[10] H. P. Geering, "Optimal control theory for nonscalar-valued performance criteria," Ph.D. dissertation, Dept. Elec. Eng., Mass. Inst. Technol., Cambridge, Aug. 1971 (available in the form of Microfiche Nr. AD 731 213, NTIS, U. S. Dept. of Comm.).
[11] H. P. Geering and M. Athans, "Optimal control theory for non-scalar-valued performance criteria," in *Proc. 5th Ann. Princeton Conf. Information Sciences and Systems*, Princeton, N.J., Mar. 1971.
[12] H. Halkin, "On the necessary condition for optimal control of non-linear systems," *J. Analyse Mathématique*, vol. 12, pp. 1–82, 1964.
[13] H. Halkin, "Topological aspects of optimal control of dynamical polysystems," *Contrib. Differential Equations*, vol. 3, pp. 377–385, 1964.
[14] E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*. New York: Wiley, 1967.
[15] L. W. Neustadt, "A general theory of extremals," *J. Comput. Sci. Syst.*, vol. 3, pp. 57–92, 1969.
[16] C. Olech, "Existence theorems for optimal problems with vector valued cost function," Center of Dynamical Systems, Brown University, Providence, R.I., Tech. Rep. 67-6, 1967.
[17] K. Ritter, "Optimization theory in linear spaces—Part I," *Math. Annal.*, vol. 182, pp. 189–206, 1969.
——, "Optimization theory in linear spaces—Part II," *Math. Annal.*, vol. 183, pp. 169–180, 1969.
——, "Optimization theory in linear spaces—Part III," *Math. Annal.*, vol. 184, pp. 133–154, 1970.
[18] R. T. Rockafellar, *Convex Analysis*. Princeton, N.J.: Princeton Univ. Press, 1970.
[19] E. Tse, "On the optimal control of linear systems with incomplete information," Ph.D. dissertation, Dep. Elec. Eng., Mass. Inst. Technol., Cambridge, Nov., 1969.
[20] B. Z. Vulikh, *Introduction to the Theory of Partially Ordered Spaces*. Groningen, The Netherlands: Wolters-Noordhoff Sc., 1967.
[21] H. S. Witsenhausen, "Minimax control of uncertain systems," Electronic Systems Lab., Mass. Inst. Technol., Cambridge, Rep. ESL-R-269, 1966.
[22] L. A. Zadeh, "Optimality and non-scalar-valued performance criteria," *IEEE Trans. Automat. Contr.* (Corresp.), vol. AC-8, pp. 59–60, 1963.

Hans P. Geering (S'70–M'71) was born on June 7, 1942. He received the degree of Diplomierter Elektroingenieur from the Eidgenössische Technische Hochschule in Zurich, Switzerland, in 1966, and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1969 and 1971, respectively.

In 1967 and 1968, he worked with Sprecher & Schuh AG in Suhr, Switzerland, and Oerlikon-Buehrle AG in Zurich, Switzerland. From September, 1968 to August, 1971 he was a Research Assistant in the M.I.T. Electronic Systems Laboratory and a Teaching Assistant in the Department of Electrical Engineering at M.I.T. He is presently with Oerlikon-Buehrle AG. His research interests are in estimation and control theory.

Dr. Geering is a member of Schweizerische Gesellschaft für Automatik (IFAC) and Schweizerischer Elektrotechnischer Verein.

Michael Athans (S'58–M'61–SM'69–F'73) for a photograph and biography see page 30 of the February issue of this TRANSACTIONS.

Dual Effect, Certainty Equivalence, and Separation in Stochastic Control

YAAKOV BAR-SHALOM, MEMBER, IEEE, AND EDISON TSE, MEMBER, IEEE

***Abstract*—In this paper the various policies in fixed end-time stochastic control are discussed first. The emphasis is on the difference between the feedback and closed-loop policies. It is shown how the closed-loop policy has the important property that it can be actively adaptive, while the feedback policy can only be passively adaptive. The feature of being actively adaptive is possible when the control has a dual effect, i.e., in addition to its effect on the state it affects the state uncertainty. The intimate connection between the neutrality (lack of dual effect) and certainty equivalence properties for a class of problems is proved. This new result is then used to widen the class of problems for which it was previously known that the certainty equivalence property holds.**

Manuscript received October 29, 1973. Paper recommended by D. D. Sworder, Past Chairman of the IEEE S-CS Stochastic Control Committee and D. G. Lainiotis, Chairman of the IEEE S-CS Adaptive and Learning Systems, Pattern Recognition Committee. This work was supported in part by the AFOSR under Contracts F44620-73-C-0028 and F44620-74-C-0026.

The authors are with Systems Control, Inc., Palo Alto, Calif. 94304.

I. Introduction

IN THE first part of this paper a discussion of the possible policies in the fixed end-time control of stochastic systems is presented. The various classes of policies differ in their information patterns—the availability of past observations and the possible usage of information about the future observations. Since the controls are nonanticipative, the only information about the future observations a controller can use is the probability distribution of the corresponding observation random variables. Loosely speaking, this is called "the future observation program and the associated statistics." It is shown that the incorporation of the future observation program and the associated statistics is what makes the difference between a feedback control policy and a closed-loop control policy. It is only the latter one that takes into account the possible benefit to be derived from the future observations—it "knows" that the loop will be closed after each future measurement. The distinction between the feedback and closed-loop policies has not been sufficiently stressed in the literature and is believed to be especially important when one is looking for suboptimal algorithms. In such a case it is most desirable to obtain, if possible, a control scheme from the closed-loop class rather than the feedback class. These control policies, along with others are defined in Section II. It is pointed out that a consequence of the above mentioned difference between a feedback policy and a closed-loop policy is the following: the former is *passively adaptive* while the latter can be *actively adaptive*. The active learning feature of a closed-loop control policy can be utilized when the control has the so-called dual effect (Feldbaum [12]).

The control is said to have a dual effect when, in addition to its effect on the state of the system, it affects the uncertainty of the system's state. If the control cannot affect this uncertainty, the system is called neutral. When such a dual effect is present, the control can be used to improve the estimation, which ultimately helps to achieve the control objective. In this way a closed-loop stochastic control regulates its learning as required by the control objective. Another widely encountered property of a stochastic control, the certainty equivalence, and a weaker version of it, the separation property are also discussed in Section II.

The connection between the neutrality and the certainty equivalence properties is the topic of Section III. It is shown that for a class of problems the optimal (closed-loop) stochastic control has the certainty equivalence property if and only if the system is neutral. This new result is then used in Section III to widen the class of problems for which it was previously known that the certainty equivalence property holds.

II. Classes of Stochastic Control Policies, the Dual Effect and Certainty Equivalence

The control problem is defined as follows. The state of the system at time k, x_k, evolves according to the equation

$$x_{k+1} = f_k(x_k,u_k,v_k), \qquad k = 0,1,\cdots,N-1 \qquad (2.1)$$

where u_k is the control applied at time k and v_k, the process noise, and x_0 are random variables. The number of stages, N, is assumed to be given.[1] At time k, prior to applying the control, an observation might be obtained. The corresponding measurement is given by the equation

$$y_k = h_k(x_k,w_k), \qquad k = 0,1,\cdots,N \qquad (2.2)$$

where w_k, the measurement noise, is a random variable. All the above variables are, in general, vector-valued with appropriate dimensions. Denoting the underlying probability space by $(\Omega,\mathcal{B},\mathcal{P})$, the above random variables are $\mathcal{B}$ measurable functions of ω, the generic element of the sample space Ω, $\mathcal{B}$ is the σ-algebra of sets in Ω, and $\mathcal{P}$ the probability measure on $\mathcal{B}$.

The cost to be minimized is

$$J_0 = E\{\mathcal{L}^N[X_0{}^N,U_0{}^{N-1}]\} \qquad (2.3)$$

where $\mathcal{L}^N$ is a real-valued function,

$$U_0{}^{N-1} \triangleq \{u_i\}_{i=0}{}^{N-1} \qquad (2.4)$$

and $X_0{}^N$ is defined similarly.

The arguments of the minimization of the cost as given above are the controls $U_0{}^{N-1}$ to be applied during the N-stage control process.

As will be seen in the sequel, the various control policies to be discussed differ in the availability of past measurements and the future observation program with its associated statistics, to be defined next.

The set of observations from time i to time j, when the input sequence $U_0{}^{j-1}$ has been applied to the system, is denoted by[2]

$$Y_i{}^j = Y_i{}^j[\omega,U_0{}^{j-1}] \triangleq \{y_k\}_{k=i}{}^j, \qquad 1 \le i \le j \le N. \quad (2.5)$$

Since the probability distribution of x_0 is available, one can assume that there is no measurement at $k = 0$ and therefore $Y_i{}^j$ is defined for $i \ge 1$. Note also that the last measurement, y_N, is irrelevant, since no more controls will be applied.

The knowledge about the dynamics is denoted

$$\mathfrak{D} \triangleq \{f_k(\cdot,\cdot,\cdot)\}_{k=0}{}^{N-1}. \qquad (2.6)$$

The knowledge about the measurement system between times i and j, called the measurement program, is

$$\mathfrak{M}_i{}^j \triangleq \{h_k(\cdot,\cdot)\}_{k=i}{}^j, \qquad 1 \le i \le j \le N-1. \quad (2.7)$$

If the lower subscript for X or U is zero and for $\mathfrak{M}$ or Y is one, it will be dropped for simplicity in the sequel.

The joint probability measure induced by the random variables $x_0,v_0,\cdots,v_{N-1},w_1,\cdots,w_k$ will be represented by the symbol

$$\mathcal{S}^k \triangleq dP(x_0,v_0,\cdots,v_{N-1},w_1,\cdots,w_k) \qquad (2.8a)$$

while

$$\mathcal{S}^0 \triangleq dP(x_0,v_0,\cdots,v_{N-1}). \qquad (2.8b)$$

Even though this is an abuse of language, one shall call $\mathcal{S}$, for the sake of conciseness, the "statistics" of the corresponding random variables.

The controller is assumed to be causal, i.e., u_k cannot be a function of $Y_{k+1}{}^{N-1}$ or any of its subsets. It will be also assumed that, when computing the present control u_k, perfect knowledge of all past controls is available. Therefore u_k can be, at most, (Y^k,U^{k-1})-measurable; this is the "real-time" information. The "off-line" kinds of information are given by (2.6)–(2.8).

[1] The reason for limiting the present discussion to fixed end-time problems is the fact that in a free end-time stochastic control problem the definition of the end time, in a probabilistic sense (stopping time), can be done in a number of ways (see, e.g., Kushner [15]) and this is beyond the scope of the present work.

[2] If there is no confusion, the arguments of $Y_i{}^j$ will be dropped for simplicity.

Witsenhausen [30] discussed the various information patterns in stochastic control related to the sharing of the available data among the various "control stations." The aspect to be discussed in the following relates to the information about the future observations. Obviously, a causal control is constrained to be a function of only the observations that have already been obtained; however, as will be seen later, knowledge of the statistics of the future observations plays a key role in stochastic control.

With the above notations one can define the following classes of policies in fixed end-time stochastic control:

1) The *open-loop* policy. In this case the control has the following form:

$$u_k^{OL} = u_k^{OL}[\mathfrak{D},\mathcal{S}^0,\mathcal{L}^N], \qquad k = 0,\cdots,N-1 \quad (2.9)$$

i.e., no measurement knowledge is available for the controller.

2) The *feedback* policy. In this case at every time k, Y^k is available for the computation of the control but no knowledge about the future measurements is available, i.e.,

$$u_k^F = u_k^F[Y^k,U^{k-1}; \mathfrak{D},\mathfrak{M}^k,\mathcal{S}^k,\mathcal{L}^N], \qquad k = 0,\cdots,N-1. \quad (2.10)$$

The open-loop-optimal feedback (OLOF) control (Dreyfus [10], also [24] and [1]) belongs to the feedback class. It assumes that no observations will be made in the future but the observations already made are used in the computation of the control. In subsequent studies this policy has also been called optimal open-loop feedback in [6] and open-loop feedback optimal (OLFO) in [28], [29]. Other control algorithms from the literature that belong to the feedback class are those in [11], [21].

3) The *m-measurement feedback* policy (Curry [9]). In this policy, in addition to using the information from the observations already made, the controller depends on the observation program and the associated statistics for the next m stages as well

$$u_k^{mF} = u_k^{mF}[Y^k,U^{k-1}; \mathfrak{D},\mathfrak{M}^{k+m},\mathcal{S}^{k+m},\mathcal{L}^N] \quad (2.11)$$

when $k > N-1-m$, $k+m$ is replaced by $N-1$.

4) The *closed-loop* policy. This policy (called also "truly feedback" in [10]) incorporates *all* the remaining observation program, i.e., the knowledge that the loop will stay *closed* through the end of the process is fully utilized. Therefore, this policy has the form

$$u_k^{CL} = u_k^{CL}[Y^k,U^{k-1}; \mathfrak{D},\mathfrak{M}^{N-1},\mathcal{S}^{N-1},\mathcal{L}^N], \qquad k = 0,\cdots,N-1. \quad (2.12)$$

Note that the same information about the past appears in (2.10)–(2.12) and they differ only in the availability of knowledge about future observations. Such a controller "causally anticipates" the future measurements.

The above definitions of classes of stochastic control policies are *independent* of the system to be controlled or the performance index. Furthermore, no optimality requirement was used in any of these definitions.

As pointed out by Feldbaum [12], in addition to its effect on the state, the present control might affect the future state uncertainty. This was called the *dual effect* of the control. As can be seen from the above definitions, only the closed-loop control (2.12)[3] takes this effect into account. This follows from the fact that the benefit from the future observations is anticipated via the future observation program and the associated statistics. Loosely speaking, the control might be used to help in getting "better" estimates of the state (or some components of it).

The rigorous definition of dual effect is as follows. Denote by $x_{k,i}$ the ith component of the state at time k and let

$$M^r_{k|k,i}[Y^k(\omega,U^{k-1}),U^{k-1}] \triangleq E\{[x_{k,i}(\omega,U^{k-1}) - E(x_{k,i}|Y^k(\omega,U^{k-1}),U^{k-1})]^r|Y^k(\omega,U^{k-1}),U^{k-1}\} \quad (2.13)$$

be the rth central moment of $x_{k,i}$ conditioned upon the σ-algebra generated by Y^k and U^{k-1}. Then, the control is said to have *no dual effect of order r* ($r \geq 2$) if[4]

$$E[M^r_{k|k,i}|Y^j(\omega,U^{j-1}),U^{k-1}] = E[M^r_{k|k,i}|\bar{Y}^j(\omega)], \quad \text{a.s. } (\omega),\ \forall i,\ \forall U^{k-1},\ \forall j \leq k \quad (2.14)$$

where

$$\bar{Y}^j(\omega) \triangleq \bar{Y}^j(\omega, U^{j-1} = 0) \triangleq \bar{Y}^j(\omega, u_0 = \cdots = u_{j-1} = 0) \quad (2.15)$$

are the measurements of the corresponding autonomous system. In other words, the expected future uncertainty is $\bar{Y}$-measurable, i.e., it is not affected by the control with probability one.

Conversely, if (2.14) does not hold for some $r \geq 2$, i.e., if the control can affect, with nonzero probability, one such central moment, then the control has a *dual effect*. The lack of dual effect has been called neutrality by Feldbaum [12].

There are two aspects in which the closed-loop policy differs from the feedback policy

1) Caution. In a stochastic control problem, due to the inherent uncertainties, the controller has to be "cautious" not to increase the effect of the existing uncertainties on the cost. However, the closed-loop controller, since it "knows" that future observations will be available and corrective actions based upon them will be taken, will exercise less "caution." It was this aspect of a stochastic controller that Dreyfus [10] used when he showed via an example a numerical improvement in the performance of the CL versus OLOF policy.

2) Probing or Active Learning. When the dual effect is present, the control can "help" in learning (estimation) by decreasing the uncertainty about the state. Therefore, the closed-loop control, which takes into account the future observation program and statistics, as pointed out in (2.12), has the capability of *active learning* [26], [27] when the dual effect exists.[5] A feedback controller, even though

[3] The m-measurement feedback does this partially.

[4] Joint moments should also be included but due to the notational complexity they are omitted.

[5] It is implicitly assumed that if a control has a dual effect this helps decrease some uncertainty. This is the case when there are unknown parameters in the system and a closed-loop type control will then "probe" to "learn" them. Otherwise, added "caution" is required.

it "learns" by using the measurements, it does not actively "help" the learning. This learning can be called, therefore, passive, or accidental, and the corresponding control policy passively adaptive, as opposed to the closed-loop control which is actively adaptive [27].

A sequential allocation procedure for a class of resource allocation problems where the decision maker can reduce the uncertainty that affects the results of the allocation has been presented in [7]. This procedure is of the one-measurement feedback class—the present decision depends upon the "value" of the information to be obtained from the next observation.

The closed-loop optimal control sequence is obtained by applying the principle of optimality (Bellman [5]) as follows:

$$J_0^{CLO} = \min_{u_0} E\{\cdots \min_{u_{N-2}} E[\min_{u_{N-1}} E\{\mathcal{L}^N[X^N,U^{N-1}]| Y^{N-1},U^{N-2}\}|Y^{N-2},U^{N-3}]\cdots\}. \quad (2.16)$$

This is the optimal stochastic control that will give the absolute minimum of the cost (2.3) (see also Meier *et al.* [17]).

In the case of an additive cost of the form

$$\mathcal{L}^N[X^N,U^{N-1}] = \mathcal{L}_N(x_N) + \sum_{i=0}^{N-1} \mathcal{L}_k(x_k,u_k) \quad (2.17)$$

the principle of optimality leads to the more usual recursive dynamic programming equation (under suitable existence conditions for the expectations and the minima—see. e.g., [14]).

$$J_k^{CLO} = \min_{u_k} E\{\mathcal{L}_k(x_k,u_k) + J_{k+1}^{CLO}[Y^{k+1},U^k]|Y^k,U^{k-1}\}, \quad k = N-1,\cdots,0 \quad (2.18)$$

with the end condition

$$J_N^{CLO} = E\{\mathcal{L}_N(x_N)|Y^N,U^{N-1}\} \quad (2.19)$$

where J_{k+1}^{CLO} is called the "cost-to-go."

The control u_k is readily seen from (2.16) to be a function of the future observation program and the associated statistics because the expression which is minimized with respect to u_k contains the future observations that are "averaged out" conditioned upon the present information. The active learning feature of the closed-loop control mentioned earlier has another important property in the case of the optimal control. This learning is done to the extent required by the overall performance. Since learning might be conflicting with the control purpose [12], the optimal stochastic control "balances" its learning and control effects such as to minimize the overall cost. This was clearly evident in [26], [27], where the "estimation" and "control proper" parts of the cost were separated and the control was obtained such as to minimize the sum of these two parts.

A type of stochastic control frequently used is the one obtained by applying the so-called certainty equivalence principle (see, e.g., [20], [25], [1]). This control is obtained by first computing the optimal deterministic control for the problem in consideration (i.e., without process noise and with complete and perfect state observations)

$$u_k = \phi_k(x_k) \quad (2.20)$$

and the, replacing x_k by its estimate

$$\hat{x}_{k|k} = E\{x_k|Y^k,U^{k-1}\} \quad (2.21)$$

i.e.,

$$u_k^{CE} = \phi_k(\hat{x}_{k|k}). \quad (2.22)$$

Note that this control is of the *feedback* type rather than *closed-loop* type—it makes use of the available observations but does not account for the future observations.

In a control problem it is said that the *certainty equivalence* (CE) property holds if the closed-loop optimal control has the same form as the deterministic optimal-control with x_k replaced by $\hat{x}_{k|k}$, i.e.,

$$u_k^{CLO} = \phi_k(\hat{x}_{k|k}). \quad (2.23)$$

In general, (2.22) is only an ad hoc control procedure (see e.g., [1], [23], [30]).

The *separation* property (see, e. g., [1], [2], [30]) is a weaker one than the certainty equivalence. The closed loop optimal control has the separation property if it depends on the data only via $\hat{x}_{k|k}$

$$u_k^{CLO} = \psi_k(\hat{x}_{k|k}) \quad (2.24)$$

where the function ψ_k can be different from ϕ_k obtained in the deterministic case (2.20). It can be easily seen that certainty equivalence implies separation but not the other way around. A problem in which the optimal control has the separation property but not the certainty equivalence was studied in [22].

The main result of this paper, to be presented in the next section, is that, for a class of problems, the certainty equivalence property holds if and only if the control has no dual effect (i.e., the system is neutral). The importance of this result lies in the fact that it gives the conditions under which the certainty equivalence control (2.22) coincides with the closed-loop optimal control. The sufficiency of neutrality for certainty equivalence has been suggested by Patchell and Jacobs [18].

III. The Connection Between the Dual Effect of the Control and the Certainty Equivalence Property

Consider the multidimensional system with linear dynamics and additive white, but not necessarily Gaussian noise

$$x_{k+1} = F_k x_k + G_k u_k + v_k \quad (3.1)$$

where F_k and G_k are known matrices of appropriate dimensions

$$Ev_k = 0 \quad (3.2a)$$

$$Ev_k v_j' = V_k \delta_{kj} \quad (3.2b)$$

and the general measurement model

$$y_k = h_k(x_k, w_k) \qquad (3.3)$$

where w_k is the measurement noise with known but arbitrary statistics. The only restrictive assumption on the measurement noise sequence is that it is independent of the process noise. The cost to be minimized is assumed to be quadratic

$$J_0 = E\left\{x_N' Q_N x_N + \sum_{i=0}^{N-1} x_i' Q_i x_i + u_i' R_i u_i\right\} \qquad (3.4)$$

where $Q_i \geq 0$ and $R_i > 0$ (this guarantees the existence of the solution). For the control problem, defined by (3.1)–(3.4), the following result can be stated.

Theorem: The optimal stochastic control (i.e., closed-loop) for the system with linear dynamics (3.1) with white process noise (3.2), measurement equation (3.3), and cost (3.4) has the certainty equivalence property for all $Q_i \geq 0$, $R_i > 0$ if and only if, *the control has no dual effect* of second order, i.e., the updated covariance $\Sigma_{k|k}$ is not a function of the past control sequence U^{k-1}, for all k.

Similarly to (2.14) this requirement can be written as follows:

$$E\{\Sigma_{k|k} | Y^j(\omega, U^{j-1}), U^{k-1}\} = E\{\Sigma_{k|k} | \bar{Y}^j(\omega)\},$$
$$\text{a. s. } (\omega), \ \forall U^{k-1}, \ \forall j \leq k \ \forall k. \qquad (3.5)$$

Proof:

Sufficiency: It will be proved by induction that the optimal cost-to-go will be of the form

$$J_{k+1}^* = E\{x_{k+1}' P_{k+1} x_{k+1} | Y^{k+1}, U^k\} + \alpha_{k+1} \qquad (3.6)$$

where P_{k+1} is a constant nonnegative definite matrix independent of U^{k-1} and α_{k+1} is independent of U^k in the sense of (2.14), i.e.,

$$E\{\alpha_{k+1} | Y^j(\omega, U^{j-1}), U^k\} = E\{\alpha_{k+1} | Y^j(\omega)\},$$
$$\text{a. s. } (\omega), \ \forall U^{k-1}, \ \forall j \leq k. \qquad (3.7)$$

For $k + 1 = N$, it is easily seen that (3.6) is true with

$$P_N = Q_N; \ \alpha_N = 0. \qquad (3.8)$$

Assuming that (3.6) and (3.7) are true and using the stochastic dynamic programming equation (2.18), one has

$$\begin{aligned}
J_k^* &= \min_{u_k} E\{x_k' Q_k x_k + u_k' R_k u_k + J_{k+1}^* | Y^k, U^{k-1}\} \\
&= \min_{u_k} \{E[x_k' Q_k x_k + u_k' R_k u_k | Y^k, U^{k-1}] \\
&\quad + E[x_{k+1}'(u_k) P_{k+1} x_{k+1}(u_k) | Y^k, U^{k-1}]\} \\
&\quad + E\{\alpha_{k+1} | Y^k, U^{k-1}\} \\
&= \min_{u_k} E\{x_k'[Q_k + F_k' P_{k+1} F_k] x_k \\
&\quad + u_k' G_k' P_{k+1} F_k x_k + x_k' F_k' P_{k+1} G_k u_k \\
&\quad + u_k'[R_k + G_k' P_{k+1} G_k] u_k | Y^k, U^{k-1}\} \\
&\quad + \operatorname{tr}(V_k P_{k+1}) + E\{\alpha_{k+1} | \bar{Y}^k\}
\end{aligned} \qquad (3.9)$$

where (3.2), the whiteness of v_k, and its independence of the measurement noises have been utilized. In view of this

$$E\{x_k v_k' | Y^k, U^{k-1}\} = 0 \qquad (3.9a)$$
$$E\{v_k v_k' | Y^k, U^{k-1}\} = V_k. \qquad (3.9b)$$

The optimal u_k^* which minimizes the right-hand side of (3.9) is

$$\begin{aligned}
u_k^* &= -[R_k + G_k' P_{k+1} G_k]^{-1} G_k' P_{k+1} F_k E\{x_k | Y^k, U^{k-1}\} \\
&\triangleq L_k \hat{x}_{k|k}.
\end{aligned} \qquad (3.10)$$

If v_k depends on, for instance, w_k, then (3.9a) would not hold, in general, and u_k^* would not be given by (3.10). Therefore, this independence is needed to have, in general, the CE property hold. Substituting (3.10) into (3.9) and going through some algebraic manipulations, one has J_k^* of the form (3.6) with P_k and α_k satisfying the following recusion relationships:

$$P_k = F_k'[P_{k+1} - P_{k+1} G_k [R_k + G_k' P_{k+1} G_k]^{-1} G_k' P_{k+1}] F_k + Q_k \qquad (3.11)$$

$$\begin{aligned}
\alpha_k &= \operatorname{tr}[F_k' P_{k+1} G_k [R_k + G_k' P_{k+1} G_k]^{-1} G_k' P_{k+1} F_k \Sigma_{k|k} \\
&\quad + V_k P_{k+1}] + E\{\alpha_{k+1} | \bar{Y}^k\} \triangleq g_k(R_k, \Sigma_{k|k}) \\
&\quad + E\{\alpha_{k+1} | \bar{Y}^k\}.
\end{aligned} \qquad (3.12)$$

From (3.8), (3.10), and (3.11), it is readily seen that L_k is the optimum feedback gain as in the deterministic version of the problem (see, e.g., [3]). Subsequently, (3.5) and (3.12) imply that α_k indeed has property (3.7), i.e., it is independent of U^{k-1}, and thus, the induction proof of sufficiency is completed. Therefore, the certainty equivalence property holds when the system is "neutral" in the sense of (3.5), i.e., there is no dual effect of second order.

Necessity: To prove necessity, one has to show that if the optimal control is given by (3.10) for $k = N - 1, \cdots, 0$, this implies that $\Sigma_{k|k}$ is independent of u_j *for all* $j < k$, for all k. Assume that u_k^* is given by (3.10) for $k = N - 1, \cdots, j + 1$. Then

$$J_{j+1}^* = E\{x_{j+1}' P_{j+1} x_{j+1} | Y^{j+1}, U^j\} + \beta_{j+1} \qquad (3.13)$$

where

$$\beta_{j+1} \triangleq \sum_{k=j+1}^{N-1} E\{g_k(R_k, \Sigma_{k|k}) | Y^{j+1}, U^j\} \qquad (3.14)$$

and g_k is as defined in (3.12). We know that u_j^* is obtained by minimizing

$$J_j = E\{x_j' Q_j x_j + u_j' R_j u_j + J_{j+1}^* | Y_j, U^{j-1}\} \qquad (3.15)$$

and it is easy to see that its expression will be given by (3.10) with $k = j$ only if

1) β_{j+1} is independent of u_j.
2) β_{j+1} is a function of u_j but assumes its minimum at u_j^* as given by (3.10) with $k = j$.

However, notice that u_j^* is a function of R_j which does not enter into β_{j+1} according to (3.14); therefore, since we assumed that certainty equivalence holds for all $R_j > 0$, 2) cannot happen. Furthermore, the only way u_j might enter into β_{j+1} is via $\Sigma_{k|k}$, $k \geq j + 1$ and, hence, $\Sigma_{k|k}$ is independent of $u_j, j > k$. Since this is true for $j = 0, \cdots, N-1$, it follows that $\Sigma_{k|k}$ is independent of U^{k-1}, i.e., it is necessary that the control have no dual effect of second order.

This completes the proof.

IV. Discussion and Examples

Consider the following properties that characterize various versions of the stochastic control problem considered in Section III.

1) Linear dynamics with zero-mean white additive process noise independent of the measurement noise.
2) Fixed end-time, quadratic cost.
3) Gaussian process noise.
4) Linear measurement with additive noise.
5) White measurement noise sequence.
6) Gaussian measurement noise.

The Linear-Quadratic-Gaussian (LQG) problem [4], definied by properties 1)–6) above, is well-known to have the certainty equivalence property (Joseph and Tou [14], Gunckel and Franklin [13]). Root [19] has proven that one can relax the Gaussian requirements on both the process and measurement noises, i.e., for the problem defined by 1), 2), 4), 5) the certainty equivalence property holds, too. It was also shown [28] that, if the noises are still Gaussian, but the measurement noise is not white, i.e., 1)–4), 6) the optimal control is certainty equivalent. It is interesting to point out that in continuous time certainty equivalence was proven (see [8]) under the usual LQG assumptions weakened by allowing correlation between the process and measurement noises.

A first new class of problems for which the theorem of Section III can be used to show that the CE property holds is the one defined by 1), 2), 4). Consider the system (3.1) with the measurements

$$y_k = H_k x_k + w_k \tag{4.1}$$

with no restrictions whatsoever on the measurement noise w_k. Similarly to Wonham [31], let

$$x_k \triangleq \tilde{x}_k + \bar{x}_k \tag{4.2}$$

where $\tilde{x}_k$ is the state of the autonomous part of system (3.1), i. e.,

$$\tilde{x}_k = F_k \tilde{x}_k + v_k \tag{4.3}$$

with the corresponding part of the observation

$$\tilde{y}_k = H_k \tilde{x}_k + w_k \tag{4.4}$$

and $\bar{x}_k$ is the state of the forced, noiseless, part of the system (3.1)

$$\bar{x}_k = F_k \bar{x}_k + G_k u_k \tag{4.5}$$

with

$$\bar{y}_k = H_k \bar{x}_k. \tag{4.6}$$

The initial state of the forced part of the system is taken as

$$\bar{x}_0 = 0. \tag{4.7}$$

It can be easily shown by superposition, that, since everything is linear, x_k obeys (3.1) and

$$y_k = \tilde{y}_k + \bar{y}_k. \tag{4.8}$$

Notice that $\bar{x}_k$ and $\bar{y}_k$ are exactly known at every time (because they are U^{k-1}-measurable). Therefore,

$$E\{x_k | Y^k, U^{k-1}\} = E\{\tilde{x}_k | Y^k, U^{k-1}\} + \bar{x}_k = E\{\tilde{x}_k | \tilde{Y}^k\} + \bar{x}_k$$

and

$$\operatorname{cov}\{x_k | Y^k, U^{k-1}\} = \operatorname{cov}\{\tilde{x}_k | \tilde{Y}^k\} \tag{4.10}$$

i.e., the control is neutral and the CE property holds.

A second example will illustrate how in a problem defined by 1), 2), i.e., with partially nonlinear measurement system, one has the certainty equivalence property. Let the system be

$$x_{k+1} \triangleq \begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \end{bmatrix} = \begin{bmatrix} F_k^{11} & F_k^{12} \\ 0 & F_k^{22} \end{bmatrix} x_k + \begin{bmatrix} G_k^1 \\ 0 \end{bmatrix} u_k + \begin{bmatrix} v_k^1 \\ v_k^2 \end{bmatrix} \tag{4.11}$$

and the measurement

$$y_k \triangleq \begin{bmatrix} y_k^1 \\ y_k^2 \end{bmatrix} = \begin{bmatrix} H_k x_k + w_k^1 \\ h_k(x_k^2, w_k^2) \end{bmatrix} \tag{4.12}$$

i.e., the part of the state denoted by x_k^2 is measured via a nonlinear device. However, since x_k^2, is not a function of the past controls—this follows from the particular form of F_k in this problem—it is readily apparent that the control is neutral. In general, for a system which is stabilizable [32] but not completely controllable, if there is a nonlinear observation on only the uncontrollable subspace, then CE holds. Therefore, even in certain systems with nonlinear observations, the certainty equivalence property holds.

V. Conclusion

There is a distinction between feedback and closed-loop policies in stochastic control and it is only the closed-loop policy that takes into account the possible estimation benefit to be derived from future observations. This leads to the active learning feature of the closed-loop control policy that can be utilized when the control has a dual effect. There is an intimate connection between the dual effect of the control and the certainty equivalence property. The class of problems for which it was previously known that the certainty equivalence property (and, hence, separation) holds has been expanded.

Acknowledgment

The authors would like to thank Dr. A. Segall for valuable comments and criticism and Prof. R. Sivan for stimulating discussions.

References

[1] M. Aoki *Optimization of Stochastic Systems*. New York: Academic, 1967.

[2] K. J. Åström, *Introduction to Stochastic Control Theory*. New York: Academic, 1970

[3] M Athans and P. Falb, *Optimal Control*. New York: McGraw-Hill, 1966.

[4] *Special Issue on Linear-Quadratic-Gaussian Problem, IEEE Trans. Automat. Contr.*, vol. AC-16, Dec. 1971.

[5] R. Bellman, *Adaptive Control Processes: A Guided Tour*. Princeton, N.J.: Princeton Univ. Press, 1961.

[6] Y. Bar-Shalom and R. Sivan, "On the optimal control of discrete-time linear systems with random parameters," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 3–8, Feb. 1969.

[7] Y. Bar-Shalom, R. E. Larson, and M A. Grossberg, "Application of stochastic control theory to resource allocation under uncertainty," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 1–7, Feb. 1974.

[8] J. S. Boland, W. B Douglass, N. P Dwivedi, and W. G. Hopkins, "Filtering and optimization of linear stochastic systems with cross-correlated noise," Dep. Elec. Eng., Auburn Univ., Auburn, Ala., Tech. Rep. 20, Apr. 1969.

[9] R. E. Curry, "A new algorithm for suboptimal stochastic control," *IEEE Trans Automat. Contr.*, vol. AC-14, pp. 533–536, Oct. 1969.

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

[11] J. G. Deshpande, T. N. Upadhyay, and D. G. Lainiotis, "Adaptive control of linear stochastic systems," *Automatica*, vol. 9, pp. 107–115, 1972.

[12] A. A. Feldbaum, *Optimal Control Systems*. New York: Academic Press, 1965.

[13] T. L Gunckel and G. F. Franklin, "A general solution for linear, sampled-data control," *Trans. ASME* (J. Basic Eng.), ser. D., vol. 85, pp. 197–203, June 1963.

[14] P. D. Joseph and J. T. Tou, "On linear control theory," *AIEE Trans. Appl. Ind.*, vol. 80, pp. 193–196, 1961.

[15] H. Kushner, *Stochastic Stability and Control*. New York: Academic, 1967

[16] ——, *Introduction to Stochastic Control*. New York: Holt, Rinehart and Winston, 1971.

[17] L. Meier, R. E. Larson, and A. J. Tether, "Dynamic programming for stochastic control of discrete systems," *IEEE Trans. Automat. Contr.* (*Special issue on Linear Quadratic Gaussian Problem*), vol. AC-16, pp. 767–775, Dec. 1971.

[18] J. W. Patchell and O. L. R. Jacobs, "Separability, neutrality and certainty equivalence," *Int. J. Contr.*, vol 13, pp. 337–342, 1971.

[19] J. G. Root, "Optimum control of non-Gaussian linear stochastic systems with inaccessible state variables," *SIAM J. Contr.*, vol. 7, pp. 317–323, May 1969.

[20] H. Simon, "Dynamic programming under uncertainty with a quadratic function," *Econometrica*, vol. 24, pp. 74–81, 1956.

[21] G. N. Saridis an R. N. Lobbia, "Parameter identification and control of linear discrete-time systems," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 52–60, Feb. 1972.

[22] J. Speyer, J. Deyst, and D. Jacobson, "Optimization of stochastic linear systems with additive measurement and process noise using exponential performance criteria," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 358–366, Aug. 1974.

[23] S. C. Sworder and D. D. Sworder, "Feedback estimation systems and the separation principle of stochastic control," *IEEE Trans. Automat. Contr.* (Short Paper), vol. AC-16, pp. 350–354, Aug. 1971.

[24] H. A. Spang, "Optimum control of an unknown linear plant using Bayesian estimation of the error," *IEEE Trans. Automat. Contr.*, vol. AC-10, pp. 80–83, Jan. 1965.

[25] H. Theil, "A note on certainty equivalence in dynamic planning," *Econometrica*, vol. 25, pp. 346–349, 1957.

[26] E. Tse, Y. Bar-Shalom, and L. Meier, III, "Wide-sense adaptive dual control of stochastic nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 98–108, Apr. 1973.

[27] E. Tse and Y. Bar-Shalom, "An actively adaptive control for linear systems with random parameters via dual control approach," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 109–117, Apr. 1973.

[28] E. Tse, "On the optimal control for linear systems with incomplete information," Electronic Systems Lab., Mass. Inst. Technol., Cambridge Mass., Rep ESL-R-412, Jan. 1970.

[29] E. Tse and M. Athans "Adaptive stochastic control for a class of linear systems," *IEEE Trans. Automat Contr.* vol. AC-17, pp. 38–51, Feb. 1972.

[30] H. S. Witsenhausen, "Separation of estimation and control for discrete-time systems," *Proc. IEEE*, vol. 59, pp. 1557–1566, Nov. 1971.

[31] W. M. Wonham, "On the separation theorem of stochastic control," *SIAM J. Contr.*, vol. 6, no. 2, pp. 312–326, 1968.

[32] ——, "On pole assignment in multiinput controllable linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 660–665, Dec. 1967.

Yaakov Bar-Shalom (S'63–M'66) for a photograph and biography see page 7 of the February issue of this TRANSACTIONS.

Edison Tse (M'70) was born in Kwangtung, China, on January 21, 1944. He received the B.S. and M.S. degrees, simultaneously, in 1967 and the Ph.D. degree in 1970, all in electrical engineering, from the Massachusetts Institute of Technology, Cambridge.

During 1966–1967 and 1968–1969 he was a Teaching Assistant at M.I.T., teaching graduate courses in stochastic system and optimal control theory, and during 1967 and 1968 he was a Research Assistant at M.I.T., doing research in nonlinear filtering. He was a consultant for the State Street Bank of Boston from 1968 to 1969, where he applied optimal control to banking problems. Since 1969 he has been with Systems Control, Inc., Palo Alto, Calif., where he is now a Senior Research Engineer. Starting January 1974, he is also a Lecturer in the Department of Engineering-Economic Systems at Stanford University, Stanford, Calif. His current research interests include opitmal control theory, dynamic allocation, and stochastic estimation, identification, and control.

Dr. Tse received the 1973 Donald P. Eckman Award for outstandin gachievement in the field of automatic control from the American Automatic Control Council. He is a member of Sigma Xi, Eta Kappa Nu, and Tau Beta Pi. He is also Chairman of the Technical Committee on Stochastic Control and an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.