The Journal of the Astronautical Sciences, Vol. 51, No. 1, January–March 2003, pp. 75–101

Constraint in Attitude Estimation Part II: Unconstrained Estimation[1]

Malcolm D. Shuster[2]

We compute for insight, not for numbers.

– Richard W. Hamming (1915–1998)

Abstract

The consequences of ignoring the norm constraint in quaternion estimation or the orthogonality constraint in the estimation of the attitude matrix are examined within the framework of batch maximum-likelihood estimation. Unconstrained estimation of the attitude matrix is shown to be a useful first step to the constrained estimation, because it confers global convergence to the estimation process. Unconstrained estimation of the quaternion, however, is shown to be fraught with problems. Apart from the fact that the unconstrained quaternion estimate can introduce errors which cannot be eliminated later, it can also lead to a singular (hence, noninvertible) inverse covariance matrix when used with a realistic measurement model. Consequently, the unconstrained quaternion estimate cannot be constructed. This has the additional consequence that there can be no recovery from an arbitrary initial condition. Certain practices associated with unconstrained quaternion estimation in the Kalman filter, in particular, the partial reset, are shown to interfere with the process of the correct restoration of the quaternion norm. Thus, one is forced to conclude that unconstrained quaternion estimation should be avoided in practice.

Introduction

As stated in the introduction to Part I [1] this work seeks to provide a more rigorous foundation for research on unconstrained attitude estimation [2–6] and to show many of the pitfalls of unconstrained quaternion estimation.

The first part of this work [1] presented the foundations of batch least-squares attitude estimation within the framework of Maximum-Likelihood Estimation and a

[1]This and the preceding article [1] are an expansion of an earlier conference report [2], presented in August 1993.

[2]Director of Research, Acme Spacecraft Company, 13017 Wisteria Drive, Box 328, Germantown, Maryland 20874. email: mdshuster@comcast.net.

detailed account of the attitude measurement sensitivity matrix. It was shown there that the unconstrained quaternion measurement sensitivity matrix was ambiguous, and, therefore, would lead to meaningless unconstrained estimates of the quaternion. In particular, it was shown that for one very physical case of quaternion estimation the estimation process broke down completely. In the present part we examine unconstrained attitude matrix and quaternion estimation in more detail.

The present Part II concentrates on specific examples of unconstrained attitude estimation. Our studies are restricted to the attitude matrix and the quaternion, since these (apart from the Cayley-Klein parameters, which differ only trivially from the quaternion) are the only higher-dimensional attitude representations in common use. Our attention is restricted again to batch attitude estimation, since batch estimation is more transparent then the jumble of equations in the Kalman filter. Thus, we limit out studies to systems without process noise, which means we will examine effectively static systems.[3] Within numerical error the Kalman filter must yield the same result as the batch estimator. Hence, a failure of batch estimation must be echoed in the failure of the Kalman filter for the same data. Our specific tool will be the covariance matrix (or inverse-covariance matrix if the former does not exist) of the unconstrained estimator, which we will frequently compare with the QUEST covariance matrix, which establishes a useful scale. This will prove to be more revealing than simulations, which only give a single sample.

The domain of the constrained quaternion is the unit sphere in four dimensions, usually denoted by S^3. (S^2 is the unit sphere in three dimensions, S^1 the unit circle.) The domain of the unconstrained quaternion is simply R^4, Thus, we may write for constrained and unconstrained quaternion estimates

$$\bar{q}^* \equiv \arg \min_{\bar{q} \epsilon S^3} J_{\bar{q}}(\bar{q}) \quad \text{and} \quad \bar{q}^*_{\text{UC}} \equiv \arg \min_{\bar{q} \epsilon R^4} J_{\bar{q}}(\bar{q}) \tag{1ab}$$

Similarly, the domain of the constrained attitude matrix is the set of elements of $SO(3)$ (which we will denote by $SO(3)$) and that of the unconstrained attitude matrix is $R^{3\times 3}$. Thus

$$A^* \equiv \arg \min_{A \epsilon SO(3)} J_A(A) \quad \text{and} \quad A^*_{\text{UC}} \equiv \arg \min_{A \epsilon R^{3\times 3}} J_A(A) \tag{2ab}$$

The same cost function is employed for both the constrained and unconstrained estimates. The unconstrained estimates[4] are obtained by Newton-Raphson iteration in the representation itself rather than in terms of $\boldsymbol{\epsilon}$. Generally, the calculation of the unconstrained estimate is trivial, compared with the constrained estimation problem.

As we have pointed out in Part I [1], unconstrained attitude estimation is meaningless, except in those unlikely cases where it leads (without tinkering) to an estimate which satisfies the constraint. Hence, the chief result of unconstrained attitude estimation is to transform the problem from estimating the constrained attitude representation given the data to estimating the constrained attitude representation given the unconstrained estimate and its covariance matrix. Thus, we will be occupied in this work largely with developing methods for restoring the proper constraint to the attitude representation.

[3]One could, of course, study measurements with process noise in a batch framework, but there is hardly any point to doing so, because the dimension of the measurement vector would be so large.

[4]*Unconstrained estimate* in this work will always refer to the attitude matrix or the quaternion and never to a three-dimensional representation, which is *ipso facto* constrained.

There are two general approaches to constrained estimation: (1) the method of Lagrange multipliers and (2) by iteration of a representation of lower dimension, which for the present work will always be $\boldsymbol{\epsilon}$-iteration. The method of Lagrange multipliers of its own has an unconstrained minimization which leads to a form of the estimate $\boldsymbol{\alpha}^*(\lambda)$, for which $\lambda = 0$ yields the unconstrained estimate and λ^* yields the properly constrained estimate. Thus, we obtain a usually continuous function connecting the unconstrained to the constrained attitude estimate. The problem of estimating the correctly constrained estimate is thereby transformed one step to finding the desired value of λ^*. The use of the method of Lagrange multipliers in attitude estimation is by no means exotic. The best known example of an (unconstrained) quaternion estimator which uses the Lagrange multiplier approach to restore the constraint is QUEST [7].[5] The $\boldsymbol{\epsilon}$-iteration is just the estimation procedure presented early in Part I of this work. In it, at each iteration, we write effectively

$$\Delta A_i = [\![\boldsymbol{\epsilon}_i]\!] A_i(-) \quad \text{and} \quad \Delta \bar{q}_i = \frac{1}{2}\,\Xi(\bar{q}_i(-))\boldsymbol{\epsilon}_i \tag{3ab}$$

In both approaches, the unconstrained estimate and its (unconstrained) covariance matrix substitute for the original data.

We begin Part II by examining the unconstrained estimation of the attitude matrix. We shall find that this unconstrained estimate converges globally in a single step and with a nonsingular 9×9 convariance matrix. There is no need for an *a priori* estimate in order to start the estimation. In addition, given the assumption of Gaussian noise, the unconstrained estimate provides a Gaussian sufficient statistic [8] for the attitude, and, therefore, an efficient and exact means for obtaining the properly constrained attitude estimate, which we illustrate by both of the two methods above. Further, since earlier workers have examined the Wahba problem [9] without constraint directly in terms of the attitude matrix [10, 11, 12], we do so also, with some new results. Unconstrained estimation of the attitude matrix offers a genuine benefit to attitude estimation.

The same cannot be said for unconstrained quaternion estimation. Unconstrained quaternion estimation is not globally convergent and requires an *a priori* starting value, and the iteration process does not converge in a single step (these two facts are not unrelated). In order to obtain useful results from our studies of unconstrained quaternion estimation, we start the iterative process at the true value of the quaternion, so that, barring highly diseased estimation problems, we can expect the iterative process to converge effectively, if not exactly, in a single step to the unconstrained maximum-likelihood estimate. This, obviously, is not a possible procedure for practical attitude estimation, but it is the most instructive. We develop the constraint restoration procedure also for the quaternion using both a Largrange multiplier and $\boldsymbol{\epsilon}$-iteration. However, because the unconstrained maximum-likelihood estimate of the quaternion is not a sufficient statistic [8], the constraint restoration cannot be exact, but if the data is of arc-second accuracy, then barring particularly diseased estimation problems, it will be close enough. Because of the low dimension of the quaternion, we can show explicitly that the Lagrange multiplier method truly does lead to the desired constrained quaternion estimate, at least to first order in the estimate error. We determine also the condition under

[5]The common property of the Lagrange-multiplier-based attitude estimation algorithms seems to be that they all have names (QUEST, FOAM, ESOQ, etc.)

which restoration of the norm constraint amounts to division by the quaternion norm. Most important of all, quaternion estimation from direction measurements alone leads to a singular 4×4 inverse covariance matrix. Thus, estimation of the unconstrained quaternion is impossible in this case no matter the quantity of data.

Unconstrained Estimation of the Attitude Matrix

The unconstrained estimation of the attitude matrix might seem a foolish task given the excess (six) of newly unconstrained parameters. It will turn out however, to be a useful exercise.

Assume that any two-dimensional focal-plane measurements have been converted to vector measurements following the prescriptions of Appendix A of Part I. Then the measurements will be either scalar or vector, each of which are linear in the elements of the attitude matrix A. We can write, therefore, in very general notation

$$\mathbf{z}_k = H^*_{\mathbf{A},k}\mathbf{A}_{\text{UC}} + \boldsymbol{\eta}_k \quad \text{with} \quad \boldsymbol{\eta}_k \sim \mathcal{N}(\mathbf{0}, R_k), \tag{4ab}$$

where the 9×1 attitude column vector is given by

$$\begin{aligned}\mathbf{A} &\equiv [A_{11}, A_{12}, A_{13}, A_{21}, A_{22}, A_{23}, A_{31}, A_{32}, A_{33}]^{\text{T}} \\ &\equiv [A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9]^{\text{T}}\end{aligned} \tag{5}$$

and

$$^{(\text{scalar})}H^*_{\mathbf{A}} = [u_1\mathbf{v}^{\text{T}}, u_2\mathbf{v}^{\text{T}}, u_3\mathbf{v}^{\text{T}}], \qquad ^{(\text{vector})}H^*_{\mathbf{A}} = \begin{bmatrix} \mathbf{v}^{\text{T}} & \mathbf{0}^{\text{T}} & \mathbf{0}^{\text{T}} \\ \mathbf{0}^{\text{T}} & \mathbf{v}^{\text{T}} & \mathbf{0}^{\text{T}} \\ \mathbf{0}^{\text{T}} & \mathbf{0}^{\text{T}} & \mathbf{v}^{\text{T}} \end{bmatrix} \tag{6ab}$$

which are 1×9 and 3×9 matrices, respectively.[6]

The weighted least-squares cost function is then

$$J(\mathbf{A}) = \frac{1}{2}\sum_{k=1}^{N}[\mathbf{z}_k - H^*_{\mathbf{A},k}\mathbf{A}]^{\text{T}}R_k^{-1}[\mathbf{z}_k - H^*_{\mathbf{A},k}\mathbf{A}] \tag{7}$$

Unconstrained minimization of $J(\mathbf{A})$ leads straightforwardly in one step to

$$(P^{\text{UC}}_{\mathbf{AA}})^{-1} = \sum_{k=1}^{N} H^{*\text{T}}_{\mathbf{A},k}R_k^{-1}H^*_{\mathbf{A},k}, \qquad \mathbf{A}^*_{\text{UC}} = P^{\text{UC}}_{\mathbf{AA}}\sum_{k=1}^{N} H^{*\text{T}}_{\mathbf{A},k}R_k^{-1}\mathbf{z}_k \tag{8ab}$$

Constraint Restoration for the Attitude Matrix I

We assume that there are sufficient measurements so that $(P^{\text{UC}}_{\mathbf{AA}})^{-1}$ is invertible (see below) and that $\mathbf{A}^*_{\text{UC}}$ is sufficiently close to the constrained estimate that it is extremely unlikely that when orthogonality is restored the determinant is not unity. If $(P^{\text{UC}}_{\mathbf{AA}})^{-1}$ were not invertible, then $\mathbf{A}^*$ could not be constructed. Since the measurement model of equation (4) is linear and Gaussian, $\mathbf{A}^*_{\text{UC}}$ is not only an effective measurement for $\mathbf{A}$ but a sufficient statistic [8] as well. Thus, $\mathbf{A}^*_{\text{UC}}$ and its 9×9 convariance matrix contain exactly the same information about $\mathbf{A}$ as the original measurements and measurement covariance matrices. This means that the weighted least-squares cost function can be written identically as

$$J(\mathbf{A}) = J'(\mathbf{A}) + Y \tag{9}$$

[6]Equation (6b) and equation (21) below appeared in the Kalman filter implementation of reference [4].

with

$$J'(\mathbf{A}) = \frac{1}{2}[\mathbf{A}^*_{\text{UC}} - \mathbf{A}]^{\text{T}}(P^{\text{UC}}_{\mathbf{AA}})^{-1}[\mathbf{A}^*_{\text{UC}} - \mathbf{A}] \tag{10a}$$

and Y is independent of $\mathbf{A}$. One shows readily that

$$Y = \frac{1}{2}\sum_{k=1}^{N} \mathbf{z}_k^{\text{T}} R_k^{-1} \mathbf{z}_k - \frac{1}{2}\mathbf{A}^*_{\text{UC}}(P^{\text{UC}}_{\mathbf{AA}})^{-1}\mathbf{A}^*_{\text{UC}} \tag{10b}$$

Since Y does not depend on $\mathbf{A}$, it plays no role in determining the constrained maximum-likelihood estimate and may be discarded.

We can now restore the orthogonality constraint by minimizing

$$J_\lambda(A) \equiv J'(\mathbf{A}) + \frac{1}{2}\sum_{i,j}\lambda_{ij}(\mathbf{a}_i^{\text{T}}\mathbf{a}_j - \delta_{ij}) \tag{11}$$

with $\lambda_{ij} = \lambda_{ji}$ so that, in reality, there are only six Lagrange multipliers. We have written

$$\mathbf{A} \equiv [\mathbf{a}_1^{\text{T}}, \mathbf{a}_2^{\text{T}}, \mathbf{a}_3^{\text{T}}]^{\text{T}} \tag{12}$$

This is identically equivalent to the minimization of

$$J'(\mathbf{A}) + \frac{1}{2}(\mathbf{A}^{\text{T}}\Lambda(\lambda)\mathbf{A} - \text{tr}\,\lambda) \tag{13}$$

where the symmetric 3×3 matrix λ and the symmetric 9×9 matrix $\Lambda(\lambda)$ is given by

$$\lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} \quad \text{and} \quad \Lambda(\lambda) = \begin{bmatrix} \lambda_{11} I_{3\times3} & \lambda_{12} I_{3\times3} & \lambda_{13} I_{3\times3} \\ \lambda_{21} I_{3\times3} & \lambda_{22} I_{3\times3} & \lambda_{23} I_{3\times3} \\ \lambda_{31} I_{3\times3} & \lambda_{32} I_{3\times3} & \lambda_{33} I_{3\times3} \end{bmatrix} \tag{14ab}$$

We minimize equation (13) without constraint and then chose the matrix λ so that the constraint is satisfied.

Straightforward minimization of $J_\lambda(A)$ without constraint leads to

$$\mathbf{A}^*(\lambda) = (I_{9\times9} + P^{\text{UC}}_{\mathbf{AA}}\Lambda(\lambda))^{-1}\mathbf{A}^*_{\text{UC}} \tag{15}$$

and we must solve for the matrix λ^* from the six equations

$$\mathbf{a}_i^{*\text{T}}(\lambda^*)\mathbf{a}_j^*(\lambda^*) = \mathbf{a}_{i\,\text{UC}}^{*\text{T}}(I_{9\times9} + P^{\text{UC}}_{\mathbf{AA}}\Lambda(\lambda^*))^{-2}\mathbf{a}^*_{j\,\text{UC}} = \delta_{ij} \quad 1 \le i \le j \le 3 \tag{16}$$

Constraint Restoration for the Attitude Matrix II

We can equally well construct the correctly constrained attitude matrix by $\boldsymbol{\epsilon}$-iteration. Write

$$\mathbf{A} \equiv [\mathbf{a}_1^{\text{T}}, \mathbf{a}_2^{\text{T}}, \mathbf{a}_3^{\text{T}}]^{\text{T}}, \qquad \mathbf{A}^*_{\text{UC}} \equiv [\mathbf{a}_{1\,\text{UC}}^{*\text{T}}, \mathbf{a}_{2\,\text{UC}}^{*\text{T}}, \mathbf{a}_{3\,\text{UC}}^{*\text{T}}]^{\text{T}} \tag{17ab}$$

Then define

$$\hat{\mathbf{a}}_1(-) \equiv \text{unit}(\mathbf{a}^*_{1\,\text{UC}})), \quad \hat{\mathbf{a}}_2(-) \equiv \text{unit}(\mathbf{a}^*_{3\,\text{UC}} \times \mathbf{a}^*_{1\,\text{UC}}), \quad \hat{\mathbf{a}}_3(-) \equiv \hat{\mathbf{a}}_1(-) \times \hat{\mathbf{a}}_2(-) \tag{18abc}$$

which is a right-hand orthonormal triad and

$$\mathbf{A}_1^*(-) = [\hat{\mathbf{a}}_1^{\text{T}}(-), \hat{\mathbf{a}}_2^{\text{T}}(-), \hat{\mathbf{a}}_3^{\text{T}}(-)]^{\text{T}} \tag{19}$$

is a first estimate of the constrained estimate of the 9×1 attitude vector. Further refinement will come from minimizing successively

$$J(\boldsymbol{\epsilon}_i) = \frac{1}{2}\sum_{k=1}^{N}[\mathbf{A}^*_{\rm UC} - \mathbf{A}^*_i(-) + H_\epsilon(\mathbf{A}^*_i(-))\boldsymbol{\epsilon}_i]^{\rm T} \times (P^{\rm UC}_{\mathbf{AA}})^{-1}[\mathbf{A}^*_{\rm UC} - \mathbf{A}^*_i(-) + H_\epsilon(\mathbf{A}^*_i(-))\boldsymbol{\epsilon}_i] \tag{20}$$

with

$$H_\epsilon(\mathbf{A}) = \begin{bmatrix} 0 & 0 & 0 & A_7 & A_8 & A_9 & -A_4 & -A_5 & -A_6 \\ -A_7 & -A_8 & -A_9 & 0 & 0 & 0 & A_1 & A_2 & A_3 \\ A_4 & A_5 & A_6 & -A_1 & -A_2 & -A_3 & 0 & 0 & 0 \end{bmatrix}^{\rm T} \tag{21}$$

One obtains by repeated iteration

$$P^{-1}_{\boldsymbol{\epsilon\epsilon},i} = H^{\rm T}_{\boldsymbol{\epsilon}}(\mathbf{A}^*_i(-))\,(P^{\rm UC}_{\mathbf{AA}})^{-1}H_{\boldsymbol{\epsilon}}(\mathbf{A}^*_i(-)) \tag{22}$$

$$\boldsymbol{\epsilon}^*_i(+) = P_{\boldsymbol{\epsilon\epsilon},i}H^{\rm T}_{\boldsymbol{\epsilon}}(\mathbf{A}^*_{\rm UC}(-))\,(P^{\rm UC}_{\mathbf{AA}})^{-1}(\mathbf{A}^*_{\rm UC} - \mathbf{A}^*_i(-)) \tag{23}$$

$$\mathbf{A}^*_{i+1}(-) = \mathbf{A}^*_i(+) = \mathbf{A}^*_i(-) + H_{\boldsymbol{\epsilon}}(\mathbf{A}^*_i(-))\boldsymbol{\epsilon}^*_i(+) \tag{24}$$

and

$$A^* = \lim_{i\to\infty} A^*_i(+), \qquad P^{-1}_{\boldsymbol{\epsilon\epsilon}} = H^{\rm T}_{\boldsymbol{\epsilon}}(\mathbf{A}^*)\,(P^{\rm UC}_{\mathbf{AA}})^{-1}H_{\boldsymbol{\epsilon}}(\mathbf{A}^*) \tag{25}$$

The properly constrained $P_{\mathbf{AA}}$ is given by

$$P_{\mathbf{AA}} = H_\epsilon(\mathbf{A}^*)P_{\epsilon\epsilon}H^{\rm T}_\epsilon(\mathbf{A}^*) \tag{26}$$

Note that we can write

$$A^*(\bar{q}(+)) = A(\bar{q}(+)) \quad \text{and} \quad A^* = A(\bar{q}^*) \tag{27ab}$$

with $A(\bar{q})$ given by equation (I-43).[7] Again, it is advisable to accumulate $\mathbf{A}^*$ as the quaternion in order to control the round-off error.

We can check the rank of $(P^{\rm UC}_{\mathbf{AA}})^{-1}$ by evaluating equation (8a) assuming that all measurements are unit vectors, distributed uniformly over the celestial sphere and obeying the QUEST measurement model, with the result

$$(P^{\rm UC}_{\mathbf{AA}})^{-1} = \frac{1}{\sigma^2_{\rm tot}}{\rm diag}(a,b,b,b,a,b,b,b,a) \tag{28}$$

where $a = 2/15$ and $b = 2/9$. Here ${\rm diag}(a,b,\ldots,z)$ denotes a diagonal matrix in terms of its diagonal elements. For the full-vector model $a = b = 1/3$. The unconstrained 9×9 covariance matrix is generally nonsingular, and the unconstrained estimation of the attitude matrix from direction and three-axis magnetometer data will succeed admirably. The above methods provide a general means for solving batch attitude estimation problems which might not otherwise be easily solvable.[8]

Unconstrained Attitude Matrix Estimation and the Wahba Problem

If we relax the orthogonality constraint on the attitude matrix, then the Wahba cost function becomes [11]

[7]Equation I-43 is to be interpreted as equation (43) of Part I.
[8]See reference [13] or [14] for examples of rather unwieldly deterministic methods.

$$J(A_{\mathrm{UC}}) = \frac{1}{2\sigma_{\mathrm{tot}}^2}\sum_{k=1}^{N} a_k[1 - 2\hat{\mathbf{W}}_k^{\mathrm{T}} A_{\mathrm{UC}}\hat{\mathbf{V}}_k + \hat{\mathbf{V}}_k A_{\mathrm{UC}}^{\mathrm{T}} A_{\mathrm{UC}}\hat{\mathbf{V}}_k]$$
$$= \frac{1}{\sigma_{\mathrm{tot}}^2}\left\{\frac{1}{2} - 2\mathrm{tr}[B^{\mathrm{T}} A_{\mathrm{UC}}] + \mathrm{tr}[E A_{\mathrm{UC}}^{\mathrm{T}} A_{\mathrm{UC}}]\right\} \tag{29}$$

where

$$B = \sum_{k=1}^{N} a_k \hat{\mathbf{W}}_k \hat{\mathbf{V}}_k^{\mathrm{T}} \quad \text{and} \quad E = \sum_{k=1}^{N} a_k \hat{\mathbf{V}}_k \hat{\mathbf{V}}_k^{\mathrm{T}} \tag{30ab}$$

Straightforward minimization of the cost function leads to

$$A_{\mathrm{UC}}^{*} = BE^{-1} \tag{31}$$

a result first due to Brock [10, 12].

The error in A_{UC}^{*} is simply

$$\Delta A_{\mathrm{UC}}^{*} = (\Delta B)E^{-1} = \sum_{k=1}^{N} a_k(\Delta\hat{\mathbf{W}}_k)\hat{\mathbf{V}}_k^{\mathrm{T}} E^{-1} \tag{32}$$

from which it is an easy matter to compute the mean orthogonality defect of A_{UC}, which is, using the QUEST measurement model

$$\Delta^{\mathrm{orthog}} \equiv E\{A_{\mathrm{UC}}^{*\mathrm{T}} A_{\mathrm{UC}}^{*} - I\} = 2\sigma_{\mathrm{tot}}^2 E^{-1} \tag{33}$$

where, as usual, $E\{\cdot\}$ denotes the expectation operation. The 9×9 covariance matrix of A_{UC}^{*} is also easy to calculate but uninstructive. To appreciate the scale of the orthogonality defect, note that the covariance method of the QUEST attitude solution (i.e., a constrained solution) is given by

$$P_{\epsilon\epsilon}^{\mathrm{QUEST}} = \sigma_{\mathrm{tot}}^2 (I - E)_{\mathrm{body}}^{-1} \tag{34}$$

where the subscript body means that the matrix has been transformed to the body frame. Note that E is resolved along space axes while $P_{\epsilon\epsilon}$ is resolved along body axes (i.e., $\hat{\mathbf{V}}_k \to \hat{\mathbf{W}}_k$).

In the isotropic case in which the measurements have equal variances and are distributed uniformly over the unit sphere, we have

$$E^{\mathrm{isotropic}} = \frac{1}{3} I_{3\times 3} \tag{35}$$

whence

$$\Delta_{\mathrm{isotropic}}^{\mathrm{orthog}} = 6\sigma_{\mathrm{tot}}^2 I_{3\times 3} \quad \text{and} \quad (P_{\epsilon\epsilon}^{\mathrm{QUEST}})_{\mathrm{isotropic}} = \frac{3}{2}\sigma_{\mathrm{tot}}^2 I_{3\times 3} \tag{36ab}$$

The error due to ignoring the constraint is four times larger than the error due to the measurement uncertainties.

More interesting is the case where the measurements (and hence the reference vectors) are confined to some small square field of view of full width α. Assuming that the measurements are spread uniformly over the field of view then the expected value of E is

$$E^{\mathrm{narrow\,FOV}} = A_{\mathrm{true}}^{\mathrm{T}} \begin{bmatrix} \beta & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & 1 - 2\beta \end{bmatrix} A_{\mathrm{true}} \tag{37}$$

with

$$\beta \approx \frac{\alpha^2}{12} \tag{38}$$

For an 8 deg by 8 deg field of view ($\beta \approx 1/600$)

$$\Delta^{\text{orthog}}_{\text{narrow FOV}} \approx \sigma^2_{\text{tot}} \begin{bmatrix} 1200. & 0 & 0 \\ 0 & 1200. & 0 \\ 0 & 0 & 2.0 \end{bmatrix} \tag{39a}$$

$$(P^{\text{QUEST}}_{\epsilon\epsilon})_{\text{narrow FOV}} \approx \sigma^2_{\text{tot}} \begin{bmatrix} 1.0 & 0 & 0 \\ 0 & 1.0 & 0 \\ 0 & 0 & 300. \end{bmatrix} \tag{39b}$$

In our "numerical" computations we suppose that the space axes, body axes, and sensor axes are all aligned, and that the sensor boresight is the z-axis. From Equations (39) we see that errors due to ignoring the constraint can overwhelm the genuine errors from the measurements.

The orthogonality defect can be significantly larger than typical attitude errors.

It is interesting to note that $B = A^*_{\text{UC}} E$ with E symmetric. Hence, the Wahba (constrained) attitude A^* will be closest to A^*_{UC} in the Schur norm [15].

$$A^* = \arg \min_{A \varepsilon SO(3)} \|A^*_{\text{UC}} - A\| \tag{40}$$

with the Schur norm for an arbitrary $n \times n$ matrix M defined by

$$\|M\| \equiv \sum_{i=1}^{n} \sum_{j=1}^{n} |M_{ij}|^2 \tag{41}$$

Equation (40) is true, because the QUEST algorithm simply effects a polar decomposition of B with A^* the orthogonal matrix of the decomposition. Thus, for A^* to be generated by the QUEST algorithm it is sufficient that $B = A^*S$, where S is a symmetric matrix. If $A^*_{\text{UC}} E$ qualifies as such a matrix, then obviously so must A^*_{UC}. Thus, A^*_{UC} inserted into the QUEST algorithm in place of B would generate A^*. However, only for $S = E$ will the QUEST algorithm yield the correct attitude estimate error covariance matrix. Note also that the mean-square estimate error in A^*_{UC} is given by

$$E\{\|A^*_{\text{UC}} - A^{\text{true}}\|\} = 2\sigma^2_{\text{tot}} \operatorname{tr}(E^{-1}) = \operatorname{tr}(\Delta^{\text{orthog}}) \tag{42}$$

Note that we cannot compute the covariance matrix for A^*_{UC} from the Hessian matrix of the Wahba cost function. This is because the result which allows us to replace the singular measurement covariance matrix $\sigma^2_k(I - \hat{\mathbf{W}}^{\text{true}}_k \hat{\mathbf{W}}^{\text{true}\,T}_k)$ by $\sigma^2_k I_{3\times3}$ is no longer true, as we have shown in Part I, if the orthogonality constraint is relaxed. Thus, we must compute the 9×9 covariance matrix of A by brute force using our result for A^*_{UC} and the QUEST measurement model.

Constraint Restoration for the Attitude Matrix III

Consider equations (8) above with $R_k = \sigma^2_k I_{3\times3}$ but with all of the measurements direction measurements. This Wahba cost function is no longer the correct MLE-derived cost function for unconstrained attitude estimation, but this really isn't important. Once the constraint of proper-orthogonality has been restored, than the Wahba cost function will be equivalent to that derived from MLE for direction measurements. In the notation of equation (30) we may write

$$(P_{\mathbf{AA}}^{\mathrm{UC}})^{-1} = \frac{1}{\sigma_{\mathrm{tot}}^2}\begin{bmatrix} E & 0 & 0 \\ 0 & E & 0 \\ 0 & 0 & E \end{bmatrix} \tag{43}$$

with E given by equation (30b). Then we can define

$$\mathbf{B} = \mathbf{A}_{\mathrm{UC}}^{*}\frac{1}{\sigma_{\mathrm{tot}}^2}(P_{\mathbf{AA}}^{\mathrm{UC}})^{-1} \tag{44}$$

Now $\mathbf{B}$ has the same relation to B as $\mathbf{A}$ has to A. It follows then that

$$A^* = \arg \min_{A \epsilon SO(3)} \| A - B(\mathbf{B}) \| \tag{45}$$

Thus, the QUEST algorithm may be used to restore the constraint to the attitude matrix (via the quaternion) and the correct attitude estimate error covariance matrix as well.

Restoration of Quaternion Constraint I

The quaternion, in order to be a quaternion of rotation, must have unit norm. We now address the problem of determining the quaternion of rotation given the unconstrained estimate $\bar{\boldsymbol{q}}_{\mathrm{UC}}^{*}$ and its (4×4) estimate error covariance matrix $P_{\bar{q}\bar{q}}^{\mathrm{UC}}$.

In the noise-free case, barring a failure of the estimation process (as in Part I), the estimate of the unconstrained quaternion should be the true quaternion of rotation. Therefore, we should have *approximately* that

$$\bar{q}_{\mathrm{UC}}^{*} = \bar{q} + \bar{\eta}_{\bar{q}} \tag{46}$$

where $\bar{q}$ is the quaternion of rotation and

$$\bar{\eta}_{\bar{q}} \sim \mathcal{N}(\mathbf{0}, P_{\bar{q}\bar{q}}^{\mathrm{UC}}) \tag{47}$$

The approximation[9] is in the assertion about $\bar{\eta}_{\bar{q}}$. The unconstrained estimate in this case is an approximate effective measurement for the quaternion of rotation. To obtain the quaternion of rotation we must constrain $\bar{q}$ to have unit norm. Thus, to estimate $\bar{q}$ we may use Lagrange's Method of Multipliers to enforce the constraint. The estimate of the quaternion of rotation, therefore, minimizes the cost function

$$J(\bar{q}) = \frac{1}{2}(\bar{q}_{\mathrm{UC}}^{*} - \bar{q})^{\mathrm{T}}(P_{\bar{q}\bar{q}}^{\mathrm{UC}})^{-1}(\bar{q}_{\mathrm{UC}}^{*} - \bar{q}) + \frac{1}{2}\lambda(\bar{q}^{\mathrm{T}}\bar{q} - 1) \tag{48}$$

without constraint, and the desired value λ^* of λ is the Lagrange multiplier which causes the constraint will be satisfied. Note that without the Lagrange term, the unconstrained minimization of the cost function would yield $\bar{q}_{\mathrm{UC}}^{*}$ with 4×4 covariance matrix $P_{\bar{q}\bar{q}}^{\mathrm{UC}}$. The minimization of the full cost function (including the Lagrange term) leads to

$$\bar{q}^*(\lambda) = (I + \lambda P_{\bar{q}\bar{q}}^{\mathrm{UC}})^{-1}\bar{q}_{\mathrm{UC}}^{*} \tag{49}$$

and λ^* is a solution of

$$f(\lambda^*) \equiv \bar{q}^{*\mathrm{T}}(\lambda^*)\bar{q}^*(\lambda^*) - 1 = \bar{q}_{\mathrm{UC}}^{*\mathrm{T}}(I + \lambda^* P_{\bar{q}\bar{q}}^{\mathrm{UC}})^{-2}\bar{q}_{\mathrm{UC}}^{*} - 1 = 0 \tag{50}$$

[9]If equation (47) were true then the quaternion cost function would be quadratic in $\bar{q}$. We know however, that it is quartic in $\bar{q}$. Thus, unlike the case for A_{UC}^{*}, $\bar{q}_{\mathrm{UC}}^{*}$ cannot be a true sufficient statistic of for the attitude. Likewise, $\bar{\eta}_{\bar{q}}$ cannot be truly Gaussian.

Since $\bar{q}^*_{\mathrm{UC}}$ should be close to the norm-constrained estimate, we expect $\lambda^* P^{\mathrm{UC}}_{\bar{q}\bar{q}}$ to be small. Therefore, it will usually be sufficient to calculate λ^*_o using one iteration of the Newton-Raphson method with vanishing initial value. Thus

$$\lambda^*_o \approx -\frac{f(0)}{f'(0)} = \frac{1}{2}(\bar{q}^{*\mathrm{T}}_{\mathrm{UC}} P^{\mathrm{UC}}_{\bar{q}\bar{q}} \bar{q}^*_{\mathrm{UC}})^{-1}(\bar{q}^{*\mathrm{T}}_{\mathrm{UC}} \bar{q}^*_{\mathrm{UC}} - 1) \tag{51}$$

If $\bar{q}(\lambda^*_o)$ is not sufficiently close to having unit norm, then the Newton-Raphson method is applied repeatedly to equation (50) to obtain further refinement.

Restoration of Quaternion Constraint II

Define as the first properly normed estimate of the quaternion

$$\bar{q}_1(-) \equiv \bar{q}^*_{\mathrm{UC}}/|\bar{q}^*_{\mathrm{UC}}| \equiv \mathrm{unit}(\bar{q}^*_{\mathrm{UC}}) \tag{52}$$

the naively normalized unconstrained quaternion estimate. Then the most general quaternion of rotation can be written at iteration i as

$$\bar{q} = \delta\bar{q}_i \otimes \bar{q}_i(-) = \bar{q}_i(-) + \Xi(\bar{q}_i(-))\boldsymbol{\epsilon}_i/2 \tag{53}$$

$\boldsymbol{\epsilon}^*_i(+)$ minimizes the cost function

$$J(\boldsymbol{\epsilon}_i) = \frac{1}{2}(\bar{q}^*_{\mathrm{UC}} - \bar{q}^*_i(-) - \Xi(\bar{q}^*_i(-))\boldsymbol{\epsilon}_i/2)^{\mathrm{T}}(P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}(\bar{q}^*_{\mathrm{UC}} - \bar{q}^*_i(-) - \Xi(\bar{q}^*_i(-))\boldsymbol{\epsilon}_i/2) \tag{54}$$

with the result

$$\boldsymbol{\epsilon}^*_i(+) = 2P_{\boldsymbol{\epsilon\epsilon},i}\Xi^{\mathrm{T}}(\bar{q}^*_i(-))\,(P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}(\bar{q}^*_{\mathrm{UC}} - \bar{q}^*_i(-)) \tag{55}$$

with

$$P^{-1}_{\boldsymbol{\epsilon\epsilon},i} = \frac{1}{4}\Xi^{\mathrm{T}}(\bar{q}^*_i(-))(P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}\Xi(\bar{q}^*_i(-)) \tag{56}$$

Thus, after the sequence of estimates converges to $\bar{q}^*$

$$P^{-1}_{\boldsymbol{\epsilon\epsilon}} = \frac{1}{4}\Xi^{\mathrm{T}}(\bar{q}^*)(P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}\Xi(\bar{q}^*) \tag{57}$$

is the inverse covariance matrix of $\boldsymbol{\epsilon}^*$. The 4×4 covariance matrix of the quaternion estimate with restored norm is clearly

$$P_{\bar{q}\bar{q}} = \Xi(\bar{q}^*)P_{\boldsymbol{\epsilon\epsilon}}\Xi^{\mathrm{T}}(\bar{q}^*)/4 \tag{58}$$

with the singularity explicit. Equation (55) shows clearly that simply dividing the unconstrained estimate of the quaternion by its norm does not provide the proper unit quaternion, in general. Note that $\boldsymbol{\epsilon}^*_1(+)$ is proportional to the norm defect

$$\boldsymbol{\epsilon}^*_1(+) = 2(|\bar{q}^*_1(-)| - 1)P_{\boldsymbol{\epsilon\epsilon},1}\Xi^{\mathrm{T}}(\bar{q}^*_1(-))(P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}\bar{q}^*_1(-) \tag{59}$$

and that it vanishes identically if $P^{\mathrm{UC}}_{\bar{q}\bar{q}}$ is a multiple of the 4×4 identity matrix. These statements need not mean, however, that $\boldsymbol{\epsilon}^*_1(+)$ is negligibly small. If $\bar{q}^*_{\mathrm{UC}}$ had been determined, say, from data from a star camera with very narrow field of view, then the eigenvalues of $P^{\mathrm{UC}}_{\bar{q}\bar{q}}$ might differ by very large factors.

Note in passing that $\bar{q}^*_{\mathrm{UC}} - \bar{q}^*_i$ can almost never vanish, because the two quantities are unlikely (with probability 1) to have the same norm. Therefore, for the right member of equation (55) to vanish, which must occur at the converged estimate $\bar{q}^*$, we must have

$$(P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}(\bar{q}^*_{\mathrm{UC}} - \bar{q}^*) = \alpha \bar{q}^* \tag{60}$$

for some value α, so that this quantity can be annihilated by the factor $\Xi^{\mathrm{T}}(\bar{q}^*)$). Solving for $\bar{q}^*$ leads to

$$\bar{q}^* = (I + \alpha P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}\bar{q}^*_{\mathrm{UC}} \tag{61}$$

which is just the Lagrange multiplier solution again with α the Lagrange multiplier. The two methods are clearly equivalent.

A Hybrid Estimator

If one insists on unconstrained estimation as a first step, then it is clear that the best method is to estimate the unconstrained attitude matrix, and use the unconstrained estimator as a sufficient statistic for estimating the constrained quaternion, either via Lagrange multipliers, by $\boldsymbol{\epsilon}$-iteration, or via QUEST.

Constraint or Restored Constraint, That is the Question!

It is possible to demonstrate explicitly to first order that the quaternion with correctly restored constraint is indeed the same as the quaternion one would obtain from carrying out a properly constrained estimation. Consider the least-squares cost function which we write as

$$J(\Delta\bar{q}_{\mathrm{UC}}) = \frac{1}{2}\sum_{k=1}^{N}(\Delta\mathbf{z}_k - H^*_k\,\Delta\bar{q}^{\mathrm{UC}}_k)^{\mathrm{T}}R_k^{-1}(\Delta\mathbf{z}_k - H^*_k\,\Delta\bar{q}^{\mathrm{UC}}_k) \tag{62}$$

Here $\Delta\mathbf{z}_k$ is the residual measurement, given by

$$\Delta\mathbf{z}_k \equiv z_k - z_{o,k} = H^*_k\,\Delta\bar{q}_{\mathrm{UC}} + \boldsymbol{\eta}_k, \quad k = 1, \ldots, N \tag{63}$$

with H^*_k given as in Part I, the constraint-insensitive measurement sensitivity matrix (note the asterisk), and R_k is the covariance matrix of the measurement noise. We assume that both the unconstrained and constrained estimates can be reached within a single iteration of the appropriate Gauss-Newton algorithm to within terms of order $|\Delta\bar{q}_{\mathrm{UC}}|^2$. This is achieved most easily by setting $\bar{q}(-) = \bar{q}_{\mathrm{true}}$. For simplicity we assume that $\bar{q}^{\mathrm{true}} = \bar{1}$, the identity quaternion. Thus, we are performing the update with respected to predicted body axes.

For the unconstrained quaternion correction, (which is not constrained to preserve the norm and treats the four components of the quaternion as independent) the estimate is given by

$$\Delta\bar{q}^*_{\mathrm{UC}} = P^{\mathrm{UC}}_{\bar{q}\bar{q}}\,\bar{p}_{\mathrm{UC}} = \begin{bmatrix}\Delta\mathbf{q}^*_{\mathrm{UC}} \\ \Delta q_{4\,\mathrm{UC}}\end{bmatrix} \tag{64}$$

where the 4×4 convariance matrix $P^{\mathrm{UC}}_{\bar{q}\bar{q}}$ and the information quaternion $\bar{p}_{\mathrm{UC}}$ are given by

$$P^{\mathrm{UC}}_{\bar{q}\bar{q}} = \left[\sum_{k=1}^{N} H^{*\mathrm{T}}_k R_k^{-1} H^*_k\right]^{-1}, \quad \bar{p}_{\mathrm{UC}} = \sum_{k=1}^{N} H^{*\mathrm{T}}_k R_k^{-1}\,\Delta\mathbf{z}_k \tag{65ab}$$

Note that implicit in equations (65) is the assumption that $P_{\bar{q}\bar{q}}^{\mathrm{UC}}$ be invertible.

For the constrained estimate, which *is* norm-preserving to first order, for the same data is (note that we estimate only the vectorial coordinates and determine the scalar component from the norm constraint) we obtain

$$\Delta\mathbf{q}^*_{\text{constrained}} = P_{\text{constrained}}\mathbf{p}_{\text{constrained}} \tag{66}$$

with

$$P_{\text{constrained}} = \left[\sum_{k=1}^{N} H_{1,k}^{\mathrm{T}} R_k^{-1} H_{1,k}\right]^{-1} = ((P_{\bar{q}\bar{q}}^{\mathrm{UC}})^{-1})_{11}^{-1} \tag{67a}$$

$$\mathbf{p}_{\text{constrained}} = \sum_{k=1}^{N} H_{1,k}^{\mathrm{T}} R_k^{-1} \Delta\mathbf{z}_k \tag{67b}$$

and we have written

$$H_k^* = [H_{1,k} \,|\, H_{2,k}^*] \tag{68}$$

where $H_{1,k}$ is the partition which includes the first three columns of H_k^*, and $H_{2,k}^*$ is the partition containing the fourth column. In our example, we have also that $H_{1,k}^* = H_{1,k}$. Had we not chosen to make the *a priori* quaternion $\bar{1}$, this simple decomposition of H_k^* would not have been obtained. It follows from the comparison of equations (65) and (67) that

$$\mathbf{p}_{\text{constrained}} = \mathbf{p}_{\text{UC}} \tag{69}$$

where $\mathbf{p}_{\text{UC}}$ denotes the vectorial components of $\bar{p}_{\text{UC}}$. Hence, we can find a relation between the constrained and unconstrained corrections to the quaternion by solving equation (64) for $\bar{p}_{\text{UC}}$ in terms of $\Delta\bar{q}^*_{\text{UC}}$ and inserting the value of $\mathbf{p}_{\text{UC}}$ from this expression in equation (66). This leads to

$$\Delta\mathbf{q}^*_{\text{constrained}} = \Delta\mathbf{q}^*_{\text{UC}} + (P_{\bar{q}\bar{q}}^{-1})_{11}^{-1}(P_{\bar{q}\bar{q}}^{-1})_{12}\,\Delta q^*_{4\,\text{UC}} \tag{70}$$

where, consistent with the partition of H_k, we have partitioned the 4×4 quaternion covariance and information matrices as

$$P_{\bar{q}\bar{q}}^{\mathrm{UC}} = \begin{bmatrix} (P_{\bar{q}\bar{q}}^{\mathrm{UC}})_{11} & (P_{\bar{q}\bar{q}}^{\mathrm{UC}})_{12} \\ (P_{\bar{q}\bar{q}}^{\mathrm{UC}})_{21} & (P_{\bar{q}\bar{q}}^{\mathrm{UC}})_{22} \end{bmatrix} \quad \text{and} \quad (P_{\bar{q}\bar{q}}^{\mathrm{UC}})^{-1} = \begin{bmatrix} ((P_{\bar{q}\bar{q}}^{\mathrm{UC}})^{-1})_{11} & ((P_{\bar{q}\bar{q}}^{\mathrm{UC}})^{-1})_{12} \\ ((P_{\bar{q}\bar{q}}^{\mathrm{UC}})^{-1})_{21} & ((P_{\bar{q}\bar{q}}^{\mathrm{UC}})^{-1})_{22} \end{bmatrix} \tag{71ab}$$

We will return to equation (70) and equations (71) soon. Note that equation (70) demonstrates that for our linearized measurements with approximately Gaussian noise $\bar{q}^*_{\text{UC}}$ is an effective measurement) for $\bar{q}$, at the least to first order. Note also that the constraint-insensitive and constraint-sensitive measurement sensitivities in our example can differ only by the value of $H_{2,k}$.

With equation (70) we are given the prescription for constructing the standard estimate and its covariance matrix from the unconstrained quaternion estimate and its covariance. From equation (51)

$$\lambda = (P_{\bar{q}\bar{q}}^{\mathrm{UC}})_{22}^{-1}\,\Delta q_{4\,\text{UC}} \tag{72}$$

Substituting this in equation (49) leads to lowest order in $\Delta\bar{q}^*_{\text{UC}}$

$$\begin{aligned}\bar{q}^* &= (I + \lambda P_{\bar{q}\bar{q}}^{\mathrm{UC}})^{-1}\bar{q}^*_{\text{UC}} \\ &\approx \bar{q}^*_{\text{UC}} - \Delta q^*_{4\,\text{UC}}(P_{\bar{q}\bar{q}}^{\mathrm{UC}})_{22}^{-1} P_{\bar{q}\bar{q}}^{\mathrm{UC}}\bar{q}^*_{\text{UC}}\end{aligned} \tag{73}$$

The vectorial component of the desired optimal quaternion is simply (to this same order)

$$\Delta\mathbf{q}^* = \Delta\mathbf{q}^*_{\mathrm{UC}} - (P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}_{22}(P^{\mathrm{UC}}_{\bar{q}\bar{q}})_{12}\,\Delta q^*_{4\mathrm{UC}} \tag{74}$$

But

$$-(P^{\mathrm{UC}}_{\bar{q}\bar{q}})_{12}(P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}_{22} = ((P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1})^{-1}_{11}((P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1})_{12} \tag{75}$$

so that, in fact, comparing equation (74) with equation (70) we have

$$\mathbf{q}^* = \mathbf{q}^*_{\mathrm{constrained}} \tag{76}$$

It follows that

$$\bar{q}^* = \bar{q}_{\mathrm{constrained}} \tag{77}$$

since $q_{4,\,\mathrm{constrained}}$ is determined from the constraint. Thus, the unconstrained estimate to the quaternion followed by the correct normalization correction is identical (within terms of order $|\Delta\bar{q}^*_{\mathrm{UC}}|^2$) to the standard estimate. Note again that this correct normalization correction is more complicated than the simple division by the quaternion norm.[10]

Unconstrained Quaternion Estimation for the Wahba Cost Function and the QUEST Measurement Model

It is now time that we examine the covariance matrix of the unconstrained quaternion estimate in detail. To avoid the problem of the unknowable initial estimate of the quaternion we shall simply assume once again that the initial estimate $\bar{q}(-)$ is simply $\bar{q}^{\mathrm{true}}$ and study the effect of one iteration of the Newton-Raphson method. On this basis we shall study the behavior of the 4×4 covariance matrix for the unconstrained quaternion. The 4×4 covariance matrix cannot be evaluated in this case as the inverse of the expectation value of the Hessian matrix (or *companion matrix*) but must be calculated by brute force. This is because, as pointed out in Part I, the Wahba problem is not the MLE-derived cost function for unconstrained estimation. This is not a great hardship.

Since the *a priori* value for the quaternion is $\bar{q}^{\mathrm{true}}$, we are led to examine a cost function of the form

$$\begin{aligned} J(\Delta\bar{q}_{\mathrm{UC}}) &= \frac{1}{2}\sum_{k=1}^{N}\frac{1}{\sigma_k^2}|\hat{\mathbf{W}}_k - A(\bar{q}^{\mathrm{true}} + \Delta\bar{q}_{\mathrm{UC}})\hat{\mathbf{V}}_k|^2 \\ &= \frac{1}{2\sigma_{\mathrm{tot}}^2}\sum_{k=1}^{N} a_k|\Delta\hat{\mathbf{W}}_k - H_k^*\,\Delta\bar{q}_{\mathrm{UC}})|^2 \end{aligned} \tag{78}$$

A single iteration of the Newton-Raphson method yields

$$\Delta\bar{q}^*_{\mathrm{UC}} = C^{-1}\frac{1}{\sigma_{\mathrm{tot}}^2}\sum_{k=1}^{N} a_k H_k^{*\mathrm{T}}\,\Delta\hat{\mathbf{W}}_k \tag{79}$$

where C is the *companion matrix*

$$C = \frac{1}{\sigma_{\mathrm{tot}}^2}\sum_{k=1}^{N} a_k H_k^{*\mathrm{T}} H_k^* \tag{80}$$

[10]The latter, of course, is the correct operation to compensate for errors arising from the finite precision of the computations (round-off error) since these are not dependent on the statistics of the estimate in any way.

The 4 × 4 quaternion covariance matrix is then given by

$$(P_{\bar{q}\bar{q}}^{\rm UC}) = C^{-1} G C^{-1} \tag{81}$$

with

$$\begin{aligned} G &= \frac{1}{(\sigma_{\rm tot}^2)^2} E\left\{ \sum_{k=1}^{N} a_k H_k^{*\rm T} \Delta\hat{\mathbf{W}}_k \sum_{m=1}^{N} a_m \Delta\hat{\mathbf{W}}_m^{\rm T} H_m^* \right\} \\ &= \frac{1}{\sigma_{\rm tot}^2} \sum_{k=1}^{N} a_k H_k^{*\rm T} (I - \hat{\mathbf{W}}_k^{\rm true} \hat{\mathbf{W}}_k^{\rm true\,T}) H_k^* \\ &= \left\{ C - \frac{1}{\sigma_{\rm tot}^2} \sum_{k=1}^{N} a_k H_k^{*\rm T} \hat{\mathbf{W}}_k^{\rm true} \hat{\mathbf{W}}_k^{\rm true\,T} H_k^* \right\} \equiv C - D \end{aligned} \tag{82}$$

The inconsistency of the cost function and measurement covariance matrix is now also clear from the presence of D, because asymptotically C would be the Fisher information matrix if J were the MLE cost function consistent with the measurement model.

The calculation of the companion matrix is straightforward. We take as the measurement sensitivity matrix

$$H_k^* = -2\Xi^{\rm T}(\bar{q} \otimes \overline{\mathbf{V}}_k) + 2a\hat{\mathbf{V}}_k \bar{q}^{\rm T} \tag{83}$$

where a is a free parameter, which may have any value (see equation (I-77)). For $a = 0$ we obtain the measurement sensitivity matrix corresponding to equation (I-43), for $a = -1$ that corresponding to equation (I-73). This is by no means the most general form for H_k^*, but it will be enough to show the strong dependence of the covariance matrix on the choice of constraint-insensitive measurement sensitivity matrix.

For an infinitesimal rotation ($\bar{q}^*(-) = \bar{1}$, or, equivalently, for a rotation from predicted body axes)

$$C = \frac{4}{\sigma_{\rm tot}^2} \begin{bmatrix} I_{3\times 3} - E & \mathbf{0} \\ \mathbf{0}^{\rm T} & (1+a)^2 \end{bmatrix}, \qquad D = \frac{4}{\sigma_{\rm tot}^2} \begin{bmatrix} \mathbf{0}_{3\times 3} & \mathbf{0} \\ \mathbf{0}^{\rm T} & (1+a)^2 \end{bmatrix} \tag{84ab}$$

with the result

$$P_{\bar{q}\bar{q}}^{\rm UC} = \frac{\sigma_{\rm tot}^2}{4} \begin{bmatrix} (I_{3\times 3} - E)^{-1} & \mathbf{0} \\ \mathbf{0}^{\rm T} & 0 \end{bmatrix} = \begin{bmatrix} P_{qq}^{\rm QUEST} & \mathbf{0} \\ \mathbf{0}^{\rm T} & 0 \end{bmatrix} \tag{85}$$

for $a \neq -1$.

For $a = 0$ in which case the attitude matrix is related to the quaternion according to equation (I-43), the 4 × 4 quaternion covariance matrix is exactly what one would expect from a properly constrained estimate. For $a = -1$, in which case the attitude matrix is related to the quaternion by equation (I-73), the quaternion covariance matrix is given by

$$P_{\bar{q}\bar{q}}^{\rm UC} = \begin{bmatrix} P_{qq}^{\rm QUEST} & \mathbf{0} \\ \mathbf{0}^{\rm T} & \infty \end{bmatrix} \tag{86}$$

It might seem that the quaternion covariance is indeterminate at this value. However, it is clear from the measurement sensitivity matrix that the measurement is insensitive to $q_4^{\rm UC}$, so that the corresponding information must be zero, and P_{44} infinite.

Unconstrained Quaternion Estimation with the Wahba Cost Function and the Full-Vector Measurement Model

The calculation here, again for $\bar{q}(-) = \bar{q}^{\text{true}} = \bar{1}$, is trivial since the result must be simply

$$(P_{\bar{q}\bar{q}}^{\text{UC}})^{-1} = C = \frac{4}{\sigma_{\text{tot}}^2}\begin{bmatrix} (I^{3\times 3} - E) & \mathbf{0} \\ \mathbf{0}^{\text{T}} & (1 + a)^2 \end{bmatrix} \tag{87}$$

with C given by equation (84a). The result is not the same as that from the Wahba problem. The fact that the result for the Wahba cost function shows no dependence on the free parameter in the measurement sensitivity matrix or even on the lack of constraint does not mean that it is more correct. The Wahba cost function is not consistent with the QUEST measurement model. Equations (87) and (88) will have important consequences for the Kalman filter.

Unconstrained Quaternion Estimation for the QUEST Measurement Model and the QMM Cost Function

In this case we have directly from the Fisher information matrix that

$$(P_{\bar{q}\bar{q}}^{\text{UC}})^{-1} = C - D = \frac{4}{\sigma_{\text{tot}}^2}\begin{bmatrix} (I_{3\times 3} - E) & \mathbf{0} \\ \mathbf{0}^{\text{T}} & 0 \end{bmatrix} \tag{88}$$

The Fisher information matrix is no longer invertible. Thus, with the more physical of the two measurement models and the correct MLE cost function, the covariance matrix is infinite. The estimation problem can be made finite only by having an *a priori* quaternion covariance matrix for which P_{44}^{UC} is nonvanishing. One could also consider having a three-axis magnetometer as one of the sensors, but that is not the purpose of this exercise.

The complexities and possible failures of estimating a quaternion in these ways as opposed to a completely constrained estimation via QUEST, SVD, FOAM, ESOQ and other algorithms [7, 16] should be noted.

Quaternion Norm Restoration Revisited

Let us reëxamine the quaternion norm restoration. We assume again for convenience that $\bar{q}(-) = \bar{q}_{\text{true}} = \bar{1}$. Hence, we may write

$$\bar{q}_{\text{UC}}^{*} = \begin{bmatrix} \Delta\mathbf{q}_{\text{UC}}^{*} \\ 1 + \Delta q_{4\,\text{UC}}^{*} \end{bmatrix} \tag{89}$$

We suppose also that the covariance has the form

$$P_{\bar{q}\bar{q}}^{\text{UC}} = \begin{bmatrix} P_{qq}^{\text{UC}} & \mathbf{0} \\ \mathbf{0}^{\text{T}} & P_{44}^{\text{UC}} \end{bmatrix} \tag{90}$$

that is, that $\Delta\mathbf{q}_{\text{UC}}^{*}$ is uncorrelated with $\Delta\mathbf{q}_{4\,\text{UC}}^{*}$. Thus far, this has been the case with the measurement model for both unit-vector and full-vector sensors. For a general model measurement covariance matrix R_k, this will not be true and the result (for the inverse unconstrained 4×4 quaternion covariance matrix) will be

$$(P_{\bar{q}\bar{q}}^{\text{UC}})^{-1} = \sum_{k=1}^{N}\begin{bmatrix} [\![\hat{\mathbf{W}}_k]\!]R_k^{-1}[\![\hat{\mathbf{W}}_k]\!]^{\text{T}} & [\![\hat{\mathbf{W}}_k]\!]R_k^{-1}\hat{\mathbf{W}}_k \\ \hat{\mathbf{W}}_k^{\text{T}}R_k^{-1}[\![\hat{\mathbf{W}}_k]\!]^{\text{T}} & \hat{\mathbf{W}}_k^{\text{T}}R_k^{-1}\hat{\mathbf{W}}_k \end{bmatrix} \tag{91}$$

Clearly, we can expect a lack of correlation between $\Delta\mathbf{q}^*_{\mathrm{UC}}$ and $\Delta q^*_{4\mathrm{UC}}$ only if

$$\sum_{k=1}^{N} [\![\mathbf{W}_k]\!] R_k^{-1} \mathbf{W}_k = \mathbf{0} \tag{92}$$

for all measurements, which will be true for all measurements only if R_k is proportional to $I_{3\times3}$ or to $I_{3\times3} - \hat{\mathbf{W}}\hat{\mathbf{W}}^{\mathrm{T}}$. In general, for more realistic systems, $\Delta\mathbf{q}^*_{\mathrm{UC}}$ and $\Delta q^*_{4\mathrm{UC}}$ or more generally

$$\Delta\mathbf{q}^*_{\|\mathrm{UC}} \equiv \bar{q}^{\mathrm{T}}(-)\,\Delta\bar{q}^*_{\mathrm{UC}} \quad \text{and} \quad \Delta\mathbf{q}^*_{\perp\mathrm{UC}} \equiv \Xi^{\mathrm{T}}(\mathbf{q}(-))\Delta\bar{q}^*_{\mathrm{UC}} \tag{93ab}$$

will be correlated. (See Appendix B of Part I.)

Assuming equation (90) we have for the Lagrange multiplier

$$\lambda = \frac{\Delta q^*_{4\mathrm{UC}}}{(P^{\mathrm{UC}}_{44})^2(1 + \Delta q^*_{4\mathrm{UC}})^2} + O(|\Delta\bar{q}^*_{\mathrm{UC}}|^2) \tag{94}$$

Hence, the norm restoration yields after a single Newton-Raphson iteration (as before)

$$\begin{aligned}\bar{q}^* &= (I + \lambda P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}\bar{q}^*_{\mathrm{UC}} + O(|\Delta\bar{q}^*_{\mathrm{UC}}|^2)\\ &= (I - \lambda P^{\mathrm{UC}}_{\bar{q}\bar{q}})\bar{q}^*_{\mathrm{UC}} + O(|\Delta\bar{q}^*_{\mathrm{UC}}|^2)\\ &= \begin{bmatrix}\Delta\mathbf{q}^*_{\mathrm{UC}}\\ 1\end{bmatrix} + O(|\Delta\bar{q}^*_{\mathrm{UC}}|^2)\end{aligned} \tag{95}$$

Thus, the norm error has been reduced from $O(|\Delta\bar{q}^*_{\mathrm{UC}}|)$ to $O(|\Delta\bar{q}^*_{\mathrm{UC}}|^2)$ as expected. We further note that after one iteration

$$\bar{q}^* = \bar{q}^*_{\mathrm{UC}}/|\bar{q}^*_{\mathrm{UC}}| + O(|\Delta\bar{q}^*_{\mathrm{UC}}|^2) \tag{96}$$

provided that $\Delta\mathbf{q}^*_{\mathrm{UC}}$ and $\Delta\mathbf{q}^*_{4\mathrm{UC}}$ are uncorrelated. Thus, if $P^{\mathrm{UC}}_{\bar{q}\bar{q}}$ has the structure of equation (90), then the correctly constrained quaternion can be obtained, at least to first order, by division by the quaternion norm.

For a general measurement covariance matrix we obtain

$$\Delta\mathbf{q}^* = \Delta\mathbf{q}^*_{\mathrm{UC}} - (P^{\mathrm{UC}}_{44})^{-1}P^{\mathrm{UC}}_{\mathbf{q}4}\,q^*_{4\mathrm{UC}} \tag{97}$$

and $\Delta\mathbf{q}^*$ will differ from $\Delta\mathbf{q}^*_{\mathrm{UC}}$ by terms of order $\Delta\bar{q}^*_{\mathrm{UC}}$, which means that the constrained quaternion cannot be determined in general by dividing the unconstrained quaternion by its norm. (See Appendix B of Part I for the general decomposition for an arbitrary frame.)

Persistence of an Initial Condition

In theoretical studies the Kalman filter is frequently initialized arbitrarily, generally with very large initial covariance, and the ability of the Kalman filter to recover from this initial condition is then part of the simulation testing. In a well-designed mission, obviously, such a procedure should not be needed, and its execution could pose a danger to the spacecraft. Nonetheless, for many practitioners of attitude estimation it seems to be an important part of performance tests of the unconstrained filters and will be examined here. Thus, for the full-vector measurement model with an initial condition we should examine (as usual, for $\bar{q}(-) = \bar{q}^{\mathrm{true}} = \bar{1}$), recalling equations (87) and (88) above

$$(P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}_{\mathrm{FVM}} = \frac{4}{\sigma^2_{\mathrm{tot}}}\begin{bmatrix}(I - E) & \mathbf{0}\\ \mathbf{0}^{\mathrm{T}} & (1 + a)^2\end{bmatrix} + \frac{4}{\sigma^2_o}\begin{bmatrix}I_{3\times3} & \mathbf{0}\\ \mathbf{0}^{\mathrm{T}} & 1\end{bmatrix} \tag{98}$$

and for the QUEST measurement model

$$(P^{\mathrm{UC}}_{\bar{q}\bar{q}})^{-1}_{\mathrm{QMM}} = \frac{4}{\sigma^2_{\mathrm{tot}}}\begin{bmatrix} (I - E) & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & 0 \end{bmatrix} + \frac{4}{\sigma^2_o}\begin{bmatrix} I_{3\times 3} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \tag{99}$$

with the initial arbitrary covariance given by $P_o = \sigma_o^2 I_{4\times 4}$. For $\sigma_k^2 = \sigma^2$ independent of k, σ^2_{tot} will be equal to σ^2/N. Thus, for the full-vector measurement model, the ambiguous term coming from the unconstrained estimation will dominate the initial condition asymptotically. However, for the QUEST measurement model, which is a realistic model for star-cameras, it is the initial condition which will be perpetuated in every frame. The solution, of course, is to use the full-vector model for the direction measurements, since this will be justified once the norm constraint is restored. This means, of course, that the unconstrained quaternion estimate becomes even more unphysical and meaningless.

The LMS Survey Paper of the Attitude Kalman Filter

It is well to reexamine the LMS paper [17]. At a distance of two decades it is clear that the importance of that paper derives in large part from the fact that it came at an important cæsura in the development of Kalman filtering for spacecraft attitude, when that topic had reached maturity and the important points of implementation had been settled. The innovations of that paper were modest, largely in details of the implementation. The paper was submitted and published, in fact, as a survey paper. The importance of LMS lies in its timing and the care and clarity with which those authors presented the then state-of-the-art, which requires little modification today.

Among the numerous references cited in the historical survey of LMS, the most important early publication for LMS is that of Toda, Heiss and Schlee [18], which was presented at the Symposium on Attitude Determination, held at the Aerospace Corporation, Los Angeles, California in October 1969 (also an important milestone in the development of the attitude Kalman filter). That work first reached the junior authors of LMS via the work of Murrell [19]. Note that LMS presented three different implementations of the attitude Kalman filter: (1) a space-referenced implementation in which the redundant quaternion variable is removed by truncation (essentially the PADS system of TRW [20]); (2) a body-referenced implementation in which the redundant quaternion variable removed is the scalar component of the quaternion of an infinitesimal rotation; and (3) a hybrid approach which is very close to (2). All three implementations have been tested by Ferraresi [21] who found them very similar in performance. LMS also presented a filter in which the quaternion update is always $\Delta\bar{q}$ and the quaternion covariance matrix is 4×4, without explicit treatment of the constraint (which is maintained to first-order by the singularity of the covariance matrix), but only as a point of reference with strong warnings not to use it.[11] Nowadays, the LMS authors lean most toward the body-referenced implementation.

There are two pillars of the attitude Kalman filters of LMS. The first is the dynamics replacement model of Farrenkopf [20, 22]. In this model the gyro read-out noise is neglected and the gyro bias vector replaces the angular velocity vector in the state vector. In this way the other sources of gyro error become process noise

[11]Nonetheless, some readers have wrongly interpreted LMS as presenting a preferred approach for implementing a 4×4 quaternion covariance matrix.

in the filter propagation. The second is the treatment of the attitude correction and attitude error in the filter. In the propagation of the attitude the four-component quaternion is used, because the unit-norm is conserved by the form of the kinematic equations. The attitude error and error correction, however, are expressed by a three-dimensional representation, which is simply $\Delta\mathbf{q} = [\Delta q_1, \Delta q_2, \Delta q_3]^{\mathrm{T}}$ for the truncation method and $\delta\mathbf{q} = [\Delta\delta q_1, \Delta\delta q_2, \Delta\delta q_3]^{\mathrm{T}}$ for the body-referenced method.

The chief advantage of the three-dimensional treatment of quaternion error and correction is that it eliminates the need to maintain the singularity of the quaternion matrix. The chief advantage of working with a body-referenced quaternion is that the choice of which component to discard becomes obvious and the measurement sensitivity matrices take on their simplest form.

The technology of the extended Kalman filter for spacecraft attitude has been extremely stable since the appearance of LMS. One notable practical advance (though a minute theoretical advance) has been the QUEST filter [23, 24, 25], in which the star camera data is preprocessed using the QUEST algorithm [7] and the resulting quaternion used as an effective measurement of the attitude in the Kalman filter update. Since a star camera might observe as many as fifty stars in a single frame, this greatly decreases the computational burden as well making the filter better behaved numerically. This is the implementation of the attitude Kalman filter used for NASA's deep space missions, where accuracy, speed of execution, and reliability are all pressing concerns. Another notable practical advance has been the use of the square-root information filter (SRIF) [26] in the attitude Kalman filter [27] instead of the covariance filter employed by LMS. The great advantage of the SRIF is that it obviates the need to initialize the Kalman filter estimate with a large arbitrary covariance matrix, which generally corrupts the filter results for a considerable time and can also lead to divergence. Unfortunately, the SRIF has not received wide application in attitude estimation. Theoretical work since LMS has experimented with different attitude representations, wider applications to spacecraft attitude systems (especially gyroless systems), and more exotic filters.

Additive and Multiplicative Corrections

This writer senses the feeling among some workers that the additive and multiplicative implementations of the Kalman filter update express different but equally valid and not necessarily equivalent Kalman filter approaches. This is not true, the two approaches are exactly equivalent and should yield the same result within round-off error, as was demonstrated by Ferraresi [21]. The differences between the "additive" and "multiplicative" approaches is really only one of frame as pointed out a decade ago in reference [2]. Nonetheless, the terminology "additive" and "multiplicative" have become ingrained, for which this writer bears some responsibility. Hopefully, the true nature of the difference will become clear in the present section.

There are, as we indicated in Part I, two approaches to representing the attitude correction, which we may understand either as: (1) the "quotient" $\delta\bar{q}_k$ of the *a posteriori* and the *a priori* values of the quaternion or (2) the arithmetic "difference" $\Delta\bar{q}_k$ between these two quantities. Thus, in LMS we contrasted $\Delta\bar{q}_k$ and $\delta\bar{q}_k$. In the following we adopt a different point of view, namely that $\bar{q}_k$ and $\delta\bar{q}_k$ are the attitude state vectors, the first referenced to inertial axes, the second to predicted body axes. The update corrections are then $\Delta\bar{q}_k$ and $\Delta(\delta\bar{q}_k)$, both with *a priori* values of $\bar{0}$. The difference is subtle, but important for establishing the exact parallelism between the additive and multiplicative approaches, which we prefer to think of as inertially-referenced and predicted-body-referenced implementations. These two implementations correspond to the "truncated" and "body-referenced" implementation of LMS, Sections X and XI, respectively. We ask the reader to banish the phrase "multiplicative correction" from his or her mind.

These two implementations may be written as

$$\bar{q}_k(+) = \bar{q}_k(-) + \Delta\bar{q}_k(+), \qquad \delta\bar{q}_k(+) = \delta\bar{q}_k(-) + \Delta(\delta\bar{q}_k)(+) \tag{100ab}$$

Because of the norm constraint

$$\mathbf{q}_k^{\mathrm{T}}\mathbf{q}_k + q_{4,k}^2 = 1, \qquad (\delta\mathbf{q}_k)^{\mathrm{T}}(\delta\mathbf{q}_k) + (\delta q_{4,k})^2 = 1 \tag{101ab}$$

only $\Delta\mathbf{q}(+)$ and $\Delta(\delta\mathbf{q})(+)$ are computed directly from the data. The respective scalar update corrections are computed from the constraints and the vector corrections according to

$$\Delta q_{4,k}(+) = -\frac{1}{q_{4,k}(-)}\mathbf{q}_k^{\mathrm{T}}(-)\,\Delta\mathbf{q}_k(+) \tag{102a}$$

$$\Delta(\delta q_{4,k})(+) = -\frac{1}{\delta q_{4,k}(-)}\delta\mathbf{q}_k^{\mathrm{T}}(-)\,\Delta(\delta\mathbf{q}_k(+)) = 0 \tag{102b}$$

The vanishing of $\Delta(\delta q_4(+))$ follows from the vanishing of $\delta\mathbf{q}(-)$. The reader should not conclude from this that the predicted-body referenced correction is only three-dimensional while that for the inertially referenced correction is four-dimensional. Equivalent to the exact vanishing of $\Delta(\delta\bar{q}_4)(+) = \delta\bar{q}_k^{\mathrm{T}}(-)\,\Delta(\delta\bar{q}_k)(+)$ is the exact vanishing of $\bar{q}_k^{\mathrm{T}}(-)\,\Delta\bar{q}_k(+)$. Thus, it would be correct to say that the correction is four-dimensional but that $\Delta(\delta q_4)(+)$ happens to vanish. The difference, again, is only one of frame.

The similar natures of $\Delta\bar{q}_k$ and $\Delta(\delta\bar{q}_k)$ can be seen also through an examination of the computation of the Kalman filter update. For the two representations

$$\Delta\mathbf{q}_k(+) = \Delta\mathbf{q}_k(-) + K_k^I\nu_k^I = K_k^I\nu_k^I \tag{103a}$$

$$\Delta(\delta\mathbf{q}_k)(+) = \Delta(\delta\mathbf{q}_k)(-) + K_k^B\nu_k^B = K_k^B\nu_k^B \tag{103b}$$

Where the superscripts I and B denote "inertial" and "predicted body," respectively.

Likewise, for the measurement sensitivity matrices. If $H_k^*(\bar{q}) = [\mathbf{h}_k^{\mathrm{T}}(\bar{q}), h_{4,k}(\bar{q})]$ is the 4×1 constraint-insensitive measurement sensitivity matrix, then the 3×1 constraint-sensitive sensitivity matrix for the vector components of the quaternion for the space-referenced and predicted-body referenced update, respectively, is

$$H_{\bar{q},k}(\bar{q}_k(-)) = \mathbf{h}_k(\bar{q}_k(-)) - \frac{h_{4,k}(\bar{q}_k(-))}{q_{4,k}(-)}\mathbf{q}_k^{\mathrm{T}}(-) \tag{104a}$$

$$\begin{aligned} H_{\delta\bar{q},k}(\delta\bar{q}_k(-)) &= \mathbf{h}_k(\delta\bar{q}_k(-)) - \frac{h_{4,k}(\delta\bar{q}_k(-))}{\delta q_{4,k}(-)}\delta\mathbf{q}_k^{\mathrm{T}}(-) \\ &= \mathbf{h}_k(\delta\bar{q}_k(-)) \end{aligned} \tag{104b}$$

where again the body-referenced expression simplifies, because $\delta\mathbf{q}_k(-) = \mathbf{0}$.

The measurement model for a scalar measurement in the two procedures is

$$\mathsf{z}_k = \mathbf{u}_k^{\mathrm{T}}A(\bar{q}_k)\mathbf{v}_k^{\mathrm{I}} + \boldsymbol{\eta}_k = \mathbf{u}_k^{\mathrm{T}}A(\delta\bar{q}_k)\mathbf{v}_k^{\mathrm{B}} + \boldsymbol{\eta}_k \tag{105}$$

where $A(\bar{q})$ simply denotes the functional relationship of equation (I-43). The observation vectors $\mathbf{u}_k$ are always in the body frame.

The two updates are equivalent and *both are additive*! The only distinction is the frame.[12]

As a last step we must write for the predicted-body-referenced update procedure

$$\bar{q}_k(+) = \delta\bar{q}_k(+) \otimes \bar{q}_k(-) \tag{106}$$

The "multiplication," from this point of view, is really not part of the update but the transformation back to an inertially-referenced quaternion, because it only makes sense to accumulate the attitude as an inertially-referenced quaternion. Otherwise, the definition of the spacecraft attitude would change at every update. The computational burden of this operation is small compared to the greater computational burden of carrying out the update with respect to inertial axes.

To emphasize yet once more that the difference between the "additive" and "multiplicative" approaches differ only in the choice of frame we write

$$\Delta\bar{q}_k = \{\bar{q}_k(-)\}_R\,\Delta(\delta\bar{q}_k), \qquad \Delta(\delta\bar{q}_k) = \{\bar{q}_k(-)\}_R^{\mathrm{T}}\,\Delta\bar{q}_k \tag{107ab}$$

Note that both update procedures preserve the norm only to first order in the correction. Hence, to correct for second-order effects and round-off error it is necessary to divide by the quaternion norm, after implementing equation (100a) for the space referenced update and after equation (106) for the predicted body-referenced update.

Consequences of the Batch Estimation Results for the Unconstrained Quaternion Kalman Filter

The Kalman filter is also a maximum-likelihood estimator [28]. To see this let $\mathbf{x}_k(-)$ and $P_k(-)$ be the *a priori* state estimate and state estimate-error covariance matrix. Then the MLE (more exactly, MAP) dictated cost function for $\mathbf{x}_k$ is

$$J(\mathbf{x}_k) = \frac{1}{2}(\mathbf{x}_k(-) - \mathbf{x}_k)^{\mathrm{T}} P_k^{-1}(-)\,(\mathbf{x}_k(-) - \mathbf{x}_k) + \frac{1}{2}(\mathbf{z}_k - \mathbf{f}_k(\mathbf{x}))^{\mathrm{T}} R_k^{-1}(\mathbf{z}_k - \mathbf{f}_k(\mathbf{x})) \tag{108}$$

As a function of the correction $\Delta\mathbf{x}_k$ this becomes

$$J(\Delta\mathbf{x}_k) = \frac{1}{2}\Delta\mathbf{x}_k^{\mathrm{T}} P_k^{-1}(-)\,\Delta\mathbf{x}_k + \frac{1}{2}(\Delta\mathbf{z}_k - H_k\,\Delta\mathbf{x})^{\mathrm{T}} R_k^{-1}(\Delta\mathbf{z}_k - H_k\,\Delta\mathbf{x}) \tag{109}$$

with $\Delta\mathbf{x}_k(-) = 0$. Minimization of $J(\Delta\mathbf{x}_k)$ leads straightforwardly to

$$P_k^{-1}(+) = P_k^{-1}(-) + H_k^{\mathrm{T}} R_k^{-1} H_k, \qquad \Delta\mathbf{x}_k(+) = P_k(+) R_k^{-1}\,\Delta\mathbf{z}_k \tag{110ab}$$

This we recognize immediately as the information form of the Kalman filter, to be more exact, the extended Kalman filter (EKF), because the measurement model has been linearized. (See [17] for references.) The derivation of the covariance form of the filter from the information form is assumed to be known to readers.

[12]Note in passing that in terms of $\boldsymbol{\epsilon}_k(+)$ the four-dimensional update equations take the form $\Delta\bar{q}_k(+) = \Xi(\bar{q}_k(-))\boldsymbol{\epsilon}_k(+)/2$ and $\Delta(\delta\bar{q}_k)(+) = \Xi(\delta\bar{q}_k(-))\boldsymbol{\epsilon}_k(+)/2$. The second equation is equivalent to equation (100b) above. The first equation, however, is the update step of the remaining implementation of LMS, the reduced-order implementation (Section IX).

It has been assumed in deriving these last results that the components of the state vector $\mathbf{x}_k$ are truly independent variables (although not necessarily statistically independent) and that the estimate error covariance matrix P_k is full rank.

The measurement sensitivity matrix has been shown to be ambiguous for unconstrained quaternion estimation, which implies that unconstrained quaternion estimates must also be ambiguous, and, therefore, meaningless. How, then should one mechanize a Kalman filter for the quaternion when the norm constraint is abandoned?

The answer to this is simple. Since only the constrained quaternion estimate has meaning, we chose the quaternion measurement sensitivity matrix so as to make the problem of getting to the constrained quaternion estimate easiest. At the moment, that would seem to mean that we choose the relationship between the attitude matrix and the quaternion to be equation (I-43), especially since we know that equation (I-73) has nasty aftereffects.

Obviously, it should make no difference, apart from the presence of an initial condition, if the norm is restored after each update or after several updates. However, it is important that the norm restoration take proper account of the statistics of the unconstrained estimation that preceded it. Unfortunately, an initial-condition is needed for unconstrained quaternion estimation if we model the measurements as direction measurements, and this initial condition will persist forever in the filter, Note that process noise will not speed the decay of the quaternion norm defect, because the kinematic equations are norm preserving. Fortunately, we can model the direction measurements as full-vector measurements (see above) as long as we do not alter the state vector in any way before performing a correct constraint restoration.

An interesting point for unrepentant EKF "adders": it should be obvious now from our batch studies that the unconstrained quaternion EKF will be simplest if the prediction and update steps are carried out in the (predicted) body frame. Thus, for both correctly constrained and blithely unconstrained EKFs the first commandment is to multiply![13]

One consequence of this work is that the unit-vector filter [23, 29] can be used in general for unconstrained attitude estimates, because the quaternion with the constraint restored will be correct.

Comparison with Earlier and Current Practice (Batch and Filter)

The earliest work (1983) known to the author in which (1) an attitude constraint is relaxed, (2) the unconstrained estimate is a sufficient statistic [8] for the correctly constrained estimate, and (3) the constraint is restored via the method of Lagrange multipliers, treated not three-axis attitude estimation but spin-axis attitude estimation [30]. This work, a batch rather than a sequential estimator, has been implemented frequently in real missions and has performed well.

This same technique has been applied to quaternions in reference [2]. This latter work was not intended to develop a new algorithm for quaternion estimation but to show that the unconstrained quaternion Kalman filter of Bar-Itzhack and

[13] *Be fruitful and multiply* (Genesis [1:18]) is the first commandment in the Bible. However, the word for multiply there has the literal sense rather of "greaten" (root: רב), that is, to make greater or more numerous. The word (in both Mishnaic and Modern Hebrew) for multiply in the arithmetic sense derives from the root כפל, equally ancient (it has a biconsonantal precursor כף) and having the meaning *fold*. Compare Latin *multiplicare*, where *plicare* means to *fold* (note also English *ply*).

Oshman [3] (see below), if the constraint were restored correctly, would necessarily yield the same results as the LMS filters to within second-order in the attitude errors but with a significantly greater computational burden than the constrained and lower-dimensional LMS filters. This author's unstated purpose in reference [2], was to turn workers away altogether from an unconstrained quaternion filter which he regarded, and still regards, as suffering from insufficient rigor and presenting results which he found troubling.

Bar-Itzhack and Reiner [4] published a Kalman filter implementation for the unconstrained direction-cosine matrix (to which equations (4) through (8) bear very close resemblance). However, no prescription was given for the restoration for the constraint, which was left to a procedure to be determined later. Professor Bar-Itzhack's previous work on the orthogonalization of a matrix [31–35], while very interesting geometrically and correct for repairing the constraint defects due to round-off error, was not statistical in nature and would not have been appropriate for norm restoration to the unconstrained attitude matrix estimate.

The foundation paper of unconstrained quaternion estimation is the work of Bar-Itzhack and Oshman [3]. Unfortunately, the algorithm for restoring the quaternion norm is not based on the statistics of the unconstrained estimate. The *partial reset* method is an *ad hoc* method supported only by limited simulations. On the basis of the work presented here for batch estimation, the chief effect of partial resets would be to eliminate the possibility of restoring the quaternion norm in a statistically correct fashion. Though not stated in their publication, Bar-Itzhack and Oshman use the full-vector model for direction measurements. Hence, their filter is spared the problem of an eternally persistent initial condition, but at the cost of using a measurement model which is invalid for most attitude sensors.

A number of alternate methods have been proposed for the restoration of the quaternion norm to the filtered unconstrained quaternions [5, 6]. All of these methods are *ad hoc* and do not use all of the data but truncate it in arbitrary ways. The methods proposed in the present work, which use all of the data rigorously (except for the approximation that the unconstrained quaternion estimate is a sufficient statistic), are not among them.

Kasdin and Weaver [36] follow a different approach. They first implement a filter in which the attitude is parameterized by an unconstrained attitude matrix. The advantage of this is that the filter converges then in one step, as we saw in our batch example, which is not true for the quaternion. Kasdin and Weaver then refilter the results of the first filter using that fact that the unconstrained attitude matrix estimate is a sufficient statistic for the quaternion. The second filter is iterated until convergence with a Lagrange multiplier for the quaternion to insure convergence to unit norm. The expression "iterated until convergence" indicates that these authors are actually performing an iterative batch estimation at each step of the filter. Thus, their work is much closer to the methods presented here than to an extended Kalman filter. It would be interesting to know how well this two-step filter performed with only a single iteration of the second step.

A different approach to estimating the initial condition of a Kalman filter state is a two-step process [37] applied to misalignment estimation. Here the first filter generates a sequence of sufficient statistics for the initial condition and appropriate sensitivity matrices and covariances, while the second step uses these quantities to estimate the initial condition. In this last application, since one focused only on the alignment estimation and not of other parameters or state variables, the second filter step was replaced by a batch estimator. This approach assumes that the initial condition enters linearly into the filter state vector.

Markley has begun a careful study of the continuous unconstrained Kalman filter which includes process noise [38]. His studies should throw light on features of the unconstrained Kalman filter different from those explorable within the context of the present work.

Means, Covariances, and Constraints

We saw in equation (I-83) that the mean of the quaternion of rotation, and, therefore, the covariance matrix, are not physically meaningful, because the mean takes on values outside the domain of definition of the quaternion of rotation. Certainly, these statistics convey some information about approximate values and range of variation, but they have no value in constructing a probability density function unless we know that the random variable is Gaussian. Note also that the nonzero variance for q_4 and its lack of correlation with the other variables does not imply that it is an independent random variable, as is made clear by equation (I-79b). The attitude has only three degrees of freedom; therefore, there can be only three distinct attitude random variables, for which, in the present example, $\mathbf{q}$ is certainly the best choice. One could, of course, write a joint pdf for the four components of the quaternion as

$$p(\mathbf{q}, q_4) = p(\mathbf{q})\delta\left(q_4 - \sqrt{1 - |\mathbf{q}|^2}\right) \tag{111}$$

in which the δ-function is really a shorthand for equation (79b). There are only three random variables required to describe a random attitude or an attitude estimate.

In fact, since we normally linearize our measurement model in the EFK, higher-order terms, like the $(4, 4)$ element of the quaternion covariance matrix, have no significance. In fact, only a 3×3 attitude estimate-error covariance matrix has mathematical significance.[14] The singular 4×4 quaternion estimate-error covariance matrix of equation (I-84) is no more or less unsuitable than the nonsingular 4×4 covariance matrix of equation (I-83). All that is important is that both preserve the quaternion norm to first order in the Kalman filter update. (The unconstrained quaternion estimate-error covariance matrix, of course, should be a nonsingular 4×4 covariance matrix, but the unconstrained quaternion cannot be interpreted as a quaternion of rotation.)

The singular nature of the QUEST measurement model covariance is nothing more than the statement that a unit-vector measurement consists effectively (and truly) of only two measurements, as in equations (I-29). To write equations (I-92) and (I-97) and then use this model to construct a different covariance matrix R_k containing higher-order terms in σ_k^2 is to stretch the model too far. We can remove the singularity of the QUEST measurement model covariance matrix in practice,

[14]Even when the domain of the representation is the domain of a vector space, so that addition and multiplication by a scalar, key ingredients of computing an expectation, are defined, there are difficulties. Consider the modified Rodrigues parameters (MRPs), which can certainly be considered the domain of a Euclidean vector space on R^3. The two MRPs $[p, 0, 0]^T$ and $[-1/p, 0, 0]^T$, represent the same attitude. However, the mean MRP is $[p/2 - 1/2p, 0, 0]^T$, which is the representation of a different attitude.

but only because asymptotically, the measurement sensitivity matrix will annihilate the unphysical component. Thus, it is important to maintain the singularity of the 4×4 covariance matrix of the constrained quaternion estimate, not to satisfy some divine mathematical commandment, but to maintain the norm constraint of the quaternion.

Summary and Discussion

Despite being largely limited to batch attitude estimation, this work has covered a lot of ground. The reader will be spared the usual triumphant list of accomplishments. However, the author feels compelled to point out two important features of this work: (1) the very careful and detailed treatment of measurement sensitivity matrices in Part I, which is key to all of the results of Part II, and (2) the clarity that comes from carrying out estimation operations in the predicted body frame. More than 150 years ago we learned that rigid-body mechanics was best studied in the body frame. It is unfortunate that some workers are reluctant to learn this same lesson for attitude estimation.

The abandonment of the quaternion norm constraint is not synonymous with the additive EKF. The TRW PADS approach [20], which appears as the truncated filter in LMS and whose update operations were reviewed above in parallel with the "multiplicative" approach, is additive and takes account of the quaternion norm constraint. It is less convenient then the body-referenced methods of LMS but no less correct. Equally well, one could have relaxed the norm constraint on $\delta\bar{q}_k$ in the update step. But who would model a quantity which changes from unity during the update step by no more than $O(\sigma_k^2)$ as a *variable* when one is routinely discarding terms of $O(\sigma_k^2)$? A reasonable person simply calls this quantity 1 (and does not introduce a fourth free parameter into the attitude). Not surprisingly, one never hears of an unconstrained MEKF[15], except in this paper, where all of the examples are "multiplicative." However, the quantity $\bar{q}_k^{\mathrm{T}}(-)\,\Delta\bar{q}_k$ has the same limited variation, but is blithely made a free variable by the nest of "adders," perhaps, because the constraint is less obvious when seen from inertial axes.

Thus, once more and for the record, AEKF does not imply an unconstrained filter, and MEKF does not imply a constrained filter. From our studies here wise practice would seem to suggest that one implement a correctly constrained Kalman filter for both AEKF and MEKF.

Arguments that the four components of the quaternion correction are unconstrained rest either on the fact that the supposed covariance matrix is nonsingular or on the fact that $\Delta\bar{q}_k$ is not a quaternion of rotation. Both of these arguments have been disposed of in Part I of this work. However, the second argument is still intoned by "adders" [39].

An important criticism of the unconstrained quaternion filter work until recently is the lack of careful theoretical study. Generally, the only justification offered for abandoning the norm constraint has been simulation results. However, these too are dissatisfying. Generally, one hopes in the time-development of a Kalman filter simulation that one will first see the decay of an arbitrary initial condition (if there is any), and then the steady decrease as $1/N$ in the filter covariances as they approach their asymptotic limits. However, if the filter estimates

[15]In references [5] and [6], which address the norm in both the AEKF and the MEKF, the AEKF is the unconstrained filter, but the MEKF is constrained.

are truly converging to unit norm, then one should see the condition number of the 4×4 quaternion covariance matrix tend toward infinity with N. This does not seem to have been tested. Nor has the whiteness of the residuals been tested. Note that the fact that the unconstrained filter converges to the true quaternion asymptotically does not mean that the norm constraint has been achieved. A counterexample (when there is no process noise) would be the case that the attitude variances decrease as σ^2/N but the variance of the norm defect decreases as $10^2(\sigma^2/N)$. Of course, if σ^2/N were truly minuscule (for example, 10^{-14}), one might not care. However, in practice such a situation will not occur.

Finally, we must ask: what is the justification for an unconstrained quaternion estimator, either as a batch estimator or as an extended Kalman filter? There is obvious benefit to an initial unconstrained estimation of the attitude matrix, because this leads (1) to a more efficient arrangement of the data and (2) to making the estimation procedure globally convergent without the introduction of an arbitrary initial condition. There does not, however, seem to be any benefit to unconstrained quaternion estimation, since it is more burdensome than constrained quaternion estimation, and, when one has finished, one must still do the work of restoring the constraint, which, by itself, is at least as difficult as constrained quaternion estimation. The unconstrained AEKF is faster than the constrained AEKF only if one accepts the unconstrained quaternion estimate as a substitute for the constrained quaternion estimate. The lack of constraint for the quaternion also brings with it special ills including the occasional failure of the estimation process. To this writer it seems that unconstrained quaternion estimation brings only extra burdens and no benefits.

Acknowledgments

The author is grateful to Itzhack Bar-Itzhack, Yang Cheng, Richard Harman, F. Landis Markley, Yaakov Oshman, Mark Pittelkau, Mark Psiaki, and Julie Thienel (formerly Deutschmann) for many interesting discussions on this topic over the years. Dr. Markley is to be especially thanked for a critical reading of the manuscript and pointing out numerous errors of presentation.

References

[1] SHUSTER, M. D. "Constraint in Attitude Estimation, Part I: Constrained Estimation," *The Journal of the Astronautical Sciences,* Vol. 51, No. 1, pp. 51–74.

[2] SHUSTER, M. D. "The Quaternion in the Kalman Filter," presented as paper No. AAS 93–62, *Proceedings,* AAS/AIAA Astrodynamics Specialists Conference, Victoria, British Columbia, Canada, August 1993; printed as "The Quaternion in the Kalman Filter," *Advances in the Astronautical Sciences*, Vol. 85, pp. 25–37, 1993.

[3] BAR-ITZHACK, I. Y. and OSHMAN, J. "Attitude Determination from Vector Observations: Quaternion Observation," *IEEE Transactions on Automatic Control*, Vol. AES-21, No. 1, January 1985, pp. 128–136.

[4] BAR-ITZHACK, I. Y. and REINER, J. "Recursive Attitude Determination from Vector Observations: DCM Identification," *Journal of Guidance, Control and Dynamics*, Vol. 7, No. 1, January-February 1984, pp. 51–56.

[5] BAR-ITZHACK, I. Y., MARKLEY, F. L., and DEUTSCHMANN, J. "Quaternion Normalization in Additive EKF for Spacecraft Attitude Determination," presented at the Flight Mechanics/Estimation Theory Symposium, NASA Goddard Space Flight Center, May 1991.

[6] BAR-ITZHACK, I. Y., DEUTSCHMANN, J., and MARKLEY, F. L. "Quaternion Normalization in Additive EKF for Spacecraft Attitude Determination," *Proceedings, AIAA Guidance and Control Conference*, New Orleans, Louisiana, August 11–14, 1991.

[7] SHUSTER, M. D. and OH, S. D. "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance*, *Control and Dynamics*, Vol. 4, No. 1. Jan.–Feb. 1981, pp. 70–77.
[8] SORENSON, H. W. *Parameter Estimation*, Marcel Dekker, New York, 1980.
[9] WAHBA, G. "A Least-Squares Estimate of Satellite Attitude," Problem 65–1, *SIAM Review*, Vol. 7, No. 3, July 1965, p. 409.
[10] BROCK, J. E. "Optimal Matrices Describing Linear Systems," *AIAA Journal*, Vol. 2, No. 7, July 1968, pp. 1292–1296.
[11] BAR-ITZHACK, I. Y. "Polar Decomposition for Attitude Determination from Vector Measurements," *Proceedings*, *Flight Mechanics/Estimation Theory Symposium*, NASA Goddard Space Flight Center, Greenbelt, Maryland, May 1992, pp. 243–257.
[12] MARKLEY, F. L. and BAR-ITZHACK, I. Y. "Unconstrained Optimal Transformation Matrix," *IEEE Transaction on Aerospace and Electronic Systems*, Vol. 34, No. 1, January 1998, pp. 338–340.
[13] SHUSTER, M. D. "A Broad Look at Deterministic Three-Axis Attitude Determination," Paper No. AIAA 2000-4239, *Proceedings,* AIAA/AAS Astrodynamics, Conference, Denver, Colorado, August 2000.
[14] SHUSTER, M. D. "Deterministic Three-Axis Attitude Determination," to appear, *The Journal of the Astronautical Sciences.*
[15] SHUSTER, M. D. "Maximum Likelihood Estimation of Spacecraft Attitude," *The Journal of the Astronautical Sciences*, Vol. 37, No. 1, January–March, 1989, pp. 79–88.
[16] MARKLEY, F. L. and MORTARI, D. "Quaternion Attitude Estimation Using Vector Observations," *The Journal of the Astronautical Sciences*, Vol. 48, Nos. 2 and 3, April–September 2000, pp. 359–380.
[17] 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, September–October 1982, pp. 417–429.
[18] TODA, N. F., HEISS, J. L., and SCHLEE, F. H. "SPARS: The System, Algorithms, and Test Results," *Proceedings of the Symposium on Spacecraft Attitude Determination*, Aerospace Corporation Report No. TR-0066(5306)-12, Vol. 1, Sept.–October 1969, pp. 361–370.
[19] MURRELL, J. W. "Precise Attitude Determination for Multimission Spacecraft," AIAA Paper 78-1248, *Proceedings*, AIAA Guidance and Control Conference, Palo Alto, California, August 1978.
[20] IWENS, R. P. and FARRENKOPF, R. L. "Performance Evaluation of a Precision Attitude Determination Scheme," AIAA Paper No. 71-964, August 1971.
[21] FERRARESI, V. A. *Utilização conjunta de sensores inerciais e não-inerciais em determinação de attitude de satélites via filtro de Kalman*, Publicação No. INPE-4313-TDL/280, Instituto Nacional de Pesquisas Espaciais, São José dos Campos (SP), Brazil, August 1987.
[22] FARRENKOPF, R. L. "Analytic Steady-State Accuracy Solutions for Two Common Attitude Estimators, *Journal of Guidance and Control*, Vol. 1, July-Aug. 1978, pp. 282–284.
[23] SHUSTER, M. D. "Kalman Filtering of Spacecraft Attitude and the QUEST Model," *The Journal of the Astronautical Sciences* Vol. 38, No. 3, July–September, 1990, pp. 377–393.
[24] VAROTTO, E. C., ORLANDO, V., and LOPES, R. V. F. "Um procedimento para determinação da atitude de satélites artificiais utilizando técnicas de estimação ótima estática e diâmica," *Atas*, 6^0. *Congresso Brasileiro de Automática*, *Belo Horizonte*, *1986*, pp. 946–951.
[25] FISHER, H. L., SHUSTER, M. D., and STRIKWERDA, T. E. "Attitude Determination for the Star Tracker Mission," presented as paper No. AAS 89-365 at the AAS/AIAA Astrodynamics Conference, Stowe, Vermont, August 6–10, 1989; printed in *Advances in the Astronautical Sciences*, Vol. 71, 1990, pp. 138–150.
[26] BIERMAN, G. J. *Factorization Methods for Discrete Sequential Estimation*, Academic Press, New York, 1977.
[27] PSIAKI, M. L. "Square-Root Information Filtering and Fixed-Interval Smoothing with Singularities," *Automatica*, Vol. 35, 1999, pp. 1323–1331.
[28] RAUCH, H. E., TUNG, F., and STRIEBEL, C. T. "Maximum Likelihood Estimation of Linear Dynamics Systems," *AIAA Journal*, Vol. 3, No. 8, 1965, pp. 1445–1450.
[29] SEDLAK, J. and CHU, D. "Kalman Filter Estimation of Attitude and Gyro Bias with the QUEST Observation Model," presented as paper AAS-93-297 at the AAS/GSFC International Symposium on Space Flight Dynamics, NASA Goddard Space Flight Center, April 26-30, 1993.

[30] SHUSTER, M. D. "Efficient Algorithms for Spin-Axis Attitude Estimation," *The Journal of the Astronautical Sciences*, Vol. 31, No. 2, April–June 1983, pp. 237–249.

[31] BAR-ITZHACK, I. Y. and FEGLEY, K. A. "Orthogonalization Techniques of a Direction Cosine Matrix," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-5, September 1969, pp. 798–804.

[32] BAR-ITZHACK, I. Y. "Iterative Optimal Orthogonalization of the Strapdown Matrix," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-11, January 1975, pp. 30–37.

[33] BAR-ITZHACK, I. Y., MEYER, J., and FUHRMANN, P. A. "Strapdown Matrix Orthogonalization: the Dual Iterative Algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-12, January 1976. pp. 32–37.

[34] BAR-ITZHACK, I. Y. and MEYER, J. "On the Convergence of Iterative Orthogonaliztion Processes," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-12, March 1976, pp. 146–151.

[35] MEYER, J. and BAR-ITZHACK, I. Y. "Practical Comparison of Iterative Matrix Orthogonalization Algorithms," *IEEE Transactions on Aerrospace and Electronic Systems*, Vol. AES-13, May 1977, pp. 230–235.

[36] KASDIN, N. J. and WEAVER, T. J. M. "Recursive Satellite Attitude Estimation with the Two-Step Optimal Estimator," Paper AIAA 2002-4462, *Proceedings,* AIAA/AAS Astrodynamics Specialists Conference, Monterey, California, August 5–8, 2002.

[37] ZANARDI, M. C. and SHUSTER, M. D. "Batch and Filter Approaches to Spacecraft Sensor Alignment Estimation," *Proceedings*, 12th International Symposium on Spaceflight Dynamics, Darmstadt, Germany, June 1997.

[38] MARKLEY, F. L. "Attitude Estimation or Quaternion Estimation," to appear, *The Journal of the Astronautical Sciences*, Vol. 52, Nos. 1 and 2, January–June 2004.

[39] BAR-ITZHACK, I. Y. and THIENEL, J. K. "On the Singularity in the Estimation of the Quaternion-of-Rotation," presented as paper AIAA 2002-4831 at the AIAA/AAS Astrodynamics Specialists Conference, Monterey, California, August 5–8, 2002.