# Performance of suboptimal linear control systems

L. Meier and B. D. O. Anderson


## Synopsis

It is shown that the calculation of the performance of a class of suboptimal linear control systems is similar to the calculation of the performance of an equivalent optimal system. The performances of both optimal and suboptimal linear systems have terms for cost owing to uncertainty about the initial state of the plant, a nonzero mean of the initial state, error in the estimate of the state, and error in control owing to additive noise at the input. A suboptimal system, in addition to having greater costs than the optimal system for each of these terms, has a cost term which vanishes if either optimal control or optimal estimation is used.


## List of symbols

$x, x_e$ = state and state estimate
$F, G, H$ = matrixes describing a finite-dimensional dynamic system
$v, w$ = additive noise at output and input
$y, u$ = output and input
$K$ = feedback law
$V$ = performance index
$Q, R$ = weighting matrixes for control problem
$M, N$ = noise-covariance matrixes
$x_0, m$ = initial state and mean initial state
$S$ = state-estimate covariance
$S_0$ = initial $S$
$L$ = feedback law for state estimator
$t_0, t_1$ = initial and final time
$z$ = state of suboptimal system
$P$ = matrix appearing in control problem
$E$ = expectation operator
Subscript $s$ indicates parameters of suboptimal system
A prime indicates transposition

## 1 Introduction

The systems under consideration are linear finite-dimensional systems, with additive Gaussian noise at the input and output, and thus can be described by the state-space equations:

$$\dot{x} = F(t)x + G(t)u + w \quad . \quad . \quad . \quad (1a)$$

$$y = H(t)x + v \quad . \quad . \quad . \quad (1b)$$

where $x$ is an $n$vector (the state), $u$ is a $p$vector (the input), $y$ is an $m$vector (the output), and the matrixes $F$, $G$ and $H$ are of the appropriate dimensions. The terms $v$ and $w$ represent the noise, and are discussed later. Under certain conditions, discussed in later Sections, the optimum performance, given by

$$V[E\{x(t_0)\}, t_0] = \min E\left\{\int_{t_0}^{t_1}(x'Qx + u'Ru)dt\right\} \quad . \quad (2)$$

can be achieved by choosing

$$u = -Kx_e \quad . \quad . \quad . \quad (3)$$

where $x_e$ is the minimum-variance estimate of $x$: i.e. a linear feedback of the best estimate of the state is used to provide the optimal input. The question now arises as to what $V$ is under this feedback law, and what $V$ results if the law is changed from the optimal law. The calculations giving these are presented as the main material of this paper.

There are, actually, two distinct ways by which eqn. 3 can be varied. The first is to change the feedback law $K$; the second consists in using a suboptimal estimate of $x$. Problems of this, or a similar, nature have been considered in References 1–4, which deal with discrete-time problems, and References 5–8, which deal with continuous-time problems. These References consider, principally, optimal control laws, while Reference 9, on the other hand, presents a detailed consideration of suboptimal discrete systems. This paper parallels much of the discussion in Reference 9. It was first shown by Sivan,[10] although only for the discrete-time case, that the only stochastic optimal-control problems leading to a linear-feedback law are those involving Gaussian noise and a quadratic loss function. For this case, References 2 and 11 demonstrate that it is valid to decouple the problems of estimation, i.e. determining $E(x)$, and control, i.e. determining the feedback law $K$. The conclusion is that the same feedback law can be used, irrespective of the noise statistics. As Reference 7 points out, however, a completely rigorous treatment, deriving this result, is still lacking for the continuous-time case. It is not our aim to present such a treatment, but rather to examine the consequences of using certain feedback laws, assuming that decoupling is valid.

The paper is arranged as follows. In Section 2 the filtering and control problem is reviewed, and in Section 3 the optimal-control problem is posed precisely, for derivation of the optimum-performance index (theorem 4). At the same time, the awkward points which remain to be cleared up by more rigorous arguments in the derivation of the optimal law are stated explicitly. The optimal law itself will not, however, be derived. In Section 4, suboptimal systems are discussed, and it is shown how the results of Section 3 may be used for the calculation of the suboptimal-performance indexes. Section 5 discusses directions for future research, which, if followed, might yield results describing simpler, and more economical, designs of large-dimensional systems than are at present possible.

## 2 Review of filtering and control problems

This discussion of the filtering problem summarises the development of Reference 12. Consider a linear system, governed by eqns. 1$a$ and 1$b$, where, at time $t_0$, the initial state $x_0$ is known only to the extent that its probability distribution is Gaussian, with mean $m$ and

$$\text{cov}\,(x_0, x_0) = S_0 \quad . \quad . \quad . \quad (4)$$

Naturally, $S_0$ is a nonnegative definite symmetric matrix. The input and output additive noises are independent and Gaussian, with mean zero, and

$$\text{cov}\,\{w(t), w(\tau)\} = M(t)\delta(t-\tau) \quad . \quad . \quad . \quad (5a)$$

$$\text{cov}\,\{v(t), v(\tau)\} = N(t)\delta(t-\tau) \quad . \quad . \quad . \quad (5b)$$

Both $M$ and $N$ are nonnegative definite symmetric matrixes; furthermore, $N$ is nonsingular, which implies that no component of the output can be measured exactly.

The problem is to design a dynamic system, with inputs $u$ and $y$, whose output is the minimum-variance estimate of $x$, the mean of $x(t)$, conditional on a knowledge of $u$ and $y$ over the interval $[t_0, t]$. It is possible to do this with a finite-dimensional linear system, as shown in Reference 12. Fig. 1 shows the form of the system, the estimate of $x(t)$ being denoted by $x_e(t)$. The initial state of the estimator is taken as $m$. The solution to the filter problem is given by the following theorem.[12]

*Theorem* 1. Consider the plant described by eqn. 1, with noise as defined in eqns. 4 and 5. Then the gain matrix of the filter (Fig. 1), optimal in the sense described, is

$$L = SH'N^{-1} \quad . \quad . \quad . \quad (6)$$

Paper 5169 C, first received 20th May 1966 and in revised form 22nd September 1966
Dr. Meier is with the Stanford Research Institute, Menlo Park, Calif., USA, and Prof. Anderson is with the Stanford Electronics Laboratory, Stanford, Calif., USA

where $S$ is the solution of the matrix Riccati differential equation

$$\dot{S} = FS + SF' - SH'N^{-1}HS + M \quad . \quad . \quad . \quad . \quad (7)$$

subject to the initial condition

$$S(t_0) = S_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

It is shown[12] that eqn. 7 has a solution, defined by $[t_0, t_1]$, for any $t_1 > t_0$. A most important fact also given in Reference 12 is lemma 1.

*Lemma* 1. Irrespective of $u_{[t_0, t]}$, i.e. the history of $u$ over the interval $[t_0, t]$,

$$\text{cov}\,\{x_e(t) - x(t), x_e(t) - x(t) | u_{[t_0, t]}, y_{[t_0, t]}\} = S(t) \quad (9)$$

In other words, $S(t)$ is the covariance of the error involved in estimating the state, conditional on a knowledge of $u$ and $y$, from the initial time up to time $t$. The state $x(t)$ is a

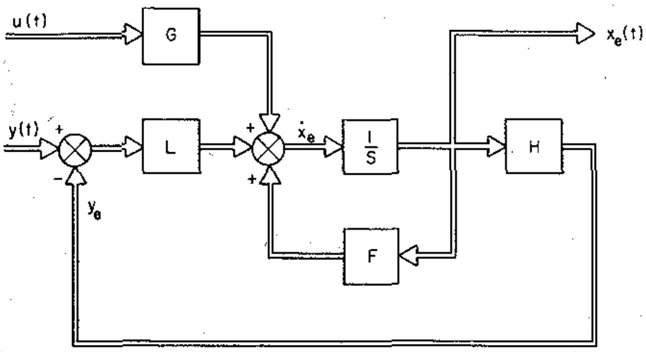


**Fig. 1**
*Optimal estimation*

Gaussian random variable, and thus knowledge of $x_e(t)$ and $S(t)$ gives the full conditional statistical distribution of $x(t)$.

For the control problem, suppose there is no noise present; so the system is described by

$$\dot{x} = Fx + Gu \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

The object is to select $u$, as a function of $x$, to find

$$V(x_0, t_0) = \min \int_{t_0}^{t_1} (x'Qx + u'Ru)dt \quad . \quad . \quad . \quad . \quad (11)$$

which, as the notation implies, is a function of the initial state. The matrixes $Q$ and $R$ are nonnegative definite symmetric, with $R$ nonsingular. The solution to the control problem is given by theorem 2.[13]

*Theorem* 2. Consider a plant (eqn. 10) with a performance index defined by eqn. 11. Optimum performance is achieved by taking

$$u = -Kx \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

$$\text{where } K = R^{-1}G'P \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

and $P$ is the solution of the matrix Riccati differential equation

$$-\dot{P} = F'P + PF - PGR^{-1}G'P + Q \quad . \quad . \quad . \quad (14)$$

subject to

$$P(t_1) = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

Then $V(x_0, t_0) = x_0'P(t_0)x_0$ . . . . . . . . (16)

It is shown[13] that eqn 14 has a unique, symmetric, nonnegative definite solution, for all $[t_0, t_1]$, $(t_0 < t_1)$, so that eqn. 16 is well defined.

The duality of the estimation and control problem is shown by the similarity between eqns. 7 and 14.

## 3 Stochastic optimal-control problem

Consider the problem: given the system described by eqns. 1*a*, 1*b*, 4, 5*a* and 5*c*, where, as before, $w$ and $v$ are independent and Gaussian with mean zero, and $x_0$ is Gaussian with mean $m$, show how to select $u$ as a function of $y$ to obtain $V$ as given in eqn. 2. As remarked in Section 1, other workers have shown that, provided that $x$ is replaced by $x_e$, the optimal control law is independent of the noise statistics. In particular, the optimal gain is identical with the optimal gain for the deterministic case. We can thus state theorem 3.

*Theorem* 3. Consider the plant (eqn. 1) with noise given statistically by eqns. 4 and 5. The performance index (eqn. 2) is obtained by choosing

$$u(t) = -K(t)x_e(t) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

where $K$ is derived from eqns. 14, 15 and 16, and is the solution to the corresponding problem of theorem 2.

The complete system of plant and controller is shown in Fig. 2. The equations of this combined system are

$$\dot{x} = Fx - GKx_e + w \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18a)$$

$$\dot{x}_e = Fx_e - GKx_e + LH(x - x_e) + Lv \quad . \quad . \quad (18b)$$

Though the feedback law (eqn. 17) is independent of the statistics of the noise, i.e. the matrixes $S_0$, $M(t)$ and $N(t)$, the optimum performance is not. This optimum performance is given by noting that

$$\begin{aligned} x'Qx + u'Ru &= x'Qx + x_e'PGR^{-1}G'Px_e' \text{ (by eqns. 17 and 13)} \\ &= x'(-\dot{P} - F'P - PF + PGR^{-1}G'P)x \\ &\quad + x_e'PGR^{-1}G'Px_e \quad \text{(by eqn. 14)} \\ &= -x'\dot{P}x - \dot{x}'Px - x'P\dot{x} - 2x'PGR^{-1}G'Px_e \\ &\quad + 2x'Pw + x'PGR^{-1}G'Px + x_e'PGR^{-1} \\ &\quad G'Px_e \quad \text{(by eqns. 18a and 13)} \\ &= -\frac{d}{dt}(x'Px) + 2x'Pw + (x' - x_e')PGR^{-1} \\ &\quad G'P(x - x_e) \quad . \quad . \quad . \quad . \quad (19) \end{aligned}$$

Consequently,

$$V(m, t_0) = E\int_{t_0}^{t_1} \left\{ -\frac{d}{dt}(x'Px) + 2x'Pw + (x' - x_e')PGR^{-1}G'P(x - x_e) \right\} dt \quad . \quad (20)$$

The first term may readily be separated out to give

$$E\left(-x'Px\Big|_{t_0}^{t_1}\right) = E\{x_0'P(t_0)x_0\}$$

using the boundary condition on $P$ (eqn. 15). The initial statistical conditions on $x(t_0)$ are also given in terms of its covariance $S_0$, and thus

$$E\left(-x'Px\Big|_{t_0}^{t_1}\right) = \text{tr}\,\{S_0P(t_0)\} + m'P(t_0)m$$

tr denoting the trace of a matrix (Appendix 8.1).

The second term in the integrand of eqn. 20 can be evaluated as follows. It follows from eqn. 18*a*, where $\phi$ is the system transition matrix, that

$$x(t) = \phi(t, t_0)x_0 + \int_{t_0}^{t} \phi(t, \tau)\{-G(\tau)K(\tau)x_e(\tau) + w(\tau)\}dt$$

$$\text{and} \quad \text{cov}\,(x, w) = \tfrac{1}{2}M(t) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (21a)$$

where the $\frac{1}{2}$ follows from properties of the delta function, and is further discussed at the end of this Section. Now a second application of the method of Appendix 8.1 yields

$$E\int_{t_0}^{t_1} 2x'Pwdt = \int_{t_0}^{t_1} \text{tr}\,(MP)dt \quad . \quad . \quad . \quad . \quad . \quad . \quad (21b)$$

since $w$ has zero mean. Lemma 1 is stated in terms of conditional, i.e. *a posteriori*, covariances, while the expectation in eqn. 20 is *a priori*; therefore to use lemma 1 to evaluate the third term of eqn. 20 requires the identity

$$\underset{x}{E}(x) \equiv \underset{y}{E}\{\underset{x}{E}(x|y)\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (22a)$$

Hence, from eqn. 9, and using Appendix 8.1 a third time,

$$E\left\{\int_{t_0}^{t_1} (x - x_e)'PGR^{-1}G'P(x - x_e)dt\right\}$$

$$= E\left[\int_{t_0}^{t_1} E\{(x - x_e)'PGR^{-1}G'P(x - x_e)|u_{[t_0,t]}y_{[t_0,t]}\}dt\right]$$

$$= E\left\{\int_{t_0}^{t_1} \mathrm{tr}\,(SPGR^{-1}G'P)dt\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (22b)$$

Because $S$, $P$, $R$ and $G$ are all deterministic, the final expectation can be dropped, and we have theorem 4.

*Theorem* 4. With the same hypothesis as theorem 3, we have

$$V(m, t_0) = \mathrm{tr}\,\{S_0P(t_0)\} + m'P(t_0)m$$

$$+ \int_{t_0}^{t_1} \mathrm{tr}\,(SPGR^{-1}G'P + MP)dt \quad (23)$$

where $S$ is as defined in theorem 1, and $P$ is as defined in theorem 2. Clearly, the second term in eqn. 23 is nonnegative, since $P(t_0)$ is a nonnegative definite matrix. The first term and the two integrands are also nonnegative (in fact, positive in nondegenerate situations), since $S_0$, $P(t_0)$, $S$, $PGR^{-1}G'P$, $M$ and $P$ are all nonnegative definite (Appendix 8.2). Thus all four terms are nonnegative; they are due, respectively, to

(*a*) uncertainty about the initial state
(*b*) a nonzero mean of the initial state
(*c*) error in the estimate of the state
(*d*) error in the control due to the additive noise at the input.

Note that, in the noisefree case, only the second term is present, and eqn. 23 reduces to eqn. 16, the solution for the deterministic problem.

Under suitable conditions for the plant, e.g. complete observability and complete controllability,[13] it is possible to examine the filtering and control problems separately, when the upper time limit $t_1$ becomes infinite. In the case of a time-invariant plant with stationary noise, the feedback law $K$ and the estimator feedback law $L$ become time-invariant. Under these conditions, it is evident from eqn. 23 that $V$ will increase without bound. The physical reasoning behind this is not hard to see; there will always be noise, perturbing the system and the estimator from the zero state, and, as a result, compensatory control. Thus, even if the system state did become zero, it would not stay there; further control, contributing to $V$, would result. Note, however, that, if there is no output or input noise ($M = N = 0$) but only uncertainty about the initial state, $V$ will be finite for the time-invariant ($t_1 = \infty$) case.

This Section closes with some remarks on the derivation of eqn. 22 and, more generally, the optimal law. The derivation by Kalman and Bucy[12] of the differential equation satisfied by $S$ (eqn. 7) includes essentially the same difficulty as the derivation of eqn. 21; integrals containing $\delta$ functions appear which have the $\delta$ function singularity at one of the endpoints of an integral. Presumably, this is due to the imperfect characterisation of the stochastic differential equation eqns. 18*a* and *b*, or the equivalent equations in the Kalman and Bucy paper. The derivation of the optimal law appears to depend on similar manipulations. There is, however, every indication that the end results are quite correct and only their derivation is open to criticism.

## 4 Calculation of suboptimal performance indexes

A suboptimal system will result if the estimator is built incorrectly or the wrong feedback law is used. Referring to Fig. 2, consider systems where $L$ is replaced by $L_s$ and $K$ is replaced by $K_s$. This Section is devoted to showing how such a system may be analysed.

Taking the plant, estimator, and feedback law realisation as one combined system, suppose it has a state vector

$$z = \begin{bmatrix} x \\ x_e - x \end{bmatrix} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

For the second set of components of $z$, $x_e$ could have been used, but $x_e - x$ proves more convenient. From eqns. 18*a* and *b*,

$$\dot{x} = (F - GK_s)x - GK_s(x_e - x) + w$$

and $\quad \dot{x}_e - \dot{x} = (F - L_sH)(x_e - x) + L_sv - w$

Consequently,

$$\dot{z} = \begin{bmatrix} F - GK_s & -GK_s \\ 0 & F - L_sH \end{bmatrix} z + \begin{bmatrix} w \\ L_s v - w \end{bmatrix} \quad (25a)$$

which may be written, in briefer form, as

$$\dot{z} = F_s z + w_s \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (25b)$$

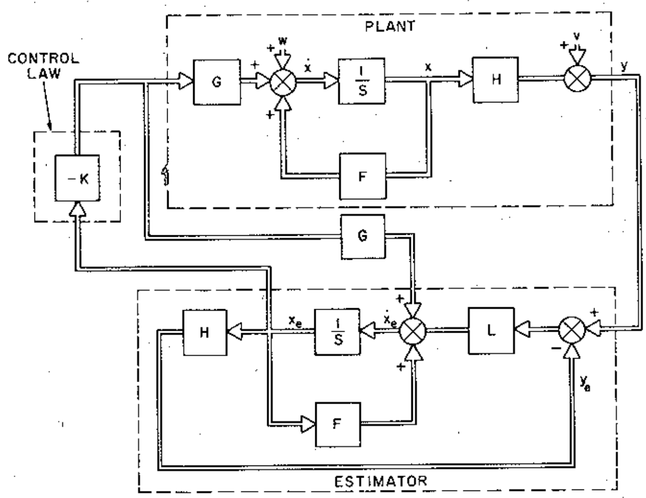


**Fig. 2**
*Linear control system*

$$\text{where } F_s = \begin{bmatrix} F - GK_s & -GK_s \\ 0 & F - L_sH \end{bmatrix} \quad . \quad . \quad . \quad . \quad . \quad (25c)$$

and the noise term $w_s$ is given by

$$w_s = \begin{bmatrix} w \\ L_s v - w \end{bmatrix} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (25d)$$

Moreover, easy calculations yield

$$\mathrm{cov}\,(w_s\ w_s) = M_s = \begin{bmatrix} M & -M \\ -M & M + L_sNL'_s \end{bmatrix} \quad (25e)$$

The cost of using laws $K_s$, $L_s$ for control of the plant (eqn. 1) is

$$V_s(m, t_0) = E\int_{t_0}^{t_1} (x'Qx + u'Ru)dt$$

$$= E\int_{t_0}^{t_1} \{x'Qx + (x' + x'_e - x')K_sRK_s(x + x_e - x)\}dt$$

$$= E\int_{t_0}^{t_1} z'Q_szdt \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (26)$$

$$\text{where } Q_s = \begin{bmatrix} Q + K'_sRK_s & K'_sRK_s \\ K'_sRK_s & K'_sRK_s \end{bmatrix} \quad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

Eqn. 26 states the cost, using suboptimal laws for the original system. Consider now the combined system, and suppose that we allow inputs $u_s$, but with the input matrix $G_s$ zero; i.e. the inputs have no effect on the system.

Eqn. 23, with $G$ replaced by $O$, $P$ replaced by $P_s$ and $S$ replaced by $S_s$, is used to calculate $V_s$:

$$V_s(m, t_0) = \mathrm{tr}\,\{S_{s0}P_s(t_0) + m_sP_s(t_0)m_s\}$$

$$+ \int_{t_0}^{t_1} \mathrm{tr}\,(M_sP_s)dt \quad . \quad (28)$$

The matrix $S_{s0}$ is the covariance at time $t_0$ of $z$, and is given by

$$S_{s0} = \begin{bmatrix} S_0 & S_0 \\ S_0 & S_0 \end{bmatrix} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

The matrix $P_s$ is the solution of the Riccati equation (eqn. 14):

$$-\dot{P}_s = F_s'P_s + P_sF_s + Q_s \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (30)$$

where $F_s$ and $Q_s$ are as given in eqns. 25$c$ and 27. The initial condition is

$$P_s(t_1) = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (31)$$

Proceeding to the second term of eqn. 28, we note that

$$m_s = \begin{bmatrix} m \\ 0 \end{bmatrix} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (32)$$

The third term derives from eqn. 25$e$, giving $M_s$, and eqn. 30, giving $P_s$. Hence, in principle, eqn. 28 may be evaluated. The expression on the right-hand side of eqn. 28 is not in its most convenient form, however, and does not exhibit separately the dependence of $V$ on the initial covariance of the state, the output noise and so on. Therefore, an alternative representation of $V$ is developed. It is convenient to partition $P_s$ in the same fashion as the other matrixes:

$$P_s = \begin{bmatrix} P_1 & P_2 \\ P_2' & P_3 \end{bmatrix} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (33)$$

Then eqn. 30 becomes

$$-\dot{P}_1 = (F - GK_s)'P_1 + P_1(F - GK_s) + Q + K_s'RK_s \quad (34a)$$

$$-\dot{P}_2 = (F - GK_s)'P_2 + P_2(F - L_sH) - P_1GK_s + K_s'RK_s \quad . \quad . \quad . \quad . \quad (34b)$$

$$-\dot{P}_3 = -(GK_s)'P_2 - P'_2GK_s + (F - L_sH)'P_3 + P_3(F - L_sH) + K'_sRK_s \quad . \quad . \quad . \quad . \quad (34c)$$

with $P_1(t_1) = P_2(t_1) = P_3(t_1) = 0$ . . . . . . . (35)

The covariance $S_s$ of $z$ can be found, by using eqn. 7 and replacing $F$ by $F_s$, $M$ by $M_s$ and setting $H = 0$. The result is

$$\dot{S}_s = F_sS_s + S_sF_s' + M_s \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (36)$$

Partitioning $S_s$, in the same way as $P_s$, allows rewriting of eqn. 36 as

$$\dot{S}_1 = (F - GK_s)S_1 + S_1(F - GK_s)' - GK_sS_2' - S_2(GK_s)' + M \quad . \quad . \quad . \quad . \quad (37a)$$

$$\dot{S}_2 = (F - GK_s)S_2 + S_2(F - L_sH)' - GK_sS_3' - M \quad . \quad (37b)$$

$$\dot{S}_3 = (F - L_sH)S_3 + S_3(F - L_sH)' + M + L_sNL_s' \quad . \quad (37c)$$

with eqn. 29

$$S_1(t_0) = S_2(t_0) = S_3(t_0) = S_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)$$

Substitution of eqns. 34, 35, 37 and 38 into eqn. 28 yields theorem 5, after much algebraic manipulation, which is given in Appendix 8.3.

*Theorem* 5. Consider the plant (eqn. 1), with noise as defined in eqns. 4 and 5. Suppose an estimator has feedback law $L_s$, while the plant has a feedback law $K_s$. Then

$$V_s(m, t_0) = \mathrm{tr}\{S_0P_1(t_0)\} + m'P_1(t_0)m + \int_{t_0}^{t_1}[\mathrm{tr}\{S_3(GK_sP_1 + P_1K_s'G' - K_sRK_s') + MP_1 + 2(L_sNL_s' - S_3H'L_s)P_1\}]dt \quad . \quad . \quad . \quad (39)$$

The various costs in this expression can be allocated to different sources. Thus, $m'P_1(t_0)m$ is the performance index of a deterministic linear regulator, when the feedback law $K_s$ is used, instead of the optimal law $K$. If the initial state $x_0$ is only known to the extent of its mean and covariance, but there is no input or output additive noise, the additional contribution to the performance index is given by $\mathrm{tr}\{S_0P_1(t_0)\}$. As for the optimal case, the first term of the integrand corresponds to error in the estimate of the state; the second term to error in the control from the additive noise at the input. The third term is new, and arises from the suboptimality; it involves both $S_3$, a modified estimation-error covariance, and a matrix $P_2$, which depends on the feedback law chosen.

If $K_s$ is set equal to the optimal gain, eqns. 14 and 34 can be used to show that $P_1 = P$, $P_2 = 0$. This yields, for eqn. 39,

$$V_s(m, t_0) = \mathrm{tr}\{S_0P(t_0)\} + m'P(t_0)m + \int_{t_0}^{t_1}\mathrm{tr}(S_3PGR^{-1}G'P + MP)dt \quad . \quad (40)$$

which is remarkably similar to the optimal result (eqn. 23), the only difference being the appearance of $S_3$ rather than $S$. Likewise, if $L_s$ is replaced by the optimal $L$, eqns. 7 and 37$c$ yield $S_3 = S$, and

$$V_s(m, t_0) = \mathrm{tr}\{S_0P_1(t_0)\} + m'P_1(t_0)m + \int_{t_0}^{t_1}\mathrm{tr}\{S(GK_sP_1 + P_1K_s'G' - K_sRK_s') + MP_1\}dt \quad (41)$$

Again the fifth term in eqn. 39 drops out; use of either the optimal-estimator law or the optimal-feedback law causes it to vanish. From eqn. 41, if both optimal-feedback and estimator laws are used, the original performance index (eqn. 23) is recovered.

## 5 Conclusions

The equations in this paper are suitable for digital computation, as discussed in Reference 12, despite the fact that they are nonlinear. Moreover, the equations required for calculating suboptimal-performance indexes are of the same nature as those required for calculating the optimal-performance index.

Further investigation of the variation of performance index resulting from simplified controller design (corresponding to zero columns of $L_s$ or $K_s$) is clearly in order. Many systems today are, no doubt, over-designed, in the sense that only small changes in the performance index would result from gross simplification of the controller. The breakdown of the performance into the sum of components, each arising from different parts of the overall system, should make investigation easier.

## 6 Acknowledgment

The work was carried out under Joint Services Contract Nonr 225(83), and was based on work done by NASA Ames Research Center on Contract NAS-2-2457. One of the authors (B.D.O.A.) would also like to acknowledge the financial assistance of the Services Canteens Trust Fund, an agency of the Australian Government, and to thank the United States Education Foundation in Australia for a Fulbright travelling grant.

## 7 References


1 GUNCKEL, T. L.: 'Optimum design of sampled data systems with random parameters', Report SEL TR 2102-2, Stanford Electronics Laboratories, Stanford, Calif., 1961

2 JOSEPH, P. D., and TOU, J. T.: 'On linear control theory', *Trans. Amer. Inst. Elect. Engrs.*, 1961, Pt. II, **80**, p. 18

3 KALMAN, R. E.: 'On the general theory of control systems', Proceedings of the 1st international congress of the IFAC, Moscow, 1960, pp. 481–492

4 SCHULTZ, P. R.: 'An optimal control problem with state vector measurement errors', *in* LEONDES, C. J. (Ed.): 'Advances in control systems', Vol. 1 (Academic Press, 1964)

5 FLORENTIN, J. J.: 'Optimal control of continuous time, Markov, stochastic systems', *J. Electronics Control*, 1961, **10**, pp. 473–488

6 FLORENTIN, J. J.: 'Optimal control of systems with generalized Poisson inputs', Joint automatic control conference, New York, 1962

7 WONHAM, W. M.: 'Stochastic problems in optimal control', *IEEE Internat. Convention Record*, 1963, pp. 114–124

8 KRASOVSKII, N. N., and LIDSKII, Z. A.: 'Analytical design of controllers in systems with random attributes', *Automn. Remote Control*, 1961, **22**, pp. 1021–1025, 1141–1146 and 1289–1294

9 MEIER, L.: 'Combined optimum control and estimation theory', Stanford Research Institute Contractor Report to NASA Ames Research Center, Contract NAS 2-2457, 1965

10 SIVAN, R.: 'The necessary and sufficient condition for the optimal controller to be linear', Joint automatic control conference, Stanford, 1964

11 FLORENTIN, J. J.: 'Partial observability and optimal control', *J. Electronics Control*, 1960, **9**, pp. 65–80

12 KALMAN, R. E., and BUCY, R. S.: 'New results in linear filtering and prediction theory', *Trans. ASME*, 1961, **83**, Series D, pp. 95–108

13 KALMAN, R. E.: 'Contributions to the theory of optimal control', *Bol. Soc. Mat. Mex.*, 1960, pp. 102–119

14 PARZEN, E.: 'Modern probability theory and its applications' (Wiley, 1960), p. 384

## 8 Appendixes

### 8.1 Evaluation of two expectations

To evaluate $E(x'Ax)$, where $A$ is any $n \times n$ square matrix,

$$E(x'Ax) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} E(x_i x_j) \quad . \quad . \quad . \quad (42)$$

Now

$$E(x_i x_j) = E(x_i)E(x_j) + E[\{x_i - E(x_i)\}\{x_j - E(x_j)\}] \quad (43)$$

and so, if

$$\bar{x}_i \stackrel{\Delta}{=} E(x_i) \quad . \quad . \quad . \quad (44)$$

and $V \stackrel{\Delta}{=} \text{cov}(x - \bar{x}, x - \bar{x})$ . . . . . . . . (45)

then

$$E(x'Ax) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij}(\bar{x}_i \bar{x}_j + V_{ij})$$

$$= \bar{x}'A\bar{x} + \text{tr}(AV) \quad . \quad . \quad . \quad (46)$$

Similarly,

$$E(x'Ay) = \bar{x}'A\bar{y} + \text{tr}(AW) \quad . \quad . \quad . \quad (47)$$

where $W = \text{cov}(x - \bar{x}, y - \bar{y})$ . . . . . . . . (48)

### 8.2 Trace of product of nonnegative definite matrixes

In this Appendix it is shown that, if $A$ and $B$ are nonnegative definite,

$$\text{tr}(AB) \geqslant 0 \quad . \quad . \quad . \quad (49)$$

Since

$$\text{tr}(AB) = \text{tr}(BA)$$

for all nonsingular $T$

$$\text{tr}(AB) = \text{tr}\{T'ATT^{-1}B(T^{-1})'\} \quad . \quad . \quad . \quad (50)$$

which means that $A$ may be taken to be diagonal. Then

$$\text{tr}(AB) = \sum_i a_{ii} b_{ii}$$

$$\geqslant 0 \quad . \quad . \quad . \quad (51)$$

since the elements on the main diagonal of a nonnegative definite matrix are themselves nonnegative.

### 8.3 Derivation of eqn. 39

The purpose of this Appendix is to present the algebraic manipulations needed to derive eqn. 39. From eqns. 25 and 23, we have

$$\text{tr}(M_s P_s) = \text{tr}(MP_1 - MP_2' - MP_2 + MP_3 + L_s N L_s P_3) \quad . \quad . \quad . \quad (52)$$

and from eqn. 37*c*,

$$\text{tr}(MP_2) = \text{tr}(\dot{S}_3 P_2) - \text{tr}(L_s N L_s' P_2) - \text{tr}\{(F - L_s H) S_3 P_2 + S_3(F - L_s H)' P_2\} \quad . \quad (53)$$

From Appendix 8.2 and eqn. 34*b*, it follows that

$$\text{tr}\{(F - L_s H) S_3 P_2\} = \text{tr}\{S_3 P_2 (F - L_s H)\}$$
$$= - \text{tr}\{S_3 \dot{P}_2\} - \text{tr}\{S_3(F - GK_s)' P_2\} + \text{tr}\{S_3(P_1 GK_s - K_s' R K_s)\} \quad . \quad (54)$$

and combining eqns. 53 and 54,

$$\text{tr}(MP_2) = \text{tr}(\dot{S}_3 P_2) - \text{tr}(L_s N L_s' P_2) + \text{tr}(S_3 \dot{P}_2) - \text{tr}(S_3 P_1 GK_s) + tr(S_3 K_s' R K_s) - \text{tr}\{S_3(GK_s - L_s H)' P_2\} \quad . \quad . \quad . \quad (55)$$

Similarly,

$$\text{tr}(MP_2') = \text{tr}(\dot{S}_3 P_2') - \text{tr}(L_s N L_s' P_2') + \text{tr}(S_3 \dot{P}_2') - \text{tr}\{S_3(GK_s)' P_1\} + \text{tr}(S_3 K_s' R K_s) - \text{tr}\{S_3 P_2'(GK_s - L_s H)\} \quad . \quad . \quad . \quad (56)$$

and from eqns. 37*c* and 34*c*,

$$\text{tr}\{M + L_s N L_s') P_3\} = \text{tr}(\dot{S}_3 P_3 + S_3 \dot{P}_3) + \text{tr}\{S_3(K_s R K_s')\} - \text{tr}[S_3\{P_2' GK_s + (GK_s)' P_2\}] \quad . \quad . \quad . \quad (57)$$

Substitution of eqns. 55, 56 and 57 in eqn. 52 yields

$$\text{tr}(M_s P_s) = \text{tr}\{\dot{S}_3(P_3 - P_2 - P_2') + S_3(\dot{P}_3 - \dot{P}_2 - \dot{P}_2')\} + \text{tr}(M_s P_1) + \text{tr}\{S_3(GK_s P_1 + P_1 K_s' G' - K_s R K_s')\} + 2\,\text{tr}(L_s N L_s' - S_3 H' L_s' P_2) \quad . \quad . \quad . \quad (58)$$

where use has been made of the fact that

$$\text{tr}(A'B') = \text{tr}(BA) = \text{tr}(AB) \quad . \quad . \quad . \quad (59)$$

But

$$\frac{d}{dt} \text{tr}\{S_3(P_3 - P_2 - P_2')\} = \text{tr}\{\dot{S}_3(P_3 - P_2 - P_2') + S_3(\dot{P}_3 - \dot{P}_2 - \dot{P}_2')\} \quad . \quad (60)$$

and therefore

$$\int_{t_0}^{t_1} \text{tr}(M_s P_s) dt = \text{tr}\{S_0(P_3 - P_2 - P_2')\} + \int_{t_0}^{t_1} \{\text{tr}(MP_1) + S_3(GK_s P_1 + P_1 K_s' G' - K_s R K_s') + 2(L_s N L_s' - S_3 H' L_s' P_2)\} dt \quad . \quad . \quad (61)$$

Substitution of this result into eqn. 28 yields eqn. 39, since $G_s = 0$ and $m_s' = [m', 0]$.