APPENDIX A

A DERIVATION OF THE DISCRETE-TIME "CENTRALIZED" KALMAN FILTER VIA THE MATRIX MINIMUM PRINCIPLE

A.1 Introduction

An application of the Matrix Minimum Principle (Ref. 1), a well-known tool for the derivation of the continuous-time Kalman filter (Ref. 2), is presented here for an exclusively discrete-time derivation of the discrete-time Kalman filtering equations. This application is demonstrated here for the following reasons:

* The approach of Refs. 3–9 to decentralized filtering is via utilization of the continuous-time version of the Matrix Minimum Principle, where the decentralization requirement is essentially included in the routine analysis as an additional structural constraint, and the filter gains are then optimized to provide the minimum variance solutions for the prescribed filter structures.

* The results of Refs. 3–9 for decentralized filtering are stated and derived only within a continuous-time framework (discrepancies that can arise between discretization of continuous-time results and exclusively discrete-time results are noted and explained on pp. 136–139 of Ref. 10, on p. 297 of Ref. 38, and other difficulties discussed in Ref. 11); however, a practical application for JTIDS ReINav appears to be better suited to the discrete-time formulation [fn. A-1] provided below.

* The ten (10) alternative decentralized filter structures considered in Ref. 9 are obtained by a different (but repeated) application of the continuous-time Matrix Minimum Principle to estimation problems [fn. A-2], so a simple
exposition of the pertinent steps (as provided below) should be useful as an instructive guide.

* The approaches of Refs. 1 to 9 neither correctly acknowledge nor properly accommodate the additional slight analytical difficulties now known (since 1977) to arise (Refs. 15, 16) in taking gradients with respect to symmetric matrices (occurring naturally in these particular applications), and in obtaining the canonical forward and backward equations associated with applying the Matrix Minimum Principle.

* The milestone application of the Matrix Minimum Principle to filtering problems (Ref. 2) unfortunately provides no clarification for why the particular cost function of just a terminal accuracy constraint was chosen. Use of the same terminal accuracy constraint persisted without clarification throughout Refs. 2, 3-9, 17, but is questioned and correctly modified [fn. A-3] both in Refs. 18, 19, 21, 22, and here.

* The canonical equations that arise in the continuous-time application of the Matrix Minimum Principle are two "coupled" [fn. A-4] matrix differential equations that comprise a Two Point Boundary Value Problem (TPBVP). Once the appropriate differential equations are specified, a solution procedure has to also be supplied for the continuous-time formulation. However, in the discrete-time application of the Matrix Minimum Principle, the canonical equations of the TPBVP are matrix difference equations, which when iterated provide a direct solution.

* There (apparently) are few previous demonstrations (to this author's knowledge, excepting Ref. 19) of the discrete-time Matrix Minimum Principle being applied to a discrete-time filter formulation, perhaps being due to the four case ambiguity that can arise between use of $P(k|k-1)$ or $P(k|k)$ and $\hat{x}(k|k-1)$ or $\hat{x}(k|k)$ in attempting to formulate a meaningful problem statement [fn. A-5]. This void in discrete-time formulations points out the need for a clear, updated exposition for this application.
A.2 Problem Statement

The following notation is used throughout this derivation.

System:

\[ x(k+1) = \phi(k+1,k)x(k) + w(k) \quad (A.2-1) \]

Measurements:

\[ z(k) = H(k)x(k) + v(k) \quad (A.2-2) \]

The process and measurement noises \( w(k) \) and \( v(k) \) are assumed to be independent, zero mean, white Gaussian noises having associated covariance matrices \( Q(k) \) and \( R(k) \), respectively, and uncorrelated with the Gaussian initial condition

\[ x(0) \sim N(\emptyset, P) \quad (A.2-3) \]

It is also assumed that \( R(k) \) is non-singular

\[ R(k) = R^T(k) > 0 \quad \text{for all} \quad 0 \leq k \quad (A.2-4) \]

Assume a filter for estimating the state \( x(k) \) (via \( x(k/k-1) \)) to be of the following linear form

\[ x(k+1|k) = \Gamma(k+1,k)x(k|k-1) + \chi(k)z(k) \quad (A.2-5) \]

Objectives:

1. Specify conditions for an unbiased estimator to reveal the structure of \( \Gamma(k+1,k) \) and appropriate initial conditions for Eq. A.2-5 to ensure that the estimator that evolves is both conditionally and unconditionally unbiased as, respectively, indicated by [fn. A-6]

\[ E[\hat{x}(k+1|k)|Z(k)]] = E[x(k+1)|Z(k)] \quad (A.2-6) \]
(c.f., p. 159 of Ref. 23) and

$$E[\hat{x}(k+1)] = E[x(k)]$$  \hspace{1cm} (A.2-7)

where $E[.]$ is the expectation operator.

2. Use the discrete-time version of the Matrix Minimum Principle to specify the optimal filter gains $[K(k)]$ in Eq. A.2-5 by considering the consequential effect on the variance of estimation error.

A.3 Filter Structure Due to Allowing Only Unbiased Estimates

Let the error of estimation of interest in this discrete-time formulation be

$$e(k+1|k) \triangleq x(k+1) - \hat{x}(k+1|k)$$  \hspace{1cm} (A.3-1)

Substituting for $x(k+1)$ and $\hat{x}(k|k)$ in Eq. A.3-1 using A.2-1 and A.2-5, respectively, yields

$$e(k+1|k) = \phi(k+1,k)x(k) + w(k) - \Gamma(k+1,k)\hat{x}(k|k-1) - K(k)z(k)$$  \hspace{1cm} (A.3-2a)

$$= \phi(k+1,k)x(k) + w(k) - \Gamma(k+1,k)\hat{x}(k|k-1) - K(k)\left[H(k)x(k) + \nu(k)\right]$$

$$+ \left[\Gamma(k+1,k)\hat{x}(k) + \Gamma(k+1,k)x(k)\right] = 0, \text{ but introduced for convenience}$$  \hspace{1cm} (A.3-2b)

$$= \Gamma(k+1,k)e(k|k-1) + \left[\phi(k+1,k) - \Gamma(k+1,k) - K(k)H(k)\right]x(k) + \left[w(k) - K(k)\nu(k)\right]$$  \hspace{1cm} (A.3-2c)

where Eq. A.2-2 was used to substitute for $\tilde{z}(k)$ in Eq. A.3-2b and the definition of $e(k|k-1)$ was used from Eq. A.3-1 in Eq. A.3-2c.

The condition of the estimator $\hat{x}(k+1|k)$ being conditionally unbiased requires that

$$E[x(k+1) - \hat{x}(k+1|k) | z(k)] = 0 \text{ for all } k \geq 0$$  \hspace{1cm} (A.3-3a)
which is equivalent to

$$E[e(k+1|k)|Z(k)] = 0 \text{ for all } k \geq 0$$  \hspace{1cm} (A.3-3b)

Taking expectations conditioned [fn. A-7] on $Z(k)$ throughout Eq. A.3-2c yields

$$E[e(k+1|k)|Z(k)] = \Gamma(k+1,k) E[e(k|k-1)|Z(k)] + [\phi(k+1) - \Gamma(k+1,k) - K(k)H(k)]E[x(k)|Z(k)] + \underbrace{E[w(k) - \Gamma(k) \nu(k)|Z(k)]}_{0}$$  \hspace{1cm} (A.3-4)

where the last term on the right side of Eq. A.3-4 is zero since the noises are of zero mean.

Conditions (obtained by examining the time evolution of Eq. A.3-4) to guarantee that the estimator of Eq. A.2-5 satisfies a conditional and unconditional [fn. A-8] unbiasedness requirement (the conditional unbiasedness requirement being the more stringent) of Eqs. A.3-3 and A.2-7, respectively, are:

1. For the driving term in Eq. A.3-4 to be identically zero

2. For the expected value of Eq. A.3-4 as [fn. A-9]

$$E[e(k+1|k)] = \Gamma(k+1,k) E[e(k|k-1)] + [\phi(k+1) - \Gamma(k+1,k) - K(k)H(k)]E[x(k)]$$  \hspace{1cm} (A.3-5)

to have zero initial condition (consistent with Eq. A.2-7 at $k=0$), hence

$$E[e(k+1|k)] = 0 \text{ for all } k \geq 0$$  \hspace{1cm} (A.3-6)

as obtained by repeated iteration of Eq. A.3-5.

The above two conditions can be represented explicitly as

$$[\phi(k+1,k) - \Gamma(k+1,k) - K(k)H(k)]E[\hat{x}(k)|Z(k)] \equiv 0$$  \hspace{1cm} (A.3-7)

and

$$0 = E[e(0|1)] = E[x(0)] - E[\hat{x}(0|1)]$$  \hspace{1cm} (A.3-8)
respectively. Since \( \mathbb{E}[x(k)|Z(k)] \) cannot realistically be assumed to always be zero, Eq. A.3-7 is satisfied only when

\[
\Gamma(k+1,k) = \Phi(k+1,k) - K(k)H(k) \tag{A.3-9}
\]

The condition of Eq. A.3-8 is satisfied when

\[
\hat{x}(0|-1) = \mathbb{E}[\hat{x}(0)] = 0 \tag{A.3-10}
\]

Thus, providing both conditional and unconditional unbiased estimates restricts Eq. A.2-5 to have the following structure

\[
\hat{x}(k+1|k) = [\Phi(k+1,k) - K(k)H(k)]\hat{x}(k|k-1) + K(k)z(k) \tag{A.3-11}
\]

with \( \hat{x}(0|-1) \) as specified in Eq. A.3-10.

### A.4 Specifying the Optimal Gain Via the Discrete-Time Matrix Minimum Principle

Now that the structure of \( \Gamma(k+1,k) \) appearing in Eq. A.2-5 has been specified in Eq. A.3-9 to yield an unbiased estimator of the form of Eq. A.3-11, it still remains to specify the gain \( \{K(k)\} \). This time-varying filter gain \( \{K(k)\} \) is specified via the following three steps.

**Step 1:**

Obtaining the difference equation for the time evolution of the covariance of the estimation error

\[
P(k+1|k) = \mathbb{E}[(e(k+1|k) - \mathbb{E}[e(k+1|k)](e(k+1|k)
\]

\[
- \mathbb{E}[e(k+1|k)])^T] \tag{A.4-1a}
\]

\[
= \mathbb{E}[e(k+1|k)e^T(k+1|k)] \tag{A.4-1b}
\]

**Step 2:**

Specifying the appropriate scalar performance measure or cost function to be the weighted-mean-squared error in estimation consisting of

\[
L(k) = \mathbb{E}[e^T(k|k-1)M(k)e(k|k-1)] \tag{A.4-2a}
\]
\[ = \text{tr}[M(k) \ E[e(k|k-1)e^T(k|k-1)]] \]  
(A.4-2b)

\[ = \text{tr}[M(k) \ P(k|k-1)] \]  
(A.4-2c)

for any fixed [fn. A-10] arbitrary \(\{M(k)\}\) such that

\[ M(k) = M^T(k) > 0 \quad \text{for} \ k \geq 0 \]  
(A.4-3)

(i.e., \(M(k)\) is positive definite.)

**Step 3:**

Minimizing a cost function \(J\) consisting of terms of the form of \(L(k)\) in Eq. A.4-2c [fn. A-11] (according to the ample historical precedents of Refs. 2-9) with respect to the filter gain matrices \(K(k)\) via a convenient approach using the Matrix Minimum Principle.

According to Step 1, from Eq. A.3-2c (with the condition of Eq. A.3-8 in force) yields the following covariance equation

\[ P(k+1|k) = E[e(k+1|k)e^T(k+1|k)] \]  
(A.4-4a)

\[ = [\Phi(k+1,k)-K(k)H(k)]P(k|k-1)[\Phi(k+1,k)-K(k)H(k)]^T + Q(k) \]

\[ + K(k)R(k)K^T(k) \]  
(A.4-4b)

with

\[ P(0|-1) = E[e(0|-1)e^T(0|-1)] = E[(x(0)-\hat{x}(0|-1))(x(0)-\hat{x}(0|-1))^T] = P \]  
(A.4-5)

Using the results of Step 2, proceed to form the cost function (using the results of Eq. A.4-2c) as

\[ J = \text{tr} \left[ \sum_{k=0}^{N_0-1} L_k M(k) P(k|k-1) \right] + \text{tr} \left[ M(N_0) P(N_0|N_0-1) \right] \]

(A.4-6a)

\[ = \text{tr} \left[ \sum_{k=0}^{N_0-1} \sum_{n=-N_0}^{N_0} M(k) P(k|k-1) \right] + \text{tr} \left[ M(N_0) P(N_0|N_0-1) \right] \]  
(A.4-6b)
Throughout this analysis, the impact of including the additional new [fn. A-12] terms in Eq. A.4-6b on subsequent equations will be identified by enclosure within a box. Therefore, results consistent with previous formulations are available at a glance by discarding the boxed terms.) From the cost function of Eq. A.4-6b and the dynamical evolution constraint of Eq. A.4-4b, application of the discrete-time Matrix Minimum Principle (p. 597 of Ref. 1 in complete analogy to the discrete-time vector formulation on pp. 132-133 of Ref. 10) requires formation of a scalar Hamiltonian function as

$$H^*_k(k), P(k | k-1), \Lambda(k+1), k) \delta \text{tr} \left\{ \left[ \phi - K(k) H(k) \right] P(k | k-1) \left[ \phi - K(k) H(k) \right]^T + Q(k) + K(k) R(k) K(k)^T \right\} + \text{tr} \left( M(k) P(k | k-1) \right) \right)$$

(A.4-7)

where

$$\Lambda (k+1)$$ is the n xn costate matrix.

According to the discrete-time formulation of the Matrix Minimum Principle (p. 597 of Ref. 1), the canonical equations that describe the necessary conditions for a minimum are:

$$O = \frac{\partial H}{\partial k} (K, P(k | k-1), \Lambda(k+1), k) \mid_\star$$

(A.4-8)

(where the \( \star \) denotes evaluation at the minimum).

$$P^*_k(k+1 | k) = \frac{\partial^2 H}{\partial k^2} (K, P(k | k-1), \Lambda(k+1), k) \mid_\star \left[ \phi - K(k) H(k) \right] P^*_k(k | k-1) \left[ \phi - K(k) H(k) \right]^T + Q(k) + K^*_k(k) R(k) K^*_k(k)$$

(A.4-9)

If \( P(k | k-1) \) were not a symmetric matrix, the costate would evolve according to [fn. A-13]

$$\Lambda^*_k(k) = -\frac{\partial H}{\partial P} (K, 0, P, \Lambda(k+1), k) \mid_\star$$

(A.4-10a)

$$= \delta \Lambda^*_k(k+1) + \phi^T_k H^*_k(k) \Lambda^*_k(k+1) + \phi^T_k \Lambda^*_k(k+1) K^*_k H(k) + H^*_k(k) K^*_k H(k) + H^*_k(k) K^*_k(k)^T$$

(A.4-10b)

$$= \left[ \phi(k+1, k) - K^*_k(k) H(k) \right]^T \Lambda^*_k(k+1) \left[ \phi(k+1, k) - K^*_k(k) H(k) \right] + M(k)$$

(A.4-10c)
with transversality (i.e., boundary) condition

\[ \Lambda^*(N_0) = \frac{3}{\partial P} \text{tr} \left[ M(N_0) P \right] = M^T(N_0) = M(N_0) > 0 \]  

(A.4-11)

Because \( P(k|k-1) \) is symmetric, the following form is indicated in Refs. 15, 16 for the evolution of the costate matrix

\[ \Lambda^*(k) + \Lambda^T(k) - \text{diag}(\Lambda^*(k)) = \frac{3}{\partial P_{\text{symmetric}}} \text{tr} \left[ M(k), P_{\text{symmetric}}, \Lambda(k+1), k \right] \]  

(A.4-12a)

\[ = \left[ \phi(k+1,k) - K(k) H(k) \right]^T \Lambda^*(k+1) \left[ \phi(k+1,k) - K(k) H(k) \right] + M(k) \]  

\[ + \left[ \phi(k+1,k) - K(k) H(k) \right] \left[ \phi(k+1,k) - K(k) H(k) \right]^T \Lambda^T(k+1) \]  

\[ - \text{diag}(\phi(k+1,k) - K(k) H(k))^T \Lambda^*(k) \left[ \phi(k+1,k) - K(k) H(k) \right] + \text{diag} M(k) \]  

(A.4-12b)

with boundary condition

\[ \Lambda^*(N_0) + \Lambda^T(N_0) - \text{diag}(\Lambda^*(N_0)) = \frac{3}{\partial P(N_0), \text{symmetric}} \text{tr} \left[ M(N_0), P(N_0), \text{symmetric} \right] \]  

(A.4-13a)

\[ = M(N_0) + M^T(N_0) - \text{diag} (M(N_0)) \]  

(A.4-13b)

It has also been noted on p. 996 of Ref. 16 (as can easily be substantiated by substitution) that the solution of Eq. A.4-10c with final condition as Eq. A.4-11 is one solution to Eq. A.4-12b with final condition as Eq. A.4-13b. Therefore, attention here will be focused on Eqs. A.4-10c and A.4-11 for describing the backwards evolution of the costate matrix.

Upon performing the indicated matrix gradients in Eq. A.4-8 (Refs. 31,32,33), the result is

\[ 0 = \Lambda^T(k+1) - \phi(k+1,k) P^* (k|k-1) H^T(k) + K^*(k) \left[ H(k) P^*(k|k-1) H^T(k) + R(k) \right] \]  

\[ + \Lambda^*(k+1) - \phi(k+1,k) P^* (k|k-1) H^T(k) + K^*(k) \left[ H(k) P^*(k|k-1) H^T(k) + R(k) \right] \]  

(A.4-14)
denoted as the "coupling equation" in Ref. 2.

The following arguments proceed as in Refs. 2, 9 (but with more justification as will be indicated parenthetically now that the boxed terms are included in Eqs. A.4-6 to A.4-12). Note from the assumption of Eq. A.4-3 and the form of Eqs. A.4-10c and A.4-11 that the $\Lambda(k)$ that evolves backward in time from $k=N_0$ to $k=0$ is symmetric and positive definite. Without the boxed terms present in this analysis, it is more difficult (if not impossible) to rigorously establish that $\Lambda(k)$ is positive definite since it is not known for sure that the matrix quantity

$$\begin{vmatrix} \phi(k+1,k) - K^*(k)H(k) \end{vmatrix}$$

(A.4-15)

that pre- and postmultiplies the first term on the right side of Eq. A.4-10c is guaranteed to be of full rank. While the form of the matrix quantity in Eq. A.4-15 is reminiscent of the feedback matrix of the Kalman filter (which is known to be nonsingular and thus of full rank), it is yet to be established that the quantity $K^*(k)$ appearing in Eq. A.4-15 is in fact the Kalman gain [fn. A-14]. Because of the presence of the boxed term in Eq. A.4-10c, it can be concluded that the $\Lambda(k)$ evolving from Eqs. A.4-10c and A.4-11 is symmetric and positive definite (independent of the matrix quantity of Eq. A.4-15 being of full rank).

Since $\Lambda(k)$ is positive definite, a consequence is that:

"$\Lambda^{-1}(k)$ exists for $0 \leq k \leq N_0$ and is symmetric"

Thus (as done in Refs. 2, 3-9), Eq. A.4-14 can be premultiplied throughout by $\Lambda^{-1}(k)$ to result in

$$0 = 2 \phi(k+1,k)P^*(k|k-1)H^T(k) + 2K^*(k)H(k)P^*(k|k-1)H^T(k) + R(k)$$

(A.4-16a)

or

$$K^*(k) = (k+1,k)P^*(k|k-1)H^T(k)(H(k)P(k|k-1)H^T(k) + R(k))^{-1}$$

(A.4-16b)

There is no worry about

$$H(k)P(k|k-1)H^T(k) + R(k)$$

(A.4-17)
not being invertible since \( R(k) \) alone is invertible via the assumption of Eq. A.2-4.

Substituting the optimal gain of Eq. A.4-16b back into Eq. A.4-9 yields

\[
P^*(k+1|k) = [\phi - \phi P^* H^T (HP^* H^T + R)^{-1} H] P^*[\phi - \phi P^* H^T (HP^* H^T + R)^{-1} H]^T
\]

\[
+ Q + \phi P^* H^T (HP^* H^T + R)^{-1} R (HP^* H^T + R)^{-1} HP^* \phi^T\tag{A.4-18a}
\]

\[
= \phi [I - \phi P^* H^T (HP^* H^T + R)^{-1} H] P^*[I - \phi P^* H^T (HP^* H^T + R)^{-1} H]^T \phi^T
\]

\[
+ Q + \phi P^* H^T (HP^* H^T + R)^{-1} R (HP^* H^T + R)^{-1} HP^* \phi^T
\]

\[
= \phi P^* \phi^T - \phi P^* H^T (HP^* H^T + R)^{-1} HP^* \phi^T
\]

\[
+ \phi P^* H^T (HP^* H^T + R)^{-1} HP^* \phi^T = \phi P^* \phi^T + Q + \phi P^* H^T (HP^* H^T + R)^{-1} HP^* \phi^T
\]

\[
= \phi P^* \phi^T - \phi P^* H^T (HP^* H^T + R)^{-1} HP^* \phi^T + Q \tag{A.4-18c}
\]

\[
+ \phi P^* H^T (HP^* H^T + R)^{-1} (HP^* H^T + R)^{-1} HP^* \phi^T + Q
\]

\[
= \phi P^* \phi^T - \phi P^* H^T (HP^* H^T + R)^{-1} HP^* \phi^T + Q
\]

\[
= \phi P^* (k+1,k) H^T(k) + P^* -1 (k|k-1)^T \phi(k+1,k) + Q(k) \tag{A.4-18g}
\]

\[
= \phi(k+1,k) [H^T(k) R^T(k) H(k) + P^* -1 (k|k-1)]^{-1} \phi(k+1,k) + Q(k) \tag{A.4-18h}
\]
The structure of Eqs. A.4-18g,h is recognized to be the standard Riccati Equation that arises in optimal "centralized" Kalman filtering.

As a consequence of Theorem 5.4 on p. 171 of Ref. 23, the optimal filtered estimate $\hat{x}(k|k)$ and corresponding covariance matrix $P(k|k)$ are related to the predicted estimate and covariance, respectively, by the following equations

$$\hat{x}(k+1|k) = \phi(k+1,k)\hat{x}(k|k)$$ (A.4-19)

and

$$P(k+1|k) = \phi(k+1,k)P(k|k)\phi^T(k+1,k)+Q(k)$$ (A.4-20)

Therefore, rearranging Eq. A.4-19 and substituting the result of Eq. A.3-11 yields

$$\hat{x}(k|k) = [\phi(k+1,k)]^{-1}\hat{x}(k+1|k)$$ (A.4-21a)

$$= \phi^{-1}(k+1,k) [ \phi(k+1,k)-K^*(k)H(k) ]\hat{x}(k|k-1)+\phi^{-1}(k+1,k)K^*(k)z(k)$$ (A.4-21b)

$$= [I-\phi^{-1}(k+1,k)K^*(k)H(k)]\hat{x}(k|k-1)K^*(k)z(k)$$ (A.4-21c)

Factoring Eq. A.4-4b yields the following form

$$P(k+1|k) = \phi(k+1,k)(I-\phi^{-1}(k+1,k)K^*(k)H(k))P(k|k-1)(I-\phi^{-1}(k+1,k)K^*(k)H(k))^T+\phi^{-1}(k+1,k)K^*(k)R(k)K^T(k)\phi^{-1}(k+1,k)$$ (A.4-22)

which by similarity to the form of Eq. A.4-20 yields the following association

$$P(k|k) = [I-K(k)H(k)]P(k|k-1)[I-K(k)H(k)]^T+K(k)R(k)K^T(k)$$ (A.4-23)

where

$$\tilde{K}^*(k) = \tilde{K}^*(k)\phi^{-1}(k+1,k)K^*(k) = \phi^{-1}(k+1,k)\phi(k+1,k)P^*(k|k-1)H^T(k)\left[H(k)P^*(k|k-1)H^T(k)+R(k)\right]^{-1}$$ (A.4-24)
Using the notation of Eq. A.4-24, Eq. A.4-21c becomes
\[
\hat{x}(k|k) = (I - \bar{K}(k)H(k)) \hat{x}(k|k-1) + \bar{K}(k)z(k)
\]  
\[
\dot{x}(k|k-1) + \bar{K}(k) [z(k) - H(k) \hat{x}(k|k-1)]
\]

Thus the approach of Appendix A, utilizing the Matrix Minimum Principle, yields Eqs. A.3-10, A.3-11, A.4-16b, A.4-5, A.4-18g, A.4-23, A.4-24, and A.4-25b to completely specify the filter structure, both before and after a measurement update.

A.5 Perspectives on Use of Matrix Minimum Principle to Generalize for Decentralized Filtering

As also noted in Refs. 2, 3-9, the costate equation of Eq. A.4-10c completely decouples from the filtering equations for this specific "optimal centralized filtering" application, so that an exact explicit solution for \( \Lambda(k) \) is not necessary here. All that is necessary here is to guarantee (fn. A-15) that the costate \( \Lambda(k) \) is symmetric and that \( \Lambda^{-1}(k) \) exists for \( \delta < k < N \); however, such is not always the case in other filtering applications since explicit solution or approximation of the costate is required in some of the generalizations to decentralized and reduced-order filtering (e.g., Refs. 9,18). In both of these generalizations, the Matrix Minimum Principle is the primary exploratory tool for specification of the appropriate constrained filtering algorithms (i.e., iteration equations).

Effort was expended here to obtain the correct formulation, correct canonical equations, and correct intermediate arguments to serve as a rigorous guide, as the same techniques are applied (as done in Chapter 2) in obtaining discrete-time decentralized results.

Remark 1:

Note that the previous formulation of Refs. 2, 3, 9 as a simple linear control problem with specified initial value, free terminal value, and fixed terminal time (known as a Mayer-type variational problem as discussed for the filtering application on p. 293 of Refs. 17, 34) is retained if the cost function of Eq. A.4-6 is forced to degenerate to just a terminal accuracy constraint. This degeneration is accomplished by weakening the conditions of Eq. A.4-3 on
M(k) to be just positive semi-definite for \( 0 \leq k \leq N_0 - 1 \) (but still positive definite at \( k = N_0 \)), then making the specific assignment that

\[
M(k) = 0 \text{ for } 0 \leq k \leq N_0 - 1
\]  

(A.5-1)

Notice that this is equivalent to ignoring or removing the boxed terms appearing in Eqs. A.4-6 to A.4-12. However, as discussed in the paragraph following Eq. A.4-14, the strong (i.e., valid) justification for concluding that \( \Lambda(k) \) is nonsingular for \( 0 \leq k \leq N_0 - 1 \) (without invoking "that which is yet to be demonstrated") is consequently removed along with the boxed terms.

**Remark 2:**

Five alternate derivations of the discrete-time Kalman filter via such routes as:

* orthogonal projections
* recursive least squares
* maximum likelihood
* minimum variance
* conditional expectations

are provided on pp. 201-209 and pp. 168, 342, 343 of Ref. 35. Yet another approach to the derivation of the optimal filter, described as "perhaps the sleekest one to date" (preface of Ref. 36), proceeds via arguments pertaining to three martingales (Ref. 36) that arise in the filtering context.

**Remark 3:**

As mentioned on p. 696 of Ref. 2, it is well-known that the minimum principle provides only necessary conditions for optimality. Ref. 2 indicates that sufficient conditions for optimality can be obtained for the continuous-time formulation by investigating the associated Hamilton-Jacobi(-Bellman) equation (p. 15 of Ref. 37)

\[
\frac{\partial J(P,t)}{\partial t} = -H[K(P(t),\frac{\partial J(P,t)}{\partial P}),P(t),\frac{\partial J(P,t)}{\partial P},t]
\]  

(A.5-2)
with boundary condition

\[ J(P, t) = 0 \]  
(A.5-3)

for all \( P(t) \) that satisfy the terminal accuracy constraint. Ref. 2 states that sufficiency is provided by demonstrating that Eqs. A.5-2 and A.5-3 are satisfied for all \( P \), independent of how the filter gain \( K(t) \) is specified. As summarized in Eq. 59 of Ref. 1, the Hamilton-Jacobi-Bellman theory requires that the costate \( \lambda(t) \) be interpreted as

\[ \lambda(t) = \frac{\partial}{\partial P} J(P, t) \]  
(A.5-4)

Ref. 2 asserts that this demonstrative proof of sufficiency has been carried out (for the continuous-time case of only a terminal accuracy constraint) and that it is "straightforward but lengthy" and so relegated to Ref. 37. (Since explicit demonstration in Ref. 37 has not yet been found, this may still be an open question. However, resolution of this issue does not affect the JTIDS application.)
Footnotes to Appendix A

[fn. A-1] An area of further research suggested on pp. 93-94 of Ref. 9 is to derive the discrete-time formulations of the decentralized filters, as needed before clear-cut operation counts can be provided (as done in Refs. 12, 13, 14 for other filter structures) for further comparisons of computational efficiency/computer burdens.

[fn. A-2] The Matrix Minimum Principle is the only tool that has been used for generalization to different areas such as in obtaining decentralized filtering results (i.e., deriving the appropriate recursive equations describing the mechanization). The Matrix Minimum Principle has also been utilized as an approach to specify recursive equations for reduced-order filtering (Refs. 18, 19) and for optimal measurement sensor utilization in Kalman filtering (Refs. 20, 21, 22).

[fn. A-3] The cost function modification is such that it can be easily reverted to the previous cost function as a special case.

[fn. A-4] In certain special cases, these equations decouple and explicit solution of the "backward" equation becomes unnecessary.

[fn. A-5] An analytically inconvenient problem statement here can give rise to analytical problems that do not become apparent until near the end of a fairly lengthy and involved derivation, thus causing the analyst to have to backtrack and alter the convention of $\hat{x}$ and $P$ usage in the original structural assumptions until a tractable form is found.

[fn. A-6] Standard notation is $Z(k) = \{z(0), z(1), z(2), \ldots, z(k)\}$. The conditional expectation $E[\cdot | Z(k)]$ has a rigorous measure-theoretic interpretation as
the Radon-Nikodym derivative of the overall probability measure with respect to its restriction on the sub-sigma algebra generated by \( z(k) \) (i.e., the smallest sigma algebra with respect to which each of the random variables \( z(0), z(1), \ldots, z(k) \) is measurable) as discussed in Refs. 24, 25, 26. This detailed measure-theoretic or axiomatic probability perspective is unnecessary for the present discussion and so is avoided here.

[fn. A-7] Recently, certain technical questions have been raised (Ref. 27) pertaining to the information contained in the measurements versus the information extracted during filtering in obtaining white residuals. These questions have been successfully resolved in Refs. 28, 29, and 79.

[fn. A-8] Special explicit consideration of both the conditional and unconditional unbiasedness requirement and its consequences for the filtering application as demonstrated here are provided on pp. 291-292 of Ref. 17 (for continuous-time) but the distinction is absent in Refs. 2, 3-9, 18, 19.

[fn. A-9] Use has been made of the well-known identity

\[
E\{E(.|\overline{Z(k)})\} = E[.]
\]

(pathological cases cautioned against in Ref. 30 are absent in the above application).

[fn. A-10] A particular sequence \( \{M(k)\} \) is not singled out here since the resulting optimal filter gains for the "centralized" Kalman filter are independent of the particular weightings (pp. 44, 45 of Ref. 25).

[fn. A-11] The gains \( \{K(k)\} \) are optimized to minimize a function of \( P(k|k-1) \) rather than a function of \( P(k|k) \) since \( P(k|k-1) \) is a worst case envelope that always dominates the corresponding \( P(k|k) \) and minimizing the worst case accomplishes the minimization of both.

[fn. A-12] The entire cost function used in the continuous-time formulation of Ref. 2 consisted of only
\[ J_2 = L_\tau [M(\tau), P(\tau)] = \text{tr}[M(\tau)P(\tau)] \]

which was indicated on p. 694 of Ref. 2 to be only a terminal-time accuracy condition (where the terminal time is equivalent to \( N_0 \) in the above discrete-time analysis). Alternatively, Ref. 9 indicated that their problem formulation focused interest on minimizing (the more appropriate)

\[ J_3 = L_{t_1}[M(t_1), P(t_1)] = \text{tr}[M(t_1)P(t_1)] \text{ for } 0 \leq t_1 \leq \tau \]

but they state that this is equivalent to only a terminal-time penalty and Ref. 9 treats the accuracy term to be minimized in the same manner as Ref. 2. However, the cost function of Eq. A.4-10b is used in this analysis because it is perceived to be more appropriate; otherwise filter gains could conceivably be specified to minimize the terminal accuracy while allowing prior loose unacceptable accuracy in estimation during an earlier time segment.

[fn. A-13] A marked difference between the discrete-time and continuous-time formulations is the presence of a plus sign in Eq. A.4-10a instead of a minus sign as occurs in the continuous-time case (cf., p. 694 of Ref. 2, p. 596 of Ref. 1, and p. 15 of Ref. 9 [where an error in signs occurs]).

[fn. A-14] Otherwise circular reasoning would occur in using what is to be established as an intermediate step is establishing the desired result.

[fn. A-15] Recall that Refs. 2, 9 utilized simpler (but less physically motivated) cost functions \( J_2 \) and \( J_3 \) (as discussed in the footnote A-12) yet obtained the correct final answers via theoretical arguments that were less than fully justified for the discrete-time formulation (as discussed in the paragraph following Eq. A.4-14). In Refs. 2, 9, the approach was motivated by the conclusion that the well-known correct results were obtained. Here the correct results are obtained by a correct approach which therefore should be more amenable to the further generalizations to be encountered.
REFERENCES


75. McGarty, T.P., Stochastic Systems and State Estimation,


IR-480
STABILITY CONDITIONS FOR
THE RELNAV COMMUNITY AS
A DECENTRALIZED ESTIMATOR
FINAL REPORT
30 SEPTEMBER 1980

Prepared for:
Naval Air Development Center
Warminster, PA 18974
[Contract No. N62269-79-D-0301
Delivery Order No. 0004
(CDRL A010)]

Prepared by:
Intermetrics, Inc.
733 Concord Ave.
Cambridge, MA 02138

Dr. Thomas H. Kerr
Lead Author
Jean Onerheim
Product Control
Alfred S. Gilman
Technical Director
Bruce McCoy
Program Manager