Effects of Strapdown Seeker Scale-Factor Uncertainty on Optimal Guidance

Warren W. Willman*
Naval Weapons Center, China Lake, California

The dominant effects of small seeker scale-factor uncertainty on optimal guidance are determined under certain limitations for a homing missile that uses a strapdown seeker to intercept a randomly maneuvering target. These effects are a change in a feedback gain and the addition of a rapidly fluctuating acceleration component. The latter mainly serves to refine the scale-factor estimate, even though uncertainty in this parameter is not explicitly penalized in the performance criterion for which the guidance law is optimized. This criterion is a standard quadratic one, which essentially maximizes a weighted average of the missile's terminal speed and its probability of intercept, and for which proportional navigation would be the optimal guidance law in the absence of scale-factor uncertainty. The numerical significance of these results is shown for a specific example.

Introduction

STRAPDOWN seekers have numerous advantages for homing missile guidance, but typically have significant scale-factor errors in addition to the noiselike errors present in conventional gimballed seekers.[1] The question of how guidance laws should be modified for the presence of such errors is investigated here by analyzing a stochastic optimal control problem that results from transforming an idealized missile guidance situation of this sort. The resulting problem has the standard linear-quadratic-Gaussian form except for quadratic and bilinear terms in the measurement equation due to the seeker scale-factor error, which is treated as an additional state vector component. For typical homing missile parameters, the noise term in these measurements is small, in a certain relative sense, which gives this derived control problem special properties.

This problem can also be considered an adaptive control problem, in which the scale-factor error is the uncertain parameter instead of a state component. There is an extensive literature on adaptive control,[2,3] but most of it concerns stability or, at most, asymptotic performance in time, whereas rapid efficiency is definitely important here. In the most closely related adaptive control work, Chen and Guo[4] consider a quadratic performance index that is normalized so that only asymptotic performance matters, and Hijab[5] treats parameters with only a finite number of possible values. Also, neither of these cases includes bilinear measurement terms when viewed as a nonlinear control problem by including the parameters as additional state components.

Related work on strapdown seeker guidance[1] uses a guidance law that is optimal for an approximating linear-quadratic-Gaussian control problem and does not include the effects of scale-factor uncertainty in this optimization. Here, the dominant effects are determined by applying an approximate dynamic programming analysis to the more refined optimal control problem formulation mentioned earlier. The dynamic programming state includes the departures of certain Kalman filter covariance matrix components from their local averages, and the analysis uses the fact that these departures are small because the measurement noise is. This analysis is not at a mathematically rigorous level, although the constructions developed here might also be helpful in a more ambitious treatment of that sort. Expressions denoting ordinary differential equations with white noise terms should be understood as the formally corresponding stochastic differential equations in the Ito sense[6,7] if a rigorous interpretation is desired.

One interesting result here is that the scale-factor uncertainty introduces a rapidly fluctuating component in the optimal guidance law, apparently because it is advantageous to dither the missile's attitude to estimate this scale factor better. Other recent work[8,9] has shown that uncertainty in the target range and closing speed can also cause deliberate oscillations to appear in an optimal guidance law for the purpose of enhancing the estimation of those parameters. However, the resemblance to the present case is only slight, because the oscillations are much slower, the missile is limited to angle measurements of the target, and the parameter uncertainty is explicitly penalized in the performance criterion being optimized, which is not done here.

Idealized Homing Missile Guidance Problem

The problem treated here is an idealization of a planar intercept problem in which a homing missile can control its acceleration perpendicular to its flight path and receives noisy angular measurements of the line of sight to a randomly maneuvering target. (This control acceleration is often called the guidance command in this context.) Only the relative motion components perpendicular to the nominal initial line of sight are considered in this idealization and the vehicles' accelerations along this line are ignored. These deleted motion components are typically unimportant to a first approximation if the two vehicles are close to being on a collision course initially and if the missile can measure the target's range and range rate with negligible error. The radial component of the relative velocity becomes a known constant (the nominal closing speed) in the idealized problem, and that of the relative position a predetermined time function. However, the corresponding control laws would be implemented in practice by using the measured closing speed and range in place of these nominal values.

The target's acceleration is approximated in the idealized problem as white noise for simplicity. For a target acceleration with an rms value A and correlation time T, the intensity q (i.e., spectral density or normalized variance) of this white noise approximation is chosen so that the mean-square velocity increment (qT) it causes during the correlation time T

Received Jan. 14, 1987; revision received June 8, 1987.

*Mathematician, Research Department. Member AIAA.

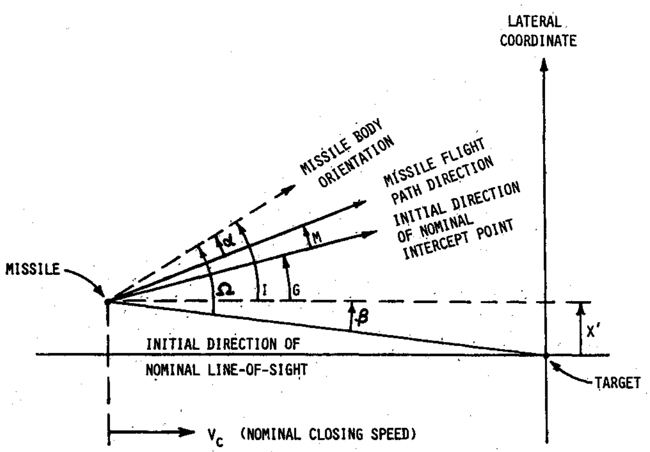

Fig. 1 Idealized intercept geometry.

equals that (A^2T^2) produced by a random constant acceleration of rms value A. This replacement only affects slightly the control behavior in the closely related example of Bryson and Ho,[10] in which the target acceleration is modeled more realistically as a normal random process with variance A^2 and an exponentially decaying autocorrelation function whose time constant is the correlation time T.

The geometry of the idealized problem is summarized in Fig. 1, where the origin is fixed to the target and the horizontal axis is aligned with the nominal initial line of sight. G will be small if the missile is fast compared to the target, and β is assumed to be small enough that the approximation

$$\beta = \frac{x'}{V_c(f-t')} \tag{1}$$

has negligible error, where t' is the current time and f the time at which the missile reaches the vertical axis. The dynamics of the lateral notion are

$$\dot{x}' = v' \tag{2a}$$

$$\dot{v}' = u' + w \tag{2b}$$

where

x' = relative lateral position,
v' = relative lateral velocity,
u' = control acceleration (guidance command),
w = negative of lateral target acceleration (zero-mean Gaussian white noise with intensity $q = A^2T$),

and the common gravitational acceleration has been subtracted throughout. The time variable t' is suppressed in the notation here, and some variables are primed because they will be transformed later. $x'(0)$ and $v'(0)$ are considered independent normal random variables a priori, with the mean of $x'(0)$ being zero.

The missile is assumed to measure $M(0)$, u' and its own inertial attitude I with negligible error, and a strapdown seeker output ζ, which is approximated in the idealized problem as

$$\zeta = (1+\epsilon)\Omega + \gamma$$

where the scale-factor error ϵ is a zero-mean normal random variable and γ is a Gaussian white noise process. Subtracting $I(t')$ from $\zeta(t')$, multiplying the result by $V_c(f-t')$, and using Eq. (1) give the equivalent derived measurement

$$z' = x' + \epsilon V_c \tau \Omega + \bar{\gamma} \tag{3}$$

where

$$\tau = f - t' \qquad \text{(time to go)}$$

and $\bar{\gamma}$ denotes $V_c\tau\gamma$, which is also a Gaussian white noise process. The intensity parameter of $\bar{\gamma}$ is

$$r = \begin{cases} \sigma_1^2 \Delta_1 V_c^2 \tau^2, & \gamma = \text{angular imprecision (e.g., thermal noise)} \\ \sigma_2^2 \Delta_2, & \gamma = \text{radar glint noise} \end{cases}$$

where

σ_1 = rms angle error represented by γ,
σ_2 = rms linear error in apparent target position due to glint,
Δ_i = correlation time of corresponding error; $i = 1, 2$.

The control acceleration u' is proportional to the missile's angle of attack α in the idealized problem, as would be true of body lift, and the term $\epsilon^2 G$ is treated as negligible compared to z'. This allows the derived measurement to be expressed as

$$z' = (1+\sigma_\epsilon\theta)\tilde{x} + (V_c/c)\sigma_\epsilon\tau\theta(u'+h') + \bar{\gamma} \tag{4}$$

for

$$\dot{h} = cu'/V_c; \qquad h(0) = cM(0) \tag{5}$$

where

c = airframe gain (u'/α, a constant),
σ_ϵ^2 = prior variance of ϵ,
$\theta = \epsilon/\sigma_\epsilon$ [prior distribution is normal (0, 1) by construction],
$\tilde{x} = x' + V_c\tau\sigma_\epsilon G\theta$.

A priori, θ and the white noises $\bar{\gamma}$ and w are taken as statistically independent of each other, and of $x'(0)$ and $v'(0)$. For $\tilde{v}$ defined as

$$\tilde{v} = v' - V_c\sigma_\epsilon G\theta$$

$\tilde{x}$ and $\tilde{v}$ have the same dynamics [Eq. (2)] as x' and $\dot{v}'$. Any constant biases in ζ or I have no effect here because their difference can be added to $\sigma_\epsilon G\theta$ in the construction of $\tilde{x}$ and $\tilde{v}$ to give the same result.

The objective is to find a control law for u' that minimizes the performance criterion

$$J' = \frac{1}{2}E[D\tilde{x}^2(f) + \int_0^f u'^2\,dt'] \tag{6}$$

where D is a positive parameter and E denotes prior expectation. As usual, a control law is defined as a decision rule that, for each t' in $[0, f]$, specifies the current control $u'(t')$ as a function of the current measurement history $\{[z'(t^*), t^*]: 0 \le t^* < t'\}$. For typical homing missile aerodynamics, the integrand in J' is approximately proportional to the induced drag force, and the first term [which equals $Dx'^2(f)$ by construction] is approximately proportional to the probability of unsuccessful intercept except in the uninteresting case where this probability is large. Thus, minimizing this criterion corresponds to the reasonable objective of maximizing a weighted average of the expected missile speed in the vicinity of the target and the probability of successful intercept. Furthermore, it happens that the optimal control law is typically very insensitive to the weights used in this average. Limitations on the missile's control authority have been assumed to be an insignificant restriction on its ability to achieve this ideal optimum.

To summarize, the idealized optimal control problem formulated here has the dynamics of Eqs. (2) and (5), with $\tilde{x}$ and $\tilde{v}$ in place of x' and v', the measurements of Eq. (4); and the criterion of Eq. (6). Of course, the intercept problem is three-dimensional for actual homing missiles. Except for bank-to-turn missiles, however, that have a preferred orientation, this can be reduced under reasonable assumptions to two independent planar problems of the type considered here, one for each of the two orthogonal control components perpendicular to the initial nominal line of sight.

Transformation to a Mathematically Convenient Form

It is convenient for analysis to use the logarithmic time variable

$$t = \ln(f/\tau) \tag{7}$$

Its final value is infinite because τ is zero then, but this anomaly can be removed by terminating the problem slightly before the time of closest approach f, say at $\tau = \delta$, and using the quantity $\tilde{x} + \tilde{v}\tau$ instead of $\tilde{x}$ in the criterion (6). This has a negligible effect on the outcome, basically because $\tilde{x} + \tilde{v}\tau$ is what the distance of closest approach would be at any given time if neither vehicle accelerated thereafter, and these accelerations would have progressively less effect on this quantity as t' approaches f. The resulting final value of t is

$$t_f = \ln(f/\delta) \tag{8}$$

The state, measurement, and control variables in the idealized problem are also transformed to

$$x = \tilde{x}/\sqrt{q\tau^3} \tag{9a}$$

$$v = (\tilde{x} + \tilde{v}\tau)/\sqrt{q\tau^3} \tag{9b}$$

$$h = h'(\sqrt{\tau/q}) \tag{9c}$$

$$z = z'/\sqrt{q\tau^3} \tag{9d}$$

$$u = u'(\sqrt{\tau/q}) \tag{9e}$$

and the performance criterion is replaced by the equivalent one

$$J = J'/q$$

It then follows from Eqs. (2-6) that this problem is converted to one with dynamics of the form

$$\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} a & 1 \\ 0 & b \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} (u + w_1) \tag{10a}$$

$$\dot{h} = (a - m)h + Fu \tag{10b}$$

state measurements of the form

$$z = x + \phi\theta(Fx + h + u) + w_2 \tag{11}$$

and performance criterion of the form

$$J = \frac{1}{2} E\,[s_f v^2(t_f) + \int_0^{t_f} u^2(t)\,dt] \tag{12}$$

where w_1 and w_2 are statistically independent zero-mean Gaussian white noise processes with respective intensities 1 and $1/K^4$, where

$$\phi = \phi_0 e^{mt}$$

$$F = F_0 e^{-mt}$$

$$K = K_0 e^{nt}$$

and where s_f, a, b, m, n, ϕ_0, F_0, and K_0 are known constants with $s_f \geq 0$, $\phi_0 \geq 0$, and $K_0 > 0$. In this case,

$$s_f = D\delta^3 \tag{13a}$$

$$\phi_0 = (V_c\sigma_\epsilon)/cf \tag{13b}$$

$$F_0 = cf/V_c \tag{13c}$$

$$K_0 = (q/r)^{1/4} f \tag{13d}$$

$$a = 1/2 \tag{13e}$$

$$b = 3/2 \tag{13f}$$

$$m = 1 \tag{13g}$$

and

$$n = \begin{cases} -1 \text{ for seeker glint noise only} \\ -1/2 \text{ for seeker angle noise only} \end{cases}$$

A priori, $h(0)$ is known to the controller, and $x(0)$, $v(0)$, and θ are random variables that have a joint probability distribution of the form

$$\begin{bmatrix} x(0) \\ v(0) \\ \theta \end{bmatrix} \sim \text{normal} \left(\begin{bmatrix} 0 \\ \bar{v}_0 \\ 0 \end{bmatrix}, \begin{bmatrix} p_{xx} & p_{xv} & p_{x\theta} \\ p_{xv} & p_{vv} & 0 \\ p_{x\theta} & 0 & 1 \end{bmatrix} \right) \tag{14}$$

and are statistically independent of w_1 and w_2. Also, $\dot{\theta} = 0$. If s_f and t_f are computed from δ by Eqs. (8) and (13), the results for this transformed problem approach the appropriate limits for any finite t (i.e., logarithmic inverse time to go) as $\delta \to 0$. However, some further constraints will be imposed shortly that will limit how small δ can be taken.

We will treat a, b, m, and n as general constants, since this adds no difficulty and might lead to further insights by showing how the problem's structural features are related to its solution. Finding the optimal control law is very difficult, however, so we only consider the problem of finding an approximation thereof that is asymptotically accurate to order ϕ^2 for small ϕ under the condition that

$$\phi \ll 1/K \ll 1 \qquad \text{for } t \text{ in} \quad [0, t_f] \tag{15}$$

and the further conditions that $\bar{v}_0$ and $h(0)$ are of order $1/\sqrt{K}$, $p_{x\theta}$ is of order $\phi/\sqrt{K}$, and s_f, a, b, m, and n are of order unity. What is meant by such an approximation is that the control law always generates a control value u that is the same to order ϕ^2 as that generated by an optimal control law, except perhaps for a set of measurement histories of negligible probability.

Also, the treatment of this problem is limited here to finding the control law associated with a cost-to-go function that has the formal appearance of satisfying the Bellman equation corresponding to Eqs. (10-12) to order ϕ^2. This control law would be the desired asymptotic approximation if the equations involved in the analysis are well-posed and the formally higher-order terms are indeed so in some appropriate sense. A rigorous verification of these conditions, such as that of Balikrishnan[11] in a deterministic context, is beyond the scope of this paper, however. In this sense, the control law obtained here is only a plausible candidate for the approximation being sought. This plausibility is enhanced, though, by the fact that the exact optimal control law is well known and rigorously justified for $\phi = 0$ (a standard linear-quadratic-Gaussian case) and the approximation derived here for small ϕ converges to this control law as $\phi \to 0$.

State Estimation

A key step in deriving useful results is to consider the variable

$$y = x + \phi\theta(Fx + h + u)$$

Since $\dot{\phi} = n\phi$ and $\dot{F} = -mF$,

$$\begin{bmatrix} \dot{y} \\ \dot{v} \\ \dot{\theta} \end{bmatrix} = \left[\begin{array}{c|c|c} a & 1 + F\phi\theta & (\dot{u} + Fu + mu - au)\phi \\ \hline 0 & b & 0 \\ \hline 0 & 0 & 0 \end{array}\right] \begin{bmatrix} y \\ v \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} (u + w_1) \tag{16}$$

and

$$z = y + w_2 \tag{17}$$

Current and past values of u, and therefore of $\dot{u}$, are known to the controller, so the current conditional probability distribution of (y, v, θ), given current and past values of z, is well defined as long as u remains differentiable. Controls with Wiener-process components will be used later; they are not actually differentiable, so these should really be interpreted as smooth approximations of the Wiener processes used formally. Such approximations would be unavoidable in practice anyway and presumably could be made in a way that amounts to local averaging over a time interval so short that the effects would be negligible to the order of accuracy retained in the analysis that follows. Even so, if white noise is then substituted as usual in $\dot{u}$ for the derivative of such a Wiener-process approximation, a deterministic correction term must be added to the resulting equation for $\dot{p}_1$ (see the following) to make it completely accurate when considered as an Ito equation in the standard way.[12] This correction term is also negligible to the accuracy retained here, however, so this distinction is not important either.

Since the prior distribution of the initial state $[y(0), v(0), \theta]$ is normal, the means and covariance matrices of the aforementioned posterior distributions are given approximately by the standard extended Kalman filter for the dynamics and measurements of Eqs. (16) and (17). These filter equations are straightforward and are not displayed here. The resulting (time-varying) conditional mean and covariance matrix are denoted componentwise as

$$\begin{bmatrix} \hat{y} \\ \hat{v} \\ \hat{\theta} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} p_1 & p_2 & p_4 \\ p_2 & p_3 & p_5 \\ p_4 & p_5 & L \end{bmatrix}$$

respectively. Approximations of higher asymptotic order in ϕ can be constructed with the use of Edgeworth expansions,[13] but it can be shown by normalizing the state components by their conditional standard deviations[14] that the measurement noise in this case is small enough to make the effects of such refinements negligible to the order of accuracy retained here. It is also convenient to define

$$\xi = K^2(z - \hat{y}) \tag{18}$$

which is the normalized innovation process for this filter. As such, ξ can be treated as a unit-intensity zero-mean Gaussian white noise process in determining the statistical behavior of $\hat{y}$, $\hat{v}$, and $\hat{\theta}$.[6]

It happens that L varies more slowly than the other covariance matrix components. Another key step here, which takes advantage of this, is to define the nominal values

$$\bar{P}_1 = \frac{\sqrt{2}}{K^3} + \frac{a + b + 2n}{K^4} \tag{19a}$$

$$\bar{P}_2 = \frac{1}{K^2} + \frac{2b + n}{\sqrt{2}\,K^3} \tag{19b}$$

$$\bar{P}_3 = \frac{\sqrt{2}}{K} + \frac{2b + 3n - 2m}{K^2} \tag{19c}$$

and the normalized variables

$$d_1 = (K/\phi)^2 (p_1 - \bar{p}_1) \tag{20a}$$

$$d_2 = (K/\phi^2)(p_2 - \bar{p}_2) \tag{20b}$$

$$d_3 = (p_3 - \bar{p}_3)/\phi^2 \tag{20c}$$

$$d_4 = (K/\phi)p_4 \tag{20d}$$

$$d_5 = p_5/\phi \tag{20e}$$

It follows from these definitions and the filter equations just mentioned that, to the accuracy being retained,

$$\dot{\hat{v}} = b\hat{v} + u + \left(1 + \frac{2b + n}{\sqrt{2}K} + K\phi^2 d_2\right)\xi \tag{21}$$

and

$$\dot{L} = -\phi^2 K^2 d_4^2 \tag{22}$$

It is also straightforward to obtain expressions for $\dot{d}_1, \dot{d}_2, \ldots, \dot{d}_5$ by differentiating Eqs. (20) and substituting from the filter equations and from Eqs. (19) and their derivatives. These expressions contain the d_i, L, u, and $\dot{u}$, but not the p_i or the $\bar{p}_i$. They are rather lengthy, however, and not worth displaying in detail. Their only subtlety concerns terms in $\dot{d}_1$, $\dot{d}_2$, and $\dot{d}_3$ that have ϕ in the denominator. These terms can be disregarded because their orders of magnitude can be reduced to any negative power in K by using more refined expressions than Eqs. (19) for $\bar{p}_1$, $\bar{p}_2$, and $\bar{p}_3$, with appropriate terms of successively higher negative powers in K. Doing so would add further terms to the $\dot{d}_i$, but these extra terms would all be of high enough negative powers in K to be negligible in determining the optimal control to the order of accuracy being considered here.

Approximate Estimator Behavior for a Restricted Class of Control Laws

If $\phi = 0$, it follows from standard results that the optimal control law is of the form $u = \bar{g}\hat{v}$, where $\bar{g}(t)$ is a certain deterministic time function such that $\bar{g}$ and $\dot{\bar{g}}$ are both of order unity. Since we are only concerned with small ϕ here, we consider control laws of the form

$$u = g\hat{v} + \lambda \tag{23}$$

where g is a deterministic time function, to be chosen for convenience later, such that g and $\dot{g}$ are of order unity, and where $|\lambda| \ll 1$, except perhaps for a negligibly improbable set of realizations. From Eqs. (21) and (23),

$$\dot{u} = (\dot{g} + gb + g^2)\hat{v} + g\lambda + g\left(1 + K\phi^2 d_2 + \frac{2b + m}{\sqrt{2}K}\right)\xi \tag{24}$$

for such a control law if $\dot{\lambda}$ is also negligible (which it will be for the optimal control law approximation next derived). For a control law with all these properties, Eqs. (23) and (24) can be substituted for u and $\dot{u}$ in Eqs. (21) and (22) and in the equations mentioned in the preceding paragraph for $\dot{d}_1, \dots,$ $\dot{d}_5$, and $\dot{\hat{\theta}}$ to obtain a system of differential equations driven by the small perturbation control λ and the white noise ξ.

If this equation system is considered over a time interval of order $1/K$, then g, ϕ, F, and K are approximately constant and, from Eqs. (21) and (22), so are $\hat{v}$ and L for $\phi \ll 1/K$ if d_2 and d_4^2 are of order K. Conditioned on these current values of $\hat{v}$ and L, and $\dot{d}_i$ can be analyzed with these approximations as an independent subsystem driven by λ and ξ to show that the mean and covariance matrix of the d_i converge with time constants of order $1/K$ to "steady-state" values such that

$$E\left(\begin{bmatrix} d_4 \\ d_5 \end{bmatrix}\right) = -[\dot{g} + F + g(g + F + m + b - a)]L\hat{v}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (25)$$

and

$$\mathrm{cov}\left(\begin{bmatrix} d_4 \\ d_5 \end{bmatrix}\right) = \frac{Kg^2L^2}{2\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (26)$$

to the orders of magnitude of the terms shown, and such that $E(d_1^2)$, $E(d_2^2)$, and $E(d_3^2)$ are of order K^2. By the Chebychev inequality, this also means that d_1, d_2, and d_3 are of order K, and d_4 and d_5 are of order $\sqrt{K}$ (except perhaps for a negligibly improbable set of realizations) since the marginal distribution of L and $\hat{v}$ is such that these last are of order unity. Also, the (matrix) correlation function for the d_i decays with time constants of order $1/K$. These approximate results are obtained by assuming the above orders of magnitude to delete the less significant terms from the $\dot{d}_i$ equations and, with g, ϕ, F, K, $\hat{v}$, and L treated as known constants in these simplified equations, using standard methods[6] to derive differential equations for the propagation of the first and second moments of the d_i. These last equations are consistent with the results stated here and verify them under the assumption that the full equation system for $\hat{v}$, $\hat{\theta}$, L, and the d_i is well-posed in some appropriate sense.

For a specified $g(t)$, it will also be convenient to approximate L by the predetermined time function $\bar{L}(t)$ satisfying

$$\dot{\bar{L}} = -(K^3\phi^2g^2\bar{L}^2)/(2\sqrt{2}); \qquad \bar{L}(0) = 1 \quad (27)$$

which results from local-averaging in Eq. (22). If $L - \bar{L}$ is defined as R, the resulting equation for $\dot{R}$ shows that $|R| \ll 1$ for $d_4(0)$ and $d_5(0)$ restricted to the assumed orders of magnitude.

Control Optimization

Since $g(t)$ in Eq. (23) is considered specified, the problem here reduces to that of finding an optimal control law for the perturbation control λ, to which we seek only an asymptotic approximation. A convenient choice of g will be used for this purpose, but one for which g and $\dot{g}$ are of order unity.

An optimal expected cost-to-go function can be defined consistently in terms of time and the conditional distribution of y, v, and θ,[15] and the principle of optimality of dynamic programming can then be applied in the usual way[16] to derive a Bellman equation for this function, the solution of which specifies the optimal control law for λ. Since the conditional distribution here is approximately normal, it happens that an order ϕ^2 approximation of such a solution can be found that has the form $H(t, \hat{v}, d_1, \dots, d_5)$, i.e., which depends only on certain first and second moments of this distribution. The derivation of the Bellman equation for this restricted class of cost functions requires the conditional expected values of the increments $\Delta\hat{v}$, ΔL, Δd_i, $i = 1, \dots, 5$ and of quadratic products thereof, over an infinitesimal time increment Δt, given the data up to the beginning of this time increment. Since this conditioning is equivalent to conditioning on the conditional distribution of y, v, and θ at that time, these expectations can be evaluated from Eqs. (21) and (22) and the equations for the $\dot{d}_i$, with $g\hat{v} + \lambda$ substituted for u. Here, $\Delta\hat{v}$ is taken as $\dot{\hat{v}}\Delta t$, etc.

In so doing, we first consider the case of $\phi \ll 1/K^2$ and retain only terms up to order ϕ^2 and $\phi^2\lambda^j$, $j = 1, 2, \dots$ in the resulting Bellman equation. Also, we further restrict consideration to possible solutions of the form

$$H = (1/2)\,[s(\hat{v}^2 + \phi^2 d_3) + B] + (1/K)\,(Q_1 d_1 + Q_2 d_2 + Q_3 d_3 + Q_4\hat{v}d_4 + Q_5\hat{v}d_5) \quad (28)$$

where s, B, and the Q_i are functions of t only, with the Q_i of order ϕ^2, and for which $\dot{\lambda}$, the time derivative of the associated optimal perturbation control, contributes only terms small compared to ϕ^2 to the Bellman equation. These restrictions and use of the fact that L is approximately the control-independent deterministic time function $\bar{L}$ result in a Bellman equation of the form

$$-\frac{\partial H}{\partial t} = \min_{\lambda}\Gamma(t, \hat{v}, d_1, d_2, d_3, d_4, d_5, \lambda) \quad (29a)$$

$$H(t_f, \dots) \equiv (1/2)\, s_f[\hat{v}^2 + \phi^2(t_f)d_3 + \bar{p}_3(t_f)] \quad (29b)$$

where Γ is a rather lengthy expression and $\bar{p}_3$ is given by Eq. (19) or a suitable refinement thereof. The control-dependent terms in Γ are

$$(1/2)\,\lambda^2 + [s + g + (g + F + m - a)Q_4\bar{L}]\lambda\hat{v} + (g + F + m - a)(2Q_1d_4 + Q_2d_5)\lambda \quad (30)$$

If $g(t)$ is chosen to make the coefficient of $\lambda\hat{v}$ in expression (30) identically zero, the minimizing λ in the Bellman equation (29), obtained by equating the λ derivative of (30) to zero, is

$$\lambda = (a - g - m - F)(2Q_1d_4 + Q_2d_5) \quad (31)$$

to order ϕ^2, and the minimized contribution of the control-dependent terms to $-\partial H/\partial t$ is negligible to order ϕ^2. Collecting the remaining terms shows that the function H of Eq. (28) satisfies the Bellman equation to order ϕ^2 for this choice of $g(t)$ if

$$-\dot{s} = g^2 + 2(b + g)s + 2[\dot{g} + F + g \times (g + F + m + b - a)]\bar{L}Q_4; \quad s(t_f) = s_f$$

$$-\dot{B} = s + 2\bar{L}Q_4; \qquad B(t_f) = s_f\bar{p}_3(t_f)$$

$$-\dot{Q}_1 = -K(2\sqrt{2}Q_1 + Q_2); \qquad Q_1(t_f) = 0$$

$$-\dot{Q}_2 = K[2Q_1 - \sqrt{2}Q_2 - 2Q_3 - \phi^2 s(2b + n)/\sqrt{2}]; \qquad Q_2(t_f) = 0$$

$$-\dot{Q}_3 = K[Q_2 - (1/2)\phi^2 g(2s + g)]; \qquad Q_3(t_f) = 0$$

$$-\dot{Q}_4 = K\{2[\dot{g} + F + g(g + F + m + b - a)] \times Q_1 - \sqrt{2}Q_4 - Q_5\}; \qquad Q_4(t_f) = 0$$

$$-\dot{Q}_5 = K\{[\dot{g} + F + g(g + F + m + b - a)]Q_2 + Q_4\}; \qquad Q_5(t_f) = 0$$

As a consequence of the dynamic programming procedure,[16] the optimal perturbation control is given to order ϕ^2 by the corresponding λ of Eq. (31).

The Q_i converge rapidly in reverse time (with time constant of order $1/K$) to fast-time steady-state values, which to order ϕ^2 are

$$Q_3 = -\frac{\phi^2}{4\sqrt{2}}[3g^2 + 2s(3g + 2b + n)] \tag{32a}$$

$$Q_2 = -2\sqrt{2}Q_1 = (1/2)\phi^2 g(2s + g) \tag{32b}$$

$$Q_4 = -\sqrt{2}\,Q_5 = -(1/2)\phi^2 g(2s + g) \times [\dot{g} + F + g(g + F + m + b - a)] \tag{32c}$$

Equating the coefficient of $\lambda\hat{v}$ in expression (30) to zero and using approximations (32) imply that

$$g = -s + (1/2)s^2(s + a - m - F) \times [F + s(a + b - m - F)]\phi^2\bar{L} \tag{33}$$

to order ϕ^2, except perhaps within order $1/K$ of the terminal time. Thus, except for these small boundary effects, which will no longer be considered,

$$-\dot{s} = 2bs - s^2 + s^2[F + s(a + b - m - F)]^2\phi^2\bar{L} \tag{34}$$

to order ϕ^2.

As required for consistency, the $g(t)$ chosen here is such that g and $\dot{g}$ are of order unity and the Q_i are of order ϕ^2. It also follows from differentiating Eq. (31), substituting for the derivatives in the resulting expression, and from previously established orders of magnitude for the quantities involved, that $\dot{\lambda}$ is small enough that it would contribute only negligibly to the Bellman equation, as was assumed.

These results can be extended to all $\phi \ll 1/K$. If ϕ is not also small compared to $1/K^2$, extra quadratic terms in the d_i appear in the approximate Bellman equation (29). However, if d_2^2, d_1d_2, d_2d_3, d_5^2, and d_4d_5 terms are added to the trial solution H of Eq. (28) in the derivation, the rapid mixing of the d_i keeps the coefficients of these extra terms small enough so that the only effect at the level of significance retained here is to change the value of the $B(t)$ term in H. This adds a deterministic time function to the expected cost to go, but does not affect the optimal control.

Implementation

Even with the preceding simplifications, the optimal control law is given (to order ϕ^2) in terms of the solutions of the coupled equations for $\dot{s}$ and $\dot{\bar{L}}$, which form a two-point boundary value problem. This coupling is weak enough, hovever, that the control law can be determined to the same order of accuracy without this difficulty by using the approximation $\bar{s}$ of s that results from deleting the order-ϕ^2 terms in Eq. (34). It then follows from Eqs. (23), (27), and (31–34) that the optimal control law can be obtained to order ϕ^2 by solving

$$-\dot{\bar{s}} = 2b\bar{s} - \bar{s}^2; \qquad \bar{s}(t_f) = s_f \tag{35}$$

$$\dot{\bar{L}} = -\frac{\phi^2K^3}{2\sqrt{2}}\bar{s}^2\bar{L}^2; \quad \bar{L}(0) = 1 \tag{36}$$

and Eq. (34) in turn, and then computing the control as

$$u = -\{s - (1/2)s^2(s + a - m - F) \times [F + s(a + b - m - F)]\phi^2\bar{L}\}\hat{v} - (1/2)\phi^2 s^2(s + a - m - F)(d_5 - d_4/\sqrt{2}) \tag{37}$$

Implementing this control law would also require computing $\hat{v}$, d_4, and d_5 in real time from the incoming z measurements. This could be done with the extended Kalman filter for the original observation system of Eqs. (10), (11), and (14), which provides $\hat{v}$ directly as a component of the conditional mean vector $(\hat{x}, \hat{v}, \hat{\theta})$ it generates. This filter uses the current control u as an input, which is available from Eq. (37), but not $\dot{u}$. Also,

$$\xi = K^2(z - \hat{y}) = K^2\{z - [\hat{x} + \phi\hat{\theta}(F\hat{x} + u + h) + F\phi\,\mathrm{cov}(x, \theta)]\} \tag{38}$$

by Eq. (18), the definition of y, and the properties of the expectation operator. As explained in the next section, the effects of the d_4 and d_5 control terms on the dynamic system are mostly averaged out because they fluctuate so rapidly. This allows the effective accuracy of Eq. (37) to be retained with the differential equations generating d_4 and d_5 truncated to

$$\dot{d}_4 = K[d_5 - \sqrt{2}d_4 + s(3s - 5b)\bar{L}\hat{v} - s\bar{L}\xi] \tag{39a}$$

$$d_4(0) = K_0[h(0) - s(0)\bar{v}_0 + P_{x\theta}/\phi_0] \tag{39b}$$

$$\dot{d}_5 = -Kd_4; \qquad d_5(0) = 0 \tag{39c}$$

Equations (33) and (34) have been used here to substitute for g and $\dot{g}$, and the initial conditions follow from Eqs. (20) and (37). As a result, the overall control law here could be implemented by using Eqs. (37–39), with the standard extended Kalman filter for the (x, v, θ) system of Eqs. (10), (11), and (14) to generate $\hat{x}$, $\hat{v}$, $\hat{\theta}$ and $\mathrm{cov}(x, \theta)$ from the z measurements.

Finally, with the definitions

$$\psi = d_5 + [F + (a + b - m - F)s]\bar{L}\hat{v} \tag{40}$$

$$\eta = (1/2)\phi^2 s^2(F + m - s - a)(\psi - d_4/\sqrt{2}) \tag{41}$$

and

$$N = s - s^2(s + a - m - F)[F + s(a + b - m - F)]\phi^2\bar{L} \tag{42}$$

the control law of Eq. (37) can be expressed as

$$u(t) = -N(t)\hat{v}(t) + \eta(t) \tag{43}$$

Also, since K, s, $\hat{v}$, $\bar{L}$, and F vary slowly compared to d_4 and d_5, it follows from Eqs. (38) and (39) that

$$\dot{d}_4 = K(\psi - \sqrt{2}d_4 - s\bar{L}\xi) \tag{44a}$$

$$d_4(0) = K_0[h(0) - s(0)\bar{v}_0 - p_{x\theta}/\phi_0] \tag{44b}$$

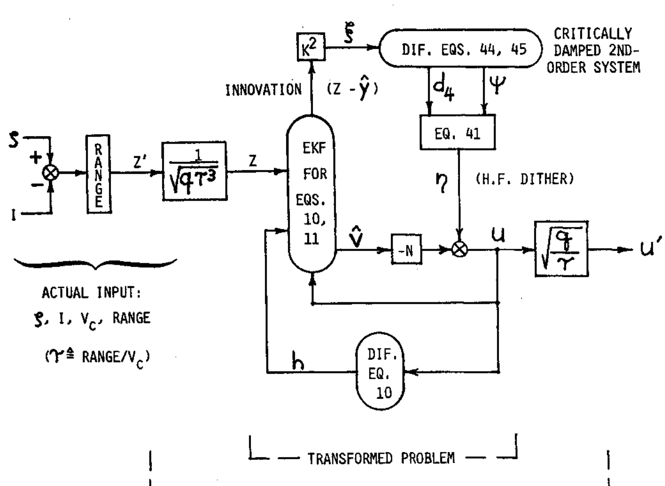

Fig. 2 Control law implementation.

and

$$\dot{\psi} = -Kd_4; \qquad \psi(0) = [F_0 + (a + b - m - F_0)s(0)]\bar{v}_0 \tag{45}$$

to the order of accuracy shown.This last implementation is shown schematically in Fig. 2. Differentiating Eq. (45) and substituting from Eqs. (44) and (45) gives

$$\ddot{\psi} + (\sqrt{2}K + n)\dot{\psi} + K^2\psi = K^2 s\bar{L}\xi$$

Since n is negligible compared to K, Eqs. (44) and (45) generating d_4 and ψ constitute a high-frequency (K) critically damped second-order system driven by the extended Kalman filter's innovation process.

Behavior Characteristics of the Optimally Controlled System

Equations (41), (44), and (45) define η as a linear output of a stable linear system driven by the white noise ξ. From standard results for such cases[10] and the fact that K, s, $\hat{v}$, and $\bar{L}$ vary relatively slowly, η is a normal random process with correlation time of order $1/K$, zero mean to order ϕ^2, and standard deviation

$$\sigma_\eta = (1/4)\,(\sqrt{3K}/\sqrt{2})s^3|s + a - m - F|\phi^2\bar{L} \tag{46}$$

to order $\sqrt{K}\phi^2$, except within order $1/K$ of the initial time. Also, the statistical dependence of η and $\hat{v}$ decays to zero after the initial time with time constant of order $1/K$.

The first control term ($-N\hat{v}$) of Eq. (43) fluctuates more slowly than η, with correlation time of order unity. Its prior mean, and also the prior mean $\bar{v}(t)$ of $\hat{v}(t)$, can be found to order ϕ^2 by substituting from Eq. (43) for u in Eq. (21), taking the conditional expectation given $\hat{v}$, and then the prior expectation of the result. This gives $\bar{v}(t)$ to this accuracy as the solution of

$$\dot{\bar{v}} = (b - N)\bar{v} + \bar{\eta}; \qquad \bar{v}(0) = \bar{v}_0 \tag{47}$$

where $\bar{\eta}(t)$ is the prior mean of $\eta(t)$. $\bar{\eta}(t)$ can be determined from Eqs. (41), (44), and (45) by the standard methods mentioned in the preceding paragraph. Taking prior expectations in Eq. (43) gives the prior mean $\bar{u}$ of the total control u as

$$\bar{u} = -N\bar{v} + \bar{\eta} \tag{48}$$

and that of the first term as $-N\bar{v}$. It is a standard result for linear transformations of random variables that the variance of this first control term is $N^2(t)\sigma_v^2(t)$, where σ_v^2 denotes the prior variance of $\hat{v}$ when control law (43) is used. This variance can be found to order unity by letting ϕ be zero and using the standard equations for the resulting linear-quadratic-Gaussian optimal control problem.[10] Determining σ_v^2 to order ϕ^2 is more difficult and is not attempted here.

If ϕ is zero, it is a standard result that $u = \bar{s}\hat{v}$. Thus, it is apparent from Eqs. (42) and (43) that changing ϕ to a nonzero value has two asymptotic effects on the optimal control law to order ϕ^2, aside from effects on the state estimation procedure. One is to change the feedback gain on $\hat{v}$ from $\bar{s}$ to N, a change of order ϕ^2. The other is to add the zero-mean, rapidly fluctuating term η to the control. The instantaneous magnitude of η is usually of order $\sqrt{K}\,\phi^2$, but it is shown next that its effects of significance for the performance criterion of Eq. (12) are also only of order ϕ^2, so it is consistent with this level of accuracy to determine the instantaneous values of η only to order $\sqrt{K}\,\phi^2$ as long as their mean is zero to order ϕ^2.

Since the correlation time of v is of order unity, the order-of-magnitude effect of the control component η on v is that of the integral of η over a unit time interval. This integral is a zero-mean random variable because η is zero-mean. Since the correlation time of η is of order $1/K$, the variance of this integral is, to an order-of-magnitude approximation, that of a sum of K independent random increments, each with the variance of η/K, which variance is of order $(\phi^4 K)/K^2$. The variance of this sum, which is the sum of these variances, is of order ϕ^4. Thus, the contribution of the rapidly fluctuating control component η to the state variable v appearing in the criterion has standard deviation of order ϕ^2, despite the fact that η itself is typically of order $\sqrt{K}\,\phi^2$. A similar argument applies to its effect on the integrated square of control.

The presence of this instantaneously large control component whose effects on v are largely self-cancelling might seem strange in a control law that is optimal for a performance criterion that penalizes control effort and otherwise depends only on the final value of v. However, it might be explained as a high-frequency dithering whose basic function is to accelerate the estimation of the parameter θ. As such, it would be an example of the probing phenomenon identified by Feldbaum,[17] which is control effort expended to reduce uncertainty in the state variable (which includes θ here) so that performance can be improved in the long run. Knowing θ more precisely is important because it would be optimal to use a control proportional to v if v were known exactly, and the uncertainty in v is sensitive to uncertainty in θ whenever the control is changing. This sensitivity arises from the fact that both $\dot{u}$ and v cause the same kind of change in the measurement z, which is apparent from Eq. (11) and the direct variation of $\dot{x}$ with v. If θ is not known exactly, the portion due to $\dot{u}$ cannot be eliminated unambiguously in estimating v from changes in z.

Furthermore, η reduces more efficiently the uncertainty in θ than the more slowly varying control component $-N\hat{v}$ in terms of the penalty incurred for control effort. For a general control history $u(t)$, the truncated equations corresponding to Eqs. (39) reduce to

$$\ddot{d}_4 + \sqrt{2}\dot{d}_4 + d_4 = \frac{\mathrm{d}}{\mathrm{d}t^*}\,[\phi(K\dot{u} + Fu + mu - au)]$$

after converting to the fast time variable $t^* = K(1 - e^{-nt})/n$ and differentiating. This means that, to a first approximation, d_4 is the output of a critically damped second-order system driven by a quantity that is zero for any constant control, and, from Eq. (22), that only *changes* in the control are effective in

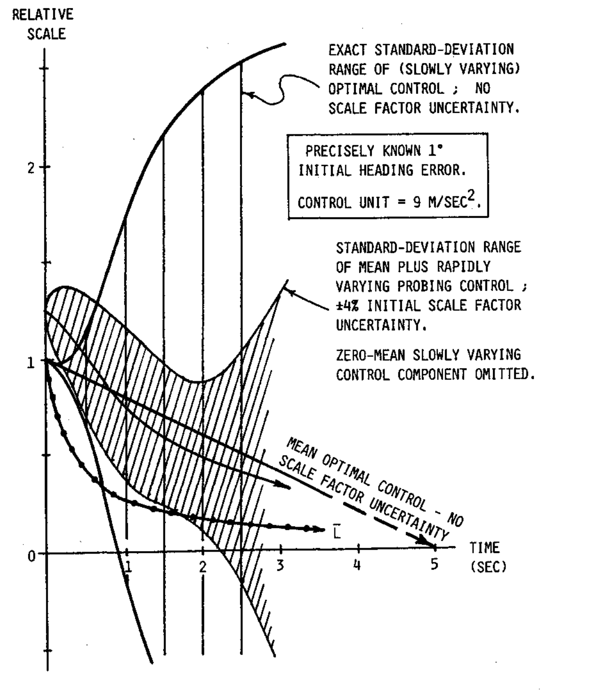

Fig. 3 Approximate statistical behavior of optimal control law: Example.

reducing L, the conditional variance of θ. Also, for a given amplitude, higher-frequency control variations reduce L faster, as long as they do not exceed the Kalman filter's response speed. In fact, a rather elaborate analysis of the frequency content of η and $\hat{v}$ under the optimal control law shows that, compared to the slow control component, the average contribution of η to $\dot{L}$ is enhanced by a factor of about $\sqrt{K}$ over its rms magnitude.

Numerical Example

Some results of applying this theory are shown in Fig. 3 for a particular example. They are expressed in terms of the original untransformed problem. In this case, the missile is radar-guided from a maximum range of only 6 km, so the noise corrupting its measurements of the target line-of-sight angle is treated as predominantly glint ($n = -1$). The other parameters are

$$f = 5 \text{ s}$$

$$V_c = 1200 \text{ m/s}$$

$$G = 0$$

$$c = 150 \text{ m/s}^2 \text{ rad}$$

$$\left.\begin{array}{l} A = 2 \text{ g} \\ T = 2 \text{ s} \end{array}\right\} \quad (q = 800 \text{ m}^2/\text{s}^3)$$

$$\left.\begin{array}{l} \sigma_2 = 10 \text{ m} \\ \Delta_2 = 1/2 \text{ s} \\ \sigma_\epsilon = 0.04 \end{array}\right\} \quad (r = 50 \text{ m}^2 \text{ s})$$

The initial relative lateral position and velocity are known precisely, with $x'(0) = 0$ by definition and $v'(0) = 15$ m/s. For an approaching target traveling at 300 m/s, this $v'(0)$ corresponds to a 1 deg initial aiming error M. As long as D is large enough to allow reasonably large rms control values to occur,[10] these results do not depend significantly on D before the last second of flight. The validity conditions (15) become $\sigma_\epsilon \ll 1/16$ and $\tau \gg 1/2$ s in this example. The first of these is met marginally by the assumed value of $\sigma_\epsilon = 4\%$. The second really amounts to a requirement that the conditional variance of the scale factor be greatly reduced (i.e., $L \ll 1$) before the last half-second of flight, so that the problem has essentially the linear-quadratic-Gaussian form thereafter. It is clear from Fig. 3 that this condition is also met.

Also shown in Fig. 3 are the prior mean and standard deviation of the optimal control in the absence of scale-factor uncertainty (i.e., for $\sigma_\epsilon = 0$). By comparison, the optimal control law modifications induced by this uncertainty are significant in magnitude only during the first second of flight. Because of the $\sqrt{K}$ enhancement factor, however, which is $\sqrt{2(5-t'/\text{s})}$ in this case, the rapidly fluctuating probing control actually makes a fairly comparable contribution to reducing the scale-factor uncertainty and corresponds to missile attitude variations of 2–5 deg here.

Conclusion

With the use of a logarithmic time variable, the missile guidance problem treated here has been transformed into one of controlling a dynamic system with meaningful time constants. In the absence of seeker scale-factor uncertainty, the optimal control law is linear feedback on a Kalman filter estimate of relative lateral velocity, which corresponds to proportional navigation (linear feedback on the line-of-sight rate) in the original problem. The main effects of such scale-factor uncertainty on the optimal control law are to change the feedback gain in this linear control and, more interestingly, to add a dithering control component. This dithering occurs at a speed typical of the filter time constants, which are much faster than those of the dynamic system because the measurements are very accurate. The dithering is produced by the filter's innovation process driving a high-frequency critically damped second-order system. It acts as a probing control and has the effect of enhancing the estimation of the uncertain scale factor by means of the accompanying attitude variations. It is more pronounced at short ranges and high closing speeds, when it is more crucial to estimate quickly this scale factor.

A number of features of the problem here might be involved in producing this dithering effect. The noise in the line-of-sight measurements is very low. The multivariate structure of this problem is such that the state component whose estimate is wanted for feedback control is the time derivative of the component that is measured. Also, the problem structure is such that changes in the control are required to estimate effectively the scale-factor error, and faster changes are more efficient than slower ones for this purpose. However, just what problem features would generally lead to a solution of this type, or some similar type, is not clear at this point.

References

[1]Verges, P.L. and McClendon, J.R., "Optimal Control and Estimation for Strapdown Seeker Guidance of Tactical Missiles," *Journal of Guidance, Control, and Dynamics*, Vol. 5, May-June 1982, pp. 225-226.

[2]Astrom, K.J., "Theory and Applications of Adaptive Control – A Survey," *Automatica*, Vol. 19, Sept. 1983, pp. 471-486.

[3]Kumar, P.R., "A Survey of Some Results in Stochastic Adaptive Control," *SIAM Journal on Control and Optimization*, Vol. 23, May 1985, pp. 329-380.

[4]Chen, H.-F. and Guo, L., "Optimal Stochastic Adaptive Control with Quadratic Index," *International Journal of Control*, Vol. 43, March 1986, pp., 869-881.

[5]Hijab, O.B., "The Adaptive LQG Problem – Part I," *IEEE Transactions on Automatic Control*, Vol. AC-28, Feb. 1983, pp. 171-178.

[6]Schweppe, F.C., *Uncertain Dynamic Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1973.

[7]Wong, E. and Hajek, L., *Stochastic Processes in Engineering Systems*, Springer Verlag, New York, 1986.

[8]Speyer, J.L., Hull D.G., and Tseng, C.Y., "Estimation Enhancement by Trajectory Modulation for Homing Missiles," *Journal of Guidance, Control, and Dynamics*, Vol. 7, March-April 1984, pp. 167-174.

[9]Hull, D.G. and Speyer, J.L., "Maximum Information Guidance for Homing Missiles," *Journal of Guidance, Control, and Dynamics*, Vol. 8, July-Aug. 1985, pp. 494-497.

[10]Bryson, A.E. and Ho, Y.-C., *Applied Optimal Control*, Hemisphere, Washington, DC, 1975.

[11]Balakrishnan, A.V., "On a New Computing Technique in Optimal Control," *SIAM Journal on Control*, Vol. 6, 1968, pp. 149-173.

[12]Wong, E., *Stochastic Processes in Information and Dynamical Systems*, McGraw-Hill, New York, 1971, p. 162.

[13]Willman, W.W., "Edgeworth Expansions in State Perturbation Estimation," *IEEE Transactions on Automatic Control*, Vol. AC-26, April 1981, pp. 493-498.

[14]Willman, W.W., "Asymptotically Optimal Adaptive Control," Naval Weapons Center, China Lake, CA, TP-6530, July 1984.

[15]Stratonovich, R.L., "On the Theory of Optimal Control: Sufficient Coordinates," *Automation and Remote Control*, Vol. 23, 1963, pp. 847-854.

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

[17]Feldbaum, A.A., "Dual Control Theory I," *Automation and Remote Control*, Vol. 21, 1961, pp. 874-880.