AIAA 2002-4831

ON THE SINGULARITY IN THE ESTIMATION OF THE QUATERNION-OF-ROTATION

Itzhack Y. Bar-Itzhack[1] and Julie K. Thienel[2]

ABSTRACT

It has been claimed in the archival literature that the covariance matrix of a Kalman filter, which is designed to estimate the quaternion-of-rotation, is necessarily rank deficient because the normality constraint of the quaternion produces dependence between the quaternion elements. In reality, though, this phenomenon does not occur. The covariance matrix is not singular, and the filter is well behaved. Several simple examples are presented that demonstrate the regularity of the covariance matrix. First, a Kalman filter is designed to estimate variables subject to a functional relationship. Then the particular problem of quaternion estimation is analyzed. It is shown that the discrepancy stems from the fact that the functional relationship exists between the elements of the quaternion but not between its estimated elements.

I. INTRODUCTION

The quaternion-of-rotation is a four-element parameterization of attitude, and since the quaternion is normal, one element is redundant. This fact has brought researchers [see e.g. Ref. 1] to the conclusion that when a Kalman filter (KF) is used to estimate all four parameters of the quaternion-of-rotation, the filter covariance matrix is necessarily singular. The argument behind this assertion is that the dependent variables cause singularities. However, as will be shown in the ensuing discussion, the assertion that the covariance matrix is singular when all four elements of the quaternion are estimated is not necessarily true, and if it happens to be singular, it is not because of the quaternion normality.

We note that there are two principal approaches to the application of an extended KF (EKF) to quaternion-of-rotation estimation; namely, the *multiplicative approach* that yields the multiplicative EKF (MEKF) and the *additive approach* that yields the additive EKF (AEKF). In the MEKF the difference between the estimated and the true quaternion is defined as a quaternion-of-rotation between the true coordinate system, and the estimated coordinate system. In the estimation process the components of this difference quaternion are estimated, and are then used to update the a-priori estimate of the full quaternion. Because the difference itself is defined as a quaternion, this update is performed through a quaternion multiplication [1, 2] hence the name *multiplicative approach*. Since this difference is a quaternion, its length is unity and, therefore, one of its components is a deterministic function of the other three.

In the AEKF [3], the difference between the true and estimated quaternions is defined as a simple subtraction of one vector from the other. Using this approach the difference is estimated and then the estimate is added to the a-priori full quaternion estimate hence the name *additive approach*. This *vector* difference does not constitute a quaternion; therefore, its length is not necessarily equal to unity. In fact, if the quaternion and its estimate are close to one another, then surely the difference quaternion is not of unit length. Naturally, the unity constraint is not imposed on the elements of the difference quaternion, which become part of the estimated state vector. Therefore, the corresponding covariance matrix is not inherently singular.

In this paper we show three realities. First we show that even if there is a functional relationship between the *true* values of states, this relationship does not necessarily exist between their *estimates*; therefore, the filter-computed covariance matrix is not necessarily singular. Second we show that even if such a relationship is imposed on the estimates, the covariance is still not necessarily singular, and, third, we show that these claims hold, in particular, for the case of quaternion estimation when the additive approach is employed.

We establish these realities in an evolutionary manner. We start our presentation with a conceptual example of estimating the vertices of a rotating square. This example illustrates the first reality; namely, even if there is a functional relationship between the *true* values of states there is not necessarily a relationship between their *estimates*. Then we present a simple linear example of estimating the position of the edges of a sliding rod. Here we demonstrate *numerically* the same reality and also the second reality; namely, even if a relationship is imposed on the estimates, the covariance is still not necessarily singular. To demonstrate that these realities

[1] Sophie and William Shamban Professor of Aerospace Engineering. Technion-Israel Institute of Technology, Haifa 32000, Israel. Member Technion Asher Space Research Institute. Tel: +972-4-829-3196, Fax: +972-4-823-1848, Email: ibaritz@tx.technion.ac.il

[2] Aerospace Engineer, NASA GSFC, Code 572, Greenbelt, MD 20771. Tel: 301 286-9033, Fax: 301 286-0369, Email: julie.thienel@gsfc.nasa.gov

Copyright

exist when the quaternion is estimated, we start with a simple linear example of estimating the four elements of a quaternion when the measurements are quaternion measurements. We choose a static case where the body is not rotating. This choice enables us to analytically prove the first reality for this case. In the final example we treat the classical case where the body rotates and the measurements are vector measurements. For this case we numerically demonstrate that both realities exist when quaternions are estimated using the AEKF; namely, the filter-computed covariance matrix is non-singular, and it remains non-singular even when normality is imposed on the filter estimates in a brut force manner. To explain the results of the latter example we analyze the operation of the ordinary EKF, which is actually an AEKF, and show that the forced normalization of the estimated quaternion has no bearing on the covariance matrix.

II. ESTIMATION OF SQUARE VERTICES

We start our presentation with a conceptual example of estimating the vertices of a rotating square. This example illustrates the fact that even if there is a functional relationship between the correct values of estimated states there is not necessarily a relationship between their estimates. Consider the system described in Fig. 1 where a square is placed on a disk that turns at an angular velocity, ω. We obtain noisy measurements of the vertices of the square and try to estimate the location of these vertices on the disk. Suppose that our initial estimate places the vertices at x_1, x_2, x_3 and x_4.

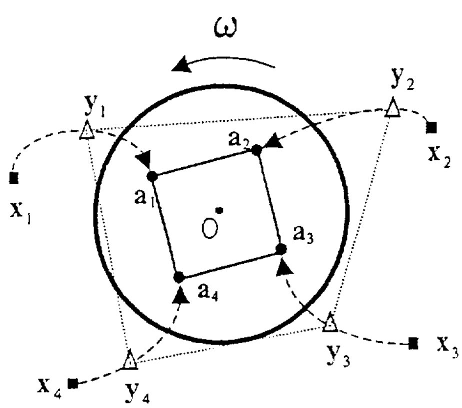


Fig. 1: Estimation of the vertices of a square

After the second measurement update, the estimates move to points y_1, y_2, y_3 and y_4 respectively and so on. The estimates keep moving along the curved trajectories until they reach an infinitesimal distance from the vertices a_1, a_2, a_3 and a_4. (Fig. 1 shows the position of the square at the end of the estimation process.) Because there is a relationship between the locations of the vertices of a square, knowing the location of three of them, say a_1, a_2 and a_3, we can find a_4, the fourth of them. This, however, does not mean that if we know y_1, y_2 and y_3 we also know y_4. As is indeed shown in Fig. 1, y_1, y_2, y_3, y_4 do not form a square. In other words, the fact that we know that there is a deterministic relationship between the four vertices of the square does not mean there is also a deterministic relationship between their estimates. Similarly, *the fact that there is a deterministic relationship between the four elements* q_1, q_2, q_3 *and* q_4 *of a quaternion does not mean there is also a deterministic relationship between their corresponding estimates* $\hat{q}_1, \hat{q}_2, \hat{q}_3$ *and* $\hat{q}_4$. Moreover, if we use the normality constraint to compute one element of the quaternion as a function of the other three estimated elements, the result will not necessarily be equal to the estimate of that element.

III. ESTIMATION OF THE EDGES OF A ROD

After having seen conceptually that there is no reason to assume that an algebraic relationship that exists between the states of a system is also carried to their estimates, we move forward to *numerically* demonstrate this fact and the assertion that even if the relationship is imposed on the estimates, the filter covariance is still not necessarily singular.

Consider the rod shown in Fig. 2. It slides along the x-axis at a constant velocity V. The coordinates of its edges are s_1 and s_2, respectively. In order to describe the equations of motion of the two edges in the state space we define the following state variables

$$x_1 = s_1 \tag{1.a}$$

$$x_2 = \dot{x}_1 = V = \text{const.} \tag{1.b}$$

$$\dot{x}_2 = 0 \tag{1.c}$$

$$x_3 = s_2 \tag{1.d}$$

$$x_4 = \dot{x}_3 = V = \text{const.} \tag{1.e}$$

$$\dot{x}_4 = 0 \tag{1.f}$$

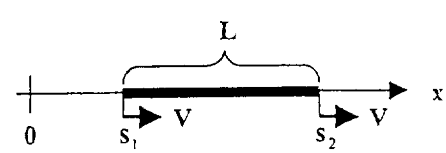


Fig. 2: Moving rod

In matrix form these equations are

$$\frac{d}{dt}\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\end{bmatrix}=\begin{bmatrix}0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} \quad (2.a)$$

and in a discrete form they are

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}_{k+1}=\begin{bmatrix}1&\Delta t&0&0\\0&1&0&0\\0&0&1&\Delta t\\0&0&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}_k \quad (2.b)$$

where Δt is the time difference between the discrete time points t_k and t_{k+1}. We assume that at the time point designated as zero, the rear end of the rod, s_1, passed by the origin, 0, at the constant velocity, V. Therefore the true initial state vector is

$$x_0=\begin{bmatrix}0\\V\\L\\V\end{bmatrix} \quad (3)$$

The measured quantities are the positions of s_1 and s_2 on the x-axis. It is assumed that these measurements are contaminated by zero mean white noise signals v_1 and v_2 respectively, thus the measurement vector is

$$\mathbf{z}_m=\begin{bmatrix}x_1+v_1\\x_3+v_2\end{bmatrix} \quad (4)$$

Using the following data

$$L=2\text{ m};\quad V=1\text{ m/sec};\quad \Delta t=1\text{ sec};\quad \sigma_{v_1}=\sigma_{v_2}=0.1\text{ m} \quad (5)$$

Eqs. (1) through (4) are used to simulate the correct state vector and the measurements. A linear Kalman filter provides estimates of the state vector. To avoid the well-known divergence phenomenon that occurs in unexcited state vector dynamics [4], zero mean white process noise is added to the dynamics equation *in the filter only.* A covariance matrix, Q_k is added to the time-propagated filter covariance matrix. The matrix is a diagonal matrix with the values

$$Q_k=\text{diag}\{10^{-1}\quad 10^{-1}\quad 10^{-1}\quad 10^{-1}\} \quad (6)$$

Our estimate of the initial state is

$$x_0=\begin{bmatrix}0+\delta_1\\V+\delta_2\\L+\delta_3\\V+\delta_4\end{bmatrix} \quad (7)$$

where δ_1 is 0.2, δ_2 is 0.1, δ_3 is -0.2, and δ_4 is -0.1. Accordingly, we set the initial covariance matrix to be the following diagonal matrix

$$P_0=\text{diag}\{(3\cdot 0.2)^2\quad(3\cdot 0.1)^2\quad(3\cdot 0.2)^2\quad(3\cdot 0.1)^2\} \quad (8)$$

The filter is run for 20 sec and in-spite of the following dependence between the states x_1 and x_2

$$x_3=x_1+L \quad (9)$$

no *singularity is observed in the filter covariance matrix.* This is evident by a simple inspection of Fig. 3 where the behavior of the filter covariance matrix eigenvalues is presented.

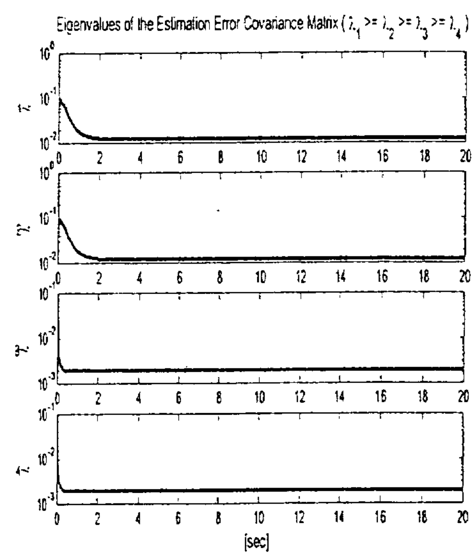


Fig. 3: Eigenvalues of the covariance matrix

It is not surprising that no singularity occurs in the filter covariance matrix because the relationship between x_3 and x_1 is not imposed in the filter model. One may speculate that imposing the relationship on the estimates $\hat{x}_3$ and $\hat{x}_1$ will cause the matrix to become singular. In order to examine this proposition, we imposed the distance constraint between the rod end-points by forcing the new a-posteriori estimates of $\hat{x}_1$ and $\hat{x}_3$ to be

$$\hat{x}_1^*(+)=\tfrac{1}{2}[\hat{x}_1(+)+\hat{x}_3(+)-L] \quad (10.a)$$

and

$$\hat{x}_3^*(+)=\tfrac{1}{2}[\hat{x}_1(+)+\hat{x}_3(+)+L] \quad (10.b)$$

(Note that as a result of this change, an additional error term must be accounted for, since the state estimates are altered. This term is included in the filter propagation and update stages according to the discussion presented in Sections VII and VIII.) No singularity in the covariance can be detected as shown in the eigenvalue history in Fig. 4.

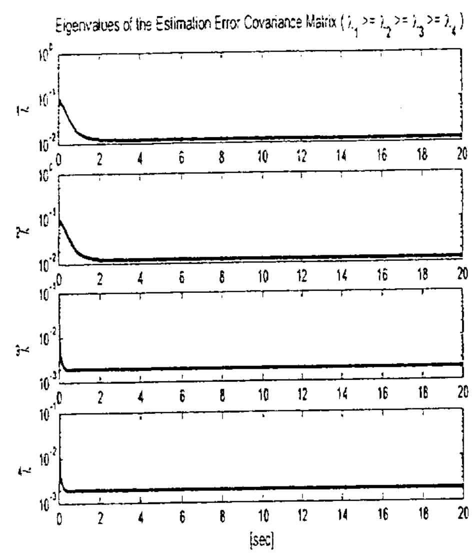


Fig. 4: Eigenvalues of the covariance matrix in the constrained case.

IV. QUATERNION ESTIMATION WITH QUATERNION MEASUREMENTS

Next in our evolutionary exposition of the singularity issue we demonstrate the lack of covariance singularity when the quaternion is estimated. We start with a simple linear example of estimating the four elements of a quaternion when the measurements are quaternion measurements. We choose a static case where the body is not rotating. This choice enables us to prove *analytically* that the filter-computed covariance matrix of the non-normalized quaternion estimate for this case is not singular.

Consider a rigid satellite body, fixed in inertial space. The system equations that describe this simple, static case are

$$q_k = I q_{k-1} \tag{11.a}$$

$$q_{m,k} = I q_k + v_k \tag{11.b}$$

where I is the fourth order identity matrix, and $q_{m,k}$ is a measurement of the quaternion at time, t_k. This system is linear; therefore, the ordinary KF can be applied in estimating $\mathbf{q}$. Consequently a covariance analysis can be carried out which is independent of the state or of its estimate. Let us assume that

$$P_0 = \sigma^2 I \tag{12.a}$$

The recurrence relations that describe the covariance propagation are

$$P_k(-) = A_{k-1} P_{k-1}(+) A_{k-1}^T + Q_{k-1} \tag{13.a}$$

$$K_k = P_k(-) H_k^T [H_k P_k(-) H_k^T + R_k]^{-1} \tag{13.b}$$

$$P_k(+) = [I - K_k H_k] P_k(-) \tag{13.c}$$

where $A_k = H_k = I$. In this example we choose $Q_{k-1} = 0$. Let us consider a case where $R_k = rI$. Eqs. (13) become

$$P_k(-) = P_{k-1}(+) \tag{14.a}$$

$$K_k = P_k(-)[P_k(-) + rI]^{-1} \tag{14.b}$$

$$P_k(+) = [I - K_k] P_k(-) \tag{14.c}$$

Using Eq. (14.a), Eqs. (14.b) and (14.c) can be written as

$$K_k = P_{k-1}(+)[P_{k-1}(+) + rI]^{-1} \tag{14.d}$$

$$P_k(+) = [I - K_k] P_{k-1}(+) \tag{14.e}$$

Since P_0 is a diagonal matrix, all the matrices in Eqs. (14) are diagonal. Therefore, we can write the last equation for any of the elements of $P_k(+)$ in terms of the same element in $P_{k-1}(+)$. Dropping the plus sign, we obtain the following recurrence relation for any element of $P_k(+)$:

$$p_k = \frac{r p_{k-1}}{p_{k-1} + r} \tag{15.a}$$

with

$$p_0 = \sigma^2 \tag{15.b}$$

It is quite obvious that p_k, which is an eigenvalue of $P_k(+)$, is not zero and thus the covariance matrix of the filter is not singular.

V. ESTIMATION OF A QUATERNION FROM VECTOR MEASUREMENTS

In this final example we treat the classical case where the body rotates and the measurements are vector measurements, we numerically demonstrate that the filter-computed covariance matrix is non-singular both when we do and do not impose normality on the estimated quaternion. First we present the simulation, followed by the filter development.

V.1 Simulation

Without loss of generality, we consider a case where the body rotates at the constant angular rate. Let the rate vector in this simulation be

$$\omega = [0.1 \quad 0.2 \quad 0.3]^T \text{ rad/sec} \tag{16}$$

The quaternion at time t_{k+1} as a function of its value at time t_k is given by [5, p. 512]

$$\mathbf{q}_{k+1} = e^{\frac{1}{2}\Omega(\omega)\cdot\Delta t}\mathbf{q}_k \qquad (17)$$

We assume that two vector measurements are obtained at each time point, t_k. They constantly point at two celestial objects located 90° apart and are given in the reference coordinates as

$$\mathbf{r}_1 = [1 \;\; 0 \;\; 0]^T \qquad (18.a)$$

$$\mathbf{r}_2 = [0 \;\; 1 \;\; 0]^T \qquad (18.b)$$

In body coordinates we simulate two corresponding measured vectors as

$$\mathbf{b}_{1,k+1} = D(\mathbf{q}_{k+1})\mathbf{r}_1 + \mathbf{v}_1 \qquad (19.a)$$

$$\mathbf{b}_{2,k+1} = D(\mathbf{q}_{k+1})\mathbf{r}_2 + \mathbf{v}_2 \qquad (19.b)$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are uncorrelated, Gaussian white sequences with covariance R each and [5, p. 414]

$$D(\mathbf{q}_{k+1}) = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\ 2(q_1q_3 + q_2q_4) & 2(q_2q_3 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix}_{k+1} \qquad (20)$$

In addition to $\mathbf{b}_{1,k+1}$ and $\mathbf{b}_{2,k+1}$ the filter is also furnished with $\mathbf{r}_1$, $\mathbf{r}_2$, ω and the covariance matrix R.

V.2 Filter development

V.2.1 Dynamics

The filter dynamics equation is simply

$$\hat{\mathbf{q}}_{k+1} = e^{\frac{1}{2}\Omega(\omega)\Delta t}\hat{\mathbf{q}}_k \qquad (21)$$

V.2.2 Measurements

The effective measurement equation for the AEKF is developed as follows

$$\mathbf{z}_{k+1} = \mathbf{b}_{k+1} - \hat{\mathbf{b}}_{k+1} = D(\mathbf{q}_{k+1})\mathbf{r} - D(\hat{\mathbf{q}}_{k+1}(-))\mathbf{r} + \mathbf{v}_{k+1} \qquad (22.a)$$

$$\mathbf{z}_{k+1} = D(\hat{\mathbf{q}}_{k+1}(-) + d\mathbf{q}_{k+1})\mathbf{r} - D(\hat{\mathbf{q}}_{k+1}(-))\mathbf{r} + \mathbf{v}_{k+1} \qquad (22.b)$$

$$\mathbf{z}_{k+1} = D(\hat{\mathbf{q}}_{k+1}(-))\mathbf{r} + \left.\frac{\partial[D(\mathbf{q})\mathbf{r}]}{\partial \mathbf{q}}\right|_{\mathbf{q}_{k+1}} d\mathbf{q}_{k+1} - D(\hat{\mathbf{q}}_{k+1}(-))\mathbf{r} + \mathbf{v}_{k+1} \qquad (22.c)$$

$$\mathbf{z}_{k+1} = \left.\frac{\partial[D(\mathbf{q})\mathbf{r}]}{\partial \mathbf{q}}\right|_{\hat{\mathbf{q}}_{k+1}(-)} d\mathbf{q}_{k+1} + \mathbf{v}_{k+1} \qquad (22.d)$$

Using Eq. (20) we obtain

$$D(\mathbf{q})\mathbf{r} = \begin{bmatrix} (q_1^2 - q_2^2 - q_3^2 + q_4^2)r_1 + 2(q_1q_2 + q_3q_4)r_2 + 2(q_1q_3 - q_2q_4)r_3 \\ 2(q_1q_2 - q_3q_4)r_1 + (-q_1^2 + q_2^2 - q_3^2 + q_4^2)r_2 + 2(q_2q_3 + q_1q_4)r_3 \\ 2(q_1q_3 + q_2q_4)r_1 + 2(q_2q_3 - q_1q_4)r_2 + (-q_1^2 - q_2^2 + q_3^2 + q_4^2)r_3 \end{bmatrix} \qquad (23)$$

where r_1, r_2 and r_3 are the components of $\mathbf{r}$. Define

$$H_{k+1} = \left.\frac{\partial[D(\mathbf{q})\mathbf{r}]}{\partial \mathbf{q}}\right|_{\hat{\mathbf{q}}_{k+1}(-)} \qquad (24)$$

then using Eq. (23) we obtain

$$H_{k+1} = 2\begin{bmatrix} (q_1r_1 + q_2r_2 + q_3r_3) & (-q_2r_1 + q_1r_2 - q_4r_3) & (-q_3r_1 + q_4r_2 + q_1r_3) & (q_4r_1 + q_3r_2 - q_2r_3) \\ (q_2r_1 - q_1r_2 + q_4r_3) & (q_1r_1 + q_2r_2 + q_3r_3) & (-q_4r_1 - q_3r_2 + q_2r_3) & (-q_3r_1 + q_4r_2 + q_1r_3) \\ (q_3r_1 - q_4r_2 - q_1r_3) & (q_4r_1 + q_3r_2 - q_2r_3) & (q_1r_1 + q_2r_2 + q_3r_3) & (q_2r_1 - q_1r_2 + q_4r_3) \end{bmatrix}_{\hat{\mathbf{q}}_{k+1}(-)} \qquad (25)$$

and Eq. (22.d) becomes

$$\mathbf{z}_{k+1} = H_{k+1}d\mathbf{q}_{k+1} + \mathbf{v}_{k+1} \qquad (26)$$

The filter is propagated and also updated every second. At every update point we have two effective vector measurements, $\mathbf{z}_{1,k+1}$ and $\mathbf{z}_{2,k+1}$. The algorithm is a standard EKF algorithm. We ran two cases, one without normalization of the estimated quaternion and one with normalization. Fig. 5 shows the attitude error when normalization is not imposed on the estimated quaternion, with an initial error of 10 degrees. Fig. 6 shows the covariance matrix eigenvalues, with the scale reduced to show the steady-state behavior. Fig. 7 shows the deviation from unity of the norm of the estimated quaternion. Figs. 8 and 9 present the results for the case where normalization is imposed on the estimated quaternion. Note that *the eigenvalues do not change as a result of the quaternion normalization*. (See Sections VII and VIII for a discussion on the changes to the EKF algorithm as a result of the normalization.)

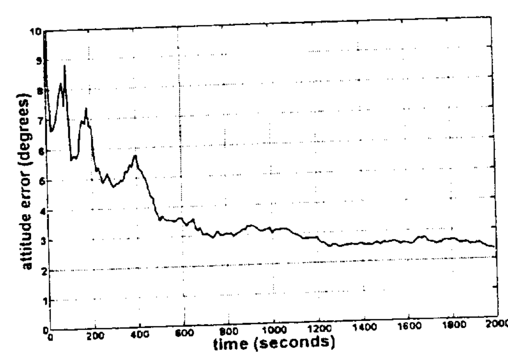


Fig. 5: Attitude error

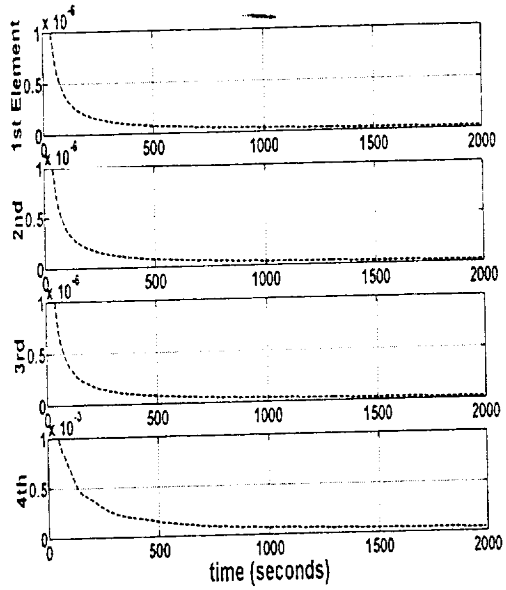


Fig. 6: Covariance matrix eigenvalues.

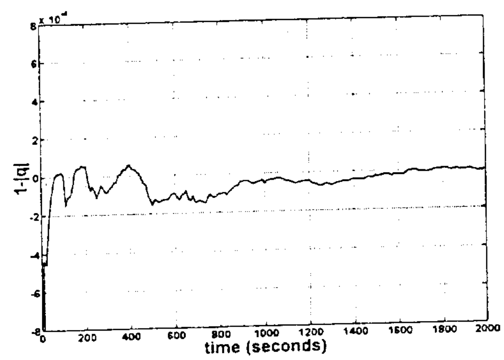


Fig. 7: Norm error of the estimated quaternion

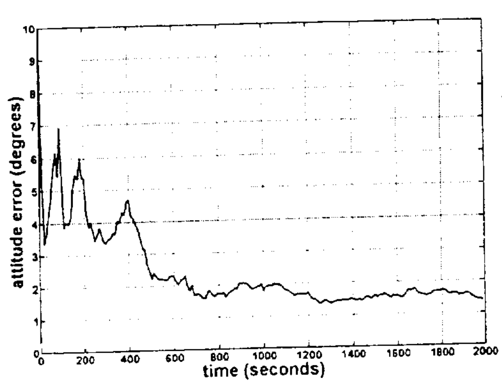


Fig. 8: Attitude error in the constrained case

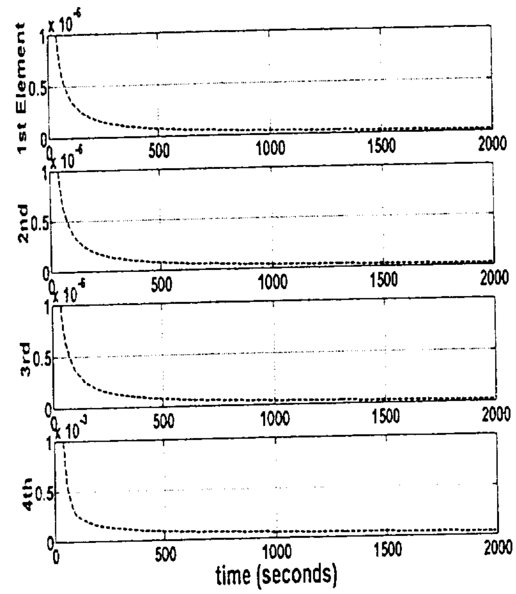


Fig. 9: Covariance matrix eigenvalues.

VI. THE FULLY RESET ADDITIVE EKF

To explain the results obtained in the preceding examples, without the imposed constraints, we need to analyze the operation of the AEKF, and for that we need to review the EKF algorithm. We do it in a manner that is somewhat different from the usual textbook algorithm development, but the resultant algorithm is the standard EKF. (We need to adopt this approach for the development that will be presented in the following section.) Here we treat the special case that is applicable to quaternion estimation, where the measurement model is nonlinear, but the dynamics model that describes the state propagation is linear.

VI.1 Measurement Update

A measurement vector $\mathbf{y}_k$, at time t_k, is related to the state vector $\mathbf{X}_k$, at time t_k, according to the following nonlinear vector function

$$\mathbf{y}_k = \mathbf{h}(\mathbf{X}_k) + \mathbf{v}_k \tag{27}$$

where $\mathbf{v}_k$ is a zero mean white noise vector. Suppose that at this time point we have an a-priori estimate, $\hat{\mathbf{X}}_k(-)$, of this state vector. We wish to use the measurement $\mathbf{y}_k$ to improve this a-priori estimate. The improved estimate is called the a-posteriori estimate and is denoted by $\hat{\mathbf{X}}_k(+)$. The KF was developed for linear measurements whereas the measurement equation, Eq. (27), is nonlinear. *However, as will be shown in the ensuing discussion, the EKF estimates the difference,* $\mathbf{e}$, *between the true state vector,* $\mathbf{X}$, *and its estimate,* $\hat{\mathbf{X}}$, *and the model treated by the EKF is, to first order, a linear model in* $\mathbf{e}$.

We recall that the state update equation in the linear KF is

$$\hat{X}_k(+) = \hat{X}_k(-) + K_k[y_k - \hat{y}_k] \quad (28.a)$$

where

$$\hat{y}_k = H_k \hat{X}_k(-) \quad (28.b)$$

which results from the linear model

$$y_k = H_k X_k + v_k \quad (28.c)$$

In EKF terminology the term $y_k - \hat{y}_k$ is known as the *effective measurement*, which we denote by z_k; that is

$$z_k = y_k - \hat{y}_k \quad (29)$$

Using Eq. (27) we can write the effective measurement as follows.

$$z_k = h(X_k) + v_k - \hat{y}_k \quad (30)$$

Note that $\hat{y}_k$, in Eq. (28.b), was obtained from the linear measurement equation (28.c) by dropping the noise vector and substituting $\hat{X}_k(-)$ for X_k. If we do the same in Eq. (27) then Eq. (30) becomes

$$z_k = h(X_k) + v_k - h(\hat{X}_k(-)) \quad (31)$$

As mentioned before, we denote by $\mathbf{e}$ the difference between the true and estimated state vector; that is,

$$\mathbf{e} = X - \hat{X} \quad (32)$$

Because $X = \hat{X} + \mathbf{e}$, we can expand $h(X_k)$ in a Taylor series about $\hat{X}$, drop the second and higher order terms in $\mathbf{e}$, and substitute the result into Eq. (31). This yields z_k as a linear function of $\mathbf{e}$. The first order expansion of $h(X_k)$ about $\hat{X}_k(-)$ yields

$$h(X_k) = h(\hat{X}_k(-)) + \left.\frac{\partial h(X)}{\partial X}\right|_{X=\hat{X}_k(-)} e_k(-) \quad (33)$$

and substitution of the last equation into Eq. (31) gives

$$z_k = \left.\frac{\partial h(X)}{\partial X}\right|_{X=\hat{X}_k(-)} e_k(-) + v_k \quad (34)$$

Let

$$H_k = \left.\frac{\partial h(X)}{\partial X}\right|_{X=\hat{X}_k(-)} \quad (35)$$

then Eq. (34) can be written as

$$z_k = H_k e_k(-) + v_k \quad (36)$$

Following the measurement-update equation of the state vector in Eq. (28.a) we write

$$\hat{e}_k(+) = \hat{e}_k(-) + K_k z_k \quad (37)$$

Substitution of Eq. (36) into Eq. (37) yields

$$e_k(+) = e_k(-) - K_k H_k e_k(-) - K_k v_k \quad (38)$$

which can be written as

$$e_k(+) = [I - K_k H_k] e_k(-) - K_k v_k \quad (39)$$

Assuming (the knowingly inaccurate assumption) that the filter is unbiased; that is, $E\{e_k(+)\} = 0$, the updated estimation error covariance is computed as follows

$$P_k(+) = E\{e_k(+) e_k^T(+)\} \quad (40)$$

Substituting Eq. (39) along with the fact that $E\{e_k(-) v_k^T\} = 0$, and the following notations

$$R_k = E\{v_k v_k^T\} \quad (41)$$

$$P_k(-) = E\{e_k(-) e_k^T(-)\} \quad (42)$$

Eq. (40) becomes

$$P_k(+) = [I - K_k H_k] P_k(-) [I - K_k H_k]^T + K_k R_k K_k^T \quad (43)$$

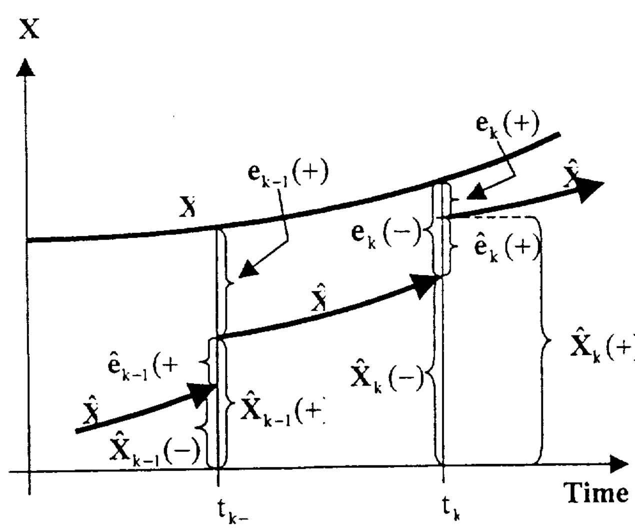


Fig. 10: The state, its estimate, the difference and its estimate at time t_k.

The sequence of events associated with the measurement update in the EKF is presented in Fig. 10 and can be summarized as follows:

1. $\hat{X}_{k-1}(+)$ is propagated from time t_{k-1} to time point t_k to become $\hat{X}_k(-)$.

2. The difference $X_k - \hat{X}_k(-)$ is $e_k(-)$.

3. The measurement update yields $\hat{e}_k(+)$, an estimate of $e_k(-)$.

4. $\hat{e}_k(+)$ is added to $\hat{X}_k(-)$ to form $\hat{X}_k(+)$ which is then propagated to time point t_{k+1} to become $\hat{X}_{k+1}(-)$. We call this operation *full reset*. Note that once $\hat{e}_k(+)$ is absorbed in $\hat{X}_k(+)$ it is not propagated separately.

Once $\hat{e}_k(+)$ is added to $\hat{X}_k(-)$ a full reset is performed, and there is no $\hat{e}_k(+)$ to be propagated forward; therefore, $\hat{e}_k(-)$, the a priori estimate of **e** at the next time update point, is also zero.

Using the preceding explanation, the EKF algorithm at the measurement update can be summarized as follows.

At a measurement update:
Signal:

$$\hat{e}_k(-) = 0 \tag{44.a}$$

$$\hat{e}_k(+) = \hat{e}_k(-) + K_k(y_k - \hat{y}_k) \tag{44.b}$$

$$\hat{X}_k(+) = \hat{X}_k(-) + \hat{e}_k(+) \tag{44.c}$$

Note that a sequential substitution of Eqs. (44.a, b, c) yields the following textbook expression for the state measurement-update equation

$$\hat{X}_k(+) = \hat{X}_k(-) + K_k(y_k - \hat{y}_k) \tag{44.d}$$

Covariance:

$$P_k(+) = [I - K_k H_k]P_k(-)[I - K_k H_k]^T + K_k R_k K_k^T \tag{45}$$

In the next section we will use the formulation of the state measurement-update given in Eqs. (44.a-c) to show the effect that normalization has on the state measurement-update algorithm. This cannot be shown when the textbook expression of Eq. (44.d) is used.

VI.2 Time Propagation

Consider the case where the discrete process equation that describes the time propagation of a general state vector is given by

$$X_k = A_{k-1}X_{k-1} + w_{k-1} \tag{46}$$

From Eq. (32) and as illustrated in Fig. 10

$$X_{k-1} = \hat{X}_{k-1}(+) + e_{k-1}(+) \tag{47.a}$$

and

$$X_k = \hat{X}_k(-) + e_k(-) \tag{47.b}$$

Substituting X_{k-1} given in Eqs. (47) into Eq. (46) yields

$$\hat{X}_k(-) + e_k(-) = A_{k-1}\hat{X}_{k-1}(+) + A_{k-1}e_{k-1}(+) + w_{k-1} \tag{58}$$

Since the dynamics model of (21) is lincar [6, p. 75]

$$\hat{X}_k(-) = A_{k-1}\hat{X}_{k-1}(+) \tag{59}$$

Subtracting the last equation from Eq. (48) yields

$$e_k(-) = A_{k-1}e_{k-1}(+) + w_{k-1} \tag{50}$$

Using these results we can now examine the time propagation of the state estimate and the covariance matrix in this particular EKF. Assuming $E\{e_k(-)\} = 0$, the propagated covariance matrix is defined as

$$P_k(-) = E\{e_k(-)e_k^T(-)\} \tag{51}$$

Substitution of $e_k(-)$, given in Eq. (50), into the last expression, given that $e_{k-1}(+)$ and w_{k-1} are uncorrelated, and using the notation $Q_{k-1} = E\{w_{k-1}w_{k-1}^T\}$ we obtain the well known result [6, p. 76]

$$P_k(-) = A_{k-1}P_{k-1}(+)A_{k-1}^T + Q_{k-1} \tag{52}$$

Using Eqs. (49) and (52) we can summarize the time propagation stage of the EKF when the dynamics model is linear.

Time update:
Signal:

$$\hat{X}_k(-) = A_{k-1}\hat{X}_{k-1}(+) \tag{53}$$

Covariance:

$$P_k(-) = A_{k-1}P_{k-1}(+)A_{k-1}^T + Q_{k-1} \tag{54}$$

We realize from the preceding development that the AEKF that is used to estimate the quaternion when no normalization takes place is the ordinary EKF algorithm. Moreover, it is obvious that the fact that the true quaternion is a normal vector has no bearing on the filter covariance matrix. This is, of course, also true for all the linear examples presented before.

VII. THE PARTIALLY RESET ADDITIVE EXTENDED KALMAN FILTER

To explain the results obtained in the numerical examples, when constraints were imposed, we need to consider a special version of the EKF. For this we will make use of the developments presented in the previous section.

VII.1 Time Propagation

Figure 11.a describes the steps of the ordinary (fully reset) EKF at the measurement update stage as described

in the preceding section and illustrated in Fig. 10. We have shown that after computing $\hat{e}_k(+)$, the estimate of e_k, it is added to $\hat{X}_k(-)$ to form $\hat{X}_k(+)$ which is then propagated to time point t_{k+1} to become $\hat{X}_{k+1}(-)$. Since $\hat{e}_k(+)$ in its entirety is added to $\hat{X}_k(+)$, this operation constitutes a full reset. Because $\hat{e}_k(+)$ is propagated forward as a part of $\hat{X}_k(+)$, it is not propagated independently. Suppose now that only a part of $\hat{e}_k(+)$, denoted in Fig. 11.b by Δ_k, is added to $\hat{X}_k(-)$ to form $\hat{X}^*_k(+)$ which is then propagated to time point t_{k+1} to become $\hat{X}_{k+1}(-)$. This *partial reset* leaves $\hat{e}^*_k(+)$ out of the propagated full state vector. Therefore $\hat{e}^*_k(+)$ has to be propagated forward separately. It is easy to show that when the dynamics model is linear, $\hat{e}^*_k(+)$ is propagated according to

$$\hat{e}_{k+1}(-) = A_k \hat{e}^*_k(+) \tag{55.a}$$

Reducing the index by 1 yields

$$\hat{e}_k(-) = A_{k-1} \hat{e}^*_{k-1}(+) \tag{55.b}$$

We note in Fig. 11.b that *the partial reset, does not change the value of the actual estimation error,* $e_k(+)$, therefore the *partial reset does not influence the covariance matrix* because the latter is the covariance matrix of **e**, the estimation error itself.

The AEKF algorithm for the time propagation stage in this partially reset case is then as follows.

At a time update

Signal:

$$\hat{X}_k(-) = A_{k-1} \hat{X}^*_{k-1}(+) \tag{56.a}$$

$$\hat{e}_k(-) = A_{k-1} \hat{e}^*_{k-1}(+) \tag{56.b}$$

Covariance:

$$P_k(-) = A_{k-1} P_{k-1}(+) A^T_{k-1} + Q_{k-1} \tag{57}$$

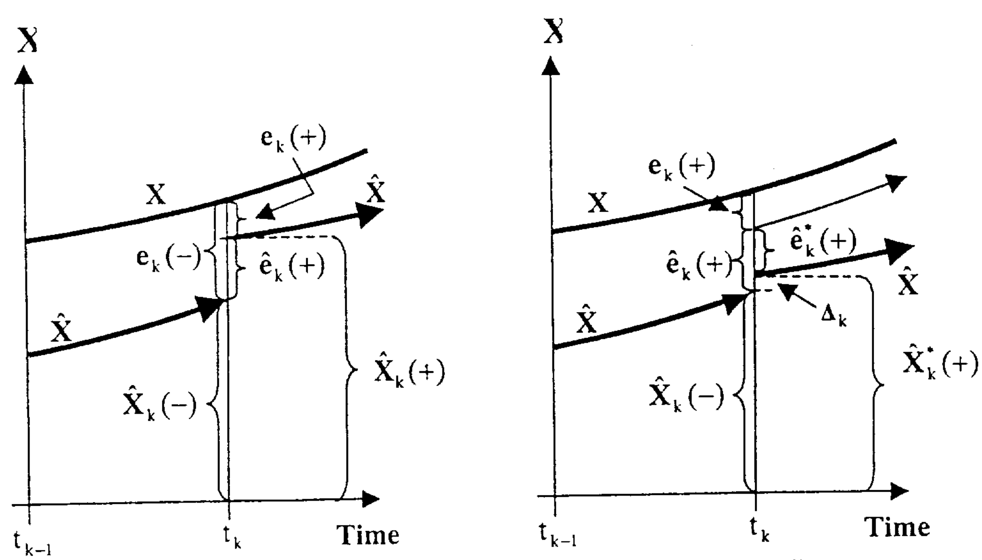
VII.2 Measurement Update

From the preceding discussion of the partial reset operation and the developments presented in Section VI, it is obvious that the measurement update is performed as follows.

Measurement update

Signal:

$$\hat{e}_k(+) = \hat{e}_k(-) + K_k(y_k - \hat{y}_k) \tag{58.a}$$



(a) The fully reset case

(b) The partially reset case

Fig. 11: Fully versus partially reset case

$$\hat{X}_k(+) = \hat{X}_k(-) + \hat{e}_k(+) \quad (58.b)$$

Covariance:

$$P_k(+) = [I - K_k H_k] P_k(-) [I - K_k H_k]^T + K_k R_k K_k^T \quad (59)$$

VII.3 Reset Update

Although the change from $\hat{X}_k(+)$ to $\hat{X}_k^*(+)$, shown in Fig. 11.b, can be of any nature, because we are dealing with the normalization operation, we chose the particular change

$$\hat{X}_k^*(+) = \hat{X}_k(+)/\|\hat{X}_k(+)\| \quad (60.a)$$

It is clear from Fig. 17.b that once $\hat{X}_k^*(+)$ is defined, the value of $\hat{e}_k^*(+)$ is also defined as

$$\hat{e}_k^*(+) = \hat{X}_k(+) - \hat{X}_k^*(+) \quad (60.b)$$

Note that the condition of partially reset state vector does not come about by adding Δ_k to $\hat{X}_k(-)$ but rather by subtracting $\hat{e}_k^*(+)$ from the fully reset state vector $\hat{X}_k(+)$. As explained before, the partial reset operation does not influence the covariance matrix; therefore, the reset update is performed as follows.

Reset update

Signal:

$$\hat{X}_k^*(+) = \hat{X}_k(+)/\|\hat{X}_k(+)\| \quad (61.a)$$

$$\hat{e}_k^*(+) = \hat{X}_k(+) - \hat{X}_k^*(+) \quad (61.b)$$

Covariance:

$$P_k^*(+) = P_k(+) \quad (62)$$

Both the fully and partially reset AEKF are summarized in Table I.

Table I: AEKF with Full and Partial Reset

System Model: Dynamics: $X_k = A_k X_{k-1} + w_{k-1}$; Measurement: $y_k = h(X_k) + v_k$	
AEKF with Linear Dynamics and Full Reset	**AEKF with Linear Dynamics and Partial Reset**
Time Propagation: $\hat{X}_k(-) = A_{k-1}\hat{X}_{k-1}(+)$ $P_{k-1}(-) = A_{k-1}P_{k-1}(+) A_{k-1}^T + Q_{k-1}$	Time Propagation: $\hat{X}_k(-) = A_{k-1}\hat{X}_{k-1}^*(+)$ $\hat{e}_k(-) = A_{k-1}\hat{e}_{k-1}^*(+)$ $P_{k-1}(-) = A_{k-1}P_{k-1}(+) A_{k-1}^T + Q_{k-1}$
Measurement Update: $\hat{e}_k(+) = K_k(y_k - \hat{y}_k)$ $\hat{X}_k(+) = \hat{X}_k(-) + \hat{e}_k(+)$ *Consequently:* $\hat{X}_k(+) = \hat{X}_k(-) + K_k(y_k - \hat{y}_k)$ $P_k(+) = [I - K_k H_k] P_k(-) \cdot [I - K_k H_k]^T + K_k R_k K_k^T$	Measurement Update: $\hat{e}_k(+) = \hat{e}_k(-) + K_k(y_k - \hat{y}_k)$ $\hat{X}_k(+) = \hat{X}_k(-) + \hat{e}_k(+)$ *Consequently:* $\hat{X}_k(+) = \hat{X}_k(-) + \hat{e}_k(-) + K_k(y_k - \hat{y}_k)$ $P_k(+) = [I - K_k H_k] P_k(-) [I - K_k H_k]^T + K_k R_k K_k^T$ Reset Update: $\hat{X}_k^*(+) = \hat{X}_k(+)/\|\hat{X}_k(+)\|$ $\hat{e}_k^*(+) = \hat{X}_k(+) - \hat{X}_k^*(+)$ $P_k^*(+) = P_k(+)$

VIII. THE REGULARITY OF THE AEKF FOR QUATERNION ESTIMATION

After presenting some illustrative examples and having prepared the theoretical background we can now analyze the results of the examples. We start with quaternion estimation using the AEKF without normalization of the estimated quaternion.

Quaternion estimation without normalization

We presented two cases of quaternion estimation; namely, a static case with quaternion measurements and a dynamic case with vector measurements. The former constituted a strictly linear estimation problem that required the use of a standard KF. In the standard KF there is no connection between the quaternion estimate and the filter covariance matrix. Therefore, theoretically, the normality quality of the true quaternion had no bearing on the covariance matrix. Indeed for this example we derived an analytic expression that exhibited no singularity.

In the more complicated case with a rotating body and vector measurements, the dynamics equation is linear but the measurement equation is nonlinear. In that case we used the AEKF and still found no singularity. The fact that the true quaternion is normal does not enter into the computation of the covariance matrix. The only special feature of this case is the orthogonality of the transition matrix that propagates **q**, which preserves the norm of the propagated state vector be it normal or not.

Quaternion estimation with normalization

Forcing normalization on the a-posteriori estimate of the quaternion does not affect the covariance matrix. The partially reset operation is performed as follows. First, the quaternion is updated as

$$\hat{q}_k(+) = \hat{q}_k(-) + \hat{e}_k(+) \tag{63}$$

where $\hat{e}_k(+)$ is the estimate of the difference $q_k - \hat{q}_k(-)$. Now, the forced normalization yields $\hat{q}_k^*(+)$ as follows

$$\hat{q}_k^*(+) = \frac{\hat{q}_k(+)}{\|\hat{q}_k(+)\|} \tag{64}$$

According to (61.b), the remaining error term is given as

$$\hat{e}_k^* = \hat{q}_k(+) - \frac{\hat{q}_k(+)}{|\hat{q}_k(+)|} \tag{65}$$

This is a realization of the partially reset case of the AEKF depicted in Fig. (11.b) and listed on the right column of Table I. We note that the only difference between the partially reset AEKF and the fully reset AEKF (which, as mentioned, is actually the ordinary EKF) is in the propagation of $\hat{X}_k^*(+)$, which here is $\hat{q}_k^*(+)$, and of $\hat{e}_k^*(+)$. (There are cases where in practice the filter performs well even when $\hat{e}_k^*(+)$ is not propagated. It is assumed that the reason for it is that $\hat{e}_k^*(+)$, which is caused by the normalization operation, is quite small and/or the measurement update subdues any divergence tendency.) Since the partially reset operation has no effect on the covariance matrix computation, the normalization operation has no effect on the covariance of the AEKF. With or without normalization, we obtain the same nonsingular covariance matrix, P.

IX. CONCLUSIONS

The purpose of this paper is to explain that using an AEKF for estimating the quaternion-of-rotation does not result in a singular covariance matrix. We started this paper by presenting a conceptual example of estimating the vertices of a rotating square where, in spite of the geometric connection that exists between three of the vertices and the fourth one, there is no reason to assume that such a connection exists between estimates of the vertices. This constituted an analogy to the case of estimating the quaternion-of-rotation where there is no reason to assume that there is a functional relationship between the estimates of the quaternion elements although there is a connection between the elements of the true quaternion.

Advancing from the qualitative example to a quantitative one, we presented a KF that estimated the position of the edges of a sliding rod the length of which was constant. In that example we checked the singularity of the covariance matrix both when the constant length between the estimated edges was not forced and when we did force it. In both cases the covariance matrix exhibited no singularity. Moreover, the matrix eigenvalues were identical in both cases.

Following these general examples we moved to the case of quaternion estimation. We started with a particular case of a quaternion estimation problem of a rigid body with constant attitude and quaternion measurements. In this simple example where normalization was not forced we proved analytically that the covariance matrix is not singular. Finally we presented a case with a rotating rigid body and vector measurements. Again, the computed covariance matrix exhibited no singularity.

In order to explain the results obtained in these examples we presented the ordinary EKF algorithm as a full reset estimation problem and, in parallel, we presented an EKF version as a partial reset estimation problem. It was shown that in either algorithm no inherent covariance matrix singularity is involved. Finally it was shown that the AEKF for estimating the quaternion with no forced normalization is an EKF with full reset, and when normalization is forced, the algorithm is an EKF with partial reset, and as mentioned, in either case there is no inherent singularity in the covariance matrix. Indeed, in numerous runs of the AEKF under diverse conditions, with simulated and real data, with and without normalization, the covariance matrix was never singular.

References

[1] 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, 1982, pp. 417-429.

[2] Deutschmann (now Thienel), J. and Bar-Itzhack, I.Y., "Evaluation of Attitude and Orbit Estimation Using Actual Earth Magnetic Field Data", *AIAA J. of Guidance, Control, and Dynamics*, Vol. 24, No. 3, May-June 2001, pp. 616-623.

[3] Bar-Itzhack, I.Y., and Oshman, Y., "Altitude Determination from Vector Measurements: Quaternion Estimation," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. AES-21, No. 1, Jan. 1985, pp. 128-136.

[4] Bryson, A.E. "Kalman Filter Divergence and Aircraft Motion Estimator", *J. of Guidance and Control*, Vol. 1, No. 1, January-February 1978, pp. 71-79.

[5] Wertz, J.R., (Ed.), *Spacecraft Attitude Determination and Control*, Reidel Publ. Co., Dordrecht, Holland, 1984.

[6] Gelb, A. (editor), *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.