The Annals of Applied Probability
1993, Vol. 3, No. 2, 364–379

OPTIMAL CONTROL AND REPLACEMENT WITH STATE-DEPENDENT FAILURE RATE: DYNAMIC PROGRAMMING[1]

BY ARTHUR C. HEINRICHER AND RICHARD H. STOCKBRIDGE

University of Kentucky

A class of stochastic control problems where the payoff depends on the running maximum of a diffusion process is described. The controller must make two kinds of decision: first, he must choose a work rate (this decision determines the rate of profit as well as the proximity of failure), and second, he must decide when to replace a deteriorated system with a new one. Preventive replacement is a realistic option if the cost for replacement after failure is larger than the cost of a preventive replacement.

We focus on the profit and replacement cost for a single work cycle and solve the problem in two stages. First, the optimal feedback control (work rate) is determined by maximizing the payoff during a single excursion of a controlled diffusion away from the running maximum. This step involves the solution of the Hamilton–Jacobi–Bellman (HJB) partial differential equation. The second step is to determine the optimal replacement set. The assumption that failure occurs only on the set where the state is increasing implies that replacement is optimal only on this set. This leads to a simple formula for the optimal replacement level in terms of the value function.

1. Introduction. This paper is devoted to stochastic control problems motivated by optimal control and replacement problems for deteriorating systems. The models are constructed from diffusions but are nonstandard because one component (the "wear") is required to be continuous and monotone. Working policies affect the rate of wear which in turn affects the rate of failure. The system is assumed to fail at a rate k which depends on the state. If replacement at or after failure is more expensive than a planned replacement, then it may be optimal to replace the system while it is still in working condition.

We focus here on a single working cycle and maximize the revenue collected minus the replacement/failure cost. We use dynamic programming techniques to give sufficient conditions for optimality. The solution for the dynamic programming partial differential equation leads to the optimal control policy (in feedback form) as well as the formula for the optimal value and a means of determining the optimal replacement level. This work can be viewed as one step toward maximizing the long-run average profit for the system when it is renewed at each failure or replacement. We describe briefly

Received January 1992; revised May 1992.
[1]Partially supported by the NSF Grant DMS-90-06674.
AMS 1991 *subject classifications*. Primary 49B60, 49C20, 93E20.
Key words and phrases. Controlled diffusion, running maximum, dynamic programming, optimal replacement, state dependent failure, stochastic wear models.

this extension in Section 5. A more powerful approach to the long-run average problem will be studied in a companion paper [Heinricher and Stockbridge (1993)].

One of the fundamental works on optimal replacement for deteriorating systems (without control between replacements) is Taylor (1975). Anderson (1988, 1990) analyzes the variational and quasivariational inequalities that arise in optimal replacement problems for general (monotone) Markov processes (again without control between replacements). Conrad and McClamroch (1987) describe an application in automated manufacturing where the rate of work controls the rate of deterioration for a drilling machine. The recent survey article by Valdez-Flores and Feldman (1989) contains a wealth of additional references.

One important application which requires a continuous and monotone stochastic process model is the optimal control of *wear*; automobile tires and drill bits do not "unwear." Çinlar (1972) has analyzed perhaps the most general class of stochastic processes suitable for modeling wear [see also Çinlar (1977, 1984)]. In these *Markov additive* processes, one coordinate of a multidimensional Markov process is monotone but not a Markov process when considered alone. The monotone component is the model for wear and the remaining coordinates model the environmental influences that cause wear.

In this paper, the Markov process in the background is a one-dimensional (controlled) diffusion. The monotone component is the running maximum of this diffusion. Throughout the paper, the *state* is given by the pair (x_t, y_t) satisfying

$$(1) \qquad \begin{aligned} dx_t &= f(x_t, y_t, u_t)\,dt + \sigma(x_t, y_t)\,dw_t, & x_0 &= x,\\ y_t &= \max\{x_s : 0 \le s \le t\} \vee y, & y_0 &= y \ge x. \end{aligned}$$

Here $w = (w_t,\ 0 \le t < \infty)$ denotes a standard, one-dimensional Brownian motion and $u = (u_t,\ 0 \le t < \infty)$ is a control process. Stochastic control problems involving the running max $y = (y_t, 0 \le t < \infty)$ are studied in Heinricher and Stockbridge (1991) [see also Barron (1991)].

Assume that the controller has complete observations of the pair (x_t, y_t) and that the system fails at a random time ζ. The failure time is defined via a random threshold Y:

$$(2) \qquad \zeta = \inf\{t \ge 0 : y_t \ge Y\}.$$

We assume that Y is a nonnegative random variable independent of $((x_t, y_t), 0 \le t < \infty)$ with distribution

$$G(y) = 1 - \exp\left(-\int_0^y k(z)\,dz\right) \qquad y \ge 0,$$

where $k(\cdot)$ is a nonnegative, right-continuous and nondecreasing function.

This implies

$$P_{xy}(\zeta > t | y < Y) = E_{xy}\left[\exp\left(-\int_y^{y(t)} k(z)\,dz\right)\right]$$
$$= E_{xy}\left[\exp\left(-\int_0^t k(y(s))\,dy(s)\right)\right],$$

where we use the pathwise continuity of the running max and the right-continuity of $k(\cdot)$. [Here and throughout the paper, a subscript on a probability or expectation of the form "xy" indicates conditioning on the initial state being (x, y). If there is no subscript, the initial condition is $(x, y) = (0, 0)$.] Notice that

$$P_{xy}(y(\zeta) > \Delta) = \exp\left(-\int_y^{\Delta} k(z)\,dz\right), \qquad y \le \Delta. \tag{3}$$

Our definition of ζ is motivated by the assumption that *failure can occur only while the running max is increasing*. In particular, if the system is working at level y, it will fail as the state increases from y to $y + \delta y$ with probability $k(y)\,\delta y + o(\delta y)$. It is important to notice that this is not a standard state-dependent failure mechanism; we have a failure rate that is *per unit wear* and not *per unit time*.

The controller has the option to perform a preventive replacement, at a cost R_2 which may depend on the present state y, and return the system to the *new* state $y = 0$. We shall consider only replacement times which are first-passage times. That is, for a *replacement level* $\Delta \ge 0$, the replacement time τ is

$$\tau = \tau(\Delta) = \inf\{t \ge 0\colon y_t \ge \Delta\}, \tag{4}$$

where $\tau = +\infty$ if the set is empty (which will happen if and only if $\Delta = +\infty$).

There is a cost $R_1 \ge R_2$ for replacement at or after failure and so the cost associated with a replacement level Δ is

$$R(\Delta) = R_1(y(\zeta))1_{\{\zeta \le \tau\}} + R_2(y(\tau))1_{\{\zeta > \tau\}}$$
$$= R_1(y(\zeta))1_{\{y(\zeta) \le \Delta\}} + R_2(\Delta)1_{\{y(\zeta) > \Delta\}}.$$

If the controller chooses to work, then revenue is accumulated at rate $h = h(x_t, y_t, u_t)$. For an initial state (x, y), control policy $u = (u_t, 0 \le t < \infty)$, and replacement level $\Delta \ge y$, the total *profit* for the cycle is

$$J(x, y; u, \Delta) - \overline{R}(\Delta) = E_{xy}\left[\int_0^{\zeta(u) \wedge \tau(\Delta)} h(x_t, y_t, u_t)\,dt - R(\Delta)\right]. \tag{5}$$

The special structure of the failure rate implies that it is always optimal to work if $x < y$ [assuming that $h(x, y, u) > 0$ for some admissible control], because y is constant in this region and failure cannot occur while y is constant. This simplifies and separates the optimal replacement decision from the optimal control decision:

1. The optimal replacement set is restricted to the main diagonal $\{(x, y): x = y\}$.
2. The optimal control policy maximizes the profit accrued up to the first time that $x_t = y_t$.

1.1. *Summary.* We use dynamic programming methods to obtain sufficient conditions for optimality and to determine optimal control policies as well as optimal replacement levels. The optimal feedback control is characterized for all (x, y) with $x < y$ by solving a simpler "auxiliary" control problem: maximize the revenue collected up to the first time that $x_t = y_t$. This involves the solution of a (simpler) HJB equation on each excursion of the controlled diffusion (below the running maximum). Our assumptions concerning the failure-rate restrict the optimal replacement set to the main diagonal $(x = y)$ and so the optimal replacement level is determined by a simple one-dimensional maximization problem.

We give sufficient conditions for optimality via the Hamilton–Jacobi–Bellman equation in Section 3. The special form of the failure mechanism introduces an oblique derivative condition along the boundary of the state space $(x = y)$. Theorem 3.4 connects the auxiliary problems with the original control problem and represents the optimal value in terms of the auxiliary value functions.

Section 4 is devoted to the replacement level Δ. The explicit formula for the optimal value in Theorem 3.4 leads to an explicit formula for the replacement level. We include simple examples to illustrate our approach.

We describe briefly in Section 5 how these methods can be adapted to the long-term average control problem. This extension is considered in detail in Heinricher and Stockbridge (1993) where we take an invariant measure approach to the optimization problem.

The approaches here and in the companion paper Heinricher and Stockbridge (1993) are complementary. The dynamic programming approach gives explicit formulae for an optimal feedback control when the HJB equation has a sufficiently smooth solution. The invariant measure approach does not require any regularity of the value function, but it does not lead directly to the optimal control policies.

2. Formulation of the problem. The technical formulation of our control problem follows the standard approach to controlled diffusions described, for instance, in the text by Fleming and Rishel (1975).

For the admissible controls, take the collection of *nonanticipative controls* as defined in Chapter VI, page 162 of Fleming and Rishel (1975), and let $\mathscr{A}$ denote the collection of admissible controls. We assume that control processes $u = (u_t, 0 \le t < \infty)$ take values in a *compact* subset U of the real numbers and that the coefficients of the problem satisfy conditions sufficient to provide polynomial growth for the value function and guarantee existence of solutions to (1).

We assume that:

CONDITION 1. h is continuous and bounded on $\{(x, y, u)\colon x \le y,\ y \ge 0, u \in U\}$.

CONDITION 2. f is a bounded C^1 function on $\{(x, y, u)\colon x \le y,\ y \ge 0, u \in U\}$ and σ is a bounded C^1 function on $\{(x, y)\colon x \le y,\ y \ge 0\}$ with

$$|f_x(x,y,u)| + |f_y(x,y,u)| + |f_u(x,y,u)| \le K, \qquad x \le y,\ y \ge 0,\ u \in U,$$
$$|\sigma_x(x,y)| + |\sigma_y(x,y)| \le K, \qquad x \le y,\ y \ge 0,$$

for a suitable constant K.

As noted in Heinricher and Stockbridge (1991a), the problem may be ill-posed (because the expected failure time is infinite) without a positive lower bound on $f(x, y, u)$, so we require that:

CONDITION 3. There is a constant α such that

$$0 < \alpha \le f(x,y,u), \qquad x \le y,\ y \ge 0,\ u \in U.$$

REMARK 2.1. These conditions are sufficient to guarantee the existence and pathwise uniqueness of solutions. Condition 3 guarantees that $E_{xy}[\tau(\Delta)] < \infty$ if $\Delta < \infty$.

We make the following assumption concerning the failure rate $k = k(y)$.

CONDITION 4. $k(y) \ge 0$ for $y \ge 0$, $k(\cdot)$ is nondecreasing and right-continuous (possibly infinite) with

$$\int_0^\infty k(z)\,dz = +\infty.$$

Recalling (3), this condition implies that

$$P(\zeta = \infty) = 0.$$

REMARK 2.2. Conditions 3 and 4 combine to give the following estimate. There is an $\varepsilon > 0$ such that

$$k_\varepsilon = \min\{y \ge 0\colon k(y) \ge \varepsilon\}$$

is finite. Define

$$\theta_\varepsilon = \inf\{t \ge 0\colon y_t \ge k_\varepsilon\},$$

which has finite expectation by Condition 3. We have then

$$E\left[\exp\left(-\int_0^T k(y_t)\,dy_t\right)\right] \le E\left[\exp\left(-\int_{\theta_\varepsilon}^T k(y_t)\,dy_t\right)\right]$$
$$\le E\left[\exp(-\varepsilon(y_T - k_\varepsilon))\right].$$

As described in the introduction, we will consider only replacement policies specified by a threshold $\Delta \ge y_0 = y$, which, with our assumptions on the failure time ζ, implies that the expected replacement cost takes the form

$$\text{(6)} \quad \bar{R}_{xy}(\Delta) = E_{xy}[R(\Delta)] = \int_y^\Delta R_1(z)k(z)e^{-\int_y^z k(s)\,ds}\,dz + R_2(\Delta)e^{-\int_y^\Delta k(s)\,ds}.$$

We make the following assumptions concerning the replacement and failure costs:

CONDITION 5. $R_1(y)$ and $R_2(y)$ are nonnegative, nondecreasing, bounded functions with

$$R_1(y) \ge R_2(y) \quad (y \ge 0).$$

When the replacement level Δ is not a control decision, this problem is related to the problems solved in Heinricher and Stockbridge (1991b, c). In fact, the problems solved in those papers are special cases of the problem considered here if we take the failure rate $k(y)$ to be identically 0 for $0 \le y < \Delta$ and identically $+\infty$ for $y \ge \Delta$, where Δ is a fixed (deterministic) failure level.

3. Solve the control problem: Determine u^*. We describe now sufficient conditions for optimality for the single-cycle problems. These are dynamic programming conditions and involve the Hamilton–Jacobi–Bellman (HJB) partial differential equation. Throughout this section, the replacement level Δ is fixed.

The objective is to choose an admissible control process $u = (u_t, 0 \le t < \infty)$ to maximize

$$\text{(7)} \qquad J(x, y; u, \Delta) = E_{xy}\int_0^{\zeta(u)\wedge\tau(\Delta)} h(x_t, y_t, u_t)\,dt,$$

where the state is the pair (x_t, y_t) defined in (1), $\zeta = \zeta(u)$ is the failure time, and $\tau = \tau(\Delta)$ is the replacement time.

Our assumptions on the failure time ζ allow us to reformulate the control problem as one for the system (1) but with an exponential discount factor in the integrand:

$$\text{(8)} \qquad J(x, y; u, \Delta) = E_{xy}\int_0^{\tau(\Delta)} e^{-\int_0^t k(y_s)\,dy_s} h(x_t, y_t, u_t)\,dt.$$

The following theorem is an extension of the standard sufficient conditions for optimality as presented in Chapter VI of Fleming Rishel (1975) [see also Heinricher and Stockbridge (1991a)].

THEOREM 3.1. *Let $V(x, y)$ be a solution of the dynamic programming equation*

$$\max_{u \in U}\left\{\tfrac{1}{2}\sigma(x, y)^2 V_{xx}(x, y) + f(x, y, u)V_x(x, y) + h(x, y, u)\right\} = 0, \tag{9}$$

in the region $x < y$, $0 < y < \Delta$, satisfying the boundary condition

$$V_y(y, y) - k(y)V(y, y) = 0, \qquad 0 < y < \Delta, \tag{10}$$

as well as the terminal condition

$$V(\Delta, \Delta) = 0. \tag{11}$$

(If $\Delta = +\infty$, the terminal condition is not enforced.) In addition, suppose $V(x,y)$ is continuous, twice continuously differentiable with respect to x, and satisfies a polynomial growth condition

$$|V(x, y)| \le C(1 + |x|^p + |y|^p), \qquad x \le y, 0 \le y \le \Delta, \tag{12}$$

for appropriate constants C and p. Then:

(a) *$V(x, y) \ge J(x, y; u, \Delta)$ for any admissible control u and any $x \le y$.*
(b) *If u^* is an admissible control which attains the maximum in* (9), *then u^* is optimal and $V(x, y) = J(x, y; u^*, \Delta)$ is the value function.*

REMARK 3.2. Note that our assumption that failure can occur only when y_t is increasing has put the "killing term" $-k(y)V(y, y)$ into the boundary condition (10). This is in contrast to the usual sort of killing which would surface as a zeroeth order term in the HJB equation (9).

PROOF. Consider first the case $\Delta = +\infty$. Let $u = (u_t, 0 \le t < \infty)$ be an admissible control process and let (x_t, y_t) denote the associated state pair. For $T > 0$ and $N > 0$, define

$$T(N) := \inf\{t > 0\colon x_t \le -N, y_t \ge N\} \wedge T.$$

Let $V(x, y)$ satisfy the smoothness and growth conditions in the statement of the theorem. The generalized Itô formula provides

$$\begin{aligned} V(x, y) = &-\int_0^{T(N)} e^{-\int_0^t k(y_s)\,dy_s}\big[f(x_t, y_t, u_t)V_x(x_t, y_t) \\ &\qquad + \tfrac{1}{2}\sigma^2(x_t, y_t)V_{xx}(x_t, y_t)\big]\,dt \\ &-\int_0^{T(N)} e^{-\int_0^t k(y_s)\,dy_s}\big[V_y(x_t, y_t) - k(y_t)V(x_t, y_t)\big]\,dy_t \\ &-\int_0^{T(N)} e^{-\int_0^t k(y_s)\,dy_s}\sigma(x_t, y_t)V_x(x_t, y_t)\,dw_t \\ &+ e^{-\int_0^{T(N)} k(y_s)\,dy_s}V(x_{T(N)}, y_{T(N)}), \end{aligned}$$

where dy is the measure associated with the monotone increasing process y. Since the process $(y_t,\ 0 \le t < \infty)$ increases only on the set $\{t\colon x_t = y_t\}$, the measure dy assigns mass only on this set. Hence the boundary condition (10) implies that the second integral is 0. Taking expectations, using (9) and the fact that the stochastic integral has zero expectation (because the integrand is bounded on the truncated region), we obtain

$$\begin{aligned} V(x,y) &= -E_{xy}\int_0^{T(N)} e^{-\int_0^t k(y_s)\,dy_s}\big[f(x_t,y_t,u_t)V_x(x_t,y_t) \\ &\qquad\qquad + \tfrac{1}{2}\sigma^2(x_t,y_t)V_{xx}(x_t,y_t)\big]\,dt \\ &\quad + E_{xy}e^{-\int_0^{T(N)}k(y_s)\,dy_s}V(x_{T(N)},y_{T(N)}) \\ &\ge E_{xy}\int_0^{T(N)} e^{-\int_0^t k(y_s)\,dy_s}h(x_t,y_t,u_t)\,dt \\ &\quad + E_{xy}e^{-\int_0^{T(N)}k(y_s)\,dy_s}V(x_{T(N)},y_{T(N)}). \end{aligned}$$

First let $N \to \infty$. Using the continuity of $V(x, y)$ and the monotone convergence theorem, we obtain

$$V(x,y) \ge E_{xy}\int_0^T e^{-\int_0^t k(y_s)\,dy_s}h(x_t,y_t,u_t)\,dt + E_{xy}e^{-\int_0^T k(y_s)\,dy_s}V(x_T,y_T).$$

Remark 2.2 and Condition 3, combined with the polynomial growth of $V(x, y)$, show that the final term vanishes in the limit as $T \to \infty$. The monotone convergence theorem allows us to conclude that $V(x, y)$ is an upper bound on the maximum value.

If the replacement level Δ is finite, the argument is the same except in the last step, where we use the terminal condition (11) to conclude that

$$\lim_{T\to\infty} V(x_{T\wedge\tau(\Delta)}, y_{T\wedge\tau(\Delta)}) = V(x_{\tau(\Delta)}, y_{\tau(\Delta)}) = 0.$$

If a control $u^* = u^*(x, y)$ achieves equality in (9), then equality holds throughout the above argument and u^* achieves the value $V(x, y)$. Hence

$$V(x,y) = \max_{u\in\mathscr{A}} J(x,y;u)$$

and u^* is an optimal control. □

The key to the actual solution of the control problem is the realization that the value function can be constructed from a family of *auxiliary problems* where y is a fixed parameter. This decomposition was introduced in Heinricher and Stockbridge (1991a) for a simpler problem (without failure or replacement) and we summarize it here.

Fix y, let $x \le y$, and define

$$\theta = \theta(x,y;u) := \inf\{t \ge 0\colon x_t = y\}. \tag{13}$$

We seek an admissible control u to maximize

(14) $$I(x, y; u) = E_{xy}\int_0^\theta h(x_t, y, u_t)\,dt.$$

Observe that the failure term does not enter this objective function because $x_t < y_t \equiv y$ for $0 \le t < \theta$. Let $W(x, y)$ denote the value function for this auxiliary problem:

$$W(x, y) = \sup_{u \in \mathscr{A}} I(x, y; u).$$

The HJB equation satisfied by $W(x, y)$ is of the standard form:

(15) $$\max_{u \in U}\{\tfrac{1}{2}\sigma(x, y)^2 W_{xx}(x, y) + f(x, y, u)W_x(x, y) + h(x, y, u)\} = 0$$

on the half line $x < y$ with the terminal condition

(16) $$W(y, y) = 0.$$

There is a simple relationship between the value functions for the auxiliary problem and the original, single-cycle control problem. The proof is based on the dynamic programming principle [see Lions (1983)].

PROPOSITION 3.3. *The value functions for the single-cycle control problem and the auxiliary problem satisfy*

(17) $$V(x, y) = W(x, y) + V(y, y), \qquad x \le y.$$

We can go one step further and represent the optimal value $V(y, y)$ on the diagonal in terms of the auxiliary value $W(x, y)$. In this way, the value function $V(x, y)$ is determined entirely in terms of the auxiliary value $W(x, y)$.

THEOREM 3.4. *For each $0 \le y \le \Delta$, let $W(x, y)$ be a solution of the dynamic programming equation* (15) *on the half line $x < y$ satisfying the terminal condition* (16). *Then the single-cycle value function is given by*

(18) $$V(x, y) = W(x, y) + \int_y^\Delta e^{-\int_y^z k(s)\,ds} W_y(z, z)\,dz, \qquad x \le y, 0 \le y \le \Delta.$$

This is valid as long as $W(x, y)$ is continuous with respect to (x, y), differentiable along $x = y$ and twice continuously differentiable with respect to x and satisfies the polynomial growth condition

$$|W(x, y)| \le C(1 + |x|^p + |y|^p), \qquad x \le y,$$

for appropriate constants C and p.

In addition, if $u^*(x, y)$ is an admissible control which attains the maximum in (15), *then $u^*(x, y)$ is an optimal control for the auxiliary problem as well as the running max problem* (7).

PROOF. Defining $V(x, y)$ as in (18), $V(x, y)$ inherits exactly the smoothness of $W(x, y)$; in particular, we have

$$V_x(x, y) = W_x(x, y), \qquad V_{xx}(x, y) = W_{xx}(x, y).$$

Since $W(x, y)$ satisfies (15), $V(x, y)$ satisfies (9) as well as the boundary condition (10); differentiating (18) with respect to y provides

$$V_y(x, y) - k(y)V(x, y) = W_y(x, y) - W_y(y, y) - k(y)W(x, y)$$

and the right-hand side vanishes along the diagonal $x = y$. Theorem 3.1 identifies $V(x, y)$ as the value function and $u^*(x, y)$ as the optimal control policy for the running max problem. □

REMARK 3.5. For initial data (x, y) with $y > \Delta$, the representation in Theorem 3.4 is extended by

$$V(x, y) = W(x, y), \qquad x \le y,\ y \ge \Delta.$$

In particular, the process is stopped as soon as the diffusion returns to the main diagonal.

One application of the previous theorem provides a simple formula for the expected failure/replacement time when the system is replaced at level Δ and the control policy is constant.

LEMMA 3.6. *Assume that the drift is given by $f(x, y, u) = f(u)$, with $u \in U$ constant, and that the diffusion coefficient $\sigma(x, y) = \sigma$ is constant. If the system starts in the initial state (x, y) with $x \le y$, a constant control policy is used, and the preventive replacement threshold is $\Delta \ge y$, then*

$$(19) \quad E_{xy}[\zeta(u) \wedge \tau(\Delta)] = \frac{(y - x)}{f(u)} + \frac{1}{f(u)} \int_y^{\Delta} \exp\left(-\int_y^z k(s)\, ds\right) dz.$$

PROOF. The representation (19) is obtained by solving a boundary value problem. The problem of interest is the dynamic programming partial differential equation (9) with the boundary condition (10) when the control set is the singleton $U = \{u\}$ and the coefficients f and σ are defined as in the lemma.

Begin by solving the auxiliary problem

$$\tfrac{1}{2}\sigma^2\psi_{xx}(x, y) + f(u)\psi_x(x, y) + 1 = 0, \qquad x < y,$$

with the terminal condition

$$\psi(y, y) = 0.$$

(The solution is the expected time to return to the diagonal.) The unique solution for this problem (satisfying a polynomial growth condition as $x \to -\infty$) is

$$\psi(x, y) = \frac{(y - x)}{f(u)}, \qquad x \le y.$$

The representation (18) in Theorem 3.4 provides the formula we seek:

$$\begin{aligned}\phi(x, y) &= \psi(x, y) + \int_y^{\Delta} \exp\left(-\int_y^z k(s)\, ds\right) \psi_y(z, z)\, dz \\ &= \frac{(y - x)}{f(u)} + \frac{1}{f(u)} \int_y^{\Delta} \exp\left(-\int_y^z k(s)\, ds\right) dz.\end{aligned}$$

It is an application of the generalized Itô formula to verify that

$$E_{xy}[\zeta(u) \wedge \tau(\Delta)] = \phi(x, y),$$

and the proof is complete. □

4. Solution of the replacement problem: Determination of Δ^*. Theorem 3.4 provides a representation for the optimal revenue for a fixed replacement level Δ (18). We now bring in the replacement cost and optimize over $\Delta \ge 0$. For this section, assume that the system starts in the new state $(x, y) = (0, 0)$.

Combining the representation (18) with the formula (6) for $\overline{R}(\Delta)$ and integrating, we have the following expression for the total profit accrued during a life cycle:

$$\begin{aligned}J(0, 0; u^*, \Delta) - \overline{R}(\Delta) &= \int_0^{\Delta} e^{-\int_0^z k(s)\, ds} W_y(z, z)\, dz - \overline{R}(\Delta) \\ &= \int_0^{\Delta} e^{-\int_0^z k(s)\, ds} \left[W_y(z, z) - (R_1(z) - R_2(\Delta)) k(z)\right] dx \\ &\quad - R_2(\Delta).\end{aligned}$$

This function of one real variable is maximized at the optimal replacement level Δ^*.

When the replacement costs $R_1 \ge R_2$ are constants, the solution to this optimization problem is simple if we know something about the monotonicity of W_y on the diagonal.

THEOREM 4.1. *Assume that R_1 and R_2 are constant and let $W = W(x, y)$ satisfying the hypotheses of Theorem 3.4 be the auxiliary value function. Assume that*

$$z \mapsto W_y(z, z) \quad \textit{is nonincreasing for } z \ge 0. \tag{20}$$

Then the optimal replacement level is given by

$$\Delta^* = \inf\{z \ge 0: W_y(z, z) - (R_1 - R_2)k(z) \le 0\}, \tag{21}$$

with $\Delta^* = +\infty$ *if the set is empty.*

PROOF. When R_1 and R_2 are constants, the function to be maximized reduces to

$$\Delta \mapsto \int_0^\Delta e^{-\int_0^z k(s)\,ds}\left[W_y(z, z) - (R_1 - R_2)k(z)\right] dz - R_2.$$

Because $k(z)$ is nondecreasing and $R_1 - R_2 \ge 0$, the integrand is nonincreasing and it is optimal to stop as soon as the integrand is less than or equal to 0. □

If the drift f and profit rate h satisfy reasonable conditions, then (20) will hold.

LEMMA 4.2. *Assume that the state equation* (1) *admits a unique, strong solution for any admissible control policy* u. *Assume also that, in addition to Conditions* 1 *and* 2, h *and* f *satisfy*:

(i) $h = h(x, y, u)$ *is nonincreasing in* x *and* y *for fixed* u.
(ii) $f = f(x, y, u)$ *is nondecreasing in* x *and* y *for fixed* u.

Then

$$z \mapsto W_y(z, z) \quad \textit{is nonincreasing}.$$

PROOF. It is sufficient to show that

$$W(z_1, z_1 + \delta) - W(z_2, z_2 + \delta) \ge 0 \tag{22}$$

for all $0 \le z_1 < z_2$ and $\delta > 0$. Let $u^2 = (u_t^2,\ 0 \le t < \infty)$ be an admissible control and let (x^2, y^2) denote the corresponding state if the initial conditions are $(x_0, y_0) = (z_2, z_2 + \delta)$. Use the same control process when the initial conditions are $(x_0, y_0) = (z_1, z_1 + \delta)$ and denote the state by (x^1, y^1).

The monotonicity assumption (ii) implies that

$$f(x + z_1, z_1 + \delta, u) \le f(x + z_2, z_2 + \delta, u) \quad \text{for all } x \text{ and } u.$$

A comparison theorem for stochastic differential equations [see, e.g., Section 5.2C in Karatzas and Shreve (1988)] implies that

$$P\left(x_t^1 - z_1 \le x_t^2 - z_2 \text{ for all } t \ge 0\right) = 1 \tag{23}$$

and hence

$$\theta_1 = \inf\{t \ge 0;\ x_t^1 - z_1 = \delta\} \ge \theta_2 = \inf\{t \ge 0;\ x_t^2 - z_2 = \delta\}.$$

The monotonicity of h then implies that

$$I(z_1, z_1 + \delta; u^2) - I(z_2, z_2 + \delta; u^2) \ge 0.$$

The control process u^2 was arbitrary, so we have shown that the original auxiliary value starting at $(z_1, z_1 + \delta)$ is at least as large as the auxiliary value starting at $(z_2, z_2 + \delta)$. This implies (22) and completes the proof. □

REMARK 4.3. The assumptions on f and h have natural interpretations for applications in controlled wear. Assumption (i) says that the revenue rate decreases as the system ages while assumption (ii) says that an older system has a higher wear rate.

4.1. *Examples.* We present two simple examples to illustrate our approach. In the first example, we assume in addition to Conditions 1–5 that:

1. the profit rate $h(x, y, u) \equiv h > 0$;
2. the failure rate $k(y) \equiv k > 0$;
3. the failure cost $R_1(y) \equiv R_1 > 0$;
4. the drift rate $f(x, y, u) \equiv u \in [u_{\min}, u_{\max}]$, with $u_{\min} > 0$;
5. the constant diffusion coefficient $\sigma \equiv 1$.

It follows from the results in Section 3 that the optimal control u^* is in fact constant for this case. We will show that the optimal replacement level Δ^* is either 0 or $+\infty$, so it may be optimal to work the system until it fails.

For this special example, the single-cycle payoff with replacement level Δ and control u is

$$J(0,0; u, \Delta) = h \cdot E[\zeta(u) \wedge \tau(\Delta)].$$

Thus it is optimal to simply maximize the survival time by choosing $u^* \equiv u_{\min}$ [because $(h > 0)$].

The optimal replacement level Δ^* is determined by

$$\begin{aligned} J(0,0; u^*, \Delta^*) &- \overline{R}(\Delta^*) \\ &= \max_{\Delta \geq 0} \big[h \cdot E[\zeta(u_{\min}) \wedge \tau(\Delta)] - R_1 \\ &\qquad + (R_1 - R_2(\Delta)) P(y(\zeta) > \Delta) \big] \\ &= \max_{\Delta \geq 0} \left[\frac{h}{k \cdot u_{\min}} (1 - e^{-k\Delta})(R_1 - R_2(\Delta))(1 - e^{-k \cdot \Delta}) - R_2(\Delta) \right] \\ &= \max_{\Delta \geq 0} \left[\left(\frac{h}{k \cdot u_{\min}} - R_1 \right)(1 - e^{-k\Delta}) - R_2(\Delta) e^{-k \cdot \Delta} \right]. \end{aligned}$$

We used Lemma 3.6 to compute the expected failure/replacement time. So, if h satisfies

$$h > k \cdot u_{\min} \cdot R_1,$$

then $\Delta^* = +\infty$ and the optimal policy is to always work at the slowest rate possible and always work until the system fails. If, on the other hand, h is too small, then $\Delta^* = 0$ and it is optimal to stop immediately.

In a second example, we obtain a finite (nonzero) preventive replacement level. Let $U = [u_{\min}, u_{\max}]$ with $u_{\min} > 0$ be the control set. Assume, in addition to Conditions 1–5, that:

1. the profit rate $h(x, y, u) = h(y, u) = a(u) - b(y)$, where a and b are positive, increasing functions;
2. the drift rate $f(x, y, u) = f(u)$, where f is a strictly positive, increasing function;
3. the ratio $a(u)/f(u)$ is increasing on U;
4. the replacement costs R_1 and R_2 are constants;
5. the diffusion coefficient $\sigma \equiv 1$.

This is an example of a "pure" running max problem described in Heinricher and Stockbridge (1991a). When the running profit and the drift and diffusion coefficients depend on only u and y, the optimal control policy is constant on excursions and the auxiliary problems are particularly easy to solve.

For this special example, it is not difficult to show that the function

$$W(x, y) = [a(u^*) - b(y)] \frac{(y - x)}{f(u^*)},$$

satisfies (15) when the feedback control u^* maximizes

$$\frac{a(u) - b(y)}{f(u)} = \frac{a(u)}{f(u)} - \frac{b(y)}{f(u)}$$

with respect to $u \in [u_{\min}, u_{\max}]$. Our assumptions imply that this ratio is increasing for each y, and hence the maximum occurs at $u^*(y) = u_{\max}$ (for all y).

By Theorem 3.4, this provides a formula for the full value function $V(x, y)$. In this case, we are interested in the value with $x = y = 0$, which gives us

$$\begin{aligned} J(0,0; u^*, \Delta) &= V(0,0) \\ &= \int_0^{\Delta} e^{-\int_0^z k(s)\,ds} \left(\frac{a(u_{\max}) - b(z)}{f(u_{\max})} \right) dz. \end{aligned}$$

The optimal replacement threshold is obtained by maximizing

$$\begin{aligned} \Delta \mapsto\ & J(0,0; u^*, \Delta) - \overline{R}(\Delta) \\ &= \int_0^{\Delta} e^{-\int_0^z k(s)\,ds} \left[\frac{a(u_{\max}) - b(z)}{f(u_{\max})} - (R_1 - R_2) k(z) \right] dz - R_2 \end{aligned}$$

over $\Delta \geq 0$. The integrand is decreasing with z and, as in Theorem 4.1, the optimal replacement policy stops as soon as the integrand is negative. That is,

$$\Delta^* = \inf \left\{ t > 0 \colon \frac{a(u_{\max}) - b(z)}{f(u_{\max})} - (R_1 - R_2) k(z) \leq 0 \right\}. \tag{24}$$

So it is always optimal to work at the maximum rate and it is optimal to replace at Δ^* defined in (24).

5. Long-run average problem. We close with a brief discussion of how to use the results of the previous sections to solve the problem where the objective is to maximize the long-run average profit. This problem will be studied in greater detail, and with different tools, in the second part of this work [Heinricher and Stockbridge (1993)].

Assume that the control process is renewed with the state. If one has a control in feedback form, this is automatic. The long-run average payoff can be represented, using the theory of renewal processes, as the reward per cycle divided by the cycle length. Let $\zeta(u)$ denote the failure time if the initial state is $(x, y) = (0, 0)$ and the control policy is u. The long-run average payoff for a policy u and replacement level Δ takes the form

$$\Lambda(u, \Delta) = \frac{E\left[\int_0^{\zeta(u) \wedge \tau(\Delta)} h(x_t, y_t, u_t)\, dt - R(\Delta)\right]}{E[\zeta(u) \wedge \tau(\Delta)]}, \tag{25}$$

and we denote the optimal value for the long-run average problem by

$$\lambda^* = \max_{u \in \mathscr{A},\, \Delta \geq 0} \Lambda(u, \Delta). \tag{26}$$

As in Heinricher and Stockbridge (1991b) [see also Taylor (1975) and Aven and Bergman (1986)], we transform the long-run average problem into a parametrized family of single-cycle problems. For any admissible pair (u, Δ), we have

$$\Lambda(u, \Delta) = \frac{E\left[\int_0^{\zeta(u) \wedge \tau(\Delta)} h(x_t, y_t, u_t)\, dt - R(\Delta)\right]}{E[\zeta(u) \wedge \tau(\Delta)]} \leq \lambda^* \tag{27}$$

and thus

$$E\left[\int_0^{\zeta(u) \wedge \tau(\Delta)} (h(x_t, y_t, u_t) - \lambda^*)\, dt\right] - \overline{R}(\Delta) \leq 0. \tag{28}$$

A long-run optimal pair (u^*, Δ^*) satisfies both (27) and (28) with equality. Thus (u^*, Δ^*) is also optimal for the problem:

$$\underset{u \in \mathscr{A},\, \Delta \geq 0}{\text{Maximize}} \left[E\int_0^{\zeta(u) \wedge \tau(\Delta)} (h(x_t, y_t, u_t) - \lambda^*)\, dt - \overline{R}(\Delta)\right], \tag{29}$$

and the optimal value is 0.

The obvious difficulty is that λ^* is not known a priori, so we work with a family of control problems parametrized by λ and determine the optimal value λ^* iteratively. For $\lambda \in \mathbb{R}$, define a single-cycle control problem

$$J(\lambda, u, \Delta) - \overline{R}(\Delta) = E\left[\int_0^{\zeta(u) \wedge \tau(\Delta)} (h(x_t, y_t, u_t) - \lambda)\, dt - R(\Delta)\right].$$

Thus for each λ, we have a single-cycle problem which can be solved using the results in Theorem 3.4 to determine $u^* = u^*(\lambda)$ and Theorem 4.1 to determine $\Delta^* = \Delta^*(\lambda)$. The optimal value λ^*, as well as the long-run optimal control and replacement level can be determined iteratively.

Acknowledgments. It is a pleasure to acknowledge helpful discussions with Steven E. Shreve during the development of this work. We would also like to thank the referee and Associate Editor for several improvements in the mathematics and the manuscript.

REFERENCES

ANDERSON, R. S. (1988). Replacement with nonconstant operating cost. *SIAM J. Control Optim.* **26** 1076–1098.

ANDERSON, R. S. (1990). Long-run average maintenance problems. Preprint.

AVEN, T. and BERGMAN, B. (1986). Optimal replacement times—A general set-up. *J. Appl Probab.* **23** 432–442.

BARRON, E. M. (1991). The Bellman equation for control of the running max of a diffusion and applications to look-back options. Preprint.

ÇINLAR, E. (1972). Markov additive processes, I, II. *Z. Wahrsch. Verw. Gebiete* **24** 85–121.

ÇINLAR, E. (1977). Shock and wear models and Markov additive processes. In *Theory and Applications of Reliability* **1**. Academic, New York.

ÇINLAR, E. (1984). Markov and semi-Markov models for deterioration. In *Reliability Theory and Models*. Academic, New York.

CONRAD, C. and MCCLAMROCH, N. H. (1987). The drilling problem: A stochastic modeling and control example in manufacturing. *IEEE Trans. Automat. Control* **32** 947–958.

FLEMING, W. H. and RISHEL, R. W. (1975). *Deterministic and Stochastic Optimal Control.* Springer, New York.

HEINRICHER, A. C. and STOCKBRIDGE, R. H. (1991a). Optimal control of the running max. *SIAM J. Control Optim.* **29** 936–953.

HEINRICHER, A. C. and STOCKBRIDGE, R. H. (1991b). Long-term average control for a continuous, monotone process. *J. Appl. Math. Optim.* To appear.

HEINRICHER, A. C. and STOCKBRIDGE, R. H. (1992). An infinite-dimensional LP solution to control of a continuous, monotone process. In *Applied Stochastic Analysis* (I. Karatzas, D. Ocone, eds.) *Lecture Notes in Control and Information Sciences* **177** 134–143. Springer, Berlin.

HEINRICHER, A. C. and STOCKBRIDGE, R. H. (1993). Optimal control and replacement with state-dependent failure rate: An invariant measure approach. *Ann. Appl. Probab.* **3** 380–402.

KARATZAS, I. and SHREVE, S. E. (1987). *Brownian Motion and Stochastic Calculus*. Springer, New York.

LIONS, P. L. (1983)). Optimal control of diffusions and Hamilton–Jacobi equations, Part I: the dynamic programming principle and applications. *Comm. Partial Differential Equations* **8** 1101–1174.

TAYLOR, H. M. (1975). Optimal replacement under additive damage and other failure models. *Naval Res. Log.* **22** 1–18.

VALDEZ-FLORES, C. and FELDMAN, R. M. (1989). A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Res. Log.* **36** 419–446.

DEPARTMENT OF MATHEMATICS
WORCESTER POLYTECHNIC INSTITUTE
WORCESTER, MASSACHUSETTS 01609

DEPARTMENT OF STATISTICS
UNIVERSITY OF KENTUCKY
LEXINGTON, KENTUCKY 40506