Root's and Rost's Embeddings: Construction, Optimality and Applications to Variance Options

submitted by

Jiajie Wang

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

2011

SUMMARY

Root's solution (Root [1969]) to the Skorokhod embedding problem can be described as the first hitting time of a space-time process (X_t, t) on a so-called barrier, characterised by certain properties, such that the stopped underlying process X has a given distribution. Recent work of Dupire [2005] and Carr and Lee [2010] has highlighted the importance of understanding the Root's solution for the model-independent hedging of variance options.

We consider the problem of finding Root's solutions when the underlying process is a time-homogeneous diffusion with a given initial distribution in one dimension. We are interested in constructing Root's solution by partial differential equations. We begin by showing that, under some mild conditions, constructing Root's solution is equivalent to solving a specialized parabolic free boundary problem in the case where the underlying process is a Brownian motion starting at 0. This result is then extended to time-homogeneous diffusions. Replacing some conditions needed in the free boundary construction, we then also consider the construction of Root's solutions by variational inequalities. Finally we consider the optimality and applications of Root's solutions. Unlike the existing proof of optimality (Rost [1976]), which relies on potential theory, an alternative proof is given by finding a path-wise inequality which has an important application for the construction of subhedging strategies in the financial context. In addition, we also consider these questions, construction and optimality, for Rost's solution, which is also known as the reverse of the Root's solution.

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor, Alexander M. G. Cox, for introducing me to this interesting topic, for his patient guidance and continuous support of my Ph.D study, and for his inspiration and enthusiasm in our research. This thesis also owes much to his thoughtful and helpful comments.

I further express my gratitude to all the members in the Prob-L@B (Probability Laboratory at Bath) for the stimulating working environment. The weekly seminars introduce me to a wide range of different topics in the frontier of the study in probability, as well as the graduate courses provide me with a deep understanding of some special subjects.

Thanks to everybody – staff, students and secretaries – in the Department of Mathematical Sciences for making it a friendly place to research. Thanks to Melina Freitag for preparing this LaTeX thesis template. I would also like to thank all my friends who make my life fun in last three years.

My deepest appreciation goes to my parents, for everything they have done for me, for their faith in me and allowing me to be as ambitious as I wanted. They make this thesis possible.

Contents

List of Figures

Chapter 1

Introduction

After Norbert Weiner showed its existence, Brownian motion, the physical phenomenon first observed and described by Robert Brown, has become an important object of study throughout the pure and applied mathematical sciences.

In this thesis, we are concerned with a particular question related to the study of Brownian motion, and moreover, the diffusion processes driven by Brownian motion. The question is the so called *Skorokhod embedding problem* named for the Ukrainian mathematician A. V. Skorokhod who first posed the question (Skorokhod [1961], and English translation Skorokhod [1965]): Suppose W is a one-dimensional Brownian motion and μ is a distribution on $\mathbb{R}$. Can we find a stopping time T such that W_T has distribution μ?

1.1 A Brief Introduction to The Problem

We start with the solution to the problem given by Skorokhod immediately after he posed it. His solution relies on a randomization external to W, and a rigorous statement of his solution can be found in Freedman [1971]. For Brownian motion W and a given centred probability distribution μ on $\mathbb{R}$, define $\rho : \mathbb{R}_+ \to \mathbb{R}$,

$$\lambda \;\mapsto\; -\inf\left\{ y \in \mathbb{R} : \int_{\mathbb{R}} x \mathbf{1}_{(-\infty,y]\cup[\lambda,\infty)} \mu(\mathrm{d}x) \;\leq\; 0 \right\},$$

Let R be an independent random variable such that for all $x \in \mathbb{R}$,

$$\mathbb{P}[R \leq x] \;=\; \int_{\infty}^{x} \left(1 + \frac{y}{\rho(y)}\right) \mu(\mathrm{d}y).$$

Then, the stopping time defined by

$$T \;=\; \inf\{t \geq 0 : \; W_t \notin (-\rho(R), R)\}$$

satisfies $W_T \sim \mu$. Here, the random variable R is the so called *external randomization* mentioned above. Moreover, we have that the process $W^T := \{W_{T\wedge t}\}_{t\geq 0}$ is uniformly integrable. By Itô's formula, $\{W_t^2 - t\}_{t\geq 0}$ is a martingale, and then by the optional stopping theorem, $\mathbb{E}[\tau] = \mathbb{E}[W_\tau^2]$ for any stopping time τ with finite expectation. A stronger result can be found in Sawyer [1974][see (19), Section 4]: if W^τ is uniformly integrable, then for all $p > 1$, there exist constants a_p and A_p such that

$$a_p \mathbb{E}[|W_\tau|^{2p}] \;\leq\; \mathbb{E}[\tau^p] \;\leq\; A_p \mathbb{E}[|W_\tau|^{2p}].$$

Therefore, $\mathbb{E}[\tau] = \mathbb{E}[W_\tau^2] < \infty$ for all centred target distribution μ with finite variance.

We introduce another solution, known as Doob's solution, to show that without some restriction, the Skorokhod embedding problem is trivial. Denote the cumulative distribution functions of the target distribution μ and the standard normal distribution $N(0,1)$ by F and Φ respectively, we then define

$$S \;=\; \inf\left\{t \geq 1 : F(W_t) \;=\; \Phi(W_1)\right\},$$

then $W_S \sim \mu$, and moreover $\mathbb{E}[S] = \infty$ unless $\mu = N(0,1)$ (one can find more details in Rogers and Williams [2000a][Section I.7]).

Compared with T, the construction of S is more straightforward. However, its deficiency is also very clear: $\mathbb{E}[S] < \infty$ if and only if $\mu = N(0,1)$, but by contrast $\mathbb{E}[T] = \mathbb{E}[W_T^2] < \infty$ for all centred target distribution with finite variance. As a usual criterion for the choice of stopping times, it is expected that the target distribution can be realized as soon as possible. Therefore, usually, we regard T as a "better" embedding than S. In fact, most solutions to the embedding problem proposed after Skorokhod are with the restriction of uniform integrability.

Besides standard Brownian motion, the same problem for more general processes has been treated. In more general case we first need to confirm the existence of embeddings. Suppose X is a Markov process with the initial probability distribution ν and the transition semi-group $\left(P_t^X; t \geq 0\right)$. Rost [1971] showed that there exists an embedding for another probability distribution μ if and only if

$$\nu U^X \ \geq \ \mu U^X,$$

where $U^X = \int_0^\infty P_t^X \mathrm{d}t$ is the potential kernel of X and νU^x could be seen as the occupation measure (on $\mathbb{R}$) for X along its trajectories where $X_0 \sim \nu$ (Obłój [2004], Section 2.2). In other words, this condition can be written as: for any positive, continuous and compactly supported function f,

$$\int_0^\infty \mathbb{E}^\nu[f(X_t)]\mathrm{d}t \ \geq \ \int_0^\infty \mathbb{E}^\mu[f(X_t)]\mathrm{d}t.$$

Now we reformulate the Skorokhod embedding problem as following:

Skorokhod Embedding Problem. *Suppose that X is a Markov process with the initial distribution ν. For some probability distribution μ, find a stopping time T such that $X_T \sim \mu$.*

Throughout this thesis, a stopping time T is written as *UI stopping time* for short if X^T is uniformly integrable.

Now the question is that, among more than one available UI embeddings for a distribution μ, which one is "better"? The answer to this question, obviously, depends on what we mean when we are talking about "better" or "worse" embeddings. The usual criterion mentioned above, the size of $\mathbb{E}[T]$, is trivial here, since there is no difference among all UI embeddings for a fixed target distribution. Depending on different applications, many criterions and corresponding optimal embeddings are posed. There have been a large number of works dedicated to the development of different solutions to the Skorokhod embedding problem and the study of their properties, especially optimality given different criterions, for examples, Dubins [1968], Root [1969], Rost [1971], Monroe [1972b], Chacon and Walsh [1976], Azéma and Yor [1979a], Vallois [1983], Perkins [1986], etc. It is impossible to include all works on the subject in such a short list, and we refer a curious reader to a more detailed survey paper, Obłój [2004].

1.2 Applications of Skorokhod Embeddings

The Skorokhod embedding problem is remarkable because it has been shown to be helpful in the study of many other subjects. We are interested in its application in finance. The use of Skorokhod embedding techniques to solve model-independent (or robust) hedging problems in finance can be traced back to the paper Hobson [1998a]. We present here some typical examples of its application.

Example 1.2.1 (Azéma and Yor [1979a])**.** The Azéma-Yor embedding is the first entrance of the joint process $(W, \overline{W})$, where $\overline{W}_t = \sup_{s \le t} W_s$, into a domain D. For any centred probability distribution μ, to find the domain, we define

$$\Psi(x) \;:=\; \frac{1}{\mu([x,\infty))} \int_{[x,\infty)} r\,\mu(\mathrm{d}r) \quad \text{and} \quad B := \{(x,y) : \, y \ge \Psi(x)\}\,.$$

We denote the hitting time of $(W, \overline{W})$ on B by T_{AY}, i.e.

$$T_{AY} \;=\; \inf\left\{t \ge 0 : \, \overline{W}_t \ge \Psi(W_t)\right\}.$$

Then T_{AY} is a UI embedding of μ. The Azéma-Yor solution is characterized by the optimality property that it maximises the law of the supreme process among the class of UI embeddings, that is, given a UI embedding of μ, denoted by ρ, we have $\mathbb{P}[\overline{W}_\rho \ge x] \le \mathbb{P}[\overline{W}_{T_{AY}} \ge x]$ for all $x \in \mathbb{R}_+$. This result was shown by Azéma and Yor [1979b] immediately after they proposed the solution. Hobson [1998b] argued the property by the discovery of a path-wise inequality, as a by-product, the upper bound is calculated explicitly. We will try to explain briefly the result.

Given an increasing function F with derivative f, our aim is to maximise $\mathbb{E}[F(\overline{W}_\rho)]$ among all UI embeddings of μ. For simplicity, we assume that the inverse of Ψ exists and is denoted by ψ, and both Ψ and ψ are continuous, this assumption also implies $\overline{W}_{T_{AY}} = \Psi(W_{T_{AY}})$ where T_{AY} is the Azéma-Yor embedding of μ. Define $g(y) := f(y)/(y - \psi(y))$ and $G(x,y) := \int_0^y g(r)(r-x)\mathrm{d}r$, then one can find that $\big(G(W_t, \overline{W}_t);\, t \ge 0\big)$ is a martingale, known as the Azéma martingale, and $F(y) = G(x,y) + H(x,y)$ always, where

$$H(x,y) \;:=\; F(0) + \int_0^y f(r)\frac{x-\psi(r)}{r-\psi(r)}\mathrm{d}r \;\le\; F(0) + \int_0^{\Psi(x)} f(r)\frac{x-\psi(r)}{r-\psi(r)}\mathrm{d}r,$$

and the equality in the inequality holds if and only if $y = \Psi(x)$. We denote the expression on the right-hand side by $\widetilde{H}(x)$. Since $\overline{W}_{T_{AY}} = \Psi(W_{T_{AY}})$, we have that

$$\mathbb{E}[F(T_{AY})] = \mathbb{E}[G(W_{T_{AY}}, \overline{W}_{T_{AY}}) + \widetilde{H}(W_{T_{AY}})] = \mathbb{E}[\widetilde{H}(W_{T_{AY}})].$$

Therefore, for any ρ in $\mathcal{T}(\mu)$, the collection of all UI embedding of μ, we obtain the path-wise inequality, and then take expectations,

$$F(\overline{W}_\rho) \;\leq\; G(W_\rho, \overline{W}_\rho) + \widetilde{H}(W_\rho) \quad\Longrightarrow\quad \mathbb{E}[F(\overline{W}_\rho)] \;\leq\; \mathbb{E}[\widetilde{H}(W_{T_{AY}})] \;=\; \mathbb{E}[F(\overline{W}_{T_{AY}})].$$

Since the result can be extended to an arbitrary increasing function F, let $F = \mathbf{1}_{[x,\infty)}$ and we have

$$\sup_{\rho\in\mathcal{T}(\mu)} \mathbb{P}[\overline{W}_\rho \geq x] \;\leq\; \mathbb{P}[\overline{W}_{T_{AY}} \geq x].$$

Now given $T > 0$, by the time change

$$M_t = W_{(t/T-t)\wedge\rho} \tag{1.2.1}$$

where $\rho \in \mathcal{T}(\mu)$, $M := (M_t;\, t \geq 0)$ is a martingale null at 0 and $M_T \sim \mu$. In options pricing theory, the martingale M can be regarded as the price process whose marginal distribution at T coincides with μ. Let $F(s) = (s-K)_+$ and $\rho = T_{AY}$, because of the optimality of T_{AY} on the maximum process, the path-wise inequality obtained above can be applied as the super-replication of a *look-back call option with fixed strike* [1] and the explicit upper bound $\int \widetilde{H}(x)\mu(\mathrm{d}x)$ is regarded as an upper bound of the price of the option.

In addition, using the Azéma-Yor solution, Dubins and Schwarz [1988] first solved the optimal stopping problem of the form $\sup_\rho \mathbb{E}\left[\phi(\overline{W}_\rho) - \int_0^\rho c(W_s)\mathrm{d}s\right]$ given the simplest case where $\phi(x) = x$ and c constant. More general cases were treated by Peskir [1998, 1999], Meilijson [2003], Obłój [2007]. The authors showed that the solutions to the optimal stopping problem is the Azéma-Yor embeddings of μ determined by ϕ and c.

Example 1.2.2 (Perkins [1986])**.** The second example we introduce is the Perkins solution, T_P, which is defined as

$$T_P \;=\; \inf\left\{t > 0 :\, W_t \notin (-\gamma_+(\overline{W}_t), \gamma_-(\underline{W}_t))\right\},$$

[1] Roughly, look-back call options with fixed strike are the options with the payoff $(S_{\max} - K)_+$, where $S_{\max}$ is the highest price of the underlying asset before the maturity and K is the strike price. For more details in look-back options, we refer a reader to Hobson [1998a]

where $\underline{W}_t = -\inf_{s\le t} W_s$, and γ_+, γ_- are two functions (see the original work for more details; also see Cox and Hobson [2004] as a generalization). This embedding has the property that it simultaneously minimises the law of the maximum process and maximises the law of the minimum process: for $\rho \in \mathcal{T}(\mu)$ and $\lambda > 0$,

$$\mathbb{P}[\overline{W}_{T_P} \ge \lambda] \;\le\; \mathbb{P}[\overline{W}_\rho \ge \lambda]; \qquad \mathbb{P}[\underline{W}_{T_P} \ge \lambda] \;\le\; \mathbb{P}[\underline{W}_\rho \ge \lambda].$$

A similar construction is given by Hobson and Pedersen [2002]. In their work the authors give the greatest lower bound on the law of the maximum process and their explicit embedding, under the time change (1.2.1) with $\rho = T_P$, can be applied to the robust hedging of a forward start digital option.

Example 1.2.3 (Vallois [1983])**.** The Vallois solution, as our last example, can be described as the first entrance of the joint process (W_t, L_t) into a domain, where L is the local time of the underlying process at 0: there exist two non-negative, non-increasing functions h_+ and h_- such that

$$T_V \;=\; \inf\{t \ge 0 : \, W_t \notin (-h_-(L_t), h_+(L_t))\}.$$

By developing a class of path-wise inequalities, Cox et al. [2008] verified that Vallois' embedding maximises $\mathbb{E}[L_\rho - K]_+$ among all UI embeddings. In addition, one can interpret the path-wise inequalities as super-replication strategies of options written on the local time. Moreover, similar to the Azéma-Yor solution, the authors solve optimal stopping problems of the form $\sup_\rho \mathbb{E}\left[\phi(L_\rho) - \int_0^\rho c(W_s)\mathrm{d}s\right]$

Other than the examples and literature mentioned above, More recent results in this direction include Cox and Obłój [2011a] and Cox and Obłój [2011b]. We also refer the reader to Hobson [2009] which is a comprehensive survey of the literature on the Skorokhod embedding problem with a specific emphasis on applications in mathematical finance. We also refer the reader to Obłój [2004] for other applications of the solutions to the Skorokhod embedding problems.

1.3 An Overview of the Thesis

The subsequent content in this thesis is concerned with Root's and Rost's solutions (normally, known as *Root's barrier* and *reversed barrier*) to the Skorokhod embedding problem. As is well known, Root's barrier is remarkable since it minimises the variance of the stopping time among all UI embeddings. Now, given a target distribution μ, can we find Root's barrier for μ explicitly in practice? Unfortunately, even for very simple target distribution μ, it seems to be very difficult.

In this thesis, we will show, under some mild assumptions, that finding Root's barrier is equivalent to finding a solution to a specified free boundary problem or a specified variational inequality. The equivalence provides us with a possible method to compute Root's solution.

Moreover, the original proof of the optimality by Rost [1976] relies heavily on notions from potential theory. In this thesis, using probabilistic techniques, we will develop a 'path-wise inequality' which encodes the optimality. We then interpret such an inequality mathematically as a hedging strategy for a variance option.

We also consider these questions in relation to Rost's barrier which can be regarded as the reverse of Root's barrier and which maximises the variance of the stopping time among all UI embeddings.

1.3.1 Chapter 2: Connection to Free Boundary Problems

We begin with a brief description of *Root's stopping time* which was first introduced by Root [1969]. In Root's original work, his embedding can be described as the first entrance of the joint process $((W_t, t); t \geq 0)$ into a closed set B with the property that, roughly, $(x, t) \in B$ implies $(x, s) \in B$ for all $s > t$. Loynes [1970] showed that it can be equivalently defined as

$$B \;=\; \{\, (x,t) : t \geq R(x)\} \,,$$

where $R : \mathbb{R} \to \overline{R}$ is a lower semi-continuous function. In their works, the closed set $B = \{(x, t) : t \geq R(x)\}$ is called Root's barrier. In our work, for convenience, we always call the open set $D = \{(x, t) : 0 < t < R(x)\}$ *Root's domain.* After that, some related results (mainly given by Loynes [1970], Rost [1971, 1976], Chacon [1977]) will also be introduced.

Beginning the discussion of the connection to free boundary problem, we assume the underlying process of the Skorokhod embedding problem is a Brownian motion with the initial distribution δ_0. Suppose that τ_D is Root's solution to the Skorokhod embedding problem of a probability distribution μ, we then show the potential process of the stopped distributions,

$$u(x,t) := -\mathbb{E}[|x - W_{t\wedge\tau_D}|],$$

solves a second-order parabolic partial differential equation initial-boundary value problem where D is the Root's domain. For $t > 0$

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{1}{2}\dfrac{\partial^2 u}{\partial x^2}, & \text{for } (x,t) \in D; \\ u(x,0) = -|x|, & \text{for } x \in \mathbb{R}; \\ u(x,t) = -\mathrm{U}\mu(x) := -\displaystyle\int_{\mathbb{R}} |x-y|\mu(\mathrm{d}y) & \text{for } (x,t) \notin D. \end{cases}$$

Conversely, consider this PDE system without *a priori* knowledge of Root's barrier. Now the system becomes a *free boundary problem,* since both the function u and the domain D are unknown. Our question is, if a couple (u, D) solves this free boundary problem, can we claim that D is the Root's domain of μ? By some standard results in the potential theory, the key point of this question essentially is the relation between the second-order derivative of u and the stopping density of Brownian motion related to the domain D. To find the relation, we pose an additional restriction on the free boundary problem, which is concerned with the limit of $\partial^2 u/\partial x^2$ on the boundary of D. We call the restriction *'vanishing second derivative on the boundary'*. With it, we can show that D is Root's domain for μ.

Now our interest is in the restriction of vanishing second derivative. Is it reasonable? We will see that this restriction is satisfied by a fairly large class of Root's barriers.

Finally, we will extend all the results to time-homogeneous diffusions and a general initial distribution ν.

1.3.2 Chapter 3: Connection to Variational Inequalities

We are interested in the variational inequalities of the form

$$\begin{cases} -\dfrac{\partial v}{\partial t} + Av - f \ \leq \ 0; \\ \Big(-\dfrac{\partial v}{\partial t} + Av - f\Big)(u - \psi) \ = \ 0; \\ u - \psi \ \leq \ 0. \end{cases}$$

where A is a suitable differential operator. By Bensoussan and Lions [1982], there is a unique solution (in suitable spaces) to a *strong form* of the variational inequality.

In this chapter, we treat the Skorokhod embedding problem of μ for the case that the diffusion process X satisfies $\mathrm{d}X_t = \sigma(X_t)\mathrm{d}W_t$, and the initial distribution is ν.

First, we assume the diffusion coefficient σ is bounded by two positive numbers. We show, using some results from Chapter 2, the potential process of the stopping distributions generated by Root's barrier of μ is a solution to the strong variational inequality with the appropriate setting of parameters. Therefore, by the uniqueness result, finding Root's barrier for μ is equivalent to finding the solution to the specified variational inequality.

Then we consider the case that the underlying process is a geometric Brownian motion which does not satisfy the condition that σ is bounded by two positive numbers. Changing the spatial variable in the potential process, it is shown that we still can find Root's barriers by solving the variational inequalities.

After that, we introduce Rost's barrier which can be regarded as the reverse of Root's barrier. Then, given the existence of Rost's barrier for μ, we show that it can be found by similar variational inequalities.

1.3.3 Chapter 4: Optimality and Applications in Finance

Since the original proof of the optimality given by Rost [1976] relies on notions from potential theory, in this chapter, we give an alternative proof of this result using probabilistic techniques and we shall be able to give a *'path-wise inequality'*: given μ and a convex function F, we can find a submartingale G_t such that $G_{t\wedge\tau_D}$ is a uniformly integrable martingale when τ_D is Root's embedding of μ, and such that $F(t)-G_t \geq H(X_t)$ for some function $H : \mathbb{R} \to \mathbb{R}$. It follows that Root's embedding minimises $\mathbb{E}[F(\tau)]$ among all embeddings for μ. We also treat the optimality of Rost's embedding. In a similar manner as in Root's case, we show Rost's embedding maximises $\mathbb{E}[F(\tau)]$ among all embeddings of μ.

After that, we apply the results observed above in the study of *variance options* which allow one to speculate on or hedge risks associated with the volatility of some underlying assets. Consider a (discounted) asset which has dynamics under the risk-neutral measure:

$$\frac{\mathrm{d}S_t}{S_t} = \sigma_t \mathrm{d}W_t,$$

where the process σ_t is not necessarily known. Our aim is to sub-replicate the variance option with the payoff $F(\langle \ln S \rangle_T)$. Given the law of S_T, denoted by μ, we will find the lower bound of the price of the variance option by the optimality of Root's embedding, and give the subhedging strategy by the path-wise inequality.

Chapter 2

Connecting Root's Barriers and Free Boundary Problems

In this chapter, we connect the embedding first observed in Root [1969] to a particular free boundary problem, which has been first suggested by Bruno Dupire.

In the original work of Dupire [2005], the author observed that the potentials of the distribution of a Brownian motion killed by a Root's barrier satisfies a free boundary problem. But, to make sure that we can generate Root's barrier from a free boundary problem, the more important result we need is the uniqueness of solutions to the free boundary problems. For example, in McConnell [1991], by parabolic potential theory, the author constructed the solutions to a class of two-sided Stefan problems from Green functions, and then some uniqueness result was given. As an important application, his results yield an independent construction of the solution (proposed by Chacon [1985]) to the Skorokhod embedding problem. In this work, we will reformulate the free boundary problem suggested by Dupire, and then show the uniqueness result.

2.1 Introduction and Preliminaries

We begin with some crucial definitions and results which will be used throughout this thesis without further explanation. As usual, $W = \{W_t\}_{t \geq 0}$ denotes a one-dimensional Brownian motion, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{F}$ is the completed natural filtration generated by the Brownian motion. For a random variable X, its distribution is denoted by $\mathcal{L}(X)$. The support of a measure μ is denoted by $supp(\mu)$. Finally, $\mu_n \Rightarrow \mu$ signifies weak convergence of measures.

Consider a time-homogeneous diffusion X which satisfies the stochastic differential equation

$$\mathrm{d}X_t = \sigma(X_t)\mathrm{d}W_t. \tag{2.1.1}$$

Given some probability distribution μ and ν on $\mathbb{R}$, $\mathcal{T}(\sigma, \nu, \mu)$ denotes all the UI stopping times solving the Skorokhod embedding problem stated in Section 1.1 with regard to the initial distribution ν, the target distribution μ and the underlying process X determined by σ. That is

$$\mathcal{T}(\sigma, \nu, \mu) = \{\tau \text{ is a UI stopping time: } \mathbb{P}^{\nu}[X_\tau \in \mathrm{d}x] = \mu(\mathrm{d}x)\}. \tag{2.1.2}$$

And we will drop σ when $\sigma(x) \equiv 1$, i.e., $X = W$, and drop ν when $\nu = \delta_0$, the Dirac point mass at 0.

2.1.1 Features of Root's Solution

Our interest is in Root's solution, also known as Root's barrier, to the Skorokhod embedding problem. Root [1969] first proposed the solution. In his work, Root showed that if W is a Brownian motion with $W_0 = 0$, and μ is a centred probability distribution with finite variance, then, considering the *co-space-time Brownian motion*, (W_t, t), there exists a stopping time $\tau \in \mathcal{T}(\mu)$, which is the first hitting time of a *barrier*, which is defined as:

Definition 2.1.1 (Root's Barrier)**.** A closed subset B of $[-\infty, +\infty] \times [0, +\infty]$ is a *barrier* if

(i). $(x, +\infty) \in B$ for all $x \in [-\infty, +\infty]$;

(ii). $(\pm\infty, t) \in B$ for all $t \in [0, \infty]$;

(iii). if $(x, t) \in B$ then $(x, s) \in B$ whenever $s > t$.

After Root's original work, there are two important papers concerning the construction of barriers. In Loynes [1970], the author proved a number of results related to the barriers. From our perspective, the most important are, firstly, that the barrier B can be written as:

$$B = \{(x, t) : t \geq R(x)\},$$

where *the barrier function* $R : \mathbb{R} \to [0, \infty]$ is a lower semi-continuous function (with the obvious extensions to cover $R(x) = \infty$); this is a representation that we will make frequent use of. And with a certain abuse of terminology, the open set $D = \{(x, t) : 0 < t < R(x)\}$, the complement of B in $\mathbb{R} \times (0, \infty)$ is sometimes also called a barrier.

In addition, the following result of Loynes [1970, Theorem 1] treated uniqueness of Root's solution.

Theorem 2.1.2. *For any centred probability distribution with finite variance, it is generated by exactly one regular barrier with finite expectation of the corresponding stopping time.*

Here, *the regular barriers* are defined as following:

Definition 2.1.3 (Regular Barrier)**.** A barrier B generated by R is *regular* if R vanishes outsider the interval $[x_-, x_+]$ where x_+ and x_- are defined as

$$\begin{aligned} x_+ &= \inf\{x > 0 : R(x) = 0\}\,; \\ x_- &= \sup\{x < 0 : R(x) = 0\}\,. \end{aligned}$$

This result tells that Root's solution is essentially unique: if there are two barriers which embed the same distribution, then their corresponding stopping times are equal with probability one.

The other important reference regarding Root's construction is Rost [1976]. This work vastly extends the generality of the results of Root and Loynes, and uses mostly potential-theoretic techniques. Rost works in the generality of a Markov process X_t on a compact metric space E (with the transition semi-group $\big(P_t^X; t \geq 0\big)$), which satisfies the strong Markov property and is right-continuous. Then Rost recalls (from an original definition of Dinges [1974] in the discrete setting) the notion of *minimal residual expectation*:

Definition 2.1.4 (Minimal Residual Expectation)**.** Suppose that the initial distribution of X is ν. We say that a stopping time τ^* is of *minimal residual expectation (m.r.e.) with respect to* μ, if $\nu P_{\tau^*}^X = \mu$ and if for all τ such that $\nu P_\tau^X = \mu$ one has

$$\mathbb{E}^\nu \left[\int_{t \wedge \tau^*}^{\tau^*} f(X_r) \mathrm{d}r \right] \;\leq\; \mathbb{E}^\nu \left[\int_{t \wedge \tau}^{\tau} f(X_r) \mathrm{d}r \right], \tag{2.1.3}$$

for all positive Borel measurable function f and $t \geq 0$.

Let $f \equiv 1$ in (2.1.3), it follows that τ^* minimises the quantity:

$$\mathbb{E}^\nu\left[(\tau - t)_+\right] \;=\; \mathbb{E}^\nu\left[\int_{\tau\wedge t}^{\tau} \mathrm{d}s\right] \;=\; \int_t^\infty \mathbb{P}^\nu[\tau > s]\mathrm{d}s,$$

over all $\tau \in \mathcal{T}(\sigma, \nu, \mu)$.

Then Rost gives the following results connecting Root's stopping times and stopping times of m.r.e..

Theorem 2.1.5. *Suppose that the initial distribution of X is ν. If a probability measure μ satisfies that , $\nu U^X \geq \mu U^X$, then we have the following result:*

(i). *There exists a stopping time of m.r.e. with respect to μ;*

(ii). *If the one-point sets are regular[1] for X, then any stopping time T of m.r.e. with respect to νP_T^X is Root's stopping time;*

(iii). *Every Root's stopping time T is of m.r.e. with respect to νP_T^X. Moreover, if S is also a stopping time of m.r.e. with respect to νP_T^X, then $S = T$, $\mathbb{P}^\nu$-a.s..*

The results (i) and (ii) above imply the existence of the Root's stopping time whenever $\nu U^X \geq \mu U^X$ and the one-point sets are regular for X. Then the uniqueness of Root's stopping times is implied by (iii).

Moreover, note that a stopping time is of minimal residual expectation if and only if, for every convex, increasing function $F(t)$ (where, without loss of generality, we take $F(0) = F'_+(0) = 0$), it minimises the quantity:

$$\mathbb{E}[F(\tau)] \;=\; \mathbb{E}\left[\int_0^\infty (\tau - t)_+ F''(\mathrm{d}t)\right],$$

the second derivative F'' of a convex function (or the difference of two convex functions) F in the sense of distributions is a positive (respectively, a signed) Radon measure denoted by $F''(\mathrm{d}t)$. This fact is a trivial consequence of the above representation.

There are a number of important properties that the Root barrier possesses. Firstly, we note that, as a consequence of the fact that B is closed and (iii) of Definition 2.1.1,

[1] The regularity of the one-point set $\{x\}$ means that almost all paths of the space-time process $(X_\cdot, \cdot)$ starting from (x, t) in an arbitrary small time interval hit the set $\{(x, s) : s > t\}$. The classes of processes satisfying the regularity include the class of time-homogeneous diffusions we consider in this thesis.

the barrier is regular for the class of processes we will consider (time-homogeneous diffusions). This will have important analytical benefits. Secondly, there are important consequences for the density of the stopped process: it is clear that if $(x,t) \not\in B$, then we have $\mathbb{P}(W_{t\wedge\tau_D} \in dx) = \mathbb{P}(W_{t\wedge\tau_D} \in dx, t < \tau_D)$ (see Lemma 2.2.2), which will also be of importance in what follows.

Finally, we note some simple examples where the barrier function can be explicitly calculated, or properties derived if the underlying process is a Brownian motion: firstly, if μ is a normal distribution $N(0,\sigma^2)$, we easily see that the barrier function $R(x) \equiv \sigma^2$ embeds μ. Secondly, if μ consists of two atoms (weighted appropriately) at $a < 0 < b$, the corresponding regular barrier is (see also Figure 2-1(a)):

$$R(x) \;=\; \begin{cases} 0\,, & x \not\in (a,b)\,; \\ \infty\,, & x \in (a,b)\,. \end{cases}$$

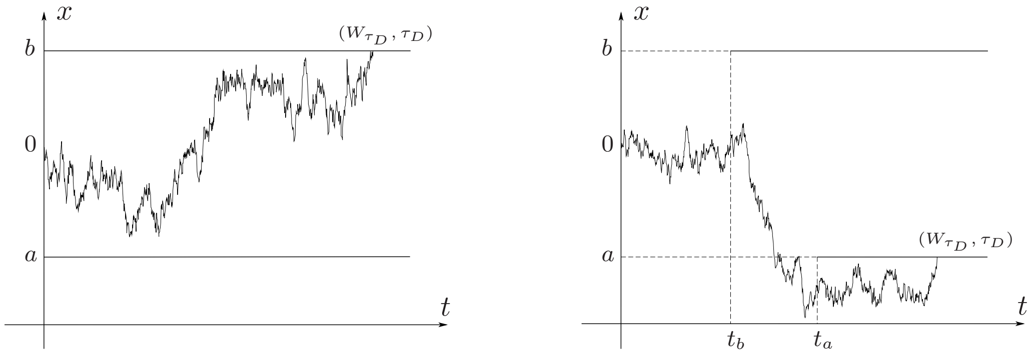
In this example, note firstly that, without the regularity defined in Definition 2.1.3, the function R is not unique: we can choose any behaviour outside $[a,b]$, and the stopping times achieved are the same. Secondly, we note that there are even more general Root's solutions to the Skorokhod embedding problem (without the uniform integrability condition) since there are also barriers of the form (see also Figure 2-1(b)):

$$R(x) \;=\; \begin{cases} t_a\,, & x = a\,; \\ t_b\,, & x = b\,; \\ \infty\,, & x \in \{a,b\}\,, \end{cases}$$

which embed the same law (provided $t_a, t_b > 0$), but which do not satisfy the uniform integrability condition.

Another example of Root's construction can be found in Huff [1975]. We will use this solution as an example of the optimality of Root's solution later (Example 4.1.2, Chapter 4).

In general, a barrier can exhibit some fairly nasty features: As an extreme example, consider the canonical measure on a middle third Cantor set C (scaled so that it is on $[-1,1]$). Root's result tells us that there exists a barrier which embeds this distribution, and clearly the resulting barrier function must be finite only on the Cantor set, however the target distribution has no atoms, so that the 'spikes' in the barrier function can



(a) Uniformly integrable embedding. (b) A 'bad' case.

Figure 2-1: Different Root's embeddings of an identical $\boldsymbol{\mu}$.

not be isolated (i.e. we must have $\liminf_{y\uparrow x} R(y) = \liminf_{y\downarrow x} R(y) = R(x)$ for all $x \in (-1,1) \cap C$).

2.1.2 On Potential

We consider X, a Markov process on $\mathbb{R}$ with the transition semi-group $\left(P_t^X; t \geq 0\right)$. In the theory of Markov processes, *the potential kernel* U^X denotes $\int_{\mathbb{R}_+} P_t^X \mathrm{d}t$, a linear operator on the space of measures on $\mathbb{R}$. Then, if $X_0 \sim \mu$, μU^X can be regarded as the occupation measure for X along its paths. Regarding potential kernels and embedding problems, Rost [1971, Theorem 4] says that, if $X_0 \sim \nu$, there exists a (possibly randomized) stopping time T (here, the term "randomized" means that the stopping time is dependent on not only the path of the underlying process, but also some external factors) such that $X_T \sim \mu$ if and only if

$$\mu U^X \ \leq \ \nu U^X. \tag{2.1.4}$$

Combining this result with Theorem 2.1.5, we see that, if the one-point sets are regular for X, then there exists a unique Root's stopping time T such that $X_T \sim \mu$ if and only if (2.1.4) holds.

Suppose $p_t^X(x,\cdot)$ exists, which is the transition density at time t of X starting at x, the condition (2.1.4) is equivalent to

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}} p_t^X(x,y)\nu(\mathrm{d}y)\mathrm{d}t \ \geq \ \int_{\mathbb{R}_+}\int_{\mathbb{R}} p_t^X(x,y)\mu(\mathrm{d}y)\mathrm{d}t, \quad \text{for all } \ x \in \mathbb{R}.$$

If X is a Brownian motion, we note that the integrals are infinite. To resolve this we use the compensated condition:

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}} \left(p_t^X(x,y) - p_0^X(x,y)\right)(\nu-\mu)(\mathrm{d}y)\mathrm{d}t \;\geq\; 0, \quad \text{for all } \; x \in \mathbb{R},$$

where the left-hand side can be writen as

$$\int_{\mathbb{R}}\int_{\mathbb{R}_+} \frac{1}{\sqrt{2\pi t}}\left[e^{-(y-x)^2/(2t)} - 1\right]\mathrm{d}t(\nu-\mu)(\mathrm{d}y) \;=\; -\int_{\mathbb{R}} |y-x|(\nu-\mu)(\mathrm{d}y). \tag{2.1.5}$$

Now we refer to

$$\mathrm{U}\mu(x) \;:=\; -\int_{\mathbb{R}} |x-y|\mu(\mathrm{d}y) \tag{2.1.6}$$

as to *the potential of* μ, then the compensated condition can be written as $\mathrm{U}\mu \leq \mathrm{U}\nu$.

Through our work, we only consider the diffusion X satisfying

$$\mathrm{d}X_t \;=\; \sigma(X_t)\mathrm{d}W_t.$$

Under the assumption $\mathrm{U}\nu \geq \mathrm{U}\mu$, according to Obłój [2004, Proposition 8.1], for any starting distribution ν, there exists UI embedding of μ for the diffusion X. Then we must have that $\nu U^X \geq \mu U^X$, and hence, Theorem 2.1.5 gives the existence of the barrier embedding μ into X under $\mathbb{P}^\nu$. Therefore, throughout this thesis, the following restriction is always assumed:

$$\boldsymbol{\mathrm{U}\nu \;\geq\; \mathrm{U}\mu\,.} \tag{2.1.7}$$

Moreover, the following results can be found in Chacon [1977]:

Theorem 2.1.6. *A distribution is integrable if and only if its potential is finite on* $\mathbb{R}$. *Now suppose* ν, μ *and* μ_n, $n = 1, 2, \cdots$, *are integrable probability distributions, we have*

(i). *If the mean* $\int_{\mathbb{R}} x\mu(\mathrm{d}x) = m$, *then* $\mathrm{U}\mu \leq \mathrm{U}\delta_m = -|x-m|$, *and if* $\mathrm{U}\mu \leq \mathrm{U}\nu$, *then* μ *and* ν *have the same mean;*

(ii). *If* μ *and* ν *have same mean, then* $\lim_{|x|\to\infty}(\mathrm{U}\mu(x) - \mathrm{U}\nu(x)) = 0$*;*

(iii). *If* μ *and* μ_n *share the same mean, then* $\mu_n \Rightarrow \mu$ *if and only if* $\lim_{n\to\infty} \mathrm{U}\mu_n(x) = \mathrm{U}\mu(x)$ *for all* $x \in \mathbb{R}$*;*

(iv). *If* $\mathrm{U}\mu \leq \mathrm{U}\nu$, *then*

$$\int_{\mathbb{R}} x^2\mu(\mathrm{d}x) - \int_{\mathbb{R}} x^2\nu(\mathrm{d}x) \;=\; \int_{\mathbb{R}} \Big[\mathrm{U}\nu(x) - \mathrm{U}\mu(x)\Big]\mathrm{d}x \;\geq\; 0;$$

(v). $\mathrm{U}\mu|_{[a,\infty)} = \mathrm{U}\nu|_{[a,\infty)}$ *if and only if* $\mu|_{(a,\infty)} = \nu|_{(a,\infty)}$*;*

(vi). $\mathrm{U}\mu$ *is differentiable almost everywhere with left and right derivatives*

$$\mathrm{U}\mu'_{-} \;=\; 1 - 2\mu((-\infty, x)), \quad \mathrm{U}\mu'_{+} \;=\; 1 - 2\mu((-\infty, x]),$$

consequently, $\mathrm{U}\mu$ *is concave and Lipschitz continuous with parameter 1.*

Because of (vi), the potential of any distribution μ can be written as the infimum of a countable number of affine functions. Using this property, Chacon and Walsh [1976] developed a solution to the Skorokhod embedding problem in a simple and general way, and their scheme is called *the balayage of potentials.*

2.2 Deriving the Free Boundary Problem from Root's Solution

Initially, we consider the construction of a barrier, for a Brownian motion starting at 0. However, we will in general be interested in this question when our underlying process X is the unique strong solution to the stochastic differential equation (2.1.1) with $X_0 = m \in \mathbb{R}$. Hence, we write the Skorokhod embedding problem as (recall the definition of $\mathcal{T}$: (2.1.2))

SEP($\boldsymbol{\sigma, \nu, \mu}$): Suppose (2.1.7) holds. Find a lower-semicontinuous function $R(x)$ such that the domain $D = \{(x,t) : 0 < t < R(x)\}$ has $\tau_D \in \mathcal{T}(\sigma, \nu, \mu)$, where
$\tau_D \;:=\; \inf\{t > 0 : \; (X_t, t) \;\notin\; D\} \;=\; \inf\{t > 0 : \; t \geq R(X_t)\}.$

By the uniform integrability of τ_D, (or by the assumption $\mathrm{U}\nu \geq \mathrm{U}\mu$), it is clear the means of μ and ν are same.

We would then like to connect **SEP($\boldsymbol{\sigma, \nu, \mu}$)** with the following free boundary problem (recall the definition of the potential (2.1.6)):

FBP(σ, ν, μ): Suppose (2.1.7) holds. Find a continuous function $u : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ and an open set

$$D = \{(x, t) : \, 0 < t < R(x)\}$$

where $R : \mathbb{R} \to \overline{\mathbb{R}}_+$ is a lower semi-continuous function, such that (u, D) satisfies

$$u \;\in\; \mathbb{C}^0(\mathbb{R} \times [0, \infty)) \quad \text{and} \quad u \;\in\; \mathbb{C}^{2,1}(D)\,; \tag{2.2.1a}$$

$$\frac{\partial u}{\partial t} \;=\; \frac{\sigma^2(x)}{2} \frac{\partial^2 u}{\partial x^2}, \quad \text{for } \; (x, t) \in D\,; \tag{2.2.1b}$$

$$u(x, 0) \;=\; \mathrm{U}\nu(x), \quad \text{for all } \; x \in \mathbb{R}; \tag{2.2.1c}$$

$$u(x, t) \;=\; \mathrm{U}\mu(x), \quad \text{if } \; t \geq R(x) \; \text{ and } \; x \in \mathbb{R}; \tag{2.2.1d}$$

$$u(x, t) \;\downarrow\; \mathrm{U}\mu(x) \;\; \text{as} \;\; t \uparrow \infty, \quad \text{if } \; R(x) = \infty; \tag{2.2.1e}$$

$$u(\cdot, t) \text{ is concave with respect to } x \in \mathbb{R}. \tag{2.2.1f}$$

Here, $\mathbb{C}^0(\mathbb{R} \times [0, \infty))$ is the collection of all functions continuous on $(\mathbb{R} \times [0, \infty))$ and $\mathbb{C}^{2,1}(D)$ is the collection of all functions having continuous second derivatives with respect to x and continuous derivatives with respect to t on D.

When we consider, as we will in this section and the next two, the case where X is a standard Brownian motion which implies $\sigma(x) \equiv 1$ and $\nu = \delta_0$, we will drop the σ from the name, and call the problems **SEP(μ)** and **FBP(μ)**.

Our first result is that we can derive a solution to the free boundary problem from Root's solution to the Skorokhod embedding problem:

Theorem 2.2.1. *Assume D solves* **SEP(μ)***, with corresponding UI stopping time τ_D. Then the couple (u, D), where $u := -\mathbb{E}[|\, x - W_{t \wedge \tau_D} |]$, solves* **FBP($\mu$)***.*

Before proving this theorem, we give two useful results concerned with the stopping distribution of Root's solution.

Lemma 2.2.2. *For any $(x, t) \in D$, $\; \mathbb{P}[\, W_{t \wedge \tau_D} \in \mathrm{d}x \,] = \mathbb{P}[\, W_t \in \mathrm{d}x, \, t < \tau_D]$.*

Proof. By lower semi-continuity of R, since $(x,t) \in D$, then there exists $h > 0$ such that $(x-h,\, x+h) \times [0,\, t+h) \subset D$, so for any $y \in (x-h,\, x+h)$, $R(y) > t$. So for $\omega \in \Omega$, if $\tau_D(\omega) \leq t$, we have $R(W_{\tau_D}(\omega)) \leq \tau_D(\omega) \leq t$, then $W_{\tau_D}(\omega) \notin (x-h,\, x+h)$. Therefore,

$$\begin{aligned}
&\mathbb{P}\big[\, W_{t\wedge\tau_D} \in \mathrm{d}x \,\big] \\
&\qquad = \mathbb{P}[\, W_t \in \mathrm{d}x,\ t < \tau_D] + \mathbb{P}[\, W_{\tau_D} \in \mathrm{d}x,\ t \geq \tau_D] \ = \ \mathbb{P}[\, W_t \in \mathrm{d}x,\, t < \tau_D].
\end{aligned}$$

□

Lemma 2.2.3. *The measure corresponding to $\mathcal{L}(W_t; t < \tau_D)$ has a density $p^D(x,t)$ with respect to Lebesgue on D, and the density is smooth and satisfies:*

$$\frac{\partial}{\partial t} p^D(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2} p^D(x,t).$$

Further, trivially, $p_D(x,t)$ is dominated by the density of a standard Brownian motion.

This result appears to be standard, but we are unable to find concise references. We give a short proof based on Rogers and Williams [2000b, V.38.5].

Proof. First note that, as a measure, $\mathcal{L}(W_t; t < \tau_D)$ is dominated by the usual transition measure, so the density $p^D(x,t)$ exists.

Let (x_0, t_0) be a point in D, and we can therefore find an $\varepsilon > 0$ such that $A = (x_0 - \varepsilon, x_0 + \varepsilon) \times (t_0 - \varepsilon, t_0 + \varepsilon)$ satisfies $\overline{A} \subseteq D$. Then let f be a function in $\mathbb{C}_K^\infty(A)$, and by Itô:

$$f(W_{t\wedge\tau_D}, t) - f(W_0, 0) \ = \ \int_0^t \frac{\partial}{\partial x} f(W_{s\wedge\tau_D}, s)\mathrm{d}W_s + \int_0^t \left(\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial t}\right) f(W_{s\wedge\tau_D}, s)\mathrm{d}s.$$

Since f is compactly supported, taking $t > t_0 + \varepsilon$, the two terms on the left disappear, and the first integral term is a martingale. Hence, taking expectations, and interchanging the order of differentiation, we get:

$$\int_0^t \int p^D(y,s) \left(\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial t}\right) f(X_{s\wedge\tau_D}, s)\mathrm{d}y\, \mathrm{d}s \ = \ 0.$$

Interpreting $p^D(y,s)$ as a distribution, we get (in a distributional sense)

$$\left(\frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial t}\right) p^D(x,t) \;=\; 0, \quad \text{for} \;\; (x,t) \in A,$$

and since the heat operator is hypo-elliptic, we conclude that $p^D(x,t)$ is smooth in A. (e.g. Friedman [1963, 11.1, Theorem 3]). $\square$

We are now able to prove that any solution to Root's embedding problem is a solution to the free boundary problem.

Proof of Theorem 2.2.1. We first show that the function u is suitably differentiable on D, and satisfies (2.2.1b). By Lemma 2.2.2 and (vi) of Theorem 2.1.6, it follows that

$$\frac{\partial u}{\partial x}(x,t) \;=\; 1 - 2\mathbb{P}(W_{t\wedge\tau_D} < x),$$

and therefore (in D) by Lemma 2.2.3 the function u has a smooth second derivative in x. Using Lemma 2.2.2, for $(x,t) \in D$, when ε is sufficiently small, we have

$$\left|\frac{\partial u}{\partial x}(x+\varepsilon,t) - \frac{\partial u}{\partial x}(x-\varepsilon,t)\right| \;\leq\; 2\mathbb{P}\left[W_t \in (x-\varepsilon, x+\varepsilon)\right].$$

Therefore, by the dominated convergence theorem and Fubini's theorem, we have

$$\begin{aligned} -\frac{1}{2}\int_0^t \frac{\partial^2}{\partial x^2} u(x,s)\mathrm{d}s \;&=\; \frac{\partial}{\partial x}\left(\int_0^t \mathbb{P}\left[W_{s\wedge\tau_D} < x\right]\mathrm{d}s\right) \\ &=\; \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[\frac{1}{2\varepsilon}\int_0^t \mathbf{1}_{[x-\varepsilon<W_{s\wedge\tau_D}<x+\varepsilon]}\mathrm{d}s\right], \end{aligned}$$

Consider the right-hand side, since $(x,t) \in D$, $(x,s) \in D$ for all $s \leq t$, and hence, for $\varepsilon > 0$ small enough, by Lemma 2.2.2

$$\mathbf{1}_{[x-\varepsilon<W_{s\wedge\tau_D}<x+\varepsilon]} \;=\; \mathbf{1}_{[x-\varepsilon<W_s<x+\varepsilon]}\mathbf{1}_{[s<\tau_D]}$$

Thus,

$$
\begin{aligned}
-\frac{1}{2}\int_0^t \frac{\partial^2}{\partial x^2}u(x,s)\mathrm{d}s &= \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[\frac{1}{2\varepsilon}\int_0^t \mathbf{1}_{[x-\varepsilon<W_s<x+\varepsilon]}\mathbf{1}_{[s<\tau_D]}\mathrm{d}s\right] \\
&= \lim_{\varepsilon\downarrow 0}\mathbb{E}\left[\frac{1}{2\varepsilon}\int_0^{t\wedge\tau_D} \mathbf{1}_{[x-\varepsilon<W_s<x+\varepsilon]}\mathrm{d}s\right] \\
&= \mathbb{E}\left[L^x_{t\wedge\tau_D}\right] \;=\; \mathbb{E}\left[|x-W_{t\wedge\tau_D}|\right]-|x|\,,
\end{aligned}
$$

where we use the dominated convergence theorem in the second equality, and L^x is the local time of Brownian motion at x. It follows that u satisfies (2.2.1b), and further that $u(x,t)$ is differentiable in t with continuous (and in fact smooth) derivative.

Since u satisfies (2.2.1c) and clearly also (2.2.1f), we need only show (2.2.1a), (2.2.1d) and (2.2.1e).

For (2.2.1a) it remains only to show that u is continuous on the whole of $\mathbb{R}\times[0,\infty)$. For any $(x,t),(y,s)\in\mathbb{R}\times[0,\infty)$, we have, as $(y,s)\to(x,t)$,

$$
\begin{aligned}
|u(y,s)-u(x,t)| &= \Big|\mathbb{E}[|\,y-W_{s\wedge\tau_D}|]-\mathbb{E}[|\,x-W_{t\wedge\tau_D}|]\Big| \\
&\le |y-x|+\mathbb{E}[|W_{t\wedge\tau_D}-W_{s\wedge\tau_D}|] \\
&\le |y-x|+\mathbb{E}\Big\{\mathbb{E}\Big[|W_t-W_s|\Big|\mathcal{F}_{\tau_D}\Big]\Big\} \\
&= |y-x|+\mathbb{E}\left[|W_t-W_s|\right] \;\longrightarrow\; 0,
\end{aligned}
$$

where we use Jensen's inequality in the third line. Therefore, u is continuous on $\mathbb{R}\times[0,\infty)$. Now (2.2.1a) is proved.

For $(x,t)\in D^{\complement}$, it follows from the definition of the barrier that if $\tau_D>t$, the Brownian path cannot cross the line $\{(x,s):\ s\ge t\}$ in the time interval $[t,\tau_D)$, and hence, we have that

$$
L^x_{t\wedge\tau_D} \;=\; L^x_t\mathbf{1}_{\tau_D>t}+L^x_{\tau_D}\mathbf{1}_{\tau_D\le t} \;=\; L^x_{\tau_D}\mathbf{1}_{\tau_D>t}+L^x_{\tau_D}\mathbf{1}_{\tau_D\le t} \;=\; L^x_{\tau_D}.
$$

Therefore, for $t\ge R(x)$,

$$
\mathbb{E}\left[|\,x-W_{t\wedge\tau_D}|\right] \;=\; |x|+\mathbb{E}\left[L^x_{t\wedge\tau_D}\right] \;=\; |x|+\mathbb{E}\left[L^x_{\tau_D}\right] \;=\; \mathbb{E}\left[|\,x-W_{\tau_D}|\right], \tag{2.2.2}
$$

where the last equality holds because τ_D is a UI stopping time. So (2.2.1d) holds. To see that (2.2.1e) holds, note that we can take the limit in (2.2.2) as $t\to\infty$, and using the fact that $W_{t\wedge\tau_D}$ is UI and $\tau_D < \infty$ a.s. to deduce

$$\lim_{t\to\infty} -\mathbb{E}|\, x - W_{t\wedge\tau_D}| = -\mathbb{E}|\, x - W_{\tau_D}| \;=\; \mathrm{U}\mu(x)$$

□

Remark 2.2.4. We note that the uniform integrability condition is used only in the proof of (2.2.1d) and (2.2.1e), and hence, without this condition, u still satisfies (2.2.1a), (2.2.1b), (2.2.1c) and (2.2.1f).

2.3 Uniqueness of Solutions to Free Boundary Problems

We have proved that we can construct a solution to the free boundary problem from a solution to Root's embedding problem. However, is the solution to the free boundary problem, (u, D), unique? Equivalently, we consider the converse problem: does any solution of $\mathbf{FBP}(\boldsymbol{\mu})$ solve $\mathbf{SEP}(\boldsymbol{\mu})$?

While it would be ideal to provide a complete converse to the result above, there appear to be a number of technical issues that prevent this. A simple approach to the issue would be to provide an analytical proof of the uniqueness to the free boundary problem, and since we know Root's solution exists, this must be same solution. Unfortunately, the exact nature and conditions do not appear to immediately allow this, at least in the degree of generality we would like to do this. We note, for example, that a barrier may include 'spikes' of a single point at which $R(x_0) < \liminf_{x\to x_0} R(x)$, at which one expects atoms of the measure to appear. Alternatively, there can be regions of space where the barrier is infinite, or where the barrier is flat in either the time or space direction. As an extreme example of a possible barrier, consider embedding the canonical measure on a middle-third Cantor set (stretched so its centre is at 0 and with extremes ± 1). Root's results tell us that there exists a corresponding barrier $R(x)$, which will necessarily take the value infinity at points which are not in the Cantor set, and therefore correspond to uncountably many spikes, however none of these can themselves give rise to atoms in the stopped law, so they must contain further structure.

The main issue here is then the problem of wild behaviour of the boundary of the domain, but Root's problem does not admit an easy way of deducing properties of the boundary directly from knowledge of the measure. To resolve this, we shall introduce an additional property which we require of our solution to the free boundary problem near the boundary. In this sense, we are moving from a free boundary problem where we only specify global properties (e.g. concavity) to a situation where we also specify some local behaviour. The exact formulation of the criterion will be given shortly, but the key idea is that particles near the barrier will strike the barrier almost immediately, so that the density of particles near the barrier should go to zero as we approach the barrier. Roughly, this says that the second derivative of a solution (which, in D, we have identified with the density of particles) should disappear close to the barrier. In fact, we shall need to be slightly more careful, as we now explain.

To verify $\mathbf{SEP}(\boldsymbol{\mu})$, we study the solutions to $\mathbf{FBP}(\boldsymbol{\mu})$. The next result allows us to make the connection between solutions to the free boundary problem and potentials of probability distributions:

Lemma 2.3.1. *If $u : \mathbb{R} \to \mathbb{R}_-$ is a concave function satisfying*

$$u(x) \leq -|x| \quad \textit{and} \quad \lim_{|x|\to\infty} (u(x) + |x|) = 0,$$

then u is the potential of a centred probability distribution ν on $\mathbb{R}$, that is $u(x) = \mathrm{U}\nu$.

Proof. By Revuz and Yor [1999, Proposition A.3.2], there exists a positive Radon measure ν such that $u(x) = -\int_{\mathbb{R}} |x-y|\nu(\mathrm{d}y)$. Moreover ν is a probability measure if $-\lim_{x\to\infty} u'_+(x) = \lim_{x\to-\infty} u'_-(x) = 1$ where the right and left derivatives of u are denoted by u'_+ and u'_-, which follows from the conditions on u. So ν is a probability measure and moreover ν is centred by (i) of Theorem 2.1.6. $\square$

By Lemma 2.3.1, if (u, D) solves the $\mathbf{FBP}(\boldsymbol{\mu})$, then for any $t \geq 0$, there exists μ_t, a centred probability distribution on $\mathbb{R}$, such that $u(\cdot, t) = \mathrm{U}\mu_t$. Since u is continuous on $\mathbb{R} \times [0, \infty)$ and for all $x \in \mathbb{R}$, $\lim_{t\to 0} u(x,t) = -|x|$, by (iii) of Theorem 2.1.6, we have

$$\{\mu_t\}_{t>0} \text{ is weakly continuous} \quad \text{and} \quad \mu_t \Rightarrow \delta_0, \quad \text{as } t \to 0. \tag{2.3.1}$$

It is easy to verify that for $(x,t) \in D$, i.e. $R(x) > t$,

$$\mu_t(\mathrm{d}x) = -\frac{1}{2}\frac{\partial^2 u}{\partial x^2}(x,t)\mathrm{d}x.$$

Now, for any $t \geq 0$ define a sub-probability measure $\hat{\mu}_t$ by

$$\hat{\mu}_t(\mathrm{d}x) := \begin{cases} q(x,t)\mathrm{d}x, & t > 0; \\ \delta_0(\mathrm{d}x), & t = 0, \end{cases} \quad \text{where} \quad q(x,t) := -\frac{1}{2}\mathbf{1}_{[(x,t)\in D]}\frac{\partial^2 u}{\partial x^2}(x,t), \quad (2.3.2)$$

noting that since the operator $\frac{1}{2}\partial_{xx} - \partial_t$ is hypo-elliptic, q is smooth on D. The following results are concerned with $\hat{\mu}$ and q. Denote $D_r := \{x \in \mathbb{R} : R(x) > r\}$. Since $D.$ is decreasing, for any $x \in D_t^{\complement}$, $q(x,t) = q(x,t+s) = 0$ where $t, s > 0$. For any $x \in D_t$, since D is open, there exists $h > 0$ such that $\{x\} \times (t, t+h) \subset D$, i.e. $(x, t+s) \in D_{t+s}$ for any $s < h$, and hence, we have $\lim_{s\downarrow 0} q(x, t+s) = q(x,t)$ for any $x \in \mathbb{R}$.

We now turn to the additional boundary condition. Consider initially the following example:

Example 2.3.2. Suppose $D = (\mathbb{R} \times \mathbb{R}_+)/\{(y,s) : y = x, s \geq t\}$ for some $(x,t) \in \mathbb{R} \times \mathbb{R}_+$. For $(y,s) \in D$ with $y < x,\ s > t$, by the reflection principle we can compute the density $\widetilde{p}(y,s)$ of $W_{s\wedge\tau_D}$ at y to be:

$$\widetilde{p}(y,s) = \frac{1}{\sqrt{2\pi s}}\left[\exp\left\{-\frac{(x-\varepsilon)^2}{2s}\right\}\Phi\left(\frac{rx}{\sqrt{rst}} + \varepsilon\sqrt{\frac{t}{rs}}\right)\right.$$
$$\left. - \exp\left\{-\frac{(x+\varepsilon)^2}{2s}\right\}\Phi\left(\frac{rx}{\sqrt{rst}} - \varepsilon\sqrt{\frac{t}{rs}}\right)\right],$$

where we have taken $r := s - t,\ \varepsilon := x - y$. Now let $\varepsilon \downarrow 0$, and $r \downarrow \eta \geq 0$. If $\eta > 0$, $\widetilde{p}(y,s) \downarrow 0$, but if $\eta = 0$, the convergence is dependent on the specific choice of r, ε: we have $\widetilde{p}(y,s) \not\to 0$ if $\varepsilon/\sqrt{r} \to \infty$ or $\varepsilon/\sqrt{r} \to c$ for some constant $c > 0$.

As a result observed from this example, we cannot expect convergence of the second derivatives to zero along all paths to the boundary. A similar issue will also arise if we consider barriers which are constant over some interval, where there is no sense in which the second derivative will vanish.

Our proof of the uniqueness of a solution will rely on constructing solutions by running diffusions backwards in time. In this case, the correct notion of convergence to zero will be that the second derivative disappears along almost every path of a reversed Brownian motion which hits the boundary. Specifically, we assume D and u satisfy: for any fixed point $(y,s) \in D$, $\frac{\partial^2}{\partial x^2}u(y + W_t, s - t)$ converges almost surely and in L^1 to 0 along trajectories of $(y + W_t, s - t)$ at $\partial_+ D := \partial D \cap \{(x,t) : t > 0\}$. Specifically,

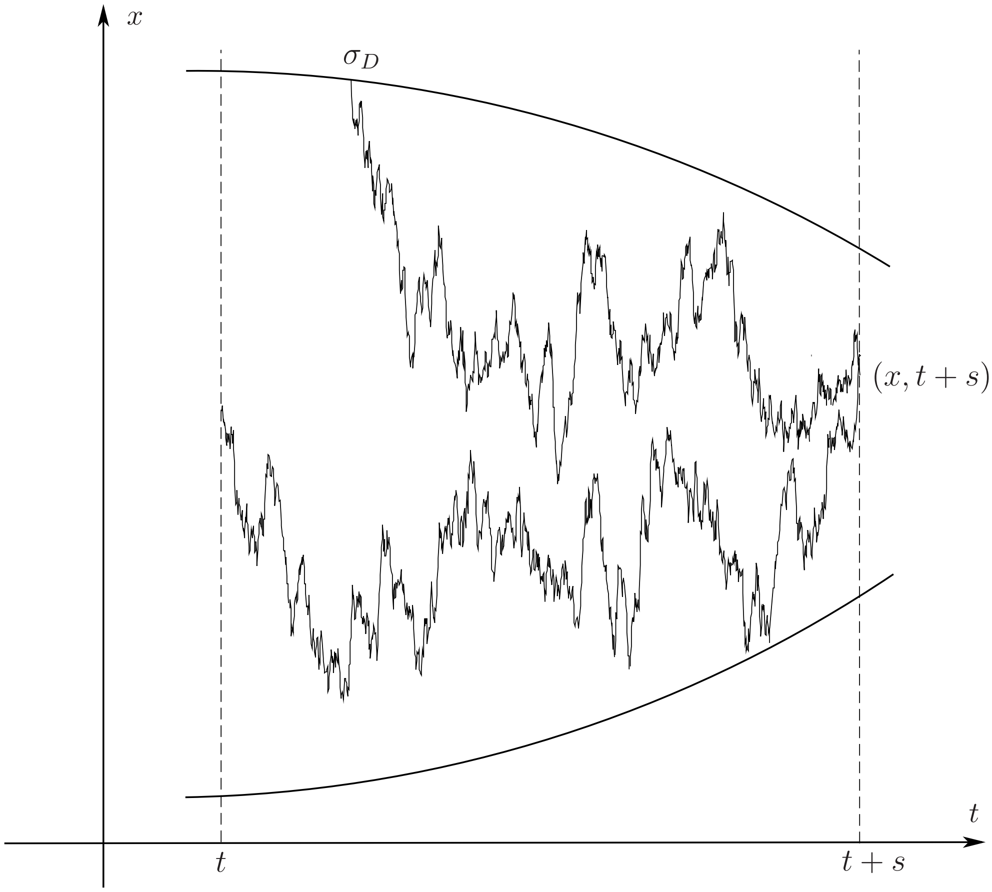


Figure 2-2: The Definition of $\boldsymbol{\sigma_D}$.

writing $\sigma_D = \inf\{t \geq 0 : (y + W_t, s - t) \in \partial_+ D\}$ as in Figure 2-2, if σ_N is a sequence of stopping times increasing to σ_D as $N \to \infty$, then

$$\frac{\partial^2}{\partial x^2} u(y + W_{\sigma_N}, s - \sigma_N)\ \mathbf{1}_{[\sigma_N < s]} \xrightarrow{\text{a.s. and in } L^1} 0. \tag{2.3.3}$$

With this additional condition, we shall be able to show the following result:

Theorem 2.3.3. *If the couple (u, D) solves* $\mathbf{FBP(\boldsymbol{\mu})}$ *and satisfies (2.3.3), then D solves* $\mathbf{SEP(\boldsymbol{\mu})}$*, and $u(x,t) = -\mathbb{E}[|x - W_{t \wedge \tau_D}|]$.*

Our proof is based on the proof of a similar result in Rost [1976, Theorem 1], which in turn hinges on a result in Rost [1971, Proposition 6]. A key step is connecting the solution to the free boundary problem with the law of the stopped process at a fixed time, and this is done in the following proposition:

Proposition 2.3.4. *Suppose (u, D) is a solution to* $\mathbf{FBP}(\boldsymbol{\mu})$ *and that $\hat{\mu}_t$ is defined as above. Moreover, suppose (u, D) satisfies (2.3.3). Then*

$$\langle \hat{\mu}_t, f \rangle \; = \; \mathbb{E}\left[f(W_t) \mathbf{1}_{[t<\tau_D]} \right], \quad \text{for any} \;\; t \geq 0, \;\; f \in \mathbb{C}_b^0, \tag{2.3.4}$$

where $\mathbb{C}_b^0$ is the system of bounded continuous functions on $\mathbb{R}$

Before proving this proposition, we recall some terminologies from Rost [1971]. Firstly, we introduce a class of functions, denoted by $\mathcal{H}$, such that for any $H \in \mathcal{H}$:

$$\begin{cases} H : \Omega \times [\,0, \, \infty\,) \to [\,0, 1\,]\,; \\ H(\omega, \cdot) \text{ is decreasing, right continuous for each } \omega \in \Omega\,; \\ H(\cdot, t) \text{ is } \mathcal{F}_t\text{-measurable for each } t \geq 0\,. \end{cases}$$

As in the same work, we define a randomized stopping time T^H for $H \in \mathcal{H}$ by setting, for $s \in (0, 1]$, and the randomized expectation with respect to T^H,

$$T_s^H \; := \; \inf\{\, t : \; H(\cdot, t) < s\}, \tag{2.3.5}$$

$$\mathbb{E}\left[f(W_t); \; t < T^H\right] := \int_0^1 \mathbb{E}\left[f(W_t) \mathbf{1}_{[t<T_s^H]}\right] \mathrm{d}s. \tag{2.3.6}$$

With this definition, we recall Rost [1971, Proposition 6]:

Lemma 2.3.5 (Rost [1971])**.** *If $\{X_t\}_{t\geq 0}$ is a strong Markov process with semigroup P^X, and $\mu, \; \mu_t, \; t \geq 0$ are finite positive measures with the property:*

$$\mu_\cdot \textit{ is weakly right continuous;} \tag{2.3.7a}$$

$$\mu_{t+s} \leq \mu_t P_s^X \leq \mu P_{t+s}^X, \quad \textit{for } t, s \geq 0, \tag{2.3.7b}$$

then there exists a randomized stopping time T such that

$$\langle \mu_t, f \rangle \; = \; \mathbb{E}^\mu[f(X_t); \; T > t], \quad t \geq 0, \;\; f \in \mathbb{C}_b^0.$$

We also sketch Rost's proof of this result. Define binary $t = m/2^k$ with $m \in \mathbb{N}$ and

$k \in \mathbb{N}$, and the functions

$$Q^{(k)}(\cdot, t) \ := \ \prod_{n=0}^{m} \left(h_n^{(k)}(X_{n/2^k}) \right)$$

where the $h_n^{(k)} : \mathbb{R} \to [0,1]$ are measurable functions such that

$$h_0^{(k)} \cdot \mathrm{d}\mu \ = \ \mathrm{d}\mu_0, \quad h_n^{(k)} \cdot \mathrm{d}(\mu_{(n-1)/2^k} P^X_{1/2^k}) \ = \ \mathrm{d}\mu_{n/2^k}, \quad n \in \mathbb{N}.$$

Then for binary t and $f \in \mathbb{C}_b^0$, if k is large enough, we have

$$\langle \mu_t, \, f \rangle \ = \ \mathbb{E}^{\mu}[f(X_t)] \cdot Q^{(k)}(\cdot, t). \tag{2.3.8}$$

Furthermore, look at L^∞ over the product of the measure spaces $(\Omega, \mathcal{F}, \mathbb{P}^\mu)$ and the space of binaries with it discrete σ-algebra and the counting measure, a function $Q \in \mathcal{H}$ can be defined as a weak* limit of $Q^{(k)}$ (the limit may be not unique), and satisfies (2.3.8) where $Q^{(k)}$ is replaced by Q. The desired result follows by defining $T = T^Q$ as (2.3.5). Thanks to this result, we then can give the proof of Proposition 2.3.4.

Proof of Proposition 2.3.4. Our proof will take the following form: first, we show that for any $t, s > 0$, we have $\hat{\mu}_{t+s} \leq \hat{\mu}_t P_s$, where $P.$ is the semi-group generated by Brownian motion. Then we show that $\{\hat{\mu}_t\}_{t \geq 0}$ satisfies (2.3.7a) and (2.3.7b), so there exists a corresponding randomised stopping time by Lemma 2.3.5; finally, we will show this randomized stopping time is in fact a normal stopping time, i.e. it does not rely on any randomization external to W, and the stopping time is equal to τ_D.

Step 1:

For any $(x, t+s) \in D$ where $t > 0, s > 0$, we consider a Brownian motion started at $(x, t+s)$ and running backwards. We denote the hitting time of the set D by this process as σ_D. Define also the stopping time:

$$\sigma_D^N = \inf\{r \geq 0 : d((x + W_r, t + s - r), D^{\complement}) \leq N^{-1}\},$$

the first time the same Brownian motion is within distance N^{-1} of the barrier.

Then on $\{\sigma_D \leq s\}$, we have $q(x + W_{\sigma_D^N}, t + s - \sigma_D^N) \to 0$ a.s. as $N \to \infty$ by (2.3.3).

Further, by Itô's formula, since q satisfies the heat equation on D,

$$q(x+W_{r\wedge\sigma_D^N}, t+s-r\wedge\sigma_D^N) = q(x,t+s) + \int_0^{r\wedge\sigma_D^N} \frac{\partial}{\partial x} q(x+W_h, t+s-h)\mathrm{d}W_h,$$

and hence, for any $N\in\mathbb{N}$, $\{q(x+W_{r\wedge\sigma_D^N\wedge T_N}, t+s-r\wedge\sigma_D^N\wedge T_N)\}_{r\in[0,s]}$ is a continuous martingale, where $T_N^+ := \inf\{t\geq 0: W_t > N\}$, $T_N^- := \inf\{t\geq 0: W_t < -N\}$ and $T_N = T_N^+\wedge T_N^-$. Thus,

$$q(x,t+s) \;=\; \mathbb{E}[q(x+W_{s\wedge\sigma_D^N\wedge T_N}, t+s-s\wedge\sigma_D^N\wedge T_N)].$$

Letting N go to infinity, we have, by Fatou's Lemma,

$$q(x,t+s) \;\geq\; \mathbb{E}[\lim_{N\to\infty} q(x+W_{s\wedge\sigma_D^N\wedge T_N}, t+s-s\wedge\sigma_D^N\wedge T_N)] \;=\; \mathbb{E}[q(x+W_s,t)\mathbf{1}_{[s<\sigma_D]}].$$

where the indicator function is a consequence of (2.3.3).

On the other hand, by the L^1 convergence in (2.3.3), we have

$$\begin{aligned}
q(x,&t+s) \;=\; \mathbb{E}\left[q(x+W_{s\wedge\sigma_D^N\wedge T_N}, t+s-s\wedge\sigma_D^N\wedge T_N)\right] \\
&\leq \mathbb{E}\left[q(x+W_{s\wedge\sigma_D^N}, t+s-s\wedge\sigma_D^N)\right] + \mathbb{E}\left[q(x+W_{T_N}, t+s-T_N)\mathbf{1}_{[s\geq T_N]}\right] \\
&\leq \lim_{N\to\infty}\mathbb{E}\left[q(x+W_s,t)\mathbf{1}_{[s<\sigma_D^N]}\right] + \lim_{N\to\infty}\mathbb{E}\left[q(x+W_{\sigma_D^N}, t+s-\sigma_D^N)\mathbf{1}_{[s\geq\sigma_D^N]}\right] \qquad (2.3.9)\\
&\qquad\qquad + \lim_{N\to\infty}\mathbb{E}\left[q\left(x+W_{T_N}, t+s-T_N\right)\mathbf{1}_{[s\geq T_N]}\right] \\
&= \mathbb{E}\left[q(x+W_s,t)\mathbf{1}_{[s<\sigma_D]}\right] + \lim_{N\to\infty}\mathbb{E}\left[q\left(x+W_{T_N}, t+s-T_N\right)\mathbf{1}_{[s\geq T_N]}\right].
\end{aligned}$$

By the reflection principle, we know that the density of T_N^+ is

$$f_N^+(r) \;=\; \frac{N}{\sqrt{2\pi r^3}}\exp\left\{-\frac{N^2}{2r}\right\} \;\longrightarrow\; 0, \quad \text{as } \; N\to 0,$$

and $f_N^+(r)$ is increasing on $(0,N^2/3)$, with respect to r. Then we can choose N_0 sufficiently large such that for any $r\in(0,s)$ and $N>N_0$, $f_N^+(r)<1$. So

$$\mathbb{E}\left[q\left(x+W_{T_N^+}, t+s-T_N^+\right)\mathbf{1}_{[s\geq T_N^+]}\right] \;=\; \int_0^s q(x+N,t+s-r)f_N^+(r)\mathrm{d}r < \; q^*(x+N),$$

where $q^*(x) := \int_t^{t+s} q(x,r)\mathrm{d}r$, and

$$s \;\geq\; \int_t^{t+s}\int_{\mathbb{R}} q(x,r)\mathrm{d}x\mathrm{d}r \;\geq\; \int_1^{\infty} q^*(x)\mathrm{d}x + \int_{-\infty}^{-1} q^*(x)\mathrm{d}x$$
$$= \sum_{N=1}^{\infty}\int_0^1 q^*(x+N)\mathrm{d}x + \sum_{N=1}^{\infty}\int_0^1 q^*(x-N)\mathrm{d}x$$
$$= \int_0^1 \sum_{N=1}^{\infty} q^*(x+N)\mathrm{d}x + \int_0^1 \sum_{N=1}^{\infty} q^*(x-N)\mathrm{d}x,$$

where the last equality holds because $q^* \geq 0$. Thus, $\sum_{N=1}^{\infty} q^*(x+N) < \infty$ and $\sum_{N=1}^{\infty} q^*(x-N) < \infty$ for (Lebesgue) almost all x in (0,1). Therefore,

$$\lim_{N\to\pm\infty} q^*(x+N) = 0, \quad \text{a.e. } x \in \mathbb{R}, \tag{2.3.10}$$

and $\lim_{N\to\infty}\mathbb{E}\left[q\left(x+W_{T_N^+}, t+s-T_N^+\right)\mathbf{1}_{[s\geq T_N^+]}\right] = 0$, a.e.. The same result holds for T^-, so the second term on the right-hand side of (2.3.9) vanishes as $N\to\infty$ and

$$q(x,t+s) \;=\; \mathbb{E}[q(x+W_s,t)\mathbf{1}_{[s<\sigma_D]}] \;\leq\; \mathbb{E}[q(x+W_s,t)], \quad \text{a.e. } x, \tag{2.3.11}$$

which implies, $\hat{\mu}_{t+s} \;\leq\; \hat{\mu}_t P_s$. Note that the smoothness of $q(x,t)$ in D means that the identity must in fact hold for all x.

Step 2:

We now check the remaining conditions of Lemma 2.3.5. We begin by showing the right continuity of $\hat{\mu}_t$ at $t=0$. Since $\mu_s \Rightarrow \mu_0 := \delta_0$, as $s\downarrow 0$, for any closed set $A\in\mathcal{B}(\mathbb{R})$, if $0\notin A$, we have

$$\limsup_{s\downarrow 0}\hat{\mu}_s(A) \;\leq\; \lim_{s\downarrow 0}\mu_s(A) \;=\; \mu_0(A) \;=\; 0, \text{ i.e. } \lim_{s\downarrow 0}\hat{\mu}_s(A) \;=\; 0.$$

For any $h>0$ sufficiently small that $(-h,h)\times(0,h)\subset D$ (such an h exists since we would otherwise we would have $R(0)=0$ and the result is trivial), $s<h$, and $f\in\mathbb{C}_b^0$,

$$\int_{\mathbb{R}} f(x)\hat{\mu}_s(\mathrm{d}x) \;=\; \int_{[h,\infty)} f(x)\hat{\mu}_s(\mathrm{d}x) + \int_{(-\infty,-h]} f(x)\hat{\mu}_s(\mathrm{d}x) + \int_{(-h,h)} f(x)\hat{\mu}_s(\mathrm{d}x),$$

where the first two terms vanish as $s \downarrow 0$. Therefore,

$$\lim_{s\downarrow 0}\int_{\mathbb{R}} f(x)\hat{\mu}_s(\mathrm{d}x) \;=\; \lim_{s\downarrow 0}\int_{-h}^{h} f(x)\mu_s(\mathrm{d}x) \;=\; \lim_{s\downarrow 0}\int_{\mathbb{R}} f(x)\mu_s(\mathrm{d}x) = f(0),$$

where the first equality holds because $\hat{\mu}_s(-h,h) = \mu_s(-h,h)$ for $s < h$. Hence $\hat{\mu}_s \Rightarrow \mu_0$ as $s \downarrow 0$.

Now recall (2.3.11), and let $t \downarrow 0$. Since q is continuous on the open set D and $\hat{\mu}_t \Rightarrow \mu_0$, we have $\hat{\mu}_s \leq P_s$, and hence $\hat{\mu}_t P_s \leq P_{t+s}$, so $\hat{\mu}_.$ satisfies (2.3.7b).

Now, to apply Lemma 2.3.5, we only need to check right weak continuity of $\{\hat{\mu}_t\}_{t>0}$. For $t > 0$, consider $f \in \mathbb{C}_b^0$, and without loss of generality, assume $f \leq 1$. Then

$$\left|\int_{\mathbb{R}} f(x)\hat{\mu}_{t+s}(\mathrm{d}x) - \int_{\mathbb{R}} f(x)\hat{\mu}_t(\mathrm{d}x)\right| \;\leq\; \int_{\mathbb{R}} |q(x,t+s) - q(x,t)|\mathrm{d}x.$$

Since $\hat{\mu}_{t+s} \leq P_{t+s}$, $q(x, t+s)$ is dominated by the density of W_{t+s}, and by the dominated convergence theorem, the right side vanishes as $s \downarrow 0$. Therefore $\hat{\mu}_.$ also satisfies (2.3.7a), and so there exists $H(\cdot, t) \in \mathcal{H}$ and a randomised stopping time T^H such that

$$\langle \hat{\mu}_t, f\rangle \;=\; \mathbb{E}[f(W_t);\ T^H > t], \quad t \geq 0,\ f \in \mathbb{C}_b^0. \tag{2.3.12}$$

Step 3:
To show T^H in (2.3.12) is equal to τ_D, we analyse Rost's proof of Lemma 2.3.5. Note that for any $(x,t) \in D$, and s sufficiently small, we must have $q(x,t), q(x,t+s) > 0$. Define $h_{t,s}(x) := \mathrm{d}\hat{\mu}_{t+s}/\mathrm{d}(\hat{\mu}_t P_s)$, and hence $0 \leq h_{t,s} \leq 1$ and we have

$$h_{t,s}(x)\mathbb{E}\left[q(x+W_s,t)\right] \;=\; q(x,t+s), \quad \text{a.e. in } x \in \mathbb{R}.$$

For $(x, t+s) \in D$ and s sufficiently small, since q is parabolic in D, then

$$\begin{cases} q(x,t+s) \;=\; q(x,t) + s\dfrac{\partial q}{\partial t}(x,t) + o(s); \\ \mathbb{E}[q(x+W_s,t)] \;=\; q(x,t) + s\dfrac{1}{2}\dfrac{\partial^2 q}{\partial x^2}(x,t) + o(s) \;=\; q(x,t) + s\dfrac{\partial q}{\partial t}(x,t) + o(s), \end{cases}$$

and, since $q(x,t) > 0$,

$$h_{t,s}(x) = 1 - o(s), \qquad \text{for almost all } x \text{ such that } (x,t+s) \in D. \tag{2.3.13}$$

For t binary, i.e. $t = 2^{-n}m$, $m, n \in \mathbb{N}$, denote

$$H^{(n)}(t) \ := \ \prod_{k=0}^{t\cdot 2^n} \Big[h_{(k-1)\cdot 2^{-n}, 2^{-n}} \left(W_{k\cdot 2^{-n}} \right) \Big].$$

Now given $\bar{\omega} \in \Omega$ and t binary such that $t < \tau_D(\bar{\omega})$, then for $k = 0, 1, 2, \cdots, t \cdot 2^n$ where n is sufficiently large that $t \cdot 2^n \in \mathbb{N}$, we have $(2^{-n}k, W_{2^{-n}k}(\bar{\omega})) \in D$. Thus, for any $t \in [0, \infty)$, by (2.3.13),

$$\big[h_{(k-1)\cdot 2^{-n}, 2^{-n}} \left(W_{k\cdot 2^{-n}}(\bar{\omega}) \right) \big]^{t\cdot 2^n} \ = \ \big[1 - o(2^{-n}) \big]^{t\cdot 2^n} \ \xrightarrow{\text{a.s.}} 1, \qquad \text{as } n \uparrow \infty. \quad (2.3.14)$$

According to Rost's proof, the function H in (2.3.12) is the right-continuous modification of a weak* limit of $H^{(n)}$ as $n \to \infty$. Therefore, by right-continuity of H, (2.3.14) implies that for any $t < \tau_D(\bar{\omega})$, then $H(\bar{\omega}, t) = 1$. Moreover, by the definition of H and T^H, we have,

$$1 \ = \ H(\bar{\omega}, t) \ = \ \int_0^1 \mathbf{1}_{[s \le H(\bar{\omega}, t)]} \mathrm{d}s \ = \ \int_0^1 \mathbf{1}_{[t < T_s^H]}(\bar{\omega}) \mathrm{d}s,$$

thus, under Lebesgue measure, for almost all $s \in (0, 1)$, $t < T_s^H$, and hence we can conclude that for any $t \ge 0$, $t < \tau_D(\bar{\omega})$ implies $t < T_s^H(\bar{\omega})$, so $\tau_D \le T_s^H$, a.s..

On the other hand, consider $g(\cdot) = \mathbf{1}_{[(\cdot, t) \in D]}$. Then (2.3.12) still holds for g by the dominated convergence theorem. By the definition of $\hat{\mu}_t$ and (2.3.12),

$$\int_0^1 \mathbb{E}\mathbf{1}_{[t < T_s^H]} \mathrm{d}s \ = \ \langle \hat{\mu}_t, 1 \rangle \ = \ \langle \hat{\mu}_t, g \rangle \ = \ \int_0^1 \mathbb{E} \left[\mathbf{1}_{[(W_t, t) \in D]} \mathbf{1}_{[t < T_s^H]} \right] \mathrm{d}s,$$

and so, for almost all s, $\mathbf{1}_{[(W_t,t)\in D]}\mathbf{1}_{[t < T_s^H]} = \mathbf{1}_{[t < T_s^H]}$, a.s.. Therefore, $t < T_s^H$ implies $(W_t, t) \in D$, and moreover for any $r < t$, $(W_r, r) \in D$, i.e. $t < \tau_D$. So $T_s^H \le \tau_D$, a.s., and we can conclude that

$$T_s^H \ = \ \tau_D, \ \text{a.s., for almost all } s \in (0, 1),$$

and hence,

$$\langle \hat{\mu}_t, f \rangle \ = \ \mathbb{E}[f(W_t); \ T^H > t] \ = \ \mathbb{E}[f(W_t)\mathbf{1}_{[t < \tau_D]}], \quad t \ge 0, \ f \in \mathbb{C}_b^0.$$

□

We are now in a position to prove the main theorem in this section.

Proof of Theorem 2.3.3. Applying Proposition 2.3.4, for $(x,t) \in D$,

$$\mu_t(\mathrm{d}x) \;=\; \hat{\mu}_t(\mathrm{d}x) \;=\; \mathbb{P}[W_t \in \mathrm{d}x;\, t < \tau_D] \;=\; \mathbb{P}[W_{t\wedge\tau_D} \in \mathrm{d}x]. \tag{2.3.15}$$

Let $v(x,t) := -\mathbb{E}[|x - W_{t\wedge\tau_D}|]$, and recall Remark 2.2.4. Then by Theorem 2.2.1 and (2.3.15), $w := u - v$ is continuous on $\mathbb{R} \times [0,\infty)$ and satisfies

$$\begin{cases} \dfrac{\partial w}{\partial t}(x,t) \;=\; \dfrac{1}{2}\dfrac{\partial^2 w}{\partial x^2}(x,t) \;=\; 0, & \text{on } D\,; \\[2ex] w(x,0) \;=\; 0, & \text{for any } x \in \mathbb{R}\,. \end{cases}$$

Noting that for any $(x,t) \in D$ and $s \le t$, $(x,s) \in D$, we have

$$w(x,t) \;=\; \int_0^t \frac{\partial w}{\partial s}(x,s)\mathrm{d}s \;=\; 0\,,$$

and therefore, $w(x,t) = 0$ on $\partial_+ D$, i.e.,

$$-\mathbb{E}[|x - W_{t\wedge\tau_D}|] \;=\; u(x,t) \;\ge\; \mathrm{U}\mu.$$

By Chacon [1977, Lemma 5.1], we conclude that W^{τ_D} is uniformly integrable, i.e. τ_D is a UI stopping time, and hence,

$$\mathrm{U}\mu(x) \;=\; -\mathbb{E}[|x - W_{\tau_D}|],$$

and (v) of Theorem 2.1.6 allows us to conclude that $\tau_D \in \mathcal{T}(\mu)$. □

2.4 Sufficiency of the Vanishing Second Derivative

So far, we do not know that solutions to **SEP(μ)** satisfy (2.3.3) in general. Indeed, to understand the connection between **FBP(μ)** and **SEP(μ)**, it is necessary to verify (2.3.3) for general barriers. We are not in fact able to verify that (2.3.3) holds for all barriers, but rather we will restrict to a fairly large class of barriers. Specifically, we will identify the types of boundary points along which the density of the process will disappear, and thus the points at which the probabilistic solution will satisfy (2.3.3).

Before looking at the probabilistic solution, we note a trivial class of points x at which we do not need to verify (2.3.3). Suppose x satisfies

$$\text{for any } \delta > 0, \text{ there exist } y_1 \in (x-\delta, x) \text{ and } y_2 \in (x, x+\delta), \\ \text{such that } R(y_1), R(y_2) \le R(x). \tag{2.4.1}$$

then any Brownian trajectory starting at $(y, s) \in D$ and running backward could never exit D at $(x, R(x))$.

Now recall Example 2.3.2. For the barrier $\widetilde{D}$ generated by the boundary function $\widetilde{R}(y) = t$, for $y = x$ and ∞ otherwise. Let $\delta > 0$ and $r > \alpha\delta^\theta$ for some constants $\alpha > 0$ and $0 < \theta < 2$. Then:

$$\begin{aligned}
\widetilde{p}(x-\delta, t+r) &= \frac{\partial}{\partial y}\mathbb{P}\left[W_{t+r} < y, t+r < \tau_{\widetilde{D}}\right]\Big|_{y=x-\delta} \\
&= \frac{1}{\sqrt{2\pi(t+r)}}\Bigg[\exp\left\{-\frac{(x-\delta)^2}{2(t+r)}\right\}\Phi\left(\frac{rx}{\sqrt{rt(t+r)}} + \frac{\delta}{\sqrt{r}}\sqrt{\frac{t}{t+r}}\right) \\
&\qquad - \exp\left\{-\frac{(x+\delta)^2}{2(t+r)}\right\}\Phi\left(\frac{rx}{\sqrt{rt(t+r)}} - \frac{\delta}{\sqrt{r}}\sqrt{\frac{t}{t+r}}\right)\Bigg] \\
&\le \frac{1}{\sqrt{2\pi t}}\Bigg\{\left[\exp\left\{-\frac{(x-\delta)^2}{2(t+r)}\right\} - \exp\left\{-\frac{(x+\delta)^2}{2(t+r)}\right\}\right] \\
&\qquad + \left[\Phi\left(\frac{rx}{\sqrt{rt(t+r)}} + \frac{\delta}{\sqrt{r}}\sqrt{\frac{t}{t+r}}\right) - \Phi\left(\frac{rx}{\sqrt{rt(t+r)}} - \frac{\delta}{\sqrt{r}}\sqrt{\frac{t}{t+r}}\right)\right]\Bigg\} \\
&\le \frac{1}{\sqrt{2\pi t}}\left\{\left[\frac{2x\delta}{t+r}\exp\left\{-\frac{x^2}{2(t+r)}\right\} + o(\delta)\right] + 2\left[\Phi\left(\frac{\delta}{\sqrt{r}}\sqrt{\frac{t}{t+r}}\right) - \Phi(0)\right]\right\} \\
&\le \sqrt{\frac{2}{\pi t}}\left[\frac{x}{t}\delta + \Phi\left(\frac{\delta^{1-\theta/2}}{\sqrt{\alpha}}\right) - \Phi(0) + o(\delta)\right] \longrightarrow 0, \quad \text{as } \delta \to 0, \text{uniformly in } r,
\end{aligned} \tag{2.4.2}$$

where we note that $\Phi(a+b) - \Phi(a-b) < \Phi(b) - \Phi(-b)$, for any $a, b > 0$. A similar result holds for $\delta < 0$, and $r > \alpha(-\delta)^\theta$.

Now consider a more general function R. By the computations in Example 2.3.2, for any boundary points $(x,t) \in \partial_+ D$, if $R(x) < t$, clearly:

$$\lim_{\substack{(y,s)\downdownarrows(x,t) \\ (y,s)\in D}} q(y,s) \;=\; 0,$$

where $q(y,s)\mathrm{d}y := \mathbb{P}\left[W_s \in \mathrm{d}y, s < \tau_D\right]$. Therefore, we need only consider $(x,t) \in \partial_+ D$ where $t = R(x)$. More generally, if we have a point on the boundary which is 'sufficiently close' to points (y,s) that we can use the estimate in (2.4.2), then we can also conclude that the density at (x,t) will converge to zero. Specifically, suppose that x does not satisfy (2.4.1) and we can find $\alpha > 0$, $\theta \in (0,2)$ such that for all $\beta > 0$, there is some $\delta \in (0,\beta)$ for which either

$$R(x) - R(x+\delta) > \alpha\delta^\theta, \text{or} \tag{2.4.3a}$$

$$R(x) - R(x-\delta) > \alpha\delta^\theta, \tag{2.4.3b}$$

we then can find a suitable sequence along which we may use the estimate (2.4.2) to get the desired conclusion as follows.

Suppose (2.4.3a) holds. We consider $(x-\delta, R(x)+r) \in D$ where $0 < \delta \to 0,\ r \downarrow 0$. For any $\varepsilon > 0$, we can choose δ^* such that for any $\delta < \delta^*$,

$$\sqrt{\frac{4}{\pi R(x)}}\left[\frac{2x+\delta}{R(x)}\delta + \Phi\left(\frac{\delta^{1-\theta/2}}{\sqrt{\alpha^*}}\right) - \Phi(0) + \delta\right] < \varepsilon,$$

where $\alpha^* = (1/2)^\theta\alpha$.

Now choose $\hat{\delta}$ such that for any $\delta < \hat{\delta}$, $R(x+\delta) \;>\; R(x)/2$. Such $\hat{\delta}$ exists on account of the lower semi-continuity of R. Take $\delta_0 \in \left(0, \frac{1}{2}\left(\delta^* \wedge \hat{\delta}\right)\right)$ such that (2.4.3a) holds. Then for any $\delta < \delta_0$ and $r > 0$, we have

$$R(x) + r - R(x+\delta_0) \;>\; R(x) - R(x+\delta_0) \;>\; \alpha\delta_0^\theta \;>\; \alpha^*(\delta+\delta_0)^\theta.$$

By the calculation in (2.4.2), (noting that the barrier is now at $(x+\delta_0, R(x+\delta_0))$ as

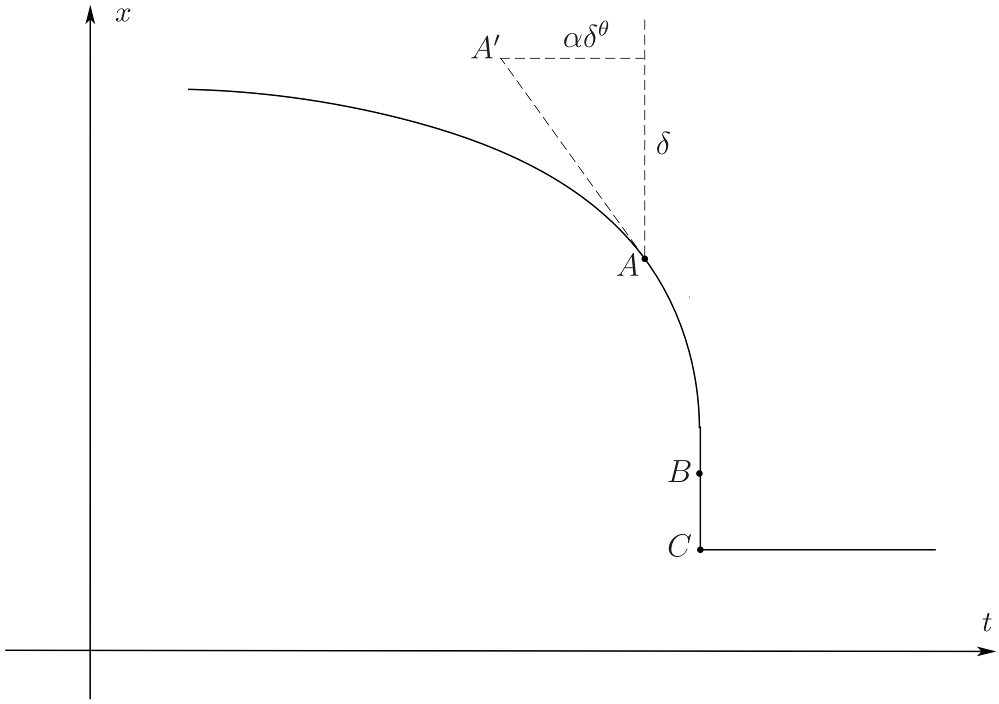


Figure 2-3: Classification of boundary points. *We can see A satisfies (2.4.3a) (take $\theta = 1$ in this figure), B satisfies (2.4.1), C is a point which does not satisfy any one among (2.4.1), (2.4.3a) and (2.4.3b). In addition, the points on the right of C satisfy $t > R(x)$.*

in Figure 2-4, and we must replace α with α^*), we have:

$$
\begin{aligned}
q(x-\delta,\, R(x)+r) \;&=\; \frac{\partial}{\partial y}\mathbb{P}\left[W_{R(x)+r} < y, R(x)+r < \tau_D\right]\Big|_{y=x-\delta} \\
&\le \sqrt{\frac{2}{\pi R(x+\delta_0)}}\left[\frac{x+\delta_0}{R(x+\delta_0)}(\delta+\delta_0) + \Phi\left(\frac{(\delta+\delta_0)^{1-\theta/2}}{\sqrt{\alpha^*}}\right) - \Phi(0) + o(\delta)\right] \\
&\le \sqrt{\frac{2}{\pi R(x)/2}}\left[\frac{x+\delta_0}{R(x)/2}(\delta+\delta_0) + \Phi\left(\frac{(\delta+\delta_0)^{1-\theta/2}}{\sqrt{\alpha^*}}\right) - \Phi(0) + o(\delta)\right] \\
&\le \sqrt{\frac{4}{\pi R(x)}}\left[\frac{2(x+\delta_0)}{R(x)}(2\delta_0) + \Phi\left(\frac{(2\delta_0)^{1-\theta/2}}{\sqrt{\alpha^*}}\right) - \Phi(0) + o(\delta)\right] \;<\; \varepsilon.
\end{aligned}
\tag{2.4.4}
$$

Then for x satisfying (2.4.3a),

$$\lim_{\substack{(y,s)\downdownarrows(x,R(x))\\(y,s)\in D}} q(y,s) \;=\; 0,$$

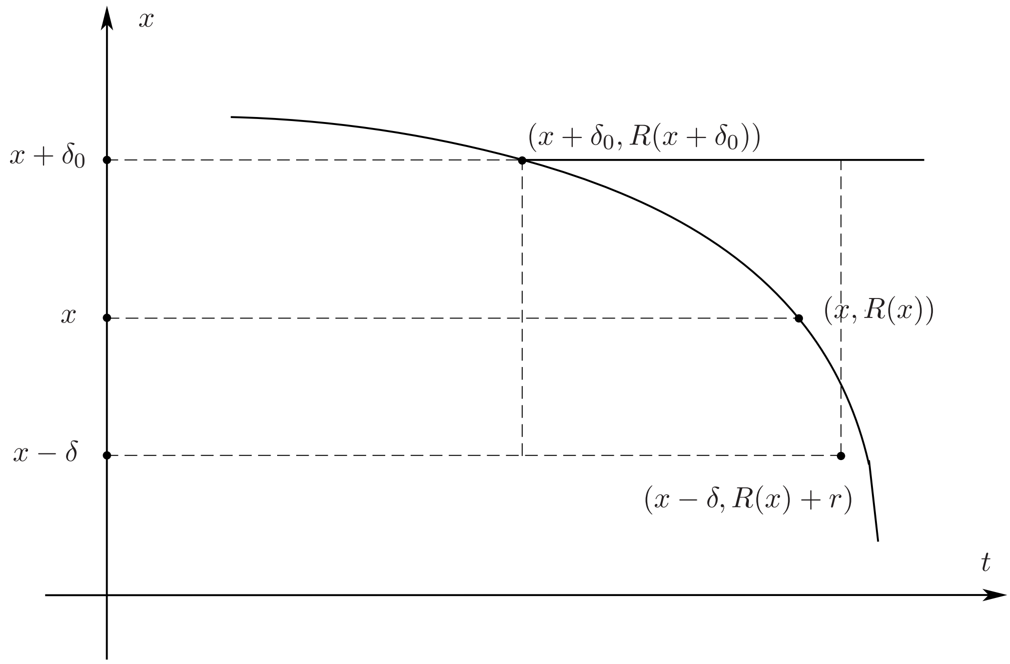


Figure 2-4: Illustration of the calculation of vanishing density. *In (2.4.4), we use the estimation obtained in (2.4.2), where we replace (x,t) by $(x+\delta_0, R(x+\delta_0))$ and replace (δ, r) by $(\delta+\delta_0, R(x)+r-R(x+\delta_0)$.*

and L^1 convergence is straightforward since $q(y,s)$ is dominated by the density of Brownian motion, which is continuous and thus bounded on any compact subset of $\{(x,t) : x \in \mathbb{R}, t > 0\}$.

A similar argument leads to the same conclusion for $(x, R(x))$ where x satisfies (2.4.3b). For the boundary function R, let U^R, V_+^R and V_-^R be the collections of points $x \in \mathbb{R}$ which satisfy (2.4.1), (2.4.3a) and (2.4.3b) respectively. Then the following theorem can be regarded as a stronger version of Theorem 2.2.1: it gives a sufficient condition that the solution to $\mathbf{SEP}(\boldsymbol{\mu})$ and the corresponding potential process satisfy (2.3.3).

Theorem 2.4.1. *Suppose D solves $\mathbf{SEP}(\boldsymbol{\mu})$, with boundary function R. Moreover, assume*

$$\mathbb{R} \backslash (U^R \cup V_+^R \cup V_-^R) \text{ is a countable set,} \tag{2.4.5}$$

then the couple (u, D) solves $\mathbf{FBP}(\boldsymbol{\mu})$ and satisfies (2.3.3) where $u(x,t) = -\mathbb{E}[\,|x - W_{t\wedge\tau_D}|]$.

Proof. This result now follows from the above identities: for any points in ∂D_+ which are in any of the sets U^R, V_+^R or V_-^R, the boundary is either inaccessible for $(y, s) \in D$ (as is the case for U^R or e.g. $x \in V_+^R$ where $y > x$), or we have shown that (2.3.3) will hold. In addition, for points $(y, s) \in \partial D_+$ where $s > R(y)$, by (2.4.2), we see that (2.3.3) holds, and therefore the only remaining points on the boundary where (2.3.3) can fail are the points $(x, R(x))$ where x is in the set $\mathbb{R} \backslash (U^R \cup V_+^R \cup V_-^R)$. However this is a countable subset of $\mathbb{R} \times \mathbb{R}_+$, and so will not be hit by the process with probability one. Hence (2.3.3) does indeed hold for the solution generated by the stopping time τ_D. □

2.5 Extensions to Diffusions

So far, we have considered $\mathbf{SEP}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ and $\mathbf{FBP}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ in the context of a standard Brownian motion. In this section, we show that our results extend to the case where the underlying process is the solution to the time-homogeneous stochastic differential equation as (2.1.1):

$$\begin{cases} \mathrm{d}X_t \;=\; \sigma(X_t)\mathrm{d}W_t\,; \\ X_0 \;=\; \xi \;\sim\; \nu\,. \end{cases} \tag{2.5.1}$$

Initially, we assume that $\sigma : \mathbb{R} \to \mathbb{R}$ satisfies, for some positive constant $K > 0$,

$$|\,\sigma(x) - \sigma(y)\,| \;\leq\; K|\,x - y|\,; \tag{2.5.2a}$$

$$0 < \sigma(x) < \sqrt{K(1 + x^2)}\,; \tag{2.5.2b}$$

$$\sigma \text{ is smooth.} \tag{2.5.2c}$$

From standard results on SDEs, (2.5.2a) and (2.5.2b) imply that the unique strong solution X of (2.5.1) is a strong Markov process with the infinitesimal generator $\frac{1}{2}\sigma^2\partial_{xx}$ for any initial value $a \in \mathbb{R}$. And moreover, (2.5.2c) implies that the operator $L := \frac{1}{2}\sigma^2\partial_{xx} - \partial_t$ is hypo-elliptic (see Stroock [2008, Theorem 3.4.1]).

We study the potential process of the stopped process X^{τ_D}, where $\tau_D \in \mathcal{T}(\sigma, \delta_m, \mu)$ is Root's stopping time:

$$u(x, t) \;:=\; -\mathbb{E}^\nu[|x - X_{t\wedge\tau_D}|].$$

It is easy to check that Lemma 2.2.2 and Lemma 2.2.3 still hold in the diffusion context, so that corresponding density $p^D(x,t)$ remains smooth. Then, following the beginning of the proof of Theorem 2.2.1, we have

$$-\frac{1}{2}\int_0^t \frac{\partial^2}{\partial x^2}u(x,s)\mathrm{d}s \;=\; \lim_{\varepsilon\downarrow 0}\mathbb{E}^\nu\left[\frac{1}{2\varepsilon}\int_0^{t\wedge\tau_D}\mathbf{1}_{[x-\varepsilon<X_s<x+\varepsilon]}\mathrm{d}s\right].$$

By (2.5.2a) and (2.5.2b), we have

$$\frac{1}{2\varepsilon}\int_0^{t\wedge\tau_D}\left|\sigma^2(x)-\sigma^2(X_s)\right|\mathbf{1}_{[x-\varepsilon<X_s<x+\varepsilon]}\mathrm{d}s \;\leq\; 2tK^2(1+(x+K\varepsilon)^2),$$

so by the dominated convergence theorem,

$$\begin{aligned}-\frac{\sigma^2(x)}{2}\int_0^t \frac{\partial^2}{\partial x^2}u(x,s)\mathrm{d}s \;&=\; \lim_{\varepsilon\downarrow 0}\mathbb{E}^\nu\left[\frac{1}{2\varepsilon}\int_0^{t\wedge\tau_D}\sigma^2(X_s)\mathbf{1}_{[x-\varepsilon<X_s<x+\varepsilon]}\mathrm{d}s\right]\\ &=\; \mathbb{E}^\nu[L^x_{t\wedge\tau_D}] \;=\; \mathbb{E}^\nu[|x-X_{t\wedge\tau_D}|]+\mathrm{U}\nu,\end{aligned}$$

here, L^x is the local time of the diffusion X, satisfying Tanaka's formula:

$$\mathrm{d}|X_t-x| \;=\; \mathrm{sgn}(X_t-x)\mathrm{d}X_t+\mathrm{d}L^x_t.$$

It follows that u satisfies (2.2.1b). The other arguments in the proof of Theorem 2.2.1 will go through in the diffusion setting without significant alteration:

Theorem 2.5.1. *Suppose the diffusion X solves (2.5.1) where (2.5.2a) – (2.5.2c) hold. If D solves* $\mathbf{SEP}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$ *with* $\mathrm{U}\mu\leq\mathrm{U}\nu$*, with corresponding UI stopping time τ_D, i.e. $\tau_D\in\mathcal{T}(\sigma,\nu,\mu)$. Then the couple (u,D), where $u:=-\mathbb{E}^\nu[|\,x-X_{t\wedge\tau_D}|]$, solves* $\mathbf{FBP}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$.

2.5.1 Uniqueness for Diffusions

We consider the uniqueness of the solution to $\mathbf{FBP}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$. As in Section 2.3, if (u,D) solves $\mathbf{FBP}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$, by Lemma 2.3.1, for any $t\geq 0$, there is a probability distribution μ_t on $\mathbb{R}$, satisfying $u(\cdot,t)=\mathrm{U}\mu_t$, and hence, (2.3.1) still holds.

Moreover, different from the Brownian case, we assume through this section that

$$\mu \text{ is a distribution with finite variance satisfying } \mathrm{U}\mu\leq\mathrm{U}\nu. \qquad (2.5.3)$$

By (iv) of Theorem 2.1.6,

$$\text{for any } t > 0, \quad \int_{\mathbb{R}} x^2 \mu_t(\mathrm{d}x) \leq \int_{\mathbb{R}} x^2 \mu(\mathrm{d}x) < \infty. \tag{2.5.4}$$

Again, under the assumptions (2.5.2a) – (2.5.2c), let X be the unique strong solution to the (2.5.1), and we have the Kolmogorov forward and backward equations with $p_t(x, y)$, the transition density of X,

$$\frac{\partial}{\partial t} p_t(x, y) \;=\; \frac{1}{2}\sigma^2(x) \frac{\partial^2}{\partial x^2} p_t(x, y) \;=\; \frac{1}{2}\frac{\partial^2}{\partial y^2}[\sigma^2(y) p_t(x, y)]\,; \tag{2.5.5}$$

$$\lim_{t \downarrow 0} p_t(\cdot, y) \;=\; \lim_{t \downarrow 0} p_t(y, \cdot) \;=\; \delta_y\,, \tag{2.5.6}$$

and the following relation is valid:

Proposition 2.5.2. *Suppose (2.5.2a) – (2.5.2c) hold, then*

$$\sigma^2(y) p_t(x, y) \;=\; \sigma^2(x) p_t(y, x), \quad \textit{for all} \;\; (x, y, t) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+.$$

This result appeared in Itô and McKean [1974], also see a more direct proof in Ekström and Tysk [2011]. The following proof was independently derived.

Proof. Fix some $y_0 \in \mathbb{R}$ and let $h(x, t) := \sigma^2(y_0) p_t(x, y_0) - \sigma^2(x) p_t(y_0, x)$, the proof is complete if we can show $h(x, t) \equiv 0$ on $\mathbb{R} \times \mathbb{R}_+$. By (2.5.5), we have

$$\frac{\partial}{\partial t} h(x, t) \;=\; \frac{1}{2}\sigma^2(x) \left[\sigma^2(y_0) \frac{\partial^2}{\partial x^2} p_t(x, y_0) - \frac{\partial^2}{\partial x^2}\big(\sigma^2(x) p_t(y_0, x)\big)\right] \;=\; \frac{1}{2}\sigma^2(x) \frac{\partial^2}{\partial x^2} h(x, t).$$

Now we consider the initial value of h, $\lim_{t \downarrow 0} g(x, t)$. For any $f \in \mathbb{C}_b^0$, as $t \downarrow 0$,

$$\int_{\mathbb{R}} f(x) \sigma^2(x) p_t(y_0, x) \mathrm{d}x \;\longrightarrow\; \sigma^2(y_0) f(y_0).$$

On the other hand, by (2.5.6), $\sigma^2(y_0) \int_{\mathbb{R}} f(x) p_t(x, y_0) \mathrm{d}x$ also converges to $\sigma^2(y_0) f(y_0)$. Therefore $h(x, t)$ satisfies

$$\begin{cases} \dfrac{\partial h}{\partial t} = \dfrac{1}{2}\sigma^2(x) \dfrac{\partial^2 h}{\partial x^2}, & \text{on } \mathbb{R} \times \mathbb{R}_+; \\[2ex] h(x, 0) \;\equiv\; 0, & \text{for any } x \in \mathbb{R}, \end{cases}$$

and hence, by Øksendal [1995, Theorem 8.1], $h(x,t) \equiv 0$ on $\mathbb{R} \times \mathbb{R}_+$. $\square$

As previously, define D_t as in Section 2.3, and

$$D_0 \;:=\; \bigcup_{t>0} D_t \;=\; \lim_{t\downarrow 0} D_t\,, \qquad \hat{\mu}_0(\mathrm{d}x) \;:=\; \mathbf{1}_{D_0}(x)\nu(\mathrm{d}x), \tag{2.5.7}$$

and then define the sub-probability measures,

$$\hat{\mu}_t(\mathrm{d}x) := \begin{cases} q(x,t)\mathrm{d}x, & t>0; \\ \hat{\mu}_0(\mathrm{d}x), & t=0, \end{cases} \quad \text{where} \quad q(x,t) := -\frac{1}{2}\mathbf{1}_{(x,t)\in D}\frac{\partial^2 u}{\partial x^2}(x,t). \tag{2.5.8}$$

In addition, we define $\tilde{q}(x,t) = \sigma^2(x)q(x,t)$. Then for $(x,t) \in D$, since σ and q are sufficiently smooth on D,

$$\frac{\sigma^2(x)}{2}\frac{\partial^2 \tilde{q}}{\partial x^2}(x,t) \;=\; -\frac{\sigma^2(x)}{2}\frac{\partial^2}{\partial x^2}\left(\frac{\partial u}{\partial t}\right) \;=\; -\frac{\partial}{\partial t}\left(\frac{\sigma^2(x)}{2}\frac{\partial^2 u}{\partial x^2}\right) \;=\; \frac{\partial \tilde{q}}{\partial t}. \tag{2.5.9}$$

As in Section 2.3, we introduce the additional condition: we assume D and q satisfy, for any fixed point $(y,s) \in D$, $q(x,t)$ converges almost surely and in L^1 to 0 along trajectories of $(X_t^y, s-t)$ at $\partial_+ D = \partial D \cap \{(x,t) : t > 0\}$. Specifically, writing $\sigma_D = \inf\{t \geq 0 : (X_t^y, s-t) \in \partial_+ D$, if σ_N is a sequence of stopping times increasing to σ_D as $N \to \infty$, then

$$q(X^y_{\sigma_N}, s-\sigma_N)\mathbf{1}_{\{\sigma_D<s\}} \xrightarrow{a.s. \text{ and in } L^1} 0. \tag{2.5.10}$$

With this condition, we are able to give a version of Theorem 2.3.3 in the diffusion setting.

Theorem 2.5.3. *Suppose the diffusion X satisfies (2.5.1) where (2.5.2a) – (2.5.2c) hold. Suppose also μ satisfies (2.5.3). If the couple (u, D) solves* $\mathbf{FBP}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ *and satisfies (2.5.10), then D solves* $\mathbf{SEP}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$*, and $u(x,t) = -\mathbb{E}^\nu[|x - X_{t\wedge\tau_D}|]$.*

Our proof will use the same ideas as the proof of Theorem 2.3.3 and Proposition 2.3.4. Below, we highlight where additional care needs to be taken.

Proof. As in the first step of Proposition 2.3.4, for any $(x, t+s) \in D$ where $t>0,\ s>0$, by Itô's formula, we have that

$$\begin{aligned}\tilde{q}(X^x_{s\wedge\sigma_D}, t+s-s\wedge\sigma_D) &= \tilde{q}(x,t+s) - \int_0^{s\wedge\sigma_D} \frac{\partial}{\partial t}\tilde{q}(X^x_r, t+s-r)\mathrm{d}r \\ &\quad + \int_0^{s\wedge\sigma_D} \frac{\partial}{\partial x}\tilde{q}(X^x_r, t+s-r)\mathrm{d}X_r \\ &\quad + \frac{1}{2}\int_0^{s\wedge\sigma_D} \frac{\partial^2}{\partial x^2}\tilde{q}(X^x_r, t+s-r)\mathrm{d}\left\langle X^x\right\rangle_r,\end{aligned}$$

where σ_D is as defined in Section 2.3, and $\langle X^x\rangle$ denotes the quadratic variation of X^x. By (2.5.9) and (2.5.10), we have, for $T_N := T_N^+ \wedge T_N^-$ where $T_N^\pm = \inf\{t \geq 0 : X_t - X_0 = \pm N\}$

$$\tilde{q}(x, t+s) = \mathbb{E}^x\left[\tilde{q}(X_{s\wedge\sigma_D\wedge T_N}, t+s-s\wedge\sigma_D\wedge T_N)\right].$$

To show

$$\tilde{q}(x, t+s) = \mathbb{E}^x\left[\tilde{q}(X_s, t)\mathbf{1}_{s<\sigma_D}\right], \quad \text{a.e. } x \in \mathbb{R}, \tag{2.5.11}$$

we only need to show

$$\lim_{N\to\infty} \mathbb{E}^x\left[\tilde{q}\left(x \pm N, t+s-T_N\right)\mathbf{1}_{s\geq T_N}\right] = 0, \quad \text{a.e. } x \in \mathbb{R}. \tag{2.5.12}$$

Since (2.5.2b) and (2.5.4) hold, we have,

$$\begin{aligned}\int_t^{t+s}\int_{\mathbb{R}} \tilde{q}(x,r)\mathrm{d}x\mathrm{d}r &\leq \int_t^{t+s}\int_{\mathbb{R}} K(1+x^2)q(x,r)\mathrm{d}x\mathrm{d}r \\ &\leq Ks\left(1+\int_{\mathbb{R}} x^2\mu_t(\mathrm{d}x)\right) \leq Ks\left(1+\int_{\mathbb{R}} x^2\mu(\mathrm{d}x)\right) < \infty,\end{aligned}$$

then, as in (2.3.10), we have

$$\int_t^{t+s} \tilde{q}(x+N, r)\mathrm{d}r \longrightarrow 0, \quad \text{as } N \to \pm\infty, \quad \text{a.e. } x \in \mathbb{R}$$

On the other hand, if we define the process A,

$$A(t) = \int_0^t \sigma^2(X_s)\mathrm{d}s, \tag{2.5.13}$$

and $\widetilde{W}_t := X_{A^{-1}(t)}$ then $\widetilde{W}$ is a Brownian motion starting at X_0. Moreover define

$a_N := |x+N| \vee |x-N|$ and

$$\widetilde{T}_N := \inf\left\{t>0 : |\widetilde{W}_t - \widetilde{W}_0| \geq N\right\}.$$

It is easy to see that $\widetilde{T}_N = A(T_N)$. Thus, for any $r>0$, $\varepsilon>0$, by (2.5.2b),

$$\begin{aligned}
&\mathbb{P}^x\left[T_N \in [r, r+\varepsilon)\right] \\
=&\mathbb{P}^x\left[T_N \in [r,r+\varepsilon);\, \widetilde{T}_N \in \left[A(r), A(r)+\int_r^{r+\varepsilon}\sigma^2(X_s)\mathrm{d}s\right)\right] \\
\leq&\mathbb{P}^x\left[T_N \in [r,r+\varepsilon);\, \widetilde{T}_N \in \left[A(r), A(r)+K\varepsilon+K\int_r^{r+\varepsilon}X_s^2\mathrm{d}s\right)\right] \\
=&\mathbb{P}^x\left[T_N \in [r,r+\varepsilon);\, \widetilde{T}_N \in \left[A(r), A(r)+K(1+a_N^2)\varepsilon+K\int_{T_N}^{T_N+\varepsilon}(X_s^2-X_{T_N}^2)\mathrm{d}s\right)\right] \\
\leq&\mathbb{P}^x\left[T_N \in [r,r+\varepsilon);\, \widetilde{T}_N \in \left[A(r), A(r)+K(1+a_N^2)\varepsilon+K\int_{T_N^+}^{T_N^++\varepsilon}(X_s^2-X_{T_N^+}^2)\mathrm{d}s\right)\right] \\
&+\mathbb{P}^x\left[T_N \in [r,r+\varepsilon);\, \widetilde{T}_N \in \left[A(r), A(r)+K(1+a_N^2)\varepsilon+K\int_{T_N^-}^{T_N^-+\varepsilon}(X_s^2-X_{T_N^-}^2)\mathrm{d}s\right)\right].
\end{aligned}$$

Noting that for any $r>0$, $T_N \geq r$ implies $X_s^2 < a_N^2$ for $s<r$, we have, by the strong Markov property,

$$\begin{aligned}
&\mathbb{P}^x\left[T_N \in [r,r+\varepsilon);\, \widetilde{T}_N \in \left[A(r), A(r)+K(1+a_N^2)\varepsilon+K\int_{T_N^+}^{T_N^++\varepsilon}(X_s^2-X_{T_N^+}^2)\mathrm{d}s\right)\right] \\
\leq&\ \mathbb{P}^x\Bigg[A(r) < K(1+a_N^2)r;\, \widetilde{T}_N \in \Big[A(r), A(r)+K(1+a_N^2)\varepsilon \\
&\qquad + 2a_N K\int_{T_N^+}^{T_N^++\varepsilon}|X_s-(x+N)|\,\mathrm{d}s + K\int_{T_N^+}^{T_N^++\varepsilon}|X_s-(x+N)|^2\,\mathrm{d}s\Big)\Bigg] \\
\leq&\ \int_0^{K(1+a_N^2)r}\mathbb{P}^x\left[\widetilde{T}_N \in [h,\, h+K(1+4a_N^2)\varepsilon)\right]\mathbb{P}^x[A(r)\in \mathrm{d}h] \\
&\qquad + \mathbb{P}^{x+N}\left[\sup_{s\in(0,\varepsilon)}|X_s-(x+N)| > |x+N|\right] \\
\leq&\ \int_0^{K(1+a_N^2)r}\mathbb{P}^x\left[\widetilde{T}_N \in [h,\, h+K(1+4a_N^2)\varepsilon)\right]\mathbb{P}^x[A(r)\in \mathrm{d}h] + \frac{\mathbb{E}^0 Y_\varepsilon^2}{(x+N)^2},
\end{aligned}$$

where Y satisfies the stochastic equation

$$\mathrm{d}Y_s = \sigma(Y_s+(x+N))\mathrm{d}W_s,$$

and then

$$\mathbb{E}^0 Y_\varepsilon^2 \;\le\; K\mathbb{E}^0\left[\int_0^\varepsilon \left[1+(Y_s+x+N)^2\right]\mathrm{d}s\right] \;\le\; K[1+2(x+N)^2]\varepsilon + 2K\int_0^\varepsilon \mathbb{E}^0 Y_s^2\mathrm{d}s,$$

by Gronwall's inequality, we have

$$\mathbb{E}^0 Y_\varepsilon^2 \;\le\; \frac{1+2(x+N)^2}{2}\left(e^{2K\varepsilon}-1\right).$$

Therefore,

$$\begin{aligned}
&\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\,\mathbb{P}^x\Bigg[T_N\in[r,r+\varepsilon);\;\widetilde{T}_N\in\Big[A(r),A(r)+K(1+a_N^2)\varepsilon\\
&\qquad\qquad\qquad\qquad\qquad\qquad + K\int_{T_N^+}^{T_N^++\varepsilon}(X_s^2-X_{T_N^+}^2)\mathrm{d}s\Big)\Bigg]\\
&\le \int_0^{K(1+a_N^2)r} K(1+4a_N^2)\tilde{f}_N(h)\,\mathbb{P}^x[A(r)\in\mathrm{d}h] + 2K\frac{1+2(x+N)^2}{2(x+N)^2}\\
&\le \int_0^{K(1+a_N^2)r} \frac{2K(1+4a_N^2)}{\sqrt{2\pi h^3}}e^{-\frac{N^2}{2h}}\,\mathbb{P}^x[A(r)\in\mathrm{d}h] + 2K\frac{1+2(x+N)^2}{2(x+N)^2},
\end{aligned}$$

where $\tilde{f}_N(h)$ is the density of $\widetilde{T}_N$. By the definition of a_N, it is easy to see that we can find r small enough and N large enough such that $K(1+a_N^2)r < N^2/3$, and then the integrand above is decreasing in h, so

$$\begin{aligned}
&\int_0^{K(1+a_N^2)r} \frac{2K(1+4a_N^2)}{\sqrt{2\pi h^3}}e^{-\frac{N^2}{2h}}\,\mathbb{P}^x[A(r)\in\mathrm{d}h] + 2K\frac{1+2(x+N)^2}{2(x+N)^2}\\
&\le \frac{2K(1+4a_N^2)e^{-3/2}}{\sqrt{2\pi(N^2/3)^3}}\int_0^{N^2/3}\mathbb{P}^x[A(r)\in\mathrm{d}h] + 2K\frac{1+2(x+N)^2}{2(x+N)^2}\\
&\le \frac{2K(1+4a_N^2)e^{-3/2}}{\sqrt{2\pi(N^2/3)^3}} + 2K\frac{1+2(x+N)^2}{2(x+N)^2} \;\longrightarrow\; 2K, \quad \text{as} \quad N\to\infty.
\end{aligned}$$

Now we can conclude that for $r < (3K)^{-1}$,

$$\lim_{N\to\infty}\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\mathbb{P}^x[T_N\in[r,r+\varepsilon)] \;\le\; 4K$$

so $f_N(r) < 4K+1$, for r small enough and N large enough. Then (2.5.12) follows, and then (2.5.11) is proved.

Using the fact that $\sigma^2(x)p_t(y,x) = \sigma^2(y)p_t(x,y)$ (Proposition 2.5.2), (2.5.11) implies that

$$\begin{aligned}\frac{\hat{\mu}_t P_s^X(\mathrm{d}x)}{\mathrm{d}x} &= \int_{\mathbb{R}} q(y,t)p_s(y,x)\mathrm{d}y = \frac{1}{\sigma^2(x)}\mathbb{E}^x[\tilde{q}(X_s,t)] \\ &\geq \frac{1}{\sigma^2(x)}\tilde{q}(x,t+s) = q(x,t+s), \qquad \text{a.e. } x \in \mathbb{R},\end{aligned} \tag{2.5.14}$$

where P^X is the transition semi-group of X. To apply Lemma 2.3.5, we then need to check the the right weak continuity of $\hat{\mu}_.$ at time 0. For any, $f \in \mathbb{C}_b^\infty$, w.l.o.g. we assume $|f| \leq 1$, the right weak continuity is implied by $\mu_t \Rightarrow \nu$ and

$$\begin{aligned}\left|\int_{\mathbb{R}} f\mathrm{d}\hat{\mu}_t - \int_{\mathbb{R}} f\mathrm{d}\hat{\mu}_0\right| &\leq \int_{D_0\backslash D_t} \mathrm{d}\mu_t + \left|\int_{D_0} f(\mathrm{d}\mu_t - \mathrm{d}\nu)\right| \\ &= \int_{D_0\backslash D_t} \mathrm{d}\mu + \left|\int_{D_0} f(\mathrm{d}\mu_t - \mathrm{d}\nu)\right|,\end{aligned}$$

where the equality holds because $\mu_t = \mu$ on $\mathbb{R}\backslash D_t$ by (v) of Theorem 2.1.6. The proof of the right weak continuity of $\{\hat{\mu}_t\}_{t>0}$ is same as in the Brownian case. We conclude $\{\hat{\mu}_t\}_{t\geq 0}$ satisfies (2.3.7a) and (2.3.7b). Therefore, there exists $H(\cdot,t) \in \mathcal{H}$ and the randomized stopping time T^H such that

$$\langle \hat{\mu}_t, f\rangle = \mathbb{E}^{\hat{\mu}_0}[f(X_t);\ T^H > t], \quad t \geq 0,\ f \in \mathbb{C}_b^0. \tag{2.5.15}$$

Now define $h_{t,s}(x) := \mathrm{d}\hat{\mu}_{t+s}/\mathrm{d}\hat{\mu}_t Q_s$. We have, noting (2.5.14), for almost all $x \in \mathbb{R}$,

$$q(x,t+s) = h_{t,s}(x)\frac{1}{\sigma^2(x)}\mathbb{E}^x[\tilde{q}(X_s,t)], \quad \text{i.e.} \quad h_{t,s}(x)\mathbb{E}^x[\tilde{q}(X_s,t)] = \tilde{q}(x,t+s),$$

and the same argument as for the Brownian case implies

$$T_s^H = \tau_D, \text{ a.s., for almost all } s \in (0,1),$$

and

$$\langle \hat{\mu}_t, f\rangle = \mathbb{E}^{\hat{\mu}_0}[f(X_t)\mathbf{1}_{[t<\tau_D]}], \quad t \geq 0,\ f \in \mathbb{C}_b^0.$$

Moreover, since the one-point sets are regular for X, if a trajectory of $X(\omega)$ starts at $y \notin D$, then $\tau_D(\omega) = 0$, and hence, the right-hand side of the equality above can be replaced by $\mathbb{E}^\nu[f(X_t)\mathbf{1}_{[t<\tau_D]}]$. Therefore, for $(x,t)\in D$,

$$\mu_t(\mathrm{d}x) \;=\; \hat{\mu}_t(\mathrm{d}x) \;=\; \mathbb{P}^\nu[X_t\in \mathrm{d}x,\, t<\tau_D] \;=\; \mathbb{P}^\nu[X_{t\wedge\tau_D}\in \mathrm{d}x],$$

and the desired result follows by repeating the proof of Theorem 2.3.3. □

2.5.2 Vanishing Second Derivative for Diffusions

We repeat the analysis of Section 2.4 for the diffusion case, and consider what additional condition would ensure the solution to $\mathbf{SEP}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$ satisfies (2.5.10).

Since (2.5.10) is a local property, without loss of generality, we study the sufficiency for the diffusion process under a stronger assumption:

$$\exists\, K > 0, \quad \text{such that} \quad \frac{1}{K} < \sigma(x) < K, \quad \text{for} \;\; x\in\mathbb{R}. \tag{2.5.16}$$

Then the following lemma is recalled (see Stroock [2008, Theorem 3.3.11] of the estimates of higher order derivatives).

Lemma 2.5.4 (Gaussian Estimates)**.** *Suppose that (2.5.2a) – (2.5.2c) and (2.5.16) hold. Then there exist constants $C_i<\infty$ such that, for $i=0,1$,*

$$\left(\frac{\partial}{\partial x}\right)^i p_t(x,y) \;\le\; \frac{C_i}{1\wedge t^{(1+i)/2}}\exp\left\{-\left(C_i t-\frac{|\,y-x|^2}{C_i t}\right)^-\right\},$$

where $p_t(x,y)$ is the transition density of X.

For the simple barrier D defined in Example 2.3.2, and any $\delta>0$ sufficiently small and $r>\alpha\delta^\theta$ for some constants $\alpha>0$ and $\theta\in(0,2)$, according to the time change (2.5.13), writing $\overline{X}_r := \sup_{h<r} X_h$ and $\overline{W}_r := \sup_{h<r} W_h$, we have, by Gaussian estimates,

$$\begin{aligned}
\widetilde{p}(x-\delta, t+r) &:= \frac{\partial}{\partial x}\mathbb{P}^0\left[X_{t+r} < x-\delta,\, t+r < \tau_D\right] \\
&= \int_{-\infty}^{x} p_t(0,z)\frac{\partial}{\partial x}\mathbb{P}^z\left[X_r < x-\delta,\, \overline{X}_r < x\right]\mathrm{d}z \\
&= \int_{-\infty}^{x} p_t(0,z)\frac{\partial}{\partial x}\mathbb{P}^z\left[W_{A(r)} < x-\delta,\, \overline{W}_{A(r)} < x\right]\mathrm{d}z \\
&= \int_{-\infty}^{x} p_t(0,z)\frac{\partial}{\partial x}\int_0^{K^2 r}\mathbb{P}^z\left[W_h < x-\delta,\, \overline{W}_h < x\right]\mathbb{P}^z[A(r)\in \mathrm{d}h]\mathrm{d}z \\
&\le \frac{C_0}{1\wedge\sqrt{t}}\int_{-\infty}^{x}\mathbb{E}^z\left[\frac{1}{\sqrt{2\pi A(r)}}\left(\exp\left\{-\frac{(x-z-\delta)^2}{2A(r)}\right\} - \exp\left\{-\frac{(x-z+\delta)^2}{2A(r)}\right\}\right)\right]\mathrm{d}z \\
&= \frac{C_0}{1\wedge\sqrt{t}}\int_{-\infty}^{x}\mathbb{E}^z\left[\int_{x-z-\delta}^{x-z+\delta}\frac{y}{\sqrt{2\pi A(r)^3}}\exp\left\{-\frac{y^2}{2A(r)}\right\}\mathrm{d}y\right]\mathrm{d}z \\
&\le \frac{C_0K^3}{1\wedge\sqrt{t}}\int_0^{\infty}\int_{z-\delta}^{z+\delta}\frac{y}{\sqrt{2\pi r^3}}\exp\left\{-\frac{y^2}{2K^2 r}\right\}\mathrm{d}y\mathrm{d}z \;=\; \left(\frac{2}{\pi}\right)^{1/2}\left(\frac{C_0K^5}{1\wedge\sqrt{t}}\right)\frac{\delta}{\sqrt{r}} \\
&\le \left(\frac{2}{\alpha\pi}\right)^{1/2}\left(\frac{C_0K^5}{1\wedge\sqrt{t}}\right)\delta^{1-\theta/2} \;\longrightarrow\; 0,\quad \text{as } \delta\longrightarrow 0,\ \text{uniformly in } r,
\end{aligned} \tag{2.5.17}$$

where C_0 is a constant depending only on K in (2.5.2a). Obviously, the same result holds for $\delta < 0$ and $r > \alpha(-\delta)^{\theta}$.

Now recall the definitions of the point sets U^R, V_+^R, V_-^R in Section 2.4. By (2.5.17), a similar argument to previously leads to the diffusion version of Theorem 2.4.1.

Theorem 2.5.5. *Suppose that (2.5.2a) – (2.5.2c) hold. If D solves* $\mathbf{SEP}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$*, and moreover, satisfies (2.4.5). Then the couple (u, D) solves* $\mathbf{FBP}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$ *and satisfies (2.5.10) where $u(x,t) = -\mathbb{E}^{\nu}\left[|x - X_{t\wedge\tau_D}|\right]$.*

Together, we can state the results of this chapter in the following form: Suppose ν, μ are two probability measures with finite expectation, satisfying (2.1.7), and the time-homogeneous diffusion X solves (2.1.1) with the diffusion coefficient σ satisfying (2.5.2a) – (2.5.2c). Then we have the following relation:

$$\left\{\begin{array}{l}\mathbf{SEP}\\ (2.4.5)\end{array}\right. \xRightarrow{\text{Theorem 2.5.5}} \left\{\begin{array}{l}\mathbf{FBP}\\ (2.5.10)\end{array}\right. \xRightarrow{\text{Theorem 2.5.3}} \mathbf{SEP} \xRightarrow{\text{Theorem 2.5.1}} \mathbf{FBP}$$

Chapter 3

Connection to Variational Inequalities

(The work regarding Root's barrier in Chapter 3 and Chapter 4 has appeared in Cox and Wang [2011].)

Now we study the relation between Root's Skorokhod embedding problem and a variational inequality. The notation and definitions and some of the key results which we will use come from Bensoussan and Lions [1982].

3.1 Notation and Elementary Results

We begin with some necessary notations and elementary results about evolutionary variational inequality. Given a constant $\lambda \geq 0$ and a finite time horizon $T > 0$, we define the Banach spaces $H^{m,\lambda} \subset L^2(\mathbb{R})$, $L^2(0,T;H^{m,\lambda})$ and $L^\infty(0,T;H^{m,\lambda})$ with the norms:

$$\|g\|^2_{H^{m,\lambda}} = \sum_{k=0}^{m} \int_{\mathbb{R}} e^{-\lambda|x|} |D^k g(x)|^2 \mathrm{d}x;$$

$$\|w\|^2_{L^2(0,T;H^{m,\lambda})} = \int_0^T \|w(\cdot,t)\|^2_{H^{m,\lambda}} \mathrm{d}t; \quad \|w\|_{L^\infty(0,T;H^{m,\lambda})} = \operatorname*{esssup}_{t\in[0,T]} \|w(\cdot,t)\|_{H^{m,\lambda}}$$

where the derivatives D^k are to be interpreted as weak derivatives – that is, $D^k g$ is defined by the requirement that

$$\int_{\mathbb{R}} \phi(x)(D^k g)(x)\mathrm{d}x \;=\; (-1)^k \int_{\mathbb{R}} g(x)\frac{\partial^k \phi}{\partial x^k}(x)\mathrm{d}x,$$

for all $\phi \in \mathbb{C}_K^\infty(\mathbb{R})$, and $\mathbb{C}_K^\infty(\mathbb{R})$ is the set of compactly supported smooth functions on $\mathbb{R}$. In particular, the spaces $H^{m,\lambda}$, $L^2(0,T;H^{m,\lambda})$ and $L^\infty(0,T;H^{m,\lambda})$ are Hilbert spaces with the obvious inner products. In addition, elements of the set $H^{1,\lambda}$ can always be taken to be continuous and $\mathbb{C}_K^\infty(\mathbb{R})$ is dense in $H^{m,\lambda}$, (see e.g. Friedman [1963, Theorem 5.5.20]).

For functions $a(x,t),\ b(x,t),\ c(x,t) \in L^\infty(\mathbb{R}\times(0,T))$, we define operators:

$$a_\lambda(t;v,w) \;=\; \int_{\mathbb{R}} e^{-2\lambda|x|}\left[a(x,t)\frac{\partial v}{\partial x}\frac{\partial w}{\partial x} + b(x,t)\frac{\partial v}{\partial x}\cdot w + c(x,t)vw\right]\mathrm{d}x$$

for $v,w \in L^2(0,T;H^{1,\lambda})$. Moreover if $\partial a/\partial x$ exits, we define, for $v \in H^{2,\lambda}$,

$$A(t)v \;=\; -\frac{\partial}{\partial x}\left(a(x,t)\frac{\partial v}{\partial x}\right) + \left[b(x,t) + 2\lambda\cdot\mathrm{sgn}(x)\,a(x,t)\right]\frac{\partial v}{\partial x} + c(x,t)v.$$

And finally, for $v,w \in H^{0,\lambda}$,

$$(v,w)_\lambda \;=\; \int_{\mathbb{R}} e^{-2\lambda|x|}vw\,\mathrm{d}x,$$

so that, for suitably differentiable test functions ϕ and $v \in H^{2,\lambda}$:

$$(\phi, A(t)v)_\lambda \;=\; a_\lambda(t;v,\phi).$$

We formulate *the evolutionary variational inequality* in which we are interested as follows: *given $T > 0$, we find a function v, in a suitable space, such that almost everywhere in $\mathbb{R}\times(0,T)$,*

$$\frac{\partial v}{\partial t} + A(t)v - f \;\geq\; 0; \tag{3.1.1a}$$

$$\frac{\partial v}{\partial t} + A(t)v - f > 0 \implies v = \psi; \tag{3.1.1b}$$

$$v(x,t) \;\geq\; \psi(x,t); \tag{3.1.1c}$$

$$v(x,0) \;=\; \bar{v}, \qquad a.e.\;\; x\in\mathbb{R}, \tag{3.1.1d}$$

hold, where ψ, $\bar{v}$ and f are given functions.

Regarding the evolutionary variational inequality, we have the following definition of *strong solutions* and the restatement of Bensoussan and Lions [1982, Theorem 2.2, and section 2.15, Chapter 3].

Definition 3.1.1 (Strong solution)**.** Given $\lambda, T > 0$, we say that v is *a strong solution* to the evolutionary variational inequality (3.1.1a) – (3.1.1d) if

$$v \in L^\infty(0,T;H^{1,\lambda}), \quad \frac{\partial v}{\partial t} \in L^2(0,T;H^{0,\lambda}); \tag{3.1.2a}$$

$$\begin{aligned} &\left(\frac{\partial v}{\partial t}, w - v\right)_\lambda + a_\lambda(t; v, w - v) \geq (f, w - v)_\lambda, \quad \text{a.e.} \quad t \in (0,T), \\ &\qquad \forall w \in H^{1,\lambda} \ \text{ such that } \ w(x) \ \geq \ \psi(x,t), \quad \text{a.e.} \quad (x,t) \in \mathbb{R} \times (0,T), \end{aligned} \tag{3.1.2b}$$

and v satisfies (3.1.1c) and (3.1.1d)

Note here, it is a slightly different from the original definition of strong solution of Bensoussan and Lions [1982], in which (3.1.2a) is replaced by

$$v \in L^2(0,T;H^{1,\lambda}), \quad \frac{\partial v}{\partial t} \in L^2(0,T;H^{0,\lambda}).$$

Since we consider our problem in one-dimension, then the diffusion coefficient σ is 'symmetric' trivially, and then we can consider the variational inequality with stronger restriction, and then the following theorem is given:

Theorem 3.1.2. *For any given $\lambda > 0$ and $T > 0$, suppose:*

(i). $a, b, c, \dfrac{\partial a}{\partial t}$ *are bounded on $\mathbb{R} \times (0,T)$ with $a(x,t) \geq \varepsilon$ a.e. in $\mathbb{R} \times (0,T)$ for some $\varepsilon > 0$;*

(ii). $f \in L^2(0,T;H^{0,\lambda})$; $\quad \psi, \dfrac{\partial \psi}{\partial t} \in L^2(0,T;H^{1,\lambda})$;

(iii). $\bar{v} \in H^{1,\lambda}$, $\ \bar{v} \geq \psi(0)$;

(iv). *The set*

$$\begin{aligned} \mathcal{X} := \Big\{ w \in L^2(0,T;H^{1,\lambda}) : \ \frac{\partial w}{\partial t} \in L^2\big(0,T;(H^{1,\lambda})^*\big), \\ w(x,t) \geq \psi(x,t) \quad \text{a.e. in } \mathbb{R} \times [0,T] \Big\} \end{aligned}$$

is non-empty, where $(H^{1,\lambda})^*$ denotes the dual space of $H^{1,\lambda}$.

Then there exists a unique strong solution to the evolutionary variational inequality (3.1.1a) – (3.1.1d). Moreover, if the strong solution v satisfies that $v \in L^2(0,T;H^{2,\lambda})$, then v satisfies (3.1.1a) and (3.1.1b).

Proof. For the most part, this theorem is a restatement of Bensoussan and Lions [1982, Theorem 2.2, and Section 2.15, Chapter 3], where we have mapped $t \mapsto T-t$ and $v \mapsto -v$.

We therefore only need to explain the last part of the result. If we suppose $v \in L^2(0,T;H^{2,\lambda})$ and $\phi \in H^{1,\lambda}$, we have

$$\begin{aligned} a_\lambda(t;v,\phi) &= \int_{\mathbb{R}} e^{-2\lambda|x|} a(x,t)\frac{\partial v}{\partial x}\mathrm{d}\phi + \int_{\mathbb{R}} e^{-2\lambda|x|}\phi\left[b(x,t)\frac{\partial v}{\partial x} + c(x,t)v\right]\mathrm{d}x \\ &= \left[e^{-2\lambda|x|}a(x,t)\frac{\partial v}{\partial x}\phi\right]_{-\infty}^{\infty} + \int_{\mathbb{R}} e^{-2\lambda|x|}\phi\cdot A(t)v\,\mathrm{d}x, \end{aligned}$$

where the first term on the right-hand side vanishes since $v \in L^2(0,t;H^{1,\lambda})$ and $\phi \in H^{1,\lambda}$. Therefore, by (3.1.2b), for any $w \in H^{1,\lambda}$ such that $w \geq \psi$ a.e. in $\mathbb{R}$,

$$\left(\frac{\partial v}{\partial t} + A(t)v - f,\, w-v\right)_\lambda \;\geq\; 0, \quad \text{a.e.} \quad t \in (0,T).$$

Taking for example $w = v + \phi$, for a positive test function ϕ, we conclude that (3.1.1a) holds. Moreover, let $w = \psi$ in the inequality above, we have

$$\int_{\mathbb{R}} e^{-2\lambda|x|}\left(\frac{\partial v}{\partial t} + A(t)v - f\right)(\psi - v)\mathrm{d}x \;\geq\; 0.$$

Then (3.1.1b) follows from (3.1.1a) and (3.1.1c). □

3.2 Connection with Skorokhod's Embedding Problems

To connect our embedding problem $\mathbf{SEP}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$ with the variational inequality, we need some assumptions on the diffusion coefficient σ, the starting distribution ν and the target distribution μ. Firstly, on $\sigma : \mathbb{R} \to \mathbb{R}_+$, we still assume (2.5.2a) – (2.5.2c) hold. In addition, we assume that (2.5.16) holds. On μ and ν, we still assume that $\mathrm{U}\mu \leq \mathrm{U}\nu$ to ensure that the existence of solution to $\mathbf{SEP}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$.

Under these assumptions, we can specify the coefficients in the evolutionary variational inequality (3.1.1a) – (3.1.1d) to be:

$$\begin{aligned} a(x,t) &= \frac{\sigma^2(x)}{2}; \quad b(x,t) \;=\; \sigma(x)\sigma'(x) - \lambda\sigma^2(x)\,\mathrm{sgn}(x); \\ c(x,t) &= f(x,t) \;=\; 0; \quad \psi(x,t) \;=\; \mathrm{U}\mu(x); \quad \bar{v} \;=\; \mathrm{U}\nu(x), \end{aligned} \tag{3.2.1}$$

then the corresponding operators are given by

$$A(t) \;=\; -\frac{\sigma^2(x)}{2}\frac{\partial^2}{\partial x^2}, \qquad \text{and}$$

$$a_\lambda(t;v,w) \;=\; \int_{\mathbb{R}} e^{-2\lambda|x|}\left[\frac{\sigma^2(x)}{2}\frac{\partial v}{\partial x}\frac{\partial w}{\partial x} + \Big(\sigma(x)\sigma'(x) - \lambda\sigma^2(x)\mathrm{sgn}(x)\Big)\frac{\partial v}{\partial x}\cdot w\right]\mathrm{d}x.$$

Now we write the evolutionary variational inequality as:

$$\begin{cases} \dfrac{\partial v}{\partial t} - \dfrac{\sigma^2(x)}{2}\dfrac{\partial^2 v}{\partial x^2} \;\geq\; 0; \\ \dfrac{\partial v}{\partial t} - \dfrac{\sigma^2(x)}{2}\dfrac{\partial^2 v}{\partial x^2} \;>\; 0 \implies v \;=\; \mathrm{U}\mu; \\ v(x,t) \;\geq\; \mathrm{U}\mu(x); \quad v(x,0) \;=\; \mathrm{U}\nu(x). \end{cases}$$

And our problem is

EVI($\boldsymbol{\sigma,\nu,\mu}$): For given $T>0$, find a function $v:\mathbb{R}\times[0,T]\to\mathbb{R}$ in a suitable space such that (3.1.1a) – (3.1.1d) hold, where all the coefficients are given in (3.2.1).

As a first observation, we try to find a solution to **EVI($\boldsymbol{\sigma,\nu,\mu}$)** from **SEP($\boldsymbol{\sigma,\nu,\mu}$)**. We denote the solution to **SEP($\boldsymbol{\sigma,\nu,\mu}$)** by D, and define the potential $u(x,t) := -\mathbb{E}^\nu[|x - X_{t\wedge\tau_D}|]$. In the preceding sections, we have already learned that,

$$\frac{\partial u}{\partial t} \;=\; \frac{\sigma^2(x)}{2}\frac{\partial^2 u}{\partial x^2}, \quad \text{in } D. \tag{3.2.2}$$

For $(x,t)\notin D$, *very roughly*, $\partial u/\partial t = 0$ and $A(t)u \geq 0$ since, in the sense of distributions,

$$\mathrm{d}\left(\frac{\partial u}{\partial x}\right) \;=\; 2\mu_t(\mathrm{d}x) \;\text{ on }\; \mathbb{R}, \tag{3.2.3}$$

where

$$\mu_t(\mathrm{d}x) \;=\; \mathbb{P}^\nu\left[\,X_{t\wedge\tau_D} \in \mathrm{d}x\,\right].$$

On the other hand, since $u = \mathrm{U}\mu$ when $(x,t) \in \mathbb{R} \times \mathbb{R}_+ \backslash D$, so we intuitively see that u solves $\mathbf{EVI}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$.

For more description in the relation between the evolutionary variational inequality and the Skorokhod embedding problem in Root's sense, we consider the strong solution to $\mathbf{EVI}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$. Our main result is then to show that finding the solution to $\mathbf{SEP}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ is equivalent to finding a (and hence the unique, by Theorem 3.1.2) strong solution to $\mathbf{EVI}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$.

Theorem 3.2.1. *Suppose (2.5.2a) – (2.5.2c), and (2.5.16) hold. Moreover, assume D is the solution to $\mathbf{SEP}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ and v is the strong solution to $\mathbf{EVI}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ respectively. Define*

$$\begin{aligned} u(x,t) &:= -\mathbb{E}^{\nu}[|x - X_{t\wedge\tau_D}|], \\ D^T &:= \big\{ (x,t) \in \mathbb{R} \times (0,T) : v(x,t) > \psi(x,t) \big\}. \end{aligned} \tag{3.2.4}$$

Then we have

$$\begin{aligned} D^T &= D \cap \big(\mathbb{R} \times [0,T]\big), \quad \textit{and} \\ u(x,t) &= v(x,t), \quad \textit{for} \ \ (x,t) \in \mathbb{R} \times [0,T], \end{aligned}$$

Proof. Let $\lambda > 0$ be fixed, and suppose D is the solution to $\mathbf{SEP}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$. We need to show u is a strong solution to $\mathbf{EVI}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$. First note that $\mathrm{U}\mu(x) + |x|$ is continuous on $\mathbb{R}$, and converges to 0 as $x \to \pm\infty$, and hence is bounded. So $(x,t) \mapsto \mathrm{U}\mu(x) + |x| \in L^{\infty}(0,T;H^{0,\lambda})$, and then $\mathrm{U}\mu \in L^{\infty}(0,T;H^{0,\lambda})$. Similarly, $\mathrm{U}\nu \in L^{\infty}(0,T;H^{0,\lambda})$. Since $0 \geq \mathrm{U}\nu(x) \geq u(x,t) \geq \mathrm{U}\mu(x)$ for all $t \in [0,T]$, we have $u \in L^{\infty}(0,T;H^{0,\lambda})$. By (vi) of Theorem 2.1.6, we also have $|\partial u/\partial x| \leq 1$ since $u(\cdot,t)$ is the potential of some probability distribution. Therefore we have $u \in L^{\infty}(0,T;H^{1,\lambda})$.

By the discussion in the preceding chapter, $|\partial u/\partial t| \leq |A(t)u| \leq K^2 p^{\nu}(x,t)$ a.e. on $\mathbb{R} \times [0,T]$ where $p^{\nu}(x,t)$ is the density of the diffusion process X with the starting distribution ν. Then by the Gaussian estimates in Lemma 2.5.4, we know there exists some constant $A > 0$, depending only on K, such that

$$\begin{aligned}
\left\|\frac{\partial u}{\partial t}\right\|_{L^2(0,T;H^{0,\lambda})} &\leq \int_{\mathbb{R}} \nu(\mathrm{d}y) \int_0^T \int_{\mathbb{R}} \frac{A}{1 \wedge t} \exp\left\{-2\left(At - \frac{(x-y)^2}{At}\right)^- - \lambda|x|\right\} \mathrm{d}x\mathrm{d}t \\
&= \int_{\mathbb{R}} \nu(\mathrm{d}y) \int_0^T \frac{A}{1 \wedge t} \int_{y-At}^{y+At} e^{-\lambda|x|} \mathrm{d}x\mathrm{d}t \\
&\qquad + \int_{\mathbb{R}} \nu(\mathrm{d}y) \int_0^T \frac{Ae^{2At}}{1 \wedge t} \int_{\mathbb{R}\setminus(y-At,y+At)} \exp\left\{-\frac{2(x-y)^2}{At} - \lambda|x|\right\} \mathrm{d}x\mathrm{d}t \\
&\leq \int_{\mathbb{R}} \nu(\mathrm{d}y) \int_0^T \frac{A}{1 \wedge t} \int_{-At}^{At} e^{-\lambda|x|} \mathrm{d}x\mathrm{d}t \\
&\qquad + \int_{\mathbb{R}} \nu(\mathrm{d}y) \int_0^T \frac{Ae^{2At}}{1 \wedge t} \int_{\mathbb{R}\setminus(y-At,y+At)} \exp\left\{-\frac{2(x-y)^2}{At}\right\} \mathrm{d}x\mathrm{d}t \\
&= \frac{2A}{\lambda} \int_0^T \frac{1}{1 \wedge t}(1 - e^{-\lambda At})\mathrm{d}t + 2A \int_0^T \frac{e^{2At}}{1 \wedge t} \int_{At}^{\infty} \exp\left\{-\frac{2x^2}{At}\right\} \mathrm{d}x\mathrm{d}t \\
&\leq \frac{2A}{\lambda}\left[\int_0^{T\wedge(A\lambda)^{-1}} \frac{A\lambda t}{1 \wedge t}\mathrm{d}t + \int_{T\wedge(A\lambda)^{-1}}^T \frac{1}{1 \wedge t}(1 - e^{-\lambda At})\mathrm{d}t\right] \\
&\qquad + \frac{A^{3/2}\pi^{1/2}}{\sqrt{2}}\left[\int_0^1 \frac{e^{2At}}{\sqrt{t}}\mathrm{d}t + \int_1^T e^{2At}\sqrt{t}\mathrm{d}t\right] \; < \; \infty,
\end{aligned}$$

where we apply Hölder's inequality in the first inequality. So $\partial u/\partial t \in L^2(0,T;H^{0,\lambda})$, and (3.1.2a) is done.

Conditions (3.1.1c) and (3.1.1d) are clear by the same arguments used in the proof of Theorem 2.5.1. Now we consider (3.1.2b).

For any $w \in H^{1,\lambda}$, we can find a sequence $\{\phi_n\} \subset \mathbb{C}_K^\infty$ such that

$$\lim_{n\to\infty} \left\|\phi_n - (w - u(\cdot,t))\right\|_{H^{1,\lambda}} \; = \; 0. \tag{3.2.5}$$

Moreover, $e^{-\lambda|x|}u(x,t)$ is bounded, and if $e^{-\lambda|x|}w(x)$ is also bounded then we can in addition find a sequence $\{\phi_n\} \subset \mathbb{C}_K^\infty$ such that $e^{-2\lambda|x|}\phi_n(x) \geq -K'$ for some positive constant K' independent of n.

For any n, we therefore have by (3.2.3),

$$\begin{aligned}\int_{\mathbb{R}} e^{-2\lambda|x|}\frac{\sigma^2}{2}\frac{\partial u}{\partial x}\frac{\partial \phi_n}{\partial x}\mathrm{d}x \\ &= \left[e^{-2\lambda|x|}\phi_n\cdot\frac{\sigma^2}{2}\frac{\partial u}{\partial x}\right]_{-\infty}^{\infty} - \int_{\mathbb{R}}\phi_n \mathrm{d}\left(e^{-2\lambda|x|}\frac{\sigma^2}{2}\frac{\partial u}{\partial x}\right) \\ &= \int_{\mathbb{R}} e^{-2\lambda|x|}\phi_n\cdot\sigma^2\mathrm{d}\mu_t - \int_{\mathbb{R}} e^{-2\lambda|x|}\left(\sigma\sigma' - \lambda\cdot\mathrm{sgn}(x)\sigma^2\right)\frac{\partial u}{\partial x}\cdot\phi_n\mathrm{d}x.\end{aligned} \tag{3.2.6}$$

On the other hand, since $(\partial u/\partial t)(\mathrm{d}x)$ vanishes outside D, and, using (3.2.2) and (3.2.3), this is equal to $-\sigma^2(x)\mu_t(\mathrm{d}x)$ for $(x,t)\in D$, and we have, for almost every $t\in[0,T]$,

$$\int_{\mathbb{R}} e^{-2\lambda|x|}\phi_n\cdot\frac{\partial u}{\partial t}\mathrm{d}x + \int_{\mathbb{R}} e^{-2\lambda|x|}\phi_n\cdot\sigma^2\mathrm{d}\mu_t \;=\; \int_{\mathbb{R}\setminus D_t} e^{-2\lambda|x|}\phi_n\cdot\sigma^2\mathrm{d}\mu_t, \tag{3.2.7}$$

where $D_t := \{x : (x,t)\in D\}$. Now, by (3.2.6) and (3.2.7),

$$\begin{aligned}\left(\frac{\partial u}{\partial t},\phi_n\right)_{\lambda} + a_{\lambda}(t;u,\phi_n) \\ &= \int_{\mathbb{R}} e^{-2\lambda|x|}\left[\frac{\partial u}{\partial t}\phi_n + \frac{\sigma^2}{2}\frac{\partial u}{\partial x}\frac{\partial\phi_n}{\partial x} + \left(\sigma\sigma' - \lambda\cdot\mathrm{sgn}(x)\sigma^2\right)\frac{\partial u}{\partial x}\phi_n\right]\mathrm{d}x \\ &= \int_{\mathbb{R}} e^{-2\lambda|x|}\phi_n\cdot\frac{\partial u}{\partial t}\mathrm{d}x + \int_{\mathbb{R}} e^{-2\lambda|x|}\phi_n\cdot\sigma^2\mathrm{d}\mu_t \\ &= \int_{\mathbb{R}\setminus D_t} -e^{-2\lambda|x|}\phi_n\cdot\sigma^2\mathrm{d}\mu_t,\end{aligned}$$

for almost every $t\in[0,T]$. Now suppose initially we have $e^{-\lambda|x|}w$ bounded, and choose a sequence ϕ_n as above. Then, we can let $n\to\infty$ and apply Fatou's Lemma and the fact that $u(\cdot,t) = \mathrm{U}\mu = \psi$ on $\mathbb{R}\setminus D_t$ and $w\geq\psi$ to get:

$$\begin{aligned}\left(\frac{\partial u}{\partial t}, w-u\right)_{\lambda} &+ a_{\lambda}(t;u,w-u) \\ &= \int_{\mathbb{R}\setminus D_t} e^{-2\lambda|x|}\cdot(w-\psi)\cdot\sigma^2\mathrm{d}\mu_{T-t} \;\geq\; 0,\end{aligned}$$

for almost every $t\in[0,t]$. So (3.1.2b) holds when $e^{-\lambda|x|}w$ is bounded. The general case follows from noting that $\max\{w,-N\}$ converges to w in $H^{1,\lambda}$. We can conclude that u is a strong solution to $\mathbf{EVI}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$.

Conversely, suppose that we have already found the solution to $\mathbf{EVI}(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})$, denoted by v. Note here, (iv) of Theorem 3.1.2 is satisfied trivially: $0\in\mathcal{X}$. Now, by

Theorem 3.1.2 and the preceding argument, we have

$$u(x,t) \;=\; v(x,t), \qquad \text{for } (x,t) \in \mathbb{R} \times [0,T].$$

In Chapter 2, we showed that $u > \mathrm{U}\mu$ on D and equal to $\mathrm{U}\mu$ outside D, so $D^T = D \cap (\mathbb{R} \times [0,T])$. □

Remark 3.2.2. The constant λ which appears in the evolutionary variational inequality can now be seen to be unimportant. In Section 2.1, we introduce the uniqueness result of Root's barrier Rost [1976, Theorem 2 and its corollary]. Now we consider two positive number $\lambda < \lambda^*$, then by Theorem 3.1.2, there exists v and v^* satisfying (3.1.1c), (3.1.1d) and (3.1.2a), (3.1.2b) with parameters λ and λ^* respectively. According to Theorem 3.2.1,

$$v(x,t) = u(x,t) = v^*(x,t).$$

Therefore, the description of Root's barrier by the strong solution to the evolutionary variational inequality is not affected as the choice of the parameter $\lambda > 0$. We do however need $\lambda > 0$, since this assumption is used in e.g. (3.2.6) to ensure we can integrate by parts.

Remark 3.2.3. As noted in Bensoussan and Lions [1982, Section 4.9, Chapter 3], and which is well known, one can connect the solution to the evolutionary variational inequality $\mathbf{EVI}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ to the solution of a particular optimal stopping problem. In our context, the function v which arises in the solution to $\mathbf{EVI}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ is also the function which arises from solving the problem:

$$v(x,t) \;=\; \sup_{\tau \le t} \mathbb{E}^x \left[\, \mathrm{U}\mu(X_\tau)\mathbf{1}_{[\tau<t]} + \mathrm{U}\nu(X_\tau)\mathbf{1}_{[\tau=t]} \, \right]. \tag{3.2.8}$$

This seems a rather interesting observation, and at one level extends a number of connections known to exist between solutions to the Skorokhod embedding problem, and solutions to optimal stopping problems: e.g. Peskir [1999], Obłój [2007] and Cox et al. [2008].

What is rather interesting, and appears to differ from these other situations, is that the above examples are all cases where the same stopping time is both a Skorokhod embedding, and a solution to the relevant optimal stopping problem. In the context here, we see that the optimal stopping problem is *not* solved by Root's stopping time.

Rather, the problem given in (3.2.8) runs 'backwards' in time: if we keep t fixed, then the solution to (3.2.8) is:

$$\tau_D \;=\; \inf\{s \geq 0 : (X_s, t-s) \notin D\} \wedge t.$$

In addition, our connection between these two problems is only through the analytic statement of the problem: it would be interesting to have a probabilistic explanation for the correspondence.

Remark 3.2.4. The above ideas also allow us to construct alternative embeddings which fail to be uniformly integrable. Consider using the variational inequality to construct the domain D in the manner described above, but with the function ψ chosen to be $\mathrm{U}\mu - c$, for some positive constant c. Then one might expect D to generate a barrier, which is non-empty, so that $\tau_D < \infty$ a.s., and the functions u and v defined in Theorem 3.2.1 to agree (for example by taking bounded approximations to D). In particular $\lim_{t\to\infty} u(\cdot, t) = \mathrm{U}\mu - c$. Since X^{τ_D} is no longer uniformly integrable, we cannot simply infer that this holds in the limit, but we can consider for example

$$u(x,t) - u(y,t) \;=\; -\mathbb{E}\left[\,|X_{t\wedge\tau_D} - x| - |X_{t\wedge\tau_D} - y|\,\right]$$

which is a bounded function. Taking the limit as $t \to \infty$, we can deduce that

$$-\mathbb{E}\left[\,|X_{t\wedge\tau_D} - x| - |X_{t\wedge\tau_D} - y|\,\right] \;=\; \mathrm{U}\mu(x) - \mathrm{U}\mu(y).$$

From this expression, we can divide through by $(x-y)$ and take the limit as $x \downarrow y$ to get $2\mathbb{P}[X_{\tau_D} > y] - 1$. The law of X_{τ_D} now follows.

Note also that there is no reason that the distribution above needed to have the same mean as ν, and this can lead to constructions where the means differ. In general, these constructions will not give rise to a uniformly integrable embedding, but if we take two general (integrable) distributions, there is a natural choice, which is to find the smallest $c \in \mathbb{R}$ such that $\mathrm{U}\nu(x) \geq \mathrm{U}\mu(x) - c$. In such a case, we would expect the resulting embedding to be *minimal* in the sense that there is no other construction of a stopping time which embeds the same distribution, and is almost surely smaller. See Monroe [1972a] and Cox [2008] for further detail regarding minimality.

3.3 Geometric Brownian Motion

An important motivating example for our study is the financial application of Root's solution which will be described in detail later in Chapter 4. In Dupire [2005] and Carr and Lee [2010], the case $\sigma : x \mapsto x$ plays a key role in both the pricing and the construction of a hedging portfolio. However, in the previous section, we only discussed the relation between Root's construction and variational inequalities with bounded diffusion coefficient σ.

In this section, we study this special case: $\sigma(x) = x$, so that X is a geometric Brownian motion. In addition, we will assume that the process is strictly positive, so that ν and μ are supported on $(0, \infty)$. We therefore consider the Skorokhod embedding problem $\mathbf{SEP(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})}$ with initial distribution ν, where ν and μ are integrable probability distributions satisfying

$$supp(\mu) \;\subset\; \mathbb{R}_+, \quad supp(\nu) \;\subset\; \mathbb{R}_+, \quad \mathrm{U}\mu \leq \mathrm{U}\nu \;\text{ and }\; \int x^2 \mathrm{d}\nu < \infty\,. \tag{3.3.1}$$

Note that the assumptions that μ and ν are integrable, and $\mathrm{U}\mu \leq \mathrm{U}\nu$ together imply that

$$m \;:=\; \int_{\mathbb{R}} x\,\nu(\mathrm{d}x) \;=\; \int_{\mathbb{R}} x\,\mu(\mathrm{d}x) \;>0.$$

The solution to the stochastic differential equation

$$\mathrm{d}X_t \;=\; X_t \mathrm{d}W_t, \quad X_0 \;=\; x_0$$

is the geometric Brownian motion $x_0 \exp\{W_t - t/2\}$, and, for $y > 0$, the transition density of the process is:

$$p_t(y, x) \;=\; \frac{1}{x}\frac{1}{\sqrt{2\pi t}}\mathbf{1}_{[x>0]} \exp\left\{-\frac{(\ln x - \ln y + t/2)^2}{2t}\right\}. \tag{3.3.2}$$

By analogy with Theorem 2.5.1, if D is the solution to $\mathbf{SEP(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})}$, then we would expect

$$\begin{cases} \dfrac{\partial u}{\partial t} \;=\; \dfrac{x^2}{2}\dfrac{\partial^2 u}{\partial x^2}, & \text{for } D; \\ u(x,t) \;=\; \mathrm{U}\mu(x), & \text{for } (x,t) \in \mathbb{R} \times \mathbb{R}_+ \backslash D, \end{cases}$$

where u is defined as before by $u(x,t) = -\mathbb{E}[|x - X_{t\wedge\tau_D}|]$. However, if we follow the arguments in Section 3.2, we find that we need to set $a(x,t) = x^2/2$ in $\mathbf{EVI}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$, which would not satisfy the first condition of Theorem 3.1.2. To avoid this we will perform a simple transformation of the problem. We set

$$v(x,t) \;=\; u(e^x, t), \quad (x,t) \in \mathbb{R} \times [0,T].$$

Define the operator

$$A(t) \;:=\; -\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}\frac{\partial}{\partial x},$$

then we have, when $(e^x, t) \in D$,

$$\frac{\partial v}{\partial t} + A(t)v \;=\; 0. \tag{3.3.3}$$

We state our main result of this section as follows:

Theorem 3.3.1. *Suppose $\sigma(x) = x$ on $\mathbb{R}_+$ and μ and ν satisfy (3.3.1). Moreover, assume D solves $\mathbf{SEP}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$, and $u := -\mathbb{E}^\nu[|x - X_{t\wedge\tau_D}|]$. Then $v := u(e^x, t)$ is the unique strong solution to $\mathbf{EVI}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ where we set*

$$\begin{aligned} a(x,t) \;=\; \frac{1}{2}; \quad b(x,t) \;=\; \frac{1}{2} - \lambda \cdot \operatorname{sgn}(x); \quad c(x,t) \;=\; f(x,t) \;=\; 0; \\ \psi(x,t) \;=\; \mathrm{U}\mu(e^x); \quad \bar{v} \;=\; \mathrm{U}\nu(e^x); \quad \lambda > \frac{1}{2}. \end{aligned} \tag{3.3.4}$$

Proof. Much of the proof will follow the proof of Theorem 3.2.1. As before, (3.1.1c) and (3.1.1d) are clear. In addition, we note that $\psi - e^x$ is continuous and converges to 0 as $x \to \infty$, so $x \mapsto \psi - e^x \in L^\infty(0,T; H^{0,\lambda})$, and hence $\psi \in L^\infty(0,T; H^{0,\lambda})$ for $\lambda > 1/2$. Thus, $v \in L^\infty(0,T; H^{0,\lambda})$. Moreover, we can easily see $|\partial v/\partial x|$ is bounded by e^x. Therefore, $v \in L^\infty(0,T; H^{1,\lambda})$ when $\lambda > 1/2$.

On the other hand, since $\partial v/\partial t$ is bounded by $e^{2x}\int p_t(y, e^x)\nu(\mathrm{d}y)$, we have, by Hölder's inequality,

$$\left|\frac{\partial v}{\partial t}\right|^2 \;\leq\; \int_{\mathbb{R}_+} \frac{1}{2\pi t} \exp\left\{-\frac{(x - \ln y + t/2)^2}{t} + 2x\right\} \nu(\mathrm{d}y),$$

and hence,

$$\left\|\frac{\partial v}{\partial t}\right\|_{L^2(0,T;H^{0,\lambda})} \le \int_{\mathbb{R}_+}\nu(\mathrm{d}y)\int_0^T\int_{\mathbb{R}}\frac{1}{2\pi t}\exp\left\{-\frac{(x-\ln y+t/2)^2}{t}+2x-\lambda|x|\right\}\mathrm{d}x\mathrm{d}t$$
$$\le \int_{\mathbb{R}_+}\nu(\mathrm{d}y)\int_0^T\int_{\mathbb{R}}\frac{1}{2\pi t}\exp\left\{-\frac{x^2}{t}+2x-t+2\ln y\right\}\mathrm{d}x\mathrm{d}t$$
$$\le \int_{\mathbb{R}_+}y^2\nu(\mathrm{d}y)\int_0^T\frac{1}{2\sqrt{\pi t}}\mathrm{d}t \;<\; \infty,$$

and we have (3.1.2a).

In the sense of distributions, we know that

$$\mathrm{d}\left(\frac{\partial u}{\partial x}\right) \;=\; -2\mu_t(\mathrm{d}x)$$

where μ_t, as in the last section, denotes the law of $X_{t\wedge\tau_D}$. If we define the measure ν_t as

$$\nu_t(\mathrm{d}x) \;:=\; e^x\mu_t(\mathrm{d}e^x),$$

we then have

$$\mathrm{d}\left(\frac{\partial v}{\partial x}\right) \;=\; \mathrm{d}\left(e^x\left.\frac{\partial u(y,t)}{\partial y}\right|_{y=e^x}\right) \;=\; -2\nu_t(\mathrm{d}x)+\frac{\partial v}{\partial x}\mathrm{d}x. \tag{3.3.5}$$

Now take any $w\in H^{1,\lambda}$, and take $\{\phi_n\}\subset\mathbb{C}_K^\infty$ satisfying (3.2.5). By (3.3.3) and (3.3.5), similar arguments to those used in the proof of Theorem 3.2.1 give

$$\int_{\mathbb{R}}\frac{1}{2}e^{-2\lambda|x|}\frac{\partial v}{\partial x}\frac{\partial\phi_n}{\partial x}\mathrm{d}x+\int_{\mathbb{R}}\frac{1}{2}e^{-2\lambda|x|}\frac{\partial v}{\partial x}\cdot\phi_n\mathrm{d}x$$
$$= \int_{\mathbb{R}}e^{-2\lambda|x|}\phi_n\nu_t(\mathrm{d}x)+\int_{\mathbb{R}}e^{-2\lambda|x|}\lambda\cdot\mathrm{sgn}(x)\frac{\partial v}{\partial x}\cdot\phi_n\mathrm{d}x,$$
$$\int_{\mathbb{R}}e^{-2\lambda|x|}\frac{\partial v}{\partial t}\phi_n\mathrm{d}x+\int_{\mathbb{R}}e^{-2\lambda|x|}\phi_n\nu_t(\mathrm{d}x)$$
$$= \int_{\mathbb{R}\setminus\widetilde{D}_t}e^{-2\lambda|x|}\phi_n\nu_t(\mathrm{d}x),$$

for almost all $t\in[0,T]$, where $\widetilde{D}_t := \{x\in\mathbb{R}:\ (e^x,t)\in D\}$. Thus, for almost every

$t \in [0, T]$,

$$\begin{aligned}\left(\frac{\partial v}{\partial t}, \phi_n\right)_\lambda &+ a_\lambda(t; v, \phi_n) \\ &= \int_{\mathbb{R}} \left[\frac{\partial v}{\partial t}\phi_n + \frac{1}{2}\frac{\partial \phi_n}{\partial x}\frac{\partial v}{\partial x} + \left(\frac{1}{2} - \lambda \cdot \operatorname{sgn}(x)\right)\phi_n \frac{\partial v}{\partial x}\right] \mathrm{d}x \\ &= \int_{\mathbb{R}\setminus\widetilde{D}_t} e^{-2\lambda|x|}\phi_n \nu_t(\mathrm{d}x).\end{aligned}$$

Finally, following the same arguments as in the proof of Theorem 3.2.1, we conclude (3.1.2b) holds. Therefore v is the strong solution to $\mathbf{EVI}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ with coefficients determined by (3.3.1). The uniqueness is clear since it is easy to check the coefficients defined in (3.3.1) satisfy the conditions in Theorem 3.1.2. □

3.4 Connection to Reversed Barriers

In this section, we are concerned with the *reversed barrier* or *Rost's barrier*, which was proposed by Hermann Rost. Explicit results regarding this construction are not well established. It appears to originate in Chacon [1985], building on work of Rost. See also the survey paper by Obłój [2004, Section 7.3].

3.4.1 Introduction of Reversed Barriers

We begin this section by introducing the definition of the reversed barrier.

Definition 3.4.1 (Reserved Barrier)**.** A closed subset B of $[-\infty, +\infty] \times [0, +\infty]$ is a *reversed barrier* if

(i). $(x, 0) \in B$ for all $x \in [-\infty, +\infty]$;

(ii). $(\pm\infty, t) \in B$ for all $t \in [0, \infty]$;

(iii). if $(x, t) \in B$ then $(x, s) \in B$ whenever $s < t$.

As in the case of barriers, for a reversed barrier, we define $D = (\mathbb{R} \times (0, \infty))\setminus B$, and the Rost's stopping time

$$\tau_D = \inf\{t > 0 : (X_t, t) \in D\} = \inf\{t > 0 : t \leq R(X_t)\},$$

where $R : \mathbb{R} \to [0, \infty]$ is an upper semi-continuous function. In fact, it is not difficult to see that, just like the one-one relation between Root's barriers and lower semi-continuous functions, for any reversed barrier B, there exists a unique upper semi-continuous function $R : \mathbb{R} \to [0, \infty]$, such that

$$B = \big\{(x, t) \in [-\infty, \infty] \times [0, +\infty] : t \leq R(x)\big\}.$$

We also cite the result given in Obłój [2004], regarding existence and optimality of Rost's solution, in the case of Brownian motion.

Theorem 3.4.2. *Given a probability distribution μ on $\mathbb{R}$ with*

$$\mu(\{0\}) = 0, \qquad \int_{\mathbb{R}} x\, \mu(\mathrm{d}x) = 0 \quad \text{and} \quad \int_{\mathbb{R}} x^2 \mu(\mathrm{d}x) < \infty,$$

then

(i). *there exists a Rost's solution τ_D solving the Skorokhod embedding problem of μ and $\mathbb{E}[\tau_D] = \int x^2 \mathrm{d}\mu$;*

(ii). *if τ is a stopping time solving the same embedding problem and $\mathbb{E}[\tau] = v$, then for any $p < 1$, $\mathbb{E}[\tau_D^p] \leq \mathbb{E}[\tau^p]$, and for any $p > 1$, $\mathbb{E}[\tau_D^p] \geq \mathbb{E}[\tau^p]$.*

Remark 3.4.3 (Obłój)**.** In the same setting of Theorem 3.4.2, we could introduce *regular reversed barriers* similar to the case of Root's barriers, which lead us to a uniqueness result by analogy to Loynes [1970, Theorem 1] (see Section 7.3 Obłój [2004] for more details).

In his Ph.D. thesis, Chacon [1985] discussed solutions to Skorokhod embedding problems by the filling scheme[1]:

Theorem 3.4.4 (Filling scheme stopping time)**.** *Given two probability measures ν and μ. Let P_t be the semigroup of the Markov process X. Then there exists a (randomized) stopping time T with $\mu = \nu P_T$ if and only if $\langle \mu, f \rangle \leq \langle \nu, f \rangle$ for all non-negative f such that $f \geq P_S f$ for any stopping time S.*

Moreover, if T is a filling scheme stopping time such that $\mu = \nu P_T$, S is a stopping time also with $\mu = \nu P_S$, and suppose $\mathbb{E}^\nu[T] = \mathbb{E}^\mu[S] < \infty$. We then have $\mathbb{E}^\nu[F(T)] \geq \mathbb{E}^\nu[F(S)]$ for any convex function F on $[0, \infty)$.

[1] We have to mention here that *barrier* defined by R.V. Chacon in his thesis is in fact *reversed barrier* in our context.

In particular, Theorem 3.24 in his work shows that under some conditions including that ν and μ are supported by disjointed sets, the filling scheme stopping time is in fact given by a reversed barrier.

The reason why we call Rost's construction as reversed barrier is twofold. Firstly, in view of its definition, if a point is in Rost's (Root's) barrier then all the points to the left (right) of the point are also in the barrier. Secondly, in view of optimality, as mentioned in Theorem 3.4.2 and Theorem 3.4.4, the reversed barrier maximises the variance of the UI embedding of μ whereas the Root's solution minimises it.

Finally, we state a simple example where Root's barrier and a reversed barrier embed an identical distribution μ when the underlying process is a Brownian motion.

Given a positive number T, we define Root's stopping time τ_{Root} with the barrier function (see Figure 3-1(a))

$$R(x) \ = \ \begin{cases} 0\,, & x \notin (-2,2)\,; \\ T\,, & x = \pm 1\,; \\ \infty\,, & \text{otherwise}\,. \end{cases}$$

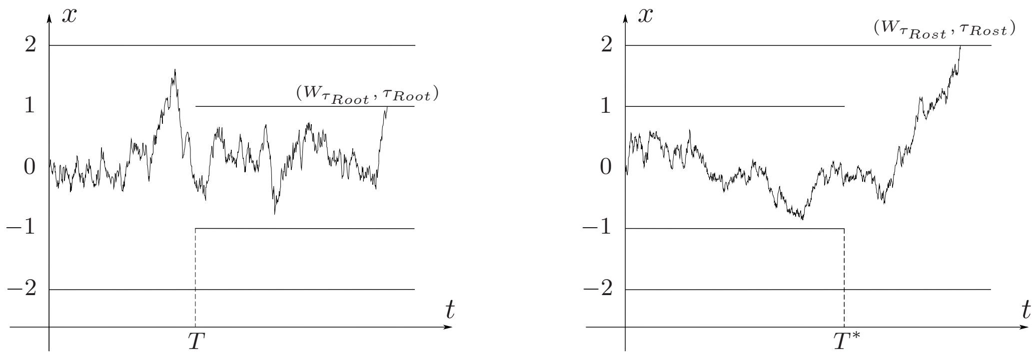
Then $\mathbb{P}[W_{\tau_{Root}} = \pm 1] = p$ and $\mathbb{P}[W_{\tau_{Root}} = \pm 2] = \frac{1}{2} - p$ for some $p \in (0, \frac{1}{2})$. Denote the law by μ.

We always can find another positive number T^*, such that $\mathbb{P}[\sup_{t<T^*} |W_t| > 1] = 2p$, and we then can construct Rost's stopping time τ_{Rost} with the reversed barrier function (see Figure 3-1(b))

$$R^*(x) \ = \ \begin{cases} \infty\,, & x \notin (-2,2)\,; \\ T^*\,, & x = \pm 1\,; \\ 0\,, & \text{otherwise}\,. \end{cases}$$

It is easy to see $W_{\tau_{Rost}}$ has also the distribution μ.

Another example is the stopping time τ_D generated by the barrier as in Figure 2-1(a). We have introduced it as Root's stopping time. However, it is also a Rost's stopping time (of course, generated by a different barrier function). Now we consider an arbitrary UI stopping time τ embedding the same distribution as τ_D, by the optimality of Root's



(a) The Root's barrier. (b) The Rost's barrier.

Figure 3-1: A Root's embedding and A Rost's embedding of an identical μ.

and Rost's embeddings, we have, for any $t > 0$,

$$\mathbb{E}[(\tau_D - t)_+] \;\leq\; \mathbb{E}[(\tau - t)_+] \;\leq\; \mathbb{E}[(\tau_D - t)_+].$$

It implies that $\mathcal{L}(\tau) = \mathcal{L}(\tau_D)$. In fact, we know that τ_D is the only UI stopping time for μ.

3.4.2 Connection with Variational Inequalities

We consider the construction of the reversed barrier. We are interested in this question when our underlying process X is the unique strong solution to the stochastic differential equation (2.1.1) and $X_0 \sim \nu$. And hence, we write the Skorokhod embedding problem as

SEP*$(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$: Find an upper-semicontinuous function $R(x)$ such that the domain $D = \{(x,t) : t > R(x)\}$ has $\tau_D \in \mathcal{T}(\sigma, \nu, \mu)$, where $\tau_D := \inf\{t > 0 : (X_t, t) \notin D\} = \inf\{t > 0 : t \leq R(X_t)\}$.

We keep the assumption $\mathrm{U}\nu \geq \mathrm{U}\mu$ as in the preceding sections. Before investigating the relation between reversed barriers and the evolutionary variational inequalities, we find the boundary problem satisfied by **SEP*$(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$**. We follow our idea used in the case of Root's barrier, and begin by considering the law $\mathcal{L}(X_{t\wedge\tau_D})$:

Lemma 3.4.5. *Suppose $D = \{(x,t) : t > R(x)\}$ where $R : \mathbb{R} \to [0, \infty]$ is an upper* *semi-continuous function. Then we have, for all* $f \in \mathbb{C}_b^0$,

$$\mathbb{E}^\nu[\, f(X_{t\wedge\tau_D})\,] - \mathbb{E}^\nu[\, f(X_{\tau_D})\,] \;=\; \begin{cases} \mathbb{E}^\nu[\,\mathbf{1}_{[t<\tau_D]} f(X_t)\,], & \textit{if } (x,t) \in D; \\ -\mathbb{E}^\nu[\,\mathbf{1}_{[t<\tau_D]} f(X_{\tau_D})\,], & \textit{if } (x,t) \notin D. \end{cases}$$

Proof. We only need to show

$$\mathbb{P}^\nu[X_{t\wedge\tau_D} \in \mathrm{d}x] - \mathbb{P}^\nu[X_{\tau_D} \in \mathrm{d}x] \;=\; \begin{cases} \mathbb{P}^\nu[X_t \in \mathrm{d}x,\, t < \tau_D], & \text{if } (x,t) \in D; \\ -\mathbb{P}^\nu[X_{\tau_D} \in \mathrm{d}x,\, t < \tau_D], & \text{if } (x,t) \notin D. \end{cases} \tag{3.4.1}$$

It is easy to see the left-hand side of (3.4.1) is equal to

$$\mathbb{P}^\nu[X_t \in \mathrm{d}x,\, t < \tau_D] - \mathbb{P}^\nu[X_{\tau_D} \in \mathrm{d}x,\, t < \tau_D]. \tag{3.4.2}$$

Suppose that $(x,t) \in D$. Then the set $\{(x,s) : s \geq t\} \subset D$, that means for any trajectory $X(\omega)$, if $\tau_D(\omega) > t$, $X_{\tau_D}(\omega) \notin \mathrm{d}x$. Therefore the second term in (3.4.2) vanishes.

On the other hand, if $(x,t) \notin D$, then the line section $\{(x,s) : s \leq t\} \subset D^\complement$, and then any trajectory $X(\omega)$ cannot cross this line section before hitting $D^\complement$. Therefore the first term in (3.4.2) vanishes. ☐

The following result connects $\mathbf{SEP^*}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ to a specialized partial differential equation with suitable initial and boundary condition.

Theorem 3.4.6. *Suppose (2.5.2a) – (2.5.2c) hold. Assume D solves* $\mathbf{SEP^*}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$, *and the corresponding UI stopping time is denoted by* τ_D. *Then the couple* (u, D), *where* $u(x,t) := \mathrm{U}\mu + \mathbb{E}^\nu[|\, x - W_{t\wedge\tau_D}|]$ *satisfies*

$$u \;\in\; \mathbb{C}^0(\mathbb{R}\times[0,\infty)) \quad \textit{and} \quad u|_D \;\in\; \mathbb{C}^{2,1}(D)\,; \tag{3.4.3a}$$

$$\frac{\partial u}{\partial t} \;=\; \frac{\sigma^2(x)}{2}\frac{\partial^2 u}{\partial x^2}, \qquad \textit{for } (x,t) \in D\,; \tag{3.4.3b}$$

$$u(x,0) \;=\; \mathrm{U}\mu(x) - \mathrm{U}\nu(x), \qquad \textit{for } x \in \mathbb{R}\,; \tag{3.4.3c}$$

$$u(x,t) \;=\; \mathrm{U}\mu(x) - \mathrm{U}\nu(x), \qquad \textit{if } t \leq R(x) \textit{ and } x \in \mathbb{R}\,. \tag{3.4.3d}$$

Proof. By (vi) of Theorem 2.1.6, u defined as in this theorem is differentiable almost everywhere with left and right derivatives in x,

$$
\begin{aligned}
u'_- &= 2\Big(\mathbb{P}^\nu\left[\,X_{t\wedge\tau_D} < x\,\right] - \mathbb{P}^\nu\left[\,X_{\tau_D} < x\,\right]\Big); \\
u'_+ &= 2\Big(\mathbb{P}^\nu\left[\,X_{t\wedge\tau_D} \le x\,\right] - \mathbb{P}^\nu\left[\,X_{\tau_D} \le x\,\right]\Big).
\end{aligned}
$$

By Lemma 3.4.5, for $(x,t) \in D$,

$$
\left(\frac{1}{2}\frac{\partial^2 u}{\partial x^2}\right)\mathrm{d}x \;=\; \mathbb{P}^\nu\left[\,X_{t\wedge\tau_D} \in \mathrm{d}x\,\right] - \mathbb{P}^\nu\left[\,X_{\tau_D} \in \mathrm{d}x\,\right] \;=\; \mathbb{P}^\nu\left[\,X_t \in \mathrm{d}x,\, t < \tau_D\,\right]. \quad (3.4.4)
$$

By Lemma 2.2.3 (which remains true here), the measure $\mathcal{L}(X_t, t < \tau_D)$ has no mass point in D, and hence, we have $u'_- = u'_+$ for $(x,t) \in D$. Moreover, $\partial u/\partial x$ is smooth on D. Similar to the case of the Root's barrier, we then have

$$
\begin{aligned}
\frac{\sigma^2}{2}\int_{R(x)}^t \frac{\partial^2 u}{\partial x^2}\mathrm{d}s &= \sigma^2(x)\int_0^t \lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\mathbb{P}^\nu\left[\,X_s \in (x-\varepsilon, x+\varepsilon),\, s < \tau_D\,\right]\mathrm{d}s \\
&= \mathbb{E}^\nu\left[L^x_{t\wedge\tau_D}\right] \;=\; \mathbb{E}^\nu\left[|x - X_{t\wedge\tau_D}|\right] - \mathbb{E}^\nu[|x - X_0|],
\end{aligned}
$$

where the first equality holds because when $s < R(x)$, $\mathbb{P}^\nu\left[\,X_s \in \mathrm{d}x,\, s < \tau_D\,\right] = 0$, and (3.4.3b) then is done. Similar to the case of Root's barrier, we can easily show that (3.4.3a), and (3.4.3c) is clear. Thus, we only need to show (3.4.3d). In fact, if $t \le R(x)$, i.e. $(x,t) \notin D$, then no trajectory of X can hit x before time $t \wedge \tau_D$, so $L^x_{t\wedge\tau_D} = 0$, a.s., and hence,

$$
u(x,t) \;=\; \mathrm{U}\mu(x) + \mathbb{E}^\nu[|x - X_0|] + \mathbb{E}^\nu\left[L^x_{t\wedge\tau_D}\right] \;=\; \mathrm{U}\mu - \mathrm{U}\nu.
$$

$\square$

We believe that, as in Chapter 2, the construction also could be connected to a free boundary problem with the initial condition and boundary condition as (3.4.3c) and (3.4.3d), and the region where u satisfies the heat equation is bounded an upper semi-continuous function from below. We do not pursue this matter here, and will construct the reversed barrier by the following evolutionary variational inequality:

$$\begin{cases} \dfrac{\partial v}{\partial t} - \dfrac{\sigma^2(x)}{2}\dfrac{\partial^2 v}{\partial x^2} \geq 0; \\ \dfrac{\partial v}{\partial t} - \dfrac{\sigma^2(x)}{2}\dfrac{\partial^2 v}{\partial x^2} > 0 \Longrightarrow v = \mathrm{U}\mu - \mathrm{U}\nu; \\ v(x,t) \geq v(x,0) = \mathrm{U}\mu(x) - \mathrm{U}\nu(x). \end{cases}$$

Our problem is then:

EVI*$(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$: For given $T > 0$, find a function $v : \mathbb{R} \times [0,T] \to \mathbb{R}$ in a suitable space such that (3.1.1a) – (3.1.1d) hold, where all the coefficients are given as

$$a(x,t) = \frac{\sigma^2(x)}{2}; \quad b(x,t) = \sigma(x)\sigma'(x) - \lambda \mathrm{sgn}(x)\sigma^2(x);$$
$$c(x,t) = f(x,t) = 0; \quad \psi = \bar{v} = \mathrm{U}\mu - \mathrm{U}\nu. \tag{3.4.5}$$

Our main result connecting **SEP*$(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$** and **EVI*$(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$** is stated as follows:

Theorem 3.4.7. *Suppose (2.5.2a) – (2.5.2c) and (2.5.16) hold. Moreover, assume D solves* **SEP*$(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$** *and v is the strong solution to* **EVI*$(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$**. *Define*

$$u(x,t) := \mathrm{U}\mu(x) + \mathbb{E}^{\nu}[|x - X_{t\wedge\tau_D}|]$$

and D^T as (3.2.4). Then we have that, $D^T = D \cap \mathbb{R} \times [0,T]$, and for all $(x,t) \in \mathbb{R} \times [0,T]$, $u(x,t) = v(x,t)$.

Proof. Similar to the case of Root's barrier, it is easy to show that (3.1.1c), (3.1.1d) and (3.1.2a) hold.

Note that, in the sense of distributions,

$$\mathrm{d}\left(\frac{\partial u}{\partial x}\right) = 2\mu_t(\mathrm{d}x) - 2\mu(\mathrm{d}x),$$

where μ_t denotes the law of $X_{t\wedge\tau_D}$.

We repeat the calculation in (3.2.6) and (3.2.7). For $\phi_n \in \mathbb{C}_K^\infty$ satisfying (3.2.5), and letting $n \to \infty$, we have a.e. in $t \in [0,T]$,

$$\begin{aligned}\left(\frac{\partial u}{\partial t}, w-u\right)_\lambda + a_\lambda(t;u,w-u) &= \int_{\mathbb{R}\backslash D_t} e^{-2\lambda|x|}(w-u)\cdot\sigma^2 \mathrm{d}(\mu-\mu_t) \\ &= \int_{\mathbb{R}\backslash D_t} e^{-\lambda|x|}(w-\psi)\cdot\sigma^2 \mathrm{d}(\mu-\mu_t)\,,\end{aligned}$$

where the last equality holds because $u = \psi$ on $\mathbb{R}\backslash D_t$ by Theorem 3.4.6. By (3.4.1), the measure $\mu - \mu_t$ is non-negative on $D_t^{\complement}$, and hence for $w \geq \psi$,

$$\int_{\mathbb{R}\backslash D_t} e^{-\lambda|x|}(w-\psi)\cdot\sigma^2 \mathrm{d}(\mu-\mu_t) \;\geq\; 0,$$

and (3.1.2b) holds. Therefore u is the unique strong solution to $\mathbf{EVI^*(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})}$, i.e. $u = v$. The result regarding D^T is straightforward and the theorem is proved. □

Remark 3.4.8. Suppose the solution to $\mathbf{SEP^*(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})}$ is unique[2]. Then, similar to Remark 3.2.2, the description of the reversed barrier by the strong solution to $\mathbf{EVI^*(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})}$ is not affected by the choice of λ.

Remark 3.4.9. As in Root's case (Remark 3.2.3), by Bensoussan and Lions [1982, Section 4.9, Chapter 3], the function v which arises in the solution to $\mathbf{EVI^*(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})}$ is also the function which arises from the optimal stopping problem:

$$v(x,t) \;=\; \sup_{\tau\leq t} \mathbb{E}^x\left[\mathrm{U}\mu(X_\tau) - \mathrm{U}\nu(X_\tau)\right]\,. \tag{3.4.6}$$

and the solution to the optimal stopping problem is

$$\tau_D \;=\; \inf\{s \geq 0 :\ (X_s, t-s) \in D^{\complement}\} \wedge t.$$

At last, regarding geometric Brownian motions, by analogy with Theorem 3.3.1, we have

Theorem 3.4.10. *Suppose $\sigma(x) = x$ on $\mathbb{R}_+$ and μ and ν satisfy (3.3.1). Moreover, assume D solves* $\mathbf{SEP^*(\boldsymbol{\sigma},\boldsymbol{\nu},\boldsymbol{\mu})}$, *and*

$$u(x,t) := \mathrm{U}\mu(x) + \mathbb{E}^\nu[|x - X_{t\wedge\tau_D}|].$$

[2] Unlike Root's case (Theorem 2.1.5), the uniqueness of the reversed barrier has never been well investigated in the general case. The only result we can find is the discussion by Jan Obłój (see Remark 3.4.3).

Then $v := u(e^x, t)$ is the unique strong solution to $\mathbf{EVI^*(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})}$ *where we set*

$$a(x,t) = \frac{1}{2}; \quad b(x,t) = \frac{1}{2} - \lambda \cdot \mathrm{sgn}(x); \quad c(x,t) = f(x,t) = 0;$$
$$\psi(x,t) = \bar{v}(x) = \mathrm{U}\mu(e^x) - \mathrm{U}\nu(e^x); \quad \lambda > \frac{1}{2}.$$

Proof. As before, (3.1.1c) and (3.1.1d) obviously hold. Since $\psi = (\mathrm{U}\mu - \mathrm{U}\nu)(e^x) \leq 0$ is continuous, and converges to 0 as x goes to infinity, then $\psi \in L^\infty(0,T; H^{0,\lambda})$ for $\lambda > 0$. Thus, v, bounded by 0 and ψ, is also of the class $L^\infty(0,T; H^{0,\lambda})$. Moreover, $|\partial v/\partial x|$ is bounded by $2e^x$. Therefore $v \in L^\infty(0,T; H^{0,\lambda})$ for $\lambda > 1/2$. In addition $\partial v/\partial t \in L^2(0,T; H^{0,\lambda})$ holds in the same way as in Theorem 3.3.1 since $\partial \mathrm{U}\mu/\partial t = 0$. (3.1.2a) is proved. In the sense of distributions, if we define $\nu_t(\mathrm{d}x) := e^x \mu_t(\mathrm{d}e^x)$ where μ_t is the law of $X_{t\wedge\tau_D}$, we have

$$\mathrm{d}\left(\frac{\partial v}{\partial x}\right) = \frac{\partial v}{\partial x}\mathrm{d}x - 2\nu_t(\mathrm{d}x).$$

Now for any $w \in H^{1,\lambda}$, and take $\{\phi_n\} \in \mathbb{C}_K^\infty$ satisfying (3.2.5). Similar arguments as in Theorem 3.3.1 and Theorem 3.4.7 give us,

$$\left(\frac{\partial v}{\partial t}, \phi_n\right)_\lambda + a_\lambda(t; v, \phi_n) = \int_{\mathbb{R}\setminus\widetilde{D}_t} e^{-2\lambda|x|}\phi_n \nu_t(\mathrm{d}x) \geq 0,$$

where $\widetilde{D}_t := \{x \in \mathbb{R} : (e^x, t) \in D\}$ and the last inequality holds because of (3.4.1). Finally, let $n \to \infty$, and repeat the same arguments as in the proof of Theorem 3.2.1, we reach (3.1.2b), and hence v is the strong solution to the variational inequality. The uniqueness is clear. $\square$

3.4.3 More Discussion on Reversed Barriers

We have discussed the construction of barriers and reversed barriers by variational inequalities. In both cases, we start with assuming the existence of the barriers for the given triple (σ, ν, μ), and then show the corresponding potential processes are the unique solution to the variational inequalities. In Root's case, existence is guaranteed by $\mathrm{U}\mu \leq \mathrm{U}\nu$ as mentioned in Section 2.1, and then the strong solution to the variational inequality with the coefficients (3.2.1) must give Root's barrier embedding μ in the process X with the initial distribution ν. In Rost's case, however, the existence of the barrier is doubtful. Therefore, without the strong assumption of existence, we cannot assert that the solution to the variational inequality with (3.4.5) gives the desired Rost's stopping time such that $\mathbb{P}^{\nu^*}(X_{\tau_{Rost}} \in \mathrm{d}x) = \mu^*(\mathrm{d}x)$ for all pairs (ν^*, μ^*) satisfying $\mathrm{U}\nu^* - \mathrm{U}\mu^* = \mathrm{U}\nu - \mathrm{U}\mu$. We illustrate it by the following example.

Example 3.4.11. Consider the Skorokhod embedding problem $\mathbf{SEP^*(\boldsymbol{\nu}, \boldsymbol{\mu})}$ (we drop the diffusion coefficient σ here because the underlying process is a Brownian motion W), where

$$\nu \;=\; \delta_0, \quad \mu \;=\; \frac{1}{3}\left(\delta_{-1} + \delta_0 + \delta_1\right).$$

The pair (ν, μ) obviously does not satisfy the condition in Theorem 3.4.2. However, we still try to construct the reversed barrier by the variational inequality $\mathbf{EVI^*(\boldsymbol{\nu}, \boldsymbol{\mu})}$. Here

$$a \;=\; \frac{1}{2}; \quad b \;=\; -\lambda\,\mathrm{sgn}(x); \quad c \;=\; f \;=\; 0; \quad \psi \;=\; \bar{v} \;=\; \frac{2}{3}(|x| - 1) \wedge 0.$$

Now we define two sub-probability measures $\nu_0 = \frac{2}{3}\delta_0$ and $\mu_0 = \frac{1}{3}(\delta_{-1} + \delta_1)$. It is easy to check, although ν_0 and μ_0 here are not probability measures, we have $\mathbb{P}^{\nu_0}[W_{\tau_D} \in \mathrm{d}x] = \mu_0(\mathrm{d}x)$, where $\tau_D = \inf\{t : |W_t| > 1\}$ is a Rost stopping time. On the other hand, since $\mathrm{U}\mu_0 - \mathrm{U}\nu_0$ is also $\frac{2}{3}(|x| - 1) \wedge 0$, $u = -\mathbb{E}^{\nu_0}[|x - W_{t\wedge\tau_D}|]$ is the strong solution to $\mathbf{EVI^*(\boldsymbol{\nu}, \boldsymbol{\mu})}$. However, the reversed barrier given by u, $B = \{(x,t) \in \bar{\mathbb{R}} \times \bar{\mathbb{R}}_+ : |x| \geq 1\}$ does not embed μ into W starting at 0 — in fact, $\mathcal{L}(W_{\tau_D}) = \frac{1}{2}(\delta_{-1} + \delta_1) \neq \mu$. Now, can we use D given by $\mathbf{EVI^*(\boldsymbol{\nu}, \boldsymbol{\mu})}$ to construct the embedding? The remainder of this section contains the answer, and we will complete this example at last.

First of all, we note that for ν, μ with disjoint support, the reversed barrier exists (Chacon [1985, Theorem 3.24]). Then the strong solution to $\mathbf{EVI^*(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})}$ gives the solution, denoted by τ_D to $\mathbf{SEP^*(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})}$ as we have discussed in last section. Now we suppose ν and μ are not disjoint supported. As in measure theory, for a signed measure m, we define, for $A \in \mathcal{F}$,

$$\begin{aligned} m_+(A) &= \sup\{m(B) : B \subset A, B \in \mathcal{F}\}; \\ m_-(A) &= \sup\{-m(B) : B \subset A, B \in \mathcal{F}\}. \end{aligned}$$

By Jordan-Hahn decomposition (see e.g. Halmos [1950, Section 29]), there exists $E \in$

$\mathcal{B}(\mathbb{R})$ such that for all $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{cases} \nu_0(A) \;:=\; (\nu-\mu)_+(A) \;=\; \nu(A\cap E)-\mu(A\cap E), & \text{and} \;\; \nu^+ \;:=\; \nu-\nu_0; \\ \mu_0(A) \;:=\; (\nu-\mu)_-(A) \;=\; \mu(A\cap E^{\complement})-\nu(A\cap E^{\complement}), & \text{and} \;\; \mu^+ \;:=\; \mu-\mu_0. \end{cases} \tag{3.4.7}$$

We then have the following easy lemma.

Lemma 3.4.12. *If ν and μ are probability distributions on $\mathbb{R}$, and the decomposition of ν and μ is defined as in (3.4.7), we have*

(i). *ν_0 and μ_0 are disjoint and non-negative;*

(ii). *$0 \le \nu_0(\mathbb{R}) = \mu_0(\mathbb{R}) \le 1$, where $\nu_0(\mathbb{R}) = 0$ if and only if $\nu = \mu$;*

(iii). *$\mathrm{U}\mu_0 - \mathrm{U}\nu_0 = \mathrm{U}\mu - \mathrm{U}\nu$;*

(iv). *$\nu^+ = \mu^+$.*

Proof. First of all, (i) is obvious. The following equality shows the first part of (ii)

$$\nu_0(\mathbb{R}) \;=\; \nu(E)-\mu(E) \;=\; \mu(E^{\complement})-\nu(E^{\complement}) \;=\mu_0(\mathbb{R}).$$

Now we suppose $\nu_0(\mathbb{R}) = 0$. Since $\nu_0(\mathbb{R}) = (\nu-\mu)_+(\mathbb{R}) = \sup\{\nu(A)-\mu(A) : A \in \mathcal{B}(\mathbb{R})\}$, we have $\nu \le \mu$. If there exists A such that $\nu(A) < \mu(A)$, we must have $\nu(A^{\complement}) > \mu(A^{\complement})$ since $\nu(\mathbb{R}) = \mu(\mathbb{R}) = 1$. This contradicts the fact that $\nu \le \mu$. Therefore $\nu(A) = \mu(A)$ for all $A \in \mathcal{B}(\mathbb{R})$. We have shown (ii) holds.

Moreover, we have

$$\begin{aligned} \mathrm{U}\mu_0 \;&=\; \int_{E^{\complement}} |x-y|\nu(\mathrm{d}y) - \int_{E^{\complement}} |x-y|\mu(\mathrm{d}y) \\ &= \mathrm{U}\mu - \mathrm{U}\nu - \int_E |x-y|\nu(\mathrm{d}y) + \int_E |x-y|\mu(\mathrm{d}y) \;= \mathrm{U}\mu - \mathrm{U}\nu + \mathrm{U}\nu_0, \end{aligned}$$

We have proved (iii). At last, (iv) holds because for any $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned} \nu^+(A) \;&=\; \nu(A) - \nu(A\cap E) + \mu(A\cap E) \\ &=\; \nu(A\cap E^{\complement}) + \mu(A) - \mu(A\cap E^{\complement}) \;=\; \mu^+(A). \end{aligned}$$

$\square$

Assume two distinct measures ν and μ such that $\mathrm{U}\mu \le \mathrm{U}\nu$, by Lemma 3.4.12 (i) – (iii), we can find the Rost's solution τ_D to $\mathbf{SEP}^*(\boldsymbol{\sigma}, \boldsymbol{\nu}^*, \boldsymbol{\mu}^*)$ where the probability distributions ν^* and μ^* are the normalization of ν_0 and μ_0, that is $\nu^*(A) = \nu_0(A)/\nu_0(\mathbb{R})$ and $\mu^*(A) = \mu_0(A)/\mu_0(\mathbb{R})$ for $A \in \mathcal{B}(\mathbb{R})$. The following theorem gives the embedding for (ν, μ) by a randomized stopping time based on τ_D.

Theorem 3.4.13. *Assume that ν and μ are two distinct probability distributions satisfying $\mathrm{U}\mu \le \mathrm{U}\nu$, and decomposed as (3.4.7). Also assume τ^* is a stopping time embedding μ^* into X with the initial distribution ν^* where ν^* and μ^* are the normalization of ν_0 and μ_0. Denote the Radon-Nikodym derivative $\mathrm{d}\nu^+/\mathrm{d}\nu$ by f, and let Z be a random variable independent of the underlying process X and uniformly distributed on $(0,1)$. Moreover, define*

$$T \;=\; \begin{cases} 0, & \textit{if } \; Z \le f(X_0); \\ \infty, & \textit{if } \; Z > f(X_0). \end{cases}$$

and $\tau = \tau^* \wedge T$. Then for any $g \in \mathbb{C}_b^0$,

$$\mathbb{E}^\nu[\, g(X_\tau)\,] \;=\; \langle\, \mu,\, g\,\rangle.$$

Proof. First of all, since ν_0 is a sub-probability measure with $\nu_0(A) \le \nu(A)$, $\forall A \in \mathcal{B}(\mathbb{R})$, the definition of f makes sense, and moreover $0 \le f \le 1$. For any $g \in \mathbb{C}_b^0$, we have

$$\mathbb{E}^\nu[\, g(X_\tau)\,] = \mathbb{E}^\nu[\, g(X_{\tau^*})\mathbf{1}_{[Z > f(X_0)]}\,] + \mathbb{E}^\nu[\, g(X_0)\mathbf{1}_{[Z \le f(X_0)]}\,],$$

where, by Lemma 3.4.12 (ii),

$$\begin{aligned} \mathbb{E}^\nu[\, g(X_{\tau^*})\mathbf{1}_{[Z > f(X_0)]}\,] \;&=\; \int_{\mathbb{R}} \int_0^1 \mathbf{1}_{[z > f(y)]} \mathbb{E}^y[\, g(X_{\tau^*})\,]\, \mathrm{d}z\, \nu(\mathrm{d}y) \\ &=\; \int_{\mathbb{R}} \mathbb{E}^y[\, g(X_{\tau^*})\,]\, \nu(\mathrm{d}y) - \int_{\mathbb{R}} \mathbb{E}^y[\, g(X_{\tau^*})\,] f(y)\, \nu(\mathrm{d}y) \\ &=\; \int_{\mathbb{R}} \mathbb{E}^y[\, g(X_{\tau^*})\,]\, (\nu - \nu^+)(\mathrm{d}y) \;=\; \mathbb{E}^{\nu_0}[\, g(X_{\tau^*})\,] \\ &=\; \nu_0(\mathbb{R}) \cdot \mathbb{E}^{\nu^*}[\, g(X_{\tau^*})\,] \;=\; \langle\, \mu_0(\mathbb{R}) \cdot \mu^*,\, g\,\rangle \;=\; \langle\, \mu_0,\, g\,\rangle, \end{aligned}$$

and, by Lemma 3.4.12 (iv),

$$\begin{aligned} \mathbb{E}^\nu[\, g(X_0)\mathbf{1}_{[Z \le f(X_0)]}\,] \;&=\; \int_{\mathbb{R}} \int_0^1 g(y)\mathbf{1}_{[z \le f(y)]}\, \mathrm{d}z\, \nu(\mathrm{d}y) \\ &=\; \int_{\mathbb{R}} g(y) f(y)\, \nu(\mathrm{d}y) \;=\; \langle\, \nu^+,\, g\,\rangle \;=\; \langle\, \mu^+,\, g\,\rangle. \end{aligned}$$

The desired result follows immediately. $\square$

Now we return to the example.

Continuation of Example 3.4.11. In fact, the measures ν_0 and μ_0 defined as $\frac{2}{3}\delta_0$ and $\frac{1}{3}(\delta_{-1}+\delta_1)$ coincide with the definition in (3.4.7). Moreover, we define $\nu^+ = \mu^+ = \frac{1}{3}\delta_0$. We have found Rost's embedding τ_D for (ν_0, μ_0). It is easy to see $f(x) = \mathrm{d}\nu^+/\mathrm{d}\nu \equiv 1/3$. The random variables Z and T are defined as in Theorem 3.4.13, then T is a random variable independent of W and distributed as $\mathbb{P}[T = 0] = 1/3$ and $\mathbb{P}[T = \infty] = 2/3$. Thus, for $\tau := \tau_D \wedge T$,

$$\begin{aligned}
\mathbb{P}^\nu[\,W_\tau = 1\,] &= \frac{2}{3}\,\mathbb{P}^\nu[\,W_{\tau_D} = 1\,] + \frac{1}{3}\,\mathbb{P}^\nu[\,W_0 = 1\,] &= \frac{1}{3};\\
\mathbb{P}^\nu[\,W_\tau = 0\,] &= \frac{2}{3}\,\mathbb{P}^\nu[\,W_{\tau_D} = 0\,] + \frac{1}{3}\,\mathbb{P}^\nu[\,W_0 = 0\,] &= \frac{1}{3};
\end{aligned}$$

similarly, $\mathbb{P}^\nu[\,W_\tau = -1\,] = 1/3$. That is $W_\tau \sim \mu$. We have completed the construction of the embedding by the randomization of the Rost's stopping time.

As conclusion, given a pair of different distributions $\{\nu, \mu\}$ with $\mathrm{U}\mu \le \mathrm{U}\nu$, without the knowledge of the existence of the corresponding reversed barrier, we can always find the solution, v, of **EVI***$(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$, and then obtain a reversed barrier D by (3.2.4) where $\psi = \mathrm{U}\mu - \mathrm{U}\nu$. Now we test that if τ_D is the desired embedding for (ν, μ) or not. If yes, then we have completed the construction of the reversed barrier. If not, by Theorem 3.4.6, we can assert that no solution to the problem **SEP***$(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ exists. So, by Theorem 3.4.13, we can construct a *randomized Rost's stopping time* to solve the Skorokhod embedding problem.

3.5 A Numerical Example

Using the results we have obtained, we can find Root's and Rost's barrier by variational inequalities, at least, numerically. In this section, we will give a numerical example to see that.

In this example, we define $\nu = \delta_0$, μ is the uniform distribution on $(-1, 1)$ and the underlying process is a Brownian motion. We will realize the numerical results in MATLAB. We consider Root's case firstly. Running the code (see Appendix A), we can get v and r as the approximation of the solutions to **EVI(μ)** and **SEP(μ)** and see them graphically (we take steps $m = 200$, $n = 10000$ and $T = 0.5$):

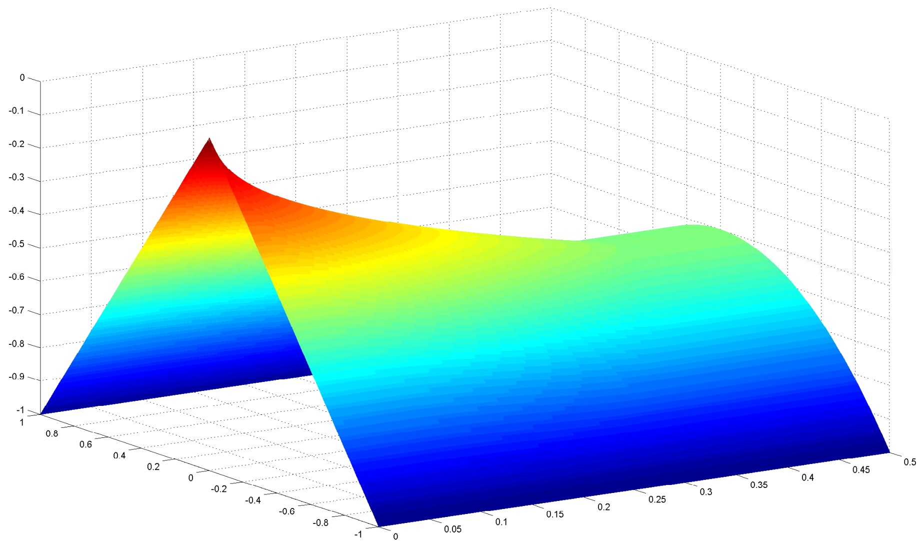


Figure 3-2: The numerical solution v to EVI(μ), ($m = 200$, $n = 10000$).

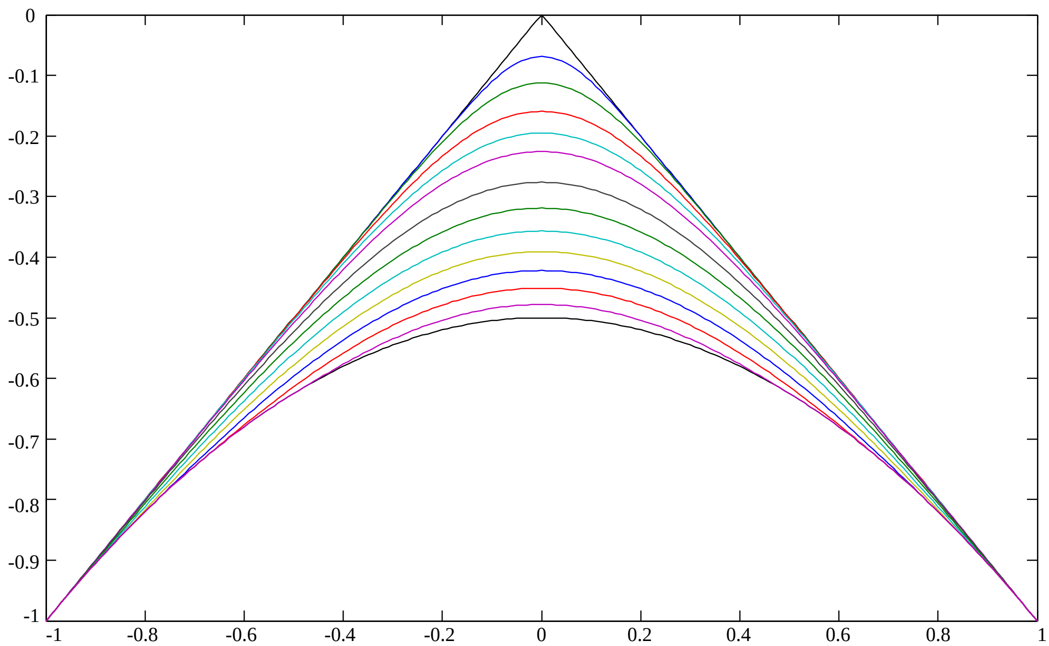


Figure 3-3: The evolution of potential (Root's case).

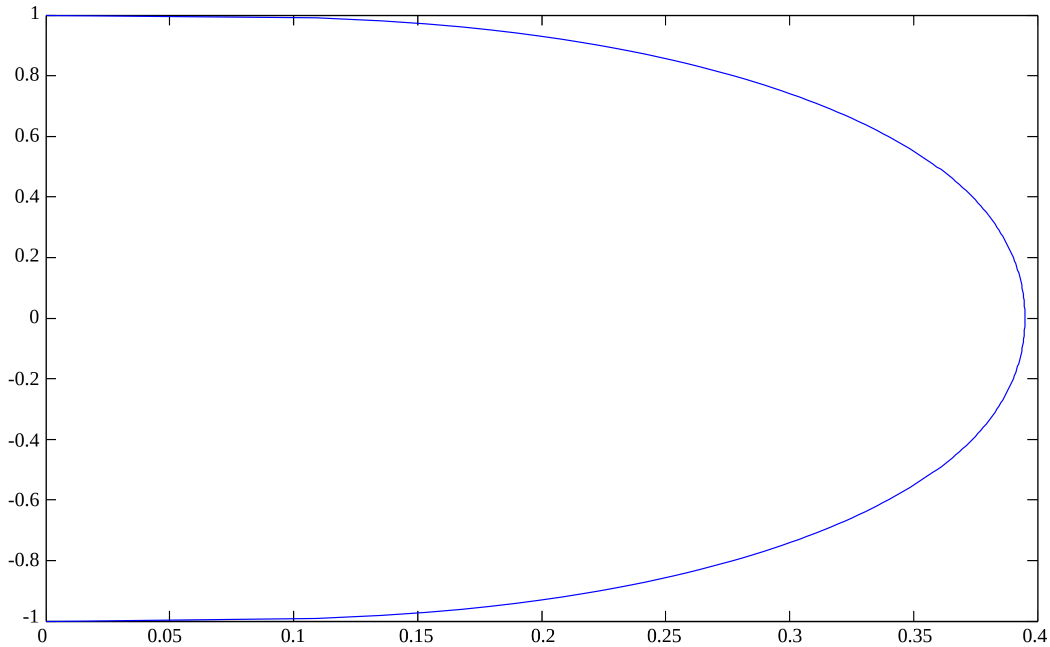


Figure 3-4: The numerical solution R to SEP(μ).

For the case of reversed barrier, we run the code in Appendix B. Taking steps $m = 200$, $n = 10000$ and $T = 2.5$, we can see the numerical solutions to $\mathbf{EVI^*(\mu)}$ and $\mathbf{SEP^*(\mu)}$ graphically (note in Rost's case that our approximation is effective only if $t < T$):

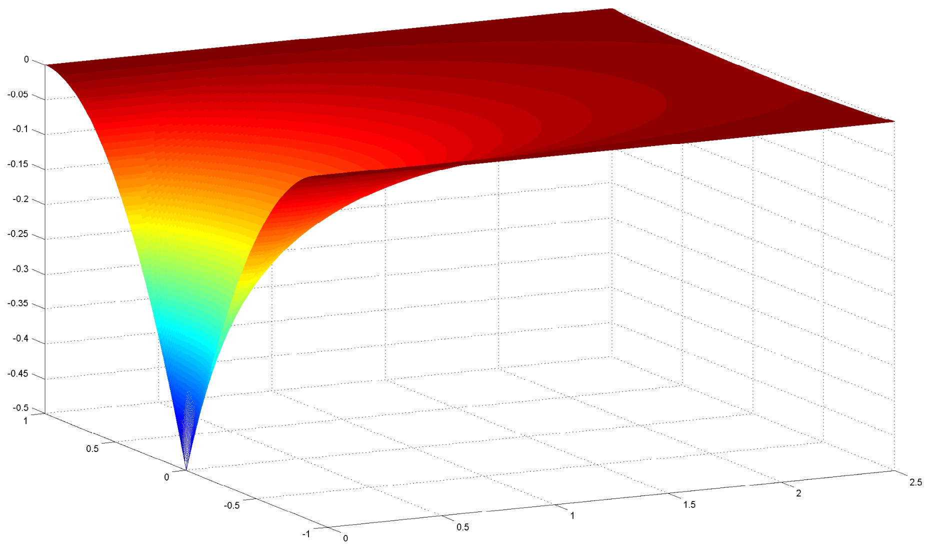


Figure 3-5: The numerical solution to EVI$^*(\mu)$, $(m = 200,\ n = 10000)$.

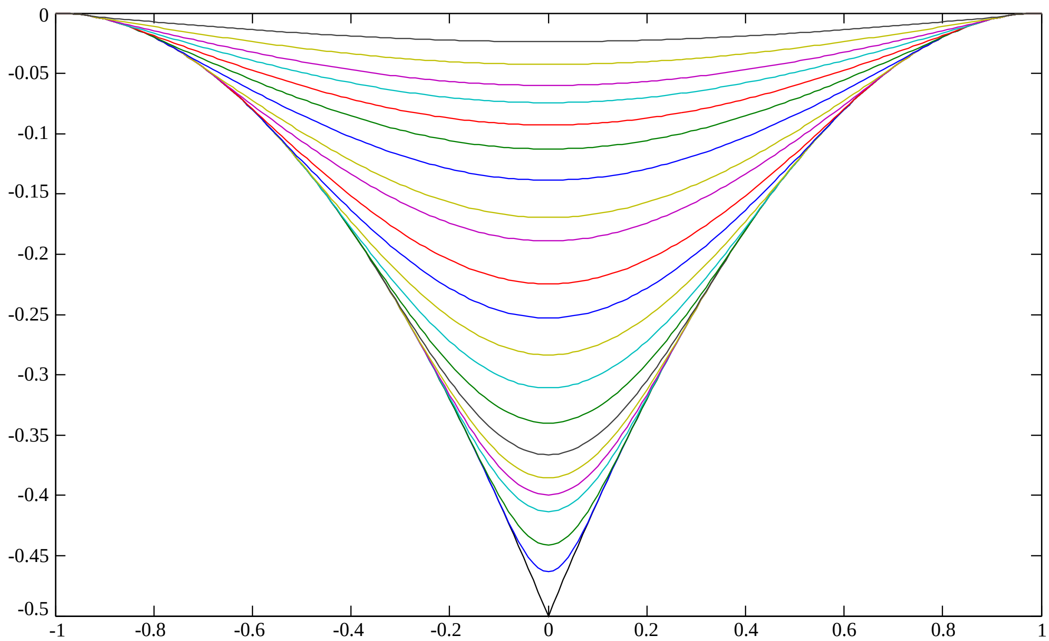


Figure 3-6: The evolution of potential (Rost's case).

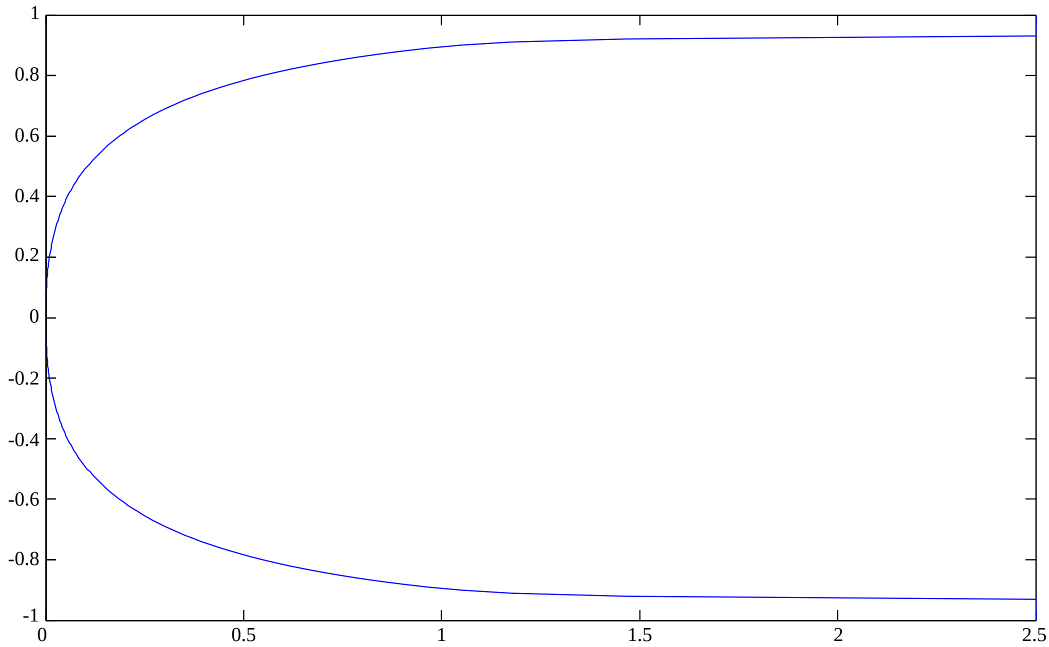


Figure 3-7: The numerical solution $\boldsymbol{R}$ to SEP$^*(\boldsymbol{\mu})$. *The approximation here is only effective when* $R(x) < 2.5$. *In fact, what we present here is the plot of* $x \mapsto R(x) \wedge 2.5$.

Chapter 4

Optimality and Applications in Finance

So far we have analysed Root's and Rost's solutions to the Skorokhod embedding problem. Our motivation for this is recent work connecting the solutions to problems arising in mathematical finance, specially, model-independent bounds for *variance options*, which has been observed by Dupire [2005], Carr and Lee [2010], Hobson [2009].

The financial motivation can be described as follows: consider a (discounted) asset which has dynamics under the risk-neutral measure

$$\frac{\mathrm{d}S_t}{S_t} \;=\; \sigma_t \mathrm{d}W_t,$$

where the process σ_t is not necessarily known. We are interested in variance options, which are contracts where the payoff depends on the realised quadratic variation of the log-price process: specifically, we have

$$\mathrm{d}(\ln S_t) \;=\; \sigma_t \mathrm{d}W_t - \frac{1}{2}\sigma_t^2 \mathrm{d}t,$$

and therefore

$$\langle \ln S \rangle_T \;=\; \int_0^T \sigma_t^2 \mathrm{d}t.$$

An option on variance is then an option with payoff $F(\langle \ln S \rangle_T)$ for some function F. Important examples include variance swaps, which pay the holder $\langle \ln S \rangle_T - K$, and variance calls which pay the holder $(\langle \ln S \rangle_T - K)_+$. We shall be particularly interested in the case of a variance call, but our results will extend to a wider class payoffs.

In fact, variance options have been a topic of much interest in recent years, both from the industrial point of view, where innovations such as the VIX index have contributed to a large growth in products which are directly dependent on quantities derived from the quadratic variation, and also on the academic side, with a number of interesting contribution in the literature. The academic results go back to work of Dupire [1993] and Neuberger [1994], who noted that a variance swap – that is a contract which pays $\langle \ln S \rangle_T$ – can be replicated model-independently using a contract paying the logarithm of the asset at maturity through the identity (from Itô's Lemma):

$$\ln(S_T) - \ln(S_0) \;=\; \int_0^T \frac{1}{S_t} \mathrm{d}S_t - \frac{1}{2} \langle \ln S \rangle_T .$$

More recently, work on options and swaps on volatility and variance, (in a model-based setting) includes Howison et al. [2004], Broadie and Jain [2008], Kallsen et al. [2010]. Other work Keller-Ressel and Muhle-Karbe [2010], Keller-Ressel [2011] has considered the differences between the theoretical ($\langle \ln S \rangle_T$) and the discrete approximation ($\sum_k \ln(S_{(k+1)\delta}/S_{k\delta})^2$) which is usually specified in the contract . Finally, several papers have considered variants on the model-independent problems, Carr and Lee [2010], Carr et al. [2011], Davis et al. [2010], or problems where the modelling assumptions are fairly weak. This latter framework is of particular interest for options on variance, since the markets since the markets for such products are still fairly young, and so making strong modelling assumptions might not be as strongly justified as it could be in a well-established market.

Now we let $\mathrm{d}\widetilde{X}_t = \widetilde{X}_t \mathrm{d}\widetilde{W}_t$ for a suitable Brownian motion $\widetilde{W}$ and we can find a continuous time change τ_t such that $S_t = \widetilde{X}_{\tau_t}$, and so:

$$\mathrm{d}\tau_t \;=\; \frac{\sigma_t^2 S_t^2}{S_t^2} \mathrm{d}t.$$

Hence,

$$\left(\widetilde{X}_{\tau_T}, \tau_T\right) \;=\; \left(S_T, \int_0^T \sigma_t^2 \mathrm{d}t\right) \;=\; \left(S_T, \langle \ln S \rangle_T\right).$$

Now suppose that we know the prices of call options on S_T with maturity T, and at all strikes (recall that σ_t is not assumed known). Then we can derive the law of S_T under the risk-neutral measure from the Breeden-Litzenberger [1978] formula. Call this law

μ. This suggests that the problem of finding a lower bound on the price of a variance call (for an unknown σ) is equivalent to:

$$\begin{aligned} &\text{find a stopping time } \tau \text{ to minimise } \mathbb{E}[\tau - K]_+, \\ &\qquad\qquad\qquad\qquad\qquad\text{subject to } \mathcal{L}(\widetilde{X}_\tau) = \mu. \end{aligned}$$

This is essentially the problem for which Rost [1976] has shown that the solution is given by Root's barrier (Theorem 2.1.5). (In fact, the result trivially extends to payoff of the form $F(\langle \ln S \rangle_T)$ where $F(\cdot)$ is a convex increasing function.)

In the following sections, we show that the lower bound which is implied by Rost's result can be enforced through a suitable hedging strategy, which allows an arbitrage whenever the price of a variance call trades below the given lower bound. To accomplish it, we will give a novel proof of the optimality of Root's barrier, and from this barrier, we will be able to derive a suitable hedging strategy.

4.1 Optimality of Root's Solution

For a given distribution μ, Rost [1976] proves that Root's construction is optimal in the sense of 'minimal residual expectation' (see Definition 2.1.4 and Theorem 2.1.5). It is easy to check that this is equivalent to the slightly more general optimisation problem:

OPT($\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu}$): Suppose (2.1.7) holds and X satisfies (2.1.1). Find a stopping time $\tau \in \mathcal{T}(\sigma, \nu, \mu)$ such that, for a given increasing convex function F with $F(0) = 0$,

$$\mathbb{E}^\nu\big[F(\tau)\big] \;=\; \inf_{\rho \in \mathcal{T}(\sigma,\nu,\mu)} \mathbb{E}^\nu\big[F(\rho)\big].$$

Our aim in this section is twofold. Firstly, since Rost's original proof relies heavily on notions from potential theory, to give a proof of this result using machinery from probability. Secondly, we shall be able to give a 'path-wise inequality' which encodes the optimality in the sense that we can find a submartingale G, and a function $H(x)$ such that

$$F(t) \;\geq\; G_t + H(X_t) \tag{4.1.1}$$

and such that, for τ_D, equality holds in (4.1.1) and $G_{t\wedge\tau_D}$ is a uniformly integrable martingale. It then follows that τ_D does indeed minimise $\mathbb{E}[F(\tau)]$ among all solutions to the Skorokhod Embedding problem. The importance of (4.1.1) is that we can characterise the submartingale G_t, which will correspond in the financial setting to a dynamic trading strategy for constructing a sub-replicating hedging strategy for call-type payoffs on variance options.

First of all, the following examples illustrate roughly our idea to find the path-wise inequality.

Example 4.1.1. Suppose we take Root's barrier $D := \{(x,t) : t < R(x)\}$ with the boundary function

$$R(x) \;=\; -\lambda(x+\alpha)(x-\beta)\mathbf{1}_{(-\alpha,\beta)},$$

where $\lambda, \alpha, \beta > 0$. Given a standard Brownian motion W and Root's stopping time

$$\tau_D \;=\; \inf\{t > 0 : \, t \geq R(W_t)\},$$

define $\mu := \mathcal{L}(W_{\tau_D})$. Now we consider the optimal stopping problem $\mathbf{OPT(\boldsymbol{\mu})}$ (we drop σ and ν when $\sigma \equiv 1$ and $\nu = \delta_0$) with $F(t) = \frac{1}{2}t^2$.

For $(x,t) \in \mathbb{R}\times\mathbb{R}_+$, define $M(x,t) = \mathbb{E}^{(x,t)}[\tau_D]$. Then if $t \geq R(x)$, $M(x,t) = t$. If $0 \leq t < R(x)$, we have

$$\begin{cases} M(x,t) \;=\; -\mathbb{E}^{(x,t)}\left[\lambda(W_{\tau_D}+\alpha)(W_{\tau_D}-\beta)\right]; & \text{(by the definition of } \tau_D) \\ M(x,t) \;=\; \mathbb{E}^{(x,t)}\left[W_{\tau_D}^2\right] - x^2 + t, & \text{(by Itô's formula)} \end{cases}$$

so

$$M(x,t) \;=\; \frac{\lambda}{1+\lambda}\left[\, t - (x+\alpha)(x-\beta)\right], \qquad \text{for } 0 \leq t < R(x).$$

It is easy to verify that $(\partial/\partial t + \frac{1}{2}\partial^2/\partial x^2)M = 0$ on D and M is continuous on $\mathbb{R}\times\mathbb{R}_+$.

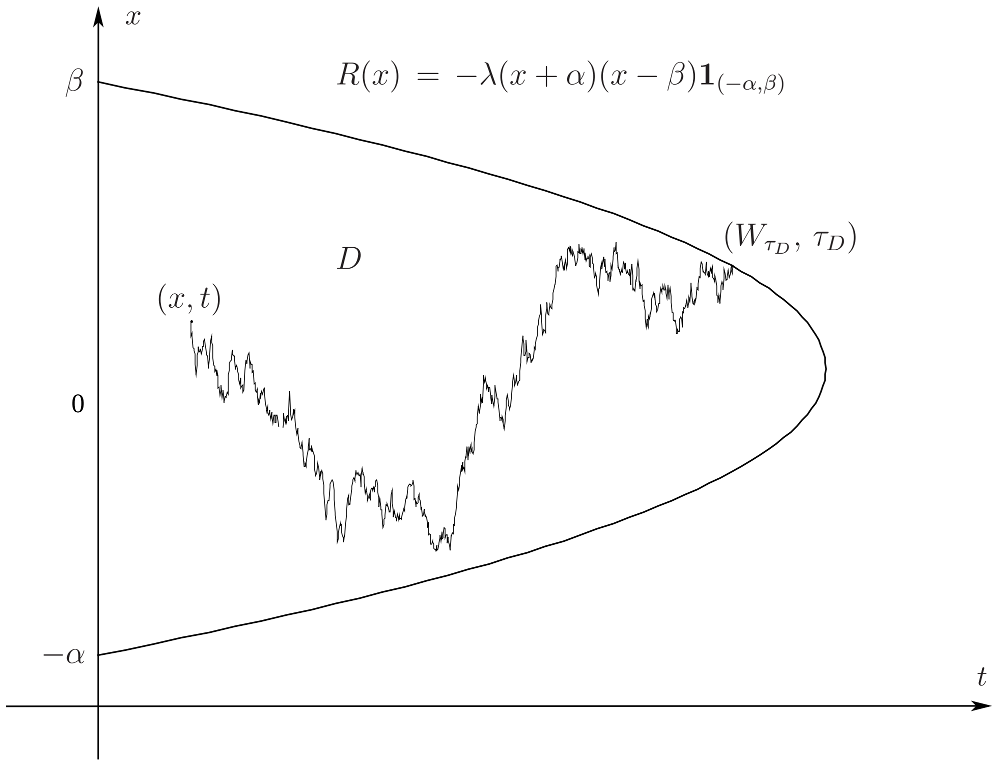


Figure 4-1: *Example 4.1.1*

Now define

$$
\begin{aligned}
Z(x) &= 2\int_0^x \int_0^y M(z,0)\mathrm{d}z\mathrm{d}y \;=\; -2\int_0^x \int_0^y \frac{\lambda}{1+\lambda}(z+\alpha)(z-\beta)\mathbf{1}_{(-\alpha,\beta)}\mathrm{d}z\mathrm{d}y, \\
G(x,t) &= \int_0^t M(x,s)\mathrm{d}s - Z(x) \\
&= \begin{cases} \dfrac{\lambda}{1+\lambda}\left[\dfrac{t^2}{2} - t(x+\alpha)(x-\beta)\right] - Z(x), & \text{if } \ 0 \le t < R(x); \\ \dfrac{R^2(x)}{2(1+\lambda)} + \dfrac{1}{2}t^2 - Z(x), & \text{if } \ t \ge R(x). \end{cases}
\end{aligned}
$$

Hence, we have

$$\frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x^2} \;=\; 0 \quad \text{on } D.$$

Therefore, $(G(W_{t\wedge\tau_D}, t\wedge\tau_D);\, t \ge 0)$ is a martingale. On the other hand, for $x \in \mathbb{R}\backslash(-\alpha,\beta)$, obviously,

$$\frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x^2} \;=\; t \;\ge\; 0.$$

For $x \in (-\alpha, \beta)$ and $t \geq R(x)$,

$$\begin{aligned} \frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x^2} &= \frac{1}{2(1+\lambda)}\left[R'(x)^2 - 2\lambda R(x)\right] - \frac{R(x)}{1+\lambda} + t \\ &= \frac{R'(x)^2}{2(1+\lambda)} - R(x) + t \;\geq\; \frac{R'(x)^2}{2(1+\lambda)} \;\geq\; 0. \end{aligned}$$

Therefore, $(G(W_t, t);\, t \geq 0)$ is a continuous submartingale, and hence, since it is a martingale before hitting the barrier, we have

$$\mathbb{E}[G(W_\tau, \tau)] \;\geq\; \mathbb{E}[G(W_{\tau_D}, \tau_D)]$$

where τ is an arbitrary stopping time.

Moreover, we have

$$\begin{aligned} &\int_0^t M(x,s)\mathrm{d}s + \int_0^{R(x)} [f(s) - M(x,s)]\mathrm{d}s - F(t) \\ &\qquad\qquad = \begin{cases} -\dfrac{[\,R(x) - t\,]^2}{2(1+\lambda)}, & \text{if } \; 0 \leq t < R(x); \\ 0, & \text{if } \; t \geq R(x). \end{cases} \end{aligned}$$

Therefore, for $(x,t) \in \mathbb{R} \times \mathbb{R}_+$,

$$\int_0^t G(x,s)\mathrm{d}s + \int_0^{R(x)} [f(s) - M(x,s)]\mathrm{d}s + Z(x) \;\leq\; F(t),$$

so for any UI stopping time τ such that $\mathcal{L}(W_\tau) = \mu = \mathcal{L}(W_{\tau_D})$,

$$\begin{aligned} \mathbb{E}\big[F(\tau)\big] \;&\geq\; \mathbb{E}\left[G(W_\tau, \tau) + F(R(W_\tau)) - \int_0^{R(W_\tau)} M(W_\tau, s)\mathrm{d}s + Z(W_\tau)\right] \\ &\geq\; \mathbb{E}\left[G(W_{\tau_D}, \tau_D)\right] + \mathbb{E}\left[F(R(W_{\tau_D})) - \int_0^{R(W_{\tau_D})} M(W_{\tau_D}, s)\mathrm{d}s + Z(W_{\tau_D})\right] \\ &=\; \mathbb{E}\left[F(R(W_{\tau_D}))\right] + \mathbb{E}\left[\int_{R(W_{\tau_D})}^{\tau_D} M(W_{\tau_D}, s)\mathrm{d}s\right] \;=\; \mathbb{E}\big[F(\tau_D)\big], \end{aligned} \tag{4.1.2}$$

which shows the optimality of Root's barrier.

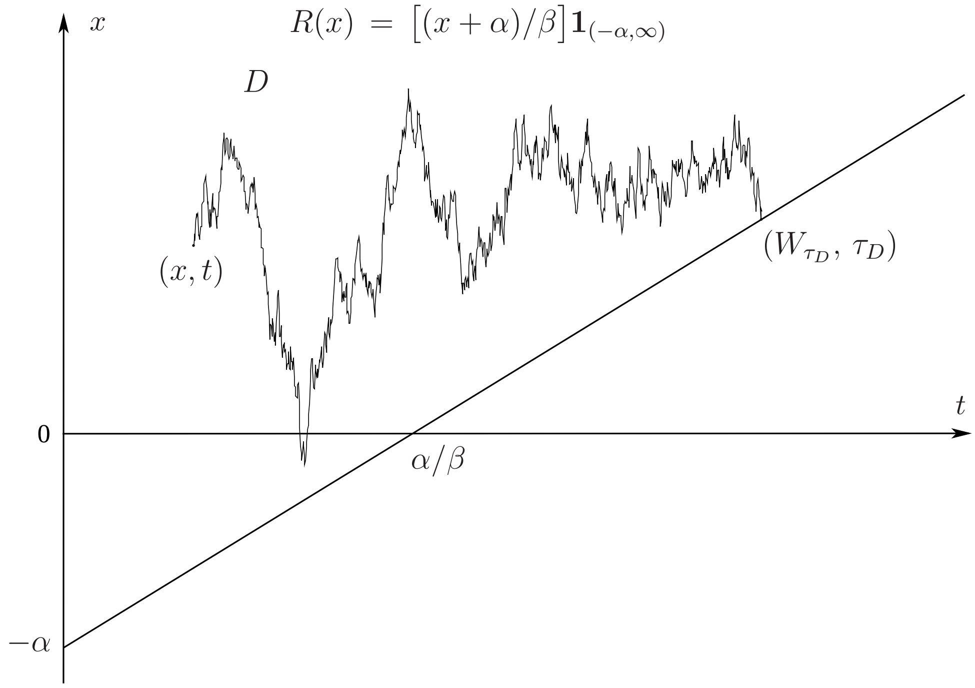


Figure 4-2: *Example 4.1.2*

The other example can be seen in Hobson [2009, Example 5.6] or Huff [1975], we will show the same result in a different way.

Example 4.1.2. Suppose we have a barrier $D := \{(x,t) : t < R(x)\}$ with the boundary function $R(x) = [(x+\alpha)/\beta]\mathbf{1}_{(-\alpha,\infty)}$, where $\alpha, \beta > 0$. Given a standard Brownian motion W and Root's stopping time $\tau_D = \inf\{t > 0 : t \geq R(W_t)\}$, define $\mu := \mathcal{L}(W_{\tau_D})$. Now we consider the optimization problem $\mathbf{OPT(\boldsymbol{\mu})}$ with $F(t) = e^{\eta t}$, $(2\eta < \beta^2)$ and then $f(t) := F'(t) = \eta e^{\eta t}$. By the calculations in Hobson [2009, Example 5.6], we have

$$M(x,t) \;:=\; \mathbb{E}^{(x,t)}[f(\tau_D)] = \begin{cases} \eta e^{\alpha\phi}\exp\{\phi x + (\eta - \beta\phi)t\}, & \text{for } 0 \leq t < R(x), \\ \eta e^{\eta t}, & \text{for } t \geq R(x), \end{cases}$$

where $\phi = \beta - \sqrt{\beta^2 - 2\eta}$. Since $\phi^2 = 2(\beta\phi - \eta)$, $(\partial/\partial t + \frac{1}{2}\partial^2/\partial x^2)M = 0$ on D.

Define G and Z in the same way as in Example 4.1.1, and we have

$$G(x,t) \;:=\; \begin{cases} \dfrac{1}{\eta - \beta\phi}\big[\,M(x,t) - M(x,0)\,\big] - Z(x), & \text{for } \; 0 \leq t < R(x); \\ \dfrac{\beta\phi}{\eta - \beta\phi}e^{\eta R(x)} - \dfrac{\eta}{\eta - \beta\phi}e^{\beta\phi R(x)} + e^{\eta t} - Z(x), & \text{for } \; t \geq R(x). \end{cases}$$

Thus, for $0 < t < R(x)$, $(\frac{1}{2}\partial^2/\partial x^2 + \partial/\partial t)G = 0$; for $t \geq R(x)$, and $x \leq -\alpha$, then $(\frac{1}{2}\partial^2/\partial x^2 + \partial/\partial t)G = \eta e^{\eta t} \geq 0$, and if $x > -\alpha$,

$$\frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x^2} \;=\; \eta e^{\eta t} - \frac{\eta^2}{\beta\phi} e^{\eta R(x)} \;\geq\; \frac{\eta}{\beta\phi}(\beta\phi - \eta)e^{\eta t} \;=\; \frac{\eta\phi}{2\beta}e^{\eta t} \;\geq\; 0.$$

Moreover, we can check, for $(x,t) \in \mathbb{R} \times \mathbb{R}_+$,

$$\int_0^t G(x,s)\mathrm{d}s + \int_0^{R(x)} [f(s) - M(x,s)]\mathrm{d}s + Z(x) \;\leq\; F(t),$$

so for any $\tau \in \mathcal{T}(\mu)$, as in (4.1.2), optimality of the Root's barrier follows.

These two examples suggest to us how to find the suitable pair (G, H) in (4.1.1). Consider the optimisation problem $\mathbf{OPT}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$ with the convex function F. We suppose firstly that we have solved $\mathbf{SEP}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$, and hence have our barrier D with the barrier function R lower semi-continuous. Define the function

$$M(x,t) \;=\; \mathbb{E}^{(x,t)}\left[f(\tau_D)\right], \tag{4.1.3}$$

where τ_D is the corresponding Root stopping time and f is the right derivative of F. In the following, we shall assume:

$$M(x,t) \;<\; \infty, \quad \forall\,(x,t) \in \mathbb{R} \times \mathbb{R}_+\,. \tag{4.1.4}$$

In fact, the assumption $\beta^2 > 2\eta$ ensures that (4.1.4) holds in Example 4.1.2. We suppose also (at least initially) that (2.5.2a) – (2.5.2c) and (2.5.16) still hold. Note that $M(x,t)$ now has the following important properties. First, since f is right-continuous, $M(x,t) = f(t)$ whenever $(x,t) \notin D$ and $t > 0$. In addition, since f is increasing, for all x and t we have $M(x,t) \geq f(t)$.

Now, as in the examples above, define

$$Z(x) = 2\int_0^x \int_0^y \frac{M(z,0)}{\sigma^2(z)}\,\mathrm{d}z\,\mathrm{d}y. \tag{4.1.5}$$

So, in particular, we have $Z''(x) = 2M(x,0)/\sigma^2(x)$ and $Z(x)$ is a convex function. Define also:

$$G(x,t) \;=\; \int_0^t M(x,s)\,\mathrm{d}s - Z(x). \tag{4.1.6}$$

and

$$H(x) \;=\; \int_0^{R(x)} [f(s) - M(x,s)]\mathrm{d}s + Z(x). \tag{4.1.7}$$

Two key results concerning these functions are then:

Proposition 4.1.3. *We have, for all $(x,t) \in \mathbb{R} \times \mathbb{R}_+$:*

$$G(x,t) + H(x) \;\leq\; F(t). \tag{4.1.8}$$

And also

Lemma 4.1.4. *Suppose that f is bounded, and, for any $T > 0$:*

$$\mathbb{E}\left[\int_0^T Z'(X_s)^2 \sigma(X_s)^2 \mathrm{d}s\right] \;<\; \infty, \quad \textit{and} \;\; \mathbb{E}\left[Z(X_0)\right] \;<\; \infty. \tag{4.1.9}$$

Then the process

$$(G(X_{t\wedge\tau_D}, t \wedge \tau_D);\, t \geq 0) \quad \textit{is a martingale,} \tag{4.1.10}$$

and

$$(G(X_t, t);\, t \geq 0) \quad \textit{is a submartingale.} \tag{4.1.11}$$

Using these results, we are able to prove the following theorem, which gives us Rost's result regarding the optimality of Root's barrier.

Theorem 4.1.5. *Suppose D solves* $\mathbf{SEP}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$*, and moreover, (4.1.4) and (4.1.9) hold. Then the corresponding Root's stopping time τ_D solves* $\mathbf{OPT}(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})$*.*

Proof. We begin by considering the case where $\mathbb{E}[\tau_D] < \infty$ and $\mathbb{E}[\tau] < \infty$, and moreover f is bounded, i.e.

$$\text{there exists } C \;<\; \infty, \;\text{ such that } f(t) \;\leq\; C. \tag{4.1.12}$$

Then, since $M(\cdot, 0)$ is also bounded by C on $\mathbb{R}$, we can apply the Meyer-Itô formula to Z (see Protter [2005, Theorem 71, Chapter IV]):

$$Z(X_t) \;=\; Z(X_0) + \int_0^t Z'(X_r)\mathrm{d}X_r + \frac{1}{2}\int_0^t Z''(X_r)\sigma^2(X_r)\mathrm{d}r.$$

By (4.1.9), we obtain

$$\mathbb{E}[Z(X_{t\wedge\tau})] \;=\; \mathbb{E}[Z(X_0)] + \mathbb{E}\left[\int_0^{t\wedge\tau} M(X_s,0)\mathrm{d}s\right] \;\leq\; \mathbb{E}[Z(X_0)] + C\mathbb{E}[\tau].$$

Applying Fatou's Lemma, we deduce that for any stopping time τ with finite expectation, $Z(X_\tau)$ is integrable. Moreover for such a stopping time, by convexity, $Z(X_{t\wedge\tau}) \leq \mathbb{E}[Z(X_\tau)|\mathcal{F}_t]$. Since, noting $Z(x) \geq 0$ for $x \in \mathbb{R}$,

$$\mathbb{E}[Z(X_\tau)|\mathcal{F}_t] \;\leq\; -Z(X_{t\wedge\tau}) \;\leq\; G(X_{t\wedge\tau}, t\wedge\tau) \;\leq\; \int_0^{t\wedge\tau} M(X_{t\wedge\tau}, s)\mathrm{d}s \;\leq\; C\tau,$$

we can see that $G(X_{t\wedge\tau}, t\wedge\tau)$ is a submartingale which is bounded below by a uniformly integrable martingale (since $Z(X_\tau)$ is integrable), and bounded above by $C\tau$. It follows that $\mathbb{E}[G(X_{t\wedge\tau}, t\wedge\tau)] \to \mathbb{E}[G(X_\tau,\tau)]$ as $t \to \infty$. The same arguments hold when we replace τ by τ_D.

Since $R(X_{\tau_D}) \leq \tau_D$ and if $t \geq R(x)$ then $\tau_D = t$, $\mathbb{P}^{(x,t)}$-a.s., so that $M(X_{\tau_D}, s) = f(s)$ for $s \geq \tau_D$, we have

$$\begin{aligned} G(X_{\tau_D},\tau_D) + &\int_0^{R(X_{\tau_D})} \Big[f(s) - M(X_{\tau_D}, s)\Big]\mathrm{d}s + Z(X_{\tau_D}) \\ &= \int_0^{\tau_D} M(X_{\tau_D}, s)\mathrm{d}s + \int_0^{R(X_{\tau_D})} \Big[f(s) - M(X_{\tau_D}, s)\Big]\mathrm{d}s \\ &= \int_0^{\tau_D} M(X_{\tau_D}, s)\mathrm{d}s + \int_0^{\tau_D} \Big[f(s) - M(X_{\tau_D}, s)\Big]\mathrm{d}s \\ &= \int_0^{\tau_D} f(s)\mathrm{d}s \;=\; F(\tau_D). \end{aligned} \tag{4.1.13}$$

On the other hand, since τ_D, $\tau \in \mathcal{T}(\sigma,\nu,\mu)$, and observing that $G(X_{\tau_D},\tau_D)$ and $F(\tau_D)$ are integrable (because f is bounded and τ_D is integrable), so too is $H(X_{\tau_D})$, and

$$\mathbb{E}\left[H(X_{\tau_D})\right] \;=\; \mathbb{E}\left[H(X_\tau)\right].$$

In addition, by Lemma 4.1.4 and the limiting behaviour deduced above, we have

$$\begin{aligned} \mathbb{E}\left[G(X_{\tau_D},\tau_D)\right] \;&=\; \lim_{t\to\infty} \mathbb{E}\left[G(X_{t\wedge\tau_D}, t\wedge\tau_D)\right] \\ &\leq\; \lim_{t\to\infty} \mathbb{E}\left[G(X_{t\wedge\tau}, t\wedge\tau)\right] \;=\; \mathbb{E}\left[G(X_\tau,\tau)\right]. \end{aligned}$$

Putting these together, we get

$$\mathbb{E}\left[F(\tau_D)\right] \;=\; \mathbb{E}\left[G(X_{\tau_D},\tau_D)+H(X_{\tau_D})\right] \;\leq\; \mathbb{E}\left[G(X_\tau,\tau)+H(X_\tau)\right] \;\leq\; \mathbb{E}\left[F(\tau)\right],$$

where the last inequality holds because of (4.1.8).

We now consider the case where at least one of τ or τ_D has infinite expectation. Note that if $F(\cdot) \not\equiv 0$, then there is some $\alpha, \beta \in \mathbb{R}$ with $\beta > 0$ such that $F(t) \geq \alpha + \beta t$, and hence we cannot have $\mathbb{E}[\tau] = \infty$ ($\mathbb{E}[\tau_D] = \infty$) without $\mathbb{E}[F(\tau)] = \infty$ ($\mathbb{E}[F(\tau_D)] = \infty$).

The only case which need concern us is the case where $\mathbb{E}[\tau] < \infty$, but $\mathbb{E}[\tau_D] = \infty$. Note however that τ_D remains a UI stopping time, so $\mathbb{E}[X_{t\wedge\tau_D}|\mathcal{F}_t] = X_t$. In addition, from the arguments applied above, we know $Z(X_\tau)$ is integrable and since $X_\tau \sim X_{\tau_D}$, so too is $Z(X_{\tau_D})$. Then $H(X_\tau)$ and $H(X_{\tau_D})$ are both bounded above by an integrable random variable, so their expectations are well defined (although possibly not finite), and equal. Then, as above,

$$-\mathbb{E}\left[Z(X_{\tau_D})|\mathcal{F}_t\right] \;\leq\; -Z(X_{t\wedge\tau_D}) \;\leq\; G(X_{t\wedge\tau_D}, t\wedge\tau_D).$$

We can deduce that

$$\mathbb{E}\left[G(X_{\tau_D},\tau_D)\right] \;\leq\; \lim_{n\to\infty} \mathbb{E}\left[G(X_{t\wedge\tau_D}, t\wedge\tau_D)\right] \;=\; G(X_0,0) \;\leq\; \mathbb{E}\left[G(X_\tau,\tau)\right].$$

The remaining steps follow as previously, and it must follow that in fact $\mathbb{E}[F(\tau_D)] \leq \mathbb{E}[F(\tau)]$.

To observe that the result still holds when f is unbounded, we define, for $N = 1, 2, 3, \cdots$,

$$f_N \;=\; f\wedge N \qquad \text{and} \qquad F_N(t) \;=\; \int_0^t f_N(s)\mathrm{d}s.$$

Then we can apply the above argument to F_N to get $\mathbb{E}\big[\,F_N(\tau)\,\big] \;\leq\; \mathbb{E}\big[\,F_N(\tau_D)\,\big]$, let $N\to\infty$ and by the monotone convergence theorem we have $\mathbb{E}\big[\,F(\tau)\,\big] \;\leq\; \mathbb{E}\big[\,F(\tau_D)\,\big]$. $\square$

We now turn to the proofs of our key results, Proposition 4.1.3 and Lemma 4.1.4:

Proof of Proposition 4.1.3. If $t \leq R(x)$, then the left-hand side of (4.1.8) is:

$$\int_0^t f(s)\,\mathrm{d}s - \int_t^{R(x)} M(x,s)\,\mathrm{d}s = F(t) - \int_t^{R(x)} M(x,s)\,\mathrm{d}s$$

and we know $M(x,s) \geq f(s) \geq 0$, so that the inequality holds.

Now consider the case where $R(x) \leq t$. Then the left-hand side of (4.1.8) becomes:

$$\int_{R(x)}^t M(x,s)\,\mathrm{d}s + \int_0^{R(x)} f(s)\,\mathrm{d}s = \int_{R(x)}^t f(s)\,\mathrm{d}s + \int_0^{R(x)} f(s)\,\mathrm{d}s = F(t).$$

□

Note that what we have done in Example 4.1.1 and Example 4.1.2 depends on the smoothness of R on its support, which makes it possible to compute $\partial G/\partial t$ and $\partial^2 G/\partial x^2$, and then the submartingale / martingale result follows by Itô's formula. However, without the smoothness assumption on R in general cases, we will prove Lemma 4.1.4 by the strong Markov property of X [1].

Proof of Lemma 4.1.4. We begin by noting that, since $M(x,0)$ is convex and bounded, and therefore, the Meyer-Itô formula (Protter [2005, Theorem 71, Chapter IV]) gives:

$$Z(X_t) - Z(X_s) = \int_s^t Z'(X_r)\,\mathrm{d}X_r + \frac{1}{2}\int_s^t Z''(X_r)\sigma^2(X_r)\,\mathrm{d}r.$$

It follows from (4.1.9) that the first integral is a martingale. So we get:

$$\mathbb{E}\left[Z(X_t) - Z(X_s)|\mathcal{F}_s\right] \;=\; \int_s^t \mathbb{E}\left[M(X_r,0)|\mathcal{F}_s\right]\,\mathrm{d}r, \quad s \leq t.$$

In addition, since $M(x,t) \geq f(t)$ and $f(t)$ is increasing, for $r, u \geq 0$ by the strong Markov property, writing $\widetilde{X}$ for an independent stochastic process with the same law as X and $\widetilde{\tau}_D$ for the corresponding hitting time of the barrier, and X^x means that the

[1] If we have an additional assumption: $R'(x)$ exists on $\mathrm{int}(supp(R))$, we can show the submartingale / martingale result in a similar manner as in Example 4.1.1 and Example 4.1.2.

process started at x, we have:

$$\begin{aligned}
\mathbb{E}^{(x,r)}\left[f(\tau_D)|\mathcal{F}_{r+u}\right] &= \mathbf{1}_{[\tau_D>r+u]}\mathbb{E}^{(x,r)}\left[f(\tau_D)|\mathcal{F}_{r+u}\right] + \mathbf{1}_{[\tau_D\le r+u]}\mathbb{E}^{(x,r)}\left[f(\tau_D)|\mathcal{F}_{r+u}\right] \\
&\le \mathbf{1}_{[\tau_D>r+u]}\mathbb{E}^{(X_u^x,r+u)}[f(\widetilde{\tau}_D)] + \mathbf{1}_{[\tau_D\le r+u]}f(r+u) \\
&\le M(X_u^x, r+u).
\end{aligned} \tag{4.1.14}$$

In particular, let $r=0$ in (4.1.14), we have $\mathbb{E}^{(x,0)}[f(\tau_D)|\mathcal{F}_u] \le M(X_u^x, u)$. Then for $s, u \in [0,t]$,

$$\begin{aligned}
\mathbb{E}\left[M(X_t,u)|\mathcal{F}_s\right] &= \mathbb{E}^{X_s}\left[M(\widetilde{X}_{t-s},u)\right] \\
&\ge \mathbb{E}^{(X_s,u-(t-s))}\left[f(\widetilde{\tau}_D)\right] \\
&\ge M(X_s, u-(t-s)),
\end{aligned} \tag{4.1.15}$$

when $u \ge t-s$. On the other hand, if $u < t-s$:

$$\begin{aligned}
\mathbb{E}\left[M(X_t,u)|\mathcal{F}_s\right] &= \mathbb{E}\Big[\mathbb{E}^{(X_{t-u},0)}\big[M(\widetilde{X}_u,u)\big]\Big|\mathcal{F}_s\Big] \\
&\ge \mathbb{E}\Big[\mathbb{E}^{(X_{t-u},0)}\big[f(\widetilde{\tau}_D)\big]\Big|\mathcal{F}_s\Big] \\
&\ge \mathbb{E}\big[M(X_{t-u},0)\big|\mathcal{F}_s\big].
\end{aligned} \tag{4.1.16}$$

Then we can write:

$$
\begin{aligned}
\mathbb{E}\big[G(X_t,t)|\mathcal{F}_s\big] \;&=\; \int_0^t \mathbb{E}\left[M(X_t,u)|\mathcal{F}_s\right]\,\mathrm{d}u - \mathbb{E}\left[Z(X_t)|\mathcal{F}_s\right] \\
&= G(X_s,s) + \int_0^t \mathbb{E}\left[M(X_t,u)|\mathcal{F}_s\right]\,\mathrm{d}u \\
&\qquad\qquad - \int_0^s M(X_s,u)\,\mathrm{d}u - \mathbb{E}\left[Z(X_t)-Z(X_s)|\mathcal{F}_s\right] \\
&\geq G(X_s,s) + \int_0^{t-s} \mathbb{E}\left[M(X_{t-u},0)|\mathcal{F}_s\right]\,\mathrm{d}u - \int_0^s M(X_s,u)\,\mathrm{d}u \\
&\qquad - \int_s^t \mathbb{E}\left[M(X_u,0)|\mathcal{F}_s\right]\,\mathrm{d}u + \int_{t-s}^t M(X_s,s-t+u)\,\mathrm{d}u \\
&\geq G(X_s,s) + \int_s^t \mathbb{E}\left[M(X_u,0)|\mathcal{F}_s\right]\,\mathrm{d}u - \int_s^t \mathbb{E}\left[M(X_u,0)|\mathcal{F}_s\right]\,\mathrm{d}u \\
&\qquad\qquad + \int_0^s M(X_s,u)\,\mathrm{d}u - \int_0^s M(X_s,u)\,\mathrm{d}u \\
&\geq G(X_s,s),
\end{aligned}
$$

where we have used (4.1.15) and (4.1.16) in the third line.

On the other hand, as a part of (4.1.14), for $r,u\geq 0$, we have

$$
\begin{cases}
\mathbf{1}_{[\tau_D>r+u]}\mathbb{E}^{(x,r)}[f(\tau_D)|\mathcal{F}_{r+u}] \;=\; \mathbf{1}_{[\tau_D>r+u]}M(X_u^x,r+u); \\
\mathbf{1}_{[\tau_D>u]}\mathbb{E}^{(x,0)}[f(\tau_D)|\mathcal{F}_u] \;=\; \mathbf{1}_{[\tau_D>u]}M(X_u^x,u).
\end{cases}
$$

Thus, on $\{\tau_D>s\}$, we get:

$$
\mathbb{E}[\,M(X_{t\wedge\tau_D},t\wedge\tau_D-u)|\mathcal{F}_s\,] \;=\; \begin{cases} M(X_s,s-u), & u\in(0,\,s); \\ \mathbb{E}[\,M(X_u,0)|\mathcal{F}_s\,], & u\in[s,\,t\wedge\tau_D]. \end{cases} \tag{4.1.17}
$$

Therefore a similar calculation to above gives, for $s \leq \tau_D$:

$$\begin{aligned}
\mathbb{E}\big[G(X_{t\wedge\tau_D}, t\wedge\tau_D)|\mathcal{F}_s\big] &= \mathbb{E}\left[\int_0^{t\wedge\tau_D} M(X_{t\wedge\tau_D}, t\wedge\tau_D - u)\,\mathrm{d}u\middle|\mathcal{F}_s\right] - \mathbb{E}\left[Z(X_{t\wedge\tau_D})|\mathcal{F}_s\right] \\
&= \int_0^s M(X_s, s-u)\,\mathrm{d}u + \mathbb{E}\left[\int_s^{t\wedge\tau_D} M(X_{t\wedge\tau_D}, t\wedge\tau_D - u)\,\mathrm{d}u\middle|\mathcal{F}_s\right] \\
&\qquad - Z(X_s) - \mathbb{E}\left[\int_s^{t\wedge\tau_D} M(X_u, 0)\,\mathrm{d}u\middle|\mathcal{F}_s\right] \\
&= \mathbb{E}\left[\int_s^t \mathbb{E}\left[M(X_{t\wedge\tau_D}, t\wedge\tau_D - u) - M(X_u, 0)\middle|\mathcal{F}_u\right]\mathbf{1}_{\{u\leq\tau_D\}}\,\mathrm{d}u\middle|\mathcal{F}_s\right] + G(X_s, s) \\
&= G(X_s, s),
\end{aligned}$$

where we have used (4.1.17). □

Remark 4.1.6. Note that the fact that our choice of D given in the solution is the domain D which arises as the Root's solution to the Skorokhod embedding problem is only used in Theorem 4.1.5 to enforce the lower bound. In fact, we could choose any barrier D' as our domain, and this would result in a lower bound, and corresponding functions G and H. The choice of Root's barrier gives the optimal lower bound, in that we can attain equality for some stopping time. In this context, it is worth recalling the lower bounds given by Carr and Lee [2010, Proposition 3.1] – here a lower bound is given which essentially corresponds to choosing the domain with $R(x) = Q$ for a constant Q. The arguments given above show that similar constructions are available for any choice of R, and the optimal choice corresponds to the Root's construction.

Remark 4.1.7. Although the preceding section is written for a diffusion on $\mathbb{R}$, it is not hard to check that the case where $\sigma : x \mapsto x$ can also be included without many changes. In this setting, we need to restrict the space variable to the half space $(0, \infty)$ (so we assume that $\tau_D < \infty$, a.s.), and consider a starting distribution which is also supported only on $(0, \infty)$, and with a corresponding change to (4.1.4).

4.2 Optimality of Rost's Solution

For a given distribution, Theorem 3.4.2 says that Rost's construction is optimal in the sense of "maximal variance", and moreover, a slight generalization of this result gives the solution to the following optimization problem:

OPT*(σ, ν, μ): Suppose (2.1.7) holds and X satisfies (2.1.1). Find a stopping time $\tau \in \mathcal{T}(\sigma, \nu, \mu)$ such that, for a given increasing convex function F with $F(0) = 0$,

$$\mathbb{E}^{\nu}\big[F(\tau)\big] \;=\; \sup_{\rho \in \mathcal{T}(\sigma,\nu,\mu)} \mathbb{E}^{\nu}\big[F(\rho)\big].$$

As in Section 4.1, our aim in this section is to give a proof of a similar result to Theorem 4.1.5, and to give a path-wise inequality which encodes the optimality.

We suppose firstly that $D = \{(x,t) : t > R(x)\}$ is the solution to **SEP*(σ, ν, μ)** where R is an upper semi-continuous function. Define the function $M = \mathbb{E}^{(x,t)}[f(\tau_D)]$, same as (4.1.3), but here we let f be the left derivative of F, and hence, $M(x,t) = f(t)$ whenever $0 \leq t < R(x)$.

Again, we begin our exploration with an example.

Example 4.2.1. Suppose we have a reversed barrier $D := \{(x,t) : t > R(x)\}$ with the boundary function

$$R(x) \;=\; 2(x^2 - 1) \vee 0.$$

Given a standard Brownian motion W, and Rost's stopping time

$$\tau_D \;=\; \inf\{t > 0 : \, t \leq R(W_t)\},$$

define $\mu := \mathcal{L}(W_{\tau_D})$. Now we consider the optimal stopping problem **OPT*(μ)** with $F(t) = \frac{1}{2}t^2$.

For $(x,t) \in \mathbb{R} \times \mathbb{R}_+$, define $M(x,t) = \mathbb{E}^{(x,t)}[\tau_D]$. Then as in Example 4.1.1, we have

$$\begin{cases} M(x,t) \;=\; 2t - 2\left(x^2 - 1\right), & \text{for } t > 2(x^2 - 1); \\ M(x,t) \;=\; t, & \text{otherwise.} \end{cases}$$

It is easy to verify that $(\partial/\partial t + \frac{1}{2}\partial^2/\partial x^2)M = 0$ on D.

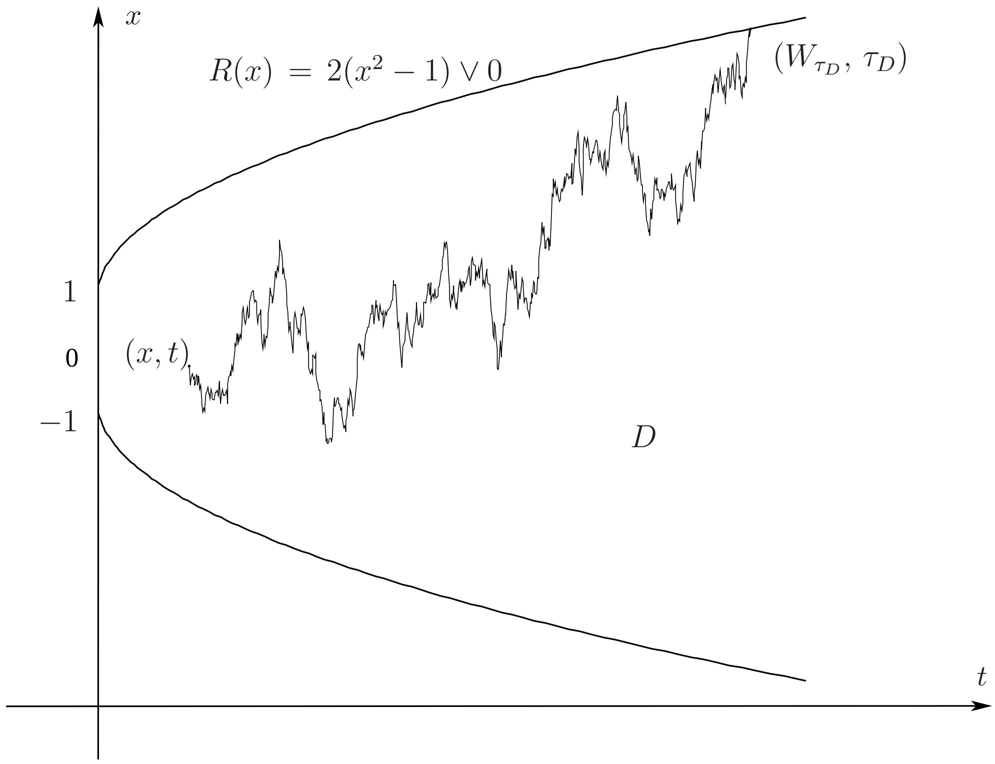


Figure 4-3: *Example 4.2.1*

Now given a fixed time level $T > 0$, we define, for $t > 0$

$$
\begin{aligned}
Z_T(x) &= 2\int_0^x \int_0^y M(z,T)\mathrm{d}z\mathrm{d}y, \\
G_T(x,t) &= F(T) - \int_t^T M(x,s)\mathrm{d}s - Z_T(x) \\
&= \begin{cases} 2(x^2-1)(T-t) + t^2 - \dfrac{T^2}{2} - Z_T(x), & \text{if } \ t > R(x); \\ R(x)T + \dfrac{t^2}{2} - \dfrac{R^2(x)}{2} - \dfrac{T^2}{2} - Z_T(x), & \text{if } \ t \le R(x) < T, \\ F(t) - Z_T(x), & \text{if } \ T \le R(x). \end{cases}
\end{aligned}
$$

Hence, if $R(x) < t \le T$, we have

$$
-\frac{1}{2}\frac{\partial^2 G_T}{\partial x^2} = M(x,T) - 2(T-t) = 2t - 2(x^2-1) = \frac{\partial G_T}{\partial t}.
$$

If $t \le T \le R(x)$, then

$$\frac{\partial G_T}{\partial t} + \frac{1}{2}\frac{\partial^2 G_T}{\partial x^2} \;=\; t - M(x,T) \;=\; t - T \;\le 0,$$

and for $0 < t \le R(x) < T$, (we must have $|x| > 1$ in this case, and hence, $R(x) = 2(x^2-1)$)

$$\begin{aligned}\frac{\partial G_T}{\partial t} + \frac{1}{2}\frac{\partial^2 G_T}{\partial x^2} \;&=\; t + \frac{R''(x)T - R''(x)R(x) - (R'(x))^2}{2} - \big(2T - R(x)\big) \\ &=\; t - R(x) - \frac{(R'(x))^2}{2} \;\le\; 0.\end{aligned}$$

Therefore, before the time T, $G_T(X_t, t)$ is a supermartingale and $G_T(X_{t\wedge\tau_D}, t\wedge\tau_D)$ is a martingale.

This example provides us with the idea to construct a (super)martingale using a reversed barrier. In fact, unlike Root's case where we integrate $M(x,t)$ from the left point, 0, to t, we take the integration from t to some fixed $T > t$. That is because, if $(x,t) \in D$, $M(x,s)$ is coparabolic for all $s \ge t$ when D is Rost, and is coparabolic for all $s \le t$ when D is Root.

Now given a fixed time level $T > 0$, we define

$$Z_T(x) \;=\; 2\int_0^x \int_0^y \frac{M(z,T)}{\sigma^2(z)}\mathrm{d}z\mathrm{d}y,$$

and in particular, $Z''(x) = 2M(x,T)/\sigma^2(x)$. Define also

$$\begin{aligned}G_T(x,t) \;&=\; F(T) - \int_t^T M(x,s)\mathrm{d}s - Z_T(x) \\ H_T(x) \;&=\; \int_{R(x)}^T \Big[M(x,s) - f(s)\Big]\mathrm{d}s + Z_T(x) \\ &=\; \int_{R(x)\wedge T}^T \Big[M(x,s) - f(s)\Big]\mathrm{d}s + Z_T(x).\end{aligned}$$

Then we claim that

Proposition 4.2.2. *For all $(x, t, T) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$, we have,*

$$\begin{cases} G_T(x,t) \ + \ H_T(x) \ \geq \ F(t), & \text{if } \ t > R(x)\,; \\ G_T(x,t) \ + \ H_T(x) \ = \ F(t), & \text{if } \ t \leq R(x)\,. \end{cases} \tag{4.2.1}$$

And also

Lemma 4.2.3. *Suppose that (4.1.12) remains true here for f. For given $T > 0$, if*

$$\Big(Q(X_t)\,;\, 0 \leq t \leq T \Big) \ \text{ is a uniformly integrable family,} \tag{4.2.2}$$

where we denote $Q : x \mapsto \int_0^x \int_0^y 2\sigma^{-2}(z)\mathrm{d}z\mathrm{d}y$. Then the process

$$\Big(G_T(X_{t\wedge\tau_D}, t \wedge \tau_D) \ ;\, 0 \leq t \leq T \Big) \ \text{ is a martingale,} \tag{4.2.3}$$

and

$$\Big(G_T(X_t, t);\, 0 \leq t \leq T \Big) \ \text{ is a supermartingale.} \tag{4.2.4}$$

Then the main result of this section follows.

Theorem 4.2.4. *Suppose τ_D solves* $\mathbf{SEP^*(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})}$, *moreover, we assume that for all $T > 0$, (4.2.2) holds. Then the corresponding Rost's stopping time τ_D solves* $\mathbf{OPT^*(\boldsymbol{\sigma}, \boldsymbol{\nu}, \boldsymbol{\mu})}$.

Proof. The first case we consider is the case where $\mathbb{E}[\tau_D] = \infty$. Since $F(t) \geq \alpha + \beta t$ for some constants $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}_+$, we must have $\mathbb{E}[F(\tau_D)] = \infty$. The result is trivial. So we always assume $\mathbb{E}[\tau_D] < \infty$.

Under the assumption $\mathbb{E}[\tau_D] < \infty$, consider $Q(\cdot)$ as defined after (4.2.2), we have

$$\mathbb{E}[Q(X_{\tau_D})] \ = \ \mathbb{E}[Q(X_0)] + \mathbb{E}[\tau_D] \ < \ \infty.$$

Therefore, for all $\tau \in \mathcal{T}(\sigma, \nu, \mu)$, since Q is convex, we have

$$\mathbb{E}[t \wedge \tau] \ = \ \mathbb{E}[Q_{t\wedge\tau}] - \mathbb{E}[Q(X_0)] \leq \mathbb{E}[Q(X_\tau)] - \mathbb{E}[Q(X_0)] \ = \ \mathbb{E}[\tau_D] < \infty,$$

which implies $\mathbb{E}[\tau] < \infty$ by the monotone convergence theorem. In the remainder of this proof, we always assume $\mathbb{E}[\tau] < \infty$ and $\mathbb{E}[\tau_D] < \infty$.

As in the proof of Theorem 4.1.5, we firstly assume f is bounded, i.e. f satisfies (4.1.12). The same arguments as in the proof of Theorem 4.1.5 give us that

$$\mathbb{E}\left[Z_T(X_{t\wedge\tau})\right] \;\le\; \mathbb{E}\left[Z_T(X_0)\right] + C\mathbb{E}[\tau] \;\le\; C\left(\mathbb{E}\left[Q(X_0)\right] + \mathbb{E}[\tau]\right) \;<\; \infty, \tag{4.2.5}$$

and the same argument implies $\mathbb{E}[Z_T(X_\tau)] < \infty$. We then have that $\mathbb{E}[Z_T(X_\tau)|\mathcal{F}_t]$ is a uniformly integrable martingale, and by convexity, $Z_T(X_{t\wedge\tau}) \le \mathbb{E}[Z_T(X_\tau)|\mathcal{F}_t]$. Therefore,

$$-C|T-(t\wedge\tau)| \;\le\; F(T) - G_T(X_{t\wedge\tau}, t\wedge\tau) \;\le\; C|T-(t\wedge\tau)| + \mathbb{E}[Z_T(X_\tau)|\mathcal{F}_t].$$

It follows that $\mathbb{E}[G_T(X_{t\wedge\tau}, t\wedge\tau)] \to \mathbb{E}[G_T(X_\tau,\tau)]$ as $t\to\infty$. On the other hand,

$$\mathbb{E}\left[H_T(X_\tau)\right] \;=\; \mathbb{E}\left[\int_{T\wedge R(X_\tau)}^{T}\left[M(X_\tau,s) - f(s)\right]\mathrm{d}s\right] + \mathbb{E}\left[Z_T(X_\tau)\right] \;<\; \infty.$$

The same arguments hold when τ is replaced by τ_D, and then we have

$$\mathbb{E}\left[H_T(X_\tau)\right] \;=\; \mathbb{E}\left[H_T(X_{\tau_D})\right] \qquad \text{and} \qquad \mathbb{E}\left[Z_T(X_\tau)\right] \;=\; \mathbb{E}\left[Z_T(X_{\tau_D})\right].$$

In addition, by Lemma 4.2.3, we have,

$$\mathbb{E}\left[G_T(X_{T\wedge\tau_D}, T\wedge\tau_D)\right] \ge \mathbb{E}\left[G_T(X_{T\wedge\tau}, T\wedge\tau)\right].$$

Combining the results above with (4.2.1), we then have,

$$
\begin{aligned}
\mathbb{E}\big[\,F(\tau)\,\big] \;&\leq\; \mathbb{E}\big[\,G_T(X_\tau,\tau)+H_T(X_\tau)\,\big]\\
&=\; \mathbb{E}\big[\,G_T(X_{T\wedge\tau},T\wedge\tau)+H_T(X_\tau)\,\big]+\mathbb{E}\big[\,G_T(X_\tau,\tau)-G_T(X_{T\wedge\tau},T\wedge\tau)\,\big]\\
&\leq\; \mathbb{E}\big[\,G_T(X_{T\wedge\tau_D},T\wedge\tau_D)+H_T(X_{\tau_D})\,\big]+\mathbb{E}\big[\,G_T(X_\tau,\tau)-G_T(X_{T\wedge\tau},T\wedge\tau)\,\big]\\
&=\; \mathbb{E}\big[\,G_T(X_{\tau_D},\tau_D)+H_T(X_{\tau_D})\,\big]+\mathbb{E}\big[\,G_T(X_\tau,\tau)-G_T(X_{T\wedge\tau},T\wedge\tau)\,\big]\\
&\qquad\qquad -\mathbb{E}\big[\,G_T(X_{\tau_D},\tau_D)-G_T(X_{T\wedge\tau_D},T\wedge\tau_D)\,\big]\\
&=\; \mathbb{E}\big[\,F(\tau_D)\,\big]+\mathbb{E}\left[\int_{T\wedge\tau}^{T}M(X_{T\wedge\tau},s)\mathrm{d}s-\int_{\tau}^{T}M(X_\tau,s)\mathrm{d}s+Z_T(X_{T\wedge\tau})\right]\\
&\qquad -\mathbb{E}\left[\int_{T\wedge\tau_D}^{T}M(X_{T\wedge\tau_D},s)\mathrm{d}s-\int_{\tau_D}^{T}M(X_{\tau_D},s)\mathrm{d}s+Z_T(X_{T\wedge\tau_D})\right]\\
&=\; \mathbb{E}\big[\,F(\tau_D)\,\big]+\mathbb{E}\left[\mathbf{1}_{[\tau>T]}\int_T^{\tau}M(X_\tau,s)\mathrm{d}s\right]-\mathbb{E}\left[\mathbf{1}_{[\tau_D>T]}\int_T^{\tau_D}M(X_{\tau_D},s)\mathrm{d}s\right]\\
&\qquad +\mathbb{E}\Big[\,Z_T(X_{T\wedge\tau})-Z_T(X_{T\wedge\tau_D})\,\Big].
\end{aligned}
\tag{4.2.6}
$$

Since $f\leq C$, we have

$$
\begin{aligned}
0 \;\leq\; \mathbb{E}\left[\mathbf{1}_{[\tau>T]}\int_T^{\tau}M(X_\tau,s)\mathrm{d}s\right] \;&\leq\; C\,\mathbb{E}\big[\,\mathbf{1}_{[\tau>T]}(\tau-T)\,\big]\\
&=\; C\,\mathbb{E}\big[\,\tau-T\wedge\tau\,\big] \;\longrightarrow\; 0, \qquad \text{as} \;\; T\to\infty.
\end{aligned}
$$

Similarly,

$$
\lim_{T\to\infty}\mathbb{E}\left[\mathbf{1}_{[\tau_D>T]}\int_T^{\tau_D}M(X_{\tau_D},s)\mathrm{d}s\right] \;=\; 0.
$$

Consider $Q(\cdot)$ as defined after (4.2.2), we have,

$$
\mathbb{E}\left[Q(X_{T\wedge\tau})\right] \;=\; \mathbb{E}\left[Q(X_0)\right]+\mathbb{E}[T\wedge\tau] \;\leq\; \mathbb{E}\left[Q(X_0)\right]+\mathbb{E}[\tau], \tag{4.2.7}
$$

and then, applying Fatou's Lemma, we deduce that $Q(X_\tau)$ is integrable. By the convexity of $Q(\cdot)$, $Q(X_{T\wedge\tau})\leq\mathbb{E}[Q(X_\tau)|\mathcal{F}_T]$, hence, $Q(X_{T\wedge\tau})\to Q(X_\tau)$ in L^1. Noting that $Z_T(X_{T\wedge\tau})\leq CQ(X_{T\wedge\tau})$ and $Z_T(X_{T\wedge\tau})\to CQ(X_\tau)$ a.s. as $T\to\infty$, we have

$$
\lim_{T\to\infty}\mathbb{E}\left[Z_T(X_{T\wedge\tau})\right] \;=\; C\mathbb{E}\left[Q(X_\tau)\right] \;<\infty.
$$

The same arguments hold when τ is replaced by τ_D, and moreover, $\mathbb{E}[Q(X_\tau)] = \mathbb{E}[Q(X_{\tau_D})]$. Now, let T go to infinity in (4.2.6), and we have

$$\mathbb{E}\big[\,F(\tau)\,\big] \;\le\; \mathbb{E}\big[\,F(\tau_D)\,\big].$$

Finally we get rid of the assumption (4.1.12) by the same arguments as in Theorem 4.1.5. □

Now we turn to the proofs of Proposition 4.2.2 and Lemma 4.2.3.

Proof of Proposition 4.2.2. We have

$$G_T(x,t) \;+\; H_T(x) \;=\; \int_{R(x)}^{t} M(x,s)\mathrm{d}s \;+\; F(R(x)).$$

If $(x,t) \in D$, i.e. $t > R(x)$,

$$\begin{aligned} G_T(x,t) \;+\; H_T(x) \;&=\; \int_{R(x)}^{t} M(x,s)\mathrm{d}s \;+\; F(R(x)) \\ &\ge\; \int_{R(x)}^{t} f(s)\mathrm{d}s \;+\; F(R(x)) \;=\; F(t). \end{aligned}$$

If $(x,t) \notin D$, i.e. $t \le R(x)$,

$$\begin{aligned} G_T(x,t) \;+\; H_T(x) \;&=\; -\int_{t}^{R(x)} M(x,s)\mathrm{d}s \;+\; F(R(x)) \\ &=\; -\int_{t}^{R(x)} f(s)\mathrm{d}s \;+\; F(R(x)) \;=\; F(t). \end{aligned}$$

□

Proof of Lemma 4.2.3. Similarly to the proof of Lemma 4.1.4, for $s \le t \le T$, by (4.2.2), the Meyer-Itô formula gives,

$$Z_T(X_t) - Z_T(X_s) \;=\; \int_s^t Z_T'(X_u)\mathrm{d}X_u + \int_s^t M(X_u,T)\mathrm{d}u\,.$$

By (4.2.2) and the fact f is bounded, it is easy to see that the family $(Z_T(X_t); 0 \le t \le T)$ is uniformly integrable. By the Doob-Meyer decomposition theorem (e.g. Karatzas and Shreve [1991, Theorem 4.10, Chapter 1]), the first term on the right-hand side is a uniformly integrable martingale,

$$\mathbb{E}\big[Z_T(X_t) - Z_T(X_s)\,\big|\,\mathcal{F}_s\big] \;=\; \int_s^t \mathbb{E}\big[\,M(X_u,T)\,\big|\,\mathcal{F}_s\big]\mathrm{d}u\,.$$

Then we have,

$$\begin{aligned}
&G_T(X_s,s) - \mathbb{E}\left[\,G_T(X_t,t)\,\middle|\,\mathcal{F}_s\right] \\
&\quad= \int_t^T \mathbb{E}\big[\,M(X_t,u)\,\big|\,\mathcal{F}_s\big]\mathrm{d}u + \int_s^t \mathbb{E}\big[\,M(X_u,T)\,\big|\,\mathcal{F}_s\big]\mathrm{d}u - \int_s^T M(X_s,u)\mathrm{d}u \\
&\quad= \int_t^T \mathbb{E}\big[\,M(X_t,u)\,\big|\,\mathcal{F}_s\big]\mathrm{d}u - \int_s^{T-t+s} M(X_s,u)\mathrm{d}u \\
&\qquad\qquad + \int_s^t \mathbb{E}\big[\,M(X_u,T)\,\big|\,\mathcal{F}_s\big]\mathrm{d}u - \int_{T-t+s}^T M(X_s,u)\mathrm{d}u \\
&\quad= \int_t^T \Big\{\mathbb{E}\big[\,M(X_t,u)\,\big|\,\mathcal{F}_s\big] - M(X_s,u-(t-s))\Big\}\mathrm{d}u \\
&\qquad\qquad + \int_s^t \Big\{\mathbb{E}\big[\,M(X_u,T)\,\big|\,\mathcal{F}_s\big] - M(X_s,T-(t-u))\Big\}\mathrm{d}u.
\end{aligned}$$

Now, for $u \in (t,T)$, define $\widetilde{X},\ \widetilde{\tau}_D,\ X^x$ as in the proof of Lemma 4.1.4,

$$\begin{aligned}
\mathbb{E}^{(x,u-(t-s))}\big[\,f(\tau_D)\,|\,\mathcal{F}_u\big] \;&\leq \mathbf{1}_{[\tau_D\leq u]}f(u) + \mathbf{1}_{[\tau_D>u]}\mathbb{E}^{(x,u-(t-s))}\big[\,f(\tau_D)\,|\,\mathcal{F}_u\big] \\
&= \mathbf{1}_{[\tau_D\leq u]}f(u) + \mathbf{1}_{[\tau_D>u]}\mathbb{E}^{(X^x_{t-s},u)}\big[\,f(\widetilde{\tau}_D)\,\big] \;\leq\; M(X^x_{t-s},u),
\end{aligned} \tag{4.2.8}$$

and hence,

$$\begin{aligned}
\mathbb{E}\big[\,M(X_t,u)\,\big|\,\mathcal{F}_s\big] \;&=\; \mathbb{E}^{X_s}M(\widetilde{X}_{t-s},u) \\
&\geq\; \mathbb{E}^{(X_s,u-(t-s))}\big[\,f(\widetilde{\tau}_D)\,\big] \;=\; M(X_s,u-(t-s)).
\end{aligned} \tag{4.2.9}$$

For $u \in (s,t)$, replacing u by T and t by u in (4.2.8) gives that

$$\mathbb{E}^{(x,T-(u-s))}\big[\,f(\tau_D)\,|\,\mathcal{F}_T\big] \;\leq\; M(X^x_{u-s},T),$$

and hence,

$$\mathbb{E}\big[\,M(X_u,T)\,\big|\,\mathcal{F}_s\big] \;\geq\; M(X_s,T-(u-s)). \tag{4.2.10}$$

It follows that

$$\begin{aligned}
&\int_s^t \Big\{ \mathbb{E}\big[\, M(X_u, T) \,\big|\, \mathcal{F}_s \big] - M(X_s, T-(t-u)) \Big\} \mathrm{d}u \\
=\ &\int_s^t \Big\{ \mathbb{E}\big[\, M(X_u, T) \,\big|\, \mathcal{F}_s \big] - M(X_s, T-(u-s)) \Big\} \mathrm{d}u \ \geq 0.
\end{aligned}$$

Therefore,

$$G_T(X_s, s) - \mathbb{E}\,[\, G_T(X_t, t) | \mathcal{F}_s] \ \geq \ 0,$$

which implies (4.2.4).

On the other hand, as a part of (4.2.8),

$$\mathbf{1}_{[T<\tau_D]} \mathbb{E}^{(x, u-(t-s))} \big[\, f(\tau_D) \,|\, \mathcal{F}_u \big] \ = \ \mathbf{1}_{[\tau_D > u]} \mathbb{E}^{(X^x_{t-s}, u)} \big[\, f(\widetilde{\tau}_D) \,\big],$$

and hence, it is easy to see that, on $\{T < \tau_D\}$, the equalities hold in the inequalities (4.2.9) and (4.2.10). Thus, (4.2.3) follows. $\square$

For bounded f, although the path-wise inequality in this section $G_T(X_t, t) + H_T(X_t) \geq F(t)$ holds for all $T, t > 0$, $G_T(X_t, t)$ is a supermartingale only on $[0, T]$. Now our question is that: can we find a *global* path-wise inequality $G^*_t + H^*(X_t) \geq F(t)$, and G^*_t is a supermartingale on $[0, \infty]$ and a martingale on $[0, \tau_D]$? We will see a special case where we can find such G^* and H^* in the following discussion.

To show that, we replace (4.1.12) by a stronger assumption: there exists some $\alpha > 1$, such that

$$\text{for } t \text{ sufficiently large,} \ \ C \ \geq \ f(t) \ \geq \ C - O(t^{-\alpha}). \tag{4.2.11}$$

Under this assumption, it is easy to check there exists a $J(x,t)$ such that

$$J(x,t) \ = \ \lim_{T\to\infty} \int_t^T [M(x,s) - f(s)]\, \mathrm{d}s, \tag{4.2.12}$$

then we define

$$\begin{cases} \widetilde{G}(x,t) \ = \ \lim\limits_{T\to\infty} G_T(x,t) \ = \ F(t) - J(x,t) - CQ(x); \\ \widetilde{H}(x) \ = \ \lim\limits_{T\to\infty} H_T(x) \ = \ J(x, R(x)) + CQ(x). \end{cases} \tag{4.2.13}$$

Then, letting $T \to \infty$ in (4.2.1),

$$\begin{cases} \widetilde{G}(x,t) \, + \, \widetilde{H}(x) \; > \; F(t), & \text{if } \; t > R(x)\,; \\ \widetilde{G}(x,t) \, + \, \widetilde{H}(x) \; = \; F(t), & \text{if } \; t \leq R(x)\,. \end{cases} \tag{4.2.14}$$

By the monotone convergence theorem, for all $t > 0$, $\mathbb{E}[\int_t^T [M(X_t, s) - f(s)]\mathrm{d}s] \to \mathbb{E}[J(X_t, t)]$ as $T \to \infty$, and then by Sheffé's Lemma, $\int_t^T [M(X_t, s) - f(s)]\mathrm{d}s \to J(X_t, t)$ in L^1. On the other hand, by (4.2.7), $Z_T(X_t) \to CQ(X_t)$ in L^1, and hence,

$$G_T(X_t, t) \;\xrightarrow{L^1}\; \widetilde{G}(X_t, t) \;\text{ and }\; H_T(X_t) \;\xrightarrow{L^1}\; \widetilde{H}(X_t).$$

It follows that $\left(\widetilde{G}(X_t, t);\, t \geq 0\right)$ is a supermartingale and $\left(\widetilde{G}(X_{t\wedge\tau_D}, t \wedge \tau_D);\, t \geq 0\right)$ is a martingale (since the conditional expectation, as an operator, is continuous in L^p for $p \geq 1$). We then can show that (if τ, τ_D is integrable),

$$\begin{aligned} \mathbb{E}\left[F(\tau)\right] \;&\leq\; \mathbb{E}\left[\widetilde{G}(X_\tau, \tau) + \widetilde{H}(X_\tau)\right] \\ &\leq\; \mathbb{E}\left[\widetilde{G}(X_{\tau_D}, \tau_D) + \widetilde{H}(X_{\tau_D})\right] \;=\; \mathbb{E}\left[F(\tau_D)\right]. \end{aligned}$$

An example where (4.2.14) holds is the payoff function stated at the beginning of this chapter: $F(t) = (t - K)_+$. We see that for $t > K$, the left derivative $f(t) = 1$, and hence

$$J(x,t) \;=\; \int_t^K [M(x,s) - f(s)]\mathrm{d}s,$$

we then repeat all arguments above to obtain the path-wise inequality and the optimality result.

Another example we note here is the case where (4.2.14) does not hold. It is simply Example 4.2.1, where $F(t) = t^2/2$. According to the computations above we have,

$$\begin{aligned} &\int_{t\vee 2(x^2-1)}^T [M(x,s) - f(s)]\mathrm{d}s \\ &\qquad = \left[\frac{T^2}{2} - 2(x^2-1)T\right] - \left[\frac{\left(t \vee 2(x^2-1)\right)^2}{2} - 2(x^2-1)\left(t \vee 2(x^2-1)\right)\right]. \end{aligned}$$

Clearly this goes to infinity as $T \to \infty$.

4.3 Financial Application

We now turn to our motivating financial problem: consider an asset price defined on a complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, with:

$$\frac{\mathrm{d}S_t}{S_t} = r_t \mathrm{d}t + \sigma_t \mathrm{d}W_t \tag{4.3.1}$$

under some probability measure $\mathbb{Q} \sim \mathbb{P}$, where $\mathbb{P}$ is the objective probability measure, and W_t a $\mathbb{Q}$-Brownian motion. In addition, we suppose r_t is the risk-free rate which we require to be known, but which need not be constant. In particular, let r_t, σ_t be locally bounded, predictable processes so that the integral is well defined in (4.3.1), and so S_t is an Itô process. We suppose that the process σ_t is not known (or more specifically, we aim to produce conclusions which hold for all σ_t in the class described). Specifically, we shall suppose:

Assumption 4.3.1. *The asset price process, under some probability measure $\mathbb{Q} \sim \mathbb{P}$ is the solution to the stochastic differential equation (4.3.1), where r_t and σ_t are locally bounded, predictable processes.*

In addition, we need to make the following assumptions regarding the set of call options which are initially traded:

Assumption 4.3.2. *We suppose that call options with maturity T, and at all strikes $\{K : K \geq 0\}$ are traded at time 0, and the prices, $C(K)$, are assumed to be known, In addition, we suppose call-put parity holds, so that the price of a put option with strike K is*

$$P(K) = K \cdot \exp\left\{-\int_0^T r_s \mathrm{d}s\right\} - S_0 + C(K).$$

We make the additional assumption that $C(K)$ is continuous, decreasing and convex function, with $C(0) = S_0$, $C'_+(0) = -\exp\{-\int_0^T r_s \mathrm{d}s\}$ and $C(K) \to 0$ as $K \to \infty$.

Many of these notions can be motivated by arbitrage concerns (see e.g. Cox and Obłój [2011b]). That there are plausible situations in which these assumptions do not hold can be seen by considering models with bubbles (e.g. Cox and Hobson [2005]), in which call-put parity fails, and $C(K) \not\to 0$ as $K \to \infty$. Let us define

$$B_t = \exp\left\{\int_0^t r_s \mathrm{d}s\right\},$$

and make the assumptions above. Following the perspective that the prices correspond to expectations under $\mathbb{Q}$, the implied law of $B_T^{-1}S_T$ (which we will denote μ) can be recovered by the Breeden-Litzenberger [1978] formula:

$$\mu((K,\infty)) \;=\; \mathbb{Q}^*\big(B_T^{-1}S_T \in (K,\infty)\big) \;=\; -2B_T C'_+(B_T K). \tag{4.3.2}$$

Here we have used $\mathbb{Q}^*$ to emphasize the fact that this is only an *implied* probability, and not necessarily the distribution under the actual measure $\mathbb{Q}$. From (4.3.2) we can deduce that $\mathrm{U}\mu(x) = S_0 - 2C(B_T x) - x$, giving an affine mapping between the function $\mathrm{U}\mu(x)$ and the call prices. We do not impose the condition that the law of $B_T^{-1}S_T$ under $\mathbb{Q}$ is μ, we merely note that this is the law implied by the traded options. We also do not assume anything about the price paths of the call options: our only assumptions are their initial prices, and that they return the usual payoff at maturity. It can now also be seen that the assumption that $C'_+(0) = -B_T^{-1}$ is equivalent to assuming that there is no atom at 0 – i.e. μ is supported on $(0,\infty)$. Finally, note also that this follows from the assumption that μ is an integrable measure with mean S_0.

Our goal is to now to use the knowledge of the call prices to find a lower bound on the price of an option which has payoff

$$F\left(\int_0^T \sigma_t^2 \mathrm{d}t\right) \;=\; F\left(\langle \ln S\rangle_T\right).$$

Consider the discounted stock price:

$$X_t \;=\; S_t \cdot \exp\left\{-\int_0^t r_s \mathrm{d}s\right\} \;=\; B_t^{-1}S_t.$$

Under Assumption 4.3.1, X_t satisfies the stochastic differential equation:

$$\mathrm{d}X_t \;=\; \sigma_t X_t \mathrm{d}W_t.$$

Defining a time change

$$\tau_t \;=\; \int_0^t \sigma_s^2 \mathrm{d}s,$$

and writing A_t for its right-continuous inverse so that $\tau_{A_t} = t$, we note that $\widetilde{W}_t = \int_0^{A_t} \sigma_s \mathrm{d}W_s$ is a Brownian motion with respect to $\mathcal{F}_{A_t}$, and if we set $\widetilde{X}_t = X_{A_t}$, we have:

$$\mathrm{d}\widetilde{X}_t \;=\; \widetilde{X}_t \mathrm{d}\widetilde{W}_t.$$

In particular, $\widetilde{X}$ is now of a form where we may apply our earlier results, using the target distribution arising from (4.3.2), and noting also that $\widetilde{X}_0 = S_0$ and $\widetilde{X}_{\tau_T} = X_T = B_T^{-1} S_T$.

We now define τ_D as the Root's embedding of μ for the diffusion $\widetilde{X}$, and define functions as in Section 4.1, so that $f(t) = F'_+(t)$ and (4.1.3) – (4.1.7) hold. Our aim is to use (4.1.8), which now reads:

$$G(X_{A_t}, t) + H(X_{A_t}) \;=\; G(\widetilde{X}_t, t) + H(\widetilde{X}_t) \;\leq\; F(t) \;=\; F\left(\int_0^{A_t} \sigma_s^2 \mathrm{d}s\right), \qquad (4.3.3)$$

to construct a sub-replicating portfolio. We shall first show that we can construct a trading strategy that sub-replicates the $G(\widetilde{X}_t, t)$ portion of the portfolio. Then we argue that we are able, using a portfolio of calls, puts, cash and the underlying, to replicate the payoff $H(X_T)$.

Since $\left(G(\widetilde{X}_t, t);\, t \geq 0\right)$ is a submartingale, we do not expect to be able to replicate this in a completely self-financing manner. However, by the Doob-Meyer decomposition theorem, and the martingale representation theorem, we can certainly find some process $\left(\widetilde{\phi}_t;\, t \geq 0\right)$ such that:

$$G(\widetilde{X}_t, t) \;\geq\; G(\widetilde{X}_0, 0) + \int_0^t \widetilde{\phi}_s \mathrm{d}\widetilde{X}_s$$

and such that there is equality at $t = \tau_D$. Moreover, since $\left(G(\widetilde{X}_{t\wedge\tau_D}, t\wedge\tau_D);\, t \geq 0\right)$ is a martingale, and G is of $\mathbb{C}^{2,1}$ class in D, we have:

$$G(\widetilde{X}_{t\wedge\tau_D}, t\wedge\tau_D) \;=\; G(\widetilde{X}_0, 0) + \int_0^{t\wedge\tau_D} \frac{\partial G}{\partial x}(\widetilde{X}_{s\wedge\tau_D}, s\wedge\tau_D)\mathrm{d}\widetilde{X}_s.$$

More generally, we would not expect $\partial G/\partial x$ to exist everywhere in $D^{\complement}$, however, if for example left and right derivatives exist, then we could choose

$$\widetilde{\phi}_t \;\in\; \left[\frac{\partial G}{\partial x}(x-, t),\, \frac{\partial G}{\partial x}(x+, t)\right]$$

as our holding of the risky asset (or alternatively, but less explicitly, take

$$\widetilde{\phi}_t \;=\; \frac{\partial}{\partial x}\left(\mathbb{E}^{(x,t)}\left[G(\widetilde{X}_{t+\delta}, t_0 + \delta)\right]\right),$$

for $t \in [t_0, t_0 + \delta]$).

It follows then that we can identify a process $\left(\widetilde{\phi}_t;\, t \geq 0\right)$ with

$$G(\widetilde{X}_{\tau_t}, \tau_t) \;\geq\; G(\widetilde{X}_0, 0) + \int_0^{\tau_t} \widetilde{\phi}_s \mathrm{d}\widetilde{X}_s \;=\; G(X_0, 0) + \int_0^t \widetilde{\phi}_{\tau_s} \mathrm{d}X_s,$$

where we have used e.g. Revuz and Yor [1999, Proposition V.1.4]. Finally, writing $\phi_t = \widetilde{\phi}_{\tau_t}$, we have:

$$G(X_t, \tau_t) \;\geq\; G(X_0, 0) + \int_0^t \phi_s \mathrm{d}X_s \;=\; G(X_0, 0) + \int_0^t \phi_s \mathrm{d}\left(B_s^{-1} S_s\right).$$

If we consider the self-financing portfolio which consists of holding $\phi_s B_T^{-1}$ units of the risky asset, and an initial investment of $G(X_0, 0) B_T^{-1} - \phi_0 S_0 B_T^{-1}$ in the risk-free asset, this has value V_t at time t, where

$$\mathrm{d}\left(B_t^{-1} V_t\right) \;=\; B_T^{-1} \phi_t \,\mathrm{d}\left(B_t^{-1} S_t\right) \quad \text{and} \quad V_0 \;=\; G(X_0, 0) B_T^{-1},$$

and therefore

$$V_T \;=\; B_T \left(V_0 B_0^{-1} + \int_0^T B_T^{-1} \phi_s \,\mathrm{d}\left(B_s^{-1} S_s\right) \right) \;=\; G(X_0, 0) + \int_0^T \phi_s \mathrm{d}X_s.$$

We now turn to the $H(X_T)$ component in (4.3.3). If $H(x)$ can be written as the difference of two convex functions (so in particular, $H''(\mathrm{d}K)$ is a well defined signed measure) we can write:

$$\begin{aligned} H(x) \;=\; & H(S_0) + H'_+(S_0)(x - S_0) + \int_{(S_0, \infty)} (x - K)_+ H''(\mathrm{d}K) \\ & + \int_{(0, S_0]} (K - x)_+ H''(\mathrm{d}K). \end{aligned}$$

Taking $x = X_T = B_T^{-1} S_T$ we get:

$$\begin{aligned} H(X_T) \;=\; & H(S_0) + H'_+(S_0)(B_T^{-1} S_T - S_0) + B_T^{-1} \int_{(S_0, \infty)} (S_T - B_T K)_+ H''(\mathrm{d}K) \\ & + B_T^{-1} \int_{(0, S_0]} (B_T K - S_T)_+ H''(\mathrm{d}K). \end{aligned}$$

This implies that the payoff $H(X_T)$ can be replicated at time T by 'holding' a portfolio of:

$$
\begin{aligned}
&B_T^{-1}\big[H(S_0) - S_0 H'_+(S_0)\big] \quad \text{in cash;}\\
&B_T^{-1} H'_+(S_0) \quad \text{units of the asset;}\\
&B_T^{-1} H''(\mathrm{d}K) \quad \text{units of the call with strike } B_T K \text{ for } K \in (S_0, \infty);\\
&B_T^{-1} H''(\mathrm{d}K) \quad \text{units of the put with strike } B_T K \text{ for } K \in (0, S_0],
\end{aligned}
\tag{4.3.4}
$$

where the final two terms should be interpreted appropriately. In practice, the function $H(\cdot)$ can typically be approximated by a piecewise linear function, where the 'kinks' in the function correspond to traded strikes of calls or puts, in which case the number of units of each option to hold is determined by the change in the gradient at the relevant strike. The initial cost of setting up such a portfolio is well defined provided the integrability condition:

$$
\int_{(0,S_0]} P(B_T K)|H''|(\mathrm{d}K) + \int_{(S_0,\infty)} C(B_T K)|H''|(\mathrm{d}K) \;<\; \infty, \tag{4.3.5}
$$

is satisfied, where $|H''|(\mathrm{d}K)$ is the total variation of the signed measure $H''(\mathrm{d}K)$. We therefore shall make the following assumption:

Assumption 4.3.3. *The payoff $H(X_T)$ can be replicated using a suitable portfolio of call and put options, cash and the underlying, with a finite price at time* 0.

We can therefore combine these to get the following theorem:

Theorem 4.3.4. *Suppose Assumption 4.3.1 and 4.3.2 hold, and suppose $F(\cdot)$ is a convex, increasing function with $F(0) = 0$ and right derivative $f(t) := F'_+(t)$ which is bounded. Let $M(x,t)$ be defined as in (4.1.3), and is determined by the solution τ_D to* $\mathbf{SEP(\boldsymbol{\sigma}, \boldsymbol{\delta}_{S_0}, \boldsymbol{\mu})}$ *for $\sigma : x \mapsto x$, and where μ is determined by (4.3.2). We also define*

$$
Z(x) \;=\; 2\int_1^x \int_1^y \frac{M(z,0)}{z^2}\mathrm{d}z,
$$

and then the functions G and H are as defined in (4.1.6) and (4.1.7). Moreover, suppose Assumption 4.3.3 holds. Then there exists an arbitrage if the price of an *option with payoff $F(\langle \ln S\rangle_T)$ is less than:*

$$B_T^{-1}\left[G(S_0,0)+H(S_0)+\int_{(S_0,\infty)} C(B_TK)H''(\mathrm{d}K)+\int_{(0,S_0]} P(B_TK)H''(\mathrm{d}K)\right]. \tag{4.3.6}$$

Moreover, this bound is optimal in the sense that there exists a model which is free of arbitrage, under which the bound can be attained.

Proof. It follows from Theorem 3.3.1 that, given μ, we can find a domain D and corresponding stopping time τ_D which solves $\mathbf{SEP}(\boldsymbol{\sigma}, \boldsymbol{\delta_{S_0}}, \boldsymbol{\mu})$. Applying Proposition 4.1.3 (and bearing in mind Remark 4.1.7), we conclude that the strategy described above will indeed sub-replicate, and we can therefore produce an arbitrage by purchasing the option, and selling short the portfolio of calls, puts and the underlying given in (4.3.4), and in addition, holding the dynamic portfolio with $-\phi_t B_T^{-1}$ units of the underlying at time t. For the process $\widetilde{X}_t$ satisfying $\mathrm{d}\widetilde{X}_t = \widetilde{X}_t\mathrm{d}W_t$, given that f is bounded and

$$\left[Z'(\widetilde{X}_s)\sigma(\widetilde{X}_s)\right]^2 \;=\; 4\widetilde{X}_s^2\left[\int_1^{1+|\widetilde{X}_s-1|}\frac{M(z,0)}{z^2}\mathrm{d}z\right]^2 \;\le\; C\widetilde{X}_s^2\left[1-\frac{1}{1+\left|\widetilde{X}_s-1\right|}\right]^2,$$

where C is a constant only dependent on the upper bound of f, it follows that (4.1.9) holds. The other condition assumed in Theorem 4.1.5, (4.1.4), also clearly holds. As a consequence, we do indeed have a subhedge.

To see that this is the best possible bound, we need to show that there is a model which satisfies Assumption 4.3.1, has law μ under $\mathbb{Q}$ at time T, and such that the subhedge is actually a hedge. But consider the stopping time τ_D for the process $\widetilde{X}_t$. Define the process $(X_t;\, 0\le t\le T)$ where

$$X_t \;=\; \widetilde{X}_{\frac{t}{T-t}\wedge\tau_D}, \quad \text{for} \quad t\in[0,T],$$

and then X_t satisfies the stochastic differential equation $\mathrm{d}X_s = \sigma_s X_s \mathrm{d}W_s$ with the choice of

$$\sigma_s^2 \;=\; \frac{T-s+1}{(T-t)^2}\mathbf{1}_{[\frac{s}{T-s}<\tau_D]}. \tag{4.3.7}$$

Since $\tau_D<\infty$, a.s., then $X_T=\widetilde{X}_{\tau_D}$, $\tau_T=\tau_D$ and $S_t=X_tB_t$ is a price process satisfying Assumption 4.3.1 with

$$F\left(\int_0^T \sigma_t^2 \mathrm{d}t\right) \;=\; F(\tau_D).$$

Finally, it follows from (4.1.13) that at time T, the value of the hedging portfolio exactly equals the payoff of the option. $\square$

Note that we can drop the condition of f being bounded provided (4.1.4) and (4.1.9) hold.

Remark 4.3.5. The above results are given in the context of an increasing, convex function, but there is also a similar result concerning increasing, concave functions which can be derived. Consider a bounded, increasing function f satisfying (4.1.12), and define the concave function

$$L(t) \;=\; \int_0^t [C - f(s)]\mathrm{d}s \;=\; Ct - F(t).$$

Using Theorem 4.3.4 and the identity:

$$\ln(S_T) - \ln(S_0) \;=\; \int_0^T \frac{1}{S_t}\mathrm{d}S_t - \frac{1}{2}\langle \ln S\rangle_T\,,$$

it is easy to see that the price of a contract with payoff $L(\langle \ln S\rangle_T)$ must be bounded above by:

$$\begin{aligned}&2CQ - 2CB_T^{-1}\ln(S_0) - B_T^{-1}G(S_0,0) - B_T^{-1}H(S_0)\\&\qquad - B_T^{-1}\int_{(S_0,\infty)} C(B_T K)H''(\mathrm{d}K) - B_T^{-1}\int_{(0,S_0]} C(B_T K)H''(\mathrm{d}K),\end{aligned}$$

where Q is the price of a log-contract (that is, an option with payoff $\ln(S_T)$). As before, this upper bound is the best possible, under a similar set of assumptions.

Remark 4.3.6. An analogous result can be shown for *forward start options*. Suppose that the option has payoff

$$F\left(\int_S^T \sigma_t^2 \mathrm{d}t\right) \;=\; F\left(\langle S\rangle_T - \langle S\rangle_S\right)$$

for fixed times $0 < S < T$. Then we can use the previous results for general starting distributions to deduce a similar result to Theorem 4.3.4 for forward start options, provided we assume that there are calls traded at both S and T.

We use essentially the same idea as above: we aim to hold a portfolio which (sub-) replicates $G(X_t, \tau_t)$ for $t \in [S, T]$, and hold the payoff $H(X_T)$ as a portfolio of calls. However, we now have $\tau_t = \int_S^t \sigma_s^2 \mathrm{d}s$, and so $\widetilde{X}_t = X_{A_t}$, gives $\widetilde{X}_0 = X_S$ (recall that A_t was assumed right-continuous). The procedure is much as above, except that we need to use the solution to Theorem 4.1.5 with a general starting distribution, and the amount $G(\widetilde{X}_0, 0)$ will be a $\mathcal{F}_S$-random variable. The initial distribution ν can be derived using the Breeden-Litzenberger formula (4.3.2) at time S. To ensure that we hold the amount $G(\widetilde{X}_0, 0)$ at time S, we observe that $G(\widetilde{X}_0, 0) = G(X_S, 0)$. Hence, if e.g. $G(x, 0)$ can be written as the difference of two convex functions, we can replicate this amount by holding a portfolio of calls and puts with maturity S in a similar manner to (4.3.4). The remaining details follow as in the hedge described in Theorem 4.3.4.

Remark 4.3.7. We can also consider variants on the realised variance. Consider a slightly different time-change: suppose we set

$$\tau_t = \int_0^t \sigma_s^2 \lambda(X_s) \mathrm{d}s,$$

for some 'nice' function $\lambda(x)$, which in particular we suppose is bounded above and below by positive constants. Then following the computations above, we see that

$$\widetilde{X}_t = X_{A_T} = \int_0^{A_T} \frac{X_t}{\sqrt{\lambda(X_t)}} \left(\sigma_t \sqrt{\lambda(X_t)} \mathrm{d}W_t \right) = \int_0^t \frac{X_{A_t}}{\sqrt{\lambda(X_{A_t})}} \mathrm{d}\widetilde{W}_t,$$

and therefore $\mathrm{d}\widetilde{X}_t = \sigma(\widetilde{X}_t)\mathrm{d}\widetilde{W}_t$, where $\sigma : x \mapsto x/\sqrt{\lambda(x)}$. It seems feasible (Theorem 3.3.1 would need to be extended, but for 'nice' λ, this should be straightforward) that the above arguments could then be extended to provide robust hedges on convex payoffs of the form:

$$F\left(\int_0^T \sigma_s^2 \lambda(X_s) \mathrm{d}s \right).$$

An interesting special case of this would then be to give robust bounds on the price of an option on *corridor variance:*

$$F\left(\int_0^T \sigma_s^2 \mathbf{1}_{[a,b]}(S_s)) \mathrm{d}s \right). \tag{4.3.8}$$

by considering $\lambda(x) = \mathbf{1}_{[a,b]}(x)$, however this would only work in the case where there are no discount rates (i.e. $B_t = 1$). In general, we can only give a right lower bound for options on:

$$F\left(\int_0^T \sigma_s^2 \mathbf{1}_{[\bar{a},\bar{b}]}(X_s))\mathrm{d}s\right),$$

although this does provide a lower bound for (4.3.8) by considering the case where $\bar{a} := a$ and $\bar{b} = B_T b$.

Remark 4.3.8. Based on the discussion above and the results obtained in Section 4.2, we can also give the upper bound of a variance option with payoff $F(\langle \ln S\rangle_T)$.

To find the upper bound, we suppose that $f(t) := F'_-(t)$ satisfies (4.2.11), and then we can define $M(x,t)$ and $J(x,t)$ as in (4.1.3) and (4.2.12), where the stopping time τ_D is the solution τ_D to $\mathbf{SEP^*}(\boldsymbol{\sigma}, \boldsymbol{\delta_{S_0}}, \boldsymbol{\mathcal{L}(S_T)})$ for $\sigma : x \mapsto x$. We also define $Q(x) = 2(x - \ln x - 1)$ (so $Q''(x) = 2/\sigma^2$). If $\widetilde{X}$ is geometric Brownian motion with initial value S_0 (without loss of generality, we assume $S_0 = 1$), then

$$Q(\widetilde{X}_t) \;=\; 2\left[e^{W_t - t/2} - \left(W_t - \frac{t}{2}\right) - 1\right].$$

It is easy to check that, for all $T > 0$, $\sup_{t\le T} \mathbb{E}[Q^2(\widetilde{X}_t)] < \infty$. It follows that $Q(\widetilde{X})$ satisfies (4.2.2). Now given the functions $\widetilde{G}$ and $\widetilde{H}$ as defined in (4.2.13), our superhedge of the variance option can be described as the combination of a static portfolio (4.3.4) where H is replaced by $\widetilde{H}$, and a self-financing dynamic portfolio which consists of $B_T^{-1}\psi_t$ units of the risky asset and $B_T^{-1}\left(\widetilde{G}(S_0,0) - \psi_t S_t\right)$ in cash. Here we identify the process $\psi_t = \widetilde{\psi}_{\tau_t}$ by

$$\widetilde{\psi}_t \;\in\; \left[\frac{\partial \widetilde{G}}{\partial x}(x+,t),\, \frac{\partial \widetilde{G}}{\partial x}(x-,t)\right].$$

It is easy to see that the total initial investment of this superhedge is

$$B_T^{-1}\left[\widetilde{G}(S_0,0) + \widetilde{H}(S_0) + \int_{(S_0,\infty)} C(B_T K)\widetilde{H}''(\mathrm{d}K) + \int_{(0,S_0]} P(B_T K)\widetilde{H}''(\mathrm{d}K)\right]. \tag{4.3.9}$$

A special case is when σ_t coincides with (4.3.7) where τ_D is the Rost's solution, the superhedge is indeed a hedge, and hence, we conclude the upper bound (4.3.9) is an optimal upper bound. And then we have the following theorem.

Theorem 4.3.9. *Suppose Assumption 4.3.1 and Assumption 4.3.2 hold, and suppose $F(\cdot)$ is a convex, increasing function with $F(0) = 0$ and left derivative $f(t) := F'_+(t)$ which is bounded and satisfies (4.2.11). Let $M(x,t)$ and $J(x,t)$ be defined as in (4.1.3)* *and (4.2.12), and are determined by the solution* τ_D *to* $\mathbf{SEP^*}(\boldsymbol{\sigma}, \boldsymbol{\delta_{S_0}}, \boldsymbol{\mu})$ *for* $\sigma : x \mapsto x$*, and where* μ *is determined by (4.3.2). We also define* $Q(x) = 2(x - \ln x - 1)$*, and then the functions* $\widetilde{G}$ *and* $\widetilde{H}$ *are as defined in (4.2.13). Moreover, suppose Assumption 4.3.3 holds. Then there exists an arbitrage if the price of an option with payoff* $F(\langle \ln S \rangle_T)$ *is higher than the amount given by (4.3.9). Moreover, this bound is optimal in the sense that there exists a model which is free of arbitrage, under which the bound can be attained.*

Chapter 5

Further Work

In this final chapter we present some further questions which have arisen from the previous work.

In Chapter 2 and Chapter 3, we discussed the construction of Root's barrier by a corresponding free boundary problem and variational inequality respectively. The most significant difference is that, in the construction by variational inequality, we have to assume that the diffusion coefficient $\sigma \geq \varepsilon > 0$, to guarantee the uniqueness of the strong solution to the variational inequality (see the proof of Bensoussan and Lions [1982, Theorem 2.2, Chapter 3]). When σ fails to satisfy this condition, we might have two options: performing a simple transformation of the variational inequality, just as we have done in Section 3.3, or constructing Root's barrier by the free boundary problem. The former method is valid in only a limited number of cases. In general, we turn to the free boundary problem, and we then find that, instead of the boundedness condition on σ, the vanishing second derivative condition (2.5.10) arises. Since it is difficult to verify, we could ask whether we can get rid of this condition, or whether we can replace this condition by some other conditions which are more practical.

The next question arises from Remark 3.2.4 where we discuss the connection to minimality and non-centred target distributions. We have assumed the existence of Root's solution in advance. Can we prove the existence directly from the variational inequality, both for centred and non-centred distributions?

In fact, given the variational inequality with $u(x, 0) = \mathrm{U}\nu$ and $\bar{u} = \mathrm{U}\mu - C$ (with the assumption, $\mathrm{U}\mu - C \leq \mathrm{U}\nu$). By optimality, $\{(x, t)|u(x, t) = \bar{u}\}$ gives a barrier, and moreover, u is a concave function. If we can show (for example, under the assumption (2.3.3))

$$-\frac{1}{2}u''(\mathrm{d}x) \;=\; \mathbb{P}^{\nu}[W_{t\wedge\tau_D} \in \mathrm{d}x] \quad \text{for} \quad (x,t) \in D = [u(x,t) > \bar{u}],$$

then $u = -\mathbb{E}|x - W_{t\wedge\tau_D}|$, and hence, as $t \to \infty$ $-\mathbb{E}|x - W_{\tau_D}| + \mathbb{E}|y - W_{\tau_D}| = \mathrm{U}\mu(x) - \mathrm{U}\mu(y)$. Then the law of W_{τ_D} is μ.

We also believe that there are interesting lines of research that now arise. The construction opens up a number of questions regarding Root's solution to the Skorokhod embedding problem: for example, what can be said about the shape of the boundary? Under what conditions on μ will the boundary be smooth? When does $R(x) \to 0$ as $x \to \pm\infty$? When is $R(x)$ bounded? Properties of free boundaries are well-studied in the analytic literature, and may be useful in answering these questions.

The connection with optimal stopping noted in Remark 3.2.3 is interesting, and obtaining a deeper understanding between optimal stopping problems and optimal Skorokhod embeddings seems to be an interesting area of research.

Other questions that arise from the practical standpoint include how to incorporate additional market information (e.g. calls at an intermediate time Brown et al. [2001]), and how to adjust for the fact that there will generally only be a finite set of quoted calls (see Davis et al. [2010] for a related question). Remark 4.3.7 also suggests open questions regarding more general choices of $\sigma(x)$.

Appendix A

The Matlab code used in Section 3.5 (Root's case)

```
function [v, r, x, t] = EVI(m,n,T)
% this function calculates the numerical solution to EVI(mu), v,
% and the numerical solution to SEP(mu), r.

dx = 2/m;  x = [0:m]'*dx-1;  dt = T/n;  t = [0:n]*dt;
r = zeros(m+1,1);  v = zeros(m+1,n+1);

% define v0 is the potential of the initial distribution,
%       psi is the potential of the target distribution.
v0 = -abs(x);  psi = zeros(size(x));
for i = 1:m+1
    if abs(x(i))<1
        psi(i) = -(1/2)*((x(i))^2+1);
    else
        psi(i) = -abs(x(i));
    end
end

% then solve the variational inequality by Crank-Nicholson method.
v(:,1) = v0;  v(1,:) = psi(1);  v(m+1,:) = psi(m+1);
```

```
a = (dt/(dx^2))/4;  d = [-2*a; zeros(m-3,1); -2*a];
P = (1-2*a)*eye(m-1);  Q = (1+2*a)*eye(m-1);
for i = 2:m-1
    P(i-1,i) = a; Q(i-1,i) = -a; P(i,i-1) = a; Q(i,i-1) = -a;
end

for k = 2:n+1
  % Q is tridiagonal, apply the tridiagonal matrix algorithm (TDMA).
    b(1) = Q(1,1); c(1) = Q(1,2); e(1) = 0;
    b(m-1) = Q(m-1,m-1); c(m-1) = 0; e(m-1) = Q(m-1,m-2);
    for i = 2:m-2
        b(i) = Q(i,i); c(i) = Q(i,i+1); e(i) = Q(i,i-1);
    end
    h = P*v(2:m,k-1)+d; c(1) = c(1)/b(1); h(1) = h(1)/b(1);
    for i = 2:m-1
        c(i) = c(i)/(b(i)-c(i-1)*e(i));
        h(i) = (h(i)-h(i-1)*e(i))/(b(i)-c(i-1)*e(i));
    end
    y(m-1) = h(m-1);
    for i = m-2:-1:1
        y(i) = h(i)-c(i)*y(i+1);
    end
  % TDMA end.
    v(2:m,k) = max(y', psi(2:m));
end

% at last compute the barrier.
for k = 1:m+1
    p = find(v(k,:) == psi(k));
    if  numel(p) == 0
        r(k) = 0;
    else
        q = p(1); r(k) = t(q);
    end
end
```

Appendix B

The Matlab code used in Section 3.5 (Rost's case)

```
function [v, r, x, t] = rEVI(m,n,T)
% this function calculates the numerical solution to EVI*(mu), v,
% and the numerical solution to SEP*(mu), r.

dx = 2/m;  x = [0:m]'*dx-1;  dt = T/n;  t = [0:n]*dt;
r = zeros(m+1,1);  v = zeros(m+1,n+1);

% define v0 = psi as the difference between the potentials
% of target and initial distribution
v0 = -abs(x);  psi = zeros(size(x));
for i = 1:m+1
    if abs(x(i))<1
        psi(i) = -(1/2)*((x(i))^2+1);
    else
        psi(i) = -abs(x(i));
    end
end
v0 = psi -v0; psi = v0;

% then solve the variational inequality by Crank-Nicholson method.
```

```
v(:,1) = v0;  v(1,:) = psi(1);  v(m+1,:) = psi(m+1);
a = (dt/(dx^2))/4;  d = [-2*a; zeros(m-3,1); -2*a];
P = (1-2*a)*eye(m-1);  Q = (1+2*a)*eye(m-1);
for i = 2:m-1
    P(i-1,i) = a; Q(i-1,i) = -a; P(i,i-1) = a; Q(i,i-1) = -a;
end
for k = 2:n+1
  % Q is tridiagonal, apply the tridiagonal matrix algorithm (TDMA).
    b(1) = Q(1,1); c(1) = Q(1,2); e(1) = 0;
    b(m-1) = Q(m-1,m-1); c(m-1) = 0; e(m-1) = Q(m-1,m-2);
    for i = 2:m-2
        b(i) = Q(i,i); c(i) = Q(i,i+1); e(i) = Q(i,i-1);
    end
    h = P*v(2:m,k-1)+d; c(1) = c(1)/b(1); h(1) = h(1)/b(1);
    for i = 2:m-1
        c(i) = c(i)/(b(i)-c(i-1)*e(i));
        h(i) = (h(i)-h(i-1)*e(i))/(b(i)-c(i-1)*e(i));
    end
    y(m-1) = h(m-1);
    for i = m-2:-1:1
        y(i) = h(i)-c(i)*y(i+1);
    end
  % TDMA end.
    v(2:m,k) = max(y', psi(2:m));
end

% at last compute the barrier.
for k = 1:m+1
    p = find(v(k,:) ~= psi(k));
    if  numel(p) == 0
        r(k) = T;
    else
        q = p(1); r(k) = t(q);
    end
end
```

Bibliography

Jacques Azéma and Marc Yor. Une solution simple au problème de Skorokhod. In *Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78)*, volume 721 of *Lecture Notes in Math.*, pages 90–115. Springer, Berlin, 1979a.

Jacques Azéma and Marc Yor. Le problème de Skorokhod: compléments à "Une solution simple au problème de Skorokhod". In *Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78)*, volume 721 of *Lecture Notes in Math.*, pages 625–633. Springer, Berlin, 1979b.

Alain Bensoussan and Jacques-Louis Lions. *Applications of variational inequalities in stochastic control*, volume 12 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1982. ISBN 0-444-86358-3. Translated from the French.

Douglas T. Breeden and Robert H. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of Business*, 51(4):621–651, October 1978.

Mark Broadie and Ashish Jain. Pricing and hedging volatility derivatives. *The Journal of Derivatives*, 15(3):7–24, 2008.

Haydyn Brown, David Hobson, and L. C. G. Rogers. Robust hedging of barrier options. *Math. Finance*, 11(3):285–314, 2001. ISSN 0960-1627. doi: 10.1111/1467-9965.00116.

Peter Carr and Roger Lee. Hedging variance options on continuous semimartingales. *Finance Stoch.*, 14(2):179–207, 2010. ISSN 0949-2984. doi: 10.1007/s00780-009-0110-3.

Peter Carr, Roger Lee, and Liuren Wu. Variance swaps on time-changed Lévy processes. *Finance and Stochastics*, pages 1–21, 2011. ISSN 0949-2984. 10.1007/s00780-011-0157-9.

R. V. Chacon. Potential processes. *Trans. Amer. Math. Soc.*, 226:39–58, 1977. ISSN 0002-9947.

R. V. Chacon and J. B. Walsh. One-dimensional potential embedding. In *Séminaire de Probabilités, X (Prèmiere partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)*, pages 19–23. Lecture Notes in Math., Vol. 511. Springer, Berlin, 1976.

Rene Manuel Chacon. *Barrier stopping times and the filling scheme*. ProQuest LLC, Ann Arbor, MI, 1985. Thesis (Ph.D.) – University of Washington.

A. M. G. Cox. Extending Chacon-Walsh: minimality and generalised starting distributions. In *Séminaire de probabilités XLI*, volume 1934 of *Lecture Notes in Math.*, pages 233–264. Springer, Berlin, 2008. doi: 10.1007/978-3-540-77913-1_12.

A. M. G. Cox and D. G. Hobson. An optimal Skorokhod embedding for diffusions. *Stochastic Process. Appl.*, 111(1):17–39, 2004. ISSN 0304-4149. doi: 10.1016/j.spa.2004.01.003.

A. M. G. Cox and D. G. Hobson. Local martingales, bubbles and option prices. *Finance Stoch.*, 9(4):477–492, 2005. ISSN 0949-2984. doi: 10.1007/s00780-005-0162-y.

A. M. G. Cox and J. Obłój. Robust pricing and hedging of double no-touch options. *Finance and Stochastics*, pages 1–33, 2011a. ISSN 0949-2984. doi: 10.1007/s00780-011-0154-z.

A. M. G. Cox and Jan Obłój. Robust hedging of double touch barrier options. *SIAM Journal on Financial Mathematics*, 2(1):141–182, 2011b. ISSN 1945-497X. doi: 10.1137/090777487.

A. M. G. Cox and Jiajie Wang. Root's barrier: Construction, optimality and applications to variance options. Preprint, arXiv:1104.3583, 2011.

A. M. G. Cox, David Hobson, and Jan Obłój. Pathwise inequalities for local time: applications to Skorokhod embeddings and optimal stopping. *Ann. Appl. Probab.*, 18(5):1870–1896, 2008. ISSN 1050-5164. doi: 10.1214/07-AAP507.

Mark H.A. Davis, Jan Obłój, and Vimal Raval. Arbitrage bounds for weighted variance swap prices. Preprint, arXiv:1001.2678, 2010.

H. Dinges. Stopping sequences. In *Séminaire de Probabilitiés, VIII (Univ. Strasbourg, année universitaire 1972–1973)*, pages 27–36. Lecture Notes in Math., Vol. 381. Springer, Berlin, 1974. Journées de la Sociéte Mathématique de France de Probabilités, Strasbourg, 25 Mai 1973.

Lester E. Dubins. On a theorem of Skorohod. *Ann. Math. Statist.*, 39:2094–2097, 1968. ISSN 0003-4851.

Lester E. Dubins and Gideon Schwarz. A sharp inequality for sub-martingales and stopping-times. *Astérisque*, (157-158):129–145, 1988. ISSN 0303-1179. Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987).

Bruno Dupire. Model art. *Risk*, 6(9):118–120, 1993.

Bruno Dupire. Arbitrage bounds for volatility derivatives as free boundary problem. Presentation at *PDE and Mathematical Finance, KTH, Stockholm*, 2005.

Erik Ekström and Johan Tysk. Boundary behaviour of densities for non-negative diffusions. Available from: `http://www2.math.uu.se/~johant/pq.pdf`, 2011.

David Freedman. *Brownian motion and diffusion*. Holden-Day, San Francisco, Calif., 1971.

Avner Friedman. *Generalized functions and partial differential equations*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1963.

Paul R. Halmos. *Measure Theory*. D. Van Nostrand Company, Inc., New York, N. Y., 1950.

David G. Hobson. Robust hedging of the lookback option. *Finance Stoch.*, 2(4):329–347, 1998a. ISSN 0949-2984. doi: 0.1007/s007800050044.

David G. Hobson. The maximum maximum of a martingale. In *Séminaire de Probabilités, XXXII*, volume 1686 of *Lecture Notes in Math.*, pages 250–263. Springer, Berlin, 1998b. doi: 10.1007/BFb0101762.

David G. Hobson. The Skorokhod embedding problem and model-independent bounds for option prices. Available from: `http://www.warwick.ac.uk/go/dhobson/`, 2009.

David G. Hobson and J. L. Pedersen. The minimum maximum of a continuous martingale with given initial and terminal laws. *Ann. Probab.*, 30(2):978–999, 2002. ISSN 0091-1798. doi: 10.1214/aop/1023481014.

S. Howison, A. Rafailidis, and H. Rasmussen. On the pricing and hedging of volatility derivatives. *Applied Mathematical Finance*, 11(4):317–346, 2004.

Barthel W. Huff. The inverse Gaussian distribution and Root's barrier construction. *Sankhyā Ser. A*, 37(3):345–353, 1975. ISSN 0581-572X.

Kiyosi Itô and Henry P. McKean, Jr. *Diffusion processes and their sample paths*. Springer-Verlag, Berlin, 1974. Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 125.

J. Kallsen, J. Muhle-Karbe, and M. Voß. Pricing options on variance in affine stochastic volatility models. *Mathematical Finance*, 2010. doi: 10.1111/j.1467-9965.2010.00447.x.ToAppear.

Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991. ISBN 0-387-97655-8.

Martin Keller-Ressel. Convex order properties of discrete realized variance and applications to variance options. Preprint, arXiv:1103.2310, 2011.

Martin Keller-Ressel and Johannes Muhle-Karbe. Asymptotic and exact pricing of options on variance. Preprint, arXiv:1003.5514, 2010.

R. M. Loynes. Stopping times on Brownian motion: Some properties of Root's construction. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 16:211–218, 1970.

Terry R. McConnell. The two-sided Stefan problem with a spatially dependent latent heat. *Trans. Amer. Math. Soc.*, 326(2):669–699, 1991. ISSN 0002-9947.

Isaac Meilijson. The time to a given drawdown in Brownian motion. In *Séminaire de Probabilités XXXVII*, volume 1832 of *Lecture Notes in Math.*, pages 94–108. Springer, Berlin, 2003.

Itrel Monroe. On embedding right continuous martingales in Brownian motion. *Ann. Math. Statist.*, 43:1293–1311, 1972a. ISSN 0003-4851.

Itrel Monroe. Using additive functionals to embed preassigned distributions in symmetric stable processes. *Trans. Amer. Math. Soc.*, 163:131–146, 1972b. ISSN 0002-9947.

Anthony Neuberger. The log contract. *The Journal of Portfolio Management*, 20(2):74–80, 1994.

Jan Obłój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390 (electronic), 2004. ISSN 1549-5787. doi: 10.1214/154957804100000060.

Jan Obłój. The maximality principle revisited: on certain optimal stopping problems. In *Séminaire de Probabilités XL*, volume 1899 of *Lecture Notes in Math.*, pages 309–328. Springer, Berlin, 2007. doi: 10.1007/978-3-540-71189-6_16.

Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, fourth edition, 1995. ISBN 3-540-60243-7. An introduction with applications.

Edwin Perkins. The Cereteli-Davis solution to the H^1-embedding problem and an optimal embedding in Brownian motion. In *Seminar on stochastic processes, 1985 (Gainesville, Fla., 1985)*, volume 12 of *Progr. Probab. Statist.*, pages 172–223. Birkhäuser Boston, Boston, MA, 1986.

Goran Peskir. Optimal stopping of the maximum process: the maximality principle. *Ann. Probab.*, 26(4):1614–1640, 1998. ISSN 0091-1798. doi: 10.1214/aop/1022855875.

Goran Peskir. Designing options given the risk: the optimal Skorokhod-embedding problem. *Stochastic Process. Appl.*, 81(1):25–38, 1999. ISSN 0304-4149. doi: 10.1016/S0304-4149(98)00097-0.

Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. ISBN 3-540-00313-4. Second edition. Version 2.1, Corrected third printing.

Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999. ISBN 3-540-64325-7.

L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 1.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000a. ISBN 0-521-77594-9. Foundations, Reprint of the second (1994) edition.

L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 2.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000b. ISBN 0-521-77593-0. Itô calculus, Reprint of the second (1994) edition.

D. H. Root. The existence of certain stopping times on Brownian motion. *Ann. Math. Statist.*, 40:715–718, 1969. ISSN 0003-4851.

H. Rost. The stopping distributions of a Markov Process. *Invent. Math.*, 14:1–16, 1971. ISSN 0020-9910.

H. Rost. Skorokhod stopping times of minimal variance. In *Séminaire de Probabilités, X (Première partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)*, pages 194–208. Lecture Notes in Math., Vol. 511. Springer, Berlin, 1976.

Stanley Sawyer. The Skorokhod representation. *Rocky Mountain J. Math.*, 4:579–596, 1974. ISSN 0035-7596. Papers arising from a Conference on Stochastic Differential Equations (Univ. Alberta, Edmonton, Alta., 1972).

A. V. Skorokhod. *Issledovaniya po teorii sluchainykh protsessov (Stokhasticheskie differentsialnye uravneniya i predelnye teoremy dlya protsessov Markova).* Izdat. Kiev. Univ., Kiev, 1961.

A. V. Skorokhod. *Studies in the theory of random processes.* Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass., 1965.

Daniel W. Stroock. *Partial differential equations for probabilists*, volume 112 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2008. ISBN 978-0-521-88651-2.

Pierre Vallois. Le problème de Skorokhod sur R: une approche avec le temps local. In *Seminar on probability, XVII*, volume 986 of *Lecture Notes in Math.*, pages 227–239. Springer, Berlin, 1983.