AIAA 81-4003

Three-Axis Attitude Determination from Vector Observations

M.D. Shuster and S.D. Oh

Reprinted from

Journal of Guidance and Control

Volume 4, Number 1, January-February 1981, Page 70.

AMERICAN INSTITUTE OF AERONAUTICS AND ASTRONAUTICS • 1290 AVENUE OF THE AMERICAS • NEW YORK, NEW YORK, N.Y. 10104

AIAA 81-4003

Three-Axis Attitude Determination from Vector Observations

M. D. Shuster* and S.D. Oh†
Computer Sciences Corporation, Silver Spring, Md.

Two computationally efficient algorithms are presented for determining three-axis attitude from two or more vector observations. The first of these, the TRIAD algorithm, provides a deterministic (i.e., nonoptimal) solution for the attitude based on two vector observations. The second, the QUEST algorithm, is an optimal algorithm which determines the attitude that achieves the best weighted overlap of an arbitrary number of reference and observation vectors. Analytical expressions are given for the covariance matrices for the two algorithms using a fairly realistic model for the measurement errors. The mathematical relationship of the two algorithms and their relative merits are discussed and numerical examples are given. The advantage of computing the covariance matrix in the body frame rather than in the inertial frame (e.g., in terms of Euler angles) is emphasized. These results are valuable when a single-frame attitude must be computed frequently. They will also be useful to the mission analyst or spacecraft engineer for the evaluation of launch-window constraints or of attitude accuracies for different attitude sensor configurations.

I. Introduction

A RECURRENT problem in spacecraft attitude determination is to determine the attitude from a set of vector measurements. Thus, an orthogonal matrix A (the attitude matrix or direction-cosine matrix) is sought which satisfies

$$A\hat{V}_i = \hat{W}_i \qquad (i=1,\ldots,n) \tag{1}$$

where $\hat{V}_1,\ldots,\hat{V}_n$ are a set of reference unit vectors, which are n known directions (e.g., the direction of the Earth, the sun, a star, or the geomagnetic field) in the reference coordinate system, and $\hat{W}_1,\ldots,\hat{W}_n$ are the observation unit vectors, which are the same n directions as measured in the spacecraft-body coordinate system. (In general, unit vectors will be denoted by carets.)

Because both the observation and the reference unit vectors are corrupted by error, a solution for A does not exist in general, not even for $n=2$. This work studies two approaches to this problem: one deterministic and the other optimal.

The deterministic method, the TRIAD algorithm,[1] determines the attitude by first discarding part of the measurements so that a solution exists. Because the algorithm is very simple, it has become the most popular method for determining three-axis attitude for spacecraft that provide complete vector information. The algorithm has been in existence for at least a decade and has been implemented in many recent missions. These have included Small Astronomy Satellite (SAS), Seasat, Atmospheric Explorer Missions (AEM), and Magsat (for coarse definitive attitude) and will include the Dynamics Explorer (DE) missions.

However, whereas the computation of the TRIAD attitude matrix is very simple and straightforward, calculations of the attitude covariance matrix as currently implemented[2] have been rather complicated, often requiring the computation of numerous partial derivatives as differences. To calculate these partials, the number of computations which must be performed is usually many times greater than the number required for determining the attitude. Therefore, a simple analytical expression for the covariance matrix of the TRIAD attitude would be very useful. The derivation of such an expression is one of the goals of this paper.

The greatest drawback of the TRIAD algorithm is that it can accommodate only two observations. When more than two measurements are available, these can be utilized only by cumbersomely combining the attitude solutions for the various observation-vector pairs. In addition, even when there are only two observations some accuracy is lost because part of the measurement is discarded.

These two drawbacks are usually not present in an optimal algorithm, which computes a best estimate of the spacecraft attitude based on a loss function which takes into account all n measurements. Optimal algorithms, however, are usually much slower than deterministic algorithms. For the present work the chosen loss function is

$$L(A) = \frac{1}{2}\sum_{i=1}^{n} a_i \,|\hat{W}_i - A\hat{V}_i|^2 \tag{2}$$

which was first proposed by Wahba[3] in 1965. Davenport (see Ref. 4) has shown that this quadratic loss function in the attitude matrix could be transformed into a quadratic loss function in the corresponding quaternion. This is a great simplification of the problem posed by Wahba since the quaternion is subject to fewer constraints than the nine elements of the attitude matrix. Davenport's substitution leads directly to an eigenvalue equation for the quaternion. This eigenvalue equation, as elaborated by Keat,[4] is the basis for the work presented in this paper.

It is possible to develop an approximation scheme[5] that permits the computation of the optimal quaternion to arbitrarily high accuracy without having to solve the complete eigenvalue problem explicitly and with a significant reduction in computation. This algorithm, QUEST (QUaternion ESTimator), maintains all the computational advantages of a fast deterministic algorithm while yielding an optimal result. A simple analytical expression also can be obtained for the QUEST covariance matrix.

The quaternion eigenvalue equation of Davenport and Keat was first implemented in ground support software for the High Energy Astronomy Observatory (HEAO). Because attitude is computed typically only ten times daily for the three spacecraft of the HEAO mission, the complete solution of a four-dimensional eigenvalue problem did not pose a significant computation burden. The QUEST algorithm was developed for the Magsat mission, where fine definitive at-

Received Oct. 7, 1979; revision received April 28, 1980.

*Senior Member of the Technical Staff, Attitude Systems Operation, Systems Science Division. Member AIAA.

†Member of the Technical Staff, Attitude Systems Operation, Systems Science Division.

titude must be computed every 0.25 s for the life of the mission using as many as three sensors. For Magsat the reduction in computation made possible by the QUEST algorithm is substantial.

It turns out that a simple connection exists between the TRIAD and the QUEST algorithms. This connection is derived rigorously, and the relative merits of the two algorithms are discussed.

An important point, which is stressed throughout this work for both the TRIAD and the QUEST algorithm (in fact, for any algorithm), is that the preferred frame for computing the attitude covariance matrix is the spacecraft-body frame. This is true not only for ease of computation but also for ease of interpretation. It is possible for the spacecraft attitude to be accurately determined while the variances, when referenced to nonbody axes, are quite large. These points will be made clearer in the two sections which follow.

II. The TRIAD Algorithm

The Attitude Matrix

Given two nonparallel reference unit vectors $\hat{V}_1$ and $\hat{V}_2$ and the corresponding observation unit vectors $\hat{W}_1$ and $\hat{W}_2$, we wish to find an orthogonal matrix A which satisfies

$$A\hat{V}_1=\hat{W}_1 \qquad A\hat{V}_2=\hat{W}_2 \tag{3}$$

Because the matrix A is overdetermined by the above equations we begin by constructing two triads of manifestly orthonormal reference and observation vectors according to

$$\hat{r}_1=\hat{V}_1 \qquad \hat{r}_2=(\hat{V}_1\times\hat{V}_2)/|\hat{V}_1\times\hat{V}_2|$$
$$\hat{r}_3=(\hat{V}_1\times(\hat{V}_1\times\hat{V}_2))/|\hat{V}_1\times\hat{V}_2| \tag{4}$$

$$\hat{s}_1=\hat{W}_1 \qquad \hat{s}_2=(\hat{W}_1\times\hat{W}_2)/|\hat{W}_1\times\hat{W}_2|$$
$$\hat{s}_3=(\hat{W}_1\times(\hat{W}_1\times\hat{W}_2))/|\hat{W}_1\times\hat{W}_2| \tag{5}$$

There exists a unique orthogonal matrix A which satisfies

$$A\hat{r}_i=\hat{s}_i \qquad (i=1,2,3) \tag{6}$$

which is given by

$$A=\sum_{i=1}^{3}\hat{s}_i\hat{r}_i^T \tag{7}$$

where T denotes the matrix transpose. ($\hat{r}_i$ is interpreted as a 3×1 matrix and $\hat{r}_i^T$ as a 1×3 matrix.) In other notation Eq. (7) is identical to

$$A=M_{\mathrm{obs}}M_{\mathrm{ref}}^T \tag{8}$$

with

$$M_{\mathrm{ref}}=[\hat{r}_1\,\vdots\,\hat{r}_2\,\vdots\,\hat{r}_3] \qquad M_{\mathrm{obs}}=[\hat{s}_1\,\vdots\,\hat{s}_2\,\vdots\,\hat{s}_3] \tag{9}$$

The right members of Eqs. (9) are 3×3 matrices labeled according to their column vectors. Equation (7) or, equivalently, Eq. (8) defines the TRIAD solution.

A necessary and sufficient condition that the attitude matrix given by Eq. (7) also satisfy Eqs. (3) is

$$\hat{V}_1\cdot\hat{V}_2=\hat{W}_1\cdot\hat{W}_2 \tag{10}$$

The TRIAD solution is not symmetric in indices 1 and 2. Clearly, because part of the information contained in the second vector is discarded, the TRIAD solution will be more accurate when $(\hat{W}_1, \hat{V}_1)$ is chosen to be the observation-reference vector pair of greater accuracy.

The TRIAD Covariance Matrix

Conventionally, the attitude covariance matrix is defined as the covariance matrix of a set of Euler angles which parameterize the attitude. This turns out to be very cumbersome to calculate and in many ways less informative than the covariance matrix of a set of angles which are referenced to the body axes. We shall develop first the formalism for the body-referenced covariance matrix and then show the connection to the covariance matrix of the Euler angles.

The error angle vector

$$\delta\boldsymbol{\theta}=(\delta\theta_1,\delta\theta_2,\delta\theta_3)^T \tag{11}$$

is defined as the set of angles (measured in radians) of the small rotation carrying the true attitude matrix into the measured attitude matrix. It is assumed that $\delta\boldsymbol{\theta}$ is unbiased so that the true attitude matrix is also the expected mean to first order in the angles, which are assumed to be small. Thus, to first order

$$A=\begin{bmatrix}1 & \delta\theta_3 & -\delta\theta_2\\ -\delta\theta_3 & 1 & \delta\theta_1\\ \delta\theta_2 & -\delta\theta_1 & 1\end{bmatrix}\langle A\rangle \tag{12}$$

where the brackets denote the expectation value.

The Cartesian attitude covariance matrix is defined as

$$P_{\theta\theta}=\langle\delta\boldsymbol{\theta}\delta\boldsymbol{\theta}^T\rangle \tag{13}$$

For the purpose of computation, however, it is convenient to examine a related covariance matrix defined as

$$P=\langle\delta A\delta A^T\rangle \tag{14}$$

where

$$\delta A=A-\langle A\rangle \tag{15}$$

From Eq. (12) it follows that P and $P_{\theta\theta}$ satisfy

$$P_{\theta\theta}=(\tfrac{1}{2}\,\mathrm{tr}\,P)I-P \tag{16}$$

where tr denotes the trace operation and I is the identity matrix. Because a simple expression is available for A, namely, Eq. (8), P is the quantity most readily calculable.

From Eq. (8) it can further be shown that

$$P=\langle\delta M_{\mathrm{obs}}\delta M_{\mathrm{obs}}^T\rangle+A\langle\delta M_{\mathrm{ref}}\delta M_{\mathrm{ref}}^T\rangle A^T \tag{17}$$

or

$$P=P_{\mathrm{obs}}+AP_{\mathrm{ref}}A^T \tag{18}$$

Similarly, P_{obs} and P_{ref} can each be written as the sum of two terms, each generated by the variation of a single observation or reference vector.

Because the vectors $\hat{V}_m$ and $\hat{W}_m$ are constrained to be unit vectors, the error in any one of them must to first order lie in the plane perpendicular to that vector. Thus, to lowest order in $\delta\hat{W}_m$ and $\delta\hat{V}_m$

$$\delta\hat{W}_m\cdot\hat{W}_m=0 \qquad \delta\hat{V}_m\cdot\hat{V}_m=0 \qquad (m=1,2) \tag{19}$$

The error vectors $\delta\hat{W}_m$ and $\delta\hat{V}_m$, therefore can have only two degrees of freedom. We make the further approximation that the error vector has an axially symmetric distribution about the respective unit vector. In terms of the covariance matrices of the error vectors (assumed to be uncorrelated with one another), this approximation reads

$$\langle \delta\hat{W}_m \delta\hat{W}_n^T \rangle = \sigma_{Wm}^2 \delta_{mn} (I - \hat{W}_m \hat{W}_m^T) \qquad (m,n=1,2)$$

$$\langle \delta\hat{V}_m \delta\hat{V}_n^T \rangle = \sigma_{Vm}^2 \delta_{mn} (I - \hat{V}_m \hat{V}_m^T) \qquad (m,n=1,2)$$

$$\langle \delta\hat{V}_m \delta\hat{W}_n^T \rangle = 0 \qquad (m,n=1,2) \quad (20)$$

where σ_X^2 is the variance of a component of $\hat{X}$ along a direction normal to $\langle \hat{X} \rangle$.

For a vector sensor with a limited field of view this assumption on the error distribution, namely, that the error along each sensor axis is roughly the same, is usually the case. For a sensor with an extended field of view (e.g., a sun sensor with a half-cone angle of 70 deg), Eqs. (20) will sometimes be a poor approximation when the measurement falls near the edge of the field of view. This is due both to geometrical effects and to deficiencies in the hardware. However, even in these limiting cases, the approximation of Eqs. (20) seldom fails badly.

Noting that

$$P_{obs} = \sum_{i=1}^{3} \langle \delta\hat{s}_i \delta\hat{s}_i^T \rangle \qquad (21)$$

the calculation of P_{obs} is simplified by using the relation

$$\delta\hat{s}_i = \frac{1}{|s_i|} (I - \hat{s}_i \hat{s}_i^T) \delta s_i \qquad (22)$$

where s_i is an unnormalized vector given by Eqs. (5) but without the denominator. The calculation of P_{obs} is lengthy but straightforward, and the expression for P_{ref} is identical in form. From these, P and $P_{\theta\theta}$ may be calculated to yield in terms of the triad vectors

$$P_{\theta\theta} = \left[(\sigma_1^2 + \sigma_2^2) \frac{1}{|\hat{W}_1 \times \hat{W}_2|^2} - \sigma_1^2 \right] \hat{s}_1 \hat{s}_1^T$$

$$+ \sigma_1^2 (\hat{s}_2 \hat{s}_2^T + \hat{s}_3 \hat{s}_3^T) - \sigma^2{}_1 \frac{(\hat{W}_1 \cdot \hat{W}_2)}{|\hat{W}_1 \times \hat{W}_2|} (\hat{s}_1 \hat{s}_3^T + \hat{s}_3 \hat{s}_1^T) \qquad (23)$$

$$\sigma_1^2 = \sigma_{V1}^2 + \sigma_{W1}^2 \qquad \sigma_2^2 = \sigma_{V2}^2 + \sigma_{W2}^2 \qquad (24)$$

and the approximation has been made that

$$\hat{V}_1 \cdot \hat{V}_2 = \hat{W}_1 \cdot \hat{W}_2 \qquad (25)$$

Equation (23) may be rewritten in terms of the observation vectors as

$$P_{\theta\theta} = \sigma_1^2 I + \frac{1}{|\hat{W}_1 \times \hat{W}_2|^2} [(\sigma_2^2 - \sigma_1^2) \hat{W}_1 \hat{W}_1^T$$

$$+ \sigma_1^2 (\hat{W}_1 \cdot \hat{W}_2)(\hat{W}_1 \hat{W}_2^T + \hat{W}_2 \hat{W}_1^T)] \qquad (26)$$

which is the desired expression.

The inverse covariance matrix, sometimes called the information matrix, has the equally simple form

$$P_{\theta\theta}^{-1} = \frac{1}{\sigma_1^2} \{ I - \hat{W}_1 \hat{W}_1^T \} + \frac{1}{\sigma_2^2} \hat{s}_4 \hat{s}_4^T \qquad (27)$$

where

$$\hat{s}_4 = (\hat{W}_2 \times (\hat{W}_1 \times \hat{W}_2)) / |\hat{W}_1 \times \hat{W}_2| \qquad (28)$$

Computation of the Covariance Matrix for the Euler Angles

In many applications, it is not the covariance matrix of the error angles $(\delta\theta_1, \delta\theta_2, \delta\theta_3)$, measured about the body coordinate axes, which is sought but rather the error covariance matrix associated with a given sequence of Euler angles[6] (ϕ_1, ϕ_2, ϕ_3), which parameterize the spacecraft attitude. The error covariance matrix for a set of Euler angles may be obtained from $P_{\theta\theta}$ by noting that the error in the attitude matrix is given by

$$\delta A = \begin{bmatrix} 0 & \delta\theta_3 & -\delta\theta_2 \\ -\delta\theta_3 & 0 & \delta\theta_1 \\ \delta\theta_2 & -\delta\theta_1 & 0 \end{bmatrix} A \qquad (29)$$

and also by

$$\delta A = \sum_{i=1}^{3} \frac{\partial A}{\partial \phi_i} \delta\phi_i \qquad (30)$$

To simplify notation, A has been written in place of $\langle A \rangle$. This is, in fact, correct to lowest nonvanishing order in δA.

Comparing the matrices of Eqs. (29) and (30) element by element and solving for $\delta\theta$ leads to

$$\delta\theta = [H(\phi_1, \phi_2, \phi_3)]^{-1} \delta\phi \qquad (31)$$

and the matrix H^{-1} is given by

$$[H^{-1}]_{ij} = \frac{1}{2} \sum_{kl} \epsilon_{ikl} \left[\frac{\partial A}{\partial \phi_j} A^T \right]_{kl} \qquad (32)$$

where ϵ_{ikl} is the Lévi-Civitá symbol. If A_k is the kth column vector of A, then the right member of Eq. (32) can be rewritten as

$$\frac{1}{2} \sum_{k=1}^{3} \left(\frac{\partial A_k}{\partial \phi_j} \times A_k \right),$$

The covariance matrix in the Euler angles is then

$$P_{\phi\phi} = \langle \delta\phi \delta\phi^T \rangle = H P_{\theta\theta} H^T \qquad (33)$$

It should be remarked that although covariance matrices are computed most often in terms of the Euler angles, the Cartesian error covariance matrix $P_{\theta\theta}$ is more useful. In particular, because the Cartesian error angles $(\delta\theta_1, \delta\theta_2, \delta\theta_3)$ form a vector, $P_{\theta\theta}$ is a tensor of rank 2. Thus, an orthogonal transformation of the observation vectors induces a corresponding orthogonal similarity transformation on $P_{\theta\theta}$, while an orthogonal transformation of the reference vectors leaves $P_{\theta\theta}$ unchanged. In comparison, the transformation of $P_{\phi\phi}$ is much more complicated. Also, the trace of $P_{\theta\theta}$ provides a convenient scalar quantity for judging the root-sum-square accuracy of the attitude solution, which is independent of the choice of representation and the attitude. $P_{\phi\phi}$ yields no comparable quantity. Finally, although $P_{\theta\theta}$ is attitude independent, $P_{\phi\phi}$ is very strongly dependent on the attitude through H. Thus, very large values of $P_{\phi\phi}$ need not mean that the attitude is poorly known.

The results of this section are completely general and are not specific to the TRIAD algorithm. Thus, Eqs. (31-33) give the connection between the Cartesian error angles and the Euler angles for any attitude determination algorithm. In particular, Eq. (33) can be used to reconstruct $P_{\theta\theta}$ from $P_{\phi\phi}$ when only the latter is known.

III. The QUEST Algorithm‡

The Quaternion Eigenvalue Equation§

We wish to find an orthogonal matrix A_{opt} that minimizes the loss function

$$L(A) = \frac{1}{2} \sum_{i=1}^{n} a_i |\hat{W}_i - A \hat{V}_i|^2 \qquad (34)$$

‡The remainder of this paper is the work of the first author.

§The derivation here follows closely that of Ref. 4.

where the a_i, $i=1,\ldots,n$ are a set of nonnegative weights. Because the loss function may be scaled without affecting the determination of A_{opt}, it is possible to set

$$\sum_{i=1}^{n} a_i = 1 \tag{35}$$

The gain function $g(A)$ is defined by

$$g(A) = 1 - L(A) = \sum_{i=1}^{n} a_i \hat{W}_i^T A \hat{V}_i \tag{36}$$

The loss function $L(A)$ will be at a minimum when the gain function $g(A)$ is at a maximum. All further discussion will be directed at finding the optimal attitude matrix A_{opt}, which maximizes $g(A)$. Interpreting the individual terms of Eq. (36) as 1×1 matrices, it follows from a well-known theorem on the trace that

$$g(A) = \sum_{i=1}^{n} a_i \mathrm{tr}[\hat{W}_i^T A \hat{V}_i] = \mathrm{tr}[AB^T] \tag{37}$$

where tr denotes the trace operation and B, the attitude profile matrix, is given by

$$B = \sum_{i=1}^{n} a_i \hat{W}_i \hat{V}_i^T \tag{38}$$

The maximization of $g(A)$ is complicated by the fact that the nine elements of A are subject to six constraints. Therefore, it is convenient to express A in terms of its related quaternion.

The quaternion $\bar{q}$ representing a rotation is given by[6]

$$\bar{q} = \begin{Bmatrix} Q \\ q \end{Bmatrix} = \begin{Bmatrix} \hat{X}\sin(\theta/2) \\ \cos(\theta/2) \end{Bmatrix} \tag{39}$$

where $\hat{X}$ is the axis of rotation and θ is the angle of rotation about $\hat{X}$. The quaternion satisfies a single constraint, which is

$$\bar{q}^T\bar{q} = |Q|^2 + q^2 = 1 \tag{40}$$

The attitude matrix A is related to the quaternion by[6]

$$A(\bar{q}) = (q^2 - Q\cdot Q)I + 2QQ^T + 2q\underline{Q} \tag{41}$$

where I is the identity matrix and $\underline{Q}$ is the antisymmetric matrix given by

$$\underline{Q} = \begin{bmatrix} 0 & Q_3 & -Q_2 \\ -Q_3 & 0 & Q_1 \\ Q_2 & -Q_1 & 0 \end{bmatrix} \tag{42}$$

Substituting Eq. (41) into Eq. (37), the gain function may be rewritten as

$$g(\bar{q}) = (q^2 - Q\cdot Q)\mathrm{tr}B^T + 2\mathrm{tr}[QQ^T B^T] + 2q\mathrm{tr}[\underline{Q}B^T] \tag{43}$$

Introducing the quantities

$$\sigma = \mathrm{tr}B = \sum_{i=1}^{n} a_i \hat{W}_i \cdot \hat{V}_i \tag{44}$$

$$S = B + B^T = \sum_{i=1}^{n} a_i(\hat{W}_i\hat{V}_i^T + \hat{V}_i\hat{W}_i^T) \tag{45}$$

$$Z = \sum_{i=1}^{n} a_i(\hat{W}_i \times \hat{V}_i) \tag{46}$$

leads to the bilinear form

$$g(\bar{q}) = \bar{q}^T K \bar{q} \tag{47}$$

where the 4×4 matrix K is given by

$$K = \left[\begin{array}{c|c} S - \sigma I & Z \\ \hline Z^T & \sigma \end{array}\right] \tag{48}$$

Equation (46) may be written alternatively as

$$\underline{Z} = B - B^T \tag{49}$$

The problem of determining the optimal attitude has been reduced to finding the quaternion that maximizes the bilinear form of Eq. (47). The constraint of Eq. (40) can be taken into account by the method of Lagrange multipliers.[7] A new gain function $g'(\bar{q})$ is defined as

$$g'(\bar{q}) = \bar{q}^T K \bar{q} - \lambda \bar{q}^T \bar{q} \tag{50}$$

which is maximized without constraint. λ is then chosen to satisfy the constraint. It may be verified by straightforward differentiation that $g'(\bar{q})$ attains a stationary value provided

$$K\bar{q} = \lambda\bar{q} \tag{51}$$

Thus, $\bar{q}_{opt}$ must be an eigenvector of K. Equation (51) is independent of the normalization of $\bar{q}$ and, therefore, Eq. (40) does not determine λ. However, λ must be an eigenvalue of K and for each eigenvector of K

$$g(\bar{q}) = \bar{q}^T K \bar{q} = \lambda \bar{q}^T \bar{q} = \lambda \tag{52}$$

Thus, $g(\bar{q})$ will be maximized if $\bar{q}_{opt}$ is chosen to be the eigenvector of K belonging to the largest eigenvalue of K. More concisely,

$$K\bar{q}_{opt} = \lambda_{max}\bar{q}_{opt} \tag{53}$$

which is the desired result.

Construction of the Optimal Quaternion

Equation (53) can be rearranged to read, for any eigenvalue λ,

$$Y = [(\lambda + \sigma)I - S]^{-1} Z \tag{54}$$

$$\lambda = \sigma + Z \cdot Y \tag{55}$$

where Y is the Gibbs vector defined as

$$Y = Q/q = \hat{X}\tan(\theta/2) \tag{56}$$

In terms of the Gibbs vector,

$$\bar{q} = \frac{1}{\sqrt{1 + |Y|^2}} \begin{Bmatrix} Y \\ 1 \end{Bmatrix} \tag{57}$$

When λ is equal to λ_{max}, Y and $\bar{q}$ are representations of the optimal attitude solution. Inserting Eq. (54) into Eq. (55) leads to an equation for the eigenvalues

$$\lambda = \sigma + Z^T \frac{1}{(\lambda + \sigma)I - S} Z \tag{58}$$

Equation (58) is equivalent to the characteristic equation for the eigenvalues of K, the explicit solution of which it is desired to avoid. However, it should be noted that

$$\lambda_{max} = 1 - \frac{1}{2}\sum_{i=1}^{n} a_i |\hat{W}_i - A_{opt}\hat{V}_i|^2 \tag{59}$$

which is very close to unity. Substituting

$$\lambda_{max} = 1 \tag{60}$$

into Eq. (54) leads to an expression for the attitude which is accurate to second order in the measurement errors, provided that the matrix

$$[(\lambda_{max} + \sigma)I - S]$$

is not singular. However, the Gibbs vector becomes infinite when the angle of rotation (in radians unless otherwise specified) is π. Hence, from Eq. (54) the above matrix must be singular there, and the approximation of Eq. (60) is not useful when the angle of rotation is close to π. The remainder of this subsection is devoted to developing more accurate methods which avoid the problems posed by this singularity.

The first step is to derive an expression that permits the computation of $\bar{q}_{opt}$ without the intermediary of the Gibbs vector.

It should be noted that an eigenvalue ξ of any square matrix S satisfies the characteristic equation

$$\det |S - \xi I| = 0 \tag{61}$$

which for a 3×3 matrix takes the form

$$-\xi^3 + 2\sigma\xi^2 - \kappa\xi + \Delta = 0 \tag{62}$$

with

$$\sigma = \tfrac{1}{2}\mathrm{tr}S \qquad \kappa = \mathrm{tr}(\mathrm{adj}\ S) \qquad \Delta = \det S \tag{63}$$

where tr, adj, and det note the trace, adjoint matrix, and determinant, respectively. By the Cayley-Hamilton theorem,[8] S satisfies this same equation in the sense that

$$S^3 = 2\sigma S^2 - \kappa S + \Delta I \tag{64}$$

Equation (64) may be used to express any meromorphic function of S as a quadratic in S, in particular,

$$[(\omega + \sigma)I - S]^{-1} = \gamma^{-1}(\alpha I + \beta S + S^2) \tag{65}$$

where

$$\alpha = \omega^2 - \sigma^2 + \kappa \qquad \beta = \omega - \sigma \qquad \gamma = (\omega + \sigma)\alpha - \Delta \tag{66}$$

Letting ω assume the value λ_{max} leads to

$$Y_{opt} = X/\gamma \tag{67}$$

where

$$X = (\alpha I + \beta S + S^2)Z \tag{68}$$

It follows then from Eqs. (57) and (67) that

$$\bar{q}_{opt} = \frac{1}{\sqrt{\gamma^2 + |X|^2}} \begin{Bmatrix} X \\ \gamma \end{Bmatrix} \tag{69}$$

in which the Gibbs vector no longer appears.

Equation (65) applied to Eq. (58) leads to a convenient expression for the characteristic equation, namely,

$$\lambda^4 - (a+b)\lambda^2 - c\lambda + (ab + c\sigma - d) = 0 \tag{70}$$

where

$$a = \sigma^2 - \kappa \qquad b = \sigma^2 + Z^T Z$$

$$c = \Delta + Z^T S Z \qquad d = Z^T S^2 Z \tag{71}$$

Because λ_{max} is known to be very close to unity, the Newton-Raphson method[8] applied to Eq. (70) with unity as a starting value allows λ_{max} to be computed to arbitrarily high accuracy. For sensor accuracies better than 1 arc-min (1 deg), the accuracy of a 64-bit word is exhausted with only one iteration (two iterations) of the Newton-Raphson method. The computational advantage of this approach as compared to that required for the complete solution of the eigenvalue problem is evident, especially when one considers that the same quantities which figure in Eqs. (71) must also be calculated for the construction of the quaternion.

For the case where there are only two observations, λ_{max} has a simple, exact closed-form expression

$$\lambda_{max} = \sqrt{a_1^2 + 2a_1 a_2 \cos(\theta_V - \theta_W) + a_2^2} \tag{72}$$

where

$$\cos(\theta_V - \theta_W) = (\hat{V}_1 \cdot \hat{V}_2)(\hat{W}_1 \cdot \hat{W}_2) + |\hat{V}_1 \times \hat{V}_2|\,|\hat{W}_1 \times \hat{W}_2| \tag{73}$$

Equation (69) can still lead to an indeterminant result if both γ and X vanish simultaneously. It is clear from Eq. (67) that γ vanishes if and only if the angle of rotation is π. In fact, γ is the determinant of the matrix whose inverse is given by Eq. (65). Unfortunately, even if X does not vanish along with γ, Eq. (69) will not be accurate when the angle of rotation is close to π because of large cancellations which occur. This problem can be eliminated completely by employing the method of sequential rotations discussed below.

It is noted without proof that a rotation through an angle greater than $\pi/2$ can be expressed as a rotation through π about one of the coordinate axes followed by a rotation about a new axis through an angle less than $\pi/2$. An initial rotation through π about one of the coordinate axes is equivalent to changing the signs of two components of each of the reference vectors. The quaternion $\bar{p} = (p_1, p_2, p_3, p_4)^T$ of the optimal rotation transforming the new reference vectors $\hat{V}_i'$, $i = 1, \ldots, n$ into the observation vectors $\hat{W}_i$, $i = 1, \ldots, n$ as calculated from Eq. (69) is related very simply to the desired optimal quaternion. The results for the three possible cases are

1) Initial rotation through π about the x axis:

$$\hat{V}_i' = (\hat{V}_{ix}, -\hat{V}_{iy}, -\hat{V}_{iz}) \qquad \bar{q} = (p_4, -p_3, p_2, -p_1)^T \tag{74}$$

2) Initial rotation through π about the y axis:

$$\hat{V}_i' = (-\hat{V}_{ix}, \hat{V}_{iy}, -\hat{V}_{iz}) \qquad \bar{q} = (p_3, p_4, -p_1, -p_2)^T \tag{75}$$

3) Initial rotation through π about the z axis:

$$\hat{V}_i' = (-\hat{V}_{ix}, -\hat{V}_{iy}, \hat{V}_{iz}) \qquad \bar{q} = (-p_2, p_1, p_4, -p_3)^T \tag{76}$$

Clearly, that initial rotation (including no initial rotation) will yield the most accurate estimate of $\bar{q}_{opt}$ for which $|\gamma|$ achieves the largest value. In any practical application, however, $|\gamma|$ can be allowed to become quite small before the method of sequential rotations need be invoked.

Approximations Near Null Attitude

When the attitude matrix is known to be very close to the identity matrix (null attitude) or, equivalently, when the angle of rotation is known to be small, simple approximations may be obtained. In that case, Z will be a small quantity of the order of the angle of rotation and to this same order $\hat{W}_i$ may be replaced by $\hat{V}_i$ in computing other expressions.

Thus, if δ is a quantity of the order of the error of observation or the angle or rotation, whichever is larger, then

$$Z = \mathcal{O}(\delta) \qquad \sigma = 1 + \mathcal{O}(\delta^2)$$

$$\lambda_{max} = 1 + \mathcal{O}(\delta^2) \qquad S = S_0 + \mathcal{O}(\delta) \tag{77}$$

where

$$S_0 = 2\sum_{i=1}^{n} a_i \hat{V}_i \hat{V}_i^T \tag{78}$$

With these approximations Eq. (54) becomes

$$Y = [2I - S_0]^{-1} Z + \mathcal{O}(\delta^2) \tag{79}$$

Likewise, near null attitude, Eqs. (67) and (68) simplify to

$$Y = \frac{1}{2\kappa_0 - \Delta_0} [\kappa_0 I + S_0^2] Z + \mathcal{O}(\delta^2) \tag{80}$$

where

$$\kappa_0 = \mathrm{tr}(\mathrm{adj}\ S_0) \qquad \Delta_0 = \det S_0 \tag{81}$$

When there are only two observations, Eq. (80) reduces to

$$\begin{aligned} Y = \tfrac{1}{2}\{ & a_1 \hat{W}_1 \times \hat{V}_1 + a_2 \hat{W}_2 \times \hat{V}_2 \\ & + a_1 [(b\hat{W}_1 - a\hat{W}_2)\cdot(\hat{V}_1 \times \hat{V}_2)]\hat{V}_1 \\ & + a_2 [(a\hat{W}_1 - b\hat{W}_2)\cdot(\hat{V}_1 \times \hat{V}_2)]\hat{V}_2\} + \mathcal{O}(\delta^2) \end{aligned} \tag{82}$$

with a and b given by

$$a = 1/|\hat{V}_1 \times \hat{V}_2|^2 \qquad b = (\hat{V}_1 \cdot \hat{V}_2)/|\hat{V}_1 \times \hat{V}_2|^2 \tag{83}$$

Equations (79-83) may be of practical use when the spacecraft has a high pointing accuracy requirement in an inertial reference coordinate system. In that case, the vectors $\hat{V}_i$, $i = 1,...,n$ and the matrix S_0 need be computed only once and the connection between the observations and the Gibbs vector is immediate.

Similarly, these small-angle algorithms may be used as optimal correctors in conjunction with a fast nonoptimal algorithm that provides a good initial estimate. However, in general, these hybrid algorithms will be computationally less economical than the more direct calculation of the optimal attitude.

The QUEST Covariance Matrix

The covariance matrix for the quaternion is defined as follows: Let $\delta\bar{q}$ be the quaternion of the small rotation that takes the true quaternion into the optimal quaternion calculated according to the procedures described earlier in this section. In terms of quaternion composition

$$\bar{q}_{\mathrm{opt}} = \bar{q}_{\mathrm{true}} \otimes \delta\bar{q} \tag{84}$$

For historical reasons, a sequence of quaternions is written in opposite order to the same sequence of rotation matrices. $\delta\bar{q}$ is assumed to be unbiased, i.e.,

$$\langle \delta\bar{q} \rangle = \left\langle \begin{Bmatrix} \delta Q \\ \delta q \end{Bmatrix} \right\rangle = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} \tag{85}$$

where the brackets denote the expectation value. Note that Eq. (85) can hold for the scalar component of $\delta\bar{q}$ only to order $\langle |\delta Q|^2 \rangle$. The 3×3 quaternion covariance matrix is defined as

$$P_{QQ} = \langle \delta Q \delta Q^T \rangle \tag{86}$$

This is related to the covariance matrix of the Cartesian error angles of Sec. II by

$$P_{QQ} = \tfrac{1}{4} P_{\theta\theta} \tag{87}$$

which follows directly from

$$\delta Q = (\delta\theta_1, \delta\theta_2, \delta\theta_3)^T / 2 \tag{88}$$

where $\delta\theta_1$, $\delta\theta_2$, and $\delta\theta_3$ are the angles characterizing the infinitesimal rotation of Sec. II. The error vectors for the observation and reference unit vectors are taken to have axially symmetric distributions as in Sec. II.

Because the definition of the quaternion covariance matrix is independent of $\bar{q}_{\mathrm{true}}$, it may be taken to be the identity quaternion. Thus, it is sufficient to compute the quaternion covariance matrix under the special assumption that the true reference and observation vectors are identical.

$$\hat{W}_i = \hat{V}_i \qquad (i = 1,...,n) \tag{89}$$

It follows then that with very high confidence $\delta\bar{q}$ is the quaternion of a small rotation and the results of the previous subsection can be used to obtain an expression for $\delta\bar{q}$. Thus, from Eq. (79) to within very small quantities, we have

$$\delta Q = M^{-1} \delta Z \tag{90}$$

where

$$M = 2I - 2\sum_{i=1}^{n} a_i \hat{W}_i \hat{W}_i^T \tag{91}$$

$$\delta Z = \sum_{i=1}^{n} a_i (\delta\hat{W}_i \times \hat{V}_i + \hat{W}_i \times \delta\hat{V}_i) \tag{92}$$

Equation (86) now becomes

$$P_{QQ} = M^{-1} \langle \delta Z \delta Z^T \rangle M^{-1} \tag{93}$$

The evaluation of the expectation value is straightforward and leads to

$$\langle \delta Z \delta Z^T \rangle = \sum_{i=1}^{n} a_i^2 \sigma_i^2 [I - \hat{W}_i \hat{W}_i^T] \tag{94}$$

with

$$\sigma_i^2 = \sigma_{Vi}^2 + \sigma_{Wi}^2 \tag{95}$$

as before.

The a_i could now be chosen to minimize, for example, the trace of the covariance matrix. It is obvious from Eqs. (91) and (94) that the choice will, in general, depend on the configuration of the observation vectors in a complicated manner. A much simpler choice is to determine the weights a_i that minimize the original loss function of Eq. (34) when this is evaluated at the true attitude. This leads to

$$a_i = \sigma_{\mathrm{tot}}^2 / \sigma_i^2 \tag{96}$$

and the constant σ_{tot}^2 is determined from Eq. (35). Hence,

$$(\sigma_{\mathrm{tot}}^2)^{-1} = \sum_{i=1}^{n} (\sigma_i^2)^{-1} \tag{97}$$

Combining the above equations leads finally to

$$P_{QQ} = \frac{1}{4}\sigma_{\mathrm{tot}}^2 \left[I - \sum_{i=1}^{n} a_i \hat{W}_i \hat{W}_i^T \right]^{-1} \tag{98}$$

A unique solution for the optimal quaternion will exist provided that the inverse of P_{QQ}, the information matrix,

$$P_{QQ}^{-1} = 4\sum_{i=1}^{n} \frac{1}{\sigma_i^2} (I - \hat{W}_i \hat{W}_i^T) \tag{99}$$

is nonsingular.

For the special case where there are only two observation vectors, Eq. (98) may be simplified to

$$P_{QQ} = \tfrac{1}{4} \{ \sigma_{tot}^2 I + |\hat{W}_1 \times \hat{W}_2|^{-2} [(\sigma_1^2 - \sigma_{tot}^2) \hat{W}_1 \hat{W}_1^T + (\sigma_1^2 - \sigma_{tot}^2) \hat{W}_2 \hat{W}_2^T + \sigma_{tot}^2 (\hat{W}_1 \cdot \hat{W}_2) (\hat{W}_1 \hat{W}_2^T + \hat{W}_2 \hat{W}_1^T)] \} \tag{100}$$

Equation (98), just as Eq. (26), gives the covariance matrix in the body system. More conventionally, one defines the quaternion covariance as a 4×4 matrix given by

$$P_{qq} = \langle \Delta \bar{q} \Delta \bar{q}^T \rangle \tag{101}$$

where $\Delta \bar{q}$ is given by

$$\bar{q}_{opt} = \bar{q}_{true} + \Delta \bar{q} \tag{102}$$

The inertially referenced quaternion covariance matrix P_{qq} can then be shown to be given by

$$P_{qq} = [\bar{q}_{opt}] \left[\begin{array}{c|c} P_{QQ} & 0 \\ \hline 0^T & 0 \end{array} \right] [\bar{q}_{opt}]^T \tag{103}$$

and the matrix $[\bar{q}]$ is

$$[\bar{q}] = \begin{bmatrix} q_4 & -q_3 & q_2 & q_1 \\ q_3 & q_4 & -q_1 & q_2 \\ -q_2 & q_1 & q_4 & q_3 \\ -q_1 & -q_2 & -q_3 & q_4 \end{bmatrix} \tag{104}$$

The singular nature of P_{qq} is manifest.

The covariance matrix in the Euler angles may be obtained easily using Eqs. (80) and (33). It may be pointed out here also that Eqs. (103) and (104) are independent of the actual algorithm used to compute the quaternion.

IV. A Comparison of the TRIAD and QUEST Algorithms

There is a simple connection between the TRIAD and QUEST algorithms. Examine the loss function of Eq. (34) for the special case where there are only two observation vectors

$$L(A) = \tfrac{1}{2} a_1 |\hat{W}_1 - A\hat{V}_1|^2 + \tfrac{1}{2} a_2 |\hat{W}_2 - A\hat{V}_2|^2 \tag{105}$$

It will now be proved that the orthogonal matrix that minimizes $L(A)$ passes in the limit $a_2/a_1 \to 0$ into the TRIAD attitude matrix.

It may be noted first that as a_2/a_1 becomes increasingly smaller, the constraint

$$A\hat{V}_1 = \hat{W}_1 \tag{106}$$

is enforced with increasingly greater severity. Therefore, it is sufficient to show that the attitude matrix that minimizes the loss function of Eq. (105) subject to the constraint of Eq. (106) is the TRIAD attitude matrix.

In terms of the TRIAD vectors of Sec. II, the constraint becomes

$$A\hat{r}_1 = \hat{s}_1 \tag{107}$$

and the constrained loss function takes the form

$$L(A) = \tfrac{1}{2} a_2 | (\hat{W}_1 \cdot \hat{W}_2 - \hat{V}_1 \cdot \hat{V}_2) \hat{s}_1 - |\hat{W}_1 \times \hat{W}_2| \hat{s}_3 + |\hat{V}_1 \times \hat{V}_2| A\hat{r}_3 |^2 \tag{108}$$

Because A is orthogonal, $A\hat{r}_3$ can have no component along $\hat{s}_1$. Therefore, the choice

$$A\hat{r}_3 = \hat{s}_3 \tag{109}$$

minimizes the loss function subject to the constraint. Because A must also be unimodular, it follows that

$$A\hat{r}_2 = \hat{s}_2 \tag{110}$$

Hence, A is the TRIAD attitude matrix. This completes the proof.

Because the QUEST attitude matrix for the special case where there are only two observations passes into the TRIAD attitude matrix as $a_2/a_1 \to 0$ or, equivalently, as $\sigma_2/\sigma_1 \to \infty$, it follows that the respective covariance matrices must become equal in the same limit. Noting Eq. (87) and comparing Eq. (100) with Eq. (26) leads to

$$P_{QQ}^{TRIAD} = P_{QQ}^{QUEST} + \Delta P_{QQ} \tag{111}$$

where

$$\Delta P_{QQ} = \tfrac{1}{4} (\sigma_1^2 - \sigma_{tot}^2) \hat{s}_2 \hat{s}_2^T \tag{112}$$

which is manifestly positive semidefinite and gives the additional covariance of the TRIAD algorithm over and above that of the QUEST algorithm. Two special cases of interest are noted in the following paragraphs.

For $\sigma_1 = \sigma_2$, the quantity $\sigma_1^2 - \sigma_{tot}^2$ is equal to σ_{tot}^2. In this case the variance of the TRIAD solution along the axis $\hat{s}_2$ is twice that of the QUEST solution. Thus, the QUEST algorithm is to be preferred in this situation if attitude accuracy about that axis is critical.

On the other hand for $\sigma_2 \gg \sigma_1$

$$\sigma_1^2 - \sigma_{tot}^2 = \sigma_1^4 / \sigma_2^2 \tag{113}$$

which tends to zero as $\sigma_2/\sigma_1 \to \infty$, as required by the previous discussion. In this case, there is little advantage to using the QUEST algorithm over the TRIAD algorithm where there are only two observations. As a practical example, imagine the combination of a precision sun sensor with an accuracy of 10 arc-sec coupled with an Earth sensor with an accuracy of 20 arc-min. For this sensor combination the accuracies of the TRIAD and QUEST solutions about the $\hat{s}_2$ axis will differ only by 0.0003 arc-sec, which may be safely neglected.

The extra covariance term given by Eq. (112) has a very simple interpretation. It can be shown that

$$\Delta P_{QQ} = \langle Q(T-Q) Q^T(T-Q) \rangle \tag{114}$$

where $Q(T-Q)$ are the vector components of the quaternion connecting the TRIAD and QUEST solutions.

V. Conclusions

Two algorithms, TRIAD and QUEST, have been presented which determine three-axis attitude from vector observations with both high computational efficiency and high numerical accuracy. The two algorithms do not require the costly evaluation of trigonometric functions nor are they very costly of computer core. Thus, they are ideally suited to attitude groud support software systems when the attitude must be computed very frequently and also to onboard attitude determination systems, should these be required to determine single-frame (i.e., nonsequential) attitude estimates. Simple but realistic analytical expressions have been presented for the covariance matrices associated with these two algorithms. These will be useful both for large-volume ground processing of data and for mission planning. The importance of computing covariance matrices in the spacecraft body coordinate system has been demonstrated.

For moderate-accuracy missions, the increased accuracy of the QUEST algorithm, compared to that of the TRIAD algorithm, is generally not sufficient to offset the additional computational burden imposed, which is approximately twice as much for QUEST as for TRIAD (in terms of the number of FORTRAN statements). For high-accuracy missions, the opposite is usually the case and the QUEST algorithm is the algorithm of choice. When more than three observations must be employed in a single frame, the QUEST algorithm becomes computationally more efficient as well. A possible drawback is the method of sequential rotations, which potentially could increase the number of computations by a factor of 4. However, these extra computations will need to be implemented only over a small segment of each orbit, and for a high-accuracy mission such as Magsat, it has been estimated[5] that they need not be implemented at all. In the ground support software for the Magsat spacecraft, which was launched on Oct. 30, 1979, the QUEST algorithm has shown itself to be highly efficient and reliable.

Acknowledgments

The authors wish to thank Dr. F.L. Markley of the Naval Research Laboratory, Dr. R.W. Byrne of Computer Sciences Corporation, and E.J. Lefferts of NASA Goddard Space Flight Center for interesting discussions and helpful criticisms. Dr. Markley, in particular, has supplied several of the equations which appear in this report. The material on the connection between the TRIAD and the QUEST algorithms was written largely in response to a question posed by Dr. J. LeMay of the Aerospace Corporation. The support and encouragement of R.D. Werking of the Attitude Determination and Control Section of NASA Goddard Space Flight Center is gratefully acknowledged. This work was supported in part by the National Aeronautics and Space Administration under Contract NAS 5-24300.

References

[1] 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, p. 420-428.

[2] D. Niebur, et al., "HCMM Attitude Analysis and Support Plan," NASA X-581-78-4, March 1978, p. 4-149.

[3] Wahba, G., "A Least Squares Estimate of Satellite Attitude, Problem 65.1," *SIAM Review*, July 1966, pp. 385-386.

[4] Keat, J., "Analysis of Least-Squares Attitude Determination Routine, DOAOP," Computer Sciences Corp., Report CSC/TM-77/6034, Feb. 1977.

[5] Shuster, M.D., "Approximate Algorithms for Fast Optimal Attitude Computation," *Proceedings, AIAA Guidance and Control Conference*, Palo Alto, Calif., Aug. 7-9, 1978.

[6] 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. 410-420.

[7] Hildebrand, F.B., *Advanced Calculus for Applications*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

[8] Hoffman, K. and Kunze, R., *Linear Algebra*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1961.