# A Predictive Attitude Determination Algorithm

John L. Crassidis
Department of Mechanical Engineering
The Catholic University of America
Washington, DC 20064

F. Landis Markley
Guidance, Navigation, and Control Branch, Code 712
Goddard Space Flight Center
Greenbelt, MD 20771

## Abstract

In this paper, a new and efficient algorithm is developed for attitude determination from vector observations. The new algorithm, called the Predictive Attitude Determination (PAD) algorithm, is derived from a general nonlinear predictive filter approach. Traditional deterministic algorithms are shown to be suboptimal for anisotropic measurement errors. The major advantage of the PAD algorithm is that it can be easily be applied to the case where anisotropic measurement errors exist. Also, an analytical expression is derived for the steady-state attitude error covariance, which is shown to be equivalent to the optimal covariance derived from maximum likelihood techniques. Simulation studies indicate that the new algorithm is able to accurately determine the attitude of a spacecraft, even for radically anisotropic measurement errors.

## Introduction

Attitude determination refers to the identification of a proper orthogonal rotation matrix so that the measured observations in the body frame equal the reference frame observations mapped by that matrix into the body frame. If all the measured and reference vectors are error free, then the rotation (attitude) matrix is the same for all sets of observations. However, if measurements errors exist, such as noise, then a least-squares type method must be used. For this case, the most common method for determining the attitude matrix uses a loss function first posed by Wahba [1]. This problem involves finding an orthogonal rotation matrix which minimizes the weighted sum of the squares of the observation residuals.

Since its origination in 1965, there have been many algorithms developed which minimize Wahba's loss function. The first practical method was given by Davenport's q-method [2], which solves for the quaternion representation of the rotation matrix. However, this method requires an eigenvalue/eigenvector decomposition of a dimension 4 matrix, which may be computationally intense. A more efficient method was proposed by Shuster [3], called the QUEST algorithm, which simplifies the q-method approach by solving for the components of a Gibbs vector, and uses the fact that any meromorphic function of a dimension 3 matrix can be represented as a quadratic in that matrix. Other methods solve for the attitude matrix directly (e.g., see Refs. [4-5]). In particular, the FOAM algorithm [5] has been shown to be comparable to QUEST in computational speed, and has also been shown to be more robust in some cases. Still other methods which address Wahba's problem can be found in Ref. [6-8].

In Wahba's problem each vector residual is weighted by a scalar number to reflect the relative importance of each sensor. Shuster [9] has shown that Wahba's problem is equivalent to a maximum likelihood estimation problem, where the scalar weight is equal to the scalar inverse variance of the measurement error process. Shuster [10] has further shown that a scalar variance is a good approximation of the actual measurement errors, except in the case where the measurement errors are radically anisotropic. An anisotropic measurement error may be produced by a single-axis failure or degradation in the sensor. This was evident for the Hubble spacecraft where single axis failures occurred on both three-axis magnetometers in 1992 (unfortunately, the failures occurred on the same axis). If these corrupt measurements are used for attitude determination, then the solution to Wahba's problem using scalar weighting is not optimal (as will be shown). Therefore, algorithms such as QUEST and FOAM can produce suboptimal attitude solutions.

In this paper, a new and efficient algorithm is derived which determines the attitude for both isotropic and anisotropic measurement errors. The algorithm is based on a predictive filtering scheme for nonlinear systems [11]. This scheme has been successfully applied to estimate the attitude of a spacecraft using a dynamic model for rate information [12]. The predictive filter developed in this paper is essentially reduced to a deterministic approach, since the corrections required to update the model are not weighted in the predictive filter loss function. Therefore, the new algorithm is known as a Predictive Attitude Determination (PAD) algorithm. Also, an analytical expression is derived for the attitude error covariance. It will be further shown that when the measurement errors are isotropic, the PAD steady-state attitude error covariance is identical to the QUEST covariance in Ref. [3].

The organization of this paper proceeds as follows. First, a summary of Wahba's problem is shown. Then, Wahba's problem is generalized for anisotropic measurement errors. Also, attitude covariance expressions are shown for the original and generalized loss functions. Next, a brief review of the predictive filter for nonlinear systems is shown. Then, the PAD algorithm and covariance expression are developed. Finally, the PAD algorithm is used to determine a simulated spacecraft's attitude using two star trackers as sensors, with a single-axis failure in one tracker.

## Background

In this section, Wahba's problem is reviewed. Also, a generalized version of Wahba's problem is shown, which involves anisotropic measurement errors. A covariance expression is also derived for the generalized problem. Wahba's original problem, modified to include the covariance weighting [9], involves finding a proper orthogonal matrix $A$ that minimizes the following loss function

$$J(A) = \frac{1}{2}\sum_{i=1}^{n} \sigma_i^{-2} \left(\underline{\tilde{b}}_i - A\underline{r}_i\right)^T \left(\underline{\tilde{b}}_i - A\underline{r}_i\right) \tag{1}$$

where $\underline{r}_i$ are the representations of the directions to some observed object, $\underline{\tilde{b}}_i$ are the measured observations in the spacecraft body frame (the tilde denotes measurement), $\sigma_i$ are the standard deviations of the corresponding measurement errors, and $n$ is the number of observations. Since the attitude matrix is assumed orthogonal, the loss function in Equation (1) can be shown to be equivalent to minimizing the following loss function

$$J(A) = -\text{trace}\left(A B^T\right) \tag{2}$$

where

$$B = \sum_{i=1}^{n} \sigma_i^{-2} \left(\underline{\tilde{b}}_i \, \underline{r}_i^T\right) \tag{3}$$

A convenient expression for the attitude matrix is the quaternion representation, defined as [13]

$$\underline{q} \equiv \begin{bmatrix} \underline{q}_{13} \\ q_4 \end{bmatrix} \tag{4}$$

with

$$\underline{q}_{13} \equiv \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \underline{\hat{e}} \sin\left(\theta / 2\right) \tag{5a}$$

$$q_4 = \cos\left(\theta / 2\right) \tag{5b}$$

where $\underline{\hat{e}}$ is a unit vector corresponding to the axis of rotation and $\theta$ is the angle of rotation. The quaternion satisfies a single constraint, given by

$$\underline{q}^T \underline{q} = \underline{q}_{13}^T \underline{q}_{13} + q_4^2 = 1 \tag{6}$$

The attitude matrix is related to the quaternion by

$$A\left(\underline{q}\right) = -\Xi^T\left(\underline{q}\right)\Psi\left(\underline{q}\right) \tag{7}$$

where

$$\Xi\left(\underline{q}\right) \equiv \begin{bmatrix} q_4 I_{3\times 3} + \left[\underline{q}_{13}\times\right] \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ -\underline{q}_{13}^T \end{bmatrix} \tag{8a}$$

$$\Psi\left(\underline{q}\right) \equiv \begin{bmatrix} -q_4 I_{3\times 3} + \left[\underline{q}_{13}\times\right] \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \underline{q}_{13}^T \end{bmatrix} \tag{8b}$$

where $I_{3\times 3}$ is a $3\times 3$ identity matrix. The $3\times 3$ matrix $\left[\underline{q}_{13}\times\right]$ is referred to as cross product matrix since $\underline{a}\times\underline{b} = \left[\underline{a}\times\right]\underline{b}$, with

$$\left[\underline{a}\times\right] \equiv \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \tag{9}$$

From Equation (2) it is clear that the quaternion representation leads to a loss function that is quadratic in the quaternions. An efficient algorithm which minimizes this loss function is given by the QUEST algorithm [3].

As seen in Equation (1), the $3\times 1$ measurement errors are assumed to be isotropic, (i.e., the covariance is assumed to be given by a scalar times the identity). The generalized version of Wahba's loss function for anisotropic errors can be derived using maximum likelihood. Assuming a Gaussian distribution for the error process leads to the following generalized loss function

$$J\left(A\right) = \frac{1}{2}\sum_{i=1}^{n}\left(\underline{\tilde{b}}_i - A\underline{r}_i\right)^T R_i^{-1}\left(\underline{\tilde{b}}_i - A\underline{r}_i\right) \tag{10}$$

where $R_i$ represents the measurement error covariance matrix of the $i^{\text{th}}$ measurement (for a detailed discussion of the measurement error model see Ref. [9]). From Equation (10), it is clear that an anisotropic covariance matrix leads to a quartic dependence if the quaternion representation is used. A general solution can be found be using a nonlinear least-squares approach, but this may be extremely computational intense. Another method involves finding a scalar value which minimizes the error between the loss functions in Equations (10) and Equation (1) [14]. Therefore, algorithms such as the q-method [2] and QUEST [3] may be used, but may produce suboptimal attitude solutions. Other methods which determine the attitude matrix, such as FOAM [5], also yield sub-optimal solutions in this case. The error introduced by using the scalar approach can be investigated by deriving its attitude covariance error, which can be shown to be given by [14]

$$P_{\text{body}} = E\left\{\delta\underline{\alpha}^T \delta\underline{\alpha}\right\} \approx \left[\sum_{i=1}^{n} a_i \left[\underline{\tilde{b}}_i \times\right]^2\right]^{-1} \sum_{i=1}^{n} a_i^2 \left[\underline{\tilde{b}}_i \times\right] R_i \left[\underline{\tilde{b}}_i \times\right]^T \left[\sum_{i=1}^{n} a_i \left[\underline{\tilde{b}}_i \times\right]^2\right]^{-1} \tag{11}$$

where $\delta\underline{\alpha}$ represents a small angle error, and $E\{\ \}$ denotes expectation. In actuality, $\underline{\tilde{b}}_i$ should be replaced with $A_{\text{true}}\, \underline{r}_i$, but Equation (11) is extremely accurate for low noise. Note that if $R_i = \sigma_i^2 I$, setting $a_i = \sigma_i^{-2}$ yields

$$P_{\text{body}} \approx -\left[\sum_{i=1}^{n} \sigma_i^{-2} \left[\underline{\tilde{b}}_i \times\right]^2\right]^{-1} \tag{12}$$

Therefore, in this case the covariance in Equation (11) would be identical to the covariance given by QUEST [3].

An attitude error covariance can also be derived from the generalized loss function in Equation (10). This is accomplished by using results from maximum likelihood estimation [9]. The Fisher information matrix for a parameter vector $\underline{x}$ is given by

$$F_{xx} = E\left\{\frac{\partial}{\partial \underline{x}\, \partial \underline{x}^T} J\left(\underline{x}\right)\right\}_{\underline{x}_{\text{true}}} \tag{13}$$

where $J\left(\underline{x}\right)$ is the negative log likelihood function, which is the loss function in this case. If the measurements are Gaussian and linear in the parameter vector, then the error covariance is given by

$$P_{xx} = F_{xx}^{-1} \tag{14}$$

Now, the attitude matrix is approximated by

$$A = e^{-[\delta\underline{\alpha}\times]} A_{\text{true}} \approx \left(I - [\delta\underline{\alpha}\times] + \frac{1}{2}[\delta\underline{\alpha}\times]^2\right) A_{\text{true}} \tag{15}$$

Equations (15) is next substituted into Equation (10) to determine the Fisher information matrix. First-order terms vanish in the partials, and third-order terms become zero since $E\left\{\delta\underline{\alpha}\right\} = \underline{0}$. Also, assuming that the quartic terms are negligible leads to the following form for the optimal covariance

$$P_{\text{opt}} \approx \left[ \sum_{i=1}^{n} \left[ \underline{\tilde{b}}_i \times \right] R_i^{-1} \left[ \underline{\tilde{b}}_i \times \right]^T \right]^{-1} \tag{16}$$

Note that the optimal covariance in Equation (16) reduces to the covariance in Equation (12) if the condition $R_i = \sigma_i^2 I$ is true. Also, the diagonal elements of covariance in Equation (16) are always smaller or equal to the corresponding elements in Equation (11). Therefore, methods which minimize Wahba's original loss function for anisotropic measurement errors can produce suboptimal results.

## Predictive Attitude Determination

In this section, the predictive attitude determination (PAD) algorithm is derived. First, a brief review of the nonlinear predictive filter is shown (see Ref. [11] for more details). Then, the filter algorithm is reduced to a deterministic-type approach for attitude determination. Finally, a covariance expression for the attitude errors using PAD are derived.

**Predictive Filtering**

In the nonlinear predictive filter it is assumed that the state and output estimates are given by a preliminary model and a to-be-determined model error vector, given by

$$\dot{\underline{\hat{x}}}(t) = \underline{f}\left(\underline{\hat{x}}(t), t\right) + G(t)\,\underline{d}(t) \tag{17a}$$

$$\underline{\hat{y}}(t) = \underline{c}\left(\underline{\hat{x}}(t), t\right) \tag{17b}$$

where $\underline{f}$ is a $p \times 1$ model vector, $\underline{\hat{x}}(t)$ is a $p \times 1$ state estimate vector, $\underline{d}(t)$ is a $l \times 1$ model error vector, $G(t)$ is a $p \times l$ model-error distribution matrix, $\underline{c}$ is a $m \times 1$ measurement vector, and $\underline{\hat{y}}(t)$ is a $m \times 1$ estimated output vector. State-observable discrete measurements are assumed for Equation (17b) in the following form

$$\underline{\tilde{y}}(t_k) = \underline{c}\left(\underline{x}(t_k), t_k\right) + \underline{v}(t_k) \tag{18}$$

where $\underline{\tilde{y}}(t_k)$ is a $m \times 1$ measurement vector at time $t_k$, $\underline{x}(t_k)$ is the true state vector, and $\underline{v}(t_k)$ is a $m \times 1$ measurement noise vector which is assumed to be a zero-mean, Gaussian white-noise distributed process with

$$E\left\{\underline{v}(t_k)\right\} = \underline{0} \tag{19a}$$

$$E\left\{\underline{v}(t_k)\,\underline{v}^T(t_{k'})\right\} = R\,\delta_{kk'} \tag{19b}$$

where $R$ is a $m \times m$ positive-definite covariance matrix.

A loss functional consisting of the weighted sum square of the measurement-minus-estimate residuals plus the weighted sum square of the model correction term is minimized, given by

$$J = \frac{1}{2}\left\{\underline{\tilde{y}}(t_{k+1}) - \underline{\hat{y}}(t_{k+1})\right\}^T R^{-1} \left\{\underline{\tilde{y}}(t_{k+1}) - \underline{\hat{y}}(t_{k+1})\right\} + \frac{1}{2}\underline{d}^T(t_k)\,W\,\underline{d}(t_k) \tag{20}$$

where $W$ is a $l \times l$ weighting matrix. The necessary conditions for the minimization of Equation (20) lead to the following model error solution

$$\underline{d}(t_k) = -\left\{\left[\Lambda(\Delta t) S(\hat{\underline{x}}_k)\right]^T R^{-1} \Lambda(\Delta t) S(\hat{\underline{x}}_k) + W\right\}^{-1} \left[\Lambda(\Delta t) S(\hat{\underline{x}}_k)\right]^T R^{-1} \left[\underline{z}(\hat{\underline{x}}_k, \Delta t) - \tilde{\underline{y}}(t_{k+1}) - \hat{\underline{y}}(t_k)\right] \tag{21}$$

where $\hat{\underline{x}}_k \equiv \hat{\underline{x}}(t_k)$, $\Delta t$ is the measurement sampling interval, $S(\hat{\underline{x}})$ is a $m \times l$ dimensional matrix, and $\Lambda(\Delta t)$ is a $m \times m$ diagonal matrix with elements given by

$$\lambda_{ii} = \frac{\Delta t^{p_i}}{p_i!}, \quad i = 1, 2, \ldots, m \tag{22}$$

where $p_i$, $i = 1, 2, \ldots, m$, is the lowest order of the derivative of $c_i(\hat{\underline{x}}(t))$ in which any component of $\underline{d}(t)$ first appears due to successive differentiation and substitution for $\dot{\hat{x}}_i(t)$ on the right side. The $i^{\text{th}}$ component of $\underline{z}(\hat{\underline{x}}, \Delta t)$ is given by

$$z_i(\hat{\underline{x}}, \Delta t) = \sum_{k=1}^{p_i} \frac{\Delta t^k}{k!} L_f^k(c_i) \tag{23}$$

where $L_f^k(c_i)$ is the $k^{\text{th}}$ Lie derivative, defined by

$$\begin{aligned} L_f^k(c_i) &= c_i && \text{for } k = 0 \\ L_f^k(c_i) &= \frac{\partial L_f^{k-1}(c_i)}{\partial \hat{\underline{x}}} \underline{f} && \text{for } k \geq 1 \end{aligned} \tag{24}$$

The $i^{\text{th}}$ row of $S(\hat{\underline{x}})$ is given by

$$s_i = \left\{L_{g_1}\left[L_f^{p_i - 1}(c_i)\right], \ldots, L_{g_l}\left[L_f^{p_i - 1}(c_i)\right]\right\}, \quad i = 1, 2, \ldots, m \tag{25}$$

where $g_j$ is the $j^{\text{th}}$ column of $G(t)$, and the Lie derivative is defined by

$$L_{g_j}\left[L_f^{p_i - 1}(c_i)\right] \equiv \frac{\partial L_f^{p_i - 1}(c_i)}{\partial \hat{\underline{x}}} g_j, \quad j = 1, 2, \ldots, l \tag{26}$$

Equation (26) is in essence a generalized sensitivity matrix for nonlinear systems. Therefore, given a state estimate at time $t_k$, then Equation (21) is used to process the measurement at time $t_{k+1}$ to find the $\underline{d}(t_k)$ to be used in $[t_k, t_{k+1}]$ to propagate the state estimate to time $t_{k+1}$. The weighting matrix $W$ serves to weight the relative importance between the propagated model and measured quantities. If this matrix is set to zero, then no weight is placed on the model corrections.

**PAD Algorithm**

In the PAD algorithm it is assumed that the model is given by the quaternion kinematics model. PAD requires no dynamics model; it assumes that the attitude rate is adequately modeled by a constant model error $\underline{d}$ between measurements, so that

$$\dot{\hat{\underline{q}}} = \frac{1}{2}\Xi\left(\hat{\underline{q}}\right)\underline{d} \tag{27}$$

where $\hat{q}$ denotes the determined quaternion. Since the body measurements $\left(\tilde{\underline{b}}_i\right)$ are used as the required tracking trajectories, the output vector in Equation (18) is given by (dropping the subscript $i$ for the moment)

$$\underline{c}\left(\hat{\underline{x}}\right) = A\left(\hat{\underline{q}}\right)\underline{r} \tag{28}$$

The lowest order time derivative of $\hat{\underline{q}}$ in Equation (28) in which any component of $\underline{d}$ first appears in Equation (27) is one, so that $p_i = 1$. Therefore, the $\Lambda$ and $\underline{z}$ quantities formed from Equations (22) and (23) are given by

$$\Lambda = \Delta t\, I \tag{29a}$$

$$\underline{z}\left(\hat{\underline{x}}, \Delta t\right) = \underline{0} \tag{29b}$$

The derivative of Equation (28) with respect to $\hat{\underline{q}}$ can be shown to be given by

$$\frac{\partial \underline{c}}{\partial \hat{\underline{q}}} = -2\,\Xi^T\left(\hat{\underline{q}}\right)\Gamma\left(\underline{r}\right) \tag{30}$$

where

$$\Gamma\left(\underline{r}\right) \equiv \begin{bmatrix} -\left[\underline{r}\times\right] & \vdots & -\underline{r} \\ \cdots\cdots & \vdots & \cdots\cdots \\ \underline{r}^T & \vdots & 0 \end{bmatrix} \tag{31}$$

Therefore, the $S$ matrix in Equation (21), which is formed using Equation (25) is given by

$$S = -\Xi^T\left(\hat{\underline{q}}\right)\Gamma\left(\underline{r}\right)\Xi\left(\hat{\underline{q}}\right) = \left[A\left(\hat{\underline{q}}\right)\underline{r}\times\right] \tag{32}$$

The $3\times 3$ matrix $\left[A\left(\hat{\underline{q}}\right)\underline{r}\times\right]$ is analogous to the sensitivity matrix used in a Kalman filter (see Ref. [15]). This matrix has rank 2, which reflects the fact that there is no information about rotations around the current measurement vector. For a deterministic attitude solution the weighting matrix $W$ is set to zero in Equation (21). Therefore, the extension for multiple vector measurement sets, assuming that the errors between vector measurement sets are uncorrelated, is given by the following model error

$$\underline{d}\left(t_k\right) = \frac{1}{\Delta t}\left\{\sum_{i=1}^{n}\left[A\left(\hat{\underline{q}}_k\right)\underline{r}_i\times\right]R_i^{-1}\left[A\left(\hat{\underline{q}}_k\right)\underline{r}_i\times\right]\right\}^{-1}\sum_{i=1}^{n}\left[A\left(\hat{\underline{q}}_k\right)\underline{r}_i\times\right]R_i^{-1}\left(\tilde{\underline{b}}_i\left(t_{k+1}\right) - A\left(\hat{\underline{q}}_k\right)\underline{r}_i\right) \tag{33}$$

The inverse expression in Equation (33) exists only if at least two of the vector observations are not parallel, which is equivalent to all methods which solve Wahba's problem. The determined quaternion can be found by integrating Equation (27) from time $t_k$ to $t_{k+1}$. Since $\underline{d}$ is constant over this interval, the discrete propagation for Equation (27) found in Ref. [16] can be used. It should be noted that Equation (33) represents an exact linearization for an interval $\Delta t$ (see Ref. [17]). However, for practical applications the sampling rate should be well below Nyquist's limit [18].

In order to derive an attitude error covariance from Equation (33), a propagated expression can be derived using a similar approach found in Ref. [15]. Assuming continuous measurements and small $\Delta t$, the propagated attitude error covariance can be shown to be given by ([15], [19])

$$\dot{P} = -\left[\underline{d}\times\right]P - P\left[\underline{d}\times\right]^T + E\left\{\underline{\eta}\,\underline{\eta}^T\right\} - P\,S^T R^{-1} S\,P/\Delta t \tag{34}$$

where

$$S = \left\{\left[A(\underline{q})\underline{r}_1\times\right]^T \;\vdots\; \left[A(\hat{\underline{q}})\underline{r}_2\times\right]^T \;\vdots\; \cdots \;\vdots\; \left[A(\hat{\underline{q}})\underline{r}_n\times\right]^T\right\}^T \tag{35}$$

The measurements in PAD play two roles: 1) dynamics model replacement, similar to gyros used in [15], and 2) actual attitude update, similar to line-of-sight observations used in [15]. Thus, measurement errors contribute to both $E\left\{\underline{\eta}\,\underline{\eta}^T\right\}$ (similar to process noise) and to the last term in Equation (34) (similar to the usual continuous-time Kalman filter). The $\Delta t$ term is due to the conversion of the discrete-time measurements to continuous time ([19], [20]) (note, in the limit $\Delta t \rightarrow 0$). The term $E\left\{\underline{\eta}\,\underline{\eta}^T\right\}$ is a covariance due to the continuous-time measurements, with

$$\underline{\eta} = \frac{1}{\Delta t}\left\{\sum_{i=1}^{n}\left[A(\hat{\underline{q}})\underline{r}_i\times\right]\bar{R}_i^{-1}\left[A(\hat{\underline{q}})\underline{r}_i\times\right]^T\right\}^{-1}\sum_{i=1}^{n}\left[A(\hat{\underline{q}})\underline{r}_i\times\right]\bar{R}_i^{-1}\underline{v}_i \tag{36}$$

where $\underline{v}_i$ is the continuous measurement noise with covariance $\bar{R}_i$. The expectation in Equation (34) is therefore given by

$$E\left\{\underline{\eta}\,\underline{\eta}^T\right\} = \frac{1}{\Delta t^2}\left\{\sum_{i=1}^{n}\left[A(\hat{\underline{q}})\underline{r}_i\times\right]\bar{R}_i^{-1}\left[A(\hat{\underline{q}})\underline{r}_i\times\right]^T\right\}^{-1} \tag{37}$$

Next, assuming that the transients in Equation (34) decay quickly leads to the following steady-state condition

$$E\left\{\underline{\eta}\,\underline{\eta}^T\right\} \approx P\,S^T R^{-1} S\,P/\Delta t \tag{38}$$

Solving for $P$ and converting to discrete-time (i.e., using the approximation $\Delta t\,P$ [21]), yields the following discrete-time steady-state covariance

$$P_{\text{pad}} \approx \left[\sum_{i=1}^{n}\left[A(\hat{\underline{q}}_k)\underline{r}_i\times\right]R_i^{-1}\left[A(\hat{\underline{q}}_k)\underline{r}_i\times\right]^T\right]^{-1} \tag{39}$$

This expression is essentially equivalent to the optimal covariance for the generalized version of Wahba's problem, shown by Equation (16). Therefore, the PAD algorithm is in essence equivalent to solving Wahba's generalized loss function. Also note that the approximation in Equation (39) is valid only for small $\Delta t$ (i.e., well below Nyquist's limit).

If the measurement errors are isotropic for each vector observation, then the model error in Equation (33) can be rewritten by setting $R_i = \sigma_i^2 I$. Noting that $[\underline{a}\times]\underline{a} = \underline{0}$ for any $\underline{a}$ leads to the following simple model error solution

$$\underline{d}(t_k) = \frac{1}{\Delta t}\left\{\sum_{i=1}^{n} \sigma_i^{-2}\left[A(\hat{\underline{q}}_k)\underline{r}_i\times\right]^2\right\}^{-1}\sum_{i=1}^{n}\sigma_i^{-2}\left[A(\hat{\underline{q}}_k)\underline{r}_i\times\right]\tilde{\underline{b}}_i(k+1) \tag{40}$$

Also, the covariance in Equation (39) reduces down to

$$P_{\text{pad}} \approx -\left[\sum_{i=1}^{n}\sigma_i^{-2}\left[A(\hat{\underline{q}}_k)\underline{r}_i\times\right]^2\right]^{-1} \tag{41}$$

which is essentially equivalent to the QUEST covariance shown in Ref. [3]. Also, since Equation (39) or (41) is used to determine the model error, the PAD algorithm determines the steady-state attitude error covariance as part of its solution.

**Error Analysis**

In this section, an error analysis is shown with respect to initial condition errors and sampling interval. The continuous output estimate can be shown to be given by [11]

$$\dot{\hat{\underline{y}}} = S\,\underline{d} \tag{42}$$

where

$$\hat{\underline{y}} = \left\{\left[A(\hat{\underline{q}})\underline{r}_1\right]^T \,\vdots\, \left[A(\hat{\underline{q}})\underline{r}_2\right]^T \,\vdots\, \cdots \,\vdots\, \left[A(\hat{\underline{q}})\underline{r}_n\right]^T\right\}^T \tag{43}$$

and $S$ is given by Equation (35). The error dynamics model between the measured and estimated observations can easily be shown to be given by

$$\dot{\underline{e}} = -\frac{1}{\Delta t}Q\,\underline{e} + [I - Q]\dot{\tilde{\underline{y}}} \tag{44}$$

where $\underline{e} = \tilde{\underline{y}} - \hat{\underline{y}}$, and

$$Q = S\,P_{\text{pad}}S^T R^{-1} \tag{45}$$

Since at least two vector observations are required in the PAD algorithm, the matrix $Q$ will always be positive semi-definite. Therefore, as long as the body observations are non-parallel and bounded, then the error in Equation (44) is also bounded for any initial condition error. Also, the error dynamics are a function of $1/\Delta t$, which means that the errors converge faster as the sampling interval decreases, which is intuitively correct.

## Spacecraft Simulation

In this section, an example is shown using PAD to determine the attitude from simulated star tracker measurements. The star tracker measures the tangent of two angles, $\alpha$ and $\beta$, resulting in a body vector given by

$$\underline{b}_i = \frac{1}{\sqrt{\left(\tan^2 a_i + \tan^2 \beta_i + 1\right)}} \begin{bmatrix} \tan a_i \\ \tan \beta_i \\ 1 \end{bmatrix} \quad (46)$$

where the z-axis of the tracker is along the boresight. The star tracker measurements are obtained by adding Gaussian noise to $\tan a_i$ and $\tan \beta_i$, with a 3σ value of 18 arc-sec. Measurements are sampled at 1 second intervals. The theoretical measurement error covariance for the model in Equation (46) is not isotropic. However, Shuster [10] has shown that if the noise variances on $\tan a_i$ and $\tan \beta_i$ are relatively equal and small, then the true covariance can be effectively replaced by an isotropic matrix.

The spacecraft is assumed to be earth-pointing with a rotation rate about the spacecraft y-axis (negative orbit normal) of 0.0011 rad/sec. The spacecraft z-axis is defined to be pointed nadir, and the x-axis completes the triad. Two trackers are used in the simulation. The first one has its boresight along the spacecraft y-axis, and the second has its boresight along the spacecraft x-axis. In the first simulation it assumed that both trackers measure two stars each with about a 0.5 degree separation between them. A plot of the attitude errors and 3σ bounds using the PAD algorithm is shown in Figure 1. These errors agree with the errors produced using the QUEST algorithm.

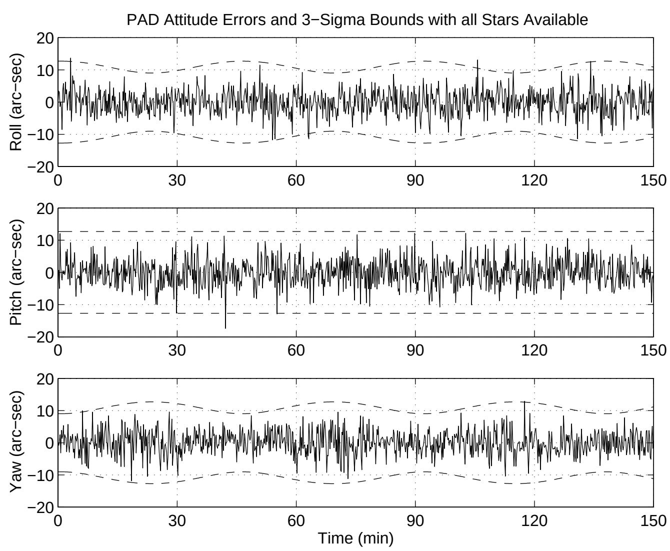


**Figure 1 PAD Attitude Errors Using Both Trackers**

In the next simulation, it is assumed that the second tracker has a single axis failure in $\tan \beta_i$. The disadvantage of methods which solve Wahba's original loss function is that they cannot use the single axis information from the second tracker. This is due to the fact that a scalar measurement error variance is assumed. Therefore, in order to use QUEST, only the first tracker measurements are used. A plot of the attitude errors and bounds using QUEST for this case is shown in Figure 2. Since the first tracker's boresight is along the spacecraft y-axis, the pitch axis has the largest error.

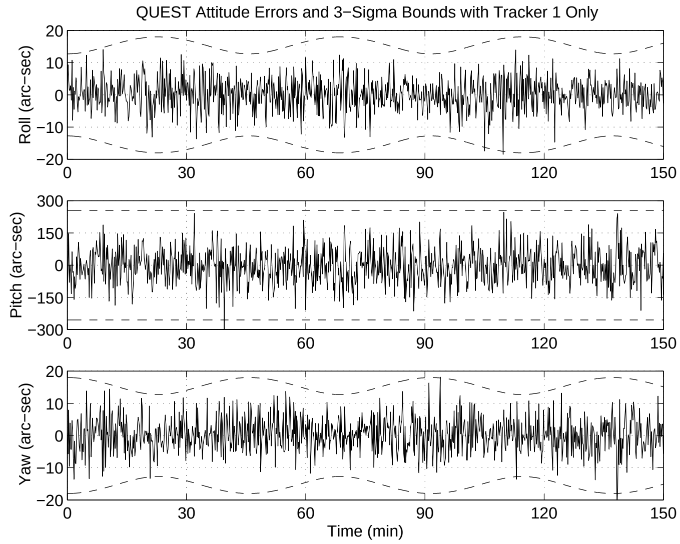


**Figure 2 QUEST Attitude Errors Using Tracker 1 Only**

The PAD algorithm can still use the single axis information from the second tracker. This is accomplished by using an inverse covariance matrix for that tracker given by

$$R_2^{-1} = \begin{bmatrix} \sigma^{-2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{47}$$

The (3,3) element is also set to zero since no information of the magnitude is known by the single axis failure. A plot of the PAD pitch error and bound using two stars in the first tracker and only one star with the single axis failure in the second tracker is shown in Figure 3 (the roll and yaw errors are approximately the same as the one-tracker case, due to the configuration of the trackers). Clearly, the pitch errors are reduced significantly compared with the QUEST solution in Figure 2.

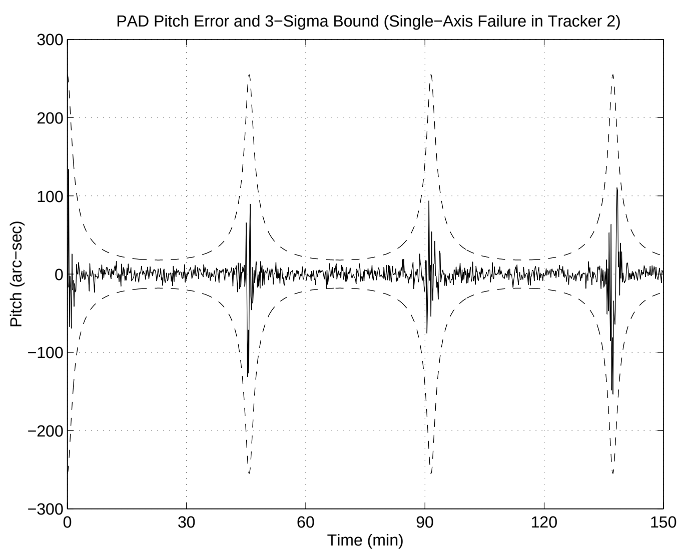


**Figure 3 PAD Pitch Error with a Single Axis Failure on Tracker 2**

A plot of the 3σ pitch error bounds from both the QUEST solution and PAD solution is shown in Figure 4. The peaks at 45, 90, and 135 minutes for PAD are due to attitude geometry, and to the fact that only one axis in Tracker 2 is used in the measurement. Depending on the attitude geometry, the

PAD algorithm reduces the error by an order of magnitude from the QUEST solution. Although this simulation represents an extreme case, it clearly proves that the PAD algorithm provides a viable approach for attitude determination when anisotropic errors exist.

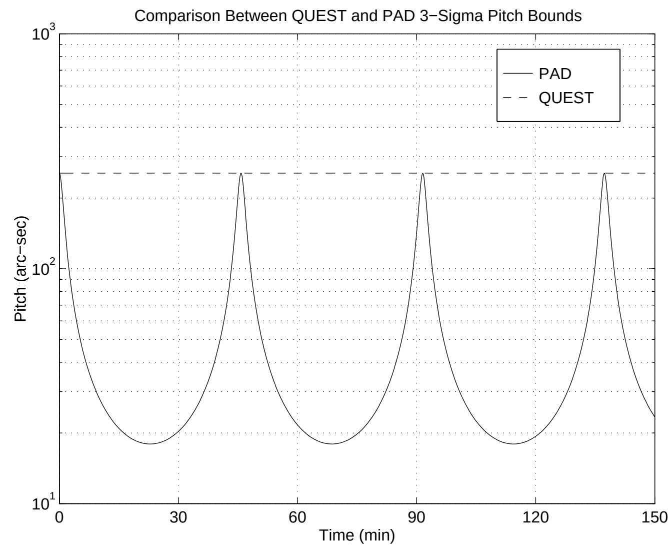


**Figure 4 QUEST and PAD Pitch Error Bounds**

Also, a test for initial condition errors was performed. A number of Monte Carlo runs was simulated using a normalized random vector for the initial quaternion. A plot of the convergence of the PAD algorithm for one case is shown in Figure 5. In each case, the PAD algorithm is able to converge within seven sampling intervals. The PAD algorithm may be initialized by picking two well-separated stars, and using TRIAD or two-observation (no iteration) QUEST or FOAM. However, this example shows that the attitude converges even for large initial condition errors.

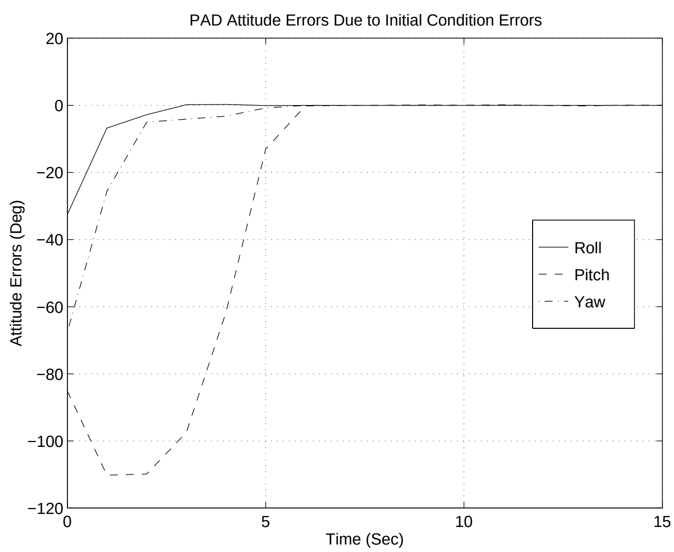


**Figure 5 PAD Convergence to Initial Condition Errors**

## Algorithm Computational Comparisons

The PAD algorithm is compared with QUEST and FOAM for computational floating point operations (FLOPS). In these comparisons, it is assumed that all vector observations are normalized and have isotropic errors only. All simulations were performed using MATLAB, and all matrix functions (such determinant, adjoint, etc.) were written out explicitly for the QUEST and FOAM algorithms. Also, the FOAM rotation matrix was converted to a quaternion, but this only requires 9 FLOPS. Since FOAM and QUEST do not require the attitude error covariance to calculate the attitude, two sets of test comparisons were made. The first one calculates the FLOPS without the covariance calculation in FOAM and QUEST, and the second calculates the FLOPS with the covariance calculation. PAD implicitly solves for the error covariance as part of its solution, so the number of FLOPS is the same in both sets. The number of FLOPS for $n$ vector observations without calculating the covariance in both QUEST and FOAM is given by

$$\begin{aligned} F_{\text{PAD}} &= 332 + 67(n-2) \\ F_{\text{QUEST}} &= 326 + 30(n-2) \\ F_{\text{FOAM}} &= 392 + 30(n-2) \end{aligned} \tag{48}$$

From Equation (48) it is clear that the QUEST algorithm requires the least number of FLOPS for any number of observations. Also, FOAM requires less number of FLOPS than PAD when three or more observations are present. The number of FLOPS for $n$ vector observations with calculating the covariance in both QUEST and FOAM is given by

$$\begin{aligned} F_{\text{PAD}} &= 332 + 67(n-2) \\ F_{\text{QUEST}} &= 442 + 57(n-2) \\ F_{\text{FOAM}} &= 482 + 30(n-2) \end{aligned} \tag{49}$$

For this case, PAD requires the least number of FLOPS until $n = 7$. When seven or more observations are present, then FOAM requires the least number of FLOPS. The QUEST algorithm overcomes PAD for the least number of FLOPS when thirteen or more observations are present. Also, FOAM overcomes QUEST for the least number of FLOPS when four or more observations are present. This is consistent with the CPU comparison shown in Ref. [5]. These case comparisons show that the PAD algorithm seems to be quite efficient in comparison to other attitude determination algorithms.

An advantage of both QUEST and FOAM is that they provide a point-by-point solution, independent of the sampling interval. As mentioned previously, PAD is a function of the sampling interval. A test was performed in order to investigate the effects of sampling interval. It is assumed that the propagation of the quaternion model in Equation (27) is performed at the sampling interval. Also, both trackers with two stars each are assumed in the simulations. The quaternion propagation frequency for an earth-point spacecraft is given by half the orbit rate. Nyquist's upper bound with a safety factor of 10 is about 500 seconds. Results for the $3\sigma$ attitude errors produced for different sampling intervals in PAD are shown in Table 1. This shows that the errors start to become significant with a sampling interval of about 100 seconds, and are quite significant with a sampling interval of 500 seconds. Although this study shows that PAD can produce large errors for large sampling intervals, the sampling intervals used for typical on-board spacecraft applications (e.g., in a Kalman filter) are well within the region where PAD provides accurate results.

**Table 1 PAD 3σ Attitude Errors for Various Sampling Intervals**

| $\Delta t$ (sec) | Roll (arc-sec) | Pitch (arc-sec) | Yaw (arc-sec) |
|---|---|---|---|
| 1 | 11 | 12 | 11 |
| 10 | 11 | 12 | 11 |
| 50 | 37 | 13 | 37 |
| 100 | 130 | 13 | 130 |
| 250 | 800 | 80 | 800 |
| 500 | 3000 | 700 | 3000 |
| 750 | 8,000 | 3,000 | 8,000 |
| 1000 | 30,000 | 600,000 | 30,000 |

## Conclusions

In this paper, a new algorithm was developed for attitude determination. The major advantage of this new algorithm over traditional algorithms, such as QUEST and FOAM, is that it is easily applicable to the case where anisotropic measurement errors exist. Also, the algorithm is computationally simpler than an extended Kalman filter approach, since no dynamics model is needed. The steady-state attitude error covariance for the new algorithm was shown to be equivalent to the optimal covariance, derived by solving a generalized form of Wahba's problem. Also, the attitude error covariance was shown to reduce to the QUEST covariance when isotropic conditions exist. Simulation studies indicated that the PAD algorithm provides a viable approach for attitude determination even when radically anisotropic errors exist. Finally, the PAD algorithm seems to be computationally comparable to both the QUEST and FOAM algorithms when isotropic measurement errors exist for all observations.

## References

[1] Wahba, G., "A Least Squares Estimate of Spacecraft Attitude," Problem 65-1, *SIAM Review*, Vol. 7, No. 3, July 1965, p. 409.

[2] Lerner, G.M., "Three-Axis Attitude Determination," *Spacecraft Attitude Determination and Control*, edited by J.R. Wertz, D. Reidel Publishing Co., Dordrecht, The Netherlands, 1978, pp. 426-428.

[3] Shuster, M.D., and Oh, S.D., "Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, Jan.-Feb. 1981, pp. 70-77.

[4] Markley, F.L., "Attitude Determination Using Vector Observations and the Singular Value Decomposition," *The Journal of the Astronautical Sciences*, Vol. 36, No. 3, July-Sept. 1988, pp. 245-258.

[5] Markley, F.L., "Attitude Determination from Vector Observations: A Fast Optimal Matrix Algorithm," *The Journal of the Astronautical Sciences*, Vol. 41, No. 2, April 1993, pp. 261-280.

[6] Bar-Itzhack, I.Y., "Polar Decomposition for Attitude Determination from Vector Observations," *Proceedings of the Flight Mechanics/Estimation Theory Symposium*, NASA-Goddard Space Flight Center, Greenbelt, MD, May 1992, pp. 243-257.

[7] Mortari, D., “Energy Approach Algorithm for Attitude Determination from Vector Observations,” *Advances in the Astronautical Sciences-Space Flight Mechanics*, Vol. 89, 1995, AAS paper #95-207, pp. 773-784.

[8] Mortari, D., “Euler-q Algorithm for Attitude Determination from Vector Observations,” *AAS/AIAA Space Flight Mechanics Meeting*, Austin, TX, Feb. 1996, AAS Paper #96-173.

[9] Shuster, M.D., “Maximum Likelihood Estimation of Spacecraft Attitude,” *The Journal of the Astronautical Sciences*, Vol. 37, No. 1, Jan.-March 1989, pp. 79-88.

[10] Shuster, M.D., “Kalman Filtering of Spacecraft Attitude and the QUEST Model,” *The Journal of the Astronautical Sciences*, Vol. 38, No. 3, July-Sept. 1990, pp. 377-393.

[11] Crassidis, J.L., and Markley, F.L., “Predictive Filtering for Nonlinear Systems,” *Journal of Guidance, Control and Dynamics*, in print, paper #G4969.

[12] Crassidis, J.L., and Markley, F.L., “Predictive Filtering for Attitude Estimation Without Rate Sensors,” *AAS/AIAA Space Flight Mechanics Meeting*, Austin, TX, Feb. 1996, AAS Paper #96-174.

[13] Markley, F.L., “Parameterization of the Attitude,” *Spacecraft Attitude Determination and Control*, edited by J.R. Wertz, D. Reidel Publishing Co., Dordrecht, The Netherlands, 1978, pp. 414-416.

[14] Crassidis, J.L., and Markley, F.L., “Attitude Determination Using Global Positioning System Signals,” *AIAA Journal of Guidance, Control, and Dynamics*, in review.

[15] 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.

[16] Spence, C.B., and Markley, F.L., “Attitude Propagation,” *Spacecraft Attitude Determination and Control*, edited by J.R. Wertz, D. Reidel Publishing Co., Dordrecht, The Netherlands, 1978, pp. 564-566.

[17] Hunt, L.R., Luksic, M., Su, R., “Exact Linearizations of Input-Output Systems,” *International Journal of Control*, Vol. 43, No. 1, 1986, pp. 247-255.

[18] Palm, W.J., *Modeling, Analysis, and Control of Dynamic Systems*, John Wiley & Sons, NY, 1983, pp. 230-231.

[19] Crassidis, J.L., “Integrated Estimation and Identification for Robust Control of Multivariable Systems,” Ph.D. Dissertation, State University of New York at Buffalo, May 1993.

[20] Gelb, A., *Applied Optimal Estimation*, MIT Press, MA, 1974.

[21] Farrenkopf, R.L., “Analytic Steady-State Accuracy Solutions for Two Common Spacecraft Attitude Estimators,” *Journal of Guidance and Control*, Vol. 1, July-Aug. 1978, pp. 282-284.