Stochastic control problems for multidimensional martingales

Benjamin Alexander Robinson

Thesis submitted for the degree of
Doctor of Philosophy

University of Bath

Department of Mathematical Sciences

September 2020

Declaration of any previous submission of the work

The material presented here for examination for the award of a higher degree by research has not been incorporated into a submission for another degree.

Candidate's signature ....

Benjamin A. Robinson

Declaration of authorship

I am the author of this thesis, and the work described therein was carried out by myself personally, in collaboration with my supervisor, Alexander M. G. Cox.

Candidate's signature ....

Benjamin A. Robinson

SUMMARY

We study a stochastic control problem for continuous multidimensional martingales, motivated by recent developments in robust finance and martingale optimal transport.

In a radially symmetric environment, we explicitly construct the solution to this problem under mild regularity conditions. We consolidate some ideas from the theory of viscosity solutions of PDEs, which we then apply to solve our problem.

Under a particular growth condition on the cost function, we solve the control problem in the two-dimensional case by proving that a weak solution of a certain SDE generates a Brownian filtration. We prove non-existence of strong solutions of this SDE and a related SDE, building on ideas from the study of Tsirelson's equation. These results lead us to conjecture that there is a gap between a Markov formulation of the control problem and a strong and weak formulation.

Finally, we draw a connection to two further control problems. We characterise each of these problems in terms of viscosity solutions of a Monge-Ampère equation, similar to that which arises in the classical theory of optimal transport.

ACKNOWLEDGEMENTS

There are many people without whom it would not have been possible to produce this thesis, or at least the process would have been much less enjoyable.

First, I would like to thank my supervisor Alexander Cox for introducing me to this area of research, and for always finding time to discuss mathematics with me and to read my work. I am also grateful to Andreas Kyprianou for his guidance and encouragement, and to Jon Warren for thoroughly examining this thesis and providing useful recommendations.

I have been supported by a scholarship from the EPSRC Centre for Doctoral Training in Statistical Applied Mathematics at Bath (SAMBa), under the project EP/L015684/1. Being a part of SAMBa has meant being in an amazing community within the department. In particular, I would like to acknowledge Susie and Jess for all they have done for us in SAMBa.

It has been a pleasure to share an office with the past and present inhabitants of 4 West 5.8, and I am very grateful to Emiko for facilitating 'Strike Office', and to everyone who has kept our coffee breaks going virtually. Thank you to Marcus and John for many useful mathematical discussions, to James for his sound advice, and to Matt for helping to reduce the number of grammatical errors in this thesis.

Thank you to everyone who has been my friend in Bath, whether singing with me in choirs, playing quizzes, standing on picket lines, organising conferences, running, dancing, or drinking coffee. Thank you to my extended household at Livingstone Terrace for many great memories, especially to Federico who stuck with me for 4 years, and to Malena and Elizabeth whose friendship has been invaluable over the years. Special thanks also go to Lizzi and Chris for putting up with me throughout lockdown and helping me to survive a pandemic whilst writing this thesis.

I would also like to thank A1, Swoop, and all my family and friends outside of Bath for their support. Finally, thank you to Hayley for always being by my side, with coffee and teddies, through the worst and best moments of my PhD journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ALGORITHMS

CHAPTER 1 INTRODUCTION

The theory of stochastic optimal control is concerned with solving optimisation problems where the dynamics of the underlying process are stochastic. Such a process is typically described via a stochastic integral or stochastic differential equation that depends on an auxiliary stochastic process. This auxiliary process is known as the control process and can be chosen from a given set. The aim of a stochastic optimal control problem is to minimise the expectation of a cost that depends on the path of the underlying stochastic process.

In this thesis we study stochastic optimal control problems and their intersection with the theories of martingale optimal transport, existence of weak and strong solutions of SDEs, and Monge-Ampère equations.

We consider problems related to the control of continuous multidimensional martingales with fixed quadratic variation. In one dimension, it is well known that any continuous martingale is a time-change of a standard Brownian motion. Martingales with fixed quadratic variation are a natural generalisation of Brownian motion to higher dimensions. Imposing this constraint will allow us to study the structure of the optimal martingales in the control problems that we consider. Further motivation for studying these problems comes from a connection with martingale optimal transport, as we explain in the following section.

1.1 Motivation: Robust finance

We begin by motivating the stochastic control problem with a discussion of recent developments in the field of robust finance.

In classical mathematical finance, one typically assumes that the price of some underlying asset evolves according to an SDE from some parametric family. Then historical data is used to estimate the parameters. Model misspecification is a potential flaw in this approach. No model can be entirely correct and, in particular, any model will fail to accurately predict the impact of a severe unexpected event. It is therefore important to quantify the uncertainty in the choice of the model. This type of uncertainty is often referred to as Knightian uncertainty.

To avoid the pitfalls that come with specifying a model, Hobson first proposed a robust method for pricing and hedging a lookback option in the 1998 paper [32]. Hobson obtains hedging strategies and bounds on the price of the option that are independent of any model and based instead on observed option prices. Since this paper, the field of *robust* or *model-independent finance* has expanded significantly. One of the bounds in [32] is obtained by relating the problem to a Skorokhod embedding problem; a connection of this type holds in more generality. The 2011 lecture notes of Hobson [33] provide a survey of results in model-independent finance that are obtained from Skorokhod embedding techniques.

More recently, a variation of the Monge-Kantorovich optimal transport problem has been applied to problems in robust finance. The *martingale optimal transport* problem, as it is known, was initially developed in the context of robust finance in papers of Beiglböck, Henry-Labordère and Penkner [3], Galichon, Henry-Labordère and Touzi [27], Hobson and Klimmek [34], and Hobson and Neuberger [35].

1.1.1 Martingale optimal transport

The problem of martingale optimal transport is a variation on the classical Monge-Kantorovich transport problem. Fix $d \in \mathbb{N}$ and suppose that μ_0 and μ_1 are probability measures on $\mathbb{R}^d$. Let $\mathcal{M}$ be the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ and define the set of *couplings* of μ_0 and μ_1 by

$$\Pi(\mu_0, \mu_1) := \left\{ \pi \in \mathcal{M} \colon \int_{\mathbb{R}^d} \pi(\cdot, \mathrm{d}y) = \mu_0(\cdot), \int_{\mathbb{R}^d} \pi(\mathrm{d}x, \cdot) = \mu_1(\cdot) \right\}.$$

Let $c : \mathbb{R}^d \times \mathbb{R}^d$ be a measurable function. Then, as defined in Villani's book [63], the optimal transport problem is to find

$$\inf_{\pi \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) \pi(\mathrm{d}x, \mathrm{d}y).$$

To define the martingale optimal transport problem we restrict the set of admissible couplings to satisfy an additional martingale condition, as in [4] for example.

Define the set of *martingale couplings* of μ_0 and μ_1 by

$$\Pi_M(\mu_0,\mu_1) := \{\pi \in \Pi(\mu_0,\mu_1) \colon \mathbb{E}^\pi(Y|X) = X, \quad \text{if} \quad \mathrm{Law}(X,Y) = \pi\}\,.$$

Then the martingale optimal transport problem is to find

$$\inf_{\pi \in \Pi_M(\mu_0,\mu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x,y)\pi(\mathrm{d}x,\mathrm{d}y).$$

Equivalently, we can express this as finding

$$\inf_{\substack{\pi \in \Pi_M(\mu_0,\mu_1) \\ \mathrm{Law}(X,Y)=\pi}} \mathbb{E}^\pi\left[c(X,Y)\right].$$

In the case that $d = 1$, Beiglböck and Juillet established results on the structure of optimal couplings in [4], and in [5] Beiglböck, Nutz and Touzi proved a duality result analogous to the classical Monge-Kantorovich duality. Namely, for c lower semicontinuous and non-negative, define a set of triples of integrable functions

$$\begin{aligned}\mathcal{H}(\mu_0,\mu_1) := \{(\phi,\psi,h) \colon \quad & c(X,Y) \geq \phi(X) + \psi(Y) + h(X)(Y-X) \\ & \pi - a.s. \quad \text{for all } \pi \in \Pi_M(\mu_0,\mu_1)\}\,.\end{aligned}$$

Then

$$\inf_{\substack{\pi \in \Pi_M(\mu_0,\mu_1) \\ \mathrm{Law}(X,Y)=\pi}} \mathbb{E}^\pi\left[c(X,Y)\right] = \sup_{(\phi,\psi,h)\in\mathcal{H}(\mu_0,\mu_1)} \left\{\mathbb{E}^{\mu_0}[\phi(X)] + \mathbb{E}^{\mu_1}[\psi(Y)]\right\}.$$

Financially, the solution of the dual problem on the right hand side corresponds to the maximum cost of a subhedging strategy, and this is independent of any choice of model.

De March extended this duality result to higher dimensions in [16], based on the structure results of De March and Touzi in [17]. Further results on the structure of optimal couplings in arbitrary dimensions are established by De March in [15] and by Ghoussoub, Kim and Lim in [29].

1.1.2 Relationship to stochastic control

In [57], Tan and Touzi present an alternative approach to solving the martingale optimal transport problem, similar to the work of Benamou and Brenier for the classical optimal transport problem in [8]. Backhoff-Veraguas, Beiglböck, Huesmann and Källblad take a similar approach in [1]. In [57], Tan and Touzi consider a dual formulation of the problem, which we rewrite as follows. Consider a process X and define the set of measures

$$\mathcal{P} := \left\{\mathbb{P} : \quad (X_t)_{t\in[0,1]} \quad \text{is a martingale under} \quad \mathbb{P}, \quad \operatorname{Law}(X_0) = \mu_0\right\}.$$

Then, defining $C_b(\mathbb{R}^d)$ to be the set of bounded continuous functions on $\mathbb{R}^d$,

$$\begin{aligned}\inf_{\substack{\mathbb{P}\in\mathcal{P}\\ \operatorname{Law}(X_1)=\mu_1}} \quad & \mathbb{E}^{\mathbb{P}}\left[\int_0^1 f(X_s)\,\mathrm{d}s\right] \\ &= \sup_{\lambda\in C_b(\mathbb{R}^d)} \left\{\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}\left[\int_0^1 f(X_s)\,\mathrm{d}s + \lambda(X_1)\right] - \int_{\mathbb{R}^d} \lambda(x)\mu_1(\mathrm{d}x)\right\}.\end{aligned}$$

The interior optimisation problem is a stochastic optimal control problem, and it is this type of problem that we will investigate in this thesis. Instead of the fixed time horizon in the above problem, we will consider martingales running up to the first exit time of a bounded domain, and we will constrain the martingales to have fixed quadratic variation. Defining the set of measures

$$\tilde{\mathcal{P}} := \{\mathbb{P} : \quad (X_t)_{t\geq 0} \text{ is a martingale under } \mathbb{P}, \quad \langle X\rangle_t = t,\ t\geq 0, \quad \operatorname{Law}(X_0) = \mu_0\},$$

and a set of stopping times $\mathcal{T}$, we are interested in finding

$$\sup_{\lambda\in C_b(\mathbb{R}^d)} \left\{\sup_{\tau\in\mathcal{T}} \inf_{\mathbb{P}\in\tilde{\mathcal{P}}} \mathbb{E}^{\mathbb{P}}\left[\int_0^\tau f(X_s)\,\mathrm{d}s + \lambda(X_\tau)\right] - \int_{\mathbb{R}^d} \lambda(x)\mu_1(\mathrm{d}x)\right\}.$$

In this thesis we focus on the interior control problem of optimising over martingales.

1.2 Literature on the stochastic control problem

An overview of stochastic control is presented by Fleming and Soner in [26], by Pham in [49], and by Touzi in [58]. Pham and Touzi both describe some of the financial applications of the theory of stochastic control in [49] and [58], respectively.

In the following section, we will take our definition for the strong formulation of the control problem from [58]. We will also consider a weak formulation of the control problem, following El Karoui and Tan in [20], where they introduced weak and relaxed forms of the control problem. We will refer frequently to [26] and [58] in the following sections as we introduce the dynamic programming principle and the Hamilton-Jacobi-Bellman (HJB) equation as tools to solve the stochastic control problem.

When considering martingales with a fixed quadratic variation, an appropriate control set is

$$U = \left\{ \sigma \in \mathbb{R}^{d,d} \colon \operatorname{Tr}(\sigma\sigma^\top) = 1 \right\}.$$

We see this by considering a martingale X that is adapted to the natural filtration of some Brownian motion B. Then there exists an adapted process σ such that X has the representation

$$\mathrm{d}X_t = \sigma_t \, \mathrm{d}B_t. \tag{1.1}$$

If $\sigma_t \in U$ for all $t \geq 0$, then the quadratic variation $\langle X \rangle$ is given by

$$\mathrm{d}\langle X \rangle_t = \operatorname{Tr}(\sigma_t \sigma_t^\top) \, \mathrm{d}t = \mathrm{d}t.$$

With the above definition of the control set U, the value of the control problem should solve the HJB equation

$$-\frac{1}{2} \inf_{\sigma \in U} \operatorname{Tr}\left(D^2 u \sigma \sigma^\top\right) = f. \tag{1.2}$$

The specific problem of stochastic control over martingales with a fixed quadratic variation has appeared recently in the two papers [40] and [41] of Larsson and Ruf. In [40] the authors consider the problem of finding the greatest almost sure lower bound on the exit time of a martingale from some domain. Larsson and Ruf apply this control problem in [41] to find the minimal time horizon over which relative arbitrage can be achieved for a market with at least two stocks. While the control set in [40] and [41] is the same as in the problem that we study, we consider a different class of cost functions.

The HJB equation (1.2) takes a similar form to the Black-Scholes-Barenblatt (BSB) equation, as studied in [61]. Compared with the PDE (1.2), the BSB equation has an additional time derivative term, and the infimum can be taken over a more general compact control set. The BSB equation is an HJB equation corresponding to a time-inhomogeneous control problem of the type discussed in Section 3.3 of [58]. In [30], the BSB equation is applied to find a super-hedging strategy for European multi-asset derivatives.

Feng and Jensen study another HJB equation in [24] that is related to the PDE (1.2). The same control set U appears in their HJB equation, but there is an additional term inside the infimum that depends directly on the element of U. In Section 5.3, we will show that the control problem corresponding to this HJB equation is related to our original problem. In [24], the authors use the equivalence of this HJB equation with a Monge-Ampère equation in order to develop a numerical scheme for the Monge-Ampère equation, but they do not discuss the associated control problem.

In [28], Gaveau studies another control problem that is related to a Monge-Ampère equation. In our original control problem, we fix the quadratic variation of the martingales, which corresponds to constraining the Frobenius norm of the volatility matrix in the martingale representation (1.1). In the control problem in [28], this constraint is replaced with a constraint on the determinant of the volatility matrix. In Section 5.5, we will show how this problem is related to our original control problem.

1.3 Outline of the thesis

In this thesis, we study control problems for martingales over two related control sets. We give a PDE characterisation of these general problems and demonstrate the relationship between them. Specialising to the case of a radially symmetric environment, we give explicit solutions and prove further properties of the control problems.

The main contributions of this thesis are the following:

1. We provide an explicit solution to a control problem for multidimensional martingales with fixed quadratic variation in a radially symmetric environment in Theorem 2.30, motivated by applications in robust finance.

2. We present two SDEs that do not admit a strong solution in Theorem 3.4 and Theorem 3.15. The SDEs arise naturally from the above problem of stochastic optimal control. Proving that a weak solution of the first SDE generates a Brownian filtration allows us to complete the proof of Theorem 2.30. The results on non-existence of strong solutions lead to Conjecture 3.5 that asserts that there is a gap between a Markov formulation of the control problem and a strong and weak formulation.

3. We characterise the value functions of two stochastic control problems in terms of viscosity solutions of a Monge-Ampère equation in Theorem 5.24 and Corollary 5.37. These control problems are equivalent to each other and related to the first problem that we study, as shown in Theorem 5.40.

The thesis is organised as follows.

1.3.1 Section 1.4: Preliminaries

In the remainder of this first chapter, we define the first stochastic control problem that we will study, and we prove some preliminary properties of the value function. We define both a strong and a weak formulation of the control problem. In Proposition 1.7 we show that, under certain assumptions, the weak and strong formulations are equivalent, by referring to a result from El Karoui and Tan's paper [20]. We show that the value function is semiconvex in Lemma 1.11 and deduce that it is locally Lipschitz in Corollary 1.13. We use this continuity property to prove directly that the value function satisfies a dynamic programming principle in Proposition 1.17. In Section 1.4.4 we introduce the HJB equation that will be key to solving the control problem explicitly in Chapter 2.

1.3.2 Chapter 2: Stochastic control of martingales in a radially symmetric environment

In this chapter we study the control problem defined in Section 1.4.1 in a radially symmetric environment. We construct a candidate for the value function explicitly. For a continuous radially symmetric cost function with sufficient regularity at the origin, we verify the candidate value function in Proposition 2.15 by showing that it is a viscosity solution of an HJB equation with appropriate boundary condition. This method of proof relies on Theorem 4.20, which shows that the value function is the unique viscosity solution of the HJB equation with a given boundary condition. We present the technical details of this theorem in Chapter 4.

To construct the value function, we reduce the problem to a one-dimensional switching problem. We identify the optimal switching points and observe that the principles of smooth and continuous fit are satisfied at these points, although the usual rationale for smooth fit is not always applicable.

We find that the optimal process switches between two behaviour regimes: the process either follows *radial motion*, moving as a Brownian motion on a line through the origin, or *tangential motion*, moving on a tangent to the current position. The latter behaviour results in a process with deterministically increasing radius. This property has been exploited by Fernholtz, Karatzas and Ruf in [25] to solve a relative arbitrage problem, as described in [41].

We extend the result of Proposition 2.15 to relax the regularity of the cost function and allow the cost to become infinite at the origin. In Theorem 2.30, we characterise the value function and determine under which growth conditions the value remains finite. For a regime of moderate growth of the cost function in dimension $d = 2$, we require a result on Brownian filtrations from Chapter 3 in order to complete the proof of Theorem 2.30. We also introduce a Markov formulation of the control problem and show that this is equivalent to the strong and weak formulations, with the possible exception of the moderate growth regime mentioned above. We consider this growth regime in the following chapter.

1.3.3 Chapter 3: SDEs with no strong solution arising from a problem of stochastic control

In this chapter, we present two SDEs that have no strong solution. Both of these SDEs arise naturally from the control problem of the previous chapter in dimension $d = 2$. We first consider an SDE describing tangential motion starting from the origin. We show that a weak solution of this SDE generates a Brownian filtration and use this to show that the strong value function is equal to the weak value function in the regime of moderate growth of the cost function. This completes the proof of Theorem 2.30 from the previous chapter.

In Theorem 3.4 we prove that this first SDE has no strong solution. This result leads us to formulate Conjecture 3.5, which asserts that there is a gap between the Markov value function and the strong and weak value functions at the origin in the moderate growth regime. We go on to consider martingales that approximate optimal behaviour in the sense that their associated value converges to the strong value function. We show in Theorem 3.15 that the SDEs describing one such possible approximating sequence also have no strong solution starting from the origin. This supports Conjecture 3.5.

The proofs of these theorems adapt techniques used in the study of the example of an SDE with no strong solution given by Tsirelson in [59]. In particular, we make use of the properties of circular Brownian motion that Émery and Schachermayer prove in their study of Tsirelson's equation in [21].

1.3.4 Chapter 4: Viscosity solutions of Hamilton-Jacobi-Bellman equations

The main result of this chapter is Theorem 4.20, which states that the value function defined in Section 1.4.1 is the unique viscosity solution of the associated HJB equation with appropriate boundary condition. This is the result that is used to find the explicit form of the value function in a radially symmetric environment in Chapter 2.

We follow standard arguments from Touzi's book [58] to show that the value function is a viscosity solution of the HJB equation. We establish uniqueness by proving a comparison principle, as is typical in the theory of viscosity solutions of PDEs. The standard proof of comparison for viscosity solutions given in Crandall, Ishii and Lions's User's Guide [13] requires coercivity of the differential operator in the zeroth derivative. Since coercivity does not hold for the HJB equation that we are considering, we use the perturbation argument from Section 5.C of [13]. The perturbation that we choose is the same as that suggested by Ishii and Lions for proving comparison for a Monge-Ampère equation in [36]. In this way, we adapt the standard proof of comparison to obtain a uniqueness result for the HJB equation in Proposition 4.19. We conclude the proof of Theorem 4.20 by verifying that the value function extends continuously to the boundary of the domain and satisfies the boundary condition pointwise. Our argument is based on similar results to those proved by Gaveau in [28] for a related control problem.

1.3.5 Chapter 5: Control problems related to a Monge-Ampère equation

In this final chapter, we study two further control problems that are related to the problem defined in Section 1.4.1. We show that the value function of each of these problems solves a Monge-Ampère equation and deduce that the two problems are equivalent.

First, we refer to the paper [24] where Feng and Jensen show that a Monge-Ampère equation has an equivalent formulation as an HJB equation in the sense of viscosity solutions. The stochastic control problem that is related to this HJB equation is not discussed in [24], and so we study this problem here. The control set is the same as for the problem defined in Section 1.4.1, but an additional cost is introduced to penalise martingales whose diffusion matrix has a small determinant. We show that the value function of this problem is the unique viscosity solution of a Monge-Ampère equation with appropriate boundary condition.

We then consider the control problem that Gaveau studies in [28]. In this problem, the constraint that the controlled martingale must have fixed quadratic variation is replaced with a constraint on the determinant of the diffusion matrix. In [28], Gaveau shows that the value function is a weak solution of a Monge-Ampère equation, but this result came prior to the introduction of viscosity solutions by Crandall and Lions in their 1983 paper [14]. We make use of the modern theory of viscosity solutions and prove that the value function is the unique viscosity solution of the Monge-Ampère equation with appropriate boundary condition.

Finally, in Theorem 5.40, we prove that the common value of the two problems introduced in this chapter is bounded below by the value function of the problem defined in Section 1.4.1. By means of examples, we show that equality may hold in this bound, although the inequality is strict in some simple examples.

1.4 Preliminaries

In this section we formulate a stochastic control problem over continuous multidimensional martingales with fixed quadratic variation. We prove preliminary results on convexity and continuity of the value function. We then prove a dynamic programming principle and heuristically derive the associated HJB equation.

1.4.1 Problem formulation

We now formulate the control problem precisely, as follows.

Fix $d \in \mathbb{N}$. We introduce the control set

$$U := \left\{\sigma \in \mathbb{R}^{d,d} \colon \operatorname{Tr}(\sigma\sigma^\top) = 1\right\}.$$

Let $D \subset \mathbb{R}^d$ be a domain and define the functions $f : D \to \mathbb{R}$ and $g : \partial D \to \mathbb{R}$, which we call the *running cost* and *boundary cost*, respectively. We make the following assumptions.

Assumption 1.1. Suppose that

1. The domain D is bounded;
2. The cost functions f and g are upper semicontinuous;
3. The running cost f is bounded above; i.e. $f \leq M$, for some $M \geq 0$;
4. The boundary cost g is bounded above; i.e. $g \leq K$ for some $K \geq 0$.

We introduce two variants of the control problem: a strong formulation and a weak formulation. We will show in Proposition 1.7 that, under Assumption 1.1, these two formulations are equivalent. We will relax our assumptions in Section 2.4 and show that this equivalence still holds. In Chapter 3, we will address a case where the theory of weak solutions of SDEs and Brownian filtrations is needed to prove the equivalence of weak and strong formulations.

Strong formulation

The strong formulation of the control problem is to find the strong value function $v^S : D \to \mathbb{R}$, which we now define as in [58]. In order to define the value function, we introduce the set of controls, which will be U-valued processes, and we describe the dynamics of the controlled martingales via the stochastic integral (1.3) below.

Let $(\Omega_0, \mathcal{F}, \mathbb{P}_0)$ be a probability space on which a d-dimensional Brownian motion B is defined with natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$.

Control: Define the set of controls

$$\mathcal{U} := \{U\text{-valued } \mathbb{F}\text{-progressively measurable processes}\} .$$

Dynamics: For any $x \in D$ and $\nu = (\nu_t)_{t\geq 0} \in \mathcal{U}$, define X^ν by the stochastic integral

$$X_t^\nu = x + \int_0^t \nu_s \,\mathrm{d}B_s, \quad t \geq 0, \tag{1.3}$$

and define the associated exit time from the domain by

$$\tau := \inf \{t \geq 0\colon X_t^\nu \notin D\} .$$

Example 1.2. Let $\sigma : D \to U$ be Lipschitz. Then there is a unique strong solution X^σ of the SDE

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t, \quad X_0 = x.$$

Define $\nu_t = \sigma(X_t^\sigma)$, for all $t \geq 0$. Then $\nu \in \mathcal{U}$ and, for any $t \geq 0$, $X_t^\sigma = x + \int_0^t \nu_s \,\mathrm{d}B_s$.

Notation. For a process Y starting from a point y and a functional F of the path of Y, we denote the expectation with respect to the law of Y by

$$\mathbb{E}^y [F(Y)] := \mathbb{E} [F(Y)|Y_0 = y] .$$

Value function: We define the *strong value function* $v^S : D \to \mathbb{R}$ by

$$v^S(x) := \inf_{\nu\in\mathcal{U}} \mathbb{E}^x \left[\int_0^\tau f(X_s^\nu)\,\mathrm{d}s + g(X_{\tau^\nu}^\nu)\right] . \tag{1.4}$$

Remark 1.3. Note that, for any $\nu \in \mathcal{U}$, the quadratic variation of a controlled martingale X^ν is given by

$$\langle X^\nu \rangle_t = \int_0^t \mathrm{Tr}(\nu_s \nu_s^\top)\,\mathrm{d}s = t,$$

for any $t \geq 0$, by the definition of the control set U.

Definition 1.4. We say that a process X has *unit quadratic variation* if its quadratic variation is given by

$$\langle X \rangle_t = t, \quad \text{for all} \quad t \geq 0.$$

A martingale with unit quadratic variation has the property that the expected exit time of the martingale from a ball is fixed. This gives a bound on the expected exit time from the domain D as follows.

Notation. Let $R > 0$ and $x \in \mathbb{R}^d$. We denote the d-dimensional open ball of radius R centred at x by

$$B_R(x) := \{y \in \mathbb{R}^d \colon\ |y - x| < R\}.$$

For a given process X, we denote the first exit time from $B_R(0)$ by

$$\tau_R := \inf \left\{t \geq 0 \colon X_t \notin B_R(0)\right\}.$$

Proposition 1.5. *Let X be a continuous martingale with initial condition $X_0 = x \in D$, and suppose that X has unit quadratic variation. Fix $R > 0$. Then*

$$\mathbb{E}^x[\tau_R] = R^2 - |x|^2 .$$

Moreover, defining $\tau := \inf \{t \geq 0 \colon X_t \notin D\}$, we have the bound

$$\mathbb{E}^x[\tau] \leq \operatorname{diam}(D)^2 - |x|^2 < \infty.$$

Proof. Applying Itô's formula to $|X_{\tau_R}|^2$ and taking expectations, we find that

$$\mathbb{E}^x\left[|X_{\tau_R}|^2\right] - |x|^2 = \mathbb{E}^x\left[\langle X \rangle_{\tau_R}\right] = \mathbb{E}^x[\tau_R],$$

since X is a martingale and has unit quadratic variation. Therefore, by continuity of the paths of X, we have

$$\mathbb{E}^x[\tau_R] = R^2 - |x|^2 .$$

Now set $R = \operatorname{diam}(D)$ so that $D \subseteq B_R(x)$. Then the inequality $\tau \leq \tau_R$ holds pointwise and, in particular,

$$\mathbb{E}^x[\tau] \leq \mathbb{E}^x[\tau_R] = \operatorname{diam}(D)^2 - |x|^2 < \infty,$$

as required. □

Weak formulation

We now introduce the weak formulation of the control problem, following El Karoui and Tan in [20]. The problem is to find the weak value function $v^W : D \to \mathbb{R}$, which we define below. In the weak formulation, the controls will take values in a set of probability measures, and the dynamics of the controlled martingales will be described as solutions of a local martingale problem.

Define the space of continuous paths $\Omega := C([0,\infty), \mathbb{R}^d)$ and denote the set of Borel measurable functions $\nu : \mathbb{R}_+ \to U$ by $\mathcal{B}(\mathbb{R}_+, U)$. Then set $\overline{\Omega} = \Omega \times \mathcal{B}(\mathbb{R}_+, U)$ and denote an element of $\overline{\Omega}$ by $\overline{\omega} = (\omega, u)$. Define the canonical process $\overline{X} = (X, \nu)$ on $\overline{\Omega}$ by $X_t(\overline{\omega}) = \omega_t$, for each $t \geq 0$, and $\nu(\overline{\omega}) = u$. We define the canonical filtration as in [20]. For $\phi \in C_b(\mathbb{R}_+ \times U)$, $s \geq 0$, define

$$M_s(\phi) := \int_0^s \phi(r, \nu_r)\,\mathrm{d}r.$$

Then define the canonical filtration $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t\geq 0}$ by

$$\overline{\mathcal{F}}_t := \sigma\left\{(X_s, M_s(\phi))\colon \phi \in C_b(\mathbb{R}_+ \times U), s \leq t\right\}, \quad t \geq 0.$$

Control: Let $\mathbb{M}$ be the set of probability measures on the set $\overline{\Omega}$. For each $x \in D$, let

$$\mathbb{M}_x = \left\{\mathbb{P} \in \mathbb{M}\colon \mathbb{P}(X_0 = x) = 1\right\}.$$

Dynamics: For $x \in D$, define

$$\begin{aligned}\mathcal{P}_x := \{\mathbb{P} \in \mathbb{M}_x : \quad & t \mapsto \phi(X_t) - \phi(X_0) - \frac{1}{2}\int_0^t \mathrm{Tr}\left(D^2\phi(X_s)\nu_s\nu_s^\top\right)\mathrm{d}s \\ & \text{is a } (\overline{\mathbb{F}}, \mathbb{P})\text{-local martingale for all } \phi \in C^2(\mathbb{R}^d)\},\end{aligned}$$

and let $\tau = \inf\{t \geq 0\colon X_t \notin D\}$.

Notation. For a process Y, a functional F of the path of Y, and a probability measure $\mathbb{P}$ on path space, we denote the expectation with respect to $\mathbb{P}$ by

$$\mathbb{E}^{\mathbb{P}}\left[F(Y)\right].$$

Value function: Define the *weak value function* $v^W : D \to \mathbb{R}$ by

$$v^W(x) = \inf_{\mathbb{P}\in\mathcal{P}_x} \mathbb{E}^{\mathbb{P}}\left[\int_0^\tau f(X_s)\,\mathrm{d}s + g(X_\tau)\right].$$

Remark 1.6. A measure $\mathbb{P} \in \mathcal{P}_x$ is a solution of a local martingale problem, as defined in Definition 4.5 of [38, Chapter 5]. As shown in Problem 4.3 and Proposition 4.6 of [38, Chapter 5], there is a correspondence between solutions of a local martingale problem and weak solutions of an SDE. In our set up, a measure $\mathbb{P} \in \mathcal{P}_x$ corresponds to a weak solution of the SDE (1.3) with initial distribution δ_x.

We will now show that, under Assumption 1.1, the weak and strong value functions are equal, by refering to Theorem 4.5 of [20].

Proposition 1.7. *Suppose that Assumption 1.1 holds. Then the weak and strong formulations of the control problem are equivalent; i.e.* $v^S = v^W$ *in* D.

Proof. We apply Theorem 4.5 of [20], which gives conditions for equality of the weak and strong value functions. Define a function $\Phi : \Omega \to \mathbb{R}$ by

$$\Phi(\omega) = \int_0^{\tau(\omega)} f(X_s(\omega))\,\mathrm{d}s + g(X_{\tau(\omega)}(\omega)),$$

and fix $x \in D$. Then, by Theorem 4.5 of [20], it is sufficient to show that Φ is upper semicontinuous and bounded above by some random variable ξ that is uniformly integrable under the family of probability measures $\mathcal{P}_x$.

Under our assumptions, $f : D \to \mathbb{R}$ and $g : \partial D \to \mathbb{R}$ are upper semicontinuous and so Φ is also upper semicontinuous.

Since we have also assumed that f and g are bounded above, we have the bound

$$\Phi(\omega) \leq M\tau(\omega) + K =: \xi(\omega).$$

Fix $\mathbb{P} \in \mathcal{P}_x$ and let (X, ν) have joint law $\mathbb{P}$. Then the process X has unit quadratic variation, and so by Proposition 1.5, we have the bound

$$\mathbb{E}^{\mathbb{P}}[\tau] \leq \mathrm{diam}(D)^2 - |x|^2 .$$

Hence

$$\mathbb{E}^{\mathbb{P}}[\xi] \leq M\,\mathrm{diam}(D)^2 - |x|^2 + K < \infty,$$

independently of the choice of measure $\mathbb{P}$. Therefore ξ is uniformly integrable under $\mathcal{P}_x$.

We apply Theorem 4.5 of [20] to conclude that $v^S(x) = v^W(x)$. □

With the result of Proposition 1.7 in hand, we will write $v = v^W = v^S$ and refer to v as the *value function*. We choose to work with the strong formulation of the control problem in the following chapters, unless we explicitly refer to the weak formulation.

1.4.2 Dynamic programming principle

The approach that we take to solving the stochastic optimal control problem defined above is to use a 'guess and verify' method. First, we conjecture an optimal strategy for a particular problem. We then calculate the value associated to following this strategy, as a function of the starting point of the controlled process. In this section, we introduce the dynamic programming principle, which provides a necessary and sufficient condition for a given function to be equal to the value function. The dynamic programming principle is a key technique in the study of stochastic control problems, as described, for example, by Fleming and Soner in Section 7 of [26, Chapter III] and by Touzi in Section 3.2 of [58]. Also known as the *Bellman principle*, the dynamic programming principle for stochastic optimisation problems dates back to the 1952 work of Bellman in [6] and [7]. We will use the dynamic programming principle in Example 2.1 and Example 2.6 of Chapter 2 to verify that a conjectured optimal strategy is indeed optimal.

We will also use the dynamic programming principle to derive the HJB equation, which is a nonlinear PDE that the value function must satisfy in a certain weak sense. In Chapter 4, we will develop the theory of viscosity solutions and see that this is the appropriate notion of weak solution in this context. Having proved uniquenss of viscosity solutions, we will deduce in Theorem 4.20 that a given function is equal to the value function if and only if it is a viscosity solution of the HJB equation with appropriate boundary condition. In Section 2.3, we will find it convenient to use this PDE characterisation to verify a candidate value function, rather than working directly with the dynamic programming principle.

We now define the dynamic programming principle, following Touzi's definition of the classical dynamic programming principle in Section 3.2 of [58].

Suppose that we follow a suboptimal strategy $\nu \in \mathcal{U}$, starting from position $x \in D$ at time 0, up until a stopping time ρ. Consider the minimum expected cost when starting from the position at time ρ, plus the cost accrued up until time ρ. We expect this total cost to be greater than the minimum expected cost when starting from position x at time 0. In the case that an optimal strategy exists and we choose to follow this strategy, we would expect the above two quantities to be equal. This means that we expect the value function v to satisfy the following principle.

Definition 1.8 (Dynamic programming principle)**.** We say that a *dynamic programming principle* holds for the value function v if, for any $x \in D$, and for any stopping time θ such that $\theta \in [0,\tau]$ almost surely, v satisfies

$$v(x) = \inf_{\sigma\in\mathcal{U}} \mathbb{E}^x \left[\int_0^\theta f(X_s^\sigma)\,\mathrm{d}s + v(X_\theta^\sigma) \right]. \tag{1.5}$$

Remark 1.9. If there exists an optimal control $\sigma^\star \in \mathcal{U}$, then this is equivalent to stating that

$$v(X_t^\sigma) + \int_0^t f(X_s^\sigma)\,\mathrm{d}s \quad \text{is} \quad \begin{cases} \text{a submartingale,} & \text{for all} \quad \sigma \in \mathcal{U}, \\ \text{a martingale,} & \text{for} \quad \sigma = \sigma^\star. \end{cases}$$

There are many references in the literature where a dynamic programming principle is proved. For example, Bouchard and Touzi prove a weak dynamic programming principle in [9] in a very general setup. A dynamic programming principle is also proved in [26] and [58].

The usual difficulty in proving a dynamic programming principle is that the value function is not necessarily continuous, as stated in [9]. However, in our case, we will be able to show a priori that the value function is in fact continuous. We can then exploit this property to prove the dynamic programming principle directly.

1.4.3 Proof of a dynamic programming principle

In this section, we give a direct proof of the dynamic programming principle (1.5) for the value function defined in Section 1.4.1.

We first prove that the value function is semiconvex and hence locally Lipschitz. We say that a function is semiconvex if it can be transformed into a convex function by the addition of a quadratic term. We give the following definition of semiconvexity, as in Section 6.7 of [58].

Definition 1.10. Let $\lambda > 0$. We say that a function $f : \mathbb{R}^d \to \mathbb{R}$ is λ-*semiconvex* if the function $f^\lambda : \mathbb{R}^d \to \mathbb{R}$, defined by

$$f^\lambda(x) = f(x) + \frac{\lambda}{2}\,|x|^2, \quad x \in \mathbb{R}^d,$$

is convex.

Lemma 1.11. *Suppose that Assumption 1.1 holds and the domain D is strictly convex. Then v is $2M$-semiconvex in D, where $M \geq 0$ is such that $f \leq M$ in D.*

Proof. Let $x_0, x_1 \in D$. Consider a martingale starting from a point y on the straight line connecting these two points; i.e. $y = \lambda x_0 + (1-\lambda)x_1 \in D$, for some $\lambda \in (0,1)$. Let $\varepsilon > 0$ and define the control $\sigma^\star$ as follows.

First, define the constant control

$$\overline{\sigma} \equiv \frac{1}{|x_1 - x_0|} \Big[x_1 - x_0; \quad 0; \quad \dots; \quad 0\Big] \in U.$$

Let W be the first component of the Brownian motion B. Then, for $t \geq 0$,

$$X_t^{\overline{\sigma}} = y + \frac{x_1 - x_0}{|x_1 - x_0|} W_t.$$

Set $\sigma_t^\star = \overline{\sigma}$ for $t \leq H_{x_0,x_1}$, where H_{x_0,x_1} is the first hitting time of either x_0 or x_1. The controlled process $X^{\sigma^\star}$ runs as a Brownian motion on the line connecting the points x_0 and x_1 until hitting one of these points.

For $i \in \{0, 1\}$, let $\sigma^{i,\varepsilon} \in \mathcal{U}$ satisfy

$$v(x_i) > \mathbb{E}^{x_i} \left[\int_0^\tau f(X_s^{\sigma^{i,\varepsilon}}) \,\mathrm{d}s + g(X_\tau^{\sigma^{i,\varepsilon}}) \right] - \varepsilon.$$

After the first hitting time of x_0 or x_1, we set the controlled process to follow one of these ε-optimal controls, choosing $\sigma^{0,\varepsilon}$ if the process hits x_0 before x_1, and $\sigma^{1,\varepsilon}$ otherwise. By construction, we have that $\sigma^\star \in \mathcal{U}$.

Let us write H_{x_0} for the first hitting time of x_0 and H_{x_1} for the first hitting time of x_1. We can condition on the value of the controlled process at the hitting time H_{x_0,x_1} to find that

$$\begin{aligned}
v(y) &\leq \mathbb{E}^y \left[\int_0^\tau f(X_s^{\sigma^\star}) \,\mathrm{d}s + g(X_\tau^{\sigma^\star}) \right] \\
&\leq M\mathbb{E}^y[H_{x_0,x_1}] + \mathbb{E}^{x_0} \left[\int_0^\tau f(X_s^{\sigma^{0,\varepsilon}}) \,\mathrm{d}s \right] \mathbb{P}^y[H_{x_0} < H_{x_1}] \\
&\quad + \mathbb{E}^{x_1} \left[\int_0^\tau f(X_s^{\sigma^{1,\varepsilon}}) \,\mathrm{d}s \right] \mathbb{P}^y[H_{x_1} < H_{x_0}] + \mathbb{E}^y \left[g(X_\tau^{\sigma^\star}) \right] \\
&= M\mathbb{E}^y[H_{x_0,x_1}] + \mathbb{E}^{x_0} \left[\int_0^\tau f(X_s^{\sigma^{0,\varepsilon}}) \,\mathrm{d}s + g(X_\tau^{\sigma^{0,\varepsilon}}) \right] \mathbb{P}^y[H_{x_0} < H_{x_1}] \\
&\quad + \mathbb{E}^{x_1} \left[\int_0^\tau f(X_s^{\sigma^{1,\varepsilon}}) \,\mathrm{d}s + g(X_\tau^{\sigma^{1,\varepsilon}}) \right] \mathbb{P}^y[H_{x_1} < H_{x_0}] \\
&< M\mathbb{E}^y[H_{x_0,x_1}] + v(x_0)\mathbb{P}^y[H_{x_0} < H_{x_1}] + v(x_1)\mathbb{P}^y[H_{x_1} < H_{x_0}] + 2\varepsilon,
\end{aligned}$$

using the upper bound on f and the definition of the ε-optimal controls $\sigma^{\varepsilon,0}$ and $\sigma^{\varepsilon,1}$. Calculating that

$$\mathbb{P}^y[H_{x_0} < H_{x_1}] = \lambda,$$

and

$$\begin{aligned}\mathbb{E}^y[H_{x_0,x_1}] &= \lambda\,|y-x_0|^2 + (1-\lambda)\,|x_1-y|^2 \\ &\leq \lambda\,|x_0|^2 + \lambda\,|y|^2 + (1-\lambda)\,|x_1|^2 + (1-\lambda)\,|y|^2 \\ &= \lambda\,|x_0|^2 + (1-\lambda)\,|x_1|^2 - |y|^2\,,\end{aligned}$$

we find that

$$v(y) < M\left(\lambda\,|x_0|^2 + (1-\lambda)\,|x_1|^2 - |y|^2\right) + \lambda v(x_0) + (1-\lambda)v(x_1) + 2\varepsilon.$$

So, taking the limit as $\varepsilon \to 0$, we have

$$v(y) \leq M\left(\lambda\,|x_0|^2 + (1-\lambda)\,|x_1|^2 - |y|^2\right) + \lambda v(x_0) + (1-\lambda)v(x_1).$$

Therefore

$$\begin{aligned}v(y) + M\,|y|^2 &\leq M\left(\lambda\,|x_0|^2 + (1-\lambda)\,|x_1|^2\right) + \lambda v(x_0) + (1-\lambda)v(x_1) \\ &= \lambda\left(v(x_0) + M\,|x_0|^2\right) + (1-\lambda)\left(v(x_1) + M\,|x_1|^2\right).\end{aligned}$$

This shows that the map $x \mapsto v(x) + M\,|x|^2$ is a convex function. Hence v is $2M$-semiconvex, as required. □

Remark 1.12. In particular, in the case that the cost function f is negative, we have shown that the value function v is convex. An intuitive justification for this is that, since it is favourable at any point $x \in D$ to run on for a short time t, we expect

$$\mathbb{E}^x[v(X_t)] \leq v(x) = v(\mathbb{E}^x[X_t]),$$

by the martingale property of X. Appealing to Jensen's inequality, this suggests that v should be convex.

If f is bounded above by some M, then running on for a short time t has a cost of at most Mt. By the unit quadratic variation condition, the process $t \mapsto |X_t|^2 - t$ is a martingale, and so we expect

$$\begin{aligned}\mathbb{E}^x[v(X_t) + M\,|X_t|^2] &= \mathbb{E}^x[v(X_t)] + Mt \\ &\leq v(x) = v(\mathbb{E}^x[X_t]) \\ &\leq v(\mathbb{E}^x[X_t]) + M\,|\mathbb{E}[X_t]|^2\,,\end{aligned}$$

using the martingale property of X in the penultimate line. Referring again to Jensen's inequality, we then expect the map $x \mapsto v(x) + M\,|x|^2$ to be convex.

For unbounded f, there is no reason to expect any convexity result for v, as running on for even a short time could incur an unbounded cost.

Suppose now that $v(x) > -\infty$ for all $x \in D$. Then Theorem 10.4 of [52] tells us that when v is convex, v is locally Lipschitz. In fact, we will see that semiconvexity is sufficient to show that v is locally Lipschitz in D.

Corollary 1.13. *Suppose that Assumption 1.1 holds, D is strictly convex, and $v(x) > -\infty$ for all $x \in D$. Then v is locally Lipschitz in D.*

Proof. By Lemma 1.11, we have that v is $2M$-semiconvex in D, for $M \geq 0$ such that $f \leq M$ in D; i.e. the function $v^M : \mathbb{R}^d \to \mathbb{R}$, defined by $v^M(x) = v(x) + M\,|x|^2$ is convex in D. Then, since the condition $v > -\infty$ implies that $v^M > -\infty$, we can apply Theorem 10.4 of [52] to see that v^M is locally Lipschitz.

It is also known that $x \mapsto |x|^2$ is locally Lipschitz. Therefore, since

$$v(x) = v^M(x) - M\,|x|^2\,, \quad x \in D,$$

v is locally Lipschitz in D. □

Remark 1.14. In the proof of continuity, we exclude the case where v takes the value $-\infty$ at some point. In this case, since we have a finite boundary condition, we would not expect the value function to be continuous on the whole domain.

Remark 1.15. Touzi gives an intuitive justification of the dynamic programming principle in Section 3.2.1 of [58], which we adapt here. We note that continuity of the value function enables us to make this argument rigorous. In particular, we know a priori that v is measurable and equal to its upper and lower semicontinuous envelopes. Moreover, the value function does not depend on time, and is of the form considered in Chapter 2 of [58] with coefficient $k \equiv 0$. These properties further simplify the proof.

We now prove the dynamic programming principle under the following strengthening of Assumption 1.1.

Assumption 1.16. Suppose that Assumption 1.1 holds and, moreover, the domain D is strictly convex and the value function v satisfies $v(x) > -\infty$, for any $x \in D$.

Proposition 1.17. *Suppose that Assumption 1.16 is satisfied. Then the following dynamic programming principle holds.*

For any $x \in D$ and for any stopping time θ such that $\theta \in [0, \tau]$ almost surely, v satisfies

$$v(x) = \inf_{\nu \in \mathcal{U}} \mathbb{E}^x \left[\int_0^\theta f(X_s^\nu)\, \mathrm{d}s + v(X_\theta^\nu) \right]. \tag{1.5}$$

Proof. Under the stated assumptions, we have that v is continuous by Corollary 1.13.

Define $\mathcal{I} : D \times \mathcal{U} \to \mathbb{R}$ by

$$\mathcal{I}(x, \nu) := \mathbb{E}^x \left[\int_0^\tau f(X_s^\nu) \, \mathrm{d}s + g(X_\tau^\nu) \right],$$

for each $x \in D$ and $\nu \in \mathcal{U}$, so that

$$v(x) = \inf_{\nu \in \mathcal{U}} \mathcal{I}(x, \nu).$$

Fix $x \in D$ and $\nu \in \mathcal{U}$. Let θ be a stopping time such that $\theta \in [0, \tau]$ almost surely, and fix $\omega \in \Omega$. By an argument similar to that in the proof of Proposition 5.4 of [9], there exists a control $\tilde{\nu}^\omega \in \mathcal{U}$ such that

$$\begin{aligned} \mathbb{E}^x \left[\int_0^\tau f(X_s^\nu) \, \mathrm{d}s + g(X_\tau^\nu) \,\middle|\, \mathcal{F}_{\theta(\omega)} \right] &= \int_0^{\theta(\omega)} f(X_s^\nu) \, \mathrm{d}s + \mathcal{I}\left(X_{\theta(\omega)}^\nu(\omega), \tilde{\nu}^\omega \right) \\ &\geq \int_0^{\theta(\omega)} f(X_s^\nu) \, \mathrm{d}s + v\left(X_{\theta(\omega)}^\nu(\omega) \right). \end{aligned}$$

Then, by the tower property for conditional expectation,

$$\mathcal{I}(x, \nu) \geq \mathbb{E}^x \left[\int_0^\theta f(X_s^\nu) \, \mathrm{d}s + v\left(X_\theta^\nu \right) \right].$$

By Assumption 1.16 and Corollary 1.13, the functions f, g and v are measurable, and so the above expressions are all well-defined. Taking the infimum over $\nu \in \mathcal{U}$ on both sides yields

$$v(x) \geq \inf_{\nu \in \mathcal{U}} \mathbb{E}^x \left[\int_0^\theta f(X_s^\nu) \, \mathrm{d}s + v(X_\theta^\nu) \right].$$

To prove the inequality in the other direction, fix an arbitrary $\mu \in \mathcal{U}$ and $\varepsilon > 0$. We wish to take $\nu^\varepsilon \in \mathcal{U}$ such that

$$\mathcal{I}(X_\theta^{\nu^\varepsilon}, \nu^\varepsilon) \leq v(X_\theta^{\nu^\varepsilon}) + \varepsilon,$$

and $\nu^\varepsilon = \mu$ on $[0, \theta]$. However, it is not clear a priori that there exists such a ν^ε that has the required measurability properties. We will return to this issue of measurable selection below.

Supposing that we can take such a ν^ε, then

$$\begin{aligned} v(x) \le \mathcal{I}(x,\nu^\varepsilon) &= \mathbb{E}^x \left[\int_0^\theta f(X_s^{\nu^\varepsilon})\,\mathrm{d}s + \mathcal{I}(X_\theta^{\nu^\varepsilon},\nu^\varepsilon) \right] \\ &\le \mathbb{E}^x \left[\int_0^\theta f(X_s^{\mu})\,\mathrm{d}s + v(X_\theta^{\mu}) \right] + \varepsilon. \end{aligned}$$

If we now take the supremum over all $\mu \in \mathcal{U}$ and let $\varepsilon \to 0$, we get

$$v(x) \le \inf_{\mu\in\mathcal{U}} \mathbb{E}^x \left[\int_0^\theta f(X_s^{\mu})\,\mathrm{d}s + v(X_\theta^{\mu}) \right].$$

Since we know a priori that v is continuous, by Corollary 1.13, we are able to make the required measurable selection argument directly. For any $x \in D$, continuity of v implies that there exists $\delta(x) > 0$ such that

$$v(x) \ge v(y) - \frac{\varepsilon}{3}, \quad \text{for all} \quad y \in B_{\delta(x)}(x). \tag{1.6}$$

Since D is open, we can choose $\delta(x)$ sufficiently small that

$$B_{\delta(x)^{\frac{1}{2}}}(x) \subset D.$$

Moreover, fix

$$c := \sup_{\xi\in D} \mathbb{E}^\xi[\tau],$$

and note that $c \le \operatorname{diam}(D)^2 < \infty$, by Proposition 1.5. Now choose $\delta(x)$ sufficiently small that

$$\delta(x)^{\frac{1}{2}} \le \frac{\varepsilon}{3\left[(c+2)M + K\right]} \wedge 1. \tag{1.7}$$

Note that we have the following open cover of $D \subset \mathbb{R}^d$:

$$D \subset \bigcup_{x\in D} B_{\delta(x)}(x).$$

There is then a countable subcover given by

$$D \subset \bigcup_{\alpha\in\mathbb{N}} B_{\delta_\alpha}(x_\alpha),$$

for points $x_\alpha \in D$ and radii $\delta_\alpha = \delta(x_\alpha) > 0$, indexed by $\alpha \in \mathbb{N}$.

Fix $\omega \in \Omega$. For $t < \theta(\omega)$, let $\nu_t^\varepsilon(\omega) = \mu_t(\omega)$. For $t \ge \theta(\omega)$, ν_t^ε must depend on the position of the controlled process at the stopping time $\theta(\omega)$, which we denote

$z := X^{\nu^\varepsilon}_{\theta(\omega)}(\omega)$.

Since the set $\{B_{\delta_\alpha}(x_\alpha)\colon \alpha \in \mathbb{N}\}$ is a countable cover of D, we have that $z \in B_{\delta_\alpha}(x_\alpha)$, for some fixed $\alpha \in \mathbb{N}$. Let us denote $\delta = \delta_\alpha$ for convenience in what follows, and define $\eta := \delta^{\frac{1}{2}}$.

Note that $B_\eta(x_\alpha) \subset D$, and define H_η to be the hitting time

$$H_\eta := \inf\{t \geq \theta\colon X^{\nu^\varepsilon}_t \in \partial B_\eta(x_\alpha) \cup \{x_\alpha\}\}.$$

Define

$$\sigma^1 = \frac{z - x_\alpha}{|z - x_\alpha|},$$

and let

$$\sigma = \Big[\sigma^1; \quad 0; \quad \dots; \quad 0\Big] \in \mathbb{R}^{d,d}.$$

Set $\nu^\varepsilon_t(\omega) = \sigma$ for $t \in [\theta(\omega), H_\eta]$. Then the controlled process X^{ν^ε} moves on the straight line connecting z to x_α until either hitting x_α or leaving the ball $B_\eta(x_\alpha)$.

We now calculate the probability p of hitting x_α before leaving the ball of radius η. Define $\lambda := |z - x_\alpha|$. Then

$$x_\alpha = z - \lambda\sigma^1,$$

and the vector σ^1 intersects the boundary of the ball of radius η at the point

$$y = z + (\eta - \lambda)\sigma^1.$$

We now have

$$X^{\nu^\varepsilon}_{H_\eta} = \begin{cases} x_\alpha, & \text{with probability} \quad p, \\ y, & \text{with probability} \quad 1 - p. \end{cases}$$

We can calculate the hitting probabilities

$$p = \frac{\eta - \lambda}{\eta} < 1 \quad \text{and} \quad 1 - p = \frac{\lambda}{\eta} < \frac{\delta}{\eta} = \delta^{\frac{1}{2}}, \tag{1.8}$$

noting that $\lambda = |z - x_\alpha| < \delta$. The expected hitting time is then

$$\begin{aligned} \mathbb{E}^z[H_\eta] &= (1-p)(\eta - \lambda)^2 + p\lambda^2 \\ &< (1-p)\eta^2 + p\lambda^2 \\ &< \delta\eta + \delta^2 = \delta^{\frac{3}{2}} + \delta^2. \end{aligned} \tag{1.9}$$

The bound (1.7) on δ then implies that the process hits x_α with high probability in a short time.

Let $\nu^{\varepsilon,\alpha}$ be an $\frac{\varepsilon}{3}$-optimal strategy starting from x_α, so that

$$\mathcal{I}(x_\alpha, \nu^{\varepsilon,\alpha}) \leq v(x_\alpha) + \frac{\varepsilon}{3}. \tag{1.10}$$

If, at the hitting time H_η, we have $X^{\nu^\varepsilon}_{H_\eta} = x_\alpha$, then set the controlled process X^{ν^ε} to follow the control $\nu^{\varepsilon,\alpha}$ from this time onwards. Otherwise set X^{ν^ε} to follow the constant control $\nu^\varepsilon_t = \sigma$, for $t \geq H_\eta$. We can then write

$$\begin{aligned}
\mathcal{I}(z, \nu^\varepsilon) &= \mathbb{E}^z \left[\int_0^\tau f(X^{\nu^\varepsilon}_s) \, \mathrm{d}s + g(X^{\nu^\varepsilon}_\tau) \right] \\
&= \mathbb{E}^z \left[\int_0^{H_\eta} f(X^\sigma_s) \, \mathrm{d}s \right] + p\mathbb{E}^{x_\alpha} \left[\int_0^\tau f(X^{\nu^{\varepsilon,\alpha}}_s) \, \mathrm{d}s + g(X^{\nu^{\varepsilon,\alpha}}_\tau) \right] \\
&\quad + (1-p)\mathbb{E}^y \left[\int_0^\tau f(X^\sigma_s) \, \mathrm{d}s + g(X^\sigma_\tau) \right].
\end{aligned}$$

We now bound this expression term by term. Making use of the estimate (1.9) on $\mathbb{E}^z[H_\eta]$, along with the upper bound $f \leq M$ from Assumption 1.16, we see that the first term is bounded by

$$\begin{aligned}
\mathbb{E}^z \left[\int_0^{H_\eta} f(X^\sigma_s) \, \mathrm{d}s \right] &\leq M\mathbb{E}^z[H_\eta] \\
&< M(\delta^{\frac{3}{2}} + \delta^2).
\end{aligned}$$

Using the estimate (1.8) on $1-p$, and the upper bounds $f \leq M$ and $g \leq K$ from Assumption 1.16, we see that the final term is bounded by

$$\begin{aligned}
(1-p)\mathbb{E}^y \left[\int_0^\tau f(X^\sigma_s) \, \mathrm{d}s + g(X^\sigma_\tau) \right] &\leq (1-p) \left(M\mathbb{E}^y[\tau] + K \right) \\
&< \delta^{\frac{1}{2}}(cM + K),
\end{aligned}$$

recalling that $c = \sup_{\xi \in D} \mathbb{E}^\xi[\tau]$.

Combining the bounds on these terms, and using the fact that $p < 1$, we have that

$$\begin{aligned}
\mathcal{I}(z, \nu^\varepsilon) &< \mathbb{E}^{x_\alpha} \left[\int_0^\tau f(X^{\nu^{\varepsilon,\alpha}}_s) \, \mathrm{d}s + g(X^{\nu^{\varepsilon,\alpha}}_\tau) \right] + M(\delta^{\frac{3}{2}} + \delta^2) + \delta^{\frac{1}{2}}(cM + K) \\
&= \mathcal{I}(x_\alpha, \nu^{\varepsilon,\alpha}) + M(\delta^{\frac{3}{2}} + \delta^2) + \delta^{\frac{1}{2}}(cM + K).
\end{aligned} \tag{1.11}$$

Now, by the bound (1.7),

$$M(\delta^{\frac{3}{2}} + \delta^2) + \delta^{\frac{1}{2}}(cM + K) < \delta^{\frac{1}{2}} \left[(c+2)M + K \right] \leq \frac{\varepsilon}{3}.$$

Inserting this bound into inequality (1.11), together with the property of $\nu^{\varepsilon,\alpha}$ from (1.10), we get

$$\begin{aligned}\mathcal{I}(z,\nu^{\varepsilon}) &< \mathcal{I}(x_{\alpha},\nu^{\varepsilon,\alpha}) + \frac{\varepsilon}{3}\\ &\leq v(x_{\alpha}) + \frac{\varepsilon}{3} + \frac{\varepsilon}{3}.\end{aligned}$$

From the property of δ in (1.6), we also have that

$$v(x_{\alpha}) \leq v(z) + \frac{\varepsilon}{3},$$

since $z \in B_{\delta}(x_{\alpha})$. Hence

$$\mathcal{I}(z,\nu^{\varepsilon}) < v(z) + \varepsilon.$$

Note that the strategy ν^{ε} defined in this way depends only on the index α, for which $X_{\theta(\omega)}(\omega) \in B_{\delta_{\alpha}}(x_{\alpha})$. Since the cover of D made up of balls of this form is countable, we conclude that the strategy has the required measurability properties.

□

Having proved that the dynamic programming principle holds, we now go on to derive the Hamilton-Jacobi-Bellman equation.

1.4.4 A Hamilton-Jacobi-Bellman equation

The control problem defined in Section 1.4.1 is associated to a second order PDE known both as a *Hamilton-Jacobi-Bellman (HJB) equation* and a *dynamic programming equation*, as described in Section 3.3 of [58] and in Section 7 of [26, Chapter VII]. Specifically, we expect the value function v to satisfy the boundary value problem

$$\begin{cases} -\frac{1}{2}\inf_{\sigma\in U}\operatorname{Tr}\left(D^2 v\sigma\sigma^{\top}\right) = f, & \text{in} \quad D,\\ v = g, & \text{on} \quad \partial D,\end{cases} \tag{1.12}$$

where $D^2 v$ is the Hessian of v. The HJB equation in (1.12) is a fully nonlinear degenerate elliptic PDE.

In general,we expect v to solve (1.12) in the viscosity sense, as defined in Chapter 4. Here, we impose sufficient smoothness conditions on the value function v and the running cost f such that v should be a classical solution of the HJB equation in (1.12). We will demonstrate the subsolution property below and, under further assumptions on the domain, we will show that the boundary condition is satisfied.

Notation. For a twice continuously differentiable function $u : D \to \mathbb{R}$, we denote the gradient Du and the Hessian D^2u.

Suppose that Assumption 1.16 holds and, moreover, that v is twice continuously differentiable and f is continuous. Under these conditions, we will show that the value function v is a classical subsolution of the HJB equation in (1.12), following the proof of Proposition 3.4 of [58].

Fix $x \in D$, let $\sigma \in U$ and define X^σ to be the process following the constant control that is equal to σ; that is

$$X_t^\sigma = x + \sigma B_t, \quad t \geq 0.$$

Under the assumption that v is twice continuously differentiable, we can apply Itô's formula to find that

$$\mathrm{d}v(X_t^\sigma) = Dv(X_t^\sigma)\sigma \,\mathrm{d}B_t + \frac{1}{2}\operatorname{Tr}\left(D^2v(X_t^\sigma)\sigma\sigma^\top\right)\mathrm{d}t. \tag{1.13}$$

Fix some $\delta > 0$ such that $B_\delta(x) \subset D$, define $\theta := \inf\{t \geq 0 \colon X_t^\sigma \notin B_\delta(x)\}$, and set $\theta_h = \theta \wedge h$, for any $h \geq 0$. Note that $\theta_h = h$ for all h sufficiently small.

Under Assumption 1.16, the dynamic programming principle (1.5) holds by Proposition 1.17. Since $v(X_t^\sigma)$ satisfies (1.13), we have

$$\begin{aligned}\mathbb{E}^x\left[\frac{1}{h}\int_0^{\theta_h}\left(\frac{1}{2}\operatorname{Tr}\left(D^2v(X_t^\sigma)\sigma\sigma^\top\right) + f(X_t^\sigma)\right)\mathrm{d}t\right] &= \frac{1}{h}\mathbb{E}^x\left[v(X_{\theta_h}^\sigma) + \int_0^{\theta_h} f(X_t^\sigma)\,\mathrm{d}t\right] \\ &\quad - v(x) \\ &\geq 0,\end{aligned}$$

where the inequality is a consequence of the dynamic programming principle (1.5).

We can then use the mean value theorem and the dominated convergence theorem to take the limit as $h \to 0$. By continuity of f, we conclude that

$$\frac{1}{2}\operatorname{Tr}\left(D^2v(x)\sigma\sigma^\top\right) + f(x) \geq 0.$$

Taking the infimum over $\sigma \in U$, we have

$$-\frac{1}{2}\inf_{\sigma\in U}\operatorname{Tr}\left(D^2v(x)\sigma\sigma^\top\right) - f(x) \leq 0.$$

Hence v is a classical subsolution of the HJB equation in (1.12).

As noted by Touzi before Proposition 3.5 of [58], the proof that v is a supersolution is more technical, and we do not present a proof here. In Theorem 4.20, we will prove that v is a viscosity solution of (1.12) under weaker conditions. In the case that v is twice continuously differentiable, this result implies that v is also a classical solution.

We now show that, under additional assumptions on the domain, the boundary condition is satisfied. In Section 4.5, we will show that the boundary condition is satisfied under weaker conditions. We take the following definition of a regular boundary point from Definition 9.2.8 of [46]. Define a process X on the domain D and denote $\tau := \inf\{t > 0\colon X \notin D\}$. We say that a point $a \in \partial D$ is *regular* for X if $\mathbb{P}^a[\tau = 0] = 1$.

Suppose that, for each control $\nu \in \mathcal{U}$, all points $a \in \partial D$ are regular for the controlled process X^ν. Let $a \in \partial D$. Then

$$\begin{aligned} v(a) &= \inf_{\nu \in \mathcal{U}} \mathbb{E}^a \left[\int_0^\tau f(X_s^\nu)\,\mathrm{d}s + g(X_\tau^\nu) \right] \\ &= \inf_{\nu \in \mathcal{U}} \mathbb{E}^a \left[g(X_0^\nu) \right] = g(a). \end{aligned}$$

Therefore we have the boundary condition

$$v = g \quad \text{on} \quad \partial D.$$

In practice, we do not expect v to be continuously differentiable, even for continuous cost functions f, so we cannot expect v to be a classical solution of the boundary value problem (1.12). In Chapter 4, we introduce viscosity solutions, which are the appropriate notion of weak solution for this context. We state the main theorem of Chapter 4 here, as we will apply this theorem in the following chapter before giving the proof.

Theorem 4.20. *Suppose that Assumption 1.16 holds, and suppose further that the domain D is uniformly convex, the running cost f is continuous in D, and the boundary cost g is uniformly continuous on ∂D.*

Then the value function $v : D \to \mathbb{R}$ defined in Section 1.4.1 extends continuously to $\overline{D}$ and is the unique viscosity solution of the HJB equation

$$-\frac{1}{2} \inf_{\sigma \in U} \operatorname{Tr}\left(D^2 v \sigma \sigma^\top\right) - f = 0$$

in D, with boundary condition

$$v = g \quad \textit{on} \quad \partial D.$$

This theorem provides a necessary and sufficient condition for a candidate function to be equal to the value function.

In Chapter 2, we make some additional assumptions on the structure of the problem that allow us to find an explicit expression for the value function, using Theorem 4.20 to prove optimality.

CHAPTER 2 STOCHASTIC CONTROL OF MARTINGALES IN A RADIALLY SYMMETRIC ENVIRONMENT

In a radially symmetric environment, we are able to find an explicit solution to the control problem for martingales with unit quadratic variation. We construct the value function by reducing the control problem to a one-dimensional switching problem between two regimes, and we observe continuous and smooth fit properties at the switching points. For continuous cost functions, we prove optimality by referring to the theory of viscosity solutions for the associated Hamilton-Jacobi-Bellman equation. We extend this result to cost functions that may become infinite at the origin. We also introduce a Markov formulation of the control problem and show that this is equivalent to the strong and weak formulations, with a possible exception depending on the growth rate of the cost function at the origin.

2.1 Introduction

Let $d \geq 2$ and $R > 0$. Define the domain to be the open ball about the origin with radius R, which we denote $D = B_R(0) \subset \mathbb{R}^d$. Throughout this chapter, we work on the domain D and consider cost functions $f : D \to \mathbb{R}$ of the form

$$f(x) = \tilde{f}(|x|), \quad x \in D,$$

for some $\tilde{f} : [0, R) \to \mathbb{R}$. We call a function f of this form *radially symmetric*. In this chapter, we consider the control problem defined in Section 1.4.1 for radially symmetric cost functions f with a constant boundary cost g.

When $\tilde{f}$ is monotonically increasing, we will see that an optimal strategy is for the controlled process to run as a one-dimensional Brownian motion on the radius that passes through the current position. We will define such a process as radial motion in Definition 2.7.

On the other hand, when $\tilde{f}$ is monotonically decreasing, we will see that any strategy under which the radius of the controlled process increases deterministically is optimal. This behaviour can be achieved by choosing to move in a direction orthogonal to the current position. Denoting by $x^\top$ a vector orthogonal to $x \in \mathbb{R}^d \setminus \{0\}$ with the same magnitude, an example of such a process is a solution of the SDE

$$\mathrm{d}X_t = \frac{1}{|X_t|}\left[X_t^\top;\quad 0;\quad \cdots;0\right]\mathrm{d}B_t. \tag{2.1}$$

We will define this process as tangential motion in Definition 2.3. The fact that solutions of the SDE (2.1) have deterministically increasing radius has been used, for example, by Fernholz, Karatzas and Ruf in [25], and by Larsson and Ruf in [41], to study a problem of relative arbitrage. In Proposition 3.21 we will derive a more general form of SDE whose solutions exhibit the same property.

We will find that switching between the two regimes of radial motion and tangential motion is optimal for a large class of radially symmetric cost functions. In Section 2.3.1, we give a heuristic argument for reducing the control problem to a one-dimensional switching problem for the radius process. By considering the generators of the radius processes corresponding to radial and tangential motion, we find ODEs that the expected cost should solve under each of the two regimes. We derive conditions for identifying the optimal switching points in Section 2.3.2. When switching into the diffusive regime of radial motion, we impose a smooth fit condition. At the points of switching into the deterministic regime of tangential motion, however, we only need to impose continuous fit. Nevertheless, these switching points exhibit the smooth fit property. We discuss this phenomenon in Section 2.3.4.

In Section 2.3.3, we use the switching points that we have identified to solve a system of ODEs and construct a candidate for the value function for a radially symmetric cost function. Under regularity conditions on the cost function given in Assumption 2.11, we prove optimality of the candidate value function in Proposition 2.15. In particular, we assume that the cost function is continuous, so Assumption 1.1 is satisfied and the weak and strong value functions are equal by Proposition 1.7. To verify that the value function is equal to our candidate, we show that the candidate function is a viscosity solution of the Hamilton-Jacobi-Bellman equation

$$-\frac{1}{2}\inf_{\sigma\in U}\operatorname{Tr}\left(D^2u\sigma\sigma^\top\right) = f, \quad \text{in} \quad D,$$

with boundary condition $u = g$ on ∂D. We then appeal to Theorem 4.20 to see that the value function is the unique viscosity solution of this boundary value problem and is therefore equal to the candidate function. We will introduce the required theory of viscosity solutions and prove Theorem 4.20 in Chapter 4.

In Section 2.4, we relax the regularity conditions on the cost function. In particular, we allow the cost function to become infinite at the origin. In this case, equality between the weak and strong value functions is no longer guaranteed a priori. We show in Theorem 2.30 that the weak and strong value functions do coincide and take the same form as the candidate that we constructed in Section 2.3.3. We also find growth conditions under which the value function remains finite while the cost function becomes infinite. We identify a regime of moderate growth to infinity at the origin where we require results on Brownian filtrations from Chapter 3 in order to complete the proof of Theorem 2.30.

Finally, we introduce Markov controls in Section 2.5. We show that, under certain growth conditions on the cost function, the Markov formulation is equivalent to the strong and weak formulations of the control problem. In the regime of moderate growth mentioned above, we conjecture that there is a gap between the Markov value function and the strong and weak value functions at the origin. This conjecture is based on the fact that (2.1) has a weak solution but no strong solution starting from the origin. We prove this fact in Chapter 3, where we also discuss the conjecture further.

We begin this chapter by considering two simple examples of minimising and maximising the expected time spent in a ball about the origin.

2.2 Occupation times

Fix $R > 0$ so that the domain is $D = B_R(0) \subset \mathbb{R}^d$.

We first consider the following example of minimising the expected time spent in a ball about the origin.

Example 2.1. Let $\rho \in (0, R)$, define $f : D \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0, & |x| \leq \rho, \\ -1, & |x| \in (\rho, R), \end{cases}$$

and fix the boundary cost $g \equiv 0$.

We seek the value function

$$v(x) = \inf_{\sigma \in \mathcal{U}} \mathbb{E}^x \left[\int_0^\tau f(X_s^\sigma)\,\mathrm{d}s \right] = \inf_{\sigma \in \mathcal{U}} \mathbb{E}^x \left[\int_0^\tau -\mathbb{1}_{\{|X_s^\sigma| \in (\rho, R)\}}\,\mathrm{d}s \right].$$

That is, we wish to maximise the expected time that the radius process $|X^\sigma|$ spends in the interval (ρ, R).

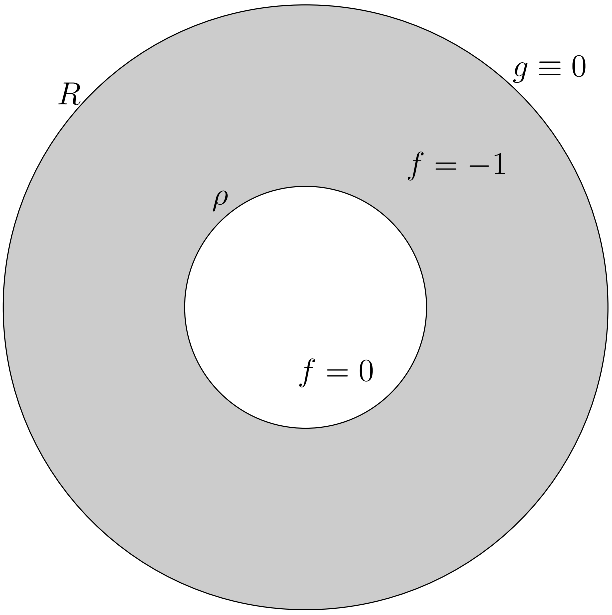


Figure 2.1: Cost function for Example 2.1

Since the problem is radially symmetric, we expect the value function v to depend only on the radius. In fact, in this example and the example that follows, it will be convenient to work with the squared radius of any controlled process. We now derive an SDE for this squared radius process.

Lemma 2.2. *Let $x \in D$, $\sigma \in \mathcal{U}$, and define X^σ by the stochastic integral*

$$X_t^\sigma = x + \int_0^t \sigma_s \,\mathrm{d}B_s, \quad t \geq 0.$$

Define the squared radius process Z^σ by $Z_t^\sigma := |X_t^\sigma|^2$, $t \geq 0$. Then Z^σ satisfies the SDE

$$\mathrm{d}Z_t^\sigma = 2X_t^\top \sigma_t \,\mathrm{d}B_t + \mathrm{d}t, \tag{2.2}$$

with initial condition $Z_0^\sigma = |x|^2$.

Proof. We apply Itô's formula to the function $g : D \to \mathbb{R}$ defined by $g(x) = |x|^2$, for all $x \in D$. For any $t > 0$, Itô's formula gives

$$\begin{aligned} \mathrm{d}Z_t^\sigma = \mathrm{d}g(X_t) &= Dg(X_t)^\top \mathrm{d}X_t + \frac{1}{2}\operatorname{Tr}(D^2 g(X_t)\sigma_t\sigma_t^\top)\,\mathrm{d}t \\ &= 2X^\top \sigma_t\,\mathrm{d}B_t + \operatorname{Tr}(\sigma_t\sigma_t^\top)\,\mathrm{d}t \\ &= 2X^\top \sigma_t\,\mathrm{d}B_t + \mathrm{d}t, \end{aligned}$$

using the constraint that $\sigma_t \in U$. □

We conjecture that, at the boundary $\{x \in D\colon |x| = \rho\}$, any optimal control for Example 2.1 must enforce motion tangential to this internal boundary. We now define a process that exhibits this behaviour.

Definition 2.3 (Tangential motion)**.** For $x \in D \setminus \{0\}$, define

$$\sigma^0(x) := \frac{1}{|x|}\left[x^\perp;\quad 0;\quad \cdots;\quad 0\right] \in \mathbb{R}^{d,d}, \tag{2.3}$$

where $x^\perp$ denotes any $x \in \mathbb{R}^d \setminus \{0\}$ such that $x^\top x^\perp = 0$. Fix $x \in D \setminus \{0\}$ and suppose that X^{σ^0} is a strong solution of the SDE

$$\mathrm{d}X_t = \sigma^0(X_t)\,\mathrm{d}B_t, \quad X_0 = x.$$

For $t \geq 0$, define

$$\sigma_t^0 := \sigma^0(X_t^{\sigma^0}),$$

so that

$$X_t^{\sigma^0} = x + \int_0^t \sigma_s^0\,\mathrm{d}B_s.$$

We say that the process X^{σ^0} follows *tangential motion.*

Note that $\sigma^0(0)$ is not defined. We investigate the existence of a process following tangential motion at the origin in Chapter 3.

For σ^0 defined in Definition 2.3, we can find a formula for the squared radius process Z^{σ^0} via Lemma 2.2, as follows.

Lemma 2.4. *Suppose that X^{σ^0} follows tangential motion, as defined in Definition 2.3, with $X_0^{\sigma^0} = x \neq 0$. Then the radius process is deterministically increasing and, for any $t \geq 0$,*

$$Z_t^{\sigma^0} = \left|X_t^{\sigma^0}\right|^2 = |x| + t.$$

Proof. For $t \geq 0$, provided that $\left|X_t^{\sigma^0}\right| \neq 0$, we see that

$$\left(X_t^{\sigma^0}\right)^\top \sigma_t^0 = \frac{1}{\left|X_t^{\sigma^0}\right|}\left[\left(X_t^{\sigma^0}\right)^\top \left(X_t^{\sigma^0}\right)^\perp, \quad 0, \quad \dots, \quad 0\right] = \left[0, \quad \dots, \quad 0\right].$$

Therefore, by Lemma 2.2, Z^{σ^0} satisfies

$$\mathrm{d}Z_t^{\sigma^0} = \mathrm{d}t.$$

Let $\xi = |x|^2 \neq 0$, so that $Z_0^{\sigma^0} = \xi$. Then Z^{σ^0} is the deterministically increasing process given by

$$Z_t^{\sigma^0} = \xi + t,$$

for $t \geq 0$. $\square$

As a consequence of the above lemma, supposing that $X_0^{\sigma^0} \neq 0$, we have that $\left|X_t^{\sigma^0}\right| > 0$ for all $t \geq 0$. Therefore the control σ^0 is well-defined when starting away from the origin. Note that, for $d \geq 3$, a control of this form is not unique, since the orthogonal vector in the definition of σ^0 can be chosen as any element of a $(d-1)$-dimensional subspace.

The observation that the process X^{σ^0} has deterministically increasing radius was made by Fernholtz, Karatzas and Ruf in Section 6.2 of [25] and again by Larsson and Ruf in Section 4.2 of [41], where they consider a problem of relative arbitrage.

In Figure 2.2, we show a simulated trajectory of a process following tangential motion in dimension $d = 2$. We note that a similar simulation is produced in Figure 2 of [41].

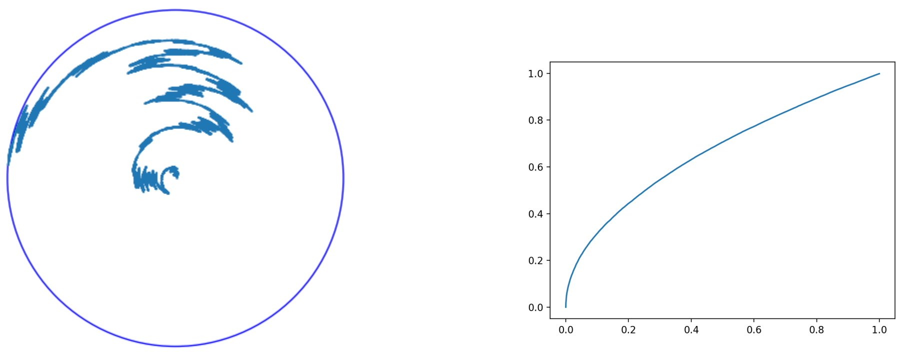


Figure 2.2: A sample path of a process $(X_t^{\sigma^0})_{t\geq 0}$ following tangential motion in dimension $d = 2$ and its radius

Having defined tangential motion and proved a key property of this process, we now construct a candidate for the value function in Example 2.1.

Fix $\xi \geq \rho^2$. Then we conjecture that the control σ^0 defined in Definition 2.3 is optimal, and we compute the expected cost

$$\begin{aligned}\mathbb{E}^x\left[\int_0^\tau -\mathbb{1}_{\{|X_s^{\sigma^0}|\in(\rho,R)\}}\,\mathrm{d}s\right] &= \mathbb{E}^\xi\left[\int_0^\tau -\mathbb{1}_{\{Z_s^{\sigma^0}\in(\rho^2,R^2)\}}\,\mathrm{d}s\right]\\ &= -\int_0^\infty \mathbb{1}_{\{s\in(0,R^2-\xi)\}}\,\mathrm{d}s\\ &= -\int_0^{R^2-\xi}\mathrm{d}s\\ &= \xi - R^2.\end{aligned}$$

Now suppose that $\xi < \rho^2$. This includes the case where the process starts at the origin, where the control σ^0 is not well-defined. However, since the cost is zero in the ball $\{x \in \mathbb{R}^d\colon\ |x| < \rho\}$, we will see that any strategy is optimal in this region. For a fixed $r \in (\sqrt{\xi}, \rho)$ and an arbitrary $\sigma \in \mathcal{U}$, define the control $\sigma^\star$ by

$$\sigma_t^\star = \begin{cases}\sigma_t, & \left|X_t^{\sigma^\star}\right| < r,\\ \sigma_t^0, & \left|X_t^{\sigma^\star}\right| \in [r, R).\end{cases}$$

Then we compute the expected cost

$$\begin{aligned}\mathbb{E}^x\left[\int_0^\tau -\mathbb{1}_{\{|X_s^{\sigma^\star}|\in(\rho,R)\}}\,\mathrm{d}s\right] &= \mathbb{E}^\xi\left[\int_0^\tau -\mathbb{1}_{\{Z_s^{\sigma^\star}\in(\rho^2,R^2)\}}\,\mathrm{d}s\right]\\ &= \mathbb{E}^{r^2}\left[\int_0^\tau -\mathbb{1}_{\{Z_s^{\sigma^0}\in(\rho^2,R^2)\}}\,\mathrm{d}s\right]\\ &= -\int_{r^2-\xi}^\infty \mathbb{1}_{\{s\in(\rho^2-\xi,R^2-\xi)\}}\,\mathrm{d}s\\ &= -\int_{\rho^2}^{R^2}\mathrm{d}s\\ &= \rho^2 - R^2.\end{aligned}$$

This calculation gives us a conjecture for the value function in Example 2.1. Using the Itô-Tanaka formula, we will show that our candidate function satisfies a dynamic programming principle, as described in Section 1.4.2, and we can then deduce that this function must be the value function.

Proposition 2.5. *Let $w : [0, R^2) \to \mathbb{R}$ be defined by*

$$w(\xi) = \begin{cases} \rho^2 - R^2, & \xi \le \rho^2, \\ \xi - R^2, & \xi \in (\rho^2, R^2), \end{cases}$$

and define $\overline{v} : D \to \mathbb{R}$ by $\overline{v}(x) = w(|x|^2)$, for $x \in D$. Then the value function for Example 2.1 is given by $v = \overline{v}$.

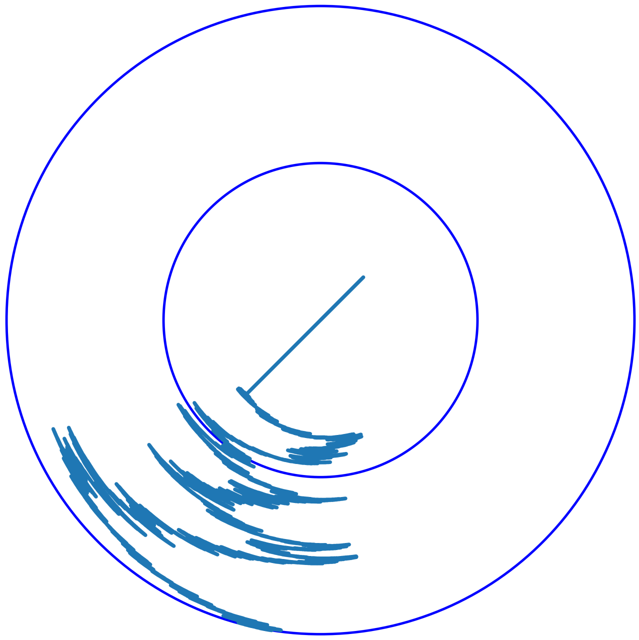


Figure 2.3: A possible trajectory for an optimal strategy in Example 2.1

Proof. We first show that $\overline{v}$ satisfies the form of the dynamic programming principle given in Remark 1.9.

Define $\tilde{f} : [0, R^2) \to \mathbb{R}$ by

$$\tilde{f}(\xi) = -\mathbb{1}_{\{\xi \in (\rho^2, R^2)\}}, \quad \xi \in [0, R^2),$$

so that

$$f(x) = \tilde{f}(|x|), \quad x \in D.$$

We seek to prove that $w(Z_t^\sigma) + \int_0^t \tilde{f}(Z_s^\sigma)\,\mathrm{d}s$ is a submartingale for all $\sigma \in \mathcal{U}$, and that $w(Z_t^{\sigma^\star}) + \int_0^t \tilde{f}(Z_s^{\sigma^\star})\,\mathrm{d}s$ is a martingale for an optimal strategy $\sigma^\star \in \mathcal{U}$.

Let $\sigma \in \mathcal{U}$. We note that w is not continuously differentiable at $\xi = \rho^2$, so we apply the Itô-Tanaka formula to write down an SDE for $w(Z_t^\sigma)$. Recall that the

Itô-Tanaka formula, given for example in Theorem 1.5 of [51, Chapter VI] states that, for all $t \geq 0$,

$$w(Z_t^\sigma) - w(\xi) = \int_0^t w'_-(Z_s^\sigma)\,\mathrm{d}Z_s^\sigma + \frac{1}{2}\int_0^{R^2} L_t^{\sigma,a} w''(\mathrm{d}a),$$

where

- w'_- is the left derivative of w, which exists everywhere;
- $w''(\mathrm{d}a)$ is the distributional derivative of w'_-; i.e. the measure on $\mathbb{R}$ such that, for all $\xi \in [0, R^2)$,
$$w'_-(\xi) = \int_{-\infty}^{\xi} w''(\mathrm{d}a);$$
- $L_t^{\sigma,a}$ is the local time spent at a by the process Z^σ up to time t.

We calculate that

$$w'_-(\xi) = \begin{cases} 0, & \text{for} \quad \xi \leq \rho^2, \\ 1, & \text{for} \quad \xi \in (\rho^2, R^2), \end{cases}$$

and

$$w''(\mathrm{d}a) = \delta_{\rho^2}(a).$$

Hence, by the Itô-Tanaka formula,

$$\begin{aligned} w(Z_t^\sigma) - w(\xi) &= \int_0^t \mathbb{1}_{\{Z_s^\sigma > \rho^2\}}\,\mathrm{d}Z_s^\sigma + \frac{1}{2}L_t^{\sigma,\rho^2} \\ &= 2\int_0^t \mathbb{1}_{\{Z_s^\sigma > \rho^2\}} X_s^\top \sigma_s\,\mathrm{d}B_s + \int_0^t \mathbb{1}_{\{Z_s^\sigma > \rho^2\}}\,\mathrm{d}s + \frac{1}{2}L_t^{\sigma,\rho^2}, \end{aligned}$$

and so

$$\begin{aligned} w(Z_t^\sigma) - w(\xi) + \int_0^t \tilde{f}(Z_s^\sigma)\,\mathrm{d}s &= 2\int_0^t \mathbb{1}_{\{Z_s^\sigma > \rho^2\}} X_s^\top \sigma_s\,\mathrm{d}B_s + \int_0^t \mathbb{1}_{\{Z_s^\sigma > \rho^2\}}\,\mathrm{d}s \\ &\quad + \frac{1}{2}L_t^{\sigma,\rho^2} - \int_0^t \mathbb{1}_{\{Z_s^\sigma > \rho^2\}}\,\mathrm{d}s \\ &= 2\int_0^t \mathbb{1}_{\{Z_s^\sigma > \rho^2\}} X_s^\top \sigma_s\,\mathrm{d}B_s + \frac{1}{2}L_t^{\sigma,\rho^2}. \end{aligned} \tag{2.4}$$

Since local time is always non-negative, we have shown that

$$w(Z_t^\sigma) + \int_0^t \tilde{f}(Z_s^\sigma)\,\mathrm{d}s$$

is a submartingale for any $\sigma \in \mathcal{U}$.

Now we note that, for any $\sigma \in \mathcal{U}$,

$$\overline{v}(X_\tau^\sigma) = w(Z_\tau^\sigma) = w(R^2) = 0,$$

by continuity of the paths of X^σ. Therefore, we can use the submartingale property and the optional sampling theorem to find that

$$\begin{aligned}
\mathbb{E}^x \left[\int_0^\tau f(X_s^\sigma) \, \mathrm{d}s \right] &= \mathbb{E}^x \left[\int_0^\tau f(X_s^\sigma) \, \mathrm{d}s + \overline{v}(X_\tau^\sigma) \right] \\
&= \mathbb{E}^\xi \left[\int_0^\tau \tilde{f}(Z_s^\sigma) \, \mathrm{d}s + w(Z_\tau^\sigma) \right] \\
&\geq w(\xi) = \overline{v}(x).
\end{aligned}$$

Now, supposing that $\xi \neq 0$, consider the control $\sigma^\star = \sigma^0$, so that $Z_t^{\sigma^\star} = \xi + t$, for any $t \geq 0$. For $\xi \in (0, \rho^2]$, we have

$$\begin{aligned}
\mathbb{E}^\xi \left[- \int_0^\tau \mathbb{1}_{\left\{ Z_s^{\sigma^\star} \in (\rho^2, R^2) \right\}} \, \mathrm{d}s \right] &= - \int_0^\infty \mathbb{1}_{\{ s \in (\rho^2 - \xi, R^2 - \xi) \}} \, \mathrm{d}s \\
&= \rho^2 - R^2 = w(\xi).
\end{aligned}$$

For $\xi \in (\rho^2, R^2)$, we have

$$\begin{aligned}
\mathbb{E}^\xi \left[- \int_0^\tau \mathbb{1}_{\left\{ Z_s^{\sigma^\star} \in (\rho^2, R^2) \right\}} \, \mathrm{d}s \right] &= - \int_0^\tau \mathbb{1}_{\left\{ Z_s^{\sigma^\star} \in (\xi, R^2) \right\}} \, \mathrm{d}s \\
&= - \int_0^\infty \mathbb{1}_{\{ s \in (0, R^2 - \xi) \}} \, \mathrm{d}s \\
&= \xi - R^2 = w(\xi).
\end{aligned}$$

In the case that $\xi = 0$, fix $r \in (0, \rho)$ and $\sigma \in \mathcal{U}$, and take

$$\sigma_t^\star = \begin{cases} \sigma_t, & \left| X_t^{\sigma^\star} \right| < r, \\ \sigma_t^0, & \left| X_t^{\sigma^\star} \right| \in [r, R). \end{cases}$$

Then

$$\begin{aligned}
\mathbb{E}^0 \left[- \int_0^\tau \mathbb{1}_{\left\{ Z_s^{\sigma^\star} \in (\rho^2, R^2) \right\}} \, \mathrm{d}s \right] &= \mathbb{E}^r \left[- \int_0^\tau \mathbb{1}_{\left\{ Z_s^{\sigma^\star} \in (\rho^2, R^2) \right\}} \, \mathrm{d}s \right] \\
&= - \int_0^\infty \mathbb{1}_{\{ s \in (\rho^2 - \xi, R^2 - \xi) \}} \, \mathrm{d}s \\
&= R^2 - \rho^2 = w(0).
\end{aligned}$$

We conclude that, for any $x \in D$,

$$\overline{v}(x) = \inf_{\sigma \in \mathcal{U}} \mathbb{E}^x \left[\int_0^\tau f(X_s^\sigma)\,\mathrm{d}s \right] = v(x).$$

Hence the conjectured function $\overline{v}$ is indeed the value function. $\square$

We now turn to a second example of maximising the expected time spent in a ball around the origin.

Example 2.6. Fix $\rho \in (0, R)$, define the cost $f : D \to \mathbb{R}$ by

$$f(x) = \begin{cases} -1, & |x| < \rho \\ 0, & |x| \in [\rho, R), \end{cases}$$

and fix the boundary cost $g \equiv 0$. We seek the value function

$$v(x) = \inf_{\sigma \in \mathcal{U}} \mathbb{E}^x \left[\int_0^\tau f(X_s^\sigma)\,\mathrm{d}s \right] = \inf_{\sigma \in \mathcal{U}} \mathbb{E}^x \left[\int_0^\tau -\mathbb{1}_{\{|X_s^\sigma| < \rho\}}\,\mathrm{d}s \right].$$

That is, we wish to maximise the expected time that the martingale spends in the ball $B_\rho(0)$.

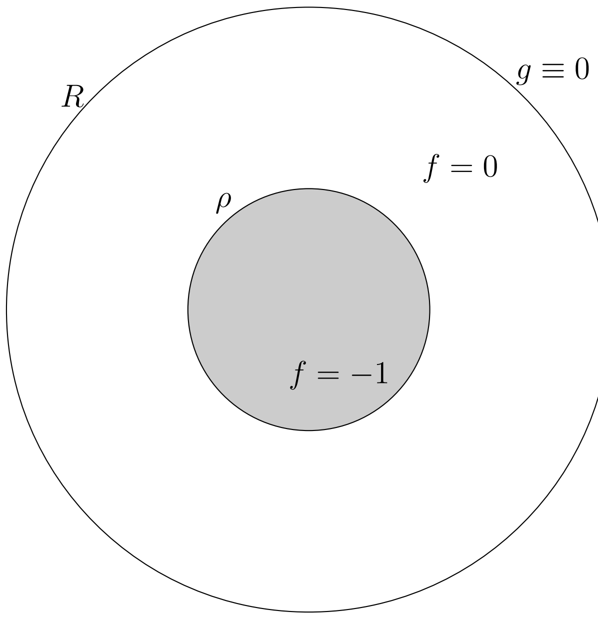


Figure 2.4: Cost function for Example 2.6

We propose that an optimal strategy is to run as a Brownian motion on the radius of the domain. We now define a process that follows this strategy.

Definition 2.7 (Radial motion)**.** Define a function $\sigma^1 : D \to \mathbb{R}$ by

$$\sigma^1(x) = \begin{cases} \frac{1}{|x|}\begin{bmatrix} x; & 0; & \cdots; & 0 \end{bmatrix}, & x \neq 0, \\ \begin{bmatrix} e_1; & 0; & \cdots; & 0 \end{bmatrix}, & x = 0, \end{cases} \tag{2.5}$$

where e_1 is the unit vector in the first coordinate direction. Fix $x \in D$ and define σ^1 to be the constant control given by $\sigma^1_t = \sigma(x)$, for all $t \geq 0$. Define X^{σ^1} by

$$X^{\sigma^1}_t = x + \int_0^t \sigma^1_s \, \mathrm{d}B_s = x + \sigma^1(x) B_t, \quad t \geq 0.$$

We say that the process X^{σ^1} follows *radial motion*.

A simulated trajectory of a process following radial motion, along with the sample path of its radius, is shown in Figure 2.5.

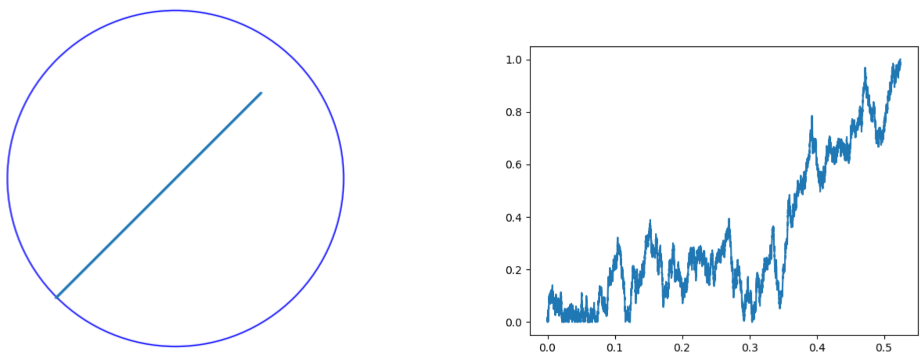


Figure 2.5: A sample path of a process $(X^{\sigma^1}_t)_{t \geq 0}$ following radial motion in dimension $d = 2$ and its radius

Let W be the first component of B, and note that W is a one-dimensional Brownian motion. Then, defining σ^1 as in Definition 2.7, we see that, for $x \neq 0$,

$$X^{\sigma^1}_t = x + \int_0^t \sigma^1_s \, \mathrm{d}B_s = x + W_t \frac{x}{|x|},$$

and, for $x = 0$,

$$X^{\sigma^1}_t = W_t e_1.$$

Hence $\left| X^{\sigma^1}_t \right| = \left| |x| + W_t \right|$, and so

$$\mathbb{E}^x \left[\int_0^\tau -\mathbb{1}_{\left\{ \left| X^{\sigma^1}_t \right| < \rho \right\}} \right] = \mathbb{E}^{|x|} \left[- \int_0^{\tau_R} \mathbb{1}_{\{W_t \in (-\rho, \rho)\}} \right],$$

where $\tau_R = \inf\{t \geq 0\colon\ |W_t| = R\}$.

We can compute this expected cost by using the Green's function for one-dimensional Brownian motion, using the results that we summarise in Appendix B, as follows.

The scale function s and speed measure m, as defined in Definition B.1 and Definition B.3 respectively, are given by

$$s(y) = y - c, \quad y \in \mathbb{R},$$

for some constant $c \in \mathbb{R}$, and

$$\int m(\mathrm{d}y) = 2 \int \mathrm{d}y.$$

Hence the Green's function G on the interval $[-\rho, \rho]$, as defined in Definition B.4, is given by

$$G(r, y) = \begin{cases} \frac{(y+R)(R-r)}{2R}, & y \leq r, \\ \frac{(r+R)(R-y)}{2R}, & y \geq r. \end{cases}$$

We can now apply Proposition B.5, which tells us that

$$\mathbb{E}^{|x|}\left[-\int_0^{\tau_R} \mathbb{1}_{\{W_t \in (-\rho,\rho)\}}\right] = -\int_{-\rho}^{\rho} G(|x|, y) m(\mathrm{d}y).$$

For $|x| \geq \rho$, we calculate

$$\begin{aligned} -\int_{-\rho}^{\rho} G(|x|, y) m(\mathrm{d}y) &= -\frac{R - |x|}{R} \int_{-\rho}^{\rho} (y + R)\,\mathrm{d}y \\ &= 2\rho\,|x| - 2\rho R, \end{aligned}$$

and, for $|x| < \rho$,

$$\begin{aligned} -\int_{-\rho}^{\rho} G(|x|, y) m(\mathrm{d}y) &= -\frac{R - |x|}{R} \int_{-\rho}^{|x|} (y + R)\,\mathrm{d}y \\ &\quad - \frac{|x| + R}{R} \int_{|x|}^{\rho} (R - y)\,\mathrm{d}y \\ &= \frac{R - |x|}{R} \left(\frac{1}{2}\left(\rho^2 - |x|^2\right) - R\,|x| - \rho R\right) \\ &\quad + \frac{|x| + R}{R} \left(\frac{1}{2}\left(\rho^2 - |x|^2\right) + R\,|x| - \rho R\right) \\ &= |x|^2 + \rho^2 - 2\rho R. \end{aligned}$$

This gives us a candidate for the value function in Example 2.6. Again, we present this function in terms of the radius squared. We can then apply Itô's formula, using the SDE for the squared radius process that we derived in Lemma 2.2.

Proposition 2.8. *Let $w : [0, R^2) \to \mathbb{R}$ be defined by*

$$w(\xi) = \begin{cases} \xi + \rho^2 - 2\rho R, & \xi \leq \rho^2, \\ 2\rho\xi^{\frac{1}{2}} - 2\rho R, & \xi \in (\rho^2, R^2), \end{cases}$$

and define $\overline{v} : D \to \mathbb{R}$ by $\overline{v}(x) = w(|x|^2)$, for $x \in D$. Then the value function for Example 2.6 is given by $v = \overline{v}$.

Notation. Throughout this thesis, I denotes the d-dimensional identity matrix.

Proof of Proposition 2.8. Again we will show that $\overline{v}$ satisfies the form of the dynamic programming principle given in Remark 1.9.

Note first that w is continuously differentiable and twice piecewise continuously differentiable, with

$$w'(\xi) = \begin{cases} 1, & \xi \leq \rho^2, \\ \rho\xi^{-\frac{1}{2}}, & \xi \in (\rho^2, R^2), \end{cases}$$

and

$$w''(\xi) = \begin{cases} 0, & \xi \leq \rho^2, \\ -\frac{1}{2}\rho\xi^{-\frac{3}{2}}, & \xi \in (\rho^2, R^2). \end{cases}$$

Hence we can apply Itô's formula to $w(Z_t^\sigma)$, for any $\sigma \in \mathcal{U}$, recalling that $Z_t^\sigma = |X_t^\sigma|^2$. Let $Z_0^\sigma = \xi \in [0, R^2)$. Then, for $t > 0$,

$$\begin{aligned} w(Z_t^\sigma) - w(\xi) = &\int_0^t \mathbb{1}_{\{Z_s^\sigma \leq \rho^2\}} \,\mathrm{d}Z_s^\sigma + \rho \int_0^t \mathbb{1}_{\{Z_s^\sigma \in (\rho^2, R^2)\}} (Z_s^\sigma)^{-\frac{1}{2}} \,\mathrm{d}Z_s^\sigma \\ &- \frac{\rho}{4} \int_0^t \mathbb{1}_{\{Z_s^\sigma \in (\rho^2, R^2)\}} (Z_s^\sigma)^{-\frac{3}{2}} \,\mathrm{d}\langle Z^\sigma \rangle_s. \end{aligned} \tag{2.6}$$

Substituting in the SDE (2.2) for Z^σ, we find that there is a square-integrable martingale M^σ such that

$$\begin{aligned} w(Z^\sigma_t) - w(\xi) &= \int_0^t \mathrm{d}M^\sigma_s + \int_0^t \mathbb{1}_{\{Z^\sigma_s \le \rho^2\}} \,\mathrm{d}s + \rho \int_0^t \mathbb{1}_{\{Z^\sigma_s \in (\rho^2, R^2)\}} (Z^\sigma_s)^{-\frac{1}{2}} \,\mathrm{d}s \\ &\quad - \rho \int_0^t \mathbb{1}_{\{Z^\sigma_s \in (\rho^2, R^2)\}} (Z^\sigma_s)^{-\frac{3}{2}} \operatorname{Tr}\left(X^\sigma_s X^{\sigma\top}_s \sigma_s \sigma_s^\top \right) \mathrm{d}s \\ &= \int_0^t \mathrm{d}M^\sigma s - \int_0^t f(X^\sigma_s) \,\mathrm{d}s \\ &\quad + \rho \int_0^t \mathbb{1}_{\{|X^\sigma_s| \in (\rho, R)\}} \left|X^\sigma_s\right|^{-3} \operatorname{Tr}\left(\left[\left|X^\sigma_s\right|^2 I - X^\sigma_s X^{\sigma\top}_s \right] \sigma_s \sigma_s^\top \right) \mathrm{d}s. \end{aligned}$$

Noting that the matrix $|x|^2 I - xx^\top$ is positive semi-definite for any $x \in \mathbb{R}^d$, we see that the final integral in the above equation is always non-negative, and so

$$\overline{v}(X^\sigma_t) + \int_0^t f(X^\sigma_s) \,\mathrm{d}s$$

is a submartingale for any $\sigma \in \mathcal{U}$.

Now take $\sigma = \sigma^1$ and let W be the first component of the Brownian motion B. Then from the SDE (2.2) for the squared radius process, we see that $Z := Z^{\sigma^1}$ is a one-dimensional squared Bessel process satisfying

$$\mathrm{d}Z_t = 2\sqrt{Z_t} \,\mathrm{d}W_t + \mathrm{d}t.$$

Substituting this SDE for Z into our calculation (2.6), and defining $X := X^{\sigma^1}$, we find that there is a square-integrable martingale M such that, for any $t > 0$,

$$\begin{aligned} w(Z_t) - w(\xi) &= \int_0^t \mathrm{d}M_s + \int_0^t \mathbb{1}_{\{Z_s \le \rho^2\}} \,\mathrm{d}s + \rho \int_0^t \mathbb{1}_{\{Z_s \in (\rho^2, R^2)\}} Z_s^{-\frac{1}{2}} \,\mathrm{d}s \\ &\quad - \frac{\rho}{4} \int_0^t \mathbb{1}_{\{Z_s \in (\rho^2, R^2)\}} Z_s^{-\frac{3}{2}} \cdot 4 Z_s \,\mathrm{d}s \\ &= \int_0^t \mathrm{d}M_s - \int_0^t f(X_s) \,\mathrm{d}s. \end{aligned}$$

Hence $\overline{v}(X_t) + \int_0^t f(X_s) \,\mathrm{d}s$ is a martingale.

By the optional sampling theorem, we see that, for any $\sigma \in \mathcal{U}$ and $x \in D$,

$$\begin{aligned} \overline{v}(x) &\le \mathbb{E}^x \left[\overline{v}(X^\sigma_\tau) + \int_0^\tau f(X^\sigma_s) \,\mathrm{d}s \right] \\ &= w(R^2) + \mathbb{E}^x \left[\int_0^\tau f(X^\sigma_s) \,\mathrm{d}s \right] = \mathbb{E}^x \left[\int_0^\tau f(X^\sigma_s) \,\mathrm{d}s \right], \end{aligned}$$

by the submartingale property. Similarly, by the martingale property for $X = X^{\sigma^1}$,

$$\overline{v}(x) = \mathbb{E}^x\left[\int_0^\tau f(X_s)\,\mathrm{d}s\right].$$

Therefore

$$\overline{v}(x) = \inf_{\sigma\in\mathcal{U}} \mathbb{E}^x\left[\int_0^\tau f(X_s^\sigma)\,\mathrm{d}s\right],$$

as required. $\square$

For the step cost functions considered above, optimal controls involve tangential and radial motion, as defined in Definition 2.3 and Definition 2.7, respectively. In the next section, we will show that the controls corresponding to tangential and radial motion are optimal for sufficiently smooth monotone cost functions. We will apply the results from the above examples to find the value functions for these monotone costs.

2.2.1 Examples with monotone costs

For a sufficiently smooth cost function $f : D \to \mathbb{R}$, and for any $x \in D$, we can write

$$\begin{aligned} f(x) = \tilde{f}(|x|) &= \int_0^{|x|} \tilde{f}'(r)\,\mathrm{d}r \\ &= \int_0^R \mathbb{1}_{\{|x|\geq r\}}\tilde{f}'(r)\,\mathrm{d}r, \end{aligned} \tag{2.7}$$

expressing the cost in terms of an indicator function. When the cost function is also monotone, we can then apply the results from Example 2.1 and Example 2.6 to find the value function, as we show in the following results.

We first show that radial motion is optimal for increasing costs.

Proposition 2.9. *Suppose that $\tilde{f}$ is a continuously differentiable increasing function with $\tilde{f}(0) = 0$. Then the control σ^1, as defined in Definition 2.7, is optimal and the value function v is given by*

$$v(x) = \mathbb{E}^x\left[\int_0^\tau f\left(X_s^{\sigma^1}\right)\mathrm{d}s\right] + g = 2\int_{|x|}^R\int_0^r \tilde{f}(s)\,\mathrm{d}s\,\mathrm{d}r + g, \quad x \in D.$$

Proof. Fix $x \in D$ and $r \in (0, R)$. Note first that, as shown in Proposition 1.5, $\mathbb{E}^x[\tau] = R^2 - |x|^2$ for any $\sigma \in \mathcal{U}$.

From Proposition 2.8, we know that the control σ^1 is optimal for an increasing step function, and so for any $\sigma \in \mathcal{U}$,

$$\begin{aligned}
\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\{|X_s^\sigma|\geq r\}}\,\mathrm{d}s\right] &= \mathbb{E}^x[\tau] + \mathbb{E}^x\left[\int_0^\tau -\mathbb{1}_{\{|X_s^\sigma|<r\}}\,\mathrm{d}s\right] \\
&\geq \mathbb{E}^x[\tau] + \inf_{\nu\in\mathcal{U}} \mathbb{E}^x\left[\int_0^\tau -\mathbb{1}_{\{|X_s^\nu|<r\}}\,\mathrm{d}s\right] \\
&= \inf_{\nu\in\mathcal{U}} \mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\{|X_s^\nu|\geq r\}}\,\mathrm{d}s\right] \\
&= \mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\left\{\left|X_s^{\sigma^1}\right|\geq r\right\}}\,\mathrm{d}s\right].
\end{aligned} \tag{2.8}$$

Since $\tilde{f}$ is continuously differentiable, we can write f as in (2.7). Therefore, for any $\sigma \in \mathcal{U}$,

$$\begin{aligned}
\mathbb{E}^x\left[\int_0^\tau f(X_s^\sigma)\,\mathrm{d}s\right] &= \mathbb{E}^x\left[\int_0^\tau\int_0^R \mathbb{1}_{\{|X_s^\sigma|\geq r\}}\tilde{f}'(r)\,\mathrm{d}r\,\mathrm{d}s\right] \\
&= \int_0^R \tilde{f}'(r)\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\{|X_s^\sigma|\geq r\}}\,\mathrm{d}s\right]\mathrm{d}r,
\end{aligned}$$

using the fact that $\tilde{f}'$ is bounded to exchange the order of integration. Now, since $\tilde{f}' \geq 0$, the inequality (2.8) yields

$$\begin{aligned}
\mathbb{E}^x\left[\int_0^\tau f(X_s^\sigma)\,\mathrm{d}s\right] &\geq \int_0^R \tilde{f}'(r)\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\left\{\left|X_s^{\sigma^1}\right|\geq r\right\}}\,\mathrm{d}s\right]\mathrm{d}r \\
&= \mathbb{E}^x\left[\int_0^\tau \tilde{f}\left(\left|X_s^{\sigma^1}\right|\right)\mathrm{d}s\right].
\end{aligned}$$

This shows that

$$\begin{aligned}
v(x) &= \int_0^R \tilde{f}'(r)\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\left\{\left|X_s^{\sigma^1}\right|\geq r\right\}}\,\mathrm{d}s\right]\mathrm{d}r + g \\
&= \left(R^2 - |x|^2\right)\tilde{f}(R) + \int_0^R \tilde{f}'(r)\mathbb{E}^x\left[\int_0^\tau -\mathbb{1}_{\left\{\left|X_s^{\sigma^1}\right|<r\right\}}\,\mathrm{d}s\right]\mathrm{d}r + g.
\end{aligned}$$

Substituting in the value function from Proposition 2.8, we calculate that

$$\begin{aligned}
v(x) = \left(R^2 - |x|^2\right)\tilde{f}(R) &- 2R\int_0^R r\tilde{f}'(r)\,\mathrm{d}r + 2\,|x|\int_0^{|x|} r\tilde{f}'(r)\,\mathrm{d}r \\
&+ |x|^2\left(\tilde{f}(R) - \tilde{f}(|x|)\right) + \int_{|x|}^R r^2\tilde{f}'(r)\,\mathrm{d}r + g.
\end{aligned} \tag{2.9}$$

We can apply integration by parts to get

$$-2R\int_0^R r\tilde{f}'(r)\,\mathrm{d}r = -2R^2\tilde{f}(R) + 2R\int_0^R \tilde{f}(r)\,\mathrm{d}r,$$

and

$$2\,|x|\int_0^{|x|} r\tilde{f}'(r)\,\mathrm{d}r = 2\,|x|^2\,\tilde{f}(|x|) - 2\,|x|\int_0^{|x|} \tilde{f}(r)\,\mathrm{d}r.$$

Now applying integration by parts twice, we see that

$$\begin{aligned}\int_{|x|}^R r^2\tilde{f}'(r)\,\mathrm{d}r &= R^2\tilde{f}(R) - |x|^2\,\tilde{f}(|x|) - 2\int_{|x|}^R r\tilde{f}(r)\,\mathrm{d}r \\ &= R^2\tilde{f}(R) - |x|^2\,\tilde{f}(|x|) - 2R\int_0^R \tilde{f}(r)\,\mathrm{d}r + 2\,|x|\int_0^{|x|} \tilde{f}(r)\,\mathrm{d}r \\ &\quad + 2\int_{|x|}^R\int_0^r \tilde{f}(s)\,\mathrm{d}s\,\mathrm{d}r.\end{aligned}$$

On substituting these expressions back into our calculation of the value function in (2.9), all but one of the terms cancel and we find that

$$v(x) = 2\int_{|x|}^R\int_0^r \tilde{f}(s)\,\mathrm{d}s\,\mathrm{d}r + g,$$

as required. □

We now show that, away from the origin, tangential motion is optimal for decreasing costs. We exclude the origin here, since we have not found a control starting from the origin that is optimal for all values of ρ in Example 2.1. In Section 2.3, when we treat more general radially symmetric costs, we will be able to find the value function at the origin. We will address the issue of the existence of optimal strategies starting from the origin in detail in Chapter 3.

Proposition 2.10. *Suppose that $\tilde{f}$ is a continuously differentiable decreasing function with $\tilde{f}(0) = 0$. Then the control σ^0, as defined in Definition 2.3, is optimal away from the origin, and the value function v is given by*

$$v(x) = \mathbb{E}^x\left[\int_0^\tau f\left(X_s^{\sigma^0}\right)\mathrm{d}s\right] + g = 2\int_{|x|}^R r\tilde{f}(r)\,\mathrm{d}r + g, \quad x \in D\setminus\{0\}.$$

Proof of Proposition 2.10. Fix $x \in D\setminus\{0\}$ and $r \in (0, R)$. By Proposition 2.5, we know that σ^0 is optimal for a decreasing step function. Therefore, for any $\sigma \in \mathcal{U}$,

$$\begin{aligned}\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\{|X_s^\sigma|\geq r\}}\,\mathrm{d}s\right] &\geq \inf_{\nu\in\mathcal{U}}\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\{|X_s^\nu|\geq r\}}\,\mathrm{d}s\right]\\ &= \mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\left\{\left|X_s^{\sigma^0}\right|\geq r\right\}}\,\mathrm{d}s\right].\end{aligned}$$

And so, similarly to the previous example, we can calculate that for any $\sigma \in \mathcal{U}$,

$$\begin{aligned}\mathbb{E}^x\left[\int_0^\tau \tilde{f}(|X_s^\sigma|)\,\mathrm{d}s\right] &= \int_0^R \tilde{f}'(r)\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\{|X_s^\sigma|\geq r\}}\,\mathrm{d}s\right]\mathrm{d}r\\ &\geq \int_0^R \tilde{f}'(r)\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\left\{\left|X_s^{\sigma^0}\right|\geq r\right\}}\,\mathrm{d}s\right]\mathrm{d}r\\ &= \mathbb{E}^x\left[\int_0^\tau \tilde{f}\left(\left|X_s^{\sigma^0}\right|\right)\mathrm{d}s\right],\end{aligned}$$

where we use the fact that $\tilde{f}' \leq 0$. Hence

$$\begin{aligned}v(x) &= \int_0^R \tilde{f}'(r)\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\left\{\left|X_s^{\sigma^0}\right|\geq r\right\}}\,\mathrm{d}s\right]\mathrm{d}r + g\\ &= -\int_0^R \tilde{f}'(r)\mathbb{E}^x\left[\int_0^\tau -\mathbb{1}_{\left\{\left|X_s^{\sigma^0}\right|\geq r\right\}}\,\mathrm{d}s\right]\mathrm{d}r + g.\end{aligned}$$

Substituting in the value function from Proposition 2.5, we calculate that

$$\begin{aligned}v(x) &= \left(R^2 - |x|^2\right)\tilde{f}(|x|) + \int_{|x|}^R \left(R^2 - r^2\right)\tilde{f}'(r)\,\mathrm{d}r + g\\ &= R^2\tilde{f}(R) - |x|^2\,\tilde{f}(|x|) - \int_{|x|}^R r^2\tilde{f}'(r)\,\mathrm{d}r + g.\end{aligned}$$

Applying integration by parts, this simplifies to

$$v(x) = 2\int_{|x|}^R r\tilde{f}(r)\,\mathrm{d}r + g,$$

as required. □

We have seen that, for smooth increasing costs, the control σ^1 which enforces radial motion is always optimal, and for smooth decreasing costs, the control σ^0 which enforces tangential motion is optimal everywhere except at the origin. In the following sections, we will show that, for a continuous radially symmetric cost function with sufficient regularity, an optimal control is to switch between radial and tangential motion.

2.3 Explicit solution in the general case

In this section, we consider the control problem for more general radially symmetric cost functions, removing the restriction of monotonicity. We make the ansatz that the optimal strategy is to switch between two extreme behaviours in the control set, namely the strategies of tangential and radial motion defined in Definition 2.3 and Definition 2.7, respectively. In this way, we reduce the control problem to a one-dimensional optimal switching problem for the radius process. We use the principles of smooth and continuous fit to identify the optimal switching points, and we provide an algorithm to construct a candidate for the value function. We are able to write this function explicitly in Definition 2.14. We refer to the theory of viscosity solutions that we develop in Chapter 4 in order to verify that the candidate function is equal to the value function.

We make the following assumptions.

Assumption 2.11. Suppose that

1. The domain is $D = B_R(0) \subset \mathbb{R}^d$, for some $R > 0$ and $d \geq 2$;
2. The cost function f is radially symmetric; i.e. $f(x) = \tilde{f}(|x|)$, for some function $\tilde{f} : [0, R) \to \mathbb{R}$;
3. The boundary cost g is constant;
4. The cost function f is continuous;
5. There exists $\eta > 0$ such that the cost function $\tilde{f}$ is monotone on the interval $(0, \eta)$;
6. The one-sided derivative $\tilde{f}'_+(r)$ exists for all $r > 0$ and changes sign only finitely many times;
7. There exists $\delta > 0$ such that $\tilde{f}$ is continuously differentiable on $(0, \delta)$ and $\lim_{r\to 0} r\tilde{f}'(r) = 0$.

Remark 2.12. In Section 2.4, we will relax the fourth condition on continuity and the seventh condition on differentiability.

We rule out the case that the cost function oscillates at the origin by imposing the fifth condition on monotonicity. We will see in the following sections that the fifth and sixth conditions allow us to find an optimal strategy that switches between two regimes finitely many times. We believe that we would still be able to solve the control problem explicitly if we relax the fifth and sixth conditions, but in this case an optimal strategy may not exist. To simplify our exposition, we do not treat this case here.

Recall the definitions of the functions σ^0 in (2.3) and σ^1 in (2.5), which are associated to tangential and radial motion, respectively.

We conjecture that, in the case that $\tilde{f}$ is increasing at the origin, there exists a sequence of points $0 = s_0 < r_1 < s_1 < \ldots < R$ such that an optimal control is of the form

$$\sigma_t^\star = \begin{cases} \sigma^1\left(X_0^{\sigma^\star}\right), & \left|X_t^{\sigma^\star}\right| \in [0, r_1), \\ \sigma^0\left(X_t^{\sigma^\star}\right), & \left|X_t^{\sigma^\star}\right| \in [r_i, s_i], \quad i \geq 1, \\ \sigma^1\left(X_{\tau_{s_i}}^{\sigma^\star}\right), & \left|X_t^{\sigma^\star}\right| \in (s_i, r_{i+1}), \quad i \geq 1, \end{cases} \tag{2.10}$$

where, for each $i \geq 1$, we define the hitting time

$$\tau_{s_i} := \inf\left\{t \geq 0\colon \left|X_t^{\sigma^\star}\right| = s_i\right\}.$$

Note that $t \mapsto \left|X_t^{\sigma^\star}\right|$ is deterministically increasing when $\left|X_t^{\sigma^\star}\right| \in [r_i, s_i]$, for any $i \geq 1$, by Lemma 2.2. Therefore, if $\left|X_0^{\sigma^\star}\right| \geq r_1$, $\left|X_t^{\sigma^\star}\right| \geq r_1$ for all $t \geq 0$.

Similarly, if $\tilde{f}$ is decreasing at the origin, we conjecture that there is a sequence of points $0 = r_0 < s_0 < r_1 < \ldots < R$ such that an optimal control is of the form

$$\sigma_t^\star = \begin{cases} \sigma^0\left(X_t^{\sigma^\star}\right), & \left|X_t^{\sigma^\star}\right| \in (0, s_0], \\ \sigma^1\left(X_{\tau_{s_i}}^{\sigma^\star}\right), & \left|X_t^{\sigma^\star}\right| \in (s_{i-1}, r_i), \quad i \geq 1, \\ \sigma^0\left(X_t^{\sigma^\star}\right), & \left|X_t^{\sigma^\star}\right| \in [r_i, s_i], \quad i \geq 1. \end{cases} \tag{2.11}$$

Note that, in this second case, we do not make any claim about the optimal behaviour at the origin. Since $\sigma^0(0)$ is not defined, we will require some approximation at the origin in this case. We explore this further in Section 2.4 where we relax Assumption 2.11.

In either case, we conjecture that, at any time, an optimally controlled process should follow either radial motion or tangential motion, depending only on the current radial position of the process. We present a simulated trajectory of such a controlled process for an example with two switching points in Figure 2.6. In Proposition 2.15, we will prove that the control $\sigma^\star$ is optimal.

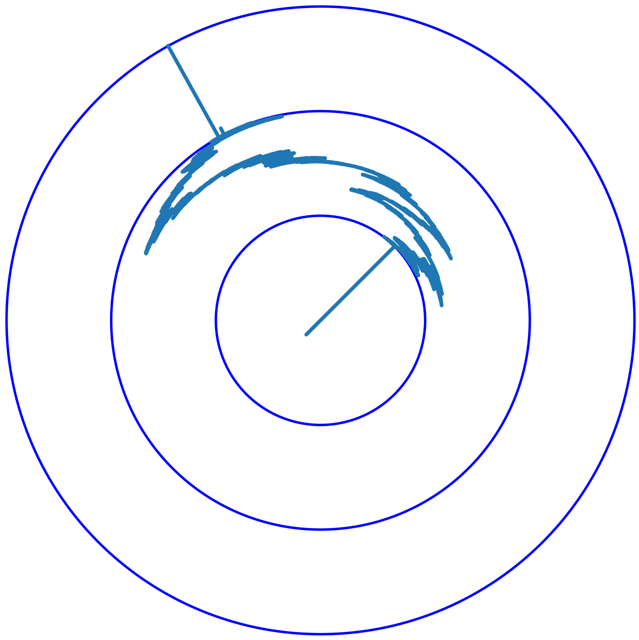


Figure 2.6: A sample path of an optimal controlled process in a case with two switching points

2.3.1 Reduction to a switching problem

Before beginning to construct a candidate for the value function, we give the following justification for our conjecture that switching between radial motion and tangential motion should be optimal. We will work with the radius of the controlled process in this section. We now derive an SDE for the radius process under some simplifying assumptions.

Proposition 2.13. *Let $\sigma \in \mathcal{U}$ be of the form*

$$\sigma_t = \left[\overline{\sigma}_t; \quad 0; \quad \ldots; \quad 0\right],$$

where $\overline{\sigma_t} \in \mathbb{R}^d$ with $|\overline{\sigma_t}| = 1$, for $t \geq 0$. Let $x \in D \setminus \{0\}$ and suppose that X^σ solves the SDE

$$\mathrm{d}X_t^\sigma = \sigma_t \,\mathrm{d}B_t,$$

with initial condition $X_0^\sigma = x$.

Set $r_0 := |x|$, let $\varepsilon \in (0, r)$, and define

$$\tau_\varepsilon := \inf\left\{t > 0\colon \left|R_t^\lambda - r_0\right| = \varepsilon\right\}.$$

Then there exists a $[0,1]$*-valued process* λ *such that* $|X_t^\sigma| = R_t^\lambda$*, where* R^λ *solves the SDE*

$$\mathrm{d}R_t^\lambda = \lambda_t \,\mathrm{d}W_t + \frac{1-\lambda_t^2}{2R_t^\lambda}\,\mathrm{d}t, \quad R_0^\lambda = r_0,$$

on the interval $[0,\tau_\varepsilon]$*, for a one-dimensional Brownian motion* W*.*

Proof. Let W be the first component of the Brownian motion B. Then the process X^σ solves

$$\mathrm{d}X_t^\sigma = \overline{\sigma}_t \,\mathrm{d}W_t,$$

with $X_0^\sigma = x$. Considering $t \in [0,\tau_\varepsilon]$, so that $X_t^\sigma \neq 0$, we can apply Itô's formula to find that the radius of X^σ satisfies the SDE

$$\mathrm{d}|X_t^\sigma| = |X_t^\sigma|^{-1}\,(X_t^\sigma)^\top \overline{\sigma}_t \,\mathrm{d}W_t + \frac{1}{2}\,|X_t^\sigma|^{-3}\,\mathrm{Tr}\left(\left[|X_t^\sigma|^2\, I - X_t^\sigma (X_t^\sigma)^\top\right] \overline{\sigma}_t \overline{\sigma}_t^\top\right)\mathrm{d}t. \quad (2.12)$$

Now let $(X_t^\sigma)^\perp$ denote the vector with norm $\left|(X_t^\sigma)^\perp\right| = |X_t^\sigma|$ that is orthogonal to the vector X_t^σ and satisfies

$$\overline{\sigma}_t = |X_t^\sigma|^{-1}\left(\lambda_t X_t^\sigma + \mu_t (X_t^\sigma)^\perp\right), \quad (2.13)$$

for some $\lambda_t, \mu_t \in \mathbb{R}$. Using the condition $|\overline{\sigma}_t| = 1$, we see that

$$1 = \lambda_t^2 + \mu_t^2,$$

and so $\lambda_t \in [0,1]$ and $\mu_t = \sqrt{1-\lambda_t^2}$.

Substituting the expression (2.13) for $\overline{\sigma}_t$ back into the SDE (2.12) for $|X^\sigma|$, and repeatedly using the identities $(X_t^\sigma)^\top X_t^\sigma = |X_t^\sigma|^2$ and $(X_t^\sigma)^\top (X_t^\sigma)^\perp = 0$, we have

$$\mathrm{d}|X_t^\sigma| = \lambda_t \,\mathrm{d}W_t + \frac{1}{2}\,|X_t^\sigma|^{-1}\left(1-\lambda_t^2\right)\mathrm{d}t.$$

Therefore, writing $R_t^\lambda = |X_t^\sigma|$, where λ_t is defined via (2.13), we arrive at the desired form of the SDE

$$\mathrm{d}R_t^\lambda = \lambda_t \,\mathrm{d}W_t + \frac{1-\lambda_t^2}{2R_t^\lambda}\,\mathrm{d}t.$$

□

Now, suppose further that the process λ in the proof of Proposition 2.13 takes the form

$$\lambda_t = \lambda(R_t^\lambda), \quad t \geq 0.$$

Then we can write down the infinitesimal generator $\mathcal{L}^\lambda$ for the process R^λ as

$$\mathcal{L}^\lambda u(r) = -\frac{1}{2}\lambda^2(r) u''(r) - \frac{1-\lambda^2(r)}{2r} u'(r),$$

for $r \in (r_0 - \varepsilon, r_0 + \varepsilon)$ and any smooth function $u \in C^2((r_0 - \varepsilon, r_0 + \varepsilon), \mathbb{R})$.

Consider the following simplification of the control problem. Restrict the control set to contain only those controls that give rise to a process λ of the form specified above. Let $v^R : D \to \mathbb{R}$ be the value function of this simplified problem. By radial symmetry, we can write

$$v^R(x) = \tilde{v}^R(|x|),$$

for some $\tilde{v}^R : [0, R) \to \mathbb{R}$. Supposing that $\tilde{v}^R$ is twice continuously differentiable, we expect $\tilde{v}^R$ to be a classical solution of a Hamilton-Jacobi-Bellman equation, as described in Section 1.4.4. By the results of Section 3.3 of [58], $\tilde{v}^R$ should solve

$$\inf_\lambda \mathcal{L}^\lambda \tilde{v}^R = \tilde{f},$$

in the interval $(r_0 - \varepsilon, r_0 + \varepsilon)$, where the infimum is taken over functions $\lambda : (r_0 - \varepsilon, r_0 + \varepsilon) \to [0, 1]$.

Note that we can rewrite the generator as

$$\mathcal{L}^\lambda \tilde{v}^R(r) = -\frac{1}{2r}(\tilde{v}^R)'(r) - \frac{r}{2}\lambda^2(r)\left[\frac{1}{r}(\tilde{v}^R)'(r)\right]'.$$

Hence, at points r such that $\left[\frac{1}{r}(\tilde{v}^R)'(r)\right]' > 0$, the infimum is attained for $\lambda(r) = 1$, while at points r such that $\left[\frac{1}{r}(\tilde{v}^R)'(r)\right]' < 0$, the infimum is attained for $\lambda(r) = 0$. At a point r such that $\left[\frac{1}{r}(\tilde{v}^R)'(r)\right]' = 0$, the infimum is attained for any value $\lambda(r) \in [0, 1]$.

Returning to the expression for $\overline{\sigma}$ in terms of λ in (2.13), we see that setting $\lambda_t = 1$ gives $\overline{\sigma}_t = \frac{X_t^\sigma}{|X_t^\sigma|}$, with generator $\mathcal{L}^1$ given by

$$\mathcal{L}^1 u(r) = -\frac{1}{2}u''(r). \tag{2.14}$$

Note that, away from the origin, a controlled process following this control has the same behaviour as radial motion, as defined in Definition 2.7. On the other hand, $\lambda = 0$ corresponds to tangential motion, as defined in Definition 2.3, with generator

$$\mathcal{L}^0 u(r) = -\frac{1}{2r}u'(r). \tag{2.15}$$

Therefore the above calculations support our claim that the optimal strategy should be to switch between these two behaviour regimes.

We note that, in the above discussion, we restricted the control set and made the strong assumption that the value function is twice continuously differentiable.

In order to prove that the behaviour described above is in fact optimal without these restrictions, we will need to refer to the theory of viscosity solutions for HJB equations that we develop in Chapter 4.

We now identify the conjectured optimal switching points and construct a candidate for the value function, before proving optimality in Proposition 2.15.

2.3.2 Optimal switching points

With the justification of the previous section, we make the ansatz that the optimal strategy is of the form described in (2.10) or in (2.11). We now seek to find the optimal switching points r_i and s_i.

We will find that we require continuous fit and a condition on the first derivative to fix the points r_i, and we will need to impose smooth fit and a condition on the second derivative to fix the points s_i. It is interesting to note that smooth fit also holds at the points r_i, although we do not enforce it.

Under the conjectured optimal behaviour, the value function is of the form

$$v(x) = \tilde{v}(|x|), \quad x \in D,$$

for some $\tilde{v} : [0, R) \to \mathbb{R}$. To identify the optimal switching points, we will assume that $\tilde{v}$ is differentiable in the interval $(0, R)$ and satisfies the boundary condition $\tilde{v}(R) = g$. Then, for any $r \in (0, R)$, we have

$$\tilde{v}(r) = g - \int_r^R \tilde{v}'(s)\,\mathrm{d}s.$$

When we verify our candidate for the value function in Proposition 2.15, we will show that v is in fact continuously differentiable in D and attains the boundary condition $v = g$ on ∂D.

By definition of the value function, the expected cost associated to any admissible control at some radius $r \in (0, R)$ is greater than the value $\tilde{v}(r)$. Therefore the derivative of such an expected cost at some $r \in (0, R)$ must be less than the derivative of the value function $\tilde{v}'(r)$. We will use this observation to determine the optimal switching points.

Let $\tilde{V} : [0, R] \to \mathbb{R}$ and define a candidate value function $V : D \to \mathbb{R}$ by $V(x) = \tilde{V}(|x|)$ for $x \in D$. The first step in constructing this function V is to find the optimal switching points, as follows.

Suppose that there exists some $i \geq 1$ such that $0 < s_{i-1} < r_i < R$. Then we expect that the optimal control switches from tangential motion to radial motion at the point s_{i-1}. In some interval (s, s_{i-1}), we set $\tilde{V} = w_{i-1}$, where w_{i-1} solves the ODE

$$\mathcal{L}^0 w_{i-1}(r) = -2r\tilde{f}(r),$$

and $\mathcal{L}^0$ is the generator associated to tangential motion that is defined in (2.15). This ODE is equivalent to the first order ODE

$$w'_{i-1}(r) = -2r\tilde{f}(r).$$

In the interval (s_{i-1}, r_i), we set $\tilde{V} = u_i$, where u_i solves the ODE

$$\mathcal{L}^1 u_i(r) = \tilde{f}(r),$$

and $\mathcal{L}^1$ is the generator associated to radial motion that is defined in (2.14). We can write this ODE as

$$u''_i(r) = -2\tilde{f}(r).$$

We fix the boundary conditions

$$u_i(s_{i-1}) = w_{i-1}(s_{i-1}), \quad \text{and} \quad u'_{i+}(s_{i-1}) = w'_{i-1}(s_{i-1}) = -2s_{i-1}\tilde{f}(s_{i-1}),$$

to define u_i uniquely.

Now, in the interval $(r_i, s_i \wedge R)$, we suppose that tangential motion is optimal and set $\tilde{V} = w_i$, where w_i solves the first order ODE

$$w'_i(r) = -2r\tilde{f}(r).$$

We then have the following free boundary problem:

$$\begin{cases} \tilde{V}''(r) = -2\tilde{f}(r), & r \in (s_{i-1}, r_i), \\ \tilde{V}'(r) = -2r\tilde{f}(r), & r \in (r_i, s_i \wedge R), \\ \tilde{V}(r_i+) = \tilde{V}(r_i-), \end{cases} \tag{2.16}$$

where the point r_i is to be found. Note that we require the continuous fit condition at r_i in order to solve the first order ODE in $(r_i, s_i \wedge R)$.

As noted above, we determine the switching point by comparing the derivatives of u_i and w_i. The point r_i should be the first point at which $w'_i(r) = -2r\tilde{f}(r)$ is greater than the first derivative of u_i. Therefore we define r_i by

$$r_i := \inf \left\{ r > s_{i-1} \colon s_{i-1}\tilde{f}(s_{i-1}) + \int_{s_{i-1}}^{r} \tilde{f}(s)\,\mathrm{d}s > r\tilde{f}(r) \right\}.$$

That is the first point after s_{i-1} at which the running average of the cost function becomes greater than its current value. Note that this point cannot be in a region where $\tilde{f}$ is increasing and so r_i is greater than or equal to the first point of decrease of $\tilde{f}$ after s_{i-1}.

In Figure 2.7b, we show an example of choosing the switching point r_1 by comparing derivatives. We see in Figure 2.7a that, for this example, the switching point r_1 is strictly greater than the turning point at which the cost function starts to decrease. Also note that, although we have only imposed continuous fit at the point r_1, we can see in Figure 2.7b that the derivatives are equal at r_1. For any continuous cost function, this smooth fit property arises in the same way; we will discuss this in detail in Section 2.3.4.

Let us now suppose that $s_i < R$. We suppose that, in the interval $(s_i, r_{i+1} \wedge R)$, radial motion is once again optimal, and we set $\tilde{V} = u_{i+1}$, where u_{i+1} solves the second order ODE

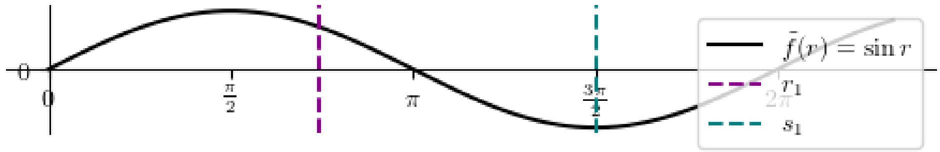
$$u''_{i+1}(r) = -2\tilde{f}(r).$$

Then we have a second free boundary problem

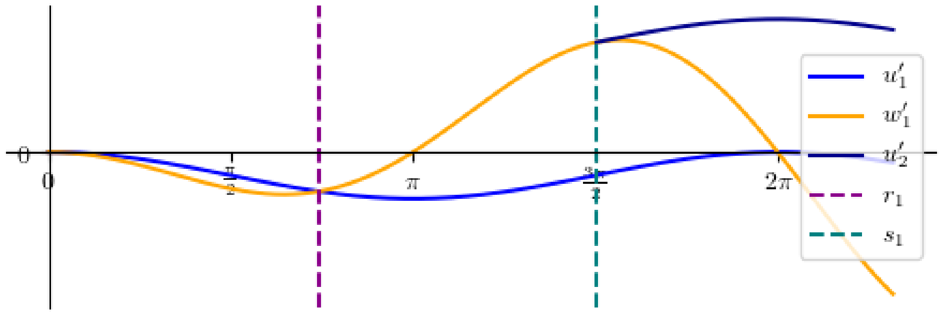
$$\begin{cases} \tilde{V}'(r) = -2r\tilde{f}(r), & r \in (r_i, s_i), \\ \tilde{V}''(r) = -2\tilde{f}(r), & r \in (s_i, r_{i+1} \wedge R), \\ \tilde{V}(s_i+) = \tilde{V}(s_i-), \\ \tilde{V}'_+(s_i) = \tilde{V}'_-(s_i), \end{cases} \tag{2.17}$$

where the point s_i is to be found. Here we require both smooth fit and continuous fit at the point s_i in order to solve the second order ODE in the interval $(s_i, r_{i+1} \wedge R)$.

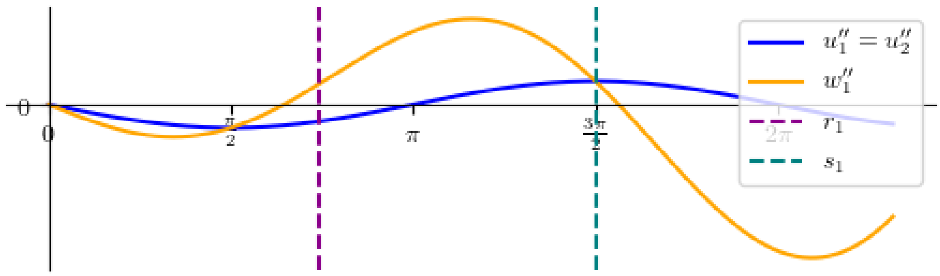
Having imposed the smooth fit condition $\tilde{V}'_+(s_i) = \tilde{V}'_-(s_i)$, the first derivatives of solutions of $w'_i(r) = -2r\tilde{f}(r)$ and $u''_{i+1}(r) = -2\tilde{f}(r)$ are equal for any choice of s_i. In order to fix the point s_i, we require a second order condition. Recall from Assumption 2.11 that we assume that the right derivative of $\tilde{f}$ exists everywhere. This allows us to define s_i to be the first point at which $u''_{i+1}(r) = -2\tilde{f}(r)$ is greater than the one-sided second derivative from the right of the solution of $w'_i(r) = -2r\tilde{f}(r)$. Thus there is an interval of positive length on which the first derivatives are in this same order.



(a) Radial part of the cost function $\tilde{f}(r) = \sin r$



(b) First derivatives of the expected costs u_1, w_1, and u_2



(c) Second derivatives of the expected costs u_1, w_1, and u_2

Figure 2.7: The first two switching points r_1, s_1 are shown for the cost function $f(x) = \sin|x|$. The switching point r_1 is the first point at which $w_1'(r) = -2r\tilde{f}(r)$ exceeds u_1', where u_1 solves $u_1''(r) = -2\tilde{f}(r)$, with ${u_1}'_+(0) = 0$, as shown in (b). The switching point s_1 is the first point after r_1 at which $u_2'' = -2\tilde{f}$ exceeds w_1'', as shown in (c). Fixing $u_2'(s_1) = w_1'(s_1)$, we see in (b) that s_1 is chosen such that u_2' remains greater than w_1' over an interval of positive length.

We can calculate the one-sided second derivative from the right of w_i as

$$w_i''{}_+(r) = -2\tilde{f}(r) - 2r\tilde{f}_+'(r).$$

This leads us to define s_i by

$$s_i := \inf\left\{s > r_i \colon \tilde{f}_+'(s) > 0\right\}.$$

In this case, the switching point is exactly the turning point at which $\tilde{f}$ starts to increase. For the example in Figure 2.7, we can see that the switching point s_1 does indeed coincide with this turning point. Figure 2.7c shows how this switching point is chosen by comparing second derivatives, and Figure 2.7b shows that the first derivatives at this point have the desired properties.

Note that the sixth condition of Assumption 2.11 implies that there are finitely many switching points s_i and thus finitely many points r_i. Taking the above definitions of r_i and s_j for all values of i, j such that $r_i, s_j < R$, we now solve the ODEs in (2.16) and (2.17) to construct a candidate for the value function.

2.3.3 Construction of the value function

In this section we construct the candidate function V, which we will go on to prove is equal to the value function. We break the construction down into two cases depending on the behaviour of the cost function at the origin, and then into two further sub-cases depending on the behaviour of the cost function at the boundary of the domain.

Case I: Increasing cost at the origin

Suppose first that $\tilde{f}$ is increasing on the interval $(0, \eta)$. We summarise the construction of the candidate value function in this case in Algorithm 1.

Fix $s_0 = 0$. Since we expect the optimal control to enforce radial motion in the ball $B_\eta(0)$, we solve the second order ODE

$$u_1''(r) = -2\tilde{f}(r), \quad r \in (0, R).$$

We require two boundary conditions in order to uniquely define the solution u_1. We impose the boundary condition ${u_1}'_+(0) = 0$ for the following reasons.

First, from the discussion in the previous section, we recall that we will define the first switching point to be

$$r_1 = \inf\left\{r > 0\colon u_1'(r) < -2r\tilde{f}(r)\right\},$$

since we are seeking to maximise the derivative of the candidate value function. Therefore, for $r \in (0, r_1)$, we must have $u_1'(r) \geq -2r\tilde{f}(r)$ and, in particular

$${u_1}'_+(0) = \lim_{r \downarrow 0} u_1'(r) \geq -2 \lim_{r \downarrow 0} r\tilde{f}(r) = 0.$$

To get the opposite inequality, fix $\delta \in (0, \eta)$ and $r \in (0, \delta)$ and apply Itô's formula

Algorithm 1 Construction of the value function in Case I

Define $s_0 = 0$.
Solve $u_1''(r) = -2\tilde{f}(r)$, with ${u_1}'_+(0) = 0$, $u_1(0) = \alpha$, for some $\alpha \in \mathbb{R}$.
Define $r_1 := \inf\left\{r > 0\colon\ \int_0^r \tilde{f}(s)\,\mathrm{d}s > r\tilde{f}(r)\right\}$.
Set $\tilde{V} = u_1$ on $(0, r_1 \wedge R]$.
if $r_1 < R$ **then**
  for $i \geq 1$ **do**
    Solve $w_i'(r) = -2r\tilde{f}(r)$, with $w_i(r_i) = u_i(r_i)$.
    Define $s_i := \inf\left\{r > r_i\colon\ \tilde{f}'_+(s) > 0\right\}$.
    Set $\tilde{V} = w_i$ on $(r_i, s_i \wedge R]$.
    if $s_i \geq R$ **then**
      break
    end if
    Solve $u_{i+1}''(r) = -2\tilde{f}(r)$, with $u_{i+1}'(s_i+) = -2s_i\tilde{f}(s_i)$ and $u_{i+1}(s_i) = w_i(s_i)$.
    Define $r_{i+1} := \inf\left\{r > s_i\colon\ s_i\tilde{f}(s_i) + \int_{s_i}^r \tilde{f}(s)\,\mathrm{d}s > r\tilde{f}(r)\right\}$.
    Set $\tilde{V} = u_{i+1}$ on $(s_i, r_{i+1} \wedge R]$.
    if $r_{i+1} \geq R$ **then**
      break
    end if
  end for
end if
Fix α such that $\tilde{V}(R) = g$.

to $u_1(\delta) = u_1\left(\left|X^{\sigma^1}_{\tau_\delta}\right|\right)$ to see that

$$\begin{aligned} u_1(\delta) - u_1(r) &= \mathbb{E}^r\left[u_1\left(\left|X^{\sigma^1}_{\tau_\delta}\right|\right)\right] - u_1(r) \\ &= \frac{1}{2}\mathbb{E}^r\left[\int_0^{\tau_\delta} u_1''\left(\left|X^{\sigma^1}_s\right|\right)\mathrm{d}s\right] = -\mathbb{E}^r\left[\int_0^{\tau_\delta} \tilde{f}\left(\left|X^{\sigma^1}_s\right|\right)\mathrm{d}s\right]. \end{aligned}$$

Then, applying dominated convergence to take the limit as $r \downarrow 0$, and using the fact that $\tilde{f}$ is increasing, we have that

$$\begin{aligned} \lim_{r\downarrow 0}\frac{1}{\delta}\left(u_1(\delta) - u_1(r)\right) &= -\frac{1}{\delta}\mathbb{E}^0\left[\int_0^{\tau_\delta} \tilde{f}\left(\left|X^{\sigma^1}_s\right|\right)\mathrm{d}s\right] \\ &\leq -\frac{1}{\delta}\tilde{f}(0)\mathbb{E}^0[\tau_\delta] \\ &= -\delta\tilde{f}(0). \end{aligned}$$

Hence

$$0 \leq {u_1}'_+(0) \leq -\lim_{\delta\downarrow 0}\delta\tilde{f}(0) = 0.$$

As well as imposing the above condition on the first derivative, we also fix an arbitrary value $u_1(0) = \alpha \in \mathbb{R}$. Having constructed the candidate value function, up to this arbitrary constant, on the whole domain, we will use the external boundary condition $\tilde{V}(R) = g$ to determine the value of α. We now have

$$u_1(r) = \alpha - 2\int_0^r \int_0^s \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s.$$

Define

$$r_1 := \inf\left\{r > 0\colon \int_0^r \tilde{f}(s)\,\mathrm{d}s > r\tilde{f}(r)\right\},$$

and set $\tilde{V}(r) = u_1(r)$ for $r \in (0, r_1 \wedge R]$.

If $r_1 < R$, we then expect the optimal control to switch to enforcing tangential motion. Therefore we solve the first order ODE

$$w_1'(r) = -2r\tilde{f}(r), \quad r \in (r_1, R).$$

In order to uniquely define the solution w_1, we impose the continuous fit condition $w_1(r_1) = \tilde{V}(r_1)$. Then we have

$$\begin{aligned} w_1(r) &= \tilde{V}(r_1) - 2\int_{r_1}^r s\tilde{f}(s)\,\mathrm{d}s \\ &= \alpha - 2\int_{r_1}^r s\tilde{f}(s)\,\mathrm{d}s - 2\int_0^{r_1}\int_0^s \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s. \end{aligned}$$

Now define

$$s_1 := \inf\left\{r > r_1\colon \tilde{f}_+'(r) > 0\right\},$$

and set $\tilde{V}(r) = w_1(r)$ for $r \in (r_1, s_1 \wedge R]$.

If $s_1 < R$, then we expect the optimal control to switch back to enforcing radial motion, and so we solve the second order ODE

$$u_2''(r) = -2\tilde{f}(r), \quad r \in (s_1, R).$$

At this point, we impose both the continuous fit condition $u_2(s_1) = \tilde{V}(s_1)$ and the smooth fit condition ${u_2}_+'(s_1) = \tilde{V}'(s_1)$ in order to uniquely define u_2. We then find that

$$u_2'(r) = \tilde{V}'(s_1) - 2\int_{s_1}^r \tilde{f}(s)\,\mathrm{d}s,$$

and so

$$\begin{aligned} u_2(r) &= \tilde{V}(s_1) + (r - s_1)\tilde{V}'(s_1) - 2\int_{s_1}^{r}\int_{s_1}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s \\ &= \alpha - 2\int_{s_1}^{r}\int_{s_1}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s - 2\int_{0}^{r_1}\int_{0}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s \\ &\quad - 2\int_{r_1}^{s_1} s\tilde{f}(s)\,\mathrm{d}s - 2(r_1 - s_1)r_1\tilde{f}(r_1). \end{aligned}$$

Defining

$$r_2 := \inf\left\{r > s_1\colon\, s_1\tilde{f}(s_1) + \int_{s_1}^{r} \tilde{f}(s)\,\mathrm{d}s > r\tilde{f}(r)\right\},$$

we set $\tilde{V}(r) = u_2(r)$ for $r \in (s_1, r_2 \wedge R]$.

We continue in this way until reaching the boundary of the domain, setting

$$\tilde{V}(r) = \begin{cases} u_i(r), & r \in (s_{i-1}, r_i \wedge R], \\ w_i(r), & r \in (r_i, s_i \wedge R], \end{cases}$$

for each $i \geq 1$.

Fixing $i \geq 2$, for $r \in (s_{i-1}, r_i \wedge R]$, we calculate that

$$\begin{aligned} u_i(r) &= w_{i-1}(s_{i-1}) + (r - s_{i-1})w_{i-1}'(s_{i-1}) - 2\int_{s_{i-1}}^{r}\int_{s_{i-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s \\ &= u_{i-1}(r_{i-1}) - 2\int_{r_{i-1}}^{s_{i-1}} s\tilde{f}(s)\,\mathrm{d}s \\ &\quad - 2(r - s_{i-1})s_{i-1}\tilde{f}(s_{i-1}) - 2\int_{s_{i-1}}^{r}\int_{s_{i-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s. \end{aligned}$$

Noting that

$$u_1(r_1) = \alpha - 2\int_{0}^{r_1}\int_{0}^{s} \tilde{f}(t)\,\mathrm{d}t,$$

we calculate recursively that

$$\begin{aligned} u_i(r) = \alpha &- 2\int_{0}^{r_1}\int_{0}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s + 2(r_i - r)s_{i-1}\tilde{f}(s_{i-1}) + 2\int_{r}^{r_i}\int_{s_{i-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s \\ &- 2\sum_{j=2}^{i}\left[(r_j - s_{j-1})s_{j-1}\tilde{f}(s_{j-1}) + \int_{r_{j-1}}^{s_{j-1}} s\tilde{f}(s)\,\mathrm{d}s + \int_{s_{j-1}}^{r_j}\int_{s_{j-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s\right]. \end{aligned} \tag{2.18}$$

Now, for $r \in (r_i, s_i \wedge R]$, we calculate that

$$\begin{aligned} w_i(r) &= u_i(r_i) - 2\int_{r_i}^{r} s\tilde{f}(s)\,\mathrm{d}s \\ &= w_{i-1}(s_{i-1}) - 2(r_i - s_{i-1})s_{i-1}\tilde{f}(s_{i-1}) - 2\int_{s_{i-1}}^{r_i}\int_{s_{i-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s \\ &\quad - 2\int_{r_i}^{r} s\tilde{f}(s)\,\mathrm{d}s. \end{aligned}$$

We note that

$$\begin{aligned} w_1(s_1) &= u_1(r_1) - 2\int_{r_1}^{s_1} s\tilde{f}(s)\,\mathrm{d}s \\ &= \alpha - 2\int_0^{r_1}\int_0^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s - 2\int_{r_1}^{s_1} s\tilde{f}(s)\,\mathrm{d}s, \end{aligned}$$

and calculate recursively that

$$\begin{aligned} w_i(r) = \alpha &- 2\int_0^{r_1}\int_0^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s - 2\int_{r_i}^{r} s\tilde{f}(s)\,\mathrm{d}s \\ &- 2\sum_{j=2}^{i}\left[(r_j - s_{j-1})s_{j-1}\tilde{f}(s_{j-1}) + \int_{s_{j-1}}^{r_j}\int_{s_{j-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s + \int_{r_{j-1}}^{s_{j-1}} s\tilde{f}(s)\,\mathrm{d}s\right]. \end{aligned} \tag{2.19}$$

In order to determine the value of α, we use the boundary condition on ∂D. Let $K \in \mathbb{N}$ be such that $R \in (s_{K-1}, s_K]$. We consider the following two sub-cases, depending on the behaviour at the boundary.

Radial motion at the boundary: Suppose that $R \in (s_{K-1}, r_K]$. Then we expect radial motion to be optimal close to the boundary of the domain, and we have $\tilde{V}(r) = u_K(r)$ for $r \in (s_{K-1}, R]$.

Imposing the boundary condition $V(x) = g$ for $x \in \partial D$, we have $u_K(R) = g$. Setting $i = K$ and $r = R$ in (2.18), we find that

$$\begin{aligned} g = \alpha - 2\int_0^{r_1}\int_0^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s + 2(r_K - R)s_{K-1}\tilde{f}(s_{K-1}) + 2\int_R^{r_K}\int_{s_{K-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s \\ - 2\sum_{j=2}^{K}\left[(r_j - s_{j-1})s_{j-1}\tilde{f}(s_{j-1}) + \int_{s_{j-1}}^{r_j}\int_{s_{j-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s + \int_{r_{j-1}}^{s_{j-1}} s\tilde{f}(s)\,\mathrm{d}s\right]. \end{aligned}$$

We can now substitute the value of α into (2.18) and (2.19) to find closed form expressions for the values

$$\begin{aligned} u_i(r) &= 2(r_i - r)s_{i-1}\tilde{f}(s_{i-1}) + 2\int_r^{r_i}\int_{s_{i-1}}^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s \\ &\quad + 2\sum_{j=i+1}^{K}\left[(r_j - s_{j-1})s_{j-1}\tilde{f}(s_{j-1}) + \int_{r_{j-1}}^{s_{j-1}} s\tilde{f}(s\,\mathrm{d}s) + \int_{s_{j-1}}^{r_j}\int_{s_{j-1}}^{s}\tilde{f}(t)\,\mathrm{d}t\right] \\ &\quad - 2(r_K - R)s_{K-1}\tilde{f}(s_{K-1}) - 2\int_R^{r_K}\int_{s_{K-1}}^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s + g, \end{aligned}$$

for $r \in (s_{i-1}, r_i \wedge R]$, $i = 1, \ldots, K$, and

$$\begin{aligned} w_i(r) &= 2\int_r^{s_i} s\tilde{f}(s)\,\mathrm{d}s + 2(R - s_{K-1})s_{K-1}\tilde{f}(s_{K-1}) + 2\int_{s_{K-1}}^{R}\int_{s_{K-1}}^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s + g \\ &\quad + 2\sum_{j=i+1}^{K-1}\left[(r_j - s_{j-1})s_{j-1}\tilde{f}(s_{j-1}) + \int_{r_j}^{s_j} s\tilde{f}(s)\,\mathrm{d}s + \int_{s_{j-1}}^{r_j}\int_{s_{j-1}}^{s}\tilde{f}(t)\,\mathrm{d}t\right], \end{aligned}$$

for $r \in (r_i, s_i]$, $i = 1, \ldots, K-1$.

Tangential motion at the boundary: Now suppose that $R \in (r_K, s_K]$, so that we expect tangential motion to be optimal close to the boundary of the domain. Then we have $\tilde{V}(r) = w_K(r)$ for $r \in (r_K, R]$.

Imposing the boundary condition $V(x) = g$ for $x \in \partial D$, we have $w_K(R) = g$. Setting $i = K$ and $r = R$ in (2.19), we find that

$$\begin{aligned} g &= \alpha - 2\int_0^{r_1}\int_0^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s - 2\int_{r_K}^{R} s\tilde{f}(s)\,\mathrm{d}s \\ &\quad - 2\sum_{i=2}^{K}\left[(r_i - s_{i-1})s_{i-1}\tilde{f}(s_{i-1}) + \int_{r_{i-1}}^{s_{i-1}} s\tilde{f}(s)\,\mathrm{d}s + \int_{s_{i-1}}^{r_i}\int_{s_{i-1}}^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s\right]. \end{aligned}$$

Having found the value of α, we can substitute this into (2.18) and (2.19) to find closed form expressions for the values

$$\begin{aligned} u_i(r) &= 2(r_i - r)s_{i-1}\tilde{f}(s_{i-1}) + 2\int_r^{r_i}\int_{s_{i-1}}^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s + 2\int_{r_K}^{R} s\tilde{f}(s)\,\mathrm{d}s + g \\ &\quad + 2\sum_{j=i+1}^{K}\left[(r_j - s_{j-1})s_{j-1}\tilde{f}(s_{j-1}) + \int_{r_{j-1}}^{s_{j-1}} s\tilde{f}(s)\,\mathrm{d}s + \int_{s_{j-1}}^{r_j}\int_{s_{j-1}}^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s\right], \end{aligned}$$

for $r \in (s_{i-1}, r_i]$, $i = 1, \cdots, K$, and

$$w_i(r) = 2\int_r^{s_i} s\tilde{f}(s)\,\mathrm{d}s - 2\int_R^{s_K} s\tilde{f}(s)\,\mathrm{d}s + g$$
$$+ 2\sum_{j=i+1}^{K}\left[(r_j - s_{j-1})s_{j-1}\tilde{f}(s_{j-1}) + \int_{r_j}^{s_j} s\tilde{f}(s)\,\mathrm{d}s + \int_{s_{j-1}}^{r_j}\int_{s_{j-1}}^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s\right],$$

for $r \in (r_i, s_i \wedge R]$, $i = 1, \cdots, K$.

We summarise the candidate value function in Definition 2.14 below.

Case II: Decreasing cost at the origin

We now turn to the second case where $\tilde{f}$ is decreasing on the interval $(0, \eta)$. We summarise the construction of the candidate value function in this case in Algorithm 2.

Algorithm 2 Construction of the value function in Case II

Define $r_0 = 0$.
Solve $w_0'(r) = -2r\tilde{f}(r)$, with $w_0(r) = \alpha$, for some $\alpha \in \mathbb{R}$.
Define $s_0 := \inf\left\{r > 0\colon \tilde{f}_+'(r) > 0\right\}$.
Set $\tilde{V} = w_0$ on $(0, s_0 \wedge R]$.
if $s_0 < R$ **then**
  for $i \geq 0$ **do**
    Solve $u_{i+1}''(r) = -2\tilde{f}(r)$, with $u_{i+1}'(s_i+) = -2s_i\tilde{f}(s_i)$ and $u_{i+1}(s_i) = w_i(s_i)$.
    Define $r_{i+1} := \inf\left\{r > s_i\colon s_i\tilde{f}(s_i) + \int_{s_i}^r \tilde{f}(s)\,\mathrm{d}s > r\tilde{f}(r)\right\}$.
    Set $\tilde{V} = u_{i+1}$ on $(s_i, r_{i+1} \wedge R]$.
    if $r_{i+1} \geq R$ **then**
      break
    end if
    Solve $w_{i+1}'(r) = -2r\tilde{f}(r)$, with $w_{i+1}(r_{i+1}) = u_{i+1}(r_{i+1})$.
    Define $s_{i+1} := \inf\left\{r > r_{i+1}\colon \tilde{f}_+'(r) > 0\right\}$.
    Set $\tilde{V}(R) = g$ on $(r_{i+1}, s_{i+1} \wedge R]$.
    if $s_{i+1} \geq R$ **then**
      break
    end if
  end for
end if
Fix α such that $\tilde{V}(R) = g$.

We expect the optimal control to enforce tangential motion in $B_\eta(0) \setminus B_\varepsilon(0)$, for any $\varepsilon \in (0, \eta)$. As we will see in Section 2.4, it will be possible to define a control at the origin whose cost approximates the cost associated to tangential motion. Without further justification here, we fix $r_0 = 0$ and seek the solution w_0 to the first order ODE

$$w_0'(r) = -2r\tilde{f}(r), \quad r \in (0, R).$$

Note that this ODE fixes the first derivative and, in particular, ${w_1}'_+(0) = 0$. In order to uniquely define w_1, we need to impose one boundary condition. As in the previous section, we will fix an arbitrary value $w_1(0) = \alpha \in \mathbb{R}$, and we will determine the value of α from the external boundary condition $\tilde{V}(R) = g$, once we have constructed the candidate value function on the whole domain.

The construction of the value function proceeds in the same way as in Case I, and we omit the details here. We state the candidate value function in both cases in the following Definition 2.14.

Definition 2.14 (Candidate value function)**.** Let the cost functions f and g be as in Assumption 2.11. For $k \in \mathbb{N}$ and $i = 0, \dots, k$, define the constant

$$\mathfrak{F}_i^k := 2 \sum_{j=i+1}^{k} \left[(r_j - s_{j-1}) s_{j-1} \tilde{f}(s_{j-1}) + \int_{s_{j-1}}^{r_j} \int_{s_{j-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s + \int_{r_j}^{s_j} s\tilde{f}(s)\,\mathrm{d}s \right].$$

Then we define the candidate value function $V : D \to \mathbb{R}$ as follows.

Case I: If $\tilde{f}$ is increasing in $(0, \eta)$, then set $s_0 = 0$ and let $K \in \mathbb{N}$ be such that $R \in (s_{K-1}, s_K]$. For $x \in D$, define

$$\begin{aligned} V(x) = g &- 2\int_{R\vee r_K}^{s_K} s\tilde{f}(s)\,\mathrm{d}s \\ &- 2(r_K - R \wedge r_K) s_{K-1} \tilde{f}(s_{K-1}) - 2 \int_{R\wedge r_K}^{r_K} \int_{s_{K-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s \\ &+ 2 \sum_{i=1}^{K} \mathbb{1}_{\{(s_{i-1}, s_i]\}}(|x|) \left[(r_i - |x| \wedge r_i) s_{i-1} \tilde{f}(s_{1-1}) + \int_{|x|\wedge r_i}^{r_i} \int_{s_{i-1}}^{s} \tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s \right. \\ &\qquad\qquad \left. + \int_{|x|\vee r_i}^{s_i} s\tilde{f}(s)\,\mathrm{d}s + \mathfrak{F}_i^K \right]. \end{aligned}$$

Case II: If $\tilde{f}$ is decreasing in $(0,\eta)$, then set $r_0 = 0$ and let $L \in \mathbb{N}$ be such that $R \in (r_L, r_{L+1}]$. For $x \in D$, define

$$\begin{aligned} V(x) &= g - 2\int_{R\wedge s_L}^{s_L} s\tilde{f}(s)\,\mathrm{d}s \\ &\quad + 2(R\vee s_L - s_L)s_L\tilde{f}(s_L) + 2\int_{s_L}^{R\vee s_L}\int_{s_L}^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s \\ &\quad + 2\sum_{i=0}^{L}\mathbb{1}_{\{(r_i,r_{i+1}]\}}(|x|)\left[\int_{|x|\wedge s_i}^{s_i} s\tilde{f}(s)\,\mathrm{d}s - (|x|\vee s_i - s_i)s_i\tilde{f}(s_i)\right. \\ &\qquad\qquad \left. - \int_{s_i}^{|x|\vee s_i}\int_{s_i}^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s + \mathfrak{F}_i^L\right]. \end{aligned}$$

Before turning to the rigorous proof of optimality in Section 2.3.5, we make a digression to discuss the smooth fit property that the candidate value function exhibits.

2.3.4 The principle of smooth fit

In the preceding construction, the smooth fit condition is required to fix the switching points s_i. It is notable, however, that we do not need to impose smooth fit to uniquely identify the points r_i, but the smooth fit condition is nevertheless satisfied at these switching points.

The principle of smooth fit is commonly used in optimal stopping problems, as described in Section 9.1 of [47]. For a continuous $\mathbb{R}$-valued diffusion process, the optimal stopping time is the first exit time of some interval. The interval is chosen such that the value function dominates the cost function everywhere and matches both the value and the first derivative of the cost function at the end points.

In [48], Pham shows how the smooth fit property arises in a one-dimensional switching problem, similar to the problem that we are studying in this chapter. Pham proves that smooth fit holds using the theory of viscosity solutions, under the assumption that the underlying stochastic process has strictly positive diffusivity in each regime.

In our problem, the controlled radius process behaves locally like a Brownian motion in the regime of radial motion. Therefore, at the points s_i, where the optimal behaviour switches to radial motion, the conditions are met for Pham's result to hold. This justifies the smooth fit condition at the switching points s_i. However, in the regime of tangential motion, the controlled radius process is deterministic. Therefore we cannot apply Pham's reasoning to justify smooth fit at the points r_i

where the optimal behaviour switches to this regime. It is interesting to note that, although Pham's justification from [48] breaks down at the points r_i, smooth fit still holds at these switching points.

2.3.5 Proof of optimality

We now turn to the proof that the candidate function that we have constructed is indeed the value function.

Proposition 2.15. *Under Assumption 2.11, the value function v is continuously differentiable and takes the form $v = V$, where V is defined in Definition 2.14.*

Moreover, there exists an optimal control $\sigma^\star \in \mathcal{U}$ in the following cases. If $\tilde{f}$ is increasing in $(0, \eta)$, then the control $\sigma^\star$ defined in (2.10) is optimal. If $\tilde{f}$ is decreasing in $(0, \eta)$ and the initial condition is $x \in D \setminus \{0\}$, then the control $\sigma^\star$ defined in (2.11) is optimal.

In order to prove this result, we refer to the theory of viscosity solutions for Hamilton-Jacobi-Bellman (HJB) equations that we develop in Chapter 4. The main result of Chapter 4 is the following theorem, which we restate here for reference.

Theorem 4.20. *Suppose that Assumption 1.16 holds, and suppose further that the domain D is uniformly convex, the running cost f is continuous in D, and the boundary cost g is uniformly continuous on ∂D.*

Then the value function $v : D \to \mathbb{R}$ defined in Section 1.4.1 extends continuously to $\overline{D}$ and is the unique viscosity solution of the HJB equation

$$-\frac{1}{2} \inf_{\sigma \in U} \operatorname{Tr}\left(D^2 v \sigma \sigma^\top\right) - f = 0$$

in D, with boundary condition

$$v = g \quad on \quad \partial D.$$

In this section, we will prove that the candidate function V is a viscosity solution of the HJB equation

$$-\frac{1}{2} \inf_{\sigma \in U} \operatorname{Tr}\left(D^2 V(x) \sigma \sigma^\top\right) = f(x), \quad x \in D, \tag{2.20}$$

with boundary condition $V = g$ on ∂D. We then appeal to Theorem 4.20, as stated above, to see that the value function v is a viscosity solution of the same boundary value problem and, moreover, such a solution is unique. From this, we conclude that the function V is equal to the value function v.

We first show that V is a classical solution of (2.20) in the regions where we expect radial motion to be optimal.

Lemma 2.16. *For each $i \geq 1$, define $u_i : (s_{i-1}, r_i \wedge R] \to \mathbb{R}$ by*

$$u_i(r) = 2 \int_r^{r_i} \int_{s_{i-1}}^{s} \tilde{f}(t) \, \mathrm{d}t \, \mathrm{d}s + 2(r_i - r) s_{i-1} \tilde{f}(s_{i-1}) + C_i^u,$$

for an arbitrary constant C_i^u, and define $U_i : D \to \mathbb{R}$ by

$$U_i(x) = u_i(|x|).$$

Then U_i is a classical solution of the PDE (2.20) *in the region*

$$\{x \in D \colon \ |x| \in (s_{i-1}, r_i \wedge R)\}.$$

Proof. Fix $i \geq 1$ and let $x \in D$ be such that $|x| \in (s_{i-1}, r_i \wedge R)$. Observe that, by definition of r_i,

$$u_i'(|x|) \geq -2 \, |x| \, \tilde{f}(|x|). \tag{2.21}$$

We have that U_i is twice continuously differentiable at x and

$$D^2 U_i(x) = |x|^{-3} \left[|x| \, u_i''(|x|) - u_i'(|x|) \right] x x^\top + |x|^{-1} \, u_i'(|x|) I.$$

Substituting in $u_i''(|x|) = -2\tilde{f}(|x|)$ and rearranging gives

$$\begin{aligned} D^2 U_i(x) &= - |x|^{-3} \left[2 \, |x| \, \tilde{f}(|x|) + u_i'(|x|) \right] x x^\top + |x|^{-1} \, u_i'(|x|) I \\ &= -2\tilde{f}(|x|) I + |x|^{-3} \left[2 \, |x| \, \tilde{f}(|x|) + u_i'(|x|) \right] \left[|x|^2 I - x x^\top \right]. \end{aligned}$$

Hence, for any $\sigma \in U$,

$$\begin{aligned} \operatorname{Tr}\left(D^2 U_i(x) \sigma \sigma^\top \right) = {} & -2\tilde{f}(|x|) \operatorname{Tr}(\sigma \sigma^\top) \\ & + |x|^{-3} \left[2 \, |x| \, \tilde{f}(|x|) + u_i'(|x|) \right] \operatorname{Tr}\left(\left[|x|^2 I - x x^\top \right] \sigma \sigma^\top \right). \end{aligned}$$

Noting that $|x|^2 I - x x^\top$ is positive semi-definite, and using (2.21), we have

$$\operatorname{Tr}\left(D^2 U_i(x) \sigma \sigma^\top \right) \geq -2\tilde{f}(|x|) \operatorname{Tr}(\sigma \sigma^\top) = -2 f(x),$$

for any $\sigma \in U$.

Taking $\sigma = \sigma^1(x)$, where $\sigma^1 : D \to \mathbb{R}$ is the function defined in Definition 2.7, we see that

$$\mathrm{Tr}\left(\left[|x|^2 I - xx^\top\right]\sigma^1(x)\sigma^1(x)^\top\right) = 0,$$

and so

$$\mathrm{Tr}\left(D^2 U_i(x)\sigma^1(x)\sigma^1(x)^\top\right) = -2f(x).$$

Hence U_i is a classical solution of the PDE (2.20) in the the region $\{x \in D\colon\ |x| \in (s_{i-1}, r_i \wedge R)\}$. □

We next show that V is a viscosity solution of (2.20) in the regions where we expect tangential motion to be optimal.

Lemma 2.17. *For each $i \geq 0$, define $w_i : (r_i, s_i \wedge R] \to \mathbb{R}$ by*

$$w_i(r) = 2\int_r^{s_i} s\tilde{f}(s)\,\mathrm{d}s + C_i^w,$$

for an arbitrary constant C_i^w, and define $W_i : D \to \mathbb{R}$ by

$$W_i(x) = w_i(|x|).$$

Then W_i is a viscosity solution of the PDE (2.20) *in the region*

$$\{x \in D\colon\ |x| \in (r_i, s_i \wedge R)\}.$$

Note that w_i is twice continuously differentiable if and only if $\tilde{f}$ is continuously differentiable. We first suppose that this is the case and prove the following lemma.

Lemma 2.18. *Fix $i \geq 0$ and suppose that $\tilde{f}$ is continuously differentiable in the interval $(r_i, s_i \wedge R)$. Then W_i defined in Lemma 2.17 is a classical solution of the PDE* (2.20) *in the region $\{x \in D\colon\ |x| \in (r_i, s_i \wedge R)\}$.*

Proof. Let $x \in D$ be such that $|x| \in (r_i, s_i \wedge R)$. Observe that, by definition of s_i,

$$w''_{i+1}(|x|) \geq -2\tilde{f}(|x|). \tag{2.22}$$

Since $\tilde{f}$ is assumed to be continuously differentiable, we have that w_i and W_i are both twice continuously differentiable, and

$$D^2 W_i(x) = |x|^{-3}\left[|x|\, w''_i(|x|) - w'_i(|x|)\right] xx^\top + |x|^{-1} w'_i(|x|) I.$$

Substituting in $w'_i(|x|) = -2\,|x|\,\tilde{f}(|x|)$, we have

$$D^2 W_i(x) = |x|^{-2}\left[w''_i(|x|) + 2\tilde{f}(|x|)\right] xx^\top - 2\tilde{f}(|x|) I.$$

Hence, for any $\sigma \in U$,

$$\begin{aligned}\operatorname{Tr}\left(D^2 W_i(x)\sigma\sigma^\top\right) &= |x|^{-2}\left[w_i''(|x|) + 2\tilde{f}(|x|)\right]\operatorname{Tr}(xx^\top\sigma\sigma^\top) - 2\tilde{f}(|x|)\operatorname{Tr}(\sigma\sigma^\top)\\ &\geq -2\tilde{f}(|x|)\operatorname{Tr}(\sigma\sigma^\top) = -2f(x),\end{aligned}$$

using the inequality (2.22).

Taking $\sigma = \sigma^0(x)$, where $\sigma^0 : D \to \mathbb{R}$ is the function defined in Definition 2.3, we see that

$$\operatorname{Tr}\left(xx^\top\sigma^0(x)\sigma^0(x)^\top\right) = 0,$$

and so

$$\operatorname{Tr}\left(D^2 W_i(x)\sigma^0(x)\sigma^0(x)^\top\right) = -2f(x).$$

Hence W_i is a classical solution of the PDE (2.20) in the region $\{x \in D\colon\ |x| \in (r_i, s_i \wedge R)\}$.

□

We can now prove Lemma 2.17, by using smooth approximations to the continuous function $\tilde{f}$ and applying a standard stability result for viscosity solutions, which can be found, for example, in Lemma 6.2 of [26, Chapter II].

Proof of Lemma 2.17. Fix $i \geq 1$. Since $\tilde{f}$ is continuous on $[r_i, s_i \wedge R]$, we can approximate $\tilde{f}$ uniformly by polynomials $(\tilde{f}^k)_{k\in\mathbb{N}}$ (see e.g. Theorem 7.26 of [54]).

For convenience, define the region $D_i := \{x \in D\colon\ |x| \in (r_i, s_i \wedge R)\}$. Let $k \in \mathbb{N}$ and define $W_i^k : D_i \to \mathbb{R}$ by

$$W_i^k(x) := -2\int_{r_i}^{|x|} \tilde{f}^k(s)s\,\mathrm{d}s + C_i^w.$$

Define $f^k : D_i \to \mathbb{R}$ by $f^k(x) = \tilde{f}^k(|x|)$, and define $F^k : D_i \times \mathbb{R}^{d,d} \to \mathbb{R}$ by

$$F^k(x, X) = -\frac{1}{2}\inf_{\sigma\in U}\operatorname{Tr}(X\sigma\sigma^\top) - f^k(x).$$

Then, since $\tilde{f}^k$ is continuously differentiable, we can apply Lemma 2.18 to see that W_i^k is a classical solution, and therefore a viscosity solution, of

$$F^k(x, D^2 W_i^k(x)) = 0 \quad \text{for} \quad x \in D_i.$$

We now show that F^k converges uniformly to $F : D_i \times \mathbb{R}^{d,d} \to \mathbb{R}$, defined by

$$F(x, X) = -\frac{1}{2}\inf_{\sigma\in U}\operatorname{Tr}(X\sigma\sigma^\top) - f(x),$$

and that W_i^k converges uniformly to W_i.

Let $\varepsilon > 0$. Then, by uniform convergence of $(\tilde{f}^k)_{k\in\mathbb{N}}$, there exists $N \in \mathbb{N}$ such that

$$\left|\tilde{f}(r) - \tilde{f}^k(r)\right| < \varepsilon, \quad \text{for all} \quad r \in [r_0, R] \quad \text{and} \quad k \geq N.$$

Let $k \geq N$, $x \in D_i$ and $X \in \mathbb{R}^{d,d}$. Then $|x| \in [r_i, s_i \wedge R]$, and so

$$\begin{aligned}\left|F(x,X) - F^k(x,X)\right| &= \left|f(x) - f^k(x)\right| \\ &= \left|\tilde{f}(|x|) - \tilde{f}^k(|x|)\right| < \varepsilon.\end{aligned}$$

Therefore $F^k \to F$ uniformly on $D_i \times \mathbb{R}^{d,d}$.

Now choose $M \in \mathbb{N}$ such that

$$\left|\tilde{f}(r) - \tilde{f}^k(r)\right| < \frac{\varepsilon}{2s_i(s_i - r_i)}, \quad \text{for all} \quad r \in [r_i, s_i] \quad \text{and} \quad k \geq M.$$

Let $k \geq M$ and $x \in D_i$. Then $|x| \in [r_i, s_i \wedge R]$, and so

$$\begin{aligned}\left|W_i(x) - W_i^k(x)\right| &= 2\left|\int_{r_i}^{|x|} \left(\tilde{f}(s) - \tilde{f}^k(s)\right) s \,\mathrm{d}s\right| \\ &\leq 2\int_{r_i}^{|x|} \left|\tilde{f}(s) - \tilde{f}^k(s)\right| |s| \,\mathrm{d}s \\ &\leq 2\int_{r_i}^{s_i} \left|\tilde{f}(s) - \tilde{f}^k(s)\right| |s| \,\mathrm{d}s \\ &\leq 2(s_i - r_i)\frac{\varepsilon}{2s_i(s_i - r_i)} s_i \\ &= \varepsilon.\end{aligned}$$

Hence $W_i^k \to W_i$ uniformly on D_i.

We can now apply the stability result given in Lemma 6.2 of [26, Chapter II], to conclude that W_i is a viscosity solution of

$$F(x, D^2 W_i(x)) = 0 \quad \text{for} \quad x \in D_i;$$

i.e. W_i is a viscosity solution of the PDE (2.20) in the region $\{x \in D\colon\ |x| \in (r_i, s_i \wedge R)\}$. □

We now combine the above lemmas to prove that V is the value function.

Proof of Proposition 2.15. We divide the domain D into disjoint regions and prove first that V is a viscosity solution of (2.20) in the interior of each region.

Step 1: Fix $i \geq 1$ such that $s_{i-1} \leq R$, if such a point exists. In the region $\{x \in D\colon |x| \in (s_{i-1}, r_i \wedge R)\}$, we have $V = U_i$, for a particular choice of constant C_i^u. So by Lemma 2.16, V is a viscosity solution of (2.20) in this region.

Now fix $i \geq 0$ such that $r_i \leq R$, if such a point exists. In the region $\{x \in D\colon |x| \in (r_i, s_i \wedge R)\}$, we have $V = W_i$ for a particular choice of constant C_i^w, and so V is a viscosity solution of (2.20) in this region, by Lemma 2.17.

Step 2: We next prove that V is a viscosity solution of (2.20) on each of the internal boundaries between the regions.

Let $i \geq 0$ be such that $r_i < R$, if such a point exists. Consider $x_i \in D$ such that $|x_i| = r_i$. Note that

$$\begin{aligned}
&\lim_{|x|\to r_i-} D^2V(x) \\
&\qquad = \lim_{|x|\to r_i-} D^2U_i(x) \\
&\qquad = -\lim_{|x|\to r_i-} \left[2\tilde{f}(|x|)I + |x|^{-3}\left(2\,|x|\,\tilde{f}(|x|) + u_i'(|x|)\right)\left[|x|^2\, I - xx^\top\right]\right] \\
&\qquad = -2\tilde{f}(r_i)I,
\end{aligned} \tag{2.23}$$

since $2r_i\tilde{f}(r_i) + u_{i}{}'_{-}(r_i) = 0$, by definition of r_i and continuity of $\tilde{f}$.

To show that V is a viscosity subsolution at x_i, let $x_i \in \arg\min(\phi - V)$, for some $\phi \in C^\infty(D)$. Since $V \in C^1(D)$, it must be the case that $D\phi(x_i) = DV(x_i)$, and that the Hessian of ϕ satisfies

$$D^2\phi(x_i) \geq \lim_{|x|\to r_i-} D^2V(x) = -2\tilde{f}(r_i)I,$$

as calculated in (2.23). Hence, for any $\sigma \in U$,

$$\operatorname{Tr}\left(D^2\phi(x_i)\sigma\sigma^\top\right) \geq -2\tilde{f}(r_i)\operatorname{Tr}(\sigma\sigma^\top) = -2f(x_i),$$

and so

$$-\frac{1}{2}\inf_{\sigma\in U}\operatorname{Tr}\left(D^2\phi(x_i)\sigma\sigma^\top\right) \leq f(x_i),$$

as required.

To show the supersolution property, let $x_i \in \arg\max(\psi - V)$, for some $\psi \in C^\infty(D)$. Then by a similar argument to the one above, we have

$$D^2\psi(x_i) \leq -2\tilde{f}(r_i)I,$$

and so

$$\operatorname{Tr}\left(D^2\psi(x_i)\sigma\sigma^\top\right) \le -2f(x_i),$$

for any $\sigma \in U$, which implies that

$$-\frac{1}{2}\inf_{\sigma\in U}\operatorname{Tr}\left(D^2\psi(x_i)\sigma\sigma^\top\right) \ge f(x_i).$$

Now let $i \ge 0$ be such that $s_i < R$, if such a point exists, and consider $x_i \in D$ such that $|x_i| = s_i$. Here, note that

$$\begin{aligned}
&\lim_{|x|\to s_i+} D^2V(x) \\
&\quad= \lim_{|x|\to s_i+} D^2U_{i+1}(x) \\
&\quad= -\lim_{|x|\to s_i+}\left[2\tilde{f}(|x|)I + |x|^{-3}\left(2\,|x|\,\tilde{f}(|x|) + u_{i+1}'(|x|)\right)\left[|x|^2\, I - xx^\top\right]\right] \\
&\quad= -2\tilde{f}(s_i)I,
\end{aligned} \tag{2.24}$$

using the fact that $2s_i\tilde{f}(s_i) + {u_{i+1}}'_+(s_i) = 0$, by definition of s_i and the smooth fit property.

To show that V is a viscosity solution at points of radius s_i, we follow the same reasoning as we did for points of radius r_i. For $x_i \in \arg\min(\phi - V)$ and $\phi \in C^\infty(D)$, we have that

$$D^2\phi(x_i) \ge \lim_{|x|\to s_i+} D^2V(x) = -2\tilde{f}(s_i)I,$$

using (2.24). So, for any $\sigma \in U$,

$$\operatorname{Tr}\left(D^2\phi(x_i)\sigma\sigma^\top\right) \ge -2f(x_i),$$

which implies that the subsolution property holds.

Similarly, for $x_i \in \arg\max(\psi - V)$ and $\psi \in C^\infty(D)$, we have

$$D^2\psi(x_i) \le -2\tilde{f}(s_i)I,$$

and so, for any $\sigma \in U$,

$$\operatorname{Tr}\left(D^2\psi(x_i)\sigma\sigma^\top\right) \le -2f(x_i),$$

which implies the supersolution property.

Step 3: We have shown that V is a viscosity solution of (2.20) in $D \setminus \{0\}$. We now consider the behaviour at the origin. Recall from Assumption 2.11 that we have assumed that $\tilde{f}$ is monotone on some interval $(0, \eta)$.

Case I: Suppose that $\tilde{f}$ is strictly increasing on $(0, \eta)$.

Then $V = U_1$ in some neighbourhood of the origin. We see that $r_1 > \eta$, and so $V = U_1$ in $B_\eta(0)$. Let $x \in B_\eta(0)$ and consider

$$D^2V(x) = -2\tilde{f}(|x|)I + |x|^{-3}\left(2\,|x|\,\tilde{f}(|x|) + u_1'(|x|)\right)\left[|x|^2\,I - xx^\top\right].$$

Since $|x| < r_1$, we have

$$2\,|x|\,\tilde{f}(|x|) + u_1'(|x|) > 0.$$

Substituting in the value of u_1' and considering a first order Taylor expansion around 0, we find that there exists $C > 0$ such that

$$\begin{aligned} 2\,|x|\,\tilde{f}(|x|) + u_1'(|x|) = -2\int_0^{|x|}\tilde{f}(s)\,\mathrm{d}s &= 2\,|x|\left(\tilde{f}(|x|) - \tilde{f}(0)\right) + o(|x|) \\ &\leq 2\,|x|\left(\tilde{f}(|x|) - \tilde{f}(0)\right) + C\,|x|^2. \end{aligned}$$

Hence, for $j, k \in \{1, \ldots, d\}$,

$$\begin{aligned} 0 &\leq |x|^{-3}\left(2\,|x|\,\tilde{f}(|x|) + u_i'(|x|)\right)\left|\left[|x|^2\,I - xx^\top\right]_{jk}\right| \\ &\leq |x|^{-1}\left(2\,|x|\,\tilde{f}(|x|) + u_i'(|x|)\right) \\ &\leq 2\left(\tilde{f}(|x|) - \tilde{f}(0)\right)C\,|x|. \end{aligned}$$

Taking the limit as $|x| \to 0+$, by continuity of $\tilde{f}$, we see that

$$\lim_{x\to 0} D^2V(x) = -2\tilde{f}(0)I.$$

We now show that V is a viscosity solution of (2.20) at 0. Let $\phi \in C^\infty(D)$ be such that $0 \in \arg\max(V - \phi)$. Since the gradient DV is continuous, we have $D\phi(0) = DV(0) = 0$ and

$$D^2\phi(0) \leq \lim_{x\to 0} D^2V(x) = -2\tilde{f}(0)I,$$

and so V is a viscosity subsolution of (2.20) at 0. On the other hand, for any

$\psi \in C^\infty(D)$ such that $0 \in \arg\min(V - \psi)$, we have $D\psi(0) = DV(0) = 0$ and

$$D^2\psi(0) \geq -2\tilde{f}(0),$$

and so V is a viscosity supersolution of (2.20) at 0.

Case II: On the other hand, if $\tilde{f}$ is decreasing in $(0, \eta)$, we have that $V = W_1$ in $B_\eta(0)$. Recall from Assumption 2.11 that we have assumed that $\tilde{f}$ is continuously differentiable on some interval $(0, \delta)$, and consider $x \in D$ such that $|x| < \delta \wedge \eta$.

We have that

$$\begin{aligned} D^2V(x) &= |x|^{-2}\left[w_1''(|x|) + 2\tilde{f}(|x|)\right] xx^\top - 2\tilde{f}(|x|)I \\ &= 2\,|x|^{-2}\left[-\,|x|\,\tilde{f}'(|x|) - \tilde{f}(|x|) + \tilde{f}(|x|)\right] xx^\top - 2\tilde{f}(|x|)I \\ &= -2\,|x|^{-1}\,\tilde{f}'(|x|)xx^\top - 2\tilde{f}(|x|)I. \end{aligned}$$

Since $\tilde{f}'(|x|) \leq 0$, we get the following bound. For $j, k \in \{1, \ldots, d\}$,

$$0 \leq -2\,|x|^{-1}\,\tilde{f}'(|x|)\,|x_j x_k| \leq -2\,|x|\,\tilde{f}'(|x|) \to 0, \quad \text{as} \quad |x| \to 0+,$$

by the fifth statement of Assumption 2.11.

Therefore $\lim_{x\to 0} D^2V(x) = -2\tilde{f}(0)I$, and so V is a viscosity solution of (2.20) at the origin, by the same argument as for Case I.

Step 4: By construction of the function V, the boundary condition $V = g$ on ∂D is satisfied. We conclude, by Theorem 4.20, that the function V is equal to the value function v. Also, by the construction of V, we have that the value function v is continuously differentiable in D.

Step 5: Finally, we turn to the proof that the control $\sigma^\star$ is optimal. It is sufficient to show that

$$t \mapsto V(X_t^{\sigma^\star}) + \int_0^t f(X_s^{\sigma^\star})\,\mathrm{d}s$$

is a martingale. We will work with the squared radius of the process $X^{\sigma^\star}$, writing $Z_t^{\sigma^\star} = \left|X_t^{\sigma^\star}\right|$, for $t \geq 0$. We also let $\overline{V} : [0, R^2) \to \mathbb{R}$ be such that $V(x) = \overline{V}(|x|)$ for all $x \in D$.

Suppose that $\tilde{f}$ is increasing on the interval $(0, \eta)$. Then $\sigma^\star$ is given by (2.10). Letting W be the first component of the Brownian motion B, Lemma 2.2 tells us that $Z^{\sigma^\star}$ satisfies the SDE

$$\mathrm{d}Z_t^{\sigma^\star} = \mathrm{d}t + 2\left(\sum_i \mathbb{1}_{\left\{Z_t^{\sigma^\star}\in(s_i^2, r_{i+1}^2\wedge R^2)\right\}} + \mathbb{1}_{\left\{Z_t^{\sigma^\star}\in[0, r_1^2\wedge R^2)\right\}}\right)\sqrt{Z_t^{\sigma^\star}}\,\mathrm{d}W_t,$$

where the index i runs from 1 to the first i such that $r_{i+1} \geq R$.

In each interval $[r_i^2, s_i^2]$, there is a constant C such that

$$\overline{V}(z) = 2\int_{\sqrt{z}}^{s_i} s\tilde{f}(s)\,\mathrm{d}s + C.$$

Therefore, since $\mathrm{d}Z_t^{\sigma^\star} = \mathrm{d}t$ when $Z_t^{\sigma^\star} \in [r_i^2, s_i^2]$, we can make a change of variables to find that

$$\mathbb{1}_{\left\{Z_t^{\sigma^\star}\in[r_i^2, s_i^2]\right\}}\,\mathrm{d}\overline{V}(Z_t^{\sigma^\star}) = -\mathbb{1}_{\left\{Z_t^{\sigma^\star}\in[r_i^2, s_i^2]\right\}}\tilde{f}\left(\sqrt{Z_t^{\sigma^\star}}\right)\mathrm{d}t. \tag{2.25}$$

Now, in each interval (s_i^2, r_{i+1}^2), there is a constant C such that

$$\overline{V}(z) = 2\int_{\sqrt{z}}^{r_{i+1}}\int_{s_i}^{s}\tilde{f}(t)\,\mathrm{d}t\,\mathrm{d}s + 2(r_{i+1} - \sqrt{z})s_i\tilde{f}(s_i) + C.$$

We see that V is twice continuously differentiable in such an interval, and so we can apply Itô's formula to $\overline{V}(Z^{\sigma^\star})$. We calculate the derivatives

$$\overline{V}'(z) = -z^{-\frac{1}{2}}\int_{s_i}^{\sqrt{z}}\tilde{f}(s)\,\mathrm{d}s - z^{-\frac{1}{2}}s_i\tilde{f}(s_i),$$

and

$$\overline{V}''(z) = \frac{1}{2}z^{-\frac{3}{2}}\int_{s_i}^{\sqrt{z}}\tilde{f}(s)\,\mathrm{d}s - \frac{1}{2}Z^{-1}\tilde{f}(\sqrt{z}) + \frac{1}{2}z^{-\frac{3}{2}}s_i\tilde{f}(s_i).$$

Then, by Itô's formula, we find that

$$\begin{aligned}\mathbb{1}_{\left\{Z_t^{\sigma^\star}\in(s_i^2, r_{i+1}^2)\right\}}\,\mathrm{d}\overline{V}(Z_t^{\sigma^\star}) = &-\mathbb{1}_{\left\{Z_t^{\sigma^\star}\in(s_i^2, r_{i+1}^2)\right\}}\tilde{f}\left(\sqrt{Z_t^{\sigma^\star}}\right)\mathrm{d}t\\ &+ 2\mathbb{1}_{\left\{Z_t^{\sigma^\star}\in(s_i^2, r_{i+1}^2)\right\}}\overline{V}'(Z_t^{\sigma^\star})\sqrt{Z_t^{\sigma^\star}}\,\mathrm{d}W_t.\end{aligned}$$

We have a similar expression for the interval $[0, r_1^2)$, and so combining this with (2.25), we have

$$\begin{aligned}V(X_t^{\sigma^\star}) - V(X_0^{\sigma^\star}) = &-\int_0^t f(X_s^{\sigma^\star})\,\mathrm{d}s\\ &+ 2\int_0^t\left(\sum_i \mathbb{1}_{\left\{Z_s^{\sigma^\star}\in(s_i^2, r_{i+1}^2)\right\}} + \mathbb{1}_{\left\{Z_s^{\sigma^\star}\in[0, r_1^2)\right\}}\right)\sqrt{Z_s^{\sigma^\star}}\,\mathrm{d}W_s,\end{aligned}$$

for any $t \geq 0$. This shows that the required martingale property holds, and so $\sigma^\star$ is an optimal control.

Now suppose that $\tilde{f}$ is decreasing on the interval $(0, \eta)$, and let $X_0^{\sigma^\star} = x$, for some $x \in D \setminus \{0\}$. In this case $\sigma^\star$ is given by (2.11), and $Z^{\sigma^\star}$ satisfies

$$\mathrm{d}Z_t^{\sigma^\star} = \mathrm{d}t + 2\sum_i \mathbb{1}_{\left\{Z_t^{\sigma^\star} \in (s_i^2, r_{i+1}^2 \wedge R^2)\right\}} \sqrt{Z_t^{\sigma^\star}}\,\mathrm{d}W_t,$$

where now the index i runs from 0 to the first i such that $r_{i+1} \geq R$. We see that $Z^{\sigma^\star}$ never hits the origin.

We can make the same calculations as above to find that, for any $t \geq 0$,

$$V(X_t^{\sigma^\star}) - V(X_0^{\sigma^\star}) = -\int_0^t f(X_s^{\sigma^\star})\,\mathrm{d}s + 2\int_0^t \sum_i \mathbb{1}_{\left\{Z_s^{\sigma^\star} \in (s_i^2, r_{i+1}^2)\right\}} \sqrt{Z_s^{\sigma^\star}}\,\mathrm{d}W_s,$$

and so the required martingale property holds once again. We conclude that $\sigma^\star$ is optimal. □

We required the smoothness conditions on the running cost f in Assumption 2.11 in order to show that the candidate value function is a viscosity solution at the origin. In Section 2.4, we will relax these assumptions and extend the above result to include cost functions that have an infinite discontinuity at the origin. In this case, we cannot define a viscosity solution of the HJB equation (2.20) at the origin, and so Theorem 4.20 will no longer be applicable.

2.4 Infinite cost at the origin

We now extend Proposition 2.15 by considering the case where the cost function is continuous on the whole domain, except at the origin where it is allowed to become infinite. We will show that the value function takes the same form as we saw in Proposition 2.15. We will also find growth conditions on the cost function under which the value function becomes infinite. We note that, in allowing the cost function to become infinite at the origin, we must take care to check that we still have equality between the strong value function v^S and the weak value function v^W, as we showed in Proposition 1.7 for the case of continuous cost functions. We will find a particular growth regime where we require results on Brownian filtrations from Chapter 3 in order to prove that $v^S(0) = v^W(0)$ in dimension $d = 2$.

We relax Assumption 2.11 to remove some of the regularity conditions on the cost function f, as follows.

Assumption 2.19. We assume that

1. The domain is $D = B_R(0) \subset \mathbb{R}^d$, for some $R > 0$ and $d \geq 2$;
2. The cost function f is radially symmetric; i.e. $f(x) = \tilde{f}(|x|)$, for some function $\tilde{f} : [0, R) \to \mathbb{R}$;
3. The boundary cost g is constant;
4. The cost function f is continuous on $D \setminus \{0\}$;
5. There exists $\eta > 0$ such that the cost function $\tilde{f}$ is monotone on the interval $(0, \eta)$;
6. The one-sided derivative $\tilde{f}'_+(r)$ exists for all $r > 0$ and changes sign only finitely many times.

Note that we retain the fifth statement in this assumption to ensure that the cost function does not oscillate as it approaches the origin, and we retain the sixth statement so that there are finitely many switching points and these are well-defined.

Having relaxed the conditions on the cost function f, we can no longer use the theory of viscosity solutions. To prove the following results, we once again treat the cases of increasing and decreasing costs separately, and we distinguish between regimes of slow and fast growth at the origin. The different growth regimes will be determined by the convergence of the integrals

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s \quad \text{and} \quad \int_0^r s\tilde{f}(s)\,\mathrm{d}s.$$

In each of the proofs in this section, we make the simplifying assumption that the boundary cost is $g = 0$. However, the results still hold for any constant boundary cost g.

2.4.1 Cost functions increasing at the origin

We first consider cost functions that are increasing in some neighbourhood around the origin. In this case, we will find that radial motion, as defined in Definition 2.7, is optimal close to the origin.

Proposition 2.20. *Suppose that Assumption 2.19 holds and there exists $\eta > 0$ such that $\tilde{f}$ is negative and increasing on the interval $(0, \eta)$. Then the strong and weak value functions defined in Section 1.4.1 are equal, and we can write $v = v^S = v^W$.*

If, for any $r > 0$,

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s > -\infty,$$

then the value function v is finite and equal to the candidate value V defined in Definition 2.14.

If, for any $r > 0$,

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s = -\infty,$$

then $v \equiv -\infty$.

Remark 2.21. Note that, since $\tilde{f}$ is increasing on $(0, \eta)$, the function V is defined in Case I of Definition 2.14, with $s_0 = 0$ and $r_1 = \inf\left\{r > 0\colon \int_0^r \tilde{f}(s)\,\mathrm{d}s > r\tilde{f}(r)\right\}$. When $\int_0^r \tilde{f}(s)\,\mathrm{d}s > -\infty$ for any $r > 0$, the switching point r_1 is well-defined.

Proof of Proposition 2.20. First suppose that, for any $r > 0$,

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s > -\infty.$$

For $N \in \mathbb{N}$, define an approximating sequence of functions $\tilde{f}_N : [0, R) \to \mathbb{R}$ by

$$\tilde{f}_N(r) = \begin{cases} \tilde{f}(\frac{1}{N}), & r \leq \frac{1}{N}, \\ \tilde{f}(r), & r > \frac{1}{N}, \end{cases}$$

and define $f_N : D \to \mathbb{R}$ by $f_N(x) = \tilde{f}_N(|x|)$ for $x \in D$. Then f_N is continuous and bounded. Moreover, fixing $N > \frac{1}{\eta}$, we have the bound $f_N \geq f$.

Now define $v_N^S : D \to \mathbb{R}$ by

$$v_N^S(x) := \inf_{\sigma \in \mathcal{U}} \mathbb{E}^x \left[\int_0^\tau f_N(X_s^\sigma)\,\mathrm{d}s\right], \quad x \in D,$$

using the same notation as in the definition of the strong value function v^S in Section 1.4.1. Note that $v_N^S \geq v^S$.

Let V_N denote the candidate value function defined in Case I of Definition 2.14 with the function $\tilde{f}$ replaced by $\tilde{f}_N$. Since Assumption 2.11 is satisfied for the value function v_N^S, we can apply Proposition 2.15 to see that $v_N^S = V_N$. We can also see that, for any $x \in D$, $\lim_{N\to\infty} V_N(x) = V(x)$ and $V(x)$ is finite, since $\int_0^r \tilde{f}(s)\,\mathrm{d}s > -\infty$ for any $r > 0$. We will show that $\lim_{N\to\infty} v_N^S(x) = v^S(x)$ and conclude that $v^S(x) = V(x)$.

Fix $\sigma \in \mathcal{U}$ and $x \in D$. We have

$$\begin{aligned}\mathbb{E}^x\left[\int_0^\tau \tilde{f}_N(|X_s^\sigma|)\,\mathrm{d}s\right] = \mathbb{E}^x&\left[\int_0^\tau \tilde{f}(|X_s^\sigma|)\mathbb{1}_{\left\{|X_s^\sigma|\in(\frac{1}{N},R)\right\}}\,\mathrm{d}s\right]\\&+\tilde{f}\left(\frac{1}{N}\right)\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\left\{|X_s^\sigma|\le\frac{1}{N}\right\}}\right].\end{aligned} \tag{2.26}$$

Define $K := \sup\{f(x) \colon x \in D\}$ and note that $K < \infty$ by continuity of f in $D \setminus \{0\}$. Then the sequence

$$\left(\int_0^\tau \tilde{f}(|X_s^\sigma|)\mathbb{1}_{\left\{|X_s^\sigma|\in(\frac{1}{N},R)\right\}}\,\mathrm{d}s\right)_{N\in\mathbb{N}}$$

is decreasing for $N > \frac{1}{\eta}$ and bounded above by τK. Since τ has finite expectation by Proposition 1.5, we can apply monotone convergence (see e.g. Theorem 1 of [55, Chapter II, §6]) to show that

$$\begin{aligned}\lim_{N\to\infty}\mathbb{E}^x\left[\int_0^\tau \tilde{f}(|X_s^\sigma|)\mathbb{1}_{\left\{|X_s^\sigma|\in(\frac{1}{N},R)\right\}}\,\mathrm{d}s\right] &= \mathbb{E}^x\left[\lim_{N\to\infty}\int_0^\tau \tilde{f}(|X_s^\sigma|)\mathbb{1}_{\left\{|X_s^\sigma|\in(\frac{1}{N},R)\right\}}\,\mathrm{d}s\right]\\&= \mathbb{E}^x\left[\int_0^\tau \tilde{f}(|X_s^\sigma|)\,\mathrm{d}s\right].\end{aligned}$$

We will show that the second term of (2.26) vanishes as $N \to \infty$ by referring to Proposition 2.8 on the control problem for a step cost function. Note that $\tilde{f}(\frac{1}{N}) < 0$. For $x \neq 0$, we can choose $N > \frac{1}{|x|}$, so that, by Proposition 2.8,

$$\begin{aligned}0 > \tilde{f}\left(\frac{1}{N}\right)\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\left\{|X_s^\sigma|\le\frac{1}{N}\right\}}\right] &= -\tilde{f}\left(\frac{1}{N}\right)\mathbb{E}^x\left[\int_0^\tau -\mathbb{1}_{\left\{|X_s^\sigma|\le\frac{1}{N}\right\}}\right]\\&\ge -\frac{2}{N}\tilde{f}\left(\frac{1}{N}\right)(R-|x|)\\&\xrightarrow{N\to\infty} 0,\end{aligned}$$

using the condition that $\int_0^r \tilde{f}(s)\,\mathrm{d}s > -\infty$ to find the limit.

For $x = 0$, Proposition 2.8 tells us that

$$\begin{aligned}0 > \tilde{f}\left(\frac{1}{N}\right)\mathbb{E}^0\left[\int_0^\tau \mathbb{1}_{\left\{|X^\sigma|\le\frac{1}{N}\right\}}\,\mathrm{d}s\right] &= -\tilde{f}\left(\frac{1}{N}\right)\mathbb{E}^0\left[-\int_0^\tau \mathbb{1}_{\left\{|X^\sigma|\le\frac{1}{N}\right\}}\,\mathrm{d}s\right]\\&\ge \tilde{f}\left(\frac{1}{N}\right)\left(\frac{2R}{N}-\frac{1}{N^2}\right)\\&\ge 2R\frac{1}{N}\tilde{f}\left(\frac{1}{N}\right)\xrightarrow{N\to\infty} 0.\end{aligned}$$

Hence

$$\lim_{N\to\infty}\mathbb{E}^x\left[\int_0^\tau \tilde{f}_N(|X_s^\sigma|)\,\mathrm{d}s\right] = \mathbb{E}^x\left[\int_0^\tau \tilde{f}(|X_s^\sigma|)\,\mathrm{d}s\right],$$

for any $\sigma \in \mathcal{U}$, $x \in D$.

Now fix $x \in D$ and $\varepsilon > 0$ and choose σ^ε to be an ε-optimal strategy for the cost function f; i.e.

$$\mathbb{E}^x \left[\int_0^\tau f(X_s^{\sigma^\varepsilon}) \, \mathrm{d}s \right] \leq v^S(x) + \varepsilon.$$

Then

$$\begin{aligned} v^S(x) + \varepsilon &\geq \mathbb{E}^x \left[\int_0^\tau f(X_s^{\sigma^\varepsilon}) \, \mathrm{d}s \right] \\ &= \lim_{N \to \infty} \mathbb{E}^x \left[\int_0^\tau f_N(X_s^{\sigma^\varepsilon}) \, \mathrm{d}s \right] \\ &\geq \lim_{N \to \infty} v_N^S(x) \geq v^S(x). \end{aligned}$$

Taking the limit as $\varepsilon \downarrow 0$, we see that

$$v^S(x) = \lim_{N \to \infty} v_N^S(x),$$

and by uniqueness of the limit, we have that $v^S(x) = V(x)$.

As in Proposition 1.7, we can apply Theorem 4.5 of [20] to see that $v^S = v^W$. Since f is continuous in $D \setminus \{0\}$, upper semicontinuous at 0, and bounded above by a constant, we can deduce that the conditions of Theorem 4.5 of [20] are met in the same way as in the proof of Proposition 1.7. Hence $v^W = v^S = V$.

Now suppose that, for any $r > 0$,

$$\int_0^r \tilde{f}(s) \, \mathrm{d}s = -\infty.$$

We will show that radial motion is an optimal strategy and that this strategy gives a negative infinite cost. Let the control σ^1 be as defined in Definition 2.7, and define X^{σ^1} by

$$X_t^{\sigma^1} = x + \int_0^t \sigma_s^1 \, \mathrm{d}B_s, \quad t \geq 0.$$

Let W be the first component of the Brownian motion B.

First suppose that $x \neq 0$. Then, for any $t \geq 0$,

$$X_t^{\sigma^1} = x + \frac{x}{|x|} W_t,$$

and so

$$\begin{aligned}\mathbb{E}^x\left[\int_0^\tau \tilde{f}(X_s^{\sigma^1})\,\mathrm{d}s\right] &= \mathbb{E}^{|x|}\left[\int_0^\tau \tilde{f}(|W_s|)\,\mathrm{d}s\right]\\ &= \mathbb{E}^{|x|}\left[\int_0^\tau \tilde{f}(W_s)\mathbb{1}_{\{W_s\geq 0\}}\,\mathrm{d}s\right] + \mathbb{E}^{|x|}\left[\int_0^\tau \tilde{f}(-W_s)\mathbb{1}_{\{W_s< 0\}}\,\mathrm{d}s\right].\end{aligned}$$

We can now use the Green's function G for the one-dimensional Brownian motion W on the interval $(-R,R)$, as calculated in Example 2.6 using the definitions in Appendix B. By Proposition B.5, we see that

$$\begin{aligned}\mathbb{E}^x\left[\int_0^\tau \tilde{f}(X_s^{\sigma^1})\,\mathrm{d}s\right] &= 2\int_0^R G(|x|\,,y)\tilde{f}(y)\,\mathrm{d}y + 2\int_{-R}^0 G(|x|\,,y)\tilde{f}(-y)\,\mathrm{d}y\\ &= \frac{|x|+R}{R}\int_{|x|}^R (R-y)\tilde{f}(y)\,\mathrm{d}y + \frac{R-|x|}{R}\int_0^{|x|}(y+R)\tilde{f}(y)\,\mathrm{d}y\\ &\quad+\frac{R-|x|}{R}\int_{-R}^0 (y+R)\tilde{f}(-y)\,\mathrm{d}y.\end{aligned}$$

Making a change of variables $y\mapsto -y$ in the last integral,

$$\frac{R-|x|}{R}\int_{-R}^0 (y+R)\tilde{f}(-y)\,\mathrm{d}y = \frac{R-|x|}{R}\int_0^R (R-y)\tilde{f}(y)\,\mathrm{d}y,$$

and so

$$\mathbb{E}^x\left[\int_0^\tau \tilde{f}(X_s^{\sigma^1})\,\mathrm{d}s\right] = 2\int_{|x|}^R (R-y)\tilde{f}(y)\,\mathrm{d}y + 2(R-|x|)\int_0^{|x|}\tilde{f}(y)\,\mathrm{d}y.$$

Since f is bounded above and

$$\int_0^{|x|}\tilde{f}(y)\,\mathrm{d}y = -\infty,$$

we have

$$\mathbb{E}^x\left[\int_0^\tau \tilde{f}(X_s^{\sigma^1})\,\mathrm{d}s\right] = -\infty.$$

Hence

$$v^W(x)\leq v^S(x)\leq \mathbb{E}^x\left[\int_0^\tau \tilde{f}(X_s^{\sigma^1})\,\mathrm{d}s\right] = -\infty.$$

Now let $x=0$. Then for $t\geq 0$,

$$X_t^{\sigma^1} = e_1 W_t.$$

By symmetry of the Green's function G for W about zero, we have

$$\begin{aligned}\mathbb{E}^0\left[\int_0^\tau \tilde{f}\left(\left|X_s^{\sigma^1}\right|\right) \mathrm{d}s\right] &= 2\mathbb{E}^0\left[\int_0^\tau \tilde{f}(W_s)\mathbb{1}_{\{W_s\geq 0\}}\,\mathrm{d}s\right] \\ &= 4\int_0^R G(0,y)\tilde{f}(y)\,\mathrm{d}y \\ &= 2\int_0^R (R-y)\tilde{f}(y)\,\mathrm{d}y \\ &= -\infty,\end{aligned} \tag{2.27}$$

by the growth condition on f. Hence

$$v^W(0) \leq v^S(0) \leq \mathbb{E}^0\left[\int_0^\tau \tilde{f}\left(\left|X_s^{\sigma^1}\right|\right) \mathrm{d}s\right] = -\infty.$$

We conclude that

$$v^W = v^S \equiv -\infty. \qquad \square$$

We have shown that, for cost functions increasing at the origin, there is a dichotomy depending on the convergence of $\int_0^r \tilde{f}(s)\,\mathrm{d}s$. When $\int_0^r \tilde{f}(s)\,\mathrm{d}s > -\infty$ for any $r > 0$, the value function is finite and equal to V, and when $\int_0^r \tilde{f}(s)\,\mathrm{d}s = -\infty$ for any $r > 0$, the value is identically equal to negative infinity.

2.4.2 Cost functions decreasing at the origin

We now consider cost functions that are decreasing in some neighbourhood around the origin. Excluding the origin from this neighbourhood, an optimal strategy is tangential motion, as defined in Definition 2.3.

We will first show that, away from the origin, the form of the value function is unchanged from the value function in Proposition 2.15.

Proposition 2.22. *Suppose that Assumption 2.19 holds and there exists $\eta > 0$ such that $\tilde{f}$ is positive and decreasing on the interval $(0, \eta)$.*

Then, for $x \in D \setminus \{0\}$, $v(x) = v^S(x) = v^W(x) = V(x) \in (-\infty, \infty)$, where V is the candidate value function defined in Proposition 2.15.

Remark 2.23. In this case, since $\tilde{f}$ is decreasing on $(0, \eta)$, V is defined in Case II of Definition 2.14.

Proof of Proposition 2.22. For $N \in \mathbb{N}$, define $\tilde{f}_N$, f_N and v_N^S as in the proof of Proposition 2.20. Now, for $N > \frac{1}{\eta}$, we have $\tilde{f}_N \leq \tilde{f}$, $f_N \leq f$, and $v_N^S \leq v_N$.

Recall that, by Proposition 2.15, $v_N^S = V_N$, where V_N is the candidate value function defined in Case II of Definition 2.14 with the cost function $\tilde{f}$ replaced by $\tilde{f}_N$.

Fix $x \neq 0$ and $N > \frac{1}{|x|} \vee \frac{1}{\eta}$. Then we can see that $V_N(x) = V(x)$, and $V(x)$ is finite. We will show that $v_N^S(x) = v^S(x)$ and conclude that $v^S(x) = V(x)$.

Let $\sigma^\star$ be the control defined in (2.11). Since $\tilde{f}^N$ is decreasing in the interval $(0, \eta)$, Proposition 2.15 shows that $\sigma^\star$ is optimal, and so

$$\begin{aligned} v_N^S(x) &= \mathbb{E}^x\left[\int_0^\tau \tilde{f}_N(\left|X_s^{\sigma^\star}\right|)\,\mathrm{d}s\right] \\ &= \tilde{f}\left(\frac{1}{N}\right)\mathbb{E}^x\left[\int_0^\tau \mathbb{1}_{\left\{\left|X_s^{\sigma^\star}\right| \le \frac{1}{N}\right\}}\,\mathrm{d}s\right] + \mathbb{E}^x\left[\int_0^\tau \tilde{f}(\left|X_s^{\sigma^\star}\right|)\mathbb{1}_{\left\{\left|X_s^{\sigma^\star}\right| \in (\frac{1}{N}, R)\right\}}\,\mathrm{d}s\right]. \end{aligned} \tag{2.28}$$

When $\left|X_t^{\sigma^\star}\right| \in (0, \eta)$, the radius process $t \mapsto \left|X_t^{\sigma^\star}\right|$ is deterministically increasing, by Lemma 2.2. Therefore, since $|x| > \frac{1}{N}$,

$$\mathbb{1}_{\left\{\left|X_t^{\sigma^\star}\right| \le \frac{1}{N}\right\}} = 0, \quad \text{for all} \quad t \ge 0.$$

Hence, by (2.28) and the definition of v^S, we have

$$v^S(x) \ge v_N^S(x) = \mathbb{E}^x\left[\int_0^\tau \tilde{f}(\left|X_s^{\sigma^\star}\right|)\,\mathrm{d}s\right] \ge v^S(x),$$

and so

$$v^S(x) = v_N^S(x) = V_N(x) = V(x).$$

Finally, for $N \in \mathbb{N}$, define $v_N^W : D \to \mathbb{R}$ by

$$v_N^W(x) := \inf_{\mathbb{P} \in \mathcal{P}_x} \mathbb{E}^{\mathbb{P}}\left[\int_0^\tau f_N(X_s)\,\mathrm{d}s\right], \quad x \in D,$$

using the same notation as in the definition of the weak value function v^W in Section 1.4.1. By Proposition 1.7, $v_N^S = v_N^W$. Once again, fix $x \neq 0$ and $N > \frac{1}{|x|} \vee \frac{1}{\eta}$, so that $v_N^W \le v^W$. Then we have

$$v_N^S(x) = v_N^W(x) \le v^W(x) \le v^S(x) = V(x) = v_N^S(x),$$

and we conclude that

$$v^W(x) = v^S(x) = V(x),$$

as required. □

At the origin, we have not shown that there exists an optimal control. The function σ^0 introduced in Definition 2.3 is not defined at the origin, and so we require an approximation to tangential motion. We consider different growth rates separately, as we did for increasing costs.

Proposition 2.24. *Suppose that Assumption 2.19 holds and there exists $\eta > 0$ such that $\tilde{f}$ is positive and decreasing on the interval $(0, \eta)$.*

Suppose further that, for any $r > 0$,

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s < \infty.$$

Then $v(0) = v^S(0) = v^W(0) = V(0) \in (-\infty, \infty)$, where V is the candidate value defined in Definition 2.14.

Proof. For $N \in \mathbb{N}$, define $\tilde{f}_N$, f_N and v_N^S as in the proof of Proposition 2.20. Letting V_N be the candidate value function in Case II of Definition 2.14 with $\tilde{f}$ replaced by $\tilde{f}_N$, we have $v_N^S(0) = V_N(0)$, by Proposition 2.15. We also see that $\lim_{N\to\infty} V_N(0) = V(0)$, and the value $V(0)$ is finite due to the growth condition on $\tilde{f}$. We will show that $v^S(0) = \lim_{N\to\infty} v_N^S(0)$ and conclude that $v^S(0) = V(0)$.

Fix $\delta \in (0, \eta)$ and $N > \frac{1}{\delta}$. Denote by e_1 the unit vector in the first coordinate direction. Let $\sigma^N \in \mathcal{U}$ be an optimal control for the cost f_N; that is

$$v_N^S(0) = \mathbb{E}^0\left[\int_0^\tau \tilde{f}_N\left(\left|X_s^{\sigma^N}\right|\right)\mathrm{d}s\right].$$

Since $\tilde{f}_N$ is constant on $(0, \frac{1}{N})$ and decreasing on $(\frac{1}{N}, \eta)$, by Proposition 2.15 we can choose σ^N such that

$$\sigma_t^N = \Big[e_1;\ \ 0;\ \ \ldots;\ \ 0\Big], \quad \text{for} \quad |X_t| < \frac{1}{N},$$

and

$$\sigma_t^N = \sigma^0(X_t), \quad \text{for} \quad |X_t| \in \left[\frac{1}{N}, \eta\right).$$

Also define a control σ^δ that coincides with σ^N except that we set

$$\sigma_t^\delta = \Big[e_1;\ \ 0;\ \ \ldots;\ \ 0\Big], \quad \text{for} \quad |X_t| \in \left[\frac{1}{N}, \delta\right).$$

Under either control σ^N or σ^δ, the process $t \mapsto |X_t|$ is deterministically increasing on the interval (δ, η), by Lemma 2.4. Therefore the error between the value $v_N^S(0)$

and the expected cost of choosing the control σ^δ with the cost f is

$$\begin{aligned}0 \leq E_N(\delta) :=&\mathbb{E}^0\left[\int_0^\tau \tilde{f}\left(\left|X_s^{\sigma^\delta}\right|\right)\mathrm{d}s\right] - v_N^S(0)\\ =&\mathbb{E}^0\left[\int_0^{\tau_\delta} \tilde{f}\left(\left|X_s^{\sigma^\delta}\right|\right)\mathrm{d}s\right] - \mathbb{E}^0\left[\int_0^{\tau_\delta} \tilde{f}_N\left(\left|X_s^{\sigma^N}\right|\right)\mathrm{d}s\right].\end{aligned}$$

In the ball $B_\delta(0)$, the process X^{σ^δ} is equal to a one-dimensional Brownian motion in the direction e_1 and so, making a calculation with the Green's function similar to (2.27) in the proof of Proposition 2.20, we find that

$$\mathbb{E}^0\left[\int_0^{\tau_\delta} \tilde{f}\left(\left|X_s^{\sigma^\delta}\right|\right)\mathrm{d}s\right] = 2\int_0^\delta (\delta - y)\,\tilde{f}(y)\,\mathrm{d}y.$$

We now compute the expected cost under the control σ^N. When $\left|X_t^{\sigma^N}\right| \in (\frac{1}{N}, \delta)$, the process X^{σ^N} follows tangential motion, and so we can calculate

$$\mathbb{E}^{\frac{1}{N}}\left[\int_0^{\tau_\delta} \tilde{f}_N\left(\left|X_s^{\sigma^{N,\varepsilon}}\right|\right)\mathrm{d}s\right] = \int_0^{\delta^2 - N^{-2}} \tilde{f}(\sqrt{N^{-2}+s})\,\mathrm{d}s = 2\int_{\frac{1}{N}}^\delta s\tilde{f}(s)\,\mathrm{d}s.$$

In the ball $B_{\frac{1}{N}}(0)$, the process X^{σ^N} is a one-dimensional Brownian motion and so, making another calculation with the Green's function, we can write

$$\begin{aligned}\mathbb{E}^0\left[\int_0^{\tau_\delta} \tilde{f}_N\left(\left|X_s^{\sigma^N}\right|\right)\mathrm{d}s\right] &= \mathbb{E}^0\left[\int_0^{\tau_{\frac{1}{N}}} \tilde{f}_N\left(\left|X_s^{\sigma^N}\right|\right)\mathrm{d}s\right] + \mathbb{E}^{\frac{1}{N}}\left[\int_0^{\tau_\delta} \tilde{f}_N\left(\left|X_s^{\sigma^N}\right|\right)\mathrm{d}s\right]\\ &= \frac{1}{N^2}\tilde{f}\left(\frac{1}{N}\right) + 2\int_{\frac{1}{N}}^\delta y\tilde{f}(y)\,\mathrm{d}y.\end{aligned}$$

Therefore the error is

$$E_N(\delta) = 2\int_0^\delta (\delta - y)\tilde{f}(y)\,\mathrm{d}y - 2\int_{\frac{1}{N}}^\delta y\tilde{f}(y)\,\mathrm{d}y - \frac{1}{N^2}\tilde{f}\left(\frac{1}{N}\right).$$

Since $\int_0^r \tilde{f}(s)\,\mathrm{d}s < \infty$ for any $r > 0$, we can take the limit as $N \to \infty$ to get

$$E(\delta) := \lim_{N\to\infty} E_N(\delta) = 2\int_0^\delta (\delta - 2y)\,\tilde{f}(y)\,\mathrm{d}y,$$

and then taking the limit as $\delta \to 0$ gives

$$0 \leq E(\delta) = 2\int_0^\delta (\delta - 2y)\,\tilde{f}(y)\,\mathrm{d}y \xrightarrow{\delta\to 0} 0. \tag{2.29}$$

Returning to the definition of $E_N(\delta)$, for fixed $\delta \in (0,\eta)$ and $N > \frac{1}{\delta}$, we recall that

$$v_N^S(0) + E_N(\delta) = \mathbb{E}^0\left[\int_0^\tau \tilde{f}\left(\left|X_s^{\sigma^\delta}\right|\right) \mathrm{d}s\right]. \tag{2.30}$$

Since $\tilde{f}_N \leq \tilde{f}$, we have

$$v^S(0) + E_N(\delta) \geq v_N^S(0) + E_N(\delta).$$

By the definition of v^S, note that

$$\mathbb{E}^0\left[\int_0^\tau \tilde{f}\left(\left|X_s^{\sigma^\delta}\right|\right) \mathrm{d}s\right] \geq v^S(0).$$

Combining these inequalities with (2.30), we see that

$$v^S(0) + E_N(\delta) \geq v_N^S(0) + E_N(\delta) \geq v^S(0).$$

Since the sequence $\left(v_N^S(0)\right)_{N\in\mathbb{N}}$ is monotone, we can take the limit as $N \to \infty$ and find that

$$v^S(0) + E(\delta) \geq \lim_{N\to\infty} v_N^S(0) + E(\delta) \geq v^S(0).$$

Having calculated that $\lim_{\delta\to 0} E(\delta) = 0$ in (2.29), we have that

$$v^S(0) = \lim_{N\to\infty} v_N^S(0),$$

and $v^S(0) = V(0)$.

Defining v_N^W as in the proof of Proposition 2.22, we have $v_N^S(0) = v_N^W(0)$ for any $N \in \mathbb{N}$, and so we can conclude that $v^S(0) = v^W(0) = V(0)$. □

Remark 2.25. Note that, if the growth rate of $\tilde{f}$ is such that, for any $r > 0$,

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s = \infty,$$

then the error $E(\delta)$ in the proof of Proposition 2.24 is infinite for all δ. Therefore the above argument does not generalise to costs with faster growth at the origin.

We now consider decreasing costs with faster growth at the origin.

Proposition 2.26. *Suppose that Assumption 2.19 holds and that there exists $\eta > 0$ such that $\tilde{f}$ is positive and decreasing on the interval $(0,\eta)$. If, for any $r > 0$,*

$$\int_0^r s\tilde{f}(s)\,\mathrm{d}s = \infty,$$

then

$$v(0) = +\infty.$$

Proof. Once again define $\tilde{f}_N$, f_N and v_N^S as in the proof of Proposition 2.20. Let $N > \frac{1}{\eta}$ and define the control σ^N as in the proof of Proposition 2.24, so that σ^N is optimal for the cost f_N.

Using the calculations of the expected cost under the control σ^N from the proof of Proposition 2.24, we find that

$$\begin{aligned} v_N^S(0) &= \mathbb{E}^0 \left[\int_0^\tau \tilde{f}_N \left(\left| X_s^{\sigma^N} \right| \right) \mathrm{d}s \right] \\ &= \mathbb{E}^0 \left[\int_0^{\tau_\eta} \tilde{f}_N \left(\left| X_s^{\sigma^N} \right| \right) \mathrm{d}s \right] + \mathbb{E}^0 \left[\int_0^\tau \tilde{f} \left(\left| X_s^{\sigma^N} \right| \right) \mathbb{1}_{\left\{ \left| X^{\sigma^N} \right| \in (\eta, R) \right\}} \mathrm{d}s \right] \\ &\geq \frac{1}{N^2} \tilde{f} \left(\frac{1}{N} \right) + 2 \int_{\frac{1}{N}}^\eta y \tilde{f}(y) \, \mathrm{d}y + (R^2 - \eta^2) \min \left\{ \tilde{f}(r) \colon r \in (\eta, R) \right\}. \end{aligned}$$

By the growth condition on $\tilde{f}$, we have

$$\lim_{N \to \infty} \int_{\frac{1}{N}}^\eta y \tilde{f}(y) \, \mathrm{d}y = +\infty.$$

Also, since $\tilde{f}$ is continuous on $(0, R)$,

$$\min \left\{ \tilde{f}(r) \colon r \in (\eta, R) \right\} > -\infty,$$

and so

$$\lim_{N \to \infty} v_N^S(0) = +\infty.$$

For any $N \in \mathbb{N}$, defining v_N^W as in the proof of Proposition 2.22, we have

$$v^S(0) \geq v^W(0) \geq v_N^W(0) = v_N^S(0).$$

Hence

$$v^S(0) = v^W(0) = +\infty. \qquad \square$$

We have now fully characterised the value function for any radially symmetric cost, except for the value at the origin when the cost function is decreasing at the origin and grows at such a rate that, for any $r > 0$,

$$\int_0^r \tilde{f}(s) \, \mathrm{d}s = \infty \quad \text{and} \quad \int_0^r s \tilde{f}(s) \, \mathrm{d}s < \infty.$$

We now state the result for this remaining growth regime.

Proposition 2.27. *Suppose that Assumption 2.19 holds and that there exists $\eta > 0$ such that $\tilde{f}$ is positive and decreasing on the interval $(0, \eta)$. If, for any $r > 0$,*

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s = \infty \quad \textit{and} \quad \int_0^r s\tilde{f}(s)\,\mathrm{d}s < \infty,$$

then

$$v(0) = v^S(0) = v^W(0) = V(0) \in (-\infty, \infty),$$

where V is the candidate value function defined in Definition 2.14.

We first prove the result for dimensions $d \geq 3$.

Lemma 2.28. *Under the conditions of Proposition 2.27 with $d \geq 3$, we have*

$$v(0) = v^S(0) = v^W(0) = V(0) \in (-\infty, \infty).$$

Proof. In this case, we can follow the same argument as in the proof of Proposition 2.24 except that we replace the constant control $\begin{bmatrix} e_1; & 0; & \dots; & 0 \end{bmatrix}$ with $\frac{1}{d}I$, where I is the d-dimensional identity matrix. Instead of following a one-dimensional Brownian motion at the origin, the controlled processes σ^N and σ^δ follow a scaled d-dimensional Brownian motion. We now verify that the approximation arguments in Proposition 2.24 hold with this change.

We will use the Green's function for the d-dimensional Brownian motion B, as defined in Section 3.3 of the book [45] of Mörters and Peres. By Theorem 3.32 and 3.33 of [45] and the radial symmetry of f, there are constants $C, C' > 0$ such that, for any $\delta \in (0, \eta)$,

$$\begin{aligned}
\mathbb{E}^0 \left[\int_0^{\tau_\delta} f(B_s)\,\mathrm{d}s \right] &= C \int_{B_\delta} |y|^{2-d} f(y)\,\mathrm{d}y \\
&= C' \int_0^\delta r^{d-2} \tilde{f}(r) r^{d-1}\,\mathrm{d}r \\
&= C' \int_0^\delta r\tilde{f}(r)\,\mathrm{d}r.
\end{aligned}$$

Therefore, since $\int_0^r s\tilde{f}(s)\,\mathrm{d}s < \infty$ for any $r > 0$, we have

$$\lim_{\delta \to 0} \mathbb{E}^0 \left[\int_0^{\tau_\delta} f(B_s)\,\mathrm{d}s \right] = 0.$$

Hence following the same arguments as in the proof Proposition 2.24 leads to the desired result. $\square$

Now suppose that $d = 2$. Note that, from the form of the Green's function for 2-dimensional Brownian motion given in Theorem 3.34 of [45], we can see that the argument used for $d \geq 3$ is no longer valid. In the following lemma, we treat the weak control problem in dimension $d = 2$, delaying the proof of the result for the strong control problem until Section 3.5.

Lemma 2.29. *Under the conditions of Proposition 2.27 with $d = 2$, the weak value function is given by*

$$v^W(0) = V(0) \in (-\infty, \infty).$$

Proof. Retaining the notation of the proof of Proposition 2.26, we have that, for any $y \in D$ with $|y| = \eta$,

$$\begin{aligned} v^W(0) &\geq \lim_{N\to\infty} v_N^S(0) \\ &= V(0) = 2\int_0^\eta \xi \tilde{f}(y)\,\mathrm{d}\xi + V(y), \end{aligned} \tag{2.31}$$

by Proposition 2.15 and the definition of V in Definition 2.14.

In Theorem 4.3 of [41], Larsson and Ruf prove that, for $d = 2$, there exists a weak solution X^{σ^0} of the SDE

$$\mathrm{d}X_t = \sigma^0(X_t)\,\mathrm{d}B_t; \quad X_0 = 0. \tag{2.32}$$

The process X^{σ^0} follows tangential motion starting from the origin, as defined in Definition 2.3. By Lemma 2.4, for any $t \geq 0$,

$$\left|X_t^{\sigma^0}\right| = \sqrt{t},$$

and so

$$\mathbb{E}^0\left[\int_0^{\tau_\eta} f(X_s^{\sigma^0})\,\mathrm{d}s\right] = \int_0^{\eta^2} \tilde{f}(\sqrt{s})\,\mathrm{d}s = 2\int_0^\eta \xi\tilde{f}(\xi)\,\mathrm{d}\xi.$$

Note that Assumption 1.16 holds, and so we can apply the dynamic programming principle from Proposition 1.17 to see that, for any $y \in D$ with $|y| = \eta$,

$$\begin{aligned} v^W(0) \leq v^S(0) &\leq \mathbb{E}^0\left[\int_0^{\tau_\eta} f(X_s^{\sigma^0})\,\mathrm{d}s + v^S(X_{\tau_\eta}^{\sigma^0})\right] \\ &= 2\int_0^\eta \xi\tilde{f}(\xi)\,\mathrm{d}\xi + V(y), \end{aligned}$$

using the result that $v^S = V$ away from the origin from Proposition 2.22.

Combining the above inequality with (2.31), we have

$$v^W(0) = V(0),$$

as required. $\square$

We summarise the preceding results in the following extension of Proposition 2.15.

Theorem 2.30. *Suppose that Assumption 2.19 is satisfied, and let $V : D \to \mathbb{R}$ be the candidate value function defined in Definition 2.14. Then the value function is*

$$v = v^S = v^W = V.$$

Moreover, we can determine when the value function is finite. If there exists $\eta > 0$ such that $\tilde{f}$ is increasing on the interval $(0, \eta)$, then

$$\begin{cases} v > -\infty, & \text{if} \quad \int_0^r \tilde{f}(s)\,\mathrm{d}s > -\infty \quad \text{for any} \quad r > 0, \\ v \equiv -\infty, & \text{if} \quad \int_0^r \tilde{f}(s)\,\mathrm{d}s = -\infty \quad \text{for any} \quad r > 0. \end{cases}$$

If there exists $\tilde{\eta} > 0$ such that $\tilde{f}$ is decreasing on the interval $(0, \tilde{\eta})$, then

$$\begin{cases} v < \infty, & \text{if} \quad \int_0^r s\tilde{f}(s)\,\mathrm{d}s < \infty \quad \text{for any} \quad r > 0, \\ v(0) = \infty \quad \text{and} \quad v(x) < \infty,\ x \in D \setminus \{0\}, & \text{if} \quad \int_0^r s\tilde{f}(s)\,\mathrm{d}s = \infty \quad \text{for any} \quad r > 0. \end{cases}$$

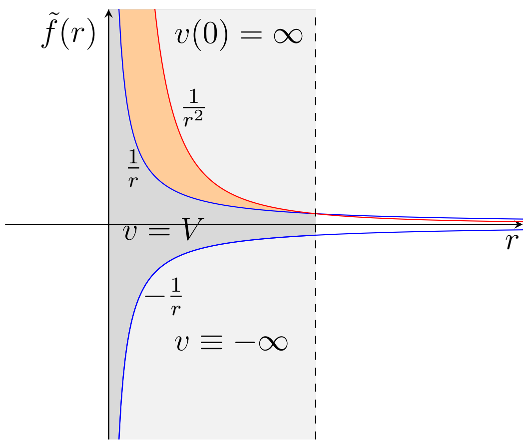


Figure 2.8: Figure showing the distinct growth regimes for the cost function in Theorem 2.30, highlighting the case where, for any $r > 0$, $\int_0^r \tilde{f}(s)\,\mathrm{d}s = \infty$ and $\int_0^r s\tilde{f}(s)\,\mathrm{d}s < \infty$, as in Proposition 2.27.

We now discuss what remains to prove Proposition 2.27 in the case $d = 2$.

Remark 2.31. Recall that, in Proposition 1.7, we appealed to Theorem 4.5 of El Karoui and Tan's paper [20] to show equality between weak and strong value functions, under the assumption that the cost function f was upper semicontinuous and bounded above by a constant.

Under the assumptions of Proposition 2.27, we cannot apply Theorem 4.5 of [20], since one of the assumptions of that theorem is no longer satisfied. Namely, in our setup, Theorem 4.5 of [20] is only applicable if the random variable

$$F_\tau := \int_0^\tau f(X_s)\,\mathrm{d}s$$

is bounded above by some random variable ξ that is uniformly integrable under the family of probability measures $\mathcal{P}_0$ defined in Section 1.4.1. We show that this condition is not satisfied as follows. Let e_1 be the unit vector in the first coordinate direction and define X^1 by

$$X_t^1 = e_1 B_t, \quad t \geq 0.$$

Then let $\mathbb{P}^{X^1}$ be the law of the process X^1 and choose $\mathbb{P} \in \mathcal{P}_0$ to be the product measure

$$\mathbb{P} := \mathbb{P}^{X^1} \times \delta_{e_1}.$$

Following the same Green's function calculation as in (2.27) in Proposition 2.20, we use Proposition B.5 to compute that

$$\mathbb{E}^{\mathbb{P}}\left[\int_0^\tau f(X_s)\,\mathrm{d}s\right] = \int_0^R (R-r)\tilde{f}(r)\,\mathrm{d}r = +\infty,$$

due to the growth condition (3.3) on $\tilde{f}$ at the origin.

Hence there does not exist any uniformly integrable upper bound on F_τ and Theorem 4.5 of [20] does not apply.

In Lemma 2.29, we found the weak value function at the origin by using the fact that there exists a weak solution of the SDE (2.32) describing tangential motion started from the origin. We will show in Theorem 3.4 that the SDE (2.32) has no strong solution. Therefore, we cannot follow the same argument as in the proof of Lemma 2.29 to find the strong value function.

In Chapter 3, we will show that the natural filtration of a weak solution of (2.32) is generated by a Brownian motion. As a consequence, we will deduce that the strong and weak value functions are in fact equal in Section 3.5.

Since the SDE (2.32) has no strong solution, the strong control cannot depend only on the current position of the controlled process. In the next section, we introduce the terminology of *Markov controls* for such controls that depend only on the current position of the controlled process.

2.5 Markov controls

We now define a Markov formulation of the control problem. This is a stronger formulation than the strong and weak formulations introduced in Section 1.4.1. Markov controls are defined similarly in Section 3 of [26, Chapter IV] and Section 3.1 of [58].

Definition 2.32. For each $x \in D$, define the set $\mathcal{U}_x^M \subset \mathcal{U}$ of *Markov controls* as follows. A control $\nu \in \mathcal{U}$ is an element of $\mathcal{U}_x^M$ if and only if, for all $t \geq 0$, $\nu_t = \sigma(X_t^{\sigma,x})$, where $X^{\sigma,x}$ is a strong solution of the SDE

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t; \quad X_0 = x,$$

for some Borel function $\sigma : D \to U$. We then write $X^\nu = X^{\sigma,x}$.

The Markov formulation of the control problem is to find the *Markov value function* $v^M : D \to \mathbb{R}$, defined by

$$v^M(x) = \inf_{\nu \in \mathcal{U}_x^M} \mathbb{E}^x \left[\int_0^\tau f(X_s^\nu)\,\mathrm{d}s + g(X_\tau^\nu) \right], \quad x \in D.$$

Proposition 2.33. *For any $x \in D$, $v^M(x) \geq v^S(x)$.*

Proof. This follows immediately from the inclusion $\mathcal{U}_x^M \subset \mathcal{U}$. □

We will now show that, under the conditions of Theorem 2.30, the Markov formulation of the control problem is equivalent to the weak and strong formulations, with one possible exception.

Proposition 2.34. *Suppose that Assumption 2.19 is satisfied, and let $V : D \to \mathbb{R}$ be the candidate value function defined in Definition 2.14. Let $x \in D$ and suppose, moreover, that one of the following conditions hold:*

(i) $\tilde{f}$ is increasing on the interval $(0, \eta)$;

(ii) $\tilde{f}$ is decreasing on the interval $(0, \eta)$ and $x \in D \setminus \{0\}$;

(iii) $\tilde{f}$ is decreasing on the interval $(0, \eta)$, $x = 0$ and, for any $r > 0$, $\int_0^r \tilde{f}(s) < \infty$;

(iv) $\tilde{f}$ is decreasing on the interval $(0, \eta)$, $x = 0$ and, for any $r > 0$, $\int_0^r s\tilde{f}(s) = \infty$;

(v) $\tilde{f}$ *is decreasing on the interval* $(0, \eta)$ *and* $d \geq 3$.

Then the Markov value function is given by

$$v^M(x) = v^S(x) = v^W(x) = V(x).$$

Proof. We treat each of the conditions from the statement of the proposition in turn.

(i) Let the control $\sigma^\star \in \mathcal{U}$ be as defined in (2.10). First note that, if $|x| \in [0, r_1)$, then $X^{\sigma^\star}$ is a strong solution of the SDE

$$\mathrm{d}X_t = \sigma^1(x)\,\mathrm{d}B_t, \quad X_0 = x,$$

up to the first hitting time of radius r_1, where the coefficient $\sigma^1(x)$ is a constant. When $\left|X_t^{\sigma^\star}\right| \in [r_i, s_i]$ for some $i \geq 1$, we have $\sigma_t^\star = \sigma^0(X_t^{\sigma^\star})$. When $\left|X_t^{\sigma^\star}\right| \in (s_i, r_{i+1})$ for some $i \geq 1$, we have $\sigma_t^\star = \sigma^1(X_{\tau_{s_i}}^{\sigma^\star})$, and we can calculate that $X^{\sigma^\star}$ actually solves

$$\mathrm{d}X_t = \sigma^1(X_t)\,\mathrm{d}B_t$$

in this region. Therefore, fixing $x \in D$, there is a Markov control $\tilde{\sigma}^\star \in \mathcal{U}_x^M$ such that

$$\mathbb{E}^x\left[\int_0^\tau f(X_s^{\tilde{\sigma}^\star})\,\mathrm{d}s + g(X_\tau^{\tilde{\sigma}^\star})\right] = \mathbb{E}^x\left[\int_0^\tau f(X_s^{\sigma^\star})\,\mathrm{d}s + g(X_\tau^{\sigma^\star})\right].$$

Under the conditions of Proposition 2.15, $\sigma^\star \in \mathcal{U}$ is optimal for the strong formulation of the control problem, and so $v^M(x) \leq v^S(x)$. By Proposition 2.33, we also have $v^M(x) \geq v^S(x)$. Hence $v^M(x) = v^S(x)$.

To complete the proof under condition *(i)*, we note that the proof of Proposition 2.20 can be adapted easily to the Markov value function rather than the strong value function. We conclude that, under condition *(i)*, $v^M = v^S$.

(ii) Let $x \in D \setminus \{0\}$. Then, taking the definition of the control $\sigma^\star \in \mathcal{U}$ from (2.11), we see that $\left|X_t^{\sigma^\star}\right| > 0$ for all $t \geq 0$. By similar reasoning as above, there is then a control $\tilde{\sigma}^\star \in \mathcal{U}_x^M$ such that

$$\mathbb{E}^x\left[\int_0^\tau f(X_s^{\tilde{\sigma}^\star})\,\mathrm{d}s + g(X_\tau^{\tilde{\sigma}^\star})\right] = \mathbb{E}^x\left[\int_0^\tau f(X_s^{\sigma^\star})\,\mathrm{d}s + g(X_\tau^{\sigma^\star})\right].$$

Once again, under the conditions of Proposition 2.15, the above expression is equal to $v^S(x)$, and so $v^M(x) \leq v^S(x)$. Combining this with Proposition 2.33, we have $v^M(x) = v^S(x)$. Noting that the proof of Proposition 2.22 can be adapted to the case of Markov controls concludes the proof under condition *(ii)*.

(iii) We note that the proof of Proposition 2.24 can be easily adapted to Markov controls, and so the result holds under condition *(iii)*.

(iv) Under condition *(iv)*, Proposition 2.26 shows that $v^S(0) = +\infty$. By Proposition 2.33, $v^M(0) \geq v^S(0)$, and so we also have $v^M(0) = +\infty$.

(v) The additional case included in condition *(v)* that has not been covered by conditions *(ii)*–*(iv)* is when $d \geq 3$, $x = 0$ and, for any $r > 0$,

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s = \infty, \quad \text{and} \quad \int_0^r s\tilde{f}(s)\,\mathrm{d}s < \infty.$$

In this case, as with condition *(iii)*, the proof of Lemma 2.28 can easily be adapted to Markov controls, and so the result holds. □

The above result does not give us the value $v^M(0)$ for dimension $d = 2$ in the case where $\tilde{f}$ is decreasing on the interval $(0, \eta)$ and, for any $r > 0$,

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s = \infty, \quad \text{and} \quad \int_0^r s\tilde{f}(s)\,\mathrm{d}s < \infty.$$

For the strong and weak formulations of the control problem, this case is treated by Proposition 2.27, which we will prove in Section 3.5. We conjecture that, in this case, there is a gap between the Markov value function and the strong and weak value functions at the origin. In Chapter 3, we will state this conjecture formally and prove two results that support the conjecture.

CHAPTER 3
SDES WITH NO STRONG SOLUTION ARISING FROM A PROBLEM OF STOCHASTIC CONTROL

In this chapter, we present two examples of SDEs that have no strong solution. These SDEs arise from considering the control problem of Chapter 2 in a regime of moderate growth at the origin in dimension $d = 2$. We prove that a weak solution of one of the SDEs generates a Brownian filtration and use this to prove equality between weak and strong value functions. Non-existence of strong solutions leads us to conjecture that there is a gap between the Markov value function and the strong and weak value functions at the origin. We prove the main results of this chapter by building on the study of an example of an SDE with no strong solution given by Tsirelson.

3.1 Introduction

We fix $d = 2$ in this chapter. We consider the control problem of Chapter 2 under Assumption 2.19 in the case that the cost function $\tilde{f}$ is decreasing on the interval $(0, \eta)$ and satisfies the growth conditions

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s = \infty, \quad \text{and} \quad \int_0^r s\tilde{f}(s)\,\mathrm{d}s, \quad \text{for any} \quad r > 0.$$

In Section 3.5, we will complete the proof of Proposition 2.27 by showing that the weak and strong value functions are equal at the origin under the above conditions.

We will also make the conjecture that there is a gap between the Markov value function and the strong value function. We will prove two results, Theorem 3.4 and Theorem 3.15, which give support to this conjecture. In Theorem 3.4, we will show that the SDE

$$\mathrm{d}X_t = \frac{1}{|X_t|}\begin{bmatrix}-X_t^2\\ X_t^1\end{bmatrix}\mathrm{d}B_t$$

has a weak solution starting from the origin but no strong solution. In Theorem 3.15, we consider SDEs whose solutions give rise to a sequence of expected costs that approximate the strong and weak value function at the origin. We show that these SDEs also have no strong solution starting from the origin, further supporting the conjecture.

We now state the definitions of weak and strong solution that we will be using in this chapter.

Notation. As different filtrations play an important role in this chapter, we will fix the following notation. For any stochastic process X, $\mathbb{F}^X = (\mathcal{F}_t^X)_{t\geq 0}$ will denote the augmentation of the natural filtration of X satisfying the usual conditions. Any other filtration introduced in this chapter will also be assumed to satisfy the usual conditions.

We take the following definitions of weak and strong solutions from Karatzas and Shreve [38, Chapter 5]. Let $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^m$ be Borel measurable functions and W an m-dimensional Brownian motion. Consider the SDE

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t. \tag{3.1}$$

Definition 3.1 (Strong solution)**.** [38, Chapter 5, Definition 2.1]
Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which an m-dimensional Brownian motion W and a random variable ξ are defined. A *strong solution* of the SDE (3.1) with *initial condition* ξ is a continuous d-dimensional process $(X_t)_{t\geq 0}$ such that

i. X is adapted to $\mathbb{F}^W$;

ii. $\mathbb{P}[X_0 = \xi] = 1$;

iii. $\mathbb{P}\left[\int_0^t \left(|b_i(s, X_s)| + |\sigma_{ij}(s, X_s)|^2\right)\mathrm{d}s < \infty\right] = 1$, for all i, j, $t \geq 0$;

iv. $X_t = X_0 + \int_0^t b(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}W_s$, for all $t \geq 0$.

Definition 3.2 (Weak solution)**.** [38, Chapter 5, Definition 3.1]
A *weak solution* of the SDE (3.1) with *initial distribution* μ is a triple

$$((X, W), (\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F}),$$

where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is a filtration satisfying the usual conditions, W is an m-dimensional $\mathbb{F}$-Brownian motion, and X is a d-dimensional continuous $\mathbb{F}$-adapted process, such that

i. X_0 has law μ;

ii. $\mathbb{P}\left[\int_0^t \left(|b_i(s, X_s)| + |\sigma_{ij}(s, X_s)|^2\right) \mathrm{d}s < \infty\right] = 1$, for all i, j, $t \geq 0$;

iii. $X_t = X_0 + \int_0^t b(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}W_s$, for all $t \geq 0$.

Remark 3.3. We emphasise that the key difference between Definition 3.1 and Definition 3.2 is that the Brownian motion W and the process X in a weak solution can be chosen together, and there is no requirement for X to be adapted to the natural filtration of W.

The first main result of this chapter is the following.

Theorem 3.4. *Let B be a real-valued Brownian motion and consider the SDE*

$$\mathrm{d}X_t = \frac{1}{|X_t|}\begin{bmatrix} -X_t^2 \\ X_t^1 \end{bmatrix} \mathrm{d}B_t. \tag{3.2}$$

Then there exists a weak solution of the SDE (3.2) *with initial distribution δ_0. However, there is no strong solution of the SDE* (3.2) *with initial condition $X_0 = 0$.*

The first statement of the theorem is proved by Larsson and Ruf in Theorem 4.3 of [41]. We will complete the proof of the theorem in Section 3.6.

In light of this result, we cannot simply adapt the proof of Lemma 2.29 in order to prove Proposition 2.27. In Section 3.5, we will use properties of the filtration generated by a weak solution of (3.2) to prove Proposition 2.27.

Theorem 3.4 also suggests that, in the case covered by Proposition 2.27, there may be a gap between the Markov value function and the strong and weak value functions. This is the assertion of the following conjecture.

Conjecture 3.5. *Fix $d = 2$. Suppose that Assumption 2.19 is satisfied and that there exists $\eta > 0$ such that $\tilde{f}$ is decreasing on $(0, \eta)$. Suppose moreover that, for any $r > 0$,*

$$\int_0^r \tilde{f}(s)\,\mathrm{d}s = \infty \quad \text{and} \quad \int_0^r s\tilde{f}(s)\,\mathrm{d}s < \infty. \tag{3.3}$$

Then

$$v^M(0) > v^S(0) = v^W(0) = V(0).$$

In Proposition 2.27, we show that $v^S(0) = v^W(0) = V(0)$, so the new assertion in this conjecture is that $v^M(0) > v^S(0)$.

We begin this chapter by presenting some examples of SDEs from the literature that have a weak solution but no strong solution. We will draw similarities between the SDE (3.2) and Tanaka and Tsirelson's examples of SDEs with no strong solution, as described, for example, in [53, Chapter V, §3].

In Section 3.3, we will introduce circular Brownian motion, which is used by Émery and Schachermayer in their study of Tsirelson's equation in [21]. We will show that the angle process of any solution of the SDE (3.2) is a regular time-change of a circular Brownian motion. We use the properties of circular Brownian motion to prove that the filtration generated by any solution of the SDE (3.2) is Brownian in Section 3.4.

In Section 3.5, we use the fact that a solution of (3.2) generates a Brownian filtration to prove Proposition 2.27.

We prove Theorem 3.4 in Section 3.6, showing that the SDE (3.2) has no strong solution starting from the origin, again making use of the properties of circular Brownian motion proved in [21].

In Section 3.7, we consider a class of SDEs whose behaviour approximates that of the SDE (3.2), in the sense that the expected costs associated to their solutions converge. In Theorem 3.15, we will prove that these SDEs also have no strong solution, adapting the proofs of some results on circular Brownian motion from [21].

We will end this chapter in Section 3.8 by discussing what remains to be shown in order to prove Conjecture 3.5.

3.2 SDEs with no strong solution in the literature

We now present two examples of SDEs that are known to have no strong solution. In this section, we collect some results from the literature on Tanaka and Tsirelson's equations. We will go on to show that the SDE (3.2) has similar properties to these SDEs, in order to prove Proposition 2.27 in Section 3.5 and to prove non-existence of strong solutions in Theorem 3.4.

3.2.1 Tanaka's example

A well-known example of an SDE with no strong solution is Tanaka's SDE, which is the following one-dimensional equation:

$$\mathrm{d}X_t = \operatorname{sign}(X_t)\,\mathrm{d}W_t. \tag{3.4}$$

The SDE (3.4) admits a weak solution but not a strong solution. The proof of this can be found, for example, in Example 3.5 of [38, Chapter 5].

To prove that there is no strong solution, the key idea is to show, using the Itô-Tanaka formula, that the inclusion

$$\mathcal{F}_t^W \subseteq \mathcal{F}_t^{|X|}$$

holds for all $t > 0$. Then it is impossible for X to be adapted to $\mathcal{F}^W$, since $\mathcal{F}_t^{|X|} \subsetneq \mathcal{F}_t^X$ for all $t > 0$.

In order to prove Theorem 3.4, we show similar inclusions to the ones above, where the increments of the solution of the SDE (3.2) play the role of the absolute value of the solution of Tanaka's SDE.

3.2.2 Tsirelson's example

A further example of an SDE with no strong solution was presented by Tsirelson in [59]. Tsirelson's example is the one-dimensional equation

$$\mathrm{d}X_t = b(t, X_\cdot)\,\mathrm{d}t + \mathrm{d}W_t, \tag{3.5}$$

with initial condition $X_0 = 0$, where b is chosen as follows.

Fix a decreasing sequence $(t_n)_{n\in -\mathbb{N}\cup\{0\}}$ such that $t_0 = 1$ and $\lim_{n\to-\infty} t_n = 0$. Denote the increments of X and t by $\Delta X_j = X_{t_j} - X_{t_{j-1}}$ and $\Delta t_j = t_j - t_{j-1}$, respectively, and define

$$b(t, X_\cdot) := \sum_{k\in -\mathbb{N}} \left(\frac{\Delta X_k}{\Delta t_k} - \left\lfloor \frac{\Delta X_k}{\Delta t_k} \right\rfloor \right) \mathbb{1}_{(t_k, t_{k+1}]}(t). \tag{3.6}$$

At a time $t \in (t_k, t_{k+1}]$, $b(t, X_\cdot)$ is the fractional part of $\frac{\Delta X_{t_k}}{\Delta t_k}$.

Any weak solution of the SDE (3.5) has the following properties, as proved, for example, in Theorem 18.3 of [53, Chapter V]:

i. At any time $t > 0$, the natural filtration of a solution X has the decomposition

$$\mathcal{F}_t^x = \mathcal{F}_t^B \vee \sigma(B(t, X_\cdot));$$

ii. For each $k \in -\mathbb{N}$, $b(t_k, X_\cdot)$ is uniformly distributed on $[0, 1)$ and independent of $\mathcal{F}_\infty^B$;

iii. The sigma-algebra $\mathcal{F}_{0+}^X$ is trivial.

Note that the drift term (3.6) in Tsirelson's SDE depends on the whole history of the process X. Therefore, to define strong and weak solutions, we need to extend Definition 3.1 and Definition 3.2 in the same way as in Definition 3.14 of [38, Chapter 5]. As remarked in [53, Chapter V], for bounded drifts depending only on the current value of the process, Zvonkin proved in [67] that a strong solution of (3.5) always exists.

In the following sections, we refer extensively to the work of Émery and Schachermayer in [21], where they demonstrate a connection between Tsirelson's example and circular Brownian motion. In [21], Émery and Schachermayer use this connection to show that the natural filtration of a weak solution of Tsirelson's equation (3.5) is generated by a Brownian motion.

3.3 Circular Brownian motion

The key observation in our proof of Theorem 3.4 is that the angle of any solution of the SDE (3.2) is a deterministic time-change of a circular Brownian motion. We take the following definition of a circular Brownian motion from the paper [21] of Émery and Schachermayer.

Definition 3.6 (Circular Brownian motion)**.** Let $(\phi_t)_{t\in\mathbb{R}}$ be a continuous $\mathbb{R}/2\pi\mathbb{Z}$-valued process. For any $s, t \in \mathbb{R}$ with $s \le t$, denote by $\int_s^t \mathrm{d}\phi_r$ the $\mathbb{R}$-valued random variable that depends continuously on t, vanishes for $t = s$, and satisfies

$$\int_s^t \mathrm{d}\phi_r \equiv \phi_t - \phi_s \mod 2\pi.$$

Let $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{R}}$ be a filtration. We say that ϕ is a *circular Brownian motion* for $\mathbb{F}$ if ϕ is adapted to $\mathbb{F}$ and, for each $s \in \mathbb{R}$, the process

$$[s, \infty) \ni t \mapsto \int_s^t \mathrm{d}\phi_r$$

is a standard Brownian motion for the filtration $(\mathcal{F}_t)_{t\in[s,\infty)}$.

With this definition in hand, we now show how a circular Brownian motion arises in our example. By Theorem 4.3 of [41], we know that there exists a weak solution $((X, B), (\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F})$ of the SDE (3.2). We can apply Itô's formula to show that there is an $\mathbb{R}/2\pi\mathbb{Z}$-valued process θ, which we call the *angle process* of the solution, such that

$$X_t = \sqrt{t}\begin{bmatrix}\cos\theta_t\\ \sin\theta_t\end{bmatrix}, \quad \text{for all} \quad t > 0,$$

and θ satisfies

$$\mathrm{d}\theta_t = t^{-\frac{1}{2}}\,\mathrm{d}B_t. \tag{3.7}$$

We now show that θ is a circular Brownian motion, up to a time-change. We define a regular time-change as in [21].

Definition 3.7. A function $a : \mathbb{R} \to (0, \infty)$ is a *regular time-change* if a is an increasing absolutely continuous bijection with absolutely continuous inverse.

Proposition 3.8. *Let* $((X, B), (\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F})$ *be a weak solution of the SDE* (3.2)*. Then the associated angle process* $(\theta_t)_{t>0}$ *is a regular time-change of a circular Brownian motion.*

Proof. Define the function $a : \mathbb{R} \to (0, \infty)$ by $a(t) = e^t$, $t \in \mathbb{R}$. Then a is a regular time-change. Define the time-changed process

$$(\tilde{\theta}_t)_{t\in\mathbb{R}} = (\theta_{a(t)})_{t\geq 0}.$$

Since, for any $t > 0$, there is a one-to-one deterministic correspondence between $X_t \in \mathbb{R}^2$ and $\theta_t \in \mathbb{R}/2\pi\mathbb{Z}$, the angle process θ is adapted to $\mathbb{F}$. Now define the time-changed filtration

$$\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{t\in\mathbb{R}} = \left(\mathcal{F}_{a(t)}\right)_{t\in\mathbb{R}}.$$

We will show that $\tilde{\theta}$ is a circular Brownian motion for $\tilde{\mathbb{F}}$.

Since a is a regular time-change, $\tilde{\theta}$ is adapted to $\tilde{\mathbb{F}}$. We also see that the $\mathbb{R}/2\pi\mathbb{Z}$-valued process $\tilde{\theta}$ is continuous. Now fix $s \in \mathbb{R}$ and consider the process

$$[s, \infty) \ni t \mapsto \int_s^t \mathrm{d}\tilde{\theta}_r = \int_s^t a(r)^{-\frac{1}{2}}\,\mathrm{d}B_{a(r)},$$

using the expression (3.7).

Since B is an $\mathbb{F}$-Brownian motion and a is a regular time-change, we have that

$$[s, \infty) \ni t \mapsto \int_s^t \mathrm{d}B_{a(r)}$$

is a $(\tilde{\mathcal{F}}_t)_{t\in[s,\infty)}$-martingale, with quadratic variation

$$\left\langle \int_s^{\cdot} \mathrm{d}B_{a(r)} \right\rangle_t = a(t) - a(s),$$

and so

$$[s, \infty) \ni t \mapsto \int_s^t \mathrm{d}\tilde{\theta}_r$$

is a continuous $(\tilde{\mathcal{F}}_t)_{t\in[s,\infty)}$-martingale. We can calculate the quadratic variation

$$\left\langle \int_s^{\cdot} \mathrm{d}\tilde{\theta}_r \right\rangle_t = \int_s^t a(r)^{-1}\,\mathrm{d}a(r) = t - s,$$

since $a(r) = e^r$, for any $r \in \mathbb{R}$.

Therefore, by Lévy's characterisation of Brownian motion, the process

$$[s,\infty) \ni t \mapsto \int_s^t \mathrm{d}\tilde{\theta}_r$$

is an $(\tilde{\mathcal{F}}_t)_{t\in[s,\infty)}$-Brownian motion. Hence $\tilde{\theta}$ is a circular Brownian motion for $\tilde{\mathbb{F}}$. $\square$

We now state two key properties of circular Brownian motion that are proved in [21]. For a circular Brownian motion ϕ, define the *innovation filtration* $\mathcal{H}$ to be the filtration generated by the increments of ϕ; i.e.

$$\mathcal{H}_t := (\{\phi_s - \phi_r\colon\ -\infty < r \le s \le t\}), \quad t \in \mathbb{R}.$$

Then Proposition 1 of [21] states that, for any $t \in \mathbb{R}$,

i. ϕ_t is uniformly distributed;

ii. ϕ_t is independent of $\mathcal{H}_\infty$.

We note the parallel between these properties of circular Brownian motion and the properties of Tsirelson's equation (3.5) stated in Section 3.2.2.

The aim of Émery and Schachermayer's paper [21] is to show that solutions of Tsirelson's equation generate a Brownian filtration, as we discuss in the following section.

3.4 Brownian filtrations

In Proposition 2 of [21], Émery and Schachermayer define a Brownian filtration as follows.

Definition 3.9 (Brownian filtration)**.** A filtration is called *Brownian* if it is the natural filtration of a real-valued Brownian motion starting from the origin.

Note that this definition agrees with the definition of a *strong Brownian filtration* given in Mansuy and Yor's book [44, Definition 6.1].

In the case of Tanaka's SDE (3.4), any weak solution is a Brownian motion, as discussed in Example 3.5 of [38, Chapter 5], and so in this case a weak solution trivially generates a Brownian filtration.

In Proposition 4 of [21], Émery and Schachermayer show that any deterministic time-change of a circular Brownian motion generates a Brownian filtration. The paper [21] concludes with a proof that there is a bijection between any weak solution of Tsirelson's equation and a circular Brownian motion. From this, the authors deduce that any weak solution of Tsirelson's equation generates a Brownian filtration.

In [19], Dubins, Feldman, Smorodinsky and Tsirelson settled an open question by presenting an example of a process that does not generate a Brownian filtration. Their proof relies on the concept of standardness, an invariant of filtrations first introduced by Vershik in the setting of ergodic theory in his doctoral thesis [62].

Another example of a process that does not generate a Brownian filtration is the diffusion that Walsh defined in [64], now known as Walsh's Brownian motion. In [60], Tsirelson proved that Walsh's Brownian motion does not generate a Brownian filtration, by introducing a new invariant of filtrations known as cosiness. Warren later used the same technique in [65] to prove that sticky Brownian motion also does not generate a Brownian filtration. In [22], Émery and Schachermayer provide a discussion of the relationship between the two invariants standardness and cosiness, along with further references to examples of their application.

We will now show that a weak solution of our SDE (3.2) generates a Brownian filtration. This filtration is therefore both standard and cosy, although we do not need to appeal to either of these notions here. The result follows directly from the relationship with circular Brownian motion that we proved in Proposition 3.8 above.

Corollary 3.10. *Let* $((X,B),(\Omega,\mathcal{F},\mathbb{P}),\mathbb{F}))$ *be a weak solution of the SDE* (3.2)*. Then* X *generates a Brownian filtration.*

Proof. Suppose that a weak solution of (3.2) exists, and write

$$X_t = \sqrt{t}\begin{bmatrix}\cos\theta_t\\ \sin\theta_t\end{bmatrix},$$

where θ is the angle process of the solution. Let $\mathbb{F}^X = (\mathcal{F}^X_t)_{t\geq 0}$ be the filtration generated by X. Then, since $X_0 = 0$ is fixed, and X_t is a deterministic bijective function of θ_t for each $t > 0$, we have

$$\mathcal{F}^X_t = \mathcal{F}^\theta_t \quad \text{for all} \quad t \geq 0,$$

where $\mathbb{F}^\theta = (\mathcal{F}^\theta_t)_{t\geq 0}$ is the filtration generated by θ.

We have seen in Proposition 3.8 that $(\theta_t)_{t>0}$ is a regular time-change of a circular Brownian motion. Propositions 2 and 3 of [21] together immediately imply that the natural filtration of any regular time-change of a circular Brownian motion is Brownian.

Hence $\mathbb{F}^\theta$ is Brownian, and it follows that $\mathbb{F}^X$ is Brownian. $\square$

We now use the fact that a weak solution of (3.2) generates a Brownian filtration to complete the proof of Proposition 2.27, showing that the weak and strong value functions are equal at the origin in the intermediate growth regime.

3.5 Proof of Proposition 2.27

Fix a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ on which a $\mathbb{R}$-valued Brownian motion B is defined with natural filtration $\mathbb{F}^B = (F_t^B)_{t\geq 0}$. We know that there exists a weak solution $((X, W), (\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F})$ of (3.2) by Theorem 4.3 of [41]. We will show that there exists an $\mathbb{F}^B$-martingale $\tilde{X}$ that is equal in law to X. This is the key step required to complete the proof of Proposition 2.27.

We will make use of the notion of *isomorphims* between filtered probability spaces in the following proof. We take the following definitions from the paper [22] of Émery and Schachermayer.

Definition 3.11 (Isomorphism)**.** Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, denote the set of random variables on that probability space by $\mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P})$. An *embedding* of $(\Omega, \mathcal{F}, \mathbb{P})$ into another probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ is a map

$$\Psi : \mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P}) \to \mathcal{L}^0(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$$

that commute with Borel operations on random variables and preserves probability laws.

An *isomorphism* from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ is an embedding that is bijective.

Remark 3.12. We follow the same convention as in [42] and also write Ψ for the map in the above definition acting on sigma-algebras, stochastic processes and filtrations.

Definition 3.13. Two filtered probability spaces $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ and $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\mathbb{F}})$, with $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ and $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t\geq 0}$, are *isomorphic* if there exists an isomorphism

$$\Psi : \mathcal{L}^0(\Omega, \mathcal{F}_\infty, \mathbb{P}) \to \mathcal{L}^0(\overline{\Omega}, \overline{\mathcal{F}}_\infty, \overline{\mathbb{P}})$$

such that $\Psi(\mathbb{F}) = \overline{\mathbb{F}}$.

In [42], Laurent gives similar definitions to the above for filtrations in discrete negative time. We will refer to results from [42] in the following proof.

Proof of Proposition 2.27. We have already proved the result for dimensions $d \geq 3$ in Lemma 2.28. We now fix $d = 2$, as in the rest of Chapter 3. In Lemma 2.29, we showed that $v^W(0) = V(0)$. It remains to show that $v^S(0) = V(0)$.

Fix a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ on which a $\mathbb{R}$-valued Brownian motion B is defined with natural filtration $\mathbb{F}^B = (F^B_t)_{t\geq 0}$, and recall the definition of the control set $\mathcal{U}$ from the definition of the strong formulation of the control problem in Section 1.4.1. We will construct an $\mathbb{F}^B$-martingale $X^{\tilde{\nu}}$ such that, for any $t \geq 0$,

$$X^{\tilde{\nu}}_t = \int_0^t \tilde{\nu}_s \,\mathrm{d}B_s,$$

for some $\tilde{\nu} \in \mathcal{U}$, and

$$\mathbb{E}^0\left[\int_0^{\tau_\eta} f(X^{\tilde{\nu}}_s)\,\mathrm{d}s\right] = 2\int_0^\eta \xi \tilde{f}(\xi)\,\mathrm{d}\xi.$$

As noted in the proof of Lemma 2.29, Theorem 4.3 of [41] gives us a weak solution $((X, B', (\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F})$ of the SDE (3.2). By Corollary 3.10, the process X generates a Brownian filtration. That is, there exists a Brownian motion W on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with natural filtration $\mathbb{F}^W = (\mathcal{F}^W_t)_{t\geq 0}$ such that the natural filtration of X is equal to $\mathbb{F}^W$.

Since B and W are both $\mathbb{R}$-valued Brownian motions, they have have the same law and so, as noted in Section 1.6 of [42], the filtered probability spaces

$$(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \mathbb{F}^B) \quad \text{and} \quad (\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F}^W)$$

are isomorphic, as defined in Definition 3.13. That is, there exists an isomorphism

$$\Psi : \mathcal{L}^0(\Omega, \mathcal{F}^W_\infty, \mathbb{P}) \to \mathcal{L}^0(\tilde{\Omega}, \tilde{\mathcal{F}}^B_\infty, \tilde{\mathbb{P}}),$$

as defined in Definition 3.11, such that

$$\Psi(\mathbb{F}^W) = \mathbb{F}^B.$$

Define a process $\tilde{X}$ on the probability space $(\tilde{\Omega}, \mathcal{F}^B_\infty, \tilde{\mathbb{P}})$ by

$$(\tilde{X}_t)_{t\geq 0} = \Psi\left((X_t)_{t\geq 0}\right).$$

For any $t \geq 0$, we have that $\Psi(\mathcal{F}_t^W) = \mathcal{F}_t^B$ and

$$\Psi : \mathcal{L}^0(\Omega, \mathcal{F}_t^W, \mathbb{P}) \to \mathcal{L}^0(\tilde{\Omega}, \tilde{\mathcal{F}}_t^B, \tilde{\mathbb{P}})$$

is an isomorphism, as noted after the definition of an isomorphism in [22]. Therefore, since X is adapted to $\mathbb{F}^W$, it follows that $\tilde{X}$ is adapted to $\mathbb{F}^B$.

Now fix $0 < s < t$. Then we can use Lemma 5.3 of [42] to apply the isomorphism Ψ to a conditional expectation and, since X is an $\mathbb{F}^W$-martingale, we see that

$$\begin{aligned}
\mathbb{E}^{\tilde{\mathbb{P}}}\left[\tilde{X}_t \mid \mathcal{F}_s^B\right] &= \mathbb{E}^{\tilde{\mathbb{P}}}\left[\Psi(X_t) \mid \Psi(\mathcal{F}_s^W)\right] \\
&= \Psi\left(\mathbb{E}^{\mathbb{P}}\left[X_t \mid \mathcal{F}_s^W\right]\right) \\
&= \Psi(X_s) = \tilde{X}_s.
\end{aligned}$$

Hence $\tilde{X}$ is an $\mathbb{F}^B$-martingale. By the definition of an isomorphism in Definition 3.11, we also have that the processes X and $\tilde{X}$ are equal in law.

We now apply the martingale representation theorem, as found for example in Theorem 3.4 of [51, Chapter 5]. This result implies that $t \mapsto \tilde{X}_t$ is continuous and there exists an $\mathbb{F}^B$-progressively measurable $\mathbb{R}$-valued process $\tilde{\nu}$ such that, for any $t \geq 0$, we have the representation

$$\tilde{X}_t = \int_0^t \tilde{\nu}_s \, \mathrm{d}B_s. \tag{3.8}$$

We can also deduce that $\tilde{X}$ has quadratic variation $t \mapsto \langle \tilde{X} \rangle_t = t$, as follows. We know that the quadratic variation of X is $t \mapsto \langle X \rangle_t = t$, and so $t \mapsto |X_t|^2 - t$ is an $\mathbb{F}^W$-martingale. Using Lemma 5.3 of [42] again, we calculate that, for any $0 < s < t$,

$$\begin{aligned}
\mathbb{E}^{\tilde{\mathbb{P}}}\left[\left|\tilde{X}_t\right|^2 - t \;\middle|\; \mathcal{F}_s^B\right] &= \mathbb{E}^{\tilde{\mathbb{P}}}\left[\Psi(|X_t|^2) \mid \Psi(\mathcal{F}_s^W)\right] - t \\
&= \Psi\left(\mathbb{E}^{\mathbb{P}}\left[|X_t|^2 \mid \mathcal{F}^W\right]\right) - t \\
&= \Psi(|X_s|^2 + t - s) - t \\
&= \left|\tilde{X}_s\right|^2 - s.
\end{aligned}$$

Hence $t \mapsto \left|\tilde{X}_t\right|^2 - t$ is an $\mathbb{F}^B$-martingale and so, for any $t \geq 0$, $\langle \tilde{X} \rangle_t = t$. From the representation (3.8), we also have that

$$t \mapsto \langle \tilde{X} \rangle_t = \int_0^t \mathrm{Tr}(\tilde{\nu}_s \tilde{\nu}_s^\top) \, \mathrm{d}s.$$

Hence $\mathrm{Tr}(\tilde{\nu}_t\tilde{\nu}_t^\top) = 1$, for any $t \geq 0$, and so $\tilde{\nu} \in \mathcal{U}$.

We can now follow the same arguments as in the proof of Lemma 2.29 to complete the proof of Proposition 2.27. Since the process $\tilde{X}$ has the same law as X, we have

$$\mathbb{E}^0\left[\int_0^{\tau_\eta} f(\tilde{X}_s)\,\mathrm{d}s\right] = \mathbb{E}^0\left[\int_0^{\tau_\eta} f(X_s)\,\mathrm{d}s\right] = 2\int_0^{\eta} \xi\tilde{f}(\xi)\,\mathrm{d}\xi.$$

Therefore, we can make a similar calculation as in the proof of Lemma 2.29 using the dynamic programming principle to see that

$$\begin{aligned} v^S(0) &\leq \mathbb{E}^0\left[\int_0^{\tau_\eta} f(\tilde{X}_s)\,\mathrm{d}s + v^S(\tilde{X}_{\tau_\eta})\right] \\ &= 2\int_0^{\eta} \xi\tilde{f}(\xi)\,\mathrm{d}\xi + V(y), \end{aligned}$$

for any $y \in D$ such that $|y| = \eta$.

Using (2.31), we also have that, for any $y \in D$ with $|y| = \eta$,

$$v^S(0) \geq v^W(0) \geq V(0) = 2\int_0^{\eta} \xi\tilde{f}(\xi)\,\mathrm{d}\xi + V(y).$$

We conclude that

$$v^S(0) = v^W(0) = V(0).$$

□

We have now shown that the strong and weak formulations of the control problem are equivalent. In the remainder of this chapter, we will prove two results that support Conjecture 3.5, which asserts that there is a gap between the strong and Markov value functions.

In the next section, we will show that the SDE (3.2) has no strong solution, and so the Brownian motion that generates the natural filtration of a weak solution cannot be the driving Brownian motion of the SDE.

3.6 Non-existence of strong solutions

The proof of non-existence of a strong solution in Theorem 3.4 will rely on the following property of the angle process that arises from the theory of circular Brownian motion discussed in Section 3.3.

Lemma 3.14. *Let W be a real-valued Brownian motion with natural filtration*

$(\mathcal{F}^W_t)_{t\geq 0}$ and let ϕ be an $\mathbb{R}/2\pi\mathbb{Z}$-valued process. Suppose that ϕ satisfies

$$\int_s^t \mathrm{d}\phi_r = \int_s^t r^{-\frac{1}{2}}\,\mathrm{d}W_r, \quad \textit{for all} \quad 0 < s \leq t,$$

where the random variables on the left-hand side are defined analogously to those in Definition 3.6.

Then ϕ cannot be adapted to $(\mathcal{F}^W_t)_{t\geq 0}$.

Proof. Suppose for contradiction that ϕ is adapted to the natural filtration of W.

Define the regular time-change $a : \mathbb{R} \to (0, \infty)$ by $a(t) = e^t$ for all $t \in \mathbb{R}$, as in the proof of Proposition 3.8, and denote the time-changed processes

$$(\tilde{\phi}_t)_{t\in\mathbb{R}} = (\phi_{a(t)})_{t>0},$$
$$(\tilde{W}_t)_{t\in\mathbb{R}} = (W_{a(t)})_{t>0}.$$

Since the time-change is deterministic, the natural filtrations $(\tilde{\mathcal{F}}^\phi_t)_{t\in\mathbb{R}}$ and $(\tilde{\mathcal{F}}^W_t)_{t\in\mathbb{R}}$ of the time-changed processes $\tilde{\phi}$ and $\tilde{W}$ are given by

$$\tilde{\mathcal{F}}^\phi_t = \mathcal{F}^\phi_{a(t)}, \quad \tilde{\mathcal{F}}^W_t = \mathcal{F}^W_{a(t)}, \quad \text{for all} \quad t \in \mathbb{R}.$$

Hence $\tilde{\phi}$ is adapted to $(\tilde{\mathcal{F}}^W_t)_{t\in\mathbb{R}}$.

By the same arugments as in the proof of Proposition 3.8, $\tilde{\phi}$ is a circular Brownian motion for $(\tilde{\mathcal{F}}^W_t)_{t\in\mathbb{R}}$ and, for any $s, t \in \mathbb{R}$ with $s \leq t$,

$$\int_s^t \mathrm{d}\tilde{\phi}_r = \int_s^t a(r)^{-\frac{1}{2}}\,\mathrm{d}\tilde{W}_r. \tag{3.9}$$

To arrive at a contradiction, we will exploit the following property of circular Brownian motion proved in Proposition 1 of [21].

Let $(\mathcal{H}_t)_{t\in\mathbb{R}}$ be the innovation filtration of $\tilde{\phi}$. Recall that, for each $t \in \mathbb{R}$, $\mathcal{H}_t$ is the sigma-algebra generated by the increments of $\tilde{\phi}$ up to time t; i.e.

$$\mathcal{H}_t := \sigma\left(\left\{\tilde{\phi}_s - \tilde{\phi}_r \colon -\infty < r \leq s \leq t\right\}\right).$$

Then we have

$$\mathcal{H}_t \subseteq \tilde{\mathcal{F}}^\phi_t \subseteq \tilde{\mathcal{F}}^W_t, \quad t \in \mathbb{R}.$$

In fact, the first inclusion must be strict, as we now show. As remarked in Section 3.3, Proposition 1 of [21] tells us that, for each $t \in \mathbb{R}$, the value of the circular Brownian motion $\tilde{\phi}_t$ is uniformly distributed on $[0, 2\pi)$ and, moreover, $\tilde{\phi}_t$ is independent of $\mathcal{H}_\infty$. Hence, for each $t \in \mathbb{R}$,

$$\mathcal{H}_t \subsetneq \tilde{\mathcal{F}}_t^\phi \subseteq \tilde{\mathcal{F}}_t^W. \tag{3.10}$$

Fix $0 < s \leq t$. Then, using the relation (3.9), we can deduce that the increment

$$\tilde{W}_t - \tilde{W}_s = \int_s^t a(r)^{\frac{1}{2}} \, \mathrm{d}\tilde{\phi}_r$$

is $\mathcal{H}_t$-measurable.

Now, taking the limit as $s \to -\infty$, $\tilde{W}_s = W_{e^s} \to 0$ almost surely, and so $\tilde{W}_t$ is $\mathcal{H}_t$-measurable. This implies that

$$\tilde{\mathcal{F}}_t^W \subseteq \mathcal{H}_t,$$

contradicting the strict inclusion in (3.10). $\square$

We are now ready to prove Theorem 3.4, showing that the SDE (3.2) has no strong solution starting from the origin.

Proof of Theorem 3.4. As noted after the statement of the theorem, the existence of a weak solution is proved by Larsson and Ruf in Theorem 4.3 of [41]. We now prove non-existence of strong solutions.

Suppose for contradiction that there exists a strong solution X of the SDE (3.2) with initial condition $X_0 = 0$. Then X is adapted to the filtration $(\mathcal{F}_t^B)_{t\geq 0}$; i.e

$$\mathcal{F}_t^X \subseteq \mathcal{F}_t^B, \quad t \geq 0. \tag{3.11}$$

Let θ be the angle process of X. Then, since θ satisfies (3.7), we have

$$\int_s^t \mathrm{d}\theta_r = \int_s^t r^{-\frac{1}{2}} \, \mathrm{d}B_r,$$

for any $0 < s \leq t$. Therefore, by Lemma 3.14, θ is not adapted to $(\mathcal{F}_t^B)_{t\geq 0}$.

We have already seen in the proof of Corollary 3.10 that

$$\mathcal{F}_t^\theta = \mathcal{F}_t^X, \quad \text{for all} \quad t \geq 0.$$

Therefore X is not adapted to $(\mathcal{F}_t^B)_{t\geq 0}$. This contradicts the inclusion (3.11). Hence the SDE (3.2) has no strong solution starting from the origin. $\square$

Having proved Theorem 3.4, we note that this supports Conjecture 3.5, as it rules out one way of constructing a Markov control that could be used to adapt the proof of Lemma 2.29. In Section 3.8, we will derive the form of other SDEs whose solutions have deterministically increasing radius and could therefore play the same role as solutions of (3.2). Proving that these SDEs also have no strong solution is a necessary step in proving the assertion of Conjecture 3.5.

Furthermore, we need to check whether there exist strong solutions of SDEs that approximate the desired behaviour at the origin. The result in the following section provides a partial negative answer to this question.

3.7 Additional SDEs without strong solutions

In order to further support Conjecture 3.5, we will show that the following SDEs have no strong solution.

Theorem 3.15. *Let B be a one-dimensional Brownian motion and let $\lambda \in (0, \frac{\sqrt{2}}{2})$ be a fixed constant. Then there is no strong solution of the SDE*

$$\mathrm{d}X_t = \frac{1}{|X_t|}\begin{bmatrix}\lambda X_t^1 - \sqrt{1-\lambda^2}X_t^2 \\ \lambda X_t^2 + \sqrt{1-\lambda^2}X_t^1\end{bmatrix}\mathrm{d}B_t; \quad X_0 = 0. \tag{3.12}$$

Before proving this result, we explain the relationship to Conjecture 3.5. As noted in Remark 2.25, the approximation used in Proposition 2.24 is not valid under the assumptions of Conjecture 3.5, since $\int_0^r \tilde{f}(s)\,\mathrm{d}s = \infty$ for any $r > 0$. However, we will show that, in the regime $\lambda \in (0, \frac{\sqrt{2}}{2})$, the expected cost associated to the process X^λ is finite and approximates the value $V(0)$ in the limit $\lambda \to 0$. We will work with the radius process to calculate this expected cost.

We first observe that the squared radius process can be rescaled to a squared Bessel process, as defined in Definition 1.1 of [51, Chapter XI]. We will show that the event of returning to the origin before leaving the domain satisfies the following zero-one law. For $\lambda \leq \frac{\sqrt{2}}{2}$, X^λ returns to the origin with probability zero; for $\lambda > \frac{\sqrt{2}}{2}$, X^λ returns to the origin with probability one. The critical value $\lambda = \frac{\sqrt{2}}{2}$ corresponds to the 2-dimensional squared Bessel process.

Proposition 3.16. *Let $\lambda \in (0,1)$ and suppose that X^λ solves the SDE (3.12). Write $Z_t^\lambda = \left|X_t^\lambda\right|^2$ for any $t \geq 0$ and define the rescaled process $\tilde{Z}^\lambda$ by $\tilde{Z}_t^\lambda = Z_{\lambda^{-2}t}$.*

Then $\tilde{Z}^\lambda$ is the square of a δ-dimensional Bessel process started from 0, where $\delta = \lambda^{-2}$.

Moreover, defining $\tau_0^\lambda := \inf\{t > 0\colon Z_t^\lambda = 0\}$, *we have*

$$\mathbb{P}^0\left[\tau_0 < \infty\right] = \begin{cases} 0, & \lambda \in (0, \frac{\sqrt{2}}{2}], \\ 1, & \lambda \in (\frac{\sqrt{2}}{2}, 1). \end{cases}$$

Proof. Applying Itô's formula, as in Lemma 2.2, we see that Z^λ satisfies

$$\mathrm{d}Z_t^\lambda = 2\lambda\sqrt{Z^\lambda}_t\,\mathrm{d}B_t + \mathrm{d}t,$$

with $Z_0^\lambda = 0$. Note that

$$t \mapsto \tilde{B}_t := \lambda B_{\lambda^{-2}t}$$

is a standard Brownian motion. Therefore, for any $t \geq 0$,

$$\tilde{Z}_t^\lambda = 2\int_0^t \sqrt{\tilde{Z}_s^\lambda}\,\mathrm{d}\tilde{B}_s + \lambda^{-2}t.$$

Set $\delta = \lambda^{-2}$. Then, referring to Definition 1.1 of [51, Chapter XI], we see that $\tilde{Z}^\lambda$ is the square of a δ-dimensional Bessel process.

Now suppose that $\lambda \in (0, \frac{\sqrt{2}}{2}]$, so that

$$\delta = \lambda^{-2} \geq 2.$$

The discussion in [51, Chapter XI] that immediately precedes Proposition 1.5 tells us that the set $\{0\}$ is *polar* for $\tilde{Z}^\lambda$. By the definition of a polar set given in Definition 2.6 of [51, Chapter V], we have that $\tilde{Z}^\lambda$ almost surely never returns to the origin in finite time, and the rescaled process Z^λ has the same property.

On the other hand, suppose that $\lambda \in (\frac{\sqrt{2}}{2}, 1)$. Then

$$\delta = \lambda^{-2} < 2,$$

and so, by the same discussion in [51, Chapter XI], $\tilde{Z}^\lambda$ returns to the origin in finite time with probability 1. Again the rescaled process Z^λ has the same property. □

Remark 3.17. Let $\lambda \in (0, \frac{\sqrt{2}}{2}]$ and Define the process R^λ by $R_t^\lambda = \left|X_t^\lambda\right|$, for $t \geq 0$. By Proposition 3.16, Z^λ almost surely never returns to the origin after time 0. Therefore, we can apply Itô's formula to $R_t^\lambda = \sqrt{Z_t^\lambda}$ to calculate that R^λ satisfies

$$\mathrm{d}R_t^\lambda = \lambda\,\mathrm{d}B_t + \frac{1-\lambda^2}{2R_t^\lambda}\,\mathrm{d}t; \quad R_0^\lambda = 0. \tag{3.13}$$

We note that the SDE (3.13) has a unique strong solution after time zero. This can be proved by adapting the standard proof of existence and uniqueness for SDEs whose coefficients are Lipschitz and have linear growth, as found for example in Theorem 2.9 of [38, Chapter 5].

We now return to calculating the expected cost in the regime $\lambda \in (0, \frac{\sqrt{2}}{2})$.

Proposition 3.18. *Suppose that the growth condition* (3.3) *holds and let* $\eta > 0$ *be such that* $\tilde{f}$ *is decreasing and positive on the interval* $(0, \eta)$.

For $\lambda \in (0, \frac{\sqrt{2}}{2})$, *if* X^λ *is a strong solution of the SDE* (3.12), *then*

$$0 \leq \mathbb{E}^0 \left[\int_0^{\tau_\eta} f(X_s^\lambda) \, \mathrm{d}s \right] < \infty,$$

and, moreover,

$$\lim_{\lambda \downarrow 0} \mathbb{E}^0 \left[\int_0^{\tau_\eta} f(X_s^\lambda) \, \mathrm{d}s \right] = 2 \int_0^\eta \xi \tilde{f}(\xi) \, \mathrm{d}\xi.$$

Proof. Fix $r, \varepsilon \in (0, \eta)$ such that $0 < \varepsilon < r < \eta$. We will calculate $\mathbb{E}^r \left[\int_0^{\tau_\varepsilon \wedge \tau_\eta} f(X_s^\lambda) \, \mathrm{d}s \right]$ and show later that

$$\mathbb{E}^0 \left[\int_0^{\tau_\eta} f(X_t^\lambda) \, \mathrm{d}t \right] = \lim_{r \to 0} \lim_{\varepsilon \to 0} \mathbb{E}^r \left[\int_0^{\tau_\varepsilon \wedge \tau_\eta} f(X_t^\lambda) \, \mathrm{d}t \right].$$

Let m be the speed measure of R^λ and G the Green's function, as defined in Definition B.3 and Definition B.4, respectively. Then, by Proposition B.5,

$$\mathbb{E}^r \left[\int_0^{\tau_\varepsilon \wedge \tau_\eta} f(X_s^\lambda) \, \mathrm{d}s \right] = \mathbb{E}^r \left[\int_0^{\tau_\varepsilon \wedge \tau_\eta} \tilde{f}(R_s^\lambda) \, \mathrm{d}s \right] = \int_\varepsilon^\eta \tilde{f}(\xi) G(r, \xi) m(\mathrm{d}\xi).$$

The speed measure m is given by

$$\int m(\mathrm{d}\xi) = \int \frac{2c^{1-\lambda^{-1}}}{\lambda^2 \xi^{1-\lambda^{-2}}} \, \mathrm{d}\xi,$$

for some constant c, and the Green's function $G : (\varepsilon, \eta)^2 \to \mathbb{R}$ is given by

$$G(r, \xi) = \begin{cases} \frac{c^{-(1-\lambda^{-1})}\lambda^2}{2\lambda^2 - 1} \frac{\left(\xi^{2-\lambda^{-2}} - \varepsilon^{2-\lambda^{-2}}\right)\left(\eta^{2-\lambda^{-2}} - r^{2-\lambda^{-2}}\right)}{\eta^{2-\lambda^{-2}} - \varepsilon^{2-\lambda^{-2}}}, & \xi \in [\varepsilon, r], \\ \frac{c^{-(1-\lambda^{-1})}\lambda^2}{2\lambda^2 - 1} \frac{\left(r^{2-\lambda^{-2}} - \varepsilon^{2-\lambda^{-2}}\right)\left(\eta^{2-\lambda^{-2}} - \xi^{2-\lambda^{-2}}\right)}{\eta^{2-\lambda^{-2}} - \varepsilon^{2-\lambda^{-2}}}, & \xi \in [r, \eta]. \end{cases}$$

Therefore, we calculate that

$$
\begin{aligned}
&\mathbb{E}^r\left[\int_0^{\tau_\varepsilon\wedge\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right] \\
&= \frac{2}{2\lambda^2-1}\frac{\eta^{2-\lambda^{-2}}-r^{2-\lambda^{-2}}}{\eta^{2-\lambda^{-2}}-\varepsilon^{2-\lambda^{-2}}}\left(\int_\varepsilon^r \xi\tilde{f}(\xi)\,\mathrm{d}\xi-\varepsilon^{2-\lambda^{-2}}\int_\varepsilon^r \xi^{\lambda^{-2}-1}\tilde{f}(\xi)\,\mathrm{d}\xi\right) \qquad (3.14)\\
&\quad+\frac{2}{2\lambda^2-1}\frac{r^{2-\lambda^{-2}}-\varepsilon^{2-\lambda^{-2}}}{\eta^{2-\lambda^{-2}}-\varepsilon^{2-\lambda^{-2}}}\left(\eta^{2-\lambda^{-2}}\int_r^\eta \xi^{\lambda^{-2}-1}\tilde{f}(\xi)\,\mathrm{d}\xi-\int_r^\eta \xi\tilde{f}(\xi)\,\mathrm{d}\xi\right).
\end{aligned}
$$

We now take the limit as $\varepsilon\to 0$. The limits

$$
\lim_{\varepsilon\downarrow 0}\int_\varepsilon^r \xi\tilde{f}(\xi)\,\mathrm{d}\xi \quad\text{and}\quad \lim_{\varepsilon\downarrow 0}\int_\varepsilon^r \xi^{\lambda^{-2}-1}\tilde{f}(\xi)\,\mathrm{d}\xi
$$

are finite under the growth condition (3.3), since $\lambda^{-2}-1>1$. Also, since $2-\lambda^{-2}<0$, we have

$$
\lim_{\varepsilon\downarrow 0}\frac{\eta^{2-\lambda^{-2}}-r^{2-\lambda^{-2}}}{\eta^{2-\lambda^{-2}}-\varepsilon^{2-\lambda^{-2}}}=0,
$$

and

$$
\lim_{\varepsilon\downarrow 0}\frac{\eta^{2-\lambda^{-2}}-r^{2-\lambda^{-2}}}{\eta^{2-\lambda^{-2}}-\varepsilon^{2-\lambda^{-2}}}\varepsilon^{2-\lambda^{-2}}=\lim_{\varepsilon\downarrow 0}\frac{\eta^{2-\lambda^{-2}}-r^{2-\lambda^{-2}}}{\left(\frac{\eta}{\varepsilon}\right)^{2-\lambda^{-2}}-1}=r^{2-\lambda^{-2}}-\eta^{2-\lambda^{-2}}.
$$

Therefore the limit of the first term in (3.14) is

$$
\begin{aligned}
&\lim_{\varepsilon\downarrow 0}\frac{2}{2\lambda^2-1}\frac{\eta^{2-\lambda^{-2}}-r^{2-\lambda^{-2}}}{\eta^{2-\lambda^{-2}}-\varepsilon^{2-\lambda^{-2}}}\left(\int_\varepsilon^r \xi\tilde{f}(\xi)\,\mathrm{d}\xi-\varepsilon^{2-\lambda^{-2}}\int_\varepsilon^r \xi^{\lambda^{-2}-1}\tilde{f}(\xi)\,\mathrm{d}\xi\right)\\
&\quad=\frac{2}{2\lambda^2-1}\left(\eta^{2-\lambda^{-2}}-r^{2-\lambda^{-2}}\right)\lim_{\varepsilon\downarrow 0}\int_\varepsilon^r \xi^{\lambda^{-2}-1}\tilde{f}(\xi)\,\mathrm{d}\xi.
\end{aligned}
$$

Under the growth condition (3.3), noting that $\beta^\star\in[1,2)$ and so $\lambda^{-2}-\beta^\star>0$, we find that there is a constant $C>0$ such that

$$
\begin{aligned}
\int_\varepsilon^r \xi^{\lambda^{-2}-1}\tilde{f}(\xi)\,\mathrm{d}\xi &\le C\int_\varepsilon^r \xi^{\lambda^{-2}-1-\beta^\star}\,\mathrm{d}\xi\\
&=\frac{C}{\lambda^{-2}-\beta^\star}\left(r^{\lambda^{-2}-\beta^\star}-\varepsilon^{\lambda^{-2}-\beta^\star}\right)\\
&\xrightarrow{\varepsilon\downarrow 0}\frac{C}{\lambda^{-2}-\beta^\star}r^{\lambda^{-2}-\beta^\star}.
\end{aligned}
$$

Considering the second term in (3.14), we have

$$
\lim_{\varepsilon\downarrow 0}\frac{r^{2-\lambda^{-2}}-\varepsilon^{2-\lambda^{-2}}}{\eta^{2-\lambda^{-2}}-\varepsilon^{2-\lambda^{-2}}}=1,
$$

and so

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}^r\left[\int_0^{\tau_\varepsilon\wedge\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right] \le \frac{2}{2\lambda^2-1}\frac{C}{\lambda^{-2}-\beta^\star}\left(\eta^{2-\lambda^{-2}}r^{\lambda^{-2}-\beta^\star}-r^{2-\beta^\star}\right) + \frac{2}{2\lambda^2-1}\left(\eta^{2-\lambda^{-2}}\int_r^\eta \xi^{\lambda^{-2}-1}\tilde f(\xi)\,\mathrm{d}\xi - \int_r^\eta \xi\tilde f(\xi)\,\mathrm{d}\xi\right).$$

Note that $\lambda^{-2}-\beta^\star>0$ and $2-\beta^\star>0$, so the first term in the sum vanishes in the limit $r\downarrow 0$.

Again, using the growth condition (3.3),

$$\lim_{r\downarrow 0}\int_r^\eta \xi\tilde f(\xi)\,\mathrm{d}\xi = \int_0^\eta \xi\tilde f(\xi)\,\mathrm{d}\xi \in(-\infty,\infty),$$

and, for some constant $C'>0$,

$$\begin{aligned}\int_r^\eta \xi^{\lambda^{-2}-1}\tilde f(\xi)\,\mathrm{d}\xi &\le C'\int_r^\eta \xi^{\lambda^{-2}-1-\beta^\star}\,\mathrm{d}\xi\\ &= \frac{C'}{\lambda^{-2}-\beta^\star}\left(\eta^{\lambda^{-2}-\beta^\star}-r^{\lambda^{-2}-\beta^\star}\right)\\ &\xrightarrow{r\downarrow 0}\frac{C'}{\lambda^{-2}-\beta^\star}\eta^{\lambda^{-2}-\beta^\star}.\end{aligned}$$

Therefore

$$\lim_{r\downarrow 0}\lim_{\varepsilon\downarrow 0}\mathbb{E}^r\left[\int_0^{\tau_\varepsilon\wedge\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right] \le \frac{2}{1-2\lambda^2}\left(\int_0^\eta \xi\tilde f(\xi)\,\mathrm{d}\xi - \frac{C'}{\lambda^{-2}-\beta^\star}\eta^{2-\beta^\star}\right). \quad (3.15)$$

We now prove that

$$\mathbb{E}^0\left[\int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right] = \lim_{r\downarrow 0}\lim_{\varepsilon\downarrow 0}\mathbb{E}^r\left[\int_0^{\tau_\varepsilon\wedge\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right].$$

First we consider the limit $\varepsilon\downarrow 0$. We showed in Proposition 3.16 that $\mathbb{P}^r\left[\tau_0<\tau_\eta\right]=0$. Therefore, with probability one,

$$\int_0^{\tau_0\wedge\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s = \int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s,$$

and so

$$\lim_{\varepsilon\downarrow 0}\int_0^{\tau_\varepsilon\wedge\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s = \int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s, \quad \text{almost surely.}$$

Since $f\ge 0$ in the ball $B_\eta(0)$, the integral inside the limit is monotone increasing as ε decreases to zero. Therefore we can apply the monotone convergence theorem to see that

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}^r\left[\int_0^{\tau_\varepsilon\wedge\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right]=\mathbb{E}^r\left[\int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right].$$

We now wish to take the limit as $r\downarrow 0$. By the Markov property, we have

$$\mathbb{E}^r\left[\int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right]=\mathbb{E}^0\left[\int_{\tau_r}^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right].$$

We also have that

$$\int_{\tau_r}^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s \xrightarrow{r\to 0} \int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s \quad \text{a.s.},$$

where the convergence is monotone, since f is positive in $B_\eta(0)$. Therefore we can apply the monotone convergence theorem again to get

$$\lim_{r\downarrow 0}\mathbb{E}^r\left[\int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right]=\mathbb{E}^0\left[\int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right].$$

Now, for any $\lambda\in(0,\frac{\sqrt{2}}{2})$, the bound in (3.15) gives us

$$0\le\mathbb{E}^0\left[\int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right]\le\frac{2}{1-2\lambda^2}\int_0^\eta \xi\tilde{f}(\xi)\,\mathrm{d}\xi-\frac{2}{1-2\lambda^2}\frac{C'}{\lambda^{-2}-\beta^\star}\eta^{2-\beta^\star}<\infty.$$

Taking the limit as $\lambda\downarrow 0$, we have

$$\lim_{\lambda\downarrow 0}\mathbb{E}^0\left[\int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right]\le 2\int_0^\eta \xi\tilde{f}(\xi)\,\mathrm{d}\xi.$$

To conclude the proof, we refer to Lemma 2.29 to see that we have the reverse inequality

$$\mathbb{E}^0\left[\int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right]\ge\inf_{\sigma\in\mathcal{U}}\mathbb{E}^0\left[\int_0^{\tau_\eta} f(X_s^\sigma)\,\mathrm{d}s\right]\ge 2\int_0^\eta \xi\tilde{f}(\xi)\,\mathrm{d}\xi,$$

for any $\lambda\in(0,\frac{\sqrt{2}}{2})$.

Hence

$$\lim_{\lambda\downarrow 0}\mathbb{E}^0\left[\int_0^{\tau_\eta} f(X_s^\lambda)\,\mathrm{d}s\right]= 2\int_0^\eta \xi\tilde{f}(\xi)\,\mathrm{d}\xi. \qquad\square$$

Proposition 3.18 shows that, if there exist strong solutions of the SDE (3.12) for all $\lambda\in(0,\frac{\sqrt{2}}{2})$, then we can take a sequence of such processes for values of λ approaching 0 such that the associated expected costs approximate $V(0)$. This would disprove Conjecture 3.5, since any strong solution of (3.12) gives rise to a Markov control.

We now turn to the proof of Theorem 3.15, where we show that such strong solutions do not exist. Here, the angle process of a solution of (3.12) is no longer a circular Brownian motion, as was the case for solutions of (3.2) in Proposition 3.8, but this process does have similar properties. We will show that, conditioned on the value of the radius, the angle process is uniformly distributed and independent of its increments. Here, we adapt Émery and Schachermayer's proof that the value of a circular Brownian motion at any time is uniformly distributed and independent of its increments from Proposition 1 of [21]. We will deduce the result of Theorem 3.15 from the following proposition.

Proposition 3.19. *Fix* $\lambda \in (0, \frac{\sqrt{2}}{2})$. *For any weak solution* $((X, W), (\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F})$ *of the SDE* (3.12)*, let* R *be the radius process and* θ *the angle process, so that we can write*

$$X_t = R_t \begin{bmatrix} \cos\theta_t \\ \sin\theta_t \end{bmatrix}, \quad t > 0.$$

Denote the hitting times of R *by*

$$\tau_\rho := \inf\{t > 0 \colon R_t = \rho\}.$$

Then, for any $\rho > 0$*,*

$$\theta_{\tau_\rho} \sim \mathrm{Unif}[0, 2\pi).$$

Moreover, θ_{τ_ρ} *is independent of*

$$\mathcal{H}_\infty := \sigma\left(\{\theta_t - \theta_s \colon 0 < s < t < \infty\}\right).$$

To prove this result, we will require the following technical lemma on the distribution of increments of the angle process.

Lemma 3.20. *Let* θ *be the angle process defined in Proposition 3.19 and fix* $\rho > 0$. *Then, for any* $\phi \in [0, 2\pi)$*,*

$$\mathbb{P}\left[\left(\theta_{\tau_\rho} - \theta_{\tau_{2^{-1}\rho}}\right) \in \{\phi + 2\pi m, \quad m \in \mathbb{Z}\}\right] < 1.$$

Proof. Suppose for contradiction that there exists $\phi \in [0, 2\pi)$ such that

$$\mathbb{P}\left[\left(\theta_{\tau_\rho} - \theta_{\tau_{2^{-1}\rho}}\right) \in \{\phi + 2\pi m, \quad m \in \mathbb{Z}\}\right] = 1. \tag{3.16}$$

Let R be the radius process defined in Proposition 3.19 and recall that R satisfies the SDE

$$\mathrm{d}R_t = \lambda\,\mathrm{d}W_t + \frac{1-\lambda^2}{2R_t}\,\mathrm{d}t; \quad R_0 = 0. \tag{3.13}$$

By Itô's formula, we calculate that θ satisfies

$$\mathrm{d}\theta_t = \sqrt{1-\lambda^2} R_t^{-1}\, \mathrm{d}W_t - \lambda\sqrt{1-\lambda^2} R_t^{-2}\, \mathrm{d}t. \tag{3.17}$$

We will use a coupling argument to arrive at a contradiction. Consider two independent weak solutions (R^1, θ^1), (R^2, θ^2) of the SDEs (3.13) and (3.17) on a common probability space. For $i = 1, 2$ and any $r \geq 0$, denote the hitting time

$$\tau_r^i := \inf\{t > 0\colon R_t^i = r\}.$$

Note that, as we observed in Remark 3.17, the SDE (3.13) for R has a unique strong solution after the first hitting time of radius $2^{-1}\rho$. Therefore, given the value of θ at this radius, the process θ is uniquely defined via the SDE (3.17). Hence all weak solutions of the SDEs (3.13) and (3.17) must have the same law after the first hitting time of radius $2^{-1}\rho$.

Fix $\psi^1, \psi^2 \in [0, 2\pi)$ such that $\psi^1 \not\equiv \psi^2 \mod 2\pi$, and shift θ^1 and θ^2 to define

$$\theta_t^{\psi^1} := \theta_t^1 + \psi^1 - \theta^1_{\tau^1_{2^{-1}\rho}} \quad \text{and} \quad \theta_t^{\psi^2} := \theta_t^2 + \psi^2 - \theta^2_{\tau^2_{2^{-1}\rho}}.$$

Then, at the first hitting time of radius $2^{-1}\rho$, the values of the processes θ^{ψ^1} and θ^{ψ^2} are almost surely equal to ψ^1 and ψ^2, respectively, and these shifted processes still satisfy the SDE (3.17).

Suppose that there exists some radius $\eta \in (2^{-1}\rho, \rho)$ such that

$$\theta^{\psi^1}_{\tau^1_\eta} = \theta^{\psi^2}_{\tau^2_\eta}.$$

Then we can couple the two processes as follows. Define $\tilde{\theta}$ by

$$\tilde{\theta}_t = \begin{cases} \theta_t^{\psi^2}, & t < \tau_\eta^2, \\ \theta^{\psi^1}_{\tau^1_\eta - \tau^2_\eta + t}, & t \geq \tau^2_\eta. \end{cases}$$

Then we see that the trajectories of (R^1, θ^{ψ^1}) and $(R^2, \tilde{\theta})$ coincide on the set $(\eta, R) \times [0, 2\pi)$. Moreover, by the Markov property, the process $\tilde{\theta}$ still satisfies the SDE (3.17). Therefore, by condition (3.16),

$$\begin{aligned} \theta^{\psi^1}_{\tau_\rho} &\equiv \psi^1 + \phi \mod 2\pi, \\ \tilde{\theta}_{\tau_\rho} &\equiv \psi^2 + \phi \mod 2\pi. \end{aligned}$$

But, by our choice of ψ^1, ψ^2, the above values are not equal, contradicting the coupling of the trajectories. This shows that, on the set $(2^{-1}\rho, \rho) \times [0, 2\pi)$, the supports of (R^1, θ^{ψ^1}) and (R^2, θ^{ψ^2}) must be disjoint.

Since our choice of the shifts ψ^1 and ψ^2 was arbitrary, the only feasible supports of (R^i, θ^{ψ^i}) are the rays connecting the points $(2^{-1}\rho, \psi^i)$ and (ρ, ψ^i), for $i = 1, 2$. This would imply that θ^{ψ^1} and θ^{ψ^2} are deterministic, but this is not the case for $\lambda < 1$.

Hence there is no $\phi \in [0, 2\pi)$ such that (3.16) holds. $\square$

We now use this lemma to prove Proposition 3.19 on the uniformity and independence properties of the angle process.

Proof of Proposition 3.19. Recall that R satisfies the SDE (3.13) and θ satisfies the SDE (3.17).

Fix $\rho > 0$. We show that θ_{τ_ρ} is uniformly distributed on $[0, 2\pi)$ by using the characteristic function of the random variable θ_{τ_ρ} on the torus, following the proof of Proposition 1 of [21]. For any $\phi \in \mathbb{R}/2\pi\mathbb{Z}$ and $k \in \mathbb{Z}$, define the characteristic function

$$e_k(\phi) := \exp\{iky\}, \quad \text{for any} \quad y \in \mathbb{R} \quad \text{such that} \quad y \equiv \phi \mod 2\pi.$$

Fix $k \in \mathbb{Z} \setminus \{0\}$ and $\rho_1 > 0$. We aim to show that $\mathbb{E}[e_k(\theta_{\tau_{\rho_1}})] = 0$.

Let $\rho_0 \in (0, \rho_1)$. Then, writing

$$\theta_{\tau_{\rho_1}} = \theta_{\tau_{\rho_0}} + \int_{\tau_{\rho_0}}^{\tau_{\rho_1}} \mathrm{d}\theta_s,$$

we have

$$\left|\mathbb{E}^{\mathbb{P}}\left[e_k(\theta_{\tau_{\rho_1}})\right]\right| = \left|\mathbb{E}^{\mathbb{P}}\left[e_k(\theta_{\tau_{\rho_0}})e_k(\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_0}})\right]\right|.$$

In order to break up the expectation on the right hand side into the product of expectations, we use the following conditional independence. We see that future increments of θ depend only on the history of θ through the current value of R, since R is Markovian. That is, for any $s < u < v$,

$$\theta_v - \theta_u \quad \text{conditioned on} \quad \sigma(R_s) \quad \text{is independent of} \quad \mathcal{F}_s^\theta.$$

Now note that, taking $s = \tau_{\rho_0}$, the σ-algebra $\sigma(R_{\tau_{\rho_o}})$ is trivial, and so future increments of θ are independent of $\mathcal{F}^\theta_{\tau_{\rho_1}}$, without any conditioning. Hence

$$\mathbb{E}^{\mathbb{P}}\left[e_k(\theta_{\tau_{\rho_0}})e_k(\theta_{\tau_{\rho_1}}-\theta_{\tau_{\rho_0}})\right]=\mathbb{E}^{\mathbb{P}}\left[e_k(\theta_{\tau_{\rho_0}})\right]\mathbb{E}^{\mathbb{P}}\left[e_k(\theta_{\tau_{\rho_1}}-\theta_{\tau_{\rho_0}})\right]. \tag{3.18}$$

We will now consider the increment $\theta_{\tau_{\rho_1}}-\theta_{\tau_{\rho_0}}$. We claim that, for small radii ρ_0, the value of this increment approaches a uniform distribution on $[0,2\pi)$. We show this by using a scaling argument, as follows.

Fix $\alpha>0$ and rescale time by defining $s:=\alpha t$ for $t\geq 0$. Then, for $t\geq 0$, define

$$\tilde{B}^\alpha_s:=\alpha^{\frac{1}{2}}B_t,\quad \tilde{R}^\alpha_s:=\alpha^{\frac{1}{2}}R_t,\quad \tilde{\theta}^\alpha_s:=\theta_t,$$

so that

$$\mathrm{d}s=\alpha\,\mathrm{d}t,\quad\text{and}\quad \mathrm{d}\tilde{B}^\alpha_s=\alpha^{\frac{1}{2}}\,\mathrm{d}B_t.$$

We can calculate

$$\begin{aligned}\mathrm{d}\tilde{R}^\alpha_s&=\alpha^{\frac{1}{2}}\left(\lambda\,\mathrm{d}B_t+\frac{1-\lambda^2}{2R_t}\,\mathrm{d}t\right)\\&=\alpha^{\frac{1}{2}}\left(\lambda\alpha^{-\frac{1}{2}}\,\mathrm{d}\tilde{B}^\alpha_s+\frac{1-\lambda^2}{2\alpha^{-\frac{1}{2}}\tilde{R}^\alpha_s}\alpha^{-1}\,\mathrm{d}s\right)\\&=\lambda\,\mathrm{d}\tilde{B}^\alpha_s+\frac{1-\lambda^2}{2\tilde{R}^\alpha_s}\,\mathrm{d}s,\end{aligned}$$

and

$$\begin{aligned}\mathrm{d}\tilde{\theta}^\alpha_s&=\sqrt{1-\lambda^2}R_t^{-1}\,\mathrm{d}B_t-\lambda\sqrt{1-\lambda^2}R_t^{-2}\,\mathrm{d}t\\&=\sqrt{1-\lambda^2}\left(\alpha^{-\frac{1}{2}}\tilde{R}^\alpha_s\right)^{-1}\alpha^{-\frac{1}{2}}\,\mathrm{d}\tilde{B}^\alpha_s-\lambda\sqrt{1-\lambda^2}\left(\alpha^{-\frac{1}{2}}\tilde{R}^\alpha_s\right)^{-2}\alpha^{-1}\,\mathrm{d}s\\&=\sqrt{1-\lambda^2}\left(\tilde{R}^\alpha_s\right)^{-1}\mathrm{d}\tilde{B}^\alpha_s-\lambda\sqrt{1-\lambda^2}\left(\tilde{R}^\alpha_s\right)^{-2}\mathrm{d}s.\end{aligned}$$

And so, after rescaling, $(\tilde{R}^\alpha,\tilde{B}^\alpha)$ and $(\tilde{\theta}^\alpha,\tilde{B}^\alpha)$ satisfy the same SDEs (3.13) and (3.17) as (R,B) and (θ,B).

For $i=0,1$, let $\tilde{\tau}^0_{\rho_i}$ be the first hitting time of ρ_i, by $\tilde{R}^\alpha_s$ starting from the origin. Then we have the following equality in distribution:

$$\theta_{\tau_{\rho_1}}-\theta_{\tau_{\rho_0}}=\tilde{\theta}^\alpha_{\tilde{\tau}^0_{\sqrt{\alpha}\rho_1}}-\tilde{\theta}^\alpha_{\tilde{\tau}^0_{\sqrt{\alpha}\rho_0}}=\theta_{\tau_{\sqrt{\alpha}\rho_1}}-\theta_{\tau_{\sqrt{\alpha}\rho_0}},$$

where the first equality holds pointwise by rescaling, and the second equality holds in distribution because the rescaled processes satisfy the same SDEs as the original processes.

Moreover, recalling our observation that increments of θ between hitting times of R are independent, we see that the increments

$$\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_0}} \quad \text{and} \quad \theta_{\tau_{\sqrt{\alpha}\rho_1}} - \theta_{\tau_{\sqrt{\alpha}\rho_0}}$$

are independent and identically distributed when $\sqrt{\alpha}\rho_1 \leq \rho_0$.

Now let $N \in \mathbb{N}$ and set $\rho_0 = 2^{-N}\rho_1$. We can write the increment of θ as a sum of i.i.d. random variables

$$\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_0}} = \sum_{k=0}^{N-1} \left(\theta_{\tau_{2^{-k}\rho_1}} - \theta_{\tau_{2^{-k+1}\rho_1}} \right),$$

and so

$$\left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_0}}\right)\right]\right| = \left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_1}} - \theta_{\tau_{2^{-1}\rho_1}}\right)\right]\right|^N.$$

By Jensen's inequality,

$$\left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_1}} - \theta_{\tau_{2^{-1}\rho_1}}\right)\right]\right|^2 \leq 1, \tag{3.19}$$

with equality if and only if there exists $\phi \in [0, 2\pi)$ such that

$$\mathbb{P}\left[\left(\theta_{\tau_{\rho_1}} - \theta_{\tau_{2^{-1}\rho_1}}\right) \in \{\phi + 2\pi m, \quad m \in \mathbb{Z}\}\right] = 1.$$

By Lemma 3.20, no such ϕ exists, and so the inequality (3.19) is strict. We then have that

$$\left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_0}}\right)\right]\right| = \left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_1}} - \theta_{\tau_{2^{-1}\rho_1}}\right)\right]\right|^N \xrightarrow{N\to\infty} 0,$$

Returning to our calculation of the characteristic function of θ_t in (3.18), we have

$$\begin{aligned}
\left|\mathbb{E}^{\mathbb{P}}\left[e_k(\theta_{\tau_{\rho_1}})\right]\right| &= \left|\mathbb{E}^{\mathbb{P}}\left[e_k(\theta_{\tau_{\rho_0}})\right]\mathbb{E}^{\mathbb{P}}\left[e_k(\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_0}})\right]\right| \\
&\leq \left|\mathbb{E}^{\mathbb{P}}\left[e_k(\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_0}})\right]\right| \\
&\xrightarrow{\rho_0\downarrow 0} 0.
\end{aligned}$$

Hence $\theta_{\tau_{\rho_1}}$ is uniformly distributed on $[0, 2\pi)$.

We now show that $\theta_{\tau_{\rho_1}}$ is independent of $\mathcal{H}_\infty$, the sigma-algebra generated by all increments of θ.

Let $(\rho_n)_{n\in\mathbb{N}}$ be a decreasing sequence such that $\rho_n > 0$, for all $n \in \mathbb{N}$, and

$\lim_{n\to\infty} \rho_n = 0$. For each $n \in \mathbb{N}$, define

$$\mathcal{H}^n := \sigma\left(\{\theta_v - \theta_u \colon \tau_{\rho_n} \leq u \leq v\}\right),$$

the sigma algebra generated by all increments of θ after the first hitting time of ρ_n.

Recalling that we are working with filtrations that satisfy the usual conditions, we have that $\mathcal{H}_\infty = \bigvee_{n\in\mathbb{N}} \mathcal{H}^n$, since $\tau_{\rho_n} \to 0$ almost surely as $n \to \infty$. Therefore, by martingale convergence (see e.g. Theorem 4.3 of [55, Chapter VII]),

$$\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_1}}\right)\middle|\mathcal{H}^n\right] \xrightarrow{n\to\infty} \mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_1}}\right)\middle|\mathcal{H}_\infty\right]$$

in $\mathcal{L}^1$ and almost surely.

We now fix $n \in \mathbb{N}$ and consider

$$\left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_1}}\right)\middle|\mathcal{H}^n\right]\right| = \left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_n}}\right) e_k\left(\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_n}}\right)\middle|\mathcal{H}^n\right]\right|.$$

By the same conditional independence arguments as we used in the proof of uniformity, $\theta_{\tau_{\rho_n}}$ is independent of $\mathcal{H}^n$. Since $\tau_{\rho_1} \geq \tau_{\rho_n}$ pointwise, $\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_n}}$ is $\mathcal{H}^n$-measurable. Therefore

$$\begin{aligned}\left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_n}}\right) e_k\left(\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_n}}\right)\middle|\mathcal{H}^n\right]\right| &= \left|e_k\left(\theta_{\tau_{\rho_1}} - \theta_{\tau_{\rho_n}}\right)\right| \left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_n}}\right)\right]\right| \\ &= \left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_n}}\right)\right]\right| \\ &= 0,\end{aligned}$$

by the uniformity of $\theta_{\tau_{\rho_n}}$.

Hence

$$\left|\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_1}}\right)\middle|\mathcal{H}^n\right]\right| = 0, \quad \text{for all} \quad n \in \mathbb{N},$$

and so, by martingale convergence,

$$\mathbb{E}^{\mathbb{P}}\left[e_k\left(\theta_{\tau_{\rho_1}}\right)\middle|\mathcal{H}_\infty\right] = 0.$$

Taking Y to be any bounded $\mathcal{H}_\infty$-measurable random variable, we then have

$$\mathbb{E}^{\mathbb{P}}\left[Y e_k\left(\theta_{\tau_{\rho_1}}\right)\right] = \mathbb{E}^{\mathbb{P}}\left[Y \mathbb{E}\left[e_k\left(\theta_{\tau_{\rho_1}}\right)\middle|\mathcal{H}_\infty\right]\right] = 0.$$

Hence $\theta_{\tau_{\rho_1}}$ is independent of $\mathcal{H}_\infty$. $\square$

We now apply the independence result in Proposition 3.19 to conclude that the SDE (3.12) has no strong solution.

Proof of Theorem 3.15. Suppose that X is a strong solution of the SDE (3.12). Then there is an $\mathbb{R}_+$-valued $\mathbb{F}^B$-adapted process R satisfying the SDE (3.13) with $R_0 = 0$, and an $\mathbb{R}/2\pi\mathbb{Z}$-valued $\mathbb{F}^B$-adapted process θ satisfying the SDE (3.17) such that

$$X_t = R_t \begin{bmatrix} \cos\theta_t \\ \sin\theta_t \end{bmatrix}, \quad t > 0.$$

Fix $\rho > 0$, and recall the definition

$$\tau_\rho := \inf\{t > 0 \colon R_t = \rho\}.$$

Then, by Proposition 3.19, θ_{τ_ρ} is independent of $\mathcal{H}_\infty$.

Under our assumption that θ is adapted to $\mathbb{F}^B$, this implies that

$$\mathcal{H}_{\tau_\rho} \subsetneq \mathcal{F}^\theta_{\tau_\rho} \subseteq \mathcal{F}^B_{\tau_\rho}. \tag{3.20}$$

However, we claim that B is adapted to $\mathcal{H}$.

To prove this claim, observe that, for any $0 < s < t$, the random variable

$$\langle\theta\rangle_t - \langle\theta\rangle_s = \int_s^t R_r^{-2}\,\mathrm{d}r$$

is $\mathcal{H}_t$-measurable. Since $R_r > 0$ almost surely for $r > 0$, as we proved in Proposition 3.16, R_t is also $\mathcal{H}_t$-measurable.

Now, from the SDE (3.13), we have that

$$R_t - R_s = \lambda(B_t - B_s) + \int_s^t \frac{1-\lambda^2}{2R_r}\,\mathrm{d}r,$$

and so $B_t - B_s$ is $\mathcal{F}^R_t$-measurable. Since $B_s \to 0$ as $s \to 0$, we can conclude that

$$\mathcal{F}^B_t \subseteq \mathcal{F}^R_t \subseteq \mathcal{H}_t. \tag{3.21}$$

Setting $t = \tau_\rho$ and combining the two inclusions (3.20) and (3.21), we arrive at the following contradiction:

$$\mathcal{F}^B_{\tau_\rho} \subseteq \mathcal{F}^R_{\tau_\rho} \subseteq \mathcal{H}_{\tau_\rho} \subsetneq \mathcal{F}^\theta_{\tau_\rho} \subseteq \mathcal{F}^B_{\tau_\rho}.$$

Hence there is no strong solution of the SDE (3.12). $\square$

3.8 Discussion of Conjecture 3.5

Theorem 3.4 shows that, for a one-dimensional Brownian motion B, the SDE

$$\mathrm{d}X_t = \frac{1}{|X_t|}\begin{bmatrix} -X_t^2 \\ X_t^1 \end{bmatrix} \mathrm{d}B_t \tag{3.2}$$

has no strong solution starting from the origin. Therefore we cannot immediately adapt the proof of Lemma 2.29 to the case of Markov controls. However, in order to prove that Conjecture 3.5 holds, we would need to show that, for any other SDE of the form

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t, \tag{3.22}$$

there is no strong solution starting from the origin that has the same expected cost.

The key property of the SDE (3.2) is that any solution of this SDE has a deterministically increasing radius, as proved in Lemma 2.4. In Proposition 3.21, we will derive the form of SDEs (3.22) that have this property and we will show in Corollary 3.22 that any such SDE can replace the SDE (3.2) in the proof of Lemma 2.29.

Proposition 3.21. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which an $\mathbb{R}^2$-valued Brownian motion is defined with natural filtration $\mathbb{F}^B$. Suppose there exists a Borel function $\sigma : D \to U$ such that the SDE*

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t, \quad X_0 = 0,$$

has a strong solution X with $t \mapsto |X_t|$ deterministically increasing. Then there exists a Borel function $\gamma : D \to \{x \in \mathbb{R}^2 \colon |x| = 1\}$ such that, for any $x = (x_1, x_2)^\top \in D$,

$$\sigma(x) = \frac{1}{|x|}\begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} \gamma(x)^\top.$$

Moreover, $|X_t| = \sqrt{t}$, for all $t \geq 0$.

Proof. Let $t \geq 0$. Since X is a continuous $\mathbb{R}^2$-valued $\mathbb{F}^B$-adapted martingale, we can write

$$X_t = R_t \begin{bmatrix} \cos\theta_t \\ \sin\theta_t \end{bmatrix},$$

for continuous $\mathbb{F}^B$-adapted semimartingales R and θ, where R takes values in $[0, \infty)$ and θ takes values in $\mathbb{R}/2\pi\mathbb{Z}$. We call R the *radius process* of X and θ the *angle* *process*. We can then write

$$\begin{aligned}
\mathrm{d}R_t &= \alpha_t^\top \,\mathrm{d}B_t + \beta_t \,\mathrm{d}t, \\
\mathrm{d}\theta_t &= \zeta_t^\top \,\mathrm{d}B_t + \xi_t \,\mathrm{d}t,
\end{aligned}$$

for some $\mathbb{F}^B$-adapted $\mathbb{R}^2$-valued processes α, ζ and $\mathbb{F}^B$-adapted $\mathbb{R}$-valued process β, ξ.

In order for R to be deterministic, we must have $\alpha \equiv (0,0)^\top$. Then, by Itô's formula, we have that

$$\begin{aligned}
\mathrm{d}X_t &= \begin{bmatrix} \left(\beta_t - \frac{1}{2}R_t\left|\zeta_t\right|^2\right)\cos\theta_t - \left(\xi_t R_t + \alpha_t^\top \zeta_t\right)\sin\theta_t \\ \left(\beta_t - \frac{1}{2}R_t\left|\zeta_t\right|^2\right)\sin\theta_t + \left(\xi_t R_t + \alpha_t^\top \zeta_t\right)\cos\theta_t \end{bmatrix} \mathrm{d}t \\
&\quad + \begin{bmatrix} \cos\theta_t \alpha_t^\top - R_t \sin\theta_t \zeta_t^\top \\ \sin\theta_t \alpha^\top + R_t\cos\theta_t\zeta_t^\top \end{bmatrix} \mathrm{d}B_t \\
&= \begin{bmatrix} \left(\beta_t - \frac{1}{2}R_t\left|\zeta_t\right|^2\right)\cos\theta_t - \xi_t R_t\sin\theta_t \\ \left(\beta_t - \frac{1}{2}R_t\left|\zeta_t\right|^2\right)\sin\theta_t + \xi_t R_t\cos\theta_t \end{bmatrix} \mathrm{d}t + \begin{bmatrix} -R_t\sin\theta_t\zeta_t^\top \\ R_t\cos\theta_t\zeta_t^\top \end{bmatrix} \mathrm{d}B_t.
\end{aligned} \tag{3.23}$$

Since X is a martingale, the drift term must vanish, and so we have the two equations

$$\begin{aligned}
\left(\beta_t - \frac{1}{2}R_t\left|\zeta_t\right|^2\right)\cos\theta_t &= \xi_t R_t \sin\theta_t, \\
\left(\beta_t - \frac{1}{2}R_t\left|\zeta_t\right|^2\right)\sin\theta_t &= -\xi_t R_t \cos\theta_t.
\end{aligned}$$

Multiplying the first equation by $\cos\theta_t$ and the second by $\sin\theta_t$ and summing, we get that

$$\beta_t - \frac{1}{2}R_t\left|\zeta_t\right|^2 = \xi_t R_t\left(\sin\theta_t\cos\theta_t - \cos\theta_t\sin\theta_t\right) = 0. \tag{3.24}$$

Hence

$$\xi_t R_t \sin\theta_t = \xi_t R_t\cos\theta_t = 0, \quad \text{for all} \quad t \geq 0,$$

and so $\xi \equiv 0$.

We now impose the unit quadratic variation condition. From (3.23), we have that

$$\sigma(X_t) = \begin{bmatrix} -R_t\sin\theta_t\zeta_t^\top \\ R_t\cos\theta_t\zeta_t^\top \end{bmatrix}.$$

And so, since $\sigma \in \mathcal{U}$,

$$\operatorname{Tr}\left(\begin{bmatrix} -R_t \sin\theta_t \zeta_t^\top \\ R_t \cos\theta_t \zeta_t^\top \end{bmatrix} \begin{bmatrix} -R_t \sin\theta_t \zeta_t^\top \\ R_t \cos\theta_t \zeta_t^\top \end{bmatrix}^\top \right) = 1.$$

This can be rewritten as

$$R_t^2 \left|\zeta_t\right|^2 \left(\sin\theta_t^2 + \cos\theta_t^2\right) = 1.$$

Therefore

$$|\zeta_t| = \frac{1}{R_t}.$$

Substituting this back into (3.24), we also have

$$\beta_t = \frac{1}{2R_t}.$$

We conclude that the radius R solves the deterministic equation

$$\mathrm{d}R_t = \frac{1}{2R_t}\,\mathrm{d}t,$$

and so, given the initial condition $R_0 = 0$, we find that

$$R_t = \sqrt{t}.$$

The angle process θ is a martingale satisfying

$$\mathrm{d}\theta_t = \zeta_t^\top\,\mathrm{d}B_t,$$

for some process ζ such that $|\zeta_t| = R_t^{-1}$.

Rewriting the matrix $\sigma(X_t)$ in terms of X, we have

$$\sigma(X_t) = \begin{bmatrix} -X_t^2 \\ X_t^1 \end{bmatrix} \zeta_t^\top,$$

for some process ζ satisfying $|\zeta_t| = |X_t|^{-1}$ for all $t \geq 0$.

Noting that ζ_t must be a Borel function of X_t, we normalise ζ and define a Borel function $\gamma : D \to \mathbb{R}^2$ such that

$$\gamma(X_t) := |X_t|\,\zeta_t, \quad t \geq 0.$$

We then have

$$\sigma(X_t) = \frac{1}{|X_t|}\begin{bmatrix} -X_t^2 \\ X_t^1 \end{bmatrix}\gamma(X_t)^\top,$$

with $|\gamma(X_t)| = 1$ for all $t \geq 0$. □

We now verify that a solution of such an SDE can play the role of tangential motion in Lemma 2.29.

Corollary 3.22. *Suppose that X is a strong solution of an SDE of the form given in Proposition 3.21. Let $R > 0$ and $\eta \in (0, R)$, and let $\tilde{f} : [0, R)$ be a continuous function. Then*

$$\mathbb{E}^0\left[\int_0^{\tau_\eta} \tilde{f}(|X_s|)\,\mathrm{d}s\right] = 2\int_0^\eta \xi\tilde{f}(\xi)\,\mathrm{d}\xi.$$

Proof. By Proposition 3.21, $|X_t| = \sqrt{t}$ for any $t \geq 0$. Therefore

$$\begin{aligned}\mathbb{E}^0\left[\int_0^{\tau_\eta} \tilde{f}(|X_s|)\,\mathrm{d}s\right] &= \int_0^{\eta^2} \tilde{f}(\sqrt{s})\,\mathrm{d}s \\ &= 2\int_0^\eta \xi\tilde{f}(\xi)\,\mathrm{d}\xi,\end{aligned}$$

making the change of variables $\xi = \sqrt{s}$ in the integral. □

If the Markov value function is equal to the strong and weak value functions and the value is attained by some Markov control, then by Proposition 3.21, taking B to be an $\mathbb{R}^2$-valued Brownian motion, there must be a strong solution of the SDE

$$\mathrm{d}X_t = \frac{1}{|X_t|}\begin{bmatrix} -X_t^2 \\ X_t^1 \end{bmatrix}\gamma(X_t)^\top\,\mathrm{d}B_t, \tag{3.25}$$

for some Borel function $\gamma : D \to \{x \in \mathbb{R}^2 \colon |x| = 1\}$, with $X_0 = 0$. A first step in completing the proof of Conjecture 3.5 would be to show that there is no strong solution of the SDE (3.25).

We would also need to show that there is no minimising sequence of Markov controls whose limiting cost is equal to the strong and weak value functions. We identified one possible minimising sequence in Proposition 3.18, by considering solutions of the SDE

$$\mathrm{d}X_t = \frac{1}{|X_t|}\begin{bmatrix} \lambda X_t^1 - \sqrt{1-\lambda^2}X_t^2 \\ \lambda X_t^2 + \sqrt{1-\lambda^2}X_t^1 \end{bmatrix}\mathrm{d}B_t, \tag{3.12}$$

starting from the origin, for a one-dimensional Brownian motion B, and taking the limit as $\lambda \downarrow 0$. In Theorem 3.15, we showed that, for small parameter values

$\lambda \in (0, \frac{\sqrt{2}}{2})$, the SDE (3.12) has no strong solution starting from the origin, thus ruling out one possible minimising sequence.

However, we could also construct minimising sequences from solutions of the more general SDE

$$\mathrm{d}X_t = \frac{1}{|X_t|} \begin{bmatrix} \lambda X_t^1 - \sqrt{1-\lambda^2} X_t^2 \\ \lambda X_t^2 + \sqrt{1-\lambda^2} X_t^1 \end{bmatrix} \gamma(X_t)^\top \, \mathrm{d}B_t,$$

for an $\mathbb{R}^2$-valued Brownian motion B and a Borel function $\gamma : D \to \{x \in \mathbb{R}^2 \colon |x| = 1\}$. We would therefore need to prove that such SDEs also have no strong solution starting from the origin.

In summary, in this chapter, we have found two SDEs that do not have a strong solution starting from the origin, as shown in Theorem 3.4 and Theorem 3.15. These theorems give support to the assertion of Conjecture 3.5 that there is a gap between the Markov and strong value functions at the origin for costs $\tilde{f}$ that satisfy the growth condition

$$\int_0^r \tilde{f}(s) \, \mathrm{d}s = \infty \quad \text{and} \quad \int_0^r s\tilde{f}(s) \, \mathrm{d}s < \infty, \quad \text{for any} \quad r > 0.$$

We have outlined above the remaining steps that would be required to prove the conjecture in future work.

CHAPTER 4

VISCOSITY SOLUTIONS OF HAMILTON-JACOBI-BELLMAN EQUATIONS

Viscosity solutions of Hamilton-Jacobi-Bellman (HJB) equations are a key tool in the study of stochastic optimal control problems. In Chapter 2, we solved a control problem for continuous radially symmetric cost functions by constructing a candidate value function and showing that it solves the appropriate Hamilton-Jacobi-Bellman equation in the viscosity sense. In order to prove that the candidate function is equal to the value function, we referred to Theorem 4.20, which states that the value function is the unique viscosity solution of a boundary value problem for the HJB equation.

In this chapter, we define the notion of viscosity solutions and prove Theorem 4.20. In order to prove the theorem, we first show that the value function is a viscosity solution of the HJB equation by applying the dynamic programming principle from Section 1.4.2. We then prove uniqueness of viscosity solutions by using a perturbation method to adapt a standard comparison principle, and finally we verify that the value function satisfies the required boundary condition. We also discuss a control problem with a discontinuous cost function and the associated viscosity theory.

4.1 Introduction

A typical approach to solving a stochastic control problem, as described for example in [26] and [58], is as follows. First conjecture an optimal control and compute the value attained by following such a control, then verify that this candidate value is in fact the value function for the control problem. The verification step can be made by directly checking that a dynamic programming principle holds, as in the examples of Section 2.2, or more commonly via a PDE characterisation. Suppose that Assumption 1.16 holds. As described in Section 1.4.4, we expect the value function v defined in Section 1.4.1 to solve the boundary value problem

$$\begin{cases} -\frac{1}{2}\inf_{\sigma\in U}\operatorname{Tr}\left(D^2 v\sigma\sigma^\top\right) = f, & \text{in} \quad D, \\ v = g, & \text{on} \quad \partial D, \end{cases} \tag{4.1}$$

as a consequence of the dynamic programming principle (1.5). The PDE in (4.1) is known as a Hamilton-Jacobi-Bellman (HJB) equation. If we have uniqueness of solutions of (4.1), then to conclude that the candidate function is equal to the value function, it is sufficient to show that the candidate solves (4.1). In this context, it is appropriate to consider solutions of the HJB equation in the viscosity sense.

Viscosity solutions were introduced by Crandall and Lions in [14] to study Hamilton-Jacobi equations; these are first order equations that arise in deterministic optimal control problems, as described in Chapter I of [26]. Lions developed the theory of viscosity solutions for second order HJB equations in [43], and viscosity solutions of more general second order equations are studied by Crandall, Ishii and Lions in their User's Guide [13] and by Ishii and Lions in [36]. In both [13] and [36], a comparison principle is proved for second order PDEs, from which uniqueness follows. Viscosity solutions of HJB equations and the related comparison principles are also presented by Fleming and Soner in [26] and by Touzi in [58].

We note that, since the HJB equation in (4.1) has no time dependence and no direct dependence on the zeroth derivative, the comparison results from the above references do not apply directly. The User's Guide [13] suggests methods for extending the comparison result that is presented there, and we will see that we can use the perturbation method suggested in Section 5.C of [13] to prove comparison for the HJB equation in (4.1).

The main result of this chapter is Theorem 4.20, which states that the value function for the control problem defined in Section 1.4.1 is the unique viscosity solution of the boundary value problem (4.1). We note that proving Theorem 4.20 will complete the proof of Proposition 2.15, where we referred to this result in order to characterise the value function of the control problem.

In this chapter, we begin by defining viscosity solutions in Section 4.2. In Section 4.3, we prove that the value function is a viscosity solution of the HJB equation. This result follows from the dynamic programming principle that we established in Section 1.4.3. We then prove a comparison principle for the HJB equation. We do this in two stages. In Lemma 4.17, we provide the details of the perturbation method suggested in Section 5.C of [13]. Then in Proposition 4.19 we complete the proof of comparison by choosing a specific perturbation that is suitable for the HJB equation. Here we use the perturbation that is suggested for proving comparison for a Monge-Ampère equation in Section V.3 of [36]. To complete the proof of Theorem 4.20, in Section 4.5 we prove that the value function extends continuously to the closure of the domain and satisfies the boundary condition $v = g$ pointwise.

We conclude this chapter by discussing viscosity solutions of PDEs with discontinuous data in Section 4.7. Extending the definition of viscosity solutions to allow discontinuous data, as in [11] or [12], our proof of comparison no longer holds. Therefore, for the control problems with step cost functions in Section 2.2, we cannot use the theory of viscosity solutions to verify the candidate value function. However, we will show that the value function does solve an HJB equation in a generalised viscosity sense.

4.2 Viscosity solutions

We first define viscosity solutions of second order PDEs, following [13] and [26].

Fix $d \in \mathbb{N}$ and let $D \subset \mathbb{R}^d$ be a domain. Denote by S_d the set of $d \times d$ symmetric matrices, and let $F : D \times \mathbb{R} \times \mathbb{R}^d \times S_d \to \mathbb{R}$ be a differential operator. We are interested in the PDE

$$F\left(x, u(x), Du(x), D^2u(x)\right) = 0. \tag{4.2}$$

As stated in the introduction of [13], we require monotonicity conditions on F in the zeroth and second order derivatives. In the following, we equip the space S_d of symmetric matrices with the usual partial ordering.

Notation. Let A and B be symmetric matrices of the same dimension. We use the notation $A \leq B$ to denote that $A - B$ is a non-positive definite matrix. Similarly, $A < B$ denotes that $A - B$ is negative definite.

Assumption 4.1. We assume that the following conditions are satisfied.

1. The operator F is continuous in each of its arguments;

2. The operator F is *proper*; i.e. for any $x \in D$, $p \in \mathbb{R}^d$ and $X \in S_d$,

$$F(x, r, p, X) \leq F(x, s, p, X) \quad \text{for} \quad r \leq s; \tag{4.3}$$

3. The operator F is *degenerate elliptic*; i.e. for any $x \in D$, $r \in \mathbb{R}$ and $p \in \mathbb{R}^d$,

$$F(x, r, p, X) \leq F(x, r, p, Y) \quad \text{for} \quad X \geq Y. \tag{4.4}$$

   Note that we allow for cases where we may have equality of the operator for some $X > Y$. These cases are known as degeneracies and include the case where F does not depend on the second order derivative.

We now motivate the definition of viscosity solutions, as in Section 2 of [13]. Suppose that there exists a classical solution u of the PDE (4.2) and that ϕ is a smooth function that sits above u at some point $x^0 \in D$; i.e. $\phi(x^0) = u(x^0)$ and $\phi \geq u$ in a neighbourhood of x^0. Then $u - \phi$ has a local maximum at x^0, which implies that $Du(x^0) = D\phi(x^0)$ and $D^2u(x^0) \leq D^2\phi(x^0)$. By ellipticity of F, we have

$$F\left(x^0, u(x^0), D\phi(x^0), D^2\phi(x^0)\right) \leq F(x^0, u(x^0), Du(x^0), D^2u(x^0)) = 0. \tag{4.5}$$

Similarly, for a smooth ψ sitting below u, we have

$$F\left(x^0, u(x^0), D\psi(x^0), D^2\psi(x^0)\right) \geq 0, \tag{4.6}$$

at the local minimum x^0 of $u - \psi$.

We use these two properties to define a weak solution of the PDE (4.2). We say that a function u is a viscosity solution of (4.2) if any smooth functions sitting above and below u satisfy the inequalities (4.5) and (4.6), respectively, without the requirement that u is twice continuously differentiable. A viscosity solution is, therefore, a generalisation of a classical solution. We give a precise definition of viscosity solutions following Definition 4.2 of [26, Chapter II].

Definition 4.2 (Viscosity solution I)**.** An upper semicontinuous function $u : \overline{D} \to \mathbb{R}$ is a *viscosity subsolution* of (4.2) if, for every smooth $\phi \in C^\infty(D)$,

$$F\left(x^0, u(x^0), D\phi(x^0), D^2\phi(x^0)\right) \leq 0,$$

at any point $x^0 \in D$ that is a local maximum of $u - \phi$.

Similarly, a lower semicontinuous function $u : \overline{D} \to \mathbb{R}$ is a *viscosity supersolution* of (4.2) if, for every smooth $\psi \in C^\infty(D)$,

$$F\left(x^0, u(x^0), D\psi(x^0), D^2\psi(x^0)\right) \geq 0,$$

at any point $x^0 \in D$ that is a local minimum of $u - \psi$.

A continuous function u that is both a viscosity subsolution and a viscosity supersolution of (4.2) is a *viscosity solution*.

We will find that a second equivalent definition of viscosity solutions will be more convenient in some cases. This definition is taken from Definition 2.2 of [13] and is given in terms of *semijets*, which we now define.

Definition 4.3 (Semijets)**.** Given a set $D \subset \mathbb{R}^d$, a function $u : D \to \mathbb{R}$ and a point $x^0 \in D$, we define $J_D^{2,+}u(x^0) \subset \mathbb{R}^d \times S_d$, the *second order superjet* of u at x^0, as follows. We say that $(p, X) \in J_D^{2,+}u(x^0)$ if and only if

$$u(x) \leq u(x^0) + p^\top(x - x^0) + \frac{1}{2}(x - x^0)^\top X(x - x^0) + o(\left|x - x^0\right|^2), \quad \text{as} \quad x \to x^0.$$

We define $J_D^{2,-}u(x^0)$, the *second order subjet* of u at x^0, similarly. We say that $(p, X) \in J_D^{2,-}u(x^0)$ if and only if

$$u(x) \geq u(x^0) + p^\top(x - x^0) + \frac{1}{2}(x - x^0)^\top X(x - x^0) + o(\left|x - x^0\right|^2), \quad \text{as} \quad x \to x^0.$$

For $x \in \operatorname{int}(D)$, we define

$$J^{2,+}u(x) = J_O^{2,+}u(x) \quad \text{and} \quad J^{2,-}u(x) = J_O^{2,-}u(x),$$

where $O \subset D$ is any neighbourhood of x.

The following definition is the main definition of a viscosity solution given in [13]. It is also given as an alternative formulation of a viscosity solution in Definition 4.1 of [26, Chapter V].

Definition 4.4 (Viscosity solution II)**.** An upper semicontinuous function $u : D \to \mathbb{R}$ is a *viscosity subsolution* of (4.2) if

$$F(x, u(x), p, X) \leq 0 \quad \text{for all} \quad x \in D,\ (p, X) \in J_D^{2,+}u(x).$$

A lower semicontinuous function $u : D \to \mathbb{R}$ is a *viscosity supersolution* of (4.2) if

$$F(x, u(x), p, X) \geq 0 \quad \text{for all} \quad x \in D,\ (p, X) \in J_D^{2,-}u(x).$$

A *viscosity solution* of (4.2) is a continuous function $u : D \to \mathbb{R}$ that is both a viscosity subsolution and a viscosity supersolution.

Remark 4.5. For F continuous in each of its arguments, the semijets in the above definition can equivalently be replaced by their closures, defined as follows.

Notation. We use the notation $\overline{J}_D^{2,+}$ to denote the closure of the set $J_D^{2,+}$ in $\mathbb{R}^d \times S_d$; i.e. $(p, X) \in \overline{J}_D^{2,+}$ if and only if there exists $(x_n, p_n, X_n)_{n\in\mathbb{N}} \subset D \times \mathbb{R}^d \times S_d$ such that

$$(p_n, X_n) \in J_D^{2,+} \quad \text{for each} \quad n \in \mathbb{N},$$

and

$$(x_n, u(x_n), p_n, X_n) \xrightarrow{n\to\infty} (x, u(x), p, X).$$

The closure of the subjet $\overline{J}_D^{2,-}$ is defined similarly.

We also introduce the notation

$$J_D^2(x) := J_D^{2,+}(x) \cap J_D^{2,-}(x),$$

as in the appendix of [13], and we denote the closure of this set as $\overline{J}_D^2$, where this is defined analogously to the closures of the semijets.

Having presented two definitions of viscosity solutions, it is necessary to check that they are equivalent. We make use of the following lemma.

Lemma 4.6. *Let $u : D \to \mathbb{R}$, $p \in \mathbb{R}^d$, $X \in S_d$, and $x^0 \in D$. Then $(p, X) \in J_D^{2,+}u(x^0)$ if and only if there exists $\phi \in C^\infty(D)$ such that*

$$x^0 \in \arg\max(u - \phi) \quad \textit{and} \quad \big(D\phi(x^0), D^2\phi(x^0)\big) = (p, X).$$

This result can be proved similarly to Lemma 4.1 of [26, Chapter V], which states the analogous result for parabolic PDEs. We omit the details of the proof here.

Proposition 4.7. *Definition 4.2 and Definition 4.4 are equivalent.*

Proof. The equivalence of the definitions is an immediate consequence of Lemma 4.6. □

4.3 Viscosity solution characterisation of the value function

The main result of this chapter is Theorem 4.20 below, which states that the value function for the control problem defined in Section 1.4.1 is the unique solution of a boundary value problem for the associated Hamilton-Jacobi-Bellman equation. Throughout this chapter, we suppose that Assumption 1.16 holds. In particular, the conditions are met for the weak and strong value functions to be equal, as shown in Proposition 1.7, and so we refer to this common function as the value function. It will be convenient to work with the strong formulation in what follows, defining the value function $v : D \to \mathbb{R}$ as in (1.4) by

$$v(x) = \inf_{\sigma \in \mathcal{U}} \mathbb{E}^x \left[\int_0^\tau f(X_s^\sigma) \, \mathrm{d}s + g(X_\tau^\sigma) \right], \quad x \in D.$$

In this section, we prove that the value function is both a viscosity subsolution and a viscosity supersolution of the HJB equation.

Recall that, under Assumption 1.16, the following dynamic programming principle holds, by Proposition 1.17. For any $x \in D$ and any stopping time θ such that $\theta \in [0, \tau]$ almost surely,

$$v(x) = \inf_{\sigma \in \mathcal{U}} \mathbb{E}^x \left[\int_0^\theta f(X_s^\sigma) \, \mathrm{d}s + v(X_\theta^\sigma) \right]. \tag{1.5}$$

We showed in Section 1.4.4 that, under sufficient smoothness conditions, the value function is a classical solution of the Hamilton-Jacobi-Bellman equation

$$-\frac{1}{2} \inf_{\sigma \in U} \operatorname{Tr} \left(D^2 v \sigma \sigma^\top \right) - f = 0, \quad x \in D. \tag{4.7}$$

We now show that, as a consequence of the dynamic programming principle, the value function is a viscosity solution of the HJB equation (4.7) under milder assumptions. We follow the same strategy of proof as in Section 7.1 of [58].

Proposition 4.8. *Suppose that Assumption 1.16 holds and that $f : D \to \mathbb{R}$ is continuous. Then the value function v is a viscosity subsolution of the HJB equation* (4.7)*.*

Proof. Let $x^0 \in D$, fix $\sigma \in U$ and let X^σ be the controlled process following the constant control σ. Let $\phi \in C^\infty(D)$ be such that $\phi(x^0) = v(x^0)$ and $\phi(x) \geq v(x)$ for all $x \in \mathcal{N}$, where $\mathcal{N} \subset D$ is an open neighbourhood of x^0.

For $h > 0$, define the stopping time

$$\theta_h := \inf \{ t \geq 0 : X_t^\sigma \in \partial \mathcal{N} \} \wedge h.$$

Since $\mathcal{N} \subset D$, we have the pointwise inequality $\theta_h < \tau$ and so, by the dynamic programming principle (1.5), we have

$$\phi(x^0) = v(x^0) \leq \mathbb{E}^{x^0} \left[\int_0^{\theta_h} f(X_s^\sigma) \, \mathrm{d}s + v(X_{\theta_h}^\sigma) \right]. \tag{4.8}$$

Applying Itô's formula to $\phi(X^\sigma_{\theta_h})$, we see that

$$\phi(x^0) = \mathbb{E}^{x^0}\left[\phi(X^\sigma_{\theta_h}) - \frac{1}{2}\int_0^{\theta_h} \operatorname{Tr}\left(D^2\phi(X^\sigma_s)\sigma\sigma^\top\right) \mathrm{d}s\right]. \tag{4.9}$$

Subtracting expression (4.8) from (4.9) gives

$$\mathbb{E}^{x^0}\left[\phi(X^\sigma_{\theta_h}) - v(X^\sigma_{\theta_h})\right] \le \mathbb{E}^{x^0}\left[\int_0^{\theta_h}\left(f(X^\sigma_s) + \frac{1}{2}\operatorname{Tr}\left(D^2\phi(X^\sigma_s)\sigma\sigma^\top\right)\right)\mathrm{d}s\right].$$

This inequality implies that

$$\mathbb{E}^{x^0}\left[\int_0^{\theta_h}\left(f(X^\sigma_s) + \frac{1}{2}\operatorname{Tr}\left(D^2\phi(X^\sigma_s)\sigma\sigma^\top\right)\right)\mathrm{d}s\right] \ge 0, \tag{4.10}$$

since $\phi - v \ge 0$ in $\mathcal{N}$, the process X^σ has continuous paths and, by Corollary 1.13, v is continuous on $\overline{\mathcal{N}}$.

Note that $\theta_h = h$ for h sufficiently small, and so, since f is continuous and ϕ is twice continuously differentiable, we can apply the mean value theorem to get

$$\lim_{h\to 0}\frac{1}{h}\int_0^{\theta_h}\left(f(X^\sigma_s) + \frac{1}{2}\operatorname{Tr}\left(D^2\phi(X^\sigma_s)\sigma\sigma^\top\right)\right)\mathrm{d}s = f(x^0) + \frac{1}{2}\operatorname{Tr}\left(D^2\phi(x^0)\sigma\sigma^\top\right).$$

The integrand is bounded and the stopping time θ_h is bounded above by τ, which has finite expectation by Proposition 1.5. Therefore the integral is bounded above by an integrable random variable independent of h. Hence we can apply dominated convergence to take the limit inside the expectation and see that

$$\lim_{h\to 0}\frac{1}{h}\mathbb{E}^{x^0}\left[\int_0^{\theta_h}\left(f(X^\sigma_s) + \frac{1}{2}\operatorname{Tr}\left(D^2\phi(X^\sigma_s)\sigma\sigma^\top\right)\right)\mathrm{d}s\right] = f(x^0) + \frac{1}{2}\operatorname{Tr}\left(D^2\phi(x^0)\sigma\sigma^\top\right).$$

Combining this with the bound (4.10) gives

$$-\frac{1}{2}\operatorname{Tr}\left(D^2\phi(x^0)\sigma\sigma^\top\right) \le f(x^0),$$

and so we have the desired result. □

Next we check the supersolution property.

Proposition 4.9. *Suppose that Assumption 1.16 holds and that $f : D \to \mathbb{R}$ is continuous. Then the value function v is a viscosity supersolution of the HJB equation (4.7).*

Proof. Let $x^0 \in D$ and let $\phi \in C^\infty(D)$ be such that $\phi(x^0) = v(x^0)$ and $\phi(x) < v(x)$

for all $x \in \mathcal{N}'$, where $\mathcal{N}' \subset D$ is some open neighbourhood of x^0.

Define $H : \mathbb{R} \times S_d \to \mathbb{R}$ by

$$H(x, X) := -\frac{1}{2} \inf_{\sigma \in U} \operatorname{Tr}\left(X\sigma\sigma^\top\right) - f(x),$$

and suppose for contradiction that

$$H(x^0, D^2\phi(x^0)) < 0.$$

Since H is continuous, there exists an open neighbourhood $\mathcal{N}$ of x^0 such that $\mathcal{N} \subset \mathcal{N}'$ and

$$H(x, D^2\phi(x)) < 0,$$

for all $x \in \mathcal{N}$.

Let $\nu \in \mathcal{U}$ be an arbitrary control and define the stopping time

$$\theta^\nu := \inf\left\{t \geq 0 : X_t^\nu \in \partial\mathcal{N}\right\}.$$

Define

$$\eta := \min_{\partial\mathcal{N}}(v - \phi) > 0.$$

Then

$$\phi(X_{\theta^\nu}^\nu) \leq v(X_{\theta^\nu}^\nu) - \eta, \tag{4.11}$$

by continuity of the paths of X^ν and continuity of v, which was shown in Corollary 1.13.

We now apply Itô's formula to $\phi(X_{\theta^\nu}^\sigma)$ to see that

$$\begin{aligned}
v(x^0) &= \phi(x^0) \\
&= \mathbb{E}^{x^0}\left[\phi(X_{\theta^\nu}^\nu) - \frac{1}{2}\int_0^{\theta^\nu} \operatorname{Tr}\left(D^2\phi(X_s^\nu)\nu_s\nu_s^\top\right) \mathrm{d}s\right] \\
&\leq \mathbb{E}^{x^0}\left[\phi(X_{\theta^\nu}^\nu) + \int_0^{\theta^\nu} \left(H(X_s^\nu, D^2\phi(X_s^\nu)) + f(X_s^\nu)\right) \mathrm{d}s\right] \\
&\leq \mathbb{E}^{x^0}\left[\phi(X_{\theta^\nu}^\nu) + \int_0^{\theta^\nu} f(X_s^\nu) \, \mathrm{d}s\right],
\end{aligned}$$

using the fact that $H \leq 0$ on $\partial\mathcal{N}$, by continuity of H.

Finally, we use the inequality (4.11) to arrive at

$$v(x^0) \leq \mathbb{E}^{x^0}\left[v(X_{\theta^\nu}^\nu) + \int_0^{\theta^\nu} f(X_s^\nu) \, \mathrm{d}s\right] - \eta.$$

Since $\eta > 0$ is independent of the arbitrary control process ν, taking the infimum over controls in $\mathcal{U}$ gives

$$v(x^0) \leq \inf_{\nu \in \mathcal{U}} \mathbb{E}^{x^0} \left[v(X^\nu_{\theta^\nu}) + \int_0^{\theta^\nu} f(X^\nu_s) \, \mathrm{d}s \right] - \eta.$$

This contradicts the dynamic programming principle (1.5).

Hence the value function v is a viscosity supersolution of the HJB equation. □

Combining the preceding two results, we see that, under Assumption 1.16 and the additional assumption that the cost function f is continuous, the value function solves the HJB equation (4.7) in the viscosity sense. We will see in Theorem 4.20 that the value function is in fact the unique viscosity solution of (4.7) that satisfies the appropriate boundary condition.

In the next section, we prove uniqueness of viscosity solutions of a boundary value problem for the HJB equation.

4.4 Comparison principle

The usual approach to proving uniqueness of viscosity solutions of a boundary value problem is to prove a comparison principle for sub- and supersolutions, as in Section 3 of [13], and deduce from this the desired uniqueness result.

Returning to the general form of the PDE (4.2), let $D \subset \mathbb{R}$ and $F : D \times \mathbb{R} \times \mathbb{R}^d \times S_d \to \mathbb{R}$. We wish to prove uniqueness of viscosity solutions of

$$F(x, u(x), Du(x)D^2u(x)) = 0, \quad x \in D, \tag{4.2}$$

that satisfy a given boundary condition. To see that a comparison principle holds for the PDE (4.2), we make the following standard assumptions, as in Section 3 of [13].

Assumption 4.10. Suppose that the following assumptions hold.

1. The domain D is open and bounded;
2. The operator F is continuous in each of its arguments;
3. The operator F is proper; i.e. F satisfies

$$F(x, r, p, X) \leq F(x, s, p, X) \quad \text{for} \quad r \leq s; \tag{4.3}$$

4. The operator F is coercive in the zeroth order derivative; i.e. there exists $\gamma > 0$ such that

$$F(x,s,p,X) - F(x,r,p,X) \geq \gamma(s-r), \quad \text{for} \quad r \leq s; \tag{4.12}$$

5. There exists a function $\omega : [0,\infty] \to [0,\infty]$, with $\omega(0+) = 0$, such that

$$F(y,r,\alpha(x-y),Y) - F(x,r,\alpha(x-y),X) \leq \omega(\alpha\,|x-y|^2 + |x-y|), \tag{4.13}$$

for any $\alpha > 0$, whenever X and Y satisfy the matrix inequality

$$-3\alpha \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \leq \begin{bmatrix} X & 0 \\ 0 & -Y \end{bmatrix} \leq 3\alpha \begin{bmatrix} I & -I \\ -I & I \end{bmatrix}. \tag{4.14}$$

Remark 4.11. Note that (4.14) implies that $X \leq Y$. Therefore the fifth statement of the above assumption is satisfied if (4.13) holds for all $X \leq Y$. In particular, as noted in Example 3.6 of [13], if $G : \mathbb{R} \times \mathbb{R}^d \times S_d \to \mathbb{R}$ is degenerate elliptic, as defined in (4.4), $f : D \to \mathbb{R}$ is continuous, and F is of the form

$$F(x,r,p,X) = G(r,p,X) - f(x),$$

then the fifth statement of Assumption 4.10 is satisfied.

We now state, but do not prove, the comparison principle that is proved in Theorem 3.3 of [13]. In Section 4.4.1, we will adapt the proof given in [13] to generalise this result.

Notation. For a domain $D \subseteq \mathbb{R}^d$, denote the sets of upper and lower semicontinuous real-valued functions on D by USC(D) and LSC(D), respectively.

Theorem 4.12 (Comparison)**.** *Let $D \subset \mathbb{R}$ and $F : D \times \mathbb{R} \times \mathbb{R}^d \times S_d \to \mathbb{R}$ satisfy Assumption 4.10. Suppose that*

$$\begin{aligned} u &\in \mathrm{USC}(\overline{D}) \quad \textit{is a viscosity subsolution of } (4.2), \\ v &\in \mathrm{LSC}(\overline{D}) \quad \textit{is a viscosity supersolution of } (4.2), \end{aligned}$$

and

$$u \leq v \quad \textit{on} \quad \partial D.$$

Then

$$u \leq v \quad \textit{on} \quad \overline{D}.$$

As a consequence of this comparison principle, we can obtain uniqueness of viscosity solutions of the Dirichlet problem

$$\begin{cases} F\left(x, u(x), Du(x)D^2u(x)\right) = 0 & \text{in} \quad D, \\ u = g & \text{on} \quad \partial D, \end{cases} \tag{4.15}$$

for some $g : \partial D \to \mathbb{R}$. We define a solution of (4.15) as follows.

Definition 4.13. We say that $u \colon \overline{D} \to \mathbb{R}$ is a viscosity solution of the Dirichlet problem (4.15) if u is a viscosity solution of the PDE (4.2) in D, u is continuous on $\overline{D}$, and

$$u(x_0) = g(x_0),$$

for all $x_0 \in \partial D$.

In this definition, we impose the boundary condition in a strict pointwise sense, as in Section 4 of [13]. This allows us to deduce uniqueness for the Dirichlet problem immediately.

Corollary 4.14 (Uniqueness)**.** *Let $D \subset \mathbb{R}$ and $F : D \times \mathbb{R} \times \mathbb{R}^d \times S_d$ satisfy Assumption 4.10, and let $g : \partial D \to \mathbb{R}$. Suppose that u and v are both viscosity solutions of the Dirichlet problem* (4.15)*. Then $u = v$ on $\overline{D}$.*

Proof. Let u and v be viscosity solutions of (4.15). Then, in particular, u is a viscosity subsolution of the PDE (4.2) and v is a viscosity supersolution of (4.2). Furthermore, $u = v = g$ on ∂D, so $u \leq v$ on ∂D. Therefore Theorem 4.12 tells us that $u \leq v$ on $\overline{D}$.

On the other hand, u is a viscosity supersolution of (4.2) and v is a viscosity subsolution of (4.2). Again, since u and v both satisfy the boundary condition in (4.15), $v \leq u$ on ∂D. Therefore, by Theorem 4.12, $v \leq u$ on $\overline{D}$.

We conclude that $u = v$ on $\overline{D}$, and so any viscosity solution of the Dirichlet problem (4.15) is unique. □

As noted in the survey paper [37] of Jensen and Smears, it is possible to prove comparison principles for viscosity solutions when the notion of boundary conditions is relaxed. For example, weaker definitions of boundary conditions for viscosity solutions are given by Barles and Souganidis in [2] and by Crandall, Ishii and Lions in Definitions 7.1 and 7.4 of the User's Guide [13]. These definitions allow for viscosity solutions that do not attain the boundary conditions continuously. However, in this thesis it will be sufficient to consider viscosity solutions that are continuous on the closure of the domain $\overline{D}$.

Remark 4.15. We make the following remarks on relaxing each statement of Assumption 4.10.

1. In Section 5.D of [13], the authors adapt the proof of the comparison principle to the case of an unbounded domain, under linear growth conditions on the sub- and supersolutions u and v. Here we will always retain the assumption that the domain D is bounded.

2. We would like to relax the continuity assumption to allow for the case where F is of the form
$$F(x, r, p, X) = G(r, p, X) - f(x),$$
where G is continuous in each of its arguments, but f may have discontinuities. In Section 4.7, we give a definition of viscosity solutions that allows for discontinuities of this type, but we have not been able to prove a comparison principle for these solutions.

3. We will always assume that F is proper.

4. To prove a comparison principle for the HJB equation (4.7), we will need to relax the coercivity condition (4.12). We can see that the HJB operator F is not coercive since, taking $r, s \in \mathbb{R}$ with $r < s$,
$$\begin{aligned} &F(x, s, p, X) - F(x, r, p, X) \\ &\quad = -\frac{1}{2}\inf_{\sigma\in U}\left\{\mathrm{Tr}(X\sigma\sigma^\top)\right\} - f(x) + \frac{1}{2}\inf_{\sigma\in U}\left\{\mathrm{Tr}(X\sigma\sigma^\top)\right\} + f(x) \\ &\quad = 0 < \gamma(s-r), \end{aligned}$$
for any $\gamma > 0$. Two methods to relax the coercivity condition (4.12) are presented in Section 5.C of [13]. In our Lemma 4.17, we verify the details of one of these methods, and we employ this method to prove a comparison principle for the HJB equation in Proposition 4.19.

5. We note that the HJB equation (4.7) satisfies the fifth statement of Assumption 4.10 when the function f is continuous. Again, we look to relax this condition by introducing viscosity solutions for (4.7) with a discontinuity in f in Section 4.7.

The key ingredient in the proof of the comparison principle for viscosity solutions is the Crandall-Ishii Lemma, which we now state, referring to Section 6.7 of [58] for a proof. We will refer to this lemma in the following section in order to prove a generalisation of the comparison principle to HJB equations.

Lemma 4.16 (Crandall-Ishii Lemma). *Let $D \subset \mathbb{R}^d$ be open and locally compact, and let $u_1, u_2 \in \mathrm{USC}(D)$. Define $w : D^2 \to \mathbb{R}$ by*

$$w(x_1, x_2) := u_1(x_1) + u_2(x_2), \quad \textit{for} \quad x_1, x_2 \in D. \tag{4.16}$$

Suppose that $x^0 \in D^2$ and $\varphi \in C^2(\overline{D^2})$ are such that

$$(w - \varphi)(x^0) = \max_{D^2}\{w - \varphi\}. \tag{4.17}$$

Then, for any $\varepsilon > 0$, there exist $X_1, X_2 \in S_d$ such that, for $i = 1, 2$,

$$(D_{x_i}\varphi(x^0), X_i) \in \overline{J}_D^{2,+} u_i(x_i^0), \tag{4.18}$$

and

$$-(\varepsilon^{-1} + \left\|D^2\varphi(x^0)\right\|)I_{2d} \leq \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \leq D^2\varphi(x^0) + \varepsilon D^2\varphi(x^0)^2, \tag{4.19}$$

where the norm $\|\cdot\|$ is defined for symmetric matrices $A \in S_{2d}$ by

$$\|A\| := \sup\left\{\left|\xi^\top A\xi\right| : \xi \in \mathbb{R}^{2d}, |\xi| \leq 1\right\}.$$

In the following section, we will use the Crandall-Ishii Lemma to show that the matrix inequality (4.14) in the fifth statement of Assumption 4.10 holds and deduce the existence of a function ω that satisfies (4.13).

4.4.1 Comparison for an HJB equation

We will now show that a comparison principle holds for the HJB equation (4.7). As noted in Remark 4.15, the coercivity condition (4.12) is not satisfied, and so we can not apply the comparison result of Theorem 4.12 directly.

We first show that the perturbation method described in Section 5.C of [13] leads to the following comparison principle without the coercivity requirement.

Lemma 4.17. *Suppose that $D \subset \mathbb{R}^d$ and $F : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times S_d \to \mathbb{R}$ satisfy statements 1, 2, 3 and 5 of Assumption 4.10.*

Let $u \in \mathrm{USC}(\overline{D})$ be a viscosity subsolution and $v \in \mathrm{LSC}(\overline{D})$ a viscosity supersolution of (4.2), and suppose that

$$u \leq v \quad \textit{on} \quad \partial D.$$

Suppose moreover that, for each $k \in \mathbb{N}$, there exists $\delta_k > 0$ and a function $\psi_k \in C^\infty(D)$ such that

$$|\psi_k| \leq \frac{1}{k},$$

and $u_k := u + \psi_k$ is a viscosity subsolution of

$$F(x, u_k, Du_k, D^2u_k) + \delta_k = 0. \tag{4.20}$$

Then

$$u \leq v \quad \textit{on} \quad \overline{D}.$$

For completeness, we give the following detailed proof, which is omitted when the result is stated in Section 5.C of [13].

Proof. We first prove that $u_k \leq v$ on ∂D implies that $u_k \leq v$ on $\overline{D}$. Then we take the limit as $k \to \infty$ to conclude that $u \leq v$ on $\overline{D}$.

Step 1: Fix $k \in \mathbb{N}$. We have that u_k is a viscosity subsolution of (4.20) and therefore also a viscosity subsolution of $F(x, u_k, Du_k, D^2u_k) = 0$, since, for any test function $\phi \in C^\infty(D)$ and $x^0 \in \arg\max(u_k - \phi)$,

$$F\left(x^0, u_k(x^0), D\phi(x^0), D^2\phi(x^0)\right) \leq -\delta_k \leq 0.$$

We also have that v is a viscosity supersolution of $F(x, v, Dv, D^2v) = 0$. Since u_k is the sum of the upper semicontinuous function u and the smooth function ψ_k, we see that u_k is upper semicontinuous.

Suppose that $u_k \leq v$ on ∂D. We will now show that we have the strict inequality $u_k < v$ on D. We broadly follow the proof of Theorem 4.12, which is given in detail as the proof of Theorem 3.3 in [13], but we note that we have not assumed that the fourth statement of Assumption 4.10 holds.

We apply the Crandall-Ishii Lemma (Lemma 4.16) to the function $\varphi : \overline{D}^2 \to \mathbb{R}$, defined by

$$\varphi(x_1, x_2) = \frac{1}{2} x^\top A x, \quad x_1, x_2 \in \overline{D},$$

where

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad A = \alpha \begin{bmatrix} I & -I \\ -I & I \end{bmatrix},$$

for some $\alpha > 0$. We can also write φ as

$$\varphi(x_1, x_2) = \frac{\alpha}{2} |x_1 - x_2|^2, \quad x_1, x_2 \in D.$$

We calculate the first and second order derivatives

$$D_{x_1}\varphi(x) = \alpha(x_1 - x_2), \quad D_{x_2}\varphi(x) = \alpha(x_2 - x_1),$$

$$D^2\varphi \equiv A, \quad (D^2\varphi)^2 = A^2 = 2\alpha A,$$

as required in Lemma 4.16. We also calculate that the norm of the Hessian of φ is

$$\left\|D^2\varphi\right\| = \inf\{\left|\xi^\top A\xi\right| : \xi \in \mathbb{R}^{2d}, |\xi| \leq 1\} = 2\alpha,$$

since, for any $\xi = (\xi_1, \xi_2)^\top \in \mathbb{R}^{2d}$ with $|\xi| \leq 1$,

$$\xi^\top A\xi = \alpha\,|\xi_1 - \xi_2|^2 \leq 2\alpha,$$

with equality when $\xi_2 = -\xi_1$ and $|\xi_1| = \frac{1}{2}$.

Now, let us suppose for contradiction that there exists $z \in D$ such that

$$u_k(z) \geq v(z). \tag{4.21}$$

We introduce the following notation. Define $m_\alpha^k \colon D^2 \to \mathbb{R}$ by

$$m_\alpha^k(x_1, x_2) := u_k(x_1) - v(x_2) - \frac{\alpha}{2}\,|x_1 - x_2|^2\,,$$

and

$$M_\alpha^k := \sup_{(x_1,x_2)\in\overline{D}^2} m_\alpha^k(x_1, x_2).$$

Note that the maximum is attained due to compactness of $\overline{D}^2$ and upper semicontinuity of $u_k - v$, and so there exists $x^\alpha \in \overline{D}^2$ such that

$$M_\alpha^k = u_k(x_1^\alpha) - v(x_2^\alpha) - \frac{\alpha}{2}\,|x_1^\alpha - x_2^\alpha|^2\,.$$

Then we see that

$$\begin{aligned} M_\alpha^k &\geq m_\alpha^k(z, z) \\ &= u_k(z) - v(z) - \frac{\alpha}{2}\,|z - z|^2 \\ &= u_k(z) - v(z) \geq 0. \end{aligned}$$

Hence

$$u_k(x_1^\alpha) - v(x_2^\alpha) \geq u_k(x_1^\alpha) - v(x_2^\alpha) - \frac{\alpha}{2}\,|x_1^\alpha - x_2^\alpha|^2 = M_\alpha^k \geq 0. \tag{4.22}$$

It can be shown that, as $\alpha \to \infty$,

$$\alpha \left| x_1^\alpha - x_2^\alpha \right|^2 \to 0 \quad \text{and so also} \quad \left| x_1^\alpha - x_2^\alpha \right| \to 0.$$

We refer to Lemma 3.1 and Proposition 3.7 of [13] for a proof of this fact. Combining the first limit with the inequality (4.22), we see that we can take α sufficiently large that $x^\alpha \in D^2$, since $u_k \leq v$ on ∂D.

Now fix $\varepsilon > 0$. Then, taking $u_1 = u_k$, $u_2 = -v$ to be the two upper semicontinuous functions in the Crandall-Ishii Lemma, we see that, as $x^\alpha \in D^2$ is a maximiser of

$$u_k(x_1) - v(x_2) - \frac{\alpha}{2} \left| x_1 - x_2 \right|^2,$$

there exist matrices $X_1^k, X_2 \in S_d$ such that

$$\left(\alpha(x_1^\alpha - x_2^\alpha), X_1^k\right) \in \overline{J}_D^{2,+} u_k(x_1^\alpha), \quad \left(-\alpha(x_1^\alpha - x_2^\alpha), X_2\right) \in \overline{J}_D^{2,+}(-v)(x_2^\alpha),$$

and

$$-\left(\varepsilon^{-1} + 2\alpha\right) \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \leq \begin{bmatrix} X_1^k & 0 \\ 0 & X_2 \end{bmatrix} \leq \alpha(1 + 2\alpha\varepsilon) \begin{bmatrix} I & -I \\ -I & I \end{bmatrix}.$$

The matrix inequality above implies the condition (4.14) in Assumption 4.10, with $X = X_1^k$, $Y = -X_2$, since by choosing $\varepsilon = \alpha^{-1}$, we have

$$-3\alpha \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \leq \begin{bmatrix} X_1^k & 0 \\ 0 & X_2 \end{bmatrix} \leq 3\alpha \begin{bmatrix} I & -I \\ -I & I \end{bmatrix}.$$

Therefore, by the fifth statement of Assumption 4.10, there exists a function $\omega\colon [0, \infty] \to [0, \infty]$, with $\omega(0+) = 0$, such that

$$\begin{aligned} &F(x_2, r, \alpha(x_1 - x_2), -X_2) - F(x_1, r, \alpha(x_1 - x_2), X_1^k) \\ &\quad \leq \omega(\alpha \left| x_1 - x_2 \right|^2 + \left| x_1 - x_2 \right|). \end{aligned} \tag{4.23}$$

Now, since u_k is a viscosity subsolution of (4.20), and $\left(\alpha(x_1^\alpha - x_2^\alpha), X_1^k\right) \in \overline{J}_D^{2,+} u_k(x_1^\alpha)$, we have

$$F\left(x_1^\alpha, u_k(x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1^k\right) \leq -\delta_k,$$

by Definition 4.4. We also note that

$$\left(-\alpha(x_1^\alpha - x_2^\alpha), X_2\right) \in \overline{J}_D^{2,+}(-v)(x_2^\alpha)$$

is equivalent to
$$(\alpha(x_1^\alpha - x_2^\alpha), -X_2) \in \overline{J}_D^{2,-} v(x_2^\alpha).$$

Since v is a viscosity supersolution of (4.2), Definition 4.4 gives us
$$F\left(x_2^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), -X_2\right) \geq 0.$$

Combining the above inequalities, we have
$$F\left(x_1^\alpha, u_k(x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1^k\right) - F\left(x_2^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), -X_2\right) \leq -\delta_k. \tag{4.24}$$

Note that coercivity would usually be used to obtain an inequality of the form (4.24), as in the proof of Theorem 3.3 of [13], but here the positive constant δ_k plays the role of the coercivity constant.

Noting that F is proper, by the third statement of Assumption 4.10, and $u_k(x_1^\alpha) \geq v(x_2^\alpha)$ by (4.22), we have that
$$0 \leq F\left(x_1^\alpha, u_k(x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1^k\right) - F\left(x_1^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1^k\right). \tag{4.25}$$

We can rewrite the right-hand side of (4.25) and apply the inequality (4.24) to get
$$\begin{aligned}
0 &\leq F\left(x_1^\alpha, u_k(x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1^k\right) - F\left(x_2^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), -X_2\right) \\
&\quad + F\left(x_2^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), -X_2\right) - F\left(x_1^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1^k\right) \\
&\leq F\left(x_2^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), -X_2\right) - F\left(x_1^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1^k\right) - \delta_k.
\end{aligned}$$

Hence, by (4.23), we have
$$\begin{aligned}
\delta_k &\leq F\left(x_2^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), -X_2\right) - F\left(x_1^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1^k\right) \\
&\leq \omega\left(\alpha\left|x_1^\alpha - x_2^\alpha\right|^2 + \left|x_1^\alpha - x_2^\alpha\right|\right).
\end{aligned}$$

Since $\alpha\left|x_1^\alpha - x_2^\alpha\right|^2 \to 0$, as $\alpha \to \infty$, and $\omega(0+) = 0$, we can take the limit as $\alpha \to \infty$ in the above inequality to arrive at
$$\delta_k \leq 0.$$

This is a contradiction, and so there cannot exist any $z \in D$ such that (4.21) holds. Hence
$$u_k < v \quad \text{on} \quad D.$$

Step 2: We have shown that the implication

$$u_k \le v \quad \text{on} \quad \partial D \quad \Rightarrow \quad u_k \le v \quad \text{on} \quad \overline{D} \tag{4.26}$$

holds for any viscosity subsolution u_k of (4.20) and any viscosity supersolution v of (4.2). We now show that this is equivalent to

$$\sup_{\overline{D}}(u_k - v) = \sup_{\partial D}(u_k - v). \tag{4.27}$$

It is clear that (4.27) implies (4.26). Now suppose that (4.26) holds and let $\xi := \sup_{\partial D}(u_k - v)$.

First suppose that $\xi > 0$. We will show that $v + \xi$ is a viscosity supersolution of (4.2). Let $x^0 \in D$ and $\phi \in C^\infty(D)$ be such that $x^0 \in \arg\min(v + \xi - \phi)$. Define $\tilde{\phi} = \phi - \xi$. Then $\tilde{\phi} \in C^\infty(D)$ with $D\phi = D\tilde{\phi}$, $D^2\phi = D^2\tilde{\phi}$ and $x^0 \in \arg\min(v - \tilde{\phi})$. Since v is a viscosity supersolution of (4.2),

$$F\left(x^0, v(x^0), D\phi(x^0), D^2\phi(x^0)\right) = F\left(x^0, v(x^0), D\tilde{\phi}(x^0), D^2\tilde{\phi}(x^0)\right) \ge 0,$$

and since $\xi > 0$ and F is proper, we get the required supersolution property

$$F(x^0, v(x^0) + \xi, D\phi(x^0), D^2\phi(x^0)) \ge F(x^0, v(x^0), D\phi(x^0), D^2\phi(x^0)) \ge 0.$$

We have $u_k \le v + \xi$ on ∂D and so, by (4.26), $u_k < v + \xi$ on $\overline{D}$.

Suppose now that $\xi < 0$. Then, similarly, $u_k - \xi$ is a viscosity subsolution of (4.20). We have $u_k - \xi \le v$ on ∂D and so, by (4.26), $u_k - \xi < v$ on $\overline{D}$.

Finally, suppose that $\xi = 0$. Then $u_k \le v$ on ∂D and so, by (4.26) again, $u_k < v$ on $\overline{D}$. We have now shown that $\sup_{\overline{D}}(u_k - v) \le \xi$. On the other hand, $\partial D \subset \overline{D}$ implies that

$$\sup_{\overline{D}}(u_k - v) \ge \sup_{\partial D}(u_k - v) = \xi,$$

and so (4.27) holds.

Step 3: The final step in the proof is to take the limit as $k \to \infty$. Let $x \in \overline{D}$. Then, combining the results of Step 1 and Step 2, we have

$$\begin{aligned} u(x) + \psi_k(x) - v(x) = u_k(x) - v(x) &\le \sup_{\overline{D}}(u_k - v) \\ &= \sup_{\partial D}(u_k - v) = \sup_{\partial D}(u + \psi_k - v). \end{aligned}$$

Recalling that $|\psi_k| \leq \frac{1}{k}$, we then have

$$u(x) - v(x) \leq \sup_{\partial D}(u - v) + 2\frac{1}{k},$$

and so, taking the limit as $k \to \infty$,

$$u(x) - v(x) \leq \sup_{\partial D}(u - v).$$

Hence $\sup_{\overline{D}}(u - v) \leq \sup_{\partial D}(u - v)$. Using the fact that $\partial D \subset \overline{D}$, once again, we conclude that

$$\sup_{\overline{D}}(u - v) = \sup_{\partial D}(u - v).$$

Therefore

$$u \leq v \quad \text{on} \quad \partial D$$

implies that

$$u \leq v \quad \text{on} \quad \overline{D},$$

as required. □

To prove that a comparison principle holds for the HJB equation (4.7), we now need to make a suitable choice of the perturbation $(\psi_k)_{k\in\mathbb{N}}$. We take inspiration from Section V.3 of [36], where Ishii and Lions use the above perturbation argument for a Monge-Ampère equation. The perturbation $(\psi_k)_{k\in\mathbb{N}}$ that we will apply to the HJB equation (4.7) is of the same form as the perturbation suggested in Section V.3 of [36] for the Monge-Ampère equation. For $k \in \mathbb{N}$, we define

$$\psi_k(x) = \frac{1}{k}\exp\left\{\frac{|x|^2}{2} - C\right\},$$

for some constant C. In Chapter 5, we will study viscosity solutions for Monge-Ampère equations and their connection to stochastic control problems. We will prove a comparison principle for a Monge-Ampère equation in Appendix A.2 using the same perturbation as defined above.

In the following proof, we make use of the fact that the control set $U \subset \mathbb{R}^{d,d}$ is compact. We now prove this assertion.

Lemma 4.18. *Let $\|\cdot\|$ be any norm on $\mathbb{R}^{d,d}$. Then*

$$U := \left\{\sigma \in \mathbb{R}^{d,d} \colon \operatorname{Tr}(\sigma\sigma^\top) = 1\right\}$$

is a compact set in the normed space $(\mathbb{R}^{d,d}, \|\cdot\|)$.

Proof. By the Heine-Borel theorem, U is compact if and only if it is a bounded and closed subset of $\mathbb{R}^{d,d}$. As all matrix norms are equivalent, it suffices to show that U is bounded and closed with respect to the Frobenius norm $\|\cdot\|_{\mathrm{F}}$, defined by

$$\|\sigma\|_{\mathrm{F}} := \sqrt{\mathrm{Tr}(\sigma\sigma^\top)}, \quad \sigma \in \mathbb{R}^{d,d}.$$

Any $\sigma \in U$ has norm $\|\sigma\|_{\mathrm{F}} = 1$. Hence U is bounded.

Now take a convergent sequence $(\sigma_n)_{n\in\mathbb{N}} \subset U$ with limit $\sigma \in \mathbb{R}^{d,d}$. We have that $\|\sigma_n\|_{\mathrm{F}} = 1$, for all $n \in \mathbb{N}$, and so

$$0 \leq |1 - \|\sigma\|_{\mathrm{F}}| = |\|\sigma_n\|_{\mathrm{F}} - \|\sigma\|_{\mathrm{F}}| \leq \|\sigma_n - \sigma\|_{\mathrm{F}} \to 0 \quad \text{as} \quad n \to \infty.$$

Therefore $\|\sigma\|_{\mathrm{F}} = 1$, and so $\mathrm{Tr}(\sigma\sigma^\top) = 1$ and $\sigma \in U$.

This shows that U is closed and completes the proof of compactness. □

We are now ready to prove the comparison principle for the HJB equation (4.7).

Proposition 4.19. *Suppose that Assumption 1.1 holds and that* $f : D \to \mathbb{R}$ *is continuous. Then we have the following comparison principle for the HJB equation* (4.7).

Suppose that

$$\begin{aligned} &u \in \mathrm{USC}(\overline{D}) \quad \textit{is a viscosity subsolution of } (4.7),\\ &v \in \mathrm{LSC}(\overline{D}) \quad \textit{is a viscosity supersolution of } (4.7), \end{aligned}$$

and

$$u \leq v \quad \textit{on} \quad \partial D.$$

Then

$$u \leq v \quad \textit{on} \quad \overline{D}.$$

Proof. First we check that conditions 1, 2, 3 and 5 of Assumption 4.10 hold.

1. By Assumption 1.1, the domain D is open and bounded.

2. We have assumed that f is continuous. We wish to argue that the operator $F : D \times \mathbb{R} \times \mathbb{R}^d \times S_d$, defined by

$$F(x, r, p, X) \equiv F(x, X) = -\frac{1}{2} \inf_{\sigma \in U} \left\{ \mathrm{Tr}\left(X\sigma\sigma^\top\right) \right\} - f(x),$$

is continuous in each of its arguments. It remains to show that the map $H : S_d \to \mathbb{R}$ defined by

$$H(X) := -\frac{1}{2} \inf_{\sigma \in U} \left\{ \operatorname{Tr}\left(X\sigma\sigma^\top \right) \right\}$$

is continuous.

Define $h : S_d \times U \to \mathbb{R}$ by

$$h(X, \sigma) := \operatorname{Tr}\left(\sigma\sigma^\top X \right),$$

so that, for any $X \in S_d$, we can write

$$H(X) = \inf_{\sigma \in U} h(X, \sigma).$$

Since both matrix multiplication and the trace operator are continuous, we have that $h : S_d \times U \to \mathbb{R}$ is continuous. Since U is compact, as proved in Lemma 4.18, it follows that $H : S_d \to \mathbb{R}$ is continuous, as the infimum over continuous functions. Hence F is continuous in each of its arguments, as required.

3. Let $r \leq s$, then, for any $x \in D$, $p \in \mathbb{R}^d$ and $X \in S_d$,

$$F(x, r, p, X) - F(x, s, p, X) = F(x, X) - F(x, X) = 0,$$

so F is proper.

5. Define $G : S_d \to \mathbb{R}$ by $G(X) = -\frac{1}{2} \inf_{\sigma \in U} \operatorname{Tr}(X\sigma\sigma^\top)$ for $X \in S_d$. Then the operator F is of the form

$$F(x, r, p, X) = G(X) - f(x),$$

with f continuous. The operator G is degenerate elliptic, since $X \leq Y$ implies that $\operatorname{Tr}(X\sigma\sigma^\top) \leq \operatorname{Tr}(Y\sigma\sigma^\top)$, for any $\sigma \in U$. Therefore, as we noted in Remark 4.11, the fifth statement of Assumption 4.10 holds.

We now apply Lemma 4.17 with the following perturbation to the subsolution. Let $m \in \mathbb{N}$ and set $C = \sup_{x \in D} \frac{|x|^2}{2}$. Define $\psi_m : D \to \mathbb{R}$ by

$$\psi_m(x) := \frac{1}{m} \exp\left\{ \frac{|x|^2}{2} - C \right\},$$

and define $u_m : D \to \mathbb{R}$

$$u_m(x) := u(x) + \psi_m(x),$$

for $x \in D$. Note that we have $|\psi_m(x)| \leq \frac{1}{m}\exp\{0\} = \frac{1}{m}$.

To conclude the proof via Lemma 4.17, we need to show that there exists $\delta_m > 0$ such that u_m is a viscosity subsolution of

$$-\frac{1}{2}\inf_{\sigma\in U}\left\{\mathrm{Tr}\left(D^2 u_m \sigma\sigma^\top\right)\right\} - f + \delta_m = 0.$$

Fix $x^0 \in D$ and let $\phi \in C^\infty(D)$ be such that $x^0 \in \arg\max(u_m - \phi)$. Then, since $\psi_m \in C^\infty(D)$ and $\psi_m \geq 0$, we have that $(\phi - \psi_m) \in C^\infty(D)$ and $x^0 \in \arg\max(u - (\phi - \psi_m))$. Since u is a viscosity subsolution of the HJB equation (4.7), this implies that

$$-\frac{1}{2}\inf_{\sigma\in U}\left\{\mathrm{Tr}\left(D^2(\phi - \psi_m)\sigma\sigma^\top\right)\right\} - f \leq 0,$$

and so

$$-\frac{1}{2}\inf_{\sigma\in U}\left\{\mathrm{Tr}\left(D^2(\phi)\sigma\sigma^\top\right)\right\} - f + \frac{1}{2}\inf_{\sigma\in U}\left\{\mathrm{Tr}\left(D^2(\psi_m)\sigma\sigma^\top\right)\right\} \leq 0. \tag{4.28}$$

Now we calculate that

$$D^2\psi_m(x) = \frac{1}{m}\exp\left\{\frac{|x|^2}{2} - C\right\}\left(I + xx^\top\right),$$

so, for any $\sigma \in U$,

$$\begin{aligned}\mathrm{Tr}\left(D^2(\psi_m)\sigma\sigma^\top\right) &= \frac{1}{m}\exp\left\{\frac{|x|^2}{2} - C\right\}\mathrm{Tr}\left([I + xx^\top]\sigma\sigma^\top\right)\\ &= \frac{1}{m}\exp\left\{\frac{|x|^2}{2} - C\right\}\left[\mathrm{Tr}(\sigma\sigma^\top) + \mathrm{Tr}(xx^\top\sigma\sigma^\top)\right].\end{aligned}$$

Since $xx^\top\sigma\sigma^\top$ is positive semi-definite and $\mathrm{Tr}(\sigma\sigma^\top) = 1$, this gives us the bound

$$\mathrm{Tr}\left(D^2(\psi_m)\sigma\sigma^\top\right) \geq \frac{1}{m}\exp\left\{\frac{|x|^2}{2} - C\right\} \geq \frac{1}{m}\exp\{-C\}.$$

Let us define $\delta_m := \frac{1}{2m}\exp\{-C\}$. Then we have that

$$\frac{1}{2}\inf_{\sigma\in U}\left\{\mathrm{Tr}(D^2(\psi_m)\sigma\sigma^\top)\right\} \geq \frac{1}{2}\frac{1}{m}\exp\{-C\} = \delta_m,$$

and so, by (4.28),

$$-\frac{1}{2}\inf_{\sigma\in U}\left\{\mathrm{Tr}(D^2(\phi)\sigma\sigma^\top)\right\} - f + \delta_m \leq 0.$$

This shows that u_m satisfies the required subsolution property and so, by Lemma 4.17, we conclude that comparison holds for the HJB equation. □

Uniqueness of viscosity solutions of the boundary value problem (4.15) for the HJB equation (4.7) follows immediately from the comparison principle, as in Corollary 4.14. We now state the main theorem that we will prove in this chapter.

Theorem 4.20. *Suppose that Assumption 1.16 holds, and suppose further that the domain D is uniformly convex, the running cost f is continuous in D, and the boundary cost g is uniformly continuous on ∂D.*

Then the value function $v : D \to \mathbb{R}$ defined in Section 1.4.1 extends continuously to $\overline{D}$ and is the unique viscosity solution of the HJB equation

$$-\frac{1}{2}\inf_{\sigma\in U}\operatorname{Tr}\left(D^2 v\sigma\sigma^\top\right) - f = 0 \tag{4.7}$$

in D, with boundary condition

$$v = g \quad on \quad \partial D.$$

We have shown that the value function v is a viscosity solution of the HJB equation (4.7) and that solutions of the boundary value problem for (4.7) are unique. To complete the proof of Theorem 4.20, it remains to show that the value function v attains the boundary condition.

4.5 Boundary condition for the value function

In this section, we will show that the value function v attains the value g on the boundary of the domain. Once again, we draw similarities with the control problem connected to a Monge-Ampère equation that we will study in Chapter 5. In [28], Gaveau shows that the value function for that control problem solves a boundary value problem for a Monge-Ampère equation in some weak sense. We adapt Gaveau's proof of attainment of the boundary condition to the case of the control problem with value function v. We adopt the following notation in this section, so that the dependence of a controlled process on the control and the initial condition is explicit.

Notation. For $\sigma \in \mathcal{U}$ and $x \in D$, denote by $X^{\sigma,x}$ a strong solution of the SDE

$$\mathrm{d}X_t = \sigma_t\,\mathrm{d}B_t,$$

with initial condition $X_0 = x$, and define the exit time

$$\tau^{\sigma,x} := \inf\{t \geq 0 \colon X_t^{\sigma,x} \in \partial D\}.$$

We will denote by $\mathbb{E}$ the expectation with respect to the law of $X^{\sigma,x}$.

We first prove an inequality for the boundary value, under the assumption that the domain is convex.

Proposition 4.21. *Suppose that Assumption 1.1 holds and, moreover, that the domain D is convex, the function f is bounded on D, and the function g is continuous on ∂D. Let $x \in D$ and $x_0 \in \partial D$. Then*

$$\limsup_{x \to x_0} v(x) \leq g(x_0).$$

Proof. Fix $\varepsilon > 0$, choose $\delta \in (0, 1)$, and let $x \in D$ be such that $|x - x_0| < \delta$.

Define the constant control $\sigma^1 \in \mathcal{U}$ by

$$\sigma_t^1 = \frac{1}{|x - x_0|} \left[x - x_0; \quad 0; \quad \cdots; \quad 0 \right], \quad t \geq 0,$$

and let W denote the first component of the Brownian motion B. Then

$$X_t^{\sigma^1,x} = \frac{W_t}{|x - x_0|}(x - x_0),$$

for any $t \geq 0$, and the controlled process $X^{\sigma^1,x}$ acts as a Brownian motion on the line connecting x to the boundary point x_0. Note that there also exists $z \in \partial D$, $z \neq x_0$, such that the line through x and x_0 intersects ∂D at z. By convexity of the domain, the line segment between x_0 and z is contained in D.

By definition of the value function,

$$\begin{aligned} v(x) \leq \mathcal{I}(x; \sigma^1) &:= \mathbb{E}\left[\int_0^{\tau^{\sigma^1,x}} f\left(X_s^{\sigma^1,x}\right) \mathrm{d}s + g\left(X_{\tau^{\sigma^1,x}}^{\sigma^1,x}\right)\right] \\ &\leq \mathbb{E}\left[\tau^{\sigma^1,x}\right] \|f\|_\infty + \mathbb{E}\left[g\left(X_{\tau^{\sigma^1,x}}^{\sigma^1,x}\right)\right]. \end{aligned} \tag{4.29}$$

For the remainder of this proof, we denote $X^x = X^{\sigma^1,x}$ and $\tau^x = \tau^{\sigma^1,x}$. We bound the first term of (4.29) as follows.

Let $p_x := \mathbb{P}\left[X_{\tau^x}^x = x_0\right]$ and $\eta_x := |x - z|$. Then, using well-known properties of the hitting times of one-dimensional Brownian motion, we have

$$\begin{aligned}\mathbb{E}\left[\tau^x\right] &= p_x\delta^2 + (1-p_x)\eta_x^2 \\ &< \delta^2 + (1-p_x)\eta_x^2.\end{aligned}$$

Since $|x - x_0| < \delta$, we also have $1 - p_x \leq \frac{\delta}{\delta+\eta_x} < \frac{\delta}{\eta_x}$, and so

$$\mathbb{E}[\tau^x] < \delta^2 + \delta\eta_x \leq \delta(1 + \mathrm{diam}(D)).$$

Choosing $\delta < \varepsilon \left(2 \left\| f \right\|_\infty [1 + \mathrm{diam}(D)]\right)^{-1}$, we can bound the first term of (4.29) by

$$\mathbb{E}\left[\tau^x\right] \left\| f \right\|_\infty < \frac{\varepsilon}{2}. \tag{4.30}$$

We bound the second term as follows. Consider

$$\begin{aligned}\left|\mathbb{E}\left[g(X^x_{\tau^x})\right] - g(x_0)\right| &= \left|p_x g(x_0) + (1-p_x)g(z) - g(x_0)\right| \\ &= (1-p_x)\left|g(z) - g(x_0)\right| \\ &\leq \min\left\{2(1-p_x)\left\|g\right\|_\infty, \left|g(x_0) - g(z)\right|\right\}.\end{aligned}$$

We treat two cases separately, depending on the ratio between the distances δ and η_x. First, suppose that $\eta_x \geq \delta^{\frac{1}{2}}$. Then

$$1 - p_x < \frac{\delta}{\eta_x} \leq \delta^{\frac{1}{2}}.$$

Choosing δ such that $\delta^{\frac{1}{2}} < \frac{\varepsilon}{4}\left\|g\right\|_\infty^{-1}$, we have

$$2(1-p_x)\left\|g\right\|_\infty < \frac{\varepsilon}{2}.$$

On the other hand, if $\eta_x < \delta^{\frac{1}{2}}$, then

$$|x_0 - z| \leq |x_0 - x| + |x - z| < \delta + \delta^{\frac{1}{2}} \leq 2\delta^{\frac{1}{2}}.$$

Since g is continuous on ∂D, we can then choose δ sufficiently small that

$$|g(x_0) - g(z)| < \frac{\varepsilon}{2}.$$

Combining the two cases, we see that we can always choose δ small enough that

$$\left|\mathbb{E}\left[g\left(X^x_{\tau^x}\right)\right] - g(x_0)\right| < \frac{\varepsilon}{2},$$

which implies that

$$\mathbb{E}\left[g\left(X^x_{\tau^x}\right)\right] < g(x_0) + \frac{\varepsilon}{2}. \tag{4.31}$$

Fixing δ sufficiently small and substituting the bounds (4.30) and (4.31) into (4.29) gives us

$$\begin{aligned} v(x) &\le \mathbb{E}\left[\int_0^{\tau^x} f\left(X^x_s\right) \mathrm{d}s + g\left(X^x_{\tau^x}\right)\right] \\ &< \frac{\varepsilon}{2} + g(x_0) + \frac{\varepsilon}{2} = g(x_0) + \varepsilon. \end{aligned}$$

Thus we have shown that $\limsup_{x\to x_0} v(x) \le g(x_0)$. □

For attainment of the boundary condition, we require a stronger convexity condition on the domain, which we define analogously to Gaveau's definition of strictly pseudoconvex subsets of the complex plane in Section 3 of [28].

Definition 4.22. A set $D \subset \mathbb{R}^d$ is *uniformly convex* if there exists $p \in C^2(\mathbb{R}^d)$ such that

$$D = \left\{x \in \mathbb{R}^d \colon p(x) < 0\right\},$$

$Dp \neq 0$ on ∂D, and p is uniformly convex; i.e. there exists $\alpha > 0$ such that the function $x \mapsto p(x) - \alpha x^2$ is convex.

In particular, this definition excludes polygonal domains in dimension $d = 2$.

We now state our result on the attainment of the boundary condition by the value function.

Proposition 4.23. *Suppose that Assumption 1.16 holds. Moreover, suppose that the domain D is uniformly convex, the function f is continuous in D, and the function g is uniformly continuous on ∂D.*

Then v extends continuously to $\overline{D}$ and, for any $x_0 \in \partial D$,

$$\lim_{x\to x_0} v(x) = g(x_0).$$

To prove this result, we use the following two lemmas, which are similar to Lemma 2 and Lemma 4 of [28]. Our first lemma gives a bound on the expected exit time from the domain.

Lemma 4.24. *Suppose that $D \subset \mathbb{R}^d$ is uniformly convex, with the function p and constant α as defined in Definition 4.22. For $\varepsilon > 0$, define a domain which contains D by*

$$D_\varepsilon := \left\{x \in \mathbb{R}^d \colon p(x) < \varepsilon\right\}.$$

Let $\sigma \in \mathcal{U}$ and $x \in D_\varepsilon$, and define the exit time

$$\tau_\varepsilon^{\sigma,x} := \inf\{t \geq 0\colon X_t^{\sigma,x} \notin D_\varepsilon\}.$$

Then $\mathbb{E}[\tau_\varepsilon^{\sigma,x}] \leq \frac{2}{\alpha}(\varepsilon - p(x))$.

Proof. Note that $\mathbb{E}\left[p\left(X_{\tau_\varepsilon}^{\sigma,x}\right)\right] = \varepsilon$, by continuity of the paths of $X^{\sigma,x}$ and continuity of the function p. Then we can apply Itô's formula to get

$$\begin{aligned}
\varepsilon = \mathbb{E}\left[p\left(X_{\tau_\varepsilon^{\sigma,x}}^{\sigma,x}\right)\right] &= p(x) + \mathbb{E}\left[\int_0^{\tau_\varepsilon^{\sigma,x}} Dp(X_s^{\sigma,x})^\top \sigma_s \, \mathrm{d}B_s\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\int_0^{\tau_\varepsilon^{\sigma,x}} \operatorname{Tr}(D^2 p(X_s^{\sigma,x})\sigma_s\sigma_s^\top)\, \mathrm{d}s\right] \qquad (4.32) \\
&= p(x) + \frac{1}{2}\mathbb{E}\left[\int_0^{\tau_\varepsilon^{\sigma,x}} \operatorname{Tr}(D^2 p(X_s^{\sigma,x})\sigma_s\sigma_s^\top)\, \mathrm{d}s\right],
\end{aligned}$$

noting that the integrand in the stochastic integral is bounded and so the integral has zero expectation.

For any $y \in \mathbb{R}^d$, the matrix $D^2p(y) - \alpha I$ is positive semi-definite, by uniform convexity of p. Therefore, for any symmetric positive semi-definite matrix A, we have the bound

$$\operatorname{Tr}(D^2 p(y)A) \geq \alpha \operatorname{Tr}(A).$$

Applying this bound to (4.32) gives

$$\varepsilon \geq p(x) + \frac{\alpha}{2}\mathbb{E}\left[\int_0^{\tau_\varepsilon^{\sigma,x}} \operatorname{Tr}(\sigma_s\sigma_s^\top)\, \mathrm{d}s\right] = p(x) + \frac{\alpha}{2}\mathbb{E}\left[\tau_\varepsilon^{\sigma,x}\right],$$

since $\sigma \in \mathcal{U}$. Therefore

$$\mathbb{E}\left[\tau_\varepsilon^{\sigma,x}\right] \leq \frac{2}{\alpha}\left(\varepsilon - p(x)\right),$$

as required. □

Corollary 4.25. *Suppose that $D \subset \mathbb{R}^d$ is uniformly convex. Then*

$$\mathbb{E}\left[\tau^{\sigma,x}\right] \leq -\frac{2}{\alpha}p(x).$$

Proof. Taking the limit $\varepsilon \downarrow 0$ in Lemma 4.24 gives the result. □

We now bound the expectation of the value that $X^{\sigma,x}$ takes on the boundary of the domain, in a similar manner to Lemma 4 of [28].

Lemma 4.26. *Let $x_0 \in \partial D$ and $x_1 \in D$ be such that $|x_0 - x_1| < 1$, and fix $\sigma \in \mathcal{U}$. Then*

$$\mathbb{E}\left|X^{\sigma,x_1}_{\tau^{\sigma,x_1}} - x_0\right| \leq C\left|x_1 - x_0\right|^{\frac{1}{2}},$$

for some constant $C \geq 0$ independent of x_0, x_1 and σ.

Proof. By definition of X^{σ,x_1}, we have that

$$\begin{aligned}\mathbb{E}\left|X^{\sigma,x_1}_{\tau^{\sigma,x_1}} - x_0\right| &= \mathbb{E}\left|x_1 + \int_0^{\tau^{\sigma,x_1}} \sigma_s \,\mathrm{d}B_s - x_0\right| \\ &\leq \left|x_1 - x_0\right| + \mathbb{E}\left|\int_0^{\tau^{\sigma,x_1}} \sigma_s \,\mathrm{d}B_s\right|.\end{aligned}$$

Applying Jensen's inequality and the Itô isometry, we can bound the above expectation by

$$\begin{aligned}\mathbb{E}\left|\int_0^{\tau^{\sigma,x_1}} \sigma_s \,\mathrm{d}B_s\right| &\leq \mathbb{E}\left[\left|\int_0^{\tau^{\sigma,x_1}} \sigma_s \,\mathrm{d}B_s\right|^2\right]^{\frac{1}{2}} \\ &= \mathbb{E}\left[\int_0^{\tau^{\sigma,x_1}} \mathrm{Tr}(\sigma_s\sigma_s^\top)\,\mathrm{d}s\right]^{\frac{1}{2}} \\ &= \mathbb{E}\left[\tau^{\sigma,x_1}\right]^{\frac{1}{2}}.\end{aligned} \tag{4.33}$$

Now, using the bound from from Corollary 4.25 and the fact that $p = 0$ on ∂D, we have

$$\mathbb{E}\left[\tau^{\sigma,x_1}\right] \leq \frac{2}{\alpha}(p(x_0) - p(x_1)).$$

Since $p \in C^2(\mathbb{R}^d)$, the function p is also Lipschitz and so, for some Lipschitz constant $L \geq 0$,

$$\mathbb{E}\left[\tau^{\sigma,x_1}\right] \leq \frac{2}{\alpha}L\left|x_0 - x_1\right|.$$

Combining this with (4.33), we have

$$\mathbb{E}\left|\int_0^{\tau^{\sigma,x_1}} \sigma_s \,\mathrm{d}B_s\right| \leq \left(\frac{2}{\alpha}L\right)^{\frac{1}{2}}\left|x_0 - x_1\right|^{\frac{1}{2}}.$$

Hence, setting $C = 1 - \left(\frac{2}{\alpha}L\right)^{\frac{1}{2}}$,

$$\begin{aligned}\mathbb{E}\left|X^{\sigma,x_1}_{\tau^{\sigma,x_1}} - x_0\right| &\leq \left|x_1 - x_0\right| + \mathbb{E}\left|\int_0^{\tau^{\sigma,x_1}} \sigma_s \,\mathrm{d}B_s\right| \\ &\leq C\left|x_0 - x_1\right|^{\frac{1}{2}}.\end{aligned}$$

□

We now use Corollary 4.25 and Lemma 4.26 to prove Proposition 4.23, following

Gaveau's proof of Theorem 1 in [28].

Proof of Proposition 4.23. Since we have assumed that Assumption 1.16 holds, we have that v is locally Lipschitz in D, by Corollary 1.13. By Proposition 4.21, we also have that

$$\limsup_{x \to x_0} v(x) \leq g(x_0).$$

We aim to show that, for $x_0 \in \partial D$,

$$\liminf_{x \to x_0} v(x) \geq g(x_0). \tag{4.34}$$

For $x \in D$ and $\sigma \in \mathcal{U}$ define

$$\mathcal{I}(x;\sigma) := \mathbb{E}\left[\int_0^{\tau^{x,\sigma}} f(X_t^{\sigma,x})\,\mathrm{d}t + g\left(X_{\tau^{\sigma,x}}^{\sigma,x}\right)\right].$$

We will first show that, for fixed $\sigma \in \mathcal{U}$,

$$\lim_{x \to x_0} \mathcal{I}(x;\sigma) = g(x_0).$$

We bound the running cost f and the boundary cost g separately.

Let p be a uniformly convex function such that $D = \{x \in \mathbb{R}^d \colon p(x) < 0\}$, and let $\alpha > 0$ be such that the function $x \mapsto p(x) - \alpha\,|x|^2$ is convex. Since f is bounded in D, we can use the bound from Corollary 4.25 to get

$$\begin{aligned}\left|\mathbb{E}\left[\int_0^{\tau} f(X_s^{\sigma,x})\,\mathrm{d}s\right]\right| &\leq \|f\|_\infty\,\mathbb{E}[\tau^{\sigma,x}] \\ &\leq -\frac{2}{\alpha}p(x)\,\|f\|_\infty\,.\end{aligned}$$

Then, since $p = 0$ on ∂D and p is Lipschitz in D with some Lipschitz constant $L > 0$, we have that

$$\begin{aligned}\left|\mathbb{E}\left[\int_0^{\tau} f(X_s^{\sigma,x})\,\mathrm{d}s\right]\right| &\leq \frac{2}{\alpha}(p(x_0) - p(x))\,\|f\|_\infty \\ &\leq \frac{2}{\alpha}L\,\|f\|_\infty\,|x - x_0|\,.\end{aligned}$$

Now, since g is uniformly continuous on ∂D, g has a modulus of continuity ψ, which we may assume to be concave and increasing, such that

$$|g(x) - g(y)| \leq \psi(|x - y|),$$

for any $x, y \in \partial D$. We can then use Jensen's inequality on the absolute value, and again on the modulus of continuity ψ, to bound the expected value of g on ∂D by

$$\begin{aligned}\left|\mathbb{E}\left[g(X_\tau^{\sigma,x})\right] - g(x_0)\right| &\leq \mathbb{E}\left|g(X_\tau^{\sigma,x}) - g(x_0)\right|, \\ &\leq \mathbb{E}\left[\psi(\left|X_\tau^{\sigma,x} - x_0\right|)\right], \\ &\leq \psi(\mathbb{E}\left|X_\tau^{\sigma,x} - x_0\right|).\end{aligned}$$

By Lemma 4.26, we have the bound $\mathbb{E}\left|X_\tau^{\sigma,x} - x_0\right| \leq C\left|x - x_0\right|^{\frac{1}{2}}$, for $C = 1 - \left(\frac{2}{\alpha}L\right)^{\frac{1}{2}}$. Using the assumption that ψ is increasing, we then have

$$\begin{aligned}\left|\mathbb{E}\left[g(X_\tau^{\sigma,x})\right] - g(x_0)\right| &\leq \psi(\mathbb{E}\left|X_\tau^{\sigma,x} - x_0\right|) \\ &\leq \psi(C\left|x - x_0\right|^{\frac{1}{2}}).\end{aligned}$$

Combining the bounds on the functions f and g, we have

$$\begin{aligned}\left|\mathcal{I}(x;\sigma) - g(x_0)\right| &\leq \left|\mathbb{E}\left[\int_0^\tau f(X_s^{\sigma,x})\,\mathrm{d}s\right]\right| + \left|\mathbb{E}\left[g(X_\tau^{\sigma,x})\right] - g(x_0)\right| \\ &\leq \frac{2}{\alpha}L\left\|f\right\|_\infty \left|x - x_0\right| + \psi(C\left|x - x_0\right|^{\frac{1}{2}}) \\ &\to 0, \quad \text{as } x \to x_0.\end{aligned} \tag{4.35}$$

We now use this limit to prove (4.34).

Fix $\varepsilon > 0$. Then, by definition of the value function, there exists $\sigma^\varepsilon \in \mathcal{U}$ such that $v(x) > \mathcal{I}(x;\sigma^\varepsilon) - \frac{\varepsilon}{2}$. Also, by (4.35), there exists $\delta > 0$ such that, for $x \in D$ with $|x - x_0| < \delta$,

$$\left|\mathcal{I}(x;\sigma^\varepsilon) - g(x_0)\right| < \frac{\varepsilon}{2}.$$

So, for $x \in D$ such that $|x - x_0| < \delta$, we have

$$\begin{aligned}v(x) - g(x_0) &> \mathcal{I}(x;\sigma^\varepsilon) - g(x_0) - \frac{\varepsilon}{2} \\ &> -\varepsilon.\end{aligned}$$

Therefore

$$\liminf_{x\to x_0} v(x) \geq g(x_0).$$

Combining this with the result of Proposition 4.21, we can conclude that

$$g(x_0) \leq \liminf_{x\to x_0} v(x) \leq \limsup_{x\to x_0} v(x) \leq g(x_0).$$

Since v is continuous in D, this implies that v extends continuously to $\overline{D}$ with

$$\lim_{x \to x_0} v(x) = g(x_0), \quad x_0 \in \partial D. \qquad \square$$

We now have all of the required tools to prove Theorem 4.20.

4.6 Proof of the main result

Proof of Theorem 4.20. Under Assumption 1.16, the results of Proposition 4.8 and Proposition 4.9 imply that v is a viscosity solution of the HJB equation (4.7) in D.

Since D is uniformly convex, f is continuous and g is uniformly continuous, Proposition 4.23 implies that v has a continuous extension to $\overline{D}$ that satisfies the boundary condition $v = g$ on ∂D.

Finally, Proposition 4.19 implies uniqueness of solutions of (4.7) with the given boundary condition, following the same argument as in Corollary 4.14.

Hence the value function is the unique viscosity solution of the HJB equation (4.7) with boundary condition $v = g$ on ∂D. $\square$

4.7 Viscosity solutions of PDEs with discontinuous data

In this section, we consider relaxing the assumption on continuity of F that we made in Assumption 4.1. Note that, in Example 2.1 and Example 2.6, where the cost function is a step function, the viscosity theory that we have developed so far does not apply, owing to the discontinuity in the cost function. Nevertheless, we were able to find the value function for each of these examples in Proposition 2.5 and Proposition 2.8.

For a cost function $f : D \to \mathbb{R}$ with a discontinuity, we would like to establish a characterisation of the value function via the HJB equation

$$-\frac{1}{2} \inf_{\sigma \in U} \operatorname{Tr}\left(D^2 v \sigma \sigma^\top\right) - f = 0. \tag{4.36}$$

We say that such a PDE has *discontinuous data*. The discontinuity in f means that the definition of viscosity solutions that we gave in Definition 4.2 does not apply to this equation.

However, viscosity solutions for PDEs with discontinuous data have been treated in the literature. In [11], Cattiaux, Dai Pra and Rœlly give a definition of viscosity solutions for a second order parabolic equation with discontinuous data, in the proof of their Proposition 2. Coclite and Risebro give a similar definition for first order equations in Definition 1.1 of [12]. We now adapt the definition given in [11], noting that we interchange the role of sub- and supersolutions here to be consistent with our Definition 4.2.

Definition 4.27 (Viscosity solution for a PDE with discontinuous data)**.** Suppose that $G : \mathbb{R} \times \mathbb{R}^d \times S_d \to \mathbb{R}$ is an elliptic operator that is continuous in each of its arguments and $f : D \to \mathbb{R}$ is an upper semicontinuous function. Define $F : D \times \mathbb{R} \times \mathbb{R}^d \times S_d \to \mathbb{R}$ by

$$F(x, r, p, X) = G(r, p, X) - f(x),$$

and consider the PDE

$$F\left(x, u(x), Du(x), D^2u(x)\right) = 0. \tag{4.37}$$

We say that a function u is a *viscosity subsolution* of (4.37) if, for any smooth function $\phi \in C^\infty(D)$ and any point $x^0 \in \arg\max(u - \phi)$,

$$G\left(u(x^0), D\phi(x^0), D^2\phi(x^0)\right) - f(x^0) \leq 0.$$

Similarly, we say that a function u is a *viscosity supersolution* of (4.37) if, for any smooth function $\psi \in C^\infty(D)$ and any point $x^0 \in \arg\min(u - \psi)$,

$$G\left(u(x^0), D\psi(x^0), D^2\psi(x^0)\right) - f_\star(x^0) \geq 0,$$

where $f_\star$ is the lower semicontinuous envelope of f.

We define a *viscosity solution* of the PDE (4.37) to be a function u that is both a viscosity subsolution and a viscosity supersolution of (4.37).

Remark 4.28. In our definition of a subsolution, we could replace f with its upper semicontinuous envelope $f^\star$, since f is assumed to be upper semicontinuous.

Remark 4.29. The above definition coincides with parts (D.3) and (D.6) of Definition 1.1 of [12]. We note that Coclite and Risebro's definition in [12] has some additional conditions that we do not enforce here.

We can show that the value function v defined in Section 1.4.1 is a viscosity solution of the HJB equation (4.36), in the sense that we have just defined, without assuming continuity of the function f.

Proposition 4.30. *Suppose that Assumption 1.1 holds. Then the value function v defined in Section 1.4.1 is a viscosity solution of the HJB equation* (4.36)*, in the sense of Definition 4.27.*

The proof of this result is a straightforward adaptation of the proofs of Proposition 4.8 and Proposition 4.9. We do not provide the details here.

We now check directly that the value functions in Example 2.1 and Example 2.6 are viscosity solutions of the appropriate HJB equations in the sense of Definition 4.27. Let $d \geq 2$ and $R > 0$. Let $D = B_R(0) \subset \mathbb{R}^d$ and fix $\rho \in (0, R)$.

Proposition 4.31. *Define $f : D \to \mathbb{R}$ as in Example 2.1 by*

$$f(x) = \begin{cases} 0, & |x| \leq \rho, \\ -1, & |x| \in (\rho, R). \end{cases}$$

Then the function $\overline{v}$ defined in Proposition 2.5 is a viscosity solution of the HJB equation (4.36) *in the sense of Definition 4.27.*

Proof. The function $\overline{v} : D \to \mathbb{R}$ is given by

$$\overline{v}(x) = \begin{cases} \rho^2 - R^2, & |x| \leq \rho, \\ |x|^2 - R^2, & |x| \in (\rho, R), \end{cases}$$

and so the Hessian at any point $x \in D$ with $|x| \neq \rho$ is

$$D^2\overline{v}(x) = \begin{cases} 0, & |x| < \rho, \\ 2I, & |x| \in (\rho, R). \end{cases}$$

Suppose that $\phi \in C^\infty(D)$ is such that the function $\overline{v} - \phi$ has a local maximum at some point x^0 with $|x^0| = \rho$. Then the Hessian of ϕ satisfies

$$D^2\phi(x^0) \geq 2I,$$

and so

$$-\frac{1}{2} \inf_{\sigma \in U} \operatorname{Tr}\left(D^2\phi(x^0)\sigma\sigma^\top\right) - f(x^0) \leq -1 < 0.$$

It is straightforward to check that the same inequality holds at any other point $x \in D$, since f is continuous there. Therefore $\overline{v}$ is a viscosity subsolution of the HJB equation (4.36).

Now suppose that $\psi \in C^\infty(D)$ is such that the function $\overline{v} - \psi$ has a local minimum at some point x^0 with $|x^0| = \rho$. Then we must have

$$D^2\psi(x^0) \leq 0,$$

and so

$$-\frac{1}{2}\inf_{\sigma\in U} \operatorname{Tr}\left(D^2\psi(x^0)\sigma\sigma^\top\right) - f_\star(x^0) \geq -f_\star(x^0) = 1 > 0.$$

On checking that the same inequality holds at all other points in D, we see that $\overline{v}$ is a viscosity supersolution. We conclude that $\overline{v}$ is a viscosity solution of the HJB equation (4.36), in the sense of Definition 4.27. □

We now treat the second example in a similar way.

Proposition 4.32. *Define $f : D \to \mathbb{R}$ as in Example 2.6 by*

$$f(x) = \begin{cases} -1, & |x| < \rho, \\ 0, & |x| \in [\rho, R). \end{cases}$$

Then the function $\overline{v}$ defined in Proposition 2.8 is a viscosity solution of the HJB equation (4.36)*, in the sense of Definition 4.27.*

Proof. The function $\overline{v} : D \to \mathbb{R}$ is now given by

$$\overline{v}(x) = \begin{cases} |x|^2 + \rho^2 - 2\rho R, & |x| \leq \rho, \\ 2\rho\,|x| - 2\rho R, & |x| \in (\rho, R), \end{cases}$$

and so the Hessian at points $x \in D$ with $|x| \neq \rho$ is

$$D^2\overline{v}(x) = \begin{cases} 2I, & |x| < \rho, \\ 2\rho\,|x|^{-3}\left[|x|^2 I - xx^\top\right], & |x| \in (\rho, R). \end{cases}$$

Suppose that $\phi \in C^2(D)$ is such that the function $\overline{v} - \phi$ has a local maximum at some point x^0 with $|x^0| = \rho$. Then the Hessian of ϕ satisfies

$$D^2\phi(x^0) \geq 2I,$$

and so

$$-\frac{1}{2}\inf_{\sigma\in U} \operatorname{Tr}\left(D^2\phi(x^0)\sigma\sigma^\top\right) - f(x^0) \leq -1 < 0.$$

Now suppose that $\psi \in C^2(D)$ is such that the function $\overline{v} - \psi$ has a local minimum at some point x^0 with $|x^0| = \rho$. Then the Hessian of ψ satisfies

$$D^2\psi(x^0) \leq 2\rho^{-2}\left[\rho^2 I - x^0(x^0)^\top\right].$$

Note that, for any $\sigma \in U$, we have $\mathrm{Tr}\left(\left[\rho^2 I - x^0(x^0)^\top\right]\sigma\sigma^\top\right) \geq 0$ and, choosing

$$\sigma^\star = \frac{1}{|x^0|}\left[x^0; 0; \cdots; 0\right],$$

we have $\mathrm{Tr}\left(\left[\rho^2 I - x^0(x^0)^\top\right]\sigma^\star(\sigma^\star)^\top\right) = 0$. Therefore

$$\inf_{\sigma\in U} \mathrm{Tr}\left(D^2\psi(x^0)\sigma\sigma^\top\right) \leq 2\rho^{-2} \inf_{\sigma\in U} \mathrm{Tr}\left(\left[\rho^2 I - x^0(x^0)^\top\right]\sigma\sigma^\top\right) = 0.$$

Hence

$$-\frac{1}{2}\inf_{\sigma\in U} \mathrm{Tr}(D^2\psi(x^0)\sigma\sigma^\top) - f_\star(x^0) \geq 1 > 0.$$

Again, it is straightforward to check that the required inequalities hold for all other points $x \in D$, and so we conclude that $\overline{v}$ is a viscosity solution of the HJB equation (4.36), in the sense of Definition 4.27. □

We have now shown that, for Example 2.1 and Example 2.6, each candidate function defined in Proposition 2.5 and Proposition 2.8 is a viscosity solution of the HJB equation (4.36). By Proposition 4.30, we have that the value function for each example is also a viscosity solution of (4.36). We have already proved in Proposition 2.5 and Proposition 2.8 that, for each example, the candidate function is in fact equal to the value function. However, we are interested in whether we could prove this result via the HJB equation, as we did for continuous costs in Proposition 2.15.

Having shown that the value function and the candidate function are viscosity solutions of the HJB equation (4.36), we would require a uniqueness theory for the HJB equation in order to conclude directly that these functions are equal, as in Theorem 4.20. However we are not aware of any uniqueness theory in the literature that is applicable to the HJB equation (4.37) with discontinuous data.

We note that the usual proof of a comparison principle for viscosity solutions breaks down when f is allowed to have a discontinuity. In this case, the fifth statement of Assumption 4.10 may not hold and so we can no longer apply the Crandall-Ishii Lemma (Lemma 4.16) as we did in Lemma 4.17 to yield a comparison principle.

In [12], Coclite and Risebro prove a uniqueness result for first order PDEs with discontinuous data under some additional regularity on the differential operator. In future work, it would be of interest to investigate whether we can extend this result to prove a comparison principle viscosity solutions of the second order HJB equation (4.36).

CHAPTER 5

CONTROL PROBLEMS RELATED TO A MONGE-AMPÈRE EQUATION

In this final chapter, we consider two further stochastic control problems that are related to the problem defined in Section 1.4.1 that we have studied so far. We will show that the two problems that we introduce below are equivalent to each other and that the associated Hamilton-Jacobi-Bellman equation is a Monge-Ampère equation.

5.1 Introduction

Fix $d \geq 2$. For a domain $D \subset \mathbb{R}^d$ and a function $f : D \to (-\infty, 0]$, consider the Monge-Ampère equation

$$\begin{cases} -\det D^2 u + (-2f)^d = 0, \\ u \quad \text{convex}. \end{cases} \tag{5.1}$$

We will define viscosity solutions of this equation over the class of convex functions in Definition 5.15 and Definition 5.17 below. We will characterise such a solution as the value function of two equivalent control problems in Theorem 5.24 and Corollary 5.37. The first of these control problems is inspired by the work of Feng and Jensen in [24], who show the equivalence of the Monge-Ampère equation (5.1) to a Hamilton-Jacobi-Bellman equation. To our knowledge, the associated control problem has not been treated in the literature. In Proposition 5.11 we prove a dynamic programming principle for this control problem. We then use this to show that the value function is the unique viscosity solution of a Dirichlet problem for the Monge-Ampère equation (5.1) in Theorem 5.24.

The second equivalent problem is based on the work of Gaveau in [28]. Gaveau characterises the value function for this problem as a weak solution of the Monge-Ampère equation (5.1). Since the paper [28] came before the introduction of viscosity solutions in the 1983 paper [14], Gaveau uses a different notion of weak solution. In Corollary 5.37, we show that the value function is once again the unique viscosity solution of a Dirichlet problem for the Monge-Ampère equation (5.1).

In this chapter, we build on the work cited above to give a complete picture of the stochastic control problems and their characterisation in terms of viscosity solutions of the Monge-Ampère equation. As a consequence, we will deduce that these two control problems are equivalent.

Both of the problems defined here are related to the control problem that we defined in Section 1.4.1. In the first problem that we introduce in this chapter, we optimise over the same control set, but add an additional penalisation on the determinant of the diffusion matrix of the controlled process, favouring those controls that give a higher determinant. In the second problem, we keep the cost function the same, but change the control set to replace the constraint on the trace of the diffusion matrix with a constraint that its determinant is bounded from below. In the following sections, we will show that the value function defined in Section 1.4.1 is a lower bound for the value functions introduced in this chapter.

As noted above, the control problems introduced in this chapter involve the determinant of the diffusion matrix. We relate this to the trace by the following result, which we will refer to several times in this chapter. This well-known result is a simple consequence of the inequality of arithmetic and geometric means (AM-GM inequality). It is proved, for example, by Krylov in Lemma 1 of [39, Section 3.2].

Lemma 5.1. *For a symmetric positive semi-definite matrix $A \in \mathbb{R}^{d,d}$,*

$$\det(A)^{\frac{1}{d}} \leq \frac{1}{d} \operatorname{Tr}(A).$$

Proof. Recall that the AM-GM inequality states that, for $n \in \mathbb{N}$ and real numbers $x_1, x_2, \dots, x_n \geq 0$,

$$\left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Since A is positive semi-definite, its eigenvalues $\lambda_1, \dots, \lambda_d$ are all non-negative. We can write the determinant and the trace of the matrix A in terms of the eigenvalues as

$$\det(A) = \prod_{i=1}^{d} \lambda_i,$$

and

$$\operatorname{Tr}(A) = \sum_{i=1}^{d} \lambda_i.$$

Then we can conclude by the AM-GM inequality that

$$\det(A)^{\frac{1}{d}} = \left(\prod_{i=1}^{d} \lambda_i \right)^{\frac{1}{d}} \leq \frac{1}{d} \sum_{i=1}^{d} \lambda_i = \frac{1}{d} \operatorname{Tr}(A). \qquad \square$$

Throughout this chapter, the following assumptions will be in force.

Assumption 5.2. Suppose that

1. The domain D is bounded and convex;
2. The cost functions f and g are upper semicontinuous;
3. The running cost f is negative; i.e. $f \leq 0$;
4. The boundary cost g is bounded above; i.e. $g \leq K$ for some $K \geq 0$.

Before introducing the control problems, we make some remarks on solutions of Monge-Ampère equations and their application in optimal transport in Section 5.2.

In Section 5.3, we study the control problem that is associated to the HJB equation studied by Feng and Jensen in [24]. We show that the weak and strong value functions for this control problem are equal and bounded below by the value function v defined in Section 1.4.1. We prove convexity and continuity of the value function and show that a dynamic programming principle is satisfied. We then prove that the value function is a viscosity solution of the associated HJB equation with appropriate boundary condition. Using a comparison principle proved in [24], we deduce uniqueness of solutions of the Dirichlet problem for this HJB equation.

In Section 5.4, we define viscosity solutions over the set of convex test functions, as required for our study of the Monge-Ampère equation (5.1). We prove uniqueness of such solutions of (5.1) with Dirichlet boundary conditions, by using the equivalence with viscosity solutions of an HJB equation proved in [24] and the uniqueness result from the previous section. We also prove a comparison principle for viscosity solutions over convex test functions in Appendix A, following the remarks of Ishii and Lions in Section V.3 of [36]. This comparison principle gives an alternative proof of uniqueness for the Monge-Ampère equation. We conclude that the value function from the previous section is the unique solution of a boundary value problem for the Monge-Ampère equation (5.1).

In Section 5.5, we consider the control problem studied by Gaveau in [28]. Once again, we show that weak and strong formulations of the problem are equivalent and that the value function is bounded below by the value function v defined in Section 1.4.1. We show that the associated HJB equation is equivalent to the Monge-Ampère equation (5.1) in the sense of viscosity solutions over convex test functions. By using the dynamic programming principle and continuity properties that are proved in [28], we characterise the value function as the unique convex viscosity solution of the Monge-Ampère equation with appropriate boundary condition. In Section 5.5.3 we give an alternative proof of attainment of the boundary condition under weaker convexity conditions on the domain.

Finally, in Section 5.6, we discuss the relationship between the three control problems that we have studied in this thesis. We see that the two control problems introduced in this chapter are equivalent, by the characterisation of the value functions in terms of convex viscosity solutions of the Monge-Ampère equation and the uniqueness of such solutions. For the examples of step cost functions in Section 2.2, we show that the value functions are in fact equal for all three control problems, by approximating the optimal strategies identified in Section 2.2. For a continuously differentiable monotone cost function, however, we show that the value function v is a strict lower bound for the two value functions defined in this chapter.

5.2 Remarks on Monge-Ampère equations

A Monge-Ampère equation, as defined, for example, in equation (5.15) of [36] and equation (4.5) of [63], is a second order fully nonlinear PDE of the form

$$\begin{cases} \det(D^2u(x)) = f(x, u(x), Du(x)), & x \in D, \\ u \quad \text{convex}, & \text{in} \quad D, \end{cases}$$

for some domain $D \subset \mathbb{R}^d$ and some function $f : D \times \mathbb{R} \times \mathbb{R}^d \to (-\infty, 0]$.

Several notions of weak solution of Monge-Ampère equations have been introduced in the literature, as described in the book [31] of Gutiérrez and in Section 4.1.4 of [63]. For example, the equivalent notions of Aleksandrov solutions and viscosity solutions of the Monge-Ampère equation are defined in Chapter 1 of [31].

The convexity constraint on the solution of the Monge-Ampère equation is required in order to ensure that the equation is degenerate elliptic, as defined in the second statement of Assumption 4.1. This is one of the conditions that we used to give sense to the definition of viscosity solutions in Definition 4.2.

Here we restrict ourselves to the Monge-Ampère equation of the form (5.1) for some $f \le 0$. We now verify that the Monge-Ampère equation is degenerate elliptic on the set of convex functions.

Lemma 5.3. *Problem* (5.1) *is a degenerate elliptic PDE problem.*

To prove this result, we will use the following lemma, which is a consequence of the Minkowski determinant inequality, as stated in [66].

Lemma 5.4. *Fix $n \in \mathbb{N}$ and let $A, B \in \mathbb{R}^{n,n}$ be symmetric positive semi-definite matrices. Then*

$$\det(A + B) \ge \det(A) + \det(B).$$

Proof. Since A and B are symmetric positive semi-definite, their eigenvalues are all non-negative reals, so $\det(A), \det(B) \ge 0$. We also have that $A + B$ is symmetric positive semi-definite and $\det(A + B) \ge 0$.

The Minkowski determinant inequality, as stated in equation (1.1) of [66], states that

$$\det(A + B)^{\frac{1}{n}} \ge \det(A)^{\frac{1}{n}} + \det(B)^{\frac{1}{n}}.$$

Since all terms in the inequality are non-negative, we can take the n^{th} power to get

$$\det(A + B) \ge \left(\det(A)^{\frac{1}{n}} + \det(B)^{\frac{1}{n}}\right)^n \ge \det(A) + \det(B),$$

as required. □

Proof of Lemma 5.3. Suppose that u is a classical solution of (5.1). Since u is convex, we have that $D^2u \ge 0$. Now let $B \ge 0$. Then

$$-\det(D^2u + B) \le -\det(D^2u) - \det(B),$$

by Lemma 5.4. Since $B \ge 0$, we know that $\det(B) \ge 0$. So

$$-\det(D^2u + B) \le -\det(D^2u).$$

Hence, since f does not depend on the Hessian matrix, (5.1) is a degenerate elliptic problem. □

We now have a degenerate elliptic problem, so one of the conditions needed to use the theory of viscosity solutions is satisfied. However, as noted in [36], our Definitions 4.2 and 4.4 of viscosity solutions cannot be applied directly to a problem of the form (5.1). Since the equation here is only elliptic for convex functions, it only makes sense to consider convex test functions in Definition 4.2, and positive semi-definite matrices in the semijets of Definition 4.4. We therefore need to adapt our definitions to suit the Monge-Ampère equation, in the same manner as in Chapter 1 of [31] and Section V.3 of [36]. In Section 5.4 below, we introduce viscosity solutions over the class of convex test functions for any PDE that is degenerate elliptic when restricted to convex solutions.

5.2.1 Monge-Ampère equations in optimal transport

In [10], Brenier showed that the following form of the Monge-Ampère equation arises in optimal transport:

$$\det(D^2u(x)) = \frac{f(x)}{g(Du(x))}, \tag{5.2}$$

where $f, g : \mathbb{R}^d \to \mathbb{R}$. More precisely, take two measures μ_0 and μ_1 on $\mathbb{R}^d$ that are both absolutely continuous with respect to Lebesgue measure, and write $\mu_0(\mathrm{d}x) = f(x)\,\mathrm{d}x$ and $\mu_1(\mathrm{d}x) = g(x)\,\mathrm{d}x$. Brenier proved that there exists a convex function $u : \mathbb{R}^d \to \mathbb{R}$ such that Du is the unique optimal map for the optimal transport problem of finding

$$\min_{\substack{T:\mathbb{R}^d\to\mathbb{R}^d\\ T_{\#}\mu_0=\mu_1}} \int_{\mathbb{R}^d} |x - T(x)|^2\,\mu_0(\mathrm{d}x),$$

which is equal to

$$\inf_{\substack{\pi\\ \int_{\mathbb{R}^d}\pi(\cdot,\mathrm{d}y)=\mu_0,\\ \int_{\mathbb{R}^d}\pi(\mathrm{d}x,\cdot)=\mu_1}} \int_{\mathbb{R}^d} |x-y|^2\,\pi(\mathrm{d}x,\mathrm{d}y).$$

In the case that μ_1 is uniform on D, the Monge-Ampère equation (5.2) becomes

$$\det(D^2u(x)) = |D|\,f(x),$$

which has the same form as (5.1).

The Brenier solution of the Monge-Ampère problem, which is related to optimal transport, is described in Section 3.2 of [18] and Section 4.1.4 of [63]. This notion of solution requires that Du maps the support of μ_0 onto the support of μ_1, and no further boundary conditions are imposed. In this chapter, we will instead consider the Monge-Ampère equation (5.1) with Dirichlet boundary conditions, which are appropriate for the control problems that we will investigate.

Recall from Section 1.1.1 that we partly motivated our study of the control problem defined in Section 1.4.1 by discussing martingale optimal transport, a variation of the classical optimal transport problem. It is notable that the related control problems introduced in this chapter are associated to a Monge-Ampère equation, which plays a role in classical optimal transport. It would be of interest to explore the connections between these optimal control problems and the two variants of optimal transport more fully, but we do not pursue that direction further in this thesis.

5.3 A stochastic control problem inspired by Feng and Jensen

The first control problem that we introduce in this chapter is inspired by the work of Feng and Jensen in their paper [24] on numerical methods for Monge-Ampère equations. In [24], the authors show that the Monge-Ampère equation (5.1) has an equivalent formulation as the HJB equation

$$- \inf_{\sigma \in U} \left\{ \frac{1}{2} \operatorname{Tr}\left(D^2 u \sigma \sigma^\top\right) + d\, f \det(\sigma \sigma^\top)^{\frac{1}{d}} \right\} = 0, \tag{5.3}$$

where $U := \{\sigma \in \mathbb{R}^{;d} :\ \operatorname{Tr}(\sigma\sigma^\top)\}$ is unchanged from the definition in Section 1.4.1. The equivalence for classical solutions can be found in Krylov's 1987 book [39] as Lemma 2 of Section 3.2, and this equivalence is described in detail in Chapter 2, §2 of the report [56] of Smears. Feng and Jensen were the first to show in [24] that the same equivalence holds for viscosity solutions. The reason for introducing the HJB formulation of the Monge-Ampère equation in [24] is that the convexity constraint in (5.1) complicates numerical methods, whereas the type of semi-Lagrangian methods presented in [24] are well-know for HJB equations.

In the context of this thesis, we expect the HJB equation (5.3) to be associated to the following stochastic control problem, which has not to our knowledge been studied in the literature. We will prove that the value function for this control problem has a characterisation in terms of the HJB equation (5.3) in the following sections. This characterisation leads to a new stochastic representation result for viscosity solutions of the Monge-Ampère equation (5.1) in Corollary 5.37.

As in Section 1.4.1, we define a strong and a weak control problem, again taking the definitions from [58] and [20], respectively.

Strong Formulation

The strong formulation of the control problem is to find the strong value function $v^S_{\mathrm{FJ}} : D \to \mathbb{R}$, defined as follows.

Let $(\Omega_0, \mathcal{F}, \mathbb{P}_0)$ be a probability space on which a d-dimensional Brownian motion B is defined with natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$.

Control: Define the set of controls

$$\mathcal{U} := \{U\text{-valued } \mathbb{F}\text{-progressively measurable processes}\}.$$

Dynamics: For any $x \in D$ and $\nu = (\nu_t)_{t\geq 0} \in \mathcal{U}$, define X^ν by the stochastic integral

$$X_t^\nu = x + \int_0^t \nu_s \,\mathrm{d}B_s, \quad t \geq 0,$$

and define the associated exit time from the domain by

$$\tau := \inf\{t \geq 0\colon X_t^\nu \notin D\}.$$

Value function: We define the *strong value function* $v_{\mathrm{FJ}}^S : D \to \mathbb{R}$ by

$$v_{\mathrm{FJ}}^S(x) := \inf_{\nu\in\mathcal{U}} \mathbb{E}^x \left[d \int_0^\tau f(X_s^\nu) \det(\sigma_s\sigma_s^\top)^{\frac{1}{d}} \,\mathrm{d}s + g(X_\tau^\nu)\right].$$

Remark 5.5. Note that the infimum in the definition of the value function v_{FJ}^S is taken over the same class of controls as in the definition of the value function v^S in (1.4). Here we have an additional term in the cost function that penalises controls σ that have a small determinant.

For example, consider a control $\sigma \in \mathcal{U}$ that can be written as

$$\sigma_t = \begin{bmatrix}\overline{\sigma}_t; & 0; & \dots; & 0\end{bmatrix}, \quad t \geq 0,$$

for some $\overline{\sigma} \in \mathbb{R}^d$. We say that σ is a *degenerate control*, since $\det(\sigma_t) = 0$, for all $t \geq 0$. A process following a degenerate control has zero running cost and therefore maximises the expected running cost for $f \leq 0$. Hence we no longer expect the degenerate optimal controls found in Chapter 2 for radially symmetric costs to be optimal here.

Weak Formulation

The weak formulation of the control problem is to find the weak value function $v_{\mathrm{FJ}}^W : D \to \mathbb{R}$, which we define below, following [20] as in Section 1.4.1.

Define the space of continuous paths $\Omega := C([0,\infty), \mathbb{R}^d)$ and denote the set of Borel measurable functions $\nu : \mathbb{R}_+ \to U$ by $\mathcal{B}(\mathbb{R}_+, U)$. Then set $\overline{\Omega} = \Omega \times \mathcal{B}(\mathbb{R}_+, U)$

and denote an element of $\overline{\Omega}$ by $\overline{\omega} = (\omega, u)$. Define the canonical process $\overline{X} = (X, \nu)$ on $\overline{\Omega}$ by $X_t(\overline{\omega}) = \omega_t$, for each $t \geq 0$, and $\nu(\overline{\omega}) = u$. For $\phi \in C_b(\mathbb{R}_+ \times U)$, $s \geq 0$, define

$$M_s(\phi) := \int_0^s \phi(r, \nu_r) \, \mathrm{d}r.$$

Then define the canonical filtration $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t \geq 0}$ by

$$\overline{\mathcal{F}}_t := \sigma \left\{ (X_s, M_s(\phi)) \colon \phi \in C_b(\mathbb{R}_+ \times U), s \leq t \right\}, \quad t \geq 0.$$

Control: Let $\mathbb{M}$ be the set of probability measures on the set $\Omega \times \mathcal{B}(\mathbb{R}_+, U)$. For each $x \in D$, let

$$\mathbb{M}_x = \left\{ \mathbb{P} \in \mathbb{M} \colon \mathbb{P}(X_0 = x) = 1 \right\}.$$

Dynamics: Define

$$\begin{aligned} \mathcal{P}_x := \{ \mathbb{P} \in \mathbb{M}_x : \quad & t \mapsto \phi(X_t) - \phi(X_0) - \frac{1}{2} \int_0^t \operatorname{Tr}\left(D^2 \phi(X_s) \nu_s \nu_s^\top \right) \mathrm{d}s \\ & \text{is a } (\overline{\mathbb{F}}, \mathbb{P})\text{-local martingale for all } \phi \in C^2(\mathbb{R}^d) \}, \end{aligned}$$

and let $\tau = \inf \{ t \geq 0 \colon X_t \notin D \}$.

Value function: We define the *weak value function* $v_{\mathrm{FJ}}^W : D \to \mathbb{R}$ by

$$v_{\mathrm{FJ}}^W(x) = \inf_{\mathbb{P} \in \mathcal{P}_x} \mathbb{E}^{\mathbb{P}} \left[d \int_0^\tau f(X_s) \det(\sigma_s \sigma_s^\top)^{\frac{1}{d}} \, \mathrm{d}s + g(X_\tau) \right].$$

Analogously to Proposition 1.7, we can show that the weak and strong control problems are equivalent.

Proposition 5.6. *Under Assumption 5.2, we have the equality $v_{\mathrm{FJ}}^S = v_{\mathrm{FJ}}^W$ in D.*

Proof. As in the proof of Proposition 1.7, we refer to Theorem 4.5 of [20]. Fix $x \in D$ and define the function $\Phi : \Omega \to \mathbb{R}$ by

$$\Phi(\omega) := d \int_0^{\tau(\omega)} f(X_s(\omega)) \det(\sigma_s \sigma_s^\top)^{\frac{1}{d}} \, \mathrm{d}s + g(X_{\tau(\omega)}(\omega)), \quad \omega \in \Omega.$$

By Theorem 4.5 of [20], it is sufficient to show that Φ is upper semicontinuous and bounded above by some random variable ξ that is uniformly integrable under the family of probability measures $\mathcal{P}_x$.

For any $\sigma \in U$, we can apply Lemma 5.1 to the symmetric positive semi-definite matrix $\sigma\sigma^\top$, to get the bound

$$d \det(\sigma\sigma^\top)^{\frac{1}{d}} \leq \mathrm{Tr}(\sigma\sigma^\top) = 1.$$

Combining this with the bounds $f \leq 0$ and $g \leq K$ from Assumption 5.2, we have the constant bound

$$\Phi(\omega) \leq K < \infty,$$

and so the uniformly integrability condition is satisfied. Moreover, since f and g are upper semicontinuous and the determinant is a continuous function, $\omega \mapsto \Phi(\omega)$ is upper semicontinuous.

Hence, by Theorem 4.5 of [20], we conclude that $v^S_{\mathrm{FJ}}(x) = v^W_{\mathrm{FJ}}(x)$ for all $x \in D$. $\square$

Given this result, we will write the common value as $v_{\mathrm{FJ}} = v^S_{\mathrm{FJ}} = v^W_{\mathrm{FJ}}$ and refer to v_{FJ} as the *value function.*

We now show that the value function v_{FJ} is bounded below by v, the value function defined in Section 1.4.1.

Proposition 5.7. *Suppose that Assumption 5.2 holds. Then $v(x) \leq v_{\mathrm{FJ}}(x)$, for all $x \in D$.*

Proof. Let $x \in D$ and $\sigma \in \mathcal{U}$. Then, for all $t \geq 0$, $\mathrm{Tr}(\sigma_t\sigma_t^\top) = 1$. By Lemma 5.1, we can bound the determinant by

$$\det(\sigma_t\sigma_t^\top)^{\frac{1}{d}} \leq \frac{1}{d}\mathrm{Tr}(\sigma_t\sigma_t^\top) = \frac{1}{d}.$$

Since $f \leq 0$, we then have

$$\mathbb{E}^x\left[d\int_0^\tau f(X^\sigma_s)\det(\sigma_s\sigma_s^\top)^{\frac{1}{d}}\,\mathrm{d}s + g(X^\sigma_\tau)\right] \geq \mathbb{E}^x\left[\int_0^\tau f(X^\sigma_s)\,\mathrm{d}s + g(X^\sigma_\tau)\right].$$

Taking the infimum over $\sigma \in \mathcal{U}$, it follows that

$$v_{FJ}(x) \geq v(x),$$

as required. $\square$

We will now prove convexity and continuity of the value function v_{FJ}, and show that v_{FJ} satisfies a dynamic programming principle. From this dynamic programming principle, it will follow that v_{FJ} is a viscosity solution of the HJB equation (5.3).

5.3.1 Dynamic programming principle

We showed in Lemma 1.11 that the value function v is convex when the cost function f is negative. We will now prove the same result for v_{FJ}, noting that we do not actually require negativity of f in the proof.

Lemma 5.8. *Suppose that Assumption 5.2 holds and that the domain D is strictly convex. Then the value function v_{FJ} is convex.*

Proof. We follow the same strategy of proof as for the proof of semiconvexity of v in Lemma 1.11, omitting many of the details here.

Let $x_0, x_1 \in D$ and fix $\lambda \in (0,1)$. Consider a martingale starting from a point $y := \lambda x_0 + (1-\lambda)x_1 \in D$, which lies on the line connecting the points x_0 and x_1. Define the control $\sigma^\star \in \mathcal{U}$ in the same way as in the proof of Lemma 1.11. In particular, for $t \le H_{x_0,x_1}$, the first hitting time of either x_0 or x_1, we have the constant degenerate control

$$\sigma_t^\star = \frac{1}{|x_1 - x_0|}\Big[x_1 - x_0; \quad 0; \quad \dots; \quad 0\Big],$$

which constrains the controlled process to the line connecting the points x_0 and x_1. Therefore $\det(\sigma_t^\star(\sigma_t^\star)^\top) = 0$, for $t \le H_{x_0,x_1}$.

Recall that, from time H_{x_0,x_1} onwards, $\sigma^\star$ coincides with one of two ε-optimal strategies, which we once again denote $\sigma^{0,\varepsilon}$ and $\sigma^{1,\varepsilon}$, as in the proof of Lemma 1.11. We then calculate that

$$\begin{aligned}
v_{FJ}(y) &\le \mathbb{E}^x\left[d\int_0^\tau f(X_s^{\sigma^\star})\det(\sigma_s^\star\sigma_s^{\star\top})^{\frac{1}{d}}\,\mathrm{d}s + g(X_\tau^{\sigma^\star})\right] \\
&= \mathbb{E}^{x_0}\left[d\int_0^\tau f(X_s^{\sigma^{0,\varepsilon}})\det(\sigma_s^{0,\varepsilon}\sigma_s^{0,\varepsilon\top})^{\frac{1}{d}}\,\mathrm{d}s + g(X_\tau^{\sigma^{0,\varepsilon}})\right]\mathbb{P}^y[H_{x_0} < H_{x_1}] \\
&\quad + \mathbb{E}^{x_1}\left[d\int_0^\tau f(X_s^{\sigma^{1,\varepsilon}})\det(\sigma_s^{1,\varepsilon}\sigma_s^{1,\varepsilon\top})^{\frac{1}{d}}\,\mathrm{d}s + g(X_\tau^{\sigma^{1,\varepsilon}})\right]\mathbb{P}^y[H_{x_1} < H_{x_0}] \\
&< \lambda v_{FJ}(x_0) + (1-\lambda)v_{FJ}(x_1) + 2\varepsilon.
\end{aligned}$$

Hence

$$v_{FJ}(y) \le \lambda v_{FJ}(x_0) + (1-\lambda)v_{FJ}(x_1),$$

and so v_{FJ} is convex, as required. □

To prove continuity of v_{FJ}, we strengthen Assumption 5.2 as follows.

Assumption 5.9. Suppose that Assumption 5.2 holds and, moreover, the domain D is strictly convex and, for any $x \in D$, $v_{\mathrm{FJ}}(x) > -\infty$.

As in Corollary 1.13, under the additional conditions of Assumption 5.9, we can deduce from the convexity result of Lemma 5.8 that the value function v_{FJ} is locally Lipschitz.

Corollary 5.10. *Suppose that Assumption 5.9 holds. Then v_{FJ} is locally Lipschitz in D.*

Proof. Local Lipschitz continuity follows directly from Lemma 5.8 by Theorem 10.4 of [52]. □

We now prove a dynamic programming principle for the value function v_{FJ}. Again, the proof follows the same strategy as the proof of Proposition 1.17, and we omit the details.

Proposition 5.11. *Suppose that Assumption 5.9 is satisfied. Then the following dynamic programming principle holds. For any $x \in D$ and for any stopping time θ such that $\theta \in [0, \tau]$ almost surely,*

$$v_{FJ}(x) = \inf_{\nu \in \mathcal{U}} \mathbb{E}^x \left[d \int_0^\theta f(X_s^\nu) \det(\nu_s \nu_s^\top)^{\frac{1}{d}} \, \mathrm{d}s + v_{FJ}(X_\theta^\nu) \right].$$

Proof. First note that, by Lemma 5.1, for any $\nu \in \mathcal{U}$,

$$0 \leq d \det(\nu_t \nu_t^\top)^{\frac{1}{d}} \leq \operatorname{Tr}(\nu_t \nu_t^\top) = 1, \tag{5.4}$$

for all $t \geq 0$. We follow the same method of proof as for Proposition 1.17, replacing the running cost $\int_0^t f(X_s^\nu) \, \mathrm{d}s$ with

$$d \int_0^t f(X_s^\nu) \det(\nu_s \nu_s^\top)^{\frac{1}{d}} \, \mathrm{d}s.$$

The bound (5.4) ensures that all of the expectations in the proof are still well-defined.

By Corollary 5.10, we have that v_{FJ} is continuous, and so we can make the same measurable selection argument as in the proof of Proposition 1.17.

As noted in Remark 5.5, any degenerate control has zero determinant. This simplifies the bound (1.11), as we do not require estimates on the running cost f or the expectation of the exit time τ.

We can follow the same arguments as in the proof of Proposition 1.17 to complete the proof. □

In the following section, we will use the dynamic programming principle to show that the value function v_{FJ} satisfies the HJB equation (5.3) in the viscosity sense.

5.3.2 Viscosity solution characterisation

In this section we will provide a characterisation of the value function as the unique viscosity solution of the Dirichlet problem

$$\begin{cases} -\inf_{\sigma\in U}\left\{\frac{1}{2}\operatorname{Tr}\left(D^2 u\sigma\sigma^\top\right) + d\,f\det(\sigma\sigma^\top)^{\frac{1}{d}}\right\} = 0, & \text{in} \quad D, \\ u = g, & \text{on} \quad \partial D, \end{cases} \tag{5.5}$$

recalling the definition of such a solution given in Definition 4.13.

As a consequence of the dynamic programming principle, we can show that, when the cost function f is continuous, the value function v_{FJ} is a viscosity solution of the HJB equation (5.3).

Proposition 5.12. *Suppose that Assumption 5.9 holds and that $f : D \to \mathbb{R}$ is continuous. Then v_{FJ} is a viscosity solution of the HJB equation* (5.3)*.*

The proof of this result proceeds exactly as the proofs of Proposition 4.8 and Proposition 4.9. Note that the bound on the determinant given by Lemma 5.1 ensures that all expectations in the proof are well-defined and that we can make the same arguments using Itô's formula. We do not give the details of the proof here.

We now check that the value function v_{FJ} attains the boundary condition g on ∂D, again following the same strategy of proof as for the value function v in Section 4.5. For this result, we require the domain to be uniformly convex, as defined in Definition 4.22.

Proposition 5.13. *Suppose that Assumption 5.9 is satisfied and, moreover, the domain D is uniformly convex, the running cost f is continuous in D, and the boundary cost g is uniformly continuous on ∂D.*

Then v_{FJ} extends continuously to $\overline{D}$ with

$$\lim_{x\to x_0} v_{\mathrm{FJ}}(x) = g(x_0),$$

for any $x_0 \in \partial D$.

Proof. From Corollary 5.10, we have that v_{FJ} is continuous in D. It remains to show that $\lim_{x\to x_0} v_{\mathrm{FJ}} = g(x_0)$ for any $x_0 \in \partial D$.

To prove this result, we will once again make use of the bound

$$d\det(\sigma\sigma^\top)^{\frac{1}{d}} \leq 1, \tag{5.4}$$

for any $\sigma \in U$, which follows from Lemma 5.1. Following the same method of proof as in Proposition 4.21, we can show that $\limsup_{x\to x_0} v_{\mathrm{FJ}}(x) \leq g(x_0)$, for any

$x_0 \in \partial D$. The bound (5.4) guarantees that the expectations are well-defined, and the fact that degenerate controls have zero determinant simplifies the proof.

We now note that the results of Lemma 4.24 and Lemma 4.26 still hold under the present assumptions. Combining these with the bound (5.4), we can follow the proof of Proposition 4.23 to find that $\liminf_{x \to x_0} v_{\mathrm{FJ}}(x) \geq g(x_0)$, for any $x_0 \in \partial D$. We conclude that $\lim_{x \to x_0} v_{\mathrm{FJ}}(x) = g(x_0)$, for any $x_0 \in \partial D$. □

We have now shown that v_{FJ} is a viscosity solution of the Dirichlet problem (5.5). In [24], Feng and Jensen prove a comparison principle for the HJB equation (5.3). This leads to the following characterisation of the value function.

Theorem 5.14. *Suppose that Assumption 5.9 holds and, moreover, the domain D is uniformly convex, the running cost $f : D \to \mathbb{R}$ is continuous, and the boundary cost $g : \partial D \to \mathbb{R}$ is uniformly continuous.*

Then v_{FJ} is the unique viscosity solution of the Dirichlet problem (5.5)*.*

Proof. In Lemma 3.6 of [24], the authors prove the following comparison principle for the HJB equation (5.3). Suppose that $u_1 : \overline{D} \to \mathbb{R}$ is a viscosity subsolution of (5.3), $u_2 : \overline{D} \to \mathbb{R}$ is a viscosity supersolution of (5.3), and that $u_1 \leq u_2$ on ∂D. Then $u_1 \leq u_2$ on $\overline{D}$. As in the proof of Corollary 4.14, we can deduce that any solution of the Dirichlet problem (5.5) must be unique.

By Proposition 5.12, we have that v_{FJ} is a viscosity solution of the HJB equation (5.3) in D. By Proposition 5.13, v_{FJ} extends continuously to $\overline{D}$ and attains the boundary condition g on ∂D. Hence v_{FJ} is a viscosity solution of the Dirichlet problem (5.5).

We conclude that v_{FJ} is the unique viscosity solution of (5.5). □

By the equivalence of viscosity solutions of the PDEs (5.1) and (5.3) that is proved in [24], we will deduce that v_{FJ} is the unique viscosity solution of a Dirichlet problem for the Monge-Ampère equation (5.1). For the Monge-Ampère equation, we require the notion of convex viscosity solutions that we define in the following section.

5.4 Convex viscosity solutions

We now define viscosity solutions over the set of convex test functions, following Section V.3 of [36].

Let $D \subset \mathbb{R}^d$ be a bounded convex domain. Consider a continuous differential operator $F : D \times \mathbb{R} \times \mathbb{R}^d \times S_d \to \mathbb{R}$ that is degenerate elliptic on the set of non-negative definite matrices; i.e.

$$F(x, r, p, X) \leq F(x, r, p, Y), \quad \text{for} \quad X, Y \geq 0, \quad \text{and} \quad X \geq Y. \tag{5.6}$$

Then we wish to define viscosity solutions of the following problem:

$$\begin{cases} F\left(x, u(x), Du(x), D^2u(x)\right) = 0 & \text{in} \quad D, \\ u \quad \text{is convex} & \text{in} \quad D, \\ u = g & \text{on} \quad \partial D. \end{cases} \tag{5.7}$$

The following definition is standard in the case of Monge-Ampère equations, as found in Definition 1.3.1 of [31], Section V.3 of [36], and Section 4.1.4 of [63]. We take the same definition for any operator F that is degenerate elliptic on the set of non-negative definite matrices. We first define solutions of the PDE

$$\begin{cases} F\left(x, u(x), Du(x), D^2u(x)\right) = 0 & \text{in} \quad D, \\ u \quad \text{is convex} & \text{in} \quad D, \end{cases} \tag{5.8}$$

before considering the boundary conditions.

Definition 5.15 (Convex viscosity solution I)**.** We say that an upper semicontinuous convex function $u : D \to \mathbb{R}$ is a *convex viscosity subsolution* of (5.8) if, for every smooth convex $\phi \in C^\infty(D)$,

$$F\left(x^0, u(x^0), D\phi(x^0), D^2\phi(x^0)\right) \leq 0,$$

at every point $x^0 \in D$ that is a local maximum of $u - \phi$.

Similarly, a lower semicontinuous convex function $u : D \to \mathbb{R}$ is a *convex viscosity supersolution* of (5.8) if, for every smooth convex $\psi \in C^\infty(D)$,

$$F\left(x^0, u(x^0), D\psi(x^0), D^2\psi(x^0)\right) \geq 0,$$

at every point $x^0 \in D$ that is a local minimum of $u - \psi$.

A continuous convex function u that is both a viscosity subsolution and a viscosity supersolution of (5.8) is a *convex viscosity solution*.

Remark 5.16. We note that we do not, in fact, alter the definition of viscosity subsolutions for the convex case. As remarked in Section 1.3 of [31], any test function approximating a convex function from above, in the sense of Definition 4.2, must itself be convex.

The requirement for the test function to be convex in the definition of a supersolution, however, is a restriction on the class of test functions. This means that we weaken the standard definition of viscosity solution. We will see in Theorem A.3 and Proposition A.5 of Appendix A that we can still obtain a comparison principle for convex viscosity solutions.

For convenience in proving the comparison principles in Appendix A, we give a reformulation of our definition, analogous to Definition 4.4, in terms of the semijets defined in Definition 4.3.

Definition 5.17 (Convex viscosity solution II)**.** An upper semicontinuous convex function $u : D \to \mathbb{R}$ is a *convex viscosity subsolution* of (5.8) if

$$F(x, u(x), p, X) \leq 0 \quad \text{for all} \quad x \in D,\, (p, X) \in J_D^{2,+}u(x) \quad \text{such that} \quad X \geq 0.$$

A lower semicontinuous convex function $u : D \to \mathbb{R}$ is a *convex viscosity supersolution* of (5.8) if

$$F(x, u(x), p, X) \geq 0 \quad \text{for all} \quad x \in D,\, (p, X) \in J_D^{2,-}u(x) \quad \text{such that} \quad X \geq 0.$$

A *convex viscosity solution* of (5.8) is a continuous convex function $u : D \to \mathbb{R}$ that is both a viscosity subsolution and a viscosity supersolution.

Remark 5.18. For F continuous in each of its arguments, the semijets in the above definition can equivalently be replaced by their closures, as is the case in Definition 4.4, since the set of non-negative definite matrices is closed.

Remark 5.19. Similarly to the previous definition, we did not need to include the requirement that X is non-negative definite in the definition of subsolution, as this is an immediate consequence of (p, X) belonging to the superjet of a convex function u, for some $p \in \mathbb{R}^d$.

Having defined convex viscosity solutions in two different ways, we now need to verify that the two definitions are equivalent.

Proposition 5.20. *Definition 5.15 and Definition 5.17 are equivalent.*

Proof. This result follows from a straightforward adaptation of Lemma 4.6.

Suppose that ψ is a convex test function with $x^0 \in \arg\min(u - \psi)$. Then

$$(p, X) := \big(D\psi(x^0), D^2\psi(x^0)\big) \in J_D^{2,-}u(x^0).$$

By convexity of ψ, we have $X \geq 0$. Therefore, if a statement is true for every $x \in D$ and $(p, X) \in J_D^{2,-}u(x)$ such that $X \geq 0$, then it is also true for every smooth convex ψ such that $x^0 \in \arg\min(u - \psi)$.

On the other hand, if $(p, X) \in J_D^{2,-}u(x^0)$, with $X \geq 0$, Lemma 4.6 allows us to construct a smooth function ψ such that $x^0 \in \arg\min(u - \psi)$, $D\psi(x^0) = p$, and $D^2\psi(x^0) = X$. Since X is non-negative definite, ψ is convex at x^0. Therefore any statement that is true for every smooth convex ψ such that $x^0 \in \arg\min(u - \psi)$ is also true for all $x \in D$ and $(p, X) \in J_D^{2,-}u(x)$.

This gives equivalence of the definitions of convex viscosity supersolutions in Definition 5.15 and Definition 5.17. By Remark 5.16 and Remark 5.19, the result for convex viscosity subsolutions follows directly from Lemma 4.6. □

Definition 5.21. We say that a convex function $u : \overline{D} \to \mathbb{R}$ is a viscosity solution of the Dirichlet problem (5.7) if u is a convex viscosity solution of (5.8) in D, in the sense of Definition 5.15 (or equivalently Definition 5.17), and $u(x) = g(x)$ for all $x \in \partial D$.

5.4.1 A Monge-Ampère equation as an HJB equation

We now have an appropriate definition of a viscosity solution of the Monge-Ampère equation (5.1). In this section, we will characterise the value function v_{FJ} from Section 5.3 as the unique convex viscosity solution of a boundary value problem for this Monge-Ampère equation. We first address the question of uniqueness.

Here we make use of the equivalence between convex viscosity solutions of the Monge-Ampère equation (5.1) and viscosity solutions of the HJB equation (5.3) that Feng and Jensen prove in [24]. Using the comparison principle from [24] for the HJB equation (5.3), as in Theorem 5.14, we deduce uniqueness of solutions of a Dirichlet problem for the Monge-Ampère equation.

Proposition 5.22. *Suppose that Assumption 5.2 holds and f is continuous in D. Then there is at most one convex viscosity solution of the Monge-Ampère problem*

$$\begin{cases} -\det(D^2u) + (-2f)^d = 0, & \text{in} \quad D, \\ u \quad \text{convex}, & \text{in} \quad D, \\ u = g, & \text{on} \quad \partial D. \end{cases} \tag{5.9}$$

Proof. In Theorem 3.3 and Theorem 3.5 of [24], Feng and Jensen show that the set of viscosity solutions of the Dirichlet problem (5.5) is equal to the set of convex viscosity solutions of the Dirichlet problem (5.9).

Lemma 3.6 of [24] gives a comparison principle for the HJB equation (5.3). This leads to uniqueness of viscosity solutions of the Dirichlet problem (5.5), as remarked in Theorem 5.14.

Therefore we have uniqueness of convex viscosity solutions of the Monge-Ampère problem (5.9). □

Remark 5.23. We note that there are other methods of proving uniqueness for the Monge-Ampère problem (5.9). We consider two of these approaches here.

1. In [31], Gutiérrez proves uniqueness of viscosity solutions of the Dirichlet problem (5.9) by using an equivalence to Aleksandrov solutions of Monge-Ampère equations. Gutiérrez proves a comparison principle for Aleksandrov solutions in Theorem 1.4.6 of [31] and shows that Aleksandrov solutions are equivalent to convex viscosity solutions in Propositions 1.3.4 and 1.7.1.

2. In [36], Ishii and Lions state a comparison principle for convex viscosity solutions of the Monge-Ampère equation in Theorem V.2, which implies uniqueness for the Dirichlet problem (5.9). While the main ideas of the proof of this result are given in Section V.3 of [36], the details are omitted. In Appendix A, we state and prove a comparison principle for a class of PDEs that are elliptic on the set of convex functions, following the ideas of [36]. We then supply the details of the proof for the particular case of the Monge-Ampère equation (5.1).

Theorem 5.24. *Suppose that Assumption 5.9 holds and, moreover, the domain D is strictly convex, the running cost f is continuous in D, and the boundary cost g is uniformly continuous on ∂D. Then the value function v_{FJ} is the unique convex viscosity solution of the Monge-Ampère problem* (5.9).

Proof. We showed in Theorem 5.14 that the value function v_{FJ} is the unique viscosity solution of the problem (5.5). Again we refer to Theorem 3.5 of [24] to see that v_{FJ} is also a convex viscosity solution of the Dirichlet problem (5.9).

By the uniqueness result of Proposition 5.22, we conclude that v_{FJ} is the unique convex viscosity solution of the Monge-Ampère problem (5.9). □

5.5 The control problem of Gaveau

We now consider a control problem studied by Gaveau in [28], which is also related to the Monge-Ampère equation (5.1).

Define the set of matrices

$$U_{\mathrm{G}} := \left\{ \sigma \in \mathbb{R}^{d,d} \colon \ \det(\sigma) \geq \frac{1}{d} \right\},$$

and let $f : D \to (-\infty, 0]$ be a continuous function. The HJB equation associated to the control problem in this section will be

$$\begin{cases} -\frac{1}{2} \inf_{\sigma \in U_{\mathrm{G}}} \operatorname{Tr}\left(D^2 u \sigma \sigma^\top\right) - f = 0, & \text{in} \quad D, \\ u \quad \text{convex} & \text{in} \quad D. \end{cases} \tag{5.10}$$

Remark 5.25. Convexity of u is necessary and sufficient for the infimum in (5.10) to be finite, since $D^2 u$ is positive semi-definite for any convex function $u \in C^2(D)$. We then see that the infimum is non-negative, and so f cannot be strictly positive when equality holds, thus justifying the conditions imposed on u and f.

In [28], Gaveau shows that (5.10) is equivalent to the Monge-Ampère equation (5.1) in the sense of classical solutions. In Section V.3 of [36], Ishii and Lions show a similar equivalence for convex viscosity solutions, with a minor modification to the definition of the set U_{G}. We now prove the following equivalence.

Lemma 5.26. *Let $u : D \to \mathbb{R}$ be a convex function. Then u is a convex viscosity solution of the HJB equation*

$$-\frac{1}{2} \inf_{\sigma \in U_{\mathrm{G}}} \operatorname{Tr}\left(D^2 u \sigma \sigma^\top\right) - f = 0, \quad \textit{in} \quad D,$$

if and only if u is a convex viscosity solution of the Monge-Ampère equation

$$-\det(D^2 u) + (-2f)^d = 0, \quad \textit{in} \quad D.$$

This result is a consequence of the following matrix identity.

Lemma 5.27. *Let A be a $d \times d$ symmetric matrix. Then*

$$\inf \left\{ \operatorname{Tr}(AB) \colon B \in S_d, B \geq 0, \det B = \frac{1}{d^d} \right\} = \begin{cases} (\det A)^{\frac{1}{d}} & \textit{if} \quad A \geq 0, \\ -\infty & \textit{otherwise.} \end{cases}$$

We note that the above identity is also stated in Section V.3 of [36], and we omit the proof here.

Proof of Lemma 5.26. In order to prove Lemma 5.26, we make two further observations. First, for any symmetric positive semi-definite matrix $B \in S_d$, there exists a matrix $\sigma \in S_d$ such that $B = \sigma\sigma^\top$. Conversely, the matrix $\sigma\sigma^\top$ is symmetric positive semi-definite for any $\sigma \in \mathbb{R}^{d,d}$. Hence the infimum over B is equal to the infimum over $\sigma \in U_{\mathrm{G}}$.

Further, we claim that we can replace the condition $\det(B) = \frac{1}{d^d}$ in Lemma 5.27 with the weaker condition $\det(B) \geq \frac{1}{d^d}$. Indeed, suppose that $\det(B) \geq \frac{1}{d^d}$. Then, by Lemma 5.1,

$$\begin{aligned}
\operatorname{Tr}(AB) &\geq d \det(AB)^{\frac{1}{d}} \\
&= d \det(A)^{\frac{1}{d}} \det(B)^{\frac{1}{d}} \\
&\geq \det(A)^{\frac{1}{d}}.
\end{aligned}$$

In the case that $\det(B) = \frac{1}{d^d}$, then the final inequality above becomes an equality. Therefore the infimum over $\left\{B \colon \det(B) \geq \frac{1}{d^d}\right\}$ is at least as large as the infimum over $\left\{B \colon \det(B) = \frac{1}{d^d}\right\}$ in Lemma 5.27. Moreover, we have the inclusion $\left\{B \colon \det(B) = \frac{1}{d^d}\right\} \subseteq \left\{B \colon \det(B) \geq \frac{1}{d^d}\right\}$, and so the two infima are actually equal.

Hence we can apply the result of Lemma 5.27 to complete the proof. □

We now define a strong and weak formulation of the control problem. For each $N \in \mathbb{N}$, define the set of matrices

$$U_{\mathrm{G}}^N := \left\{\sigma \in \mathbb{R}^{d,d} \colon \det(\sigma) \geq \frac{1}{d}, \quad \sigma \leq NI\right\} \subset U_{\mathrm{G}}.$$

Following Gaveau in [28], we define the strong formulation of the control problem as follows.

Strong Formulation

The strong formulation of the control problem is to find the strong value function $v_{\mathrm{G}}^S : D \to \mathbb{R}$, which we now define.

Let $(\Omega_0, \mathcal{F}, \mathbb{P}_0)$ be a probability space on which a d-dimensional Brownian motion B is defined, with natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$.

Control: For each $N \in \mathbb{N}$, define the set of processes

$$\mathcal{U}_{\mathrm{G}}^N := \left\{U_{\mathrm{G}}^N\text{-valued } \mathbb{F}\text{-progressively measurable processes}\right\}.$$

Then we define the set of controls by

$$\mathcal{U}_{\mathrm{G}} := \bigcup_{N\in\mathbb{N}} \mathcal{U}_{\mathrm{G}}^N.$$

Dynamics: For any $x \in D$ and $\nu = (\nu_t)_{t \geq 0} \in \mathcal{U}_{\mathrm{G}}$, define X^ν by the stochastic integral

$$X_t^\nu = x + \int_0^t \nu_s \, \mathrm{d}B_s, \quad t \geq 0,$$

and define the exit time of the domain by

$$\tau := \inf \{ t \geq 0 \colon X_t^\nu \notin D \} .$$

Value function: We define the *strong value function* $v_{\mathrm{G}}^S : D \to \mathbb{R}$ by

$$v_{\mathrm{G}}^S(x) := \inf_{\nu \in \mathcal{U}_{\mathrm{G}}} \mathbb{E}^x \left[\int_0^\tau f(X_s^\nu) \, \mathrm{d}s + g(X_\tau^\nu) \right] .$$

Remark 5.28. Note that this formulation requires that, for each control $\sigma \in \mathcal{U}_{\mathrm{G}}$, there is some $N \in \mathbb{N}$ such that $\sigma_t \leq NI$, for all $t \geq 0$. In [28], Gaveau remarks that this bound is only needed in order to show that the value function solves a Monge-Ampère equation. It would be of interest to study the problem of optimising over the set of U_{G}-valued controls, without imposing an upper bound, but we do not treat that problem here.

We now define a weak formulation of the control problem, using the setup of El Karoui and Tan in [20].

Weak Formulation

The weak control problem is to find the weak value function $v_{\mathrm{G}}^W : D \to \mathbb{R}$, defined as follows.

Define the space of continuous paths $\Omega := C([0, \infty), \mathbb{R}^d)$ and denote the set of Borel measurable functions $\nu : \mathbb{R}_+ \to U$ by $\mathcal{B}(\mathbb{R}_+, U)$. Then set $\overline{\Omega} = \Omega \times \mathcal{B}(\mathbb{R}_+, U)$ and denote an element of $\overline{\Omega}$ by $\overline{\omega} = (\omega, u)$. Define the canonical process $\overline{X} = (X, \nu)$ on $\overline{\Omega}$ by $X_t(\overline{\omega}) = \omega_t$, for each $t \geq 0$, and $\nu(\overline{\omega}) = u$. For $\phi \in C_b(\mathbb{R}_+ \times U)$, $s \geq 0$, define

$$M_s(\phi) := \int_0^s \phi(r, \nu_r) \, \mathrm{d}r.$$

Then define the canonical filtration $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t \geq 0}$ by

$$\overline{\mathcal{F}}_t := \sigma \{ (X_s, M_s(\phi)) \colon \phi \in C_b(\mathbb{R}_+ \times U), s \leq t \} , \quad t \geq 0.$$

Control: For each $N \in \mathbb{N}$, let $\mathbb{M}_{\mathrm{G}}^N$ be the set of probability measures on the set $\Omega \times \mathcal{B}(\mathbb{R}_+, U_{\mathrm{G}}^N)$. For each $x \in D$, let

$$\mathbb{M}_{\mathrm{G},x}^N = \left\{ \mathbb{P} \in \mathbb{M}_{\mathrm{G}}^N \colon \mathbb{P}(X_0 = x) = 1 \right\}.$$

Dynamics: For each $N \in \mathbb{N}$, define

$$\begin{aligned} \mathcal{P}_{\mathrm{G},x}^N := \{ \mathbb{P} \in \mathbb{M}_{\mathrm{G},x}^N \colon \quad & t \mapsto \phi(X_t) - \phi(X_0) - \frac{1}{2} \int_0^t \mathrm{Tr}\left(D^2\phi(X_s)\nu_s\nu_s^\top \right) \mathrm{d}s \\ & \text{is a } (\overline{\mathbb{F}}, \mathbb{P})\text{-local martingale for all } \phi \in C^2(\mathbb{R}^d) \}. \end{aligned}$$

Then define

$$\mathcal{P}_{\mathrm{G},x} := \bigcup_{N \in \mathbb{N}} \mathcal{P}_{\mathrm{G},x}^N.$$

Let $\tau = \inf \{ t \geq 0 \colon X_t \notin D \}$.

Value function: We define the *weak value function* $v_{\mathrm{G}}^W : D \to \mathbb{R}$ by

$$v_{\mathrm{G}}^W(x) = \inf_{\mathbb{P} \in \mathcal{P}_{\mathrm{G},x}} \mathbb{E}^{\mathbb{P}} \left[\int_0^\tau f(X_s) \, \mathrm{d}s + g(X_\tau) \right].$$

To prove that the weak and strong formulations are equivalent, we take a similar approach to that in Proposition 1.7 and Proposition 5.6, making use of the form of the control set $\mathcal{U}_{\mathrm{G}} = \bigcup_{N \in \mathbb{N}} \mathcal{U}_{\mathrm{G}}^N$.

Proposition 5.29. *Under Assumption 5.2, we have the equality $v_{\mathrm{G}}^S = v_{\mathrm{G}}^W$ in D.*

Proof. We first note that $v_{\mathrm{G}}^W \leq v_{\mathrm{G}}^S$, by definition of the weak and strong value functions. To see this, fix $x \in D$ and define

$$\mathcal{P}_{\mathrm{G},x}^S := \{ \mathbb{P} \in \mathcal{P}_{\mathrm{G},x} \colon \mathbb{P} = \mathbb{P}^{X^\nu} \times \delta_{\nu_\cdot}, \text{ for some } \nu \in \mathcal{U}_{\mathrm{G}} \},$$

where $\mathbb{P}^X$ denotes the law of a process X. Then

$$v_{\mathrm{G}}^S(x) = \inf_{\mathbb{P} \in \mathcal{P}_{\mathrm{G},x}^S} \mathbb{E}^{\mathbb{P}} \left[\int_0^\tau f(X_s) \, \mathrm{d}s + g(X_\tau) \right] \geq v_{\mathrm{G}}^W(x),$$

since $\mathcal{P}_{\mathrm{G},x}^S \subseteq \mathcal{P}_{\mathrm{G},x}$.

We now show that $v_{\mathrm{G}}^S \leq v_{\mathrm{G}}^W$, by considering the following approximations to the value functions. For each $N \in \mathbb{N}$, define the functions $v_{\mathrm{G}}^{S,N}, v_{\mathrm{G}}^{W,N} : D \to \mathbb{R}$ by

$$v_{\mathrm{G}}^{S,N}(x) := \inf_{\nu \in \mathcal{U}_{\mathrm{G}}^N} \mathbb{E}^x \left[\int_0^\tau f(X_s^\nu) \, \mathrm{d}s + g(X_\tau^\nu) \right],$$

and

$$v_{\mathrm{G}}^{W,N}(x) = \inf_{\mathbb{P} \in \mathcal{P}_{\mathrm{G},x}^{N}} \mathbb{E}^{\mathbb{P}} \left[\int_0^\tau f(X_s)\,\mathrm{d}s + g(X_\tau) \right],$$

for $x \in D$. We claim that $v_{\mathrm{G}}^{S,N} = v_{\mathrm{G}}^{W,N}$, for each $N \in \mathbb{N}$, and show that this implies the result, before proving this claim.

Fix $x \in D$. By definition of the weak value function v_{G}^{W}, there exists a sequence $(\mathbb{P}_k)_{k \in \mathbb{N}} \subset \mathcal{P}_{\mathrm{G},x}$ such that

$$v_{\mathrm{G}}^{W} = \lim_{k \to \infty} \mathbb{E}^{\mathbb{P}_k} \left[\int_0^\tau f(X_s)\,\mathrm{d}s + g(X_\tau) \right].$$

Fix $k \in \mathbb{N}$. Then, since $\mathcal{P}_{\mathrm{G},x} = \bigcup_{N \in \mathbb{N}} \mathcal{P}_{\mathrm{G},x}^{N}$, there exists $N(k) \in \mathbb{N}$ such that $\mathbb{P}_k \in \mathcal{P}_{\mathrm{G},x}^{N(k)}$.

Therefore, by definition of $v_{\mathrm{G}}^{W,N(k)}$, we have

$$\mathbb{E}^{\mathbb{P}_k} \left[\int_0^\tau f(X_s)\,\mathrm{d}s + g(X_\tau) \right] \geq v_{\mathrm{G}}^{W,N(k)}(x).$$

Supposing that $v_{\mathrm{G}}^{W,N(k)} = v_{\mathrm{G}}^{S,N(k)}$, we then have

$$\mathbb{E}^{\mathbb{P}_k} \left[\int_0^\tau f(X_s)\,\mathrm{d}s + g(X_\tau) \right] \geq v_{\mathrm{G}}^{S,N(k)}(x) \geq v_{\mathrm{G}}^{S}(x),$$

since $\mathcal{U}_{\mathrm{G}}^{N(k)} \subset \mathcal{U}_{\mathrm{G}}$. Taking the limit as $k \to \infty$, we get the desired inequality

$$v_{\mathrm{G}}^{W}(x) \geq v_{\mathrm{G}}^{S}(x).$$

We now fix $N \in \mathbb{N}$ and verify our claim that $v_{\mathrm{G}}^{S,N} = v_{\mathrm{G}}^{W,N}$. We will apply Theorem 4.5 of [20], as in the proofs of Proposition 1.7 and Proposition 5.6, making use of the boundedness of the set U_{G}^{N}.

Define the function $\Phi : \Omega \to \mathbb{R}$ by

$$\Phi(\omega) := \int_0^{\tau(\omega)} f(X_s(\omega))\,\mathrm{d}s + g\left(X_{\tau(\omega)}(\omega)\right),$$

and fix $x \in D$. Then, to show that the conditions of Theorem 4.5 of [20] are satisfied, we need to check that Φ is upper semicontinuous and bounded above by some random variable that is uniformly integrable under the family of probability measures $\mathcal{P}_{\mathrm{G},x}^{N}$

Upper semicontinuity of Φ follows from the assumption that f and g are both upper semicontinuous in Assumption 5.2. By Assumption 5.2, we also have that f

is negative and g is bounded above by some constant K. Hence

$$\Phi(\omega) \leq K, \quad \text{for all} \quad \omega \in \Omega,$$

and so the uniform integrability condition is satisfied. We now apply Theorem 4.5 of [20] to conclude that $v_{\mathrm{G}}^{S,N} = v_{\mathrm{G}}^{W,N}$.

We have therefore shown that $v_{\mathrm{G}}^{S} = v_{\mathrm{G}}^{W}$, as required. □

Having proved that the weak and strong value functions are equal, we denote the common value by $v_{\mathrm{G}} = v_{\mathrm{G}}^{S} = v_{\mathrm{G}}^{W}$ and refer to v_{G} as the *value function*. We will use the weak formulation in the proof of the next result, but we will find it convenient to work with the strong value function thereafter.

We now show that the value function v_{G} is bounded below by v.

Proposition 5.30. *Suppose that Assumption 5.2 holds. Then $v \leq v_{\mathrm{G}}$ in D.*

Proof. We will work with the weak formulations of the control problems in this proof. Note that, for each control problem, we have equality between the strong and weak value functions, by Proposition 1.7 and Proposition 5.29, respectively.

Let $x \in D$ and suppose that $\tilde{\mathbb{P}} \in \mathcal{P}_{\mathrm{G},x}$. We will use a time-change argument to find a measure $\mathbb{P} \in \mathcal{P}_x$ under which the expected cost in the control problem is less than the expected cost under $\tilde{\mathbb{P}}$.

Let (X, ν) have joint law $\tilde{\mathbb{P}}$. Then, by Proposition 4.6 of [38, Chapter 5] on the relationship between solutions of local martingale problems and weak solutions of SDEs, there exists a d-dimensional Brownian motion B such that, for any $t \geq 0$,

$$X_t = x + \int_0^t \nu_s \, \mathrm{d}B_s.$$

By definition of $\mathcal{P}_{\mathrm{G},x}$, there exists $N \in \mathbb{N}$ such that, for each $t \geq 0$, $\nu_t \in U_{\mathrm{G}}^N$. Therefore, by Lemma 5.1,

$$\mathrm{Tr}(\nu_t \nu_t^\top) \geq d \det(\nu_t \nu_t^\top)^{\frac{1}{d}} \geq 1. \tag{5.11}$$

Define A to be the quadratic variation process associated to X. Then, by (5.11),

$$A_t = \langle X \rangle_t = \int_0^t \mathrm{Tr}(\nu_s \nu_s^\top) \, \mathrm{d}s \geq t.$$

Now define $\hat{X}$ to be the time-changed process

$$\hat{X}_t = X_{A_t^{-1}},$$

where

$$A_t^{-1} := \inf\{u \geq 0\colon A_u > t\}.$$

Also define $A_{t-}^{-1} := \inf\{u \geq 0\colon A_u \geq t\}$. Referring to Section 1 of [51, Chapter IV] on quadratic variations, we show that the time-change $t \mapsto A_t^{-1}$ has the following properties.

1. The processes $t \mapsto A_t$ and $t \mapsto A_t^{-1}$ are both continuous and strictly increasing.

   To prove this, first note that $t \mapsto A_t$ is continuous and increasing by definition of quadratic variation of a continuous local martingale (see Theorem 1.8 of [51, Chapter IV]). Then, by (5.11), we have that

   $$A_{t_2} - A_{t_1} = \int_{t_1}^{t_2} \mathrm{Tr}(\nu_s \nu_s^\top)\,\mathrm{d}s \geq t_2 - t_1.$$

   Hence $t \mapsto A_t$ is strictly increasing.

   From its definition, we see that $t \mapsto A_t^{-1}$ is continuous and strictly increasing when the same properties hold for $t \mapsto A_t$.

2. $A_{A_t^{-1}} = t$, for any $t \geq 0$.

   This follows from the fact that A is strictly increasing, as discussed in Section 4 of [51, Chapter 0].

3. A^{-1} is almost surely finite.

   Note that $A_t \geq t$ implies that $A_\infty = \infty$. Suppose that $A_t^{-1} = \infty$. Then $t = A_\infty = \infty$. Hence A^{-1} is almost surely finite.

4. A and X are constant on the same intervals.

   This is a property of quadratic variation that is proved in Proposition 1.13 of [51, Chapter IV].

5. X is A^{-1}-continuous; i.e. X is constant on each interval $[A_{t-}^{-1}, A_t^{-1}]$.

   This follows from the fact that $A_{t-}^{-1} = A_t^{-1}$ for all $t \geq 0$, since A is strictly increasing.

Let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be the natural filtration of X, and denote the time-changed filtration by $\hat{\mathbb{F}} = (\mathcal{F}_{A_t^{-1}})_{t\geq 0}$. Since A^{-1} is almost surely finite and X is A^{-1}-continuous, we can apply Proposition 1.5 of [51, Chapter 5]. This result implies that the time-changed process $\hat{X}$ is an $\hat{\mathbb{F}}$-martingale with quadratic variation given by

$$\langle \hat{X} \rangle_t = \langle X \rangle_{A_t^{-1}} = A_{A_t^{-1}} = t,$$

where we use property 2 above to get the final equality.

We now wish to write

$$\hat{X}_t = x + \int_0^t \tilde{\nu}_s \,\mathrm{d}W_s,$$

for some $\hat{\mathbb{F}}$-progressively measurable $\tilde{\nu}$ and an $\hat{\mathbb{F}}$-Brownian motion W.

Define $\hat{\nu}_t := \nu_{A_t^{-1}}$ and $\hat{B}_t := B_{A_t^{-1}}$. Then, by Proposition 1.5 of [51, Chapter V],

$$\hat{X}_t = x + \int_0^t \hat{\nu}_s \,\mathrm{d}\hat{B}_s.$$

For any $t \geq 0$, define

$$\alpha_t := \mathrm{Tr}(\hat{\nu}_t \hat{\nu}_t^\top)^{-\frac{1}{2}}, \quad \text{and} \quad W_t := \int_0^t \alpha_s^{-1} \,\mathrm{d}\hat{B}_s.$$

Then, by associativity of the stochastic integral (see Proposition IV.2.4 of [51]), we have

$$\int_0^t \alpha_s \,\mathrm{d}W_s = \int_0^t \alpha_s \alpha_s^{-1} \,\mathrm{d}\hat{B}_s = \hat{B}_t.$$

This gives us

$$\begin{aligned}
\hat{X}_t &= x + \int_0^t \hat{\nu}_s \,\mathrm{d}\hat{B}_s = x + \int_0^t \alpha_s \hat{\nu}_s \,\mathrm{d}W_s \\
&= x + \int_0^t \mathrm{Tr}(\hat{\nu}_s \hat{\nu}_s^\top)^{-\frac{1}{2}} \hat{\nu}_s \,\mathrm{d}W_s \\
&= x + \int_0^t \tilde{\nu}_s \,\mathrm{d}W_s,
\end{aligned}$$

defining $\tilde{\nu}_s := \mathrm{Tr}(\hat{\nu}_s \hat{\nu}_s^\top)^{-\frac{1}{2}} \hat{\nu}_s$.

For any $s \geq 0$

$$\mathrm{Tr}(\tilde{\nu}_s \tilde{\nu}_s^\top) = \left(\mathrm{Tr}(\hat{\nu}_s \hat{\nu}_s^\top)^{-\frac{1}{2}}\right)^2 \mathrm{Tr}(\hat{\nu}_s \hat{\nu}_s^\top) = 1,$$

and so $\tilde{\nu}_s \in U$.

We now check that W is a Brownian motion, using Lévy's characterisation (see Theorem 3.6 of [51, Chapter IV]). For any i, j,

$$\langle W^i, W^j \rangle_t = \left\langle \int_0^t \alpha_s^{-1} \,\mathrm{d}\hat{B}_s^i, \int_0^t \alpha_s^{-1} \,\mathrm{d}\hat{B}_s^j \right\rangle_t = \int_0^t \alpha_s^{-2} \,\mathrm{d}\langle \hat{B}_s^i, \hat{B}_s^j \rangle,$$

where the second equality follows from Proposition 2.17 of [38, Chapter 3].

Now, by Proposition 1.5 of [51, Chapter V], we know that $\langle \hat{B}^i \rangle_t = \widehat{\langle B^i \rangle}_t$, for all i, and that $\langle \hat{B}^i - \hat{B}^j \rangle_t = \langle \widehat{B^i - B^j} \rangle_t$, for all i, j. And so, by expanding the expressions in the second equality, we can deduce that, for all i, j,

$$\langle \hat{B}^i, \hat{B}^j \rangle_t = \langle \widehat{B^i, B^j} \rangle_t = \widehat{\delta_{ij} t} = \delta_{ij} A_t^{-1}.$$

Note that

$$\begin{aligned}
\int_0^t \alpha_s^2 \, \mathrm{d}s &= \int_0^t \mathrm{Tr}(\hat{\nu}_s \hat{\nu}_s^\top)^{-1} \, \mathrm{d}s \\
&= \int_0^t \mathrm{Tr}\left(\nu_{A_s^{-1}} \nu_{A_s^{-1}}^\top\right)^{-1} \mathrm{d}s \\
&= \int_0^{A_t^{-1}} \mathrm{Tr}(\nu_u \nu_u^\top)^{-1} \, \mathrm{d}A_u \\
&= \int_0^{A_t^{-1}} \mathrm{Tr}(\nu_u \nu_u^\top)^{-1} \, \mathrm{Tr}(\nu_u \nu_u^\top) \, \mathrm{d}u = A_t^{-1},
\end{aligned}$$

making the change of variables $u = A_s^{-1}$ in the penultimate line.

So, for any i, j, we have

$$\langle W^i, W^j \rangle_t = \int_0^t \alpha_s^{-2} \, \mathrm{d}\langle \hat{B}^i, \hat{B}^j \rangle_t = \int_0^t \alpha_s^{-2} \delta_{ij} \, \mathrm{d}A_t^{-1} = \int_0^t \alpha_s^{-2} \delta_{ij} \alpha_s^2 \, \mathrm{d}s = \delta_{ij} t.$$

This shows that W is indeed a standard Brownian motion, by Lévy's characterisation.

Now define a probability measure on $\Omega \times \mathcal{B}(\mathbb{R}_+, U)$ by

$$\mathbb{P} = \mathbb{P}^{\hat{X}} \otimes \delta_{\hat{\nu}_\cdot},$$

where $\mathbb{P}^{\hat{X}}$ is the law of $\hat{X}$. Then $\mathbb{P} \in \mathcal{P}_x$.

Define $\hat{\tau} := \inf\{t \geq 0 \colon \hat{X}_t \notin D\}$ and consider

$$\mathbb{E}^{\mathbb{P}} \left[\int_0^{\hat{\tau}} f(\hat{X}_s) \, \mathrm{d}s + g(\hat{X}_{\hat{\tau}}) \right].$$

Note that, since A^{-1} is strictly increasing and

$$\hat{\tau} = \inf\{t \geq 0 \colon X_{A_t^{-1}} \notin D\},$$

we have

$$A_{\hat{\tau}}^{-1} = \tau. \tag{5.12}$$

Now we calculate

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[\int_0^{\hat{\tau}} f(\hat{X}_s)\,\mathrm{d}s + g(\hat{X}_{\hat{\tau}})\right] &= \mathbb{E}^{\mathbb{P}}\left[\int_0^{A^{-1}_{\hat{\tau}}} f(\hat{X}_{A_t})\,\mathrm{d}A_t + g(\hat{X}_{\hat{\tau}})\right]\\ &= \mathbb{E}^{\tilde{\mathbb{P}}}\left[\int_0^{A^{-1}_{\hat{\tau}}} f(X_{A^{-1}_{A_t}})\operatorname{Tr}(\hat{\nu}_t\hat{\nu}_t^{\top})\,\mathrm{d}t + g(X_{A^{-1}_{\hat{\tau}}})\right],\end{aligned}$$

making the substitution $s = A_t$ in the first line, and then using the definitions of A and $\hat{X}$. Using (5.12) and the fact that $A^{-1}_{A_t} = t$ for any $t \geq 0$, we then have

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\left[\int_0^{\hat{\tau}} f(\hat{X}_s)\,\mathrm{d}s + g(\hat{X}_{\hat{\tau}})\right] &= \mathbb{E}^{\tilde{\mathbb{P}}}\left[\int_0^{\tau} f(X_t)\operatorname{Tr}(\hat{\nu}_t\hat{\nu}_t^{\top})\,\mathrm{d}t + g(X_\tau)\right]\\ &\leq \mathbb{E}^{\tilde{\mathbb{P}}}\left[\int_0^{\tau} f(X_t)\,\mathrm{d}t + g(X_\tau)\right],\end{aligned}$$

where the final inequality follows from (5.11) and negativity of f.

We have shown that, for any $\tilde{\mathbb{P}} \in \mathcal{P}_{\mathrm{G},x}$, we can find a $\mathbb{P} \in \mathcal{P}_x$ such that

$$\mathbb{E}^{\mathbb{P}}\left[\int_0^{\hat{\tau}} f(\hat{X}_s)\,\mathrm{d}s + g(\hat{X}_{\hat{\tau}})\right] \leq \mathbb{E}^{\tilde{\mathbb{P}}}\left[\int_0^{\tau} f(X_t)\,\mathrm{d}t + g(X_\tau)\right].$$

Hence $v(x) \leq v_{\mathrm{G}}(x)$. □

In Example 5.41, we will show that equality holds in the above inequality for the cost function defined in Example 2.1. However, equality does not hold in general. We will show in Proposition 5.44 that we have the strict inequality $v < v_{\mathrm{G}}$ for a particular class of cost functions.

5.5.1 Dynamic programming principle

We now refer to the work of Gaveau in [28] to show that the value function v_{G} satisfies a dynamic programming principle and is continuous and convex. In this section we will work with the strong formulation of the control problem. We make the following strengthening of Assumption 5.2.

Assumption 5.31. Suppose that Assumption 5.2 holds and, moreover, the domain D is uniformly convex, the running cost f is uniformly continuous in D, and the boundary cost g is continuous on ∂D.

We first show that the value function has a continuous extension on $\overline{D}$ that attains the value g on the boundary ∂D, using Theorem 1 of [28].

Proposition 5.32. *Suppose that Assumption 5.31 holds. For each $\sigma \in \mathcal{U}_{\mathrm{G}}$, the function $\mathcal{I}_{\mathrm{G}}(\cdot;\sigma) : \overline{D} \to \mathbb{R}$, defined by*

$$\mathcal{I}_{\mathrm{G}}(x;\sigma) := \mathbb{E}^x \left[\int_0^\tau f(X_s^\sigma)\,\mathrm{d}s + g(X_\tau^\sigma) \right],$$

is continuous. Moreover, the value function v_{G} has a continuous extension on $\overline{D}$ with

$$\lim_{x \to x_0} v_{\mathrm{G}}(x) = g(x_0),$$

for any $x_0 \in \partial D$.

Proof. First fix $\sigma \in \mathcal{U}_{\mathrm{G}}$. Under the given assumptions, Lemmas 2–4 of [28] hold. We can follow the proof of Theorem 1 of [28] to prove estimates on the function $\mathcal{I}_{\mathrm{G}}(\cdot,\sigma) : \overline{D} \to \mathbb{R}$, similar to statements (2) and (3) of Theorem 1 of [28], where the constants are independent of the choice of σ. These estimates imply continuity of $\mathcal{I}_{\mathrm{G}}(\cdot,\sigma)$.

Then, following the proof of Theorem 1 of [28], we take the infimum over $\sigma \in \mathcal{U}_{\mathrm{G}}$ to conclude that v_{G} is continuous in D and that $\lim_{x \to x_0} v_{\mathrm{G}}(x) = g(x_0)$, for any $x_0 \in \partial D$. □

We now refer to Theorem 3 of [28] for a proof of the dynamic programming principle.

Proposition 5.33. *Suppose that Assumption 5.31 holds. Then we have the following dynamic programming principle.*

For any $x \in D$, let θ be the exit time of some domain $D' \subset D$ with $x \in D'$. Then, for any $t \geq 0$,

$$v_{\mathrm{G}}(x) = \inf_{\nu \in \mathcal{U}_{\mathrm{G}}} \mathbb{E}^x \left[\int_0^{\theta \wedge t} f(X_s^\nu)\,\mathrm{d}s + v_{\mathrm{G}}(X_{\theta \wedge t}^\nu) \right]. \tag{5.13}$$

Proof. Under the given assumptions, the continuity result of Proposition 5.32 holds, and so we can apply Lemma 5 of [28]. We note that continuity of v_{G} is used in the proof of this lemma to make a measurable selection argument, in a similar way as in our proof of Proposition 1.17.

We can then follow the proof of Theorem 3 of [28] to conclude that the dynamic programming principle (5.13) holds. □

As a corollary to this result, Gaveau shows that, under the same conditions, the value function is convex in D.

Corollary 5.34. *Suppose that Assumption 5.31 holds. Then the value function v_{G} is convex in D.*

Proof. Under the given assumptions, we have the continuity results of Proposition 5.32 and the dynamic programming principle of Proposition 5.33. Then Theorem 2 of [28] implies that v_{G} is convex in D. □

Remark 5.35. In contrast to the two control problems that we have studied so far, we have proved the dynamic programming principle for v_{G} and then deduced convexity as a corollary. In Lemma 1.11 and Lemma 5.8, we were able to show that the value functions v and v_{FJ} are convex a priori. We went on to deduce continuity and use this to prove a dynamic programming principle, under weaker conditions. We conjecture that these proofs can be adapted to show that the value function v_{G} is convex a priori without the strict conditions of Assumption 5.31.

Having established a dynamic programming principle, we will use this to show that the value function is the unique convex viscosity solution of the associated HJB equation with appropriate boundary condition.

5.5.2 Viscosity solution characterisation

We now show that the value function v_{G} solves the HJB equation (5.10) in D with boundary condition $v_{\mathrm{G}} = g$ on ∂D. We will also deduce uniqueness for this boundary value problem, by combining the uniqueness result for the Monge-Ampère equation given in Proposition 5.22 with the equivalence result proved in Lemma 5.26.

Theorem 5.36. *Suppose that Assumption 5.31 holds. Then v_{G} is the unique convex viscosity solution of the Dirichlet problem*

$$\begin{cases} -\frac{1}{2}\inf_{\sigma\in U_{\mathrm{G}}}\mathrm{Tr}(D^2u\sigma\sigma^\top) - f = 0, & \textit{in} \quad D, \\ u \quad \textit{convex} & \textit{in} \quad D, \\ u = g, & \textit{on} \quad \partial D. \end{cases} \tag{5.14}$$

Proof. Under the given assumptions, a dynamic programming principle holds for v_{G}, by Proposition 5.33. Using the fact that, for any $\sigma \in U_{\mathrm{G}}$, there exists $N \in \mathbb{N}$ such that $\sigma \leq NI$, we can follow the same arguments as in the proofs of Proposition 4.8 and Proposition 4.9, to deduce that v_{G} is a convex viscosity solution of the HJB equation (5.10) in D.

From Proposition 5.32, we also have that

$$\lim_{x\to x_0} v_{\mathrm{G}}(x) = g(x_0),$$

for any $x_0 \in \partial D$. Hence v_{G} is a convex viscosity solution of the problem (5.14)

Uniqueness will follow by equivalence of the HJB equation to a Monge-Ampère equation. By Lemma 5.26, any convex viscosity solution of the HJB equation (5.10) is also a convex viscosity solution of the Monge-Ampère equation (5.1). Convex viscosity solutions of the Monge-Ampère equation with a given Dirichlet boundary condition are unique by Proposition 5.22. Therefore there is at most one convex viscosity solution of the Dirichlet problem (5.14).

Hence v_{G} is the unique convex viscosity solution of (5.14). □

Corollary 5.37. *Suppose that Assumption 5.31 holds. Then v_{G} is the unique convex viscosity solution of the Monge-Ampère problem* (5.9).

Proof. From Theorem 5.36, we have that v_{G} is a convex viscosity solution of the Dirichlet problem (5.14). Then, by Lemma 5.26, v_{G} is also a convex viscosity solution of the Monge-Ampère problem (5.9). We have uniqueness by Proposition 5.22. □

5.5.3 Alternative proof of attainment of the boundary condition

Part of the statement of Proposition 5.32 is that the value function v_{G} attains the boundary value g on ∂D. We proved this by following the work of Gaveau in [28]. In the following lemma, we prove the attainment of the boundary condition under slightly weaker conditions.

Lemma 5.38. *Suppose that Assumption 5.2 holds and that $f : D \to \mathbb{R}$ is bounded and $g : \partial D \to \mathbb{R}$ is continuous. Let $x_0 \in \partial D$. Then*

$$\lim_{x \to x_0} v_{\mathrm{G}}(x) = g(x_0).$$

Proof. Fix $\sigma \in U_{\mathrm{G}}$ and recall the definition of $\mathcal{I}_{\mathrm{G}}(\cdot;\sigma) : \overline{D} \to \mathbb{R}$ from Proposition 5.32. We first show that

$$\lim_{x \to x_0} \mathcal{I}_{\mathrm{G}}(x;\sigma) = g(x_0).$$

For any $t > 0$, we have both $\sigma_t \sigma_t^\top \geq 0$ and $\det(\sigma_t \sigma_t^\top) > 0$. Hence $\sigma_t \sigma_t^\top > 0$. Also, since D is a convex domain, it satisfies an exterior sphere condition. We can then check that all of the conditions are satisfied in order to apply Theorem 3.3 of [50, Chapter 2]. The first result of this theorem is that

$$\lim_{x \to x_0} \mathbb{P}^x \left[\tau > t\right] = 0, \quad \text{for all} \quad t > 0.$$

This result is proved by showing that

$$\lim_{x\to x_0} \mathbb{E}^x[\tau] = 0, \tag{5.15}$$

and then applying Chebyshev's inequality. Combining (5.15) with the fact that f is bounded, we see that

$$\lim_{x\to x_0} \left| \mathbb{E}^x \left[\int_0^\tau f(X_s^\sigma)\, \mathrm{d}s \right] \right| \le \|f\|_\infty \lim_{x\to x_0} \mathbb{E}^x[\tau] = 0. \tag{5.16}$$

We now use the second result of Theorem 3.3 of [50, Chapter 2], which gives us that, for any $t > 0$,

$$\lim_{x\to x_0} \mathbb{E}^x\left[g(X_{t\wedge\tau}^\sigma)\right] = g(x_0). \tag{5.17}$$

Letting $t\to\infty$, we conclude from (5.16) and (5.17) that

$$\lim_{x\to x_0} \mathcal{I}_{\mathrm{G}}(x;\sigma) = \lim_{x\to x_0} \mathbb{E}^x \left[\int_0^\tau f(X_s^\sigma)\, \mathrm{d}s + g(X_\tau^\sigma) \right] = g(x_0). \tag{5.18}$$

We now consider the infimum, $v_{\mathrm{G}}^S(x) = \inf_{\sigma\in\mathcal{U}_{\mathrm{G}}} \mathcal{I}_{\mathrm{G}}(x;\sigma)$, for $x \in D$.

Let $\delta > 0$ and fix $x \in B_\delta(x_0) \cap D$. By definition of the infimum, there exists $\sigma^\varepsilon \in \mathcal{U}_{\mathrm{G}}$ such that

$$v_{\mathrm{G}}(x) > \mathcal{I}_{\mathrm{G}}(x;\sigma^\varepsilon) - \frac{\varepsilon}{2}.$$

By (5.18), we can choose δ such that

$$|\mathcal{I}_{\mathrm{G}}(x;\sigma^\varepsilon) - g(x_0)| < \frac{\varepsilon}{2}. \tag{5.19}$$

Therefore $\mathcal{I}_{\mathrm{G}}(x;\sigma^\varepsilon) > g(x_0) - \frac{\varepsilon}{2}$, and so

$$v_{\mathrm{G}}(x) > g(x_0) - \varepsilon.$$

We also have that $v_{\mathrm{G}}(x) \le \mathcal{I}_{\mathrm{G}}(x;\sigma^\varepsilon)$, by definition of the infimum. Using (5.19) again, we get $\mathcal{I}_{\mathrm{G}}(x;\sigma^\varepsilon) < g(x_0) + \frac{\varepsilon}{2}$, and so

$$v_{\mathrm{G}}(x) \le g(x_0) + \frac{\varepsilon}{2} < g(x_0) + \varepsilon.$$

We conclude that

$$|v_{\mathrm{G}}(x) - g(x_0)| < \varepsilon.$$

Hence

$$\lim_{x\to x_0} v_{\mathrm{G}}(x) = g(x_0),$$

as required. $\square$

Remark 5.39. A key point in the above proof is that, for any $\sigma \in U_{\mathrm{G}}$, the matrix $\sigma\sigma^\top$ is positive definite. This allows us to apply the result from [50] for any convex domain. To prove a similar result for the value function v in Section 4.3, we could not apply the same result from [50], since $\sigma\sigma^\top$ may be degenerate for $\sigma \in U$. In dimension $d = 2$, if the boundary of a domain has a straight edge, then this allows for controlled processes which are constrained to move on a line parallel to that edge. Therefore we do not expect the boundary condition to be attained. In order to prove attainment of the boundary condition in Proposition 4.23, we needed to restrict ourselves to domains satisfying the same uniform convexity condition that Gaveau imposes in [28].

5.6 Relationship between value functions

In this section, we obtain an ordering of the value functions for all of the control problems that we have considered in this thesis.

Theorem 5.40. *Suppose that each assumption holds from Assumption 1.1, Assumption 5.9 and Assumption 5.31, and suppose that the boundary cost g is uniformly continuous on ∂D.*

Then we have the following ordering between the value functions:

$$v^S = v^W \leq v^W_{\mathrm{G}} = v^S_{\mathrm{G}} = v^S_{\mathrm{FJ}} = v^W_{\mathrm{FJ}}.$$

Proof. We prove each of the relations in turn.

1. The equality $v^S = v^W$ is the result of Proposition 1.7, which holds under Assumption 1.1.

2. We proved that $v^W \leq v^W_{\mathrm{G}}$ in Proposition 5.30, under weaker conditions than Assumption 5.9.

3. The equality $v^W_{\mathrm{G}} = v^S_{\mathrm{G}}$ is the result of Proposition 5.29, which holds under Assumption 5.9.

4. Under Assumption 5.31, Corollary 5.37 implies that v^S_{G} is the unique convex viscosity solution of the Monge-Ampère problem (5.9). Combining Assumption 5.9 and Assumption 5.31 with the assumption that g is uniformly continuous on ∂D, we have the required conditions for Theorem 5.24 to hold.

Hence v^S_{FJ} is the unique convex viscosity solution of the same problem (5.9). By uniqueness of solutions of (5.9), we obtain $v^S_{\mathrm{G}} = v^S_{\mathrm{FJ}}$.

5. The final equality $v^S_{\mathrm{FJ}} = v^W_{\mathrm{FJ}}$ is the result of Proposition 5.6, which holds under Assumption 5.9. □

We now show by means of examples that, while all of the value functions may coincide for some cost functions, this is not always the case.

We first revisit Example 2.1 and show that, for the step cost function in this example, all of the value functions in Theorem 5.40 are equal.

Example 5.41. Fix $R > 0$ and let $D = B_R(0) \subset \mathbb{R}^d$. Let $\rho \in (0, R)$ and define $f : D \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0, & |x| \leq \rho, \\ -1, & |x| \in (\rho, R). \end{cases}$$

We seek the value functions $v_{\mathrm{FJ}}, v_{\mathrm{G}} : D \to \mathbb{R}$, given by

$$\begin{aligned} v_{\mathrm{FJ}}(x) &= \inf_{\sigma \in \mathcal{U}} \mathbb{E}^x \left[d \int_0^\tau f(X^\sigma_s) \det(\sigma_s \sigma_s^\top)^{\frac{1}{d}} \, \mathrm{d}s \right] \\ &= \inf_{\sigma \in \mathcal{U}} \mathbb{E}^x \left[d \int_0^\tau -\mathbb{1}_{\{|X^\sigma_s| \in (\rho, R)\}} \det(\sigma_s \sigma_s^\top)^{\frac{1}{d}} \, \mathrm{d}s \right], \end{aligned}$$

and

$$v_{\mathrm{G}}(x) = \inf_{\sigma \in \mathcal{U}_{\mathrm{G}}} \mathbb{E}^x \left[\int_0^\tau -\mathbb{1}_{\{|X^\sigma_s| \in (\rho, R)\}} \, \mathrm{d}s \right].$$

In this example, the cost function f has a discontinuity, and so we do not have a PDE characterisation for either of the value functions. Therefore, we do not know a priori that $v_{\mathrm{FJ}} = v_{\mathrm{G}}$. We will find each value function in turn and deduce that they are in fact both equal to the value function v. Note that Assumption 5.2 is satisfied, and so, for each control problem, the weak and strong formulations are equivalent.

Proposition 5.42. *In Example 5.41, the value function v_{FJ} is given by*

$$v_{\mathrm{FJ}}(x) = \begin{cases} \rho^2 - R^2, & |x| \leq \rho, \\ |x|^2 - R^2, & |x| \in (\rho, R). \end{cases}$$

Proof. Define $w : (0, R^2) \to \mathbb{R}$ by

$$w(\xi) := \begin{cases} \rho^2 - R^2, & \xi \leq \rho^2, \\ |x|^2 - R^2, & \xi \in (\rho^2, R^2), \end{cases}$$

as in Proposition 2.5. Also let $Z_t^\sigma = |X_t^\sigma|^2$ for any $\sigma \in \mathcal{U}$ and $t \geq 0$, and recall from (2.2) that Z^σ satisfies

$$\mathrm{d}Z_t^\sigma = 2X_t^\top \sigma_t \,\mathrm{d}B_t + \mathrm{d}t.$$

Then, similarly to (2.4), we can apply the Itô-Tanaka formula to get

$$\begin{aligned} w(Z_t^\sigma) - w(\xi) + d\int_0^t f(X_s^\sigma)\det(\sigma_s\sigma_s^\top)^{\frac{1}{d}}\,\mathrm{d}s = 2\int_0^t &\mathbb{1}_{\{Z_s^\sigma>\rho^2\}} X_s^\top \sigma_s\,\mathrm{d}B_s \\ &+ \int_0^t \mathbb{1}_{\{Z_s^\sigma>\rho^2\}}\,\mathrm{d}s + \frac{1}{2}L_t^{\sigma,\rho^2} \\ &- d\int_0^t \mathbb{1}_{\{Z_s^\sigma>\rho^2\}} \det(\sigma_s\sigma_s^\top)^{\frac{1}{d}}\,\mathrm{d}s. \end{aligned} \tag{5.20}$$

By Lemma 5.1, we have the bound

$$d\det(\sigma_s\sigma_s^\top)^{\frac{1}{d}} \leq \mathrm{Tr}(\sigma_s\sigma_s^\top) = 1, \tag{5.21}$$

for any $\sigma \in \mathcal{U}$, and so by non-negativity of the local time, we have

$$w(Z_t^\sigma) - w(\xi) + d\int_0^t f(X_s^\sigma)\det(\sigma_s\sigma_s^\top)^{\frac{1}{d}}\,\mathrm{d}s \geq 2\int_0^t \mathbb{1}_{\{Z_s^\sigma>\rho^2\}} X_s^\top \sigma_s\,\mathrm{d}B_s.$$

By the optional sampling theorem,

$$\begin{aligned} w(\xi) &\leq \mathbb{E}^\xi\left[w(Z_\tau^\sigma) + d\int_0^\tau f(X_s^\sigma)\det(\sigma_s\sigma_s^\top)^{\frac{1}{d}}\,\mathrm{d}s\right] \\ &= \mathbb{E}^\xi\left[d\int_0^\tau f(X_s^\sigma)\det(\sigma_s\sigma_s^\top)^{\frac{1}{d}}\,\mathrm{d}s\right], \end{aligned}$$

for any $\sigma \in \mathcal{U}$.

We now find a minimising sequence of controls. Recall from Example 2.1, that we seek controls that have zero local time on the internal boundary $|x| = \rho$. However, the penalisation of small determinants now forces us to choose a control that has as large a determinant as possible within the constraint (5.21).

Let $\sigma \in \mathcal{U}$ be any control such that $\det(\sigma_t\sigma_t^\top) = \frac{1}{d^d}$ for all $t \geq 0$. For $\varepsilon > 0$, define the control $\sigma^\varepsilon \in \mathcal{U}$ by

$$\sigma_t^\varepsilon = \begin{cases} \sigma_t, & |X_t| \in (0, \rho-\varepsilon] \cup [\rho+\varepsilon, R), \\ \frac{1}{|X_t|}\begin{bmatrix} X_t^\perp; & 0; & \cdots; & 0 \end{bmatrix}, & |X_t| \in (\rho-\varepsilon, \rho+\varepsilon), \end{cases}$$

for some $X^\perp$ that satisfies $X_t^\top X_t^\perp = 0$, for all $t \geq 0$.

This control corresponds to following any strategy with high determinant except in an annulus of width 2ε. In this annulus, the controlled process follows tangential motion, as defined in Definition 2.3, and has a deterministically increasing radius, ensuring that the process does not return to the inner ball.

Since Z^{σ^ε} is deterministically increasing in the interval $((\rho-\varepsilon)^2, (\rho+\varepsilon)^2)$, the local time at ρ^2 is $L_t^{\sigma^\varepsilon,\rho^2} = 0$, for all $t \geq 0$. Also note that, for $|X_t| \in (\rho-\varepsilon, \rho+\varepsilon)$, we have

$$\det\left(\sigma_t^\varepsilon \sigma_t^{\varepsilon\top}\right) = 0.$$

Hence, applying the optional sampling theorem to (5.20), for any $\varepsilon > 0$, we have that

$$\begin{aligned} w(\xi) = \mathbb{E}^\xi &\left[w\left(Z_\tau^{\sigma^\varepsilon}\right) + d\int_0^\tau f\left(X_s^{\sigma^\varepsilon}\right) \det\left(\sigma_s^\varepsilon \sigma_s^{\varepsilon\top}\right)^{\frac{1}{d}} \mathrm{d}s \right] \\ &- \mathbb{E}^\xi \left[\int_0^\tau \mathbb{1}_{\left\{Z_s^{\sigma^\varepsilon} \in ((\rho-\varepsilon)^2, (\rho+\varepsilon)^2)\right\}} \mathrm{d}s \right]. \end{aligned}$$

We can make similar calculations to those in the proof of Proposition 2.5 to find that, for any $t \geq 0$,

$$\begin{aligned} 0 \leq \mathbb{E}^\xi \left[\int_0^t \mathbb{1}_{\left\{Z_s^{\sigma^\varepsilon} \in ((\rho-\varepsilon)^2, (\rho+\varepsilon)^2)\right\}} \mathrm{d}s \right] \leq \int_0^{(\rho+\varepsilon)^2-(\rho-\varepsilon^2)} \mathrm{d}s &= (\rho+\varepsilon)^2 - (\rho-\varepsilon)^2 \\ &= 4\rho\varepsilon \xrightarrow{\varepsilon\downarrow 0} 0. \end{aligned}$$

Therefore taking the limit as $\varepsilon \downarrow 0$ gives

$$\begin{aligned} w(\xi) &= \lim_{\varepsilon\downarrow 0} \mathbb{E}^\xi \left[w\left(Z_\tau^{\sigma^\varepsilon}\right) + d\int_0^\tau f\left(X_s^{\sigma^\varepsilon}\right) \det\left(\sigma_s^\varepsilon \sigma_s^{\varepsilon\top}\right)^{\frac{1}{d}} \mathrm{d}s \right] \\ &= \lim_{\varepsilon\downarrow 0} \mathbb{E}^\xi \left[d\int_0^\tau f\left(X_s^{\sigma^\varepsilon}\right) \det\left(\sigma_s^\varepsilon \sigma_s^{\varepsilon\top}\right)^{\frac{1}{d}} \mathrm{d}s \right], \end{aligned}$$

and we conclude that

$$v_{\mathrm{FJ}}(x) = w(|x|^2). \qquad \square$$

We now show that the value function v_{G} coincides with v_{FJ} and v.

Proposition 5.43. *In Example 5.41, the value function v_{G} is given by*

$$v_{\mathrm{G}}(x) = \begin{cases} \rho^2 - R^2, & |x| \leq \rho, \\ |x|^x - R^2, & |x| \in (\rho, R), \end{cases}$$

and so $v^S = v^W = v_{\mathrm{G}}^W = v_{\mathrm{G}}^S = v_{\mathrm{FJ}}^S = v_{\mathrm{FJ}}^W$ in D.

We aim to replicate tangential motion on the internal boundary $\{x \in D \colon |x| = \rho\}$ with some control $\sigma \in \mathcal{U}_{\mathrm{G}}$. To satisfy the determinant constraint, we consider controls of the following form.

Let $x \in D$ and let $y_1, \dots, y_d \in \mathbb{R}^d$ be orthogonal vectors with $y_1 = x$ and $y_i^\top y_j = \delta_{ij} |x|^2$, for each $i, j = 1, \cdots, d$. Let $\lambda = (\lambda_1, \dots, \lambda_d)^\top \in (0, \infty)^d$ satisfy $\prod_{i=1}^d \lambda_i^2 = \frac{1}{d^d}$, and define

$$\sigma^\lambda(x) = \frac{1}{|x|} \Big[\lambda_1 y_1; \quad \cdots; \quad \lambda_d y_d \Big].$$

Then $\det\left(\sigma^\lambda(x)\sigma^\lambda(x)^\top\right) = \frac{1}{d^d}$, and so $\sigma^\lambda(x) \in U_{\mathrm{G}}$. Note also that we have

$$\operatorname{Tr}\left(\sigma^\lambda(x)\sigma^\lambda(x)^\top\right) = \sum_{i=1}^d \lambda_i^2. \tag{5.22}$$

Taking λ_1 small, the control $\sigma^\lambda(X_t)$ concentrates the controlled process around the subspace orthogonal to its radius, thus approximating tangential motion. However, the constraint $\prod_{i=1}^d \lambda_i^2 = \frac{1}{d^d}$ combined with the AM-GM inequality implies that the trace given in (5.22) becomes large. This means that the process has high quadratic variation and leaves the domain in a shorter time. Since the cost f is negative, spending less time in the domain results in a higher cost, which is undesirable.

We therefore need a trade-off between how well we approximate tangential motion and how low we keep the quadratic variation. To do this, we fix an annulus around the internal boundary, inside which we take a small value of λ_1, as shown in Figure 5.1. Outside of this annulus we will take all elements of λ to be equal, in order to minimise the quadratic variation. The scaling that leads to an optimising sequence of controls is to take the width of the annulus and the parameter λ_1 to zero at the same rate.

Proof of Proposition 5.43. Once again, define $w : (0, R^2) \to \mathbb{R}$ by

$$w(\xi) = \begin{cases} \rho^2 - R^2, & |x| \leq \rho, \\ \xi - R^2, & |x| \in (\rho, R). \end{cases}$$

Fix $\sigma \in \mathcal{U}_{\mathrm{G}}$. Following the proof of Lemma 2.2, we find that the squared radius process defined by $Z_t^\sigma := |X_t^\sigma|^2$, for $t \geq 0$, satisfies the SDE

$$\mathrm{d}Z_t^\sigma = 2X_t^\top \sigma_t \,\mathrm{d}B_t + \operatorname{Tr}(\sigma_t \sigma_t^\top)\,\mathrm{d}t.$$

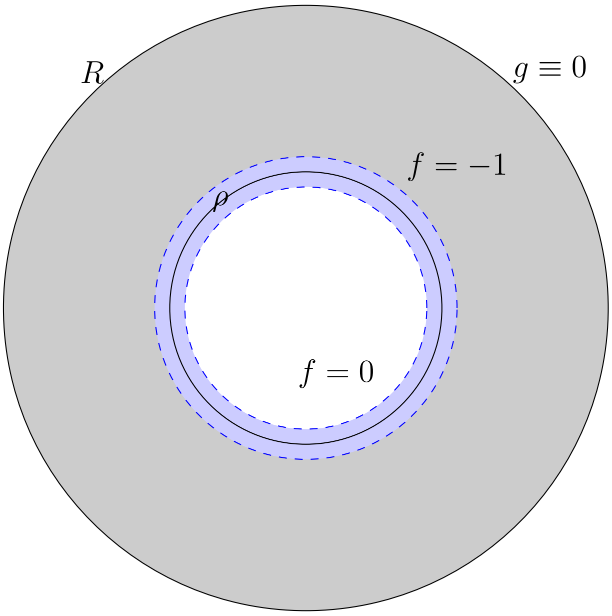


Figure 5.1: Cost function for Example 5.41 with the annulus used to define a minimising sequence of controls highlighted

Fixing $\xi \in [0, R^2)$, we can apply the Itô-Tanaka formula to find that, for any $t \geq 0$,

$$w(\xi) = \mathbb{E}^{\xi}\left[w(Z_t^{\sigma}) - \int_0^t \mathbb{1}_{\{Z_t^{\sigma} \in (\rho^2, R^2)\}} \operatorname{Tr}(\sigma_s \sigma_s^{\top}) \, \mathrm{d}s + \frac{1}{2} L_t^{\sigma, \rho^2}\right]. \tag{5.23}$$

By Lemma 5.1, we have $\operatorname{Tr}(\sigma_s \sigma_s^{\top}) \geq 1$, since $\sigma_s \in U_{\mathrm{G}}$, for each $s \geq 0$. Therefore, for any $t \geq 0$,

$$w(\xi) \leq \mathbb{E}^{\xi}\left[w(Z_t^{\sigma}) - \int_0^t \mathbb{1}_{\{Z_t^{\sigma} \in (\rho^2, R^2)\}} \, \mathrm{d}s\right].$$

We now seek a minimising sequence of controls. Fix $\delta \in (0, \min\{\rho^2, R^2 - \rho^2\})$ and define the control $\nu^{\delta} \in \mathcal{U}_{\mathrm{G}}$ as follows. Define $\lambda = (\lambda_1, \ldots, \lambda_d)^{\top}$ by

$$\lambda_1 = \delta, \quad \lambda_i = \left(\frac{1}{\delta^2 d^d}\right)^{\frac{1}{2(d-1)}},$$

and define $\hat{\lambda} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_d)^{\top}$ by

$$\hat{\lambda}_i = \frac{\sqrt{d}}{d}, \quad i = 1, \ldots, d.$$

Then let

$$\nu_t^{\delta} = \begin{cases} \sigma^{\lambda}(X_t), & |X_t|^2 \in (\rho^2 - \delta, \rho^2 + \delta), \\ \sigma^{\hat{\lambda}}(X_t), & |X_t|^2 \in (0, \rho^2 - \delta] \cup [\rho^2 + \delta). \end{cases}$$

This choice of control corresponds to speeding up the process and concentrating its path around the subspace orthogonal to the radius when the process is close to the internal boundary.

Fix $t \geq 0$. We claim that

$$\lim_{\delta \downarrow 0} \mathbb{E}^{\xi} \left[w(Z_t^{\nu^\delta}) - \int_0^t \mathbb{1}_{\left\{ Z_s^{\nu^\delta} \in (\rho^2, R^2) \right\}} \, \mathrm{d}s \right] = w(\xi).$$

Starting from (5.23), we need to show that

$$\lim_{\delta \downarrow 0} \mathbb{E}^{\xi} \left[\int_0^t \mathbb{1}_{\left\{ Z_s^{\nu^\delta} \in (\rho^2, R^2) \right\}} \left(\mathrm{Tr} \left(\nu_s^\delta \nu_s^{\delta \top} \right) - 1 \right) \mathrm{d}s \right] = 0,$$

and

$$\lim_{\delta \downarrow 0} \mathbb{E}^{\xi} \left[L_t^{\nu^\delta, \rho^2} \right] = 0.$$

Define the Green's function G and speed measure m for the process Z^{σ^δ} on the interval $[\rho^2 - \delta, \rho^2 + \delta]$, as in Definition B.4 and Definition B.3. Then, by Proposition B.5, we can write

$$\mathbb{E}^{\xi} \left[\int_0^t \mathbb{1}_{\left\{ Z_s^{\nu^\delta} \in (\rho^2, R^2) \right\}} \left(\mathrm{Tr} \left(\nu_s^\delta \nu_s^{\delta \top} \right) - 1 \right) \mathrm{d}s \right] = \left(\sum_{i=1}^d \lambda_i^2 - 1 \right) \int_{\rho^2}^{\rho^2 + \delta} G(\xi, y) m(\mathrm{d}y).$$

Define $\overline{\lambda} = \frac{2\lambda_1^2}{\sum_{i=1}^d \lambda_i^2}$. Then the speed measure m is given by

$$\int m(\mathrm{d}y) = \frac{1}{2\lambda_1^2} \int \exp\left\{ -\frac{c - y}{\overline{\lambda}} \right\} \mathrm{d}y,$$

and the Green's function G is given by

$$G(\xi, y) =$$
$$\begin{cases} \overline{\lambda} \left(\exp\left\{ -\frac{\rho^2 - \delta - c}{\overline{\lambda}} \right\} - \exp\left\{ -\frac{y - c}{\overline{\lambda}} \right\} \right) \frac{\exp\left\{ -\frac{\xi}{\overline{\lambda}} \right\} - \exp\left\{ -\frac{\rho^2 + \delta}{\overline{\lambda}} \right\}}{\exp\left\{ -\frac{\rho^2 - \delta}{\overline{\lambda}} \right\} - \exp\left\{ -\frac{\rho^2 + \delta}{\overline{\lambda}} \right\}}, & y \in (\rho^2 - \delta, \xi], \\ \overline{\lambda} \left(\exp\left\{ -\frac{\rho^2 - \delta - c}{\overline{\lambda}} \right\} - \exp\left\{ -\frac{\xi - c}{\overline{\lambda}} \right\} \right) \frac{\exp\left\{ -\frac{y}{\overline{\lambda}} \right\} - \exp\left\{ -\frac{\rho^2 + \delta}{\overline{\lambda}} \right\}}{\exp\left\{ -\frac{\rho^2 - \delta}{\overline{\lambda}} \right\} - \exp\left\{ -\frac{\rho^2 + \delta}{\overline{\lambda}} \right\}}, & y \in [\xi, \rho^2 + \delta). \end{cases}$$

Suppose that $\xi < \rho^2$. Then we find that

$$\int_{\rho^2}^{\rho^2 + \delta} G(\xi, y) m(\mathrm{d}y) = \frac{1}{\sum_{i=1}^d \lambda_i^2} \frac{1 - \exp\left\{ \frac{\rho^2 - \delta - \xi}{\overline{\lambda}} \right\}}{1 - \exp\left\{ -\frac{2\delta}{\overline{\lambda}} \right\}} \left[\delta - \overline{\lambda} \left(1 - \exp\left\{ -\frac{\delta}{\overline{\lambda}} \right\} \right) \right]. \tag{5.24}$$

We now make use of the choice to scale the parameter λ_1 and the width of the annulus δ so that they go to zero at the same rate. In particular, since we have chosen $\lambda_1 = \delta$, we have the limits

$$\begin{aligned} &\sum_{i=1}^{d} \lambda_i^2 = \delta^2 + (d-1)\left(\frac{1}{d^d \delta^2}\right)^{\frac{1}{d-1}} \xrightarrow{\delta \downarrow 0} +\infty, \\ &\overline{\lambda} = \frac{2\delta^2}{\sum_{i=1}^{d} \lambda_i^2} \xrightarrow{\delta \downarrow 0} 0, \\ \text{and} \qquad &\frac{2\delta}{\overline{\lambda}} = \frac{2\delta \sum_{i=1}^{d} \lambda_i^2}{2\delta^2} = \frac{\sum_{i=1}^{d} \lambda_i^2}{\delta} \xrightarrow{\delta \downarrow 0} +\infty. \end{aligned}$$

Therefore taking the limit $\delta \downarrow 0$ in (5.24) gives

$$\lim_{\delta \downarrow 0} \int_{\rho^2}^{\rho^2+\delta} G(\xi, y) m(\mathrm{d}y) = 0,$$

and so

$$\lim_{\delta \downarrow 0} \mathbb{E}^\xi \left[\int_0^t \mathbb{1}_{\left\{Z_s^{\nu^\delta} \in (\rho^2, R^2)\right\}} \left(\mathrm{Tr}\left(\nu_s^\delta \nu_s^{\delta \top} \right) - 1 \right) \mathrm{d}s \right] = 0.$$

We can make a similar calculation for any value of ξ.

We now consider the local time term. Note that, in the interval $(\rho^2 - \delta, \rho^2 + \delta)$, the quadratic variation of Z^{ν^δ} is given by

$$\mathrm{d}\langle Z^{\nu^\delta} \rangle_t = 4\delta^2 Z_t^{\nu^\delta} \, \mathrm{d}t.$$

Then, by the expression for local time given in Corollary 1.9 of [51, Chapter VI], we have

$$0 \le \mathbb{E}^\xi \left[L_t^{\nu^\delta, \rho^2} \right] = \lim_{\varepsilon \downarrow 0} \frac{4\delta^2}{\varepsilon} \mathbb{E}^\xi \left[\int_0^t \mathbb{1}_{\left\{Z_s^{\nu^\delta} \in [\rho^2, \rho^2+\varepsilon)\right\}} Z_s^{\nu^\delta} \, \mathrm{d}s \right].$$

Using Proposition B.5 again, we can rewrite this as

$$0 \le \mathbb{E}^\xi \left[L_t^{\nu^\delta, \rho^2} \right] = \lim_{\varepsilon \downarrow 0} \frac{4\delta^2}{\varepsilon} \int_{\rho^2}^{\rho^2+\varepsilon} G(\xi, y) y \, m(\mathrm{d}y).$$

For $\xi < \rho^2$, we calculate that

$$\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_{\rho^2}^{\rho^2+\varepsilon} G(\xi, y) y \, m(\mathrm{d}y) = \frac{\rho^2}{\sum_{i=1}^{d} \lambda_i^2} \frac{1 - \exp\left\{ \frac{\rho^2 - \delta - \xi}{\overline{\lambda}} \right\}}{1 - \exp\left\{ -\frac{2\delta}{\overline{\lambda}} \right\}},$$

and so, taking the limit as $\delta \downarrow 0$, we have

$$\lim_{\delta \downarrow 0} \mathbb{E}^{\xi}\left[L_t^{\nu^{\delta},\rho^2}\right] = 0.$$

Once again, we can make a similar calculation for any value of ξ.

Applying the optional sampling theorem to (5.23), we conclude that

$$\begin{aligned} w(\xi) &= \inf_{\sigma \in \mathcal{U}_{\mathrm{G}}} \mathbb{E}^{\xi}\left[w(Z_\tau^\sigma) - \int_0^\tau \mathbb{1}_{\{Z_\tau^\sigma \in (\rho^2, R^2)\}}\,\mathrm{d}s\right] \\ &= \inf_{\sigma \in \mathcal{U}_{\mathrm{G}}} \mathbb{E}^{\xi}\left[\int_0^\tau -\mathbb{1}_{\{Z_\tau^\sigma \in (\rho^2, R^2)\}}\,\mathrm{d}s\right], \end{aligned}$$

and so

$$v_{\mathrm{G}}(x) = w(|x|^2).$$

In Proposition 2.5, we proved that $v(x) = w(|x|^2)$, and in Proposition 5.42 we also proved that $v_{\mathrm{FJ}}(x) = w(|x|^2)$. Hence we have the equality

$$v = v_{\mathrm{FJ}} = v_{\mathrm{G}}.$$

□

We now show that equality between the value functions does not always hold. Restricting ourselves to two dimensions, the next example shows that, for a smooth decreasing cost function, we cannot have equality unless the cost is constant.

Proposition 5.44. *Fix $d = 2$ and $R > 0$, and let $D = B_R(0) \subset \mathbb{R}^2$. Let $\tilde{f} : [0, R) \to \mathbb{R}_-$ be a continuously differentiable decreasing function and define $f : D \to \mathbb{R}$ by $f(x) = \tilde{f}(|x|)$, for $x \in D$. Suppose moreover that f is not constant on D, and set $g \equiv 0$. Then, for any $x \in D$,*

$$v(x) = 2\int_{|x|}^{R} \tilde{f}(s)s\,\mathrm{d}s,$$

and there exists $x' \in D$ such that

$$v_{\mathrm{FJ}}(x') = v_{\mathrm{G}}(x') > v(x').$$

Proof. We first verify the form of the value function v, noting that we can apply Proposition 2.15 to see that $v = V$, where V is the candidate value function defined in Definition 2.14.

Since the cost function $\tilde{f}$ is decreasing on the whole interval $(0, R)$, the function V is defined in Case II of Definition 2.14, with $r_0 = 0$ and $R \in (0, s_0)$. Therefore, by Proposition 2.15 we have that

$$v(x) = V(x) = 2\int_{|x|}^{R} \tilde{f}(s)s\,\mathrm{d}s,$$

for $x \in D$.

We now show that the value function v_{G} is not equal to v. Corollary 5.37 states that v_{G} is a convex viscosity solution of the Monge-Ampère equation (5.1). We will show that v does not solve the Monge-Ampère equation.

Under the assumption that $\tilde{f}$ is continuously differentiable, v is twice continuously differentiable and we can calculate

$$\begin{aligned} D^2 v(x) &= -2\tilde{f}(|x|)I - 2\tilde{f}'(|x|)\,|x|^{-1}\,xx^\top \\ &= -\frac{2}{|x|^2}\begin{bmatrix} |x|\,\tilde{f}(|x|) + \tilde{f}'(|x|)x_1^2 & \tilde{f}'(|x|)x_1x_2 \\ \tilde{f}'(|x|)x_1x_2 & |x|\,\tilde{f}(|x|) + \tilde{f}'(|x|)x_2^2 \end{bmatrix}. \end{aligned}$$

Therefore

$$\begin{aligned} \det\big(D^2 v(x)\big) &= [-2\tilde{f}(|x|)]^2 + 4\,|x|\,\tilde{f}(|x|)\tilde{f}'(|x|) \\ &\geq [-2\tilde{f}(|x|)]^2, \end{aligned}$$

with equality if and only if either $\tilde{f} = 0$ or $\tilde{f}' = 0$. Under our assumption that f is not constant, there exists $x \in D$ for which we have the strict inequality

$$\det\big(D^2 v(x)\big) > [-2\tilde{f}(|x|)]^2.$$

Therefore v is not a classical solution of the Monge-Ampère equation (5.1) in D. Since v is twice continuously differentiable, this implies that v is not a convex viscosity solution of the Monge-Ampère equation in D. Hence, by Corollary 5.37, there exists $x' \in D$ such that $v_{\mathrm{G}}(x') \neq v(x')$.

Referring to the ordering proved in Theorem 5.40, we have that

$$v(x') < v_{\mathrm{G}}(x') = v_{\mathrm{FJ}}(x'). \qquad \square$$

We conclude by presenting a specific example of a linear cost function that fits into the setup of Proposition 5.44. For this example, we can compute the value function $v_{\mathrm{G}} = v_{\mathrm{FJ}}$ explicitly.

Example 5.45. Fix $d = 2$ and $R > 0$, and let $D = B_R(0) \subset \mathbb{R}^2$. Set $g \equiv 0$ and define $f : D \to \mathbb{R}$ by $f(x) = -\frac{|x|}{R}$, for all $x \in D$. Then

$$v(x) = \frac{2}{3R}\left(|x|^3 - R^3\right), \quad x \in D,$$

and

$$v_{\mathrm{FJ}}(x) = v_{\mathrm{G}}(x) = \frac{\sqrt{2}}{3R}\left(|x|^3 - R^3\right) > v(x), \quad x \in D.$$

Proof. Substituting the form of the cost function f into the value function from Proposition 5.44 above, we find that

$$v(x) = \frac{2}{3R}\left(|x|^3 - R^3\right), \quad x \in D.$$

Now define $V : D \to \mathbb{R}$ to be our candidate for the value function v_{G},

$$V(x) := \frac{\sqrt{2}}{3R}\left(|x|^3 - R^3\right), \quad x \in D.$$

We will show that V is a classical solution of the Monge-Ampère problem (5.14) and then appeal to Corollary 5.37 to prove that $v_{\mathrm{G}} = V$.

For $x \in D$, we calculate

$$D^2V(x) = \frac{\sqrt{2}}{R\,|x|}\left[xx^\top + |x|^2 I\right],$$

and so

$$\begin{aligned}\det(D^2V(x)) &= \frac{2}{R^2\,|x|^2}\left((2x_1^2 + x_2^2)(x_1^2 + 2x_2^2) - x_1^2x_2^2\right)\\ &= \frac{4}{R^2}\,|x|^2 = \left(-2f(x)\right)^2.\end{aligned}$$

For $x_0 \in \partial D$, we have $|x_0| = R$, and so

$$\lim_{x \to x_0} V(x) = 0.$$

Therefore V is a classical solution, and hence a convex viscosity solution, of the Monge-Ampère problem (5.14).

By Corollary 5.37, which states that the value function v_{G} is the unique viscosity solution of (5.14), we conclude that $v_G = V$.

We have now shown that, for $x \in D$,

$$v(x) = \frac{2}{3R}\left(|x|^3 - R^3\right) < \frac{\sqrt{2}}{3R}\left(|x|^3 - R^3\right) = v_{\mathrm{G}}(x) = v_{\mathrm{FJ}}(x),$$

where the final equality follows from Theorem 5.40. $\square$

We have shown that the value functions defined in this chapter are equal to each other and bounded from below by the value function v defined in Section 1.4.1. From Example 5.41, we see that there are instances where all of the value functions coincide. However, Proposition 5.44 shows that this is not always the case.

Appendices

APPENDIX A

COMPARISON PRINCIPLES FOR CONVEX VISCOSITY SOLUTIONS

In Proposition 5.22, we proved uniqueness for the Dirichlet problem (5.9) for the Monge-Ampère equation. We took the comparison principle for the HJB equation (5.3) from Feng and Jensen's paper [24], and then used the equivalence between viscosity solutions of (5.3) and convex viscosity solutions of (5.1), which is also proved in [24]. We noted in Remark 5.23 that alternative methods of proof are possible.

In this appendix, we state and prove two comparison principles for convex viscosity solutions. The first result that we prove is a comparison principle for PDEs that are elliptic on the set of convex functions and satisfy standard assumptions, including coercivity of the differential operator in the zeroth order derivative. The proof of this result requires an adaptation to the standard proof of comparison for viscosity solutions and, in particular, depends on a convex version of the Crandall-Ishii Lemma (Lemma 4.16), which we prove in Appendix A.3. We will take the key idea for this proof from Section V.3 of [36], where Ishii and Lions state a comparison principle for a Monge-Ampère equation.

Our second result is a comparison principle for the Monge-Ampère equation (5.1). This is a special case of Theorem V.2 of [36]. As for the comparison principle for the HJB equation in Proposition 4.19, we relax the coercivity assumption via a perturbation argument, using the perturbation that is suggested in Section V.3 of [36].

A.1 Comparison for PDEs that are elliptic on the set of convex functions

Fix $d \geq 2$ and let $D \subset \mathbb{R}^d$. Consider a differential operator $F : D \times \mathbb{R} \times \mathbb{R}^d \times S_d \to \mathbb{R}$ that satisfies

$$F(x, r, p, X) \leq F(x, r, p, Y) \quad \text{for} \quad X \geq Y \geq 0; \tag{A.1}$$

i.e. the operator F is degenerate elliptic on the set of positive semi-definite matrices. We say that the PDE

$$F\left(x, u(x), Du(x), D^2u(x)\right) = 0 \tag{A.2}$$

is degenerate elliptic on the set of convex functions.

We first show that a comparison principle holds for convex viscosity solutions of the PDE (A.2) under the following assumptions.

Assumption A.1. Suppose that the following assumptions hold.

1. The domain D is open, bounded and *convex*;
2. The operator F is continuous in each of its arguments;
3. The operator F is proper; i.e.
$$F(x, r, p, X) \leq F(x, s, p, X) \quad \text{for} \quad r \leq s;$$
4. The operator F is coercive in the zeroth order derivative; i.e. there exists $\gamma > 0$ such that
$$F(x, s, p, X) - F(x, r, p, X) \geq \gamma(s - r), \quad \text{for} \quad r \leq s; \tag{A.3}$$
5. There exists a function $\omega : [0, \infty] \to [0, \infty]$, with $\omega(0+) = 0$, such that
$$F(y, r, \alpha(x - y), Y) - F(x, r, \alpha(x - y), X) \leq \omega(\alpha\, |x - y|^2 + |x - y|),$$
for any $\alpha > 0$, whenever X and Y are *non-negative definite* and satisfy the following matrix inequality:
$$-3\alpha \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \leq \begin{bmatrix} X & 0 \\ 0 & -Y \end{bmatrix} \leq 3\alpha \begin{bmatrix} I & -I \\ -I & I \end{bmatrix}. \tag{A.4}$$

These assumptions agree with Assumption 4.10, except for the additional requirement of convexity of the domain in the first statement, and the relaxation of the fifth statement to consider only non-negative definite matrices.

The comparison principle that we prove is a consequence of the following adaptation to the Crandall-Ishii Lemma (Lemma 4.16).

Lemma A.2. *Let $D \subset \mathbb{R}^d$ be open, convex and locally compact, let $u_1, u_2 \in \mathrm{USC}(D)$, with u_1* convex *and u_2* concave*, and suppose that the assumptions of Lemma 4.16 are satisfied. Then, for every $\varepsilon > 0$, there exist matrices $X_1, X_2 \in S_d$, with X_1* positive semi-definite *and X_2* negative semi-definite*, such that conditions* (4.18), (4.19) *in Lemma 4.16 hold.*

Compared with Lemma 4.16, we require the additional assumption that u_1 is convex and u_2 is concave, and we get the additional result that X_1 is positive semi-definite and X_2 is negative semi-definite.

We now show that this result is exactly what we require to prove comparison, delaying the proof of the lemma until Appendix A.3. The following theorem is the convex analogue of Theorem 4.12.

Theorem A.3 (Comparison for convex solutions)**.** *Let $D \subset \mathbb{R}$ and $F : D \times \mathbb{R} \times \mathbb{R}^d \times S_d$ be such that Assumption A.1 is satisfied and the ellipticity condition* (A.1) *for positive definite matrices holds. Suppose that*

$$
\begin{aligned}
u &\in \mathrm{USC}(\overline{D}) \quad \textit{is a convex viscosity subsolution of } (\mathrm{A.2}),\\
v &\in \mathrm{LSC}(\overline{D}) \quad \textit{is a convex viscosity supersolution of } (\mathrm{A.2}),
\end{aligned}
$$

and

$$u \le v \quad on \quad \partial D.$$

Then

$$u \le v \quad on \quad \overline{D}.$$

Proof. This is a straightforward adaptation of the proof of the standard comparison principle for viscosity solutions given in Theorem 4.12, which is presented in detail as the proof of Theorem 3.3 in [13]. The following proof therefore has much in common with the proof of Lemma 4.17 and we omit some of the repetitive details.

The proof relies on the variation of the Crandall-Ishii Lemma given by Lemma A.2. We apply this lemma to the function $\varphi : \overline{D}^2 \to \mathbb{R}$, defined as in the proof of Lemma 4.17 by

$$\varphi(x_1, x_2) = \frac{1}{2} x^\top A x, \quad x_1, x_2 \in \overline{D},$$

where

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad A = \alpha \begin{bmatrix} I & -I \\ -I & I \end{bmatrix},$$

for some $\alpha > 0$. We recall that

$$D_{x_1}\varphi(x) = \alpha(x_1 - x_2), \quad D_{x_2}\varphi(x) = \alpha(x_2 - x_1),$$

$$D^2\varphi \equiv A, \quad (D^2\varphi)^2 = A^2 = 2\alpha A,$$

and

$$\left\| D^2\varphi \right\| = \inf \left\{ \left| \xi^\top A \xi \right| : \xi \in \mathbb{R}^{2d}, |\xi| \leq 1 \right\} = 2\alpha.$$

We suppose for contradiction that there exists $z \in D$ such that

$$u(z) - v(z) = \delta, \quad \text{for some} \quad \delta > 0. \tag{A.5}$$

Let $x^\alpha \in \overline{D}^2$ be a maximiser of

$$u(x_1^\alpha) - v(x_2^\alpha) - \frac{\alpha}{2} \left| x_1^\alpha - x_2^\alpha \right|^2,$$

which is guaranteed to exist by compactness of $\overline{D}^2$ and upper semicontinuity of $u - v$. Note that, as shown in [13],

$$\alpha \left| x_1^\alpha - x_2^\alpha \right|^2 \xrightarrow{\alpha \to \infty} 0 \quad \text{and} \quad \left| x_1^\alpha - x_2^\alpha \right| \xrightarrow{\alpha \to \infty} 0.$$

By the same argument as in the proof of Lemma 4.17, we can take α sufficiently large that $x^\alpha \in D^2$.

Now let $\varepsilon > 0$ and set $u_1 = u$, $u_2 = -v$. Since u and v are convex functions, it follows that u_1 is convex and u_2 is concave. Therefore, we can apply Lemma A.2, to see that there exist $X_1, X_2 \in S_d$ such that

$$(\alpha(x_1^\alpha - x_2^\alpha), X_1) \in \overline{J}_D^{2,+} u(x_1^\alpha), \quad (-\alpha(x_1^\alpha - x_2^\alpha), X_2) \in \overline{J}_D^{2,+}(-v)(x_2^\alpha),$$

and

$$-3\alpha \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \leq \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \leq 3\alpha \begin{bmatrix} I & -I \\ -I & I \end{bmatrix},$$

where we have chosen $\varepsilon = \alpha^{-1}$.

Furthermore, Lemma A.2 tells us that $X_1 \geq 0$ and $X_2 \leq 0$. This is the key additional property that we require for the case of convex viscosity solutions. Condition (A.4) in Assumption A.1 is therefore satisfied with $X = X_1$ and $Y = -X_2$ both non-negative definite matrices. So, by the fifth statement of Assumption A.1, there exists a function $\omega\colon [0,\infty] \to [0,\infty]$, with $\omega(0+) = 0$, such that

$$\begin{aligned} &F(x_2, r, \alpha(x_1 - x_2), -X_2) - F(x_1, r, \alpha(x_1 - x_2), X_1) \\ &\qquad \leq \omega\left(\alpha\,|x_1 - x_2|^2 + |x_1 - x_2|\right). \end{aligned} \tag{A.6}$$

Let $\gamma > 0$ be the coercivity constant in (A.3). Then, by the fourth statement of Assumption A.1,

$$\gamma(u(x_1^\alpha) - v(x_2^\alpha)) \leq F(x_1^\alpha, u(x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1) - F(x_1^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1).$$

Since

$$\delta = u(z) - v(z) \leq u(x_1^\alpha) - v(x_2^\alpha) - \frac{\alpha}{2}\,|x_1^\alpha - x_2^\alpha| \leq u(x_1^\alpha) - v(x_2^\alpha),$$

we have

$$\delta\gamma \leq F(x_1^\alpha, u(x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1) - F(x_1^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1). \tag{A.7}$$

Now, since u is a convex viscosity subsolution of (A.2), and $(\alpha(x_1^\alpha - x_2^\alpha), X_1) \in \overline{J}_D^{2,+}u(x_1^\alpha)$, with $X_1 \geq 0$, we have

$$F(x_1^\alpha, u(x_1^\alpha), \alpha(x_1^\alpha - x_2^\alpha), X_1) \leq 0,$$

by Definition 5.17. We also have that

$$(\alpha(x_1^\alpha - x_2^\alpha), -X_2) \in \overline{J}_D^{2,-}v(x_2^\alpha),$$

and so, as v is a convex viscosity supersolution of (A.2) and $-X_2 \geq 0$, Definition 5.17 gives the inequality

$$F(x_2^\alpha, v(x_2^\alpha), \alpha(x_1^\alpha - x_2^\alpha), -X_2) \geq 0.$$

We can now substitute the above inequalities into (A.7) and then apply (A.6), as in the proof of Lemma 4.17, to arrive at

$$\delta\gamma \leq \omega\left(\alpha\,|x_1^\alpha - x_2^\alpha|^2 + |x_1^\alpha - x_2^\alpha|\right).$$

Since $\alpha\,|x_1^\alpha - x_2^\alpha|^2 \to 0$, as $\alpha \to \infty$, and $\omega(0+) = 0$, we can take the limit as $\alpha \to \infty$ to find that

$$\delta\gamma \leq 0.$$

This contradicts our assumption (A.5), and so we conclude that $u \leq v$ on $\overline{D}$. $\square$

A.2 Comparison for a Monge-Ampère equation

We now turn to proving a comparison principle for convex viscosity solutions of the Monge-Ampère equation (5.1). Let $f : D \to (-\infty, 0]$ be a continuous function. The Monge-Ampère operator F, defined by

$$F(x, u, p, X) \equiv F(x, X) := -\det(X) + (-2f(x))^d,$$

does not satisfy the coercivity condition (A.3) in Assumption A.1, since there is no dependence on the zeroth order derivative. Therefore we cannot apply Theorem A.3 directly to the Monge-Ampère equation (5.1).

In order to prove comparison for the Monge-Ampère equation, we use the same perturbation technique that we used to prove comparison for the HJB equation (4.7) in Proposition 4.19. In Lemma 4.17, we proved a comparison principle that does not require the coercivity assumption, using the method outlined in Section 5.C of [13]. The following analogue of this result holds for convex viscosity solutions.

Lemma A.4. *Let $D \subset \mathbb{R}^d$ and $F : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times S_d \to \mathbb{R}$ satisfy statements 1, 2, 3 and 5 of Assumption A.1.*

Let $u \in \mathrm{USC}(\overline{D})$ be a convex viscosity subsolution and $v \in \mathrm{LSC}(\overline{D})$ a convex viscosity supersolution of (A.2), *and suppose that*

$$u \leq v \quad \text{on} \quad \partial D.$$

Suppose moreover that, for each $k \in \mathbb{N}$, there exists $\delta_k > 0$ and a function $\psi_k \in C^2(D)$ such that

$$|\psi_k| \leq \frac{1}{k},$$

and $u_k := u + \psi_k$ is a convex viscosity subsolution of

$$F(x, u_k, Du_k, D^2 u_k) + \delta_k = 0.$$

Then

$$u \leq v \quad \text{on} \quad \overline{D}.$$

Proof. The proof of this result is identical to the proof of Lemma 4.17, except for two modifications. Viscosity solutions are replaced by convex viscosity solutions, and Lemma A.2, the convex variant of the Crandall-Ishii Lemma that we prove in Appendix A.3, replaces Lemma 4.16. We do not reproduce the full proof here. □

The following result is a special case of the comparison principle for a Monge-Ampère equation that is stated in Theorem V.2 of [36]. In our result, the function f depends only on the spatial variable $x \in D$, whereas in [36] Ishii and Lions allow dependence on the value and gradient of the solution. We prove the result by applying Lemma A.4 with $(\psi_k)_{k\in\mathbb{N}}$ chosen to be a slight simplification of the perturbation suggested in [36]. We note that the choice of perturbation will be the same as in the proof of Proposition 4.19.

Proposition A.5. *Suppose that Assumption 5.2 holds and that the function $f : D \to \mathbb{R}$ is continuous. Then we have the following comparison principle for the Monge-Ampère equation* (5.1).

Suppose that

$$\begin{aligned} u &\in \mathrm{USC}(\overline{D}) \quad \textit{is a convex viscosity subsolution of } (5.1), \\ v &\in \mathrm{LSC}(\overline{D}) \quad \textit{is a convex viscosity supersolution of } (5.1), \end{aligned}$$

and

$$u \le v \quad \textit{on} \quad \partial D.$$

Then

$$u \le v \quad \textit{on} \quad \overline{D}.$$

The proof follows the same reasoning as the proof of Proposition 4.19.

Proof. We first check that conditions 1, 2, 3 and 5 of Assumption A.1 hold.

1. The domain D is open and bounded by Assumption A.1.

2. We have assumed that f is continuous and so, since the determinant is a continuous function, the operator $F : D \times \mathbb{R} \times \mathbb{R}^d \times S_d$ defined by

$$F(x, r, p, X) \equiv F(x, X) = -\det X + (-2f(x))^d$$

is continuous in each of its arguments.

3. Let $r \le s$, then

$$F(x, r, p, X) - F(x, s, p, X) \equiv F(x, X) - F(x, X) = 0,$$

and so F is proper.

5. Define $G : S_d \to \mathbb{R}$ by $G(X) = -\det X$. Then the operator F is of the form

$$F(x, r, p, X) = G(X) + (-2f(x))^d.$$

By Lemma 5.3, F is degenerate elliptic on the set of non-negative definite matrices, in the sense of (5.6). Then, by the same reasoning as in Remark 4.11, the fifth statement of Assumption A.1 holds.

Now, since the coercivity condition in the fourth statement of Assumption A.1 does not hold, we appeal to Lemma A.4. We define the following perturbation to the subsolution, which is the same perturbation used in Proposition 4.19, as suggested in Section V.3 of [36].

Let $m \in \mathbb{N}$ and set $C := \sup_{x \in D} \frac{|x|^2}{2}$. Define $\psi_m : D \to \mathbb{R}$ by

$$\psi_m(x) := \frac{1}{m} \exp\left\{ \frac{|x|^2}{2} - C \right\},$$

and $u_m : D \to \mathbb{R}$ by

$$u_m(x) := u(x) + \psi_m(x),$$

for $x \in D$. Then

$$|\psi_m(x)| \leq \frac{1}{m} \exp\{0\} = \frac{1}{m}.$$

We now need to show that there exists $\delta_m > 0$ such that u_m is a convex viscosity subsolution of the PDE

$$-\det(D^2 u_m) + (-2f)^d + \delta_m = 0.$$

The following section of the proof differs from that of Proposition 4.19. Fix $x^0 \in D$ and let $\phi \in C^\infty(D)$ be such that $x^0 \in \arg\max(u_m - \phi)$.

Then, since $\psi_m \in C^\infty$ and $\psi_m \geq 0$, we have that $(\phi - \psi_m) \in C^\infty(D)$ and $x^0 \in \arg\max(u - (\phi - \psi_m))$. As noted in Remark 5.16, $(\phi - \psi_m)$ is necessarily a convex function, and so, since u is a convex viscosity subsolution of (5.1), we have that

$$-\det\left(D^2\phi(x^0) - D^2\psi_m(x^0)\right) + \left(-2f(x^0)\right)^d \leq 0, \tag{A.8}$$

by Definition 5.15.

Now recall that Lemma 5.4 states that, for any $d \times d$ symmetric positive semi-definite matrices A and B,

$$\det(A+B) \geq \det(A) + \det(B). \tag{A.9}$$

We can take A and B to be the Hessian matrices of the functions ψ_m and $(\phi - \psi_m)$, respectively, since both functions are smooth and convex. Then, by (A.9),

$$\begin{aligned}\det\left(D^2\phi(x^0)\right) &= \det\left(D^2\left(\phi(x^0) - \psi_m(x^0)\right) + D^2\psi_m(x^0)\right) \\ &\geq \det\left(D^2\left(\phi(x^0) - \psi_m(x^0)\right)\right) + \det\left(D^2\psi_m(x^0)\right),\end{aligned}$$

and so

$$-\det\left(D^2\phi(x^0) - D^2\psi_m(x^0)\right) \geq -\det\left(D^2\phi(x^0)\right) + \det\left(D^2\psi_m(x^0)\right).$$

Hence, by (A.8), we have

$$-\det\left(D^2\phi(x^0)\right) + \left(-2f(x^0)\right)^d + \det\left(D^2\psi_m(x^0)\right) \leq 0. \tag{A.10}$$

In the proof of Proposition 4.19, we caclulated that, for any $x \in D$,

$$D^2\psi_m(x) = \frac{1}{m}\exp\left\{\frac{|x|^2}{2} - C\right\}\left(I + xx^\top\right),$$

and so, using the inequality (A.9) once more, we have

$$\begin{aligned}\det\left(D^2\psi_m(x^0)\right) &\geq \frac{1}{m^d}\exp\left\{d\left(\frac{|x^0|^2}{2} - C\right)\right\}\left(1 + \det\left(x^0(x^0)^\top\right)\right) \\ &\geq \frac{1}{m^d}\exp\left\{-dC\right\}.\end{aligned}$$

Defining $\delta_m := \frac{1}{m^d}\exp\{-dC\}$, the inequality (A.10) implies that

$$-\det\left(D^2\phi(x^0)\right) + \left(-2f(x^0)\right)^d + \delta_m \leq 0.$$

Hence u_m satisfies the required subsolution property.

By Lemma A.4, we conclude that comparison holds. □

Note that the above comparison principle depends on Lemma A.4, which depends in turn on Lemma A.2, the convex variant of the Crandall-Ishii Lemma. We prove Lemma A.2 in the following section.

A.3 Proof of a convex Crandall-Ishii Lemma

We now prove Lemma A.2, the adaptation of the Crandall-Ishii Lemma to the case of convex viscosity solutions. This will complete the proofs of the preceding results in this appendix. We follow the proof of the standard Crandall-Ishii Lemma (Lemma 4.16) that is given in the appendix of Crandall, Ishii and Lions's User's Guide [13]. To adapt this proof to convex viscosity solutions, we use the observation from Section V.3 of [36] on concavity of the sup-convolution of a concave function. We now define the sup-convolution, as in the appendix of [13], and state some of its important properties.

Definition A.6 (Sup-convolution)**.** Let $u : \mathbb{R}^d \to \mathbb{R}$ and $\lambda > 0$. We define the λ-sup-convolution $\hat{u}^\lambda : \mathbb{R}^d \to \mathbb{R}$ by

$$\hat{u}^\lambda(x) := \sup_{y\in\mathbb{R}^d} \left\{ u(y) - \frac{\lambda}{2} |x-y|^2 \right\}.$$

We give the straightforward proof of semiconvexity of the sup-convolution, as in [13], in the following lemma. Recall from Definition 1.10 that, for $\lambda > 0$, we say that a function $u : \mathbb{R}^d \to \mathbb{R}$ is λ-semiconvex if the map $x \mapsto f(x) + \frac{\lambda}{2} |x|^2$ is convex.

Lemma A.7. *For $\lambda > 0$, the sup-convolution $\hat{u}^\lambda$ is λ-semiconvex.*

Proof. For any $x \in D$, we can write

$$\begin{aligned} \hat{u}^\lambda(x) + \frac{\lambda}{2} |x|^2 &= \sup_{y\in\mathbb{R}^d} \left\{ u(y) + \frac{\lambda}{2} \left(|x|^2 - |x-y|^2\right) \right\} \\ &= \sup_{y\in\mathbb{R}^d} \left\{ u(y) + \frac{\lambda}{2} |y|^2 - \lambda x \cdot y \right\}. \end{aligned}$$

The right-hand side is the supremum over functions that are linear in x, and is therefore a convex function. This proves that $\hat{u}^\lambda$ is λ-semiconvex. ☐

We now state two lemmas from the appendix of [13] without proof.

The following lemma on semiconvexity is taken from Lemma A.4 of [13], where a proof is given using Aleksandrov's Theorem and Jensen's Lemma on semiconvex functions. These additional lemmas are proved in Lemma A.2 and Lemma A.3 of [13], respectively.

Lemma A.8. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be λ-semiconvex, for some $\lambda > 0$. Let $B \in S_d$ and suppose that*

$$\max_{x\in\mathbb{R}^d} \left\{ f(x) - \frac{1}{2} x^\top B x \right\} = f(0).$$

Then there exists $X \in S_d$ *such that*

$$(0, X) \in \overline{J}^2 f(0) \quad \text{and} \quad -\lambda I \leq X \leq B.$$

The next lemma is taken from Lemma A.5 of [13], where it is referred to as the *magical property of the sup-convolution.* Again we refer to [13] for the proof.

Lemma A.9 (Magical property of the sup-convolution)**.** *Fix* $\lambda > 0$. *Let* $x, p \in \mathbb{R}^d$ *and* $X \in S_d$, *and let* $u \in \mathrm{USC}(\mathbb{R}^d)$ *be bounded above. Suppose that* $(p, X) \in J^{2,+}\hat{u}^\lambda(x)$. *Then*

$$(p, X) \in J^{2,+}u\left(x + \frac{p}{\lambda}\right) \quad \text{and} \quad \hat{u}^\lambda(x) = u\left(x + \frac{p}{\lambda}\right) - \frac{1}{2\lambda}|p|^2.$$

As a consequence,

$$(0, X) \in \overline{J}^{2,+}\hat{u}^\lambda(0) \quad \text{implies that} \quad (0, X) \in \overline{J}^{2,+}u(0).$$

Before turning to the proof of the convex variant of the Crandall-Ishii Lemma, we prove the following lemma, which we will apply to a sum of sup-convolutions in the proof of Lemma A.2.

Lemma A.10. *Let* $u_1, u_2 : D \to \mathbb{R}$ *and define* $w : D^2 \to \mathbb{R}$ *by* $w(x) := u_1(x_1) + u_2(x_2)$, *for* $x = (x_1, x_2)^\top \in D^2$.

Suppose that $X \in S_{2d}$ *is such that*

$$(0, X) \in \overline{J}^2 w(0).$$

Then X *is block-diagonal with*

$$X = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix},$$

and

$$(0, X_1) \in \overline{J}^2 u_1(0) \quad \text{and} \quad (0, X_2) \in \overline{J}^2 u_2(0).$$

Proof. Let $(0, X) \in \overline{J}^2 w(0)$. Then there exist sequences $(x^n)_{n\in\mathbb{N}}$, $(p^n)_{n\in\mathbb{N}}$ and $(X^n)_{n\in\mathbb{N}}$ such that

$$(p^n, X^n) \in J^2 w(x^n), \quad \text{for all} \quad n \in \mathbb{N},$$

and

$$(x^n, w(x^n), p^n, X^n) \xrightarrow{n\to\infty} (0, w(0), 0, X).$$

For a given $n \in \mathbb{N}$, let $x_1^n, x_2^n, p_1^n, p_2^n \in \mathbb{R}^d$, $X_1^n, X_2^n \in S_d$ and $X_{12}^n \in \mathbb{R}^{d,d}$ be such that

$$x^n = \begin{bmatrix} x_1^n \\ x_2^n \end{bmatrix}, \quad p^n = \begin{bmatrix} p_1^n \\ p_2^n \end{bmatrix}, \quad X^n = \begin{bmatrix} X_1^n & X_{12}^n \\ X_{12}^{n\top} & X_2^n \end{bmatrix}.$$

Also write

$$X = \begin{bmatrix} X_1 & X_{12} \\ X_{12}^\top & X_2 \end{bmatrix},$$

for matrices $X_1, X_2 \in S_d$ and $X_{12} \in \mathbb{R}^{d,d}$.

Noting that all matrix norms are equivalent, and making use of the Frobenius norm, $\lim_{n\to\infty} X^n = X$ implies that

$$\left\| \begin{bmatrix} X_1 - X_1^n & X_{12} - X_{12}^n \\ X_{12}^\top - X_{12}^{n\top} & X_2 - X_2^n \end{bmatrix} \right\|^2 = \|X_1 - X_1^n\|^2 + \|X_2 - X_2^n\|^2 + 2\,\|X_{12} - X_{12}^n\|^2$$
$$\xrightarrow{n\to\infty} 0,$$

and so

$$\lim_{n\to\infty} X_1^n = X_1, \quad \lim_{n\to\infty} X_2^n = X_2, \quad \text{and} \quad \lim_{n\to\infty} X_{12}^n = X_{12}.$$

Now, for each $n \in \mathbb{N}$, since $(p^n, X^n) \in J^2 w(x^n)$, we have that

$$w(x) = w(x^n) + (p^n)^\top (x - x^n) + (x - x^n)^\top X^n (x - x^n) + o(|x - x^n|^2),$$

as $x \to x^n$. We can write this as

$$\begin{aligned} u_1(x_1) + u_2(x_2) = u_1(x_1^n) &+ u_2(x_2^n) + (p_1^n)^\top (x_1 - x_1^n) + (p_2^n)^\top (x_2 - x_2^n) \\ &+ \frac{1}{2}(x_1 - x_1^n)^\top X_1^n (x_1 - x_1^n) + \frac{1}{2}(x_2 - x_2^n)^\top X_2^n (x_2 - x_2^n) \\ &+ (x_1 - x_1^n)^\top X_{12}^n (x_2 - x_2^n) + o(|x_1 - x_1^n|^2 + |x_2 - x_2^n|^2), \end{aligned} \tag{A.11}$$

as $x_1 \to x_1^n$ and $x_2 \to x_2^n$.

Choosing $x_2 = x_2^n$, the above equation gives us

$$u_1(x_1) = u_1(x_1^n) + (p_1^n)^\top (x_1 - x_1^n) + \frac{1}{2}(x_1 - x_1^n)^\top X_1^n (x_1 - x_1^n) + o(|x_1 - x_1^n|^2), \tag{A.12}$$

as $x_1 \to x_1^n$. Therefore $(p_1^n, X_1^n) \in J^2 u_1(x_1^n)$.

We know that $(x_1^n, u_1(x_1^n), p_1^n, X_1^n) \to (0, u_1(0), 0, X_1)$, as $n \to \infty$, and so

$$(0, X_1) \in \overline{J}^2 u_1(0).$$

Similarly, choosing $x_1 = x_1^n$, we have

$$u_2(x_2) = u_2(x_2^n) + (p_2^n)^\top (x_2 - x_2^n) + \frac{1}{2}(x_2 - x_2^n)^\top X_2^n (x_2 - x_2^n) + o(|x_2 - x_2^n|^2), \quad \text{(A.13)}$$

as $x_2 \to x_2^n$, and so $(p_2^n, X_2^n) \in J^2 u_2(x_2^n)$.

Since $(x_1^n, u_2(x_2^n), p_2^n, X_2^n) \to (0, u_2(0), 0, X_2)$, as $n \to \infty$, we have

$$(0, X_2) \in \overline{J}^2 u_2(0).$$

Finally, we verify that X is block diagonal. Combining (A.12) and (A.13), we have

$$\begin{aligned} u_1(x_1) + u_1(x_2) = {} & u_1(x_1^n) + u_2(x_2^n) + (p_1^n)^\top (x_1 - x_1^n) + (p_2^n)^\top (x_2 - x_2^n) \\ & + \frac{1}{2}(x_1 - x_1^n)^\top X_1^n (x_1 - x_1^n) + \frac{1}{2}(x_2 - x_2^n)^\top X_2^n (x_2 - x_2^n) \\ & + o(|x_1 - x_1^n|^2) + o(|x_2 - x_2^n|^2), \end{aligned}$$

as $x_1 \to x_1^n$ and $x_2 \to x_2^n$. Comparing this to (A.11), we must have that

$$(x_1 - x_1^n)^\top X_{12}^n (x_2 - x_2^n) = o(|x_1 - x_1^n|^2 + |x_2 - x_2^n|^2),$$

as $x_1 \to x_1^n$ and $x_2 \to x_2^n$, which only holds for $X_{12}^n = 0$.

Since $X_{12}^n = 0$ for all $n \in \mathbb{N}$, and $\lim_{n\to\infty} X_{12}^n = X_{12}$, we have that $X_{12} = 0$. Hence

$$X = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix},$$

as required. □

We now use the preceding lemmas to prove the adaptation of the Crandall-Ishii Lemma to the case of convex viscosity solutions. We follow the proof of the standard Crandall-Ishii Lemma given in the appendix of [13].

Proof of Lemma A.2. The key step in adapting the proof to the case of convex viscosity solutions is the observation made in Section V.3 of [36] that, since u_2 is concave, the sup-convolution $\hat{u}_2^\lambda$ defined in Definition A.6 is also concave, for any $\lambda > 0$. We first prove this assertion.

Let $u : \mathbb{R}^d \to \mathbb{R}$ be a concave function and let $\lambda > 0$. Recall from Definition A.6 that the λ-sup-convolution $\hat{u}^\lambda : \mathbb{R}^d \to \mathbb{R}$ is defined by

$$\hat{u}^\lambda(\xi) := \sup_{y \in D} \left\{ u(y) - \frac{\lambda}{2} |\xi - y|^2 \right\}, \quad \text{for} \quad \xi \in \mathbb{R}^d.$$

Take $\xi_1, \xi_2 \in \mathbb{R}^d$ and let $\rho \in [0,1]$. Set $\overline{\xi} := \rho\xi_1 + (1-\rho)\xi_2$. We wish to show that

$$\rho\hat{u}^\lambda(\xi_1) + (1-\rho)\hat{u}^\lambda(\xi_2) \leq \hat{u}^\lambda(\overline{\xi}).$$

Let $y_1, y_2 \in \mathbb{R}^d$ and define $\overline{y} := \rho y_1 + (1-\rho)y_2$. Since u is concave, we know that

$$\rho u(y_1) + (1-\rho)u(y_2) \leq u(\overline{y}). \tag{A.14}$$

Note that

$$\overline{\xi} - \overline{y} = \rho(\xi_1 - y_1) + (1-\rho)(\xi_2 - y_2).$$

Then, since the map $|\cdot|^2 : \mathbb{R}^d \to \mathbb{R}^+$ is convex, we have

$$\rho\,|\xi_1 - y_1|^2 + (1-\rho)\,|\xi_2 - y_2|^2 \geq \left|\overline{\xi} - \overline{y}\right|^2. \tag{A.15}$$

Now, by definition,

$$\hat{u}^\lambda(\overline{\xi}) \geq u(\overline{y}) - \frac{\lambda}{2}\left|\overline{\xi} - \overline{y}\right|^2.$$

Using the concavity and convexity conditions (A.14) and (A.15), we can bound the right hand side of the above inequality by

$$\begin{aligned} u(\overline{y}) - \frac{\lambda}{2}\left|\overline{\xi} - \overline{y}\right|^2 &\geq \rho u(y_1) + (1-\rho)u(y_2) - \frac{\lambda}{2}\left(\rho\,|\xi_1 - y_1|^2 + (1-\rho)\,|\xi_2 - y_2|^2\right) \\ &= \rho\left(u(y_1) - \frac{\lambda}{2}\,|\xi_1 - y_1|^2\right) + (1-\rho)\left(u(y_2) - \frac{\lambda}{2}\,|\xi_2 - y_2|^2\right). \end{aligned}$$

Since this inequality holds for all $y_1, y_2 \in \mathbb{R}^d$, we can take the supremum over $y_1 \in \mathbb{R}^d$ and $y_2 \in \mathbb{R}^d$ on the right hand side, and we conclude that

$$\hat{u}^\lambda(\overline{\xi}) \geq \rho\hat{u}^\lambda(\xi_1) + (1-\rho)\hat{u}^\lambda(\xi_2),$$

as required.

We now use this fact to prove the lemma, following the same method as the proof of the Crandall-Ishii Lemma (Lemma 4.16) given in the appendix of [13].

Let $u_1, u_2 \in \mathrm{USC}(D)$, with u_1 convex and u_2 concave, and define $w : D^2 \to \mathbb{R}$ by

$$w(x_1, x_2) = u_1(x_1) + u_2(x_2), \quad x_1, x_2 \in D,$$

as in (4.16). Let $x^0 \in D^2$ and $\varphi \in C^2(\overline{D}^2)$ be such that $x^0 \in \arg\max_{D^2}(w - \varphi)$, as in (4.17). As noted in [13], we may assume, without loss of generality, that $D = \mathbb{R}^d$,

$x^0 = 0$, $u_1(0) = u_2(0) = 0$, and

$$\varphi(x) = \frac{1}{2} x^\top A x, \quad \text{for some} \quad A \in S_{2d}.$$

Then we have

$$w(x) - \frac{1}{2} x^\top A x \leq (w - \varphi)(0) = u_1(0) + u_2(0) = 0, \tag{A.16}$$

for any $x \in D^2$.

Fix $\varepsilon > 0$. We aim to find a positive semi-definite matrix X_1 and a negative semi-definite matrix X_2 such that the conditions (4.18) and (4.19) stated in Lemma 4.16 hold.

Let $x, y \in \mathbb{R}^{2d}$. Then, writing

$$\begin{aligned} x^\top A x &= (x-y)^\top A (x-y) - y^\top A y + 2 y^\top A x \\ &= (x-y)^\top A (x-y) + y^\top A y + 2 y^\top A (x-y), \end{aligned}$$

we can use the Cauchy-Schwarz inequality to calculate

$$x^\top A x \leq (x-y)^\top A (x-y) + y^\top A y + 2 \sqrt{\left| \sqrt{\varepsilon} A y \right|^2 \left| \frac{1}{\sqrt{\varepsilon}} (x-y) \right|^2}.$$

Noting that, for any $a, b \in \mathbb{R}_+$, $2\sqrt{ab} \leq a + b$, we then have

$$x^\top A x \leq (x-y)^\top A (x-y) + y^\top A y + \varepsilon y^\top A^2 y + \frac{1}{\varepsilon} |x-y|^2.$$

Using the Cauchy-Schwarz inequality again, we see that

$$(x-y)^\top A (x-y) \leq |A(x-y)| \, |x-y| \leq \|A\| \, |x-y|^2.$$

Therefore, defining $\lambda := \varepsilon^{-1} + \|A\|$, we have

$$x^\top A x \leq \lambda |x-y|^2 + y^\top (A + \varepsilon A^2) y.$$

Hence, by (A.16),

$$w(x) - \frac{\lambda}{2} |x-y|^2 \leq \frac{1}{2} y^\top (A + \varepsilon A^2) y.$$

Now define $B := A + \varepsilon A^2$. Then we can take the supremum over $x \in \mathbb{R}^d$ in the above inequality to see that, for any $y \in \mathbb{R}^d$,

$$\hat{w}^\lambda(y) \leq \frac{1}{2} y^\top B y,$$

and so

$$\hat{w}^\lambda(y) - \frac{1}{2} y^\top B y \leq 0 = \hat{w}^\lambda(0).$$

By Lemma A.7, $\hat{w}^\lambda$ is λ-semiconvex, so we can apply Lemma A.8 to see that there exists $X \in S_{2d}$ such that

$$(0, X) \in \overline{J}^2 \hat{w}^\lambda(0) \quad \text{and} \quad -\lambda I \leq X \leq B. \tag{A.17}$$

We now note that $\hat{w}^\lambda(y) = \hat{u}_1^\lambda(y_1) + \hat{u}_2^\lambda(y_2)$, for $y_1, y_2 \in D$. Therefore, by Lemma A.10, we have

$$(0, X_1) \in \overline{J}^2 \hat{u}_1^\lambda(0) \quad \text{and} \quad (0, X_2) \in \overline{J}^2 \hat{u}_2^\lambda(0),$$

where

$$X = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}$$

is block-diagonal.

We now exploit the fact that $\hat{u}_2^\lambda$ is concave to prove that $X_2 \leq 0$. This is one of the additional statements in the lemma, not present in the classical Crandall-Ishii Lemma stated in Lemma 4.16, that will allow us to apply this result to convex viscosity solutions.

Since $(0, X_2) \in \overline{J}^{2,-} \hat{u}_2^\lambda(0)$, we know that there exist sequences $(x_2^n)_{n\in\mathbb{N}}$, $(p_2^n)_{n\in\mathbb{N}}$ and $(X_2^n)_{n\in\mathbb{N}}$ such that $(p_2^n, X_2^n) \in J^{2,-} \hat{u}_2^\lambda(x_2^n)$, for each $n \in \mathbb{N}$, and

$$\left(x_2^n, \hat{u}_2^\lambda(x_2^n), p_2^n, X_2^n\right) \xrightarrow{n\to\infty} \left(0, \hat{u}_2^\lambda(0), 0, X_2\right).$$

By Proposition 5.20 on equivalence of the two definitions of convex viscosity solutions, there exists $\phi_2^n \in C^\infty(D)$, for each $n \in \mathbb{N}$, such that

$$x_2^n \in \arg\min\left(\hat{u}_2^\lambda - \phi_2^n\right) \quad \text{and} \quad \left(D\phi_2^n(x_2^n), D^2\phi_2^n(x_2^n)\right) = \left(p_2^n, X_2^n\right).$$

This means that each ϕ_2^n sits below the concave function $\hat{u}_2^\lambda$, coming closest at x_2^n. Therefore ϕ_2^n is itself concave at x_2^n. Hence

$$X_2^n = D^2\phi_2^n(x_2^n) \leq 0.$$

We have that $X_2^n \to X_2$, and the set of non-positive definite matrices is closed.

Therefore $X_2 \leq 0$. We also require that X_1 is non-negative definite for this adaptation of the Crandall-Ishii Lemma. We will see later that this property follows directly from convexity of u_1 by a similar argument.

By the magical property of the sup-convolution stated in Lemma A.9, we have that

$$(0, X_1) \in \overline{J}^{2,+} u_1(0) \quad \text{and} \quad (0, X_2) \in \overline{J}^{2,+} u_2(0).$$

Noting that $D_{x_i}\varphi(0) = 0$ for $i = 1, 2$, we have shown that condition (4.18) of Lemma 4.16 holds.

We can now show that $X_1 \geq 0$. Since $(0, X_1) \in \overline{J}^{2,+} u_1(0)$, there exist sequences $(x_1^n)_{n\in\mathbb{N}}$, $(p_1^n)_{n\in\mathbb{N}}$ and $(X_1^n)_{n\in\mathbb{N}}$ such that $(p^n, X^n) \in J^{2,+}u_1(0)$, for each $n \in \mathbb{N}$, and

$$(x_1^n, u_1(x_1^n), p_1^n, X_1^n) \xrightarrow{n\to\infty} (0, u_1(0), 0, X_1).$$

By Proposition 5.20, there exists $\phi_1^n \in C^\infty(D)$, for each $n \in \mathbb{N}$, such that

$$x_1^n \in \arg\max(u_1 - \phi_1^n) \quad \text{and} \quad \left(D\phi_1^n(x_1^n), D^2\phi_1^n(x_1^n)\right) = (p_1^n, X_1^n)\,.$$

This means that each ϕ_1^n sits above the convex function u_1, coming closest at x_1^n, and so ϕ_1^n is itself convex at x_1^n. Hence

$$X_1^n = D^2\phi_1^n(x_1^n) \geq 0.$$

Since the set of non-negative definite matrices is closed and $\lim_{n\to\infty} X_1^n = X_1$, we have that $X_1 \geq 0$.

Finally, using the block-diagonal structure of X, and recalling the definitions $\lambda = \varepsilon^{-1} + \|A\|$ and $B = A + \varepsilon A^2$, the inequality in (A.17) becomes

$$-\left(\varepsilon^{-1} + \|A\|\right) I_{2d} \leq \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \leq A + \varepsilon A^2.$$

Noting that $D^2\varphi \equiv A$, we have shown that condition (4.19) of Lemma 4.16 holds.

We have now shown that (4.18) and (4.19) hold, and that X_1 is non-negative definite and X_2 is non-positive definite, as required. □

APPENDIX B

SCALE FUNCTIONS AND SPEED MEASURES

In Chapter 2 and Chapter 5, we make use of the theory of scale functions and speed measures, as set out, for example, in Section 3 of [51, Chapter VII] and in Section 6 of the lecture notes [23]. In this appendix, we summarise the definitions and results that we use.

Let W be a standard one-dimensional Brownian motion, and let $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}$ be Lipschitz functions. Let X be a one-dimensional diffusion satisfying

$$\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t.$$

The *scale function* of the diffusion X, as defined in Definition 3.3 of [51, Chapter VII], describes how likely the diffusion is to move in either direction. As shown in Exercise 3.20 of [51, Chapter VII], the scale function of X can be written in the following form. Following Definition 6.1 of [23], we take this to be our definition of the scale function.

Definition B.1 (Scale function)**.** Define the scale function $s : \mathbb{R} \to \mathbb{R}$ of the diffusion X by

$$s(x) := \int_c^x \exp\left\{ -\int_c^y 2\mu(z)\sigma^{-2}(z)\,\mathrm{d}z \right\},$$

where $c \in \mathbb{R}$ is arbitrary.

Note that the scale function is defined uniquely up to an arbitrary constant c. The choice of this constant will not play any role in the following results.

We will use the scale function to compute hitting probabilities via the following result, as stated in Lemma 6.7 of [23], which is a reformulation of Proposition 3.2 of [51, Chapter VII] for our definition of the scale function.

Proposition B.2. *Let $s : \mathbb{R} \to \mathbb{R}$ be the scale function of the diffusion X. Then, for any $a, b, x \in \mathbb{R}$,with $a < x < b$,*

$$\mathbb{P}^x [\tau_b < \tau_a] = \frac{s(x) - s(a)}{s(b) - s(a)},$$

where τ_a and τ_b are defined to be the first hitting times of levels a and b, respectively, by the diffusion X.

From the scale function, we can derive the *speed measure* of the process X, which is defined in Definition 3.7 of [51, Chapter VII] and describes the time-change needed to transform X into a Brownian motion. Again, we take our definition to be the form found in Exercise 3.20 of [51, Chapter VII], which agrees with Definition 6.3 of [23] up to a multiplicative constant.

Definition B.3 (Speed measure)**.** Define the speed measure m of the diffusion X by

$$\int_A m(\mathrm{d}x) := \int_A \frac{2}{s'(x)\sigma^2(x)} \, \mathrm{d}x,$$

for any Borel set $A \subseteq \mathbb{R}$, where s is the scale function of X.

We now introduce the *Green's function*, which is defined in Definition 6.12 of [23] using the scale function and speed measure. Here, we take our definition to be consistent with the one used in Corollary 3.8 of [51, Chapter VII], which does not include the speed measure in the form of the Green's function.

Definition B.4 (Green's function)**.** Let $I = (a, b) \subset \mathbb{R}$. Then the Green's function $G_I : I \times I \to \mathbb{R}$ associated to X on the interval I is given by

$$G_I(x, y) := \begin{cases} \frac{(s(x)-s(a))(s(b)-s(y))}{s(b)-s(a)}, & a \leq x \leq y \leq b, \\ \frac{(s(y)-s(a))(s(b)-s(x))}{s(b)-s(a)}, & a \leq y \leq x \leq b. \end{cases}$$

The main result of this section is the following adaptation of both Corollary 3.8 of [51, Chapter VII] and Theorem 6.11 of [23]. We use this result several times to calculate expected costs in Chapter 2 and Chapter 5.

Proposition B.5. *Let I be an open interval and define $\tau := \inf\{t \geq 0 \colon X_t \notin I\}$. Then, for any Borel function $f : I \to \mathbb{R}$ that is bounded either from above or below,*

$$\mathbb{E}^x \left[\int_0^\tau f(X_s) \, \mathrm{d}s \right] = \int_I G_I(x, y) f(y) m(\mathrm{d}y).$$

BIBLIOGRAPHY

[1] Julio Backhoff-Veraguas, Mathias Beiglböck, Martin Huesmann, and Sigrid Källblad, *Martingale Benamou–Brenier: A probabilistic perspective*, arXiv:1708.04869 [math, q-fin] (2019) (en).

[2] Guy Barles and Panagiotis E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptotic analysis **4** (1991), no. 3, 271–283.

[3] Mathias Beiglböck, Pierre Henry-Labordère, and Friedrich Penkner, *Model-independent bounds for option prices—a mass transport approach*, Finance and Stochastics **17** (2013), no. 3, 477–501 (en).

[4] Mathias Beiglböck and Nicolas Juillet, *On a problem of optimal transport under marginal martingale constraints*, The Annals of Probability **44** (2016), no. 1, 42–106.

[5] Mathias Beiglböck, Marcel Nutz, and Nizar Touzi, *Complete duality for martingale optimal transport on the line*, The Annals of Probability **45** (2017), no. 5, 3038–3074 (en).

[6] Richard Bellman, *On the Theory of Dynamic Programming*, Proceedings of the National Academy of Sciences **38** (1952), no. 8, 716–719 (en).

[7] ———, *An introduction to the theory of dynamic programming*, RAND Corp Santa Monica CA **245** (1953) (en).

[8] Jean-David Benamou and Yann Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numerische Mathematik **84** (2000), no. 3, 375–393 (en).

[9] Bruno Bouchard and Nizar Touzi, *Weak Dynamic Programming Principle for Viscosity Solutions*, SIAM Journal on Control and Optimization **49** (2011), no. 3, 948–962 (en).

[10] Yann Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Communications on Pure and Applied Mathematics **44** (1991), no. 4, 375–417.

[11] Patrick Cattiaux, Paolo Dai Pra, and Sylvie Rœlly, *A Constructive Approach to a Class of Ergodic HJB Equations with Unbounded and Nonsmooth Cost*, SIAM Journal on Control and Optimization **47** (2008), no. 5, 2598–2615 (en).

[12] Giuseppe Maria Coclite and Nils Henrik Risebro, *Viscosity solutions of Hamilton–Jacobi equations with discontinuous coefficients*, Journal of Hyperbolic Differential Equations **4** (2007), no. 04, 771–795.

[13] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bulletin of the American Mathematical Society **27** (1992), no. 1, 1–67.

[14] Michael G. Crandall and Pierre-Louis Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Transactions of the American Mathematical Society **277** (1983), no. 1, 1–42.

[15] Hadrien De March, *Local structure of multi-dimensional martingale optimal transport*, arXiv:1805.09469 [math] (2018) (en).

[16] ———, *Quasi-sure duality for multi-dimensional martingale optimal transport*, arXiv:1805.01757 [math] (2018) (en).

[17] Hadrien De March and Nizar Touzi, *Irreducible convex paving for decomposition of multi-dimensional martingale transport plans*, The Annals of Probability **47** (2019), no. 3, 1726–1774 (en).

[18] Guido De Philippis and Alessio Figalli, *The Monge–Ampère equation and its link to optimal transportation*, Bulletin of the American Mathematical Society **51** (2014), no. 4, 527–580.

[19] Lester Dubins, Jacob Feldman, Meir Smorodinsky, and Boris Tsirelson, *Decreasing Sequences of Sigma-Fields and a Measure Change for Brownian Motion*, The Annals of Probability **24** (1996), no. 2, 882–904.

[20] Nicole El Karoui and Xiaolu Tan, *Capacities, Measurable Selection and Dynamic Programming Part II: Application in Stochastic Control Problems*, arXiv:1310.3364 [math] (2013).

[21] Michel Émery and Walter Schachermayer, *A remark on Tsirelson's stochastic differential equation*, Séminaire de Probabilités XXXIII (Jacques Azéma, Michel Émery, Michel Ledoux, and Marc Yor, eds.), vol. 1709, Springer Berlin Heidelberg, Berlin, Heidelberg, 1999, pp. 291–303 (en).

[22] ______, *On Vershik's Standardness Criterion and Tsirelson's Notion of Cosiness*, Séminaire de Probabilités XXXV (Jacques Azéma, Michel Émery, Michel Ledoux, and Marc Yor, eds.), vol. 1755, Springer Berlin Heidelberg, Berlin, Heidelberg, 2001, pp. 265–305 (en).

[23] Alison Etheridge, *C8.2: Stochastic analysis and PDEs*, https://courses.maths.ox.ac.uk/node/view_material/35260, 2016, Accessed: 2020-06-20.

[24] Xiaobing Feng and Max Jensen, *Convergent Semi-Lagrangian Methods for the Monge–Ampère Equation on Unstructured Grids*, SIAM Journal on Numerical Analysis **55** (2017), no. 2, 691–712 (en).

[25] E. Robert Fernholz, Ioannis Karatzas, and Johannes Ruf, *Volatility and arbitrage*, The Annals of Applied Probability **28** (2018), no. 1, 378–417 (en).

[26] Wendell H. Fleming and Halil Mete Soner, *Controlled Markov processes and viscosity solutions*, vol. 25, Springer Science & Business Media, 2006.

[27] A. Galichon, P. Henry-Labordère, and N. Touzi, *A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options*, The Annals of Applied Probability **24** (2014), no. 1, 312–336 (en).

[28] Bernard Gaveau, *Méthodes de contrôle optimal en analyse complexe. I. Résolution d'équations de Monge Ampère*, Journal of Functional Analysis **25** (1977), 391–411.

[29] Nassif Ghoussoub, Young-Heon Kim, and Tongseok Lim, *Structure of optimal martingale transport plans in general dimensions*, Annals of Probability **47** (2019), no. 1, 109–164 (en).

[30] Fausto Gozzi and Tiziano Vargiolu, *Superreplication of European multiasset derivatives with bounded stochastic volatility*, Mathematical Methods of Operations Research (ZOR) **55** (2002), no. 1, 69–91 (en).

[31] Cristian E. Gutiérrez, *The Monge-Ampère Equation*, second ed., Progress in Nonlinear Differential Equations and Their Applications, vol. 89, Springer International Publishing, 2016.

[32] David Hobson, *Robust hedging of the lookback option*, Finance and Stochastics **2** (1998), no. 4, 329–347 (en).

[33] ———, *The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices*, Paris-Princeton Lectures on Mathematical Finance 2010, vol. 2003, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 267–318 (en).

[34] David Hobson and Martin Klimmek, *Robust price bounds for the forward starting straddle*, Finance and Stochastics **19** (2015), no. 1, 189–214.

[35] David Hobson and Anthony Neuberger, *Robust Bounds for Forward Start Options*, Mathematical Finance **22** (2012), no. 1, 31–56 (en).

[36] Hitoshi Ishii and Pierre-Louis Lions, *Viscosity solutions of fully nonlinear second-order elliptic partial differential equations*, Journal of Differential equations **83** (1990), no. 1, 26–78.

[37] Max Jensen and Iain Smears, *On the notion of boundary conditions in comparison principles for viscosity solutions*, arXiv:1703.07313 [math] (2017).

[38] Ioannis Karatzas and Steven E. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, vol. 113, Springer New York, New York, NY, 1998.

[39] Nikolai V. Krylov, *Nonlinear elliptic and parabolic equations of the second order*, Mathematics and Its Applications (Soviet Series), vol. 7, Springer, 1987.

[40] Martin Larsson and Johannes Ruf, *Minimum curvature flow and martingale exit times*, arXiv:2003.13611 [math] (2020) (en).

[41] ———, *Relative Arbitrage: Sharp Time Horizons and Motion by Curvature*, arXiv:2003.13601 [math, q-fin] (2020) (en).

[42] Stéphane Laurent, *On Standardness and I-cosiness*, Séminaire de Probabilités XLIII (Catherine Donati-Martin, Antoine Lejay, and Alain Rouault, eds.), vol. 2006, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 127–186 (en).

[43] Pierre-Louis Lions, *Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations part 2: Viscosity solutions and uniqueness*, Communications in Partial Differential Equations **8** (1983), no. 11, 1229–1276.

[44] Roger Mansuy and Marc Yor, *Random Times and Enlargements of Filtrations in a Brownian Setting*, Lecture Notes in Mathematics, vol. 1873, Springer-Verlag, Berlin/Heidelberg, 2006 (en).

[45] Peter Mörters and Yuval Peres, *Brownian motion*, vol. 30, Cambridge University Press, 2010.

[46] Bernt Øksendal, *Stochastic differential equations: An introduction with applications*, fifth ed., Springer Science & Business Media, 2013.

[47] Goran Peskir and Albert Shiryaev, *Optimal Stopping and Free-Boundary Problems*, Birkhäuser Basel, 2006.

[48] Huyen Pham, *On the Smooth-Fit Property for One-Dimensional Optimal Switching Problem*, Séminaire de Probabilités XL (Catherine Donati-Martin, Michel Émery, Alain Rouault, and Christophe Stricker, eds.), vol. 1899, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, pp. 187–199 (en).

[49] Huyên Pham, *Continuous-time Stochastic Control and Optimization with Financial Applications*, Stochastic Modelling and Applied Probability, vol. 61, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[50] Ross Pinsky, *Positive harmonic functions and diffusion*, Cambridge Studies in Advanced Mathematics, no. 45, Cambridge University Press, New York, 1995.

[51] Daniel Revuz and Marc Yor, *Continuous Martingales and Brownian Motion*, Grundlehren Der Mathematischen Wissenschaften, vol. 293, Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

[52] Ralph Tyrell Rockafellar, *Convex analysis*, Princeton University Press, 2015.

[53] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes and martingales. Vol.2. Ito calculus*, second ed., Cambridge University Press, 2000.

[54] Walter Rudin, *Principles of Mathematical Analysis*, third ed., McGraw-Hill, 1976.

[55] Albert Nikolaevich Shiryaev, *Probability*, second ed., Graduate Texts in Mathematics, vol. 95, Springer Science & Business Media, 1996 (English).

[56] Iain Smears, *Hamilton-Jacobi-Bellman Equations: Analysis and Numerical Analysis*, maths.dur.ac.uk/Ug/projects/highlights/PR4/Smears_HJB_report.pdf, 2018, Accessed: 2020-06-20.

[57] Xiaolu Tan and Nizar Touzi, *Optimal Transportation Under Controlled Stochastic Dynamics*, The Annals of Probability **41** (2013), no. 5, 3201–3240.

[58] Nizar Touzi, *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE*, Fields Institute Monographs, vol. 29, Springer New York, New York, NY, 2013.

[59] B. Tsirelson, *An Example of a Stochastic Differential Equation Having No Strong Solution*, Theory of Probability & Its Applications **20** (1976), no. 2, 416–418 (en).

[60] ———, *Triple Points: From Non-Brownian Filtrations to Harmonic Measures*, Geometric And Functional Analysis **7** (1997), no. 6, 1096–1142 (en).

[61] Tiziano Vargiolu, *Existence, uniqueness and smoothness for the Black-Scholes-Barenblatt equation*, Tech. report, Universita di Padova, 2001.

[62] AM Vershik, *Approximation in Measure Theory*, Doctoral Thesis, 1973.

[63] Cédric Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, no. 58, American Mathematical Society, Providence, Rhode Island, 2003.

[64] John B. Walsh, *A diffusion with a discontinuous local time*, Astérisque **52-53** (1978), 37–45 (en).

[65] J. Warren, *On the joining of sticky Brownian motion*, Séminaire de Probabilités XXXIII (Jacques Azéma, Michel Émery, Michel Ledoux, and Marc Yor, eds.), vol. 1709, Springer Berlin Heidelberg, Berlin, Heidelberg, 1999, pp. 257–266 (en).

[66] Shilin Zhan, *On the Determinantal Inequalities*, Journal of Inequalities in Pure and Applied Mathematics **6** (2005), no. 4.

[67] A K Zvonkin, *A transformation of the phase space of a diffusion process that removes the drift*, Mathematics of the USSR-Sbornik **22** (1974), no. 1, 129–149 (en).