

ATTITUDE ESTIMATION OR QUATERNION ESTIMATION?

F. Landis Markley*

The attitude of spacecraft is represented by a 3×3 orthogonal matrix with unity determinant, which belongs to the three-dimensional special orthogonal group $SO(3)$. The fact that all three-parameter representations of $SO(3)$ are singular or discontinuous for certain attitudes has led to the use of higher-dimensional nonsingular parameterizations, especially the four-component quaternion. In attitude estimation, we are faced with the alternatives of using an attitude representation that is either singular or redundant. Estimation procedures fall into three broad classes. The first estimates a three-dimensional representation of attitude deviations from a reference attitude parameterized by a higher-dimensional nonsingular parameterization. The deviations from the reference are assumed to be small enough to avoid any singularity or discontinuity of the three-dimensional parameterization. The second class, which estimates a higher-dimensional representation subject to enough constraints to leave only three degrees of freedom, is difficult to formulate and apply consistently. The third class estimates a representation of $SO(3)$ with more than three dimensions, treating the parameters as independent. We refer to the most common member of this class as quaternion estimation, to contrast it with attitude estimation. We analyze the first and third of these approaches in the context of an extended Kalman filter with simplified kinematics and measurement models.

INTRODUCTION

Real-time spacecraft attitude estimation generally employs an Extended Kalman Filter (EKF) [1, 2]. Although the 3×3 orthogonal attitude matrix is the fundamental representation of the spacecraft's attitude, the orthogonality requirement imposes six constraints on its nine elements, reflecting the fact that the special orthogonal group $SO(3)$ of rotation matrices has dimension three. Therefore, employing the nine elements of the attitude matrix as components of a state vector in a Kalman filter leads to some complexity in enforcing the constraints [3, 4]. Most EKFs use lower-dimensional parameterizations of $SO(3)$, and several have used minimal three-dimensional parameterizations [5–7]. However, it is a well-known fact that all three-parameter representations of $SO(3)$ are singular or discontinuous for certain attitudes [8]. This has led to the use of higher-dimensional nonsingular parameterizations in EKFs, especially the four-component quaternion [9–13]. Reference 12 presents an overview of Kalman filtering for spacecraft attitude estimation, emphasizing the quaternion representation, with a complete list of references through 1981.

The quaternion has the lowest dimensionality possible for a globally non-singular representation of $SO(3)$, but it still has one superfluous degree of freedom. Thus we are faced with the alternatives of using an attitude representation that is either singular or redundant. Various strategies to avoid or evade this dilemma have been proposed and analyzed [14–18], and we will consider two in this paper. Our preferred strategy, which we refer to as the Multiplicative EKF (MEKF), uses a nonsingular representation for a reference attitude and a three-component representation for the deviations from this reference. This method, which was implicit in some early EKFs for estimating spacecraft attitude [9–11], has been discussed in detail in Reference 18. The second strategy we discuss treats the four components of the quaternion as independent. We also mention, but we do not analyze in detail, methods that estimate a constrained four-component quaternion with three degrees of freedom. These methods have been treated extensively elsewhere [12–17].

*Guidance, Navigation and Control Systems Engineering Branch, NASA's Goddard Space Flight Center, Greenbelt, MD 20771, landis.markley@nasa.gov

We begin with a review of the attitude parameterizations we will use [19, 20]. The MEKF is then discussed in a model with simplified kinematics and measurement processing, in order to bare its essentials. We next turn to the concept of a quaternion probability distribution function, which is not as straightforward as one might expect. This is followed by an analysis of three versions of an EKF for a quaternion with four independent components, using the same models as for the MEKF, and by a summary of our conclusions,

QUATERNIONS

A quaternion is a four-component object with a three-vector part and a scalar part

$$\bar{q} = \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix}. \quad (1)$$

The four components of \bar{q} , known as the Euler symmetric parameters or the Euler-Rodrigues parameters, first appeared in a paper by Euler [21] and in unpublished notes by Gauss [22]; but Rodrigues' classic paper of 1840 first demonstrated their general usefulness [23]. Hamilton introduced the quaternion as an abstract mathematical object in 1844 [24], but there is some question as to whether he correctly understood its relation to rotations [25]. The attitude matrix is generally written as a homogeneous quadratic function of the components of a quaternion,

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

where $I_{3 \times 3}$ is the 3×3 identity matrix, and the cross product matrix is

$$[\mathbf{q} \times] \equiv \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix}. \quad (3)$$

We see from Eq. (2) that \bar{q} and $-\bar{q}$ represent the same attitude, and that this equation gives an orthogonal matrix only if the quaternion \bar{q} has unit length;

$$|\bar{q}|^2 \equiv |\mathbf{q}|^2 + q_4^2 = 1. \quad (4)$$

In this case we can write

$$\bar{q} = \pm \begin{bmatrix} \mathbf{e} \sin(\phi/2) \\ \cos(\phi/2) \end{bmatrix}, \quad (5)$$

where \mathbf{e} and ϕ are the Euler axis and angle of rotation, respectively. We follow References 12 and 20 in writing the quaternion product as

$$\bar{p} \otimes \bar{q} \equiv \begin{bmatrix} p_4 \mathbf{q} + q_4 \mathbf{p} - \mathbf{p} \times \mathbf{q} \\ p_4 q_4 - \mathbf{p} \cdot \mathbf{q} \end{bmatrix}. \quad (6)$$

This differs from the historical multiplication convention [19, 24], denoted by $\bar{p}\bar{q}$ without an infix operator, by the sign of the cross product in the vector part. The two products are related by $\bar{p} \otimes \bar{q} = \bar{q}\bar{p}$. The convention adopted here has the useful property that

$$A(\bar{p})A(\bar{q}) = A(\bar{p} \otimes \bar{q}). \quad (7)$$

With the historical convention, the quaternion ordering on the right side of the above equation would be the reverse of the order on the left side. With either convention, though, the product of two quaternions is bilinear in the elements of the component quaternions, a property shared with the direction cosine matrix but no other attitude representation. In fact, Eq. (7) means that the rotation group and the quaternion group are almost isomorphic, the qualifier "almost" owing to the 2:1 nature of the mapping [26].

The quaternion obeys the kinematic equation

$$\dot{\bar{q}} = \frac{1}{2} \begin{bmatrix} \bar{\omega} \\ 0 \end{bmatrix} \otimes \bar{q} \equiv \frac{1}{2} \bar{\omega} \otimes \bar{q}, \quad (8)$$

where $\bar{\omega}$ is the angular velocity vector in the body frame, and we use a boldface symbol with an overbar to denote a quaternion with zero scalar part.

MULTIPLICATIVE EXTENDED KALMAN FILTER

The MEKF represents the attitude as the quaternion product

$$\bar{q} = \delta\bar{q}(\mathbf{a}) \otimes \bar{q}_{ref}, \quad (9)$$

where \bar{q}_{ref} is some unit reference quaternion and $\delta\bar{q}(\mathbf{a})$ is a unit quaternion representing the rotation from \bar{q}_{ref} to the true attitude \bar{q} , parameterized by a three-component vector \mathbf{a} . Although several choices for \mathbf{a} are possible [18], in this paper we choose it to be two times the vector of Rodrigues parameters [19, 20, 23],

$$\mathbf{a} \equiv 2\delta\mathbf{q}/\delta q_4 = 2\mathbf{e} \tan(\phi/2). \quad (10)$$

The factor of two is chosen so that $|\mathbf{a}| \approx \phi$ when the rotation angle between \bar{q} and \bar{q}_{ref} is small. The inverse of Eq. (10) is

$$\delta\bar{q}(\mathbf{a}) = \frac{1}{\sqrt{4 + |\mathbf{a}|^2}} \begin{bmatrix} \mathbf{a} \\ 2 \end{bmatrix}, \quad (11)$$

The three Rodrigues parameters are often referred to as the Gibbs vector, since they were arrayed by Gibbs in his “vector semitangent of version” [27]. They are infinite for 180° rotations, which is undesirable for a global representation of rotations; but they provide an excellent representation of rotations less than 180°. Since Eq. (10) maps \bar{q} and $-\bar{q}$ to the same point in three-dimensional Euclidean space E3, these parameters provide a 1:1 representation of the rotations less than 180° onto E3, which provides support for Gaussian distributions with infinitely long tails.

The two attitude representations $\delta\bar{q}$ and \bar{q}_{ref} in Eq. (9) are clearly redundant. The basic idea of the MEKF is to compute an unconstrained estimate of the three-component \mathbf{a} while using the correctly normalized four-component \bar{q}_{ref} to provide a globally nonsingular attitude representation. Given an estimate $\hat{\mathbf{a}}$, where a caret will denote the expectation of a random variable, Eq. (9) says that the corresponding estimate of the true attitude quaternion is $\delta\bar{q}(\hat{\mathbf{a}}) \otimes \bar{q}_{ref}$. We remove the redundancy in the attitude representation by choosing the reference quaternion \bar{q}_{ref} so that $\hat{\mathbf{a}}$ is identically zero. Since $\delta\bar{q}(0)$ is the identity quaternion, this means that the reference quaternion is the best estimate of the true quaternion. Note that the reference quaternion in the MEKF is not considered to be a random variable, and its estimate is not an expectation. The fundamental conceptual advantage of the MEKF is that \bar{q}_{ref} is a unit quaternion by definition.

The identification of \bar{q}_{ref} as the attitude estimate means in turn that \mathbf{a} is a three-component representation of the attitude error, the difference between the truth and our estimate. This provides a consistent treatment of the attitude error statistics, with the covariance of the attitude error angles in the body frame (in radians squared) represented by the covariance of \mathbf{a} .

An alternative formulation, which has some advantages, reverses the order of multiplication in Eq. (9) so that \mathbf{a} represents the attitude errors in the inertial reference frame rather than in the body frame [28]. It is also possible to represent the reference attitude by a reference attitude matrix A_{ref} rather than by a reference quaternion [9, 28]. This requires more parameters, but may save computations if the attitude matrix is explicitly required. An argument in favor of the quaternion is that it is easy to restore normalization that may be lost due to numerical errors, while restoring the orthogonality of A_{ref} is nontrivial. Gray has argued that this argument is not compelling if reasonable computational care is taken [28].

Kinematics

Since \bar{q}_{ref} is a unit quaternion, it must obey a kinematic equation of the form

$$\dot{\bar{q}}_{ref} = \frac{1}{2} \bar{\omega}_{ref} \otimes \bar{q}_{ref}, \quad (12)$$

where ω_{ref} has the obvious interpretation as the angular velocity of the reference attitude. We now show how ω_{ref} is determined by the requirement that $\hat{\mathbf{a}}$ be identically zero, which is the condition that the reference attitude be the optimal attitude estimate. Computing the time derivative of Eq. (9), using Eqs. (8) and (12), gives

$$\frac{1}{2} \bar{\omega} \otimes \bar{q} = \delta \dot{\bar{q}} \otimes \bar{q}_{ref} + \frac{1}{2} \delta \bar{q} \otimes \bar{\omega}_{ref} \otimes \bar{q}_{ref}. \quad (13)$$

Substitute Eq. (9) for \bar{q} on the left side, right-multiply the entire equation by the inverse of \bar{q}_{ref} , and rearrange to get

$$\delta \dot{\bar{q}} = \frac{1}{2} (\bar{\omega} \otimes \delta \bar{q} - \delta \bar{q} \otimes \bar{\omega}_{ref}). \quad (14)$$

Note that ω_{ref} is not a random variable. Substituting Eq. (14) into the time derivative of Eq. (10) gives

$$\dot{\mathbf{a}} = (I_{3 \times 3} + \frac{1}{4} \mathbf{a} \mathbf{a}^T) (\omega - \omega_{ref}) - \frac{1}{2} (\omega + \omega_{ref}) \times \mathbf{a}, \quad (15)$$

after some straightforward quaternion algebra. We now assume the simple kinematic model

$$\omega(t) = \hat{\omega}(t) + \mathbf{n}_\omega(t), \quad (16)$$

where $\hat{\omega}(t)$ is the nominal angular velocity and $\mathbf{n}_\omega(t)$ is a white noise process, satisfying

$$E\{\mathbf{n}_\omega(t) \mathbf{n}_\omega^T(t')\} = Q \delta(t - t') = \sigma_\omega^2 I_{3 \times 3} \delta(t - t'), \quad (17)$$

with $\delta(t - t')$ denoting the Dirac delta function. Linearizing Eq. (15) in \mathbf{a} and \mathbf{n}_ω gives

$$\dot{\mathbf{a}} = \hat{\omega} - \omega_{ref} + F \mathbf{a} + G \mathbf{n}_\omega \quad (18)$$

with

$$F \equiv -\frac{1}{2} [(\hat{\omega} + \omega_{ref}) \times] \quad (19)$$

and

$$G \equiv I_{3 \times 3}. \quad (20)$$

Measurements

We assume discrete measurements with covariance R . For ease of analysis, though, we will assume that the measurement interval τ is so much shorter than any filter time constant that we can treat the measurements as continuous with covariance $R\tau$ [1]. The attitude error estimate $\hat{\mathbf{a}}$ conditioned on the measurements is then given by the expectation of Eq. (18) augmented by a measurement term [1],

$$\dot{\hat{\mathbf{a}}} = \hat{\omega} - \omega_{ref} + F \hat{\mathbf{a}} + P_a H_M^T (\tau R)^{-1} [\mathbf{z} - \mathbf{h}(\hat{\mathbf{a}})], \quad (21)$$

where the subscript M denotes the MEKF and the 3×3 covariance matrix of the attitude error vector satisfies

$$\dot{P}_a = F P_a + P_a F^T + G Q G^T - P_a H_M^T (\tau R)^{-1} H_M P_a. \quad (22)$$

We model the measurement of a vector \mathbf{v}_b in the body coordinate frame by

$$\mathbf{z} = \mathbf{h}_M(\mathbf{a}) + \mathbf{n}_z = \mathbf{v}_b - A(\bar{q}) \mathbf{v}_l + \mathbf{n}_z, \quad (23)$$

where \mathbf{v}_I is the corresponding vector in an inertial frame and \mathbf{n}_z is Gaussian measurement noise. To first order in the attitude error vector \mathbf{a} , we have

$$A(\bar{q}) = A(\delta\bar{q} \otimes \bar{q}_{ref}) = A(\delta\bar{q})A(\bar{q}_{ref}) \approx (I_{3 \times 3} - [\mathbf{a} \times])A(\bar{q}_{ref}), \quad (24)$$

where the last step follows from Eqs. (2) and (11). Substituting into Eq. (23) gives

$$\mathbf{h}_M(\mathbf{a}) \approx \mathbf{v}_B - (I_{3 \times 3} - [\mathbf{a} \times])A(\bar{q}_{ref})\mathbf{v}_I = \mathbf{v}_B - \hat{\mathbf{v}}_M - \hat{\mathbf{v}}_M \times \mathbf{a} \quad (25)$$

with $\hat{\mathbf{v}}_M$ denoting the measurement in the body frame predicted by the reference quaternion:

$$\hat{\mathbf{v}}_M \equiv A(\bar{q}_{ref})\mathbf{v}_I. \quad (26)$$

It follows that the measurement sensitivity matrix is

$$H_M \equiv \partial \mathbf{h}_M / \partial \mathbf{a} \big|_{\hat{\mathbf{a}}} = -[\hat{\mathbf{v}}_M \times]. \quad (27)$$

If we assume that the vector measurement errors are isotropic, as is appropriate for measurement of a full vector such as the magnetic field, we have

$$R = \sigma_z^2 I_{3 \times 3}. \quad (28)$$

Making appropriate substitutions into Eq. (21) gives

$$\dot{\hat{\mathbf{a}}} = \hat{\boldsymbol{\omega}} - \boldsymbol{\omega}_{ref} - \frac{1}{2}(\hat{\boldsymbol{\omega}} + \boldsymbol{\omega}_{ref}) \times \hat{\mathbf{a}} - \tau^{-1} \sigma_z^{-2} P_a [\hat{\mathbf{v}}_M \times] (\mathbf{v}_B - \hat{\mathbf{v}}_M), \quad (29)$$

where \mathbf{v}_B is the measured vector in the body and we've made use of the fact that the observed value of \mathbf{z} is identically zero for our measurement models. Since the reference quaternion \bar{q}_{ref} is specified in the MEKF by the requirement that $\hat{\mathbf{a}}$ be identically zero, $\boldsymbol{\omega}_{ref}$ is given as

$$\boldsymbol{\omega}_{ref} = \hat{\boldsymbol{\omega}} + \tau^{-1} \sigma_z^{-2} P_a (\mathbf{v}_B \times \hat{\mathbf{v}}_M). \quad (30)$$

The reference quaternion propagation including the effect of measurements is given by Eq. (12) with this angular rate. We see that the unity norm of the reference quaternion is maintained exactly by this propagation. No "reset" operation, as required in the discrete measurement case [18], is needed in the continuous measurement approximation. We emphasize that it is not necessary to integrate Eq. (29) in the MEKF; this equation is only used to derive Eq. (30).

Substituting F , G , $\boldsymbol{\omega}_{ref}$, and H_M into Eq. (22) for the covariance matrix gives

$$\begin{aligned} \dot{P}_a = & -[\hat{\boldsymbol{\omega}} \times] P_a - P_a [\hat{\boldsymbol{\omega}} \times]^T + \sigma_\omega^2 I_{3 \times 3} - \tau^{-1} \sigma_z^{-2} P_a \left(\left| \hat{\mathbf{v}}_M \right|^2 I_{3 \times 3} - \hat{\mathbf{v}}_M \hat{\mathbf{v}}_M^T \right) P_a \\ & - \frac{1}{2} \tau^{-1} \sigma_z^{-2} \left\{ [P_a (\mathbf{v}_B \times \hat{\mathbf{v}}_M) \times] P_a + P_a [P_a (\mathbf{v}_B \times \hat{\mathbf{v}}_M) \times]^T \right\}. \end{aligned} \quad (31)$$

The last term in this equation depends on the measurement residuals. Since the covariance is defined as the expectation over the distribution of measurement errors, which are assumed to have zero mean, this term should be omitted, giving

$$\dot{P}_a = -[\hat{\boldsymbol{\omega}} \times] P_a - P_a [\hat{\boldsymbol{\omega}} \times]^T + \sigma_\omega^2 I_{3 \times 3} - \tau^{-1} \sigma_z^{-2} P_a \left(\left| \hat{\mathbf{v}}_M \right|^2 I_{3 \times 3} - \hat{\mathbf{v}}_M \hat{\mathbf{v}}_M^T \right) P_a. \quad (32)$$

If this argument isn't convincing, we can simply note that the discarded term is negligibly small compared to the terms retained in Eq. (32).

The state propagation of Eqs. (12) and (30) and the covariance propagation of Eq. (32) are the results we need from the MEKF.

Unit Vector Measurements

The errors of unit vector measurements are not isotropic, so Eq. (28) no longer applies. Consider a distribution of unit vectors about the z axis:

$$\mathbf{u} = \begin{bmatrix} x \\ y \\ \sqrt{1-x^2-y^2} \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 - \frac{1}{2}(x^2 + y^2) - \frac{1}{8}(x^2 + y^2)^2 + \dots \end{bmatrix}. \quad (33)$$

We'll assume that x and y both have Gaussian pdfs with zero mean and variance σ_z^2 . This is an idealization, since both x and y have magnitude less than unity, but it is a reasonable approximation if $\sigma_z \ll 1$. Now

$$\hat{\mathbf{u}} \equiv E\{\mathbf{u}\} = \begin{bmatrix} 0 \\ 0 \\ 1 - \sigma_z^2 - \sigma_z^4 + O(\sigma_z^6) \end{bmatrix} \quad (34)$$

and so the measurement covariance for unit vector measurements is

$$\begin{aligned} R_{unit} &\equiv E\{(\mathbf{u} - \hat{\mathbf{u}})(\mathbf{u} - \hat{\mathbf{u}})^T\} = E\left\{ \begin{bmatrix} x^2 & xy & x\sqrt{1-x^2-y^2} \\ xy & y^2 & y\sqrt{1-x^2-y^2} \\ x\sqrt{1-x^2-y^2} & y\sqrt{1-x^2-y^2} & 1-x^2-y^2 \end{bmatrix} \right\} - \hat{\mathbf{u}}\hat{\mathbf{u}}^T \\ &= \sigma_z^2 \{ I_{3 \times 3} - \mathbf{u}_0 \mathbf{u}_0^T + \sigma_z^2 [1 + O(\sigma_z^2)] \mathbf{u}_0 \mathbf{u}_0^T \}, \end{aligned} \quad (35)$$

where $\mathbf{u}_0 = [001]^T$ denotes the value of the unit vector for zero errors. The inverse of this is, keeping only the dominant term in each of the two orthogonal projections $\mathbf{u}_0 \mathbf{u}_0^T$ and $I_{3 \times 3} - \mathbf{u}_0 \mathbf{u}_0^T$,

$$R_{unit}^{-1} = \sigma_z^{-2} (I_{3 \times 3} - \mathbf{u}_0 \mathbf{u}_0^T + \sigma_z^2 \mathbf{u}_0 \mathbf{u}_0^T). \quad (36)$$

When this is substituted into Eq. (21) or (22), using $\hat{\mathbf{v}}_M$ for \mathbf{u}_0 , only the first term in the parentheses on the right side of Eq. (36) gives a non-zero contribution because of the form of H_M in Eq. (27). Thus unit vector measurements give the same result in the MEKF as vector measurements with isotropic errors.

QUATERNION ESTIMATION

It is natural to define the estimate of an unconstrained quaternion as an expectation value, defined as the integral over four-dimensional Euclidean space $E4$ weighted by a probability distribution function (pdf) $\rho(\bar{q} | Z)$ conditioned on the measurements Z :

$$\hat{\bar{q}} \equiv E\{\bar{q} | Z\} = \int_{E4} \bar{q} \rho(\bar{q} | Z) d^4 \bar{q}. \quad (37)$$

Figure 1 gives several examples of possible quaternion distributions in four-dimensional Euclidean space. Of course, only a two-dimensional cross-section of $E4$ can be shown in the figure, and one quadrant of such a cross-section is displayed in each plot with the origin of $E4$ at the lower left-hand corner.

Unit quaternions reside on the three-dimensional sphere $S3$ embedded in four-dimensional Euclidean space $E4$, indicated by the arcs in Fig. 1. Thus the pdf for a distribution of unit quaternions is given by

$$\rho(\bar{q} | Z) = \tilde{\rho}(\bar{q} | Z) \delta(\|\bar{q}\|^2 - 1), \quad (38)$$

where $\tilde{\rho}(\bar{q} | Z)$ is the pdf on $S3$ and the Dirac delta function enforces the unit norm constraint. Equations (37) and (38) provide an unsatisfactory definition of $\hat{\bar{q}}$, however, since restricting the pdf in quaternion

space to the surface of a unit sphere means that its expectation value must be *inside* the sphere, as illustrated in Fig. 1a, unless the pdf is concentrated at a point. The unity norm constraint violation will be on the order of the variance of the attitude errors, like the unit vector expectation of Eq. (34). The definition of the covariance of a distribution restricted to S3 raises similar problems.

It can be argued that the expectation value should not be naively performed in Cartesian coordinates in E4 as in Eqs. (37) and (38), but rather in terms some non-Euclidean parameterization of S3. In fact, this is exactly what the MEKF effectively accomplishes, although the MEKF entirely avoids the concept of a quaternion pdf. Having already presented this approach, we will not discuss quaternion pdfs restricted to S3 any further.

An alternative is not to assume the quaternion to be normalized. In this case an orthogonal attitude matrix is formed by using the normalized quaternion $\bar{q}/|\bar{q}|$ in the quadratic form of Eq. (2), now indicated explicitly by a subscript Q ,

$$A_R(\bar{q}) = A_Q(\bar{q}/|\bar{q}|) = |\bar{q}|^{-2} \left\{ (q_4^2 - |\mathbf{q}|^2) I_{3 \times 3} + 2\mathbf{q}\mathbf{q}^T - 2q_4[\mathbf{q} \times] \right\}. \quad (39)$$

The subscript R identifies this as the ray representation model, since any quaternion along a ray in E4 (a straight line through the origin) represents the same attitude in this representation, with the exception of the zero quaternion at the origin. The twofold ambiguity of the unit quaternion representation corresponds to the two points where the ray pierces the unit sphere S3. This was referred to as the “linearized orthogonalized matrix” (LOM) model in References 14 and 15. Note that Eq. (5) does not hold for a non-unit quaternion.

The 4x4 quaternion covariance matrix P_q reveals its structure with the aid of the matrix

$$\Xi(\bar{q}) \equiv \begin{bmatrix} q_4 I_{3 \times 3} + [\mathbf{q} \times] \\ -\mathbf{q}^T \end{bmatrix}, \quad (40)$$

which obeys the identities

$$\Xi^T(\bar{q}) \bar{q} = \mathbf{0}, \quad (41)$$

$$\Xi^T(\bar{q}) \Xi(\bar{q}) = |\bar{q}|^2 I_{3 \times 3}, \quad (42)$$

and

$$\Xi(\bar{q}) \Xi^T(\bar{q}) + \bar{q} \bar{q}^T = |\bar{q}|^2 I_{4 \times 4}. \quad (43)$$

These identities mean that the 4x4 matrix $|\bar{q}|^{-1} [\Xi(\bar{q}) \quad \bar{q}]$ is an orthogonal matrix whether or not \bar{q} is a unit quaternion, so the quaternion covariance can be parameterized without loss of generality as

$$P_q \equiv E\{(\bar{q} - \hat{\bar{q}})(\bar{q} - \hat{\bar{q}})^T\} = \begin{bmatrix} \Xi(\hat{\bar{q}}) & \hat{\bar{q}} \end{bmatrix} \begin{bmatrix} \frac{1}{4} P_a & \frac{1}{2} \mathbf{p}_c \\ \frac{1}{2} \mathbf{p}_c^T & p_n \end{bmatrix} \begin{bmatrix} \Xi^T(\hat{\bar{q}}) \\ \hat{\bar{q}}^T \end{bmatrix}. \quad (44)$$

The interpretation of this partitioning is that P_a is the 3x3 attitude covariance, p_n is the quaternion norm variance, and \mathbf{p}_c is the vector of covariance between the attitude and norm errors. This factorization is not used in unconstrained quaternion estimation, but gives insight into what is happening behind the scenes.

Figure 1b shows a distribution for which the mean has unit norm and for which the scatter in norm (i.e. along the ray from the origin through the mean) is less than the scatter in angle. Figure 1c shows a case for which the angular errors are smaller than the norm errors, and Fig. 1d illustrates a pdf with isotropic errors. Since the expected value of the quaternion is not required to be on S3 in a ray representation, Figs. 1c and 1d illustrate cases where the norm is $\hat{\bar{q}} = 1.1$. All the pdfs illustrated in Fig. 1 have $\mathbf{p}_c = \mathbf{0}$, with the exception of Fig. 1e, which is to say that is the only figure exhibiting a correlation between norm and angular errors.

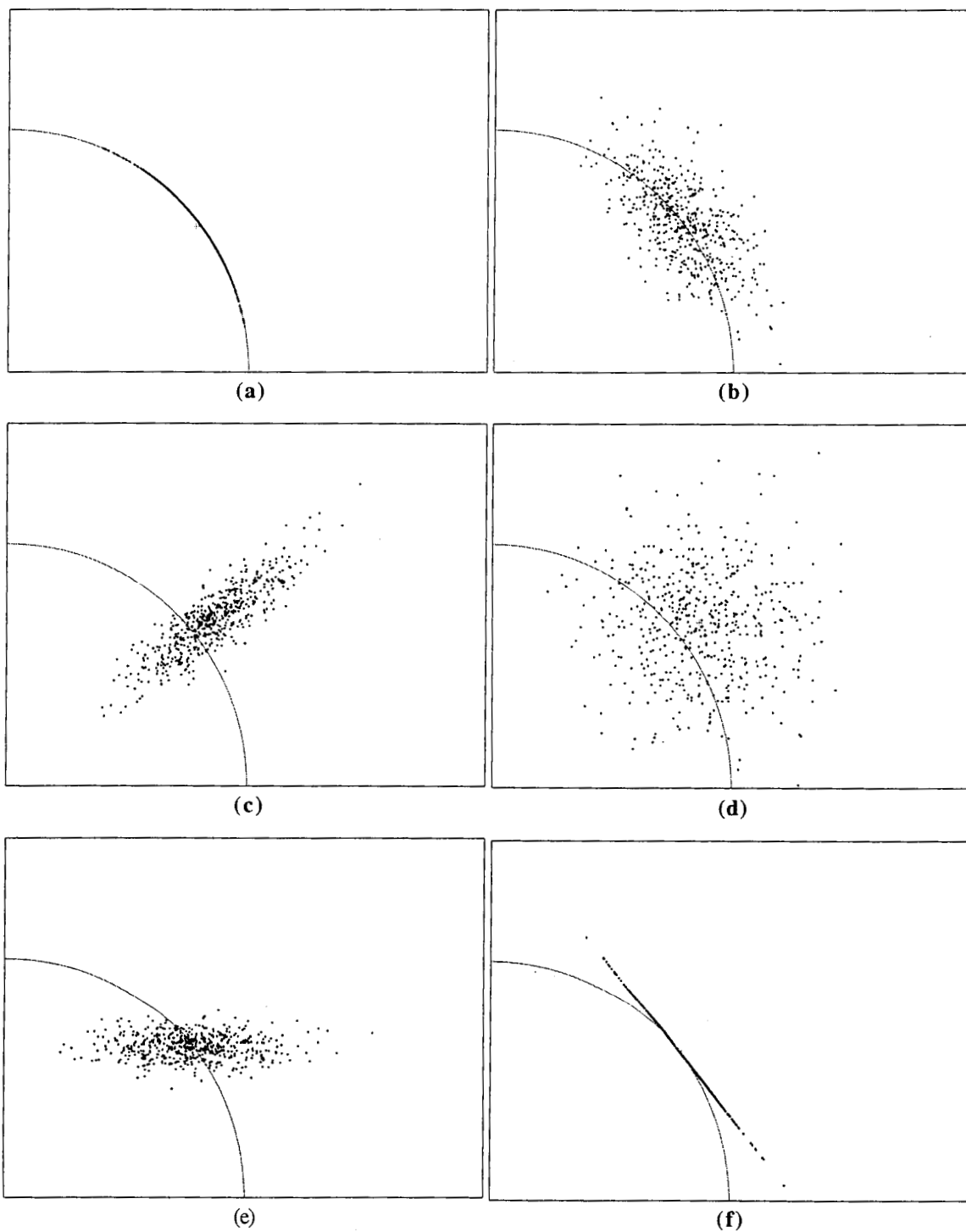


Figure 1 Representative Quaternion Probability Distributions:
The arc is one quadrant of the projection of unit sphere S^3 onto the plane.
The centroid of the distribution in each subplot is denoted by +.

Finally, Fig. 1f shows a pdf restricted to a three-dimensional manifold in E4, similarly to Fig. 1a. However, this manifold in Fig. 1f is not S3, but rather is the hyperplane tangent to S3 at \hat{q} . The covariance of the pdf illustrated by this figure has both $\mathbf{p}_c = \mathbf{0}$ and $p_n = 0$, since all the errors are orthogonal to \hat{q} . This also means that that \hat{q} is a null eigenvector of P_q . This may seem like a peculiar pdf, but it is the form implicitly assumed in Reference 12, and one of the EKFs we consider below will be shown to drive the error distribution to this limit.

Quaternion propagation with the kinematics model of Eq. (16) can be written as

$$\dot{\bar{q}} = \frac{1}{2}(\hat{\omega} + \bar{\mathbf{n}}_\omega) \otimes \bar{q} = \frac{1}{2}\Omega(\hat{\omega})\bar{q} + \frac{1}{2}\Xi(\bar{q})\mathbf{n}_\omega, \quad (45)$$

with $\Xi(\bar{q})$ given by Eq. (40) and

$$\Omega(\omega) = \begin{bmatrix} -[\omega \times] & \omega \\ -\omega^T & 0 \end{bmatrix}. \quad (46)$$

Equation (45) is independent of quaternion normalization, because Eq. (8) is a homogenous in \bar{q} . Letting $\Delta\bar{q} \equiv \bar{q} - \hat{q}$ and linearizing gives the EKF kinematics

$$\Delta\dot{\bar{q}} = F\Delta\bar{q} + G\mathbf{n}_\omega, \quad (47)$$

with

$$F = \frac{1}{2}\Omega(\hat{\omega}) \quad (48)$$

and

$$G = \frac{1}{2}\Xi(\hat{q}) \equiv \frac{1}{2}\hat{\Xi} \quad (49)$$

Substituting Eq. (44) into the quaternion propagation equation

$$\dot{P}_q = FP_q + P_qF^T + GQG^T - P_qH^T(\tau R)^{-1}HP_q. \quad (50)$$

and performing some algebra using the quaternion kinematics and Eq. (A1) of Reference 12,

$$\dot{\hat{\Xi}} = \frac{1}{2}\Omega(\hat{\omega})\hat{\Xi} + \hat{\Xi}[\hat{\omega} \times], \quad (51)$$

gives

$$\begin{bmatrix} \frac{1}{4}\dot{P}_a & \frac{1}{2}\dot{\mathbf{p}}_c \\ \frac{1}{2}\dot{\mathbf{p}}_c^T & \dot{P}_n \end{bmatrix} = \begin{bmatrix} -\frac{1}{4}([\hat{\omega} \times]P_a + P_a[\hat{\omega} \times]^T) + \frac{1}{4}\sigma_\omega^2 & -\frac{1}{2}\hat{\omega} \times \mathbf{p}_c \\ -\frac{1}{2}(\hat{\omega} \times \mathbf{p}_c)^T & 0 \end{bmatrix} - \begin{bmatrix} \frac{1}{4}P_a & \frac{1}{2}\mathbf{p}_c \\ \frac{1}{2}\mathbf{p}_c^T & P_n \end{bmatrix} \begin{bmatrix} \hat{\Xi}^T \\ \hat{q}^T \end{bmatrix} H^T(\tau R)^{-1}H \begin{bmatrix} \hat{\Xi} \\ \hat{q} \end{bmatrix} \begin{bmatrix} \frac{1}{4}P_a & \frac{1}{2}\mathbf{p}_c \\ \frac{1}{2}\mathbf{p}_c^T & P_n \end{bmatrix}. \quad (52)$$

We now consider three implementations of the quaternion EKF.

Ray Representation EKF

This uses the orthogonal attitude matrix of Eq. (39) in the measurement model

$$\mathbf{z} = \mathbf{h}_R(\bar{q}) + \mathbf{n}_R = \mathbf{v}_B - A_R(\bar{q})\mathbf{v}_I + \mathbf{n}_z. \quad (53)$$

After a good deal of algebra, we find the measurement sensitivity matrix

$$H_R \equiv \partial \mathbf{h}_R / \partial \bar{q} \Big|_{\hat{q}} = -2|\hat{q}|^{-2}[\hat{\mathbf{v}}_R \times] \hat{\Xi}^T, \quad (54)$$

with $\hat{\mathbf{v}}_R$ denoting the measurement in the body frame predicted by the expected quaternion:

$$\hat{\mathbf{v}}_R \equiv A_R(\hat{q})\mathbf{v}_I. \quad (55)$$

Isotropic vector measurements with measurement covariance obeying Eq. (28) give

$$\begin{bmatrix} \hat{\Xi}^T \\ \hat{q}^T \end{bmatrix} H_R^T (\tau R)^{-1} H_R \begin{bmatrix} \hat{\Xi} \\ \hat{q} \end{bmatrix} = 4\tau^{-1} \sigma_z^{-2} \begin{bmatrix} |\hat{\mathbf{v}}_R|^2 I_{3 \times 3} - \hat{\mathbf{v}}_R \hat{\mathbf{v}}_R^T & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}, \quad (56)$$

which leads to the covariance propagation equations

$$\dot{P}_a = -[\hat{\boldsymbol{\omega}} \times] P_a - P_a [\hat{\boldsymbol{\omega}} \times]^T + \sigma_\omega^2 I_{3 \times 3} - \tau^{-1} \sigma_z^{-2} P_a \left(|\hat{\mathbf{v}}_R|^2 I_{3 \times 3} - \hat{\mathbf{v}}_R \hat{\mathbf{v}}_R^T \right) P_a, \quad (57a)$$

$$\dot{\mathbf{p}}_c = -\hat{\boldsymbol{\omega}} \times \mathbf{p}_c - \tau^{-1} \sigma_z^{-2} P_a \left(|\hat{\mathbf{v}}_R|^2 I_{3 \times 3} - \hat{\mathbf{v}}_R \hat{\mathbf{v}}_R^T \right) \mathbf{p}_c, \quad (57b)$$

$$\dot{p}_n = -\tau^{-1} \sigma_z^{-2} |\hat{\mathbf{v}}_R \times \mathbf{p}_c|^2. \quad (57c)$$

The 3×3 attitude covariance P_a obeys the same equation as in the MEKF, Eq. (32). We see from Eq. (44) that $\mathbf{p}_c(t)$ is zero initially if the initial estimate \hat{q} is an eigenvector of the initial covariance; this includes the case that the initial P_{qq} is a multiple of the identity. In this case, Eqs. (57b) and (57c) show that $\mathbf{p}_c(t)$ is zero for all time and the norm variance p_n is constant. The norm variance is mixed in with the attitude covariance in P_{qq} , which may cause loss of significance of the latter if the attitude estimate converges to small errors and the probability distribution tends to be concentrated along a ray, as in Figure 1c.

The expectation of Eq. (45) augmented by a measurement term gives the equation for the time dependence of the quaternion estimate

$$\dot{\hat{q}} = \frac{1}{2} \Omega(\hat{\boldsymbol{\omega}}) \hat{q} - P H_R^T (\tau R)^{-1} [\mathbf{v}_B - A_R(\hat{q}) \mathbf{v}_I] = \frac{1}{2} \Omega(\hat{\boldsymbol{\omega}}) \hat{q} + 2\tau^{-1} \sigma_z^{-2} \begin{bmatrix} \hat{\Xi} \\ \hat{q} \end{bmatrix} \begin{bmatrix} \frac{1}{4} P_a & \frac{1}{2} \mathbf{p}_c \\ \frac{1}{2} \mathbf{p}_c^T & p_n \end{bmatrix} \begin{bmatrix} [\hat{\mathbf{v}}_R \times]^T \\ \mathbf{0}^T \end{bmatrix} \mathbf{v}_B, \quad (58)$$

where we have made use of the fact that $[\hat{\mathbf{v}}_R \times]^T A_R(\hat{q}) \mathbf{v}_I = -\hat{\mathbf{v}}_R \times \hat{\mathbf{v}}_R \equiv \mathbf{0}$. This can be expressed as

$$\dot{\hat{q}} = \frac{1}{2} \bar{\boldsymbol{\omega}}_{\text{eff}} \otimes \hat{q} + \tau^{-1} \sigma_z^{-2} [(\hat{\mathbf{v}}_R \times \mathbf{p}_c) \cdot \mathbf{v}_B] \hat{q}, \quad (59)$$

with

$$\boldsymbol{\omega}_{\text{eff}} \equiv \hat{\boldsymbol{\omega}} + \tau^{-1} \sigma_z^{-2} P_a (\mathbf{v}_B \times \hat{\mathbf{v}}_R). \quad (60)$$

The norm of the quaternion estimate satisfies

$$\frac{d}{dt} |\hat{q}|^2 = 2\tau^{-1} \sigma_z^{-2} |\hat{q}|^2 (\hat{\mathbf{v}}_R \times \mathbf{p}_c) \cdot \mathbf{v}_B. \quad (61)$$

We see that the norm is constant if $\hat{\mathbf{v}}_R \times \mathbf{p}_c(t) \equiv \mathbf{0}$, which is consistent with Eq. (57c). The time dependence of the normalized quaternion is found by combining Eqs. (59) and (61):

$$\frac{d}{dt} \left(\frac{\hat{q}}{|\hat{q}|} \right) = \frac{\dot{\hat{q}}}{|\hat{q}|} - \frac{\hat{q}}{2|\hat{q}|^3} \frac{d}{dt} |\hat{q}|^2 = \frac{1}{2} \bar{\boldsymbol{\omega}}_{\text{eff}} \otimes \frac{\hat{q}}{|\hat{q}|}. \quad (62)$$

Thus the attitude estimate is the same for the ray representation EKF as for the MEKF.

Using Eq. (36) for unit vector measurement errors gives, remembering that $\hat{\mathbf{v}}_R = \mathbf{u}_0$ is a unit vector,

$$H_R^T R_{\text{unit}}^{-1} = -2\sigma_z^{-2} |\hat{q}|^{-2} \hat{\Xi} [\hat{\mathbf{v}}_R \times]^T (I_{3 \times 3} - \hat{\mathbf{v}}_R \hat{\mathbf{v}}_R^T + \sigma_z^{-2} \hat{\mathbf{v}}_R \hat{\mathbf{v}}_R^T) = -2\sigma_z^{-2} |\hat{q}|^{-2} \hat{\Xi} [\hat{\mathbf{v}}_R \times]^T, \quad (63)$$

which is the same as the form derived above for vector measurements with isotropic errors. Thus the ray representation EKF gives the same covariance and state updates with unit vectors as with vectors having isotropic errors.

The ray representation form of quaternion EKF, which we have seen to be effectively equivalent to the MEKF, has been applied to attitude estimation of the ALEXIS and CAPER spacecraft [29, 30].

Quadratic Representation EKF

This uses the homogenous quadratic expression for the attitude matrix of Eq. (2) in the measurement model

$$\mathbf{z} = \mathbf{h}_Q(\bar{\mathbf{q}}) + \mathbf{n}_Q = \mathbf{v}_B - A_Q(\bar{\mathbf{q}}) \mathbf{v}_I + \mathbf{n}_z. \quad (64)$$

We note that the attitude matrix is not orthogonal unless the quaternion has unit norm. We will see that these measurements drive the norm to unity, however. The measurement sensitivity matrix is

$$H_Q \equiv \partial \mathbf{h}_Q / \partial \bar{\mathbf{q}} \Big|_{\hat{\mathbf{q}}} = -2 \left| \hat{\mathbf{q}} \right|^{-2} \left((\hat{\mathbf{v}}_Q \times) \hat{\Xi}^T + \hat{\mathbf{v}}_Q \hat{\hat{\mathbf{q}}}^T \right), \quad (65)$$

where

$$\hat{\mathbf{v}}_Q \equiv A_Q(\hat{\mathbf{q}}) \mathbf{v}_I, \quad (66)$$

is the predicted measurement in the body frame. For isotropic measurement errors, we have

$$\begin{bmatrix} \hat{\Xi}^T \\ \hat{\hat{\mathbf{q}}}^T \end{bmatrix} H_Q^T (\tau R)^{-1} H_Q \begin{bmatrix} \hat{\Xi} \\ \hat{\hat{\mathbf{q}}} \end{bmatrix} = 4\tau^{-1} \sigma_z^{-2} \begin{bmatrix} \left| \hat{\mathbf{v}}_Q \right|^2 I_{3 \times 3} - \hat{\mathbf{v}}_Q \hat{\mathbf{v}}_Q^T & \mathbf{0} \\ \mathbf{0}^T & \left| \hat{\mathbf{v}}_Q \right|^2 \end{bmatrix}, \quad (67)$$

which leads to the covariance propagation equations

$$\dot{P}_a = -[\hat{\boldsymbol{\omega}} \times] P_a - P_a [\hat{\boldsymbol{\omega}} \times]^T + \sigma_\omega^2 I_{3 \times 3} - \tau^{-1} \sigma_z^{-2} \left[P_a \left(\left| \hat{\mathbf{v}}_Q \right|^2 I_{3 \times 3} - \hat{\mathbf{v}}_Q \hat{\mathbf{v}}_Q^T \right) P_a + 4 \left| \hat{\mathbf{v}}_Q \right|^2 \mathbf{p}_c \mathbf{p}_c^T \right], \quad (68a)$$

$$\dot{\mathbf{p}}_c = -\hat{\boldsymbol{\omega}} \times \mathbf{p}_c - \tau^{-1} \sigma_z^{-2} \left[P_a \left(\left| \hat{\mathbf{v}}_Q \right|^2 I_{3 \times 3} - \hat{\mathbf{v}}_Q \hat{\mathbf{v}}_Q^T \right) + 4 \left| \hat{\mathbf{v}}_Q \right|^2 P_n \right] \mathbf{p}_c, \quad (68b)$$

$$\dot{P}_n = -\tau^{-1} \sigma_z^{-2} \left(\left| \hat{\mathbf{v}}_Q \times \mathbf{p}_c \right|^2 + 4 \left| \hat{\mathbf{v}}_Q \right|^2 P_n^2 \right). \quad (68c)$$

The last term in Eq. (68c) drives p_n asymptotically to zero, and the last term in Eq. (68b) drives \mathbf{p}_c to zero. Thus the 4x4 covariance matrix becomes singular asymptotically with null vector $\hat{\hat{\mathbf{q}}}$, and the probability distribution approaches the form illustrated in Figure 1f. The resultant potential instability of the quadratic representation EKF can only be avoided by adding unphysical process noise to the quaternion magnitude.

The propagation equation for the state estimate is

$$\dot{\hat{\mathbf{q}}} = \frac{1}{2} \Omega(\hat{\boldsymbol{\omega}}) \hat{\mathbf{q}} + P H_Q^T (\tau R)^{-1} [\mathbf{z}_Q - \mathbf{h}_Q(\hat{\mathbf{q}})] = \frac{1}{2} \Omega(\hat{\boldsymbol{\omega}}) \hat{\mathbf{q}} - P H_Q^T (\tau R_Q)^{-1} [\mathbf{v}_B - A_Q(\hat{\mathbf{q}}) \mathbf{v}_I]. \quad (69)$$

This gives

$$\dot{\hat{\mathbf{q}}} = \frac{1}{2} \Omega(\hat{\boldsymbol{\omega}}) \hat{\mathbf{q}} + 2\tau^{-1} \sigma_z^{-2} \begin{bmatrix} \frac{1}{4} P_a & \frac{1}{2} \mathbf{p}_c \\ \frac{1}{2} \mathbf{p}_c^T & P_n \end{bmatrix} \begin{bmatrix} [\hat{\mathbf{v}}_Q \times]^T \\ \hat{\mathbf{v}}_Q^T \end{bmatrix} (\mathbf{v}_B - \hat{\mathbf{v}}_Q), \quad (70)$$

which can be written as

$$\dot{\hat{\mathbf{q}}} = \frac{1}{2} \bar{\boldsymbol{\omega}}_{eff} \otimes \hat{\mathbf{q}} + \tau^{-1} \sigma_z^{-2} [(\hat{\mathbf{v}}_Q \times \mathbf{p}_c + 2P_n \hat{\mathbf{v}}_Q) \cdot (\mathbf{v}_B - \hat{\mathbf{v}}_Q)] \hat{\mathbf{q}}, \quad (71)$$

with

$$\bar{\boldsymbol{\omega}}_{eff} \equiv \hat{\boldsymbol{\omega}} + \tau^{-1} \sigma_z^{-2} \left\{ P_a (\mathbf{v}_B \times \hat{\mathbf{v}}_Q) + 2[\hat{\mathbf{v}}_Q \cdot (\mathbf{v}_B - \hat{\mathbf{v}}_Q)] \mathbf{p}_c \right\}. \quad (72)$$

The quaternion norm satisfies

$$\frac{d}{dt} |\hat{q}|^2 = 2 \hat{q}^T \dot{\hat{q}} = 2 \tau^{-1} \sigma_z^{-2} |\hat{q}|^2 (\hat{v}_Q \times \mathbf{p}_c + 2 p_n \hat{v}_Q) \cdot (\mathbf{v}_B - \hat{v}_Q). \quad (73)$$

If \hat{q} is not a unit quaternion, $\mathbf{v}_B - \hat{v}_Q$ will have an error along \hat{v}_Q that gives rise to a correction driving the norm of \hat{q} toward unity as long as p_n is not zero. If $\mathbf{p}_c(t)$ is zero initially (and thus for all time), the angular behavior given by Eq. (72) is the same as for the ray representation EKF and MEKF; but Eq. (71) shows a non-zero norm update parallel to \hat{q} , as is expected from the norm variance. If $\mathbf{p}_c(t)$ is not zero, the attitude behavior will also differ from the ray representation EKF and MEKF, and it is difficult to avoid the conclusion that the quadratic representation EKF will give incorrect results in this case.

For unit vector measurements, the predicted body vector \hat{v}_Q is not guaranteed to be a unit vector in the quadratic measurement EKF. However, \hat{v}_Q is parallel to \mathbf{u}_0 , the unit vector in Eq. (36), so

$$H_Q^T R_{unit}^{-1} = -2 \sigma_z^{-2} |\hat{q}|^{-2} (\hat{\Xi} [\hat{v}_Q \times]^T + \sigma_z^{-2} \hat{q} \hat{v}_Q^T). \quad (74)$$

This differs from the result for vector measurements with isotropic errors

$$H_Q^T R^{-1} = -2 \sigma_z^{-2} |\hat{q}|^{-2} (\hat{\Xi} [\hat{v}_Q \times]^T + \hat{q} \hat{v}_Q^T). \quad (75)$$

The result is to multiply the (4,4) element of the matrix on the right side of Eq. (67) by σ_z^{-2} for unit vector measurements. Both \dot{p}_n and $d|\hat{q}|^2/dt$ gain an extra factor of σ_z^{-2} when \mathbf{p}_c is zero. Since σ_z can be very small, the approach to quaternion normalization and also to covariance matrix singularity can be much more rapid for unit vector measurements than for measurements of vectors with isotropic errors. The explicit time dependence of p_n for $\mathbf{p}_c(t) \equiv \mathbf{0}$ and $|\hat{v}_Q|$ constant is

$$p_n(t) = \begin{cases} \left[p_n^{-1}(0) + 4 \sigma_z^{-2} |\hat{v}_Q|^2 (t/\tau) \right]^{-1} & \text{for vectors with isotropic errors} \\ \left[p_n^{-1}(0) + 4 \sigma_z^{-4} |\hat{v}_Q|^2 (t/\tau) \right]^{-1} & \text{for unit vectors.} \end{cases} \quad (76)$$

Note that the convergence of $|\hat{q}|$ to unity and of p_n to zero is a result purely of measurement processing.

Neither “brute force” normalization of the quaternion nor norm-enforcing pseudo-measurements is required, in principle. Such subterfuges have been found useful in the presence of norm errors arising from numerical and discretization effects in actual implementations, however [13–15].

The Linear Measurement EKF

Choukroun, Bar-Ithack, and Oshman formulated a measurement model that is linear in the quaternion [31]:

$$\mathbf{z} = \mathbf{h}_L(\bar{q}) + \mathbf{n}_L = \Xi(\bar{q}) \mathbf{v}_B - \Psi(\bar{q}) \mathbf{v}_I + \Xi(\bar{q}) \mathbf{n}_z, \quad (77)$$

where

$$\Psi(\bar{q}) \equiv \begin{bmatrix} q_4 I_{3 \times 3} - [\mathbf{q} \times] \\ -\mathbf{q}^T \end{bmatrix}, \quad (78)$$

Equations (41)–(43) and the identity

$$\Xi^T(\bar{q}) \Psi(\bar{q}) = A_Q(\bar{q}) = |\bar{q}|^2 A_R(\bar{q}) \quad (79)$$

indicate that this model is closely related to the ray representation EKF.

It is easily seen that

$$\mathbf{h}_L(\bar{q}) = \Xi(\bar{q}) \mathbf{v}_B - \Psi(\bar{q}) \mathbf{v}_I = \begin{bmatrix} -[(\mathbf{v}_B + \mathbf{v}_I) \times] & (\mathbf{v}_B - \mathbf{v}_I) \\ -(\mathbf{v}_B - \mathbf{v}_I)^T & 0 \end{bmatrix} \bar{q} \equiv H_L \bar{q}. \quad (80)$$

In practice, we must use the predicted measurements to compute the sensitivity matrix. Substituting $\mathbf{v}_B = \hat{\mathbf{v}}_R$ and $\mathbf{v}_I \equiv A_R^T(\hat{q})\hat{\mathbf{v}}_R$ into the above equation gives, after some algebra,

$$H_L = -2|\hat{q}|^{-2} \hat{\Xi}[\hat{\mathbf{v}}_R \times] \hat{\Xi}^T = \hat{\Xi} H_R. \quad (81)$$

We also use \hat{q} in the measurement noise covariance, giving the singular measurement covariance

$$R_L = \hat{\Xi} E\{\mathbf{n}_z \mathbf{n}_z^T\} \hat{\Xi}^T = \sigma_z^2 \hat{\Xi} \hat{\Xi}^T = \sigma_z^2 \lim_{\varepsilon \rightarrow 0} (\hat{\Xi} \hat{\Xi}^T + \varepsilon \hat{q} \hat{q}^T). \quad (82)$$

However, Eqs. (41)–(43) and (81) give

$$H_L^T R_L^{-1} = H_R^T \hat{\Xi}^T \left[\sigma_z^{-2} |\hat{q}|^{-4} \lim_{\varepsilon \rightarrow 0} (\hat{\Xi} \hat{\Xi}^T + \varepsilon^{-1} \hat{q} \hat{q}^T) \right] = \sigma_z^{-2} |\hat{q}|^{-2} H_R^T \hat{\Xi}^T. \quad (83)$$

Comparing with Eqs. (28) and (54) shows that this model gives the same covariance and state propagation as the ray representation EKF. It is not difficult to see that this equivalence also holds for unit vector measurements. The apparent simplicity of the linear dependence of the measurements on the state vector given by Eq. (77) is compensated by the appearance of state-dependent measurement noise, and the result is mathematically equivalent to ray representation EKF with its nonlinear measurement model.

DISCUSSION

Of the four methods considered in this paper, the MEKF requires the least computational effort due to the lower dimensionality of its state vector and covariance matrix. The MEKF is also the most satisfying conceptually, since it respects the dimensionality of the rotation group, and its attitude estimate is a unit quaternion by definition.

The ray representation EKF and the equivalent linear measurement EKF give the same attitude covariance and attitude estimates as the MEKF, at the cost of some additional computational burden. They are conceptually simpler than the MEKF, since they require no attitude parameterization other than the quaternion. Some numerical significance issues may arise from the unobservable norm degree of freedom of the quaternion, but these have not caused difficulties in practical applications.

The quadratic representation EKF rests on the least secure foundation. We have shown that the quaternion approaches unit norm, and thus the attitude matrix approaches an orthogonal matrix, as a natural result of measurement processing, without resorting to “brute force” normalization of the quaternion or norm-enforcing pseudo-measurements. However, the quaternion norm variance and attitude-norm correlations tend asymptotically to zero for this filter, so the 4×4 covariance matrix becomes singular. This singularity, which may lead to instability of the Kalman filter, can only be avoided by adding unphysical process noise to the quaternion magnitude. The singularity is approached more rapidly for unit vector measurements than for vector measurements with isotropic errors.

If the initial quaternion estimate is an eigenvector of the initial 4×4 quaternion covariance in the quadratic representation EKF, the ray representation EKF, or the linear measurement EKF, the correlation between the attitude errors and quaternion norm errors will be zero initially and for all time. For this initialization, the time dependence of the attitude covariance in these three filters is the same as the MEKF, independent of the quaternion norm variance. If the correlation between the attitude errors and quaternion norm errors is not zero, the attitude covariance of the quadratic representation EKF differs from that of the other filters. Also, the attitude matrix provided by quadratic representation EKF is not guaranteed to be orthogonal, so the estimates provided by this method should be regarded with caution.

The motivation for considering the four-component quaternion EKF's despite the conceptual and computational advantages of the MEKF appears to be the superficial resemblance of a quaternion EKF to a linear Kalman filter. This resemblance is deceiving, because the process noise and any dynamic parameters to be estimated enter multiplicatively rather than additively in the quaternion kinematics equation. Most measurement models are also inherently nonlinear, since we have seen that the apparent simplicity of the linear measurement model is compensated by the appearance of state-dependent measurement noise, and the result is mathematically equivalent to ray representation EKF with its nonlinear measurement model. Thus we see no valid reason to prefer a quaternion estimator to the MEKF.

ACKNOWLEDGEMENTS

I would like to acknowledge several important communications with Mark Psiaki, Mark Pittelkau, and Malcolm Shuster.

REFERENCES

1. Gelb, Arthur, ed., *Applied Optimal Estimation*, Cambridge, MA, the MIT Press, 1974
2. Brown, R. G., and Hwang, P. Y. C., *Introduction to Random Signals and Applied Kalman Filtering*, Third Edition, John Wiley & Sons, New York, pp. 343–347, 1994
3. 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, pp. 51–56, 1984
4. Kasdin, N. J. and Weaver, T. J. M. "Recursive Satellite Attitude Estimation with the Two-Step Optimal Estimator," submitted to the *Journal of Guidance, Control, and Dynamics*,
5. Farrell, J. L., "Attitude Determination by Kalman Filtering," *Automatica*, Vol. 6, 1970, pp. 419–430
6. Idan, M. "Estimation of Rodrigues Parameters from Vector Observations," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-32, No. 2, pp. 578–586, 1996
7. Crassidis, J. L., and Markley, F. L. "Attitude Estimation Using Modified Rodrigues Parameters," *Flight Mechanics/Estimation Theory Symposium 1996*, Goddard Space Flight Center, Greenbelt, MD, NASA Conference Publication 3333, May 1996, pp. 71–83
8. Stuelpnagel, John, "On the Parameterization of the Three-Dimensional Rotation Group," *SIAM Review*, Vol. 6, No. 4, pp. 422–430, 1964
9. Paulson, D. C., Jackson, D. B., and Brown, C. D., "SPARS Algorithms and Simulation Results," *Proceedings of the Symposium on Spacecraft Attitude Determination*, Aerospace Corp. Report TR-0066 (5306)-12, Vol. 1, pp. 293–317, Sept.–Oct. 1969.
10. Toda, N. F., Heiss, J. L., and Schlee, F. H., "SPARS: the System, Algorithm, and Test Results," *Proceedings of the Symposium on Spacecraft Attitude Determination*, Aerospace Corp. Report TR-0066 (5306)-12, Vol. 1, pp. 361–370, Sept.–Oct. 1969
11. Murrell, James W., "Precision Attitude Determination for Multimission Spacecraft," AIAA Paper 78-1248, *AIAA Guidance and Control Conference*, Palo Alto, CA, August 1978
12. 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, pp. 417–429, 1982
13. Bar-Itzhack, I. Y., and Oshman, Y. "Attitude Determination from Vector Observations: Quaternion Estimation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-21, pp. 128–136, 1985
14. Bar-Itzhack, I. Y., Deutschmann, J., and Markley, F. L., "Quaternion Normalization in Additive EKF for Spacecraft Attitude Determination," AIAA Paper 91-2706, *AIAA Guidance, Navigation, and Control Conference*, New Orleans, LA, August 1991

15. Deutschmann, J., Markley, F. L., and Bar-Itzhack, I., "Quaternion Normalization in Spacecraft Attitude Determination," *Flight Mechanics/Estimation Theory Symposium*, NASA Conference Publication 3186, Goddard Space Flight Center, Greenbelt, MD, May 1992
16. Shuster, Malcolm D., "The Quaternion in the Kalman Filter," *Advances in the Astronautical Sciences*, Vol. 85, pp. 25–37, 1993
17. Pittelkau, M. E., "An Analysis of the Quaternion Attitude Determination Filter," AAS Paper 03-194, *AAS/AIAA Spaceflight Mechanics Meeting*, Ponce, Puerto Rico, February 2003
18. Markley, F. L., "Attitude Error Representations for Kalman Filtering," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 2, pp. 311–317, 2003
19. Wertz, James R., ed., *Spacecraft Attitude Determination and Control*, Dordrecht, Holland, D. Reidel, 1978
20. Shuster, Malcolm D., "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, pp. 439–517, 1993
21. Euler, L., "Problema Algebraicum Ob Affectiones Prorsus Singulares Memorabile," *Novi Commentarii Academiae Scientiarum Petropolitanae*, Vol. 15, Section 33, p. 101, 1770
22. Gauss, C. F., *Werke*, Vol. VIII, pp. 357–362, Göttingen, *Königliche Gesellschaft der Wissenschaften*, 1900
23. Rodrigues, O., "Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire," *Journal de Mathématiques*, Vol. 5, pp. 380–440, 1840
24. Hamilton, W. R., "On Quaternions; or a New System of Imaginaries in Algebra," *Philosophical Magazine*, 3rd Series, Vol. 25, pp. 489–495, 1844
25. Altmann, Simon L., "Hamilton, Rodrigues, and the Quaternion Scandal," *Mathematics Magazine*, Vol. 62, No. 5, pp. 291–308, 1989
26. Curtis, Morton L., *Matrix Groups*, 2nd ed., New York, NY, Springer-Verlag, 1984, Chapter 5
27. Gibbs, J. W., *Scientific Papers*, Volume II, Dover, New York, pp. 65, 66, 1961
28. Gray, C. W., "Star Tracker/IRU Attitude Determination Filter," *Guidance and Control 2001, Advances in the Astronautical Sciences*, Vol. 107, Paper AAS 01-039, Univelt, San Diego, pp. 459–476, 2001
29. Psiaki, M.L., Theiler, J., Bloch, J., Ryan, S., Dill, R.W., and Warner, R.E., "ALEXIS Spacecraft Attitude Reconstruction with Thermal/Flexible Motions Due to Launch Damage," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 5, pp. 1033–1041, 1997
30. Psiaki, M.L., Klatt, E.M., Kintner, P.M. Jr., and Powell, S.P., "Attitude Estimation for a Flexible Spacecraft in an Unstable Spin," *Journal of Guidance, Control, and Dynamics*, Vol. 25, No. 1, pp. 88–95, 2002
31. Choukroun, D., Bar-Itzhack, I. Y., and Oshman, Y., "A Novel Quaternion Filter," AIAA Paper 2002-4460, *AIAA Guidance, Navigation, and Control Conference*, Monterey, CA, August 2002