On the Convergence from Discrete to Continuous Time in an Optimal Stopping Problem*

Paul Dupuis and Hui Wang
Lefschetz Center for Dynamical Systems
Brown University
Providence, R.I. 02912
USA

Abstract

We consider the problem of optimal stopping for a one dimensional diffusion process. Two classes of admissible stopping times are considered. The first class consists of all non-anticipating stopping times that take values in $[0, \infty]$, while the second class further restricts the set of allowed values to the discrete grid $\{nh : n = 0, 1, 2, \cdots, \infty\}$ for some parameter $h > 0$. The value functions for the two problems are denoted by $V(x)$ and $V^h(x)$, respectively. We identify the rate of convergence of $V^h(x)$ to $V(x)$ and the rate of convergence of the stopping regions, and provide simple formulas for the rate coefficients.

Keywords. Optimal stopping, continuous time, discrete time, diffusion process, rate of convergence, local time.

1 Introduction

One of the classical formulations of stochastic optimal control is that of optimal stopping. In optimal stopping, the only decision to be made is when to stop the process. When the process is stopped, a benefit is received (or a cost is paid), and the objective is to maximize the expected benefit (or minimize the expected cost). Although the problem formulation is very simple, this optimization problem has many practical applications. Examples include the pricing problems in investment theory, the valuation of American options, the development of natural resources, etc.; see, e.g., [1, 2, 4, 5, 6, 9, 10, 15, 16, 17, 18].

The formulation of the optimal stopping problem requires the specification of the class of allowed stopping times. Typically, one assumes these to be non-anticipative in an appropriate sense, so that the control does not have knowledge of the future. Another important restriction is with regard to the actual time values at which one can stop, and here there are two important cases: continuous time and discrete time. In the first case, the stopping time is allowed to take values in the interval $[0, \infty]$, with ∞ corresponding to the decision to never stop. In the second case there is a fixed discrete set of times $D \subset [0, \infty]$, and the

*Research supported in part by the National Science Foundation (NSF-DMS-0072004) and the Army Research Office (ARO-DAAD19-99-1-0223).

stopping time must be selected from this set. Typically, this discrete set is a regular grid, e.g., $D^h \doteq \{nh : n \in \mathbb{N}_0\}$, where $h > 0$ is the grid spacing.

In the present paper we focus exclusively on the one dimensional case. Although the statement of precise assumptions is deferred to Section 2.1, a rough description of the continuous and discrete time problems we consider is as follows.

Continuous time optimal stopping. We use the stochastic process model

$$\frac{dS_t}{S_t} = b(S_t)\,dt + \sigma(S_t)\,dW_t$$

where b and σ are bounded continuous functions from $\mathbb{R}$ to $\mathbb{R}$. Although the results can be extended to cover other diffusion models as well, we focus on this model because of its wide use in optimal stopping problems that occur in economics and finance. We consider a payoff defined in terms of a convex nondecreasing function $\phi : \mathbb{R} \to [0, \infty)$. The payoff from stopping at time t is $\phi(S_t)$, and the decision maker wants to maximize the expected present value by judiciously choosing a stopping time. This is modeled by the optimal stopping problem with value function

$$V(x) = \sup_{\tau \in \mathcal{S}} \mathbb{E}\left[e^{-r\tau}\phi(S_\tau) \;\middle|\; S_0 = x\right],$$

where r is the discount rate and $\mathcal{S}$ is the set of all admissible stopping times, which are allowed to take values in $[0, \infty]$. The dynamic programming equation for this problem is as follows. Let

$$\mathcal{L}V(x) = \frac{1}{2}\sigma^2(x)x^2V''(x) + b(x)xV'(x).$$

Then

$$\max\left[\phi(x) - V(x), \mathcal{L}V(x) - rV(x)\right] = 0. \tag{1.1}$$

In the case where ϕ is convex and nondecreasing, it is often optimal to stop when the process S_t first exceeds some fixed threshold x_*. In this case, the value function $V(x)$ equals $\phi(x)$ for $x \geq x_*$, and it satisfies the ordinary differential equation $-rV(x) + \mathcal{L}V(x) = 0$ for $x < x_*$. For the case where σ and b are constants, $V(x)$ takes the form Ax^β for $x < x_*$. Here β is the positive root of some quadratic equation, and (A, x_*) are constants that can be computed explicitly using the principle of smooth fit, i.e., the value function is $\mathcal{C}^1$ across the optimal exercise boundary x_*.

Discrete time optimal stopping. In this case the process model is the same as before, but the set of possible stopping times is restricted to those that take values in the time grid $D^h \doteq \{nh, n \in \mathbb{N}_0\}$. The optimal strategy is often similar to the continuous time case: stop the first time S_{nh} exceeds some fixed threshold x_*^h. Let $V^h(x)$ denote the value function. The pair $\left(V^h(x), x_*^h\right)$ satisfy the dynamic programming equation [21]

$$V^h(x) = \begin{cases} \phi(x) & , \quad x \in [x_*^h, \infty) \\ e^{-rh}\mathbb{E}\left[V^h(S_h) | S_0 = x\right], & x \in (0, x_*^h). \end{cases}$$

Closed-form solutions to this dynamic programming equation are not usually available.

The aim of the present paper is to examine the connection between these two optimal stopping problems as $h \to 0$. There are two questions of main interest:

- *What is the convergence rate of the optimal exercise boundary x_*^h to x_*, and what is the rate coefficient?*
- *What is the convergence rate of the value function $V^h(x)$ to $V(x)$, and what is the rate coefficient?*

As we will see in Section 2, the optimal exercise boundaries converge with rate $\sqrt{h}$, while the value functions converge with rate h. In both cases there is a well defined rate coefficient. The coefficient in the case of the exercise boundary is defined in terms of the expected value of a functional of local time of Brownian motion, while the coefficient for the value function involves both local time and excursions of Brownian motion. For problems where the continuous time problem can be more or less solved explicitly (e.g., the one dimensional problems considered in the present work), these results allow one to explicitly compute accurate approximations for the discrete time problem. For problems where the continuous time problem does not have an explicit solution (e.g., multidimensional problems), the analogous information could possibly be used to improve the quality of approximation obtained using numerical approximations.

Few existing results are concerned with the rate of convergence of approximations for this class of problems. Lamberton [14] considers the binomial tree approximation for pricing American options and obtains upper and lower bounds (though not a rate of convergence) for the value function. In his approximation both the time and state variables are discretized. References to a few papers giving qualitatively similar results also appear in [14].

The outline of the paper is as follows. In Section 2 we introduce notation and define the basic optimization problems. We state the main result, give an illustrative example, and then lay out the main steps in the proof of the approximation theorem. The proofs of two key approximations which are intimately connected with the local time and excursions of Brownian motion are given in Section 3. The paper concludes with an Appendix in which a result on a conditional distribution of the exit time is proved.

2 The Approximation Theorem

2.1 Notation, assumptions and background

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{F})$ with filtration $\mathbb{F} = (\mathcal{F}_t)$ satisfying the usual conditions: right-continuity and completion by $\mathbb{P}$-negligible sets. The *state process* $S = (S_t, \mathcal{F}_t)$ is modeled by

$$\frac{dS_t}{S_t} = b(S_t)\,dt + \sigma(S_t)\,dW_t, \qquad S_0 \equiv x. \tag{2.1}$$

Here $W = (W_t, \mathcal{F}_t)$ is a standard $\mathbb{F}$-Brownian motion.

Define value function

$$V(x) = \sup_{\tau \in \mathcal{S}} \mathbb{E}\left[e^{-r\tau}\phi(S_\tau) \mid S_0 = x\right],$$

where the supremum is over all stopping times with respect to the filtration $\mathbb{F}$. Define

$$V^h(x) = \sup_{\tau \in \mathcal{S}^h} \mathbb{E}\left[e^{-r\tau}\phi(S_\tau) \mid S_0 = x\right],$$

where $\mathcal{S}^h$ is the set of all stopping times that take values in D^h.

The following assumptions will be used throughout the paper.

Condition 2.1. *1. The coefficients $b : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}$ are bounded and continuous, with $\inf_{x \in \mathbb{R}} \sigma(x) > 0$. Furthermore $xb(x)$ and $x\sigma(x)$ are Lipschitz continuous.*

2. $\phi : \mathbb{R} \to [0, \infty)$ is non-decreasing, and both ϕ and its derivative ϕ' are of polynomial growth. Furthermore

$$\sup_{t \geq 0} e^{-rt}\phi(S_t) \in \mathbb{L}^1, \qquad \lim_{t \to \infty} e^{-rt}\phi(S_t) = 0, \quad a.s.$$

3. The "continuation" region for the continuous-time optimal stopping problem takes the form $\{x : V(x) > \phi(x)\} = (0, x_*)$.

4. The "continuation" region for the discrete-time optimal stopping problem takes the form $\{x : V^h(x) > \phi(x)\} = (0, x_*^h)$.

5. The payoff function ϕ is twice continuously differentiable in a neighborhood of x_*.

6. The smooth-fit-principle holds, that is, the value function V is $\mathcal{C}^1$ across the optimal exercise boundary x_*.

As noted in the Introduction, V satisfies the dynamic programming equation

$$\max\left[\phi(x) - V(x), \mathcal{L}V(x) - rV(x)\right] = 0.$$

Note that usually V is only once continuously differentiable across the optimal exercise boundary $x = x_*$. Since $\phi(x) = V(x)$ if $x \in [x_*, \infty)$ and $\phi(x) < V(x)$ if $x \in (0, x_*)$, it follows that $V''(x_*-) \geq \phi''(x_*)$, where the $-$ denotes limit from the left. Define

$$A \doteq \frac{V''(x_*-) - \phi''(x_*)}{\phi(x_*)} \geq 0. \tag{2.2}$$

Although one can construct examples where $A = 0$, it is typically the case that $A > 0$. We will assume this condition below, and merely note that the rate of convergence of the optimal threshold does not depend on A at all.

Remark 2.1. The change of variable $t = -\log x$ can be used to transform the ordinary differential equation (ODE) $\mathcal{L}f(x) - rf(x) = 0$ on $(0, \infty)$ into the ODE

$$\frac{1}{2}\sigma(e^{-t})W''(t) + \left[\frac{1}{2}\sigma(e^{-t}) - b(e^{-t})\right] W'(t) - rW(t) = 0$$

on $\mathbb{R}$. Since $\sigma(x) > 0$ for $x > 0$, the classical theory for solutions of ODEs [3] can be used to show that the general solution to $\mathcal{L}V(x) - rV(x) = 0$ can be written in the form $c_1 f_1(x) + c_2 f_2(x)$, where $f_1(x)$ is positive and bounded as $x \downarrow 0$ and $f_2(x)$ is unbounded as $x \downarrow 0$. Under Condition 2.1, the function f_1 is twice continuously differentiable on $(0, \infty)$. $V(x)$ is then equal to $c_1 f_1(x)$ for $x \in (0, x_*]$ and equal to $\phi(x)$ for $x \in [x_*, \infty)$, where c_1 and x_* are determined by the principle of smooth fit, i.e.,

$$c_1 f_1(x_*) = \phi(x_*) \text{ and } c_1 f_1'(x_*) = \phi'(x_*).$$

Remark 2.2. In the case that S is a geometric Brownian motion with $b(x) \equiv b$ and $\sigma(x) \equiv \sigma$, and $\phi(x) = (x-k)^+$ for some constant k, then Condition 2.1 holds when $r > b$. For $r \le b$, the value function for the optimal stopping problem is $+\infty$, and there is no optimal stopping time; see [6].

Remark 2.3. It is usually not *a prior* clear if parts 3 and 4 of Condition 2.1 hold for a general state process. Here we give a sufficient condition that is very easy to verify in the case $\phi(x) = (x-k)^+$. Suppose parts 1 and 2 of Condition 2.1 hold, and in addition that

$$r \ge \sup_{x \in (0,\infty)} \left\{ b(x) + x b'(x) \right\}.$$

We claim that parts 3 and 4 of Condition 2.1 hold. We will show that part 3 holds and omit the analogous proof for 4. Define $V_T(x) \doteq \sup_{\tau \le T} \mathbb{E}^x \left[e^{-r\tau} \phi(S_\tau) \right]$. Let S^x stand for the state process starting from $S_0 \equiv x$. A small modification of the proof of Theorem 5.2 [7] shows that the collection of random variables

$$\sup_{0 \le t \le T} \left| \frac{S^x(t) - S^y(t)}{x - y} \right|, \qquad y \in (x - \delta, x + \delta)$$

is uniformly integrable for fixed x, small $\delta > 0$ and terminal time T. It follows immediately that $V_T(x)$ is a continuous function, since for any stopping $\tau \le T$, we have

$$\mathbb{E} \left| e^{-r\tau} \phi\big(S^x(\tau)\big) - e^{-r\tau} \phi\big(S^y(\tau)\big) \right| \le \mathbb{E} \sup_{0 \le t \le T} |S^x(t) - S^y(t)| \le c|x - y|$$

for some constant c. Furthermore, the upper left Dini derivate of V_T is always bounded by one, i.e.,

$$\limsup_{y \uparrow x} \frac{V_T(x) - V_T(y)}{x - y} \le 1.$$

To see this, let τ_* be the optimal stopping time when $S_0 \equiv x$. The existence of τ_* is guaranteed by the classical theory of Snell envelop; see [11]. We have

$$V_T(y) \ge \mathbb{E} \left[e^{-r\tau_*} \phi(S^y_{\tau_*}) \right],$$

which implies that

$$\frac{V_T(x) - V_T(y)}{x - y} \le \mathbb{E} \left[\frac{e^{-r\tau_*} \phi(S^x_{\tau_*}) - e^{-r\tau^*} \phi(S^y_{\tau_*})}{x - y} \right].$$

However, if $x \geq y$, then $S^x(t) \geq S^y(t)$ for all t by strong uniqueness. Since ϕ is non-decreasing and

$$\phi(x) - \phi(y) = (x-k)^+ - (y-k)^+ \leq x - y \qquad \forall\, x \geq y,$$

we have

$$\frac{V_T(x) - V_T(y)}{x-y} \leq \mathbb{E}\left[\frac{e^{-r\tau_*}\left(S^x_{\tau_*} - S^y_{\tau_*}\right)}{x-y}\right].$$

Using the uniform integrability, it follows that

$$\limsup_{y \uparrow x} \frac{V_T(x) - V_T(y)}{x-y} \leq \mathbb{E}\left[\lim_{y\uparrow x}\frac{e^{-r\tau_*}\left(S^x_{\tau_*} - S^y_{\tau_*}\right)}{x-y}\right] = \mathbb{E}\left[e^{-r\tau^*} D^x(\tau_*)\right],$$

where $D^x(t) \doteq \frac{\partial}{\partial x} S^x(t)$ satisfies the SDE

$$\frac{dD^x(t)}{D^x(t)} = \left[b(S^x_t) + S^x_t b'(S^x_t)\right]\, dt + \left[\sigma(S^x_t) + S^x_t \sigma'(S^x_t)\right]\, dW_t, \qquad D^x(0) = 1.$$

See, e.g., [12, 19]. Since by assumption $r \geq \sup_{x\in(0,\infty)} [b(x) + xb'(x)]$, it is easy to check that

$$\limsup_{y \uparrow x} \frac{V_T(x) - V_T(y)}{x-y} \leq \mathbb{E}\left[e^{-r\tau^*} D^x(\tau_*)\right] \leq 1.$$

It follows from a standard result in real analysis (see [20, Proposition 5.1.2]) that

$$V_T(x) - V_T(y) \leq x - y \qquad \forall\, x \geq y.$$

This implies that

$$\{x : V_T(x) = \phi(x)\} = [x^*_T, \infty)$$

for some real number x^*_T. Indeed, if $V_T(y) = \phi(y) = (y-k)^+$, then since $V_T > 0$ we must have $y > k$. It follows that for all $x \geq y$,

$$V_T(x) \leq V_T(y) + (x-y) = (y-k) + (x-y) = x - k = \phi(x).$$

But $V_T \geq \phi$ trivially, whence $V_T(x) = \phi(x)$ for all $x \geq y$.

It remains to show that part 3 of Condition 2.1 holds. It suffices to observe that for all x

$$V_T(x) \uparrow V(x)$$

as $T \to \infty$. This completes the proof.

Remark 2.4. If S is a geometric Brownian motion with $b(x) \equiv b$, and $\sigma(x) \equiv \sigma$, and $\phi(x) = (\sum_i A_i x^{\alpha_i} - k)^+$ for some positive constants (A_i, α_i) and $k \geq 0$, then one can show that $V(x) - \phi(x)$ is decreasing, which in turn implies that parts 3 and 4 of Condition 2.1 hold. A similar argument can be found in [9].

2.2 Rates of convergence: value function and stopping region

Let $u \in [0,1)$, let W be a Brownian motion with $W_u = 0$, and define $N \doteq \inf\{n \in \mathbb{N} : W_n \geq 0\}$. Note that $N < \infty$ w.p.1. Define

$$H(u) \doteq \mathbb{E}W_N^2 \qquad \text{and} \qquad M(u) \doteq \mathbb{E}W_N. \tag{2.3}$$

In terms of these functions we define the constants

$$C = \int_0^1 H(u)\,du \qquad \text{and} \qquad K = \int_0^1 M(u)\,du. \tag{2.4}$$

These quantities are shown to be finite in Lemma 3.2. Note that $H(u) > M^2(u)$, and therefore

$$C = \int_0^1 H(u)\,du > \int_0^1 M^2(u)\,du \geq \left(\int_0^1 M(u)\,du\right)^2 = K^2.$$

Our main result is the following.

Theorem 2.1. *Assume Condition 2.1, and define the constants A, C, and K by (2.2) and (2.4). Assume that $A > 0$. The following conclusions hold for all $x \in (0, x_*)$.*

1.

$$\frac{V^h(x) - V(x)}{V(x)} = -\frac{1}{2}Ax_*^2\sigma^2(x_*)(C - K^2)h + o(h)$$

2.

$$x_*^h = x_* - Kx_*\sigma(x_*)\sqrt{h} + o(\sqrt{h}).$$

Remark 2.5. Using an elementary argument by contradiction, one can show that the asymptotic expansion in the preceding theorem holds uniformly in any compact subset of $(0, x_*)$.

Example: Consider the special case where $b(x) \equiv b$ and $\sigma(x) \equiv \sigma$. Assume $r > b$ and $\phi(x) = (x-k)^+$ for some constant $k > 0$. It follows that the value function for the continuous time optimal stopping problem is

$$V(x) = \begin{cases} Bx^\alpha & ; \quad x < x_* \\ x - k & ; \quad x \geq x_* \end{cases}$$

where

$$\alpha = \left(\frac{1}{2} - \frac{b}{\sigma^2}\right) + \sqrt{\left(\frac{1}{2} - \frac{b}{\sigma^2}\right)^2 + \frac{2r}{\sigma^2}} \qquad\qquad B = \frac{x_* - k}{x_*^\alpha},$$

and

$$x_* = \frac{\alpha}{\alpha - 1}k.$$

It follows that

$$x_*^h = x_*(1 - K\sigma\sqrt{h}) + o(\sqrt{h}) = \frac{\alpha k}{\alpha - 1}(1 - K\sigma\sqrt{h}) + o(\sqrt{h})$$

and

$$A = \frac{V''(x_*-) - \phi''(x_*)}{x_* - k} = \frac{B\alpha(\alpha-1)(x_*)^{\alpha-2}}{x_* - k},$$

which implies that

$$\frac{V^h(x) - V(x)}{V(x)} = -\frac{1}{2}Ax_*^2\sigma^2(C - K^2)h + o(h) = -\frac{1}{2}\alpha(\alpha-1)(C-K^2)\sigma^2 h + o(h).$$

2.3 Overview of the proof

In this subsection we outline and prove the main steps in the proof of Theorem 2.1. The proofs of two key asymptotic expansions are deferred to the next section.

For the simplicity of future analysis, we first introduce a bounded modification of the payoff function ϕ. This modification will not affect the asymptotics at all; see Proposition 2.1.

Let $\bar{\phi} \le \phi$ be an increasing function satisfying

$$\bar{\phi}(x) = \begin{cases} \phi(x), & \text{if } x \le x_* + a \\ k\;, & \text{if } x \ge x_* + 2a. \end{cases} \tag{2.5}$$

Here a and k are two positive constants, whose specific values are not important. Without loss of generality, we assume that $\bar{\phi}$ is twice continuously differentiable in the region $[x_*, \infty)$. Suppose h and δ are two positive constants, and let $x_\delta \doteq x_* - \delta$. We consider the quantities

$$\bar{W}_\delta(x) \doteq \mathbb{E}^x\left[e^{-r\tau_\delta}\bar{\phi}(S_{\tau_\delta})\right] \quad\text{and}\quad W_\delta(x) = \mathbb{E}^x\left[e^{-r\tau_\delta}\phi(S_{\tau_\delta})\right],$$

where

$$\tau_\delta \doteq \inf\{t \ge 0 : \; S_t \ge x_\delta\},$$

and $\mathbb{E}^x$ denotes expectation conditioned on $S_0 = x$. Note that $W_\delta(x) = \bar{W}_\delta(x)$ for all $x \le x_\delta$. We also define

$$\bar{W}_\delta^h(x) \doteq \mathbb{E}^x\left[e^{-r\tau_\delta^h}\bar{\phi}(S_{\tau_\delta^h})\right] \quad\text{and}\quad W_\delta^h(x) \doteq \mathbb{E}^x\left[e^{-r\tau_\delta^h}\phi(S_{\tau_\delta^h})\right],$$

where

$$\tau_\delta^h \doteq \inf\{nh \ge 0 : S_{nh} \ge x_\delta\}.$$

Main idea of the proof. The main idea for proving the rates of convergence is as follows. Write

$$W_\delta^h(x) - V(x) = \left[W_\delta^h(x) - \bar{W}_\delta^h(x)\right] + \left[\bar{W}_\delta^h(x) - \bar{W}_\delta(x)\right] + \left[\bar{W}_\delta(x) - V(x)\right].$$

For each term, we will obtain approximations as h and δ tend to zero. It turns out the leading term has the following form:

$$-\frac{1}{2}a_1\delta^2 + a_2\delta\sqrt{h} + a_3 h + \text{higher order term}.$$

Here a_1, a_2, and a_3 are constants (some of which depend on x), with $a_1 > 0$. Since the function $W_\delta^h(x)$ attains its maximum at $\delta_* = x_* - x_*^h$, one would expect that δ_* approximately maximizes the leading term, or

$$\delta_* = \frac{a_2}{a_1}\sqrt{h} + o(\sqrt{h}). \tag{2.6}$$

Furthermore, substituting this back in one would expect

$$V^h(x) = W_{\delta_*}^h(x) = \left(-\frac{a_2^2}{a_1} + a_3\right) h + o(h).$$

This is in fact how the argument will proceed. We begin with the estimation of the first term, which turns out to be negligible for small h and δ. Define the quantity

$$\triangle_{\delta,h} \doteq W_\delta^h(x) - \bar{W}_\delta^h(x) = \mathbb{E}^x\left[e^{-r\tau_\delta^h}\left(\phi(S_{\tau_\delta^h}) - \bar{\phi}(S_{\tau_\delta^h})\right)\right]. \tag{2.7}$$

We have the following result.

Proposition 2.1. *Define $\triangle_{\delta,h}$ by (2.7). There exist constants $L < \infty$ and $\varepsilon > 0$ such that*

$$|\triangle_{\delta,h}| \le Le^{-\frac{\varepsilon}{h}}$$

for all sufficiently small δ and h.

Proof. The proof of the proposition is based on the following bound. Let a be as in the characterization (2.5) of $\bar{\phi}$. Then for any $x \le x_*$ and $y \ge x_* + a$,

$$\mathbb{P}\left(S_h > y \,\middle|\, S_0 = x\right) \le \exp\left\{-\left[\log\frac{y}{x_*} - c_1 h\right]^2 \middle/ c_2 h\right\}, \tag{2.8}$$

where the finite constants c_1, c_2 depend only on the coefficients b, σ. To prove this bound we use the expression

$$S_h = S_0 \exp\left\{\int_0^h \left[b(S_t) - \frac{1}{2}\sigma^2(S_t)\right] dt + \int_0^h \sigma(S_t)\, dW_t\right\}.$$

Define $c_1 \doteq \|b\|_\infty + \frac{1}{2}\|\sigma^2\|_\infty$. Since $x \le x_*$

$$p \doteq \mathbb{P}^x\left(S_h > y\right) \le \mathbb{P}^x\left(e^{\int_0^h \sigma(S_t)\, dW_t} \ge e^{\log\frac{y}{x_*} - c_1 h}\right).$$

However, if $B \doteq \log\frac{y}{x_*} - c_1 h$ and $c_2 \doteq 2\|\sigma^2\|_\infty$, then for $\theta > 0$

$$\begin{aligned} p &\le \mathbb{P}^x\left(e^{\theta\int_0^h \sigma(S_t)\, dW_t} \ge e^{\theta B}\right) \\ &\le \mathbb{P}^x\left(e^{\theta\int_0^h \sigma(S_t)\, dW_t - \frac{1}{2}\theta^2\int_0^h \sigma(S_t)^2\, dt} \ge e^{\theta B - \frac{1}{4}\theta^2 c_2 h}\right) \\ &\le e^{-\theta B + \frac{1}{4}\theta^2 c_2 h}. \end{aligned}$$

The last inequality follows from Chebychev's inequality. Minimizing the right hand side over θ completes the proof of (2.8).

We now complete the proof of the proposition. To ease the exposition, we use τ in lieu of τ_δ^h throughout the proof. We have

$$\begin{aligned}\triangle_{\delta,h} &= \sum_{n=1}^{\infty} e^{-rnh}\mathbb{E}^x\left[\left(\phi(S_{nh}) - \bar{\phi}(S_{nh})\right) 1_{\{\tau=n\}}\right] \\ &\leq \sum_{n=1}^{\infty} e^{-rnh}\int_{a+x_*}^{\infty} \left|\phi'(y) - \bar{\phi}'(y)\right| \mathbb{P}^x(S_{nh} > y, \tau = n)\, dy.\end{aligned}$$

We also have, for any $y > x_* + a$, that

$$\begin{aligned}&\mathbb{P}^x(S_{nh} > y, \tau = n) \\ &\quad= \mathbb{P}^x(S_{nh} > y \,|\, \tau = n)\mathbb{P}^x(\tau = n) \\ &\quad= \mathbb{P}^x(S_{nh} > y \,|\, S_{nh} \geq x_\delta, S_{(n-1)h} < x_\delta, \cdots, S_0 < x_\delta)\mathbb{P}^x(\tau = n).\end{aligned}$$

Define the stopping time

$$\sigma \doteq \inf\{t \geq (n-1)h : S_t \geq x_\delta\}.$$

Then

$$\begin{aligned}&\mathbb{P}^x(S_{nh} > y \,|\, \tau = n) \\ &\quad= \mathbb{P}^x(S_{nh} > y \,|\, \sigma \leq nh, S_{nh} \geq x_\delta, S_{(n-1)h} < x_\delta, \cdots, S_0 < x_\delta) \\ &\quad= \int_0^h \mathbb{P}^x(S_{nh} > y \,|\, \sigma = nh - t, S_{nh} \geq x_\delta, S_{(n-1)h} < x_\delta, \cdots, S_0 < x_\delta) \\ &\qquad\qquad \cdot \mathbb{P}^x(\sigma \in nh - dt \,|\, \sigma \leq nh, S_{nh} \geq x_\delta, S_{(n-1)h} < x_\delta, \cdots, S_0 < x_\delta).\end{aligned}$$

However, the strong Markov property implies for all $t \in [0, h]$ that

$$\begin{aligned}&\mathbb{P}^x(S_{nh} > y \,|\, \sigma = nh - t, S_{nh} \geq x_\delta, S_{(n-1)h} < x_\delta, \cdots, S_0 < x_\delta) \\ &\quad= \mathbb{P}(S_t > y \,|\, S_0 = x_\delta, S_t \geq x_\delta) \\ &\quad= \frac{\mathbb{P}(S_t > y \,|\, S_0 = x_\delta)}{\mathbb{P}(S_t \geq x_\delta \,|\, S_0 = x_\delta)}.\end{aligned}$$

The denominator in this display is uniformly bounded from below away from zero for $t \in [0, 1]$:

$$\mathbb{P}(S_t \geq x_\delta \,|\, S_0 = x_\delta) \geq \alpha > 0, \qquad \forall\, t \in\ [0, 1];$$

see Lemma 3.4 for a proof. Using (2.8), for all small $h > 0$ and $t \in (0, h)$

$$\mathbb{P}(S_t > y \,|\, S_0 = x_\delta) \leq \exp\left\{ - \left[\log\frac{y}{x_*} - c_1 t\right]^2 \bigg/ c_2 t\right\} \leq \exp\left\{ - \left[\log\frac{y}{x_*} - c_1 h\right]^2 \bigg/ c_2 h\right\}.$$

Now since ϕ' is of polynomial growth and $\bar{\phi}'(x)$ is zero for large x, it follows that there are finite constants R and m such that

$$\left|\phi'(y) - \bar{\phi}'(y)\right| \leq R y^{m-1} \text{ for all } y > x_* + a.$$

Hence, for all small $\delta > 0$, the change of variable $x = \log \frac{y}{x_*} - c_1 h$ gives

$$\begin{aligned}
\triangle_{\delta,h} &\leq \frac{R}{\alpha} \sum_{n=1}^{\infty} e^{-rnh} \mathbb{P}^x(\tau = n) \int_{a+x_*}^{\infty} y^{m-1} \exp\left\{ - \left[\log \frac{y}{x_*} - c_1 h \right]^2 \Big/ c_2 h \right\} dy \\
&= \frac{R}{\alpha} (x_*)^m e^{mc_1 h} \sum_{n=1}^{\infty} e^{-rnh} \mathbb{P}^x(\tau = n) \int_{\log\left(1+\frac{a}{x_*}\right) - c_1 h}^{\infty} e^{mx - \frac{x^2}{c_2 h}} \, dx.
\end{aligned}$$

For h small enough, there exists positive numbers $\bar{a}, \bar{C}, \bar{c}$ such that

$$\begin{aligned}
\triangle_{\delta,h} &\leq \bar{C} \sum_{n=1}^{\infty} e^{-rnh} \mathbb{P}^x(\tau = n) \int_{\log\left(1+\frac{\bar{a}}{x_*}\right)}^{\infty} e^{-\frac{x^2}{\bar{c}h}} \, dx \\
&= \bar{C}\sqrt{2\pi\bar{c}} \cdot \Phi\left(- \log\left(1 + \frac{\bar{a}}{x_*}\right) \Big/ \sqrt{\bar{c}h} \right) \cdot \sqrt{h} \sum_{n=1}^{\infty} e^{-rnh} \mathbb{P}^x(\tau = n) \\
&\leq \bar{C}\sqrt{2\pi\bar{c}} \cdot \Phi\left(- \log\left(1 + \frac{\bar{a}}{x_*}\right) \Big/ \sqrt{\bar{c}h} \right) \cdot \sqrt{h}.
\end{aligned}$$

Here Φ is the cumulative distribution function for the standard normal distribution. We complete the proof of the proposition by using the asymptotic relation

$$\Phi(-x) \sim \frac{1}{\sqrt{2\pi}x} e^{-\frac{x^2}{2}}$$

as $x \to \infty$. □

The bound just proved shows that $\left[W_\delta^h(x) - \bar{W}_\delta^h(x)\right]$ is exponentially small as $h \to 0$, uniformly for all small $\delta > 0$. We now consider the terms $\left[\bar{W}_\delta^h(x) - \bar{W}_\delta(x)\right]$ and $\left[\bar{W}_\delta(x) - V(x)\right]$. When considering the asymptotic behavior of these terms, it is often convenient to scale δ with h as $h \to 0$ in the manner suggested by (2.6). For the remainder of this proof, unless explicitly stated otherwise, we will assume that

$$\delta = c\sqrt{h} + o(\sqrt{h}) \text{ as } h \to 0 \tag{2.9}$$

for a non-negative parameter c. With an abuse of notation, the quantities $\bar{W}_\delta^h(x)$ and $\bar{W}_\delta(x)$ will be denoted by $\bar{W}_c^h(x)$ and $\bar{W}_c(x)$ when the relation (2.9) holds.

We next estimate $\left[\bar{W}_c(x) - V(x)\right]$ as $h \to 0$.

Proposition 2.2. *Assume Condition 2.1 and define A by (2.2). Assume also that $A > 0$. Then*

$$\bar{W}_c(x) - V(x) = \left[-\frac{1}{2} A c^2 h + o(h) \right] V(x).$$

Proof. Recall that $V(x)$ can be characterized, for $x \leq x_*$, as a multiple of the bounded (in a neighborhood of zero) solution f_1 to $\mathcal{L}f(x) - rf(x) = 0$; see Remark 2.1. $\bar{W}_c(x)$ can be likewise characterized, with the constant determined by the boundary condition $\bar{W}_c(x_\delta) = \bar{\phi}(x_\delta)$. Thus

$$\bar{W}_c(x) = \frac{\bar{\phi}(x_\delta)}{V(x_\delta)} V(x) \text{ for all } x \in (0, x_\delta].$$

We now expand for small $\delta \geq 0$, and use $x_\delta \doteq x_* - \delta, V(x_*) = \bar{\phi}(x_*), V'(x_*) = \bar{\phi}'(x_*)$, and the definition of A to obtain

$$\frac{\bar{W}_c(x) - V(x)}{V(x)} = \left(\frac{\bar{\phi}}{V}\right)'\Bigg|_{x_*} \cdot (-\delta) + \frac{1}{2}\left(\frac{\bar{\phi}}{V}\right)''\Bigg|_{x_*} \cdot (-\delta)^2 + o(\delta^2) = -\frac{1}{2}A\delta^2 + o(\delta^2).$$

The proof is completed by using (2.9). □

In the next proposition we state the expansion for $\left[\bar{W}^h_\delta(x) - \bar{W}_\delta(x)\right]$. This estimate deals with the critical comparison between the discrete and continuous time problems. The proof of this expansion is detailed, and therefore deferred to the next section.

Proposition 2.3. *Assume Condition 2.1 and define A, C, and K by (2.2) and (2.4). Assume also that $A > 0$. Then*

$$\bar{W}^h_c(x) - \bar{W}_c(x) = \left[Kx_*\sigma(x_*)Ach - \frac{1}{2}ACx_*^2\sigma^2(x_*)h + o(h)\right]V(x).$$

Proof of Theorem 2.1. Recall that x^h_* is the optimal boundary for the stopping problem with value function V^h. On the stopping region, we always have $V^h(x) = \phi(x)$. Also, since $V^h(x)$ is defined by supremizing over a subset of the stopping times allowed in the definition of $V(x)$, it follows that $V^h(x) \leq V(x)$. Since $V(x) \geq \phi(x)$ for all x, it follows that $x^h_* \leq x_*$.

According to Propositions 2.1, 2.2 and 2.3, for each fixed $c \in [0,\infty)$

$$\frac{W^h_c(x) - V(x)}{V(x)} = \left[-\frac{1}{2}Ac^2h + Kx_*\sigma(x_*)Ach - \frac{1}{2}ACx_*^2\sigma^2(x_*)h + o(h)\right].$$

This suggests the choice $c_* \doteq Kx_*\sigma(x_*)$. Inserting this into the last display gives

$$\frac{W^h_{c_*}(x) - V(x)}{V(x)} = \left[\frac{1}{2}A(K^2 - C)x_*^2\sigma(x_*)^2h + o(h)\right],$$

and since $V^h(x) \geq W^h_{c_*}(x)$ it follows that

$$\liminf_{h\downarrow 0} \frac{V^h(x) - V(x)}{V(x)h} \geq \frac{1}{2}A(K^2 - C)x_*^2\sigma(x_*)^2. \tag{2.10}$$

Now define c^h by $x^h_* = x_* - c^h\sqrt{h}$. Since $x^h_* \leq x_*$ we know that $c^h \in [0,\infty)$. By taking a convergent subsequence, we can assume that $c^h \to \bar{c} \in [0,\infty]$. Using an elementary weak convergence argument, one can show that $x^h_* \to x_*$. First assume that $\bar{c} \in (0,\infty)$. If $\bar{c} \neq Kx_*\sigma(x_*)$, then by Propositions 2.1, 2.2 and 2.3 we have

$$\limsup_{h\downarrow 0} \frac{V^h(x) - V(x)}{V(x)h} < \frac{1}{2}A(K^2 - C)x_*^2\sigma(x_*)^2,$$

which contradicts (2.10). If $\bar{c} = \infty$, then Propositions 2.1 and 2.3 and an argument analogous to the one used in Proposition 2.2 shows that

$$\frac{V^h(x) - V(x)}{V(x)} = -A(c^h)^2h[1 + o(1)].$$

Since $(c^h)^2 \to +\infty$, this again contradicts (2.10), and thus $\bar{c} = Kx_*\sigma(x_*)$. We extend to the original sequence by the standard argument by contradiction, and Theorem 2.1 follows. □

3 Approximations and Expansions in Terms of Local Time and the Excursions of a Brownian Motion

In this section we prove Proposition 2.3, which is the expansion $\bar{W}_\delta^h(x) - \bar{W}_\delta(x)$ for small $h > 0$. We will use the fact that $\bar{W}_\delta$ has the representation

$$\bar{W}_\delta(x) = \begin{cases} \bar{\phi}(x) & , \quad \text{for} \quad x \geq x_\delta \\ -\varepsilon(x) + e^{-rh}\mathbb{E}\left[\bar{W}_\delta(S_h) \,\middle|\, S_0 = x\right], & \text{for} \quad x < x_\delta. \end{cases}$$

Here $\varepsilon(x)$ is an error term that can be given explicitly in terms of the value function $\bar{W}_\delta$ and the transition probabilities of S_{nh}, and $x_\delta \doteq x_* - \delta$. From the last display, we have

$$\varepsilon(x) = \mathbb{E}^x\left[e^{-rh}\bar{W}_\delta(S_h) - \bar{W}_\delta(x)\right], \qquad \forall\, x < x_\delta.$$

It follows from the generalized Itô formula that

$$\begin{aligned} e^{-rh}\bar{W}_\delta(S_h) - \bar{W}_\delta(x) &= \int_0^h e^{-rt}[-r\bar{\phi}(S_t) + \mathcal{L}\bar{\phi}(S_t)]1_{\{S_t \geq x_\delta\}}\, dt \\ &+ \triangle \bar{W}_\delta'(x_\delta)\int_0^h e^{-rt} dL_t^S(x_\delta) + \int_0^h e^{-rt}\bar{W}_\delta'(S_t)S_t\sigma(S_t)\, dW_t. \end{aligned} \tag{3.1}$$

Here L^S is the local time for process S, and

$$\triangle \bar{W}_\delta'(x_\delta) \doteq \bar{W}_\delta'(x_\delta+) - \bar{W}_\delta'(x_\delta-).$$

Lemma 3.1. *For every $x \in (0, x_\delta)$,*

$$\begin{aligned} \varepsilon(x) &= \mathbb{E}^x e^{-rh}\bar{W}_\delta(S_h) - \bar{W}_\delta(x) \\ &= \mathbb{E}^x \int_0^h e^{-rt}[-r\bar{\phi}(S_t) + \mathcal{L}\bar{\phi}(S_t)]1_{\{S_t \geq x_\delta\}}\, dt + \mathbb{E}^x \triangle \bar{W}_\delta'(x_\delta)\int_0^h e^{-rt} dL_t^S(x_\delta). \end{aligned}$$

Proof. The result follows from (3.1) if the stochastic integral is zero. We recall that $\bar{W}_\delta(x)$ is equal to $\bar{\phi}(x)$ for $x \geq x_\delta$ and $V(x)\bar{\phi}(x_\delta)/V(x_\delta)$ for $x \leq x_\delta$. Since $\bar{W}_\delta(x)$ is equal to $\bar{\phi}(x)$ for large x and hence constant, $\bar{W}_\delta'(x) = 0$ for all large x. Also, $\bar{W}_\delta'(x)$ is clearly bounded in a neighborhood of x_δ. For $\mu > 0$, let $\sigma_\mu \doteq \inf\{t : S_t \leq \mu\} \wedge h$. Then the boundedness of $x\bar{W}_\delta'(x)$ for $x \geq \mu$ implies

$$\begin{aligned} \mathbb{E}^x e^{-r\sigma_\mu}\bar{W}_\delta(S_{\sigma_\mu}) - \bar{W}_\delta(x) &= \mathbb{E}^x \int_0^{\sigma_\mu} e^{-rt}[-r\bar{\phi}(S_t) + \mathcal{L}\bar{\phi}(S_t)]1_{\{S_t \geq x_\delta\}}\, dt \\ &\quad + \mathbb{E}^x \triangle \bar{W}_\delta'(x_\delta)\int_0^{\sigma_\mu} e^{-rt} dL_t^S(x_\delta). \end{aligned}$$

The lemma follows by letting $\mu \downarrow 0$, and using dominated convergence for all terms but the last, which uses monotone convergence. □

Let $\tau_\delta^h \doteq \inf\{nh : S_{nh} \geq x_\delta\}$. Then the discounting and (3.1) imply the formula

$$\begin{aligned}
\bar{W}_\delta(x) &= -\mathbb{E}^x \sum_{n=0}^{\tau_\delta^h/h-1} e^{-rnh}\varepsilon(S_{nh}) + \mathbb{E}^x e^{-r\tau_\delta^h}\bar{\phi}(S_{\tau_\delta^h}) \\
&= -\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt}[-r\bar{\phi}(S_t) + \mathcal{L}\bar{\phi}(S_t)]1_{\{S_t \geq x_\delta\}}\, dt \\
&\quad - \triangle \bar{W}_\delta'(x_\delta)\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt} dL_t^S(x_\delta) + \mathbb{E}^x e^{-r\tau_\delta^h}\bar{\phi}(S_{\tau_\delta^h})
\end{aligned}$$

for all $x < x_\delta$. We recall the definition

$$\bar{W}_\delta^h(x) \doteq \mathbb{E}^x e^{-r\tau_\delta^h}\bar{\phi}(S_{\tau_\delta^h}).$$

It follows that

$$\begin{aligned}
\bar{W}_\delta^h(x) - \bar{W}_\delta(x) &= \mathbb{E}^x \int_0^{\tau_\delta} e^{-rt}[-r\bar{\phi}(S_t) + \mathcal{L}\bar{\phi}(S_t)]1_{\{S_t \geq x_\delta\}}\, dt \\
&\quad + \triangle \bar{W}_\delta'(x_\delta)\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt} dL_t^S(x_\delta).
\end{aligned}$$

The proof of Proposition 2.3 is thereby reduced to proving the following two results.

Proposition 3.1. *Assume Condition 2.1 and define A and C by (2.2) and (2.4). Assume also that $A > 0$. Then*

$$\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt}[-r\bar{\phi}(S_t) + \mathcal{L}\bar{\phi}(S_t)]1_{\{S_t \geq x_\delta\}}\, dt = \left[-\frac{1}{2}ACx_*^2\sigma^2(x_*)h + o(h)\right]V(x). \tag{3.2}$$

Proposition 3.2. *Assume Condition 2.1 and define A and K by (2.2) and (2.4). Assume also that $A > 0$. Then*

$$\triangle \bar{W}_\delta'(x_\delta)\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt} dL_t^S(x_\delta) = [Kx_*\sigma(x_*)Ach + o(h)]\,V(x). \tag{3.3}$$

The proofs of Propositions 3.1 and 3.2 use estimates on the excursions and local time of Brownian motion, respectively, and are given in the next two subsections. We will need to relate the constants that appear in these approximations to the simple constants defined by (2.3) and (2.4). The following lemma gives this relationship.

Lemma 3.2. *Let W be a standard Brownian motion that satisfies $W_u = 0$ for some $u \in [0,1)$. Define $H(u)$ and $M(u)$ by (2.3). Then*

$$H(u) = \mathbb{E}\int_u^N 1_{\{W_t \geq 0\}}\, dt \qquad \text{and} \qquad M(u) = \mathbb{E}L_{u,N}^W(0),$$

where $N \doteq \inf\{n \in \mathbb{N} : \ W_n \geq 0\}$ and $L_{u,N}^W(0)$ is the local time of W on the interval $[u, N]$.

Proof. We first prove $H(u) = \mathbb{E}\int_u^N 1_{\{W_t \geq 0\}}\, dt$. Consider the continuously differentiable convex function

$$f(x) \doteq \frac{1}{2}(x^+)^2 = \begin{cases} 0\ , & \text{if } x \leq 0 \\ \frac{1}{2}x^2, & \text{if } x \geq 0. \end{cases}$$

It follows from Itô's formula that

$$\begin{aligned} f(W_{N\wedge n}) &= f(W_u) + \int_u^{N\wedge n} f'(W_t)\, dW_t + \frac{1}{2}\int_u^{N\wedge n} f''(W_t)\, dt \\ &= \int_u^{N\wedge n} W_t 1_{\{W_t\geq 0\}}\, dW_t + \frac{1}{2}\int_u^{N\wedge n} 1_{\{W_t \geq 0\}}\, dt \end{aligned}$$

for all integers $n \in \mathbb{N}$. This yields

$$\mathbb{E}\left(W^+_{N\wedge n}\right)^2 = \mathbb{E}\int_u^{N\wedge n} 1_{\{W_t\geq 0\}}\, dt \qquad \forall\, n \in \mathbb{N}.$$

Letting $n \to \infty$, the right hand side converges to $\mathbb{E}\int_u^N 1_{\{W_t\geq 0\}}\, dt$ by the monotone convergence theorem. Since $W^+_{N\wedge n} \leq W_N$, the result will follow by dominated convergence if $\mathbb{E}W_N^2$ is finite.

To show $\mathbb{E}W_N^2$ is finite, we consider the conditional probability

$$P_{n+1}(x) \doteq \mathbb{P}\left(W_N \geq x \,\middle|\, N = n+1\right) = \mathbb{P}\left(W_{n+1} \geq x \,\middle|\, W_{n+1} \geq 0, W_n < 0, \cdots, W_1 < 0\right)$$

for all $n \in \mathbb{N}_0$ and $x \geq 0$. Define the following stopping time

$$\sigma \doteq \inf\left\{t \geq n :\ W_t = 0\right\}.$$

Then

$$\begin{aligned} P_{n+1}(x) &= \mathbb{P}\left(W_{n+1} \geq x \,\middle|\, \sigma \leq n+1, W_{n+1} \geq 0, W_n < 0, \cdots, W_1 < 0\right) \\ &= \int_0^1 \mathbb{P}\left(W_{n+1} \geq x \,\middle|\, \sigma = n+t, W_{n+1} \geq 0, W_n < 0, \cdots, W_1 < 0\right) \\ &\qquad \cdot \mathbb{P}\left(\sigma \in n+dt \,\middle|\, \sigma \leq n+1, W_{n+1} \geq 0, W_n < 0, \cdots, W_1 < 0\right). \end{aligned}$$

However, by strong Markov Property, for all $t \in [0,1]$

$$\begin{aligned} &\mathbb{P}\left(W_{n+1} \geq x \,\middle|\, \sigma = n+t, W_{n+1} \geq 0, W_n < 0, \cdots, W_1 < 0\right) \\ &\quad = \mathbb{P}\left(W_1 \geq x \,\middle|\, W_t = 0, W_1 \geq 0\right) \\ &\quad = 2\Phi\left(-\frac{x}{\sqrt{1-t}}\right) \\ &\quad \leq 2\Phi(-x). \end{aligned}$$

Here Φ is the cumulative distribution function for the standard normal. Hence,

$$\begin{aligned} P_{n+1}(x) &\leq 2\Phi(-x)\int_0^1 \mathbb{P}\left(\sigma \in n+dt \,\middle|\, \sigma \leq n+1, W_{n+1} \geq 0, W_n < 0, \cdots, W_1 < 0\right) \\ &= 2\Phi(-x). \end{aligned}$$

It follows that

$$\begin{aligned}
\mathbb{E}W_N^2 &= \sum_{n=1}^{\infty} \mathbb{E}(W_N^2 1_{\{N=n\}}) \\
&= \sum_{n=1}^{\infty} \int_0^{\infty} 2x\mathbb{P}(W_N \geq x, N=n)\, dx \\
&= \sum_{n=1}^{\infty} \int_0^{\infty} 2x\mathbb{P}(W_N \geq x \,|\, N=n)\mathbb{P}(N=n)\, dx \\
&\leq \sum_{n=1}^{\infty} \mathbb{P}(N=n) \int_0^{\infty} 4x\Phi(-x)\, dx \\
&= \int_0^{\infty} 4x\Phi(-x)\, dx \\
&< \infty.
\end{aligned}$$

As for the equality $M(u) = \mathbb{E}L_{u,N}^W(0)$, it follows from Tanaka's formula that

$$W_N = W_N^+ = \int_u^N 1_{\{W_t \geq 0\}}\, dW_t + L_{u,N}^W(0).$$

However, since the preceding proof already implies that $\mathbb{E}\int_u^N 1_{\{W_t \geq 0\}}\, dt < \infty$,

$$\mathbb{E}W_N = \mathbb{E}L_{u,N}^W(0) = M(u).$$

This completes the proof. □

3.1 Proof of Proposition 3.1

In this subsection we prove

$$\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt}[-r\bar{\phi}(S_t) + \mathcal{L}\bar{\phi}(S_t)]1_{\{S_t \geq x_\delta\}}\, dt = \left[-\frac{1}{2}ACx_*^2\sigma^2(x_*)h + o(h) \right] V(x).$$

We recall the definition

$$H(u) \doteq \mathbb{E}\int_u^N 1_{\{W_t \geq 0\}}\, dt,$$

where W is a standard Brownian motion with $W_u = 0$, and $N \doteq \inf\{n \in \mathbb{N} : W_n \geq 0\}$.

Lemma 3.3. *$H(u)$ is continuous and bounded on the interval* $[0,1)$.

Proof. Define

$$Z_u \doteq \int_u^N 1_{\{W_t \geq 0\}}\, dt$$

where W is a Brownian motion with $W_u = 0$. We first show that the family $\{Z_u, u \in [0,1)\}$ is uniformly integrable (and in particular, that $H(u)$ is bounded). Indeed, define

$$c_0 \doteq \int_u^1 1_{\{W_t \geq 0\}}\, dt, \qquad c_j \doteq \int_j^{j+1} 1_{\{W_t \geq 0\}}\, dt, j \in \mathbb{N}.$$

The key observation is that if $c_j > 0$, then W must spend some time during the interval $[j, j+1]$ to the right of zero, therefore the probability that $W_{j+1} > 0$ is at least half. Thus for all $j \in \mathbb{N}_0$

$$\mathbb{P}(N = j+1 \,|\, N > j, c_j > 0) \geq \frac{1}{2}.$$

Let $X_u \doteq \sum_{j=0}^{N-1} 1_{\{c_j > 0\}}$. Clearly X_u dominates Z_u. Furthermore, the strong Markov property implies that

$$\mathbb{P}(X_u \geq j+1 \,|\, X_u \geq j) \leq \left(1 - \frac{1}{2}\right) = \frac{1}{2}.$$

This, in turn, implies that

$$\mathbb{P}(X_u \geq n) \leq \frac{1}{2^{n-1}},$$

and thus

$$\mathbb{E}(X_u^2) = \sum_{n=1}^{\infty} 2n\mathbb{P}(X_u \geq n) \leq \sum_{n=1}^{\infty} \frac{n}{2^{n-2}} < \infty.$$

Therefore $\{Z_u, u \in [0,1)\}$ is uniformly integrable.

As for the continuity, we write

$$H(u) = \mathbb{E} \int_u^{N_u} 1_{\{B_t - B_u \geq 0\}} \, dt = \mathbb{E} Z_u.$$

where B is some standard Brownian motion with $B_0 \equiv 0$ and

$$N_u \doteq \inf \{n \in \mathbb{N}_0 : \; B_n - B_u \geq 0\}.$$

Let $u \in [0,1)$ and let $\{u_n\}$ be an arbitrary sequence in $[0,1)$ with $u_n \to u$. Since for any fixed n $\mathbb{P}(B_n - B_u = 0) = 0$,

$$Z_{u_n} \to Z_u$$

with probability one. Since the Z_{u_n} are uniformly integrable, we have

$$H(u_n) \to H(u),$$

which completes the proof. □

Now for any $u \in [0,1)$ and $h > 0$, define the function

$$G(h; u) \doteq \mathbb{E} \int_{uh}^{N^h h} e^{-r(t-uh)} \left[-r\bar{\phi}(S_t) + \mathcal{L}\bar{\phi}(S_t)\right] 1_{\{S_t \geq x_\delta\}} \, dt,$$

where

$$\frac{dS_t}{S_t} = b(S_t)\, dt + \sigma(S_t)\, dW_t, \qquad S_{uh} = 0$$

and

$$N^h \doteq \inf \{n \in \mathbb{N} : \; S_{nh} \geq x_\delta\}.$$

Let $\lfloor a \rfloor$ denote the integer part of a. It follows from strong Markov property that

$$\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt}[-r\bar{\phi}(S_t) + \mathcal{L}\bar{\phi}(S_t)] 1_{\{S_t \geq x_\delta\}}\, dt = \mathbb{E}^x \left[e^{-r\tau_\delta} G\left(h; \frac{\tau_\delta}{h} - \left\lfloor \frac{\tau_\delta}{h} \right\rfloor \right)\right].$$

The change of variable $t \mapsto th$ and the transformation

$$Y_t^{(h)} \doteq \frac{S_{th} - x_\delta}{\sqrt{h}}$$

yield

$$G(h;u) = hF(h;u) \doteq h\mathbb{E} \int_u^{N^h} e^{-r(t-u)h} \big[-r\bar{\phi} + \mathcal{L}\bar{\phi}\big](\sqrt{h} Y_t^{(h)} + x_\delta) 1_{\{Y_t^{(h)} \geq 0\}}\, dt,$$

where $Y^{(h)}$ follows the dynamics

$$dY_t^{(h)} = (\sqrt{h} Y_t^{(h)} + x_\delta) \left[\sqrt{h} b(\sqrt{h} Y_t^{(h)} + x_\delta) + \sigma(\sqrt{h} Y_t^{(h)} + x_\delta)\, dW_t \right], \qquad Y_u^{(h)} = 0.$$

Therefore

$$\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt}[-r\bar{\phi}(S_t) + \mathcal{L}\bar{\phi}(S_t)] 1_{\{S_t \geq x_\delta\}}\, dt = h\mathbb{E}^x \left[e^{-r\tau_\delta} F\left(h; \frac{\tau_\delta}{h} - \left\lfloor \frac{\tau_\delta}{h} \right\rfloor \right)\right]. \tag{3.4}$$

We have the following result regarding $F(h;u)$. Although part of the proof is similar to that of Lemma 3.3, we provide the details for completeness.

Lemma 3.4. *1. $F(h;u)$ is uniformly bounded for small h and all $u \in [0,1)$.*

2.

$$\lim_{h \to 0} F(h;u) = [-r\bar{\phi} + \mathcal{L}\bar{\phi}](x_*) H(u),$$

and the convergence is uniform on any compact subset of $[0,1)$.

Proof. Consider the family of random variables $\{Z_{h,u} :\ u \in [0,1), h \in (0,1)\}$ where

$$Z_{h,u} \doteq \int_u^{N^h} e^{-r(t-u)h} \big[-r\bar{\phi} + \mathcal{L}\bar{\phi}\big](\sqrt{h} Y_t^{(h)} + x_\delta) 1_{\{Y_t^{(h)} \geq 0\}}\, dt$$

and

$$dY_t^{(h)} = (\sqrt{h} Y_t^{(h)} + x_\delta) \left[\sqrt{h} b(\sqrt{h} Y_t^{(h)} + x_\delta) + \sigma(\sqrt{h} Y_t^{(h)} + x_\delta)\, dW_t \right], \qquad Y_u^{(h)} = 0.$$

We first show this family is uniformly integrable. Since $(-r\bar{\phi} + \mathcal{L}\bar{\phi})$ is bounded, it is sufficient to show that

$$X_{h,u} \doteq \int_u^{N^h} 1_{\{Y_t^{(h)} \geq 0\}}\, dt \tag{3.5}$$

are uniformly integrable. Define

$$c_0^{(h)} \doteq \int_u^1 1_{\{Y_t^{(h)} \geq 0\}}\, dt, \qquad c_j^{(h)} \doteq \int_j^{j+1} 1_{\{Y_t^{(h)} \geq 0\}}\, dt, j \in \mathbb{N}$$

As in the proof of Lemma 3.3, if $c_j^{(h)} > 0$, then $Y_t^{(h)}$ spends some time to the right of zero in interval $[j, j+1]$. Therefore the probability $Y_{j+1}^{(h)} \geq 0$ is bounded from below by a positive constant:

$$\mathbb{P}\left(N^h = j+1 \,\middle|\, N^h > j, c_j^{(h)} > 0\right) \geq \alpha > 0, \qquad \forall\, u \in [0,1),\ h \in (0,1).$$

A proof is as follows. To show a lower bound $\alpha > 0$ exists, it suffices to show that for some $\alpha > 0$

$$p_{t,h} \doteq \mathbb{P}(Y_t^{(h)} \geq 0 \,|\, Y_0^{(h)} \geq 0) \geq \alpha > 0, \qquad \forall\, t \in [0,1].$$

However, it is easy to see that

$$\begin{aligned} p_{t,h} &= \mathbb{P}(S_{th} \geq x_\delta \,|\, S_0 \geq x_\delta) \\ &\geq \mathbb{P}\left(\exp\left\{\int_0^{th} \left[b(S_u) - \frac{1}{2}\sigma^2(S_u)\right] du + \int_0^{th} \sigma(S_u)\, dW_u\right\} \geq 1\right) \\ &\geq \mathbb{P}\left(\int_0^{th} \sigma(S_u)\, dW_u \geq c_1 th\right), \end{aligned}$$

where $c_1 \doteq \|b\|_\infty + \frac{1}{2}\|\sigma^2\|_\infty$. We can view the stochastic integral $Q_t \doteq \int_0^t \sigma(S_u)\, dW_u$ a time-changed Brownian motion. Indeed, there exists a Brownian motion B such that

$$Q_t = B_{\langle Q \rangle_t}.$$

Let

$$\underline{\sigma} \doteq \inf_x \sigma(x) \qquad \bar{\sigma} \doteq \sup_x \sigma(x).$$

Then

$$\underline{\sigma}^2 h \leq \langle Q \rangle_t \leq \bar{\sigma}^2 h.$$

It follows that

$$p_{t,h} \geq \mathbb{P}\left(Q_{th} \geq c_1 th\right) \geq \mathbb{P}\left(\min_{\underline{\sigma}^2 th \leq s \leq \bar{\sigma}^2 th} B_s \geq c_1 th\right) = \mathbb{P}\left(\min_{\underline{\sigma}^2 \leq s \leq \bar{\sigma}^2} B_s \geq c_1 \sqrt{th}\right),$$

where the last equality follows since $\left\{\frac{1}{\sqrt{th}} B_{ths},\ s \geq 0\right\}$ is still a standard Brownian motion. For $h \in (0,1)$, we can choose

$$\alpha = \mathbb{P}\left(\min_{\underline{\sigma}^2 \leq s \leq \bar{\sigma}^2} B_s \geq c_1\right) > 0,$$

which will serve as a lower bound.

Now define

$$M_{h,u} \doteq \sum_{j=0}^{N^h-1} 1_{\{c_j^{(h)}>0\}},$$

which clearly dominates $X_{h,u}$. By the strong Markov property,

$$\mathbb{P}(M_{h,u} > j+1 \,\big|\, M_{h,u} > j) \leq 1-\alpha,$$

and thus

$$\mathbb{P}(M_{h,u} \geq j) \leq (1-\alpha)^{j-1}.$$

This implies that

$$\mathbb{E}(M_{h,u}^2) = \sum_{j=0}^{\infty} 2j\mathbb{P}(M_{h,u} \geq j) \leq \sum_{j=0}^{\infty} 2j(1-\alpha)^{j-1} < \infty,$$

which implies the uniform integrability of $\{Z_{h,u},\ u \in [0,1),\ h \in (0,1)\}$. In particular, $F(h;u)$ is uniformly bounded for $u \in [0,1)$ and $h \in (0,1)$.

For the uniform convergence, it suffices to show that for any $u \in [0,1)$ and any sequence $u^h \in [0,1)$ converging to u,

$$F(h;u^h) = \mathbb{E}Z_{h,u^h} \ \to\ [-r\bar{\phi} + \mathcal{L}\bar{\phi}](x_*)H(u).$$

Let $Y^{(h)}$ be the process with $Y^{(h)}_{u^h} = 0$. As $h \to 0$, we have that $Y^{(h)}$ converges weakly to Y, where Y is defined as

$$Y_t = x_*\sigma(x_*)(W_t - W_u).$$

By the Skorohod representation, we can assume $Y^{(h)} \to Y$ with probability one. Using the uniform integrability, it suffices to show that

$$Z_{h,u^h} \to Z \doteq \int_u^N 1_{\{Y_t \geq 0\}}\, dt$$

with probability one. Note N is almost surely finite, and that

$$N^h \to N$$

with probability one. The almost sure convergence of Z_{h,u^h} to Z then follows from the dominated convergence theorem, which completes the proof. □

Returning to the proof of Proposition 3.1, we claim that

$$\lim_{h\to 0} \mathbb{E}^x \left[e^{-r\tau_\delta} F\left(h; \frac{\tau_\delta}{h} - \left\lfloor \frac{\tau_\delta}{h} \right\rfloor\right)\right] = C[-r\bar{\phi} + \mathcal{L}\bar{\phi}](x_*)\mathbb{E}^x\left[e^{-r\tau_*}\right] \tag{3.6}$$

where $C \doteq \int_0^1 H(u)\, du$. To ease notation, let

$$U_h \doteq \frac{\tau_\delta}{h} - \left\lfloor \frac{\tau_\delta}{h} \right\rfloor.$$

It suffices to show that

$$\lim_{h\to 0} \mathbb{E}^x\left[e^{-r\tau_\delta} F(h;U_h)\right] = C[-r\bar{\phi} + \mathcal{L}\bar{\phi}](x_*)\mathbb{E}^x\left[e^{-r\tau_*}\right].$$

We can write

$$\begin{aligned}\mathbb{E}^x\left[e^{-r\tau_\delta} F(h;U_h)\right] &= \mathbb{E}^x\left[e^{-r\tau_\delta} F(h;U_h)\right] - [-r\bar{\phi} + \mathcal{L}\bar{\phi}](x_*)\mathbb{E}^x\left[e^{-r\tau_\delta} H(U_h)\right] \\ &\quad + [-r\bar{\phi} + \mathcal{L}\bar{\phi}](x_*)\mathbb{E}^x\left[e^{-r\tau_\delta} H(U_h)\right].\end{aligned}$$

In Proposition 4.1 in the Appendix we show, roughly speaking, that (τ_δ, U_h) converges in distribution to (τ_*, U) as h and δ tend to zero, where U is uniformly distributed and independent of τ_*. This is not strictly true, in that we ignore what happens on the unimportant event $\tau_* = \infty$. It follows from Proposition 4.1 that

$$\mathbb{E}^x\left[e^{-r\tau_\delta} H(U_h)\right] \;\to\; \mathbb{E}^x\left[e^{-r\tau_*}\right]\int_0^1 H(u)\,du = C\mathbb{E}^x\left[e^{-r\tau_*}\right].$$

Therefore, to prove (3.6) we must show that

$$\triangle \doteq \mathbb{E}^x\left[e^{-r\tau_\delta} F(h;U_h)\right] - [-r\bar{\phi} + \mathcal{L}\bar{\phi}](x_*)\mathbb{E}^x\left[e^{-r\tau_\delta} H(U_h)\right] \;\to\; 0.$$

Due to the uniform boundedness of F and H, there exists $R \in (0,\infty)$ such that

$$\left|F(h,u)\right| + \left|[-r\bar{\phi} + \mathcal{L}\bar{\phi}](x_*)H(u)\right| \le R \qquad \forall\, u \in [0,1)$$

when h is small enough. Since $U_h \Rightarrow U$, for h small enough,

$$\mathbb{P}\left(U_h > 1-\varepsilon\right) \le 2\varepsilon.$$

Also, by Lemma 3.4 for h small enough

$$\sup_{u\in[0,1-\varepsilon]} \left|F(h,u) - [-r\bar{\phi} + \mathcal{L}\bar{\phi}](x_*)H(u)\right| \le \varepsilon.$$

It follows that, for h small enough,

$$\triangle \le \varepsilon\mathbb{P}(U_h \le 1-\varepsilon) + R\mathbb{P}(U_h > 1-\varepsilon) \le (2R+1)\varepsilon,$$

which completes the proof of (3.6).

It follows directly from the definitions of $V(x)$ and τ_* that

$$\mathbb{E}^x\left[e^{-r\tau_*}\right] = V(x)/\bar{\phi}(x_*). \tag{3.7}$$

Also, the definition of A in (2.2) and the fact that $(-rV + \mathcal{L}V)(x_*-) = 0$ imply that

$$\begin{aligned}(-r\bar{\phi} + \mathcal{L}\bar{\phi})(x_*) &= (-rV + \mathcal{L}V)(x_*-) + \frac{1}{2}\sigma^2(x_*)x_*^2\left[\bar{\phi}''(x_*) - V''(x_*-)\right] \\ &= \frac{1}{2}\sigma^2(x_*)x_*^2\left[\bar{\phi}''(x_*) - V''(x_*-)\right] \\ &= \frac{1}{2}\sigma^2(x_*)x_*^2 A\bar{\phi}(x_*).\end{aligned}$$

The proposition follows by combining the last display with (3.4), (3.6), and (3.7).

3.2 Proof of Proposition 3.2

In this subsection we verify equation (3.3):

$$\triangle \bar{W}'_\delta(x_\delta)\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt} dL_t^S(x_\delta) = [Kx_*\sigma(x_*)Ach + o(h)]\, V(x).$$

We recall the notation $x_\delta = x_* - \delta$, where $\delta = c\sqrt{h} + o(\sqrt{h})$. It follows from the definition (2.2) of A that

$$\begin{aligned}
\triangle \bar{W}'_\delta(x_\delta) &= \bar{\phi}'(x_\delta) - \frac{\bar{\phi}(x_\delta)}{V(x_\delta)} V'(x_\delta) \\
&= \left(\bar{\phi}' - \frac{\bar{\phi}}{V} V'\right)'\Bigg|_{x_*} \cdot (-\delta) + o(\delta) \\
&= [V''(x_*-) - \bar{\phi}''(x_*)]\delta + o(\delta) \\
&= cA\phi(x_*)\sqrt{h} + o(\sqrt{h}).
\end{aligned}$$

As a consequence, the main difficulty in proving (3.3) lies with the term

$$\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt} dL_t^S(x_\delta).$$

As in the previous subsection, we consider the transformation

$$Y_t^{(h)} \doteq \frac{S_{th} - x_\delta}{\sqrt{h}}.$$

Then $Y^{(h)}$ satisfies the SDE

$$dY_t^{(h)} = (\sqrt{h}Y_t^{(h)} + x_\delta)\left[\sqrt{h}b(\sqrt{h}Y_t^{(h)} + x_\delta) + \sigma(\sqrt{h}Y_t^{(h)} + x_\delta)\, dW_t\right].$$

We have the following lemma, whose proof is trivial from the definition of the local time and thus omitted.

Lemma 3.5. *Suppose X is a semimartingale, and $Y_t \doteq aX_{bt} + v$ where $a > 0, b > 0, v$ are arbitrary constants. Let L^Y and L^X denote the local times for Y and X, respectively. Then for all $t \geq 0$*

$$L_t^Y(ax + v) = aL_{bt}^X(x).$$

It follows from the lemma that

$$\mathbb{E}^x \int_0^{\tau_\delta^h} e^{-rt} dL_t^S(x_\delta) = \sqrt{h}\mathbb{E}^x \int_0^{N^h} e^{-rth} dL_t^{Y^{(h)}}(0).$$

For any $u \in [0, 1)$, define the process

$$Y_t^* = x_*\sigma(x_*)W_t, \qquad Y_u^* = 0.$$

Also define

$$Q(u) \doteq \mathbb{E}L_{u,N}^{Y^*}(0), \quad \text{where} \quad N \doteq \inf\{n \in \mathbb{N} : \; Y_n^* \geq 0\}$$

We have the following result.

Proposition 3.3.

$$\lim_{h\to 0}\mathbb{E}^x\int_0^{N^h} e^{-rth}dL_t^{Y^{(h)}}(0)=\mathbb{E}^x\left[e^{-r\tau_*}\right]\int_0^1 Q(u)\,du$$

Before giving the proof, we show how the proposition will follow from Proposition 3.3. We have $\mathbb{E}^x e^{-r\tau_*}=V(x)/\bar{\phi}(x_*)$, and the definitions of Q and M imply $\int_0^1 Q(u)du=(x_*)^2\sigma(x_*)^2\int_0^1 M(u)du$. When combined with the expansion given above for $\Delta W_\delta'(x_\delta)$, the left hand side of (3.3) is equal to

$$cA\bar{\phi}(x_*)\frac{V(x)}{\bar{\phi}(x_*)}(x_*)^2\sigma(x_*)^2\int_0^1 M(u)du+o(h),$$

and thus (3.3) follows from Lemma 3.2.

Proof of Proposition 3.3. We consider the test function

$$f(x)\doteq\begin{cases}0, & \text{if } x\le 0\\ x, & \text{if } 0\le x\le 1\\ k, & \text{if } x\ge 2.\end{cases}$$

We require $f(x)$ to be increasing, and smooth except at the point $x=0$ (the specific choice of k is not important). It follows from the generalized Itô formula and the integration by parts formula that

$$\begin{aligned} d\left[e^{-rth}f(Y_t^{(h)})\right] &= -rhe^{-rth}f(Y_t^{(h)})\,dt+e^{-rth}D^-f(Y_t^{(h)})\,dY_t^{(h)}\\ &\quad+\frac{1}{2}e^{-rth}f''(Y_t^{(h)})dY_t^{(h)}\cdot dY_t^{(h)}+e^{-rth}dL_t^{X^{(h)}}(0).\end{aligned}$$

Without loss of generality, we let $f''(0)=0$. Now we integrate both sides from 0 to N^h and take expected value. The stochastic integral on the right hand side

$$\int_0^{N^h} e^{-rth}D^-f(Y_t^{(h)})(\sqrt{h}Y_t^{(h)}+x_\delta)\sigma(\sqrt{h}Y_t^{(h)}+x_\delta)\,dW_t$$

has expectation 0 since

$$D^-f(Y_t^{(h)})(\sqrt{h}Y_t^{(h)}+x_\delta)$$

is bounded (note that $D^-f(x)=0$ for $x\ge 2$) and σ is bounded by assumption.

The first term on the right hand side will contribute

$$-rh\mathbb{E}^x\int_0^{N^h} e^{-rth}f(Y_t^{(h)})\,dt=-rh\mathbb{E}^x\int_{\frac{\tau_\delta}{h}}^{N^h} e^{-rth}f(Y_t^{(h)})\,dt$$

since $f(x)=0$ for $x\le 0$. We recall the definition (3.5) of $X_{h,u}$. It follows from the strong Markov property that

$$\mathbb{E}^x\int_{\frac{\tau_\delta}{h}}^{N^h} e^{-rth}f(Y_t^{(h)})\,dt\le k\mathbb{E}^x\int_{\frac{\tau_\delta}{h}}^{N^h} 1_{\{Y_t^{(h)}\ge 0\}}\,dt=k\mathbb{E}^xG(h,U_h)$$

where

$$U_h \doteq \frac{\tau_\delta}{h} - \left\lfloor \frac{\tau_\delta}{h} \right\rfloor$$

and

$$G(h,u) \doteq \mathbb{E}X_{h,u}.$$

By the uniform integrability of $X_{h,u}$ for small h and $u \in [0,1)$, $\mathbb{E}^x G(h,U_h)$ is uniformly bounded for small h. Therefore the expectation of the first term in the right hand side goes to zero as $h \to 0$. The second term in the right hand side contributes

$$\sqrt{h}\mathbb{E}^x \int_0^{Nh} e^{-rth} D^- f(Y_t^{(h)})(\sqrt{h}Y_t^{(h)} + x_\delta) b(\sqrt{h}Y_t^{(h)} + x_\delta)\, dt.$$

Note that the integrand is bounded by $1_{\{Y_t^{(h)} \geq 0\}}$ up to a proportional constant. It follows exactly as in the case of the first term that the contribution of the second term goes to zero. The third term in the right hand side contributes

$$\mathbb{E}^x \int_0^{Nh} \frac{1}{2} e^{-rth} f''(Y_t^{(h)})(\sqrt{h}Y_t^{(h)} + x_\delta)^2 \sigma^2(\sqrt{h}Y_t^{(h)} + x_\delta)\, dt.$$

Since $f''(x) = 0$ for $x < 0$, expected value equals

$$\mathbb{E}^x \int_{\frac{\tau_\delta}{h}}^{Nh} \frac{1}{2} e^{-rth} f''(Y_t^{(h)})(\sqrt{h}Y_t^{(h)} + x_\delta)^2 \sigma^2(\sqrt{h}Y_t^{(h)} + x_\delta)\, dt.$$

It follows from strong Markov property that the expectation can also be written

$$\mathbb{E}^x \left[e^{-r\tau_\delta} F(h; U_h) \right],$$

where

$$F(h;u) \doteq \mathbb{E} \int_u^{Nh} \frac{1}{2} e^{-r(t-u)h} f''(Y_t^{(h)})(\sqrt{h}Y_t^{(h)} + x_\delta)^2 \sigma^2(\sqrt{h}Y_t^{(h)} + x_\delta) 1_{\{Y_t^{(h)} \geq 0\}}\, dt$$

and where $Y^{(h)}$ satisfies the same dynamics with $Y_u^{(h)} = 0$. Since the integrand is bounded due to the fact that $f''(x) = 0$ for all $x \geq 2$, it follows from an analogous argument to the one given in the proof of Lemma 3.4 that:

1. $F(h;u)$ is uniformly bounded for small h and all $u \in [0,1)$;
2. $$J(u) \doteq \lim_{h \to 0} F(h;u) = \frac{1}{2}\mathbb{E} \int_u^N f''(Y_t^*) x_*^2 \sigma^2(x_*)\, dt$$

   and the convergence is uniform on any compact subset of $[0,1)$.

The uniform convergence (on compact sets) of F and Proposition 4.1 in the Appendix imply that the expectation of the third term converges to

$$\mathbb{E}^x \left[e^{-r\tau_*} \right] \int_0^1 J(u)\, du.$$

We omit the details here since an analogous argument is used in the proof of Proposition 3.1.

It remains to calculate the contribution from the term

$$\mathbb{E}^x\left[e^{-r\tau_\delta^h}f(Y_{N^h}^{(h)})\right] = \mathbb{E}^x\left[e^{-r\tau_\delta}K(h,U_h)\right]$$

where

$$K(h;u) \doteq \mathbb{E}\left[e^{-r(N^h-u)h}f(Y_{N^h}^{(h)})\right]$$

with $Y_u^{(h)} = 0$. However, the boundedness and continuity of f ensure that

1. $K(h;u)$ is uniformly bounded for all h and all $u \in [0,1)$.

2. $$I(u) \doteq \lim_{h\to 0} K(h;u) = \mathbb{E}[f(Y_N^*)|Y_u^* = 0]$$

   and the convergence is uniform on any compact subset of $[0,1)$.

Indeed, the first claim is trivial. As for the second claim, let $u^h \to u$. Then as $h \to 0$, $Y^{(h)} \Rightarrow Y^*$. By the Skorohod representation theorem we can assume $Y^{(h)} \to Y^*$ with probability one, which also implies that $N^h \to N$ with probability one. Therefore,

$$Y_{N^h}^{(h)} \quad\to\quad Y_N^*$$

with probability one. The claim now follows from the dominated convergence theorem. Hence similarly we have

$$\mathbb{E}^x\left[e^{-r\tau_\delta^h}f(Y_{N^h}^{(h)})\right] \quad\to\quad \mathbb{E}^x\left[e^{-r\tau_*}\right]\int_0^1 I(u)\,du,$$

as $h \to 0$. It is now sufficient to prove

$$I(u) - J(u) = Q(u) \qquad \forall\ u \in [0,1).$$

This is the same showing

$$\mathbb{E}\left[f(Y_N^*) - \frac{1}{2}\int_u^N f''(Y_t^*)x_*^2\sigma^2(x_*)\,dt - L_{u,N}^{Y^*}(0)\right] = 0,$$

where

$$Y_t^* = x_*\sigma(x_*)W_t, \qquad Y_u^* = 0.$$

To prove this, we apply Itô's formula to $f(Y^*)$ to obtain

$$f(Y_N^*) - f(Y_u^*) = \int_u^N D^- f(Y_t^*)x_*\sigma(x_*)\,dW_t + \frac{1}{2}\int_u^N f''(Y_t^*)x_*^2\sigma^2(x_*)\,dt + L_{u,N}^{Y^*}(0).$$

But $f(Y_u) = f(0) = 0$. Furthermore the integrand of the stochastic integral is dominated by $1_{\{Y_t^*\geq 0\}}$ up to a proportional constant, which implies that the stochastic integral has expectation 0. This completes the proof. □

4 Appendix: Weak convergence of (τ_δ, U_h)

For an arbitrary $y > 0$, define the function

$$P^y(x,t) \doteq \mathbb{P}\left(\max_{0\le u\le t} S_u \ge y \,\middle|\, S_0 = x \right).$$

We have the following lemmata.

Lemma 4.1. *For every fixed $y > 0$, function $P^y \in \mathcal{C}^{1,2}\big((0,y)\times(0,\infty)\big) \cap \mathcal{C}\big((0,y)\times[0,\infty)\big)$ and satisfies the parabolic equation*

$$-\frac{\partial P^y}{\partial t}(x,t) + \mathcal{L}P^y(x,t) = 0 \qquad (x,t) \in (0,y)\times(0,\infty).$$

Proof: It follows from a standard weak convergence argument that P^y is a continuous function; see, e.g., [13]. Let $(x_0,t_0) \in (0,y)\times(0,\infty)$ and define the region

$$D \doteq (x_0-\varepsilon, x_0+\varepsilon)\times(t_0-\varepsilon, t_0).$$

Consider the parabolic equation

$$-\frac{\partial u}{\partial t}(x,t) + \mathcal{L}u(x,t) = 0 \quad (x,t)\in D,$$

with boundary condition $u = P^y$ on its parabolic boundary. It follows from standard PDE theory that there exists a classical solution u [8]. It remains to show that $u = P^y$ in the domain D. Define the stopping time

$$\tau \doteq \inf\left\{t \ge 0: \ (t_0 - t, S_t) \notin D\right\}.$$

It follows that the process $u(S_t, t_0 - t)$ is a (bounded) martingale. In particular,

$$u(x_0,t_0) = \mathbb{E}^{x_0} u(S_\tau, t_0-\tau) = \mathbb{E}^{x_0} P^y(S_\tau, t_0-\tau) = P^y(x_0,t_0).$$

Here the last equality follows from strong Markov property. □

For fixed $0 < x < y$, the "density" of the hitting time τ_y is defined as

$$p^y(x,t) \doteq \frac{\partial P^y}{\partial t}(x,t).$$

According to the preceding lemma, p^y is continuous in the domain $(0,y)\times(0,\infty)$.

Lemma 4.2. *Suppose $y_n \to y_*$, then $P^{y_n}(x,t) \to P^{y_*}(x,t)$ and $p^{y_n}(x,t) \to p^{y^*}(x,t)$ uniformly on any compact subset of $(0,y_*)\times(0,\infty)$.*

Proof: It suffices to show that $P^{y_n}(x,t) \to P^{y^*}(x,t)$ uniformly on any compact subset. The uniform convergence of p^{y_n} then follows from Friedman [8, Section 3.6]. Suppose $D \doteq [x_0,x_1]\times[t_0,t_1] \subseteq (0,y_*)\times(0,\infty)$ is a compact subset. In the following, we will denote

P^{y_n} and P^{y_*} by P_n and P respectively. Also, without loss of generality, we assume $y_n \le y_*$ for all n, which implies that

$$P_n(x,t) \ge P(x,t).$$

For any $\varepsilon > 0$, we want to show that for large enough n,

$$0 \le P_n(x,t) - P(x,t) \le \varepsilon \qquad \forall\ (x,t) \in D.$$

Define

$$\tau \doteq \inf\{t \ge 0 :\ S_t \ge y_*\}\,; \qquad \tau_n \doteq \inf\{t \ge 0 :\ S_t \ge y_n\}$$

Since P is continuous, it is uniformly continuous on the compact subset D. It follows that there exists a number h such that

$$\mathbb{P}\left(t < \tau \le t+h \,\middle|\, S_0 = x\right) = P(x,t+h) - P(x,t) \le \frac{\varepsilon}{2} \qquad \forall\ (x,t) \in D,$$

and thus for all $(x,t) \in D$

$$\begin{aligned} &P_n(x,t) - P(x,t) \\ &\quad= \mathbb{P}^x(\tau_n \le t,\ \tau > t) \\ &\quad= \mathbb{P}^x(\tau_n \le t,\ \tau > t+h) + \mathbb{P}^x(\tau_n \le t,\ t < \tau \le t+h) \\ &\quad\le \mathbb{P}(\tau_n \le t,\ \tau > t+h) + \frac{\varepsilon}{2}. \end{aligned}$$

However, it follows from strong Markov property that

$$\mathbb{P}(\tau_n \le t,\ \tau > t+h \,|\, S_0 = x) \le \mathbb{P}\left(\max_{0\le u\le h} S_t \le y_* \,\middle|\, S_0 = y_n\right) \quad \forall\ (x,t) \in D.$$

Note that the right hand side is independent of $(x,t) \in D$. Define

$$c_1 \doteq \|b\|_\infty + \frac{1}{2}\|\sigma^2\|.$$

Then the right hand side is dominated by

$$\mathbb{P}\left(\max_{0\le u\le h}\left[-c_1 u + Q_u\right] \le \log\frac{y_*}{y_n}\right),$$

where $Q_u \doteq \int_0^u \sigma(S_t)\,dW_t$. Since $\underline{\sigma}^2 u \le \langle Q\rangle_u \le \bar{\sigma}^2 u$ and $Q_u = B_{\langle Q\rangle_u}$ for some standard Brownian motion B, the probability is in turn dominated by

$$\mathbb{P}\left(\max_{0\le t\le \bar{\sigma}^2 h}\left[-\frac{c_1}{\underline{\sigma}} t + B_t\right] \le \log\frac{y_*}{y_n}\right).$$

For n big enough, this probability is at most $\frac{\varepsilon}{2}$ since $y_n \to y$. This completes the proof. □

Proposition 4.1. *Suppose $f :\ [0,\infty) \to \mathbb{R}$ is a bounded continuous function with*

$$\lim_{x\to\infty} f(x) = 0,$$

and $g :\ [0,1) \to \mathbb{R}$ a continuous, bounded function. Then

$$\lim_{h,\delta\to 0} \mathbb{E}^x\left[f(\tau_\delta)\, g\left(\frac{\tau_\delta}{h} - \left\lfloor\frac{\tau_\delta}{h}\right\rfloor\right)\right] = \mathbb{E}^x f(\tau_*) \cdot \int_0^1 g(u)\, du.$$

for all $x \in (0, x_*)$.

Proof: Fix $x \in (0, x^*)$. Let p_δ and p denote the "density" of τ_δ and τ_* respectively. We can assume that all x_δ are close to x_* in the sense that $x_* - x_\delta \le \delta_0$ for some δ_0, and $x < x_\delta$. Since

$$\lim_{x \to \infty} f(x) = 0,$$

we have

$$\mathbb{E}^x \left[f(\tau_\delta) g\left(\frac{\tau_\delta}{h} - \left\lfloor \frac{\tau_\delta}{h} \right\rfloor \right) \right] = \int_0^\infty f(s) g\left(\frac{s}{h} - \left\lfloor \frac{s}{h} \right\rfloor \right) p_\delta(s)\, ds.$$

For any $\varepsilon > 0$, there exists $0 < a < M < \infty$ such that

$$\int_{[0,a]} f(s) g\left(\frac{s}{h} - \left\lfloor \frac{s}{h} \right\rfloor \right) p_\delta(s)\, ds \le \|f\|_\infty \cdot \|g\|_\infty \mathbb{P}(\tau_\delta \le a) \le \|f\|_\infty \cdot \|g\|_\infty \mathbb{P}(\tau_{\delta_0} \le a) \le \varepsilon.$$

and

$$\int_{[M,\infty]} f(s) g\left(\frac{s}{h} - \left\lfloor \frac{s}{h} \right\rfloor \right) p_\delta(s)\, ds \le \max_{M \le x} |f(x)| \cdot \|g\|_\infty \le \varepsilon.$$

Note that such choices of (a, M) also make the above inequalities hold when p_δ is replaced by p. Also since $p_\delta \to p$ uniformly on the compact interval $[\varepsilon, M]$, we have

$$\int_a^M f(s) g\left(\frac{s}{h} - \left\lfloor \frac{s}{h} \right\rfloor \right) |p_\delta(s) - p(s)|\, ds \le \varepsilon$$

for δ small enough. It remains to show that, for h small enough,

$$\left| \int_a^M f(s) g\left(\frac{s}{h} - \left\lfloor \frac{s}{h} \right\rfloor \right) p(s)\, ds - \int_a^M f(s) p(s)\, ds \cdot \int_0^1 g(u)\, du \right| \le \varepsilon.$$

However, since f, p and $f \cdot p$ are all uniformly continuous on compact intervals, we have

$$\begin{aligned}
\int_a^M f(s) g\left(\frac{s}{h} - \left\lfloor \frac{s}{h} \right\rfloor \right) p(s)\, ds &= \sum_{n=\lfloor \frac{a}{h} \rfloor}^{\lfloor \frac{M}{h} \rfloor} \int_{nh}^{(n+1)h} f(nh) g\left(\frac{s}{h} - n \right) p(nh)\, ds + o(1) \\
&= \int_0^1 g(u)\, du \sum_{n=\lfloor \frac{a}{h} \rfloor}^{\lfloor \frac{M}{h} \rfloor} f(nh) p(nh) \cdot h + o(1) \\
&= \int_0^1 g(u)\, du \int_a^M f(s) p(s)\, ds + o(1)
\end{aligned}$$

This completes the proof. □

References

[1] BATHER, J.A. & CHERNOFF, H. (1966) Sequential decisions in the control of a spaceship. *Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability* **3**, 181-207.

[2] BENSOUSSAN, A. (1984) On the theory of option pricing. *Acta Appl. Math.* **2**, 139-158.

[3] BRAUN, M. (1993) *Differential Equations and Their Applications.* Springer-Verlag, New York.

[4] BRENNAN, M.J. & SCHWARTZ, E.S. (1985) Evaluating natural rescource investments. *Journal of Business* **58**, 135-157.

[5] BROADIE, M. & GLASSERMAN, P. (1997) Pricing American-style securities using simulation. *J. Econ. Dynam. Control* **21**, 1323-1352.

[6] DIXIT, A.K. & PINDYCK, R.S. (1994) *Investment Under Uncertainty.* Princeton University Press.

[7] EL KAROUI, N., JEANBLANC-PICQUÉ, M. & SHREVE, S.E. (1998) Robustness of the Black and Scholes formula. *Math. Finance* **8**, 93-126.

[8] FRIEDMAN, A. (1964) *Partial Differential Equations of Parabolic Type.* Prentice-hall, Inc. Englewood Cliffs, NJ.

[9] JACKA, S.D. (1991) Optimal stopping and the American put. *Math. Finance* **1**, 1-14.

[10] KARATZAS, I. (1988) On the pricing of American options. *Appl. Math. Optim.* **17**, 37-60.

[11] KARATZAS, I. & SHREVE, S.E. (1998) *Methods of Mathematical Finance.* Springer, Berlin Heidelberg New York.

[12] KUNITA, H. (1990) *Stochastic Flows and Stochastic Differential Equations.* Cambridge, UK; Cambridge University Press.

[13] KUSHNER, H. (1984) *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic System Theory.* MIT Press, Cambridge, MA.

[14] LAMBERTON, D. (1998) Error estimates for the binomial approximation of American put options. *Ann. of Appl. Probab.* **8**, 206-233.

[15] MCDONALD, R. & SIEGEL, D. (1986) The value of waiting to invest. *Quarterly J. Econ.* **101**, 707-728.

[16] MERTON, R.C. (1990) *Continuous-Time Finance.* Cambridge-Oxford: Blackwell.

[17] MUSIELA, M. & RUTKOWSKI, M. (1997) *Martingale Methods in Financial Modelling.* Springer, Berlin Heidelberg New York.

[18] MYNENI, R. (1992) The pricing of the American option. *Ann. Appl. Probab.* **2**, 1-23.

[19] PROTTER, P. (1995) *Stochastic Integration and Differential Equations.* Springer-Verlag, Berlin.

[20] ROYDEN, H.L. (1968) *Real Analysis.* Collier MacMillan, London.

[21] SHIRYAYEV, A.N. (1978) *Optimal Stopping Rules.* Springer-Verlag, New York.