AAS/AIAA Astrodynamics Conference, Victoria, British Columbia, August 16–19, 1993
Advances in the Astronautical Sciences, Vol. 85, 1993, pp. 25–37

AAS 93-553

The Quaternion in Kalman Filtering

Malcolm D. Shuster*

Simple results are presented for the sensitivity matrix of general attitude measurements to the quaternion, which are both constrained and unconstrained with respect to the quaternion unit norm. It is shown that for unconstrained maximum likelihood estimation, optimally restoring the quaternion norm is equivalent to estimating only the three independent parameters.

INTRODUCTION

Generally, an attitude measurement consists of a collection of scalars, each of which can be interpreted as an inner product, or as a function of these inner products. This occurs because an attitude measurement is always dependent on a sensed direction and the inner products are simply the components of this direction with respect to a coordinate system, usually fixed in the spacecraft body if the direction is that of a vector external to the spacecraft.

Consider the following example. Let v be an arbitrary vector with representations $\mathbf{v}_{\mathcal{B}}$ and $\mathbf{v}_{\mathcal{I}}$ with respect to the body and inertial coordinate systems, respectively. If A, the attitude matrix, transforms representations from the inertial to the body system, then

$$\mathbf{v}_{\mathcal{B}} = A\,\mathbf{v}_{\mathcal{I}} \,. \tag{1}$$

For a magnetometer, the measurement consists ideally of the three components of the magnetic field in the magnetometer frame. Hence, if v is the magnetic field vector, we may write the vector measurement as

$$z = \begin{bmatrix} \hat{x}\cdot v \\ \hat{y}\cdot v \\ \hat{z}\cdot v \end{bmatrix} + \varepsilon = \begin{bmatrix} \mathbf{x}_B^T A \mathbf{v}_{\mathcal{I}} \\ \mathbf{y}_B^T A \mathbf{v}_{\mathcal{I}} \\ \mathbf{z}_B^T A \mathbf{v}_{\mathcal{I}} \end{bmatrix} + \varepsilon \,, \tag{2}$$

where $\hat{x}$, $\hat{y}$ and $\hat{z}$ are the three coordinate axes of the magnetometer, and ε is the measurement noise. For a focal-plane sensor, typified by vector Sun sensors and star trackers, the measurement

Typeset with different fonts than the original. Hence, line and page breaks may not match the original publication.

*Dr. Shuster is on the Senior Professional Staff, Space Department, The Johns Hopkins University Applied Physics Laboratory, Johns Hopkins Road, Laurel Maryland, 20723-5000.

takes the form of two scalar measurements

$$z_1 = \frac{\mathbf{x}_B^T A \mathbf{v}_{\mathcal{I}}}{\mathbf{z}_B^T A \mathbf{v}_{\mathcal{I}}} + \varepsilon_1 , \quad \text{and} \quad z_2 = \frac{\mathbf{y}_B^T A \mathbf{v}_{\mathcal{I}}}{\mathbf{z}_B^T A \mathbf{v}_{\mathcal{I}}} + \varepsilon_2 , \tag{3}$$

where now v is the unit vector of the sensed direction. Thus, in calculating a sensitivity matrix, we are led to examine scalar measurements the form

$$z = \mathbf{u}^T A \, \mathbf{v} + \varepsilon , \tag{4}$$

where $\mathbf{u}$ and $\mathbf{v}$ are two arbitrary 3×1 column vectors and A is the attitude matrix, a 3×3 orthogonal matrix. It might seem that Eqs. (3) are not compatible with this form. However, we will see that more general measurement models, such as that of Eqs. (3) can be accommodated, as far as the sensitivity matrix is concerned by a suitable definition of the column vector $\mathbf{u}$, as we demonstrate below.

Since the attitude matrix A depends on only three independent parameters, it follows that z is sensitive to only three independent parameters. One could choose these three independent parameters to be some set of the three Euler angles or the Rodrigues (or Gibbs) vector.[1,2] These, however, are not regular everywhere and are inconvenient for practical calculations because of the complicated functions which must be differentiated. We therefore choose a different parameterization. If we have previous information about the attitude, so that we know that A must be close to some *a priori* value $A(-)$, then we can write

$$A = (\delta A) \, A(-) = \exp\{ \, [[\Delta \boldsymbol{\xi} \,]] \, \} \, A(-) , \tag{5}$$

where $\Delta\boldsymbol{\xi}$ is the rotation vector of a very small rotation δA. Here, $[[\Delta\boldsymbol{\xi}\,]]$ is the 3×3 antisymmetric matrix

$$[[\Delta \boldsymbol{\xi} \,]] \equiv \begin{bmatrix} 0 & \Delta\xi_3 & -\Delta\xi_2 \\ -\Delta\xi_3 & 0 & \Delta\xi_1 \\ \Delta\xi_2 & -\Delta\xi_1 & 0 \end{bmatrix} , \tag{6}$$

and $\exp\{\,\cdot\,\}$ is the matrix exponential function. Euler's formula in terms of $\Delta\boldsymbol{\xi}$ becomes

$$\delta A = \exp\{ \, [[\Delta \boldsymbol{\xi} \,]] \, \} = \cos |\Delta \boldsymbol{\xi}| \, I_{3\times 3} + \frac{1 - \cos |\Delta \boldsymbol{\xi}|}{|\Delta \boldsymbol{\xi}|^2} \, \Delta\boldsymbol{\xi} \Delta\boldsymbol{\xi}^T + \frac{\sin |\Delta \boldsymbol{\xi}|}{|\Delta \boldsymbol{\xi}|} \, [[\, \Delta \boldsymbol{\xi} \,]] , \tag{7}$$

which, because $\Delta\boldsymbol{\xi}$ is very small, we may write approximately as

$$\delta A = I_{3\times 3} + [[\, \Delta \boldsymbol{\xi} \,]] , \tag{8}$$

where the matrix $I_{3\times 3}$ is the 3×3 identity matrix. Note that the computation of δA will be simpler if it is defined not as a small rotation vector as in Eq. (5) but as twice the Rodrigues-Gibbs vector.[1,2] Equation (8) will still hold.

Substituting Eqs. (5) and (8) into Eq. (4) leads to

$$z = \mathbf{u}^T A(-) \, \mathbf{v} + \mathbf{u}^T [[\, \Delta \boldsymbol{\xi} \,]] \, A(-) \, \mathbf{v} + \varepsilon \tag{9a}$$

$$= \mathbf{u}^T A(-) \, \mathbf{v} + \big(\mathbf{u} \times A(-) \mathbf{v} \big)^T \Delta \boldsymbol{\xi} + \varepsilon \tag{9b}$$

$$= z_o + H_\xi \big(A(-) \big) \Delta \boldsymbol{\xi} + \varepsilon , \tag{9c}$$

and we have used the fact that

$$[[\mathbf{a}]]\,\mathbf{b} = -\mathbf{a} \times \mathbf{b} = \mathbf{b} \times \mathbf{a} = -[[\mathbf{b}]]\,\mathbf{a}\,. \tag{10}$$

Thus, the sensitivity matrix $H_\xi(A(-))$ is given by

$$H_\xi\big(A(-)\big) = \big(\mathbf{u} \times A(-)\mathbf{v}\big)^T\,. \tag{11}$$

Searches for an optimal A which best satisfies a set of measurement equations of the form given by Eq. (4) are best carried out by successively estimating $\Delta\boldsymbol{\xi}$ and updating A using Eq. (5) to obtain $A(-)$ for the next iteration. Generally, some non-optimal method supplies the initial value of $A(-)$.

If the measurement is an arbitrary scalar function of the representation of the measured vector in body coordinates, then we may write successively

$$\begin{aligned} z &= f(A\mathbf{v}) + \varepsilon && (12a)\\ &= f\big(A(-)\mathbf{v} + [[\,\Delta\boldsymbol{\xi}\,]]\,A(-)\mathbf{v}\big) + \varepsilon && (12b)\\ &= f\big(A(-)\mathbf{v} - [[\,A(-)\mathbf{v}\,]]\,\Delta\boldsymbol{\xi}\big) + \varepsilon && (12c)\\ &= f\big(A(-)\mathbf{v}\big) - \big[\big(A(-)\mathbf{v}\big) \times \nabla f\big(A(-)\mathbf{v}\big)\big]^T\,\Delta\boldsymbol{\xi} + \varepsilon && (12d)\\ &= f\big(A(-)\mathbf{v}\big) + H_\xi\,\Delta\boldsymbol{\xi} + \varepsilon && (12e) \end{aligned}$$

which amounts to replacing $\mathbf{u}$ by $\nabla f(A(-)\mathbf{v})$ in the earlier equations. This allows us to accommodate focal-plane measurements as given by Eqs. (3) within our model.

For the special case that $\mathbf{z}$ is a 3×1 matrix of the three components of $A\mathbf{v}$ in the body frame, then

$$\begin{aligned} \mathbf{z} &= A\mathbf{v} + \boldsymbol{\varepsilon} && (13a)\\ &= A(-)\mathbf{v} - [[\,A(-)\mathbf{v}\,]]\,\Delta\boldsymbol{\xi} + \boldsymbol{\varepsilon}\,, && (13b) \end{aligned}$$

a form which has been used to advantage in other studies.[3–5]

Likewise, Eq. (5) can be expressed in terms of the quaternion as

$$\begin{aligned} \bar{q} = \delta\bar{q} \otimes \bar{q}(-) &= \begin{bmatrix} \dfrac{\sin(|\Delta\boldsymbol{\xi}|/2)}{|\Delta\boldsymbol{\xi}|/2}\,\dfrac{\Delta\boldsymbol{\xi}}{2} \\ \cos(|\Delta\boldsymbol{\xi}|/2) \end{bmatrix} \otimes \bar{q}(-) && (14a)\\ &\approx \begin{bmatrix} \Delta\boldsymbol{\xi}/2 \\ 1 \end{bmatrix} \otimes \bar{q}(-)\,, && (14b) \end{aligned}$$

which is the parameterization for the Kalman filter update advocated by Lefferts et al.[6]

Not all workers follow this approach. Bar-Itzhack,[7] for example, prefers to write the correction of the direction-cosine matrix in the form

$$A = A(-) + \Delta A\,. \tag{15}$$

and the quaternion correction in the form[8]

$$\bar{q} = \bar{q}(-) + \Delta\bar{q}\,. \tag{16}$$

The purpose of this paper is to develop sensitivity matrices corresponding to both Eq. (14) and Eq. (16) and to discuss their implementation in estimation problems.

Estimation strategies using the additive corrections suffer from several drawbacks. Firstly, some of the elements of the quaternion and the direction-cosine matrix are constrained. Hence, some combinations of these elements must have exact values and, therefore, the associated covariance matrix of the full quaternion or direction-cosine matrix is singular. This is a serious drawback, since most practical applications require that these matrices be invertible. The practice[7,8] has been effectively to replace the zero eigenvalue of the covariance matrix by an infinite value and trusting that the constrained combinations, because they are not observable, will not mix into the estimate of the physically meaningful combinations of elements. Several numerical experiments have been carried out of these *ad hoc* approaches[9,10], which seem to yield reasonable results. A careful and convincing theoretical analysis, however, remains to be performed.

Though this work does not advocate procedures for estimating the quaternion without taking account of the unit-norm constraint, it does present forms for the sensitivity matrix which are much more transparent and more compact than those presented in Ref. 8. At the same time, by exploiting the result of Eqs. (11) and (13) it is possible to arrive at alternate forms for the sensitivity matrix which are norm-preserving within the linear approximation and numerical round-off.

The problem of all of these unconstrained correction algorithms is that they have no mathematical justification for neglecting the constraint. We must therefore ask the question: under what circumstances can we understand these unconstrained corrections rigorously and in a rigorous fashion impose the constraint on them? The answer is to recognize that the unconstrained correction is a sufficient statistic for the estimate of the properly normalized quaternion and then to use this sufficient statistic as an effective measurement for the properly constrained quaternion. It turns out that this rigorous approach yields the same result as a correction which acts only on the three independent parameters of the attitude correction. Thus, a result of this work is that the *ad hoc* schemes which do not address the unit norm of the attitude quaternion at every step, should simply be discarded.

In the present work, we first develop simple expressions for the sensitivity of scalar and vector measurements to the quaternion which ignore the norm constraint. We compare these with earlier expressions which take account of the quaternion norm constraint to first order. We then show the connection of the multiplicative and additive schemes for quaternion correction and demonstrate that these are distinguished more by the treatment of the norm constraint than by the multiplicative or additive nature of the correction. Finally, we show how the norm constraint may be restored to the quaternion in a rigorous fashion if it has been ignored in the construction of the quaternion correction.

QUATERNION MEASUREMENT SENSITIVITY MATRIX

To develop a measurement sensitivity matrix in terms of the quaternion, we write

$$z = \mathbf{u}^T A(\bar{q}) \mathbf{v} + \varepsilon \,, \tag{17}$$

with now

$$A(\bar{q}) = (q_4^2 - |\mathbf{q}|^2)\, I_{3\times 3} + 2\mathbf{q}\mathbf{q}^T + 2q_4\, [[\,\mathbf{q}\,]] \,, \tag{18}$$

and

$$\bar{q} = \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix} . \tag{19}$$

Then

$$H_q\big(\bar{q}(-)\big) \equiv \left.\frac{\partial z}{\partial \bar{q}}\right|_{\bar{q}(-)} = \mathbf{u}^T \left.\frac{\partial A(\bar{q})}{\partial \bar{q}}\right|_{\bar{q}(-)} \mathbf{v} \,. \tag{20}$$

Explicit differentiation leads to

$$\begin{aligned} H_q(\bar{q})\Delta\bar{q} = 2\,\big[(\mathbf{u}\cdot\mathbf{v})(q_4\Delta q_4 - \mathbf{q}^T\Delta\mathbf{q} + \mathbf{q}^T(\mathbf{u}\mathbf{v}^T + \mathbf{v}\mathbf{u}^T)\Delta\mathbf{q} \\ + \,(\mathbf{u}\times\mathbf{v})\cdot\mathbf{q}\Delta q_4 + q_4\,(\mathbf{u}\times\mathbf{v})^T\Delta\mathbf{q}\big] \,, \end{aligned} \tag{21}$$

where for convenience, we have discarded the designation $(-)$ of $\bar{q}(-)$ for the moment. If we write now

$$H_q(\bar{q}) = \big[\, \mathbf{h}^T(\bar{q}) \mid h_4(\bar{q}) \,\big] \equiv \bar{h}^T(\bar{q}) \,, \tag{22}$$

then

$$\bar{h}(\bar{q}) = 2 \begin{bmatrix} -(\mathbf{u}\cdot\mathbf{v})\mathbf{q} + (\mathbf{u}\mathbf{v}^T + \mathbf{v}\mathbf{u})\mathbf{q} + q_4(\mathbf{u}\times\mathbf{v}) \\ (\mathbf{u}\cdot\mathbf{v})q_4 + (\mathbf{u}\times\mathbf{v})\cdot\mathbf{q} \end{bmatrix} = -2\, M(\mathbf{u},\, \mathbf{v})\, \bar{q} \,. \tag{23}$$

The matrix M is symmetric and traceless and can be factored as

$$M(\mathbf{u},\, \mathbf{v}) = \begin{bmatrix} [[\,\mathbf{u}\,]] & \mathbf{u} \\ -\mathbf{u}^T & 0 \end{bmatrix} \begin{bmatrix} -[[\,\mathbf{v}\,]] & \mathbf{v} \\ -\mathbf{v}^T & 0 \end{bmatrix} . \tag{24}$$

If we define now

$$\bar{\mathbf{u}} = \begin{bmatrix} \mathbf{u} \\ 0 \end{bmatrix} \quad \text{and} \quad \bar{\mathbf{v}} = \begin{bmatrix} \mathbf{v} \\ 0 \end{bmatrix} , \tag{25}$$

then we can write

$$M(\mathbf{u},\, \mathbf{v}) = \{\, \bar{\mathbf{u}} \,\}_L \, \{\, \bar{\mathbf{v}} \,\}_R \,. \tag{26}$$

where $\{\,\cdot\,\}_L$ and $\{\,\cdot\,\}_R$ are the matrix representations of quaternion multiplication[2]

$$\bar{p} \otimes \bar{q} = \{\, \bar{p} \,\}_L \, \bar{q} = \{\, \bar{q} \,\}_R \, \bar{p} \,. \tag{27}$$

Combining Eqs. (22) through (27) leads finally to

$$H_q\big(\bar{q}(-)\big) = -2\,[\, \bar{\mathbf{u}} \otimes \bar{q}(-) \otimes \bar{\mathbf{v}} \,]^T \,. \tag{28}$$

This result may be obtained equally easily by noting that

$$\{\,\bar{q}\,\}_L\,\{\,\tilde{\bar{q}}\,\}_R^T = \begin{bmatrix} A(\bar{q}) & 0 \\ 0 & 1 \end{bmatrix}, \tag{29}$$

and differentiating

$$z = \bar{\mathbf{u}}^T\{\,\bar{q}\,\}_L\,\{\,\tilde{\bar{q}}\,\}_R^T\,\bar{\mathbf{v}} + \varepsilon \tag{30}$$

For the case of a complete vector measurement, we have

$$\mathbf{z} = A(\bar{q})\mathbf{v} + \boldsymbol{\varepsilon}\,. \tag{31}$$

we may regard this as a 3×1 matrix of measurements of the form given by Eq. (4). Thus,

$$H_q\big(\bar{q}(-)\big) = -2\begin{bmatrix} \big(\bar{\mathbf{1}} \otimes \bar{q}(-) \otimes \bar{\mathbf{v}}\big)^T \\ \big(\bar{\mathbf{2}} \otimes \bar{q}(-) \otimes \bar{\mathbf{v}}\big)^T \\ \big(\bar{\mathbf{3}} \otimes \bar{q}(-) \otimes \bar{\mathbf{v}}\big)^T \end{bmatrix} \tag{32a}$$

$$= -2\,\big[\,\bar{\mathbf{1}} \otimes \bar{q}(-) \otimes \bar{\mathbf{v}} \mid \bar{\mathbf{2}} \otimes \bar{q}(-) \otimes \bar{\mathbf{v}} \mid \bar{\mathbf{3}} \otimes \bar{q}(-) \otimes \bar{\mathbf{v}}\,\big]^T, \tag{32b}$$

where

$$\mathbf{1} \equiv \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{2} \equiv \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{3} \equiv \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{33}$$

It then follows that

$$H^T\big(\bar{q}(-)\big) = -2\,\{\,\bar{q}(-) \times \bar{\mathbf{v}}\,\}_R\,[\,\bar{\mathbf{1}} \mid \bar{\mathbf{2}} \mid \bar{\mathbf{3}}\,] \tag{34a}$$

$$= -2\,\{\,\bar{q}(-) \times \bar{\mathbf{v}}\,\}_R \begin{bmatrix} I_{3\times3} \\ \mathbf{0}^T \end{bmatrix} \tag{34b}$$

$$= -2\,\Xi\big(\bar{q}(-) \otimes \bar{\mathbf{v}}\big)\,. \tag{34c}$$

The matrix $\Xi(\bar{q})$, given by

$$\Xi(\bar{q}) = \begin{bmatrix} q_4 I_{3\times3} - [[\,\mathbf{q}\,]] \\ -\mathbf{q}^T \end{bmatrix}, \tag{35}$$

is familiar from the kinematic equation for the quaternion,[2,6]

$$\frac{d}{dt}\,\bar{q} = \frac{1}{2}\,\Xi(\bar{q})\,\boldsymbol{\omega}\,, \tag{36}$$

where $\boldsymbol{\omega}$ is the body-referenced angular velocity. Thus,

$$H_q\big(\bar{q}(-)\big) = -2\,\Xi^T\big(\bar{q}(-) \otimes \bar{\mathbf{v}}\big)\,. \tag{37}$$

The results of Eqs. (28) and (37) may be compared with the less compact (but equally correct) expressions in Ref. 8. For a more general measurement model, it suffices to replace $\mathbf{u}$ by $\nabla f(A(-)\mathbf{v})$ as before.

A SENSITIVITY MATRIX WITH NORM CONSTRAINT

Equation (36) provides the first step of the proper path for obtaining a sensitivity matrix which embodies the norm constraint. The change in the quaternion is related to the change in the error vector $\Delta\boldsymbol{\xi}$ in a manner similar to the quaternion kinematic equation according to

$$\Delta\bar{q} = \frac{1}{2}\,\Xi(\bar{q})\,\Delta\boldsymbol{\xi}\,. \tag{38a}$$

Solving for $\Delta\boldsymbol{\xi}$,

$$\Delta\boldsymbol{\xi} = 2\,\Xi^T(\bar{q})\,\Delta\bar{q}\,. \tag{38b}$$

Combining this with Eq. (9b) yields

$$z = \mathbf{u}^T A(-)\,\mathbf{v} + 2\left(\mathbf{u}\times A(-)\mathbf{v}\right)^T\,\Xi^T\!\left(\bar{q}(-)\right)\Delta\bar{q} + \varepsilon\,. \tag{39}$$

Hence, the sensitivity matrix for a scalar measurement must be[6]

$$H_q^{\text{constrained}}\left(\bar{q}(-)\right) = 2\left(\mathbf{u}\times A(-)\mathbf{v}\right)^T\,\Xi^T\!\left(\bar{q}(-)\right) \tag{40a}$$

$$= 2\left[\Xi\left(\bar{q}(-)\right)\left(\mathbf{u}\times A(-)\mathbf{v}\right)\right]^T\,. \tag{40b}$$

As an alternative approach to including the constraint, straightforward differentiation of

$$z' = (\bar{q}^T\bar{q})^{-1}\mathbf{u}^T A(\bar{q})\mathbf{v} + \varepsilon\,, \tag{41}$$

however, where the quaternion norm is maintained explicitly, leads to

$$H_q^{\text{constrained}}\left(\bar{q}(-)\right) = 2[\,\bar{\mathbf{u}}\otimes\bar{q}(-)\otimes\bar{\mathbf{v}}\,]^T[\,I_{4\times4} - \bar{q}(-)\bar{q}(-)^T\,]\,, \tag{42}$$

and it has been assumed that $\bar{q}^T(-)\bar{q}(-) = 1$. The equivalence of Eqs. (40) and (42) is not obvious.

Such a sensitivity matrix maintains the unit-norm constraint to first order in $\Delta\bar{q}$ because of the structure of $\Xi(\bar{q})$. Note that use of this sensitivity matrix does not guarantee the preservation of quaternion normalization due to the inadequacy of the linearization which motivates the sensitivity matrix and the effect of numerical round-off error. It does, however, reduce the growth of the norm errors.

A similar result can be used to obtain a sensitivity matrix in terms of the Euler angles. If the body-referenced angular velocity is related to the Euler angle rates by[2]

$$\boldsymbol{\omega} = M(\varphi,\,\vartheta,\,\psi)\begin{bmatrix}\dot{\varphi}\\ \dot{\vartheta}\\ \dot{\psi}\end{bmatrix} \equiv M(\boldsymbol{\phi})\,\frac{d}{dt}\,\boldsymbol{\phi}\,, \tag{43}$$

then by trivial inspection, the sensitivity matrix to changes in the Euler angles for a scalar measurement is

$$H_\phi\left(\boldsymbol{\phi}(-)\right) = \left(\mathbf{u}\times A(-)\mathbf{v}\right)^T M\left(\boldsymbol{\phi}(-)\right)\,. \tag{44}$$

Simple expressions[2,11] exist for the matrix $M(\phi)$. The expression of Eq. (44) may be compared to an equivalent formula in Ref. 12.

The sensitivity to the multiplicative correction of Eq. (14) is, in fact, within a factor of 2 the same as that for $\Delta\boldsymbol{\xi}$ developed earlier. We need only note that

$$\delta\mathbf{q} = \frac{1}{2}\,\Delta\boldsymbol{\xi}\,, \tag{45}$$

where again $\delta\mathbf{q}$ is the vectorial part of $\delta\bar{q}$.

ADDITIVE AND MULTIPLICATIVE CORRECTIONS

Much effort has been expended on contrasting the additive and the multiplicative correction of a quaternion.[9,10] This distinction is artificial and misleading, as we shall now show.

Let us write the relation between the updated and predicted quaternions as

$$\bar{q}(+) = \bar{q}(-) + \Delta\bar{q}(+)\,, \tag{46}$$

the so-called additive approach. The components are all resolved with respect to inertial axes. Let us examine the same equation expressed with respect to the predicted spacecraft body frame, i.e., we express all rotations as rotations from the predicted spacecraft body frame. Denoting the quaternions of rotation with respect to this frame by $\bar{q}'$ where

$$\bar{q}' = \bar{q} \otimes \bar{q}^{-1}(-)\,, \tag{47}$$

it follows that

$$\bar{q}'(+) = \bar{1} + \Delta\bar{q}'(+)\,, \tag{48}$$

where $\bar{1} \equiv [0 \quad 0 \quad 0 \quad 1]^T$ is the identity quaternion. If we write now

$$\delta\bar{q}(+) \equiv \bar{q}'(+)\,, \tag{49}$$

then it follows that

$$\bar{q}(+) = \delta\bar{q}(+) \otimes \bar{q}(-)\,, \tag{50}$$

which is the so-called multiplicative correction, which we have computed "additively." Thus, the distinction between the additive and the multiplicative formulations of the Kalman filter is not one of the fundamental mechanization of the filter but simply of the frame in which it is desired to compute the update. These two formulations are both present in Ref. 6, and have been studied numerically by Ferraresi.[13]

Where the important distinctions do lie is in how $\Delta\bar{q}$ or $\Delta\bar{q}'$ is calculated, that is, whether it has three or four independent parameters, and, consequently, whether $\delta\bar{q}(+)$ has unit norm. Thus, one should speak more correctly of a "three-dimensional" and a "four-dimensional" update. The misleading nomenclature "additive" and "multiplicative" correction seems to be ingrained, however. Note that if the "additive" correction is done in a manner consistent with the true degrees of freedom, then it follows that[6]

$$\Delta\bar{q} = \Xi\big(\bar{q}(-)\big)\,\delta\mathbf{q}\,, \tag{51}$$

where $\delta\mathbf{q}$ is the vectorial part of $\delta\bar{q}$.

From the earlier discussion it is clear that a correct approach is obtained by expressing this quantity in terms of some representation of the attitude of minimal degree. In this case it is clearly advantageous to work from the spacecraft body frame so that this minimal-dimensional representation will be far from a singularity, and it will be most revealing to compare the results of Ref. 6 and Refs. 9 and 10 in that frame. The results of Refs. 9 and 10, however, are not directly comparable to Ref. 6 because the former rest on the attitude Kalman filter of Bar-Itzhack and Oshman.[8] However, many points of commonality will be apparent.

RESTORATION OF QUATERNION NORMALIZATION

Consider now the estimation of a constant quaternion from scalar measurements of the form given by Eq. (4). We wish to compute the attitude estimate from these measurements, using an approximate estimate of the attitude as a point of departure. Thus, we measure the quaternion from the *a priori* estimate of the body frame, so that $\bar{q}(-) = \bar{1}$. The treatment here follows a schematic treatment done earlier for the restricted case of a two-dimensional world.[14]

The measurement model leads to a residual of the form

$$\nu_k \equiv z_k - z_{o,k} = H_k\,\Delta\bar{q} + \boldsymbol{\varepsilon}_k\,, \quad k = 1, \ldots, N\,, \tag{52}$$

with H_k given as above. Our discussions will be true also for batch estimation, and holds for both scalar and vector measurements with the appropriate modification of the column dimensions of z_k, ν_k and H_k. For simplicity of notation, we shall examine batch estimation. For linear Gaussian measurements, the Kalman filter is just a sequential mechanization of the batch estimator.

The maximum likelihood estimate of $\Delta\bar{q}$ (for the additive quaternion correction, which is not constrained to preserve the norm) is given by

$$\Delta\bar{q}^*_{\rm add} = P_{qq}\,\bar{p}_{\rm add} = \begin{bmatrix} \Delta\mathbf{q}_{\rm add} \\ \Delta q_{\rm add} \end{bmatrix}, \tag{53}$$

where the 4×4 covariance matrix, P_{qq}, and the information quaternion, $\bar{p}$, are given by

$$P_{qq} = \left[\sum_{k=1}^{N} H_k^T R_k^{-1} H_k\right]^{-1}, \quad \bar{p} = \sum_{k=1}^{N} H_k^T R_k^{-1} \nu_k\,. \tag{54}$$

and R_k is the covariance matrix of the measurement noise. As stated above, we will assume that all representations are defined with respect to the *a priori* body frame, so that, effectively, $\bar{q}(-) = \bar{1}$.

For the "multiplicative" correction (which is norm-preserving) the estimate for the same data is (note that we estimate only the vectorial coordinates and determine the scalar component from the norm condition)

$$\Delta\mathbf{q}^*_{\rm mult} = P_{\rm mult}\,\mathbf{p}_{\rm mult}\,, \tag{55}$$

with

$$P_{\text{mult}} = \left[\sum_{k=1}^{N} H_{1,k}^T \, R_k^{-1} H_{1,k} \right]^{-1} = \left(P_{qq}^{-1}\right)_{11}^{-1} , \tag{56a}$$

$$\mathbf{p}_{\text{mult}} = \sum_{k=1}^{N} H_{1,k}^T \, R_k^{-1} \nu_k \, , \tag{56b}$$

and we have written

$$H_k = \left[\, H_{1,k} \,|\, H_{2,k} \, \right] , \tag{57}$$

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 remaining column. It follows that

$$\mathbf{p}_{\text{mult}} = \mathbf{p}_{\text{add}} \, , \tag{58}$$

where $\mathbf{p}_{\text{add}}$ denotes the vectorial components of $\bar{p}_{\text{add}}$. Hence, we can find a relation between the "additive" and the "multiplicative" corrections to the quaternion by solving Eq. (53) for $\bar{p}_{\text{add}}$ in terms of $\Delta\bar{q}^*_{\text{add}}$ and using the value of $\mathbf{p}_{\text{add}}$ from this expression in Eq. (56). This leads to

$$\Delta\mathbf{q}^*_{\text{mult}} = \Delta\mathbf{q}^*_{\text{add}} + \left(P_{qq}^{-1}\right)_{11}^{-1} \left(P_{qq}^{-1}\right)_{12} \Delta q^*_{4\,\text{add}} \, , \tag{59}$$

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

$$P_{qq} = \begin{bmatrix} (P_{qq})_{11} & (P_{qq})_{12} \\ (P_{qq})_{21} & (P_{qq})_{22} \end{bmatrix} \quad \text{and} \quad P_{qq}^{-1} = \begin{bmatrix} (P_{qq}^{-1})_{11} & (P_{qq}^{-1})_{12} \\ (P_{qq}^{-1})_{21} & (P_{qq}^{-1})_{22} \end{bmatrix} . \tag{60}$$

We will return to this equation soon.

The additive correction $\Delta\bar{q}^*_{\text{add}}$ allows us to construct an optimal quaternion $\bar{q}^*_{\text{add}}$,

$$\bar{q}^*_{\text{add}} = \bar{1} + \Delta\bar{q}^*_{\text{add}} \, . \tag{72}$$

Because it does not necessarily have unit norm, $\bar{q}^*_{\text{add}}$ does not without further effort have an unambiguous connection to the attitude. However, we note that although $\bar{q}^*_{\text{add}}$ is not a "quaternion of rotation," for an assumed linear Gaussian measurement model, it is nonetheless a sufficient statistic[15] for the attitude quaternion, certainly within the linear approximation of Eq. (52). It is, in fact, an estimate of the quaternion of rotation, and we know also that were the measurement noise covariance to vanish (perfect measurements), $\bar{q}^*_{\text{add}}$ would have unit norm and be the desired quaternion. Thus, denoting the desired quaternion of rotation by $\bar{\eta}$, ($\bar{\eta}$ always has unit norm) we have that

$$\bar{q}^*_{\text{add}} = \bar{\eta} + \Delta\bar{\eta}_{\text{add}} \, , \tag{62}$$

and

$$\Delta\bar{\eta}_{\text{add}} \sim \mathcal{N}(\bar{0}, P_{qq}) \, , \tag{63}$$

with P_{qq} given by Eq. (54). Hence, the negative log-likelihood function of $\bar{q}^*_{\text{add}}$ given $\bar{\eta}$ is

$$J(\bar{q}^*_{\text{add}} \mid \bar{\eta}) = \frac{1}{2} \left[(\bar{q}^*_{\text{add}} - \bar{\eta})^T P_{qq}^{-1} \, (\bar{q}^*_{\text{add}} - \bar{\eta}) + \log \det P_{qq} + 4 \log 2\pi \right] . \tag{64}$$

and the maximum likelihood estimate of $\bar{\eta}$ is simply

$$\bar{\eta}^* = \arg \max_{\bar{\eta}\,:\,\bar{\eta}^T\bar{\eta}=1} J(\bar{q}^*_{\text{add}} \mid \bar{\eta}) , \tag{65}$$

where, since we know that the true quaternion must lie on the manifold of unit four-vectors, we must maximize the negative log-likelihood subject to the norm constraint.

We handle the constraint in the usual way, using Lagrange's method of multipliers, optimizing the quantity

$$J(\bar{q}^*_{\text{add}} \mid \bar{\eta}) + \frac{1}{2}\,\lambda \bar{\eta}^T \bar{\eta}$$

without constraint, and then choosing the value of the Lagrange multiplier, λ, for which the constraint is satisfied. Differentiating the above expression with respect to $\bar{\eta}$ and setting the derivative equal to zero leads to

$$\bar{\eta}^* = \left(I + \lambda P_{qq}\right)^{-1} \bar{q}^*_{\text{add}} , \tag{66}$$

and λ is a solution of

$$f(\lambda) \equiv \bar{\eta}^{*\,T}(\lambda)\,\bar{\eta}(\lambda) = \bar{q}^{*\,T}_{\text{add}} \left(I + \lambda P_{qq}\right)^{-2} \bar{q}^*_{\text{add}} = 1 . \tag{67}$$

We expect λP_{qq} to be small. Therefore, it will usually be sufficient to calculate λ using one iteration of the Newton–Raphson method with vanishing initial value. Thus,

$$\lambda \approx \frac{1 - f(0)}{f'(0)} = \frac{1}{2} \left(\bar{q}^{*\,T}_{\text{add}}\, P_{qq}\, \bar{q}^*_{\text{add}}\right)^{-1} \left(\bar{q}^{*\,T}_{\text{add}}\, \bar{q}^*_{\text{add}} - 1\right) . \tag{68}$$

To lowest nonvanishing order

$$\bar{q}^*_{\text{add}} \approx \bar{1} , \quad \text{and} \quad \bar{q}^{*\,T}_{\text{add}}\, \bar{q}^*_{\text{add}} - 1 \approx 2\, \Delta q_{4\,\text{add}} . \tag{69}$$

Hence,

$$\lambda = \left(P_{qq}\right)^{-1}_{22}\, \Delta q_{4\,\text{add}} . \tag{70}$$

Substituting this in Eq. (66) leads to lowest order in $\Delta\bar{q}^*_{\text{add}}$

$$\bar{\eta}^* = \left(I + \lambda P_{qq}\right)^{-1} \bar{q}^*_{\text{add}} \tag{71a}$$

$$\approx \left(I - \lambda P_{qq}\right) \bar{q}^*_{\text{add}} \tag{71b}$$

$$= \bar{q}^*_{\text{add}} - \Delta q^*_{4\,\text{add}} \left(P_{qq}\right)^{-1}_{22} P_{qq}\, \bar{q}^*_{\text{add}} . \tag{71c}$$

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

$$\boldsymbol{\eta}^* = \Delta \mathbf{q}_{\text{add}} - \left(P_{qq}\right)^{-1}_{22} \left(P_{qq}\right)_{12} \Delta q^*_{4\,\text{add}} . \tag{72}$$

But

$$-\left(P_{qq}\right)_{12} \left(P_{qq}\right)^{-1}_{22} = \left(P^{-1}_{qq}\right)^{-1}_{11} \left(P^{-1}_{qq}\right)_{12} , \tag{73}$$

so that, in fact, comparing Eq. (83) with Eq. (70) we have

$$\boldsymbol{\eta}^* = \Delta\mathbf{q}^*_{\rm mult} \,. \tag{74}$$

Since the other component must also agree to linear order in $\Delta\bar{q}^*_{\rm mult}$, it follows that

$$\bar{\eta}^* = \delta\bar{q}_{\rm mult} \,. \tag{75}$$

Thus, the additive correction to the quaternion, followed by the normalization correction *dictated unambiguously by the maximum likelihood criterion*, is identical (at least up to linear terms in $\Delta\bar{q}^*$) to the so-called multiplicative correction. It is hard to imagine that any other answer could have been possible. Obviously, it is less burdensome to calculate the multiplicative correction directly. Identical arguments hold for sequential correction of the quaternion in the Kalman filter.

DISCUSSION

Simple, compact expressions have been obtained for the sensitivity of attitude measurements to the quaternion, which both ignore and take account of the unit-norm constraint.

The implementation of estimation problems in terms of the four components of the quaternion have been examined, and the additive and multiplication formulations of the quaternion correction studied in detail. It has been shown that the distinction between these two processes does not lie chiefly in whether the correction of the *a priori* quaternion is carried out through matrix addition or quaternion multiplication, but whether the unit-norm constraint is respected during the correction process. Adherents of the additive formulation, generally, do not respect the unit-norm constraint.

The additive correction, if done correctly, is identical to the multiplicative correction but is much more burdensome. The first commandment of quaternion correction, therefore, is to multiply. We emphasize that this result is not the product of some heuristic argument or arbitrary *ad hoc* procedure to be "justified" by experiment but the unavoidable conclusion to which one is led unambiguously and rigorously by the estimation criterion.

Although *Kalman filter* appears in the title, there has been no explicit mention of the Kalman filter in this work. The maximum-likelihood estimation techniques presented here are equivalent to the Kalman filter, however, provided we assume that the noise is Gaussian,[16] which is almost always the case. The sensitivity matrices here are those which are used in the update steps of the filter to compute the residual, the residual covariance, and the Kalman filter gain. The negative-log-likelihood of Eq. (64) is not dependent on the specific form of the measurement given by Eq. (4). The treatment here is therefore very general provided the measurement is Gaussian.

ACKNOWLEDGMENTS

The author is grateful to F. Landis Markley and Itzhack Bar-Itzhack for many interesting discussions on this topic over the years. Dr. Markley, in particular, is to be thanked for a very critical reading of the manuscript.

REFERENCES

1. MARKLEY, F. L. "Parameterization of the Attitude," in WERTZ, J. R., (ed.), *Spacecraft Attitude Determination and Control*, Kluwer Academic, Dordrecht, 1978.
2. SHUSTER, M. D. "A Survey of Attitude Representations,"*Journal of the Astronautical Sciences*, Vol. 41, No. 4, pp. 439–517, October–December 1993.
3. SHUSTER, M. D. "Maximum Likelihood Estimation of Spacecraft Attitude," *Journal of the Astronautical Sciences*, Vol. 37, No. 1, January–March, 1989, pp. 79–88.
4. SHUSTER, M. D. "A Simple Kalman Filter and Smoother for Spacecraft Attitude," *Journal of the Astronautical Sciences*, Vol. 37, No. 1, January–March, 1989, pp. 89–106.
5. SHUSTER, M. D. "Kalman Filtering of Spacecraft Attitude and the QUEST Model," *Journal of the Astronautical Sciences* Vol. 38, No. 3, July–September, 1990, pp. 377–393.
6. LEFFERTS, E. J., MARKLEY, F. L. and SHUSTER, M. D. "Kalman Filtering for Spacecraft Attitude Estimation (survey paper)," *Journal of Guidance, Control and Dynamics*, Vol. 5, No. 5, September-October 1982, pp. 417–429.
7. 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.
8. 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. 126–135.
9. BAR-ITZHACK, I. Y., MARKLEY, F. L., and DEUTSCHMAN, J. "Quaternion Normalization in Additive EKF for Spacecraft Attitude Determination," *Flight Mechanics/Estimation Theory Symposium*, NASA Goddard Space Flight Center, May 1991.
10. DEUTSCHMAN, J., MARKLEY, F. L., and BAR-ITZHACK, I. Y. "Quaternion Normalization in Spacecraft Attitude Determination," *Flight Mechanics/Estimation Theory Symposium*, NASA Goddard Space Flight Center, May 1992.
11. BROUCKE, R. A. "Equations of Motion of a Rotating Rigid Body," *Journal of Guidance, Control and Dynamics*, Vol. 13, 1990, pp. 1150–1152.
12. BAR-ITZHACK, I. Y., and IDAN, M. "Recursive Attitude Determination from Vector Observations: Euler Angle Identification," *Journal of Guidance, Control and Dynamics*, Vol. 10, No. 2, March–April 1987, pp. 152–157.
13. FERRARESI, V. A. *Utilização conjunta de sensores inerciais e não-inerciais em determinação de atitude 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.
14. SHUSTER, M. D. "Attitude Analysis in Flatland: The Plane Truth," *Proceedings, Flight Mechanics/Estimation Theory Symposium*, NASA Goddard Space Flight Center, Greenbelt, Maryland, May 1992, pp. 259–274.
15. SORENSON, H. W. *Parameter Estimation*, Marcel Dekker, New York, 1980.
16. 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.