\partial \hat{\mathbf{n}} \partial \hat{\mathbf{n}}^T}(\boldsymbol{\theta}) M(\boldsymbol{\theta}) + \sum_{j=1}^3 \frac{\partial J}{\partial \hat{\mathbf{n}}_j}(\boldsymbol{\theta}) \frac{\partial^2 \hat{\mathbf{n}}_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}(\boldsymbol{\theta}) \\ &= M^T(\boldsymbol{\theta}) \mathbf{F} M(\boldsymbol{\theta}) + \sum_{j=1}^3 (\mathbf{G} + \mathbf{F}\mathbf{n}(\boldsymbol{\theta}))_j \frac{\partial^2 \hat{\mathbf{n}}_j}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}(\boldsymbol{\theta}) \end{aligned} \quad (78)$$

with

$$\begin{aligned} M(\boldsymbol{\theta}) &\equiv \begin{bmatrix} \frac{\partial \hat{\mathbf{n}}(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial \hat{\mathbf{n}}(\boldsymbol{\theta})}{\partial \theta_2} \end{bmatrix} \\ &= \begin{bmatrix} \cos \theta_1 \cos \theta_2 & \cos \theta_1 \sin \theta_2 & -\sin \theta_1 \\ -\sin \theta_1 \sin \theta_2 & \sin \theta_1 \cos \theta_2 & 0 \end{bmatrix}^T \end{aligned} \quad (79)$$

The second term in equation (78) will nearly vanish when we evaluate the expectation to compute the Fisher information matrix.²⁰ Thus,

$$F_{\boldsymbol{\theta}\boldsymbol{\theta}} = M^T(\boldsymbol{\theta}^{\text{true}}) \mathbf{F} M(\boldsymbol{\theta}^{\text{true}}) \quad (80)$$

(c.f. equation (68)) and the Gauss-Newton estimation sequence for $\boldsymbol{\theta}$ becomes

$$\theta_1^{(0)} = \cos^{-1}((n_{\text{uc}}^*)_3 / |\mathbf{n}_{\text{uc}}^*|), \quad \theta_2^{(0)} = \arctan_2((n_{\text{uc}}^*)_2, (n_{\text{uc}}^*)_1) \quad (81a)$$

$$\boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}^{(i-1)} - [M^T(\boldsymbol{\theta}^{(i-1)}) \mathbf{F} M(\boldsymbol{\theta}^{(i-1)})]^{-1} M^T(\boldsymbol{\theta}^{(i-1)}) (\mathbf{G} + \mathbf{F}\hat{\mathbf{n}}(\boldsymbol{\theta}^{(i-1)})) \quad (81b)$$

$$\boldsymbol{\theta}^* = \lim_{i \rightarrow \infty} \boldsymbol{\theta}^{(i)} \quad (81c)$$

where $\hat{\mathbf{n}}^*$ is the unconstrained estimate of the spin-axis attitude presented earlier.²¹ Equation (81b) is almost identical to equation (61) except that $M(\boldsymbol{\theta}^{(i)})$ appears instead of C_i . As before, the covariance matrix of $\boldsymbol{\theta}^*$, $P_{\boldsymbol{\theta}\boldsymbol{\theta}}$, is given by the inverse of the Fisher information matrix. Hence,

²⁰In this regard, recall equation (64) above. If \mathbf{F} is on the order of 10^6 and λ is on the order of 10^3 , then the two terms in equation (78) will be in the ratio 1000:1. The converged value of the Newton-Raphson iteration will not be altered at all by the approximation of the companion matrix. Only the value of the related Fisher information matrix will be altered by 0.1 percent, which is not significant. In any event, the discarding of the second term in equation (78) may be more explicit, but it is not less important than the terms discarded in the linearization approximation of the incremental-vector approach.

²¹Which exists only if \mathbf{F} is nonsingular and \mathbf{G} is nonvanishing.

$$P_{\theta\theta} = [M^T(\boldsymbol{\theta}^{\text{true}}) \mathbf{F} M(\boldsymbol{\theta}^{\text{true}})]^{-1} \quad (82a)$$

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} = M(\boldsymbol{\theta}^{\text{true}}) P_{\theta\theta} M^T(\boldsymbol{\theta}^{\text{true}}) \quad (82b)$$

The incremental-angle approach, obviously, has much in common with the incremental-vector method. This should not be surprising, because θ_1 and θ_2 are simply the polar and azimuthal angles of $\hat{\mathbf{n}}$, equivalently in more geographic terms, the latitude (really colatitude) and longitude of $\hat{\mathbf{n}}$. At least locally, the line of latitude and the line of longitude passing through the point of the unit sphere corresponding to $\hat{\mathbf{n}}$ are the basis vectors of the tangent space to $\hat{\mathbf{n}}$. However, because a line of latitude is not an arc of a great circle, it is $\Delta\theta_1$ and $\sin\theta_1\Delta\theta_2$ which are the equivalent increments in the tangent space rather than $\Delta\theta_1$ and $\Delta\theta_2$. Other than that, the incremental-vector treatment will be virtually the same as the incremental-angle treatment provided one chooses

$$\hat{\mathbf{b}}_i = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{n}}_i}{|\hat{\mathbf{z}} \times \hat{\mathbf{n}}_i|}, \quad \hat{\mathbf{a}}_i = \hat{\mathbf{b}}_i \times \hat{\mathbf{n}}_i \quad \text{and} \quad \boldsymbol{\epsilon}_i = \begin{bmatrix} -\Delta\theta_{1i} \\ \sin\theta_{1i}\Delta\theta_{2i} \end{bmatrix} \quad (83abc)$$

with $\hat{\mathbf{z}} \equiv [0, 0, 1]^T$.²²

Note also that $\hat{\mathbf{n}}$, $\partial\hat{\mathbf{n}}/\partial\theta_1$, and $(\sin\theta_1)^{-1}\partial\hat{\mathbf{n}}/\partial\theta_2$ also form a set of three orthonormal axes, which are related, obviously, to the column vectors of $M(\boldsymbol{\theta})$.

It follows that

$$P_{\hat{\boldsymbol{\epsilon}}\hat{\boldsymbol{\epsilon}}} = \begin{bmatrix} 1 & 0 \\ 0 & \sin\theta_1 \end{bmatrix} P_{\theta\theta} \begin{bmatrix} 1 & 0 \\ 0 & \sin\theta_1 \end{bmatrix}^T \quad (84)$$

The matrix $M(\boldsymbol{\theta})$ is singular for certain values of the spherical angles θ_1 and θ_2 . If the spacecraft spin axis is parallel or antiparallel to the z -axis, then $\theta_1 = 0$ and the matrix $M(\boldsymbol{\theta})$ becomes rank one. One avoids this situation by choosing a different set of polar and azimuthal axes. In the case above, the polar axis was the z -axis, following the common convention. If instead the x -axis is chosen, then

$$\hat{\mathbf{n}}(\boldsymbol{\theta}) = [\cos\theta_1 \quad \sin\theta_1\cos\theta_2 \quad \sin\theta_1\sin\theta_2]^T \quad (85a)$$

$$M(\boldsymbol{\theta}) = \begin{bmatrix} -\sin\theta_1 & \cos\theta_1\cos\theta_2 & \cos\theta_1\sin\theta_2 \\ 0 & -\sin\theta_1\sin\theta_2 & \sin\theta_1\cos\theta_2 \end{bmatrix}^T \quad (85b)$$

$$\theta_1^{(0)} = \cos^{-1}((n_{uc}^*)_1/|\mathbf{n}_{uc}^*|), \quad \theta_2^{(0)} = \arctan_2((n_{uc}^*)_3, (n_{uc}^*)_2) \quad (85c)$$

If the y -axis is chosen as the polar axis, then

$$\hat{\mathbf{n}}(\boldsymbol{\theta}) = [\sin\theta_1\sin\theta_2 \quad \cos\theta_1 \quad \sin\theta_1\cos\theta_2]^T \quad (86a)$$

$$M(\boldsymbol{\theta}) = \begin{bmatrix} \cos\theta_1\sin\theta_2 & -\sin\theta_1 & \cos\theta_1\cos\theta_2 \\ \sin\theta_1\cos\theta_2 & 0 & -\sin\theta_1\sin\theta_2 \end{bmatrix}^T \quad (86b)$$

$$\theta_1^{(0)} = \cos^{-1}((n_{uc}^*)_2/|\mathbf{n}_{uc}^*|), \quad \theta_2^{(0)} = \arctan_2((n_{uc}^*)_1, (n_{uc}^*)_3) \quad (86c)$$

In deriving equations (85) and (86) we have simply permuted the indices of equations (75), (79), and (81a) cyclically.²³

²²The sign occurs in $\boldsymbol{\epsilon}_{1i}$, because our (local) coordinate system is south-east-up, rather than the more common north-east-down.

²³Equations (83) hold specifically when equation (75) is true. For axes consistent with equations (85) or (86), we must replace $\hat{\mathbf{z}}$ with $\hat{\mathbf{i}}$ or $\hat{\mathbf{j}}$, respectively.

Alternate Mechanization of the Lagrange-Multiplier Solution

The Lagrange-multiplier solution would seem to be the most complicated of the solutions, but it can be made very efficient. The matrix

$$D_{i-1} \equiv (\mathbf{F} + \lambda_{i-1} I_{3 \times 3})^{-1} \quad (87)$$

need not be computed separately, since it appears only in combination with the column matrix \mathbf{G} . Thus, one need compute only the two vectors \mathbf{V}_1 and \mathbf{V}_2 , which are defined to be the solutions of

$$(\mathbf{F} + \lambda_{i-1} I_{3 \times 3})\mathbf{V}_{1,i-1} = \mathbf{G} \quad \text{and} \quad (\mathbf{F} + \lambda_{i-1} I_{3 \times 3})\mathbf{V}_{2,i-1} = \mathbf{V}_{1,i-1} \quad (88ab)$$

respectively. Many high-level computer languages, such as Matlab[®], contain an operation which solves these linear equations directly without the intermediate computation of the matrix inverse. From these two vectors we have

$$\mathbf{n}_{i-1}^T \mathbf{n}_{i-1} = \mathbf{G}^T \mathbf{V}_{2,i-1} = \mathbf{V}_{1,i-1}^T \mathbf{V}_{1,i-1} \quad (89a)$$

$$\mathbf{n}_{i-1}^T D_{i-1} \mathbf{n}_{i-1} = \mathbf{V}_{1,i-1}^T \mathbf{V}_{2,i-1} \quad (89b)$$

which are the quantities that appear in equations (30), and

$$\hat{\mathbf{n}}^* = \lim_{i \rightarrow \infty} \mathbf{V}_{1,i} \quad (90)$$

The Importance of Constraint

To appreciate the ill effect of neglecting the constraint more clearly we examine the following analytical example, based on our numerical examples. Write the inverse unconstrained Fisher information matrix and the true spin-axis vector as

$$\mathbf{P} \equiv \mathbf{F}^{-1} = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix} \quad \text{and} \quad \hat{\mathbf{n}}^{\text{true}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (91ab)$$

then it follows from equations (43), (44), (55) and (56) that

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} = \begin{bmatrix} a - e^2/c & d - ef/c & 0 \\ d - ef/c & b - f^2/c & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad P_{\hat{\mathbf{n}}_{\text{uc}}\hat{\mathbf{n}}_{\text{uc}}} = \begin{bmatrix} a & d & 0 \\ d & b & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (92ab)$$

The third row and column in each case must vanish because of equations (43) and (56). The variances of the components of the correctly constrained spin-axis vector, however, are certainly smaller.

Consider the special case

$$\mathbf{P} \equiv \mathbf{F}^{-1} = P_o \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{bmatrix} \quad (93)$$

with ρ_{ij} the correlation

$$\rho_{ij} \equiv \mathbf{P}_{ij} / \sqrt{\mathbf{P}_{ii} \mathbf{P}_{jj}} = \rho_{ji}, \quad 1 \leq i, j \leq 3 \quad (94)$$

Then from equations (92)

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} = P_o \begin{bmatrix} 1 - \rho_{13}\rho_{31} & \rho_{12} - \rho_{13}\rho_{32} & 0 \\ \rho_{21} - \rho_{23}\rho_{31} & 1 - \rho_{23}\rho_{32} & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (95a)$$

and

$$P_{\hat{\mathbf{n}}_{uc}\hat{\mathbf{n}}_{uc}} = P_o \begin{bmatrix} 1 & \rho_{12} & 0 \\ \rho_{21} & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (95b)$$

Should we have, say, $\rho_{13} = \rho_{23} = 0.95$, an extreme case, then the standard deviations of the estimation errors for the brute-force estimate would be three times larger than that for the correctly constrained estimate. The importance of treating the constraint correctly cannot be overemphasized. The above analysis is easily extended to any Cartesian spin-axis attitude covariance matrix, because there always exists a proper orthogonal transformation which carries $\{\hat{\mathbf{n}}^{\text{true}}, \hat{\mathbf{a}}^{\text{true}}, \hat{\mathbf{b}}^{\text{true}}\}$ into $\{\hat{\mathbf{1}}, \hat{\mathbf{2}}, \hat{\mathbf{3}}\}$.

One must be careful not to confuse the correlations of \mathbf{P} with those of the unit-norm estimates. For Numerical Example 2, for example, the final covariance matrices are

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} = \begin{bmatrix} 0.685 & -1.193 & 0 \\ -1.193 & 6.253 & 0 \\ 0 & 0 & 0 \end{bmatrix} \times 10^{-6} \quad (96a)$$

and

$$P_{\hat{\mathbf{n}}_{uc}\hat{\mathbf{n}}_{uc}} = \begin{bmatrix} 2.879 & -5.015 & 0 \\ -5.015 & 12.909 & 0 \\ 0 & 0 & 0 \end{bmatrix} \times 10^{-6} \quad (96b)$$

These should be compared with \mathbf{P} in equation (125c).

The brute-force method performed well in Numerical Example 1, because the correlations were small, the largest smaller than 0.2 in magnitude, the others being very close to zero. In Numerical Example 2, on the other hand, all three correlations were fairly large, with a consequent reduction in estimation accuracy. Reference [2] was too quick to declare the constraint of only small importance.

Singular Cases

Of special interest are the cases in which \mathbf{F} is singular. This special case does not lead to continuous degeneracies for the estimate of the spin-axis attitude, because the latter has only two degrees of freedom, and except for the silly case where all of the physical vectors (other than the spin-axis vector) are parallel or antiparallel to one another, \mathbf{F} will always have rank two. We shall find, however, as we did in reference [2], that we will encounter discrete degeneracies in this case.

The physical vectors from which the spin-axis aspect is measured can be coplanar if the spacecraft is in an orbit in the ecliptic plane, and relies for spin-axis attitude determination on measurements of the Sun and nadir angles. This case was prominent in reference [2], in which considerable space was devoted to the determination of the spin-axis attitude from the single simultaneous measurements (the single-frame solution) of the nadir and Sun angles. The treatment here is a simple

generalization of the algorithm of reference [2]. We call the new algorithm developed here the *pseudo-inverse method*.

If \mathbf{F} is singular, then we can write

$$\mathbf{F} = \tilde{\mathbf{U}} \tilde{\mathbf{F}} \tilde{\mathbf{U}}^T \quad (97)$$

where the 2×2 matrix $\tilde{\mathbf{F}}$ is positive-definite and the 3×2 matrix $\tilde{\mathbf{U}}$ satisfies

$$\tilde{\mathbf{U}}^T \tilde{\mathbf{U}} = \mathbf{I}_{2 \times 2}, \quad \tilde{\mathbf{U}} \tilde{\mathbf{U}}^T = \mathbf{I}_{3 \times 3} - \hat{\mathbf{u}}_3 \hat{\mathbf{u}}_3^T \quad \text{and} \quad \tilde{\mathbf{U}}^T \hat{\mathbf{u}}_3 = \mathbf{0}_{2 \times 1} \quad (98abc)$$

with $\hat{\mathbf{u}}_3$ the null vector of \mathbf{F} . We can obtain $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{F}}$ from the spectral decomposition of \mathbf{F} , but there is no requirement in the present work that $\tilde{\mathbf{F}}$ be diagonal. If $\tilde{\mathbf{F}}$ is diagonal, then

$$\tilde{\mathbf{U}} = [\hat{\mathbf{u}}_1 \quad \hat{\mathbf{u}}_2] \quad (99)$$

where $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ are the other two characteristic vectors of \mathbf{F} . From the structure of \mathbf{F} and \mathbf{G} it follows also that

$$\mathbf{F} \hat{\mathbf{u}}_3 = \mathbf{0} \Rightarrow \mathbf{G}^T \hat{\mathbf{u}}_3 = 0 \quad (100)$$

Likewise, we must have

$$\mathbf{G} = \mathbf{0} \Rightarrow \det \mathbf{F} = 0 \quad (101)$$

Defining the 2×1 matrix $\tilde{\mathbf{G}}$ by

$$\tilde{\mathbf{G}} \equiv \tilde{\mathbf{U}}^T \mathbf{G} \quad (102)$$

we can write the cost function $J(\hat{\mathbf{n}})$ of equation (24) as

$$J(\hat{\mathbf{n}}) = \mathbf{J} + \tilde{\mathbf{G}}^T \tilde{\mathbf{n}} + \frac{1}{2} \tilde{\mathbf{n}}^T \tilde{\mathbf{F}} \tilde{\mathbf{n}} = \tilde{J}(\tilde{\mathbf{n}}) \quad (103)$$

with

$$\tilde{\mathbf{n}} \equiv \tilde{\mathbf{U}}^T \hat{\mathbf{n}} \quad (104)$$

Since $\tilde{\mathbf{n}}$ is not constrained, the minimization of $J(\hat{\mathbf{n}})$ leads straightforwardly to

$$\tilde{\mathbf{n}}^* = -\tilde{\mathbf{F}}^{-1} \tilde{\mathbf{G}} \quad \text{or} \quad \hat{\mathbf{n}}^* = \pm \hat{\mathbf{u}}_3 \quad (105)$$

Noting equation (98b) it follows that the two possible solutions can be combined as

$$\hat{\mathbf{n}} = \tilde{\mathbf{U}} \tilde{\mathbf{n}} + n_{\perp} \hat{\mathbf{u}}_3 \quad \text{and} \quad \hat{\mathbf{n}}^* = \tilde{\mathbf{U}} \tilde{\mathbf{n}}^* + n_{\perp}^* \hat{\mathbf{u}}_3 \quad (106)$$

and

$$|\hat{\mathbf{n}}|^2 = |\tilde{\mathbf{n}}|^2 + |n_{\perp}|^2 = 1 \quad (107)$$

where n_{\perp} is the (ambiguous) component along $\hat{\mathbf{u}}_3$, i.e., the component perpendicular to the measurement plane.

The two solutions for the spin-axis estimate are, therefore,

$$\hat{\mathbf{n}}_{\pm}^* = \tilde{\mathbf{U}} \tilde{\mathbf{n}}^* \pm \sqrt{1 - |\tilde{\mathbf{n}}^*|^2} \hat{\mathbf{u}}_3 \quad (108)$$

Now

$$J(\hat{\mathbf{n}}_{\pm}^*) = J_o + \mathbf{G}^T \hat{\mathbf{n}}_{\pm}^* + \frac{1}{2} \hat{\mathbf{n}}_{\pm}^{*T} \mathbf{F} \hat{\mathbf{n}}_{\pm}^* = \tilde{J}(\tilde{\mathbf{n}}^*) \quad (109)$$

and so the two solutions cannot be distinguished. In order to determine which of $\hat{\mathbf{n}}_{\pm}$ is the correct estimate, we must use other data, for example, the Earth-Sun dihedral angle data described by equations (10) through (16). For the case of a spacecraft in an ecliptic orbit relying on Sun and nadir angle measurements, these additional measurements are sensitive to the component of $\hat{\mathbf{n}}$ normal to the measurement plane.

We may write the solution also as

$$\hat{\mathbf{n}}_{\pm}^* = -\mathbf{F}^{\#}\mathbf{G} \pm \sqrt{1 - |\mathbf{F}^{\#}\mathbf{G}|^2} \hat{\mathbf{u}}_3 \quad (110)$$

with

$$\mathbf{F}^{\#} = \tilde{U} \tilde{F}^{-1} \tilde{U}^T \quad (111)$$

the Moore-Penrose pseudo-inverse of \mathbf{F} ,

$$\mathbf{F}\mathbf{F}^{\#}\mathbf{F} = \mathbf{F} \quad \text{and} \quad \mathbf{F}^{\#}\mathbf{F}\mathbf{F}^{\#} = \mathbf{F}^{\#} \quad (112ab)$$

and, in this case,

$$\mathbf{F}\mathbf{F}^{\#} = \mathbf{F}^{\#}\mathbf{F} = I_{3 \times 3} - \hat{\mathbf{u}}_3 \hat{\mathbf{u}}_3^T \quad (113)$$

These expressions should be compared with those for the single-frame case with two measurements in reference [2].

Covariance Analysis for Singular \mathbf{F}

Clearly, defining

$$\Delta \tilde{\mathbf{n}} = \tilde{F}^{-1} \Delta \tilde{\mathbf{G}} \quad (114)$$

it follows that

$$P_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}} = \tilde{F}^{-1} \quad (115)$$

and

$$\begin{aligned} \Delta \tilde{\mathbf{n}} &= \left[\tilde{U} \mp \frac{1}{\sqrt{1 - |\tilde{\mathbf{n}}_{\text{true}}|^2}} \hat{\mathbf{u}}_3 \tilde{\mathbf{n}}_{\text{true}}^T \right] \Delta \tilde{\mathbf{n}} = \left[\tilde{U} - \frac{1}{n_3^{\text{true}}} \hat{\mathbf{u}}_3 \tilde{\mathbf{n}}_{\text{true}}^T \right] \Delta \tilde{\mathbf{n}} \\ &\equiv \Lambda_{\text{coplanar}} \Delta \tilde{\mathbf{n}} \end{aligned} \quad (116)$$

and finally,

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} = \Lambda_{\text{coplanar}} P_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}} \Lambda_{\text{coplanar}}^T \quad (117)$$

Note that for singular \mathbf{F} , the method is no longer iterative. Note also that the covariance matrix becomes infinite when the spin-axis is normal to the measurement plane, or, equivalently, a null vector of \mathbf{F} . The methods of this section may be useful in cases where \mathbf{F} is only nearly singular.

Numerical Results

Example 1: Good Observability

We examine these algorithms for a spacecraft in a 100-minute circular equatorial Earth orbit. The spacecraft is Earth-locked with z -axis parallel to the spin axis of the Earth (the inertial z -axis) and the spacecraft x -axis pointing toward the nadir. The spacecraft is equipped with a coarse vector magnetometer, a coarse vector Sun sensor, and a (coarse) Earth horizon sensor, all three attitude sensors with angle

equivalent accuracy σ equal to 0.5 deg. Data are taken once per minute. Such a scenario was typical of the time of publication of reference [2].

For simplicity in our simulations we assume that the geomagnetic field at the equator is constant and directed along the inertial z -axis.²⁴

$$\hat{\mathbf{B}}_I = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (118)$$

with the subscript I denoting the inertial frame. The Sun direction is taken to be

$$\hat{\mathbf{S}}_I = \begin{bmatrix} \cos \alpha \\ 0 \\ \sin \alpha \end{bmatrix} \quad (119)$$

where $\alpha = 23$ deg. This particular model implies a nonstandard choice for the direction of the inertial x - and y -axes. We assume also that the Sun will be observable for orbit argument θ , measured from the inertial x -axis, in the interval -90 deg $\leq \theta \leq +90$ deg. The nadir vector is simply the negative direction of the spacecraft position vector. Thus,

$$\mathbf{Z}_k = \begin{bmatrix} z_{B,k} \\ z_{S,k} \\ z_{E,k} \end{bmatrix} = H_k \hat{\mathbf{n}} + \mathbf{v}_k \quad (120)$$

with H_k constructed according to the model of equation (17) and the measurement error covariance matrix is taken as

$$R_k = \sigma^2 I_{3 \times 3} \quad (121)$$

when all three measurements are present. When only magnetic-field and nadir measurements are available because of occultation of the Sun by the Earth, the \mathbf{Z}_k will have only two components and we will assume $R_k = \sigma^2 I_{2 \times 2}$. Such a model is exceedingly simplified, but will be adequate for our studies and will make it easier to see the effects of the unit-norm constraint.

On the basis of this model we have generated the 3×3 matrix \mathbf{F} and the 3×1 matrix \mathbf{G} for one full orbit of data. These gave the results

$$\mathbf{F} = \begin{bmatrix} 1.231 & 0 & 0.241 \\ 0 & 0.650 & 0 \\ 0.241 & 0 & 1.415 \end{bmatrix} \times 10^6, \quad \mathbf{G} = \begin{bmatrix} -0.241 \\ -0.001 \\ -1.416 \end{bmatrix} \times 10^6 \quad (122ab)$$

$$\mathbf{P} \equiv \mathbf{F}^{-1} = \begin{bmatrix} 0.841 & 0 & -0.143 \\ 0 & 1.538 & 0 \\ -0.143 & 0 & 0.731 \end{bmatrix} \times 10^{-6} \quad (122c)$$

The near vanishing of $\mathbf{F}(1, 2)$ and $\mathbf{F}(2, 3)$ as well as of $\mathbf{G}(2)$ are the result of the special and very simple choices that were made for the test scenario.

²⁴Actually, the equatorial magnetic field at the Earth's surface has a declination varying from 2 deg to 10 deg and an inclination varying roughly from -30 deg to $+20$ deg. Our magnetic field model is not very realistic, but adequate for our simulation needs.

For this example we have computed the spin-axis attitude estimate and displayed those results in Table 1,²⁵ which shows results for the Lagrange-multiplier method (Table 1A), the incremental-vector method (Table 1B), the incremental-angle method (Table 1C), and the unconstrained brute-force method (Table 1D). The iteration was terminated in each case when the last change in the result was smaller than 0.000001. The first four columns show the iteration index and the three components of the estimates of $\hat{\mathbf{n}}$. The second iteration for each method shown in Table 1 was obviously superfluous.

Note that in all three mechanizations of the constrained estimation only one iteration was required beyond the initial unconstrained approximation. Without the use

TABLE 1. Comparison of the Four Methods for Numerical Example 1

A. Lagrange-Multiplier Method				
Iteration	n_1	n_2	n_3	λ
0	0.000241	0.000156	1.000103	0
1	0.000261	0.000156	0.999999	143
2	0.000261 ± 0.000901	0.000156 ± 0.001240	0.999999 ± 0	143
B. Incremental-Vector Method				
Iteration	n_1	n_2	n_3	$ \hat{\mathbf{n}}_i - \hat{\mathbf{n}}_{i-1} $
0	0.000241	0.000156	0.999999	—
1	0.000261	0.000156	0.999999	2.0×10^{-5}
2	0.000261 ± 0.000901	0.000156 ± 0.001240	0.999999 ± 0	2.3×10^{-9}
C. Incremental-Angle Method				
Iteration	n_1	n_2	n_3	$ \hat{\mathbf{n}}_i - \hat{\mathbf{n}}_{i-1} $
0	0.000241	0.000156	0.999999	—
1	0.000261	0.000156	0.999999	2.0×10^{-5}
2	0.000261 ± 0.000901	0.000156 ± 0.001240	0.999999 ± 0	1.4×10^{-7}
D. Unconstrained Brute-Force Method				
Iteration	n_1	n_2	n_3	
0	0.000241 ± 0.000917	0.001560 ± 0.001240	0.999999 ± 0	

²⁵In Table 1 and Table 2, numbers which are formally zero are displayed as zero. We do the same for elements of the covariance matrix, notably in equations (122a) and (125a) which are intrinsically zero or extraordinarily minute in the specific example.

of $\hat{\mathbf{n}}_{uc}$ as an initial value as many as a dozen iterations were found to be necessary. (Note that the initial value $\hat{\mathbf{n}}_{uc}$ is built into the Lagrange-multiplier method.) The errors in n_3 are much smaller than the others due to the influence of the norm constraint. Had we chosen $\hat{\mathbf{n}}^{true}$ to be different from a coordinate axis, this would not have been the case. Note also that the zero-th iteration of the Lagrange-multiplier method is different from the other two, because the former does not incorporate a normalization step. By the first iteration, however, it seems to have accomplished its mission to six significant digits. For the unconstrained brute-force method the standard deviation of the estimate of n_1 was larger than that of the other estimates by about 2 percent. This is consistent with the correlation ρ_{13} of \mathbf{P} being roughly -0.2 . This is explained fully in the following section on the importance of constraint.

The one-sigma confidence intervals for the spin-axis attitude estimates were each calculated using the expression for the spin-axis attitude covariance matrix derived for each method. Not surprisingly, the numerical values are the same for the first three cases. The large value of λ in our numerical example was not unexpected. Given equation (41b) we anticipate

$$\lambda = 0 \pm 1170 \quad (123)$$

In the present case, λ is positive, because $|\mathbf{n}_{uc}^*| > 1$.

Example 2: Poorer Observability

The measurement vector was chosen this time to be

$$\mathbf{Z}_k = [z_{E,k}, z_{S,k}]^T \quad (124)$$

We have again chosen the orbit plane to be equatorial and the x -axis of the orbit plane to be the projection of the Sun direction. The data interval spanned the orbit angular interval from 0 deg to 45 deg measured from the orbit x -axis. The time interval between measurements has been adjusted so that the estimate remains based on 100 measurements, as in the previous example. The measurements consist solely of the Sun angle and the nadir angle. For this case, we find

$$\mathbf{F} = \begin{bmatrix} 2.186 & 0.417 & 0.472 \\ 0.147 & 0.239 & 0 \\ 0.472 & 0 & 0.200 \end{bmatrix} \times 10^6, \quad \mathbf{G} = \begin{bmatrix} -0.471 \\ 0.001 \\ -0.201 \end{bmatrix} \times 10^6 \quad (125ab)$$

$$\mathbf{P} \equiv \mathbf{F}^{-1} = \begin{bmatrix} 2.879 & -5.015 & -6.784 \\ -5.015 & 12.909 & 11.814 \\ -6.784 & 11.814 & 20.969 \end{bmatrix} \times 10^{-6} \quad (125c)$$

The unconstrained covariance matrix \mathbf{P} now shows correlations $\rho_{12} = -0.862$, $\rho_{13} = -0.873$ and $\rho_{23} = -0.718$, which are considerably greater in magnitude than those in Example 1.²⁶ The three correctly constrained methods all yield the same result to six decimal places for the estimate of the spin-axis vector, but that from the unconstrained brute-force method is decidedly poorer in quality. The results are shown in Table 2.

²⁶Had we chosen to make one of the sensor accuracies larger than the other, the correlation would have been greater still.

TABLE 2. Comparison of the Four Methods for Numerical Example 2

A. Lagrange-Multiplier Method				
Iteration	n_1	n_2	n_3	λ
0	-0.000602	-0.002632	1.003112	0
1	0.000400	-0.004375	1.000010	148
2	0.000407	-0.004387	0.999990	149
4	0.000407 ± 0.000828	-0.004387 ± 0.002501	0.999990 ± 0	149
B. Incremental-Vector Method				
Iteration	n_1	n_2	n_3	$ \hat{\mathbf{n}}_i - \hat{\mathbf{n}}_{i-1} $
0	-0.000601	-0.002624	0.999996	-
1	0.000407	-0.004388	0.999990	0.0020
2	0.000407 ± 0.000828	-0.004388 ± 0.002501	0.999990 ± 0	7×10^{-7}
C. Incremental-Angle Method				
Iteration	n_1	n_2	n_3	$ \hat{\mathbf{n}}_i - \hat{\mathbf{n}}_{i-1} $
0	-0.000601	-0.002624	0.999996	-
1	0.001180	-0.004017	0.999991	0.0022
2	0.000361	-0.004309	0.999991	0.00087
3	0.000408	-0.004387	0.999990	0.000091
4	0.000407 ± 0.000827	-0.004387 ± 0.002501	0.999990 ± 0	7×10^{-7}
D. Unconstrained Brute-Force Method				
Iteration	n_1	n_2	n_3	
0	-0.000600 ± 0.001697	-0.00262 ± 0.003593	0.999996 ± 0	

The convergence of the iterative algorithms in this case is certainly poorer but better than the more than a dozen iterations which would be the case if one did not have a good initial value available. However, if we make our convergence criterion one-tenth of the standard deviation of the estimate error rather than an arbitrary 0.000001, then one iteration still provides sufficient accuracy for all four implementations of the iterative algorithms. Note also that the standard deviation of the estimate of n_1 is twice as large for the brute-force unconstrained estimate than for any of the constrained estimates. For the estimate of n_2 the brute-force error is 50 percent larger. In terms of a rotationally-invariant error measure, the trace of the Cartesian spin-axis attitude covariance matrix, the trace for the brute-force estimate is larger than that for the properly-constrained estimate by a factor 2.3.

Execution Times

Table 3 shows the relative execution times of the different methods obtained using Matlab[®]. All five results are for the first numerical example and each iterative method was terminated after a single iteration beyond the initial brute-force initialization (in the implementations of the Lagrange-multiplier method, equivalently, the estimate uses λ_1). In all cases, the code has been optimized as much as possible. Surprisingly, the use of the linear-equation solver increases the computational burden rather than decreasing it. Based on the results of Tables 2 and 3 the incremental-vector method would seem to offer the best value.²⁷

Comparison with Sampled Covariance Matrices

As a check on the consistency of our calculations, we have compared the results for the model covariances, as given by equations (43), (55), (74a), and (82b), with the sampled covariance matrices for the spin-axis attitude estimate

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{sampled}} \equiv \frac{1}{N} \sum_{m=1}^N (\hat{\mathbf{n}}_m^* - \hat{\mathbf{n}}^{\text{true}}) (\hat{\mathbf{n}}_m^* - \hat{\mathbf{n}}^{\text{true}})^T \quad (126)$$

and with a similar definition for the sampled covariance matrix $P_{\hat{\mathbf{n}}_{\text{oc}}\hat{\mathbf{n}}_{\text{oc}}}^{\text{sampled}}$. Here, $\hat{\mathbf{n}}_m^*$ is the estimate of the spin-axis vector for the m -th sampled data set, $m = 1, \dots, N$.

$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{sampled}}$ is a random matrix and satisfies

$$P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{sampled}} = P_{\hat{\mathbf{n}}\hat{\mathbf{n}}} + \Delta P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{sampled}} \quad (127)$$

For N very large, $(\Delta P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{sampled}})_{ij}$ will be approximately Gaussian and zero-mean with variance given by

$$\text{Var}\{(P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{sampled}})_{ij}\} = \frac{1}{N} [(P_{\hat{\mathbf{n}}\hat{\mathbf{n}}})_{ii} (P_{\hat{\mathbf{n}}\hat{\mathbf{n}}})_{jj} + (P_{\hat{\mathbf{n}}\hat{\mathbf{n}}})_{ij}^2] \quad (128)$$

We have computed $P_{\hat{\mathbf{n}}\hat{\mathbf{n}}}^{\text{sampled}}$, the model covariance matrix, and the confidence bounds for all four spin-axis attitude estimation methods and both numerical examples for 100 sample tests ($N = 100$) and found agreement within the anticipated confidence bounds. As an example, for the iterative algorithms of Numerical Example 2, the results were

TABLE 3. Execution Times for Spin-Axis Attitude Estimation

Method	Relative Execution Time
Iterative Optimal Methods	
Lagrange-multiplier (matrix inverse)	76
Lagrange-multiplier (linear equation)	88
Incremental-vector	66
Incremental-angle	100
Non-iterative Approximate Method	
Brute-force	34

²⁷As demonstrated recently [25], execution times or counts of floating-point operations in Matlab[®] can be a very poor indicator of algorithm speed or efficiency.

$$\begin{aligned} & \begin{bmatrix} 0.707 & -1.530 \\ -1.530 & 7.549 \end{bmatrix} \times 10^{-6} \\ &= \begin{bmatrix} 0.685 & -1.193 \\ -1.193 & 6.253 \end{bmatrix} \times 10^{-6} \pm \begin{bmatrix} 0.116 & 0.231 \\ 0.231 & 0.500 \end{bmatrix} \times 10^{-6} \quad (129) \end{aligned}$$

where, for reasons of space, we have deleted the uninteresting third row and third column, and the three matrices are in the same order as in equation (127). The errors in the sampled covariances are 0.3σ , -1.7σ and 0.2σ , where σ is the appropriate standard deviation for each covariance.

A Further Test

All of the algorithms for the nonsingular case, except for the approximate brute-force algorithm, take into account at a fundamental level the unit-norm constraint of the spin-axis attitude, although the temporary neglect of the norm constraint can clearly provide a useful first step. The performance of the brute-force algorithm has been mixed. In our first example, the unconstrained initial estimate when normalized turned out to provide all the accuracy that was needed. In our second example, however, this was not the case, and the cavalier treatment of the unit-norm constraint for the spin-axis vector led to a distinctly inferior result. Proponents of the problematic unconstrained (three-axis) quaternion Kalman filter should find a lesson here.

One can also see this by computing the figure of merit

$$\mu(\hat{\mathbf{m}}) \equiv (\hat{\mathbf{m}} - \hat{\mathbf{n}}^{\text{true}})^T P_{\hat{\mathbf{m}}}^{-1} (\hat{\mathbf{m}} - \hat{\mathbf{n}}^{\text{true}}) \quad (130)$$

For the three iterative constrained algorithms in Case 1, we found that μ had a mean value 1.771 and a standard deviation of 1.458, very close to what one would expect of a chi-square variable with two degrees of freedom (mean = 2, variance = 4).²⁸ When this same quantity is evaluated for the brute-force estimate, we found a mean of 2.073 and a standard deviation of 1.812, not very different. For Case 2, however, where the correlations played a large role in the accuracy of the estimate, we found for the iterative algorithms that $\mu = 1.830 \pm 1.948$, still compatible with a chi-square distribution, but for the brute-force method $\mu = 5.143 \pm 7.476$, very different from chi-square behavior. For Case 2, the mean of μ for the brute-force method differed from that for the constrained iterative methods by 3.313, a very significant deviation when the goodness of fit is judged by proximity to the value 2 ± 0.2 .

The Singular Case

To illustrate the efficacy of the special algorithm developed for the singular case, we have compared the Lagrange-multiplier and pseudo-inverse methods for a scenario like that of Table 2 but with $\alpha = 0$, so that the Sun vector lies in the orbit plane (whose coordinates, therefore, are ecliptic coordinates). We have also chosen

$$\hat{\mathbf{n}}^{\text{true}} = [3/5, 0, 4/5]^T \quad (131)$$

²⁸Since these sampled means and variances were for 100 tests, we anticipate an error level of 10 percent for the mean of μ and 14 percent for the standard deviation.

so that $\Lambda_{\text{coplanar}}$ will be finite. For the pseudo-inverse algorithm

$$\hat{\mathbf{n}}_{\pm}^* = \begin{bmatrix} 0.600389 \\ 0.005466 \\ \pm 0.799690 \end{bmatrix} \pm \begin{bmatrix} 0.000779 \\ 0.138917 \\ 0.000585 \end{bmatrix} \quad (132)$$

The solution $\hat{\mathbf{n}}_+^*$ is in very close agreement with the true value.²⁹ In practice, the two solutions must be distinguished by examining other data, for example, the Earth-Sun dihedral angle.

Similarities to Three-Axis Attitude Estimation

There are numerous similarities between spin-axis and three-axis attitude estimation. The use of a quadratic cost function with constant coefficients was present already in reference [2]. In three-axis attitude estimation, the 3×1 column vector $\hat{\mathbf{n}}$ must be replaced by the 9×1 column vector \mathbf{A} composed of the nine elements of the 3×3 attitude matrix A . Thus, \mathbf{G}_A is also a 9×1 column vector, and \mathbf{F}_{AA} is a 9×9 positive-semidefinite matrix (positive definite if the three-axis attitude is observable).³⁰ For vanishing measurement noise one has necessarily

$$\mathbf{A}^{\text{true}} = -\mathbf{F}_{AA}^{-1} \mathbf{G}_A^{\text{true}} \quad (133)$$

analogously to equation (32). The unconstrained brute-force estimate of the three-axis attitude “vector” is likewise

$$\mathbf{A}_{\text{uc}}^* = -\mathbf{F}_{AA}^{-1} \mathbf{G}_A^{\text{uc}} \quad (134)$$

We cannot orthogonalize \mathbf{A}_{uc}^* as simply as we could normalize \mathbf{n}_{uc}^* earlier. The simplest way to orthogonalize \mathbf{A}_{uc}^* is to use Markley’s modification [22] of Shepperd’s algorithm [23] for extracting the quaternion \bar{q} from \mathbf{A}_{uc}^* and then to recompute A from this \bar{q} . A more cumbersome approach is to maximize $\text{tr}[A_{\text{uc}}^{*T}A]$, which entails using a solution to the Wahba problem [24].³¹ These are both arbitrary orthogonalizations and do not yield the optimal estimate although either is adequate for generating an initial value for an iterative process which converges to the optimal estimate. Part II of reference [9] used an equally arbitrary but simpler method.

The Lagrange-multiplier method, which, effectively, provides an optimal orthogonalization of the brute-force estimate of the three-axis attitude, was present already in reference [2] and has found other echoes in three-axis attitude estimation. Kasdin and Weaver [26] have applied it to three-axis attitude estimation in the context of the Kalman filter. Part II of reference [9] presents it within the context of batch three-axis attitude estimation. The Lagrangian-multiplier method for three-axis attitude is complicated, because there are six independent Lagrange multipliers. Therefore, methods similar to the incremental-vector and incremental-angle methods for three-axis attitude are to be preferred.³²

In the three-axis equivalent to the incremental-vector and incremental-angle methods, one works to greater advantage with 3×3 matrices rather than with

²⁹The \pm sign preceding the column vector of confidence bounds is not linked to the sign label of the solution.

³⁰We do not make comparisons with quaternion estimation, in which the constraint is one of norm, because the measurements are usually linear in A but not in \bar{q} .

³¹Reference [24] presents an excellent review of the solutions to the Wahba problem. For its numerical results and their interpretation see references [21] and [25].

³²And of course, both are overshadowed by special solutions to the Wahba problem [24], especially the QUEST algorithm [27].

9×1 column vectors. Instead of equation (57) one uses instead

$$A = \delta A(\epsilon_i) A_{i-1} = A_{i-1} + [[\epsilon_i]] A_{i-1} \quad (134)$$

with

$$[[\mathbf{v}]] \equiv \begin{bmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{bmatrix} \quad (135)$$

This is the incremental-vector method presented in Part I of reference [9]. In general, one initializes the Newton-Raphson sequence for ϵ_i not with the brute-force attitude estimate above but with the result of the TRIAD algorithm [27], whose computation is much simpler and is always proper orthogonal. This methodology assumes that at least a subset of the data consists of direction measurements, required by the TRIAD algorithm, which is generally the case. Despite its attractiveness, this incremental-vector method for three-axis attitude has not, to the authors' knowledge, ever been implemented in actual mission support, probably because it had not been suggested before the advent of efficient solutions to the Wahba problem [24], especially QUEST [27]. Attitude determination analysts before then preferred to work with Euler angles.

For the incremental-angle method, the analogy is the differential correction methods for attitude estimation in terms of Euler angles, which were common in the early decades of the space age. Typically, one would define initial values of the Euler angles based on the TRIAD attitude solution [27] of a subset of the data (assuming the data consisted of measured directions) and then carry out the Newton-Raphson sequence in these Euler angles. This method disappeared almost overnight with the advent of the QUEST algorithm [27].

The ambiguity of the singular case occurs, because one wishes to estimate a spin-axis vector from two cosine measurements, $\hat{\mathbf{v}}_1 \cdot \hat{\mathbf{n}}$ and $\hat{\mathbf{v}}_2 \cdot \hat{\mathbf{n}}$, so that one has no way of removing the sign ambiguity of $(\hat{\mathbf{v}}_1 \times \hat{\mathbf{v}}_2) \cdot \hat{\mathbf{n}}$. One is left with a similar phenomenon when one attempts to determine the three-axis attitude from *three* cosine measurements, with the arc lengths being between three directions fixed in the spacecraft and three corresponding directions fixed in the primary reference frame [28]. Since the three-axis attitude can be described by only three parameters, it would seem that the measurement of three scalars would be enough. However, careful analysis shows that depending on the disposition of the vectors, there will be either a four-fold or an eight-fold degeneracy in the solutions. Attempts to economize on the inputs to the TRIAD algorithm have a similar effect [29].

The need to single out a particular axis in the incremental-angle and also in the incremental-vector method (one must introduce a known axis in order to construct $\hat{\mathbf{a}}_i$ and $\hat{\mathbf{b}}_i$ normal to $\hat{\mathbf{n}}_i$) is reminiscent of the problem in the fast solutions to the Wahba problem, which require the interventions of the method of sequential rotations [24, 27]. The problem arose in three-axis attitude estimation when one estimates a three-parameter set characterizing a unit four-dimensional column vector (the quaternion). In spin-axis attitude one equivalently estimates a two-parameter set characterizing a unit three-dimensional column vector.

There is nothing new under the Sun.³³

³³What was is what shall be, and what has been done is what shall be done, and there is nothing new under the sun (Ecclesiastes 1 : 9).

Discussion and Conclusions

A number of algorithms have been presented for spin-axis attitude estimation. All of these algorithms ultimately estimate a spin-axis vector, which is a representation of the spacecraft spin-axis (generally, the spacecraft body z -axis) with respect to some fiducial space-fixed coordinate system, typically inertial, because the spacecraft spin will cause the spin-axis to be inertially stabilized.

Note that the 3×3 Cartesian spin-axis Fisher information matrix, $P_{\mathbf{a}\mathbf{a}}^{-1}$, can only be of rank two, and, therefore, the 3×3 spin-axis attitude covariance matrix is not generally defined directly in terms of the 3×3 spin-axis attitude covariance matrix. The 3×3 spin-axis attitude covariance matrix in this work is in reality the pseudo-inverse of the 3×3 spin-axis Fisher information matrix and *vice versa*. Nonetheless, the error bounds in Table 1 are meaningful, because they reflect the true variation of these estimates, even though they provide no information on their mutual correlation.

The 3×3 Cartesian spin-axis attitude covariance has one excellent quality, it is always defined with respect to the same inertial axes. In a sense, this 3×3 covariance can be regarded almost as a representation of both the spin-axis attitude and the spin-axis attitude covariance matrix in the same way that the attitude profile matrix B of the Wahba problem [24, 27] contains both the (three-axis) attitude estimate and its covariance matrix [20]. Unfortunately, the null vector of this 3×3 matrix gives us the spin-axis vector only within a sign.

Execution times are often not very significant in an interpreted language such as Matlab[®] where the computation of a simple operation, such as a vector product, by means of an external function may take considerably longer than the same operation coded explicitly without a function call. In all cases we have removed such function calls when the removal resulted in a faster algorithm. Likewise, in the incremental-angle method we have avoided duplicate evaluations of the sine and cosine functions. The incremental-angle method, the “classical” method for spin-axis attitude estimation, turns out to be the slowpoke of our algorithms. It would seem that the evaluation of the trigonometric functions in the incremental-angle method imposes a greater computational burden than the vectorial apparatus of the incremental-vector method. The incremental-vector method also converges in far fewer steps, making it the clear winner. Even to achieve an error in n_1^* significantly smaller than the confidence bound (0.0087 rather than the arbitrarily more stringent 0.000001), the incremental-angle method required two iterations beyond the brute force initialization. These, of course, were the motivation for abandoning a trigonometric approach in reference [2]. The brute-force computation is fastest, as was to be anticipated, but this is insufficient recommendation for its use.

We have given special attention to singular cases, that is, when the unconstrained spin-axis attitude information matrix, \mathbf{F} , is singular. Such algorithms are useful when the physical vectors (other than the spin-axis itself) on which the cosine measurements are based are coplanar. This can occur for example, either for the case of only two cosine measurements, the principal topic of reference [2], and for the case where the spacecraft orbit lies in the ecliptic plane and the spin-axis attitude is estimated from measurements of the Sun angle and nadir angle. A recent example, showing that the methods of reference [2] are still useful, is the estimation of the spin-axis attitude from Sun aspect alone [17].

While our pseudo-inverse algorithm generates both solutions to the spin-axis attitude estimation problem for singular F , it does not tell us which one is correct. To accomplish that, we must examine other data, which can be consistent with only one of the solutions. It is better, however, to include such data (such as dihedral-angle data) directly in the construction of the spin-axis attitude estimate, so that F will not be singular, and the estimate will be more accurate.

An important outcome of this work is the importance of treating the norm constraint correctly. When the estimation error shows strong correlations, the improper treatment of the unit-norm constraint can lead to unacceptable error levels. This should also serve as a warning to analysts who insist on ignoring the unit-norm constraint of the quaternion in the attitude Kalman filter. Part II of reference [9] presented numerous cases where neglect of the norm constraint led not only to less accurate but even silly results. Reference [2], as we have said, was too hasty in discounting the importance of the norm constraint.

This work sought to bring greater sophistication to the estimation of spin-axis attitude. We must keep in mind, however, that spin-axis attitude is a largely unsophisticated quantity. Unless the spacecraft is spinning about the major principal axis, it will nutate or drift, and the spin-axis attitude will not be constant in time. Energy-dissipation devices, such as nutation dampers, will cause the spacecraft to spin about the major axis of the inertia tensor, but the instrumented "spin" axis may be offset from the major axis, owing to the difficulty of measuring the inertia tensor, especially the products of inertia, before launch. Hence, the supposed spin axis will cone in space, and further accuracy can be obtained only by estimating the direction of the major axis in the body frame as well. When a single-axis attitude estimate, say of the optical axis of a spacecraft payload, is needed at very high accuracy, it may best be accomplished using a star-tracker and estimating the full three-axis attitude.

Acknowledgment

The authors are grateful to Douglas Bender of Boeing Aerospace, Luc Fraiture of ESA, Douglas C. Freesland of ACS Engineering, Inc., Frank Janssens of ESA, F. Landis Markley of NASA Goddard Space Flight Center, Milton Phenneger of A. K. Space, James R. Wertz of Microcosm, Inc., and especially Jozef C. van der Ha of Kyushu University, for many interesting comments and helpful criticisms.

References

- [1] WERTZ, J.R. (ed.) *Spacecraft Attitude Determination and Control*, Springer, Berlin and New York, 1978.
- [2] SHUSTER, M. D. "Efficient Algorithms for Spin-Axis Attitude Estimation," *The Journal of the Astronautical Sciences*, Vol. 31, No. 2, April–June 1983, pp. 237–249; Errata, Vol. 51, No. 1, January–March 2003, p. 121.
- [3] THOMPSON, R. H., NEAL, G. F., and SHUSTER, M. D. "Magnetometer Bias Determination and Spin-Axis Attitude Estimation for the AMPTE Mission," *Journal of Guidance, Control, and Dynamics*, 1984, Vol. 7, pp. 505–507.
- [4] WERKING, R. D. "A Generalized Technique for Using Cones and Dihedral Angles in Attitude Determination," NASA X-581-73-292, September 1973.
- [5] WERTZ, J.R. and CHEN, L. C. "Geometrical Procedures for the Analysis of Spacecraft Attitude and Bias Determinability," presented as paper AAS-75-047 at the AAS/AIAA Astrodynamics Specialist Conference, Nassau, Bahamas, July 28–30, 1975.

- [6] WERTZ, J. R. and CHEN, L. C. "Geometric Limitations on Attitude Determination of Spinning Spacecraft," *Journal of Spacecraft and Rockets*, Vol. 13, 1976, pp. 564–571.
- [7] CHEN, L. C. and WERTZ, J. R. "Single-Axis Attitude Determination Accuracy," presented at the AAS/AIAA Astrodynamics Specialist Conference, Grand Teton National Park, Wyoming, September 7–9, 1977.
- [8] GRUBIN, C. "Simple Algorithm for Intersecting Two Conical Surfaces," *Journal of Spacecraft and Rockets*, Vol. 14, 1977, pp. 251–252.
- [9] SHUSTER, M. D. "Constraint in Attitude Estimation," *The Journal of the Astronautical Sciences*, Vol. 51, No. 1, January–March 2003, Part I, pp. 51–74, Part II, pp. 75–101.
- [10] SORENSON, H. W. *Parameter Estimation*, New York, Marcel Dekker, 1980.
- [11] SHUSTER, M. D. "A Survey of Attitude Representations," *The Journal of the Astronautical Sciences*, Vol. 41, No. 4, October–December 1993, pp. 439–517.
- [12] VAN DER HA, J. C. "Attitude Determination Covariance Analysis for Geostationary Transfer Orbits," *Journal of Guidance, Control, and Dynamics*, Vol. 9, No. 2, March–April 1986, pp. 156–163.
- [13] EMARA-SHABAIAK, H. E. "Spacecraft Spin Axis Attitude Determination," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 28, No. 2, April 1992, pp. 529–534.
- [14] FRAITURE, L. "Spin-Axis Attitude Estimation by a Controlled Correlation Method," *Journal of Spacecraft and Rockets*, Vol. 42, No. 1, January–February 2005, pp. 58–65.
- [15] VAN DER HA, J. C. "Equal-Chord Attitude Determination Method for Spinning Spacecraft," *Journal of Guidance, Control, and Dynamics*, Vol. 28, No. 5, September–October, 2005, pp. 997–1005.
- [16] VAN DER HA, J. C. "Spin Axis Attitude Determination Accuracy Model in Presence of Biases," *Journal of Guidance, Control, and Dynamics*, Vol. 29, No. 4, July–August 2006, pp. 799–809.
- [17] VAN DER HA, J. C. "The Two Sun-Cones Attitude Determination Method," presented as paper AAS 07-106 at the 17th Space Flight Mechanics Meeting, Sedona, Arizona, January 28–February 2, 2007; proceedings: *Advances in the Astronautical Sciences*, Vol. 127, 2007, pp. 117–130.
- [18] PHENNEGER, M. C., SINGHAL, S. P., LEE, T. W., and STENGLE, T. H. "Infrared Horizon Sensor Modeling for Attitude Determination and Control: Analysis and Mission Experience," NASA TM 86181, 1986.
- [19] VAN DER HA, J., ROGERS, G., DELLINGER, W., and STRATTON, J. "CONTOUR Phasing Orbits: Attitude Determination & Control Analysis and Flight Results," presented as paper AAS 03-150 at the 13th Space Flight Mechanics Meeting, Ponce, Puerto Rico, February 9–13, 2003; proceedings: *Advances in the Astronautical Sciences*, Vol. 114, 2003, pp. 787–781.
- [20] SHUSTER, M. D. "Maximum Likelihood Estimation of Spacecraft Attitude," *The Journal of the Astronautical Sciences*, Vol. 37, No. 1, January–March 1989, pp. 79–88.
- [21] CHENG, Y. and SHUSTER, M. D. "Robustness and Accuracy of the QUEST Algorithm," presented as paper AAS 07-102 at the 17th Space Flight Mechanics Meeting, Sedona, Arizona, January 28–February 2, 2007; proceedings: *Advances in the Astronautical Sciences*, Vol. 127, 2007, pp. 41–61.
- [22] MARKLEY, F. L. "Unit Quaternion from Rotation Matrix," *Journal of Guidance, Control, and Dynamics*, (to appear).
- [23] SHEPPERD, S. W. "Quaternion from Rotation Matrix," *Journal of Guidance and Control*, Vol. 1, 1978, pp. 223–224.
- [24] MARKLEY, F. L. and MORTARI, M. "Quaternion Attitude Estimation Using Vector Measurements," *The Journal of the Astronautical Sciences*, Vol. 48, Nos. 2 and 3, April–September 2000, pp. 359–380.
- [25] CHENG, Y. and SHUSTER, M. D. "The Speed of Attitude Estimation," presented as paper AAS 07-105 at the 17th Space Flight Mechanics Meeting, Sedona, Arizona, January 28–February 2, 2007; proceedings: *Advances in the Astronautical Sciences*, Vol. 127, 2007, pp. 101–116.
- [26] KASDIN, N. J. and WEAVER, T. J. M. "Recursive Satellite Attitude Estimation with the Two-Step Optimal Estimator," presented as paper AIAA 2002-4462 at the AIAA/AAS Astrodynamics Specialists Conference, Monterey, California, August 5–8, 2002.
- [27] SHUSTER, M. D. and OH, S. D. "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, Jan.–Feb. 1981, pp. 70–77.
- [28] SHUSTER, M. D. "Deterministic Three-Axis Attitude Determination," *The Journal of the Astronautical Sciences*, Vol. 52, No. 3, July–September 2004, pp. 405–419.

[29] SHUSTER, M. D. "The TRIAD Algorithm as Maximum-Likelihood Estimation," *The Journal of the Astronautical Sciences*, Vol. 54, No. 1, January–March 2006, pp. 113–123.

Appendix A: On the Possible Values of the Lagrange Multiplier

Equation (29) for the Lagrange multiplier is a nonlinear equation and may have multiple roots. We show in this appendix that it can have no more than six roots only one of which can be physical.

Define the function

$$\begin{aligned} g(\lambda) &\equiv \mathbf{G}^T \frac{1}{(\mathbf{F} + \lambda \mathbf{I}_{3 \times 3})^2} \mathbf{G} \\ &= \mathbf{n}_{uc}^{*T} \frac{1}{(\mathbf{I}_{3 \times 3} + \lambda \mathbf{F}^{-1})^2} \mathbf{n}_{uc}^* \end{aligned} \quad (\text{A1})$$

If the coordinate axes are chosen so that \mathbf{F} is diagonal with eigenvalues $0 \leq F_1 \leq F_2 \leq F_3$, then we may write $g(\lambda)$ as

$$g(\lambda) = \sum_{i=1}^3 \frac{G_i^2}{(\lambda + F_i)^2} = \sum_{i=1}^3 \frac{n_{uc,i}^2}{(1 + \lambda/F_i)^2} \quad (\text{A2})$$

We plot $g(\lambda)$ in Figure A-1 using the value of \mathbf{F} for Numerical Example 1. Clearly, $g(\lambda) > 0$ and becomes infinite at $\lambda = -F_i, i = 1, 2, 3$. The values of λ which lead to a solution of equation (A1) occur when the line with unit ordinate intersects $g(\lambda)$.

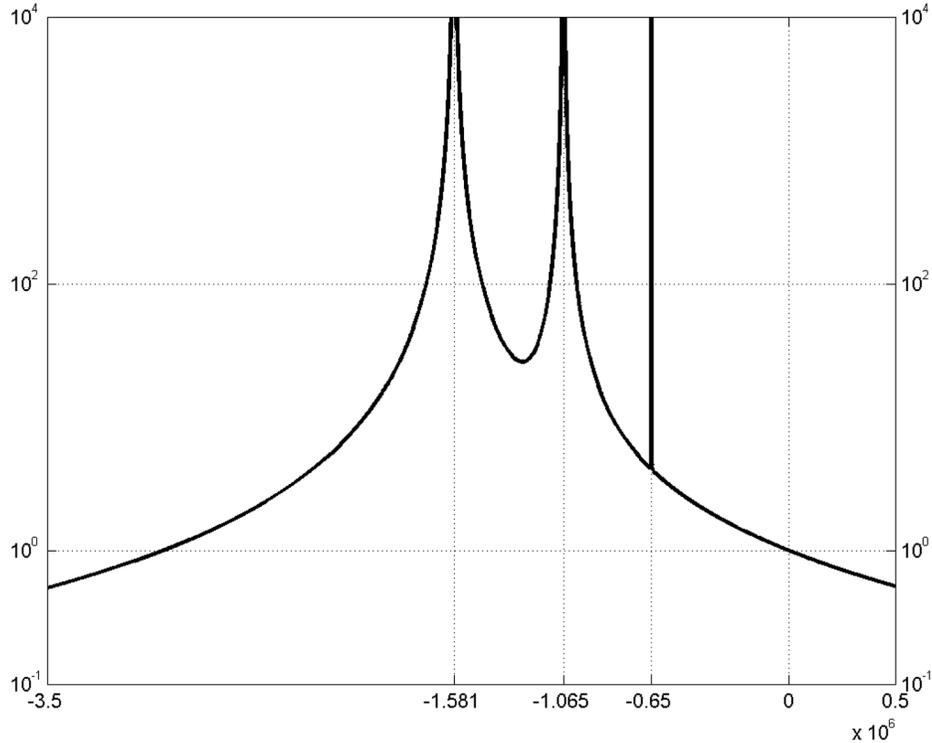


FIG. A-1. Possible Values for the Lagrange Multiplier.

In general, there can be as many as six intersections: one for $\lambda = \lambda_{\max} > -F_1$; zero, one or two in the interval $(-F_2, -F_1)$; zero, one or two in the interval $(-F_3, -F_2)$; and one at $\lambda = \lambda_{\min} < -F_3$. Since $g(\lambda) \rightarrow 0$ as $\lambda \rightarrow \pm\infty$, the solutions λ_{\max} and λ_{\min} will always exist. (The solution λ_{\min} is not visible in Figure A-1, because the abscissa does not extend sufficiently far to the left.)

There are only two solutions for λ for Numerical Example 1, because $g(\lambda) > 1$ for $-F_3 < \lambda < -F_1$. The relative proximity of the asymptotes to the vertical at $\lambda = -F_1$ reflects the smallness of $(n_{uc}^*)_1$ relative to the other two components.

Consider the simple example of $\mathbf{F} = FI_{3 \times 3}$. Then

$$g(\lambda) = \frac{|\mathbf{G}|^2}{(\lambda + F)^2} = \frac{|\mathbf{n}_{uc}|^2}{(1 + \lambda/F)^2} = \frac{1 + \delta}{(1 + \lambda/F)^2} \quad (\text{A3})$$

where δ is the norm excess of the unconstrained unnormalized solution. For our two numerical examples δ had values of approximately 0.0001 and 0.003. The solution of $g(\lambda) = 1$ for our simple examples are

$$\lambda = (-1 \pm \sqrt{1 + \delta})F \quad (\text{A4})$$

and for $|\delta| \ll 1$

$$\lambda = \lambda_{\max} \approx (\delta/2)F \quad \text{and} \quad \lambda = \lambda_{\min} \approx (-2 - \delta/2)F \quad (\text{A5})$$

The multiplier λ_{\max} will be positive, zero or negative according to whether δ is positive, zero or negative. For our simple example

$$\hat{\mathbf{n}}^*(\lambda_{\min}) = -\hat{\mathbf{n}}^*(\lambda_{\max}) \quad (\text{A6})$$

From equations (41) we know that λ_{\max} in our degenerate example will be zero-mean with root-mean-square value

$$\lambda_{\max}^{\text{rms}} = F^{1/2} \quad (\text{A7})$$

Our simple example for \mathbf{F} will arise if we choose $H_k = I_{3 \times 3}$, and $R_k = \sigma^2 I_{3 \times 3}$, $k = 1, \dots, N$. Then

$$F = N/\sigma^2 \quad (\text{A8})$$

For $N = 100$ and $\sigma = 0.01$ rad we obtain

$$F = 10^6 \quad \text{and} \quad \lambda_{\max}^{\text{rms}} = 10^3 \quad (\text{A9})$$

Thus $\lambda = \lambda_{\max}$ will be a $1000\text{-}\sigma$ event! The probability of a $1000\text{-}\sigma$ event is approximately 0.8×10^{-438} . Of course, this probability assumes that the linearization of equation (34) is correct, which won't be the case if λ_{\min} is an acceptable value for the multiplier.

We claim that only λ_{\max} can have physical significance. Consider a continuum of systems with the measurement covariance matrix given now by

$$R_k(\eta) = \eta\sigma^2 I_{3 \times 3}, \quad k = 1, \dots, N \quad (\text{A10})$$

and consider $\hat{\mathbf{n}}^*$ as a function of η . (We assume that the realizations of the noise are identical, except for the scaling by η .) For $\eta = 0$, the noise-free case, we know for any positive-definite \mathbf{F} that $J(\hat{\mathbf{n}})$ has a unique zero, which is just $-\mathbf{F}^{-1}\mathbf{G}$ and corresponds to $\lambda = 0$, which can only be λ_{\max} . (The multiplier λ_{\min} in the degenerate

noise-free example has the value $-2F$ and leads to the value $+\mathbf{F}^{-1}\mathbf{G}$, which is not a root of $J(\hat{\mathbf{n}})$.) Since λ must be a continuous function of η , it follows that the Lagrange multiplier for $\eta = 1$, the actual data, must also be λ_{\max} .

From a practical standpoint, because the second derivative of $g(\lambda)$ becomes infinite at $\lambda = -F_1$, it is extremely unlikely that a Newton-Raphson iteration could lead to λ_2 as a solution unless F_1 is minute. Thus, the possibility that the Lagrangian-multiplier method could have multiple solutions is an observability problem, similar to that of other values of λ being proximate to λ_{\max} in Davenport's q-method [14]. As we saw in reference [20], the separation of the other eigenvalues of Davenport's K -matrix from λ_{\max} was connected directly to the Fisher information matrix of the three-axis attitude. In the present application, if the spin-axis is well observable, then we must have that $F_1 \gg 1$, and λ_2 will be enormously larger in magnitude than λ_{\max} and negative.