Guidance and Control 1990
Advances in the Astronautical Sciences, Vol. 72 **AAS 90-041**

INFLIGHT ESTIMATION OF SPACECRAFT SENSOR ALIGNMENT

Malcolm D. Shuster *

The Johns Hopkins University, Applied Physics Laboratory,
Laurel, Maryland 20723-6099

Several different approaches are reviewed which are applicable to the inflight estimation of attitude sensor alignments. These include: batch alignment estimation, which uses an attitude-independent derived measurement but which requires that data from different sensors be simultaneous within a given frame; sequential alignment estimation using the Kalman filter, which does not have this requirement but is much more sensitive to the inherent nonlinear dependence of the measurement on the attitude; and batch methods in which the derived measurements are generated using a Kalman filter and which have the best properties of both of the previous methods with a smaller computational burden. These three approaches are presented in a common context without numerical examples.

INTRODUCTION

Alignment estimation forms an important part of many missions since the alignment estimation accuracy directly effects the accuracy with which the payload attitude can be determined. Although the estimation of spacecraft attitude has reached a very high level of development, the estimation of misalignments (i.e., the "difference" between the prelaunch and postlaunch alignments) and other calibration parameters is often carried out in a much less rigorous fashion. Not so long ago, inflight alignment estimation was largely neglected. The well-known book of Wertz,[1] for example, which summarizes the technology of attitude determination up to early 1978, does not discuss this topic.

*Senior Professional Staff, Guidance and Control Group, Space Department.

Earliest efforts to re-estimate spacecraft alignments inflight relied on simple least-squares procedures which did not take account of the sensor statistics and sometimes made very strong assumptions about the misalignment of one of the sensors relative to the spacecraft payload.[2–4] These efforts have avoided the need to consider attitude dynamics, firstly, by examining only frames of simultaneous sensor data, and, secondly, by using derived attitude-independent measurements[3–4] (essentially the cosine of the angle included between supposed body-referenced measurements) rather than the measurements themselves. Generally, to skirt a fundamental lack of observability, they have also chosen arbitrarily one set of postlaunch sensor misalignments to vanish, an approximation which can lead to significant errors. Because of the lack of observability, it is often impossible to test the correctness of the inflight estimates. In one case where the project scientist has been able to perform an absolute inflight alignment calibration based on the scientific data,[5] the errors in the naive least-square approach turned out to be large.

The earliest improvement[6] to this procedure came first from an attempt to treat the measurement noise but not the dynamics of the system. However, the measurement noise was treated only approximately. The variances of the derived attitude-independent measurements were calculated correctly but the correlations between these derived measurements was neglected. In addition, the derived measurements were redundant for more than three sensors, making that approach unusable without some additional machinery except in the very restricted case of a spacecraft with three or fewer attitude sensors. For the application of that approach to the Solar Maximum Mission (SMM), where only three sensors were considered at a time, the algorithm was workable although flawed. Nonetheless, the algorithm has been popular and successful in applications to spacecraft supported by NASA Goddard Space Flight Center and continues to be used, although one should treat the confidence bounds on the misalignments calculated using that algorithm with some suspicion. It is, perhaps, worth noting that that work also attempted to take account of the prelaunch calibration results in the on-orbit calibration, although the proper way to do this was not correctly understood at that time. A recent application of that algorithm has been carried out by Snow *et al.*[7] for the UARS spacecraft.

The drawbacks of that method were removed in later work,[8–9] which treated the statistics properly and made no approximations other than the small-angle approximation for the misalignments and the sensor measurement noise. That work included also the estimation of launch-shock parameters from inflight data, which now made the batch attitude-independent approach both statistically complete and rigorous. This methodology has been used also to explore some general questions of misalignment estimation, such as the correctness of estimating only coalignments, the dependence of the inflight alignment calibration accuracies on the sensor field of view, and the effect of unobservable alignment errors on the attitude. This work has been extended recently[10] to determine the temperature dependence of the attitude sensor alignments for the Solar Maximum Mission, with interesting results. Lerner[11] has extended this work to estimate other parameters besides alignments.

The batch attitude-independent approach discussed above is not useful if the data in each frame is not simultaneous. If some reliable method of determining spacecraft attitude rates is available, say from three-axis gyros or from fitted attitude solutions, the data can, of course, be made simultaneous, at least approximately. Such an approach has been used for attitude estimation when the data is not truly simultaneous[2,12]; hence, it would seem reasonable to use this altered data in estimating misalignments. However, for missions with very high attitude accuracy requirements, where gyro errors must be balanced against attitude sensor errors in a Kalman filter, such a method will no longer be adequate. Thus, it is necessary in this case to treat the attitude sensor alignments together with the attitude within the Kalman filter. An early application of Kalman filtering to attitude and alignment estimation was carried out by Murrell,[13] who applied this technique to the multimission spacecraft typified by Landsat and the Solar Maximum Mission. The Kalman filter approach has been presented recently within the context of a particular measurement model,[14] although the method of treating the alignments is, in fact, quite general. A more extensive application has been carried out by Deutschmann and Bar-Itzhack[15] using Bar-Itzhack's implementation of the Kalman filter.[16]

If the attitude sensor measurements were truly linear functions of the attitude (or attitude increment) *without constraint*, the Kalman filter approaches above, in which the sensor alignments along with the attitude were part of the state vector, would be adequate. However, the attitude sensor measurements generally depend non-linearly on the unconstrained attitude variables[†]. It is this non-linear dependence which is the cause of the attitude Kalman filter's sometimes poor convergence. When the filter converges, i. e., when the estimated attitude becomes so close to truth that the attitude sensor measurements can be truly approximated as linear inhomogeneous functions of the attitude errors, the Kalman filter becomes an accurate attitude estimator, or an accurate estimator of attitude and sensor alignments. This convergence, however, may be slow, requiring hundreds of measurement samples before convergence is attained. If data segments are shorter than this, it would not be possible to estimate attitude and misalignments accurately. Thus, it would be helpful to have a method which is less sensitive to the non-linearities of the measurements.

This problem affects not just misalignments but parameter estimation in general. There exist more powerful techniques which have been developed for estimating the parameters of stochastic systems, which are based on maximum-likelihood estimation. These depend on using the Kalman filter not to estimate the parameters themselves but to generate derived measurements whose statistical properties are superior to those of the original measurements. The new derived measurements are then used in a batch estimator to estimate the misalignments. Since the entire data segment is analyzed with a single value of the sensor alignments, the alignment estimates are much less sensitive to the convergence properties of the filter. These methods are based on techniques originally proposed by Gupta and Mehra.[17] When combined with the two-tier filter technique of Friedland,[18] that method leads to very powerful estimation techniques, which have been extended to treat a wide range of parameter types.[19] A recent application to attitude and alignment estimation (as well as orbit determination) has been carried out recently by Maute and Defonte,[20] who applied these techniques to autonomous navigation of geosynchronous spacecraft.

In the present report we examine several approaches in a common framework. By carefully defining the alignments and other quantities and taking advantage of the algebraic and geometric properties of the attitude and alignments, a great reduction of the computational

[†]Thus, three-axis magnetometer measurements may be a linear (but inhomogeneous) function of the attitude matrix (which is subject to six constraints) but a very non-linear function of the Euler angles, the rotation vector, the vector components of the quaternion, the Gibbs vector, or any other unconstrained (three-parameter) representation of the attitude.

burden is possible. The goal of this work is not to make numerical comparisons but to compare the complexities of the different approaches. Our attention is directed largely toward the filter-based algorithms since the static algorithm has been described in great detail.[8–9] However, for the sake of comparison and because it offers a convenient context for reviewing the geometrical character of alignment estimation, that work is reviewed briefly.

BASIC DEFINITIONS

Sensor Referenced Measurements

A spacecraft line-of-sight sensor such as a vector Sun sensor or star tracker measures a direction, $\hat{\mathbf{U}}_{i,k}$, in sensor coordinates, defined to be directed outward from the sensor, which is describable statistically as

$$\hat{\mathbf{U}}_{i,k} = \hat{\mathbf{U}}_{i,k}^{\text{true}} + \Delta\hat{\mathbf{U}}_{i,k} \quad , \tag{1}$$

where $\hat{\mathbf{U}}_{i,k}^{\text{true}}$ is the true value of the direction and $\Delta\hat{\mathbf{U}}_{i,k}$ is the measurement noise. Here i is the sensor index, $i = 1, \ldots, n$, and k is the temporal index, $k = 1, \ldots, N$. We assume that $\Delta\hat{\mathbf{U}}_{i,k}$ is Gaussian, zero-mean, and white, with covariance $R_{\hat{\mathbf{U}}_{i,k}}$. In more compact notation

$$\Delta\hat{\mathbf{U}}_{i,k} \sim \mathcal{N}(\mathbf{0}, R_{\hat{\mathbf{U}}_{i,k}}) \quad . \tag{2}$$

We assume more generally, in fact, that the measurements from different sensors are statistically independent, thus

$$E\{\Delta\hat{\mathbf{U}}_{i,k}\, \Delta\hat{\mathbf{U}}_{i'k'}^T\} = \delta_{ii'}\, \delta_{kk'}\, R_{\hat{\mathbf{U}}_{i,k}} \quad . \tag{3}$$

Here, $E\{\,\cdot\,\}$ denotes the expectation operator.

Because the observations are constrained to be unit vectors, $R_{\hat{\mathbf{U}}_{i,k}}$ must be singular. In particular,

$$R_{\hat{\mathbf{U}}_{i,k}}\, \hat{\mathbf{U}}_{i,k}^{\text{true}} = \mathbf{0} \quad . \tag{4}$$

Clearly, Eqs. (2) through (4) can be true only to lowest order in R. Since R is generally quite small, this level of approximation will be adequate for the purpose of alignment estimation.

That the covariance matrix $R_{\hat{\mathbf{U}}_{i,k}}$ is singular is indicative of the fact that the sensor-referenced unit vectors are derived quantities and not measured directly by the sensors. Instead one actually observes rather complicated functions of the $\hat{\mathbf{U}}_{i,k}$ from which the $\hat{\mathbf{U}}_{i,k}$ are then computed. The unit vector, however, if somewhat artificial, is the most universal and convenient quantity for describing the sensor.

Body-Referenced Vectors and Alignments

If $\hat{\mathbf{W}}_{i,k}$ denotes the measured direction in the spacecraft body frame, then the alignment matrix, S_i , is the proper orthogonal matrix defined by

$$\hat{\mathbf{W}}_{i,k} = S_i\, \hat{\mathbf{U}}_{i,k} \quad , \tag{5}$$

and, therefore,

$$\hat{\mathbf{W}}_{i,k} = S_i\, \hat{\mathbf{U}}^{\rm true}_{i,k} + S_i\, \Delta\hat{\mathbf{U}}_{i,k} \quad , \tag{6}$$

$$\equiv \hat{\mathbf{W}}^{\rm true}_{i,k} + \Delta\hat{\mathbf{W}}_{i,k} \quad . \tag{7}$$

Thus, the body-referenced observations have an error model given by

$$\Delta\hat{\mathbf{W}}_{i,k} \sim \mathcal{N}(\mathbf{0}, R_{\hat{\mathbf{W}}_{i,k}}) \quad , \tag{8}$$

with

$$R_{\hat{\mathbf{W}}_{i,k}} = S_i\; R_{\hat{\mathbf{U}}_{i,k}}\, S_i^T \quad , \tag{9}$$

and

$$R_{\hat{\mathbf{W}}_{i,k}}\, \hat{\mathbf{W}}^{\rm true}_{i,k} = \mathbf{0} \quad . \tag{10}$$

Misalignments

In general, the alignment matrix S_i is not known exactly. Instead, what is known is S_i^o , the alignment matrix determined by the prelaunch alignment calibration. Thus, we are led to define the misalignment matrix, M_i , according to

$$S_i = M_i\, S_i^o \quad . \tag{11}$$

M_i is necessarily orthogonal. Therefore, we define the misalignment vectors, $\boldsymbol{\theta}_i$, according to

$$\begin{aligned} M_i &\equiv e^{[[\boldsymbol{\theta}_i]]} \quad , \\ &= I + \left(\frac{\sin|\boldsymbol{\theta}_i|}{|\boldsymbol{\theta}_i|}\right)[[\boldsymbol{\theta}_i]] + \left(\frac{1-\cos|\boldsymbol{\theta}_i|}{|\boldsymbol{\theta}_i|^2}\right)[[\boldsymbol{\theta}_i]]^2 \quad , \end{aligned} \tag{12}$$

where $e^{\{\cdot\}}$ denotes matrix exponentiation, and $[[\boldsymbol{\theta}]]$ denotes the usual antisymmetric matrix,

$$[[\boldsymbol{\theta}]] \equiv \begin{bmatrix} 0 & \theta_3 & -\theta_2 \\ -\theta_3 & 0 & \theta_1 \\ \theta_2 & -\theta_1 & 0 \end{bmatrix} \quad . \tag{13}$$

Equation (12) is just Euler's formula for the rotation matrix recast as a function of the rotation vector. The angles θ_1, θ_2, θ_3 are the misalignment angles or simply the misalignments[†]. Since the misalignment matrix is generally a very small rotation, the misalignments will be small and we can write

$$M_i = I + [[\boldsymbol{\theta}_i]] + O(|\boldsymbol{\theta}_i|^2) \quad , \tag{14}$$

[†]Do not confuse the subscripts on θ, which label components, with those on $\boldsymbol{\theta}$, which label sensors. To be more consistent, we should write $\boldsymbol{\theta}_i = [\theta_{i1}, \theta_{i2}, \theta_{i3}]^T$. Whenever possible, however, we will avoid such a cumbersome notation, which invites confusion of the component index with the temporal index.

As a rule, we will keep only first-order terms. The measurement equation now becomes finally

$$\hat{\mathbf{U}}_{i,k} = S_i^{o\,T} M_i^T \hat{\mathbf{W}}_{i,k}^{\text{true}} + \Delta\hat{\mathbf{U}}_{i,k} \quad . \tag{15}$$

Dependence of the Measurements on the Attitude

If $\hat{\mathbf{V}}_{i,k}$ denotes the reference vector, i. e., the representation of the measured vector in the primary reference system (for example, geocentric inertial), then the attitude matrix A_k is defined according to

$$\hat{\mathbf{W}}_{i,k}^{\text{true}} = A_k \, \hat{\mathbf{V}}_{i,k}^{\text{true}} \quad , \tag{16}$$

whence,

$$\hat{\mathbf{W}}_{i,k} = A_k \hat{\mathbf{V}}_{i,k} + \Delta\hat{\mathbf{W}}_{i,k} - A_k \Delta\hat{\mathbf{V}}_{i,k} \quad , \tag{17}$$

where $\Delta\hat{\mathbf{V}}_{i,k}$ is the uncertainty in the reference vector, which we assume to be Gaussian, zero-mean, and white. Hence,

$$E\{\Delta\hat{\mathbf{V}}_{i,k} \, \Delta\hat{\mathbf{V}}_{i'k'}^T\} = \delta_{ii'} \, \delta_{kk'} \, R_{\hat{\mathbf{V}}_{i,k}} \quad . \tag{18}$$

From this it follows that the actual sensor measurements are related to the reference vectors by

$$\hat{\mathbf{U}}_{i,k} = S_i^T A_k \hat{\mathbf{V}}_{i,k} + \Delta\hat{\mathbf{U}}_{i,k} - S_i^T A_k \Delta\hat{\mathbf{V}}_{i,k} \quad . \tag{19}$$

We note immediately from Eq. (19) that the values of the measurement vectors are unchanged by the simultaneous transformations

$$S_i \rightarrow T \, S_i \quad , \qquad i = 1, \ldots, n \quad , \tag{20a}$$

$$A_k \rightarrow T \, A_k \quad , \quad k = 1, \ldots, N \quad , \tag{20b}$$

where T is an arbitrary proper orthogonal matrix. Thus, it is impossible from inflight sensor measurements to distinguish a common misalignment of the sensors from a change in the attitude. It is, therefore, impossible to estimate the sensor alignments and the attitude unambiguously from the spacecraft sensor measurements alone, and some additional measurement, e. g., the prelaunch alignment calibration, is needed in order to obtain separate estimates of these quantities. In terms of the misalignments, Eq. (19) becomes

$$\hat{\mathbf{U}}_{i,k} = S_i^{o\,T} M_i^T A_k \hat{\mathbf{V}}_{i,k} + \Delta\hat{\mathbf{U}}_{i,k} - S_i^{o\,T} M_i^T A_k \Delta\hat{\mathbf{V}}_{i,k} \quad . \tag{21}$$

This is the point of departure of the alignment estimation algorithms which we will now consider.

MAXIMUM LIKELIHOOD ESTIMATION

Given a measurement model, i. e., a probability density function of the measurement which also depends on the vector $\mathbf{x}$ which we wish to estimate, the maximum likelihood estimation of $\mathbf{x}$ is straightforward. If $\mathbf{Z}_k'$, $k = 1, \ldots, N$, is a sequence of measurements and $p_{\mathbf{z}_1,\ldots,\mathbf{z}_N}(\mathbf{Z}'_1,\ldots,\mathbf{Z}'_N;\mathbf{x})$ is the joint probability distribution of the measurements as a function of a parameter vector $\mathbf{x}$, then the maximum likelihood estimate[21−22] of $\mathbf{x}$ is given by

$$\mathbf{x}^{*\prime}_{ML} = \arg\max_{\mathbf{x}} p_{\mathbf{z}_1,\ldots,\mathbf{z}_N}(\mathbf{Z}'_1,\ldots,\mathbf{Z}'_N;\mathbf{x}) \quad , \tag{22}$$

that is, the value of $\mathbf{x}$ at which $p_{\mathbf{z}_1,\ldots,\mathbf{z}_N}(\mathbf{Z}'_1,\ldots,\mathbf{Z}'_N;\mathbf{x})$ achieves its maximum.† The maximum likelihood estimate, $\mathbf{x}^{*\prime}_{ML}$, is a function of the values of the measurements, $\mathbf{Z}'_1,\ldots,\mathbf{Z}'_N$. Generally, we reserve the notation $\mathbf{x}^{*}_{ML}$ without the prime for the maximum likelihood estimator, a random variable, which depends on the measurement random variables, $\mathbf{Z}_1,\ldots,\mathbf{Z}_N$, in the same way as the maximum likelihood estimate depends on the values (realizations) of the measurements. The prime will often be discarded in the remaining sections when no confusion will result in order to achieve a less cumbersome notation.

Defining now

$$J(\mathbf{x}) = -\log p_{\mathbf{z}_1,\ldots,\mathbf{z}_N}(\mathbf{Z}'_1,\ldots,\mathbf{Z}'_N;\mathbf{x}) \quad , \tag{23}$$

the negative-log-likelihood function, it follows that the likelihood function is a maximum when the negative-log-likelihood function is a minimum. If $J(\mathbf{x})$ is a differentiable function whose minimum does not lie on the boundary of its domain, then we have also that

$$\frac{\partial J}{\partial \mathbf{x}}(\mathbf{x}^{*\prime}_{ML}) = \mathbf{0} \quad . \tag{24}$$

A sequence of approximations to $\mathbf{x}^{*\prime}_{ML}$ may be obtained by straightforward application of the Newton-Raphson method to yield

$$\mathbf{x}(i+1) = \mathbf{x}(i) - \left[\frac{\partial^2 J}{\partial \mathbf{x}\partial \mathbf{x}^T}(\mathbf{x}(i))\right]^{-1} \frac{\partial J}{\partial \mathbf{x}}(\mathbf{x}(i)) \quad , \tag{25}$$

and under not very restrictive conditions

$$\lim_{i\to\infty} \mathbf{x}(i) = \mathbf{x}^{*\prime}_{ML} \quad . \tag{26}$$

Asymptotically, i. e., as the amount of data increases without bound we may replace the Hessian matrix in Eq. (25) with the Fisher information matrix defined as

$$F_{\mathbf{xx}} \equiv E\left\{\frac{\partial^2 J}{\partial \mathbf{x}\partial \mathbf{x}^T}\right\} \quad . \tag{27}$$

In this same limit the covariance matrix of the estimate error of $\mathbf{x}^{*\prime}_{ML}$ is given by

$$P_{\mathbf{xx}} = F_{\mathbf{xx}}^{-1} \quad . \tag{28}$$

Generally, the Fisher information matrix can be calculated in closed form.

---

†Note that an asterisk is used to designate the estimate or estimator so as not to be confused with the caret used to designate a unit vector.

The above treatment is completely general. The application of maximum likelihood methods to attitude and alignment estimation occupies the remainder of this report.

BATCH ATTITUDE-INDEPENDENT ALIGNMENT ESTIMATION

Attitude-Independent Alignment Measurements

Equation (21) is the starting point for processing the inflight data. We begin by defining an "uncalibrated" body-referenced observation vector, $\hat{\mathbf{W}}^o_{i,k}$, according to

$$\hat{\mathbf{W}}^o_{i,k} \equiv S^o_i \, \hat{\mathbf{U}}_{i,k} = M^T_i \, \hat{\mathbf{W}}_{i,k} \quad , \tag{29}$$

so that

$$\hat{\mathbf{W}}^o_{i,k} = M^T_i \, \hat{\mathbf{W}}^{\rm true}_{i,k} + \Delta\hat{\mathbf{W}}^o_{i,k} \quad , \tag{30}$$

with

$$\Delta\hat{\mathbf{W}}^o_{i,k} \sim \mathcal{N}(\mathbf{0}, R_{\hat{\mathbf{W}}^o_{i,k}}) \quad , \tag{31}$$

and

$$R_{\hat{\mathbf{W}}^o_{i,k}} = S^o_i \; R_{\hat{\mathbf{U}}_{i,k}} \, S^{o\,T}_i + M^T_i A_k R_{\hat{\mathbf{V}}_{i,k}} A^T_k M_i \quad , \tag{32}$$

If we now expand M_i to first order in $\boldsymbol{\theta}_i$, we obtain

$$\begin{aligned} \hat{\mathbf{W}}^o_{i,k} &\simeq (I - [[\,\boldsymbol{\theta}_i\,]])\, \hat{\mathbf{W}}^{\rm true}_{i,k} + \Delta\hat{\mathbf{W}}^o_{i,k} \quad , \\ &= \hat{\mathbf{W}}^{\rm true}_{i,k} + [[\,\hat{\mathbf{W}}^{\rm true}_{i,k}\,]]\, \boldsymbol{\theta}_i + \Delta\hat{\mathbf{W}}^o_{i,k} \quad . \end{aligned} \tag{33}$$

The uncalibrated body-referenced observation, $\hat{\mathbf{W}}^o_{i,k}$, as a function of the misalignments depends even to lowest order on the attitude through $\hat{\mathbf{W}}^{\rm true}_{i,k}$. We would prefer not to solve for $3N$ attitude parameters. For $N = 1000$, a not unreasonable amount of data, this leads to a problem of very high dimension. To process this number of parameters in a batch framework would be quite overwhelming computationally. Thus, we look for a means of removing the attitude dependence of the measurements. To accomplish this we note that to first order in $\boldsymbol{\theta}_i$, $\boldsymbol{\theta}_j$, and the measurement noise terms

$$\begin{aligned} \hat{\mathbf{W}}^o_{i,k} \cdot \hat{\mathbf{W}}^o_{j,k} = \hat{\mathbf{V}}_{i,k} \cdot \hat{\mathbf{V}}_{j,k} &+ (\hat{\mathbf{W}}^o_{i,k} \times \hat{\mathbf{W}}^o_{j,k}) \cdot (\boldsymbol{\theta}_i - \boldsymbol{\theta}_j) \\ &+ \hat{\mathbf{W}}^{\rm true}_{i,k} \cdot \Delta\hat{\mathbf{W}}^o_{j,k} + \hat{\mathbf{W}}^{\rm true}_{j,k} \cdot \Delta\hat{\mathbf{W}}^o_{i,k} \quad , \end{aligned} \tag{34}$$

which is independent of the attitude. Thus, we define for $i \neq j$

$$z_{ij,k} \equiv \hat{\mathbf{W}}^o_{i,k} \cdot \hat{\mathbf{W}}^o_{j,k} - \hat{\mathbf{V}}_{i,k} \cdot \hat{\mathbf{V}}_{j,k} \quad , \tag{35}$$

whence,

$$z_{ij,k} = (\hat{\mathbf{W}}^o_{i,k} \times \hat{\mathbf{W}}^o_{j,k}) \cdot (\boldsymbol{\theta}_i - \boldsymbol{\theta}_j) + \Delta z_{ij,k} \quad , \tag{36}$$

with

$$\Delta z_{ij,k} \simeq \hat{\mathbf{W}}^o_{i,k} \cdot \Delta\hat{\mathbf{W}}^o_{j,k} + \hat{\mathbf{W}}^o_{j,k} \cdot \Delta\hat{\mathbf{W}}^o_{i,k} \quad . \tag{37}$$

In Eq. (37) we have replaced $\hat{\mathbf{W}}^{\text{true}}_{i,k}$ by $\hat{\mathbf{W}}^o_{i,k}$ since this leads to no errors to lowest order in the covariance. The derived measurements, which are the observed cosine errors, are independent of the attitude to first order in the misalignments. The covariance of the $\Delta z_{ij,k}$ can be calculated easily[8] from Eqs. (37) and (32).

The above derived measurements $z_{ij,k}$ are necessarily redundant for $n > 3$ since there are $n(n-1)/2$ pairs of sensors but only $2n-3$ independent attitude-independent measurements. Thus, only a subset of these $n(n-1)/2$ pseudo-measurements can be statistically independent. A suitable subset is given by

$$\mathbf{Z}_k \equiv [z_{12,k}, \, \ldots \, , z_{1n.k}, z_{23,k}, \, \ldots \, , z_{2n,k}]^T \quad , \tag{38}$$

provided that $\hat{\mathbf{W}}^o_{1,k}$ and $\hat{\mathbf{W}}^o_{2,k}$ are not mutually parallel or parallel to any of the remaining unit-vector measurements. A superior method for constructing $\mathbf{Z}_k$, based on the singular value decomposition,[23] is given in Ref. 10. Thus, we write,

$$\mathbf{Z}_k = H_k \,\boldsymbol{\Theta} + \Delta\mathbf{Z}_k \quad , \tag{39}$$

where

$$\boldsymbol{\Theta} \equiv [\,\boldsymbol{\theta}_1^T, \, \ldots \, , \boldsymbol{\theta}_n^T]^T \quad , \tag{40}$$

is the total alignment vector and has dimension $3n$.

Thus, $\Delta\mathbf{Z}_k$ is a white Gaussian sequence with covariance matrix $P_{\mathbf{Z}_k}$. The matrices H_k and $P_{\mathbf{Z}_k}$ are obtained directly from Eqs. (36) and (37). The *a posteriori* inflight estimate of the misalignments, $\boldsymbol{\Theta}^*(+)$, together with the *a posteriori* estimate error covariance, $P_{\Theta\Theta}(+)$, may be obtained straightforwardly by maximum likelihood estimation. The prelaunch (i. e., *a priori*) estimate of the misalignments $\boldsymbol{\Theta}^*(-)$ has by definition the value $\mathbf{0}$. We denote the prelaunch estimate error covariance corrected for launch shock by $P_{\Theta\Theta}(-)$. Then the negative-log-likelihood function†, which in maximum likelihood estimation serves as a cost function, is given by

$$\begin{aligned} J_\Theta(\boldsymbol{\Theta}) = \tfrac{1}{2}\,\big[\,(\boldsymbol{\Theta}^*(-) - \boldsymbol{\Theta})^T P^{-1}_{\Theta\Theta}(-)\,(\boldsymbol{\Theta}^*(-) - \boldsymbol{\Theta}) + \log\,\det\,P_{\Theta\Theta}(-) + 3n\,\log\,2\pi\,\big] \\ + \tfrac{1}{2}\sum_{k=1}^{N}\big[\,(\mathbf{Z}_k - H_k\,\boldsymbol{\Theta})^T\,P^{-1}_{\mathbf{Z}_k}\,(\mathbf{Z}_k - H_k\,\boldsymbol{\Theta}) \\ + \log\,\det P_{\mathbf{Z}_k} + (2n_k - 3)\,\log\,2\pi\,\big] \end{aligned} \tag{41}$$

where $2n_k - 3$ is the dimension of $\mathbf{Z}_k$ if there are fewer than the full complement of sensors active at a given time t_k. Minimizing $J_\Theta(\boldsymbol{\Theta})$ leads to the usual normal equations:

$$P^{-1}_{\Theta\Theta}(+)\;\boldsymbol{\Theta}^*(+) = \sum_{k=1}^{N} H_k^T\,P^{-1}_{\mathbf{Z}_k}\,\mathbf{Z}_k \quad , \tag{42}$$

†This particular form of maximum likelihood estimation is frequently termed maximum *a posteriori* (MAP) estimation

$$P^{-1}_{\Theta\Theta}(+) = P^{-1}_{\Theta\Theta}(-) + \sum_{k=1}^{N} H_k^T \, P^{-1}_{\mathbf{Z}_k} \, H_k \quad . \tag{43}$$

The solution of these equations leads to the batch estimate of the postlaunch alignments.

Note that three linear combinations of the components of $\mathbf{\Theta}^*(+)$ must be given by the prelaunch values (i. e., 0). For this reason one has often estimated instead the relative alignment vector defined by

$$\mathbf{\Theta}^{\mathrm{rel}} \equiv [\,\boldsymbol{\theta}_2^T - \boldsymbol{\theta}_1^T, \, \ldots \, , \boldsymbol{\theta}_n^T - \boldsymbol{\theta}_1^T]^T \quad , \tag{44}$$

a vector of dimension $3n-3$. The estimate of this quantity, sometimes called the coalignment vector, with the additional condition $\boldsymbol{\theta}_1 = \mathbf{0}$ has often been used in place of the full alignments. This approach is justified only if it is known that the postlaunch alignment of sensor 1 relative to the payload should not differ appreciably from the prelaunch value. This, however, is sometimes true, as seems to be the case for the Solar Maximum Mission,[10] but was not true in the case of Magsat.[5]

The prior-free estimate of the relative misalignments, i. e., the estimate of $\mathbf{\Theta}^{\mathrm{rel}}$ which minimizes the summation term of Eq. (41), is important for determining launch shock. Obviously, from Eq. (36), a prior-free estimate of the absolute misalignments does not exist, and it is for this reason that most early works have been content to estimate only prior-free relative alignments.

ESTIMATION OF ALIGNMENTS AS KALMAN-FILTER STATE VARIABLES

Since the Kalman filter can be formulated as a maximum likelihood estimator,[24] the Kalman filter estimate of the alignments is also a maximum likelihood estimate, the one which takes account of the spacecraft dynamical degrees of freedom.

Assume again that the spacecraft is equipped with n vector sensors for which we wish to estimate alignments using the Kalman filter. The complete state vector, $\mathbf{X}(t)$, in the context of combined attitude and alignment estimation is

$$\mathbf{X}(t) = \left[\, \bar{q}^T(t) \, , \boldsymbol{\varepsilon}^T(t) \, , \bar{q}_1^T(t) \, , \, \ldots \, , \bar{q}_n^T(t) \,\right]^T \quad , \tag{45}$$

where $\bar{q}(t)$ is the attitude quaternion, $\bar{q}_i(t)$, $i = 1, \, \ldots \, , n$, are the alignment quaternia, which have the same relation to the respective alignment matrices as the attitude quaternion has to the attitude matrix, and $\boldsymbol{\varepsilon}(t)$ is the gyro bias vector. The inclusion of additional degrees of freedom in the state vector is straightforward but needlessly complicates the present discussion.

The state equations for the attitude and the gyro biases are usually modeled as[13,25]

$$\frac{d}{dt}\,\bar{q}(t) = \frac{1}{2}\Omega(\mathbf{g}(t) - \boldsymbol{\varepsilon}(t) - \boldsymbol{\eta}_1(t))\,\bar{q}(t) \quad , \tag{46a}$$

$$\frac{d}{dt}\,\boldsymbol{\varepsilon}(t) = \boldsymbol{\eta}_2(t) \quad , \tag{46b}$$

where $\mathbf{g}(t)$ is the gyro reading, and $\boldsymbol{\eta}_1(t)$ and $\boldsymbol{\eta}_2(t)$ are white Gaussian processes with power spectral density matrices $Q_1(t)$ and $Q_2(t)$, respectively. $\Omega(\boldsymbol{\omega})$ is the 4×4 matrix

$$\Omega(\boldsymbol{\omega}) \equiv \begin{bmatrix} 0 & \omega_3 & -\omega_2 & \omega_1 \\ -\omega_3 & 0 & \omega_1 & \omega_2 \\ \omega_2 & -\omega_1 & 0 & \omega_3 \\ -\omega_1 & -\omega_2 & -\omega_3 & 0 \end{bmatrix} \quad . \tag{47}$$

The gyro-referenced attitude, $\bar{q}_{\text{ref}}(t)$ satisfies

$$\frac{d}{dt}\bar{q}_{\text{ref}}(t) = \frac{1}{2}\Omega(\mathbf{g}(t) - \boldsymbol{\varepsilon}(0))\,\bar{q}(t) \quad , \tag{48}$$

and the complete state vector based on the gyro-referenced attitude and the prelaunch alignments is, in obvious notation,

$$\mathbf{X}^{\text{ref}}(t) \equiv \left[\,\bar{q}_{\text{ref}}^T(t)\,,\,\boldsymbol{\varepsilon}^T(0)\,,\,\bar{q}_1^{o\,T}(t)\,,\,\ldots\,,\,\bar{q}_n^{o\,T}(t)\,\right]^T \quad , \tag{49}$$

where for uniformity we have written time arguments for the alignment quaternia.

The incremental attitude quaternion is given by

$$\delta\bar{q}(t) = \bar{q}(t) \otimes (\bar{q}_{\text{ref}}(t))^{-1} \quad . \tag{50}$$

In general, $\delta\bar{q}(t)$ will be the quaternion of an infinitesimal rotation, which we may write as

$$\delta\bar{q}(t) = \begin{bmatrix} \boldsymbol{\xi}(t)/2 \\ 1 \end{bmatrix} + O(|\,\boldsymbol{\xi}(t)|^2) \quad . \tag{51}$$

Defining the gyro-bias increment vector by

$$\Delta\,\boldsymbol{\varepsilon}(t) = \boldsymbol{\varepsilon}(t) - \boldsymbol{\varepsilon}(0) \tag{52}$$

leads to the incremental equations

$$\frac{d}{dt}\,\boldsymbol{\xi}(t) = -\big(\mathbf{g}(t) - \boldsymbol{\varepsilon}(0)\big) \times \boldsymbol{\xi}(t) - \Delta\boldsymbol{\varepsilon}(t) - \boldsymbol{\eta}_1(t) \quad , \tag{53a}$$

$$\frac{d}{dt}\,\Delta\boldsymbol{\varepsilon}(t) = \boldsymbol{\eta}_2(t) \quad . \tag{53b}$$

Likewise, we assume that the spacecraft is rigid and the misalignments satisfy

$$\frac{d}{dt}\,\boldsymbol{\theta}_i(t) = 0 \quad , \quad i = 1, \ldots, n \quad . \tag{53c}$$

Thus, we define the incremental state vector as

$$\mathbf{x}(t) \equiv \left[\boldsymbol{\xi}^T(t)\,,\,\Delta\boldsymbol{\varepsilon}^T(t)\,,\,\boldsymbol{\theta}_1^T(t)\,,\,\ldots\,,\,\boldsymbol{\theta}_n^T(t)\right]^T \quad . \tag{54}$$

Hence, the complete state vector has dimension $4(n+1)+3$, while the incremental state vector has dimension $3(n+1)+3$. Note the the composition of the reference complete state vector with the incremental state vector is not simple addition. Note also that in the above formulation of the Kalman filter, the gyro measurements have replaced the dynamical model and the gyro measurement noise has become state process noise.

The discretised incremental state vector satisfies a state equation of the form

$$\mathbf{x}_{k+1} = \Phi_k\, \mathbf{x}_k + \mathbf{w}_k \quad , \tag{55}$$

where $\mathbf{w}_k$ is a discrete white noise process calculated from $\boldsymbol{\eta}_1(t)$ and $\boldsymbol{\eta}_2(t)$ and with covariance matrix Q_k. Φ_k and $\mathbf{w}_k$ must be such that

$$\boldsymbol{\theta}_{i,k+1} = \boldsymbol{\theta}_{i,k} \quad , \quad i = 1, \ldots, n \quad . \tag{56}$$

The state covariance is defined in terms of the incremental state vector

$$P_k = E\{\, \mathbf{x}_k\, \mathbf{x}_k^T \,\} - E\{\, \mathbf{x}_k \,\}\, E\{\, \mathbf{x}_k^T \,\} \quad . \tag{57}$$

and the Kalman filter is mechanized in terms of $\mathbf{x}_k$. The prediction equations are†

$$\mathbf{x}_{k|k-1} = \Phi_{k-1}\, \mathbf{x}_{k-1|k-1} \quad , \tag{58a}$$

$$P_{k|k-1} = \Phi'_{k-1}\, P_{k-1|k-1}\, \Phi'^{T}_{k-1} + Q_{k-1} \quad . \tag{58b}$$

The prediction of the misalignments as given by Eq. (58a) is necessarily

$$\boldsymbol{\theta}_{i,\,k|k-1} = \boldsymbol{\theta}_{i,\,k-1|k-1} \quad . \tag{59}$$

The primes on the transition matrices in Eq. (58b) are a result of the basic non-linearity of the combined attitude-gyro-bias dynamics, which leads to different transition matrices for the incremental state vectors and the incremental state errors.[25] Note that $\boldsymbol{\xi}_k$ is related to the attitude matrix according to

$$A_k = e^{[[\,\boldsymbol{\xi}_k\,]]}\, A_k^{\mathrm{ref}} \quad , \tag{60}$$

which is similar to the equivalent relation for the misalignment vectors,

$$S_{i,k} = e^{[[\,\boldsymbol{\theta}_{i,k}\,]]}\, S_i^o \quad . \tag{61}$$

Thus, we may write the measurement equation as

$$\begin{aligned}
\hat{\mathbf{U}}_{i,k} &= S_{i,k}^T\, A_k\, \hat{\mathbf{V}}_{i,k} + \Delta\hat{\mathbf{U}}_{i,k} - S_{i,k}^T\, A_k\, \Delta\hat{\mathbf{V}}_{i,k} \\
&= S_i^{o\,T}\, e^{-[[\,\boldsymbol{\theta}_{i,k}\,]]}\, e^{[[\,\boldsymbol{\xi}_k\,]]}\, A_k^{\mathrm{ref}}\, \hat{\mathbf{V}}_{i,k} + \Delta\hat{\mathbf{U}}_{i,k} - S_{i,k}^T\, A_k\, \Delta\hat{\mathbf{V}}_{i,k} \\
&\simeq S_i^{o\,T} \left(I + [[\, \boldsymbol{\xi}_k - \boldsymbol{\theta}_{i,k} \,]] \right) A_k^{\mathrm{ref}}\, \hat{\mathbf{V}}_{i,k} + \Delta\hat{\mathbf{U}}_{i,k} - S_i^{o\,T}\, A_k^{\mathrm{ref}}\, \Delta\hat{\mathbf{V}}_{i,k} \\
&= \hat{\mathbf{U}}_{i,k}^{\mathrm{ref}} + S_i^{o\,T}\, [[\, \boldsymbol{\xi}_k - \boldsymbol{\theta}_{i,k} \,]]\, \hat{\mathbf{W}}_{i,k}^{\mathrm{ref}} + \Delta\hat{\mathbf{U}}_{i,k} - S_i^{o\,T}\, A_k^{\mathrm{ref}}\, \Delta\hat{\mathbf{V}}_{i,k} \\
&= \hat{\mathbf{U}}_{i,k}^{\mathrm{ref}} - S_i^{o\,T}\, [[\, \hat{\mathbf{W}}_{i,k}^{\mathrm{ref}} \,]]\, (\boldsymbol{\xi}_k - \boldsymbol{\theta}_{i,k}) + \Delta\hat{\mathbf{U}}_{i,k} - S_i^{o\,T}\, A_k^{\mathrm{ref}}\, \Delta\hat{\mathbf{V}}_{i,k} \\
&= \hat{\mathbf{U}}_{i,k}^{\mathrm{ref}} + C_{i,k}^{\xi}\, \boldsymbol{\xi}_k + C_{i,k}^{\theta}\, \boldsymbol{\theta}_{i,k} + \Delta\hat{\mathbf{U}}_{i,k} - S_i^{o\,T}\, A_k^{\mathrm{ref}}\, \Delta\hat{\mathbf{V}}_{i,k} \quad .
\end{aligned} \tag{62}$$

†To simplify the notation, we do not write an asterisk to denote the estimate or estimator when the subscript makes this identification clear.

where

$$\hat{\mathbf{U}}_{i,k}^{\mathrm{ref}} \equiv S_i^{o\,T}\,\hat{\mathbf{W}}_{i,k}^{\mathrm{ref}} \quad , \tag{63}$$

$$\hat{\mathbf{W}}_{i,k}^{\mathrm{ref}} \equiv A_k^{\mathrm{ref}}\,\hat{\mathbf{V}}_{i,k} \quad , \tag{64}$$

are the sensor-referenced and body-referenced measurements for the reference trajectory,

$$C_{i,k}^{\theta} = -C_{i,k}^{\xi} = S_i^{o\,T}\,[[\,\hat{\mathbf{W}}_{i,k}^{\mathrm{ref}}\,]] \quad . \tag{65}$$

are the measurement sensitivity matrices, and

$$\hat{\mathbf{U}}_{i,k} = \hat{\mathbf{U}}_{i,k}^{\mathrm{ref}} + C_{i,k}\,\mathbf{x}_k + \Delta\hat{\mathbf{U}}_{i,k} - S_i^{o\,T} A_k^{\mathrm{ref}}\,\Delta\hat{\mathbf{V}}_{i,k} \quad , \tag{66}$$

where the submatrices of $C_{i,k}$ vanish except for those which multiply to $\boldsymbol{\xi}$ and $\boldsymbol{\theta}_i$.

In general, the measurements are not the $\hat{\mathbf{U}}_{i,k}$ themselves but scalar functions of the $\hat{\mathbf{U}}_{i,k}$, which we denote by $f_{i,k}(\hat{\mathbf{U}}_{i,k})$. Thus, we define the equivalent scalar measurements as

$$z_k = f_{i,k}(\hat{\mathbf{U}}_{i,k}) - f_{i,k}(\hat{\mathbf{U}}_{i,k}^{\mathrm{ref}}) \quad , \tag{67}$$

$$\simeq H_k\,\mathbf{x}_k + v_k \quad , \tag{68}$$

where, we have expanded Eq. (68) in a Taylor series about $\hat{\mathbf{U}}_{i,k}^{\mathrm{ref}}$ to obtain†

$$H_k = \left(\frac{\partial f_k}{\partial \hat{\mathbf{U}}_{i,k}}\,(\hat{\mathbf{U}}_{i,k}^{\mathrm{ref}}) \right)^T C_{i,k} \quad , \tag{69}$$

$$v_k = \left(\frac{\partial f_k}{\partial \hat{\mathbf{U}}_{i,k}}\,(\hat{\mathbf{U}}_{i,k}^{\mathrm{ref}}) \right)^T \left(\Delta\hat{\mathbf{U}}_{i,k} - S_i^{o\,T} A_k^{\mathrm{ref}}\,\Delta\hat{\mathbf{V}}_{i,k} \right) \quad . \tag{70}$$

The measurement equation is now in a form familiar to us. In principal, we can neglect the index i in labeling the measurements, as we have done in Eqs. (67) through (70) if we choose the temporal index so that each scalar measurement corresponds to a different value of k (the order of truly simultaneous measurements is unimportant). Thus, ideally we should write i_k in place of i and be aware that t_k will sometimes assume the same value for successive values of k. In predicting between equal times the transition matrices will be identity matrices and no process noise will be accumulated.

Thus, we may write the Kalman filter equations for the update step as

$$B_k = H_k\,P_{k|k-1}\,H_k^T + R_k \quad , \tag{71}$$

$$K_k = P_{k|k-1}\,H_k^T\,B_k^{-1} \quad , \tag{72}$$

$$\nu_k = z_k - H_k\,\mathbf{x}_{k|k-1} \quad , \tag{73}$$

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + K_k\,\nu_k \quad , \tag{74}$$

$$P_{k|k} = (I - K_k H_k)\,P_{k|k-1}\,(I - K_k H_k)^T + K_k\,R_k\,K_k^T \quad , \tag{75}$$

---

†We use throughout the convention that the matrix of partial derivatives of a scalar with respect to a column vector is again a column vector; hence, the superscript T in Eqs. (69) and (70).

and R_k is the variance of v_k. The *a posteriori* estimate of the misalignment vector and its covariance are given by

$$\boldsymbol{\Theta}^*(+) = \boldsymbol{\Theta}_{N|N} \quad , \quad P_{\Theta\Theta}(+) = \left(P_{N|N}\right)_{\Theta\Theta} \quad . \tag{76}$$

A word should said about the initialization of the Kalman filter. Single sensor measurements, particularly if they are processed as scalars, are not sufficient to determine the attitude. When the Kalman filter is started at an arbitrary attitude and simply allowed to converge, the lack of observability often leads the state errors to persist at large values for significant times before the state vector finally wanders into the linear region and quadratic convergence is observed. When additional alignment degrees of freedom are available for the peregrinations of the state vector, the rate of convergence will be worse still if the filter is not started off intelligently. For the attitude filter, this convergence problem can often be avoided completely by initializing the attitude and the attitude covariance using some less optimal batch method which is more robust than the Kalman filter. While it is comforting to know that the filter equations are sufficiently robust to converge from any initial condition, this should not be interpreted as the best implementation of the filter in actual data processing.

The initialization of the attitude Kalman filter is complicated by the fact that the reference trajectory must also be determined from the measurements, although once determined it becomes a deterministic trajectory whose stochastic origins are dismissed. Thus, some batch method should used to obtain good initial values for the attitude. Since the measurements are almost always equivalent to unit vectors, several reliable methods[26] are available for this. The gyro measurements are then used to determine the reference trajectory from this initial attitude (no longer regarded as an estimate). In this case the initial value of all components of $\mathbf{x}_{o|o}$ are then zero. Batch algorithms can also be used to obtain $(P_{\xi\xi})_{o|o}$, and the manufacturer's specifications provide at least a bound on the initial covariance for the gyro-bias vector.

A complication to this prescription occurs when we consider the initialization of the alignment covariance, since this depends on launch-shock corrections,[8] which must be computed from the inflight alignment estimates themselves. If one of the sensors is mounted to the same optical bench as the spacecraft payload (as was the case for the Solar Maximum Mission) and it is not expected that inflight degradation of the sensor will be important (this is less likely to be true), then it is justifiable to neglect the misalignment of that sensor relative to the payload and simply determine the relative alignments of other sensors with respect to it. Since the estimation of relative alignments will generally be dominated by the inflight data, the computation of launch-shock error levels would not be important. For Magsat,[2] however, this was not the case, and the value of the launch-shock error levels would have significantly affected the estimates of the absolute alignments, had these techniques been available to apply to that spacecraft. However, the launch-shock error levels are inconvenient to calculate except from the prior-free estimates of the relative alignments[8] and the computation of prior-free estimates with a Kalman filter is tantamount to using infinite initial covariances, which contribute to the convergence problems and loss of numerical significance. We suggest *for the calculation of launch-shock error levels only* that the filter be initialized with large initial covariances for the misalignments (say, 20 or 50 times the anticipated covariances from inflight data) but not with numbers so fantastically large that they result in a tremendous loss of numerical significance in the filter. This sets limits of five or two per cent, respectively, for the accuracy with which the launch-shock error levels can be estimated. It would be very unlikely, however, that these error levels would ever be (or need to be) determined with greater accuracy.

KALMAN-FILTER BASED BATCH ESTIMATION OF SPACECRAFT SENSOR ALIGNMENTS

Even if the measurement and process noise are small, the Kalman filter for the attitude and alignments may converge slowly because of the non-linear dependence of the measurements on the attitude. Also, the filter will be very sensitive to outliers at the beginning of a data segment. Batch algorithms, which process all of the data at once, are less sensitive to outliers and to the non-linear dependence of the negative-log-likelihood function on the parameters being estimated. However, from Eq. (68) we see that all of the measurements are correlated with one another through the correlations in $\mathbf{x}_k$. Thus, not only will the parameter set in a batch estimation procedure be very large because of the large number of attitudes to be computed, but the measurement covariance matrix, if all of the measurements were stacked into one large measurement vector, would be very large and non-diagonal, hence, very difficult to invert.

A method of removing this difficulty was developed by Gupta and Mehra.[17] These authors noted that although the measurements, z_k, are correlated, the innovations, ν_k computed by the Kalman filter are always a white sequence. Hence, instead of finding the value of $\boldsymbol{\Theta}$ which minimizes $J(z_1, \ldots, z_N ; \boldsymbol{\Theta})$ it is sufficient to find the value which minimizes $J(\nu_1, \ldots, \nu_N ; \boldsymbol{\Theta})$. Gupta and Mehra noted also that the Jacobian determinant of the (very high-dimensional) transformation matrix which transforms the column vector containing all the z_k into the column vector containing the corresponding ν_k will be unity. Hence, the two negative-log-likelihood functions will yield the same Fisher information matrix. Thus, we are led to estimate $\boldsymbol{\Theta}$ by minimizing the *a posteriori* negative-log-likelihood function,

$$J(\nu_1, \ldots, \nu_N ; \boldsymbol{\Theta}) = \frac{1}{2} \boldsymbol{\Theta}^T P_{\Theta\Theta}^{-1}(-)\, \boldsymbol{\Theta} + \frac{1}{2} \sum_{k=1}^{N} \left\{ \nu_k^T(\boldsymbol{\Theta})\, B_k(\boldsymbol{\Theta})\, \nu_k(\boldsymbol{\Theta}) + \log \det\, B_k(\boldsymbol{\Theta}) + \log 2\pi \right\} \quad , \tag{77}$$

instead of the negative-log-likelihood function given directly in terms of the z_k, although the two are formally equivalent.

In the present instance the total alignment vector, $\boldsymbol{\Theta}$, is no longer a state variable but a *constant* parameter of the system. The state vector, therefore, is now much reduced in dimension and simply

$$\mathbf{X}_k = \begin{bmatrix} \bar{q}_k \\ \boldsymbol{\varepsilon}_k \end{bmatrix} \quad , \tag{78}$$

and the incremental state vector

$$\mathbf{x}_k = \begin{bmatrix} \boldsymbol{\xi}_k \\ \Delta\, \boldsymbol{\varepsilon}_k \end{bmatrix} \quad , \tag{79}$$

Thus, in implementing Eq. (25), the gradient of the *a posteriori* negative-log-likelihood function in terms of the innovations process and the residual covariance B_k is[17]

$$\frac{\partial J}{\partial \Theta_m} = \left(P_{\Theta\Theta}^{-1}(-)\,\mathbf{\Theta}\right)_m + \sum_{k=1}^{N} \left\{ \frac{\partial \nu_k^T(\mathbf{\Theta})}{\partial \Theta_m} B_k^{-1}(\mathbf{\Theta})\,\nu_k(\mathbf{\Theta}) - \frac{1}{2}\,\nu_k^T(\mathbf{\Theta}) B_k^{-1}(\mathbf{\Theta})\,\frac{\partial B_k(\mathbf{\Theta})}{\partial \Theta_m}\,B_k^{-1}(\mathbf{\Theta})\,\nu_k(\mathbf{\Theta}) + \frac{1}{2}\,\mathrm{tr}\left[B_k^{-1}(\mathbf{\Theta})\,\frac{\partial B_k(\mathbf{\Theta})}{\partial \Theta_m}\right]\right\} \quad , \tag{80}$$

and the corresponding Fisher information matrix is given by[†]

$$F_{\ell m} = \left(P_{\Theta\Theta}^{-1}(-)\right)_{\ell m} + \sum_{k=1}^{N} \left\{ \frac{1}{2}\,\mathrm{tr}\left[B_k^{-1}(\mathbf{\Theta})\,\frac{\partial B_k(\mathbf{\Theta})}{\partial \Theta_\ell}\,B_k^{-1}(\mathbf{\Theta})\,\frac{\partial B_k(\mathbf{\Theta})}{\partial \Theta_m} \right] + E\left\{\left[\frac{\partial \nu_k^T(\mathbf{\Theta})}{\partial \Theta_\ell}\,B_k^{-1}(\mathbf{\Theta})\,\frac{\partial \nu_k(\mathbf{\Theta})}{\partial \Theta_m} \right]\right\}\right\} . \tag{81}$$

The mechanization of the filter now proceeds as before but without the components related to the misalignments, which are now simply constant parameters in the measurements. Equation (61) is now replaced by

$$S_i = M_i\,S_i^o = e^{[[\boldsymbol{\theta}_i]]}\,S_i^o \quad , \tag{82}$$

which is the same as in the batch estimator presented earlier. Equations (67) and (68) now become

$$z_k = f_{i,k}(\hat{\mathbf{U}}_{i,k}) - f_{i,k}(\hat{\mathbf{U}}_{i,k}^{\mathrm{ref}}) \quad , \tag{83}$$

$$\simeq H_k^I\,\mathbf{x}_k + C_k\,\mathbf{\Theta} + v_k \quad , \tag{84}$$

where

$$H_k^I = \left(\frac{\partial f_k}{\partial \hat{\mathbf{U}}_{i,k}}\,(\hat{\mathbf{U}}_{i,k}^{\mathrm{ref}}) \right)^T \left[\, C_{i,k}^{\xi} \;\;\vdots\;\; O_{3\times 3} \,\right] \quad , \tag{85}$$

$$C_k = \left(\frac{\partial f_k}{\partial \hat{\mathbf{U}}_{i,k}}\,(\hat{\mathbf{U}}_{i,k}^{\mathrm{ref}}) \right)^T \left[\, O_{3\times 3} \quad \cdots \quad C_{i,k}^{\theta} \quad \cdots \quad O_{3\times 3} \,\right] \quad , \tag{86}$$

with $C_{i,k}^{\theta}$ and $C_{i,k}^{\xi}$ given still by Eq. (65), and the non-zero entries in C_k occur in the submatrix which multiplies $\boldsymbol{\theta}_i$. The superscript I, distinguishes the measurement sensitivity matrix in Eq. (84) from the related quantity in Eq. (68) *et seq.* and denotes that it represents that component of the measurement which is insensitive to the alignments.

[†]Note that Gupta and Mehra make an error in their derivation of Eq. (81) leading them to include an extraneous term.

To calculate the dependence of ν_k on $\boldsymbol{\Theta}$ we note that because the Kalman filter consists only of linear operations on the state variables we may write

$$\mathbf{x}_{k|k-1} = \mathbf{x}^I_{k|k-1} - T_{k|k-1}\,\boldsymbol{\Theta} \quad , \tag{87a}$$

$$\mathbf{x}_{k|k} = \mathbf{x}^I_{k|k} - T_{k|k}\,\boldsymbol{\Theta} \quad , \tag{87b}$$

where $\mathbf{x}^I_{k|k-1}$ and $\mathbf{x}^I_{k|k}$ are independent of $\boldsymbol{\Theta}$. To determine these alignment-independent state estimates and the alignment sensitivity matrices $T_{k|k-1}$ and $T_{k|k}$ we substitute these expressions into the Kalman filter equations to obtain new filter equations of the form

$$\mathbf{x}^I_{k|k-1} = \Phi^I_{k-1}\,\mathbf{x}^I_{k-1|k-1} \quad , \tag{88}$$

$$P^I_{k|k-1} = \Phi'^{\,I}_{k-1}\,P^I_{k-1|k-1}\,(\Phi'^{\,I}_{k-1})^T + Q^I_{k-1} \quad , \tag{89}$$

$$B^I_k = H^I_k\,P^I_{k|k-1}\,(H^I_k)^T + R_k \quad , \tag{90}$$

$$K^I_k = P^I_{k|k-1}\,(H^I_k)^T\,\left(B^I_k\right)^{-1} \quad , \tag{91}$$

$$\nu^I_k = z_k - H^I_k\,\mathbf{x}^I_{k|k-1} \quad , \tag{92}$$

$$\mathbf{x}^I_{k|k} = \mathbf{x}^I_{k|k-1} + K^I_k\,\nu^I_k \quad , \tag{93}$$

$$P^I_{k|k} = (I - K^I_k H^I_k)\,P^I_{k|k-1}\,(I - K^I_k H^I_k)^T + K^I_k\,R_k\,(K^I_k)^T \quad , \tag{94}$$

and the superscript I on Φ_k and Q_k denote again that these quantities have been likewise truncated. The alignment sensitivity matrices are given by the recursion relations

$$T_{o|o} = 0 \quad , \tag{95a}$$

$$T_{k|k-1} = \Phi^I_{k-1}\,T_{k-1|k-1} \quad , \tag{95b}$$

$$T_{k|k} = (I - K^I_k H^I_k)\,T_{k|k-1} - K^I_k C_k \quad . \tag{95c}$$

The innovation is thus given by

$$\nu_k = z_k - H^I_k \mathbf{x}_{k|k-1} - C_k\,\boldsymbol{\Theta} \quad , \tag{96}$$

$$= \nu^I_k - F_k\,\boldsymbol{\Theta} \quad , \tag{97}$$

where

$$F_k = H^I_k\,T_{k|k-1} + C_k \quad . \tag{98}$$

Thus, the prior-free negative-log-likelihood function for the misalignments is given by[†]

$$J^{\text{prior-free}}(\boldsymbol{\Theta}) = \frac{1}{2}\sum_{k=1}^{N}\left\{\left(\nu^I_k - F_k\,\boldsymbol{\Theta}\right)^T (B^I_k)^{-1}\left(\nu^I_k - F_k\,\boldsymbol{\Theta}\right) + \log\,\det\,B^I_k + \log\,2\pi\right\} \quad . \tag{99}$$

[†]For clarity we write Eq. (99) in matrix form even though the three factors are each scalars.

From this negative-log-likelihood function we may estimate the prior-free relative alignments and the launch-shock error levels as in Ref. 9. The *a posteriori* estimates of the alignments taking into account both the *a priori* estimate and Eq. (99) is then given by the usual normal equations

$$P_{\Theta\Theta}^{-1}(+)\,\mathbf{\Theta}^*(+) = \sum_{k=1}^{N} F_k^T \left(B_k^I\right)^{-1} \nu_k^I \quad , \tag{100a}$$

$$P_{\Theta\Theta}^{-1}(+) = P_{\Theta\Theta}^{-1}(-) + \sum_{k=1}^{N} F_k^T (B_k^I)^{-1} F_k \quad , \tag{100b}$$

which are equivalent to Eqs. (80) and (81) if we note that B_k is independent of $\mathbf{\Theta}$. The values of $\mathbf{\Theta}^*(+)$ and $P_{\Theta\Theta}^{-1}(+)$ from Eqs. (100) correspond exactly to $\mathbf{\Theta}_{N|N}$ and $(P_{N|N})_{\Theta\Theta}$ which would have been obtained using the larger Kalman filter presented in the previous section.

Equivalently, we may instead interpret the truncated innovations, ν_k^I, as effective measurements of $\mathbf{\Theta}$ of the form

$$\nu_k^I = F_k\,\mathbf{\Theta} + \Delta\nu_k^I \quad , \tag{101}$$

with

$$\Delta\nu_k^I \sim \mathcal{N}(0,\, B_k^I) \quad , \tag{102}$$

and estimate the *a posteriori* misalignments in a Kalman filter. Thus,

$$\mathbf{\Theta}_{o|o} = \mathbf{0} \quad , \qquad\qquad (P_{\Theta\Theta})_{o|o} = P_{\Theta\Theta}(-) \quad , \tag{103}$$

$$\mathbf{\Theta}_{k|k-1} = \mathbf{\Theta}_{k-1|k-1} \quad , \qquad (P_{\Theta\Theta})_{k|k-1} = (P_{\Theta\Theta})_{k-1|k-1} \tag{104}$$

for the initialization and prediction steps, and

$$\mathcal{B}_k = F_k\,(P_{\Theta\Theta})_{k|k-1}\,F_k^T + B_k^I \quad , \tag{105}$$

$$\mathcal{K}_k = (P_{\Theta\Theta})_{k|k-1}\,F_k^T\,\mathcal{B}_k^{-1} \quad , \tag{106}$$

$$\nu_k = \nu_k^I - F_k\,\mathbf{\Theta}_{k|k-1} \quad , \tag{107}$$

$$\mathbf{\Theta}_{k|k} = \mathbf{\Theta}_{k|k-1} + \mathcal{K}_k\,\nu_k \quad , \tag{108}$$

$$(P_{\Theta\Theta})_{k|k} = (I - \mathcal{K}_k F_k)\,(P_{\Theta\Theta})_{k|k-1}\,(I - \mathcal{K}_k F_k)^T + \mathcal{K}_k\,B_k^I\,\mathcal{K}_k^T \quad , \tag{109}$$

for the update step. Friedland[18] has shown that the estimates of $\mathbf{\Theta}$ in this last filter correspond at each step to the estimates that would have been obtained from the complete filter (attitude plus misalignments) of the last section. Then, Eq. (76) is true also for these estimates. In addition, the innovation process for the two filters, which we have labelled ν_k in both instances, is, in fact, the same and

$$\mathcal{B}_k = B_k \quad , \tag{110}$$

and

$$\mathbf{x}_{k|k-1} = \mathbf{x}^I_{k|k-1} + T_{k|k-1}\,\boldsymbol{\Theta}_{k|k-1} \quad , \tag{111a}$$

$$\mathbf{x}_{k|k} = \mathbf{x}^I_{k|k} + T_{k|k}\,\boldsymbol{\Theta}_{k|k} \quad , \tag{111b}$$

Thus, the two-tier filter of Friedland permits us to decompose the $\big(3(n+1)+3\big)$-dimensional filter for both attitude and alignments into a 6-dimensional filter for the attitude alone and a $3n$-dimensional filter for the alignments. The computationally intensive prediction step takes place only in the much smaller filter, while the larger filter for the alignments is rather simple in execution. However, to avoid sensitivity to outliers, it is, perhaps, preferable in practice to use Eqs. (100) and recompute the attitudes with $\Theta^*(+)$ as a fixed parameter. Note that Eqs. (87) and (111) correspond to different filters: the first in which $\boldsymbol{\Theta}$ is simply a constant vector of parameters, and the second in which $\mathbf{x}_k$ and $\boldsymbol{\Theta}_k$ together constitute a larger filter state vector.

DISCUSSION

In the two Kalman filter approaches above we have restricted our attention to linearizations about a reference trajectory rather than the extended Kalman filter (EKF), which is linearized about the most recent estimate. While there is certainly sufficient experience that the extended Kalman filter yields reasonable estimates, the covariance simulation of the extended Kalman filter does not necessarily yield the true estimate error covariance, because part of the true estimate error is hidden in the linearization point. In addition, the innovations of the extended Kalman filter are not necessarily white, making the connection with the method of Gupta and Mehra less transparent. Nonetheless, the EKF will almost always be a very practical approximation to the methodology presented here, the only difference in execution being the choice of the point of expansion for the prediction step and the measurements. An example of an EKF implementation for the alignments is given in Ref. 15.

Since the reference trajectory is determined by the initial attitude and the gyros alone, it is necessary for these methods that the gyros be very accurate. The most important error from the gyros will be random walk errors which grow (for a three-axis system) as

$$|\,\boldsymbol{\xi}(t)|_{\rm rms} \equiv [{\rm tr}\, P_{\xi\xi}(t)]^{1/2} \quad , \tag{112}$$

$$\simeq \sqrt{3}\,\alpha\, t^{1/2} \quad , \tag{113}$$

For the linearization to be correct within the limit of accuracy of interest, we must have for the predicted attitude increment

$$|\,\boldsymbol{\xi}(t)^{\rm pred}|^2_{\rm rms} << \sigma_{\rm crit} \quad , \tag{114}$$

where $\sigma_{\rm crit}$ is the critical accuracy required for the attitude or alignments. If Eq. (114) is not true, then the quadratic terms neglected in the filter will be at least comparable to the level of accuracy desired, which is unacceptable. Comparing Eqs. (113) and (114) leads to a condition on α

$$\alpha << \sqrt{\frac{\sigma_{\rm crit}}{3t}} \quad , \tag{115}$$

and t is now the length of the data segment. Taking as typical values

$$t = 1.\ \text{hr} \qquad , \qquad \sigma_{\rm crit} = 1.\ \text{arc sec} \quad , \tag{116}$$

leads to the condition

$$\alpha << 0.07\ \text{deg hr}^{-1/2} \quad , \tag{117}$$

which is within the accuracy of modern gyros. For a recently reported example,[27] α is two orders of magnitude smaller than this value. The methodologies presented here are therefore very applicable.

Much of the simplicity of the formulae which appear in this work is due to the definition of the misalignments and attitude increment. It is this choice, for example, which leads to the measurement sensitivity submatrices for these two quantities differing only by a sign. Arbitrary definitions of these quantities lead to expressions of enormous and unnecessary complexity.

Since the alignments do not appear in the state vector in the Gupta-Mehra method it is a simple manner to compute prior-free estimates of the relative alignments. If the filter is initialized with the attitude covariance being quite large but finite, then the prior-free estimate of the misalignments calculated by the Gupta-Mehra method will exist and have roughly this covariance, although in the limit that the initial attitude covariance is infinite, and with infinite-precision arithmetic, this would not be so. Thus, caution must be exercised lest our approximations or numerical round-off lead us to false conclusions.

The methods of Gupta and Mehra[17] and Friedland[18] can, of course, be applied to the estimation of any bias, not simply misalignments. Thus, in estimating magnetometer biases, we might write

$$\mathbf{B}_k = (I + L)\mathbf{B}_k^{\rm true} + \mathbf{b} \quad , \tag{118}$$

where L is a matrix of scale-factor, misalignment, and nonorthogonality corrections and $\mathbf{b}$ is a constant additive bias. The twelve parameters in L and $\mathbf{b}$ are then joined to the parameter vector $\boldsymbol{\Theta}$. Likewise, we might consider similar corrections to the gyro measurements,

$$\mathbf{g}_k(t) = (I + G)\,\boldsymbol{\omega}_k^{\rm true}(t) + \boldsymbol{\varepsilon}(t) + \boldsymbol{\eta}_1(t) \quad , \tag{119a}$$

$$\frac{d}{dt}\,\boldsymbol{\varepsilon}(t) = \boldsymbol{\eta}_2(t) \quad , \tag{119b}$$

which, but for the matrix G, is identical to the equations used in the text. In the above case, the paramters of the matrix G enter the prediction rather than the measurement equations of the filter, a case also considered by Friedland. To treat this case we must modify Eq. (55) to read

$$\mathbf{x}_{k+1} = \Phi_k\,\mathbf{x}_k + D_k\,\boldsymbol{\Theta} + \mathbf{w}_k \quad , \tag{120}$$

and Eq. (95b) becomes in turn

$$T_{k|k-1} = \Phi_{k-1} T_{k-1|k-1} + D_{k-1} \quad . \tag{121}$$

The calculations then proceed as before.

REFERENCES

1. WERTZ, James R., (ed.), *Spacecraft Attitude Determination and Control*, D. Reidel Publishing Company, Dordrecht, the Netherlands, 1978.
2. ABSHIRE, G., McCUTCHEON, R., SUMMERS, G., and VANLANDINGHAM, F. G, "High Precision Attitude Determination for Magsat," *Proceedings, ESA International Symposium on Spaceflight Dynamics,* Darmstadt, Federal Republic of Germany, April 1981.
3. NIEBUR, D., et al., "HCMM Attitude Analysis and Support Plan," NASA X-581-78-4, March 1978.
4. des JARDINS, R., "In-Orbit Startracker Misalignment Estimates on the OAO," *Proceedings, Symposium of Spacecraft Attitude Determination*, El Segundo, Calif., Sept. 30–Oct. 2, 1969,
5. LANGEL, R., BERBERT, J., JENNINGS, T., and HORNER, R., "Magsat Data Processing: A Report for Investigators," NASA Technical Memorandum 82160, November 1981.
6. SHUSTER, M. D., CHITRE, D. M., and NIEBUR, D. P., "In-Flight Estimation of Spacecraft Attitude Sensor Accuracies and Alignments," *Journal of Guidance and Control*, Vol. 5, No. 4, pp. 339–343, 1982.
7. SNOW, F., KRACK, K., SHEU, Y., and BOSL, W., "Accuracy Study of the Upper Atmosphere Research Satellite (UARS) Definitive Attitude Determination," *Proceedings, Flight Mechanics/Estimation Theory Symposium,* NASA Goddard Space Flight Center, Greenbelt, Maryland, pp. 26–41, May 10–11, 1988.
8. SHUSTER, M. D., and PITONE, D. S., "Batch Estimation of Spacecraft Sensor Alignments, I. Basic Results," submitted to the *Journal of the Astronautical Sciences.*
9. SHUSTER, M. D., BIERMAN, G. J., and PITONE, D. S., "Batch Estimation of Spacecraft Sensor Alignments, II. Factorized Methods and General Properties," submitted to the *Journal of the Astronautical Sciences.*
10. PITONE, D. S., and SHUSTER, M. D., "Attitude Sensor Alignment for the Solar Maximum Mission," to be presented at the *Flight Mechanics/Estimation Theory Symposium,* NASA Goddard Space Flight Center, Greenbelt, Maryland, May 22–23, 1990.
11. LERNER, G. M., "Attitude Sensor Calibration Using Scalar Observations," to appear in the *Journal of the Astronautical Sciences.*
12. FALLON, L. III, HARROP, I. H., and STURCH, C. R., "Ground Attitude Determination and Gyro Calibration Procedures for the HEAO Missions," *Proceedings, AIAA 17th Aerospace Sciences Meeting,* New Orleans, Louisiana, January 1979.
13. MURRELL, J. W., "Precision Attitude Determination for Multimission Spacecraft," Paper No. 78-1248, *Proceedings, AIAA Guidance and Control Conference,* Palo Alto, California, August 7–9, 1978.

14. SHUSTER, M. D., "Kalman Filtering of Spacecraft Attitude and the QUEST Model," to appear in the *Journal of the Astronautical Sciences.*

15. DEUTSCHMANN, J., and BAR-ITZHACK, I. Y., "Extended Kalman Filter for Attitude Estimation of the Earth Radiation Budget Satellite," *Proceedings, Flight Mechanics/Estimation Theory Symposium,* NASA Goddard Space Flight Center, Greenbelt, Maryland, pp. 319–332, May 10–11, 1989.

16. BAR-ITZHACK, I. Y., and OSHMAN, Y., "Recursive Attitude Determination from Vector Observations: Quaternion Estimation," *IEEE Transactions on Aerospace and Electronics Systems,* Vol. AES-21, pp. 128–135, January 1985.

17. GUPTA, N. K., AND MEHRA, R. K., "Computational Aspects of Maximum Likelihood Estimation and Reduction in Sensitivity Function Calculations," *IEEE Transactions in Automatic Control,* Vol. AC-19, No. 6, pp. 774–783, 1974.

18. FRIEDLAND, B., "Treatment of Bias in Recursive Filtering," *IEEE Transactions on Automatic Control,* Vol. AC-17, No. 1, pp. 359–367, 1969.

19. PORTER, D. W., SHUSTER, M. D., GIBBS, B. P., and LEVINE, W. S., "A Partitioned Recursive Algorithm for the Estimation of Dynamical Parameters from Cross-Sectional Data," *Proceedings, 22nd IEEE Conference on Decision and Control,* San Antonio, Texas, 1983.

20. MAUTE, P., and DEFONTE, O., "A System for Autonomous Navigation and Attitude Determination in Geostationary Orbit," *Proceedings, International Symposium on Space Mechanics,* Toulouse, France, November 6–10, 1989.

21. SORENSON, H. W., *Parameter Estimation.* Marcel Dekker, New York, 1980.

22. NAHI, N. E., *Estimation Theory and Applications*, John Wiley & Sons, New York, 1969 (reprint: Robert K. Krieger Publishing Co., New York, 1976).

23. GOLUB, G. H., and Van LOAN, C. F., *Matrix Computations,* The Johns Hopkins University Press, Baltimore, 1983.

24. RAUCH, H. E., TUNG, F., and STRIEBEL, C. T., "Maximum Likelihood Estimation of Linear Dynamic Systems," *AIAA Journal,* Vol. 3, No. 8, pp. 1445–1450, 1965.

25. LEFFERTS, E. J., MARKLEY, F. L., and SHUSTER, M. D., "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control and Dynamics,* Vol. 5, No. 5, Sept.–Oct. 1982, pp. 417–429.

26. SHUSTER, M. D., and OH, S. D, "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control,* Vol. 4, No. 1, pp. 70–77, January-February, 1981.

27. FERNANDEZ, M., EBNER, B., AND DAHLEN, N., "Zero-Lock Laser Gyro," Paper AAS 89–024, *12th Annual AAS Guidance and Control Conference,* Keystone, Colorado, February 4–8, 1989; reprinted in *Advances in the Astronautical Sciences,* Vol. 68, pp. 235–241, 1990.