A METHOD FOR COMPUTING DOUBLE BAND POLICIES FOR SWITCHING BETWEEN TWO DIFFUSIONS

Revised Version, March 1996

Running Title: Optimal Switching Between Two Diffusions

Florin Avram†
and
Fikri Karaesmen

† To whom all correspondence should be sent
Department of Mathematics
567 Lake Hall
Northeastern University
Boston, MA 02115
E-mail: avram@neu.edu
Phone: (617) 373-5231

Abstract: We develop a method for computing the optimal double band $[b, B]$ policy for switching between two diffusions with continuous rewards and switching costs. The two switch levels $[b, B]$ are obtained as perturbations of the single optimal switching point a of the control problem with no switching costs. More precisely, we find that in the case of average reward problems the optimal switch levels can be obtained by intersecting two curves: a) the function, $\gamma(a)$, which represents the long run average reward if we were to switch between the two diffusions at a and switches were free and b) an horizontal line whose height depends on the size of the transaction costs. Our semi-analytical approach reduces, for example, the solution of a problem recently posed by Perry and Bar-Lev [20] to the solution of one non-linear equation.

A Method for Computing Double Band Policies for Switching Between Two Diffusions

Florin Avram[1] Fikri Karaesmen[2]

[1]Department of Mathematics, Northeastern University, Boston, MA 02115

[2]Department of Mech., Ind. and Manuf. Eng., Northeastern University, Boston, MA 02115

March 1996

1 Introduction

Consider a bounded storage system whose level is governed at any time by one of two Markov processes. A controller decides which of the two mechanisms will be used, with the goal of optimizing some reward: for example there might be continuous rewards for keeping the level in the interior of some state space and penalties for hitting its boundaries. Problems of this nature are common in both inventory control and in queueing systems. Continuous diffusion models have been used quite successfully as approximations of the discrete Markov models under conditions of heavy traffic for the above queueing/storage problems. Some early examples are in the work of Newell [19], Harrison [9], Reiman [23] and Foschini [7].

Our motivation for considering diffusions is however more modest: we study them just as an example of one instance of a process for which the optimal policy may be obtained explicitly, yielding thus insight into the interplay between the various parameters of the problem: means, variances and costs. The first attempts to use explicitly solvable diffusion control models in optimization problems arising in operations research, statistics or finance are due to Bather [1],[2], Chernoff [4], Benes and Shepp [3] and Samuelson and Mc Kean [24]. Diffusion methods have been ever since a hot area in the finance literature. For some recent examples, see Taksar et al. [28], Davis and Norman [6], Shreve et. al [26] and Shepp and Shiryaev[25].

One dimensional diffusion models have also been successfully employed to gain insights in inventory control models, for example by Harrison and Taksar [11] and Taksar [27]. Harrison and Reiman have pioneered the use of multidimensional reflected Brownian motion to study queueing network problems in Harrison [9], Harrison and Reiman [10] and Reiman [23]. Multidimensional reflected brownian motion problems display lots of intriguing features missing in the one dimensional problems, as revealed in the work of Williams (for example see Varadhan and Williams [29]).

The one dimensional problem of *drift control* with *switching costs* we consider below is similar to those studied by Rath [21], [22], Chernoff and Petkau [4], Krichagina et al. [16], [17] and by Perry and Bar-Lev [20]. The methods we use; the dynamic programming approach, the equation for the differential rewards and the "smooth fit" equation were first used in this context by Bather [1],[2].

Rath [21] studied the following decision problem in a queueing system: any one of two servers can be selected to serve customers at any time, there are holding costs associated with the customers waiting in the queue and there are switching costs associated with changing between servers. He showed that under heavy traffic conditions, the queue length process for this queueing system converges to a reflected Brownian Motion process with two different sets of parameters corresponding to each server. Furthermore, the costs also converge, therefore a discrete state space control problem motivates a control problem with a continuous state space.

In a later paper, Rath [22] obtained the optimal policy for the continuous time control problem by approximating the Brownian motion processes using the corresponding random walks. The policy is determined by a band of type (b, B) and it is optimal to use the first Brownian motion starting from the time the queue length reaches B from below and continue with it until the queue length falls to b. Similarly, use the second Brownian motion from the time the queue length falls to b until the time the queue length builds up to B.

Chernoff and Petkau [4] noted that Rath's problem of switching between two Brownian motions can be treated in continuous time using dynamic programming. Perry and Bar-Lev [20] also considered a similar drift control problem in the context of inventory control. They studied a bounded storage system where there are penalties for hitting the boundary and the inventory level is controlled by controlling the drift of the process.

Finally, Krichagina et al. [16] studied a manufacturing control problem which converges to a one dimensional stochastic control problem that can be solved explicitly. In a later paper, Krichagina et al. [17], studied another manufacturing control problem that involves setup costs. The limiting case of this problem is a one dimensional impulse control problem which also yields explicit solutions. Furthermore, the optimal policy for this problem is shown to be of double band type.

Our main result in this paper, given in Section 3, is an easily implementable method for finding the optimal double band policies for the drift control problems of Rath [21] and Perry and Bar Lev [20], for the case of long run average reward optimization. (We also report in the Appendix on a characterization of the switch levels for the case of optimizing the total cost until absorption on the boundaries. Unfortunately, the result in this case does not lead to significant simplifications over the extensive search for the optimal levels proposed in earlier papers.)

Our results are similar to those obtained in [12] for a different problem (inventory impulse control with linear holding costs and discounting). In Section 2, as a warm up, we show how our method works in a setup similar to the one in [12]; that of (s, S) inventory impulse control.

2 (s, S) Inventory Control

Under quite general assumptions, the optimal inventory control policy is to order $(S - s)$ units from an outside supplier whenever the inventory level drops to s. We examine first our method on this well known problem. Suppose a diffusion process $\{X(t), t \geq 0\}$ with generator $G = \mu(x)\frac{d}{dx} + \frac{\sigma^2(x)d^2}{2dx^2}$ (i.e. with drift $\mu(x)$ and variance $\sigma^2(x)$) accumulates rewards of $h(x)\, dt$ while running, where h is a positive continuous function. We control the movement by impulse control: whenever the inventory hits a certain level s, we bring it up to a higher level S, while paying a fee of k. The levels (s, S) are to be chosen so that the long run average cost is minimized. For a fixed policy $\pi = (s, S)$, let $c(x, \pi, t)$ denote the total cost up to time t

$$c(x, \pi, t) = E_x \left[\int_0^t \; h(X(s))ds + kN(t) \right] \tag{1}$$

where $N(t)$ is the number of switches up to time t. The controller's objective is to choose (s, S) to minimize the long run average cost:

$$\gamma(\pi) = \lim_{t \to \infty} \frac{c(x, \pi, t)}{t}.$$

(the limit is independent of the starting point x).

We will approach the problem in two steps. In the first step, we assume the switching cost is zero and obtain the single optimal switching point. In the next step, we expand the single point obtained in step 1 into the optimal double band.

STEP 1: In the case of no switching costs, the optimal policy is known (cf. (11) and (12)) to be instantaneous control, or "regulation" at some point a, which informally means that all negative increments of $X(t)$ occurring while $X(t)$ is at a are cancelled. Suppose a fixed regulating boundary a is chosen. The invariant measure of the resulting process on $[a, \infty)$ is proportional to $\exp(M(x))$, where M is an antiderivative of $2\mu/\sigma^2$ (see, for example [15]). Let

$$\hat{t}(x) = \int_x^\infty \exp(M(y))\, dy \tag{2}$$

and

$$\hat{f}(x) = \int_x^\infty h(y) \exp(M(y))\, dy. \tag{3}$$

The average cost when switching at a $\gamma(a)$ is thus

$$\gamma(a) = \hat{f}(a)/\hat{t}(a) \tag{4}$$

(assuming the integrals in (2) and (3) are well defined) and the optimal a^* must be chosen so that it minimizes $\hat{f}/\hat{t}$. For example, in the case of constant variance $\sigma^2 = 2$ and constant negative drift μ we find $M(y) = \exp(\mu(y-a))$ and so we have to minimize $\gamma(a) = Eh(a+Z)$, where Z is exponentially distributed with parameter $-\mu$.

Remark: A different interpretation of the functions $\hat{f}, \hat{t}$ is provided by the heuristic method of differential rewards pioneered by Bather [2], used in the next step.

STEP 2. In the presence of transaction costs, the optimal policy is of double band type (i.e. π is determined by the pair (s, S). Due to the transaction cost, the controller lets the level drop to a level s below the instantaneous control point a^* before restocking.)

The dynamic programming approach, which can be justified rigorously by an application of the generalized Ito Lemma (see [5], Theorem 8.2, pg. 161) implies that the cost $c(x, \pi, t)$ satisfies the partial differential equation

$$\frac{\partial}{\partial t}\, c = Gc(x) + h(x) \tag{5}$$

with the boundary condition

$$c(s, \pi, t) = c(S, \pi, t) + k. \tag{6}$$

.

Bather's approach constitutes in assuming that

$$c(x, \pi, t) \approx \gamma t + V(x), \tag{7}$$

where γ is the long run average cost per time unit and $V(x)$, called the differential cost starting at x, and measures the relative differences of order $O(1)$ between the various starting points after the common long run average γt is subtracted. Plugging this expansion in the dynamic programming equations for $c(x, \pi, t)$ shows that $V(x)$ satisfies the differential equation

$$GV = -h + \gamma \tag{8}$$

and the boundary condition

$$V(s) = V(S) - k. \tag{9}$$

A rigorous but less direct approach to (7) is to replace long run averages by discounted averages and let the discounting approach 0. Since the focus of our paper is to simplify the actual numerical computing of optimal double band policies, we will remain at the informal level.

Letting now $v(x)$ denote the derivative of V we replace the second order equation (8) by the first order equation

$$gv = -h + \gamma, \tag{10}$$

where g denotes the operator $\frac{\sigma^2 d}{2dx} + \mu$. Finally we decompose $v(x)$ as $v = f - \gamma t$, where t and f are solutions of

$$g\,f \quad = \quad -h \quad \text{and} \tag{11}$$

$$g\,t \quad = \quad -1 \tag{12}$$

respectively. The one boundary condition needed to determine f and t uniquely is provided by their behaviour at ∞. By choosing a large truncation boundary B, specifying the boundary conditions $f'(B) = 0$, $t'(B) = 0$ (corresponding to reflection) and letting $B \to \infty$ we find that the exponentially growing homogeneous solutions to (11) and (12) fall down. The particular solutions are $f = 2\hat{f}/\sigma^2$ and $t = 2\hat{t}/\sigma^2$ where $\hat{f}$ and $\hat{t}$ were defined in STEP 1 in (2) and (3). Thus, f and t may be viewed both as averages of x and 1 with respect to the stationary measure and as a decomposition of the derivative v of Bather's differential cost. For fixed S, s, γ and k we have determined now v and also V up to a constant and the extra constraint (9) determines γ. When S and s vary, the optimal values of (s^*, S^*) will be such that the partial derivatives of γ with respect to (s, S) are 0. Taking this into account and differentiating (9) with respect to S and s yields the equations

$$v(s) = v(S) = 0, \tag{13}$$

which may also be put in the form

$$f(s)/t(s) = f(S)/t(S) = \gamma. \tag{14}$$

Note that the function $f(a)/t(a)$ was shown in STEP 1 to be precisely the average cost when regulating at a and no switching costs are charged.

Definition: Two points satisfying equation (14) (i.e. which lead to equal long run averages in the problem with no transaction costs) will be called conjugate points.

Theorem 1: Suppose the function $\gamma = f/t$ is unimodal. Then each value: $\gamma \in [\gamma(a^*), \min\{\gamma(-\infty), \gamma(\infty)\}]$ determines a unique pair of conjugate points which are optimal for the problem with transaction cost k which is related to γ by (9)

$$\int_s^S (f - \gamma\, t)dy = k. \tag{15}$$

The optimal switch levels can thus be determined by four integrations (the computations of f, t and their integrals and the solution of one nonlinear equation, the conjugacy equation). For simple holding cost functions, these computations may be performed analytically as demonstrated in the example below.

Example 1: This example demonstrates how to find the optimal regulation point and then compute the optimal (s, S) band for a storage system with quadratic holding and backordering costs $h(x) = hx^2$. Solving the differential equations (11) and (12) with $\mu < 0$ and $\sigma^2 = 2$ yields:

$$f(x) = \frac{-2h}{\mu^3} + \frac{2hx}{\mu^2} - \frac{hx^2}{\mu}$$

and

$$t(x) = -\frac{1}{\mu}.$$

The optimal regulation point a is the point where the function $f(x)/t(x)$ is minimized. Taking derivatives, it is seen that

$$a^* = \frac{1}{\mu}.$$

The optimal cost per unit time $\gamma = f(a^*)/t(a^*)$ is

$$\gamma = \frac{h}{\mu^2}.$$

The integration in (15) may also be performed explicitly and the conjugacy equation for $d = S - a = a - s$ is just :

$$d^3 = \frac{3k(-\mu)}{4h}.$$

k	s	S	γ
0	-1	-1	1
0.001	-1.1	-0.9	1.01
0.167	-1.5	-0.5	1.25
1.333	-2	0	2
10.667	-3	1	5
36.000	-4	2	10

Table 1: Results of Example 1

Table 1 gives the optimal band (s, S) and the corresponding transaction cost k, for a few values of γ, for the parameter values $\mu = -1$ and $h = 1$. Figure 1 displays the function $f(x)/t(x) = \gamma(x)$.

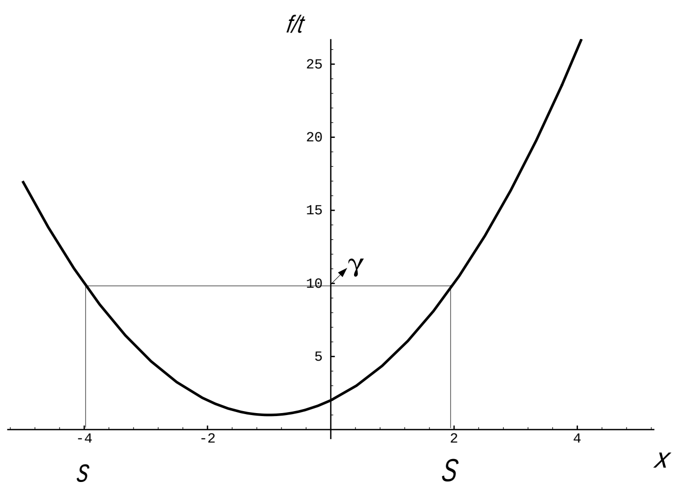

Figure 1: Optimal (s, S) pairs

3 Optimal Drift Control to Maximize Expected Reward per Unit Time

In this section we study the problem of optimally switching between two diffusion processes to maximize the expected reward per unit time. The general setup is as follows: a controller decides at every time t to use one of two available diffusion operators with constant coefficients $G_i = \frac{\sigma^2 d^2}{2\,dx^2} + \mu_i \frac{d}{dx}$, $i = 0, 1$, depending on which of two control regions the diffusion process $Y(t)$ lies in. An important simplifying assumption made throughout the paper will be that in the case of no switching costs the optimal policy uses exactly one switchpoint a. We may then call, without loss of generality, by G_i the diffusion mechanism used near the boundary point i.

We assume constant variance and piecewise constant drift, depending on which side of the switch point "a" the process lies. The resulting process is thus a diffusion with "bang-bang" drift, described by the following stochastic differential equation:

$$dY(t) = \mu(Y(t))dt + \sigma dB(t). \tag{16}$$

where $B(t)$ is a Brownian motion with $\mu = 0$ and $\sigma^2 = 1$.

The existence and uniqueness of the solution of this equation follows for example from Theorem 10.4 of [5], where an explicit computation of the density in a particular case may be also found in Exercise 4, pg. 214. An interesting fact about this process is that its distribution is uniquely determined by the controller's decision off the control boundary, and thus the value of the drift at the infinite set of times at which the process crosses from one region to the other is irrelevant. For a discussion of this in a similar context we refer to [18].

For the double band policies, the construction of the process is simpler: the diffusion mechanism is switched at each of the alternating sequence of hitting times of B and b. Finally, the process is regulated both at 0 and 1, which means that we observe the process $X(t)$ restricted to the interval $[0, 1]$ modelled by the equation

$$X(t) = Y(t) + L_0(t) - L_1(t) \tag{17}$$

where $L_i(t)$, $i = 0, 1$, are the unique minimal nondecreasing continuous functions which may increase only when $X(t)$ is at the respective boundaries. (The fact that the minimal "regulators" L_i are uniquely defined "pathwise" on the space of continuous functions, due to Harrison and Reiman, may be found for example on page 22 of [13]).

We assume the following reward structure: using the i'th generator yields a reward of $h_i(x)\ dt$ per time interval dt, switching from one mechanism to the other costs $k/2$ and the regulation leads to a penalty of α_i per unit of regulation at i, that is a regulation cost of $\sum_{i=1}^{2} \alpha_i L_i(t)$ is incurred up to time t. The total reward associated with a partition

$\pi = (\pi_0, \pi_1)$ of the state space is thus

$$U_{x,\pi,t} = E_{i,x,\pi}\left[\int_0^t \sum_{i=1}^{2} h_i(X(s))1_{[X(s)\in\pi_i]}ds - \sum_{i=1}^{2}\alpha_i L_i(t) - kN(t)/2\right] \tag{18}$$

where $N(t)$ is the number of policy switches prior to time t. The objective is to maximize the long run average reward

$$\gamma = sup_\pi \lim_{t\to\infty} \frac{U_{x,\pi,t}}{t} \tag{19}$$

(which is independent of the starting point x).

Example 2: Assume that both the switching costs and the boundary costs are 0. The stationary density of the process is truncated exponential on both sides of the switch point a, with parameters $r_i = 2\mu_i/\sigma^2$, $i = 0, 1$ respectively. Since the density has to be continuous at a it follows that

$$\gamma(a) = f(a)/t(a), \tag{20}$$

where $f(a) = \int_0^a h_0(x)e^{r_0(x-a)}2dx/\sigma^2 + \int_a^1 h_1(x)e^{r_1(x-a)}2dx/\sigma^2$ and $t(a) = \int_0^a e^{r_0(x-a)}2dx/\sigma^2 + \int_a^1 e^{r_1(x-a)}2dx/\sigma^2$ (the normalization factor $2/\sigma^2$ is convenient in the sequel) .

Thus, the optimum switch point is the maximum of the function f/t (for example, in the case $\mu_0 = \mu_1$ the optimum switch point reduces to precisely the root of the equation $h_0(x) = h_1(x)$, the "myopic" solution). We will show again that the functions f, t may also be used for locating the optimum switch points.

Theorem 2: a) The policy of switching between the two diffusions at a fixed point x achieves the long run average value of

$$\gamma(x) = \frac{f(x)}{t(x)}. \tag{21}$$

b) If a double band policy (b, B) is optimal for some transaction cost k, then the switch levels (b, B) and the corresponding long run average γ satisfy:

$$\frac{f(B)}{t(B)} = \frac{f(b)}{t(b)} = \gamma \quad \text{(conjugacy equation)} \tag{22}$$

and

$$\int_b^B (f(x) - \gamma t(x))\, dx = k \quad \text{(cost equation)} \tag{23}$$

where f, t (which take now into account also the boundary costs) are defined below.

A graphical interpretation is provided in Figure 2. $\gamma(a)$ is the average reward when switching at a the absence of switching costs. Note that if $\gamma(a)$ is unimodal, for any value

$s \in [\max[\gamma(0), \gamma(1)], \gamma(a)]$ Theorem 2 provides a unique $k > 0$ determined by (23) and a double band $[b, B]$ which is optimal for that k.

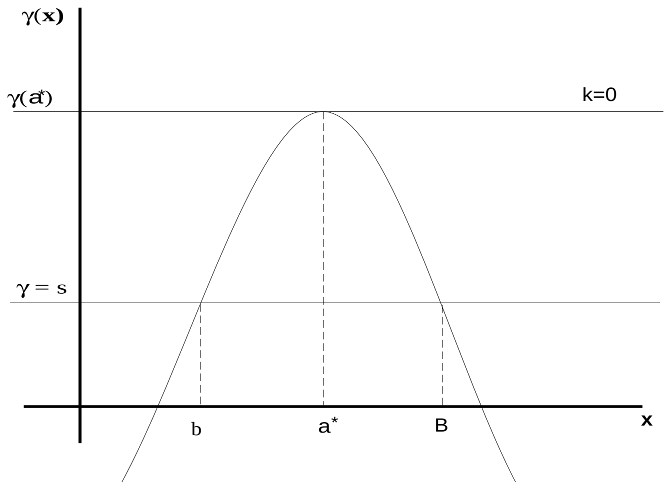

Figure 2: Graphical Interpretation of Theorem 2

Proof of Theorem 2:

a)We will extend first the definitions of f, t from the previous example so that they take into account the long run boundary costs. The most convenient way is via Bather's approach.

For a fixed switch point a, on each side of a the dynamic programming equations, obtained by plugging Bather's approximation for large t in the reward function $U(x, \pi, t)$ are

$$G_i V_i = -h_i + \gamma \tag{24}$$

$$V_i' = (-1)^i \alpha_i \tag{25}$$

$i = 0, 1$ (the boundary conditions follow from the generalized Ito formula (4.7.4) in ([13]); see also ([27], theorem 4.7).

By letting $v_i = V_i'$ denote the derivatives of Bather's differential costs, these become:

$$g_i v_i = -h_i + \gamma \tag{26}$$

$$v_i = (-1)^i \alpha_i \tag{27}$$

where g_i denote the first order differential operators $\frac{\sigma^2}{2}\frac{d}{dx} + \mu_i$.

Let now f_i and t_i be decompositions of v_i in the form $v_i = \gamma t_i - f_i$. We choose f_i, t_i as solutions of the following differential equations:

$$\left\{ \begin{array}{rcl} g_i\, f_i & = & h_i \\ f_i(i) & = & (-1)^{i+1}\alpha_i \end{array} \right. \tag{28}$$

and

$$\left\{ \begin{array}{rcl} g_i\, t_i & = & 1 \\ t_i(i) & = & 0. \end{array} \right. \tag{29}$$

It is easy to check then that $\gamma t_i - f_i$ satisfy (26). Finally, let $f = f_0 - f_1$ and $t = t_0 - t_1$. In the absence of boundary costs f and t are precisely the integrals defined in Example 2.

We now use the continuity of v_i at the switch point $\gamma t_0 - f_0 = \gamma t_1 - f_1$ to conclude that the long run average reward is $\gamma = f/t$.

b) We fix a double band (b, B) which is optimal for some cost k and note that the control policy translates into the boundary conditions:

$$\left\{ \begin{array}{rcl} V_0(B) & = & V_1(B) - k/2 \\ V_1(b) & = & V_0(b) - k/2 \end{array} \right. \tag{30}$$

which when added yield the cost equation:

$$\int_b^B (v_1 - v_0)dx = \int_b^B \left(f(x) - \gamma t(x)\right) dx = k \tag{31}$$

Differentiating (31) with respect to (b, B) and taking into account that the partial derivatives of γ with respect to B, b have to be 0 yields

$$\left\{ \begin{array}{rcl} v_0(b) & = & v_1(b) \\ v_0(B) & = & v_1(B), \end{array} \right. \tag{32}$$

also called *smooth fit* equations. Using the decomposition of v_i they can also be put in the form

$$\gamma t(b) - f(b) = 0 \tag{33}$$

and

$$\gamma t(B) - f(B) = 0 \tag{34}$$

which yields the conjugacy equations (22).

□

Remarks: 1)Besides locating the maximum of f/t, another method to obtain the optimum switchpoint a^* is to equate the common value m of $v_i(a)$ obtained by two methods:
A) by plugging the value $\gamma = f(a)/t(a)$ in $v_i = \gamma t_i - f_i$, which yields:

$$m = v_i(a) = \frac{f_0 t_1 - f_1 t_0}{t_1 - t_0} \tag{35}$$

and
B) heuristically, the smooth fit of v_i at $[b, B]$ will result in a double fit as $[b, B]$ reduces to the singlepoint a, and thus we expect that the derivatives of v_i have also to be equal at a. Letting $m = v_i(a)$, $n = v_i'(a)$, and using the differential equations $g_i v_i = -h_i + \gamma$ yields a linear system for m, n, whose solution is

$$m = \frac{h_0 - h_1}{\mu_1 - \mu_0}. \tag{36}$$

Equating the two expressions for m in the equations (35) and (36) yields an alternative equation for determining a^*.
2)Let k_a denote the switching cost for which the corresponding long run average γ_a equals $\max[\gamma(0), \gamma(1)]$. For transaction costs larger than k_a the optimal policy becomes to always use diffusion 0 or 1 depending on which of the rewards $\gamma(0)$ or $\gamma(1)$ is larger.

We end this section with some numerical examples that utilize the above results.

Example 3 (the symmetric case): Consider the following parameters: $\mu_0 = -1$, $\mu_1 = 1$, $\sigma_0^2 = \sigma_1^2 = 2$, $h_0(x) = 1 - 2x$, $h_1(x) = 2x - 1$ and $\alpha_0 = \alpha_1 = 0$. The reward structure is symmetric around 0.5 which results in 0.5 being the switching point in the absence of switching costs. The optimal (b, B) (also symmetric) and the associated cost k corresponding to a list of values γ are summarized in Table 2. In this case a no switching policy (using either one of the two servers all the time) results in $\gamma = 0.164$. This implies that a double band type policy is better than a *no switch* policy only when $k < 0.8$.
Example 4 (small variances): The specific set of parameters is: $\mu_0 = 1$, $\mu_1 = -1$ and $\sigma_0^2 = \sigma_1^2 = 1/2$, $h_0(x) = 4 - x^2$, $h_1(x) = 5x$, $\alpha_0 = 0$, $\alpha_1 = 0$ and the state space is the interval [0,4]. Figure 3 displays the function $\gamma_\sigma = f/t$ as well as the limit γ_0 obtained when the variances are 0. In fact, an easy explicit expression may be obtained for the zero variance case; and: $\gamma_0 = p_0 h_0 + p_1 h_1$ where $p_0 = -\mu_1/(\mu_0 - \mu_1)$ and $p_1 = \mu_0/(\mu_0 - \mu_1)$. The goodness of the fit (when $\sigma = 0.1$, the two functions are indistinguishable) raises the possibility of getting good results by perturbation analysis.
Remark: Let a_0^*, be the optimal switching point in the previous example for $\sigma = 0$ (i.e a_0^* maximizes $\gamma_0 = p_0 h_0 + p_1 h_1$). Using perturbation arguments as in Hopkins and Blankenship

k	b	B	γ
0	0.5	0.5	0.542
0.028	0.49	0.51	0.541
0.275	0.4	0.6	0.522
0.519	0.3	0.7	0.467
0.701	0.2	0.8	0.382
0.800	0.1	0.9	0.278
0.800	0.01	0.99	0.176

Table 2: Results of Example 3

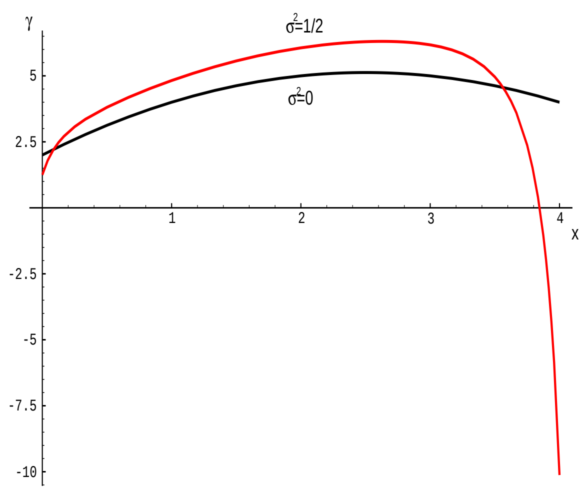

Figure 3: The Case with Zero and Small Variances

[14] we expect to see $a_\epsilon^* = a_0^* + \epsilon\kappa$ for a small variance of ϵ. In fact, an explicit expression for the perturbation paramater κ can be obtained:

$$\kappa = \frac{p_1 \frac{h_1''(a_0^*)}{\mu_1} + p_0 \frac{h_0''(a_0^*)}{\mu_0}}{p_1 h_1''(a_0^*) + p_0 h_0''(a_0^*)} \tag{37}$$

Note that, this expression agrees with equation (5.4a) in [14] for the case of $\mu_0 = \mu_1$.

Example 5: (inventory with linear holding cost and restocking reward) As a last example which is analytically tractable, consider the problem of switching between $\mu_1 = -1$ (demand) and $\mu_0 = 1$ (restocking) on the half line $[0, \infty)$ to minimize the long run average cost induced by the linear holding costs $h_1 = x$ and $h_0 = x - r_s$ (r_s represents the restocking reward). We will choose $r_s = 4e^{a^*} - 4 - 2a^*$ which has the effect of making switching at a^* optimal (this is the "Gittins index" of a^*). Letting $w = \gamma(x) - \gamma(a^*)$ we find that $B - a^*, b - a^*$ are the roots of the equation $(de^d - e^d - 1)/(e^d - e^{-a^*}/2) = w$. The plot of w is provided in Figure 4, for $a^* = 2$.

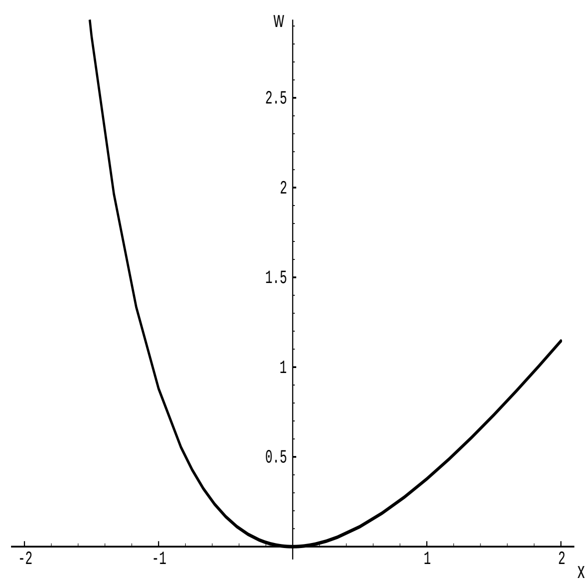

Figure 4: (b, B) values for Example 5

4 Conclusions and Future Research

In this paper we developed an efficient procedure to handle the problem of switching between two diffusion processes in the presence of transaction costs. The continuous state space makes the analysis considerably easier than the discrete analog problem which has been studied a lot in queuing theory. The optimal switching points found may be used as efficient approximations for the analogous queueing problems.

We are currently investigating the application of similar ideas for optimal switching problems in multi-dimensional diffusion processes. The multi-dimensional case is of special interest as the optimization of the discrete models of multi-class queueing networks is typically very hard.

References

[1] Bather, J., "A Continuous Time Inventory Model", *Journal of Applied Probability*, Vol. 3, pp. 538-549, 1966.

[2] Bather J., "A Diffusion Model for the Control of a Dam", *Journal of Applied Probability*, Vol. 5, pp. 55-71, 1968.

[3] Benes, V.E., L.A. Shepp and H.S. Witsenhausen,"Some Solvable Stochastic Control Problems", *Stochastics*,Vol. 4, pp. 39-83, 1980.

[4] Chernoff H. and A. Petkau, "Optimal Control of Brownian Motion", *SIAM Journal on Applied Mathematics*, Vol. 34, pp. 717-731, 1978.

[5] Chung K.J. and R.J. Williams , "Introduction to Stochastic Integration", *Birkhäuser*, Boston, 1990.

[6] Davis, M.H.A. and A.R. Norman, "Portfolio Selection with Transaction Costs", *Mathematics of Operations Research*, Vol. 19, pp. 676-713, 1990.

[7] Foschini, G.J., "On Heavy Traffic Diffusion Analysis and Dynamic Routing in Packet Switched Networks", *Computer Performance*, North Holland, Amsterdam, pp. 419-514, 1977.

[8] Harrison, J.M. and A.J. Taylor, "Optimal Control of a Brownian Storage System", *Stochastic Processes with Applications*, Vol. 6, pp. 179-194, 1978.

[9] Harrison, J.M., "The Diffusion Approximation for Tandem Queues in Heavy Traffic", *Advances in Applied Probability*, Vol. 10, pp. 886-905, 1978.

[10] Harrison, J.M. and M.I. Reiman, "Reflected Brownian Motion in an Orthant", *Annals of Probability*, Vol. 9, pp. 302-308, 1981.

[11] Harrison, J.M. and M.I. Taksar, "Instantaneous Control of Brownian Motion", *Mathematics of Operations Research*, Vol. 8, pp. 439-453, 1983.

[12] Harrison, J.M., T.M. Sellke and A.J. Taylor, "Impulse Control of Brownian Motion", *Mathematics of Operations Research*, Vol. 8, pp. 454-466, 1983.

[13] Harrison, J.M., *Brownian Motion and Stochastic Flow Systems*, John Wiley, New York, 1985.

[14] Hopkins, W.E. and G.L. Blankenship, "Perturbation Analysis of a System of Quasi-Variational Inequalities for Optimal Stochastic Scheduling", *IEEE Transactions on Automatic Control*, Vol. AC-26, pp. 1054-1070, 1981.

[15] Karlin,S and Taylor,H. *A Second Course in Stochastic Processes*, Academic Press Inc., San Diego, 1981.

[16] Krichagina, E.V., S.X.C. Lou, S.P. Sethi and M.I.Taksar, "Production Control in a Failure-Prone Manufacturing System", *Annals of Applied Probability*, Vol. 3, pp. 421-453, 1993.

[17] Krichagina, E.V., S.X.C. Lou and M.I. Taksar, "Double Band Policy for Stochastic Manufacturing Systems in Heavy Traffic", *Mathematics of Operations Research*, Vol. 19, pp. 560-596, 1994.

[18] Mandelbaum, A. "Continuous Multiarmed Bandits", *Annals of Probability*, Vol. 15, pp 1527-1556, 1987.

[19] Newell, G.F., *Approximate Behavior of Tandem Queues, Lecture Notes in Economics and Mathematical Systems, 171*, Springer-Verlag, New York, 1979.

[20] Perry, D. and S.K. Bar-Lev, "A Control of a Brownian Motion Storage System with Two Switchover Drifts", *Stochastic Analysis and Applications*, Vol. 7, pp. 103-115, 1989.

[21] Rath, J., "Controlled Queues in Heavy Traffic", *Advances in Applied Probability*, Vol. 7, pp. 656-671, 1975.

[22] Rath, J., "The Optimal Policy for a Controlled Brownian Motion Process", *SIAM Journal on Applied Mathematics*, Vol. 32, pp. 115-125, 1977.

[23] Reiman, M.I., “Open Queueing Networks in Heavy Traffic”, *Mathematics of Operations Research*, Vol. 9, pp. 451-458, 1984.

[24] Samuelson, P.A. and H.P. Mc Kean, Jr, “Rational Theory of Warrant Pricing”, *Industrial Management Review* Vol. 6, pp 13-39, 1965.

[25] Shepp, L. and A.N. Shiryaev, “The Russion Option: Reduced Regret”, *Advances in Applied Probability*, Vol. 3, pp. 631-640, 1993.

[26] Shreve, S.E., H.M Soner and G.L. Xu, “Optimal Investment and Consumption with Transaction Costs”, *Mathematical Finance*, Vol. 1, pp. 53-84, 1991.

[27] Taksar, M., “Average Optimal Singular Control and a Related Stopping Problem”, *Mathematics of Operations Research*, Vol. 10, pp. 63-81, 1985.

[28] Taksar M., M.J. Klass and D. Assaf, “A Diffusion Model for Optimal Portfolio Selection in the Presence of Brokerage Fees”, *Mathematics of Operations Research*, Vol. 13, pp. 277-294, 1988.

[29] Varadhan S. and R.J. Williams , “Brownian Motion in a Wedge with Oblique Reflection”, *Communications on Pure and Applied Mathematics*, Vol. 38 , pp 405-443, 1985.

Appendix

Optimal Switching to Maximize the Total Reward Until Reaching the Boundaries

In the previous sections, an ergodic reward problem for switching between two diffusion processes was considered. Another related problem of interest is thecase where the processes run until reaching one of the boundaries and receive final penalties depending on the boundary reached. Unfortunately, the method introduced before for ergodic problems does not lead to significant simplification in the transient case.

This time the problem is alternating between two diffusion operators, G_i,$i = 0, 1$, which are stopped at both 0 and 1. Stopping at i leads to a final penalty α_i. Once again, each switch costs k and while running the i'th diffusion rewards are incurred at the rate of $h_i(x)\, dt$. Then, the problem is maximizing with respect to π:

$$U_{x,\pi} = E_{x,\pi}\left[\int_0^A \sum_{i=1}^{2} h_i(X_i(T_i(s)))dT_i(s) - \sum_{i=1}^{2} \alpha_i I_{\{X(A)=i\}} - kN(A)\right], \tag{38}$$

where $\pi = (T_1(t), T_2(t))$ is a *time allocation*, A is the first exit time from the interval $[0,1]$, and $N(A)$ is the number of switches until reaching 0 or 1.

We start with the case $k = 0$. We assume again that there is only one switch point, a^* so that the optimal solution is to use the i'th diffusion on the interval I_i, where $I_0 = [0, a^*]$ and $I_1 = [a^*, 1]$. We will also assume that μ_0 and $-\mu_1$ are nonnegative, which ensures that both diffusions will reach a^* almost surely. We consider first an auxilliary problem, in which, upon reaching some point a, the i-th diffusion is stopped and a reward M_i is received. Running the i-th diffusion starting from x produces then an expected reward $U_i(x)$:

$$U_i(x) = E_x\left[\int_0^{\overline{A}_i} (h_i\ X_i(s)\ ds) + M_i I_{\{X(\overline{A}_i)=a\}}] - \alpha_i I_{\{X(\overline{A}_i=i)\}}\right], \tag{39}$$

where $\overline{A}_i = \min(A_i, A_a)$. Thus, U_i is the solution of

$$\begin{cases} G_i\ U_i &= -h_i \\ U_i(i) &= -\alpha_i \\ U_i(a) &= M_i \end{cases} \tag{40}$$

Now, note that, U_i may be written as

$$U_i(x) = H_i(x) + \frac{S_i(x)}{S_i(a)}\,(M_i - H_i(a)) \tag{41}$$

where H_i, S_i are any solutions of

$$\begin{cases} G_i\ H_i &= -h_i \\ H_i(i) &= -\alpha_i \end{cases} \tag{42}$$

and

$$\begin{cases} G_i\ S_i &= 0 \\ S_i(i) &= 0 \\ S_i(x) &\geq 0, \forall x \in I_i \end{cases} \tag{43}$$

(Letting $W_i = \exp[-\int^x \mu_i]\, dy$, it follows that $S_0(x) = \int_0^x W_0(x)dx,\ S_1(x) = \int_x^1 W_1(x)dx$, $H_0(x) = -\alpha_0 - W_0^{-1}\int_0^x h_0 W_0\, dy$ and $H_1(x) = -\alpha_1 + W_1^{-1}\int_x^1 h_1 W_1\, dy$.) The principle of smooth fit states that at the optimal switch point a^* of the original problem, the expected total rewards U_i, as well as their first two derivatives have to be equal. We will determine now final rewards M_i for the auxiliary problem, stopped at an arbitrary point a, so that the first two derivatives of U_i (but not necessarily U_i themselves) are equal. Letting $U_0'(a) = U_1'(a) := m$, we note (using $U_i''(a) = -\mu_i U_i' - h_i$) that $U_0''(a) = U_1''(a)$ is equivalent with

$$m = m(a) = \frac{h_1 - h_0}{\mu_0 - \mu_1} \tag{44}$$

Since

$$U_i'(a) = H_i'(a) + \frac{S_i'(a)}{S_i(a)} \left(M_i - H_i(a)\right) = m(a) \tag{45}$$

we get finally that

$$M_i = M_i(a) = (m - H_i')\frac{S_i}{S_i'} + H_i. \tag{46}$$

$M_i(a)$ may be interpreted as *subsidies* we would have to pay for stopping at a in such a way that the first two derivatives of the value functions fit smoothly. If, in addition, we had $M_0(a_*) = M_1(a_*)$ at some point a^*,then a^* could be a switch point. If not, it turns out that it is better to use the diffusion requiring a lower subsidy:

Theorem 3: The policy which maximizes the reward in the case of no switching costs is to use the diffusion for which M_i is minimal. For example, if $M_0(0) < M_1(0)$ and $M_1(1) < M_0(1)$ and M_0, M_1 intersect at a unique point a^*, then the optimal policy is to use the 0 diffusion on $[0, a^*]$ and the 1 diffusion on $[a^*, 1]$. Unfortunately, the proof of the theorem does not exploit the intuition above.

Proof of Theorem 3: Let

$$U_i(x) = H_i(x) + \frac{S_i(x)}{S_i(a)} \left(\theta - H_i(a)\right) \tag{47}$$

Below, we will use only the *smooth fit* of the first derivatives $U_0'(a) = U_1'(a)$, and then optimize with respect to a. Letting $f_i = S_i'/S_i$, we note first that the equation above implies that

$$\begin{aligned} \theta\,(f_1 - f_0) &= H_0' - H_1' + f_1\,H_1 - f_0\,H_0 && (48)\\ \text{and}\;\;(\theta - H_i)\,(f_1 - f_0) &= -H_1' + H_0' + f_{1-i}\,(H_1 - H_0) && (49) \end{aligned}$$

(from here on whenever the argument of a function is missing, it will mean the function is evaluated at a). Now, we can simplify the expression of U_i to:

$$\begin{aligned} U_i &= H_i(x) + \frac{S_i(x)\;[H_0' - H_1' + f_{1-i}\,(H_1 - H_0)]}{S_i(a)\,(f_1 - f_0)} && (50)\\ &\triangleq H_i(x) + \frac{S_i(x)}{S_i(a)}\,\frac{A_i(a)}{f_1 - f_0} && (51) \end{aligned}$$

where we denoted by A_i the numerator above. Before taking partial derivatives with respect to a, we note that

$$
\begin{aligned}
f_i' &= \frac{S_i''}{S} - \left(\frac{S_i'}{S_i}\right)^2 && (52)\\
&= -f_i\,(\mu_i + f_i) && (53)
\end{aligned}
$$

(by substituting $-\mu_i\, S_i'$ for S_i'', from the equation). Thus,

$$
f_1' - f_0' = \mu_0\, f_0 - \mu_1\, f_1 - (f_0 + f_1)\,(f_1 - f_0) = R - (f_0 + f_1)\,(f_1 - f_0) \qquad (54)
$$

where we introduce $R = \mu_0\, f_0 - \mu_1\, f_1$. Now, we make the key observation that the domain in which the 0 (1) diffusion is used is the domain in which the condition

$$
\frac{\partial U_i}{\partial a} \ \geq\ 0(\leq 0)\ \text{ if } i = 0\ (i = 1) \qquad (55)
$$

is satisfied. Letting now $H_1 - H_0 = D$, (55) becomes:

$$
\begin{aligned}
\left[-D'' - f_{1-i}\,(\mu_{1-i} + f_{1-i})\,D + f_{1-i}\,D'\right]\ S_i\,(f_1 - f_0) &\ \geq\ (\leq)\\
(-D' + f_{1-i}\,D)\ \left[S_i'\,(f_1 - f_0) + S_i\,[-(f_0 + f_1)\,(f_0 - f_1) + R]\right] & \qquad \Longleftrightarrow \qquad (56)\\
\left[-D'' + f_{1-i}\,(D' - D\,f_{1-i}) - D\,f_{1-i}\,\mu_{1-i}\right]\ (f_1 - f_0) &\ \geq\ (\leq)\\
(-D' + f_{1-i}\,D)\ \left[-f_{1-i}\,(f_1 - f_0) + R\right] & \qquad \Longleftrightarrow \qquad (57)\\
(f_1 - f_0)\,(-D'' - D\,f_{1-i}\,\mu_{1-i}) &\ \geq\ (\leq)\\
-R\,(-D' + f_{1-i}\,D) & \qquad \Longleftrightarrow \qquad (58)\\
D''\,(f_0 - f_1) + R\,D' &\ \geq\ (\leq)\\
D\,f_{1-i}\,f_i\,(\mu_1 - \mu_0) & \qquad (59)
\end{aligned}
$$

Finally, noting that $D'' = h_0 - h_1 - \mu_1\, H_1' + \mu_0\, H_0'$, we have:

$$
\begin{aligned}
(f_0 - f_1)\,(h_0 - h_1 - \mu_1\,H_1' + \mu_0\,H_0') &\ \leq\ (\geq)\\
(H_1 - H_0)\,f_0\,f_1\,(\mu_1 - \mu_0) - \mu_0\,f_0\,H_1' - &\\
\mu_1\,f_1\,H_0' + \mu_0\,f_0\,H_0' + \mu_1\,f_1\,H_1' & \qquad \Longleftrightarrow \qquad (60)\\
(f_0 - f_1)\,(h_1 - h_0) &\ \geq\ (\leq)\\
(\mu_0 - \mu_1)\ \left[f_0\,f_1\,(H_1 - H_0) - f_0\,H_1' + f_1 H_0'\right] & \qquad (61)
\end{aligned}
$$

and since $f_0 > 0$ and $f_1 < 0$, dividing by $f_0 f_1$,and letting $k_i = f_i^{-1}$ we get :

$$
\begin{aligned}
-m\,(k_1 - k_0) &\ \leq\ (\geq)\ \ H_1 - H_0 - k_1\,H_1' + k_0\,H_0' \ \Longleftrightarrow && (62)\\
M_1 &\ \geq\ (\leq)\ \ M_0 && (63)
\end{aligned}
$$

□

Theorem 4: If a double band policy is optimal for a certain switching cost k, then the switching points (b, B) must satisfy the system:

$$a_2 = A_2 A_1^{-1} a_1 \tag{64}$$

where A_1 is the matrix

$$\begin{pmatrix} 1 & -\frac{S_0(b)}{S_0(B)} \\ -\frac{S_1(B)}{S_1(b)} & 1 \end{pmatrix}$$

a_1 is the vector $(H_1(b) - H_0(b) - k, \;\; H_0(B) - H_1(B) - k)'$, A_2 is the matrix

$$\begin{pmatrix} \frac{S_1'(b)}{S_1(b)} & -\frac{S_0'(b)}{S_0(B)} \\ -\frac{S_1'(B)}{S_1(b)} & \frac{S_0'(B)}{S_0(B)} \end{pmatrix}$$

and a_2 is the derivative of a_1. The proof is omitted.

Remark: It is also possible in this case to eliminate k and retrieve the conjugacy equation

$$\begin{aligned} [H_1(B) - H_1(b) - H_0(B) + H_0(b)][S_1'(B)S_0'(b) - S_0'(B)S_1'(b)] &= \\ S_0[-S_1'(b)d(B) + S_1'(B)d(b)] &+ \\ S_1[-S_0'(B)d(b) + S_0'(b)d(B)] & \end{aligned} \tag{65}$$

where $d = H_0' - H_1'$. However, note that solving the conjugacy equation (65) is not in this case significantly easier than solving the system (64).

We conclude this section with two examples. Example 4, is a numerical application of the procedure. Example 5 considers a special case where the indices (M_i) are explicitly computable.

Example 5: Let $\mu_0 = 2$, $\mu_1 = -1$, $\sigma_0^2 = \sigma_1^2 = 2$, $h_0(x) = 4$, $h_1(x) = 1$ and $\alpha_0 = 9$, $\alpha_1 = 10$. In this case, the indices are given by: $M_0(x) = 1.16667e^x - 2.5x - 10.1667$ and $M_1(x) = 7.69403e^{-0.5x} + x - 15.6667$. Figure 5 displays the index functions; the switching point a is 0.7143 and it is optimal to use diffusion 0 when $x < 0.7143$ and to use diffusion 1 otherwise.

Example 6: Suppose that $\mu_0 = \lambda > 0$, $\mu_1 = \mu < 0$, h_0 and h_1 are constants. Then, $H_0(x) = -\alpha_0 - (h_0 x)/\lambda$, $H_1(x) = -\alpha_1 + (h_1(1-x))/\mu$, $S_0(x) = (1 - e^{-\lambda x})/\lambda$, $S_1(x) = (e^{\mu(1-x)} - 1)/\mu$ and $m = (h_0 - h_1)/(\mu - \lambda)$. The indices are:

$$M_0 = -\frac{h_0}{\lambda}x - \alpha_0 + \left(\frac{e^{\lambda x} - 1}{\lambda}\right)\left(m + \frac{h_0}{\lambda}\right), \tag{66}$$

$$M_1 = \frac{h_1}{\mu}(1-x) - \alpha_1 + \left(\frac{e^{-\mu(1-x)} - 1}{\mu}\right)\left(m + \frac{h_1}{\mu}\right) \tag{67}$$

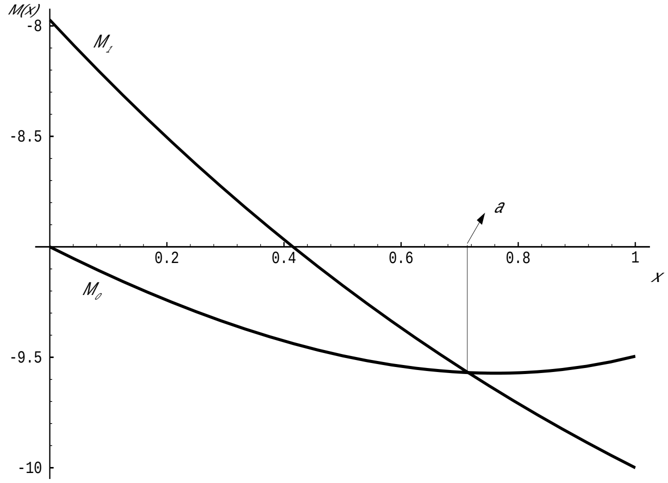

Figure 5: Optimal Switching Point for Example 4

The conditions for the existence of a switch point can be obtained by noting that $(\partial U_0/\partial a)(0) > 0$ and $(\partial U_1/\partial a)(1) < 0$ must hold. These two conditions translate into:

$$\begin{cases} -\alpha_0 < \frac{h_1}{\mu}[\frac{e^{-\mu}-1}{\mu} + 1] - \alpha_1 + m\frac{e^{-\mu}-1}{\mu} \\ \frac{h_0}{\lambda}[\frac{e^{\lambda}-1}{\lambda} - 1] - \alpha_0 + m\frac{e^{\lambda}-1}{\lambda} > -\alpha_1 \end{cases} \tag{68}$$

$$\Longleftrightarrow \begin{cases} \alpha_1 - \alpha_0 < \lambda\frac{e^{\mu}-1}{\mu}\ \frac{(h_1/\mu+h_0/\lambda)}{(\mu+\lambda)} - \frac{h_1}{\mu} \\ \alpha_0 - \alpha_1 < \mu\frac{e^{\lambda}-1}{\lambda}\ \frac{(h_1/\mu+h_0/\lambda)}{(\mu+\lambda)} - \frac{h_0}{\lambda} \end{cases} \tag{69}$$

Letting, $k = [(h_0/\lambda) - (h_1/\mu)]/(\lambda - \mu)$ yields

$$\frac{\lambda}{\mu}(e^{-\mu} - 1)k - \frac{h_1}{\mu} \leq \alpha_0 - \alpha_1 \leq -\frac{h_0}{\lambda} - \frac{\mu}{\lambda}(e^{\lambda} - 1)k \tag{70}$$

Now,

$$\frac{\lambda}{\mu}(e^{-\mu} - 1)k - \frac{h_1}{\mu} \leq -\frac{h_0}{\lambda} - \frac{\mu}{\lambda}(e^{\lambda} - 1)k \Longleftrightarrow$$

$$\frac{\lambda}{\mu}(e^{-\mu} - 1)k \leq k(\mu - \lambda) - \frac{\mu}{\lambda}(e^{\lambda} - 1)k \Longleftrightarrow$$

Since $\lambda > 0$ and $-\mu > 0$, we get:

$$k\left[\frac{-\mu}{(\lambda - \mu)}\frac{e^{\lambda} - 1}{\lambda} + \frac{\lambda}{(\lambda - \mu)}\frac{e^{-\mu} - 1}{-\mu} - 1\right] > 0$$

which reduces to

$$k > 0 \tag{71}$$

since the multiplier of k in (71) above is positive.The result of this example will be summarized in the following corollary.

Corollary: Consider the set of parameters in Example 5. A switch point a^* exists if and only if $h_0/\lambda - h_1/\mu > 0$ and if $\alpha_1 - \alpha_0$ satisfies (70) in which case a^* is given by

$$M_0(a^*) = M_1(a^*) \tag{72}$$

For example, if $\lambda = -\mu$ and $h_0 = h_1 = h$, (71) is satisfied and (70) is equivalent to

$$|\alpha_1 - \alpha_0| < \frac{h(\exp(\lambda) - 1 - \lambda)}{\lambda^2}$$

The switch condition (72) becomes

$$\frac{h(e^{\lambda a} - 1 - \lambda a)}{\lambda^2} = \frac{h(e^{\lambda(1-a)} - 1 - \lambda(1-a))}{\lambda^2} + \alpha_0 - \alpha_1$$

If furthermore $\alpha_0 = \alpha_1$, the solution is $a = 1/2$, and if $\alpha_0 - \alpha_1 = h(\exp(\lambda) - 1 - \lambda)\lambda^2$ the solution is $a = 1$.