Stochastic Processes and their Applications 8 (1978) 211–227.

CONTROLLED ONE DIMENSIONAL DIFFUSIONS WITH SWITCHING COSTS—AVERAGE COST CRITERION

Bharat T. DOSHI

Department of Statistics, Rutgers University, New Brunswick, N.J. 08903, U.S.A.

Received 12 April 1978
Revised 1 September 1978

This paper deals with a one-dimensional controlled diffusion process on a compact interval with reflecting boundaries. The set of available actions is finite and the action can be changed only at countably many stopping times. The cost structure includes both a continuous movement cost rate depending on the state and the action, and a switching cost when the action is changed. The policies are evaluated with respect to the average cost criterion. The problem is solved by looking at, for each stationary policy, an embedded stochastic process corresponding to the state intervals visited in the sequence of switching times. The communicating classes of this process are classified into closed and transient groups and a method of calculating the average cost for the closed and transient classes is given. Also given are conditions to guarantee the optimality of a stationary policy. A Brownian motion control problem with quadratic cost is worked out in detail and the form of an optimal policy is established.

Controlled diffusion   reflecting boundaries
switching costs

1. Introduction

In a recent paper Doshi [4] studied a controlled one dimensional diffusion process on a compact interval of real line with reflecting boundaries. The cost structure included both, a continuous movement cost rate depending on the current state and action, and a switching cost incurred when the action is changed. The set of available actions is finite and any change in action is allowed only at countably many stopping times. Thus a policy specifies a sequence of stopping times and also the action to be used between two consecutive stopping times as a function of the history of the process. For the continuously discounted costs Doshi [4] obtained the necessary and sufficient conditions for a stationary policy to be optimal and also a set of inequalities satisfied by the discounted cost function of an optimal policy. A method of calculating the discounted cost function from any stationary policy is also given. In this paper we study the same model with respect to the long run expected cost per unit time (average cost) as the economic criterion. There are two basic goals:

(a) to compute the average cost function for any stationary policy, and

(b) to derive conditions which are sufficient for a stationary policy to be optimal.

The model is formulated in Section 2. Section 3 deals with the computation of the average cost function for any stationary policy. A discrete-time stochastic process is embedded into the diffusion process operating under a stationary policy and the communicating classes of this process are classified into closed and transient categories. It is shown that the average cost of the diffusion process is independent of the initial conditions in a given closed class of this process and that this average cost can be obtained as the unique solution of a functional equation involving the average cost and a bias function. We also give a formal interpretation of this bias function. For the initial conditions in a transient class we obtain the average cost function in terms of the average costs of the closed classes. It is shown that in most common situations a stationary optimal policy can have only one closed class or two or more closed classes with the same average cost. The functional equations are considerably simplified in this case and naturally lead to the optimality conditions of Section 4. In Section 5 we consider an example of a controlled Brownian motion with reflecting boundaries and show that the intuitively optimal policy satisfies the optimality conditions, thus proving its optimality.

2. Model and basic assumptions

The controlled diffusion problem mentioned in section 1 has been formulated as a continuous-time Markov decision problem in Doshi [4]. For completeness, however, we state the basic elements of this model. For details and motivations behind various definitions the reader is referred to Doshi [4].

The state space $\mathscr{X} = [r_0, r_1]$, $-\infty < r_0 < r_1 < \infty$ with the Borel σ-algebra $\beta_{\mathscr{X}}$.

The action space $\mathscr{A}$ is a finite set $\{1, 2, 3, 4, \ldots, M\}$.

The sample space Ω is the set of all functions $\omega : [0, \infty) \to \mathscr{X}$ such that ω is continuous on $t \geq 0$.

Let $\mathscr{F}_t = \sigma(\omega(s); s \leq t)$ and $\mathscr{F} = \mathscr{F}_\infty$. Let $X_t(\omega) = \omega(t)$, $t \geq 0$ be the co-ordinate mappings and $\{\theta_t; t \geq 0\}$ be the translation operators on Ω. Also for any $\{\mathscr{F}_t\}$-stopping time τ let $\mathscr{F}_\tau$ be the usual σ-algebra associated with a stopping time.

Assumption 1. For each $a \in \mathscr{A}$, there exists a diffusion process $\{X_t^a; t \geq 0\}$, with state space $\mathscr{X}$, such that

(i) The diffusion coefficient $d(x, a)$ and the drift coefficient $b(x, a)$ are continuous in $x \in (r_0, r_1)$ for each $a \in \mathscr{A}$, and there exist $0 < M_1 < M_2 > \infty$ and $M_3 < \infty$ such that $0 < M_1 \leq d(x, a) \leq M_2$, $|b(x, a)| \leq M_3$ for all $x \in \mathscr{X}$, $a \in \mathscr{A}$.

(ii) Both boundaries of $\{X_t^a\}$ are reflecting.

Definition 1. An *admissible policy* π is an alternating sequence $(\tau_1, a_1\tau_2, a_2, \ldots)$ such that

(i) For $n \geq 1$, τ_n is an $\{\mathscr{F}_t\}$-stopping time depending on a_{n-1}.

(ii) $0=\tau_0\leqslant\tau_1\leqslant\tau_2\cdots$ a.s.

(iii) $a_0=a$ is given. For $n\geqslant 1$ a_n is an $\mathcal{F}_{\tau_n}\times\beta_{\mathcal{A}}$-measurable function taking value in $\mathcal{A}$.

(iv) For $t\geqslant 0$, let $N_t=\sum_0^\infty I_{\{\tau_n\in[t,t+1)\}}$. Then there exists an $N<\infty$ such that $E_{\pi,x,a}[N_t]\leqslant N$ for all $t\geqslant 0$.

(v) $a_t=a_n$ is the action used at time t for $t\in[\tau_n,\tau_{n+1})$, $n\geqslant 0$.

(vi) $X_{\tau_n^-}=X_{\tau_n}=X_{\tau_n^+}$ for all n.

Let D denote the set of all admissible policies.

Stationary policies

Definition 2. Let $f:\mathcal{X}\times\mathcal{A}\to\mathcal{A}$ be a piecewise constant function on $\mathcal{X}$, for each $a\in\mathcal{A}$, such that for each $a\in\mathcal{A}$ the following hold:

(i) $\Lambda^a=\{x\in\mathcal{X};f(x,a)=a\}$ is an open subset of $[r_0,r_1]$ for each $a\in\mathcal{A}$.

(ii) If x is on the boundary of Λ^a, then $f(x,a)=f(y,f(x,a))$ for all y in some neighborhood of x.

Then f defines a *stationary policy* π by

(iii) $\tau_0=0$, $a_0=a$ are given,

(iv) $\tau_{n+1}=\tau_n+\tau^{a_n}\cdot\theta_{\tau_n}$ $(n\geqslant 1)$, where $\tau^a=\inf\{t\geqslant 0;X_t\notin\Lambda^a\}=\inf\{t\geqslant 0;f(X_t,a)\neq a\}$, $a\in\mathcal{A}$, and

(v) $a_{n+1}=f(X_{\tau_{n+1}},a_n)$, $n\geqslant 0$.

Using Assumption 1 it can be easily verified that a stationary policy π as defined above is admissible. The sets $\{\Lambda^a;a\in\mathcal{A}\}$ are called its *continuation sets* and the function f, its *action selecting function*. Let D_S denote the set of all stationary admissible policies.

Cost structure. There are two types of costs involved, *continuous* and *lump*. Let $c_a(x)\equiv c(x,a)$ be the cost rate when action a is used in state x. We assume that for each $a\in\mathcal{A}$, $c(\cdot,a)$ is continuous in $x\in\mathcal{X}$. Let $R(x,a,a')$ denote the non-negative lump cost incurred when the action is changed from a to a' while in state x. Assume that $R(x,a,a)=0$ for all $x\in\mathcal{X}$, $a\in\mathcal{A}$, $R(x,a,a')>0$ for all $x\in\mathcal{X}$ and $a\neq a'$, and

$$R(x,a,a'')\leqslant R(x,a,a')+R(x,a',a'') \tag{2.1}$$

for all $x\in\mathcal{X}$, a, a', $a''\in\mathcal{A}$. Also assume that $R(\cdot,a,a')$ is twice continuously differentiable for each a, $a'\in\mathcal{A}$.

Economic criterion. For $x\in\mathcal{X}$, $a\in\mathcal{A}$, $n\geqslant 0$ and $\pi\in D$, define

$$V_{\pi,a}(x;n)=E_{\pi,x,a}\left[\int_0^{\tau_n}c(X_t,a_t)\,\mathrm{d}t+\sum_{k=1}^{n}R(X_{\tau_k},a_{k-1},a_k)\right]. \tag{2.2}$$

If $E_{\pi,x,a}[\tau_n]<\infty$ for all $n\geq 1$, then we define the *long run expected cost per unit time* (*average cost* for short) by

$$V_{\pi,a}(x)=\limsup_{n\to\infty}\frac{V_{\pi,a}(x;n)}{E_{\pi,x,a}[\tau_n]}. \tag{2.3}$$

On the other hand, if $E_{\pi,x,a}[\tau_n]=\infty$ for some $n\geq 1$, then we define the average cost by

$$V_{\pi,a}(x)=\limsup_{t\to\infty}\left\{E_{\pi,x,a}\left[\int_0^t c(X_s,a_s)\,\mathrm{d}s+\sum_{\tau_k\leq t}R(X_{\tau_k},a_{k-1},a_k)\right]\Big/t\right\}. \tag{2.4}$$

When $E_{\pi,x,a}[\tau_n]<\infty$ for all $n\geq 1$, it is clear that $V_{\pi,a}(x)$ defined by (2.3) and (2.4) are the same. A function $V(\cdot)$ is called *optimal average cost* if $V_a(x)=\inf_{\pi\in D}V_{\pi,a}(x)$ $(x\in\mathcal{A},\ a\in\mathcal{A})$. A policy π^* *is optimal if* $V_{\pi^*,a}(x)=V_a(x)$ for all $x\in\mathscr{X}$, $a\in\mathcal{A}$.

Infinitesimal operator and its domain

Definition 3. Let $\mathscr{D}$ be the set of functions $g:\mathscr{X}\to R$ satisfying the following:
(i) g is continuous on $[r_0,r_1]$ except at finitely many points.
(ii) g' and g'' exist and are continuous on (r_0,r_1) except at finitely many points.
(iii) $g'(r_i)=0$, $i=0,1$.

For $g\in\mathscr{D}$, define the *infinitesimal operator* of the diffusion process under action $a\in\mathcal{A}$ by

$$A_a g(x)=d(x,a)g''(x)+b(x,a)g'(x) \tag{2.5}$$

if g' and g'' are continuous at x, and by

$$A_a g(x)=\tfrac{1}{2}d(x,a)[g''(x^+)+g''(x^-)]+b(x,a)g'(x) \tag{2.6}$$

if g' is continuous at x. $A_a g$ is not defined at the points of discontinuity of g'.

Let $\mathscr{D}'$ be the subset of $\mathscr{D}$ such that for $g\in\mathscr{D}'$, g' is continuous on (r_0,r_1), that is, $A_a g$ is defined for all $x\in(r_0,r_1)$.

3. Average cost for stationary policies

The main content of this section is the calculation of the average cost and the bias function (to be defined below) for an arbitrary stationary policy π. We will also give a simple interpretation of the bias function which will motivate the sufficient conditions of Section 4.

For a given stationary policy, π, determined by $\{\Lambda^a;\ a\in\mathcal{A}\}$ and f, each Λ^a is the union of a finite number of disjoint open subintervals [note that r_0 and r_1 are interior points of $[r_0,r_1]$). Let $\lambda(i,a)$ denote the ith subinterval of Λ^a, and let n_a denote the number of subintervals of Λ^a.

Now consider the embedded stochastic process $\{Y_n;\ n=0,1,2,\ldots\}$ with the state space $\{\lambda(i,a);\ i=1,2,\ldots,n_a,\ a=1,2,\ldots,M\}$. Although Y is not necessarily a

Markov chain, the state space can be decomposed into closed communicating classes and transient states in exactly the same way as for Markov chains. To see this, note that Y_n is the nth subinterval visited by the Y process; so $Y_n = \lambda(i, a_n)$ for some i. Moreover, $a_{n+1} = f(X_{\tau_{n+1}}, a_n)$ where $X_{\tau_{n+1}}$ is the endpoint of $\lambda(i, a_n)$ reached first; so there are at most two possibilities for the value of Y_{n+1}. This situation is analogous to a finite state Markov chain with at most two positive elements in each row of the transition matrix. Using this analogy, one can decompose the state space into P, say, closed classes $C_1, C_2, \ldots, C_p$ of communicating states and a set T of transient states. To simplify notations below, we write $(x, a) \in C$ for some class C if there exists some subinterval $\lambda(i, a) \in C$ such that $x \in \lambda(i, a)$.

We now show that the average cost $V_{\pi,a}(x)$ is constant for all $(x, a) \in C$, where C is a closed class, and can be obtained as a solution to a differential equation which also gives a family $\{h_a;\, a \in \mathscr{A}\}$ of functions on $\mathscr{X}$. For $a \in \mathscr{A}$, let $C^a \equiv \{x \in \mathscr{X};\, (x, a) \in C\}$ be the a-section of C.

Theorem 3.1. *Let C be a closed class for a stationary policy π determined by $\{\Lambda^a;\, a \in \mathscr{A}\}$ and f. Then there exists a unique constant g and a family $\{h_a;\, a \in \mathscr{A}\}$ with h_a defined on C^a into R, such that the following hold:*

(i) *h_a is continuous on C^a for each $a \in \mathscr{A}$.*

(i) *h'_a and h''_a exist and are continuous on C^a for each $a \in \mathscr{A}$.*

(iii) *For $(x, a) \in C$*

$$g = c(x, a) + A_a h_a(x). \tag{3.2}$$

(iv) *If $\lambda(i, a) \in C$ and $y = A(i, a)$, $r_0 \notin \lambda(i, a)$ or $y = B(i, a)$, $r_1 \notin \lambda(i, a)$, then*

$$h_a(y) = \lim_{\substack{x \to y \\ x \in \Lambda^a}} h_a(x) = R(y, a, f(y, a)) + h_{f(y,a)}(y). \tag{3.3}$$

(v) $$h'_a(r_i) = 0 \quad \text{if } r_i \in C^a,\ i = 0, 1. \tag{3.4}$$

(vi) $$g = V_{\pi,a}(x) \quad ((x, a) \in C).$$

Proof. We first prove that there exist a constant g and a family $\{h_a;\, a \in \mathscr{A}\}$ satisfying (i)–(v) above. Let $M(x, a) = \int^x b(y, a)/d(y, a)\,\mathrm{d}y$. Consider equation (3.2) for $x \in \lambda(i, a) \in C$, that is, the ith subinterval in Λ^a. This may be rewritten as,

$$d(x, a)\,\mathrm{e}^{-M(x,a)} \frac{\mathrm{d}}{\mathrm{d}x}[\mathrm{e}^{M(x,a)} h'_a(x)] = g - c(x, a) \tag{3.5}$$

the solution of which is

$$h_a(x) = c_{i,a} + d_{i,a} \int_{A(i,a)}^{x} \mathrm{e}^{-M(y,a)}\,\mathrm{d}y + \int_{A(i,a)}^{x} \mathrm{e}^{-M(y,a)} \int_{A(i,a)}^{y} \frac{g - c(s, a)}{d(s, a)}\,\mathrm{e}^{M(s,a)}\,\mathrm{d}s\,\mathrm{d}y \tag{3.6}$$

for $x \in \lambda(i,a) = (A(i,a), B(i,a))$ and some constants $c_{i,a}$ and $d_{i,a}$. Suppose $f(A(i,a),a) = a'$, $f(B(i,a),a) = a''$, $A(i,a) \in \lambda(j,a')$ and $B(i,a) \in \lambda(k,a'')$. Then equation (3.3) gives

$$c_{i,a} - c_{j,a'} - d_{j,a'} \int_{A(j,a')}^{A(i,a)} e^{-M(y,a')}\,dy - g \int_{A(j,a')}^{A(i,a)} e^{-M(y,a')} \int_{A(j,a')}^{y} \frac{e^{M(s,a')}}{d(s,a')}\,ds\,dy$$

$$= R(A(i,a),a,a') - \int_{A(j,a')}^{A(i,a)} e^{-M(y,a')} \int_{A(j,a')}^{y} \frac{c(s,a')\,e^{M(s,a')}}{d(s,a')}\,ds\,dy, \qquad (3.7)$$

and

$$c_{i,a} + d_{i,a} \int_{A(i,a)}^{B(i,a)} e^{-M(y,a)}\,dy - c_{k,a''} - d_{k,a''} \int_{A(k,a'')}^{B(i,a)} e^{-M(y,a'')}\,dy$$

$$+ g\left\{ \int_{A(i,a)}^{B(i,a)} e^{-M(y,a)} \int_{A(i,a)}^{y} \frac{e^{M(s,a)}}{d(s,a)}\,ds\,dy \right.$$

$$\left. - \int_{A(k,a'')}^{B(i,a)} e^{-M(y,a'')} \int_{A(k,a'')}^{y} \frac{e^{M(s,a'')}}{d(s,a'')}\,ds\,dy \right\} =$$

$$= R(B(i,a),a,a'') + \int_{A(i,a)}^{B(i,a)} e^{-M(y,a)} \int_{A(i,a)}^{y} \frac{c(s,a)}{d(s,a)}\,ds\,dy$$

$$- \int_{A(k,a'')}^{B(i,a)} e^{-M(y,a'')} \int_{A(k,a'')}^{y} \frac{c(s,a'')}{d(s,a'')} e^{M(s,a'')}\,ds\,dy. \qquad (3.8)$$

If $r_0 \in \lambda(i,a)$, then equation (3.4) requires that

$$d_{i,a} = 0. \qquad (3.9)$$

Finally, if $r_1 \in \lambda(i,a)$, then equation (3.4) reduces to

$$d_{i,a}\, e^{-M(r_1,a)} - g\, e^{-M(r_1,a)} \int_{A(i,a)}^{r_1} \frac{e^{M(s,a)}}{d(s,a)}\,ds$$

$$= e^{-M(r_1,a)} \int_{A(i,a)}^{r_1} \frac{c(s,a)}{d(s,a)} e^{M(s,a)}\,ds. \qquad (3.10)$$

Suppose the closed class C has n_C elements. Then (3.7)–(3.10) give $2n_C$ linear equations in $2n_C+1$ unknowns $\{c_{i,a}, d_{i,a}, g\}$. Consider the matrix formed by the n_C columns corresponding to the coefficients of $c_{i,a}$. Since C is a closed communicating class this matrix has rank $n_C - 1$. Other n_C+1 columns are linearly independent. So the rank of the coefficient matrix is $2n_C$. This implies that the above $2n_C$ equations have a solution; g and $d_{i,a}$'s are uniquely determined by these equations and $c_{i,a}$'s are determined up to a constant. So equations (3.2)–(3.4) have a solution $(g, \{h_a;\, a \in \mathscr{A}\})$ such that

(i) h_a is continuous on C^a for $a \in \mathscr{A}$.
(ii) h_a' and h_a'' exist and are continuous on C^a for $a \in \mathscr{A}$.
(iii) g is uniquely determined by equations (3.2)–(3.4).
(iv) h_a are determined up to an additive constant.

This proves parts (i)-(v) of the theorem. We now show part (vi). Suppose (g, h_a) be any solution to equations (3.2)-(3.4). Suppose $(x, a) \in C$ and C contains at least two elements. Then

$$V_{\pi,a}(x; n) = E_{\pi,x,a}\left[\sum_{k=1}^{n} \int_{\tau_{k-1}}^{\tau_k} c_{a_{k-1}}(X_t)\,dt + \sum_{k=1}^{n} R(X_{\tau_k}, a_{k-1}, a_k)\right]$$

$$= E_{\pi,x,a}\left[\sum_{k=1}^{n} \int_{\tau_{k-1}}^{\tau_k} (g - A_{a_{k-1}} h_{a_{k-1}}(X_t))\,dt + \sum_{k=1}^{n} R(X_{\tau_k}, a_{k-1}, a_k)\right]. \tag{3.11}$$

By Dynkin [6, Corollary to Theorem 5.1]

$$E_{\pi,x,a}\left[\int_{\tau_{k-1}}^{\tau_k} A_{a_{k-1}} h_{a_{k-1}}(X_t)\,dt\right]$$

$$= E_{\pi,x,a}[h_{a_{k-1}}(X_{\tau_k}) - h_{a_{k-1}}(X_{\tau_{k-1}})]$$

$$= E_{\pi,x,a}[h_{a_k}(X_{\tau_k}) - h_{a_{k-1}}(X_{\tau_{k-1}}) + R(X_{\tau_k}, a_{k-1}, a_k)]. \tag{3.12}$$

Substituting from (3.12) into (3.11) we obtain

$$\frac{V_{\pi,a}(x; n)}{E_{\pi,x,a}[\tau_n]} = g - \frac{E_{\pi,x,a}[h_{a_n}(X_{\tau_n}) - h_a(x)]}{E_{\pi,x,a}[\tau_n]}. \tag{3.13}$$

Letting $n \to \infty$

$$V_{\pi,a}(x) = g. \tag{3.14}$$

Similarly, when C contains only one element $\lambda(1, a) = [r_0, r_1]$ it is easy to show that for $x \in [r_0, r_1]$,

$$V_{\pi,a}(x) = g = \lim_{t\to\infty} \frac{E_{\pi,x,a}\left[\int_0^t c_a(X_t)\,dt\right]}{t}, \tag{3.15}$$

thus proving the theorem.

The intent of the next corollary is to interpret the family $\{h_a; a \in \mathcal{A}\}$ with h_a defined on C^a.

Corollary 3.1. *For* $(x, a) \in C$ *let* $T(x, a) = \inf\{t \geq 0; X_t = x, a_t = a\}$. *Let* $(x_0, a_0) \in C$ *be fixed. Then for any initial conditions* $(x, a) \in C$

$$h_a(x) - h_{a_0}(x_0) = E_{\pi,x,a}[V_{\pi,a}(x; T(x_0, a_0))] - g E_{\pi,x,a}[T(x_0, a_0)], \tag{3.16}$$

where, for any $T > 0$,

$$V_{\pi,a}(x; T) = E_{\pi,x,a}\left[\int_0^T c(X_t, a_t)\,dt + \sum_{\tau_k \leq T} R(X_{\tau_k}, a_{k-1}, a_k)\right]. \tag{3.17}$$

Proof. Let $N=\max\{n;\tau_n\leq T(x_0,a_0)\}$, where τ_n is the nth stopping time in the definition of π. Then, as in the proof of part (vi) of Theorem 3.1, we have

$$
\begin{aligned}
&E_{\pi,x,a}[V_{\pi,a}(x;T(x_0,a_0))]\\
&=E_{\pi,x,a}\left[\sum_{k=1}^{N}\int_{\tau_{k-1}}^{\tau_k}c(X_t,a_{k-1})\,\mathrm{d}t+\int_{\tau_N}^{T(x_0,a_0)}C(X_t,a_N)\,\mathrm{d}t+\sum_{k=1}^{N}R(X_{\tau_k},a_{k-1},a_k)\right]\\
&=E_{\pi,x,a}\left[\sum_{k=1}^{N}\int_{\tau_{k-1}}^{\tau_k}[g-A_{a_{k-1}}h_{a_{k-1}}(X_t)]\,\mathrm{d}t+\int_{\tau_N}^{T(x_0,a_0)}[g-A_{a_N}h_{a_N}(X_t)]\,\mathrm{d}t\right.\\
&\qquad\left.+\sum_{k=1}^{N}R(X_{\tau_k},a_{k-1},a_k)\right]\\
&=gE_{\pi,x,a}[T(x_0,a_0)]-E_{\pi,x,a}[h_{a_N}(X_{T(x_0,a_0)})]+h_a(x)\\
&=gE_{\pi,x,a}[T(x_0,a_0)]-h_{a_0}(x_0)+h_a(x).
\end{aligned}
\tag{3.18}
$$

This proves the Corollary.

The above corollary suggests two different interpretations of $h_a(x)$:

(a) Since $h_a(x)$ is determined up to an additive constant it is possible to take $h_{a_0}(x_0)=0$ Then $h_a(x)=E_{\pi,x,a}[V_{\pi,a}(x;T(x_0,a_0)]-gE_{\pi,x,a}[T(x_0,a_0)]$. The first term on the right is the expected total cost in time interval $[0,T(x_0,a_0)]$ if the initial condition is (x,a). The second term is the expected cost in time interval $[0,T(x_0,a_0)]$ if the cost is incurred at the steady state rate g. Thus $h_a(x)$ is the *bias* due to the initial condition (x,a). In fact, using Corollary 3.1 it can be shown that

$$h_a(x)=\lim_{t\to\infty}[V_{\pi,a}(x;t)-gt]+\text{constant}\quad(x\in\mathscr{X},a\in\mathscr{A}).\tag{3.19}$$

Thus $h_a(x)$ is the *limiting bias* due to the initial condition (x,a). This justifies calling $\{h_a;a\in\mathscr{A}\}$ the *family of bias functions.*

(b) From equation (3.19) we get, for $(x,a)\in C$ and $(x',a')\in C$,

$$h_a(x)-h_{a'}(x')=\lim_{t\to\infty}[V_{\pi,a}(x,t)-V_{\pi,a'}(x';t)].\tag{3.20}$$

Thus $[h_a(x)-h_{a'}(x')]$ can be interpreted as the limiting relative disadvantage of starting in condition (x,a) rather than in (x',a'). This justifies calling $\{h_a;a\in\mathscr{A}\}$ the *family of disadvantage functions* or as customary in literature [see [1, 2]] the *family of potential functions.*

Theorem 3.1 gives a method of deriving the average cost corresponding to any closed class C. Suppose there are P closed classes $C_1,C_2,\ldots,C_p$ and let T be the collection of all transient states. Then from Theorem 3.1 we can derive the average costs $g_1,g_2,\ldots,g_p$ corresponding to P closed classes. We also get $\{h_a;a\in\mathscr{A}\}$ where h_a is a twice continuously differentiable function defined on $\bigcup_{i=1}^{P}C_i^a$. The following Theorem can be used to calculate $V_{\pi,a}(x)$ for $(x,a)\in T$, thus completely specifying the average cost function $V_{\pi,a}(x)$ for all $x\in\mathscr{X}$ and $a\in\mathscr{A}$.

Theorem 3.2. *$\{V_{\pi,a};\, a \in \mathcal{A}\}$ is the unique family of functions in $\mathcal{D}$ satisfying*

(i) $A_a V_{\pi,a}(x) = 0$ for $(x, a) \in T \cap \Lambda^a$,

(ii) $V_{\pi,a}(x) = V_{\pi,f(x,a)}(x)$ for $x \notin \Lambda^a$, $(x, f(x, a)) \in T$,

(iii) $V_{\pi,a}(x) = V_{\pi,f(x,a)}(x) = g_k$ for $x \notin \Lambda^a$, $(x, f(x, a)) \in C_k$.

Proof. The fact that there exists a unique solution to (i) and (ii) satisfying (iii) can be proved by the method used in the proof of Theorem 3.1. So it only remains to show that this solution equals $V_{\pi,a}(\cdot)$. Let $\{V_a;\, a \in \mathcal{A}\}$ be any family of functions in $\mathcal{D}$ satisfying (i), (ii) and (iii) above. Let $\tau_1, \tau_2, \ldots$ denote the stopping times associated with policy π, and let $N = \min\{n \geq 1,\, (X_{\tau_n}, a_n) \in C_k$ for some $k = 1, 2, \ldots, P\}$. That is, τ_N is the time of first absorption into a closed class. Then from (i) and (ii) we have for $(x, a) \in T$,

$$
\begin{aligned}
0 &= E_{\pi,x,a}\left[\sum_{k=1}^{N} \int_{\tau_{k-1}}^{\tau_k} A_{a_{k-1}} V_{a_{k-1}}(X_t)\, \mathrm{d}t\right] \\
&= E_{\pi,x,a}\left[\sum_{k=1}^{N} \{V_{a_{k-1}}(X_{\tau_k}) - V_{a_{k-1}}(X_{\tau_{k-1}})\}\right] \\
&= E_{\pi,x,a}\left[\sum_{k=1}^{N} \{V_{a_k}(X_{\tau_k}) - V_{a_{k-1}}(X_{\tau_{k-1}})\}\right] \\
&= E_{\pi,x,a}[V_{a_N}(X_{\tau_N})] - V_a(x).
\end{aligned} \tag{3.21}
$$

So for $(x, a) \in T$,

$$
\begin{aligned}
V_a(x) &= E_{\pi,x,a}[V_{a_N}(X_{\tau_N})] \\
&= \sum_{k=1}^{P} g_k P_{\pi,x,a}(\{\text{absorption in class } C_k\}) = V_{\pi,a}(x).
\end{aligned} \tag{3.22}
$$

This proves the theorem.

Remark. Theorem 3.2 can be interpreted as follows. Consider a diffusion process with starts in $(x, a) \in T$ and stops at the first instant it hits $(x', a') \in C$ for some (x', a') and a closed class C. Until then it behaves exactly like the original diffusion process. No costs are incurred until the process stops. A terminal cost $V_{\pi,a'}(x')$ is incurred if the process stops in (x', a'). Since $C(\cdot, \cdot)$ is bounded and the expected time to termination is bounded, $V_{\pi,a}(x)$ must equal the expected termination cost in the modified process. Theorem 3.2 formally states this fact.

4. Optimality conditions

In this section we derive conditions which are sufficient for the existence of a stationary optimal policy. When these conditions are satisfied, it is possible to use them to derive a stationary optimal policy using Theorem 4.2 below.

Theorems 3.1 and 3.2 enable us to calculate the average cost function from any stationary policy having P closed classes and Q transient classes. A little reflection shows that if $R(x, a, a')$ is finite for all $x \in \mathscr{X}$, and if $P \geqslant 2$ with $g_i \neq g_j$ for some (i, j), then the stationary policy under consideration cannot be optimal. To see this, suppose $g_i < g_j$. Then starting in any state $(x, a) \in C_j$ one can change to action a' such that $(x, a') \in C_i$. If $R(x, a, a') < \infty$, then for the new policy π' $V_{\pi', a}(x) = g_i < g_j = V_{\pi, a}(x)$, and so π is not optimal. It is conceivable, however, that certain changes in action are not permissible in some states; that is, $R(x, a, a') = \infty$ for some $x \in \mathscr{X}$, $a, a' \in \mathscr{A}$. In this case it is possible that a stationary optimal policy has two or more closed classes with different average costs. With our assumption of the boundedness of $R(\cdot, \cdot, \cdot)$, this situation will not arise. So for optimality consideration we may restrict our attention to stationary policies with only one closed class or with two or more closed classes having the same average cost. By Theorem 3.2 we now have $V_{\pi, a}(x) = g$ for all $x \in \mathscr{X}$, $a \in \mathscr{A}$, where g is the common average cost starting in any initial conditions. In this case we have the following extension of Theorem 3.1. The proof is similar to that of Theorem 3.1.

Theorem 4.1. *Let π be any stationary policy with P closed classes $C_1, C_2, \ldots, C_p$ and the family T of transient states. If $g_1 = g_2 = \cdots = g_p = g$, then there exists a unique family $\{h_a; a \in \mathscr{A}\}$ of functions in $\mathscr{D}$ such that*

(i) *h_a is continuous on Λ^a for each $a \in \mathscr{A}$,*

(ii) *h'_a and h''_a exist and are continuous on Λ^a for each $a \in \mathscr{A}$,*

(iii) $g = c(x, a) + A_a h_a(x) \quad (x \in \Lambda^a)$,

(iv) $h_a(y) = R(y, a, f(y, a)) + h_{f(y,a)}(y) \quad (y \notin \Lambda^a)$,

(v) $h'_a(r_i) = 0,\ i = 0, 1,\ a \in \mathscr{A}$.

Moreover, the constant g is uniquely determined by (iii)–(v) *above, and h_a is determined up to an additive constant.*

The following theorem gives a lower bound on the optimal average cost $V_a(x)$. It also gives sufficient conditions for a stationary policy π^* to be optimal.

Theorem 4.2. *Suppose there exists a constant g and a family $\{h_a; a \in \mathscr{A}\}$ of functions on $\mathscr{X}$ satisfying the following:*

(i) *$h_a \in \mathscr{D}'$ for $a \in \mathscr{A}$.*

(ii) $$g \leqslant c(x, a) + A_a h_a(x) \tag{4.1}$$

at each $x \in (r_0, r_1)$ which is a continuity point of $A_a h_a$.

(iii) *For all $x \in \mathscr{X}$ and $a \in \mathscr{A}$*

$$h_a(x) \leqslant \min_{a' \in \mathscr{A}} \{R(x, a, a') + h_{a'}(x)\}.$$

Then $g \leqslant V_a(x)$ $(x \in \mathscr{X}, a \in \mathscr{A})$. *Moreover, if* g *and* $\{h_a; a \in \mathscr{A}\}$ *are associated with a stationary policy* π^* *as in Theorem* 4.1, *then*

$$g = V_a(x) = V_{\pi^*,a}(x) \quad (x \in \mathscr{X}, a \in \mathscr{A}),$$

and π^* *is optimal.*

Proof. Suppose (g, h_a) satisfy (i), (ii) and (iii) above. Let π be any admissible policy. Then following the arguments that led to equation (3.13) we obtain

$$g \leqslant \frac{V_{\pi,a}(x; n)}{E_{\pi,x,a}[\tau_n]} + \frac{E_{\pi,x,a}[h_a(X_{\tau_n}) - h_a(x)]}{E_{\pi,x,a}[\tau_n]}$$

for $x \in \mathscr{X}$, $a \in \mathscr{A}$ and $n \geqslant 1$. Letting $n \to \infty$

$$g \leqslant V_{\pi,a}(x) \quad (x \in \mathscr{X}, a \in \mathscr{A}, \pi \in D).$$

Since π is arbitrary

$$g \leqslant V_a(x) \quad (x \in \mathscr{X}, a \in \mathscr{A}).$$

This proves the first part of the theorem. The second part is obvious from Theorem 4.1.

Remarks. (a) The above theorem may be used to verify if a given stationary policy π^* which is suspected to be optimal. We can use Theorem 4.1 to find g and $\{h_a; a \in \mathscr{A}\}$ for this policy and then verify that they satisfy the hypothesis of Theorem 4.2. In Section 5 we use this approach to show the optimality of a two-switching-levels policy for a controlled Brownian motion with quadratic costs and reflecting boundaries.

(b) Note that in Theorem 4.2 we require that h_a belong to $\mathscr{D}'$ rather than just $\mathscr{D}$. That is, h'_a is required to be continuous. Using the arguments of Chernoff and Petkau [3] it can be shown that continuity of h'_a is necessary for a stationary policy π^* to be optimal. So this requirement is not superfluous or arbitrary.

(c) Let $D^* \subset D$ consist of policies under which $\{(X_t, a_t); t \geqslant 0\}$ has the same steady state distribution as under π^*. The transient behaviour may be different under different $\pi \in D^*$. Then all $\pi \in D^*$ have the same associated average cost and all of them are optimal in the average cost sense. For $\pi \in D^*$, define

$$W_{\pi,a}(x) = \lim_{t \to \infty} [V_{\pi,a}(x; t) - gt] \quad (x \in \mathscr{X}, a \in \mathscr{A}). \tag{4.2}$$

Then it follows from the proof of Theorem 4.2 and the remarks following Corollary 3.1 that

$$W_{\pi^*,a}(x) = \min_{\pi \in D^*} W_{\pi,a}(x) \quad (x \in \mathscr{X}, a \in \mathscr{A}). \tag{4.3}$$

Thus π^* defined in Theorem 4.2 lexicographically minimizes $V_{\pi,a}(x)$ over all $\pi \in D$ and minimizes $W_{\pi,a}(x)$ over all $\pi \in D^*$.

(d) For controlled Brownian motion Chernoff and Petkau [3] have shown that conditions (i)–(iii) in Theorem 4.2 are not only sufficient but are also necessary for π^* to be optimal. Their arguments can be readily applied to our general diffusion model. However, a careful analysis of their arguments reveals that the conditions (i)–(iii) are necessary only when π^* is optimal in the stricter sense discussed in Remark (c) above. For example see Section 6 of Chernoff and Petkau [3]. Here any switching point $b \geq 0$ will give the same average cost and hence an average cost optimal policy. However, only one b satisfies conditions (i)–(iii) of Theorem 4.2.

The following theorem deals with the existence of a stationary optimal policy.

Theorem 4.3. *Suppose there exists a constant g and a family $\{h_a; a \in \mathscr{A}\}$ of functions satisfying the following*:

(i) $h_a \in \mathscr{D}'$ *for* $a \in \mathscr{A}$.

(ii) *For each* $a \in \mathscr{A}$, h_a *is continuous on* $[r_0, r_1]$.

(iii) *No discontinuity point of* $A_a h_a$ *belongs to the set*

$$\Lambda^a = \{x \in \mathscr{X}; h_a(x) < \min_{a' \in \mathscr{A}-\{a\}} \{R(x, a, a') + h_{a'}(x)\}\}.$$

(iv) $g \leq c(x, a) + A_a h_a(x)$ *for all* $x \in (r_0, r_1)$ *at which* $A_a h_a$ *is continuous, with equality on* Λ^a.

(v) $h_a(x) \leq \min_{a' \in \mathscr{A}} \{R(x, a, a') + h_{a'}(x)\}$ *for all* $x \in \mathscr{X}$, $a \in \mathscr{A}$, *with equality on* $\mathscr{X} - \Lambda^a$.

Then $g = V_a(x)$ *for all* $x \in \mathscr{X}$, $a \in \mathscr{A}$, *and there exists a stationary policy* $\pi^* \in D_S$ *which is optimal.* $\{\Lambda^a; a \in \mathscr{A}\}$ *defined in* (iii) *above are the continuation sets of* π^* *and the action selecting function f of* π^* *can be chosen so as to minimize the right hand side in* (v).

Proof. The proof is similar to that of Theorem 1 in Doshi [4] and so the details are omitted.

5. A Brownian motion control problem

Controlled Brownian motion has been used by Bather [1, 2], Chernoff and Petkau [3] and Puterman [10] to model control problems in queuing, inventory and storage systems. Quite frequently such systems have finite capacity and overflow occurs when the state goes beyond this capacity. Also, as in storage and queuing models, there is a lower limit on the possible states of the system. In such cases it is appropriate to model the system as a controlled Brownian motion with two reflecting boundaries.

In this section we consider a simple controlled Brownian motion with two reflecting boundaries at $r_0 = -1$ and $r_1 = 1$. By the translation of the coordinates and change of scale a model with any reflecting boundaries r_0 and r_1 can be transformed into this framework. We assume that our model is symmetric about 0. Specifically, $\mathscr{A} = \{1, 2\}$, $b(x, 1) \equiv \mu > 0$ and $b(x, 2) \equiv -\mu < 0$, $d(x, a) \equiv \frac{1}{2}\sigma^2$, $c(x, a) = cx^2$ for all $x \in (-1, 1)$ and $a \in \mathscr{A}$ and $R(x, 1, 2) = R(x, 2, 1) = k > 0$.

From the symmetry of the problem it is intuitive that one of the following two types of policies is optimal:

(I) Never change the action. That is, $f(x, a) = a$ for $a = 1, 2, x \in [-1, 1]$, and $\Lambda^1 = \Lambda^2 = [-1, 1]$.

(II) For some M, $0 < M < 1$, $f(x, 1) = 1$ $(-1 \leq x < M)$, $f(x, 1) = 2$ $(M \leq x \leq 1)$, $f(x, 2) = 1$ $(-1 \leq x \leq -M)$ and $f(x, 2) = 2$ $(-M < x \leq 1)$. In this case $\Lambda^1 = [-1, M)$ and $\Lambda^2 = (-M, 1]$.

It is also intuitive that I is optimal when $k \geq k_1$ and II is optimal when $k < k_1$ for some constant k_1 depending on μ, σ^2 and c. In what follows we derive this constant k_1 and show that policies I and II are in fact optimal when $k \geq k_1$ and $k < k_1$, respectively.

First we investigate the situation when policy I is optimal. Let

$$k_1 = \frac{\sigma^4 c}{\mu^3} + \frac{4c}{3\mu} - \frac{2\sigma^2 c}{\mu^2}\left[\frac{\exp(2\mu/\sigma^2) + \exp(-2\mu/\sigma^2)}{\exp(2\mu/\sigma^2) - \exp(-2\mu/\sigma^2)}\right]. \tag{5.1}$$

Theorem 5.1. *Suppose $k \geq k_1$. Then the policy π^* described by* I *is optimal.*

Proof. Under policy π^* the resulting Markov chain has two closed classes ($\lambda(1, 1) = [-1, 1]$ and $\lambda(1, 2) = [-1, 1]$) with the same average cost. By solving the equations of Theorem 4.1 we get

$$h_1(x) = A\exp(-2\mu x/\sigma^2) - \frac{cx^3}{3\mu} + \frac{\sigma^2 cx^2}{2\mu^2} - \frac{\sigma^4 cx}{2\mu^3} + \frac{gx}{\mu} \quad x \in [-1, 1] \tag{5.2}$$

and

$$h_2(x) = A\exp(2\mu x/\sigma^2) + \frac{cx^3}{3\mu} + \frac{\sigma^2 cx^2}{2\mu^2} + \frac{\sigma^4 cx}{2\mu^3} - \frac{gx}{\mu} \quad x \in [-1, 1], \tag{5.3}$$

where

$$g = \frac{\sigma^4 c}{2\mu^2} + c - \frac{\sigma^2 c}{\mu}\left[\frac{\exp(2\mu/\sigma^2) + \exp(-2\mu/\sigma^2)}{\exp(2\mu/\sigma^2) + \exp(2\mu/\sigma^2)}\right], \tag{5.4}$$

and

$$A = -\frac{\sigma^4 c}{\mu^3}\,\frac{1}{\exp(2\mu/\sigma^2) - \exp(-2\mu/\sigma^2)} < 0. \tag{5.5}$$

Functions h_1 and h_2, and constant $g = V_{\pi^*,a}(x)$ ($x \in [-1, 1]$, $a \in \mathcal{A}$) clearly satisfy hypotheses (i) and (ii) of Theorem 4.2. So it remains to show that they also satisfy hypothesis (iii) of Theorem 4.2. That is,

$$-k \leqslant h_1(x) - h_2(x) \leqslant k \quad (x \in [-1, 1]). \tag{5.6}$$

Let $J(x) = h_1(x) - h_2(x)$. Then a routine analysis of the derivatives of J implies that $J'(x) \geqslant 0$ for $x \in [-1, 1]$ and $J'(-1) = J'(1) = 0$. Thus, inequality (5.6) holds iff $J(-1) \geqslant -k$ and $J(1) \leqslant k$. But from equations (5.2)–(5.5)

$$J(1) = \frac{\sigma^4 c}{\mu^3} - \frac{2c}{3\mu} - \frac{\sigma^4 c}{\mu^3} + \frac{2}{\mu}\left[\frac{\sigma^4 c}{2\mu^2} + c - \frac{\sigma^2 c}{\mu}\,\frac{\exp(2\mu/\sigma^2) + \exp(-2\mu/\sigma^2)}{\exp(2\mu/\sigma^2) - \exp(-2\mu/\sigma^2)}\right]$$

$$= \frac{\sigma^4 c}{\mu^3} + \frac{4c}{3\mu} - \frac{2\sigma^2 c}{\mu^2}\left[\frac{\exp(2\mu/\sigma^2) + \exp(-2\mu/\sigma^2)}{\exp(2\mu/\sigma^2) + \exp(-2\mu/\sigma^2)}\right] = k_1. \tag{5.7}$$

Similarly, $J(-1) = -k_1$. So $-k \leqslant h_1(x) - h_2(x) \leqslant k$ for all $x \in [-1, 1]$ iff $k_1 \leqslant k$. This proves the theorem.

Next we show that a policy of type II is optimal if $k < k_1$. Let π_M be the policy defined by II, $0 < M < 1$. Then solving equations in Theorem 4.1 we get

$$h_1(x) = B\exp(-2\mu x/\sigma^2) - \frac{cx^3}{3\mu} + \frac{\sigma^2 cx^2}{2\mu^2} - \frac{\sigma^4 cx}{2\mu^3} + \frac{gx}{\mu} \quad x \in [-1, M), \tag{5.8}$$

$$h_2(x) = B\exp(2\mu x/\sigma^2) + \frac{cx^3}{3\mu} + \frac{\sigma^2 cx^2}{2\mu^2} + \frac{\sigma^4 cx}{2\mu^3} - \frac{gx}{\mu} \quad x \in (-M, 1], \tag{5.9}$$

$$h_1(x) = k + h_2(x) \quad x \in [M, 1], \tag{5.10}$$

and

$$h_2(x) = k + h_1(x) \quad x \in [-1, -M], \tag{5.11}$$

where

$$g = c + \frac{\sigma^2 c}{\mu} + \frac{\sigma^4 c}{2\mu^2} - \frac{k + \dfrac{2cM^3}{3\mu} + \dfrac{\sigma^4 cM}{\mu^3} + \dfrac{2M}{\mu}\left(c + \dfrac{\sigma^2 c}{\mu} + \dfrac{\sigma^4 c}{2\mu^2}\right)}{\dfrac{2M}{\mu} + \dfrac{\sigma^2}{2\mu^2}\exp(-2\mu/\sigma^2)((\exp(2\mu M/\sigma^2) - \exp(-2\mu M/\sigma^2))}, \tag{5.12}$$

and

$$B = \frac{\sigma^2}{2\mu^2}\exp(-2\mu/\sigma^2)\,\frac{k + \dfrac{2cM^3}{3\mu} - \dfrac{2cM}{\mu} - \dfrac{2\sigma^2 cM}{\mu^2}}{\dfrac{2M}{\mu} - \dfrac{\sigma^2}{2\mu^2}\exp(-2\mu/\sigma^2)((\exp(2\mu M/\sigma^2) - \exp(-2\mu M/\sigma^2))}. \tag{5.13}$$

Clearly, the average cost g here should be less than the average cost when we have the reflecting boundaries at -1 and 1 replaced by natural boundaries [see [7] for definition] at $-\infty$ and $+\infty$. So from Doshi [5]

$$g<\frac{k\mu}{2M}+\frac{cM^2}{\mu}+\frac{\sigma^4 c}{2\mu^2} \tag{5.14}$$

and hence

$$B=\frac{\frac{2gM}{\mu}-k-\frac{2cM^3}{3\mu}-\frac{\sigma^4 cM}{\mu^3}}{\exp(2\mu M/\sigma^2)-\exp(-2\mu M/\sigma^2)}<0. \tag{5.15}$$

The average cost g defined by (5.12) is a function of M. It is easy to show that

$$\left.\frac{\partial g}{\partial M}\right|_{M=0}<0 \quad\text{and}\quad \left.\frac{\partial g}{\partial M}\right|_{M=1}$$

has the sign of k_1-k. So when $k<k_1$ we have

$$\left.\frac{\partial g}{\partial M}\right|_{M=0}<0 \quad\text{and}\quad \left.\frac{\partial g}{\partial M}\right|_{M=1}>0. \tag{5.16}$$

This implies that there exists a M^*, $0<M^*<1$ such that π_{M^*} is optimal among all policies π_M, $0<M<1$. Also

$$\left.\frac{\partial g}{\partial M}\right|_{M=M^*}=0. \tag{5.17}$$

In fact, as we show in Theorem 5.2 below, the policy π_{M^*} is optimal among all policies in D.

Theorem 5.2. *When $k<k_1$, there exists an M^*, $0<M^*<1$ minimizing g over all policies of form* II. *Moreover, π_{M^*} is optimal in D.*

Proof. We verify that g, h_1 and h_2, for policy π_{M^*}, satisfy the hypotheses of Theorem 4.1. Let

$$J(x)=B[\exp(-2\mu x/\sigma^2)-\exp(2\mu x/\sigma^2)]-\frac{2cx^3}{3\mu}-\frac{\sigma^4 cx}{\mu^3}+\frac{2gx}{\mu} \quad x\in[-1,1], \tag{5.18}$$

$$W_1(x)=cx^2+\mu h_1'(x)+\frac{\sigma^2}{2}h_1''(x)-g \quad x\in[-1,1], \tag{5.19}$$

and

$$W_2(x)=cx^2-\mu h_2'(x)+\frac{\sigma^2}{2}h_2''(x)-g \quad x\in[-1,1]. \tag{5.20}$$

Then $J(x)=h_1(x)-h_2(x)$ for $x\in(-M^*,M^*)$. Also from Theorem 3.3 $W_1(x)=0$ ($x\in[-1,M)$), and $W_2(x)=0$ ($x\in(-M,1]$). So it suffices to show that h_1' and h_2' are continuous, and

$$-k\leqslant J(x)\leqslant k \quad x\in(-M,M), \tag{5.21}$$

$$W_1(x)\geqslant 0 \quad x\in[M,1], \tag{5.22}$$

$$W_2(x)\geqslant 0 \quad x\in[-1,-M]. \tag{5.23}$$

From equations (5.12), (5.13) and (5.17) we have

$$J'(-M^*)=J'(M^*)=0. \tag{5.24}$$

So h_1' and h_2' are continuous at $-M^*$ and M^* and hence on $[-1,1]$. Also, since $B<0$

$$J'''(x)=\frac{-8\mu^3}{\sigma^2}B[\exp(2\mu x/\sigma^2)+\exp(-2\mu x/\sigma^2)]-\frac{4c}{\mu} \tag{5.25}$$

is increasing in x. This, together with equation (5.24) leads to

$$J'(x)\geqslant 0 \quad \text{for } x\in(-M^*,M^*), \tag{5.26}$$

$$\begin{aligned} J'(-1)=J'(1) &=\frac{-2\mu B}{\sigma^2}[\exp(2\mu/\sigma^2)+\exp(-2\mu/\sigma^2)]-\frac{2c}{\mu}-\frac{\sigma^4 c}{\mu^3}+\frac{2g}{\mu} \\ &<\frac{-4\mu B}{\sigma^2}\exp(2\mu/\sigma^2)-\frac{2c}{\mu}-\frac{\sigma^4 c}{\mu^3}+\frac{2g}{\mu}=0, \end{aligned} \tag{5.27}$$

where the last equality follows from equations (5.12) and (5.13). Equation (5.26) implies that

$$-k=J(-M^*)\leqslant J(x)\leqslant J(M^*)=k \quad x\in(-M^*,M^*) \tag{5.28}$$

thus proving (5.21). Next we prove (5.22). The proof of inequality (5.23) is similar. For $x\in(M^*,1]$

$$\begin{aligned} W(x)&=cx^2+\mu h_1'(x)+\frac{\sigma^2}{2}h_1''(x)-g \\ &=cx^2+\mu h_2'(x)+\frac{\sigma^2}{2}h_2''(x)-g \\ &=2\mu h_2'(x). \end{aligned} \tag{5.29}$$

So it suffices to show that $h_2'(x)\geqslant 0$ for $x\in[M^*,1]$. $h_2'''(x)<0$ for $x\in(M^*,1]$. So $h_2'(x)$ is concave in $[M^*,1]$. Also $h_2'(1)=0$. So we need to show that $h_2'(M^*)\geqslant 0$.

But

$$h_2'(M^*) = \frac{2\mu B}{\sigma^2}\exp(-2\mu M^*/\sigma^2) - \frac{cM^{*3}}{3\mu} + \frac{\sigma^2 cM^{*2}}{2\mu^2} - \frac{\sigma^4 cM^*}{2\mu^3} + \frac{gM^*}{\mu}$$

$$= -\frac{\sigma^2}{4\mu}\left[\frac{-4\mu^2 B}{\sigma^4}((\exp(2\mu M^*/\sigma^2) - \exp(-2\mu M^*/\sigma^2)) - \frac{4cM^*}{\mu}\right]$$

$$= -\frac{\sigma^2}{4\mu} J''(M^*) > 0$$

since J' is decreasing at M^*. This proves (5.22) and hence the theorem.

References

[1] J. Bather, A continuous time inventory model, J. Appl. Prob. 3 (1966) 538–549.
[2] J. Bather, A diffusion model for the control of a dam. J. Appl. Prob. 5 (1968) 55–71.
[3] H. Chernoff and J. Petkau, Optimal control of a Brownian motion, working paper (1977).
[4] B.T. Doshi, Optimal control of a diffusion process with reflecting boundaries and both continuous and lump costs, to appear in the Proceedings of International Conference on Dynamic Programming (Academic Press, New York, 1977).
[5] B.T. Doshi, Two-mode control of a Brownian motion with quadratic loss and switching costs. To appear in Stochastic Processes Appl. (1978).
[6] E.B. Dynkin, Markov Processes I–II (Academic Press, 1965, New York).
[7] P. Mandl, Analytic Treatment of One-Dimensional Markov Processes (Springer-Verlag, 1968, New York).
[8] S. Pliska, Single person controlled diffusions with discounted costs, J. Opt. Theory and Appl. 12 (1973).
[9] M.L. Puterman, Sensitive discount optimality for one-dimensional diffusions, Ann. Prob. 2 (1974) 408–419.
[10] M.L. Puterman, A diffusion model for a storage system in: Logistics, a special issue of Management Sci. (1976).