Recursive Attitude Determination from Vector Observations: Direction Cosine Matrix Identification

Itzhack Y. Bar-Itzhack* and Jacob Reiner†
Technion—Israel Institute of Technology, Haifa, Israel

This work presents two recursive estimation algorithms that use pairs of measured vectors to yield minimum variance estimates of the direction cosine matrix (DCM). Both algorithms are based on a parameter identification method of a linear dynamic system. One of the algorithms is derived from a straightforward application of this identification method. In the other algorithm use is also made of the orthogonality property of the DCM to achieve a faster convergence rate to an orthogonal estimate of the DCM. Monte Carlo simulation runs were made that demonstrated the efficiency of the algorithms.

I. Introduction

DETERMINATION of the attitude difference between a Cartesian coordinate system v attached to the body of a vehicle and a reference Cartesian coordinate system u is a necessary and significant stage in the guidance and control of aerospace vehicles. There are cases in which this attitude difference has to be extracted from the measurement of vectors in the two coordinate systems. For example, this is the case when the attitude of a satellite is estimated using direction cosines of objects as observed in the satellite fixed coordinate system and direction cosines of the same objects in a known reference coordinate system. In those cases a sequence of vectors $\{\bar{r}_i\}$ $i=1,2,...,N$ is measured in the two coordinate systems u and v. The measurements in system u result in the corresponding sequence $\{u_i\}$ and the measurements in system v result in the sequence $\{v_i\}$ where the column matrices u_i as well as $v_i \epsilon R^3$. We wish to find the estimate D of the direction cosine matrix (DCM) which is the transformation from coordinate system u to coordinate system v. (Note that $\bar{r}_i$ are not necessarily unit vectors.)

This problem was posed by Wahba[1] who was the first to choose a least square criterion to define the best estimate. That is, she was looking for the orthogonal matrix D that minimizes the cost function

$$L(\hat{D}) = \sum_{i=1}^{N} \|v_i - \hat{D}u_i\|^2$$

where $\|\cdot\|$ denotes the Euclidean norm. The solution to this problem was proposed by Wahba et al.[2] Brock,[3] for example, collected the two sequences of measured vectors $\{v_i\}$ and $\{u_i\}$ into two matrices as follows

$$V \triangleq [v_1, v_2, \ldots, v_N]$$

$$U \triangleq [u_1, u_2, \ldots, u_N]$$

and derived the batch solution

$$\hat{D} = (VU^TUV^T)^{1/2}(UV^T)^{-1}$$

in which T denotes the matrix transpose. In fact, Brock showed that the last expression is the solution to a more general problem; namely, it minimizes the cost function

$$L'(\hat{D}) = tr[(V-\hat{D}U)^T W(V-\hat{D}U)]$$

in which W is some positive-definite weighting matrix.

Carta and Lackowski[4] derived a recursive formulation to Brock's batch solution. In addition, they formulated a fading-weight least square criterion from which they derived a batch as well as a recursive solution.

Shuster and Oh[5] discussed two solutions to the attitude determination problem. The first solution called the TRIAD algorithim[6] is a deterministic solution that discards all but two pairs of measured vectors. The second solution, termed QUEST algorithm,[7] is a solution of Wahba's problem, yielding an estimated quaternion that is, therefore, optimal in the least square sense.

Podgorski et al.[8] were concerned with a similar problem. They, however, were mainly concerned with the operation of the complete system rather than with the filter development. In their work, attitude is described by a quaternion and their filter estimates the deviation of the true quaternion from the assumed one.

The problem of attitude determination is a common problem in inertial navigation too. There, however, the approach is different on two accounts. First, the measurements are usually velocity errors.[9] Second, the attitude determination is based on the estimation of miaslignment angles that are used to correct, repeatedly, the estimated DCMD.[10,11]

In the present work, as in Refs. 1-8, attitude determination is based on two sequences of corresponding measured column matrices. The whole DCM (rather than the DCM error) is estimated using a minimum variance estimator. The estimation problem is actually posed as a system identification problem and the orthogonality property of the DCM is used to enhance the identification algorithm. In fact, by incorporating the DCM orthogonality requirement into the identification algorithm, this work, in a way, furnishes a solution to the long-standing problem posed in Ref. 4: namely, is there a convenient recursive form that utilizes the orthogonality property of the DCM to yield an orthogonal estimate with no further orthogonalization of the resultant DCM?

The novelty of this work is manifested in the application of an identification algorithm to the DCM estimation problem and in the successful utilization of the DCM orthgonality property for the achievement of an improved identification algorithm.

Received Nov. 22, 1982; revision received March 20, 1983.

*Associate Professor, Department of Aeronautical Engineering. Member AIAA.

†Graduate Student, Department of Aeronautical Engineering.

In the following section the DCM identification algorithm which does not employ orthogonalization will be presented. The algorithm that includes orthogonalization will then be presented in Sec. III. Section IV will present the indices of performance that were selected to rate the performance of the algorithms. Monte Carlo simulations of the two algorithms will also be presented in Sec. IV. The conclusions are discussed in Sec. V.

II. Direction Cosine Matrix Identification

The Static Case

Let $u_{0,i}$ and $v_{0,i}$ denote column matrices whose components are the components of the vector $\bar{r}_i$, coordinatized in the u and v coordinate systems, respectively. Obviously,

$$v_{0,i} = Du_{0,i} \tag{1}$$

It is assumed that the measurements u_i and v_i of $u_{0,i}$, and $v_{0,i}$ are contaminated by the noises $n_{u,i}$ and $n_{v,i}$, respectively, such that

$$u_i = u_{0,i} + n_{u,i} \tag{2a}$$

$$v_i = v_{0,i} + n_{v,i} \tag{2b}$$

where $n_{u,i}$ and $n_{v,i}$ are zero mean white noise sequences with the covariance matrices

$$\mathrm{Cov}\{n_{u,i}\} = R_{u,i}$$

$$\mathrm{Cov}\{n_{v,i}\} = R_{v,i}$$

$n_{u,i}$ and $n_{v,i}$ are uncorrelated with one another or with either $u_{0,i}$ or $v_{0,i}$. Substitution of Eqs. (2) into Eq. (1) yields

$$v_i = Du_i + n_{v,i} - Dn_{u,i} \tag{3}$$

Following Mayne's system identification algorithm[12] we reverse the role of D and u_i as follows. Let o denote a 3×1 zero column matrix, then define

$$H_i \triangleq \begin{bmatrix} u_i^T & o^T & o^T \\ o^T & u_i^T & o^T \\ o^T & o^T & u_i^T \end{bmatrix} \tag{4}$$

In addition, let d_j^T, $j=1,2,3$, denote the three *rows* of D and define the column matrix d such that

$$d = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \tag{5}$$

Finally, define n_i

$$n_i \triangleq n_{v,i} - Dn_{u,i} \tag{6}$$

Then, from Eqs. (3-6), it is clear that

$$v_i = H_i d + n_i \tag{7}$$

It is obvious, from Eq. (6), that n_i is a zero mean white noise sequence for which

$$R_i \triangleq \mathrm{Cov}\{n_i\} = R_{v,i} + DR_{u,i}D^T \tag{8}$$

With Eq. (7) on hand, we can now use the Kalman filter (KF) algorithm in order to obtain at step i the minimum variance estimate $\hat{d}_i$ of d or, in view of Eq. (5), to obtain the minimum variance estimate $\hat{D}_i$ of D. This estimate is updated with each new available measured pair u_i, v_i. The KF algorithm for this case is as follows

between measurements:

$$\hat{d}_{i+1/i} = \hat{d}_{i/i} \tag{9a}$$

$$P_{i+1/i} = P_{i/i} \tag{9b}$$

across measurements:

$$K_{i+1} = P_{i+1/i}H_{i+1}^T P_{i+1}(H_{i+1}P_{i+1/i}H_i^T + R_{i+1}]^{-1} \tag{9c}$$

$$\hat{d}_{i+1/i+1} = \hat{d}_{i+1/i} + K_{i+1}(u_{i+1} - H_{i+1}\hat{d}_{i+1/i}) \tag{9d}$$

$$P_{i+1/i+1} = P_{i+1/i} - K_{i+1}H_{i+1}P_{i+1/i} \tag{9e}$$

where the subscripts of the form m/n mean the value at step m given the measurements up to and including step n. Note from Eq. (8) that D is needed to compute R_{i+1}, which is used in Eq. (9c) of the algorithm. Having no knowledge of D itself, $\hat{D}_i$ is used instead, which is a feature of the extended Kalman filter (EKF).

The Dynamic Case

Suppose that coordinate system v is rotating with respect to coordinate system u at an angular rate $\bar{\omega}$. Then, the value of D changes between measurements and the algorithm given in Eqs. (9) has to be modified accordingly. Let the components of $\bar{\omega}$, when coordinatized in system v, be denoted by ω_x, ω_y, and ω_z.

It is well known that the rate of change of D is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} D = \Omega D \tag{10}$$

where

$$\Omega = \begin{bmatrix} o & \omega_z & -\omega_y \\ -\omega_z & o & \omega_x \\ \omega_y & -\omega_x & o \end{bmatrix} \tag{11}$$

Equation (10) can be written in the following form, which suits us better:

$$\dot{d} = W(t)d \tag{12}$$

where W is defined as follows. Define

$$W_q = \begin{bmatrix} \omega_q & o & o \\ o & \omega_q & o \\ o & o & \omega_q \end{bmatrix} \qquad q = x,y,z$$

and let O_3 be a 3×3 zero matrix then $W(t)$ is given by

$$W(t) = \begin{bmatrix} O_3 & W_z & -W_y \\ -W_z & O_3 & W_x \\ W_y & -W_x & O_3 \end{bmatrix} \tag{13}$$

Equation (12) describes the dynamic behavior of the estimated column matrix d. The discrete expression for the change in d between measurements is given by

$$d_{i+1} = \phi_i d_i \tag{14}$$

where ϕ_i is the transition matrix which corresponds to $W(t)$. In accordance with Eq. (14), the algorithm of Eqs. (9) can be modified to account for the changing D matrix by the replacement of Eqs. (9a) and (9b) to form another algorithm between measurements as follows:

between measurements:

$$\hat{d}_{i+1/i} = \phi_i \hat{d}_{i/i} \tag{15a}$$

$$P_{i+1/i} = \phi_i P_{i/i} \phi_i^T \tag{15b}$$

Due to inaccuracies encountered in the implementation of the algorithm, it is advisable[13] to add a process-noise covariance matrix to $P_{i+1/i}$ in Eq. (15b). In our case, though, such an addition stems automatically from the nature of the problem; namely, the rotation rate vector contains additive noise. The latter is due to the fact that $\tilde{\omega}$ is not known to us perfectly, but rather is a result of gyro measurements that contain noise. The measured angular velocity vector $\tilde{\omega}$ can be written as

$$\tilde{\omega} = \omega + n_\omega \tag{16}$$

where ω is the correct value and n_ω is an additive noise. Let us assume that n_ω can be characterized as a zero mean white noise column matrix with

$$\mathrm{Cov}\{n_\omega\} = Q_\omega$$

Now Eq. (10) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t} D = -[\tilde{\Omega} - N_\omega] D \tag{17}$$

where

$$N_\omega = \begin{bmatrix} o & n\omega_z & -n_{\omega y} \\ -n_{\omega z} & o & n_{\omega x} \\ n_{\omega y} & -n_{\omega x} & o \end{bmatrix} n_{\omega z} \tag{18}$$

From Eq. (17) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} D = -\tilde{\Omega} D + N_\omega D \tag{19}$$

which can be written as

$$\dot{d} = \tilde{W}(t) d + G n_\omega \tag{20}$$

where $\tilde{W}(t)$ is constructed from $\tilde{\omega}$ in the same way $W(t)$ was constructed from ω and where

$$G = \begin{bmatrix} o & -d_{13} & d_{12} \\ d_{13} & o & -d_{11} \\ -d_{21} & d_{11} & o \\ o & -d_{23} & d_{22} \\ d_{23} & o & -d_{21} \\ -d_{22} & d_{21} & o \\ o & -d_{33} & d_{32} \\ d_{33} & o & -d_{31} \\ -d_{32} & d_{31} & o \end{bmatrix} \tag{21}$$

Note that although $\tilde{W}(t)$ is not equal to the nominal matrix $W(t)$, Eq (20) still describes the true dynamics of d. Thus, a discrete version of Eq. (20) can be used to derive the KF algorithm between measurements. When Eq. (20) is discretized, the following is obtained:

$$d_{i+1} = \tilde{\phi}_i d_i + G_i n_{\omega i} \tag{22}$$

Note that G_i is a function of d that varies between the i and $(i+1)$ measurements. However, in the algorithm we adopt two simplifications, first, and only in the discretization of Gn_ω, we assume d to be constant between measurements. Second, we substitute d by $\hat{d}_i$. Consequently, the continuous expression Gn_ω can be transformed into the discrete form $G_i n_{\omega i}$ where G_i is a function of $\hat{d}_i$ and $n_{\omega i}$ is a sample of a zero mean white noise sequence $\{n_{\omega i}\}$ with

$$\mathrm{cov}\{n_{\omega i}\} = Q_i \tag{23}$$

which can be easily computed[14] using $\tilde{\Omega}$ and Q_ω. The algorithm between measurements follows immediately:

between measurements:

$$\hat{d}_{i+1/i} = \tilde{\phi}_i \hat{d}_{i/i} \tag{24a}$$

$$P_{i+1/i} = \tilde{\phi}_i P_{i/i} \tilde{\phi}_i^T + G_i Q_i G_i^T \tag{24b}$$

The algorithm across measurements is, of course, unaffected by the dynamic nature of the problem and is given in Eqs. (9).

III. Orthogonalized Direction Cosine Matrix Identification

The preceding algorithm can be improved by utilizing the orthogonality property of the DCM. A DCM satisfies the orthogonality condition

$$D^T D = I_3$$

where I_3 is a 3×3 identity matrix. Thus, when $\hat{D}_i$ converges to D_i, the matrix $\hat{D}_i^T \hat{D}_i$ converges to I_3. The convergence of $\hat{D}_i$ to D_i can be accelerated when, in addition to the preceding identification process, $\hat{D}_i$ is also forced to approach orthogonality. One way of doing it is to replace $\hat{D}_{i/i}$ by the orthogonal matrix that is the closest to it in the Euclidean sense. The computation of this particular matrix is rather cumbersome. There are, however, efficient iterative processes to compute the sought matrix. One, rather simple, iterative process is[15]

$$X_0 = \hat{D}_{i/i} \tag{25a}$$

$$X_{j+1} = 1.5X_j - 0.5X_j X_j^T X_j \tag{25b}$$

To improve the preceding algorithms we suggest the application of the first iteration of Eqs. (25) after each update of $\hat{d}_i$. To show how this is integrated into the preceding algorithm, define

$$N_{i+1} = 1.5I_3 - 0.5\hat{D}_{i+1/i+1} \hat{D}_{i+1/i+1}^T \tag{26}$$

and denote the matrix resulting from a single application of Eqs. (25) on $\hat{D}_{i+1/i+1}$ by $\hat{D}^*_{i+1/i+1}$. Then it can be seen that

$$\hat{D}^*_{i+1/i+1} = N_{i+1} \hat{D}_{i+1/i+1} \tag{27}$$

It can be shown that, in terms of the corresponding column matrices $\hat{d}_{i+1/i+1}$ and $\hat{d}^*_{i+1/i+1}$ ($\hat{d}^*_{i+1/i+1}$ corresponds to $\hat{D}^*_{i+1/i+1}$ and $\hat{d}_{i+1/i+1}$ corresponds to $\hat{D}_{i+1/i+1}$), Eq. (27) can be written as

$$\hat{d}^*_{i+1/i+1} = M_{i+1} \hat{d}_{i+1/i+1} \tag{28}$$

where

$$M_{i+1} \triangleq \begin{bmatrix} [n_{11}] & [n_{12}] & [n_{13}] \\ [n_{21}] & [n_{22}] & [n_{23}] \\ [n_{31}] & [n_{32}] & [n_{33}] \end{bmatrix} \tag{29}$$

and where $[n_{ij}]$ is a diagonal matrix whose three elements are the n_{ij} elements of N_{i+1}. The column matrix $\hat{d}^*_{i+1/i+1}$ becomes the latest orthogonalized estimate of d_{i+1}, which is then propagated in time according to Eqs. (24).

The orthogonalization stage interferes with the stochastic identification process. Thus, in order for the process to converge to a minimum variance estimate, the covariance computation has to be modified appropriately, in accordance with the operation indicated in Eq. (28); that is, in parallel to the orthogonalization operation on the estimate itself, a corresponding operation has to be performed on the covariance matrix P.

In the ensuring it will be shown that, although the exact computation which P has to undergo is cumbersome, it can be replaced by another computation that is very easy to implement in the algorithm. (Note that the following discussion is included here in order to explain the considerations which led to the simpler computation and is not a part of the algorithm itself.)

In order to examine the required operation on P, let us investigate how P is changed following the orthogonalization operation performed on the estimate. Denote the estimation error by e, then after a measurement update

$$e_{i+1/i+1} \triangleq d_{i+1} - \hat{d}_{i+1/i+1} \tag{30}$$

and, noting that $e_{i+1/i+1}$ is unbiased

$$P_{i+1/i+1} \triangleq E\{e_{i+1/i+1} e^T_{i+1/i+1}\} \tag{31}$$

After the application of the orthogonalization operation of Eq. (28), the estimation error, which is now denoted by $e^*_{i+1/i+1}$, becomes

$$e^*_{i+1/i+1} = d_{i+1} - \hat{d}^*_{i+1/i+1} \tag{32}$$

or using Eq. (28) in Eq. (32)

$$e^*_{i+1/i+1} = d_{i+1} - M_{i+1}\hat{d}_{i+1/i+1} \tag{33}$$

Adding and subtracing $d_{i+1/i+1}$ to the right-hand side of Eq. (33), $e^*_{i+1/i+1}$ can be written as

$$e^*_{i+1/i+1} = e_{i+1/i+1} + (I_9 - M_{i+1})\hat{d}_{i+1/i+1} \tag{34}$$

where I_9 is the 9×9 identity matrix. Noting that the expected value of $e^*_{i+1/i+1}$ is deterministic, $P^*_{i+1/i+1}$, the covariance matrix of $e^*_{i+1/i+1}$, can be computed as

$$P^*_{i+1/i+1} = E\{e^*_{i+1/i+1} e^{*T}_{i+1/i+1}\} - E\{e^*_{i+1/i+1}\}E\{e^{*T}_{i+1/i+1}\} \tag{35}$$

Substituting Eq. (34) into Eq. (35), the expression for $P^*_{i+1/i+1}$ can be shown to be

$$\begin{aligned} P^*_{i+1/i+1} &= E\{e_{i+1/i+1} e^T_{i+1/i+1}\} \\ &+ E\{e_{i+1/i+1}\hat{d}^T_{i+1/i+1}(I_9 - M_{i+1})^T\} \\ &+ E\{(I_9 - M_{i+1})\hat{d}_{i+1/i+1} e^T_{i+1/i+1}\} \\ &+ E\{(I_9 - M_{i+1})\hat{d}_{i+1/i+1}\hat{d}^T_{i+1/i+1}(I_9 - M_{i+1})^T\} \\ &- E\{(I_9 - M_{i+1})\hat{d}_{i+1/i+1}\}E\{\hat{d}^T_{i+1/i+1}(I_9 - M_{i+1})^T\} \end{aligned} \tag{36}$$

We notice immediately that the first term on the right-hand side of Eq. (36) is $P_{i+1/i+1}$. As M_{i+1} is a function of $\hat{d}_{i+1/i+1}$, the computation of the second, third, and fourth terms is cumbersome, if not impossible; therefore, some approximation is needed. To compute the second and third terms we assume, as our *first* approximation, that M_{i+1} is deterministic, thus, we may write for the second term

$$\begin{aligned} &E\{e_{i+1/i+1}\hat{d}^T_{i+1/i+1}(I_9 - M_{i+1})^T\} \\ &= E\{e_{i+1/i+1}\hat{d}^T_{i+1/i+1}\}(I_9 - M_{i+1})^T \end{aligned}$$

But the optimal estimate and its error are uncorrelated (see Ref. 14, p. 112); therefore, this term vanishes, as well as the third term, which is its transpose. Equation (36) can then be

Table 1 Summary of the orthogonalized DCM identification algorithm

Stage	Algorithm	Reference for variable
Initialization	*State:* $\hat{D}_0$ is computed using the first three measured vectors and then transformed into $\hat{d}_0$ or $\hat{D}_0 = I_3$ and is then transformed into $\hat{d}_0$ *Covariance matrix*: $P_0 = I_9 \cdot a$ (a is a large number)	$\hat{D}_0$ is transformed to $\hat{d}_0$ according to Eq. (5)
Between measurements	*State*: $\hat{d}_{i+1/i} = \tilde{\phi}_i \hat{d}^*_{i/i}$ *Covariance Matrix*: $P_{i+1/i} = \tilde{\phi}_i P^*_{i/i} \tilde{\phi}^T_i + G_i Q_i G^T_i$	For $\tilde{\phi}_i$ see Eqs. (20) and (22). G_i in Eq. (21) Q_i in Eq. (23)
Across a measurement	*Gain*: $K_{i+1} = P_{i+1/i} H^T_{i+1}(H_{i+1} P_{i+1/i} H^T_{i+1} + R_{i+1})^{-1}$ *State*: $\hat{d}_{i+1/i+1} = \hat{d}_{i+1/i} + K_{i+1}(u_{i+1} - H_{i+1}\hat{d}_{i+1/i})$ *Covariance*: $P_{i+1/i+1} = P_{i+1/i} - K_{i+1} H_{i+1} P_{i+1/i}$	H_{i+1} in Eq. (4) R_{i+1} in Eq. (8)
Orthogonalization	*State*: $\hat{d}^*_{i+1/i+1} = M_{i+1}\hat{d}_{i+1/i+1}$ *Covariance*: $P^*_{i+1/i+1} = P_{i+1/i+1} + (I_9 - M_{i+1})\hat{d}_{i+1/i+1}\hat{d}^T_{i+1/i+1}(I_9 - M_{i+1})^T$	M_{i+1} in Eqs. (26) and (29)

written as

$$P^*_{i+1/i+1} = P_{i+1/i+1} + E\{(I_9 - M_{i+1})\hat{d}_{i+1/i+1}\hat{d}^T_{i+1/i+1}(I_9 - M_{i+1})^T\} - E\{(I_9 - M_{i+1})\hat{d}_{i+1/i+1}\}E\{\hat{d}^T_{i+1/i+1}(I_9 - M_{i+1})^T\} \quad (37)$$

Before considering cumbersome algorithms for computing the second and third elements on the right-hand side of Eq. (37), Monte Carlo simulation runs were made determining that the elements of the third term are by an order of magnitude smaller than those of the second term; therefore, as a *second* approximation, the third term is dropped altogether. Again using our first approximation, we can then write the second term as

$$E\{(I_9 - M_{i+1})\hat{d}_{i+1/i+1}\hat{d}_{i+1/i+1}(I_9 - M_{i+1})^T\} = (I_9 - M_{i+1})E\{\hat{d}_{i+1/i+1}\hat{d}^T_{i+1/i+1}\}(I_9 - M_{i+1})^T \quad (38)$$

Several recursive algorithms for computing $E\{\hat{d}_{i+1/i+1}\hat{d}^T_{i+1/i+1}\}$ were considered. They were all too complicated to be included in a viable algorithm for computing d, therefore a *third* approximation was made according to which

$$E\{\hat{d}_{i+1/i+1}\hat{d}^T_{i+1/i+1}\} = \hat{d}_{i+1/i+1}\hat{d}^T_{i+1/i+1} \quad (39)$$

In this approximation we substitute the expected value of the quantity $\hat{d}_{i+1/i+1}\hat{d}^T_{i+1/i+1}$ by the quantity itself. In so doing, we assume that the random quantity $\hat{d}_{i+1/i+1}\hat{d}^T_{i+1/i+1}$ is not too far from its mean. This assumption becomes more and more justified as the algorithm converges and $\hat{d}_{i+1}$ approaches d_{i+1}. When Eq. (39) is substituted into Eq. (38) and the result is used in Eq. (37), the latter becomes

$$P^*_{i+1/i+1} = P_{i+1/i+1} + (I_9 - M_{i+1})\hat{d}_{i+1/i+1}\hat{d}^T_{i+1/i+1}(I_9 - M_{i+1})^T \quad (40)$$

As will be shown in the next section, this simple algorithm for computing $P^*_{i+1/i+1}$ yields very good results. This simple expression is the sole computation that the covariance matrix has to undergo due to the application of the orthogonalization cycle. A summary of the Orthogonalized DCM Identificational algorithm is given in Table 1. The simplicity of the added orthgonalization computation is obvious.

IV. Monte Carlo Simulation Results

To examine the efficiency of the algorithm with and without orthogonalization, two figures of merit were defined. The first figure of merit, denoted by J^*_{i+1}, is called *convergence index* and is defined as

$$J^*_{i+1} = \text{trace}[(\hat{D}^*_{i+1/i+1} - D_{i+1})^T(\hat{D}^*_{i+1/i+1} - D_{i+1})] \quad (41a)$$

The covergence index is equal to the sum of the square of the difference between the elements of $\hat{D}^*_{i+1/i+1}$ and $D_{i+1/i+1}$. Obviously, this index ia always positive and becomes zero only when $\hat{D}^*_{i+1/i+1}$ convergence exactly to D_{i+1}. For comparison between the two algorithms (i.e., with and without orthogonalization), we define J_{i+1}, the convergence index when no orthogonalization is applied, as

$$J_{i+1} = \text{trace}[(\hat{D}_{i+1/i+1} - D_{i+1})^T(\hat{D}_{i+1/i+1} - D_{i+1})] \quad (41b)$$

In addition, we define the *orthogonality index* F^*_{i+1} as follows

$$F^*_{i+1} = \text{trace}[(\hat{D}^{*T}_{i+1/i+1}\hat{D}^*_{i+1/i+1} - I_3)^T \times (\hat{D}^{*T}_{i+1/i+1}\hat{D}^*_{i+1/i+1} - I_3)] \quad (42a)$$

and similarly

$$F_{i+1} = \text{trace}[(\hat{D}^T_{i+1/i+1}\hat{D}_{i+1/i+1} - I_3)^T \times (\hat{D}^T_{i+1/i+1}\hat{D}_{i+1/i+1} - I_3)] \quad (42b)$$

The indices F^*_{i+1} and F_{i+1} express the closeness of $\hat{D}^*_{i+1/i+1}$, and $\hat{D}_{i+1/i+1}$, respectively, to orthogonality. They are always positive and become zero only when the respective matrices are orthogonal.

The data presented in the ensuring were obtained from Monte Carlo simulation runs. In each run, the initial DCM D_0 was

$$D_0 = \begin{bmatrix} 0.875000 & 0.433013 & -0.216506 \\ -0.216506 & 0.750000 & 0.625000 \\ 0.433013 & -0.500000 & 0.750000 \end{bmatrix}$$

The nominal angular rate was a constant vector whose components in body axes were 0.628 rad/s along each axis. The gyro measurement noises [see Eq. (16] that constituted n_ω were three zero mean white noise components whose spectral density was $0.01°/h^{1/2}$. [Note that in reality the DCM dynamics were noiseless while the dynamics used in the estimator were noisy, i.e., in the filter we had to use $\tilde{W}(t)$ rather than $W(t)$.] A measurement pair, u_i and v_i, was obtained and processed every 0.1 s. The measurement noises of u_i and v_i were zero mean and white. Their standard deviation corresponded to a random angular error of 100 arc-s.

The initial estimate of d was obtained as follows. The first three measured pairs u_i and v_i where grouped into the following matrix equation that defined $\hat{D}_0$, the estimate of the initial DCM:

$$[v_1, v_2, v_3] = \hat{D}_0[u_1, u_2, u_3] \quad (43)$$

When u_1, u_2, and u_3 were linearly independent, Eq. (43) could be solved to yield

$$\hat{D}_0 = [v_1, v_2, v_3][u_1, u_2, u_3]^{-1} \quad (44)$$

and $\hat{d}_0$ was obtained from $\hat{D}_0$ according to the operation spelled out by Eq. (5).

One hundred Monte Carlo simulation runs were made with these data using 200 measurements. Estimates of the expected values of the convergene indices and of the orthogonality indices were computed for each measurement point i using

$$\bar{J}^*_i \triangleq E\{J^*_i\} \sim \frac{1}{100}\sum_{k=1}^{100} J^*_i(k) \quad (45a)$$

$$\bar{J}_i \triangleq E\{J_i\} \sim \frac{1}{100}\sum_{k=1}^{100} J_i(k) \quad (45b)$$

$$\bar{F}^*_i \triangleq E\{F^*_i\} \sim \frac{1}{100}\sum_{k=1}^{100} F^*_i(k) \quad (46a)$$

$$\bar{F}_i \triangleq E\{F_i\} \sim \frac{1}{100}\sum_{k=1}^{100} F_i(k) \quad (46b)$$

Plots of the convergence indices $\bar{J}^*_i$ and $\bar{J}_i$ are presented in Fig. 1. The superiority of the orthogonalized DCM identification algorithm, which yields $\bar{J}^*_n$, is obvious. Moreover, as we might expect, the orthogonality index of this algorithm is better than that of the other algorithm by about nine orders of magnitude. This is clearly demonstrated in Fig. 2 where $\bar{F}^*_i$ and $\bar{F}_i$ are plotted.

Another way to initialize $\hat{d}$ was examined. $\hat{D}_0$ was simply set to the identity matrix I and then, using Eq. (5), this $\hat{D}_0$ was

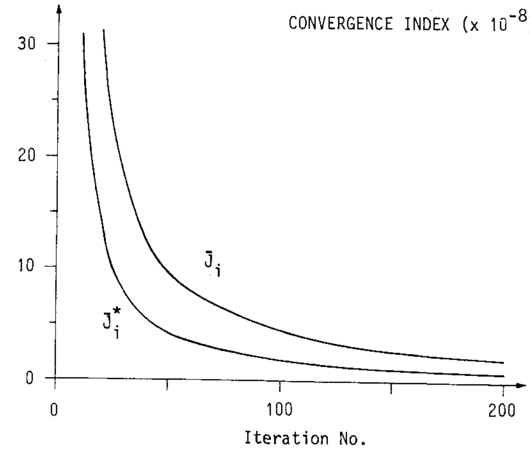

Fig. 1 Ensemble average of the convergence indices of the two DCM identification algorithms.

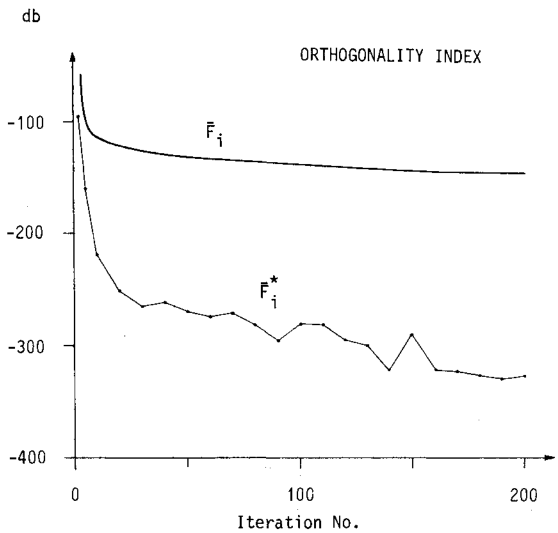

Fig. 2 Ensemble average of the orthogonality indices of the two DCM identification algorithms.

transformed into $\hat{d}_0$. One hundred Monte Carlo simulation runs were made using both algorithms. It was found that, although initially the convergence indices were worse, by far, than the corresponding indices shown in Fig. 1, eventually they, almost always, reached the corresponding values. Since the convergence rate when $\hat{D}_0 = I$, is slower than the convergence rate when $\hat{D}_0$ is calculated using Eq. (44), and since convergence is not always assured, it is recommended that the latter filter initialization method be used.

V. Conclusions

Two algorithms to estimate the direction cosine matrix using two sets of vector measurements were presented. Both algorithms are based on a stochastic parameter identification process of linear systems. The first algorithm is a result of a straightforward application of the identification process. The second algorithm uses, in addition, the orthogonality property of the direction cosine matrix to yield a faster converging algorithm, which also produces an orthogonal matrix.

A series of Monte Carlo simulation runs was performed. These runs produced the ensemble average of two indices that quantify the convergence and orthogonality qualities of the algorithms. Two methods for obtaining an initial estimate of the direction cosine matrix were considered. One method was to assume that the initial direction cosine matrix was the identity matrix. This initialization method requires no additional computation; on the other hand, two (out of the few hundred) of the simulation runs diverged. The second estimate initialization used the first three pairs of vector measurements to obtain the initial estimate. This method requires a matrix inversion; however, since the matrix is a 3×3 matrix, this constitutes no problem. With the second intialization method the estimate converged, at the very outset, quite rapidly to the correct direction cosine matrix.

In view of these findings, it is recommended that the orthogonalized direction cosine matrix identification version be used and the initial estimate be obtained from the first three pairs of vector measurement. This choice yields a rapidly converging orthogonal estimate of the direction cosine matrix.

References

[1] Wahba, G., "Problem 65-1, A Least Squares Estimate of Satellite Attitude," *SIAM Review,* Vol. 7, 1965, p. 409.

[2] Wahba, G. et al., "Problem 65-1 (Solution)," *SIAM Review,* Vol. 8, 1966, pp. 384-386.

[3] Brock, J.E., "Optimal Matrices Describing Linear Systems," *AIAA Journal,* Vol. 6, July 1968, pp. 1292-1296.

[4] Carta, D.G. and Lackowski, D.H., "Estimation of Orthogonal Transformations in Strapdown Inertial Systems," *IEEE Transactions on Automatic Control,* Vol. AC-17, Feb. 1972, pp. 97-100.

[5] Shuster, M.D. and Oh, S.D., "Three Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control,* Vol. 4, Jan.-Feb. 1981, pp. 70-77.

[6] Lerner, G.M., "Three-Axis Attitude Determination," *Spacecraft Attitude Determination and Control,* edited by J.R. Wertz, D. Reidel Publishing Co., Dordrecht, the Netherlands, 1978, pp. 420-428.

[7] Shuster, M.D., "Approximate Algorithms for Fast Optimal Attitude Computation," *Proceedings of AIAA Guidance and Control Conference,* Palo Alto, Calif., Aug. 7-9, 1978, paper 78-1249.

[8] Podgorski, W.A., Lemos, L.K., Cheng, J., and Daly, K.C., "Gyroless Attitude Determination and Control System for Advanced Environmental Satellites," *Proceedings of AIAA Guidance and Control Conference,* San Diego, Calif., Aug. 9-11, 1982, Paper 82-1614.

[9] Doty, R.L. and Nease, R.F., "Initial Conditions and Alignment," *Inertial Gudiance,* edited by G.R. Pitman Jr., John Wiley & Sons, New York, 1962, Chap. 8.

[10] Jurenka, F.D. and Leondes, C.T., "Optimum Alignment of Inertial Autonavigator," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. AES-2, Nov. 1967, pp. 880-888.

[11] Crooker, E.B. and Robins, L., "Application fo Kalman Filtering Techniques to Strapdown System Initial Alignment," *Theory and Application of Kalman Filtering,* AGARDograph 139, Feb. 1970, Ch. 7.

[12] Mayne, D.Q., "Optimal Non-Stationary Estimation of the Parameters of a Linear System with Gaussian Inputs," *Journal of Electronics and Control,* Vol. 14, Jan. 1963, pp. 101-112.

[13] Anderson, B.D.O. and Moore, J.B., *Optimal Filtering,* Prentice-Hall, Englewood Cliffs, N.J., 1979, pp. 50-51.

[14] Gelb, A., *Applied Optimal Estimation,* MIT Press, Cambridge, Mass., 1974, pp. 298-299.

[15] Bar-Itzhack, I.Y. and Meyer, J., "On the Convergene of Iterative Orthogonalization Processes," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. AES-12, March 1976, pp. 146-151.