Discrete-Time Interrupted Stochastic Control Processes

J. H. Eaton*

Department of Electrical Engineering University of California, Berkeley, Calif.

Submitted by Lotfi Zadeh

This paper is concerned with the determination of optimal policies for applying inputs to discrete-time processes which are interrupted in the sense that at each time \( t, t = 1, 2, \ldots \), there is a probability \( p_t \) (called the probability of interruption) that the state vector of the system cannot be observed. The policy is to be optimal with respect to a specified loss functional. It is shown that the optimal policy for controlling such processes can be expressed in terms of functional equations obtained by the usual procedures of dynamic programming. An interrupted process involving a linear system and quadratic loss functional is examined in detail and the optimal policy as well as the expected cost of the process under the optimal policy is expressed in analytic form. The optimal policy is found to be independent of the probability of interruption. Asymptotic properties of the optimal policy and the expected cost of the process under the optimal policy are also obtained.

I. INTRODUCTION

The class of sequential control processes considered in this paper can be described in the following fashion. Let \( x_n, n = 0, 1, 2, \ldots \) denote the state vector of a system \( B \) at time \( n \). We assume that \( B \) is characterized by the relation

\[
x_{n+1} = g_n(x_n, y_n, r_n), \quad x_0 = c
\]

(1.1)

Here \( g_n \) is a known function, \( y_n \) is the input applied by an observer at time \( n \), and the \( r_n \) are independent, identically distributed random vectors. The \( r_n \) are not assumed to be observable. Note that if the \( r_n \) were not independent the \( x_n \) would not qualify as the states of \( B \). However, the results derived later in the paper can be applied with minor modifications to the case where the \( r_n \) are not independent.

The process is interrupted in the sense that at each time \( n, (n = 1, 2, \ldots, N) \), there is a probability \( p_n \) that the state of the system cannot be observed.

* This work was performed while the author was engaged as a consultant to the Mathematics Division of The RAND Corporation.
Following Bellman and Kalaba [1], processes of this type will be called interrupted stochastic control processes. The probability of interruption $p_n$ may or may not be controllable by the observer. A partial list of the possibilities concerning the nature of the mechanism governing the probability of interruption might be:

(a) $p_n$ is a fixed scalar $p_0$ not controllable by the observer.

(b) $p_n$ is controllable by the observer in a sense to be defined later.

(c) The probability of observing $x_n$ may be increased at the expense of introducing additional noise into the system at time $n$. More specifically we shall assume that a decrease in $p_n$ changes (1.1) to

$$x_{n+1} = g_n[x_n, y_n, r_n, s_n(p_n)], \quad x_0 = c \quad (1.2)$$

Here we make the additional assumption that the $s_n(p_n)$ are independent normally distributed random vectors with mean zero and covariance matrix $\sigma^2(p_n)$.

(d) $p_n$ is a fixed vector $p_0$ whose $i$th component is the probability that the $i$th component of the state vector $x_n$ will not be observed.

In the event that the observer is able to influence the probability of interruption $p_n$, as in case (b) above, we will define a system $B^*$ which is characterized by the relations

$$x_{n+1} = g_n(x_n, y_n, r_n), \quad x_0 = c \quad (1.1)$$

and

$$p_{n+1} = h(p_n), \quad p_0 = 0 \quad (1.3)$$

Here $p_n$ is the input applied by the observer at time $n$ in order to influence the probability of being able to observe $x_{n+1}$, and $h$ is a known function. It will be convenient to denote the composite input $(y_n, p_n)$ by $v_n$.

For a process of specified length $N$ (in the sense that $n = 0, 1, 2, \ldots, N$) it is desired that feedback control as represented by $v_n$ be applied in such a manner as to minimize an expected cost associated with the process. The minimization is to be carried over the class of allowable policies to be used in the sequential choice of the inputs $v_n$. The class of allowable policies will be defined later. The cost of the process is assumed to be measured by a specified scalar valued function

$$\varphi(x_N, v_0, v_1, \ldots, v_{N-1}) \quad (1.4)$$

where $x_N$ is the terminal state of the system and $v_n$ is the input applied by the observer at time $n$. Thus the loss functional for the process is

$$E[\varphi(x_N, v_0, v_1, \ldots, v_{N-1})]. \quad (1.5)$$
The expectation $E$ is to be taken over the $r_n$ and the random variables governing the stochastic interruptions.

Processes of this type have previously been considered by Bellman and Kalaba [1]. (For a discussion without the feature of interruption see Bellman [2].) The technique of analysis used by Bellman and Kalaba leads to functional equations that have an implicit structure of a type not amenable to solution by conventional iterative techniques of dynamic programming. Bellman and Kalaba suggested the use of a method of successive approximations for computational purposes and cited certain special cases for which explicit analytic results can be obtained.

Using the information pattern of the process, which is discussed in the next section, the optimal policy for controlling an interrupted process can be written in terms of functional equations that can be obtained by the usual techniques of dynamic programming. These equations can be solved explicitly in the case of linear systems with quadratic loss functionals.\(^1\) The results presented in the sequel can also be extended to fairly broad classes of continuous-time interrupted stochastic control processes.

II. Information Pattern

The information pattern of a process has previously been considered by Bellman [6]. The definition given here will differ slightly from that used by Bellman, but the concept is essentially the same. Roughly speaking, let us consider an observer viewing a process over which he has some form of control which he wishes to exert in an optimal fashion. It is clear that at a given time $t$, whatever action the observer takes must be based on data that is available to him at that time. In this sense, the class of policies based on all the available data is complete. It will usually turn out that there are classes of policies using only proper subsets of the available data that are also complete. In other words, insofar as the minimization of a particular loss functional is concerned, the observer can usually do as well by considering only a subset of the available data. A subset of the available data constitutes an information pattern if it does not contain a proper subset over which a complete class of policies can be defined. For the purposes of this paper the information pattern of a process can be defined more precisely in terms of the ordering of allowable policies under a proper loss functional. The terms allowable policy and proper loss functional are defined below.

\(^1\) For treatments of such processes without the feature of interruption see, for example, Bellman [2], Adorno [3], Freimer [4], and Kalman and Koepcke [5].
DEFINITION. A policy \( \pi_{n,N} = (\pi_n, \pi_{n+1}, \ldots, \pi_N) \) is a mapping of the data available at each time \( t, t = n, n + 1, \ldots, N \), into an input \( v_t \) to be applied to the system. Here \( \pi_t \) is a mapping of the data available at time \( t \) into an input to be applied to the process at time \( t \). A policy \( \pi_{n,N} \) is an allowable policy if all values of \( v_t, t = n, n + 1, \ldots, N \), assumed under it are admissible. (An input \( v_t \) may fail to be admissible in the case where constraints are imposed on the inputs.)

DEFINITION. A scalar valued function \( \varphi \) is a proper loss functional for a process of length \( N \) if and only if at each time \( t, t = 0, 1, 2, \ldots, N - 1 \), its expected value (over the \( r_n \) and the random variables governing the interruptions) conditioned with respect to \( S_t \), the data available at time \( t \), can be used to produce an ordering of allowable policies \( \pi_{t,N-1} \). In other words, \( \varphi \) is a proper loss functional if for any two allowable policies \( \pi'_{n,N-1} \) and \( \pi''_{n,N-1} \), \( \varphi \) can be used to determine whether \( \pi'_{n,N-1} \) is better than, equivalent to, or inferior to \( \pi''_{n,N-1} \) according as \( E(\varphi \mid S_n, \pi'_{n,N-1}) \) is less than, equal to, or greater than \( E(\varphi \mid S_n, \pi''_{n,N-1}) \).

A subset \( I_n \) of the available data \( S_n \) constitutes an information pattern of the process at time \( n \) if the expected value of \( \varphi \) produces the same ordering of allowable policies \( \{\pi_{n,N-1}\} \) when conditioned with respect to \( I_n \) as it does when conditioned with respect to \( S_n \), and if there is no proper subset of \( I_n \) that has this property. In many cases it is apparent on intuitive grounds whether or not a particular subset of the available data constitutes an information pattern of the process. This is particularly true in the context of the functional equations of dynamic programming. The utility of the concept of information pattern lies mainly in the fact that the principle of optimality can be stated in terms of it in exactly the same manner as it is usually stated in terms of the state of a system.

The concept of information pattern is intimately related to the notion of process state used by Zadeh [7]. (See also his remark on the paper by Bellman and Kalaba [1], ref. 8.) In general there exists a many-to-one mapping from \( I_n \) into \( I_n^* \) such that when the expected value of \( \varphi \) is conditioned with respect to \( I_n^* \) the same ordering of allowable policies is produced as when conditioned with respect to \( I_n \). Essentially, if there does not exist a many-to-one mapping of \( I_n^* \) into \( I_n^{**} \) such that \( I_n^{**} \) produces the same ordering of allowable policies as \( I_n^* \), \( I_n^* \) corresponds to the process state.

III. FORMULATION IN DYNAMIC PROGRAMMING TERMS

Let \( I_n \) denote the information pattern of a discrete-time process at time \( n \). The process is assumed to start at time \( n = 0 \) and to terminate at time \( n = N \). \( N \) will be referred to as the length of the process. Let \( \pi_n \) denote a mapping of
the information pattern of the process at time \( n \) into an admissible input \( v_n \). \( \pi_n \) will be called the \( n \)th component of the allowable policy

\[
\pi_{0,N-1} = (\pi_0, \pi_1, \ldots, \pi_{N-1})
\]

to be used in the sequential choice of the inputs \( v_n, n = 0, 1, \ldots, N - 1 \). The expected cost of a process of length \( N \) will depend on both the initial information pattern \( I_0 \) and the policy used for the sequential choice of the inputs \( v_n \). The expected cost using a policy \( \pi_{0,N-1} \) of a process of length \( N \) whose initial information pattern is \( I_0 \) will be denoted by

\[
C_N(I_0, \pi_{0,N-1}) = E[\varphi(x_N, v_0, v_1, \ldots, v_{N-1}) | I_0, \pi_{0,N-1}].
\]  
(3.1)

Thus

\[
F_N(I_0) = \min_{\pi_{0,N-1}} C_N(I_0, \pi_{0,N-1})
\]  
(3.2)

is the expected cost using an optimal policy for a process of length \( N \) when the initial information pattern is \( I_0 \). In analyzing processes of specified length \( N \), it will be convenient to state the principle of optimality in the following form:

For a process of specified length \( N \), an optimal policy \( \pi_{0,N-1}^* \) has the property that whatever the previous information patterns and previous inputs were, the input chosen at time \( N - 1 \) using \( \pi_{N-1}^* \) must be optimal with respect to a one-stage process (i.e., a process of length one starting at time \( N - 1 \)) whose initial information pattern is \( I_{N-1} \).

Consider the problem faced by the observer at time \( N - 1 \), one stage before the termination of the process. Using the principle of optimality we see that if the observer's policy is to be optimal, the input \( v_{N-1} \) chosen using the \((N - 1)\)th component of his policy must be optimal with respect to a one-stage process whose initial information pattern is \( I_{N-1} \), thus for a specific \( I_{N-1} \) the input chosen using the last component of the observer's policy must minimize

\[
C_1(I_{N-1}, v_{N-1}) = E[\varphi(x_N, v_0, v_1, \ldots, v_{N-1}) | I_{N-1}]
\]

\[
= E[\{g_{N-1}(x_{N-1}, r_{N-1}, y_{N-1}), v_0, v_1, \ldots, v_{N-1}\} | I_{N-1}] \tag{3.3}
\]

with respect to \( v_{N-1} \). Here \( v_{N-1} \) is not to be considered a random variable since \( C_1(I_{N-1}, v_{N-1}) \) is a function of \( I_{N-1} \) and \( v_{N-1} \) only, and hence the minimizing value for \( v_{N-1} \) is determined by \( I_{N-1} \). The minimization of \( C_1(I_{N-1}, v_{N-1}) \) with respect to \( v_{N-1} \) for all possible \( I_{N-1} \) will yield the optimal \( v_{N-1} \) as a function of \( I_{N-1} \). Hence the minimization will yield \( \pi_{N-1}^* \), the last component of the optimal policy and also

\[
F_1(I_{N-1}) = \min_{v_{N-1}} E[\varphi(x_N, v_0, v_1, \ldots, v_{N-1}) | I_{N-1}].
\]  
(3.4)
$F_N(I_0)$ can now be written in terms of a loss function defined at time $N - 1$ rather than at time $N$, i.e.,

$$F_N(I_0) = \min_{\pi_{0,N-2}} E[F_1(I_{N-1}) \mid I_0]. \quad (3.5)$$

Therefore the process can be considered an $(N - 1)$-stage process where $E[F_1(I_{N-1})]$ is the loss functional to be minimized. $F_2(I_{N-2})$ and $\pi_{N-3}$ can be determined from $E[F_1(I_{N-1})]$ in the same manner as $F_1(I_{N-1})$ and $\pi_{N-2}$ were determined from $E[\varphi(x_N, v_0, v_1, \ldots, v_{N-1})]$. Thus

$$F_2(I_{N-2}) = \min_{v_{N-2}} E[F_1(I_{N-1}) \mid I_{N-2}]. \quad (3.6)$$

More generally, $F_n(I_{N-n})$ can be determined from $F_{n-1}[I_{N-(n-1)}]$ via the relation

$$F_n(I_{N-n}) = \min_{v_{N-n}} E[F_{n-1}(I_{N-(n-1)}) \mid I_{N-n}], \quad n = 1, 2, \ldots, N \quad (3.7)$$

Hence the optimal policy $\pi_{0,N-1}$ and $F_N(I_0)$, the expected cost of an $N$-stage process under an optimal policy when the initial information pattern is $I_0$ can be determined from functional equations which are obtained by the usual iterative procedure of dynamic programming.

**IV. An Example**

In general, straightforward application of the procedure sketched in the previous section entails the storage of excessively large amounts of data. However, for certain classes of processes, such as those involving linear systems and quadratic loss functionals, analytic results can be obtained. We will consider such a process. To be more specific, consider a discrete-time process in which at each time $n$, $n = 0, 1, 2, \ldots$ the state of the system to be controlled is determined by the recurrence relation.

$$x_{n+1} = Ax_n + r_n + Hy_n. \quad x_0 = c \quad (4.1)$$

Here $x_n$ is an $s \times 1$ vector describing the state of the system at time $n$, $A$ is a constant $s \times s$ matrix, $y_n$ is an $s \times 1$ vector representing the input to the system at time $n$, $H$ is a constant nonsingular $s \times s$ matrix, and $\{r_n\}$ is a sequence of independent random vectors with a common known distribution. For simplicity it is assumed that the vectors $r_n$ have mean zero and covariance matrix $\sigma^2 I$. The process is interrupted in the sense that at each time $n$, $n = 1, 2, \ldots$ there is a fixed probability $p > 0$ that the state of the system
cannot be observed. For a process of specified length \( N \), it is desired to determine a policy to be used in the sequential choice of the inputs \( y_n \) which minimizes

\[
C(I_0, \pi_{0,N-1}) = E\left\{ \left[ (x_N, B_0x_N) + \sum_{n=0}^{N-1} (y_n, G_y) \right] \left| I_0 \right. \right\}. \tag{4.2}
\]

Here \((x_N, B_0x_N)\) is a symmetric positive semidefinite quadratic form, \( x_N \) is the terminal state of the system, \( y_n \) is the input applied to the system at time \( n \), and \((y_n, G_y)\) is a symmetric positive definite quadratic form. The expectation \( E \) is over the \( r_n \) and the random variables governing the interruptions, and the minimization is over the class of allowable policies defined in Section II. In what follows we shall use the technique sketched in the previous section to find an optimal policy for controlling the process. The same technique can be applied to the more general problem of finding an optimal policy for controlling a time-varying system with a loss functional of the form

\[
E\left\{ \sum_{n=0}^{N} [(x_n, B_0x_n) + (y_n, G_y)] \left| I_0 \right. \right\}. \tag{4.3}
\]

For the sake of simplicity, we shall confine our attention to the time-independent terminal control problem.

Following the notation of the previous section let us denote the expected cost under an optimal policy of a process of length \( N \) when the initial information pattern is \( I_0 \) by

\[
F_N(I_0) = \min_{\pi_{0,N-1}} E\left\{ \left[ (x_N, B_0x_N) + \sum_{n=0}^{N-1} (y_n, G_y) \right] \left| I_0 \right. \right\}. \tag{4.4}
\]

Here \( \pi_{0,N-1} \) denotes the policy used in the sequential choice of the inputs \( y_n \) at times \( n = 0, 1, 2, \ldots, N - 1 \). Let us further define

\[
F_n(I_{N-n}) = \min_{\pi_{N-n,N-1}} C_n(I_0, \pi_{N-n,N-1})
\]

\[
= \min_{\pi_{N-n,N-1}} E\left\{ \left[ (x_N, B_0x_N) + \sum_{k=0}^{N-1} (y_k, G_y_k) \right] \left| I_0 \right. \right\}. \tag{4.5}
\]

Thus \( F_n(I_{N-n}) \) is the expected cost of a process of length \( N \) using an optimal policy at times \( n = N - n, \ldots, N - 1 \) when the information pattern at time \( N - n \) is \( I_{N-n} \). Using the results of the previous section we see that

\[
F_{n+1}(I_{N-(n+1)}) \min_{y_{N-(n+1)}} E[F_n(I_{N-n}) \mid I_{N-(n+1)}]. \tag{4.6}
\]
and that

$$F_N(I_0) = \min_{\pi_{0,N-n}} E[F_n(I_{N-n}) \mid I_0].$$ (4.7)

Equations (4.4)-(4.7) will form the basis for our analysis of the process.

In the following it will be convenient to use the notation

$$\mathcal{E}_n(\psi_m) = E[\psi_m \mid I_n]$$ (4.8)

where $\mathcal{E}_n(\psi_m)$ denotes the expected value of a variable $\psi_m$, conditioned with respect to $I_n$. When the time subscript $m$ on the variable $\psi_m$ is the same as the time subscript $n$ on $\mathcal{E}_n$, the subscript $n$ will be omitted from $\mathcal{E}_n$.

The optimal policy for controlling the process under consideration follows as a corollary from the following lemma.

**Lemma.** The expected value of $F_n(I_{N-n})$ defined in Eq. (4.5) can be expressed in the form:

$$E[F_n(I_{N-n})] = E \left[ (x_{N-n}, B_n x_{N-n}) + \sum_{k=0}^{N-(m+1)} (y_k, G y_k) + f_m \right]$$ (4.9)

where $B_n$ is a symmetric positive semidefinite matrix, and $f_m$ is independent of the policy used in choosing the inputs $y_m$ prior to time $N - n$.

**Proof.** To prove the lemma we shall use induction on $n$. Clearly (4.9) is true for $n = 0$ since

$$E[F_0(I_N)] = E \left[ (x_N, B_0 x_N) + \sum_{n=0}^{N-1} (y_n, G y_n) \right]$$ (4.10)

is of the required form. Suppose the lemma is true for $n = m$. Then, as far as the observer is concerned, he can consider the process to be an $(N - m)$-stage process where $E[F_m(I_{N-m})]$ is the loss functional to be minimized. Using the principle of optimality we see that if the observer’s policy is to be optimal, the value of $y_{N-(m+1)}$ chosen using his policy when the information pattern is $I_{N-(m+1)}$ must be optimal with respect to a one-stage process whose initial information pattern is $I_{N-(m+1)}$ and where $E[F_m(I_{N-m})]$ is the loss functional to be minimized. Consequently the observer’s policy must minimize

$$Q_{m+1} = E \left\{ (x_{N-m}, B_m x_{N-m}) + \sum_{k=0}^{N-(m+1)} (y_k, G y_k) + f_m \mid I_{N-(m+1)} \right\}$$ (4.11)

with respect to $y_{N-(m+1)}$. Using Eq. (4.1) to write $x_{N-m}$ in terms of $x_{N-(m+1)}$, $r_{N-(m+1)}$, and $y_{N-(m+1)}$ and splitting the sum $\sum_{k=0}^{N-(m+1)} (y_k, G y_k)$ into

$$(y_{N-(m+1)}, G y_{N-(m+1)}) + \sum_{k=0}^{N-(m+2)} (y_k, G y_k)$$
we have
\[ Q_{m+1} = E \left[ \left( [Ax_N^{(m+1)} + r_N^{(m+1)} + Hy_N^{(m+1)}] \right) \right. \\
\left. + B_m[Ax_N^{(m+1)} + r_N^{(m+1)} + Hy_N^{(m+1)}] \\
+ (y_N^{(m+1)}, Gy_N^{(m+1)}) + \sum_{k=0}^{N-(m+2)} (y_k, Gy_k) + f_m \right| I_{N-(m+1)} \right], \tag{4.12} \]

Here \( y_N^{(m+1)} \) is not to be considered a random variable since its minimizing value is determined by \( I_{N-(m+1)} \), [see Eq. (3.3) et. seq.). \( Q_{m+1} \) can be expressed after expansion as
\[
Q_{m+1} = \varepsilon(Ax_N^{-(m+1)}, B_mAx_N^{-(m+1)}) \\
+ 2(Hy_N^{-(m+1)}, B_mAe[x_N^{-(m+1)}]) + (y_N^{-(m+1)}, [G + H'B_mH])y_N^{-(m+1)} \\
+ \varepsilon(r_N^{-(m+1)}, B_mr_N^{-(m+1)}) + \sum_{k=0}^{N-(m+2)} (y_k, Gy_k) + E[f_m \mid I_{N-(m+1)}]. \tag{4.13} \]

Here we have taken advantage of the fact that \( \varepsilon(y_N^{-(m+1)}, r_N^{-(m+1)}) \) and \( \varepsilon(x_N^{-(m+1)}, r_N^{-(m+1)}) \) are equal to zero, and have combined like terms in \( y_N^{-(m+1)}. \) (Note that \( y_N^{-(m+1)} \) and \( x_N^{-(m+1)} \) are correlated with \( r_N^{-(m+2)} \) but not with \( r_N^{-(m+1)} \) by virtue of the independence of the \( r_n \)’s.) \( Q_{m+1} \) is a quadratic function of \( y_N^{-(m+1)}. \) Straightforward minimization of \( Q_{m+1} \) with respect to \( y_N^{-(m+1)} \) yields
\[
y_N^{-(m+1)} = -(H'B_mH + G)^{-1} H'B_mA \varepsilon[x_N^{-(m+1)}] \tag{4.14} \]
as the optimal value for \( y_N^{-(m+1)}. \) Substitution of the above expression for \( y_N^{-(m+1)} \) into Equation 4.13
\[
F_{m+1}[I_{N-(m+1)}] = \varepsilon(Ax_N^{-(m+1)}, B_mAx_N^{-(m+1)}) \\
- 2(H'H_BHH + G)^{-1} H'B_mA \varepsilon[x_N^{-(m+1)}], B_mA \varepsilon[x_N^{-(m+1)}]) \\
+ ([H'H_BHH + G]^{-1} H'B_mB_mA \varepsilon[x_N^{-(m+1)}], H'B_mB_mA \varepsilon[x_N^{-(m+1)}]) \\
+ \varepsilon(r_N^{-(m+1)}, B_mr_N^{-(m+1)}) + E[f_m \mid I_{N-(m+1)}] + \sum_{k=0}^{N-(m+2)} (y_k, Gy_k). \tag{4.15} \]

Noting that the second and third terms on the right-hand side of Eq. (4.15) can be combined, and observing that for an arbitrary matrix \( M \)
\[
-(M \varepsilon[x_n], \varepsilon[x_n]) = (M[\varepsilon(x_n) - x_n], [\varepsilon(x_n) - x_n]) - \varepsilon(Mx_n, x_n), \tag{4.16} \]
Eq. (4.15) can be reduced to
\[
F_{m+1}[I_{N-(m+1)}] = \mathcal{E}(x_{N-(m+1)}, A'B_m[I - H(H'B_mH + G)^{-1}H'B_m]Ax_{N-(m+1)}) \\
+ \sum_{k=0}^{N-(m+2)} (y_k, Gy_k) + \mathcal{E}(e_{m+1}, A'B_mH(H'B_mH + G)^{-1}H'B_mAe_{m+1}) \\
+ \mathcal{E}(r_{N-(m+1)}, B_m^r x_{N-(m+1)}) + E[f_m | I_{N-(m+1)}] \tag{4.17}
\]
Here \( e_{m+1} = \mathcal{E}(x_{N-(m+1)} - x_{N-(m+1)}) \) is the error in the observer's minimum variance estimate of the state of the system at time \( N - (m + 1) \).

The expected value of the last three terms on the right-hand side of Eq. (4.17) does not depend on the policy used in the choice of the inputs \( y_k \) prior to time \( N - (m + 1) \), therefore the expected value of \( F_{m+1}[I_{N-(m+1)}] \) can be written
\[
E[F_{m+1}[I_{N-(m+1)}]] = E[(x_{N-(m+1)}, B_m^{r+1}x_{N-(m+1)}) + \sum_{k=0}^{N-(m+2)} (y_k, Gy_k) + f_{m+1}]
\tag{4.18}
\]
where \( f_{m+1} \) is independent of the policy used in choosing the inputs prior to time \( N - (m + 1) \) and
\[
B_{m+1} = A'B_m[I - H(H'B_mH + G)^{-1}H'B_mA] \\
= A'B_m(HG^{-1}H'B_m + I)^{-1}H'B_mA \tag{4.19}
\]
is a real symmetric positive semidefinite matrix. Hence the validity of the assertion is established. Note the fact that \( F_{m+1}[I_{N-(m+1)}] \) is quadratic in \( x_{N-(m+1)} \) is a result of the linearity of the system.

**Corollary.** The inputs under the optimal policy for the process are given by
\[
y_{N-(m+1)} = -(H'B_mH + G)^{-1}H'B_mA \mathcal{E}(x_{N-(m+1)}) \quad m = 0, 1, \ldots, N - 1 \tag{4.20}
\]
where \( B_m \) is determined by the recurrence relation
\[
B_{m+1} = A'B_m(HG^{-1}H'B_m + I)^{-1}H'B_mA \quad m = 0, 1, 2, \ldots, N - 1, \quad B_0 = B_0 \tag{4.21}
\]
The proof of this corollary follows immediately upon inspection of Eq. (4.14) and (4.19).

The information pattern of the process at time \( n = N - (m + 1) \) includes the description of the process as given in Eq. (4.1) and (4.2) as well as the number of stages of the process remaining at time \( n = N - (m + 1) \). This
part of the information pattern is deterministically related to \( I_0 \), the information pattern at time \( n = 0 \). If we suppress that part of the information pattern which is deterministically related to \( I_0 \), the set of data which constitutes \( I_{N-(m+1)} \) is just that set which is useful in obtaining a minimum variance estimate of \( x_{N-(m+1)} \), the state of the system at time \( n = N - (m + 1) \). (Note that \( \delta(x_{N-(m+1)}) = E[x_{N-(m+1)} | I_{N-(m+1)}] \) is the only quantity in (4.20) which depends on data which are not deterministically related to \( I_0 \).) For the case under consideration in which the \( r_n \) are uncorrelated the set of available data which is useful in obtaining a minimum variance estimate of the state of the system at a given time consists of the last observed state together with the inputs applied since the last observation. As an illustration of the relation between information pattern and process state note that \( \delta(x_{N-(m+1)}) \) qualifies as a process state in the sense described at the end of Section II.

In the following we shall sketch, without giving proof, some of the more general results which can be obtained by the above techniques in the case where the \( r_n \) are correlated. Referring to the lemma, replace \( x_{N-n} \) by \( z_{N-n} \) where

\[
z_{N-n} = x_{N-n} + A^{-n} \sum_{k=0}^{n-1} A^k r_{N-(k+1)}.
\]  

(4.22)

The proof of the lemma remains the same until we reach Eq. (4.13), where it can no longer be assumed that \( \delta(x_{N-(m+1)}, r_{N-(m+1)}) \) and \( \delta(y_{N-(m+1)}, r_{N-(m+1)}) \) are equal to zero since \( r_{N-(m+1)} \) is now correlated with \( x_{N-(m+1)} \) and \( y_{N-(m+1)} \). Carrying these additional terms along will establish the validity of the lemma with \( x_{N-n} \) replaced by \( z_{N-n} \). Consequently, when the \( r_n \) are correlated, the inputs under the optimal policy are given by

\[
y_{N-(m+1)} = - (H'B_mA \delta(z_{N-(m+1)}))^{-1} H'B_mA \delta(z_{N-(m+1)}) \quad m = 0, 1, \ldots, N - 1
\]  

(4.23)

where \( B_m \) is given by the recurrence relation in Eq. (4.21).

Note that in the above we have assumed that the matrix \( A \) is nonsingular. In the event that \( A \) is singular the optimal policy can be expressed in a form that does not depend on the existence of \( A^{-1} \). The form in Eq. (4.22) and (4.23) was chosen here because of its similarly to the optimal policy for the case where the \( r_n \) are not correlated which is given in Eq. (4.20). Supressing the deterministic part of the information pattern, the set of data that constitutes \( I_{N-(m+1)} \) is just that set which is relevant in obtaining a minimum variance estimate of \( z_{N-(m+1)} \) given the information pattern at time \( N - (m + 1) \). The set of data which is relevant depends on the correlation of the \( r_n \) 's. For example, if we assume that

\[
E(r_{n+1} | r_n) = br_n = E(r_{n+1} | r_n, r_{n-1}, \ldots, r_0)
\]
it can be verified that the relevant set of data consists of all observations and all inputs since the last consecutive pair of observations (a consecutive pair of observations $x_m, x_{m+1}$ determines $r_m$ uniquely). In the event that no consecutive pair of observations has occurred the information pattern consists of all inputs and all observations since the start of the process. Note that in this case $\delta x_{N-(m+1)}$ does not qualify as a process state since $\delta x_{N-m}$ cannot be determined from $\delta x_{N-(m+1)}$ and the input $y_{N-(m+1)}$.

V. AN ALTERNATE EXPRESSION FOR THE OPTIMAL POLICY

In the previous section it was shown that under an optimal policy the inputs for controlling the process under considerations are given by

$$y_{N-(m+1)} = - (H'B_mH + G)^{-1} H'B_mA \delta[x_{N-(m+1)}] \quad m = 0, 1, 2, \cdots, N - 1$$

where $B_m$ is determined by the recurrence relation

$$B_{m+1} = A'B_m(HG^{-1}H'B_m + I)^{-1} A \quad m = 0, 1, 2, \cdots, N - 1$$

If $A$ and $B_0$ are nonsingular we see from Eq. (4.21) that $B_{m+1}$ is nonsingular. Assuming that $B_{m+1}$ is nonsingular let $W_m = B_m^{-1}$. Substituting $W_m^{-1}$ into Eq. (4.21) for $B_m$, we obtain after slight simplification

$$W_{m+1} = A^{-1}(HG^{-1}H'W_m) A^{-1}. \quad W_0 = B_0^{-1}$$

(5.1)

The solution of this difference equation yields

$$W_m = \sum_{k=0}^{m-1} A^{-k}HG^{-1}H'A^{-k} + A^{-m}W_0A^{-m}. \quad m = 1, 2, \cdots$$

(5.2)

Replacing $B_m$ in Eq. (4.20) by $W_m^{-1}$ the expression for inputs under the optimal policy is found to be

$$y_{N-(m+1)} = - (H + W_mH'^{-1}G)^{-1} A \delta[x_{N-(m+1)}]. \quad m = 0, 1, \cdots$$

(5.3)

Note that the validity of the above expression depends on the nonsingularity of the matrices $A$ and $B_0$.

VI. ASYMPTOTIC PROPERTIES OF THE OPTIMAL POLICY

From Eq. (4.20) it can be seen that the asymptotic properties of the inputs under the optimal policy are determined by those of the positive semi-definite matrix $B_m$. Therefore our discussion of the asymptotic properties
of the optimal policy will be limited to a discussion of the asymptotic properties of $B_m$.

Define the norm $\| B_m \|$ of the matrix $B_m$ by

$$\| B_m \| = \max_i \lambda_i \quad (6.1)$$

where the $\lambda_i$ are the eigenvalues of the matrix $B_m$. First, we wish to show that

$$\| B_m \| \leq \| A' H^{-1} G H^{-1} A \|, \quad m = 1, 2, \ldots \quad (6.2)$$

**PROOF.** From Eq. (4.21) we have

$$\| B_{m+1} \| = \| A' B_m (B_m + C)^{-1} CA \| \quad m = 0, 1, 2, \ldots \quad (6.3)$$

where $C = H^{-1} G H^{-1}$ is a symmetric nonsingular positive definite matrix. Notice that

$$\| B_m (B_m + C)^{-1} \| \leq \| B_m \| \| (B_m + C)^{-1} \|$$

implies that

$$\| B_m (B_m + C)^{-1} \| < 1 \quad (6.5)$$

since $C$ is a positive definite matrix. Consequently

$$\| B_{m+1} \| \leq \| A' C A \| = \| A' H^{-1} G H^{-1} A \| \quad m = 0, 1, 2, \ldots \quad (6.6)$$

which is the desired result.

Secondly, we wish to show that if $\| A \| < 1$

$$\lim_{m \to \infty} \| B_m \| = 0. \quad (6.7)$$

**PROOF.** Since $H G^{-1} H'$ is positive definite and $B_m$ is positive semidefinite, $H G^{-1} H' B_m$ is positive semidefinite. Therefore

$$\| B_{m+1} \| = \| A' B_m (H G^{-1} H' B_m + I)^{-1} A \| \leq \| A' B_m A \| \leq \| A \|^2 \| B_m \| \quad (6.8)$$

Thus, Eq. (6.7) follows when $\| A \| < 1$.

**VII. AN EXPLICIT EXPRESSION FOR $F_N(I_0)$**

From Eq. (4.18) we can write $F_N(I_0)$, the expected cost of an $N$-stage process using an optimal policy when the initial information pattern is $I_0$, as

$$F_N(I_0) = E\{ [x_0, B_N x_0] + f_N \mid I_0 \}. \quad (7.1)$$

Here $x_0$ is the initial state of the system and $B_N$ is determined by the recur-
rence relation in Eq. (4.21). Equations (4.10), (4.11), and (4.18) yield the relation

\[ E[f_{m+1} \mid I_0] = E\left[ f_m + \left( e_{m+1}, A'B_m[H'B_mH + G]^{-1}H'B_mAe_{m+1} \right) + (r_{N-(m+1)}, B_m r_{N-(m+1)}) \right] \mid I_0 \]  

which can be used to determine \( E(f_N \mid I_0) \).

From Eq. (7.2) we have

\[
E(f_N \mid I_0) = E \left( \sum_{j=1}^{N} [e_j, Q_j e_j] + (r_{N-j}, B_j r_{N-j}) \right) \mid I_0 \tag{7.3}
\]

where \( Q_j \) denotes

\[ A'B_jH(H'B_jH + G)^{-1}H'B_jA \]

and

\[ e_j = [E(x_{N-j} \mid I_{N-j}) - x_{N-j}] \]

is the error in the observer’s minimum variance estimate of the state of the system at time \( N - j \) given \( I_{N-j} \).

On inspection of Eq. (4.1) we see that

\[
e_j = \begin{cases} 
0 & \text{if } x_{N-j} \text{ was observed} \\
r_{N-(j+1)} & \text{if } x_{N-(j+1)} \text{ was the last state observed} \\
A r_{N-(j+2)} + r_{N-(j+1)} & \text{if } x_{N-(j+2)} \text{ was the last state observed} \\
& \vdots \\
\sum_{k=1}^{m} A^{k-1} r_{N-(j+k)} & \text{if } x_{N-(j+m)} \text{ was the last state observed} \\
\end{cases} \tag{7.4}
\]

Here \( j + m \leq N \) since the initial state of the system is known.

Since the \( r_n \) and the random variables governing the interruptions are independent, on taking into account the probability of interruption we find

\[
E[(e_j, Q_j e_j) \mid I_0] = \left(1 - p\right) \sum_{n=1}^{N-(j+1)} p^n E \left( \sum_{k=1}^{n} A^{k-1} r_{N-(j+k)}, Q_j \sum_{k=1}^{n} A^{k-1} r_{N-(j+k)} \right) \mid I_0
\]

\[+ p^{N-j} E \left( \sum_{k=1}^{N-j} A^{k-1} r_{N-(j+k)}, Q_j \sum_{k=1}^{N-j} A^{k-1} r_{N-(j+k)} \right) \mid I_0 \right]. \tag{7.5}
\]

The second term on the right-hand side of Eq. (7.5) arises because the probability of “observing” the initial state of the system is 1, not \((1 - p)\).
Substituting the above expression for $E[(e_j, Q_je_j) \mid I_0]$ into Eq. (7.2) we have

$$F_N(I_0) = (x_0, B_Nx_0) + \sum_{j=1}^{N} E[(r_{N-j}, B_j r_{N-j}) \mid I_0]$$

$$+ \sum_{j=1}^{N-1} \left\{ (1 - p) \sum_{n=1}^{N-(j+1)} p^n E \left[ \left( \sum_{k=1}^{n} r_{N-(j+k)}, \sum_{k=1}^{n} A^{k-1}Q_sA^{k-1}r_{N-(j+k)} \right) \mid I_0 \right] \right\}$$

$$+ p^{N-j} E \left[ \left( \sum_{k=1}^{N-j} r_{N-(j+k)}, \sum_{k=1}^{N-j} A^{k-1}Q_sA^{k-1}r_{N-(j+k)} \right) \mid I_0 \right] \right\}. \quad (7.6)$$

On noting that $\hat{\sum} r_a = \sigma^2 I$ we see that for an arbitrary $s \times s$ matrix $M$, the expected value of the inner products can be evaluated as

$$E[(r, Mr) \mid I_0] = \sum_{i=1}^{s} \left[ E \left( r_i \sum_{k=1}^{s} m_{ik} r_k \right) \mid I_0 \right] = \sum_{i=1}^{s} \left[ \sum_{k=1}^{s} m_{ik} E[r_i r_k \mid I_0] \right]$$

$$= \sigma^2 \sum_{i=1}^{s} m_{ii} = \sigma^2 \text{Tr} (M) \quad (7.7)$$

where $\text{Tr} (M)$ denotes the trace of the matrix $M$. Consequently $F_N(I_0)$ can be written

$$F_N(I_0) = (x_0, B_Nx_0) + \sigma^2 \sum_{j=1}^{N} \text{Tr} (B_j)$$

$$+ \sigma^2 \sum_{j=1}^{N-1} \left\{ (1 - p) \sum_{n=1}^{N-(j+1)} p^n \left[ \sum_{k=1}^{n} \text{Tr} \left( A^{k-1}Q_sA^{k-1} \right) \right] \right\}$$

$$+ p^{N-j} \sum_{k=1}^{N-j} \text{Tr} \left( A^{k-1}Q_sA^{k-1} \right). \quad (7.8)$$

It is interesting to note that:

1. The first term on the right-hand side of Eq. (7.8) is the cost that would be associated with a deterministic process of length $N$ using an optimal policy.

2. The sum of the first two terms on the right-hand side of Eq. (7.8) is the expected cost that would be associated with a noninterrupted process of length $N$ under an optimal policy.

3. The third term on the right-hand side of Eq. (7.8) is the additional expected cost introduced by the interruptions.
VIII. ASYMPTOTIC PROPERTIES OF $F_N(I_0)$

From Eq. (7.8) we obtain
\[
F_N(I_0) \leq (x_0, B_N x_0) + \sigma^2 \sum_{j=1}^{N} \text{Tr} (B_j) + \sigma^2 \sum_{j=1}^{N-1} \sum_{n=1}^{N-j} \sum_{k=1}^{p^n} \left[ \sum_{k=1}^{N} \text{Tr} \left( A'^{k-1} Q_j A^{k-1} \right) \right].
\]

(8.1)

$A$ and $Q$ are $s \times s$ matrices, therefore

\[
\text{Tr} \left( A'^{k-1} Q_j A^{k-1} \right) = \text{Tr} [Q_j (AA')^{k-1}] \leq s \| Q_j (AA')^{k-1} \| \leq s \| Q_j \| \| AA' \|^{k-1}.
\]

(8.2)

Let us denote $\| A \|$ by $a$, $\| AA' \|$ by $a^*$, $\| B_j \|$ by $b_j$ and $\| Q_j \|$ by $q_j$. Thus

\[
F_N(I_0) \leq (x_0, B_N x_0) + \sigma^2 s \sum_{j=1}^{N} b_j + \sigma^2 s \sum_{j=1}^{N-1} q_j \sum_{n=1}^{N-j} \sum_{k=1}^{p^n} a^{*k-1}
\]

(8.3)

Consider the last two sums on the right-hand side of Eq. (8.3). If $a^* < 1$

\[
\sum_{n=1}^{N-j} \sum_{k=1}^{p^n} a^{*k-1} \leq \left[ \frac{1}{1-a^*} \right] \sum_{n=1}^{N-j} p^n \leq \frac{p}{(1-p) (1-a^*)}.
\]

(8.4)

If $pa^* < 1 < a^*$

\[
\sum_{n=1}^{N-j} p^n \sum_{k=1}^{n} a^{*k-1} \leq a^{*1} \sum_{n=1}^{N-j} (pa^*)^n \frac{a^*}{a^* - 1} \leq \frac{pa^*}{(a^* - 1) (1 - pa^*)}.
\]

(8.5)

From the expression for $Q_j$ (Section VII) we see that

\[
q_j = \| A'B_j H (H'B_j H + G)^{-1} H'B_j A \|
\]

\[
\leq a \| B_j (H'B_j H + G)^{-1} H' \| \| B_j A \|.
\]

(8.6)

But since $G$ is positive definite

\[
\| B_j (H'B_j H + G)^{-1} H' \| \leq 1.
\]

(8.7)

Hence

\[
q_j \leq a^2 b_j.
\]

(8.8)

From Eq. (6.2) and (6.8) we have $b_j \leq \| A'H'^{-1} GH^{-1} A \|$ and $b_j \leq a^2 b_0$. Thus

\[
b_j \leq \min (a^2 b_0, \| A'H'^{-1} GH^{-1} A \|)
\]

(8.9)
If $a < 1$ we have from Eq. (8.4), (8.5), and (8.9)

$$F_N(I_0) \leq (x_0, a^{2N} b_0 x_0) + \sigma^2 s \sum_{j=1}^{N} a^{2j} b_0 + \sigma^2 s \sum_{j=1}^{N-1} a^{2j} b_0 \frac{p}{(1 - p)} (1 - a^*)$$

$$\leq a^{2N} b_0(x_0, x_0) + \frac{\sigma^2 s b_0 a^2}{1 - a^2} + \frac{\sigma^2 s b_0 a^2 p}{1 - p} (1 - a^*)$$

$$N = 1, 2, \ldots \quad (8.10)$$

Thus if $\| A \| < 1$, $F_N(I_0)$ is bounded for all $N$.

In a similar manner, if $pa^* < 1 < a^*$, we find from Eq. (8.5), (8.8), and (8.9) that

$$F_N(I_0) \leq d(x_0, x_0) + \sigma^2 s [b_0 + (N - 1) d] + \frac{\sigma^2 s p a^* b_0}{(a^* - 1) (1 - pa^*)} N \quad (8.11)$$

where $d = \| A' H' -1 GH -1 A \|$. Hence, if $pa^* < 1 < a^*$,

$$F_N(I_0) \leq K_1 + K_2 N \quad (8.12)$$

where $K_1$ and $K_2$ are bounded positive constants obtainable from Eq. (8.11).

### IX. Summary of Results

In Section III it was shown that an optimal policy for controlling in interrupted stochastic control process can be expressed in terms of functional equations obtained by the standard iterative procedure of dynamic programming. The results summarized here will concern the interrupted process described in Section IV which involves a linear system and a quadratic loss functional.

#### Results Concerning Optimal Policy

For the interrupted process described in Eq. (4.1), (4.2), et. seq. the input $y_{N-(m+1)}$ applied at time $N - (m + 1)$ under an optimal policy for a process that terminates at time $n = N$ is given by

$$y_{N-(m+1)} = -(H' B_m H + G)^{-1} H' B_m A \mathcal{E}[x_{N-(m+1)}]. \quad m = 0, 1, \ldots \quad (4.20)$$

Here $\mathcal{E}[x_{N-(m+1)}]$ is the minimum variance estimate of $x_{N-(m+1)}$, the state of the system at time $N - (m + 1)$, given the information pattern (see Section II) at time $N - (m + 1)$, and $B_m$ is determined by the recurrence relation

$$B_{m+1} = A' B_m (H' G^{-1} H' B_m + I) H' B_m A. \quad (4.21)$$

Thus the inputs applied under an optimal policy are independent of the probability of interruption. In the event that the $r_n$ in Eq. (4.1) are correlated,
the inputs applied under an optimal policy are given by Eq. (4.20) with

\[ x_{N-n} = x_{N-n} + A^{-n} \sum_{k=0}^{n-1} A^k r_{N-(k+1)} \quad n = 1, 2, \ldots \]  

(4.22)

When \( A \) and \( B_0 \) are nonsingular matrices, the inputs under the optimal policy may be expressed in the form

\[ y_{N-(m+1)} = -(H + W_m H^{-1} G) A \delta_{N-(m+1)} \quad m = 0, 1, 2, \ldots \]  

(5.3)

where

\[ W_m = \sum_{k=0}^{m-1} A^{-k} H G^{-1} H' A'^{-k} + A^{-m} W_0 A'^{-m}, \quad W_0 = B_0^{-1} \]  

(5.2)

The asymptotic properties of the optimal policy are determined by those of the matrix \( B_m \). Defining the norm \( \| B_m \| \) of the matrix \( B_m \) by

\[ \| B_m \| = \max_i | \lambda_i | \]  

(6.1)

where the \( \lambda_i \) are the eigenvalues of the matrix \( B_m \), we have

\[ \| B_m \| \leq \| A' H^{-1} G H^{-1} A \| \quad m = 1, 2, \ldots \]  

(6.2)

and

\[ \| B_m \| \leq \| A \|^2 \| B_0 \|. \quad m = 1, 2, \ldots \]  

(6.3)

Hence if \( \| A \| < 1 \)

\[ \lim_{m \to \infty} \| B_m \| = 0 \quad \text{for all } m \]  

(6.7)

Thus the input applied under the optimal policy at time \( N - m \) is a bounded function of the minimum variance estimate of the state of the system at time \( N - m \), and in particular, if \( \| A \| < 1 \), \( y_{N-m} \) approaches zero as \( m \) approaches infinity.

**Results Concerning the Expected Cost of the Process Under an Optimal Policy**

The expected cost under an optimal policy for a process of length \( N \) when the initial information pattern is \( I_0 \) is

\[ F_N(I_0) = (x_0, B_N x_0) + \sigma^2 \sum_{j=1}^{N} \text{Tr} (B_j) \]

\[ + \sigma^2 \sum_{j=1}^{N-1} \left( (1 - p) \sum_{n=1}^{N-(j+1)} p^n \sum_{k=1}^{n} \text{Tr} (A^{k-1} Q_j A^{k-1}) \right) \]

\[ + p^{N-j} \sum_{k=1}^{N-j} \text{Tr} (A^{k-1} Q_j A^{k-1}) \quad N = 1, 2, \ldots \]  

(7.8)
Here \( Q_j = A'B_jH(H'B_jH + G)^{-1}H'B_jA \), \( x_0 \) is the initial state of the system, (assumed known) and \( \text{Tr} (B_j) \) denotes the trace of the matrix \( B_j \).

If \( \| A \| < 1 \) the expected cost under an optimal policy for a process of length \( N \) is bounded for all \( N \) [see Eq. (8.10)]. If \( p \| A \| < 1 < \| A \| \) the expected cost under an optimal policy for a process of length \( N \) is bounded by

\[
F_N(I_0) \leq K_1 + K_2N
\]

(8.12)

where \( K_1 \) and \( K_2 \) are bounded positive constants obtainable from Eq. (8.11).

Acknowledgment

The author wishes to acknowledge the aid obtained concerning this paper in useful discussions with Dr. Richard Bellman of The RAND Corporation, who suggested the problem, and with Prof. L. A. Zadeh of The University of California at Berkeley.

References