AD-A037 065 NAVAL RESEARCH LAB WASHINGTON D C F/G 12/1

DUAL PERTURBATION CONTROL.(U)

FEB 77 W W WILLMAN

UNCLASSIFIED NRL-8071 NL

1 OF 1

AD
A037065

END

DATE
FILMED
4-77

NRL Report 8071

Dual Perturbation Control

WARREN W. WILLMAN

Systems Research Branch
Space Sciences Division

February 15, 1977

NAVAL RESEARCH LABORATORY
Washington, D.C.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
NRL [redacted] 8071

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
DUAL PERTURBATION CONTROL

5. TYPE OF REPORT & PERIOD COVERED
Interim report on one phase of a continuing NRL Problem.

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Warren W. Willman

8. CONTRACT OR GRANT NUMBER(s)
B01-10

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Naval Research Laboratory
Washington, D.C. 20375

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
RR003-02-41-6152

11. CONTROLLING OFFICE NAME AND ADDRESS
Department of the Navy
Office of Naval Research
Arlington, Va. 22217

12. REPORT DATE
Feb[redacted] 1977

13. NUMBER OF PAGES
48

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
(12) 49p.

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited.
(17) RR00302141

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Dual control	Perturbation feedback control
Gradient algorithm	State-dependent noise
Optimal control	Stochastic optimal control

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Optimal control laws are investigated for minimizing the expected value of a quadratic criterion with linear dynamics and state measurements, both perturbed by additive white Gaussian noise processes whose intensities depend weakly and linearly on the instantaneous state and control. A dynamic programing approach is used to derive expressions for the optimal control and cost function that are accurate to first order in the noise intensity variations and are given in terms of initial value systems of ordinary differential equations. As compared to the classical "linear-quadratic-Gaussian"

(Continued)

→ next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

251 950

LB

20. (Continued)

cont → control problem without noise-intensity variations, the first-order optimal control law modifications here are found to be the inclusion of the state covariance matrix as a measurement-driven variable in the state estimator, the appearance of deterministic skewness variables in this estimator, and the addition of a deterministic term to the control. Part of the additive control term can be interpreted as the "dual control" effect, and it is coupled to the control through a matrix time function whose properties are investigated. A useful refinement of the certainty-equivalence principle is made. When the measurement noise is state-dependent, the differential equations for state estimation have a random driving term containing the scatter matrix of the measurements, which imposes some additional restrictions on the validity of the analysis. Some aspects of the results are shown to generalize to the case of a quadratic exponential criterion, although that situation is more complicated. A method for including the effects of noise intensity gradients in iterative optimization algorithms is described, and a numerical example is given.

CONTENTS

DUAL PERTURBATION CONTROL

INTRODUCTION

One problem of considerable interest in stochastic optimal control theory is that of minimizing the expected value of a quadratic criterion in the presence of linear dynamics and state measurements, both of which are perturbed by additive Gaussian white noise processes whose parameters are known a priori. This classical "linear-quadratic-Gaussian" case is important because it is both analytically tractable and descriptive of noise-induced perturbations from nominal behavior in a more general class of optimal control problems [1]. However, it has the simplifying but degenerate property that the optimal control law is the functional composition of the solutions to a deterministic optimal control problem and a state estimation problem that are essentially independent of each other (the certainty-equivalence property). The only effect the choice of control has on the state estimation results is to shift, by a known amount, the mean of the conditional state distribution. Heuristically, this means that in this case the acquisition and exploitation of state information are independent.

Some analogous results have been obtained recently [2] for a variant of this problem in which the criterion is changed to a quadratic exponential, as it might be in minimizing a terminal miss distance and the probability of a control-dependent Poisson failure. The certainty-equivalence property does not extend to this case, but there is still no conflict between the acquisition and exploitation of state information because the estimation results are control-independent here also. There is such a conflict in the general stochastic optimal control problem, however, where the "quality" of the state estimate can be influenced by the choice of control. The optimal control law in such cases can therefore be interpreted as having a dual character [3]; it represents an optimal compromise between acquiring and exploiting state information for the ultimate purpose of minimizing the criterion.

This dual character is investigated here by considering an extension of the linear-quadratic-Gaussian problem, in which the noise "covariance" matrixes vary as functions of the instantaneous state and control. An exact solution is not attempted, but a dynamic programing approach provides an explicit expression for the optimal control law—in terms of initial value systems of ordinary differential equations—which is accurate to first order in the covariance matrix variations under the restriction that they remain small (and linear as functions of the state and control). Such results at least show how the optimal control law *starts* to be affected in this particular context when the choice of control *begins* to influence the quality of the state estimate, and this provides a starting point for speculation about these effects in a more general context. Hence, this might be called a "linear-quadratic-Gaussian infinitesimal" control problem. For such problems arising from a perturbation analysis of a more general situation, however, the restriction of smallness here presents little additional loss of generality, and the level of accuracy is

Manuscript submitted September 20, 1976.

compatible with that of the original analysis. Furthermore, some of the phenomena appear to generalize at least to the case of the quadratic exponential criterion, although the lack of a certainty-equivalence property complicates their interpretation in that case.

NOTATION

Unless otherwise noted, lower case letters are used here to denote (real) column vectors or scalars, and capital Roman letters denote matrixes. For a matrix A, A^T denotes the transpose of A. If A is square, $|A|$ denotes its determinant, $tr(A)$ denotes its trace

$$\left(\sum_i A_{ii}\right),$$

and $adj(A)$ denotes its adjoint, a square matrix whose i, jth component is the j, ith cofactor of A. If a is a vector, each component of which is a function of another vector x, then a_x denotes the matrix of partial derivatives such that $(a_x)_{ij} = \partial a_i / \partial x_j$. If a is a scalar, then a_A denotes the matrix of partial derivatives such that $(a_A)_{ij} = \partial a / \partial A_{ji}$.

It will also be necessary here to manipulate three-way matrixes of real numbers, which will always be denoted by capital Greek letters. For continuity of notation, we adopt the following definitions for such a matrix $\Gamma = \{\Gamma_{ijk} : i = 1, ..., I; j = 1, ..., J; k = 1, ..., K\}$:

Postmultiplication by a column K-vector x gives an $I \times J$ matrix such that

$$(\Gamma x)_{ij} = \sum_{k=1}^{K} \Gamma_{ijk} x_k .$$

Premultiplication by an $N \times I$ matrix A gives an $N \times J \times K$ three-way matrix such that

$$(A\Gamma)_{njk} = \sum_{i=1}^{I} A_{ni} \Gamma_{ijk} .$$

Postmultiplication by a $K \times N$ matrix B gives an $I \times J \times N$ three-way matrix such that

$$(\Gamma B)_{ijn} = \sum_{k=1}^{K} \Gamma_{ijk} B_{kn} .$$

The transpose of Γ is a $K \times I \times J$ three-way matrix Γ' such that $(\Gamma')_{kij} = \Gamma_{ijk}$. If $I = K$, $Tr(\Gamma)$ is a column J-vector such that

$$[Tr(\Gamma)]_j = \sum_{i=1}^{I} \Gamma_{iji} .$$

Γ is called symmetric if $\Gamma = \Gamma' = \Gamma''$ and $\Gamma_{ijk} = \Gamma_{jik}$ ($\Rightarrow I = J = K$).

Ax^T denotes a three-way matrix such that $(Ax^T)_{ijk} = A_{ij}x_k$.

With these definitions, the expressions $A\Gamma Bx$, BAx^T, Ax^TB and Ax^Ty are unambiguous. Many other consequences are obvious. Some useful properties that are not so obvious are listed below.

$$tr(\Gamma'x) = [Tr(\Gamma)]^T x \text{ and } \frac{\partial}{\partial x}[tr(\Gamma'x)] = [Tr(\Gamma)]^T$$

$$ATr(\Gamma') = Tr[(A\Gamma)'] \text{ and } B^T Tr(\Gamma'') = Tr[(\Gamma B)'']$$

$$Tr(A\Gamma) = Tr(\Gamma A)$$

$$(\Gamma B)' = B^T\Gamma' \text{ and } \Gamma''B = (B^T\Gamma)''$$

$$\Gamma'xB = (B^T\Gamma)'x$$

$$\Gamma xx = Tr(\Gamma'xx^T)$$

$$\Gamma \text{ symmetric} \Rightarrow (A^T\Gamma A)'A \text{ and } B(B\Gamma B^T)'' \text{ symmetric}$$

Parentheses are omitted in this notation if the order of multiplicative association is immaterial or if the interpretation is unambiguous; for example, ΓxB must mean $(\Gamma x)B$ because xB is not defined. If x is a vector and a is a scalar, then a_{Ax} denotes the three-way matrix of second partial derivatives such that

$$(a_{Ax})_{ijk} = \partial^2 a/\partial A_{ji}\partial x_k.$$

The probability density function of a random variable x is denoted by $p_x(\cdot)$ and the corresponding expectation operator by E_x. Where the meaning is clear from the context, $p(x)$, $E(x)$, and $E(x/y)$ are often used as abbreviations for $p_x(x)$, $E_x(x)$ and $E_{x/y}(x, y)$. The covariance of x is denoted by cov(x).

PROBLEM FORMULATION AND MOTIVATION

The problem of primary interest here is the following extension of the familiar linear-quadratic-Gaussian optimal control problem, in which the covariance matrixes of the process and measurement noises are allowed to have a certain kind of dependence on the state and control vectors:

$$\left.\begin{aligned} &\dot{x} = Fx + Gu + w, \\ &x(t_0) \text{ is Normal } (\hat{x}_0, P_0) \text{ a priori} \end{aligned}\right\} \text{ (dynamics)} \quad (1)$$

$$z = Hx + v \quad \text{(state measurements)} \quad (2)$$

$$J = \frac{1}{2} E\left[x^T(t_f)S_f x(t_f) + \int_{t_0}^{t_f} (x^T Ax + u^T Bu + 2c^T u)\, dt\right] \quad \text{(scalar criterion)} \qquad (3)$$

where time argument t is suppressed in the notation and

E denotes prior expected value

x is an n-dimensional state vector

u is an m-dimensional control vector

z is a k-dimensional state measurement

w and v are independent zero-mean Gaussian white noise processes with respective covariance matrix parameters

$$\left.\begin{array}{l} Q + 2\Gamma' u + 2\Psi' x, \\ R + 2\Omega' u \end{array}\right\} \text{given } u \text{ and } x$$

A, Q, and S_f are symmetric positive-semidefinite matrixes

B and R are symmetric positive-definite matrixes

All components of Ψ, Γ, *and* Ω *(which may be time-varying) are approximately infinitesimal*—let us say of order h; all other quantities are of order unity, including B^{-1} and R^{-1}

$\Gamma'_{ijk} = \Gamma'_{jik}$, $\Psi'_{ijk} = \Psi'_{jik}$, and $\Omega'_{ijk} = \Omega'_{jik}$ (to retain covariance matrix symmetry).

An alternate criterion of exponential form is also considered for comparison, but discussion of this is deferred until later. Including state dependence in the measurement noise covariance matrix presents special difficulties and is considered separately.

The objective here is to determine, at least to first order in h, the control law that minimizes criterion J. As usual, a control law is defined as a decision rule that determines the control $u(t)$ as a function of the available measurement history $Z(t) = \{[s, z(s)] : s\epsilon[t_0, t]\}$. Since white noise processes do not really exist except as a kind of shorthand notation for sequences of approximating step-function processes, the control law sought here should really be interpreted as a limiting form of the solutions to a sequence of restricted optimal control problems in which the control and noise values change only at a finite number of specified intermediate times, where the maximum time interval between such changes goes to zero in the control problem sequence.

The development here is formal, however, in the sense that no investigation is made of the conditions under which such a limit concept is meaningful. The reason for treating the problem in continuous time here is the more concise form of the results, together with the fact that they can serve as a single approximation to the results for any approximating discrete-time problem with a short enough discretization interval. As usual, the process noise covariance matrix for such a discretization interval of generic length Δ and index i is normalized as

$$\frac{1}{\Delta}\left[Q(t_0 + i\Delta) + 2\Gamma'(t_0 + i\Delta)u_i + 2\Psi'(t_0 + i\Delta)x_i\right]$$

so that the statistics of the noise increments on that interval are asymptotically the same; i.e.,

$$\mathrm{cov}[w_i\Delta] \to \mathrm{cov}\left\{\int_{t_0+i\Delta}^{t_0+(i+1)\Delta} w(t)\,dt\right\} \quad \text{as } \Delta \to 0$$

and similarly for the measurement noise covariance matrix.

Aside from their conceptual interest, problems of this class can arise in the following way. Assume that a solution to a nominal deterministic optimal control problem is available and that perturbations about this nominal path are observable and controllable. Suppose also that there are process and measurement noises, ignored in the nominal solution, whose covariance matrixes possibly depend on the state and control. A common approach to minimizing the actual expected value of the criterion in such a case is to seek a feedback solution to an "accessory minimum problem" for the perturbations, under the assumption that they remain approximately infinitesimal. In this context, the accessory minimum problem would be constructed by linearizing the dynamics *and* the noise covariance dependences about the nominal path. If the resulting problem is rescaled so that the state and control perturbations are of order unity, it is often reducible to a stochastic optimal control problem of the above form, where the covariance dependence coefficients become the small quantities. Of course, the case of state-dependent measurement noise cannot be accommodated under the present restriction. Moreover, since linear terms in the controls are included in the criterion of this formulation, a problem of this class could also represent an iteration in a corresponding second-order gradient algorithm, in which the linearizations and second-order expansion of the criterion are constructed about a trajectory that is not optimal in the deterministic problem. The importance of this lies in the possibility of iteratively modifying the nominal path to account optimally for noise-intensity gradients.

STATE ESTIMATION

For state-independent process noise ($\Psi = 0$), both noise covariance matrixes can be regarded as known, since the current control values are assumed known. In this case, therefore, it follows from well-known results for the Kalman-Bucy filter [4] that the conditional probability density of the state given the available measurements is Gaussian, and that its mean $\hat{x}$ and covariance matrix $\overline{P}$ obey the equations

$$\dot{\hat{x}} = F\hat{x} + Gu + \overline{P}H^T\overline{R}^{-1}(z - H\hat{x}); \qquad \hat{x}(t_0) = \hat{x}_0 \tag{4}$$

and

$$\dot{\overline{P}} + F\overline{P} + \overline{P}F^T + \overline{Q} - \overline{P}H^T\overline{R}^{-1}H\overline{P}; \qquad \overline{P}(t_0) = \overline{P}_0 \tag{5}$$

where

$$\overline{Q} = Q + 2\Gamma' u \tag{6}$$

$$\overline{R} = R + 2\Omega' u. \tag{7}$$

For nonzero Ψ, it is shown below that this conditional density no longer remains Gaussian to first order, but rather is of the form

$$\left(1 + \lambda^T(x - \overline{x}) + tr\left\{\frac{1}{2} L[(x - \overline{x})(x - \overline{x})^T - V] + \frac{1}{3}(x - \overline{x})(x - \overline{x})^T \Lambda (x - \overline{x})\right\}\right)\left[\frac{\exp - \frac{1}{2}(x - \overline{x})^T V^{-1}(x - \overline{x})}{(2\pi)^{1/2n} |V|^{1/2}}\right] \tag{8}$$

where V, L, and Λ are symmetric, V is positive definite, and the components of λ, L, and Λ are all of order h. In general, Eq. (8) can assume negative values for large enough magnitudes of $(x - \overline{x})$ and must be modified slightly to be a proper probability density. Because of the rapid decay of the exponential factor, however, these modifications can be confined to a region whose probability mass is negligible to arbitrary order in h for sufficiently small h, so Eq. (8) will be treated as a proper density in the following. Since it has the form of a Gaussian density function multiplied by a polynomial, it follows directly from standard results for Gaussian moments that the integral of Eq. (8) over R^n is unity and that

$$E(x) = \overline{x} + V\lambda + VTr(\Lambda V) \equiv \mu \tag{9}$$

$$\text{cov}(x) = E[(x - \mu)(x - \mu)^T] = V + VLV - (\mu - \overline{x})(\mu - \overline{x})^T \equiv U. \tag{10}$$

Assume now that the conditional density of the state x at time t is of the form of Eq. (8). After a short time increment Δ has elapsed, the conditional density of the state at time $(t + \Delta)$ can be determined to first order in Δ by first finding the density of y, where

$$y = (I + F\Delta)x + Gu\Delta + \omega,$$

and ω is a random variable whose distribution given x is zero-mean Gaussian with covariance matrix $(\overline{Q} + 2\Psi' x)\Delta$, and then finding the conditional density of y given z, where

$$z = Hy + \xi$$

and ξ is an independent zero-mean Gaussian random variable with covariance R/Δ. Since w itself is regarded as a step-function approximation to white noise in Eq. (1), no correction term of the sort described by Wong and Zakai [5] is needed here to compensate for the state dependence of the process noise.

If x has the density function of Eq. (8) and $s = Kx + b$, where K and b are constants and K^{-1} exists, it is straightforward but tedious to show from standard results for the transformation of probability densities that s also has a density of the same form, namely

$$p(s) = \left(1 + \lambda^T K^{-1}(s - \bar{s}) + tr\left\{\frac{1}{2} K^{-1^T} L K^{-1}\left[(s - \bar{s})(s - \bar{s})^T - KVK^T\right]\right.\right.$$
$$\left.\left. + \frac{1}{3}(s - \bar{s})(s - \bar{s})^T (K^{-1^T} \Lambda K^{-1})' K^{-1}(s - \bar{s})\right\}\right)\left[\frac{e^{-1/2(s-\bar{s})^T (KVK^T)^{-1}(s-\bar{s})}}{(2\pi)^{1/2n} |KVK^T|^{1/2}}\right],$$
$$\bar{s} = K\bar{x} + b. \tag{11}$$

If s is used now to denote $(I + F\Delta)x$ in particular, it follows that s has a density of the form of Eq. (8), since $(I + F\Delta)^{-1}$ always exists for sufficiently small Δ. Furthermore, its "λ," "L," and "Λ" parameters differ from those of the density of x only by order $h\Delta$. From general results for means and covariances,

$$E(s) = (I + F\Delta)\mu \equiv \mu_2$$
$$\mathrm{cov}(s) = (I + F\Delta)U(I + F^T\Delta) \equiv U_2.$$

Since knowing s is equivalent to knowing x, the probability density of $r = s + \omega$ is given by the equation

$$p(r) = p(s + \omega) = \int_{R^n} p_s(r - \omega) p_{\omega/x}[\omega, (I + F\Delta)^{-1}(r - \omega)]\, d\omega. \tag{12}$$

It is assumed initially that $\bar{Q}^{-1}$ exists, in which case the density $p_{\omega/x}[\omega, (I + F\Delta)^{-1}x]$ can be approximated to order h as

$$p_{\omega/x}[\omega, (I + F\Delta)^{-1}x] = \left\{1 - tr\left[\bar{Q}^{-1}\bar{\Psi}' x\left(I - \frac{1}{\Delta}\bar{Q}^{-1}\omega\omega^T\right)\right]\right\}\left[\frac{\exp -\frac{1}{2\Delta}\omega^T \bar{Q}^{-1}\omega}{(2\pi)^{1/2n}|\bar{Q}\Delta|^{1/2}}\right],$$
$$\bar{\Psi} = \Psi(I + F\Delta)^{-1}. \tag{13}$$

Using Eq. (8) to represent $p_s(x)$ and substituting it and Eq. (13) in Eq. (12) gives

$$p(r) = \int_{R^n} \frac{k(\omega)\exp\left\{-\frac{1}{2}\left[(r - \omega - \bar{x})^T V^{-1}(r - \omega - \bar{x}) + \frac{1}{\Delta}\omega^T \bar{Q}^{-1}\omega\right]\right\} d\omega}{(2\pi)^n |V\bar{Q}\Delta|^{1/2}}$$

to order $h\Delta$, where

$$k(\omega) = 1 + \lambda^T(r - \omega - \bar{x}) + tr\left\{\frac{1}{2} L[(r - \omega - \bar{x})(r - \omega - \bar{x})^T - V]\right.$$
$$+ \frac{1}{3} (r - \omega - \bar{x})(r - \omega - \bar{x})^T \Lambda (r - \omega - \bar{x})$$
$$\left. - \bar{Q}^{-1}\Psi'(r - \omega)\left(I - \frac{1}{\Delta} \bar{Q}^{-1}\omega\omega^T\right)\right\}.$$

Completing the square in the exponent gives

$$p(r) = \frac{\exp\left[-\frac{1}{2} (r - \bar{x})^T (V + \bar{Q}\Delta)^{-1} (r - \bar{x})\right]}{(2\pi)^{1/2n} |V + \bar{Q}\Delta|^{1/2}} \int_{R^n} k(\omega)$$

$$\times \frac{\exp\left\{-\frac{1}{2} [\omega - \bar{Q}V^{-1}(r - \bar{x})\Delta]^T [\bar{Q}\Delta - \bar{Q}V^{-1}\bar{Q}\Delta^2]^{-1} [\omega - \bar{Q}V^{-1}(r - \bar{x})\Delta]\right\}}{(2\pi)^{1/2n} |\bar{Q}\Delta - \bar{Q}V^{-1}\bar{Q}\Delta^2|^{1/2}} \, d\omega.$$

The integrand in this expression is a Gaussian density multiplied by a polynomial, so it is straightforward to verify that, to first order in Δ,

$$p(r) = \left(1 + \bar{\lambda}^T(r - \bar{x}) + tr\left\{\frac{1}{2} \bar{L}[(r - \bar{x})(r - \bar{x})^T - (V + \bar{Q}\Delta)] + \frac{1}{3} (r - \bar{x})(r - \bar{x})^T \bar{\Lambda} (r - \bar{x})\right\}\right)$$
$$\times \left[\frac{\exp - \frac{1}{2} (r - \bar{x})^T (V + \bar{Q}\Delta)^{-1} (r - \bar{x})}{(2\pi)^{1/2n} |V + \bar{Q}\Delta|^{1/2}}\right] \tag{14}$$

where

$$\bar{\lambda} = \lambda + \Delta[Tr(\bar{Q}\Lambda) - V^{-1}(\bar{Q}\lambda + 2Tr\Psi')] \tag{15}$$

$$\bar{L} = L + \Delta\{[(V^{-1}\bar{\Psi}V^{-1})'\bar{x} - L\bar{Q}V^{-1}] + [(V^{-1}\bar{\Psi}V^{-1})'\bar{x} - L\bar{Q}V^{-1}]^T\} \tag{16}$$

$$\bar{\Lambda} = \Lambda + \Delta[(V^{-1}\bar{\Psi}V^{-1} - \Lambda\bar{Q}V^{-1}) + (V^{-1}\bar{\Psi}V^{-1} - \Lambda\bar{Q}V^{-1})' + (V^{-1}\bar{\Psi}V^{-1} - \Lambda\bar{Q}V^{-1})'']. \tag{17}$$

This is again a density of the form of Eq. (8). From the definition of s and Eqs. (11) and (15) through (17), the components of λ, L, and Λ for x and r differ only by order $h\Delta$. Applying Eqs. (9) and (10) to Eqs. (14) through (17) shows that, since $y = r + Gu\Delta$,

$$E(y) = \mu_2 + Gu\Delta \equiv \mu_3 \tag{18}$$

$$\text{cov}(y) = U_2 + (\bar{Q} + 2\Psi'\mu)\Delta \equiv M \tag{19}$$

which can also be verified directly by decomposing the expectations. The same result holds for singular $\overline{Q}$ by continuity, since it does not involve $\overline{Q}^{-1}$. Equation (11) implies that y has the same density parameters as r, except that $\overline{x}$ increases by $Gu\Delta$.

As a function of y, conditional density $p(y/z)$ is proportional to $p(z/y)p(y)$. Completing the square in the exponent of this product shows that

$$p(y/z) = g\left(1 + \overline{\lambda}^T(y - \overline{x}) + tr\left\{\frac{1}{2}\overline{L}[(y - \overline{x})(y - \overline{x})^T - (V + \overline{Q}\Delta)]\right.\right.$$
$$\left.\left. + \frac{1}{3}(y - \overline{x})(y - \overline{x})^T\Lambda(y - \overline{x})\right\}\right)\exp\left[-\frac{1}{2}(y - \overline{y})^T\overline{V}^{-1}(y - \overline{y})\right]$$

where g is a constant of proportionality, $\overline{x}$ now denotes the parameter in Eq. (8) for $p(y)$, and

$$\overline{V} = V + \overline{Q}\Delta - VH^T\overline{R}^{-1}HV\Delta$$

$$\overline{y} = \overline{x} + VH^T\overline{R}^{-1}(z - H\overline{x})\Delta.$$

The polynomial factor in $p(y/z)$ can be expressed as

$$a + b^T(y - \overline{y}) + tr\left\{\frac{1}{2}L^*[(y - \overline{y})(y - \overline{y})^T - \overline{V}] + \frac{1}{3}(y - \overline{y})(y - \overline{y})^T\Lambda^*(y - \overline{y})\right\}$$

where

$$a = 1 + \Delta\left[\overline{\lambda}^T VH^T\overline{R}^{-1}(z - H\overline{x}) + tr\left(V\overline{\Lambda}VH^T\overline{R}^{-1}(z - H\overline{x})\right.\right.$$
$$\left.\left. + \frac{1}{2}\overline{L}\{VH^T\overline{R}^{-1}[(z - H\overline{x})(z - H\overline{x})^T\Delta - \overline{R}]\overline{R}^{-1}HV\}\right)\right]$$

$$b = \overline{\lambda} + \Delta[\overline{L} + \overline{\Lambda}VH^T\overline{R}^{-1}(z - H\overline{x})\Delta]\,VH^T\overline{R}^{-1}(z - H\overline{x})$$

$$L^* = \overline{L} + 2\overline{\Lambda}VH^T\overline{R}^{-1}(z - H\overline{x})\Delta$$

$$\Lambda^* = \overline{\Lambda}$$

to first order in Δ. The quantity $(z - H\overline{x})\Delta$ is regarded as a term of order $\Delta^{1/2}$ here. Since $p(y/z)$ is a probability density and must integrate to unity, a can be absorbed into the proportionality constant g to express this density in the form of Eq. (8) such that the λ, L, and Λ components differ from those of $\overline{\lambda}$, $\overline{L}$, and $\overline{\Lambda}$ only by terms of order $h\Delta$ and zero-mean random terms of order $h\Delta^{1/2}$. Carrying out the details to order Δ (only to order $\Delta^{1/2}$ for zero-mean random terms) and using Eqs. (9) and (10) show that

$$E(y/z) = \mu_3 + MH^T\overline{R}^{-1}(z - H\mu_3)\Delta + MTr\{MH^T\overline{R}^{-1}[(z - H\mu_3)(z - H\mu_3)^T\Delta - \overline{R}]\overline{R}^{-1}HM\overline{\Lambda}\}\Delta + \text{terms of order } h^2\Delta \qquad (20)$$

$$\text{cov}(y/z) = M - MH^T\overline{R}^{-1}HM + 2(M\overline{\Lambda}M)'MH^T\overline{R}^{-1}(z - H\mu_3)\Delta + \text{terms of order } h^2\Delta + \text{zero-mean random terms of order } h^2\Delta^{1/2} \qquad (21)$$

$$\Lambda = \overline{\Lambda} + \text{terms of order } h^2\Delta + \text{zero-mean random terms of order } h^2\Delta^{1/2}. \qquad (22)$$

Efforts to obtain similar results in this way with state-dependent measurement noise have been unsuccessful at this point, possibly because more information about the state is given in this case by the scatter of a series of measurements over a short period of time than by their average value. Since the zero-mean random terms are statistically independent for disjoint time increments and since the third term in Eq. (20) is a zero-mean random term of order Δ, the last two types of terms in each of Eqs. (20) through (22) can be neglected because they only contribute effects of order h^2 or smaller when "integrated" over a time interval of order unity.

If we return to the notation of Eq. (4) and (5), the overall result is that

$$\hat{x}(t + \Delta) = \hat{x}(t) + \Delta\{F(t)\hat{x}(t) + G(t)u(t) + \overline{P}(t)H^T(t)\overline{R}^{-1}(t)[z(t) - H(t)\hat{x}(t)]\},$$

$$\overline{P}(t + \Delta) = \overline{P}(t) + \Delta\{F(t)\overline{P}(t) + \overline{P}(t)F^T(t) + \overline{Q}(t) + 2\Psi'(t)\hat{x}(t) - \overline{P}(t)H^T(t)\overline{R}^{-1}(t)H(t)\overline{P}(t) + 2[\overline{P}(t)\Lambda(t)\overline{P}(t)]'\overline{P}(t)H^T(t)\overline{R}^{-1}(t)[z(t) - H(t)\hat{x}(t)]\},$$

and

$$\Lambda(t + \Delta) = \Lambda(t) + \Delta(\overline{P}^{-1}(t)\Psi(t)\overline{P}^{-1}(t) - \Lambda(t)(F(t) + \overline{Q}(t)\overline{P}^{-1}(t)) + \{\overline{P}^{-1}(t)\Psi(t)\overline{P}^{-1}(t) - \Lambda(t)[F(t) + \overline{Q}(t)\overline{P}^{-1}(t)]\}' + [\overline{P}^{-1}(t)\Psi(t)\overline{P}^{-1}(t) - \Lambda(t)[F(t) + \overline{Q}(t)\overline{P}^{-1}(t)]]'')$$

to first order in Δ, except for terms contributing effects of order Δ but of second order in h. Furthermore, the λ, L, and Λ components describing the conditional density of x in the notation of Eq. (8) change only by amounts of order $h\Delta$ in this interval, so they remain of order h; also, L and Λ remain symmetric. It is convenient at this point to express the conditional covariance matrix as the sum

$$\overline{P} = P + 2D$$

where P is a "nominal covariance matrix" defined as a deterministic time function by the classical Kalman-Bucy filter equation:

$$\dot{P} = FP + PF^T + Q - PH^T R^{-1} HP; \qquad P(t_0) = P_0. \tag{23}$$

Since $\overline{R}^{-1} = (R + 2\Omega' u)^{-1} = R^{-1}(R - 2\Omega' u)R^{-1}$ to first order in h and since the components of D are of order h, the mean $\hat{x}$ and covariance matrix $P + 2D$ of the conditional state distribution are determined to first order in h in the limit as $\Delta \to 0$, by Eq. (23) and the equations

$$\dot{\hat{x}} = F\hat{x} + Gu + [PH^T(I - 2R^{-1}\Omega' u) + 2DH^T]R^{-1}(z - H\hat{x}); \qquad \hat{x}(t_0) = \hat{x}_0 \tag{24}$$

$$\dot{D} = (F - PH^T R^{-1} H)D + D(F^T - H^T R^{-1} HP) + \Psi' \hat{x} + (\Gamma + PH^T R^{-1} \Omega R^{-1} HP)' u$$
$$+ (P\Lambda P)' PH^T R^{-1}(z - H\hat{x}); \qquad D(t_0) = 0 \tag{25}$$

$$\dot{\Lambda} = \Theta + \Theta' + \Theta'', \qquad \Theta \triangleq P^{-1}\Psi P^{-1} - \Lambda(F + QP^{-1}); \qquad \Lambda(t_0) = 0 \tag{26}$$

where the "t" argument is suppressed in the notation. Three-way matrix Λ is a deterministic time function related to the skewness of the conditional state distribution, and is identically zero in the case of state-independent process noise, when $\Psi(t)$ is identically zero. To first order in h, therefore, the so-called *information state* for this estimation process consists of both $\hat{x}$ and D, not just $\hat{x}$, and hence differs significantly from the *system state* x.

OPTIMIZATION

It follows from the arguments of Stratonovich [6] and Striebel [7] that an "optimal cost function" can be defined consistently here in terms of t and the current conditional distribution of the state given the preceding measurements. It is assumed that conditions are such that the solution to the corresponding Bellman equation and boundary condition is unique and that it is sufficiently regular that second-order changes in the equation produce only second-order changes in the solution. It is convenient to proceed by considering the possibility of such cost functions depending only on $\hat{x}$, D, and t to first order in h, in which case there exists a (scalar) function $J(\hat{x}, D, t)$ such that

$$J(\hat{x}, D, t) = \text{conditional expected "cost-to-go,"}$$

that is,

$$E\left\{\frac{1}{2}\left[x^T(t_f)S_f x(t_f) + \int_t^{t_f}(x^T A x + u^T B u + 2c^T u)\,dt\right]\right\},$$

using an optimal control law, given that $\hat{x}(t) = \hat{x}$ and $\overline{P}(t) = P(t) + 2D$, plus terms of second order in h or smaller. (The question of the possible nonexistence of an optimal control law is not examined here.) The usual invariant imbedding formalism of dynamic programming (see Dreyfus [8], for example), using $\hat{x}$ and D as state variables and neglecting second-order terms, shows that the Bellman equation reduces to the following equation for J in this case:

$$0 = \min_u E\left[\frac{1}{2}(\hat{x}^T A\hat{x} + u^T Bu) + c^T u + J_{\hat{x}}\dot{\hat{x}} + J_t + tr\left(J_D\dot{D} + \frac{dDJ_{D\hat{x}}d\hat{x}}{dt} + \frac{1}{2}J_{\hat{x}\hat{x}}\frac{d\hat{x}d\hat{x}^T}{dt}\right)\right] \tag{27}$$

where the expectation is conditioned on the event $\hat{x}(t) = \hat{x}$ and $\overline{P}(t) = P(t) + 2D$, or, equivalently, on $Z(t)$. Evaluating the conditional expected cost-to-go at the terminal time shows that J must also satisfy the boundary condition

$$J(\hat{x}, D, t_f) = \frac{1}{2}\hat{x}^T S_f\hat{x} + tr\ \frac{1}{2}\left[S_f P(t_f) + S_f D\right]. \tag{28}$$

If a function $J(\hat{x}, D, t)$ that satisfies (27) and (28) to first order in h can be found, then it is a first-order approximation to the optimal cost function and determines the optimal control law to first order by the regularity assumption.

Taking expected values in Eqs. (27) and using Eqs. (24) and (25) gives, to first order in h,

$$\begin{aligned}0 = \min_u \Bigg(&\frac{1}{2}\{\hat{x}^T A\hat{x} + tr[A(P+2D)] + u^T Bu\} + c^T u + J_{\hat{x}}(F\hat{x} + Gu) + J_t + tr\Big[J_D(FD + DF^T \\ &+ \Gamma' u + \Psi'\hat{x} - DH^T R^{-1}HP - PH^T R^{-1}HD + PH^T R^{-1}\Omega' uR^{-1}HP) + J_{\hat{x}\hat{x}}\Big(\frac{1}{2}PH^T R^{-1}HP \\ &+ DH^T R^{-1}HP + PH^T R^{-1}HD - PH^T R^{-1}\Omega' uR^{-1}HP\Big)\Big] \\ &+ \sum_{ijk}(P\Lambda PH^T R^{-1}HP)_{ijk}(PJ_{D\hat{x}})_{jik}\Bigg).\end{aligned} \tag{29}$$

Collecting terms and cyclically permuting matrix products in the trace operand gives

$$\begin{aligned}0 = \min_u \Bigg(&\frac{1}{2}[\hat{x}^T A\hat{x} + u^T Bu] + c^T u + J_{\hat{x}}[F\hat{x} + Gu] + J_t + tr\Big\{\frac{1}{2}AP + \frac{1}{2}J_{\hat{x}\hat{x}}PH^T R^{-1}HP \\ &+ [A + J_D F + F^T J_D - (J_D - J_{\hat{x}\hat{x}})PH^T R^{-1}H - H^T R^{-1}HP(J_D - J_{\hat{x}\hat{x}})]D \\ &+ [R^{-1}HP(J_D - J_{\hat{x}\hat{x}})PH^T R^{-1}\Omega' + J_D\Gamma']u\Big\} + [Tr(J_D\Psi)]^T\hat{x} \\ &+ \sum_{ijk}(P\Lambda PH^T R^{-1}HP)_{ijk}(PJ_{Dx})_{jik}\Bigg).\end{aligned} \tag{30}$$

Equating the u-derivative of the left-hand side of Eq. (30) to zero specifies the minimizing control as

$$u = -B^{-1}\{G^T J_{\hat{x}}^T + c + Tr[R^{-1}HP(J_D - J_{\hat{x}\hat{x}})PH^T R^{-1}\Omega + J_D \Gamma]\}. \tag{31}$$

Substituting Eq. (31) into Eq. (30) to eliminate the minimization operation gives

$$\frac{1}{2}\hat{x}^T A\hat{x} + J_{\hat{x}}F\hat{x} + [Tr(J_D\Psi)]^T\hat{x} + J_t + tr\left\{\frac{1}{2}AP + \frac{1}{2}J_{\hat{x}\hat{x}}PH^TR^{-1}HP\right.$$
$$\left. + [A + J_DF + F^TJ_D - (J_D - J_{\hat{x}\hat{x}})PH^TR^{-1}H - H^TR^{-1}HP(J_D - J_{\hat{x}\hat{x}})]D\right\}$$
$$- \frac{1}{2}\left[G^TJ_{\hat{x}}^T + c + Tr\left(J_D\Gamma + R^{-1}HP\left(J_D\right.\right.\right.$$
$$\left.\left.\left. - \frac{1}{2}J_{\hat{x}\hat{x}}\right)PH^TR^{-1}\Omega\right)\right]^T B^{-1}\left[G^TJ_{\hat{x}}^T + c + Tr\left(J_D\Gamma + R^{-1}HP\left(J_D\right.\right.\right.$$
$$\left.\left.\left. - \frac{1}{2}J_{\hat{x}\hat{x}}\right)PH^TR^{-1}\Omega\right)\right] + \sum_{ijk}(P\Lambda PH^TR^{-1}HP)_{ijk}(PJ_{D\hat{x}})_{jik} = 0. \tag{32}$$

For the function J defined as

$$J(\hat{x}, D, t) = \frac{1}{2}\hat{x}^TS\hat{x} + \phi^T\hat{x} + tr(ND) + \frac{1}{2}\epsilon \tag{33}$$

where S, ϕ, N, and ϵ are (deterministic) functions of t only, the partial derivatives are

$$\left.\begin{aligned} J_{\hat{x}} &= \hat{x}^TS + \phi^T \\ J_D &= N \\ J_{\hat{x}\hat{x}} &= S \\ J_{D\hat{x}} &= 0 \\ J_t &= \frac{1}{2}\hat{x}^T\dot{S}\hat{x} + \dot{\phi}^T\hat{x} + tr(\dot{N}D) + \frac{1}{2}\dot{\epsilon} \end{aligned}\right\}. \tag{34}$$

It can be verified by substituting Eq. (33) in Eq. (28) and Eq. (34) in Eq. (32) that Eqs. (28) and (32) are satisfied by Eq. (33) if

$$\dot{S} = -SF - F^TS - A + SGB^{-1}G^TS; \quad S(t_f) = S_f \tag{35}$$

$$\dot{N} = -NF - F^TN - A + H^TR^{-1}HP(N - S) + (N - S)PH^TR^{-1}H; \quad N(t_f) = S_f \tag{36}$$

$$\dot{\phi} = (SGB^{-1}G^T - F^T)\phi + SGB^{-1}\{c + Tr[N\Gamma + R^{-1}HP(N - S)PH^T R^{-1}\Omega]\} - Tr(N\Psi);$$

$$\phi(t_f) = 0 \tag{37}$$

and

$$\dot{\epsilon} = \left[G^T\phi + c + Tr[N\Gamma + R^{-1}HP(N - S)PH^T R^{-1}\Omega]\right]^T B^{-1}\left[G^T\phi + c + Tr[N\Gamma + R^{-1}HP(N - S)PH^T R^{-1}\Omega]\right] - tr(AP + SPH^T R^{-1}HP);$$

$$\epsilon(t_f) = tr[S_f P(t_f)]. \tag{38}$$

Therefore, it follows from Eqs. (31) and (34) that the optimal control law can be expressed to first order in h in terms of the solution of the terminal value system of *ordinary* differential equations (Eqs. (35) - (37)) as

$$u = -B^{-1}\{G^T S\hat{x} + c + G^T\phi + Tr[R^{-1}HP(N - S)PH^T R^{-1}\Omega + N\Gamma]\} \tag{39}$$

where P and $\hat{x}$ are given by the initial value system of ordinary differential equations (23) through (26). The implementation of this control law requires that Eqs. (24) and (25) be integrated in real time—a total of $(n/2)(n + 3)$ independent components for n state variables—to provide the current values of $\hat{x}$; the other differential equations can be solved beforehand by integrating Eqs. (23) and (26) forward, then Eqs. (35) through (37) backward.

When Γ, Ω, and Ψ are identically zero, Eqs. (23) through (26), (35), (37), and (39) reduce to the well-known solution

$$u = -B^{-1}[G^T(S\hat{x} + \overline{\phi}) + c]$$

$$\dot{\hat{x}} = F\hat{x} + Gu + PH^T R^{-1}(z - H\hat{x}); \quad \hat{x}(t_0) = \hat{x}_0 \tag{40}$$

$$\dot{\overline{\phi}} = (SGB^{-1}G^T - F^T)\overline{\phi} + SGB^{-1}c; \quad \overline{\phi}(t_f) = 0 \tag{41}$$

of the corresponding classical linear-quadratic-Gaussian problem. Two first-order departures from this classical solution are induced in the optimal control law by first-order nonzero values of these quantities. One is the augmentation of the state-estimation equations by using Eqs. (24) through (26) instead of Eq. (40) for determining $\hat{x}$. The other is the addition of the *deterministic* time function

$$\delta\bar{u} = -B^{-1}\{G^T(\phi - \overline{\phi}) + Tr[R^{-1}HP(N - S)PH^T R^{-1}\Omega + N\Gamma]\} \tag{42}$$

to the control. This structure is displayed schematically in Fig. 1. Since the conditional expected value of the driving term $(z - H\hat{x})$ in Eq. (24) is always zero, so is its prior expected value, and it follows from Eqs. (1) and (24) that the prior expected values of $x(t)$ and $\hat{x}(t)$ are always the same. Therefore, the mean sample trajectory of the optimally controlled system can be determined by Eqs. (23), (35) through (37), (41), (42) and the equations

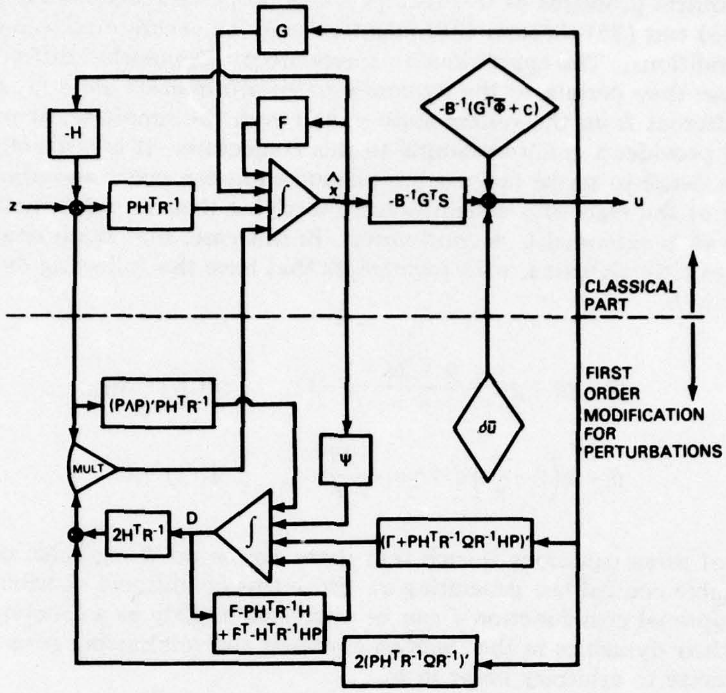

Fig. 1 — Optimal control law structure

$$\dot{\bar{x}} = F\bar{x} + Gu; \qquad \bar{x}(t_0) = \hat{x}_0 \tag{43}$$

$$\bar{u} = -B^{-1}\left[G^T(S\bar{x} + \bar{\phi}) + c\right] + \delta\bar{u} \tag{44}$$

where $\bar{x}(t)$ and $\bar{u}(t)$ denote the (prior) expected values of the state and control at time t under the optimal control law. If the control problem here represents a second-order description of the effects of perturbations about a nominal path in an iteration of a gradient algorithm, this mean sample trajectory is a natural candidate for the nominal path in the next iteration. Such a gradient algorithm would converge in general to a nominal path that is different from the deterministic optimal (with $c + G^T\bar{\phi} = B\delta\bar{u}$ instead of zero), the result representing a compromise between the deterministic optimal and a path encountering the lowest expected noise intensities. Thus the noise statistics in this context enter into the optimization of the *nominal* path as well as the correction of noise-induced deviations from this nominal.

It is interesting to note from Eqs. (33), (37), and (38) that the introduction of first-order values of Γ, Ω, and Ψ only changes the optimal cost function from the classical value by second order when c, $\hat{x}$, and D are all zero, a condition of particular interest when the problem arises from a perturbation analysis. However, this is easily shown to be the case for the cost function associated with *any* control law which differs only by first order from the classical optimum. Since the only first-order approximations used in the derivation here were the dynamics for $\hat{x}$ and D in the Bellman equation, and since first-order accuracy in the dynamics is sufficient for second-order accuracy in the cost in deterministic control problems of this sort [1], this property suggests that the cost function of Eqs. (33) and (35) through (38) might actually be accurate to second order in h under these conditions. The approximations here are of a somewhat different character, however, because they pertain to the dynamics of an *information state* $(\hat{x}, D)$ that is significantly different from the system state x, and even the simple scalar problem with $c = \Omega = \Psi = 0$ provides a counterexample to this conjecture. It is instructive to examine this example in detail to reveal how such a phenomenon can occur and also to enhance the plausibility of the regularity assumption by verifying that no contradiction arises here when the analysis is extended to second order. In this case, the conditional distribution of the state is exactly Gaussian, with parameters that have the following dynamics to arbitrary order in h:

$$\dot{\hat{x}} = f\hat{x} + gu + \frac{p + 2d}{r}(z - \hat{x}); \qquad \hat{x}(t_0) = \hat{x}_0$$

$$\dot{d} = 2\left(f - \frac{p}{r}\right)d + \gamma u - \frac{2d^2}{r}; \qquad d(t_0) = 0.$$

The structure of these equations is such that there can be no Wong-Zakai correction terms for any reasonable control law generating u. Since this conditional distribution is always Gaussian, the optimal cost function J can be expressed exactly as a function of only $\hat{x}$, d, and t. Using their dynamics in the Bellman equation and minimizing gives the following equations, accurate to arbitrary order in h:

$$u = -\frac{1}{b}\left[gJ_{\hat{x}} + \gamma J_d\right]$$

$$\frac{1}{2}a(\hat{x}^2 + p + 2d) + J_{\hat{x}}f\hat{x} + J_d\left[\left(f - \frac{p}{r}\right)d - \frac{2d^2}{r}\right] + J_t + \frac{(p + 2d)^2}{2r}J_{\hat{x}\hat{x}} - \frac{1}{2b}(gJ_{\hat{x}} + \gamma J_d)^2 = 0$$

$$J(\hat{x}, d, t_f) = \frac{1}{2}s_f(\hat{x}^2 + p(t_f) + 2d).$$

An exact solution to these equations is not available in closed form, but J can be imagined as a power series in $\hat{x}$ and d with time-dependent coefficients, and such a function of the form

$$J = \frac{1}{2}s\hat{x}^2 + \phi\hat{x} + \frac{1}{2}\nu d + \mu d^2 + \lambda\hat{x}d + \frac{1}{2}\epsilon + \text{third-order terms}$$

can be examined as a possible solution. Substituting this expression and its derivatives into the partial differential equation and collecting coefficients of like powers of $\hat{x}$ and d show that all but third-order terms vanish on the left-hand side (remember that d itself is of order h) and that the boundary condition is satisfied exactly if $\nu = s + y$ and if

$$\dot{s} = -2sf - a + \frac{s^2g^2}{b} + \frac{sg\gamma\lambda}{b}; \qquad s(t_f) = s_f$$

$$\dot{y} = 2\left(\frac{p}{r} - f\right)y - \frac{1}{b}(sg\gamma\lambda) - \frac{s^2g^2}{b}; \qquad y(t_f) = 0$$

$$\dot{\phi} = \left(\frac{sg^2}{b} - f\right)\phi + \frac{sg\gamma\nu}{b}; \qquad \phi(t_f) = 0$$

$$\dot{\mu} = 4\left(\frac{p}{r} - f\right)\mu + \frac{y}{r}; \qquad \mu(t_f) = 0$$

$$\dot{\lambda} = \left(2\frac{p}{r} - 3f\right)\lambda + \frac{sg}{b}(g\lambda + \gamma\mu); \qquad \lambda(t_f) = 0$$

$$\dot{\epsilon} = -ap - \frac{sp^2}{r} + \frac{1}{b}(g\phi + \gamma\nu)^2; \qquad \epsilon(t_f) = s_f p(t_f).$$

To second order in h, therefore, these equations determine the optimal cost function and thereby give the optimal control law as

$$u = -\frac{1}{b}\left[g(s\hat{x} + \phi) + \gamma\nu + \frac{1}{2}\gamma\lambda\hat{x} + g\lambda d + 2\gamma\nu d\right].$$

Because $-s^2g^2/b$ is the only zeroth-order driving term for $\dot{y}$, y is of order unity, and since y/r is the driving term for $\dot{\mu}$, so is μ. But $\gamma\mu$ appears as the driving term for $\dot{\lambda}$, so λ is of order h, which implies that s here, and therefore ϵ as well, differ from those given by Eq. (35) and (38) by order h^2. Thus the systematic inclusion of second-order effects introduces second-order changes in both the optimal cost function and the optimal control law, even when c, $\hat{x}$, and d are all zero. Nevertheless, the optimal control law remains unchanged to *first* order. This phenomenon appears to depend on the coefficient λ of the $\hat{x}d$ term in the optimal cost function, and hence on a property of the information state, correlation between the dynamics of $\hat{x}$ and d, which has no counterpart of the system state x.

PERFECT STATE MEASUREMENTS

In the limiting case in which current state x is known exactly and can be used in the control law, an optimal cost function $J(x, t)$ can be defined as the conditional expected cost-to-go under an optimal control law given that $x(t) = x$. In this case the Bellman equation corresponding to Eqs. (1) and (3) can be derived in the usual way to give

$$\min_u \left\{J_x(Fx + Gu) + J_t + tr\left[J_{xx}\left(\frac{1}{2}Q + \Gamma'u + \Psi'x\right)\right] + \frac{1}{2}(x^TAx + u^TBu) + c^Tu\right\} = 0 \qquad (45)$$

to first order in h. Differentiating to determine the minimizing control gives

$$u = -B^{-1}[G^T J_x^T + c + Tr(J_{xx}\Gamma)]. \tag{46}$$

If Eq. (46) is substituted into Eq. (45) to eliminate the minimization operation, and if the function

$$J(x, t) = \frac{1}{2} x^T S(t) x + \eta^T(t) x + \frac{1}{2} \delta(t) \tag{47}$$

and its partial derivatives are substituted into the resulting equation, the left-hand side is a quadratic polynomial in x whose coefficients are all identically zero if S satisfies Eq. (35) and if

$$\dot{\eta} = (SGB^{-1}G^T - F^T)\eta + SGB^{-1}[c + Tr(S\Gamma)] - Tr(S\Psi); \qquad \eta(t_f) = 0 \tag{48}$$

and

$$\dot{\delta} = [G^T\eta + c + Tr(S\Gamma)]^T B^{-1}[G^T\eta + c + Tr(S\Gamma)] - tr(SQ); \qquad \delta(t_f) = 0. \tag{49}$$

Furthermore, the cost function of Eq. (47) satisfies the boundary condition

$$J(x, t_f) = \frac{1}{2} x^T S_f x \tag{50}$$

for the terminal values given in Eqs. (35), (48), and (49). Substituting into Eq. (46) and assuming uniqueness and sufficient regularity of solutions to the Bellman equation gives the optimal control law here to first order in h as

$$u = -B^{-1}[G^T(Sx + \eta) + c + Tr(S\Gamma)]. \tag{51}$$

This control law differs from the classical optimum only by the addition of the deterministic time function $-B^{-1}[G^T(\eta - \bar{\phi}) + Tr(S\Gamma)]$. The mean sample trajectory of the optimally controlled system is given by Eqs. (35), (43), (44), and (48), except that this time function replaces $\delta\bar{u}$ in Eq. (44). Again, the covariance matrix perturbations cause only a second-order change in the optimal expected cost if $x(t_0) = 0$ and $c = 0$.

THE ROLE OF MEASUREMENT NOISE

The effects of measurement noise in this context can be clarified by considering the variable

$$Y(t) = N(t) - S(t). \tag{52}$$

The optimal control law of Eq. (39) can then be expressed in terms of the equations

$$u = -B^{-1}[G^T(S\hat{x} + \eta) + c + Tr(S\Gamma) + G^T\theta + Tr(Y\Gamma + R^{-1}HPYPH^TR^{-1}\Omega)] \quad (53)$$

$$\dot{Y} = Y(PH^TR^{-1}H - F) + (H^TR^{-1}HP - F^T)Y - SGB^{-1}G^TS; \quad Y(t_f) = 0 \quad (54)$$

$$\dot{\theta} = (SGB^{-1}G^T - F^T)\theta + SGB^{-1}Tr(Y\Gamma + R^{-1}HPYPH^TR^{-1}\Omega) - Tr(Y\Psi); \quad \theta(t_f) = 0 \quad (55)$$

where P, $\hat{x}$, S, and η are as given earlier by Eqs. (23) through (26), (35), and (48). Comparing this realization with Eq. (51) shows that it is the same control law as the optimum for the case of perfect measurements except for the replacement of x by $\hat{x}$ *and* the addition of the deterministic quantity $-B^{-1}[G^T\theta + Tr(Y\Gamma + R^{-1}HPYPH^TR^{-1}\Omega)]$, all of whose terms are coupled to the control through the matrix Y (indirectly in the case of θ, where Y appears in the driving term of differential Eq. (55) defining θ). This structure suggests that the concept of certainty-equivalence here should refer to the replacement of x by $\hat{x}$ in the optimal control law for the case of perfect measurements with the *same* process noise, *not* the completely deterministic case (the two concepts coincide in the classical perturbation-free problem). With this interpretation of certainty-equivalence, the other additive terms in the control law can be regarded as the "dual control" phenomenon identified by Feldbaum [3]. This phenomenon is the deviation of the optimal control from that which exploits the current state information optimally (interpreted here as certainty-equivalent control) for the purpose of improving the quality of this information for future exploitation. Although the influence of the noise covariance matrix perturbations in this dual control phenomenon is mediated by the matrix Y, the values of Y itself are determined entirely by the corresponding classical problem without such perturbations. Hence, Y might be regarded here as a coefficient matrix governing the sensitivity of this classical problem to dual effects caused by noise covariance perturbations of this sort. This matrix plays no role in the classical problem, however, because the conditional covariance of the state cannot be affected by the control there; i.e., there is no interference between the acquisition and exploitation of state information. Also, it follows from Eq. (34) and the definitions of Y and D that

$$Y = J_D - J_{\hat{x}\hat{x}} = 2J_{\overline{P}} - J_{\hat{x}\hat{x}} \quad (56)$$

in the control problem *with* noise covariance matrix perturbations.

Although the filter and control gains can be determined separately in the corresponding classical problem by the independent Riccati equations (Equs. (23) and (35)) for P and S, both of these variables enter into the Eq. (54), which determines Y. This last equation can be regarded as a symmetric linear differential equation in Y with driving term $-SGB^{-1}G^TS$ and zero terminal value. Since B is positive-definite by assumption, this driving term is always at least negative-semidefinite, so $Y(t)$ is symmetric and positive-semidefinite for all $t \leqslant t_f$. The fact that $Y(t_f)$ is zero indicates that dual phenomena are unimportant in control problems of sufficiently short duration, which is intuitively reasonable because there is too little time in such cases to take enough advantage of an improved state estimate to justify the cost of achieving it by nonoptimal explotiation of the current estimate. The estimation error in the classical system obeys the differential equation

$$(\dot{x} - \dot{\hat{x}}) = (F - PH^T R^{-1} H)(x - \hat{x}) + (w - PH^T R^{-1} v). \tag{57}$$

If the system (F, H) is observable, then Eq. (57) is stable, which implies that Eq. (54), for Y, is stable in reverse time.

There is also a connection between the behavior of Y and the information theory of Shannon [9]. The entropy of the state vector in the classical problem can be determined by standard methods in units of "nats" as

$$\frac{1}{2} [n \, \ell n(2\pi e) + \ell n|P|] .$$

If the measurement process is discretized in small time increments of length Δ, and if M is used to denote the value of P immediately before a measurement, the amount of information that measurement provides the controller about the state is given by the resulting reduction in the entropy, which is asymptotically

$$\frac{1}{2} \ell n|M| - \frac{1}{2} \ell n|M - MH^T R^{-1} HM\Delta| .$$

Taking the limit of this difference as $\Delta \to 0$ and dividing by Δ gives the information rate of the measurements in nats per unit time as

$$\frac{1}{2} tr(PH^T R^{-1} H).$$

On the other hand, it follows from Eq. (54) that

$$\frac{d}{dt} |Y| = 2tr(PH^T R^{-1} H)|Y| - 2tr(F)|Y| - tr[SGB^{-1} G^T S \, adj(Y)] . \tag{58}$$

This implies that at least the determinant of Y will remain close to zero if this information rate is high, which again is consistent with the intuitive interpretation of dual control phenomena. Whether the values of all the Y components remain small, however, will also depend on the structure of the observation system in the general multivariate case.

The dual aspect of the optimal control of Eq. (53) arises in two ways. One is by the direct addition of a term depending only on the current values of Y and of the control-dependent noise perturbation coefficients Γ and Ω. The other is through the current value of θ, which in turn depends on all future values of these quantities and of the state-dependent noise perturbation coefficient Ψ as well. These two effects correspond roughly to the phenomena of "caution" and "probing" identified by Bar-Shalom and Tse [10] in connection with their and Meier's more general "wide-sense adaptive" approach [11,12] to dual control problems.

AN ALTERNATE CRITERION

To what extent do the preceding results depend on the special nature of the underlying linear-quadratic-Gaussian control problem? It is instructive to consider a variant of this problem, solved recently by Speyer, Deyst and Jacobson [2], in which the quadratic performance index (3) is replaced by the exponential criterion

$$J = E\left\{\mu \exp\left[\frac{1}{2}\mu\left(x_f^T S_f x_f + \int_{t_0}^{t_f} u^T B u\, dt\right)\right]\right\} \tag{59}$$

where μ is a scalar. If the preceding state and control dependences of the noise covariance matrixes are introduced in this context, the state estimation results are the same as before. Hence it is meaningful to consider an optimal expected cost-to-go function defined to first order in h as

$$J(\hat{x}, D, t) = E\left\{\mu \exp\left[\frac{1}{2}\mu\left(x_f^T S_f x_f + \int_t^{t_f} u^T B u\, dt\right)\right] \Big/ x(t) = \hat{x}, \bar{P}(t) = P(t) + 2D\right\} \tag{60}$$

where u is generated by an optimal control law.

It follows from this definition that

$$J(\hat{x}, D, t_f) = E_{x/\hat{x},D}[\mu e^{1/2\mu x^T S_f x}]. \tag{61}$$

Assuming that $P(t_f)$ from Eq. (23) is invertible and that $P^{-1}(t_f) > \mu S_f$, this expectation can be evaluated for the class of conditional state distributions encountered here by completing the square in the exponent and using standard results for the moments of Gaussian distributions. With much manipulation this result can be expanded to first order in h as

$$\begin{aligned} J(\hat{x}, D, t_f) = {} & \frac{\mu}{|I - \mu P_f^{-1} S_f|^{1/2}} \exp\left\{\mu \hat{x}^T\left(\frac{1}{2} M_f + \mu M_f D M_f\right)\hat{x} + \mu tr(M_f D)\right. \\ & + \mu^2 \hat{x}^T S_f (P_f^{-1} - \mu S_f)^{-1} Tr[S_f (P_f^{-1} - \mu S_f)^{-1} \Lambda_f P_f] \\ & \left. + \frac{\mu^3}{3} tr[\hat{x}\hat{x}^T [S_f (P_f^{-1} - \mu S_f)^{-1} \Lambda_f (P_f^{-1} - \mu S_f)^{-1} S_f]' (P_f^{-1} - \mu S_f)^{-1} S_f \hat{x}]\right\} \end{aligned} \tag{62}$$

where P_f and Λ_f denote $P(t_f)$ and $\Lambda(t_f)$, and

$$M_f = S_f + \mu S_f (P_f^{-1} - \mu S_f)^{-1} S_f. \tag{63}$$

For a small time increment Δ, J obeys the following recursion relation to first order in Δ:

$$J(\hat{x}, D, t) = \min_u E_{x,w,v/\hat{x},D}[e^{1/2\mu\Delta u^T B u} J(\hat{x} + \Delta\hat{x}, D + \Delta D, t + \Delta)], \tag{64}$$

where $\Delta\hat{x}$ and ΔD are the increments in $\hat{x}$ and D that occur in the time interval $[t, t + \Delta]$ when control u is used. Expanding J to first order in Δ with a Taylor series about $(\hat{x}, D, t)$, using the dynamics of Eqs. (24) and (25) and taking expectations in the usual way gives the Bellman equation

$$J(\hat{x}, D, t) = \min_u e^{1/2\mu\Delta u^T B u} \Bigg(J(\hat{x}, D, t) + \Delta\Bigg[J_t + J_{\hat{x}}(F\hat{x} + Gu) + tr \Bigg\{ J_D [(F - PH^T R^{-1} H)D$$

$$+ D(F^T - H^T R^{-1} HP) + \Psi'\hat{x} + \Gamma' u + PH^T R^{-1} \Omega' u R^{-1} HP] + J_{\hat{x}\hat{x}} \Bigg[\frac{1}{2} PH^T R^{-1} HP$$

$$+ DH^T R^{-1} HP + PH^T R^{-1} HD - PH^T R^{-1} \Omega' u R^{-1} HP \Bigg]\Bigg\}$$

$$+ \sum_{ijk} (P\Lambda PH^T R^{-1} HP)_{ijk} (PJ_{D\hat{x}})_{jik} \Bigg]\Bigg). \tag{65}$$

Cyclically permuting matrix products in the trace operand, expanding the exponential factor in a Taylor series, subtracting $J(\hat{x}, D, t)$ from both sides of the resulting equation, dividing by Δ, and neglecting higher order terms in Δ gives

$$\min_u \Bigg(\frac{1}{2} \mu(u^T Bu) J + J_t + J_{\hat{x}}(F\hat{x} + Gu) + u^T Tr[PH^T R^{-1} \Omega R^{-1} HP(J_D - J_{\hat{x}\hat{x}}) + \Gamma J_D]$$

$$+ \hat{x}^T Tr(\Psi J_D) + tr \Bigg\{ [J_D F + F^T J_D - (J_D - J_{\hat{x}\hat{x}}) PH^T R^{-1} H - H^T R^{-1} HP(J_D - J_{\hat{x}\hat{x}})] D$$

$$+ \frac{1}{2} J_{\hat{x}\hat{x}} PH^T R^{-1} HP \Bigg\} + \sum_{ijk} (P\Lambda PH^T R^{-1} HP)_{ijk} (PJ_{D\hat{x}})_{jik} \Bigg) = 0. \tag{66}$$

Equating the u-derivative of Eq. (66) to zero gives the minimizing control as

$$u = -\frac{B^{-1}}{\mu J} \{G^T J_{\hat{x}}^T + Tr[PH^T R^{-1} \Omega R^{-1} HP(J_D - J_{\hat{x}\hat{x}}) + \Gamma J_D]\}. \tag{67}$$

Substituting Eq. (67) into Eq. (66) to eliminate the minimization operation gives the following partial differential equation for J:

$$J_t + [J_{\hat{x}} F + Tr^T(\Psi J_D)]\hat{x} + tr \Bigg\{ [J_D F + F^T J_D - (J_D - J_{\hat{x}\hat{x}}) PH^T R^{-1} H$$

$$- H^T R^{-1} HP(J_D - J_{\hat{x}\hat{x}})] D + \frac{1}{2} J_{\hat{x}\hat{x}} PH^T R^{-1} HP \Bigg\} - \{J_{\hat{x}} G$$

$$+ Tr^T [PH^T R^{-1} \Omega R^{-1} HP(J_D - J_{\hat{x}\hat{x}}) + \Gamma J_D]\} \frac{B^{-1}}{2\mu J} \{G^T J_{\hat{x}}^T \tag{68}$$

(Continued)

$$+ Tr[PH^TR^{-1}\Omega R^{-1}HP(J_D - J_{\hat{x}\hat{x}}) + \Gamma J_D]\} + \sum_{ijk} (P\Lambda PH^TR^{-1}HP)_{ijk}$$

$$(PJ_{D\hat{x}})_{jik} = 0. \qquad (68)$$

An exact solution to Eqs. (68) and (62) is not known. However, for a cost function of the form

$$J(\hat{x}, D, t) = \frac{\mu}{\sqrt{\alpha}} \exp\left[\mu\hat{x}^T\left(\frac{1}{2}M + \mu MDM\right)\hat{x} + \mu\phi^T\hat{x} + \mu tr(ND) + \frac{\mu^2}{3} tr(\hat{x}\hat{x}^T\Pi\hat{x})\right] \qquad (69)$$

where M, N, and Π are symmetric and the components of ϕ and Π are all of order h, it can be verified that

$$J_{\hat{x}} = \mu[\hat{x}^T(M + 2\mu MDM) + \phi^T + \mu\hat{x}^T\Pi\hat{x}]J \qquad (70)$$

$$J_D = \mu(N + \mu M\hat{x}\hat{x}^TM)J \qquad (71)$$

$$J_{\hat{x}\hat{x}} = \mu\{M + 2\mu MDM + 2\mu\Pi\hat{x} + \mu[(M + 2\mu MDM)\hat{x} + \phi + \mu\Pi\hat{x}\hat{x}][\hat{x}^T(M + 2\mu MDM) + \phi^T + \mu\hat{x}^T\Pi\hat{x}]\}J \qquad (72)$$

$$J_{D\hat{x}} = \mu^2[(M\hat{x}^TM)' + (M\hat{x}^TM)'' + (N + \mu M\hat{x}\hat{x}^TM)\hat{x}^TM]J + \text{terms of order } h \qquad (73)$$

$$J_t = \mu\left[\hat{x}^T\left(\frac{1}{2}\dot{M} + \mu\dot{M}DM + \mu MD\dot{M}\right)\hat{x} + \dot{\phi}^T\hat{x} + tr(\dot{N}D) + \frac{\mu}{3}tr(\hat{x}\hat{x}^T\dot{\Pi}\hat{x}) - \frac{\dot{\alpha}}{2\mu\alpha}\right]J. \qquad (74)$$

Substituting Eqs. (69) through (74) into Eqs. (62) and (68), neglecting terms of second order in h, and equating coefficients of like powers of $\hat{x}$ and D shows after much manipulation that the Bellman equation and boundary condition are satisfied to first order in h if

$$N = M + Y \qquad (75)$$

and

$$\dot{M} = -MF - F^TM + M(GB^{-1}G^T - \mu PH^TR^{-1}HP)M; \quad M(t_f) = M_f \qquad (76)$$

$$\dot{Y} = Y(PH^TR^{-1}H - F) + (H^TR^{-1}HP - F^T)Y - MGB^{-1}G^TM; \quad Y(t_f) = 0 \qquad (77)$$

$$\dot{\phi} = [M(GB^{-1}G^T - \mu PH^TR^{-1}HP) - F^T]\phi + MGB^{-1}Tr(\Gamma N + PH^TR^{-1}\Omega R^{-1}HPY) - \mu MPH^TR^{-1}HP\,Tr(P\Lambda PY) - Tr(\Psi N + \mu PH^TR^{-1}HP\Pi) - \mu M\,Tr(\{(P\Lambda PH^TR^{-1}HP)'P + [(P\Lambda PH^TR^{-1}HP)'P]' + [(P\Lambda PH^TR^{-1}HP)'P]''\}M); \quad \phi(t_f) = \mu S_f(P_f^{-1} - \mu S_f)^{-1}Tr[S_f(P_f^{-1} - \mu S_f)^{-1}\Lambda_f P_f] \qquad (78)$$

$$\left.\begin{aligned}&\dot{\Pi} = \Theta + \Theta' + \Theta''; \quad \Pi(t_f) = \mu[S_f(P_f^{-1} - \mu S_f)^{-1}\Lambda_f(P_f^{-1} - \mu S_f)^{-1}S_f]'(P_f^{-1} - \mu S_f)^{-1}S_f \\ &\Theta = \Pi[(GB^{-1}G^T - \mu PH^T R^{-1}HP)M - F] + [(M\Gamma M)'B^{-1}G^T - M\Psi - \mu(MP\Lambda PM)' \\ &\quad PH^T R^{-1}HP]M\end{aligned}\right\} \quad (79)$$

$$\dot{\alpha} = \mu\alpha\, tr(MPH^T R^{-1}HP); \quad \alpha(t_f) = |I - \mu P_f S_f|. \quad (80)$$

Furthermore, the solutions to Eqs. (75) through (80) are such that M, N, and Π are symmetric and ϕ and Π are of order h. Under the appropriate uniqueness and regularity conditions, therefore, they and Eq. (69) determine the optimal cost function to first order in h. From Eq. (67), the optimal control law is

$$u = - B^{-1}\{G^T[M + \mu(\Pi\hat{x} + 2MDM)]\hat{x} + G^T\phi + Tr[PH^T R^{-1}\Omega R^{-1}HPY + \Gamma(N + \mu M\hat{x}\hat{x}^T M)]\} \quad (81)$$

to first order. The variable M here corresponds to the variable $\widetilde{Q}$ in Speyer et al. [2], and these results reduce to theirs when Γ, Ω, and Ψ are all zero.

Role of Measurement Noise

In the limiting case of perfect state measurements, an optimal expected cost-to-go function $J(x, t)$ can be unambiguously defined as the conditional expected cost-to-go under an optimal control law, given $x(t) = x$. A similar derivation shows that the Bellman equation for this case is

$$\left.\begin{aligned}&J_x Fx + tr\left[J_{xx}\left(\frac{1}{2}Q + \Psi'x\right)\right] - \frac{1}{2\mu J}[J_x G + Tr^T(J_{xx}\Gamma)]B^{-1}[G^T J_x^T + Tr(J_{xx}\Gamma)] \\ &\quad + J_t = 0; \quad J(x, t_f) = \mu e^{1/2\mu x^T(t_f)S_f x(t_f)}\end{aligned}\right\} \quad (82)$$

to first order in h, and that the corresponding optimal control law is

$$u = -\frac{B^{-1}}{\mu J}[G^T J_x^T + Tr(J_{xx}\Gamma)]. \quad (83)$$

It is a matter of straightforward substitution to verify that this Bellman equation is satisfied to first order by the function

$$J(x, t) = \frac{\mu}{\sqrt{\overline{\alpha}}}\exp\left[\frac{1}{2}\mu x^T\overline{M}x + \mu\eta^T x + \frac{\mu^2}{3}tr(xx^T\overline{\Pi}x)\right] \quad (84)$$

if

$$\dot{\overline{M}} = -\overline{M}F - F^T\overline{M} + \overline{M}(GB^{-1}G^T - \mu Q)\overline{M}; \quad \overline{M}(t_f) = S_f \tag{85}$$

$$\dot{\eta} = [\overline{M}(GB^{-1}G^T - \mu Q) - F^T]\eta + \overline{M}GB^{-1}Tr(\Gamma\overline{M}) - Tr(\Psi\overline{M} + \mu\overline{\Pi}Q); \quad \eta(t_f) = 0 \tag{86}$$

$$\left.\begin{aligned} &\dot{\overline{\Pi}} = \overline{\Theta} + \overline{\Theta}' + \overline{\Theta}''; \quad \overline{\Pi}(t_f) = 0, \text{ where} \\ &\overline{\Theta} = \overline{\Pi}[(GB^{-1}G^T - \mu Q)\overline{M} - F] + (\overline{M}\Gamma\overline{M})'B^{-1}G^T\overline{M} - \overline{M}\Psi\overline{M} \end{aligned}\right\} \tag{87}$$

$$\dot{\overline{\alpha}} = \mu\overline{\alpha}\, tr(\overline{M}Q); \quad \overline{\alpha}(t_f) = 1 \tag{88}$$

and that the optimal control law can be expressed as

$$u = -B^{-1}\{G^T(\overline{M} + \mu\overline{\Pi}x)x + G^T\eta + Tr[\Gamma(\overline{M} + \mu\overline{M}xx^T\overline{M})]\}. \tag{89}$$

This reduces to the result of Jacobson [13] when Γ and Ψ are identically zero.

To investigate the dual control phenomena here, it is necessary to determine first which part of the optimal control law (Eq. (81)) constitutes "optimal exploitation of current state information." It was natural to interpret the latter as an extended form of certainty-equivalent control in the case of the quadratic criterion, but it is clear from the results of Jacobson [13] and Speyer et al. [2] that this form of certainty-equivalence does not even hold for the exponential criterion in the classical case without noise covariance perturbations, because M differs from $\overline{M}$. However, a natural extension of this property does hold in this case. *Comparing derivatives and boundary conditions shows* that

$$M = S[I - \mu(K+P)S]^{-1} = [I - \mu S(K+P)]^{-1}S = S + \mu S[(K+P)^{-1} - \mu S]^{-1}S \tag{90}$$

where S is as given by Eq. (35) with $A = 0$ and

$$\dot{K} = FK + KF^T - Q; \quad K(t_f) = 0 \tag{91}$$

and that

$$\overline{M} = S(I - \mu KS)^{-1} = (I - \mu SK)^{-1}S.$$

Since S and K are independent of the measurement process parameters, this means that the instantaneous value of the optimal control for both noisy and perfect measurements can be realized as the functional composition

$$u = -B^{-1}G^T[I - \mu S(K+P)]^{-1}S\hat{x} \tag{92}$$

of a control law determined entirely by the problem with perfect measurements operating on the mean $\hat{x}$ and covariance matrix P of the current conditional state distribution, where these parameters are taken respectively as x and 0 in the case of perfect measurements. This decomposition therefore shares the essential properties of the refined certainty-equivalence concept described in the preceding section. The main difference is that the construction here is slightly more elaborate and involves the covariance matrix generated by the state estimator as well as the mean.

This idea can be extended to the context of noise covariance perturbations by showing from Eq. (90) that

$$M + 2\mu MDM = [I - \mu S(K + P + 2D)]^{-1}S$$

to first order in h, and showing from Eqs. (90), (87), (79), (35), (26), and (23) that

$$\Pi = \{[I - \mu S(K + P)]^{-1}[\Upsilon + \mu(SP\Lambda PS)'PS][I - \mu(K + P)S]^{-1}\}'[I - \mu(K + P)S]^{-1}, \quad (93)$$

where

$$\dot{\Upsilon} = \Theta + \Theta' + \Theta''; \qquad \Theta = \Upsilon(GB^{-1}G^TS - F) + (S\Gamma S)'B^{-1}G^TS$$
$$- (S\Psi S)'(I - \mu KS); \qquad \Upsilon(t_f) = 0 \quad (94)$$

and that $\overline{\Pi}$ is given by Eq. (93) with $P = 0$ and $\Lambda = 0$. Since Eq. (94) is also independent of the measurement process, it follows from Eq. (81) that a similar realization of the optimal control law here can be constructed to first order in h as

$$u = -B^{-1}(G^T[I - \mu S(K + P + 2D)]^{-1}S\hat{x} + \mu Tr\{[(\Pi G)'' + M\Gamma M]\hat{x}\hat{x}^T\} + Tr(\Gamma M)$$
$$+ G^T\phi + Tr[(\Gamma + PH^TR^{-1}\Omega R^{-1}HP)Y]) \quad (95)$$

where M and Π now denote the expressions in Eqs. (90) and (93). The optimal control for the case of perfect measurements is given by Eq. (95) with $\hat{x} = x$ and $P = D = 0$ there and in Eqs. (90) and (93), and with the last two terms of Eq. (95) replaced by $G^T\eta$.

This construction shows that the optimal control law can be realized as the sum of a certainty-equivalent control law, in the extended sense proposed here, and a residual deterministic term. If "optimal exploitation of current state information" is interpreted as certainty-equivalent control in this sense, then the dual control phenomenon here is this residual term, an additive deterministic time function as in the case of the quadratic criterion. The portion of the deterministic terms in Eq. (95) to be included in the certainty-equivalent control law is somewhat arbitrary, however, because this form of certainty-equivalence allows the use of the deterministic time functions P and Λ as arguments in the control law. It would be ideal if ϕ in Eq. (95) could be decomposed as

$$\phi(t) = f_1[P(t), \Lambda(t), t] + Y(t)f_2[P(t), \Lambda(t), t] + f_3[P(t), t]\theta(t) \quad (96)$$

with

$$\dot{\theta}(t) = L(t)\theta(t) + Y(t)f_4[P(t), \Lambda(t), t]; \qquad \theta(t_f) = 0$$
$$\eta(t) = f_1(0, 0, t)$$

where f_1 through f_4 are independent of the measurement process and L depends only on the completely deterministic problem. The dual control could then be reasonably identified in analogy with the case of the quadratic criterion as the deterministic term

$$-B^{-1}\{G^T[Yf_2(P,\Lambda) + f_3(P)\theta] + tr[(\Gamma + PH^TR^{-1}\Omega R^{-1}HP)Y]\}$$

coupled to the control through the Y matrix.

Such a decomposition of ϕ has not been found for the general case. In the special case of classical process noise ($\Gamma = 0$, $\Psi = 0$), however, Λ, Υ, and η are all zero, and the optimal control law can be realized as

$$u = -B^{-1}\{G^T[I - \mu S(K + P + 2D)]^{-1}S\hat{x} + G^T[I - \mu S(K + P)]^{-1}\theta + Tr(PH^TR^{-1}\Omega R^{-1}HPY)\}$$

where θ is given by Eq. (55) with $\Gamma = 0$, $\Psi = 0$, and Y as given by Eq. (77) rather than Eq. (54). The dual control in this case is therefore

$$-B^{-1}\{G^T[I - \mu S(K + P)]^{-1}\theta + Tr(PH^TR^{-1}\Omega R^{-1}HPY)\}.$$

The variable Y in this context differs from that for the quadratic criterion because M replaces S in the driving term of the defining differential equation, Eq. (77). Also, it follows from Eqs. (71) and (72) that

$$Y = \frac{J_D - J_{\hat{x}\hat{x}}}{\mu J}$$

except for approximately infinitesimal terms. The qualitative behavior of this Y remains the same as that of Eq. (54), however.

MEASUREMENT NOISE STATE DEPENDENCE

The case of state-dependent measurement noise covariance matrixes is analyzed in discrete time because of difficulties described earlier. For this purpose, the following control problem is considered:

$$x_{i+1} = F_ix_i + G_iu_i + w_i; \quad x_0 \text{ (Normal } (\hat{x}_0, P_0)\text{)} \quad \text{(dynamics)}$$

$$z_i = H_ix_i + v_i \qquad \text{(state measurements)}$$

$$J = \frac{1}{2} E\left[x_N^TS_fx_N + \sum_{i=0}^{N-1} (x_i^TA_ix_i + u_i^TB_iu_i)\right] \quad \text{(criterion to be minimized)}$$

where F_i^{-1} exists, and $\{w_i\}$ and $\{v_i\}$ are independent zero-mean normal random variables, given the current state and control history such that

$$\text{cov}(w_i) = Q_i + 2\Gamma_i'u_i + 2\Psi_i'x_i; \qquad i = 0, ..., N-1$$

$$\text{cov}(v_i) = R_i + 2\Omega_{i-1}'u_{i-1} + 2\Upsilon_i'x_i; \qquad i = 1, ..., N$$

and the components of Γ_i, Ψ_i, Ω_{i-1}, and Υ_i are all of order h. The convention here is that measurement z_i is available at epoch i before control u_i is chosen, except at the initial epoch, when there is no measurement and u_0 is chosen on the basis of the prior state distribution. The linear term in the control is excluded from the criterion for simplicity, but this is otherwise the discrete-time analog of the control problem given by Eqs. (1) through (3).

State Estimation

Suppose that the conditional density of the state at epoch i after the receipt of z_i is of the form of Eq. (8), with parameters $\bar{x}_i$, V_i, λ_i, L_i, Λ_i and with corresponding mean and covariance denoted by $\hat{x}_i$ and $\bar{P}_i$. The density of w_i given x_i and u_i is zero-mean Normal with covariance matrix $Q + 2\Psi_i' x_i$, where Q denotes $Q_i + 2\Gamma_i' u_i$. Letting s denote $F_i x_i$ implies that s has a density of the form of Eq. (8), such that

$$\bar{x} = F_i \bar{x}_i$$

$$V = F_i V_i F_i^T$$

$$\lambda^T = \lambda_i^T F_i^{-1}$$

$$L = (F_i^{-1})^T L_i F_i^{-1}$$

$$\Lambda = [(F_i^{-1})^T \Lambda_i F_i^{-1}] F_i^{-1}$$

by Eq. (11). Again assuming for convenience that Q_i^{-1} exists, it follows that

$$p_{w_i/s}(w, s) = p_{w_i/x_i}(w, F_i^{-1} s)$$

$$= \{1 - tr[Q^{-1}\Psi' s(I - Q^{-1} w w^T)]\} \frac{e^{-1/2 w^T Q^{-1} w}}{(2\pi)^{1/2n} |Q|^{1/2}}$$

to first order in h, where $\Psi' = \Psi_i' F_i^{-1}$. If $r = s + w_i$, then

$$p(r) = p(s + w_i) = \int_{R^n} p_s(r - w) p_{w_i/s}(w, r - w)\, dw$$

$$= \int_{R^n} \frac{k(w)}{(2\pi)^n |VQ|^{1/2}} e^{-1/2[w^T Q^{-1} w + (r-w-\bar{x})^T V^{-1} (r-w-\bar{x})]}\, dw$$

to first order in h, where

$$k(w) = 1 + \lambda^T (r - w - \bar{x}) + tr\left\{\frac{1}{2} L[(r - w - \bar{x})(r - w - \bar{x})^T - V]\right.$$

$$\left. + \frac{1}{3} (r - w - \bar{x})(r - w - \bar{x})^T \Lambda (r - w - \bar{x}) - Q^{-1}\Psi'(r - w)[I - Q^{-1} w w^T]\right\}.$$

Completing the square in the exponent gives

$$p(r) = \frac{e^{-1/2(r-\bar{x})^T M^{-1}(r-\bar{x})}}{(2\pi)^{1/2n}|M|^{1/2}} \int_{R^n} k(w) \frac{e^{-1/2[w-(I-VM^{-1})(r-\bar{x})]^T(V-VM^{-1}V)^{-1}[w-(I-VM^{-1}(r-\bar{x})]}}{(2\pi)^{1/2n}|V - VM^{-1}V|^{1/2}}\, dw$$

where $M = V + Q$. The integral is the expected value of a third-degree polynomial in w and $(r - \bar{x})$ with respect to a Normal distribution whose mean is proportional to $(r - \bar{x})$. Therefore, it can be expressed in the form

$$\text{constant} + \bar{\lambda}^T(r-\bar{x}) + tr\left\{\frac{1}{2}\bar{L}[(r-\bar{x})(r-\bar{x})^T - M] + \frac{1}{3}(r-\bar{x})(r-\bar{x})^T\bar{\Lambda}(r-\bar{x})\right\},$$

where the constant is independent of $(r - \bar{x})$. $\bar{L}$ can be taken as symmetric because M is, and $\bar{L}$ and $\bar{\lambda}$ are of order h because λ, L, and Λ are. Since $p(r)$ is a probability density, the constant term must be unity, by Eq. (8). Carrying out the details of the third-degree terms in this expectation shows that

$$\bar{\Lambda} = (M^{-1}V\Lambda VM^{-1})'VM^{-1} + (M^{-1}\Psi M^{-1})'VM^{-1} + [(M^{-1}\Psi M^{-1})'VM^{-1}]'$$
$$+ [(M^{-1}\Psi M^{-1})'VM^{-1}]''$$

which is also symmetric and of order h. Therefore, $p(r)$ is of the form of Eq. (8). Decomposing expectations into marginals of conditionals over s shows that

$$E(r) = E(s) = F_i\hat{x}_i$$

and

$$\text{cov}(r) = \text{cov}(s) + Q + 2\Psi'\bar{x} = F_i\bar{P}_iF_i^T + Q_i + 2\Gamma_i'u_i + 2\Psi_i'\hat{x}_i$$

to first order in h.

If x_{i+1} is denoted by y for convenience, then $\bar{y} = r + G_iu_i$ and $\bar{y}$ has a density of the form of Eq. (8) with parameters $(\bar{x}, M, \bar{\lambda}, \bar{L}, \bar{\Lambda})$, where $\bar{M}$ and $\bar{\Lambda}$ are as used previously, $\bar{x}$ denotes the preceding variable $\bar{x}$ plus G_iu_i, and $\bar{\lambda}$ and $\bar{L}$ are such that

$$E(y) = F_i\hat{x}_i + G_iu_i \triangleq \bar{x}_{i+1} \quad (97)$$

$$\text{cov}(y) = \text{cov}(r) \triangleq N$$

to first order in h. Denoting z_{i+1}, Υ_{i+1}, and H_{i+1} by z, Υ, and H implies that

$$p(z/y) = [1 - tr\{R^{-1}\Upsilon'y[I - R^{-1}(z - Hy)(z - Hy)^T]\}]\,\frac{e^{-1/2(z-Hy)^T R^{-1}(z-Hy)}}{(2\pi)^{1/2k}|R|^{1/2}}$$

to first order in h for a specified u_i, where $R = R_{i+1} + 2\Omega_i' u_i$. As a function of y, the conditional density $p(y/z)$ is proportional to $p(z/y)p(y)$. Completing the square in the exponent of this product and using the "matrix inversion lemma" gives

$$p(y/z) = g\left(1 + \bar{\lambda}^T(y - \bar{x}) + tr\left\{\frac{1}{2}\,\bar{L}[(y - \bar{x})(y - \bar{x})^T - M] + \frac{1}{3}\,(y - \bar{x})(y - \bar{x})^T\bar{\Lambda}(y - \bar{x}) - R^{-1}\Upsilon' y[I - R^{-1}(z - Hy)(z - Hy)^T]\right\}\right)\frac{\exp -\frac{1}{2}\,(y - \bar{y})^T\bar{V}^{-1}(y - \bar{y})}{(2\pi)^{1/2n}|\bar{V}|^{1/2}}$$

where g is a constant of proportionality and

$$\bar{V} = M - MH^T(R + HMH^T)^{-1}HM$$

$$\bar{y} = \bar{x} + \bar{V}H^TR^{-1}(z - H\bar{x}).$$

The polynomial factor in $p(y/z)$ can be expressed to first order in h as

$$1 + b^T(y - \bar{y}) + tr\left\{\frac{1}{2}\,L^*[(y - \bar{y})(y - \bar{y})^T - V] + \frac{1}{3}\,(y - \bar{y})(y - \bar{y})^T\Lambda^*(y - \bar{y})\right\}$$

where

$$\begin{aligned} b = \bar{\lambda} &+ [\bar{L}\bar{V}H^TR^{-1} - 2H^TR^{-1}\Upsilon'\bar{x}(R + HMH^T)^{-1}](z - H\bar{x}) \\ &+ Tr\,(z - H\bar{x})(z - H\bar{x})^T\{R^{-1}H\bar{V}\bar{\Lambda}\bar{V}H^TR^{-1} - 2R^{-1}H\bar{V}[(R + HMH^T)^{-1}\Upsilon R^{-1}H]'' \\ &+ (R + HMH^T)^{-1}\Upsilon(R + HMH^T)^{-1}\} - \Upsilon R^{-1} \end{aligned}$$

$$\begin{aligned} L^* = \bar{L} &+ 2(H^TR^{-1}\Upsilon R^{-1}H)'\bar{x} + 2\{[\bar{\Lambda} + (H^TR^{-1}\Upsilon R^{-1}H)']MH^T - H^TR^{-1}\Upsilon \\ &- (\Upsilon R^{-1}H)''\}(R + HMH^T)^{-1}(z - H\bar{x}) \end{aligned}$$

$$\Lambda^* = \bar{\Lambda} + H^TR^{-1}\Upsilon R^{-1}H + (H^TR^{-1}\Upsilon R^{-1}H)' + (H^TR^{-1}\Upsilon R^{-1}H)''.$$

Since L^* and Λ^* are symmetric, $p(y/z)$ is a density of the form of Eq. (8). With the use of earlier definitions and results, it follows from Eqs. (9) and (10) that its mean $\hat{x}_{i+1}$ and covariance matrix $\bar{P}_{i+1}$ are given to first order in h by

$$\begin{aligned} \hat{x}_{i+1} = \bar{x}_{i+1} &+ NH^T(R + 2\Upsilon'\bar{x}_{i+1} + HNH^T)^{-1}(z - H\bar{x}_{i+1}) + \bar{V}\,Tr\{(R + HNH^T)^{-1}[HN\bar{\Lambda}NH^T \\ &+ \Upsilon - 2(H^TR^{-1}\Upsilon'NH)'](R + HNH^T)^{-1}[(z - H\bar{x}_{i+1})(z - H\bar{x}_{i+1})^T - (R + HNH^T)]\} \end{aligned}$$

$$\begin{aligned} \bar{P}_{i+1} = N &- NH^T(R + 2\Upsilon'\bar{x}_{i+1} + HNH^T)^{-1}HN + 2\{[\bar{V}(\bar{\Lambda} + H^TR^{-1}\Upsilon R^{-1}H)\bar{V}]'NH^T \\ &- [\bar{V}(H^TR^{-1}\Upsilon' + \Upsilon''R^{-1}H)\bar{V}]'\}(R + HNH^T)^{-1}(z - H\bar{x}_{i+1}). \end{aligned}$$

If the covariance matrix perturbations D_i are defined as $(\overline{P}_i - P_i)/2$ where $\overline{P}_i$ denotes the conditional state covariance at epoch i and $\{P_i\}$ is the sequence of nominal covariance matrixes defined recursively as

$$P_i = M_i - M_i H_i^T (R_i + H_i M_i H_i^T)^{-1} H_i M_i; \qquad P_0 = P_0 \tag{98}$$

$$M_{i+1} = F_i P_i F_i^T + Q_i \tag{99}$$

then it follows by induction on i that the components of D_i are of order h. It is a straightforward but lengthy matter of substitution into the original definitions to show by induction that conditional mean $\hat{x}_i$ and covariance matrix $(P_i + 2D_i)$ of the state at epoch i is determined to first order in h by Eqs. (97) through (99) and the following equations:

$$\begin{aligned}
\hat{x}_{i+1} = \bar{x}_{i+1} &+ \{[M_{i+1} + 2(I - P_{i+1} H_{i+1}^T R_{i+1}^{-1} H_{i+1})(F_i D_i F_i^T + \Gamma_i' u_i + \Psi_i' \hat{x}_i)] H_{i+1}^T \\
&- 2P_{i+1} H_{i+1}^T R_{i+1}^{-1} [(\Omega_i' + \Upsilon_{i+1}' G_i) u_i + \Upsilon_{i+1}' F_i \hat{x}_i]\}(R_{i+1} + H_{i+1} M_{i+1} H_{i+1}^T)^{-1} \\
&\times (z_{i+1} - H_{i+1}\bar{x}_{i+1}) + P_{i+1} Tr\{(R_{i+1} + H_{i+1} M_{i+1} H_{i+1}^T)^{-1} [\Upsilon_{i+1} \\
&+ H_{i+1} M_{i+1} \overline{\Lambda}_{i+1} M_{i+1} H_{i+1}^T - 2(H_{i+1}^T R_{i+1}^{-1} \Upsilon_{i+1}' M_{i+1} H_{i+1}^T)'](R_{i+1} \\
&+ H_{i+1} M_{i+1} H_{i+1}^T)^{-1} [(z_{i+1} - H_{i+1}\bar{x}_{i+1})(z_{i+1} - H_{i+1}\bar{x}_{i+1})^T - (R_{i+1} \\
&+ H_{i+1} M_{i+1} H_{i+1}^T)]\}; \qquad \hat{x}_0 = \hat{x}_0
\end{aligned} \tag{100}$$

$$\begin{aligned}
D_{i+1} = (I - P_{i+1} H_{i+1}^T R_{i+1}^{-1} H_{i+1})&(F_i D_i F_i^T + \Gamma_i' u_i + \Psi_i' \hat{x}_i)(I - H_{i+1}^T R_{i+1}^{-1} H_{i+1} P_{i+1}) \\
&+ P_{i+1} H_{i+1}^T R_{i+1}^{-1} [\Omega_i' u_i + \Upsilon_{i+1}' (F_i \hat{x}_i + G_i u_i)] R_{i+1}^{-1} H_{i+1} P_{i+1} \\
&+ \{[P_{i+1}(\overline{\Lambda}_{i+1} + H_{i+1}^T R_{i+1}^{-1} \Upsilon_{i+1} R_{i+1}^{-1} H_{i+1}) P_{i+1}]' M_{i+1} H_{i+1}^T - [P_{i+1}(H_{i+1}^T R_{i+1}^{-1} \Upsilon_{i+1}' \\
&+ \Upsilon_{i+1}'' R_{i+1}^{-1} H_{i+1}) P_{i+1}]'\}(R_{i+1} + H_{i+1} M_{i+1} H_{i+1}^T)^{-1}(z_{i+1} - H_{i+1}\bar{x}_{i+1}); \\
D_0 = 0 &
\end{aligned} \tag{101}$$

$$\begin{aligned}
\overline{\Lambda}_{i+1} = (M_{i+1}^{-1} F_i P_i \Lambda_i P_i F_i^T M_{i+1}^{-1})' P_i F_i^T M_{i+1}^{-1} &+ \{M_{i+1}^{-1} [\Psi_i' P_i F_i^T + (\Psi_i' P_i F_i^T)' \\
&+ (\Psi_i' P_i F_i^T)''] M_{i+1}^{-1}\}' M_{i+1}^{-1}
\end{aligned} \tag{102}$$

$$\Lambda_i = \overline{\Lambda}_i + H_i^T R_i^{-1} \Upsilon_i R_i^{-1} H_i + (H_i^T R_i^{-1} \Upsilon_i R_i^{-1} H_i)' + (H_i^T R_i^{-1} \Upsilon_i R_i^{-1} H_i)''; \qquad \Lambda_0 = 0. \tag{103}$$

The main conceptual distinction between the state estimation results here and the discrete-time analogs derived earlier for the case of state-independent measurement noise is the appearance of a driving term in Eq. (100) containing the difference between the observed and expected scatter matrix of "innovation vector" $(z_{i+1} - H_{i+1}\bar{x}_{i+1})$. This term is present in the discrete-time version even if the Υ_i are zero, but it vanishes in the continuous-time limit in this case. However, this does not happen for nonzero Υ.

Optimization

If the optimal expected cost-to-go function at epoch $i + 1$ (after z_{i+1} is available) is of the form

$$J(\hat{x}, D, i+1) = \frac{1}{2}\hat{x}^T S_{i+1}\hat{x} + (\eta_{i+1}^T + \theta_{i+1}^T)\hat{x} + tr[(S_{i+1} + Y_{i+1})D] + \frac{1}{2}\epsilon_{i+1}$$

to first order in h for $\hat{x}_{i+1} = \hat{x}$ and $D_{i+1} = D$, where S_{i+1} and Y_{i+1} are symmetric, then the Principle of Optimality implies that the optimal expected cost-to-go at epoch i is

$$J(\hat{x}, D, i) = \min_u E\left\{\frac{1}{2}(x_i^T A_i x_i + u^T B_i u + \hat{x}_{i+1}^T S_{i+1}\hat{x}_{i+1} + \epsilon_{i+1}) + (\eta_{i+1}^T + \theta_{i+1}^T)\hat{x}_{i+1} + tr(S_{i+1} + Y_{i+1})D_{i+1}\right\} \tag{104}$$

to first order in h, given that $\hat{x}_i = \hat{x}$, $u_i = u$, and $D_i = D$. The expectation in Eq. (104) can be evaluated from the dynamics of $\hat{x}$ and D to first order in h as

$$\begin{aligned}
&\frac{1}{2}[\hat{x}^T A_i\hat{x} + u^T(B_i + G_i^T S_{i+1} G_i)u + \epsilon_{i+1} + tr(A_i P_i + S_{i+1}P_{i+1}H_{i+1}^T R_{i+1}^{-1} H_{i+1} M_{i+1})] \\
&\quad + (\eta_{i+1}^T + \theta_{i+1}^T)(F_i\hat{x} + G_i u) + \hat{x}^T F_i^T S_{i+1} G_i u + tr[(I - H_{i+1}^T R_{i+1}^{-1} H_{i+1} P_{i+1}) \\
&\quad \times Y_{i+1}(I - P_{i+1}H_{i+1}^T R_{i+1}^{-1} H_{i+1})(F_i D F_i^T + \Gamma_i' u + \Psi_i'\hat{x}) \\
&\quad + R_{i+1}^{-1} H_{i+1} P_{i+1} Y_{i+1} P_{i+1} H_{i+1}^T R_{i+1}^{-1}(\Omega_i' u + \Upsilon_{i+1}' G_i u + \Upsilon_{i+1}' F_i\hat{x}) \\
&\quad + S_{i+1}(F_i D F_i^T + \Gamma_i' u + \Psi_i'\hat{x}) + A_i D] + \frac{1}{2}\hat{x}^T F_i^T S_{i+1} F_i\hat{x}.
\end{aligned} \tag{105}$$

Equating the u-derivative of Eq. (105) to zero shows that this expectation is minimized if

$$\begin{aligned}
u = -(B_i + G_i^T S_{i+1} G_i)^{-1}\{&G_i^T[S_{i+1}F_i\hat{x} + \eta_{i+1} + \theta_{i+1} + Tr(P_{i+1}H_{i+1}^T R_{i+1}^{-1}\Upsilon_{i+1} R_{i+1}^{-1} H_{i+1} \\
&\times P_{i+1} Y_{i+1})] + Tr[\Gamma_i S_{i+1} + (I - P_{i+1}H_{i+1}^T R_{i+1}^{-1} H_{i+1})\Gamma_i(I - H_{i+1}^T R_{i+1}^{-1} H_{i+1} P_{i+1})Y_{i+1} \\
&+ P_{i+1}H_{i+1}^T R_{i+1}^{-1}\Omega_i R_{i+1}^{-1} H_{i+1} P_{i+1} Y_{i+1}]\}.
\end{aligned} \tag{106}$$

Substituting Eq. (106) into Eq. (105) to eliminate the minimization operation in Eq. (104), and equating coefficients of like powers of $\hat{x}$ and D in the resulting equation shows by induction on $N - i$ that

$$J(\hat{x}, D, i) = \frac{1}{2}\hat{x}^T S_i\hat{x} + (\eta_i^T + \theta_i^T)\hat{x} + tr[(S_i + Y_i)D] + \frac{1}{2}\epsilon_i \tag{107}$$

to first order in h if

$$S_i = A_i + F_i^T[S_{i+1} - S_{i+1}G_i(B_i + G_i^T S_{i+1} G_i)^{-1} G_i^T S_{i+1}]F_i; \qquad S_N = S_f \tag{108}$$

$$\begin{aligned} Y_i = F_i^T[&(I - H_{i+1}^T R_{i+1}^{-1} H_{i+1} P_{i+1}) Y_i (I - P_{i+1} H_{i+1}^T R_{i+1}^{-1} H_{i+1}) \\ &+ S_{i+1} G_i (B_i + G_i^T S_{i+1} G_i)^{-1} G_i^T S_{i+1}] F_i; \qquad Y_N = 0 \end{aligned} \tag{109}$$

$$\begin{aligned} \eta_i = F_i^T \eta_{i+1} - F_i^T S_{i+1} G_i (B_i + G_i^T S_{i+1} G_i)^{-1} &[G_i^T \eta_{i+1} + Tr(\Gamma_i S_{i+1})] \\ &+ Tr(\Psi_i S_{i+1}); \qquad \eta_N = 0 \end{aligned} \tag{110}$$

$$\begin{aligned} \theta_i = F_i^T [&I - S_{i+1} G_i (B_i + G_i^T S_{i+1} G_i)^{-1} G_i^T] [\theta_{i+1} \\ &+ Tr(P_{i+1} H_{i+1}^T R_{i+1}^{-1} \Upsilon_{i+1} R_{i+1}^{-1} H_{i+1} P_{i+1} Y_{i+1})] - F_i^T S_{i+1} G_i (B_i \\ &+ G_i^T S_{i+1} G_i)^{-1} Tr[(I - P_{i+1} H_{i+1}^T R_{i+1}^{-1} H_{i+1}) \Gamma_i (I - H_{i+1}^T R_{i+1}^{-1} H_{i+1} P_{i+1}) Y_{i+1} \\ &+ P_{i+1} H_{i+1}^T R_{i+1}^{-1} \Omega_i R_{i+1}^{-1} H_{i+1} P_{i+1} Y_{i+1}] + Tr[(I - P_{i+1} H_{i+1}^T R_{i+1}^{-1} H_{i+1}) \Psi_i (I \\ &- H_{i+1}^T R_{i+1}^{-1} H_{i+1} P_{i+1}) Y_{i+1}]; \qquad \theta_N = 0 \end{aligned} \tag{111}$$

$$\epsilon_i = \epsilon_{i+1} + tr(A_i P_i + S_{i+1} P_{i+1} H_{i+1}^T R_{i+1}^{-1} H_{i+1} M_{i+1}); \qquad \epsilon_N = tr(S_f P_N) \tag{112}$$

since S_i and Y_i are symmetric and η_i and θ_i are of order h. This implies that the optimal control law here is given to first order in h by Eqs. (97) through (103), (106), and (108) through (111), with u_i and $\hat{x}_i$ replacing u and $\hat{x}$ in Eq. (106).

Role of Measurement Noise

If exact measurements of system state x are available to the controller, then the optimal expected cost-to-go function can be defined directly in terms of the system state. If this function is of the form

$$J(x, i+1) = \frac{1}{2} x^T S_{i+1} x + \eta_{i+1}^T x + \frac{1}{2} \delta_{i+1}$$

to first order in h at epoch $(i + 1)$ for $x_{i+1} = x$, where S_{i+1} is symmetric, then the optimal expected cost-to-go at epoch i is

$$J(x, i) = \min_u E\left[\frac{1}{2} (x^T A_i x + u^T B_i u + x_{i+1}^T S_{i+1} x_{i+1} + \delta_{i+1}) + \eta_{i+1}^T x_{i+1}\right] \tag{113}$$

to first order in h, given that $x_i = x$ and $u_i = u$. It follows from the dynamics of x that the expectation in Eq. (113) is

$$\frac{1}{2}\,[x^T(A_i + F_i^T S_{i+1} F_i)x + u^T(B_i + G_i^T S_{i+1} G_i)u + \delta_{i+1} + tr(S_{i+1}Q_i)]$$
$$+ [G_i^T(S_{i+1}F_i x + \eta_{i+1}) + Tr(\Gamma_i S_{i+1})]^T u + [F_i^T \eta_{i+1}$$
$$+ Tr(\Psi_i S_{i+1})]^T x. \tag{114}$$

Equating the u-derivative of Eq. (114) to zero shows that this expectation is minimized if

$$u = -(B_i + G_i^T S_{i+1} G_i)^{-1}[G_i^T(S_{i+1}F_i x + \eta_{i+1}) + Tr(\Gamma_i S_{i+1})]. \tag{115}$$

Substituting Eq. (115) into Eq. (114) to eliminate the minimization operation of Eq. (113), and equating coefficients of like powers of x in the resulting equation shows by induction on $N - i$ that

$$J(x, i) = \frac{1}{2}\, x^T S_i x + \eta_i^T x + \frac{1}{2}\,\delta_i \tag{116}$$

to first order in h if S_i and η_i are as given by Eqs. (108) and (110), and δ_i is given by the recursion

$$\delta_i = \delta_{i+1} + tr(S_{i+1}Q_i); \qquad \delta_N = 0. \tag{117}$$

Therefore the optimal control law here is as specified by Eqs. (108), (110), and (115) with the formal replacement of u and x by u_i and x_i. Comparing Eq. (106) with Eq. (115) shows that the optimal control laws for noisy and perfect state measurements are related to each other in the same way as their continuous-time counterparts in the case of state-independent measurement noise, with the Y_i here serving as a sequence of coupling matrixes for the dual control terms.

Asymptotic Formulas

A control problem of the form considered in this section can also serve as a discrete-time approximation to an extension of the continuous-time problem of Eqs. (1) through (3) with state-dependent measurement noise, with covariance parameter $R(t) + 2\Omega'(t)u + 2\Upsilon'(t)x$, if t and i are related such that $t = t_0 + i\Delta$, where Δ is the (constant) discretization interval, and if

$$F_i = I + F(t)\Delta$$

$$G_i = G(t)\Delta$$

$$A_i = A(t)\Delta$$

$$B_i = B(t)\Delta$$

$$Q_i = Q(t)\Delta$$

$$\Gamma_i = \Gamma(t)\Delta$$

$$\Psi_i = \Psi(t)\Delta$$

$$R_i = \frac{1}{\Delta} R(t)$$

$$\Omega_i = \frac{1}{\Delta} \Omega(t)$$

$$\Upsilon_i = \frac{1}{\Delta} \Upsilon(t).$$

The case of nonzero $c(t)$ in Eq. (3) is omitted here. If terms of second order in Δ are neglected, and if $\dot{x}(t)$ is used formally to denote the difference

$$\frac{x_{i+1} - x_i}{\Delta}$$

and similarly for other such differences, the results of this section reduce to the following asymptotic form for small Δ, where the "t" argument is suppressed in the following notation:

Filter

$$\dot{\hat{x}} = F\hat{x} + Gu + \{PH^T[I - 2R^{-1}(\Omega' u + \Upsilon'\hat{x})] + 2DH^T\}R^{-1}(z - H\hat{x})$$
$$+ P\, Tr\left\{R^{-1}\Upsilon R^{-1}\left[(z - H\hat{x})(z - H\hat{x})^T - \frac{1}{\Delta} R\right]\right\}; \quad \hat{x}(t_0) = \hat{x}_0 \tag{118}$$

$$\dot{P} = FP + PF^T + Q - PH^T R^{-1} HP; \quad P(t_0) = P_0 \tag{119}$$

$$\dot{D} = (F - PH^T R^{-1} H)D + D(F^T - H^T R^{-1} HP) + (\Gamma + PH^T R^{-1} \Omega R^{-1} HP)'u$$
$$+ (\Psi + PH^T R^{-1} \Upsilon R^{-1} HP)'\hat{x} + \{(P\Lambda P)' PH^T - [P(H^T R^{-1} \Upsilon' + \Upsilon'' R^{-1} H)P]'\}$$
$$\times R^{-1}(z - H\hat{x}); \quad D(t_0) = 0 \tag{120}$$

$$\dot{\Lambda} = \Theta + \Theta' + \Theta'', \Theta \triangleq P^{-1}\Psi P^{-1} + H^T R^{-1} \Upsilon R^{-1} H - \Lambda(F + QP^{-1}); \qquad \Lambda(t_0) = 0 \quad (121)$$

Controller

$$u = -B^{-1}\{G^T(S\hat{x} + \eta + \theta) + Tr[\Gamma S + (\Gamma + PH^T R^{-1}\Omega R^{-1} HP)Y]\} \quad (122)$$

$$\dot{S} = -SF - F^T S - A + SGB^{-1}G^T S; \qquad S(t_f) = S_f \quad (123)$$

$$\dot{Y} = Y(PH^T R^{-1} H - F) + (H^T R^{-1} HP - F^T)Y - SGB^{-1}G^T S; \qquad Y(t_f) = 0 \quad (124)$$

$$\dot{\eta} = -F^T \eta + SGB^{-1}[G^T \eta + Tr(\Gamma S)] - Tr(\Psi S); \qquad \eta(t_f) = 0 \quad (125)$$

$$\dot{\theta} = -F^T\theta + SGB^{-1}[G^T\theta + Tr(\Gamma Y + PH^T R^{-1}\Omega R^{-1} HPY)]$$
$$- Tr(\Psi Y + PH^T R^{-1}\Upsilon R^{-1} HPY); \qquad \theta(t_f) = 0 \quad (126)$$

Optimal expected cost-to-go

$$J(\hat{x}, D, t) = \frac{1}{2}\hat{x}^T S\hat{x} + (\eta + \theta)^T \hat{x} + tr[(S + Y)D] + \frac{1}{2}\epsilon \quad (127)$$

$$\dot{\epsilon} = -tr(AP + SPH^T R^{-1} HP); \qquad \epsilon(t_f) = tr[S_f P(t_f)] \quad (128)$$

Perfect measurements

$$u = -B^{-1}[G^T(Sx + \eta) + Tr(\Gamma S)] \quad (129)$$

$$J(x, t) = \frac{1}{2}x^T Sx + \eta^T x + \frac{1}{2}\delta \quad (130)$$

$$\dot{\delta} = -tr(SQ); \qquad \delta(t_f) = 0 \quad (131)$$

These equations agree with those derived earlier for Υ identically zero. A basically new phenomenon arises for nonzero Υ, however, when a driving term for $\dot{\hat{x}}$ appears that contains the scatter matrix of the measurement vector about its current expected value and depends explicitly on the length Δ of the discretization interval.

In either case, there is a conceptual difference from the continuous-time results derived earlier. As before, the discretization increment Δ must be small enough that $\Delta << h$ in order to justify the retention of terms of order h but not of order Δ in the asymptotic "differential" equations. Since terms of order h^2 were neglected in the underlying discrete-time analysis, however, these asymptotic equations are also only meaningful if $\Delta >> h^2$, an additional constraint that was absent from the earlier continuous-time results. Since the standard deviations of the measurement noise components are of order $\Delta^{-1/2}$ for small Δ, this additional constraint is equivalent to the condition that the measurement noise magnitude be small compared to $1/h$ with high probability, or equivalently that $\Upsilon' H^T z$ as well as $\Upsilon' x$ remain of order h. Furthermore, for a short discretization increment Δ, the random variables

$$\left[(z - H\hat{x})(z - H\hat{x})^T - \frac{1}{\Delta} R\right]\Delta$$

are statistically independent at different time steps to the degree of accuracy of the analysis here, and have zero mean and covariances of order unity. Hence, the cumulative contribution over an interval of order unity of the scatter matrix driving term

$$P\, Tr\left\{R^{-1}\Upsilon R^{-1}\left[(z - H\hat{x})(z - H\hat{x})^T - \frac{1}{\Delta} R\right]\right\}$$

in one filter equation, Eq. (118), is approximately a zero-mean random variable with covariance of order h^2/Δ, since this interval contains the sum of $1/\Delta$ such increments. This means that the constraint $\Delta >> h^2$ is also equivalent to the requirement that the effect of the scatter matrix driving term on the state estimate $\hat{x}$ remain small compared to unity (with high probability). If this inequality is reversed, in fact, the scatter of the state measurements dominates all the other statistics in the state estimate $\hat{x}$ generated by the filter equations, Eqs. (118) through (121), for nonzero Υ, which seems suspicious for realistic applications. This phenomenon suggests that the additional constraint in this context reflects a practical limitation in constructing an appropriate measurement noise representation.

The measurement noise in actual applications is never exactly white anyway, but rather has limited bandwidth and nonzero relaxation time. Thus it is more realistic to *imagine state measurements* that are ordinarily approximated as being corrupted by white noise as having been averaged by some kind of "prefilter" (say a sample-and-hold filter) before reaching the controller. As long as the noise is state-independent, however, the state-estimation results do not depend significantly on the exact form of this prefilter as long as its sampling period is short compared to the system time constants (and to h in the present context), and no serious error is introduced by disregarding its effects and treating the measurement noise as white.

This ceases to be true for the sort of measurement noise state-dependence considered here, where the state-estimation equation, Eq. (118), would depend explicitly on the sampling period Δ of the prefilter. This means that an additional parameter of the measurement process, normally unimportant in practice (namely a time constant equivalent to the sampling period of a sample-and-hold prefilter), must be specified in this context to achieve state-estimation results accurate to order h. Such a time constant may be readily available, however, in applications that are truly digital. Also, the results here would be valid only for measurement noise state dependence sufficiently weak that the dependency parameter h is small compared to the square root of this prefilter time constant Δ in properly adjusted units. In fact, these results suggest that if this state dependence is strong enough and if the measurement noise values become independent over a short enough time interval, then the scatter of the measurements really does contain more information about the state than their average value, which would be a drastic departure from the usual filtering situation. The analyses here break down at this point, however, and do not verify this conjecture.

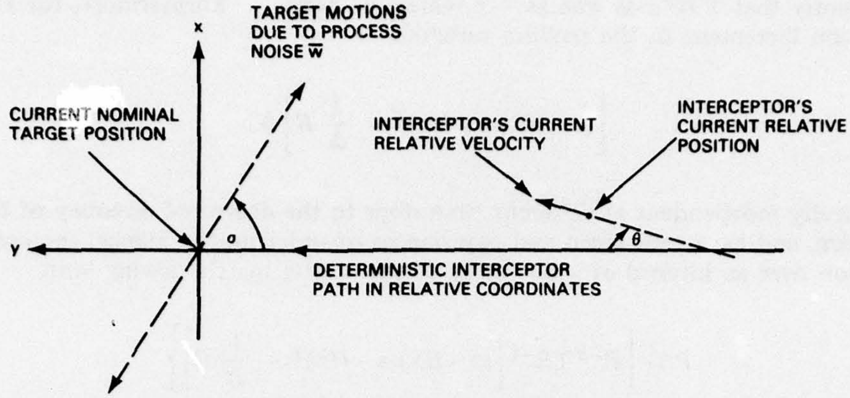

Fig. 2 — Relative motion coordinates

A NUMERICAL EXAMPLE

A numerical illustration of some of the foregoing ideas can be obtained from a planar free-space interception problem in which a homing interceptor has noisy measurements of a target's relative angular position. Any out-of-plane motions are assumed to be controlled independently. The problem developed here is too highly idealized to serve any useful design purpose, but hopefully is still indicative of the basic character of a realistic intercept situation.

The interceptor is assumed to be initially on a collision course with the target, which is subsequently perturbed by a white-noise acceleration along its trajectory, perhaps representing random drag fluctuations. The goal of the interceptor is to minimize a weighted sum of the integrated square of its maneuvering thrust and the square of the distance of closest approach to the target. It is convenient to adopt the relative coordinate system shown in Fig. 2, with the origin fixed at the nominal target position. Random forces acting on the interceptor are disregarded here. Such forces would also be significant in reality, but their inclusion here would only complicate the problem without changing its basic character. Also, the interceptor's control acceleration u is constrained for simplicity to be perpendicular to its current relative velocity (not quite optimal for non-infinitesimal θ). With this constraint, u can be regarded as a scalar, the interceptor's speed is a constant s in relative coordinates, and the interceptor path $[\bar{x}(t), \bar{y}(t)]$ generated by an otherwise general nominal control $\bar{u}(t)$ obeys the equations

$$\dot{\bar{x}} = \bar{u} \sin \bar{\theta}$$

$$\dot{\bar{y}} = -\bar{u} \cos \bar{\theta}$$

$$\dot{\bar{\theta}} = \bar{u}/s.$$

With reference to this nominal path, let t_f be the time of its closest approach to the origin, and for a general realization of the interception process define

$m(t)$ = distance the actual target location at time t_f would be from the closest point on the interceptor's path if no force were applied to either vehicle after time t

$T(t)$ = time at which the interceptor would be at this point on this path.

It follows from these definitions that

$$\dot{m} = (t_f - t)\left[\frac{T - t}{t_f - t}\right]u + (t_f - t)\sin(\sigma + \theta)\overline{w}$$

$$\dot{\theta} = u/s$$

where $\overline{w}$ is the random in-track target acceleration, taken to be a zero-mean Gaussian white noise (GWN) process with constant intensity parameter q. If these dynamics are approximated by neglecting the departures of the ratio $(T - t)/(t_f - t)$ from unity and if $\overline{m}(t)$ denotes the history of m generated by $\bar{u}$, then the deviations from the nominal path reduce to

$$\dot{\tilde{m}} = (t_f - t)\tilde{u} + (t_f - t)\sin(\sigma + \overline{\theta} + \tilde{\theta})\overline{w}$$

$$\dot{\tilde{\theta}} = \tilde{u}/s$$

where $\tilde{m} = m - \overline{m} + \overline{m}(t_f)$, $\tilde{\theta} = \theta - \overline{\theta}$, and $\tilde{u} = u - \bar{u}$. It is assumed that the actual time and distance of closest approach are approximately $T(t_f)$ and $\overline{m}(t_f)$ for a reasonable nominal path generating t_f and $\overline{m}$. If the criterion to be minimized is of the form

$$\overline{J} = \frac{1}{2}E\left\{am^2[T(t_f)] + \int_{t_0}^{T(t_f)} |u|^2\,dt\right\}; \qquad a > 0$$

and the deviations from the nominal are small, then

$$\overline{J} \approx \frac{1}{2}E\left[a\tilde{m}^2(t_f) + \int_{t_0}^{t_f} \tilde{u}^2\,dt\right] + \frac{1}{2}[T(t_f) - t_f]\bar{u}^2(t_f).$$

Assuming that deviations of $T(t_f)$ from t_f are negligible compared to other deviations from the nominal makes this equivalent to minimizing the criterion

$$J = \frac{1}{2}E\left[a\tilde{m}^2(t_f) + \int_{t_0}^{t_f} (\tilde{u}^2 + 2\bar{u}\tilde{u})\,dt\right]$$

if second-order terms in the deviations are disregarded.

For initial conditions it is assumed that $m(t_0) = \overline{m}(t_0)$ and $\theta(t_0) = \overline{\theta}(t_0)$, so the initial values of the state variables are specified as $\tilde{m}(t_0) = \overline{m}(t_f)$ and $\tilde{\theta}(t_0) = 0$. It also follows that $\tilde{\theta}(t)$ remains known exactly. The effects of noisy angular position measurements on the estimate of $\tilde{m}$ can be represented only approximately by noisy measurements of $\tilde{m}$ itself. To make such an approximation, we first assume that the major source of error here is the uncertainty in the relative velocity derived from the angle measurements. Next we consider the one-coordinate free-space system of lateral motions

$$\dot{\lambda} = \nu$$

$$\dot{\nu} = \xi$$

$$\xi \approx \mathrm{GWN}(0, \bar{q}), \qquad (\bar{q} \text{ a constant})$$

with closing speed s and exactly specified initial conditions, for which the state covariance matrix components evolve with elapsed time τ as

$$p_{\nu\nu} = \bar{q}\tau$$

$$p_{\lambda\nu} = \frac{1}{2}\bar{q}\tau^2$$

$$p_{\lambda\lambda} = \frac{1}{3}\bar{q}\tau^3 .$$

On the other hand, if terminal (lateral) position $\lambda(t_f)$ is estimated solely from noisy measurements of λ during a time interval $(t, t + \tau)$ short enough that the process noise disturbances are negligible, it is routine to show that almost the same accuracy is obtained for $\tau << t_f - t$ by lumping the observations in the outer two quarters of this interval at the corresponding endpoints. If the λ measurements are derived from line-of-sight data with noise intensity r, each lumped position observation has a linear variance of

$$\frac{4rs^2(t_f - t)^2}{\tau}$$

since the range is $s(t_f - t)$. The corresponding variance of the terminal position estimate derived from these observations is therefore

$$2\left(\frac{t_f - t}{\tau}\right)^2 \frac{4rs^2(t_f - t)^2}{\tau}$$

if only errors due to velocity uncertainty are considered. Choosing τ to match the variances of the lumped position measurements and the disturbances (of position) from the neglected process noise during the same observation interval gives

$$\frac{1}{3}\bar{q}\tau^3 = \frac{4rs^2(t_f - t)^2}{\tau} \quad \text{or} \quad \tau^2 = s\sqrt{\frac{12r}{\bar{q}}}\,(t_f - t).$$

The variance of $\lambda(t_f)$ then becomes $4s/\tau\,\sqrt{\bar{q}r/3}\,(t_f - t)^3$. But in the absence of process noise, the same variance would be obtained from noisy observations of $\lambda(t_f)$ itself over this time interval if the noise intensity were

$$4s\sqrt{\frac{\bar{q}r}{3}}\,(t_f - t)^3 .$$

In the interception problem, then, it is a reasonable approximation to endow the interceptor with a continuous measurement z of state variable $\tilde{m}$, which is basically a predicted terminal miss distance, such that

$$z(t) = \tilde{m}(t) + v(t), \qquad v \text{ is GWN}\left[0,\ 4s\sqrt{\frac{\bar{q}r}{3}}\ (t_f - t)^3\right]$$

where r is the noise intensity of the line-of-sight measurements and $\bar{q}$ now denotes the intensity of the process noise component *lateral* to the current nominal interceptor velocity in relative coordinates.

With these approximations, the intercept problem reduces to the following with respect to the postulated nominal trajectory:

$$\left.\begin{array}{ll} \dot{\tilde{m}} = (t_f - t)\tilde{u} + w; & \tilde{m}(t_0) = \overline{m}(t_f) \\ \dot{\tilde{\theta}} = \tilde{u}/s; & \tilde{\theta}(t_0) = 0 \end{array}\right\} \text{dynamics}$$

$$J = \frac{1}{2} E\left[am^2(t_f) + \int_{t_0}^{t_f} (\tilde{u}^2 + 2\bar{u}\tilde{u})\, dt\right] \quad \text{criterion to be minimized}$$

$$z = \tilde{m} + v \quad \text{state measurements}$$

where

$$w \approx \text{GWN}[0,\ q(t_f - t)^2 \sin^2(\sigma + \bar{\theta}) + 2\psi\tilde{\theta}]$$

$$\psi = q(t_f - t)^2 \sin(\sigma + \bar{\theta}) \cos(\sigma + \bar{\theta})$$

$$v \approx \text{GWN}\left[0,\ 4s\sqrt{\frac{qr}{3}}\ \sin(\sigma + \bar{\theta})(t_f - t)^3\right]$$

with $\dot{\bar{\theta}} = \bar{u}/s$, and where second-order terms in deviations from the nominal have been neglected. This is a control problem of the form considered above, with state variables $\tilde{m}$ and $\tilde{\theta}$. The only nonzero covariance perturbation parameter is

$$\Psi_{\tilde{m}\tilde{m}\tilde{\theta}} = q(t_f - t)^2 \sin(\sigma + \bar{\theta}) \cos(\sigma + \bar{\theta}).$$

Most of the equation components determining the optimal control in this case are trivial; the rest reduce to the following, where tildes are suppressed in the notation:

$$u = -(t_f - t)(S_{mm}m + \phi_m) - \frac{\phi_\theta}{s} - \bar{u}$$

$$S_{mm} = \frac{a}{1 + \frac{1}{3}a(t_f - t)^3}$$

$$\dot{N}_{mm} = \frac{P_{mm}(N_{mm} - S_{mm})}{2s\sqrt{\frac{qr}{3}}\sin(\sigma + \bar{\theta})(t_f - t)^3}; \quad N_{mm}(t_f) = a$$

$$\dot{P}_{mm} = q(t_f - t)^2 \sin^2(\sigma + \bar{\theta}) - \frac{P_{mm}^2}{4s\sqrt{\frac{qr}{3}}\sin(\sigma + \bar{\theta})(t_f - t)^3}; \quad P_{mm}(t_0) = 0$$

$$\dot{\phi}_m = S_{mm}(t_f - t)^2\phi_m + S_{mm}(t_f - t)\left(\frac{\phi_\theta}{s} - \bar{u}\right); \quad \phi_m(t_f) = 0$$

$$\dot{\phi}_\theta = -N_{mm}(t_f - t)\Psi_{mm\theta}; \quad \phi_\theta(t_f) = 0$$

$$\dot{\hat{m}} = (t_f - t)u + \frac{(P_{mm} + 2D_{mm})(z - \hat{m})}{4s\sqrt{\frac{qr}{3}}\sin(\sigma + \bar{\theta})(t_f - t)^3}; \quad \hat{m}(t_0) = \overline{m}(t_f)$$

$$\dot{\hat{\theta}} = u/s; \quad \hat{\theta}(t_0) = 0$$

$$\dot{D}_{mm} = \hat{\theta}\Psi_{mm\theta} - \frac{D_{mm}P_{mm}}{2s\sqrt{\frac{qr}{3}}\sin(\sigma + \bar{\theta})(t_f - t)^3}; \quad D_{mm}(t_0) = 0.$$

This solution can be incorporated in an iterative algorithm that gives the new nominal control generating the nominal path for the next iteration as

$$\bar{u}_{NEW} = \bar{u}_{OLD} + E(\tilde{u}),$$

where $\tilde{u}$ is generated by the optimal control law of the current iteration. It is helpful for this purpose to refine the values of $\overline{m}(t_f)$ and t_f for the next iteration by generating $\overline{m}(t_f)$ as the distance from the origin to the tangent to the interceptor's old nominal trajectory at time t_f, where this nominal is generated by the exact equations

$$\dot{\bar{x}} = \bar{u}_{OLD}\sin\bar{\theta}_{OLD}; \quad \bar{x}(t_0) = 0$$

$$\dot{\bar{y}} = -\bar{u}_{OLD}\cos\bar{\theta}_{OLD}; \quad \bar{y}(t_0) \text{ specified}$$

and then replacing the value of t_f by the time of closest approach to the origin on this tangent, assuming that the interceptor traverses it at speed s.

Figure 3 shows some numerical results of this iterative procedure for a nominally right-angle interception in absolute coordinates. The deterministic intercept trajectory (zero control) was used as the nominal for the initial iteration. Only mean sample path results are shown here, which would be the pertinent information for nominal trajectory analysis. The mean sample path does what one might expect; it departs from the deterministic intercept path for a more nearly head-on terminal approach, which appears most clearly in relative coordinates. The certainty-equivalent mean sample path is also shown to display the contribution of the dual-control effect here. The corresponding

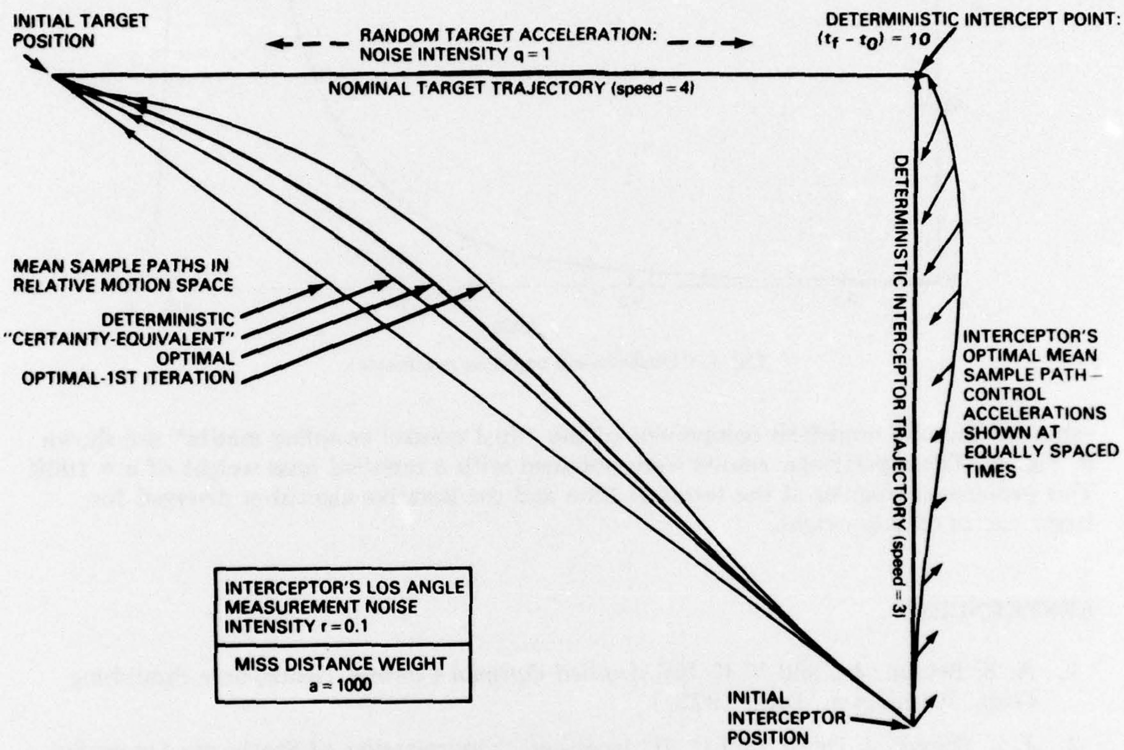

Fig. 3 — Mean sample (nominal) trajectories

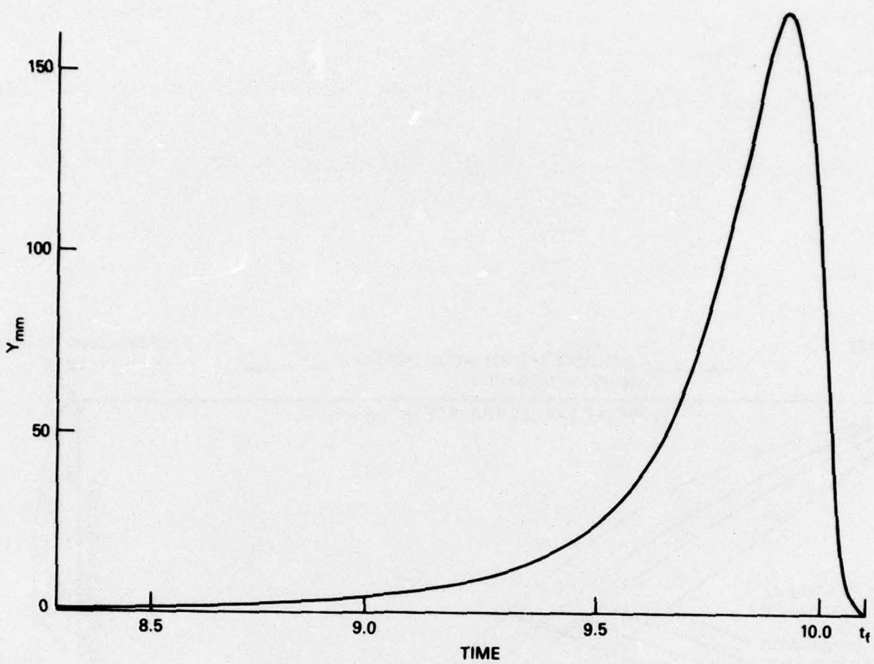

Fig. 4 — Dual-control coupling parameter

values of the one nontrivial component of the "dual control coupling matrix" are shown in Fig. 4. These particular results were obtained with a terminal miss weight of $a = 1000$. This problem is singular at the terminal time and the iterative algorithm diverged for larger values of this weight.

REFERENCES

1. A. E. Bryson, Jr., and Y.-C. Ho, *Applied Optimal Control*, Hemisphere Publishing Corp., Washington, D.C., 1975.
2. J. L. Speyer, J. Deyst, and D. H. Jacobson, "Optimization of Stochastic Linear Systems With Additive Measurement and Process Noise Using Exponential Performance Criteria," *IEEE Trans. Automat. Contr.* AC-19, 358-366 (Aug. 1974).
3. A. A. Feldbaum, "Dual Control Theory I, II, III, IV," *Automation Remote Contr.* 21, 874-880 and 1033-39 1960; 22, 1-12 and 109-121 (1961).
4. R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," *Trans. ASME, J. Basic Eng.* 83, ser. D, 95-108 (Mar. 1961).
5. E. Wong and M. Zakai, "On the Relation Between Ordinary and Stochastic Differential Equations," *Int. J. Eng. Sci.* 3, 213-229 (July 1965).
6. R. L. Stratonovich, "On the Theory of Optimal Control: Sufficient Coordinates," *Automation Remote Contr.* 23, 847-854 (1963).

7. C. Striebel, "Sufficient Statistics in the Optimum Control of Stochastic Systems," *J. Math. Anal. Appl.* 12, 576-592 (1965).

8. S. E. Dreyfus, *Dynamic Programming and the Calculus of Variations*, Academic Press, New York, 1965.

9. C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Tech. J.* 27, 379-423 and 623-656 (1948).

10. Y. Bar-Shalom and E. Tse, "Dual Effect, Certainty Equivalence, and Separation in Stochastic Control," *IEEE Trans. Automat. Contr.* AC-19, 494-500 (Oct. 1974).

11. E. Tse, Y. Bar-Shalom, and L. Meier, III, "Wide-Sense Adaptive Dual Control for Nonlinear Stochastic Systems," *IEEE Trans. Automat. Contr.* AC-18, 98-108 (Apr. 1973).

12. E. Tse and Y. Bar-Shalom, "An Actively Adaptive Control for Linear Systems With Random Parameters via the Dual Control Approach," *IEEE Trans. Automat. Contr.* AC-18, 109-117 (Apr. 1973).

13. D. H. Jacobson, "Optimal Stochastic Linear Systems with Exponential Performance Criteria and Their Relation to Deterministic Differential Games," *IEEE Trans. Automat. Contr.* AC-18, 12-131 (Apr. 1973).