AN EXPLICIT SOLUTION TO AN OPTIMAL STOPPING PROBLEM WITH REGIME SWITCHING

XIN GUO,* *IBM Research Division and University of Alberta*

Abstract

We investigate an optimal stopping time problem which arises from pricing Russian options (i.e. perpetual look-back options) on a stock whose price fluctuations are modelled by adjoining a hidden Markov process to the classical Black–Scholes geometric Brownian motion model. By extending the technique of smooth fit to allow jump discontinuities, we obtain an explicit closed-form solution. It gives a non-standard application of the well-known smooth fit principle where the optimal strategy involves jumping over the optimal boundary and by an arbitrary overshoot. Based on the optimal stopping analysis, an arbitrage-free price for Russian options under the hidden Markov model is derived.

Keywords: Hidden Markov process; martingale; optimal stopping; regime switching; Russian options

AMS 2000 Subject Classification: Primary 60G40; 62L10; 93E20

1. Introduction

In [6], we proposed a model for the fluctuations of a single stock price X_t by incorporating the existence of inside information in the following form

$$\mathrm{d}X_t = X_t \mu_{\varepsilon(t)} \, \mathrm{d}t + X_t \sigma_{\varepsilon(t)} \, \mathrm{d}W_t, \tag{1.1}$$

where W_t is the standard Wiener process, $\varepsilon(t)$ is a Markov process, independent of W_t, which represents the state of information in the investor community. For each state i, there is a *known* drift parameter μ_i and a *known* volatility parameter σ_i. The pair $(\mu_{\varepsilon(t)}, \sigma_{\varepsilon(t)})$ takes different values when $\varepsilon(t)$ is in different states.

In this paper we consider the problem of pricing Russian options, based on this model. The Russian option was coined by Shepp and Shiryayev in [12]. It is a perpetual look-back option. The owner of the option can choose any exercise date, represented by the stopping time τ ($0 \leq \tau \leq \infty$) and gets a payoff of either s (a fixed constant) or the maximum stock price achieved up to the exercise date, whichever is larger, discounted by $\mathrm{e}^{-r\tau}$, where r is a fixed number.

To price Russian options, there is a closely related optimal stopping time problem. Let $X = \{X_t, \ t \geq 0\}$ be the price process for a stock with $X_0 = x > 0$, and

$$S_t = \max \left\{ s, \ \sup_{0 \leq u \leq t} X_u \right\},$$

where $s > x$ is a given constant. The problem is to compute the value of V,

$$V = \sup_{\tau} \mathrm{E}_{x,s} \, \mathrm{e}^{-r\tau} S_{\tau}, \tag{1.2}$$

where τ is a stopping time with respect to the filtration $\mathcal{F}_{X_t} = \{X(s), \ s \leq t\}$, meaning that no clairvoyance is allowed.

When $\mu_{\varepsilon(t)}$ and $\sigma_{\varepsilon(t)}$ are constants, equation (1.1) for stock price fluctuations X_t reduces to the Black–Scholes geometric Brownian motion model, for which the corresponding optimal stopping time problem was explicitly solved by Shepp and Shiryayev in [12]. Building on their result, Duffie and Harrison derived a unique arbitrage-free price for the Russian option [3]: the value is finite when the dividend payout rate is strictly positive, but is infinite otherwise.

For the general case of the market model, valuing Russian options is more complicated. The model is incomplete (see [6]) in the sense that there are not enough financial instruments to duplicate a portfolio and there does not exist a unique martingale measure under which the price of the option can be derived by taking the corresponding discounted expectation. Therefore, we will take two steps to solve the pricing problem.

First, we start our exploration by computing V in (1.2). This is an optimal stopping problem with an infinite time horizon and with state space $\{(\varepsilon, x, s) \mid x \leq s\}$. The key is to find the so called 'free boundary' $x = f(s, \mu, \sigma, \varepsilon)$ such that if $x \leq f(s, \mu_0, \mu_0, \sigma_0, \sigma_1, \lambda_0, \lambda_1)$ we should stop immediately and exercise the option, while if $s \geq x \geq f(s, \mu_0, \mu_0, \sigma_0, \sigma_1, \lambda_0, \lambda_1)$, we should keep observing the underlying stock fluctuations. By extending the technique of the 'principle of smooth fit', which should be attributed to Kolmogorov and Chernoff (see, for instance, [1]), to allow discontinuous jumps, we obtain a closed-form solution. We will show that, when the hidden Markov process $\varepsilon(t)$ switches from one state to another, there is a discontinuous jump over the boundary, which is also called 'regime switching'. The proof of the result is via martingale theory (Section 3).

Secondly, based on this result, we investigate the problem of pricing Russian options. By introducing the idea of a 'ticket', we succeed in completing the market model and in finding one equivalent martingale measure, under which the 'fair price' of the Russian option under the new model finds its home in economics (Section 6).

Not surprisingly, the success in obtaining an explicit closed-form solution relies heavily on the Markov structure of $(X_t/S_t, \varepsilon(t))$. Therefore, we discuss, via the methodology of the Markov dynamic programming, a detailed derivation of the closed-form solution. Out of curiosity, we will also provide examples where Markov structure provides no clue to any explicit solutions (Sections 4 and 5).

2. Problems and solutions

2.1. Problems

Throughout the paper, for explicitness, we will concentrate on the two-state case in which $\varepsilon(t)$ alternates between 0 and 1 such that

$$\mu_{\varepsilon(t)} = \begin{cases} \mu_0, & \varepsilon(t) = 0, \\ \mu_1, & \varepsilon(t) = 1, \end{cases}$$

$$\sigma_{\varepsilon(t)} = \begin{cases} \sigma_0, & \varepsilon(t) = 0, \\ \sigma_1, & \varepsilon(t) = 1, \end{cases}$$

where $\sigma_0 \neq \sigma_1$.

Moreover, let λ_i denote the rate of leaving state i and τ_i the time to leave state i; then

$$P(\tau_i > t) = e^{-\lambda_i t}, \qquad i = 0, 1. \tag{2.1}$$

We will illustrate that for a K-state case ($K > 2$), an algebraic equation of order $2K$ must be solved. Let

$$\mathrm{d}X_t = X_t \mu_{\varepsilon(t)}\,\mathrm{d}t + X_t \sigma_{\varepsilon(t)}\,\mathrm{d}W_t,$$

where $\mu_{\varepsilon(t)}, \sigma_{\varepsilon(t)}, W(t)$ are as defined above, and

$$S_t = \max\left\{s, \sup_{0 \le u \le t} X_u\right\},$$

with $s \ge X_0 = x$ fixed. Equation (1.2) can then be rewritten as

$$V = V^*(x, s, \mu_0, \mu_1, \sigma_0, \sigma_1, \lambda) = \sup_\tau \mathrm{E}_{x,s}\,\mathrm{e}^{-r\tau} S_\tau. \tag{2.2}$$

2.2. Solutions

Now define

$$V^*|_{\varepsilon(0)=i} = \sup_\tau \mathrm{E}_{x,s}[\mathrm{e}^{-r\tau} S_\tau \mid \varepsilon(0) = i];$$

let

$$\boldsymbol{G}_0(c_0, c_1) = \begin{bmatrix} (1-\beta_1)c_0^{\beta_1} & (1-\beta_2)c_0^{\beta_2} & (1-\beta_3)c_0^{\beta_3} & (1-\beta_4)c_0^{\beta_4} \\ (1-\beta_1)l_1 c_0^{\beta_1} & (1-\beta_2)l_2 c_0^{\beta_2} & (1-\beta_3)l_3 c_0^{\beta_3} & (1-\beta_4)l_4 c_0^{\beta_4} \end{bmatrix}$$

$$\times \begin{bmatrix} \beta_2\beta_3\beta_4\Delta_{(2,3,4)} & (\beta_2\beta_3+\beta_2\beta_4+\beta_3\beta_4)\Delta_{(2,3,4)} & -(\beta_2+\beta_3+\beta_4)\Delta_{(2,3,4)} & -\Delta_{(2,3,4)} \\ -\beta_1\beta_3\beta_4\Delta_{(1,3,4)} & -(\beta_1\beta_3+\beta_1\beta_4+\beta_3\beta_4)\Delta_{(1,3,4)} & (\beta_1+\beta_3+\beta_4)\Delta_{(1,3,4)} & \Delta_{(1,3,4)} \\ \beta_2\beta_1\beta_4\Delta_{(1,2,4)} & (\beta_2\beta_1+\beta_1\beta_4+\beta_2\beta_4)\Delta_{(1,2,4)} & -(\beta_2+\beta_1+\beta_4)\Delta_{(1,2,4)} & -\Delta_{(1,2,4)} \\ -\beta_2\beta_1\beta_3\Delta_{(1,2,3)} & -(\beta_2\beta_1+\beta_1\beta_3+\beta_2\beta_3)\Delta_{(1,2,3)} & (\beta_2+\beta_1+\beta_3)\Delta_{(1,2,3)} & \Delta_{(1,2,3)} \end{bmatrix}$$

$$\times \begin{bmatrix} 1 \\ 0 \\ g_1 \\ g_2 \end{bmatrix},$$

and

$$\boldsymbol{G}_1(c_1, c_0) = \begin{bmatrix} (1-\beta_1)c_1^{\beta_1} & (1-\beta_2)c_1^{\beta_2} & (1-\beta_3)c_1^{\beta_3} & (1-\beta_4)c_1^{\beta_4} \\ (1-\beta_1)l_1 c_1^{\beta_1} & (1-\beta_2)l_2 c_1^{\beta_2} & (1-\beta_3)l_3 c_1^{\beta_3} & (1-\beta_4)l_4 c_1^{\beta_4} \end{bmatrix}$$

$$\times \begin{bmatrix} \beta_2\beta_3\beta_4\Delta_{(2,3,4)} & (\beta_2\beta_3+\beta_2\beta_4+\beta_3\beta_4)\Delta_{(2,3,4)} & -(\beta_2+\beta_3+\beta_4)\Delta_{(2,3,4)} & -\Delta_{(2,3,4)} \\ -\beta_1\beta_3\beta_4\Delta_{(1,3,4)} & -(\beta_1\beta_3+\beta_1\beta_4+\beta_3\beta_4)\Delta_{(1,3,4)} & (\beta_1+\beta_3+\beta_4)\Delta_{(1,3,4)} & \Delta_{(1,3,4)} \\ \beta_2\beta_1\beta_4\Delta_{(1,2,4)} & (\beta_2\beta_1+\beta_1\beta_4+\beta_2\beta_4)\Delta_{(1,2,4)} & -(\beta_2+\beta_1+\beta_4)\Delta_{(1,2,4)} & -\Delta_{(1,2,4)} \\ -\beta_2\beta_1\beta_3\Delta_{(1,2,3)} & -(\beta_2\beta_1+\beta_1\beta_3+\beta_2\beta_3)\Delta_{(1,2,3)} & (\beta_2+\beta_1+\beta_3)\Delta_{(1,2,3)} & \Delta_{(1,2,3)} \end{bmatrix}$$

$$\times \begin{bmatrix} 1 \\ 0 \\ \hat{g}_1 \\ \hat{g}_2 \end{bmatrix},$$

where $\gamma_1 < \gamma_2$ are the (real) roots of the equation

$$(r + \lambda_1) = \gamma\mu_1 + \tfrac{1}{2}\gamma(\gamma - 1)\sigma_1^2,$$

$\hat{\gamma}_1 < \hat{\gamma}_2$ are the (real) roots of the equation

$$(r + \lambda_0) = \gamma\mu_0 + \tfrac{1}{2}\gamma(\gamma - 1)\sigma_0^2,$$

the β_i are the (real) roots of the quartic equation

$$[(r + \lambda_0) - \beta\mu_0 - \tfrac{1}{2}\beta(\beta - 1)\sigma_0^2][(r + \lambda_1) - \beta\mu_1 - \tfrac{1}{2}\beta(\beta - 1)\sigma_1^2] = \lambda_0\lambda_1, \qquad (2.3)$$

for $i < j < k$

$$\Delta_{(i,j,k)} = (\beta_i - \beta_j)(\beta_i - \beta_k)(\beta_j - \beta_k)$$

and

$$g_1 = -\frac{2}{\sigma_0^2}\left(\lambda_0\left(\frac{r}{(r + \lambda_1)(\gamma_2 - \gamma_1)}\left(\gamma_2\left(\frac{c_0}{c_1}\right)^{\gamma_1} - \gamma_1\left(\frac{c_0}{c_1}\right)^{\gamma_2}\right) + \frac{\lambda_1}{r + \lambda_1}\right) - r - \lambda_0\right),$$

$$g_2 = -\frac{2}{\sigma_0^2}\left(\lambda_0\left(\frac{r\gamma_1\gamma_2}{(r + \lambda_1)(\gamma_2 - \gamma_1)}\left(\left(\frac{c_0}{c_1}\right)^{\gamma_1} - \left(\frac{c_0}{c_1}\right)^{\gamma_2}\right)\right)\right.$$
$$-\frac{2\mu_0 + \sigma_0^2}{\sigma_0}\left[\lambda_0\left(\frac{r}{(r + \lambda_1)(\gamma_2 - \gamma_1)}\left(\gamma_2\left(\frac{c_0}{c_1}\right)^{\gamma_1} - \gamma_1\left(\frac{c_0}{c_1}\right)^{\gamma_2}\right)\right.\right.$$
$$\left.\left.\left. + \frac{\lambda_1}{r + \lambda_1}\right) - r - \lambda_0\right]\right),$$

$$\hat{g}_1 = -\frac{2}{\sigma_1^2}\left(\lambda_1\left(\frac{r}{(r + \lambda_0)(\hat{\gamma}_2 - \hat{\gamma}_1)}\left(\hat{\gamma}_2\left(\frac{c_1}{c_0}\right)^{\hat{\gamma}_1} - \hat{\gamma}_1\left(\frac{c_1}{c_0}\right)^{\hat{\gamma}_2}\right) + \frac{\lambda_0}{r + \lambda_0}\right) - r - \lambda_1\right)$$

$$\hat{g}_2 = -\frac{2}{\sigma_1^2}\left(\lambda_1\left(\frac{r\hat{\gamma}_1\hat{\gamma}_2}{(r + \lambda_0)(\hat{\gamma}_2 - \hat{\gamma}_1)}\left(\left(\frac{c_1}{c_0}\right)^{\hat{\gamma}_1} - \left(\frac{c_1}{c_0}\right)^{\hat{\gamma}_2}\right)\right)\right.$$
$$-\frac{2\mu_1 + \sigma_1^2}{\sigma_1}\left[\lambda_1\left(\frac{r}{(r + \lambda_0)(\hat{\gamma}_2 - \hat{\gamma}_1)}\left(\hat{\gamma}_2\left(\frac{c_1}{c_0}\right)^{\hat{\gamma}_1} - \hat{\gamma}_1\left(\frac{c_1}{c_0}\right)^{\hat{\gamma}_2}\right)\right.\right.$$
$$\left.\left.\left. + \frac{\lambda_0}{r + \lambda_0}\right) - r - \lambda_1\right]\right).$$

Theorem 2.1. (i) *Assume that $r > \max\{(0, \mu_0, \mu_1\}$, $\lambda_i \neq 0$, and suppose there exist $0 < c_1 < c_0 < 1$ such that $G_0(c_0, c_1) = 0$, then*

$$V^*|_{\varepsilon(0)=i} = V_i(x, s) \geq s$$

such that

$$V_0(x, s) = \begin{cases} s, & x < c_0 s, \\ \displaystyle\sum_{i=1}^{4} a_i x^{\beta_i} s^{1-\beta_i}, & c_0 s < x < s, \end{cases} \qquad (2.4)$$

$$V_1(x, s) = \begin{cases} s, & 0 < x \leq c_1 s, \\ \displaystyle\frac{rs}{(r + \lambda_1)(\gamma_2 - \gamma_1)}\left[\gamma_2\left(\frac{x}{c_1 s}\right)^{\gamma_1} - \gamma_1\left(\frac{x}{c_1 s}\right)^{\gamma_2}\right] + \frac{\lambda_1 s}{r + \lambda_1}, & c_1 s < x < c_0 s, \\ \displaystyle\sum_{i=1}^{4} b_i x^{\beta_i} s^{1-\beta_i}, & c_0 s < x < s, \end{cases}$$

$$(2.5)$$

where the a_i satisfy

$$\begin{bmatrix} l_1 & l_2 & l_3 & l_4 \\ \beta_1 l_1 & \beta_2 l_2 & \beta_3 l_3 & \beta_4 l_4 \\ 1 & 1 & 1 & 1 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{r}{(r+\lambda_1)(\gamma_2-\gamma_1)}\left(\gamma_2\left(\dfrac{c_0}{c_1}\right)^{\gamma_1} - \gamma_1\left(\dfrac{c_0}{c_1}\right)^{\gamma_2}\right) + \dfrac{\lambda_1}{r+\lambda_1} \\[4mm] \dfrac{r\gamma_1\gamma_2}{(r+\lambda_1)(\gamma_2-\gamma_1)}\left(\left(\dfrac{c_0}{c_1}\right)^{\gamma_1} - \left(\dfrac{c_0}{c_1}\right)^{\gamma_2}\right) \\[4mm] 1 \\[2mm] 0 \end{bmatrix} \quad (2.6)$$

and $b_i = l_i a_i$, where

$$l_i = \frac{r + \lambda_0 - \beta_i \mu_0 - \frac{1}{2}\beta_i(\beta_i-1)\sigma_0^2}{\lambda_0}.$$

Otherwise, there exist $0 < c_0 < c_1 < 1$ so that $\mathbf{G}_1(c_1, c_0) = \mathbf{0}$ and

$$V^*|_{\varepsilon(0)=i} = V_i(x,s) \geq s$$

such that

$$V_1(x,s) = \begin{cases} s, & x < c_1 s, \\ \displaystyle\sum_{i=1}^{4} a_i x^{\beta_i} s^{1-\beta_i}, & c_1 s < x < s, \end{cases} \quad (2.7)$$

$$V_0(x,s) = \begin{cases} s, & 0 < x \leq c_0 s, \\[3mm] \dfrac{rs}{(r+\lambda_0)(\hat{\gamma}_2-\hat{\gamma}_1)}\left[\hat{\gamma}_2\left(\dfrac{x}{c_0 s}v\right)^{\hat{\gamma}_1} - \hat{\gamma}_1\left(\dfrac{x}{c_0 s}\right)^{\hat{\gamma}_2}\right] + \dfrac{\lambda_0 s}{r+\lambda_0}, & c_0 s < x < c_1 s, \\[3mm] \displaystyle\sum_{i=1}^{4} b_i x^{\beta_i} s^{1-\beta_i}, & c_1 s < x < s, \end{cases}$$

$$(2.8)$$

with the a_i given by

$$\begin{bmatrix} \hat{l}_1 & \hat{l}_2 & \hat{l}_3 & \hat{l}_4 \\ \beta_1 \hat{l}_1 & \beta_2 \hat{l}_2 & \beta_3 \hat{l}_3 & \beta_4 \hat{l}_4 \\ 1 & 1 & 1 & 1 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{r}{(r+\lambda_0)(\hat{\gamma}_2-\hat{\gamma}_1)}\left(\hat{\gamma}_2\left(\dfrac{c_1}{c_0}\right)^{\hat{\gamma}_1} - \hat{\gamma}_1\left(\dfrac{c_1}{c_0}\right)^{\hat{\gamma}_2}\right) + \dfrac{\lambda_0}{r+\lambda_0} \\[4mm] \dfrac{r\hat{\gamma}_1\hat{\gamma}_2}{(r+\lambda_0)(\hat{\gamma}_2-\hat{\gamma}_1)}\left(\left(\dfrac{c_1}{c_0}\right)^{\hat{\gamma}_1} - \left(\dfrac{c_1}{c_0}\right)^{\hat{\gamma}_2}\right) \\[4mm] 1 \\[2mm] 0 \end{bmatrix} \quad (2.9)$$
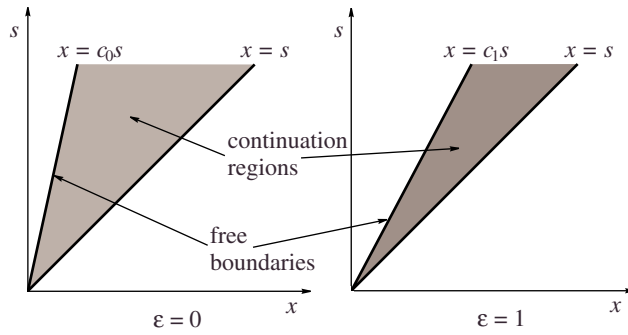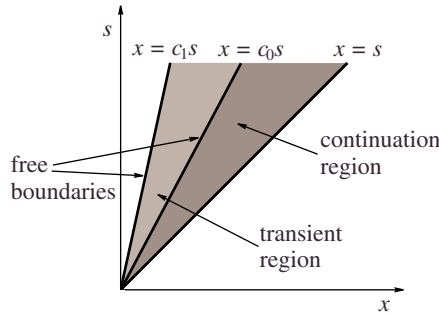
FIGURE 1.

FIGURE 2.

and $b_i = \hat{l}_i a_i$, where

$$l_i = \frac{r + \lambda_1 - \beta_i \mu_1 - \frac{1}{2}\beta_i(\beta_i - 1)\sigma_1^2}{\lambda_1} = \hat{l}_i a_i.$$

(ii) *Define $\Omega_i = \{(x, s) \mid c_i s \le x \le s\}$. Then the optimal stopping time τ_i will be the first time to leave Ω_i at state i, that is,*

$$\tau_i = \inf\left\{t \ge 0 \;\middle|\; \frac{X_t}{S_t} = c_i, \, \varepsilon(t) = i\right\}.$$

The region Ω_i is called the 'continuation region'; see Figure 1.

The interesting part of Theorem 2.1 is the so-called 'regime switching': namely, when $\varepsilon(t)$ changes state, there is an instantaneous jump. We define the 'transient region' to be the region $\{(x, s) \mid c_1 s \le x \le c_0 s\}$ when $c_1 < c_0$, and, when $c_0 < s_1$, to be the region $\{(x, s) \mid c_0 s \le x \le c_1 s\}$; see Figure 2.

When $l_i = 1$, and consequently $\hat{l}_i = 1$, we have $\lambda_0 \lambda_1 = 0$, and Theorem 2.1 is the same as in [12]. Evidently, the a_is and b_is, and hence V_0, V_1 are uniquely determined by c_0 and c_1, because the left most 4×4 matrix in (2.6) is a Vandermonde matrix which is invertible.

With some calculation, it is not hard to see that the equations (2.4) and (2.5) satisfy the following: on the region $\Omega_0 = \{(x, s) \mid c_0 s < x < s, \, 0 < x < s\}$,

$$\begin{aligned}
(r + \lambda_0)V_0 &= x\mu_0 V_0' + \tfrac{1}{2}x^2\sigma_0^2 V_0'' + \lambda_0 V_1, \\
(r + \lambda_1)V_1 &= x\mu_1 V_1' + \tfrac{1}{2}x^2\sigma_1^2 V_1'' + \lambda_1 V_0,
\end{aligned} \tag{2.10}$$

on the region $\Omega_1 = \{(x, s) \mid 0 < x \le s, \ c_1 s < x < c_0 s\}$,

$$(r + \lambda_1)V_1 = x\mu_1 V_1' + \tfrac{1}{2}x^2\sigma_1^2 V_1'' + \lambda_1 s,$$

and

$$V_1(x, s)|_{x=c_0 s} = \frac{rs}{(r+\lambda_1)(\gamma_2 - \gamma_1)}\left(\gamma_2\left(\frac{c_0}{c_1}\right)^{\gamma_1} - \gamma_1\left(\frac{c_0}{c_1}\right)^{\gamma_2}\right) + \frac{\lambda_1 s}{r + \lambda_1},$$

$$\frac{\partial V_1}{\partial x}\bigg|_{x=c_0 s} = \frac{r\gamma_1\gamma_2}{c_0(r+\lambda_1)(\gamma_2 - \gamma_1)}\left(\left(\frac{c_0}{c_1}\right)^{\gamma_1} - \left(\frac{c_0}{c_1}\right)^{\gamma_2}\right),$$

$$V_0(x, s)|_{x=c_0 s} = s \quad \text{(smooth fit)},$$

$$\frac{\partial V_0}{\partial x}\bigg|_{x=c_0 s} = 0 \quad \text{(smooth fit)},$$

$$\frac{\partial V_1(x, s)}{\partial x}\bigg|_{x=s} = 0 \quad \text{(smooth fit)},$$

$$\frac{\partial V_0(x, s)}{\partial x}\bigg|_{x=s} = 0 \quad \text{(smooth fit)}.$$

This is crucial in the derivation of the equations (2.4) and (2.5). The basic idea is the 'principle of smooth fit'.

Remark 2.1. If $\lambda_0, \lambda_1, \sigma_0, \sigma_1$ are positive, then the equation

$$[(r + \lambda_0) - \beta\mu_0 - \tfrac{1}{2}\beta(\beta - 1)\sigma_0^2][(r + \lambda_1) - \beta\mu_1 - \tfrac{1}{2}\beta(\beta - 1)\sigma_1^2] = \lambda_0\lambda_1$$

has four real roots.

The proof of this is straightforward: let

$$f(\beta) = [(r + \lambda_0) - \beta\mu_0 - \tfrac{1}{2}\beta(\beta - 1)\sigma_0^2][(r + \lambda_1) - \beta\mu_1 - \tfrac{1}{2}\beta(\beta - 1)\sigma_1^2] - \lambda_0\lambda_1,$$

and let θ_1, θ_2 be the roots of the corresponding quadratic equation,

$$(r + \lambda_0) - \theta\mu_0 - \tfrac{1}{2}\theta(\theta - 1)\sigma_0^2 = 0.$$

Clearly, the continuous function $f(\beta)$ satisfies $f(0) > 0$, $f(-\infty) > 0$, $f(\infty) > 0$, and $f(\theta_i) = -\lambda_1\lambda_2 < 0$, for $i = 1, 2$. Since $\theta_1\theta_2 = -2(r + \lambda_0)/\sigma_0^2 < 0$, it follows that the equation $f(\beta) = 0$ has four different real roots.

3. Proof of Theorem 2.1

In order to prove that V_0, V_1 given in Theorem 2.1 are indeed optimal, it suffices to verify that

$$Y_t = e^{-rt} V(X_t, S_t, \varepsilon(t))$$

is a uniformly integrable supermartingale and that the optimal stopping time is finite with probability 1. To see this, observe that if Y_t is a supermartingale and $V_i(x, s) \ge s$, then we have

$$\begin{aligned} E_{x,s}\, e^{-r\tau} S_\tau &\le E_{x,s}\, e^{-r\tau} V(X_\tau, S_\tau, \varepsilon(\tau)) = E_{x,s} Y_\tau \\ &\le E_{x,s} Y_0 = V(X_0, S_0, \varepsilon(0)) \\ &= V_i(x, s), \end{aligned}$$

and, if we take the supremum over all such τ, we obtain for all $0 \le x \le s$,

$$V^*|_{\varepsilon(0)=i} \le V_i(x, s).$$

On the other hand, the fact that Y_t is a uniformly integrable martingale implies that for any fixed t,

$$\mathrm{E}_{x,s} Y_{t \wedge \tau_i} = \mathrm{E}_{x,s} Y_0.$$

Hence, Y_t is a positive uniformly integrable local martingale.

In fact, Y_t is a uniformly integrable martingale. To see this, recall that any continuous local martingale with an integrable quadratic variation is a martingale (see [9, 1.5.24]). Therefore, it is sufficient to check that

$$\mathrm{E}\langle Y_t \rangle_\infty < \infty.$$

Since S_t increases only when $X_t = S_t$ with $V_s(s, s) = 0$, and $\mathrm{d}S_t = 0$ for all $X_t < S_t$, by Itô's formula and Fubini's theorem, we have

$$\begin{aligned}
\mathrm{E}\langle Y_t \rangle_\infty &= \mathrm{E} \int_0^\infty \mathrm{e}^{-2rt} V_x^2(X_t, S_t, \varepsilon(t)) X_t^2 \sigma_\varepsilon^2 \, \mathrm{d}t \\
&\le K^2 \sigma^2 \mathrm{E} \int_0^\infty \mathrm{e}^{-2rt} X_t^2 \, \mathrm{d}t \\
&\le K^2 \sigma^2 \int_0^\infty \mathrm{e}^{-2\zeta t} \, \mathrm{d}t \\
&< \infty,
\end{aligned}$$

where $\sigma = \max \sigma_\varepsilon$ and $K < \infty$ is an upper bound of $V_x(X_t, S_t, \varepsilon_t)$, which is a linear combination of $(X_t/S_t)^{\alpha - 1}$ for some constant α_1 and hence is bounded on the compact set $0 < c_i \le X_t/S_t \le 1$ (from (2.7) and (2.9)). The third inequality uses the fact that $\zeta = r - \max\{\mu_0, \mu_1\} > 0$ and $\mathrm{E}[X_t^2] < \mathrm{e}^{2\mu t}$ with $\mu = \max\{\mu_0, \mu_1\}$.

Now using the optional stopping theorem [10] and taking the supremum over all choices of the optimal stopping time, we get

$$V^*|_{\varepsilon(0)=i}(x, s) \ge V_i(x, s).$$

Proposition 3.1. *The martingale*

$$Y_t = \mathrm{e}^{-rt} V(X_t, S_t, \varepsilon(t))$$

is a supermartingale.

The following differentiation rule will facilitate our proof (see [4]).

Lemma 3.1. *Suppose that X is process with values in $\mathbb{R}^n$, each of whose components X^i is a semimartingale. Suppose F is a real valued twice continuously differentiable function on $\mathbb{R}^n$,*

then $F(X_t)$ is a semimartingale and

$$F(X_t) = F(X_0) + \sum_{i=1}^{n} \int_{(0,t]} \frac{\partial}{\partial x_i} F(X_{s-}) \, dX_s^i$$

$$+ \frac{1}{2} \sum_{i,j=1}^{n} \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} F(X_{s-}) \, d\langle X^{ic}, X^{jc} \rangle_s$$

$$+ \sum_{0 < s \le t} \left(F(X_s) - F(X_{s-}) - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} F(X_{s-}) \Delta X_s^i \right).$$

Proof of Proposition 3.1. The idea of the proof is as follows. Let $t > 0$, $h > 0$ and $\Delta Y_t = Y_{h+t} - Y_t$. Notice that

$$\Delta X_t = \mu_{\varepsilon(t)} X_t h + \sigma_{\varepsilon(t)} \Delta W_h.$$

Then

$$\mathrm{E}^{F_t} \Delta Y_t = \mathrm{E}[Y_{t+h} - Y_t \mid F_t^X],$$

which is actually

$$\mathrm{E}[\mathrm{e}^{-r(t+h)} V(X_{t+h}, s_{t+h}, \varepsilon_{t+h}) - \mathrm{e}^{-rt} V(X_t, S_t, \varepsilon_t) \mid F_t]$$

$$= \mathrm{E}[\mathrm{e}^{-r(t+h)} V(X_{t+h}, s_{t+h}, \varepsilon_{t+h}) \mathbf{1}_{(\varepsilon_{t+h} = \varepsilon_t)} - \mathrm{e}^{-rt} V(X_t, S_t, \varepsilon_t) \mid F_t]$$

$$\quad + \mathrm{E}[\mathrm{e}^{-r(t+h)} V(X_{t+h}, s_{t+h}, \varepsilon_{t+h}) \mathbf{1}_{(\varepsilon_{t+h} \ne \varepsilon_t)} - \mathrm{e}^{-rt} V(X_t, S_t, \varepsilon_t) \mid F_t]$$

$$= \mathrm{E}[\mathrm{e}^{-rt} \mathrm{e}^{-\lambda_t h} \mathrm{e}^{-rh} (V(X_t, S_t, \varepsilon_t) + V'(X_t, S_t, \varepsilon_t) \Delta X_t + \tfrac{1}{2} V''(X_t, S_t, \varepsilon_t) \Delta X_t^2)$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad - \mathrm{e}^{-rt} V(X_t, S_t, \varepsilon_t) \mid F_t]$$

$$\quad + \mathrm{E}[\mathrm{e}^{-rt} (1 - \mathrm{e}^{-\lambda_t h}) \mathrm{e}^{-rh} (V(X_t, S_t, \varepsilon_t')$$

$$\qquad\qquad + V'(X_t, S_t, \varepsilon_t') \Delta X_t + \tfrac{1}{2} V''(X_t, S_t, \varepsilon_t') \Delta X_t^2) \mid F_t]$$

$$= \mathrm{E}[\mathrm{e}^{-rt} (1 - \lambda_t h - rh)(V(X_t, S_t, \varepsilon_t) + V'(X_t, S_t, \varepsilon_t) \mu_{\varepsilon(t)} X_t h$$

$$\qquad + \tfrac{1}{2} V''(X_t, S_t, \varepsilon_t) \sigma_{\varepsilon(t)}^2 X_t^2 h + o(h^2)) - \mathrm{e}^{-rt} V(X_t, S_t, \varepsilon_t) \mid F_t]$$

$$\quad + \mathrm{E}[\mathrm{e}^{-rt} \mathrm{e}^{-rh} \lambda_t h (V'(X_t, S_t, \varepsilon_t) + V'(X_t, S_t, \varepsilon_{(t)}') \mu_{\varepsilon(t)} X_t h$$

$$\qquad\qquad\qquad\qquad + \tfrac{1}{2} V''(X_t, S_t, \varepsilon_{(t)}') \sigma_{\varepsilon_{(t)}'}^2 X_t^2 h) \mid F_t]$$

$$= \mathrm{E}[h \, \mathrm{e}^{-rt} [-(\lambda_t + r) V(X_t, S_t, \varepsilon_t) + V'(X_t, S_t, \varepsilon_t) \mu_{\varepsilon(t)} X_t$$

$$\qquad\qquad\qquad + \tfrac{1}{2} V''(X_t, S_t, \varepsilon_t) \sigma_{\varepsilon(t)}^2 X_t^2]$$

$$\qquad\qquad\qquad + \lambda_t h V(X_t, S_t, \varepsilon_{(t)}') + o(h^2) \mid F_t],$$

where the partial derivative is with respect to the first variable.

More rigorously, recall that for any Markov process $Z(t)$ with generator A, $f(Z(t)) - \int_0^t A(f(z(s))) \, ds$ is a martingale when f is twice continuously differentiable. Notice that $\varepsilon(t)$ has a generator

$$\begin{bmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & -\lambda_1 \end{bmatrix},$$

and $f(\varepsilon) = \varepsilon$, so that $f(0) = 0$, $f(1) = 1$, and therefore

$$M(t) = \varepsilon(t) - \int_0^t \begin{bmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & -\lambda_1 \end{bmatrix} \varepsilon(s) \, \mathrm{d}s$$

is a martingale. In consequence, we can write

$$\varepsilon(t) = \int_0^t \{\lambda_0 \chi_0(\varepsilon(s)) - \lambda_1 \chi_1(\varepsilon(s))\} \, \mathrm{d}s + M(t).$$

Therefore, applying the differentiation rule, we get

$$\mathrm{d}V(X_t, S_t, \varepsilon(t)) \, \mathrm{e}^{-rt}$$

$$= \mathrm{e}^{-rt} \left[-rV \, \mathrm{d}t + V_x(X_t, S_t, \varepsilon(t)) \, \mathrm{d}x + \tfrac{1}{2}\sigma_\varepsilon^2 V_{xx}(X_t, S_t, \varepsilon(t)) \, \mathrm{d}t \right.$$

$$\left. + V_s(X_t, S_t, \varepsilon(t)) \, \mathrm{d}S_t + \sum_{s \leq t} V(X_s, S_s, \varepsilon(s)) - V(X_s, S_s, \varepsilon(s^-)) \right].$$

Furthermore, since $\mathrm{d}\varepsilon(s) = -1$ when $\varepsilon(s) = 0$ and $\varepsilon(s^-) = 1$, we can rewrite the sum in the above equation as

$$\sum_{s \leq t} V(X_s, S_s, \varepsilon(s)) - V(X_s, S_s, \varepsilon(s^-))$$

$$= \int_0^t V(X_s, S_s, \varepsilon(s)) - V(X_s, S_s, \varepsilon(s^-))(-1)^{\varepsilon(s^-)} \, \mathrm{d}\varepsilon(s).$$

Thus, effectively, there is another martingale, say $\bar{M}_t$, such that

$$\bar{M}_t = \int_0^t -rV + V_x u_\varepsilon + \tfrac{1}{2}V_{xx}\sigma_\varepsilon^2 + (-1)^{\varepsilon(s^-)}[\lambda_0(1 - \varepsilon(s)) - \lambda_1 \varepsilon(s)]$$

$$\times \, [V(X_s, S_s, 1 - \varepsilon(s^-)) - V(X_s, S_s, \varepsilon(s^-))] \, \mathrm{d}s.$$

Together with (2.10), we see that whenever we start from time t, $\mathrm{d}Y_t \leq 0$, hence Y_t is a supermartingale in the region $0 < X_t \leq S_t c_i$. (Here c_i is chosen according to the starting state of $\varepsilon(t)$.) In the region $S_t c_i \leq X_t \leq S_t$, because S_t grows only when $X_t = S_t$, and $V_s(s, s) = 0$, we see that Y_t is a positive local martingale.

Proposition 3.2. *The martingale Y_t is uniformly integrable.*

The following lemma gives us the key (see [10]).

Lemma 3.2. (Doob's L^p-inequality.) *If X is a right-continuous martingale or positive submartingale indexed by an interval T of R, then if $X^* = \sup_t |X_t|$, for $p > 1$,*

$$\lambda^p P[X^* \geq \lambda] \leq \sup_t E[|X_t|^p].$$

Proof of Proposition 3.2. By (2.4), (2.5), (2.7) and (2.9), the V_is are bounded from above when s is fixed and therefore the uniform integrability of Y_t is equivalent to that of $Z_t = \mathrm{e}^{-rt} S_t$. Thus, it is sufficient to show that there exists a $p > 1$ such that

$$\sup_{0 \leq t < \infty} E|Z_t|^p < \infty.$$

For fixed t, using integration by parts,

$$\mathrm{E}|Z_t|^p = \int_0^\infty \mathrm{P}(Z_t^p > y)\,\mathrm{d}y = \int_0^\infty \mathrm{P}(Z_t > y^{1/p})\,\mathrm{d}y.$$

Let $\zeta = r - \max\{\mu_0, \mu_1\}$, recalling that $r > \max\{0, \mu_0, \mu_1\}$, we have $\zeta > 0$, and

$$
\begin{aligned}
\mathrm{P}(Z_t > y^{1/p}) &= \mathrm{P}\left\{ e^{-rt} \sup_{0 \le s \le t} \exp\left(\int_0^s \sigma_\varepsilon\,\mathrm{d}W_s + \int_0^s \mu_\varepsilon\,\mathrm{d}s - \int_0^s \tfrac{1}{2}\sigma_\varepsilon^2\,\mathrm{d}s \right) > y^{1/p} \right\} \\
&\le \mathrm{P}\left\{ \sup_{0 \le s \le t} \exp\left(\int_0^s \sigma_\varepsilon\,\mathrm{d}W_s - \int_0^s \sigma_\varepsilon^2\,\mathrm{d}s - t\zeta \right) > y^{1/p} \right\} \\
&= \mathrm{P}\left\{ \sup_{0 \le s \le t} \exp\left(\int_0^s \sigma_\varepsilon\,\mathrm{d}W_s - \int_0^s \tfrac{1}{2}\sigma_\varepsilon^2\,\mathrm{d}s \right) > e^{\zeta t} y^{1/p} \right\} \\
&\le \frac{\mathrm{E}\left| \exp\left(\int_0^t \sigma_\varepsilon\,\mathrm{d}W_t - \tfrac{1}{2}\sigma_\varepsilon^2\,\mathrm{d}t \right) \right|^{p'}}{e^{p'\zeta t} y^{p'/p}} \\
&= \frac{\exp(tp'(p'\sigma^2 - \sigma^2 - 2\zeta)/2)}{y^{p'/p}},
\end{aligned}
$$

where $1 < p'$ is some constant to be chosen. The first inequality holds because, if $\mathrm{P}\{X \ge Y\} = 1$, then $\mathrm{P}\{X > y\} \ge \mathrm{P}\{Y \ge y\}$. The second inequality is from Doob's L^p inequality. When $\zeta > 0$, we can choose a $p' > 1$ such that $((p')^2 - p')\sigma^2 - 2\zeta < 0$ and choose an appropriate $1 < p < p'$ such that $\mathrm{E}|Z_t|^p$ is dominated by $\int_0^\infty y^{p'/p}\,\mathrm{d}y$ which is finite. Hence,

$$\sup_{0 \le t < \infty} \mathrm{E}|Z_t|^p < \infty.$$

Proposition 3.3. *We have*

$$\mathrm{P}(\tau_i < \infty) = 1.$$

Proof. Since the stopping time is the first time at which the process (X_t, S_t) ever reaches the line $x = c_i s$ whenever starting from state i, we can actually prove a slightly more general statement, viz. $\mathrm{P}\{\tau = \infty\} = 0$, where τ is the first time t such that

$$\frac{\max_{0 \le u \le t} X_u}{X_t} \ge \frac{1}{c}$$

with $0 < c < 1$. This is equivalent to proving that $\mathrm{P}\{\tau > T\} \to 0$ as $T \to \infty$.

Noticing that

$$\mathrm{P}\{\tau > T\} = \mathrm{P}\left\{ \int_u^t \mu_\varepsilon - \frac{\sigma_\varepsilon^2}{2}\,\mathrm{d}s + \int_u^t \sigma_\varepsilon\,\mathrm{d}B_s \ge \log c,\ 0 \le u \le t \le T \right\},$$

the independence of $(X_s, X_t/X_s)$ implies that

$$
\begin{aligned}
\mathrm{P}\{\tau > 2T\} &\le \mathrm{P}\left\{ \int_u^t \mu_\varepsilon - \frac{\sigma_\varepsilon^2}{2}\,\mathrm{d}s + \int_u^t \sigma_\varepsilon\,\mathrm{d}B_s \ge \log c,\ 0 \le u \le t \le T \right\} \\
&\quad + \mathrm{P}\left\{ \int_u^t \mu_\varepsilon - \frac{\sigma_\varepsilon^2}{2}\,\mathrm{d}s + \int_u^t \sigma_\varepsilon\,\mathrm{d}B_s \ge \log c,\ T \le u \le t \le 2T \right\} \\
&\le (\mathrm{P}\{\tau > T\})^2.
\end{aligned}
$$

Since a mixture of Gaussian processes is unbounded, the propositions follows.

4. Markov structure of $(X_t/S_t, \varepsilon(t))$

In [13], by the observation that (X_t/S_t) is a Markov process when X_t is a geometric Brownian motion, the derivation of the Russian option pricing problem was simplified and the free boundary problem amounted to finding a threshold of the Markov process (X_t/S_t).

Few modifications are needed to apply in a straightforward way the same methodology to our case in validating that $(X_t/S_t, \varepsilon(t))$ is Markovian. We need only replace the classical Wiener space in [13] generated by $(\Omega, \mathcal{F}, F^W = (\mathcal{F}_t^W)_{t \geq 0}, \mathrm{P})$ with $(\Omega, \mathcal{F}, F^X = (\mathcal{F}_t^X)_{t \geq 0}, \mathrm{P})$.

Proposition 4.1. *The process $\{\psi_t^{\varepsilon(t)}\}$ is Markovian with respect to $(\Omega, \mathcal{F}, F^X, \mathrm{P})$, where*

$$\psi_t^{\varepsilon(t)} = \frac{\max\{\max_{u \leq t} X_u^{\varepsilon(u)}, \; X_0^{\varepsilon(0)} \psi_0^{\varepsilon(0)}\}}{X_t^{\varepsilon(t)}}.$$

For the detailed proof of Proposition 4.1, interested readers are referred to [13].

Although Proposition 4.1 is true, it is not clear in general how far the Markovian property can carry us. Shepp [11] gave an example in which he considers the optimal stopping problem

$$V(w, m) = \sup_{\tau} \mathrm{E}_{w,m} M_\tau \, \mathrm{e}^{-r\tau},$$

where W_t is standard Brownian motion and $M_t = \max\{s, \sup_{0 \leq u \leq t} W_u\}$. In this case, that $W_t - M_t$ is Markovian offers no clue to any explicit solution. We conclude that the tractability of our problem is contingent upon the linearity of the payoff function as well. Gerber and Shiu addressed in [5] various pricing problems with linear payoff functions.

5. Derivation of the solution—revisiting the problem in a discrete model

The martingale proof provided in Section 3 is stunningly simple, yet unrevealing. To gain a deeper understanding of the structure of the solution, we resort to a seemingly crude numerical simulation, through which the path to the answer is painlessly unveiled.

In [6], we provided several ways of discretizing the market model, one of which we now describe. Let $(X_n, \varepsilon(n))$ represent the two-dimensional Markov processes of stock price at time n and the state of the market at time n. It then satisfies the recursion

$$(X_n, \varepsilon(n)) = \eta_n^{(\varepsilon(n), \varepsilon(n-1))}(X_{n-1}, \varepsilon(n-1)),$$

where the $\eta_n^{i,j}$ are i.i.d. random variables taking values u_j with probability $p_j(\chi_{i,1-j} + (-1)^{\chi_{i,1-j}} \mathrm{e}^{-\lambda_j h})$ and $1/u_j$ with probability $(1 - p_j)(\chi_{i,1-j} + (-1)^{\chi_{i,1-j}} \mathrm{e}^{-\lambda_j h})$ where $i, j \in \{0, 1\}$ and

$$\chi(i, j) = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

For fixed T, we divide the interval $[0, T]$ into N sub-intervals such that $T = Nh$ and let

$$u_i = \mathrm{e}^{\sigma_i \sqrt{h}}, \qquad p_i = \frac{\mu_i h + \sigma_i \sqrt{h} - \frac{1}{2}\sigma_i^2 h}{2\sigma_i \sqrt{h}}, \qquad d = \mathrm{e}^{-rh}, \qquad l_i = \mathrm{e}^{-\lambda_i h}.$$

In other words, $(X_n, \varepsilon(n))$ is a random walk taking values on the set $(u_0^m u_1^n, i)$ with $i \in \{0, 1\}$, $m, n \in \{0, \pm 1, \pm 2, \dots\}$.

Now consider the Markov chain $X = ((X_n, \varepsilon(n)), \mathcal{F}_n, P)$. We can restate the Russian option pricing problem by the following dynamic programming problem: let

$$W_0(x, s) = s,$$
$$Z_0(x, s) = s,$$
$$W_m(x, s) = \max \Big\{ W_{m-1}(x, s),$$

$$dp_0 l_0 W_{m-1}(u_0 x, \max\{u_0 x, s\}) + d l_0 q_0 W_{m-1}\left(\frac{x}{u_0}, \max\left\{\frac{x}{u_0}, s\right\}\right)$$

$$+ d(1 - l_0) p_1 W_{m-1}(u_1 x, \max\{u_1 x, s\})$$

$$+ d(1 - l_0) q_1 W_{m-1}\left(\frac{x}{u_1}, \max\left\{\frac{x}{u_1}, s\right\}\right) \Big\},$$

$$Z_m(x, s) = \max \Big\{ Z_{m-1}(x, s),$$

$$dp_1 l_1 Z_{m-1}(u_1 x, \max\{u_1 x, s\}) + d l_1 q_1 Z_{m-1}\left(\frac{x}{u_1}, \max\left\{\frac{x}{u_1}, s\right\}\right)$$

$$+ (1 - l_1) d p_0 W_{m-1}(u_0 x, \max\{u_0 x, s\})$$

$$+ (1 - l_1) d q_0 W_{m-1}\left(\frac{x}{u_0}, \max\left\{\frac{x}{u_0}, s\right\}\right) \Big\}.$$

Theorem 5.1. *When $r > \max(0, \mu_1, \mu_0)$, the solution to problem (2.2) is*

$$V_0(x, s) = \lim_{h \to 0} \lim_{n \to \infty} W_n(x, s),$$
$$V_1(x, s) = \lim_{h \to 0} \lim_{n \to \infty} Z_n(x, s).$$

The proof of Theorem 5.1 is by several steps:

Proof. The proof requires the following four propositions, an application of the principle of smooth fit, and reference to [6, Theorem 4.2].

Proposition 5.1. *We have*

$$W_m(xu, su) = u W_m(x, s),$$
$$Z_m(xu, su) = u Z_m(x, s).$$

This simple, yet critical, proposition requires no proof.

Proposition 5.2. *The functions V_0 and V_1 are increasing with respect to x, and $V_0(x, s) \geq s$, $V_1(x, s) \geq s$.*

Proposition 5.3. *In the common continuation region where V_0, V_1 are C^2 smooth, V_0, V_1 satisfy*

$$(r + \lambda_0) V_0(x, s) = x \mu_0 V_0'(x, s) + \tfrac{1}{2} x^2 \sigma_0^2 V_0''(x, s) + \lambda_0 V_1(x, s),$$
$$(r + \lambda_1) V_1(x, s) = x \mu_1 V_1'(x, s) + \tfrac{1}{2} x^2 \sigma_1^2 V_1''(x, s) + \lambda_1 V_0(x, s).$$
$$\tag{5.1}$$

Proof. Notice that if, for some $m > 1$,

$$dp_0 l_0 W_{m-1}(u_0 x, \max\{u_0 x, s\}) + dl_0 q_0 W_{m-1}\left(\frac{x}{u_0}, \max\left\{\frac{x}{u_0}, s\right\}\right)$$
$$+ d(1 - l_0) p_1 Z_{m-1}(u_1 x, \max\{u_1 x, s\})$$
$$+ d(1 - l_0) q_1 Z_{m-1}\left(\frac{x}{u_1}, \max\left\{\frac{x}{u_1}, s\right\}\right) \geq W_{m-1}(x, s),$$

$$dp_1 l_1 Z_{m-1}(u_1 x, \max\{u_1 x, s\}) + dl_1 q_1 Z_{m-1}\left(\frac{x}{u_1}, \max\left\{\frac{x}{u_1}, s\right\}\right)$$
$$+ (1 - l_1) dp_0 W_{m-1}(u_0 x, \max\{u_0 x, s\})$$
$$+ d(1 - l_1) q_0 W_{m-1}\left(\frac{x}{u_0}, \max\left\{\frac{x}{u_0}, s\right\}\right) \geq Z_{m-1}(x, s),$$

then it is true for all $n > m$. Therefore, if the limit function $V_0(x, s)$ is larger than s, then there exists some m such that

$$dp_0 l_0 W_{m-1}(u_0 x, \max\{u_0 x, s\}) + dl_0 q_0 W_{m-1}\left(\frac{x}{u_0}, \max\left\{\frac{x}{u_0}, s\right\}\right)$$
$$+ d(1 - l_0) p_1 Z_{m-1}(u_1 x, \max\{u_1 x, s\})$$
$$+ d(1 - l_0) q_1 Z_{m-1}\left(\frac{x}{u_1}, \max\left\{\frac{x}{u_1}, s\right\}\right) \geq W_{m-1}(x, s).$$

In consequence, we have

$$dp_0 l_0 V_0(u_0 x, \max\{u_0 x, s\}) + dl_0 q_0 V_0\left(\frac{x}{u_0}, \max\left\{\frac{x}{u_0}, s\right\}\right)$$
$$+ d(1 - l_0) p_1 V_1(u_1 x, \max(u_1 x, s))$$
$$+ d(1 - l_0) q_1 V_1\left(\frac{x}{u_1}, \max\left\{\frac{x}{u_1}, s\right\}\right) = V_0(x, s)$$

for all x, where $V_0(x, s) > s$. Similarly, for those (x, s) such that $V_1(x, s) > s$, we have

$$dp_1 l_1 V_1(u_1 x, \max\{u_1 x, s\}) + dl_1 q_1 V_1\left(\frac{x}{u_1}, \max\left\{\frac{x}{u_1}, s\right\}\right)$$
$$+ (1 - l_1) dp_0 V_0(u_0 x, \max\{u_0 x, s\})$$
$$+ d(1 - l_1) q_0 V_0\left(\frac{x}{u_0}, \max\left\{\frac{x}{u_0}, s\right\}\right) = V_1(x, s).$$

Since V_0 and V_1 are C^2 on the continuation region, by Taylor expansions,

$$V(e^{\sigma \sqrt{h}} x, s) = V(1 + \sigma \sqrt{h} + \tfrac{1}{2}\sigma^2 h + o(h)) x, s),$$
$$V((1 + \tfrac{1}{2}\sigma^2 h + \sigma \sqrt{h}) x, s) = V(x + \Delta x, s)$$
$$= V(x, s) + \Delta x V_x(x, s) + \tfrac{1}{2}(\Delta x)^2 V_{xx}(x, s) + o(\delta x)^2 + \cdots$$
$$= V(x, s) + (\sigma \sqrt{h} x + \tfrac{1}{2}\sigma^2 h x) V_x(x, s) + \tfrac{1}{2}\sigma^2 h V_{xx}(x, s),$$

and

$$V(x, s) = (1 - rh)p[V(x, s) + (\tfrac{1}{2}\sigma^2 h - \sigma\sqrt{h})x V_x(x, s) + \tfrac{1}{2}x^2\sigma^2 h V_{xx}(x, s)]$$
$$+ (1 - rh)q[V(x, s) + (\tfrac{1}{2}\sigma^2 h + \sigma\sqrt{h})x V_x(x, s) + \tfrac{1}{2}x^2\sigma^2 h V_{xx}(x, s)]$$
$$= (1 - rh)[V(x, s) + (q - p)\sigma\sqrt{h}x V_x(x, s) + \tfrac{1}{2}x^2\sigma^2 h V_{xx}(x, s)].$$

Recalling that

$$q_i - p_i = \frac{\mu_i h - \tfrac{1}{2}\sigma_i^2 h}{\sigma_i \sqrt{h}},$$

we obtain, on the continuation region, the following equations as anticipated:

$$(r + \lambda_0)V_0(x, s) = x\mu_0 V_0'(x, s) + \tfrac{1}{2}x^2\sigma_0^2 V_0''(x, s) + \lambda_0 V_1(x, s),$$
$$(r + \lambda_1)V_1(x, s) = x\mu_1 V_1'(x, s) + \tfrac{1}{2}x^2\sigma_1^2 V_1''(x, s) + \lambda_1 V_0(x, s).$$

Proposition 5.4. *If $r > \max\{0, \mu_1, \mu_0\}$, and μ_i, λ_i and σ_1 are positive, then there exist $0 < c_0, c_1 < 1$ such that $V_i(x, s) = s$, when $x < c_i s$, and $V_1(x, s) > s$ when $c_i s < x < s$, $(i = 0, 1)$.*

Proof. We give a proof by contradiction. Evidently, $V_0(x, s) \geq s$ and $V_1(x, s) \geq s$. If, for all $0 < x \leq s$, $V_0(x, s) > s$ and $V_1(x, s) > s$, then by Proposition 5.3, in the continuation region V_0 and V_1 are of the form:

$$V_0(x, s) = \sum_{i=1}^{4} a_i(s)x^{\beta_i}, \qquad V_1(x, s) = \sum_{i=1}^{4} b_i(s)x^{\beta_i},$$

where the β_i satisfy (2.3).

Moreover, it is not hard to verify that if

$$V_0(x, 1) = \sum_{i=1}^{4} a_i x^{\beta_i}, \qquad V_1(x, 1) = \sum_{i=1}^{4} b_i x^{\beta_i},$$

are solutions to (5.1) at $s = 1$, then so are $s V_0(x/s, 1)$, $s V_1(x/s, 1)$. Hence, $a_i(s)$ is actually of the form $a_i s^{1-\beta_i}$ and $b_i(s) = b_i s^{1-\beta_i}$. We conclude that if the boundaries between the continuation region and the stopping region exist, they should be straight lines $x = c_i s$. Furthermore,

$$V_0(x, s) \geq s \quad \text{and} \quad \frac{\partial V_0}{\partial x} \geq 0,$$

therefore, $a_1 a_2 \neq 0$. Otherwise, when $x \to 0$, $V_0 \to 0$; but since $V_0 \geq s$, this would be a contradiction.

Now suppose that $a_1 \neq 0$. Since

$$V_0(x, s) = \sum_{i=1}^{4} a_i x^{\beta_i},$$

when $x \to 0$, the leading term is x^{β_1} and $V_0 > s$, which implies that a_1 is positive. However, since $V' \geq 0$ almost everywhere, the same argument applied to $V_0'(x, s) = \sum_{i=1}^{4} a_i \beta_i x^{\beta_i}$

indicates that $a_1 < 0$, a contradiction. Therefore, at least one of V_0 and V_1 should be equal to s when x is close to 0. In other words, at least one of c_0 and c_1 exists.

Now suppose that c_0 exists, but c_1 does not, i.e. when $0 < x < c_0 s$, $V_0(x, s) = s$, $V_1(x, s) > s$, then

$$dp_1 l_1 V_1(u_1 x, \max\{u_1 x, s\}) + dl_1 q_1 V_1 \left(\frac{x}{u_1}, \max\left\{ \frac{x}{u_1}, s \right\} \right)$$
$$+ (1 - l_1) dp_0 V_0(u_0 x, \max\{u_0 x, s\}) + d(1 - l_1) q_0 V_0 \left(\frac{x}{u_0}, \max\left\{ \frac{x}{u_0}, s \right\} \right)$$
$$= V_1(x, s).$$

By a Taylor expansion,

$$(r + \lambda_1) V_1 = x \mu_1 V_1' + \tfrac{1}{2} x^2 \sigma_1^2 V_1'' + \lambda_1 s,$$

which implies that

$$V_1(x, s) = A_1 x^{\gamma_1} + A_2 x^{\gamma_2} + \frac{\lambda_1}{r + \lambda_1},$$

where γ_1, γ_2 are the solution of the corresponding characteristic equation such that

$$(r + \lambda_1) = \gamma \mu_1 + \tfrac{1}{2} \gamma (\gamma - 1) \sigma_1^2.$$

Hence, $\gamma_1 \gamma_2 < 0$. Repeating the same argument, we conclude that it is impossible to have $V_1(x, s) \geq s$ and $\partial V_1 / \partial x \geq 0$. Therefore, c_1 also exists. So the two boundaries $x = c_0 s$ and $x = c_1 s$ divide Ω into three parts: what we have termed the stopping region, the common continuation region and the transient region.

Continuing the proof of Theorem 5.1, we now apply the principle of smooth fit (see [14] for details). Consider the transient region

$$\Omega_1 = \{(x, s) \mid c_1 s \leq x \leq c_0 s\},$$

where $V_0(x, s) \equiv s$. The principle of smooth fit implies that

$$\left. \frac{\partial V_1(x, s)}{\partial x} \right|_{x = c_1 s} = 0, \qquad \left. V_1(x, s) \right|_{x = c_1 s} = s,$$

and

$$(r + \lambda_1) V_1 = x \mu_1 V_1' + \tfrac{1}{2} x^2 \sigma_1^2 V_1'' + \lambda_1 s.$$

Thus, we can explicitly derive V_1 in Ω_1 in terms of c_1:

$$V_1(x, s) = \frac{rs}{(r + \lambda_1)(\gamma_2 - \gamma_1)} \left[\gamma_2 \left(\frac{x}{c_1 s} \right)^{\gamma_1} - \gamma_1 \left(\frac{x}{c_1 s} \right)^{\gamma_2} \right] + \frac{\lambda_1 s}{r + \lambda_1}, \qquad (5.2)$$

where γ_1, γ_2 satisfy

$$(r + \lambda_1) = \gamma \mu_1 + \tfrac{1}{2} \gamma (\gamma - 1) \sigma_1^2;$$

hence,

$$\gamma_1 = \frac{-(\mu_1 - \frac{1}{2}\sigma_1^2) - \sqrt{(\mu_1 - \frac{1}{2}\sigma_1^2)^2 + 2\sigma_1^2(r + \lambda_1)}}{\sigma_1^2},$$

$$\gamma_2 = \frac{-(\mu_1 - \frac{1}{2}\sigma_1^2) + \sqrt{(\mu_1 - \frac{1}{2}\sigma_1^2)^2 + 2\sigma_1^2(r + \lambda_1)}}{\sigma_1^2}.$$

Therefore,

$$V_1(x, s)|_{x=c_0 s} = \frac{rs}{(r + \lambda_1)(\gamma_2 - \gamma_1)} \left(\gamma_2 \left(\frac{c_0}{c_1} \right)^{\gamma_1} - \gamma_1 \left(\frac{c_0}{c_1} \right)^{\gamma_2} \right) + \frac{\lambda_1 s}{r + \lambda_1},$$

$$\left. \frac{\partial V_1}{\partial x} \right|_{x=c_0 s} = \frac{r\gamma_1\gamma_2}{c_0(r + \lambda_1)(\gamma_2 - \gamma_1)} \left(\left(\frac{c_0}{c_1} \right)^{\gamma_1} - \left(\frac{c_0}{c_1} \right)^{\gamma_2} \right),$$

$$V_0(x, s)|_{x=c_0 s} = s,$$

$$\left. \frac{\partial V_0}{\partial x} \right|_{x=c_0 s} = 0.$$

With respect to the boundary conditions on $x = s$, it is conceivable (see also [5]) that if the current stock price x is very close to s, almost surely X_t will exceed $se^{\delta h}$ before it is optimal for the payee to demand payment. Therefore, $V(x, s)$ does not depend on the precise value of s, which implies that

$$\left. \frac{\partial V_0}{\partial s}(x, s) \right|_{x=s} = 0$$

and

$$\left. \frac{\partial V_1}{\partial s}(x, s) \right|_{x=s} = 0.$$

Combining this with Propositions 5.1–5.4 and [6, Theorem 4.2], we reach the conclusion in Theorem 5.1.

Note. The last two boundary conditions on $x = s$ effectively guarantee that Y_t is a submartingale.

6. Arbitrage-free price for Russian options

In [6], we proposed one way to complete the market model. At each time t, there is a market for 'ticket', a security that pays one unit of account (say, dollar) at the next time $\tau(t) = \inf\{u > t : \varepsilon(u) \neq \varepsilon(t)\}$ that the Markov chain ε changes state. That contract then becomes worthless (has no future dividends), and a new contract is issued that pays at the next change of state, and so on.

It is not hard to see that, under natural pricing, this will complete the market, and provide unique arbitrage-free prices to the Russian (and European, and any other) options on the underlying risky asset (see [2], [7] and [8]). Moreover, the Russian option price will be given by our solution in Theorem 2.1 if we re-interpret μ_i by supposing that $\mu(i) = r - \delta(i)$, where $\delta(i)$ is the rate of continual dividend payout, per dollar of underlying price.

Acknowledgements

I am grateful to Larry Shepp for suggesting this project to me. Lester Dubins, Darrell Duffie, Richard Gundy, Michael Harrison, Chris Heyde and Dan Ocone provided valuable comments and advice. Thanks to the referee for kindly pointing out a missing step in the original proof of the theorem. I gratefully acknowledge support from DIMACS and the hospitality of the University of California at Berkeley.

References

[1] CHERNOFF, H. (1961). Sequential tests for the mean of a normal distribution. In *Proc. 4th. Berkeley Symp. Statist. Prob.* Vol. 1. University of California Press, Berkeley, pp. 79–91.

[2] DELBAEN, F. AND SCHACHERMAYER, W. (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300,** 463–520.

[3] DUFFIE, D. AND HARRISON, M. (1993). Arbitrage pricing of perpetual lookback options. *Ann. Appl. Prob.* **3,** 641–651.

[4] ELLIOTT, R. (1982). *Stochastic calculus and applications*. Springer, Berlin.

[5] GERBER, H. U. AND SHIU, E. S. W. (1996). Martingale approach to pricing perpetual American options on two stocks. *Math. Finance* **6,** 303–322.

[6] GUO, X. (2001). Information and option pricing. *Quant. Finance* **1,** 38–44.

[7] HARRISON, M. AND KREPS, D. (1979). Martingales and arbitrage in multiperiod securities markets. *J. Econom. Theory* **20,** 381–408.

[8] HARRISON, M. AND PLISKA, S. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stoch. Proc. Appl.* **11,** 215–260.

[9] KARATZAS, I. AND SHREVE, S. (1998). *Brownian Motion and Stochastic Calculus*. Springer, Berlin.

[10] REVUZ, D. AND YOR, M. (1991). *Continuous Martingale and Brownian Motion*. Springer, Berlin.

[11] SHEPP, L. A. (1999). A model for stock price fluctuations based on information. Preprint, Rutgers University.

[12] SHEPP, L. A. AND SHIRYAEV, A. N. (1993). The Russian option: reduced regret. *Ann. Appl. Prob.* **3,** 631–640.

[13] SHEPP, L. A. AND SHIRYAEV, A. N. (1994). A new look at pricing of the 'Russian options'. *Theory Prob. Appl.* **39,** 103–119.

[14] VAN MOERBEKE, P. L. J. (1976). On optimal stopping and free boundary problems. *Arch. Rational Mechanics Anal.* **60,** 101–148.