Abstract

The absence of a globally nonsingualar three-parameter representation of rotations forces attitude Kalman filters to estimate either a singular or a redundant attitude representation. We compare two filtering strategies using simplified kinematics and measurement models. Our favored strategy estimates a three-parameter representation of attitude deviations from a reference attitude specified 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. We point out some disadvantages of the other strategy, which directly estimates the four-parameter quaternion representation.

Introduction

Real-time spacecraft attitude estimation generally employs an Extended Kalman Filter (EKF) [1, 2]. Although the $3 \times 3$ 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, most EKFs use lower-dimensional parameterizations of $SO(3)$, with the earliest using a minimal three-dimensional parameterization [3]; but higher-dimensional parameterizations can avoid the singularities or discontinuities present in all three-parameter representations [4]. The four-component 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 face the dilemma of using an attitude representation that is either singular or redundant. Our preferred strategy for evading this dilemma uses a nonsingular representation for a reference attitude and a three-component representation for the deviations from this reference. This method, which we refer to as the Multiplicative EKF (MEKF), was implicit in some early spacecraft attitude estimators [5–8] and has been discussed in detail in [9]. An alternative strategy, the Additive EKF (AEKF), treats the four components of the quaternion as independent parameters [10].

We begin with a brief overview of quaternion estimation, emphasizing the conceptual difficulties. The MEKF is then discussed in a model with simplified kinematics and measurement models. The analysis is simplified by assuming the measurements to be spaced closely enough that they can be treated as continuous. The MEKF discussion is followed by an analysis of two versions of the AEKF, using the same model as the MEKF, and by a summary of our conclusions. Reference [11] is an extended version of this paper, and [12] presents a parallel analysis for the case of discrete measurements.
Quaternion estimation

The attitude matrix is generally written as a homogeneous quadratic function of the quaternion [13],

\[ A(q) = A \begin{bmatrix} q & q_4 \end{bmatrix} = (q_4^2 - |q|^2)I + 2qq^T - 2q_4[q \times], \] (1)

where \( I \) denotes the \( 3 \times 3 \) identity matrix, and the cross product matrix is

\[ [q \times] = \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix}. \] (2)

We use the quaternion product convention of References 8 and 13 so that

\[ A(p \odot q) = A \begin{bmatrix} p_4 & p \end{bmatrix} \odot A \begin{bmatrix} q & q_4 \end{bmatrix} = A \begin{bmatrix} p_4q + q_4p - p \times q \\ p_4q_4 - p \cdot q \end{bmatrix} = A(p)A(q). \] (3)

A conventional “additive” quaternion filter defines the estimate \( \hat{q} \) and error \( \Delta q \) by

\[ \hat{q} = E\{q\} \quad \text{and} \quad \Delta q = q - \hat{q}, \] (4)

where \( E\{\cdot\} \) denotes the expectation. This means that that

\[ E\{|q|^2\} = E\{|\hat{q} + \Delta q|^2\} = |\hat{q}|^2 + E\{|\Delta q|^2\} \geq |\hat{q}|^2, \] (5)

with equality only if \( \Delta q \) is identically zero. Equation (1) gives an orthogonal matrix only if the quaternion has unit norm. Equation (5) shows that if the random variable \( q \) has unit norm and is not error-free, the norm of its expectation must be less than unity. Although the unity norm constraint violation will be on the order of the variance of the attitude errors, we regard this as a serious conceptual problem with additive quaternion EKFs, so we turn first to the MEKF.

Multiplicative EKF

The MEKF represents the attitude as the quaternion product

\[ \bar{q} = \delta \bar{q}(a) \otimes \bar{q}_{ref}, \] (6)

where \( \bar{q}_{ref} \) is some unit reference quaternion and the rotation \( \delta \bar{q}(a) \) from \( \bar{q}_{ref} \) to the true attitude \( \bar{q} \) is parameterised by a three-component vector \( a \). Although several choices for \( a \) are possible [9], it is only important for this paper that for small rotations

\[ \delta \bar{q}(a) = \overline{1} + \frac{1}{2} a + \text{order}(|a|^2), \] (7)

where \( \overline{1} \) denotes the identity quaternion and \( \overline{a} \) denotes a quaternion with vector part \( a \) and scalar part 0. The two attitude representations \( \delta \bar{q} \) and \( \bar{q}_{ref} \) in Eq. (6) are clearly redundant. The basic idea of the MEKF is to estimate the three-vector \( a \) while using the correctly normalized four-component \( \bar{q}_{ref} \) to provide a globally nonsingular attitude representation. Given an estimate \( \hat{a} = E\{a\} \) of \( a \), Eq. (6) says
that \( \delta \overline{q}(\hat{a}) \otimes q_{ref} \) is the corresponding estimate of the true attitude quaternion \( \overline{q} \). We remove the redundancy in the attitude representation by choosing the reference quaternion \( \overline{q}_{ref} \) so that \( \hat{a} \) is identically zero, which avoids any singularity or discontinuity of the three-component representation. Since \( \delta \overline{q}(0) \) is the identity quaternion, this choice results in the reference quaternion being the estimate of the true quaternion. The identification of \( \overline{q}_{ref} \) as the attitude estimate means in turn that \( a \) is a three-component representation of the attitude error, providing a consistent treatment of the attitude error statistics with the covariance of \( a \) representing the covariance of the attitude error in the body frame. The fundamental conceptual advantage of the MEKF is that \( \overline{q}_{ref} \) is a unit quaternion by definition.

**Kinematics**

The quaternion kinematic equation is

\[
\dot{\overline{q}} = \frac{1}{2} \hat{\omega} \otimes \overline{q}, \tag{8}
\]

where \( \hat{\omega} \) is the angular velocity vector in the body frame. Since \( \overline{q}_{ref} \) is also a unit quaternion, it must obey

\[
\dot{\overline{q}}_{ref} = \frac{1}{2} \hat{\omega}_{ref} \otimes \overline{q}_{ref}, \tag{9}
\]

where \( \hat{\omega}_{ref} \), the angular velocity of the reference attitude, must be determined by the requirement that \( \hat{a} \) be identically zero. Note that \( \overline{q}_{ref} \) and \( \hat{\omega}_{ref} \) are not random variables. Computing the time derivative of Eq. (6), using Eqs. (8) and (9), gives

\[
\delta \overline{q} = \frac{1}{2} (\hat{\omega} \otimes \delta \overline{q} - \delta \overline{q} \otimes \hat{\omega}_{ref}), \tag{10}
\]

after right-multiplying the entire equation by the inverse of \( \overline{q}_{ref} \) and rearranging.

We now assume the simple kinematic model

\[
\omega(t) = \hat{\omega}(t) + n_\omega(t), \tag{11}
\]

where \( \hat{\omega}(t) \) is the nominal angular velocity, and the zero-mean Gaussian process noise \( n_\omega(t) \) obeys

\[
E\{n_\omega(t)n_\omega^T(t')\} = \delta(t - t')Q = \delta(t - t')\sigma_\omega^2I, \tag{12}
\]

with \( \delta(t - t') \) denoting the Dirac delta function. Substituting Eq. (10) into the time derivative of Eq. (7) and linearising in \( a \) and \( n_\omega \) gives

\[
\dot{a} = \dot{\omega} - \omega_{ref} + F_a a + G_a n_\omega \tag{13}
\]

with

\[
F_a = -\frac{1}{2} [(\hat{\omega} - \omega_{ref}) \times] \quad \text{and} \quad G_a = I. \tag{14}
\]

**Measurements**

We model the measurement of a vector \( v_B \) in the body coordinate frame by

\[
z = v_B = h(a) + n_z = A(\overline{q}) v_I + n_z, \tag{15}
\]
where \( \mathbf{v}_I \) is the corresponding vector in the inertial frame, and the zero-mean Gaussian measurement noise \( \mathbf{n}_z \) is assumed to be isotropic,

\[
R = E\{\mathbf{n}_z \mathbf{n}_z^T\} = \sigma_z^2 I.
\]

Substituting Eqs. (1), (3), (6), and (7) into Eq. (15) and linearising in \( \mathbf{a} \) gives

\[
\mathbf{h}(\mathbf{a}) = \hat{\mathbf{v}}_B + H_a \mathbf{a},
\]

where \( \hat{\mathbf{v}}_B \) is the predicted measurement in the body frame,

\[
\hat{\mathbf{v}}_B \equiv \mathbf{A}(\overline{q}_{\text{ref}}) \mathbf{v}_I,
\]

and the measurement sensitivity matrix is

\[
H_a = [\hat{\mathbf{v}}_B \times].
\]

The attitude error estimate \( \hat{\mathbf{a}} \) is propagated by the expectation of Eq. (13) with \( \hat{\mathbf{a}} = \mathbf{0} \), conditioned on the measurements, which are assumed to be continuous [1],

\[
\hat{\mathbf{a}} = \hat{\omega} - \omega_{\text{ref}} + P_a H_a^T R^{-1} [\mathbf{z} - \mathbf{h}(\mathbf{0})] = \hat{\omega} - \omega_{\text{ref}} - \sigma_z^{-2} P_a \hat{\mathbf{v}}_B \times \mathbf{v}_B,
\]

where \( P_a \) is the covariance of the error vector \( \mathbf{a} \). This is propagated using the approximation \( \omega_{\text{ref}} \approx \hat{\omega} \), which neglects only terms of higher order than are customarily retained in an EKF and gives

\[
\hat{P}_a \hat{Y} = F_a P_a + P_a F_a^T + G_a Q G_a^T - P_a H_a^T H_a P_a
\]

\[
= -[\hat{\omega} \times] P_a - P_a [\hat{\omega} \times]^T + \sigma_a^2 I - \sigma_z^{-2} P_a [\hat{\mathbf{v}}_B \times][\hat{\mathbf{v}}_B \times] P_a.
\]

Since the MEKF requires that \( \hat{\mathbf{a}} \) be identically zero, Eq. (20) gives \( \omega_{\text{ref}} \) as

\[
\omega_{\text{ref}} = \hat{\omega} + \sigma_z^{-2} P_a (\mathbf{v}_B \times \hat{\mathbf{v}}_B).
\]

The first term embodies the assumed kinematics, while the second incorporates the measurement updates. Eqs. (9) and (22) preserve the unity norm of the reference quaternion. No reset operation is needed with continuous measurements, in contrast to the discrete measurement case [9]. We emphasize that the MEKF does not require integration of Eq. (20), which is only used to derive Eq. (22).

**Additive EKF**

The AEKF relaxes the quaternion normalization condition and treats the four components of the quaternion as independent parameters, with the quaternion estimate and error given by Eq. (4). Linearising Eqs. (8) and (11) gives the error vector kinematics

\[
\Delta \hat{\mathbf{q}} = \frac{1}{2} (\hat{\mathbf{w}} + \mathbf{n}_\omega) \otimes (\hat{\mathbf{q}} + \Delta \mathbf{q}) - \frac{1}{2} \hat{\mathbf{w}} \otimes \hat{\mathbf{q}} = F_q \Delta \mathbf{q} + G_q \mathbf{n}_\omega,
\]

with

\[
F_q = \frac{1}{2} \begin{bmatrix} -[\omega \times] & \omega \\ -\omega^T & 0 \end{bmatrix} \quad \text{and} \quad G_q = \frac{1}{2} \begin{bmatrix} \hat{\mathbf{q}}_4 I + [\hat{\mathbf{q}} \times] \\ -\hat{\mathbf{q}}^T \end{bmatrix} = \frac{1}{2} \hat{\Omega}.
\]
The quaternion in an AEKF has uncertainties in both angle and norm. Angle uncertainties have the form given by Eq. (6), except that the role of $\overline{q}_{\text{ref}}$ in the MEKF is taken by $\hat{q}$ in an AEKF. Norm errors are parallel to $\hat{q}$, so we can write

$$\overline{q} = \delta q(a) \otimes \hat{q} + \beta \hat{q} \approx \left[ \begin{array}{c} \hat{\Xi} \\ \hat{q} \end{array} \right] \left[ \begin{array}{c} \frac{a}{2} \\ 1 \end{array} \right] + \beta \hat{q} = \hat{q} + \left[ \begin{array}{c} \hat{\Xi} \\ \hat{q} \end{array} \right] \left[ \begin{array}{c} \frac{a}{2} \\ \beta \end{array} \right].$$  \hspace{1cm} (25)

The approximate equality follows from the product rule and Eq. (7). Note that the norm error $\beta$ is a relative error (multiplying the norm of $\hat{q}$ by $1 + \beta$) rather than an absolute error (adding $\beta$ to the norm of $\hat{q}$). Then the $4 \times 4$ quaternion covariance is

$$P_q = E\{\Delta \overline{q} \Delta \overline{q}^T\} = \left[ \begin{array}{cc} \hat{\Xi} & \hat{q} \\ \hat{q}^T & \frac{1}{4} P_a + \frac{1}{2} P_c + P_n \end{array} \right],$$ \hspace{1cm} (26)

where $P_a$ is the $3 \times 3$ attitude covariance, $P_n$ is the quaternion norm variance, and $P_c$ is the vector of attitude/norm covariances. This factorization is not used in unconstrained quaternion estimators, which directly estimate the four-component $\Delta \overline{q}$ and the $4 \times 4 P_q$, but it illuminates the relation of the AEKF to the MEKF.

**Measurements**

The AEKF measurement model is very similar to the model in the MEKF, given by Eq. (15). The AEKF uses a different parameterization of the attitude matrix $A(\overline{q})$, however, and $h$ is regarded as a function of the quaternion $\overline{q}$ rather than of the attitude error vector $a$. We will consider two different implementations of the AEKF. The first, which we call the quadratic AEKF, uses Eq. (1) for $A(\overline{q})$. We have seen that this attitude matrix is not orthogonal unless the quaternion has unit norm, but we will show that these measurements drive the norm to unity in the quadratic AEKF. The second implementation uses the orthogonal attitude matrix formed by using the normalized quaternion $\overline{q}/|\overline{q}|$ in Eq. (1)

$$A_R(\overline{q}) \equiv A(\overline{q}/|\overline{q}|) = |\overline{q}|^{-2} A(\overline{q}).$$ \hspace{1cm} (27)

We refer to this as the ray representation AEKF because any non-zero quaternion along a ray in quaternion space (a straight line through the origin) represents the same attitude. The twofold ambiguity of the unit quaternion representation corresponds to the two points where the ray pierces the unit sphere. After a good deal of algebra, we find the measurement sensitivity matrix

$$H_q = \frac{\partial h}{\partial \overline{q}}|_{\overline{q}} = 2|\overline{q}|^{-2} \left( [\hat{\nu}_B^T \times] \hat{\nu}_B^T + k \hat{\nu}_B^T \right),$$ \hspace{1cm} (28)

where $\hat{\nu}_B$ is the predicted measurement given by Eq. (18) with $\overline{q}_{\text{ref}}$ replaced by $\hat{q}$, and where $k = 1$ for the quadratic AEKF and $k = 0$ for ray representation AEKF. The form of the measurement sensitivity matrix implies that quaternion norm errors, which are parallel to $\hat{q}$, are observable in the quadratic AEKF, but not in the ray representation AEKF.
The expectation of Eq. (8) conditioned on the measurements is
\[ \hat{\dot{q}} = \frac{i}{2} \hat{\omega} \otimes \hat{q} + P_q H_q \trans [z - h(\hat{q})] \]
\[ = \frac{i}{2} \hat{\omega}_{\text{eff}} \otimes \hat{q} + \sigma_z^2 \{ (\hat{v}_B \times \hat{p}_c) + 2k_2 \hat{v}_B \cdot (\hat{v}_B - \hat{v}_B) \} \hat{q}, \]
where \( \omega_{\text{eff}} \equiv \hat{\omega} + \sigma_z^2 (v_B \times v_B) + 4k_2 \hat{v}_B \cdot (v_B - \hat{v}_B) \).

The term parallel to \( \hat{\dot{q}} \) on the right side of Eq. (29) modifies the norm of the quaternion estimate, but it is not difficult to show that
\[ d(\hat{q}^T \hat{q})/dt = \frac{i}{2} \hat{\omega}_{\text{eff}} \otimes (\hat{q}^T \hat{q}) \].

Since the ray representation AEKF \( (k = 0) \) uses the normalized quaternion to compute the attitude via Eq. (27) and since its \( \omega_{\text{eff}} \) is the same as \( \omega_{\text{ref}} \) in the MEKF, it gives the same attitude estimate as the MEKF. The ray representation AEKF has been used to estimate the attitude of the ALEXIS and CAPER spacecraft [14, 15].

The quadratic AEKF \( (k = 1) \) is different. Since the attitude matrix depends on the quaternion norm in this representation, the term parallel to \( \hat{\dot{q}} \) on the right side of Eq. (29) is significant. If \( \hat{\dot{q}} \) is not a unit quaternion and \( P_n \) is not zero, \( v_B - \hat{v}_B \) will have an error along \( \hat{v}_B \) that drives the norm of \( \hat{q} \) toward unity. The angular update given by Eq. (30) is the same as for the ray representation AEKF and MEKF if \( p_c \) is zero, but is different if \( p_c \) is not zero. It is difficult to avoid the conclusion that the quadratic representation AEKF gives incorrect estimates in the latter case.

Substituting Eqs. (26) and (28) into the covariance propagation equation
\[ \hat{\dot{P}}_a = F_a P_a + P_a F_a^\trans + G_q Q G_q^\trans - P_q H_q \trans H_q P_q. \]
gives, after some algebra,
\[ \hat{\dot{P}}_a = -[\hat{\omega} \times]P_a - P_a[\hat{\omega} \times]^\trans + \sigma_z^2 I - \sigma_z^2 (P_a[\hat{v}_B \times] [\hat{v}_B \times] P_a + 4k_2 \hat{v}_B^2 p_c^T), \]
\[ \hat{\dot{p}}_c = -\hat{\omega} \times p_c - \sigma_z^2 (P_a[\hat{v}_B \times] [\hat{v}_B \times] + 4k_2 \hat{v}_B^2 p_a) p_c, \]
\[ \hat{\dot{p}_n} = -\sigma_z^2 (\hat{v}_B \times p_c)^2 + 4k_2 \hat{v}_B^2 p_c^2. \]

Eq. (33a) shows that the 3x3 attitude covariance \( P_a \) obeys the same equation in the ray representation AEKF \( (k = 0) \) as in the MEKF. Eq. (33b) shows that \( p_c \) is zero for all time in either AEKF if its initial value is zero. Eq. (26) shows that this is the case if the initial estimate \( \hat{q} \) is an eigenvector of the initial covariance; this includes the case that the initial \( P_q \) is a multiple of the identity. The measurement term in Eq. (33b) will drive \( p_c \) to zero if it is not initially zero. In the ray representation AEKF, the norm variance \( p_n \) is constant if \( p_c \) is zero, which may cause loss of significance of the attitude covariance if the attitude estimate converges to small errors, since \( P_a \) is mixed with \( p_n \) in the quaternion covariance \( P_q \).
The attitude covariance $P_n$ in the quadratic AEKF ($k = 1$) is different from the other two filters unless $p_c$ is zero. The last term in Eq. (33c) drives $p_n$ to zero in the quadratic AEKF, so the $4 \times 4$ covariance matrix becomes singular with null vector $\hat{q}$. The resultant potential instability of the quadratic AEKF can be avoided only by adding unphysical process noise to the quaternion magnitude. Note that the convergence of $\left| \hat{q} \right|$ to unity and of $p_n$ to zero is a result purely of measurement processing, requiring neither “brute force” normalization of the quaternion nor norm-enforcing pseudo-measurements. Such subterfuges have been found useful in the presence of norm errors arising from numerical and discretization effects in AEKF implementations, however [10].

**Discussion**

Of the methods considered in this paper, the MEKF requires the least computational effort due to the lower dimensionality of its 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 AEKF is conceptually simpler than the MEKF, since it requires no attitude parameterization other than the quaternion, and it gives the same attitude estimates and covariance at the cost of some additional computational burden. Numerical significance issues may arise from the unobservable quaternion norm degree of freedom, but these have not caused difficulties in applications.

The quadratic AEKF rests on less secure foundations. 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. If the initial quaternion estimate is an eigenvector of the initial $4 \times 4$ quaternion covariance, the attitude errors and quaternion norm errors are uncorrelated for all time, and the attitude covariance and attitude estimates are the same as the other filters. If the correlation between the attitude errors and quaternion norm errors is not zero, the attitude update and attitude covariance differ from those of the other filters. The estimates provided by this method should be regarded with caution in either case, since the attitude matrix is not exactly orthogonal and the $4 \times 4$ covariance matrix becomes singular, potentially leading to instability that can be avoided only by adding unphysical process noise to the quaternion magnitude.

The motivation for considering the additive EKFs, despite the conceptual and computational advantages of the MEKF, appears to be their superficial resemblance 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. Realistic measurement models are also nonlinear, so we see no valid reason to prefer an AEKF to the MEKF.

**Acknowledgements**

I would like to acknowledge several important exchanges with Russell Carpenter, Mark Psiaki, Mark Pittelkau, and Malcolm Shuster.
References


