



HAL
open science

Impulse Control of Standard Brownian Motion: Long-Term Average Criterion

Kurt Helmes, Richard H. Stockbridge, Chao Zhu

► **To cite this version:**

Kurt Helmes, Richard H. Stockbridge, Chao Zhu. Impulse Control of Standard Brownian Motion: Long-Term Average Criterion. 26th Conference on System Modeling and Optimization (CSMO), Sep 2013, Klagenfurt, Austria. pp.148-157, 10.1007/978-3-662-45504-3_14 . hal-01286407

HAL Id: hal-01286407

<https://inria.hal.science/hal-01286407v1>

Submitted on 10 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Impulse control of standard Brownian motion: Long-term average criterion^{*}

Kurt Helmes¹, Richard H. Stockbridge², and Chao Zhu²

¹ Institut für Operations Research, Humboldt-Universität zu Berlin
helmes@wiwi.hu-berlin.de

² Department of Mathematical Sciences, University of Wisconsin – Milwaukee
Milwaukee, WI 53201, USA
{stockbri,zhu}@uwm.edu

Abstract. This paper examines the impulse control of a standard Brownian motion under a long-term average criterion. In contrast with the dynamic programming approach, this paper first imbeds the stochastic control problem into an infinite-dimensional linear program over a space of measures and then reduces the problem to a simpler nonlinear optimization that has a familiar interpretation. One is able to easily identify the optimal cost and a family of optimal impulse control policies.

Keywords: impulse control, long-term average criterion, infinite dimensional linear programming, expected occupation and impulse measures.

1 Introduction

When one seeks to control a stochastic process and every intervention incurs a strictly positive cost, one must select a sequence of separate intervention times and amounts. The resulting stochastic problem is therefore an impulse control problem in which the decision maker seeks to either maximize a reward or minimize a cost. This paper examines the impulse control of the prototypical process Brownian motion under a long-term average cost criterion; a companion paper studies the impulse control of Brownian motion under a discounted criterion. The aim of the paper is to illustrate a solution approach which first imbeds the stochastic control problem into an infinite-dimensional linear program over a space of measures and then reduces the linear program to a simpler nonlinear optimization. This approach provides a new method for determining an optimal impulse control policy.

Let W be a standard Brownian motion process with natural filtration $\{\mathcal{F}_t\}$. An impulse control policy consists of a pair of sequences $(\tau, Y) := \{(\tau_k, Y_k) : k \in \mathbb{N}\}$ in which τ_k is the $\{\mathcal{F}_t\}$ -stopping time of the k th impulse and the \mathcal{F}_{τ_k} -measurable variable Y_k gives the k th impulse size. The sequence $\{\tau_k : k \in \mathbb{N}\}$ is required to be non-decreasing, a natural assumption in that intervention $k + 1$

^{*} This research was supported in part by National Science Foundation under grant DMS-1108782 and by grant award 246271 from the Simons Foundation.

must occur no earlier than intervention k . For a policy (τ, Y) , the impulse-controlled Brownian motion process is given by

$$X(t) = x_0 + W(t) + \sum_{k=1}^{\infty} I_{\{\tau_k \leq t\}} Y_k.$$

The goal of the decision maker is to keep the process close to zero and to minimize the long-term average cost incurred by the impulse policy. Define the running/deviation cost rate function by $c_0(x) = x^2$ and set the impulse costs to be $c_1(y, z) = k_1 + k_2|y - z|$ for $(y, z) \in \mathbb{R}^2$, in which $k_1 > 0$ and $k_2 \geq 0$ and y denotes the pre-intervention location which will typically be far away from zero, while z denotes the post-intervention location of the process X , which should be close to zero. Note, in particular, that there is a strictly positive cost for every impulse, even one in which $Y_k = 0$ which does not affect the value of X . Let (τ, Y) be an impulse control policy. The quantity to be minimized is

$$J(\tau, Y) := \limsup_{t \rightarrow \infty} t^{-1} \mathbb{E} \left[\int_0^t c_0(X(s)) ds + \sum_{k=1}^{\infty} I_{\{\tau_k \leq t\}} c_1(X(\tau_k^-), X(\tau_k)) \right]. \quad (1)$$

Clearly any policy (τ, Y) for which $J(\tau, Y) = \infty$ is undesirable so to be *admissible*, we require (τ, Y) to have a finite cost; the nonnegativity of c_0 and strict positivity of c_1 indicates that every policy will have nonnegative long-term average cost. The collection of all admissible impulse policies is denoted by \mathcal{A} .

Similar type of problems have been extensively investigated in the literature. An incomplete list includes the now classical works on general stochastic impulse problems [1, 4, 8, 13] as well as their applications in various areas such as portfolio optimization, inventory control, risk management, control of a dam and exchange rate intervention [2, 3, 10–12]. In particular, [11] explains the adoption of a Brownian motion model.

Unlike the aforementioned references, in which the primary tool is the dynamic programming principle and its associated quasi-variational inequalities, this paper aims to illustrate the utility of a different methodology, namely, the linear programming approach. In such an approach, we embed the stochastic impulse control problem into an infinite-dimensional linear program over appropriate measures. Further, the linear program, with the aid of an auxiliary linear program, is transformed into a nonlinear optimization problem. Then both the value of the impulse control problem and an optimal impulse control policy are easily determined. The linear programming approach toward stochastic control problem can be dated back to [9] for discrete time and a finite state space and to [14, 15] for regular stochastic control problems in continuous time with general state space. It has been further developed in [5–7] for optimal stopping and singular control problems. This paper aims to expand the utility of such a methodology to impulse control problems as well.

We make four important observations about impulse policies. Firstly, the “no-intervention policy” in which the process is $X(t) = x_0 + W(t)$ for all $t \geq 0$ incurs an infinite long-term average cost so is inadmissible. Secondly, let (τ, Y)

be an impulse policy and define $\tau_\infty := \lim_{k \rightarrow \infty} \tau_k$. Should $\tau_\infty < \infty$ on a set of positive probability, then the fixed cost $k_1 > 0$ per intervention also results in an infinite long-term average cost. Thus for every admissible policy $\tau_k \rightarrow \infty$ a.s. as $k \rightarrow \infty$. Thirdly, let (τ, Y) be a policy for which there is some k such that $\tau_k = \tau_{k+1}$ on a set of positive probability. Again due to the presence of the fixed intervention cost k_1 , the total cost up to time τ_{k+1} will be at least $k_1 \mathbb{P}(\tau_k = \tau_{k+1})$ smaller by combining these interventions into a single intervention on this set. Hence we may restrict policies to those for which $\tau_k < \tau_{k+1}$ almost surely for each k .

The final observation is similar. Suppose (τ, Y) is a policy such that on a set G of positive probability $\tau_k < \infty$ and $|X(\tau_k)| > |X(\tau_k -)|$ for some k . Consider a modification of this impulse policy and resulting process \tilde{X} which simply fails to implement this impulse on G . Define the stopping time $\sigma = \inf\{t > \tau_k : |X(t)| \leq |\tilde{X}(t)|\}$. Notice that the running costs accrued by \tilde{X} over $[\tau_k, \sigma)$ are smaller than those accrued by X . Finally, at time σ , introduce an intervention on the set G which moves the \tilde{X} process so that $\tilde{X}(\sigma) = X(\sigma)$. This intervention will incur a cost which is no greater than the cost for the process X at time τ_k . As a result, we may restrict the impulse control policies to those for which no impulse increases the distance of the process from the origin (an intuitively obvious observation).

2 Restricted problem and measure formulation

The initial analysis considers a restricted collection of impulse policies.

Condition 1 *Let \mathcal{A}_1 denote the set of policies such that for $(\tau, Y) \in \mathcal{A}_1$ the resulting process X remains bounded; that is for some $M < \infty$, $|X(t)| \leq M$ for all $t \geq 0$.*

Intuitively, Condition 1 is not much of a restriction since unbounded processes occur by allowing the Brownian motion to diffuse which incurs an expensive running cost. This restriction, however, is needed so that a transversality condition is satisfied and a stochastic integral is a martingale. Following the initial solution, the general class of impulse policies will be analyzed.

We capture the expected behavior of the process and impulses with measures. Arbitrarily fix $(\tau, Y) \in \mathcal{A}_1$ and let M be as given in Condition 1. For each $t > 0$, define the average expected occupation and average expected impulse measures $\mu_0^{(t)}$ and $\mu_1^{(t)}$, respectively, such that for each $G, G_1, G_2 \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned} \mu_0^{(t)}(G) &= t^{-1} \mathbb{E} \left[\int_0^t I_G(X(s)) ds \right], \text{ and} \\ \mu_1^{(t)}(G_1 \times G_2) &= t^{-1} \mathbb{E} \left[\sum_{k=1}^{\infty} I_{\{\tau_k \leq t\}} I_{G_1 \times G_2}(X(\tau_k -), X(\tau_k)) \right]. \end{aligned} \tag{2}$$

It is immediate that $\mu_0^{(t)}$ is a probability measure for each $t > 0$ and since $(\tau, Y) \in \mathcal{A}_1 \subset \mathcal{A}$, the finiteness of $J(\tau, Y)$ implies that the collection of measures

$\{\mu_1^{(t)} : t > 0\}$ is uniformly bounded above. We also note that the second component being measured by $\mu_1^{(t)}$ is the post-jump location so $\mu_1^{(t)}$ is a measure on the product space of (pre-jump, post-jump) pairs. In light of Condition 1, the support of $\mu_1^{(t)}$ is contained in the compact set $\mathcal{R} := \{(y, z) : |z| \leq |y| \leq M\}$. Similarly, each $\mu_0^{(t)}$ has support in the compact interval $[-M, M]$. As a result, these collections are tight and hence relatively compact.

Now notice the objective function (1) can be expressed as

$$J(\tau, Y) = \limsup_{t \rightarrow \infty} t^{-1} \left[\int c_0(x) \mu_0^{(t)}(dx) + \int c_1(y, z) \mu_1^{(t)}(dy \times dz) \right].$$

Let $\{t_j : j \in \mathbb{N}\}$ be a sequence with $t_j \rightarrow \infty$ as $j \rightarrow \infty$ such that

$$t_j^{-1} \left[\int c_0(x) \mu_0^{(t_j)}(dx) + \int c_1(y, z) \mu_1^{(t_j)}(dy \times dz) \right] \rightarrow J(\tau, Y).$$

For $i = 0, 1$, the relative compactness of $\{\mu_i^{(t_j)} : j \in \mathbb{N}\}$ implies that there exist weak limits μ_0 and μ_1 . Note μ_0 is a probability measure whereas μ_1 is a finite measure. Since c_0 and c_1 are bounded and continuous on $[-M, M]$ and \mathcal{R} , respectively,

$$J(\tau, Y) = \int c_0(x) \mu_0(dx) + \int c_1(y, z) \mu_1(dy \times dz). \quad (3)$$

It is now helpful to characterize the value of functions of the process. For $f \in \mathcal{D} = C^2(\mathbb{R})$,

$$\begin{aligned} f(X(t)) &= f(x_0) + \int_0^t f'(X(s)) dW(s) \\ &\quad + \int_0^t \frac{1}{2} f''(X(s)) ds + \sum_{k=1}^{\infty} [f(X(\tau_k)) - f(X(\tau_k-))] I_{\{\tau_k \leq t\}}. \end{aligned} \quad (4)$$

The generator A of the Brownian motion process is $Af(x) = \frac{1}{2} f''(x)$; define the jump operator B by $Bf(y, z) = f(z) - f(y)$. First taking expectations, then dividing by t and letting $t \rightarrow \infty$ in (4) results in

$$\limsup_{t \rightarrow \infty} t^{-1} \mathbb{E} \left[\int_0^t Af(X(s)) ds + \sum_{k=1}^{\infty} I_{\{\tau_k \leq t\}} Bf(X(\tau_k-), X(\tau_k)) \right] = 0; \quad (5)$$

note that the boundedness of $X(t)$ along with $f \in C^2(\mathbb{R})$ implies both the transversality condition $\lim_{t \rightarrow \infty} t^{-1} \mathbb{E}[f(X(t))] = 0$ holds and that the stochastic integral exists and has mean 0. (The same argument applies by taking the limit inferior in (4), so in fact, left-hand side of (5) is a limit.) The fact that f and f'' are continuous and bounded on $[-M, M]$ means that

$$\int Af(x) \mu_0^{(t_k)} \rightarrow \int Af(x) \mu_0(dx)$$

and

$$\int Bf(y, z) \mu_1^{(t_k)}(dy \times dz) \rightarrow \int B(y, z) \mu_1(dy \times dz)$$

and hence (5) can be written in terms of these measures as

$$\int Af(x) \mu_0(dx) + \int Bf(y, z) \mu_1(dy \times dz) = 0. \tag{6}$$

The restricted impulse control problem is therefore imbedded in the linear program of minimizing (3) over pairs of measures (μ_0, μ_1) satisfying the constraints (6) for every $f \in \mathcal{D}$. Since there may be pairs (μ_0, μ_1) which do not correspond to any $(\tau, Y) \in \mathcal{A}_1$, it follows that the value of the linear program is a lower bound on the minimal long-run average cost of the impulse control problem. Observe also that by further restricting the collection of functions for which the constraint is required to be satisfied, the corresponding “auxiliary” linear program may have even more feasible pairs and hence will provide an even lower bound on the value of the impulse control problem. These observations are summarized in the following theorem.

Theorem 2 *Let V denote the optimal value of the long-term average impulse control problem, V_{lp} denote the optimal value of the linear program which seeks to minimize (3) over measures satisfying (6) and V_{aux} be the optimal value of an auxiliary linear program which limits (6) to a smaller collection of functions $f \in \mathcal{D}_1$ for some $\mathcal{D}_1 \subset \mathcal{D}$. Then $V_{aux} \leq V_{lp} \leq V$.*

3 Partial solution: First auxiliary linear program

It would be helpful to reduce the complexity of the linear program by reducing the number of constraints. We first consider the constraints (6). The intuition is rather straightforward. Consider a function ϕ for which $A\phi \equiv -1$. Then since μ_0 is a probability measure for each feasible pair (μ_0, μ_1) , the identity (6) becomes

$$\int_{\{|z| \leq |y|\}} B\phi(y, z) \mu_1(dy \times dz) = 1. \tag{7}$$

The general solution to the equation $A\phi \equiv -1$ is $\phi(x) = -x^2 + ax + b$, in which a, b are constants. We shall select a solution ϕ so that $B\phi(x, y) = \mathbb{E}_x[\tau_y]$, where $|x| < |y|$ and $\tau_y := \inf\{t \geq 0 : |X(t)| = |x + W(t)| = |y|\}$. To this end, we notice that the function $u(x) := -x^2 + y^2$ solves the boundary value problem

$$\begin{cases} Au(x) = -1, & x \in (-|y|, |y|), \\ u(-|y|) = u(|y|) = 0. \end{cases} \tag{8}$$

Therefore the optional sampling theorem implies that for any $x \in (-|y|, |y|)$ and $t \geq 0$, we have

$$\mathbb{E}_x [u(X(t \wedge \tau_y))] = u(x) + \mathbb{E}_x \left[\int_0^{t \wedge \tau_y} Au(X(s)) ds \right] = u(x) - \mathbb{E}_x [t \wedge \tau_y].$$

Since $X(t \wedge \tau_y)$ is bounded and $\tau_y < \infty$ a.s., utilizing the boundary conditions in (8), letting $t \rightarrow \infty$ in the above equation yields $\mathbb{E}_x[\tau_y] = u(x) = -x^2 + y^2$. Therefore by selecting $\phi(x) = -x^2, x \in \mathbb{R}$, the term $B\phi(x, y)$ in (7) gives the expected time it takes the Brownian motion process starting at x to hit the set $\{-|y|, |y|\}$.

In a similar manner, by considering the boundary value problem

$$\begin{cases} Au(x) = -c_0(x), & x \in (-|y|, |y|), \\ u(-|y|) = u(|y|) = 0, \end{cases} \quad (9)$$

it follows that the function $g_0(x) := -\frac{1}{6}x^4$ satisfies

$$Bg_0(x, y) = g_0(y) - g_0(x) = \mathbb{E}_x \left[\int_0^{\tau_y} c_0(X(s)) ds \right], \quad |x| < |y|. \quad (10)$$

The following proposition establishes the required identity.

Proposition 3 *Let $(\tau, Y) \in \mathcal{A}_1$ and let X denote the resulting impulse controlled process. Recall $\{t_j : j \in \mathbb{N}\}$ is a set of times such that*

$$t_j^{-1} \left[\int c_0 d\mu_0^{(t_j)} + \int c_1 d\mu_1^{(t_j)} \right] \rightarrow J(\tau, Y).$$

Let (μ_0, μ_1) be a weak limit of $(\mu_0^{(t_j)}, \mu_1^{(t_j)})$ as $j \rightarrow \infty$. Then

$$\int c_0(x) \mu_0(dx) = \int Bg_0(y, z) \mu_1(dy \times dz). \quad (11)$$

Proof. Without loss of generality, assume that $\mu_0^{(t_j)} \Rightarrow \mu_0$ and similarly $\mu_1^{(t_j)} \Rightarrow \mu_1$. Using g_0 in (4) and taking expectations yields for each t_j

$$\begin{aligned} \mathbb{E}_{x_0}[g_0(X(t_j))] &= g_0(x_0) + \mathbb{E}_{x_0} \left[\int_0^{t_j} Ag_0(X(s)) ds \right. \\ &\quad \left. + \sum_{k=0}^{\infty} I_{\{\tau_k \leq t_j\}} Bg_0(X(\tau_k-), X(\tau_k)) \right]. \end{aligned}$$

Since for $(\tau, Y) \in \mathcal{A}_1$, $X(t)$ remains bounded, dividing by t_j , using the definitions of $\mu_0^{(t_j)}$ and $\mu_1^{(t_j)}$ in (2) and letting $j \rightarrow \infty$ establishes the result. \square

We are now ready to define the first auxiliary linear program. Restrict the constraint to the single function $\phi(x) = -x^2$ and use Proposition 3 to rewrite the objective function. The resulting linear program is

$$\begin{cases} \text{Min.} & \int [c_1(y, z) + Bg_0(y, z)] \mu_1(dy \times dz) \\ \text{S.t.} & \int B\phi(y, z) \mu_1(dy \times dz) = 1. \end{cases} \quad (12)$$

Notice the support of each feasible μ_1 is in the set $\{(y, z) \in \mathbb{R}^2 : |z| \leq |y|\}$. Thus $B\phi(y, z) = y^2 - z^2 \geq 0$ and the constraint of (12) implies that $B\phi(y, z) = y^2 - z^2$ is a probability density for every feasible measure μ_1 of (12).

We emphasize that the auxiliary linear program includes the cost of any impulse policy $(\tau, Y) \in \mathcal{A}_1$. These policies are not required to be of feedback type or even stationary. It is only required that the resulting controlled process remain bounded.

Observe that the constraint does not impose any mass restrictions on the set $\{(y, z) : B\psi(y, z) = 0\}$. However, the goal is to minimize the objective function and since the impulse cost function $c_1 > k_1 > 0$ any mass placed on this set will only increase the cost. We may therefore restrict the optimization to those measures μ_1 having support in $\{(y, z) : B\phi(y, z) > 0\} = \{(y, z) : |z| < |y|\}$. As a result the objective function can be rewritten as

$$\begin{aligned} & \int [c_1(y, z) + Bg_0(y, z)] \mu_1(dy \times dz) \\ &= \int \left(\frac{c_1(y, z) + Bg_0(y, z)}{B\phi(y, z)} \right) \cdot B\phi(y, z) \mu_1(dy \times dz) \end{aligned}$$

and so the problem reduces to the minimization of

$$F(y, z) = \frac{c_1(y, z) + Bg_0(y, z)}{B\phi(y, z)} \quad (13)$$

over $\{(y, z) : |z| \leq |y|\}$. Observe that $F(-y, -z) = F(y, z)$.

Remark 4 *The minimization of F in (13) has a very familiar interpretation. Let y and z be such that $0 \leq |z| < |y|$. Let $\tau_y = \inf\{t \geq 0 : |X(t)| = |y|\}$. Recall from (10) that $Bg_0(y, z)$ represents the expected running cost for the cycle $[0, \tau_y)$. Now consider the impulse policy in which impulses occur only when the process X hits either y or $-y$ and then jumps to z or $-z$, respectively. The symmetry of the fixed cost function c_1 means that $c_1(y, z) = c_1(-y, -z)$ and this cost is assessed at the end of the cycle. Note also $B\phi(y, z)$ gives the expected time it takes the Brownian motion starting from $-z$ or z to reach $-y$ or y . Hence the function F represents the ratio of the expected cost per cycle over the expected cycle length taken for such impulse control policies. The significance of this reformulation is that the linear programming imbedding allows arbitrary impulse policies in the class \mathcal{A}_1 yet the resulting nonlinear optimization corresponds to minimizing the cost over a subclass of these policies.*

We now determine an optimal impulse control policy.

Theorem 5 *There exist values $y_* > z_* > 0$ such that*

$$F(y_*, z_*) = F(-y_*, -z_*) = \inf_{(y, z) \in \mathcal{R}} F(y, z).$$

Proof. First since $B\phi(y, z) = y^2 - z^2$ and $Bg_0(y, z) = \frac{1}{6}(y^4 - z^4)$, the function $F(y, z) = \frac{k_1 + k_2|y-z|}{y^2 - z^2} + \frac{1}{6}(y^2 + z^2)$. Observe that as $y \rightarrow \infty$ or $z \rightarrow \infty$ or $|y-z| \rightarrow 0$, $F(y, z) \rightarrow \infty$. Hence F achieves its minimal value at some point (y_*, z_*) which by optimality requires y_* and z_* to have the same sign. \square

Theorem 6 *An optimizing pair (y_*, z_*) having positive components is the level set of the function $h(x) = \frac{k_2}{2x} + \frac{x^2}{3}$ at the level $F(y_*, z_*)$.*

Proof. Consider the function F on the set $\{(y, z) : 0 < z < y\}$ so $|y - z| = y - z$. the first-order optimality conditions are

$$\frac{[y^2 - z^2](k_2 + (2/3)y^3) - [k_1 + k_2(y - z) + (1/6)(y^4 - z^4)](2y)}{[y^2 - z^2]^2} = 0, \quad (14)$$

$$\frac{[y^2 - z^2](-k_2 - (2/3)z^3) + [k_1 + k_2(y - z) + (1/6)(y^4 - z^4)](2z)}{[y^2 - z^2]^2} = 0, \quad (15)$$

which are satisfied for pairs (y, z) such that

$$F(y, z) = h(y) = h(z), \quad (16)$$

where the function h is defined in the statement of the theorem. We have $h'(x) = \frac{4x^3 - 3k_2}{6x^2}$. When $x > 0$, one observes that h strictly decreases from $+\infty$ until it reaches a minimum at $\sqrt[3]{3k_2/4}$ after which it strictly increases to $+\infty$. For $x < 0$, h is strictly decreasing from $+\infty$ to $-\infty$. The level sets of h consist of either a single value with $x < 0$ when the level lies below the minimum of h over the positive reals or three points with $x < 0 < z < y$ when the level is above this minimum value. \square

It is still necessary to connect the optimal value of the linear program (12) to the optimal value of the long-term average impulse control problem. Examine the relative magnitudes of the roots given in the proof of Theorem 6; by definition $0 < z < y$. Observe the root $x < 0$ is such that $|x| > y$ and hence $x < -y < -z$.

Theorem 7 *The optimal long-term average cost for the restricted impulse control problem is $F(y_*, z_*)$ and an optimal impulse control policy (τ^*, Y^*) is defined by:*

$$\begin{cases} \tau_1^* = 0, \\ Y_1^* = z_* - x_0, \end{cases} \quad \begin{cases} \tau_k^* = \inf\{t > \tau_{k-1}^* : X(t-) = \pm y_*\}, \\ Y_k^* = \text{sgn}(X(\tau_k^* -))z_* - X(\tau_k^* -), \end{cases} \quad k \geq 2. \quad (17)$$

Remark 8 *Due to the nature of the long-term average criterion, many other optimal policies exist. In fact, any impulse control policy $(\tau, Y) \in \mathcal{A}_1$ can be used for a finite length of time so long as after some point the process is in the interval $[-y_*, y_*]$ and the policy of jumping to z_* when the process hits y_* and jumping to $-z_*$ at the time of hitting $-y_*$ is adopted.*

Proof. Consider the impulse control policy (17) and observe that $(\tau^*, Y^*) \in \mathcal{A}_1$. Also for each $t > 0$ by (2), $\mu_1^{(t)}$ has its support on the set

$$\{(x_0, z_*), (-y_*, -z_*), (y_*, z_*)\}$$

but the limiting measure μ_1 only has mass on $\{(-y_*, -z_*), (y_*, z_*)\}$. As a result, $J(\tau^*, Y^*) = F(y_*, z_*)$ and hence by Theorem 2 is optimal. \square

4 General solution

The solution obtained in the previous section does not include impulse control policies which allow the process X to become unbounded in either direction. It is therefore necessary to show that such policies cannot provide a lower cost.

Theorem 9 *The optimal long-term average cost over the collection of all admissible impulse control policies is $F(y_*, z_*)$ and any policy which eventually only impulses when the process reaches $\pm y_*$, with impulses respectively to $\pm z_*$, is optimal.*

Proof. Let (τ, Y) be an arbitrary admissible impulse control policy in \mathcal{A} and let X denote the resulting process. We shall prove that F_* is a lower bound on $J(\tau, Y)$ and since F_* is achieved by the impulse control policy (17), this will establish the result.

We use a localization argument. For each $n \in \mathbb{N}$, define the stopping time $\sigma_n = \inf\{t \geq 0 : |X(t)| \geq n\}$ and to simplify notation, let $F_* = F(y_*, z_*)$. We consider the function $f(x) = F_*\phi(x) - g_0(x) = \frac{1}{6}x^4 - F_*x^2$, $x \in \mathbb{R}$. Due to the choices of ϕ and g_0 , we have $Af(x) = c_0(x) - F_*$, and for any $|z| < |y|$,

$$\begin{aligned} Bf(y, z) &= F_*(\phi(z) - \phi(y)) - (g_0(z) - g_0(y)) \\ &\leq \frac{c_1(y, z) + g_0(z) - g_0(y)}{\phi(z) - \phi(y)} \cdot (\phi(z) - \phi(y)) - (g_0(z) - g_0(y)) = c_1(y, z). \end{aligned}$$

Then applying Itô's formula yields

$$\begin{aligned} f(X(t \wedge \sigma_n)) &= f(x_0) + \int_0^{t \wedge \sigma_n} Af(X(s)) ds + \int_0^{t \wedge \sigma_n} f'(X(s)) dW(s) \\ &\quad + \sum_{k=1}^{\infty} I_{\{\tau_k \leq t \wedge \sigma_n\}} Bf(X(\tau_k-), X(\tau_k)) \\ &\leq f(x_0) + \int_0^{t \wedge \sigma_n} [c_0(X(s)) - F_*] ds + \int_0^{t \wedge \sigma_n} f'(X(s)) dW(s) \\ &\quad + \sum_{k=1}^{\infty} I_{\{\tau_k \leq t \wedge \sigma_n\}} c_1(X(\tau_k-), X(\tau_k)). \end{aligned}$$

Taking expectations on both sides, and rearranging the terms, it follows that

$$\begin{aligned} F_*\mathbb{E}_{x_0}[t \wedge \sigma_n] &\leq f(x_0) - \mathbb{E}_{x_0}[f(X(t \wedge \sigma_n))] + \mathbb{E}_{x_0} \left[\int_0^{t \wedge \sigma_n} c_0(X(s)) ds \right] \\ &\quad + \mathbb{E}_{x_0} \left[\sum_{k=1}^{\infty} I_{\{\tau_k \leq t \wedge \sigma_n\}} c_1(X(\tau_k-), X(\tau_k)) \right] \\ &\leq f(x_0) - K + \mathbb{E}_{x_0} \left[\int_0^{t \wedge \sigma_n} c_0(X(s)) ds \right] \\ &\quad + \mathbb{E}_{x_0} \left[\sum_{k=0}^{\infty} c_1(X(\tau_k-), X(\tau_k)) I_{\{\tau_k \leq t \wedge \sigma_n\}} \right], \end{aligned}$$

where the last inequality follows from the observation that $f(x) = \frac{1}{6}x^4 - F_*x^2 \geq K > -\infty$ for some constant K . Letting $n \rightarrow \infty$, we know $\sigma_n \rightarrow \infty$ almost surely so the monotone convergence theorem yields

$$F_*t \leq f(x_0) - K + \mathbb{E}_{x_0} \left[\int_0^t c_0(X(s)) ds + \sum_{k=0}^{\infty} c_1(X(\tau_k-), X(\tau_k)) I_{\{\tau_k \leq t\}} \right].$$

Dividing by t and letting $t \rightarrow \infty$, we obtain

$$\begin{aligned} F_* &\leq \limsup_{t \rightarrow \infty} t^{-1} \mathbb{E}_{x_0} \left[\int_0^t c_0(X(s)) ds + \sum_{k=1}^{\infty} c_1(X(\tau_k-), X(\tau_k)) I_{\{\tau_k \leq t\}} \right] \\ &= J(\tau, Y). \quad \square \end{aligned}$$

References

1. Bensoussan, A., Lions, J.-L.: Impulse control and quasi-variational inequalities. Boreas Editions (1984)
2. Chen, X., Sim, M., Simchi-Levi, D., Sun, P.: Risk aversion in inventory management. *Oper. Res.* 55 (5), 828–842 (2007)
3. Eastham, J.F., Hastings, K.J.: Optimal impulse control of portfolios. *Math. Oper. Res.* 13 (4), 588–605 (1988)
4. Harrison, J.M., Sellke, T.M., Taylor, A.J.: Impulse control of brownian motion. *Math. Oper. Res.* 8 (3), 454–466 (1983)
5. Helmes, K., Stockbridge, R.H.: Construction of the value function and optimal rules in optimal stopping of one-dimensional diffusions. *Adv. in Appl. Probab.* 42 (1), 158–182 (2010)
6. Kurtz, T.G., Stockbridge, R.H.: Existence of Markov controls and characterization of optimal Markov controls. *SIAM J. Control Optim.* 36 (2), 609–653 (electronic) (1998)
7. Kurtz, T.G., Stockbridge, R.H.: Stationary solutions and forward equations for controlled and singular martingale problems. *Electron. J. Probab.* 6, 1–52 (2001). Paper No. 14.
8. Lions, P.L., Perthame, B.: Quasi-variational inequalities and ergodic impulse control. *SIAM J. Control Optim.* 24 (4), 604–615 (1986)
9. Manne, A.S.: Linear Programming and Sequential Decisions. *Management Sci.* 6, 259–267 (1960)
10. Melas, D., Zervos, M.: An ergodic impulse control model with applications. In: Piunovskiy, A.B. (ed.) *Modern Trends in Controlled Stochastic Processes*. Luniver Press, Frome, UK, 161–180 (2010)
11. Ormeci, M., Dai, J.G., Vate, J.V.: Impulse control of brownian motion: the constrained average cost case. *Oper. Res.* 56 (3), 618–629 (2008)
12. Piunovskiy, A.B.: *Examples in Markov Decision Processes*. Imperial College Press, London (2013)
13. Robin, M.: On some impulse control problems with long run average cost. *SIAM J. Control Optim.* 19 (3), 333–358 (1981)
14. Stockbridge, R.H.: Time-average control of martingale problems: Existence of a stationary solution. *Ann. Probab.* 18 (1), 190–205 (1990)
15. Stockbridge, R.H.: Time-average control of martingale problems: A linear programming formulation. *Ann. Probab.* 18 (1), 206–217 (1990)