Stochastic Control in Continuous Time

Kevin Ross

E-mail address: kjross@stat.stanford.edu

Department of Statistics, Stanford University, Stanford, CA 94305.

Contents

CHAPTER 1

Controlled Diffusion Processes

In preparation for our study of stochastic control we recall in this chapter some basic theory of diffusion processes. In Section 1.1 we present a brief review of stochastic differential equations. We then recall in Section 1.2 the basic theory of continuous time Markov processes. We study diffusion processes and the important notion of a generator in 1.3. Finally, in Section 1.4 we introduce controlled diffusion processes which will play a central role in stochastic control theory. For more details, see for example [**Oks03**], [**KS97**], [**EK86**], [**Dem07**], [**Bor89**].

1.1. Stochastic Differential Equations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a filtration[1] $\{\mathcal{F}_t\}$. As usual, we interpret $\{\mathcal{F}_t\}$ as representing the flow of information over time, with $\mathcal{F}_t$ being the information available at time t. A stochastic process W defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $\mathbb{R}^d$ is a *d-dimensional* $\{\mathcal{F}_t\}$*-Brownian motion* if:

(a) W is $\{\mathcal{F}_t\}$-adapted,
(b) For all $0 \le s \le t$, $W_t - W_s$ is independent of $\mathcal{F}_s$,
(c) $W_0 = 0$ almost surely,
(d) For all $0 \le s \le t$, $W_t - W_s \sim N(0, t-s)$,
(e) With probability one, the sample paths of W are continuous.

We seek a solution of the stochastic differential equation (SDE):

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \ , \quad X_0 = x, \tag{1.1.1}$$

where $x \in \mathbb{R}^n$, $b : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{R}^n$, and $\sigma : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{R}^{n \times d}$, the space of real-valued $n \times d$ matrices. With b_i, x_i, $X_i(t)$, $W_j(t)$ denoting the components of the vectors b, x, $X(t)$, $W(t)$ respectively, and σ_{ij}, $i = 1, \dots, n, j = 1, \dots, d$ denoting the entries of the matrix σ, (1.1.1) can be written componentwise as

$$dX_i(t) = b_i(t, X(t))dt + \sum_{j=1}^{d} \sigma_{ij}(t, X(t))dW_j(t) \ , \quad X_i(0) = x_i \ , \quad i = 1, \dots, n.$$

[1]Throughout we will assume that any filtration satisfies the *usual conditions* of right-continuity and completeness.

Let $a(t,x) = \sigma(t,x)\sigma^T(t,x)$ and note that a is a symmetric, nonnegative definitive $n \times n$ matrix; the entries of a will be denoted a_{ij}.

We assume that the coefficients b and σ satisfy the following global Lipschitz and linear growth conditions: There exists a constant $K \in (0,\infty)$ such that for all $t \in [0,\infty)$ and $x, y \in \mathbb{R}^n$ we have

$$|b(t,x) - b(t,y)| + ||\sigma(t,x) - \sigma(t,y)|| \leq K|x-y|, \tag{1.1.2}$$

$$|b(t,x)|^2 + ||\sigma(t,x)||^2 \leq K^2(1+|x|^2). \tag{1.1.3}$$

In the above notation, $||\sigma||^2 = \sum_{i=1}^n \sum_{j=1}^d \sigma_{ij}^2$. We refer to (1.1.2) and (1.1.3) as the "Ito conditions". The following theorem shows that (1.1.2) and (1.1.3) provide sufficient conditions on the coefficients b and σ under which there exists a unique solution to SDE (1.1.1).

THEOREM 1.1.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a filtration $\{\mathcal{F}_t\}$ and a d-dimensional $\{\mathcal{F}_t\}$-Brownian motion W. Assume that b and σ satisfy* (1.1.2) *and* (1.1.3). *Then there exists a* unique (strong) solution *to* (1.1.1). *That is, there is a stochastic process $\{X_t\}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ adapted to $\{\mathcal{F}_t\}$, with continuous sample paths a.s., which satisfies the stochastic integral equation*

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s\ , \quad 0 \leq t < \infty;$$

written componentwise, for $i = 1, \ldots, d$,

$$X_i(t) = x_i + \int_0^t b_i(s, X(s))ds + \sum_{j=1}^d \int_0^t \sigma_{ij}(s, X(s))dW_j(s)\ , \ 0 \leq t < \infty.$$

The solution is unique in the sense that if X_t and $\tilde{X}_t$ are both solutions of (1.1.1) *(as described above) then $\mathbb{P}(X_t = \tilde{X}_t, 0 \leq t < \infty) = 1$.*

The unique solution X of (1.1.1) is called an *Ito process*.

One of the main tools used in dealing with solutions of stochastic differential equations is the following version of Itô's rule.

THEOREM 1.1.2. *Let X be a solution of (1.1.1) and $f : [0,\infty) \times \mathbb{R}^n \mapsto \mathbb{R}$ be of class $C^{1,2}$. Then for all $t \in [0,\infty)$, almost surely*

$$\begin{aligned} f(t,X_t) = f(0,x) &+ \int_0^t \frac{\partial f}{\partial t}(s,X_s)ds + \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(s,X_s)b_i(s,X_s)ds \\ &+ \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(s,X_s) \sum_{j=1}^d \sigma_{ij}(s,X_s)dW_s^{(j)} \\ &+ \frac{1}{2}\sum_{i=1}^n \sum_{k=1}^n \frac{\partial^2 f}{\partial x_i \partial x_k}(s,X_s)a_{ik}(s,X_s)ds. \end{aligned} \tag{1.1.4}$$

For f of class $C^{1,2}$ let $\mathcal{L}f$ be given by

$$\mathcal{L}f(t,x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t,x)b_i(t,x) + \frac{1}{2}\sum_{i=1}^n \sum_{k=1}^n \frac{\partial^2 f}{\partial x_i \partial x_k}(t,x)a_{ik}(t,x) \tag{1.1.5}$$

With this notation, (1.1.4) can be written more compactly as

$$\begin{aligned} f(t,X_t) = f(0,x) &+ \int_0^t \Big[\frac{\partial f}{\partial t}(s,X_s)ds + \mathcal{L}f(s,X_s)\Big]ds \\ &+ \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(s,X_s) \sum_{j=1}^d \sigma_{ij}(s,X_s)dW_j(s). \end{aligned}$$

Suppose that for each i,j and all $t \in [0,\infty)$ we have

$$\mathbb{E}\int_0^t \Big|\frac{\partial f}{\partial x_i}(s,X_s)\sigma_{ij}(s,X_s)\Big|^2 ds < \infty.$$

Then the stochastic integral in the above Ito expansion (1.1.4) is a martingale and thus has zero expected value. In this case, taking expected values on both sides of the above Ito expansion yields

$$\mathbb{E}[f(t,X_t) - f(0,x)] = \mathbb{E}\int_0^t \Big[\frac{\partial f}{\partial t}(s,X_s)ds + \mathcal{L}f(s,X_s)\Big]ds.$$

This equation if often referred to as (a version of) Dynkin's formula. It allows us to replace the average at a fixed time t (i.e. $\mathbb{E}[f(t,X_t)]$ on the LHS) by an average over a fixed time interval $[0,t]$ (the RHS).

Before giving some examples, we state the 1-dimensional form of (1.1.4). Suppose that X is a real-valued solution of (1.1.1) (with b and σ real-valued) and $f : [0,\infty) \times \mathbb{R} \mapsto \mathbb{R}$ is of class $C^{1,2}$. Letting f_t denote the derivative of f with respect to the time variable and f_x, f_{xx} the first and second derivatives

of f in the state variable, we have

$$f(t,X_t) = f(0,x) + \int_0^t f_x(s,X_s)\sigma(s,X_s)dW_s + \int_0^t [f_t(s,X_s) + f_x(s,X_s)b(s,X_s) + (1/2)f_{xx}(s,X_s)\sigma^2(s,X_s)]ds.$$

EXAMPLE 1.1.3. *If b and σ are constant then the solution of* (1.1.1) *is $X_t = bt + \sigma W_t$, a* Brownian motion with drift.

EXAMPLE 1.1.4. *If $b(t,x) = \tilde{b}x$ and $\sigma(t,x) = \tilde{\sigma}x$ for constants $\tilde{b} \in \mathbb{R}$, $\tilde{\sigma} > 0$, then the solution of* (1.1.1) *is $X_t = x\exp((\tilde{b} - \tilde{\sigma}^2/2)t + \tilde{\sigma}W_t)$, a* geometric Brownian motion.

EXAMPLE 1.1.5. *If $b(t,x) = -\gamma x$ and $\sigma(t,x) = \tilde{\sigma}$ for constants $\gamma > 0$, $\tilde{\sigma} > 0$, then the solution of* (1.1.1) *can be expressed as $X_t = e^{-\gamma t}(x + B_{\hat{t}})$ with $\hat{t} = \tilde{\sigma}^2(e^{2\gamma t} - 1)/(2\gamma)$. This process is known as an Ornstein-Uhlenbeck process.*

1.2. Markov Processes

One of the most important properties of stochastic differential equations is that their solutions exhibit the Markov property. For a stochastic process with the Markov property, roughly speaking, given the present the future is independent of the past. Before making this definition rigorous, we introduce the notion of a state space. We say that a stochastic process X has *state space* $\mathbb{S}$ if $\mathbb{P}(X_t \in \mathbb{S}, 0 \le t < \infty) = 1$. We will assume that $\mathbb{S} \subseteq \mathbb{R}^n$ and will denote by $\mathcal{B}_{\mathbb{S}}$ the corresponding Borel σ-field.

DEFINITION 1.2.1. *A stochastic process $\{X_t\}$ with state space $\mathbb{S}$ is a* Markov process *with respect to $\{\mathcal{F}_t\}$ (or $\{\mathcal{F}_t\}$-Markov process) if it is $\{\mathcal{F}_t\}$-adapted and for any $t, u \ge 0$ and $A \in \mathcal{B}_{\mathbb{S}}$ we have that almost surely*

$$\mathbb{P}(X(t+u) \in A|\mathcal{F}_t) = \mathbb{P}(X(t+u) \in A|X(t)).$$

Equivalently, an $\{\mathcal{F}_t\}$-adapted process $\{X_t\}$ is an $\{\mathcal{F}_t\}$-Markov process if for any $t, u \ge 0$ and any bounded measurable function $f : \mathbb{S} \mapsto \mathbb{R}$, almost surely,

$$\mathbb{E}[f(X(t+u))|\mathcal{F}_t] = \mathbb{E}[f(X(t+u))|X(t)]. \tag{1.2.1}$$

We often write "$\{X_t\}$ is a Markov process", in which case the filtration is understood to be the canonical filtration $\{\mathcal{F}_t^X\}$ (where $\mathcal{F}_t^X = \sigma(\{X_s, 0 \le s \le t\})$). Clearly, if X is a $\{\mathcal{F}_t\}$-Markov process then it is also a Markov process with respect to its canonical filtration.

EXAMPLE 1.2.2. *We say that an $\{\mathcal{F}_t\}$-adapted stochastic process has* independent increments *if $X_t - X_s$ is independent of $\mathcal{F}_s$ for all $0 \le s < t < \infty$. It can be shown that any process with independent increments is an $\{\mathcal{F}_t\}$-Markov process. In particular $\{\mathcal{F}_t\}$-Brownian motion is an $\{\mathcal{F}_t\}$-Markov process.*

The definition of a Markov process is stated in terms of each fixed time $t+u$ in the future. The following lemma shows that the Markov property still holds even if we consider multiple future values simultaneously.

LEMMA 1.2.3. *Let $\{X_t\}$ be an $\{\mathcal{F}_t\}$-Markov process. Then $\mathbb{E}[Y|\mathcal{F}_t] = \mathbb{E}[Y|X_t]$ for any $t \geq 0$ and integrable random variable Y which is measurable with respect to the σ-field generated by the collection $\{X_s, t \leq s < \infty\}$.*

An important property of any Markov process is that its distributional properties are uniquely determined by its initial distribution and its transition probabilities, which we now define. The *initial distribution* of a Markov process is the probability measure π on $(\mathbb{S}, \mathcal{B}_\mathbb{S})$ given by $\pi(A) = \mathbb{P}(X_0 \in A)$.

DEFINITION 1.2.4. *Let $\{X(t)\}$ be a Markov process. For each $0 \leq s < t$ and fixed $x \in \mathbb{S}$ there exists a probability measure $p_{t,s}(\cdot|x)$ on $(\mathbb{S}, \mathcal{B}_\mathbb{S})$ such that for each fixed $A \in \mathcal{B}_\mathbb{S}$, the function $p_{t,s}(A|\cdot)$ is $\mathcal{B}_\mathbb{S}$-measurable and*

$$\mathbb{P}(X(t) \in A|X(s)) = \mathbb{E}(I_{X(t)\in A}|X(s)) = p_{t,s}(A|X(s)) \qquad a.s.$$

The collection $\{p_{t,s}(A|x)\}$ is called the transition probabilities *for the Markov process $\{X(t)\}$.*

Intuitively, $p_{t,s}(A|X(s))$ represents the probability that, starting from state $X(s) = x$ at time s, the process takes a value in the set A at time t. It can be shown that the collection $\{p_{t,s}(A|x)\}$ provides the transition probabilities for a Markov process X if and only if

$$\mathbb{E}[f((X(t))|\mathcal{F}_s] = \int_\mathbb{S} f(y) p_{t,s}(dy|X(s)) \qquad a.s. \tag{1.2.2}$$

for all $0 \leq s \leq t$ and all bounded $\mathcal{B}_\mathbb{S}$-measurable functions $f : \mathbb{S} \mapsto \mathbb{R}$.

The transition probabilities $p_{t,s}(\cdot|\cdot)$, together with the initial distribution $\pi(\cdot)$ determine the *finite dimensional distributions* of the Markov process. Indeed, for each nonnegative integer k, every $0 = t_0 < t_1 < \cdots < t_k$ and $A_0, \ldots, A_k \in \mathcal{B}_\mathbb{S}$ we have that

$$\begin{aligned}&\mathbb{P}(X(t_k) \in A_k, \ldots, X(t_0) \in A_0)\\ &\quad = \int_{A_k} \cdots \int_{A_0} p_{t_k,t_{k-1}}(dx_k|x_{k-1}) \cdots p_{t_1,t_0}(dx_1|x_0)\pi(dx_0),\end{aligned} \tag{1.2.3}$$

using Lesbegue integrals in the above.

Using the tower property and the definition of a Markov process, it can be shown that the transition probabilities of a Markov process satisfy the *Chapman-Kolmogorov equations*

$$p_{t,s}(A|x) = \int_\mathbb{S} p_{t,u}(A|y) p_{u,s}(dy|x)\,, \tag{1.2.4}$$

for all $t > u > s \geq 0$, $A \in \mathcal{B}_{\mathbb{S}}$ and all[2] $x \in \mathbb{S}$.

For modeling purposes it is more effective to specify the collection $p_{t,s}(A|x)$ rather than starting with the Markov process $X(t)$ and its probability space. When doing so, we obviously wish to know that there exists a Markov process whose transition probabilities coincide with those we specify. To this end we have the following definition.

DEFINITION 1.2.5. *We say that $p_{t,s}(A|x)$ for $t > s \geq 0$, $x \in \mathbb{S}$ and $A \in \mathcal{B}_{\mathbb{S}}$ are* (regular) transition probabilities *if $p_{t,s}(\cdot|x)$ are probability measures on $(\mathbb{S}, \mathcal{B}_{\mathbb{S}})$, the functions $p_{t,s}(A|\cdot)$ are Borel measurable and the Chapman-Kolmogorov equations* (1.2.4) *hold for every $t > u > s \geq 0$, $x \in \mathbb{S}$ and $A \in \mathcal{B}$.*

THEOREM 1.2.6. *Let $\{p_{t,s}(A|x) : t, s \in [0, \infty), A \in \mathcal{B}_{\mathbb{S}}, x \in \mathbb{S}\}$ be a collection of regular transition probabilities and π a probability measure on $(\mathbb{S}, \mathcal{B}_{\mathbb{S}})$. Then there is some probability space on which is defined a Markov process with initial distribution π and whose transition probabilities are $\{p_{t,s}(A|x)\}$ (and hence its finite dimensional distributions are given by* (1.2.3)*).*

Of special importance is the class of *(time) homogeneous* Markov processes, associated with *stationary* regular transition probabilities.

DEFINITION 1.2.7. *A* homogeneous $\{\mathcal{F}_t\}$-Markov process *is a $\{\mathcal{F}_t\}$-Markov process with regular transition probabilities of the form $p_{t,s}(A|x) = p_{t-s}(A|x)$, which in turn are called* stationary regular transition probabilities.

For a homogeneous Markov process the transition probabilities do not depend explicitly on the starting or ending times, only on the length of time between them. (For example, for a homogeneous Markov process, $p_{1,0}(A|x) = p_{t+1,t}(A|x)$, but this equality might not hold for a general Markov process.) It is easy to verify that for a homogeneous Markov process the Chapman-Kolmogorov equations (1.2.4) simplify to

$$p_{t+s}(A|x) = \int_{\mathbb{S}} p_t(A|y) p_s(dy|x) \quad \text{for all } t, s \geq 0, A \in \mathcal{B}_{\mathbb{S}}, x \in \mathbb{S}.$$

EXAMPLE 1.2.8. *A stochastic process X has* stationary increments *if $X_{t+h} - X_t$ and $X_{s+h} - X_s$ have the same law for all $h, s, t \geq 0$. It can be shown that if an $\{\mathcal{F}_t\}$-adapted stochastic process X has stationary and independent increments then it is a homogeneous $\{\mathcal{F}_t\}$-Markov process. In particular, d-dimensional Brownian motion is a homogeneous Markov process (with state space $\mathbb{S} = \mathbb{R}^d$). It corresponds to initial distribution $\pi(\{0\}) = 1$ and stationary*

[2]Technically, this calculation will only show that (1.2.4) only holds almost surely in x chosen according to the law of $X(s)$. However, we will always assume that (1.2.4) does hold for all $x \in \mathbb{S}$; the corresponding transition probabilities are then called *regular*.

regular transition probabilities

$$p_t(A|x) = \int_A \frac{1}{(2\pi t)^{d/2}} e^{-|y-x|^2/(2t)} dy.$$

Recall that at the start of this section we claimed that the solution of the SDE (1.1.1) is a Markov process. The following theorem provides this result.

THEOREM 1.2.9. *Under the assumptions of Theorem 1.1.1, the unique solution* $\{X_t\}$ *of* (1.1.1) *is an* $\{\mathcal{F}_t\}$*-Markov process.*

PROOF. Fix $t, s \geq 0$ and a bounded measurable function $f : \mathbb{R}^n \mapsto \mathbb{R}$. We claim that $\mathbb{E}[f(X_{t+s})|\mathcal{F}_t] = g(X_t)$ for some measurable function $g : \mathbb{R}^n \mapsto \mathbb{R}$. In this case, $\mathbb{E}[f(X_{t+s})|X_t] = \mathbb{E}[\mathbb{E}[f(X_{t+s})|\mathcal{F}_t]|X_t] = \mathbb{E}[g(X_t)|X_t] = g(X_t)$. Therefore $\mathbb{E}[f(X_{t+s})|\mathcal{F}_t] = \mathbb{E}[f(X_{t+s})|X_t]$ which proves the Markov property. Thus it suffices to prove the above claim.

On noting that

$$X_{t+s} = X_t + \int_t^{t+s} b(u, X_u)du + \int_t^{t+s} \sigma(u, X_u)dW_u$$

it is easy to check that X_{t+s} is measurable with respect to σ-field generated by the collection $(X_t, \{W_{t+u} - W_t, 0 \leq u \leq s\})$. It follows that $X_{t+s} = F(X_t, W_{t+\cdot} - W_t)$ for some measurable $F : \mathbb{R}^n \times C([t, t+s] : \mathbb{R}^d) \mapsto \mathbb{R}^n$. Let $A \in \mathcal{F}_t$ and consider the collection $(I_A, X_t, \{W_{t+u} - W_t, 0 \leq u \leq s\})$, which induces a probability law on $\mathbb{R} \times \mathbb{R}^n \times C([t, t+s] : \mathbb{R}^d)$ (with its Borel σ-field) which we denote $\mathcal{P}(x, y, z)$. Since (I_A, X_t) is $\mathcal{F}_t$-measurable and Brownian motion has independent increments, we have that (I_A, X_t) and $\{W_{t+u} - W_t, 0 \leq u \leq s\})$ are independent. Thus, letting $\mu(x, y)$ denote the law of (I_A, X_t) on $\mathbb{R} \times \mathbb{R}^n$ and $\nu(z)$ the law of $\{W_{t+u} - W_t, 0 \leq u \leq s\})$ on $C([t, t+s] : \mathbb{R}^d)$ we have that $\mathcal{P}(x, y, z) = \mu(x, y) \times \nu(z)$.

Define a Borel measurable function $g : \mathbb{R}^n \mapsto \mathbb{R}$ by

$$g(y) = \int_{C([t,t+s]:\mathbb{R}^d)} f(F(y, z))d\nu(z),$$

which is well defined and finite since f is bounded. Then using the change of variables formula and Fubini's Theorem, we have

$$\begin{aligned}
\mathbb{E}[I_A f(X_{t+s})] &= \mathbb{E}[I_A f(F(X_t, W_{t+\cdot} - W_t))] \\
&= \int_{\mathbb{R}\times\mathbb{R}^n\times C([t,t+s]:\mathbb{R}^d)} x f(F(y, z))d\mathcal{P}(x, y, z) \\
&= \int_{\mathbb{R}\times\mathbb{R}^n} \Big(\int_{C([t,t+s]:\mathbb{R}^d)} f(F(y, z))d\nu(z) \Big) x d\mu(x, y) \\
&= \int_{\mathbb{R}\times\mathbb{R}^n} g(y) x d\mu(x, y) = \mathbb{E}[I_A g(X_t)].
\end{aligned}$$

Therefore, $\mathbb{E}[f(X_{t+s})|\mathcal{F}_t] = g(X_t)$ with g as defined above. ∎

1.3. Diffusion Processes

In this section we consider the case where the coefficients b and σ (and $a = \sigma\sigma^T$) in (1.1.1) do not depend on t. That is, we consider the solution to

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t\ , \quad X_0 = x. \tag{1.3.1}$$

As usual, we assume that b and σ satisfy (1.1.2), (1.1.3) now expressed in time independent form as:

$$|b(x) - b(y)| + ||\sigma(x) - \sigma(y)|| \leq K|x - y|, \tag{1.3.2}$$

$$|b(x)|^2 + ||\sigma(x)||^2 \leq K^2(1 + |x|^2). \tag{1.3.3}$$

The unique solution to X (1.3.1) is called an *Ito diffusion*. Roughly speaking, such processes behave locally like Brownian motion. For example, with probability one, the paths of X are continuous everywhere but differentiable nowhere. Also, (again, roughly speaking) given $X(t) = x$ the increment $X(t+\delta) - X(t)$ for small δ is conditionally independent of $\mathcal{F}_t$ and has a $N(b(x)\delta, a(x)\delta)$ distribution.

It can be shown that an Ito diffusion is a *homogeneous* Markov process. Thus its distributional properties can be characterized by its initial distribution and its transition probabilities $\{p_t(A|x)\}$. However, we need to specify the transition probabilities $p_t(\cdot|\cdot)$ for all $t > 0$ which is usually a difficult task. In view of the Chapman-Kolmogorov relationship, using functional analysis one may often express $p_t(\cdot|\cdot)$ in terms of a single operator, which is called the "generator" of the Markov process. The idea is that if we can somehow specify or characterize $p_t(\cdot, \cdot)$ for small values of t then we can use the Chapman-Kolmogorov equations to obtain the transition probabilities for larger values of t. The generator provides the connection we need for infinitesimal values of t. We provide next a heuristic motivation for considering the generator and then provide a result concerning the generator of an Ito diffusion.

For simplicity, we assume that b and σ are bounded, smooth, and real-valued. Let $\mathbb{E}_x$ denoted the expected value given that $X_0 = x$ for $x \in \mathbb{R}$. From (1.2.2) we know that for suitable functions f

$$\frac{1}{t}\mathbb{E}_x[f(X_t) - f(x)] = \frac{1}{t}\int [f(y) - f(x)]p_t(dy|x).$$

Thus if we can analyze the limit as $t \downarrow 0$ of the LHS of the above equation, we provide an implicit relationship that the transition probabilities must satisfy for small values of t. Each "test function" f provides an equation that the transition probabilities must satisfy. The goal is to specify a set of test functions

which is small enough so that the limits exist, but large enough to completely determine the collection $\{p_t(A|x)\}$. With this in mind we define

$$\lim_{t\downarrow 0} \frac{1}{t}\mathbb{E}_x[f(X_t) - f(x)] \doteq Af(x) \tag{1.3.4}$$

We denote by $\mathcal{D}(A)$ the set of all functions f for which the limit on the LHS of (1.3.4) exists for all $x \in \mathbb{R}$. The LHS of (1.3.4) is called the *infinitesimal generator* of the process X applied to the test function f. Note that a priori $Af(x)$ is just a notation representing the LHS of (1.3.4). We now wish to specify: a suitable class of functions f for which the limit in (1.3.4) is defined for all $x \in \mathbb{R}$, and the form of Af for this class of functions. Before considering arbitrary f, it is instructive to consider two special cases.

Consider $f(y) = y$. With σ bounded, $\int_0^t \sigma(X_s)dW_s$ is a martingale with expected value zero. Thus

$$\mathbb{E}_x[X_t - x] = \mathbb{E}_x \int_0^t b(X_s)ds \approx b(x)t,$$

if b is smooth and t is small. Thus we expect $\lim_{t\downarrow 0} \mathbb{E}_x[(X_t - x)/t] = b(x)$. This formula gives an interpretation of the coefficient b: $b(x)$ measures the local (at x) mean velocity of the process X.

Now consider $f(y) = (y - x)^2$. Then

$$\mathbb{E}_x[f(X_t) - f(x)] = \mathbb{E}_x[(X_t - x)^2] = \mathbb{E}_x(\int_0^t b(X_s)ds)^2 + \mathbb{E}_x(\int_0^t \sigma(X_s)dW_s)^2,$$

where the cross-product terms vanish due to boundedness of b and σ and the martingale property of the stochastic integral term. Now for small t with b smooth we have $\mathbb{E}_x(\int_0^t b(X_s)ds)^2 \approx (b(x)t)^2$. On the other hand, using the Ito isometry, $\mathbb{E}_x(\int_0^t \sigma(X_s)dW_s)^2 = \mathbb{E}_x \int_0^t \sigma^2(X_s)ds \approx \sigma^2(x)t = a(x)t$, for small t and smooth σ. Thus we expect $\lim_{t\downarrow 0} \mathbb{E}_x[(X_t - x)^2/t] = a(x)$. Note that $a(x)$ represents the local rate of change in the variance (more generally, covariance matrix) of the increment $X_t - x$.

We now consider (1.3.4) where $f \in C^2(\mathbb{R})$ is bounded with bounded first and second derivatives. Applying Ito's rule to f and taking expected values (upon which the term corresponding to the stochastic integral vanishes) we obtain

$$\mathbb{E}_x[f(X_t) - f(x)] = \mathbb{E}_x \int_0^t [f'(X_s)b(X_s) + (1/2)f''(X_s)a(X_s)]ds.$$

We now divide both sides by t and take limits as $t \downarrow 0$ to obtain

$$\lim_{t\downarrow 0} \mathbb{E}_x[f(X_t) - f(x)] = f'(x)b(x) + (1/2)f''(x)a(x).$$

Thus, such an f is in $\mathcal{D}(A)$ and A on the RHS of (1.3.4) is given by

$$Af(x) = f'(x)b(x) + (1/2)f''(x)a(x).$$

If the class of f for which we define Af is large enough, then by specifying the operator A we have effectively characterized the Markov process X and hence its transition probabilities. The question then remains: what is a large enough class of functions? We do not attempt to answer this question here. Rather, throughout these notes we will assume that the class of functions f for which we define Af is sufficient to determine the distributional properties of the corresponding Markov process.

In the motivating discussion above we have made several simplifying assumptions on b and σ and the test functions f. In particular, we assumed that the various functions and derivatives were bounded so that the stochastic integral term resulting from applying Ito's rule was a martingale and hence its expectation vanished. This property is a key characteristic of $\mathcal{D}(A)$; whatever class of f we ultimately choose will need to preserve it.

We now return to the general n-dimensional setting of (1.3.1) and attempt to put the preceding discussion on a more rigorous footing.

DEFINITION 1.3.1. *Let X be a homogeneous Markov process taking values in $\mathbb{R}^n$. The* infinitesimal generator *of X applied to the test function f is denoted by Af and given by*

$$\lim_{t\downarrow 0} \frac{1}{t}\mathbb{E}_x[f(X_t) - f(x)] \doteq Af(x), \quad x \in \mathbb{R}^n. \tag{1.3.5}$$

We denote by $\mathcal{D}(A)$ the set of all functions f for which $Af(x)$ exists for all $x \in \mathbb{R}^n$.

We say that $f : \mathbb{R}^n \mapsto \mathbb{R}$ is of class C_p^2 if f is of class C^2 and it, along with its first and second partial derivatives, satisfy a polynomial growth condition: There exists constants $q \geq 1$, and $K_q \in (0, \infty)$ such that for all $x \in \mathbb{R}^n$,

$$|f(x)| \leq K_q(1 + |x|^q).$$

THEOREM 1.3.2. *Let X be the solution to* (1.3.1) *with b and σ satisfying* (1.3.2) *and* (1.3.3). *Let f be of class $C_p^2(\mathbb{R}^n)$. Then $f \in \mathcal{D}(A)$ and*

$$\begin{aligned} Af(x) &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) b_i(x) + \frac{1}{2}\sum_{i=1}^n \sum_{k=1}^n \frac{\partial^2 f}{\partial x_i \partial x_k}(x) a_{ik}(x) \\ &= \nabla f(x) \cdot b(x) + \frac{1}{2} tr(D^2 f(x) a(x)). \end{aligned}$$

In the above notation ∇ represents the gradient vector, D^2 the Hessian matrix of second partial derivatives, and tr the trace of a square matrix. Note that in the time homogeneous case $Af = \mathcal{L}f$ for $f \in C_p^2(\mathbb{R}^n)$, where $\mathcal{L}f$ is defined in (1.1.5).

REMARK 1.3.3. There is a notion of a diffusion process which is more general than what we have presented here. Roughly speaking, a diffusion process is

a Markov process with continuous sample paths that can be characterized in terms of its infinitesimal generator. Such a process is not necessarily a solution of a stochastic differential equation of the form (1.3.1). However, the Ito diffusions we have presented make up a very large class of (general) diffusions and are sufficient for most applications.

The term Ito diffusion is typically reserved for the time homogeneous case (1.3.1), while a solution of (1.1.1) is simply called an Ito process. However, many of the above considerations apply in the more general setting of (1.1.1) as well. For this reason, we often use the term "diffusion" loosely to describe a solution of (1.1.1).

Let X be the solution of (1.1.1) with b and σ satisfying the Ito conditions (1.1.2) and (1.1.3). We denote by $\mathbb{E}_{tx}$ the expected value given that $X_t = x$ for $x \in \mathbb{R}$, $t \geq 0$. For suitable functions $f : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{R}$ we define

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}_{tx}[f(t+h, X(t+h)) - f(t,x)] \doteq Af(t,x). \tag{1.3.6}$$

For reasons we do not describe here, A is called the *backward evolution operator* of X. We denote by $\mathcal{D}(A)$ the class of all function f such that the above limit is defined for all $t \in [0, \infty)$, $x \in \mathbb{R}^n$. (Note that in the time homogeneous case we only consider f independent of t and thus (1.3.6) simplifies to (1.3.5).) As before, we want $\mathcal{D}(A)$ to be large enough so that A characterizes the Markov process X.

THEOREM 1.3.4. *Let X be the solution to* (1.1.1) *with b and σ satisfying* (1.1.2) *and* (1.1.2). *Let f be of class $C_p^{1,2}([0, \infty) \times \mathbb{R}^n)$. Then $f \in \mathcal{D}(A)$ and*

$$Af(t,x) = \frac{\partial f}{\partial t}(t,x) + \mathcal{L}f(t,x),$$

where $\mathcal{L}f$ is defined by (1.1.5).

1.4. Controlled Diffusion Processes

One of the main objects of interest in stochastic control theory is a controlled diffusion, described by the (controlled) stochastic differential equation,

$$dX_t = b(t, X_t, U_t)dt + \sigma(t, X_t, U_t)dW_t \ , \quad X_0 = x. \tag{1.4.1}$$

Here, the stochastic process U represents a *control*. Given a control process U, the *controlled process* X evolves like a diffusion, but now the "controller" can influence the behavior of X by modifying its drift or diffusion coefficient at each time t through choice of the value U_t. We assume that U takes values in a set $\mathbb{U} \subset \mathbb{R}^m$ called the *control space*.

Notice that (1.4.1) is a nontrivial generalization of (1.1.1). The coefficients in (1.1.1) depend on the underlying ω only through the value $X_t(\omega)$. For example, if $X_t(\omega) = X_t(\tilde{\omega})$ then b and σ take the same values at time t. However, this is

not necessarily the case for (1.4.1), where dependence on ω is present through the value $U_t(\omega)$. Thus, even if $X_t(\omega) = X_t(\tilde{\omega})$ the values of b and σ at time t corresponding to ω and $\tilde{\omega}$ may differ (if $U_t(\omega) \neq U_t(\tilde{\omega})$). Due to this "extra dependence" on the underlying ω we need to strengthen the conditions (1.1.2) and (1.1.3) to guarantee the existence of a unique (strong) solution of (1.4.1). One way to accomplish this is to require the Ito conditions to hold uniformly over all $\alpha \in \mathbb{U}$. Thus we assume there exists a constant $K \in (0, \infty)$ such that for all $t \in [0, \infty)$, $\alpha \in \mathbb{U}$, and $x, y \in \mathbb{R}^n$ we have

$$|b(t,x,\alpha) - b(t,y,\alpha)| + ||\sigma(t,x,\alpha) - \sigma(t,y,\alpha)|| \leq K|x-y|, \tag{1.4.2}$$

$$|b(t,x,\alpha)|^2 + ||\sigma(t,x,\alpha)||^2 \leq K^2(1+|x|^2). \tag{1.4.3}$$

THEOREM 1.4.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a filtration $\{\mathcal{F}_t\}$ and a d-dimensional $\{\mathcal{F}_t\}$-Brownian motion W. Let U be an $\{\mathcal{F}_t\}$-adapted process taking values in $\mathbb{U}$. Assume that b and σ satisfy* (1.4.2) *and* (1.4.3). *Then there exists a unique (strong) solution to* (1.4.1) *(in the sense introduced in Theorem 1.1.1).*

With a slight abuse of terminology, we refer to the solution X as a controlled diffusion process. Associated with a controlled diffusion is a family of operators $\{A^\alpha : \alpha \in \mathbb{U}\}$ where, for suitable functions $f : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{R}$,

$$A^\alpha f(t,x) = \frac{\partial f}{\partial t}(t,x) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t,x) b_i(t,x,\alpha) + \frac{1}{2}\sum_{i=1}^n \sum_{k=1}^n \frac{\partial^2 f}{\partial x_i \partial x_k}(t,x) a_{ik}(t,x,\alpha)$$

We say that a process U is *non-anticipative* with respect to B if $B(t+h) - B(t)$ is independent of $\sigma(\{U_s : 0 \leq s \leq t\})$ for all $t, h \geq 0$. Such a process U does not "anticipate" or "see the future" of B. Note that in Theorem 1.4.1 U is a non-anticipative control with respect to the Brownian motion B.

One can think of the solution X of (1.4.1) as being constructed in the following manner. First, the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is specified, along with the filtration $\{\mathcal{F}_t\}$ and the Brownian motion B. Then a stochastic process U is defined on this space which is non-anticipative with respect to the given B. Finally, given B and U, the process X is constructed using (1.4.1). Note that here we have specified the control U in advance; we did not need to know any values of X to determine U.

The main assumption on U in Theorem 1.4.1 is that it is $\{\mathcal{F}_t\}$-adapted. We will be interested in several special cases of such U. A $\mathbb{U}$-valued stochastic process U is a *feedback control* if U is adapted[3] to the canonical filtration of

[3]Actually, U should be progressively measurable with respect to the canonical filtration of X but we ignore this technical distinction. In particular, if U is adapted and has right-continuous or left-continuous then it is progressively measurable.

X (i.e. the filtration $\{\mathcal{F}_t^X\}$ where $\mathcal{F}_t^X = \sigma(\{X_s, 0 \leq s \leq t\})$). (Note that in general $\mathcal{F}_t^X$ is smaller than $\mathcal{F}_t$ and thus feedback controls are a strict subset of $\{\mathcal{F}_t\}$-adapted controls.) The terminology should be intuitive; in this case, the control chosen at time t can depend on current and past values of the system (since it is adapted). Since the chosen control influences future values, there is feedback in the system.

Note that if U is a feedback control then there exists a measurable function $f : [0, \infty) \times C([0, \infty) : \mathbb{R}^n) \mapsto \mathbb{U}$ for which $U_t = f(t, X(\cdot))$ where $X(\cdot)$ represents the sample path of X. Thus, in order to compute U we must first know X and hence a feedback control cannot be specified exogeneously as in the statement of Theorem 1.4.1. Instead we must consider (1.4.1) with $U(\cdot)$ replaced by $f(\cdot, X(\cdot))$. Unfortunately, the resulting equation may not have a strong solution. That is, we might not be able to construct a process X satisfying this equation on a given probability space with a given Brownian motion. The mathematical difficulty lies in the filtration. In Theorem 1.4.1 we only require U to be adapted to $\{\mathcal{F}_t\}$, which is specified in advance. But for a feedback control we require U to be adapted to the smaller filtration $\{\mathcal{F}_t^X\}$, a more restrictive condition. But this filtration depends on the process X — which is what we are trying to construct — and so we cannot specify $\{\mathcal{F}_t^X\}$ in advance. This circularity (X depends on U which depends on X) does not seem to fit into our notion of solution introduced so far.

Thus for a feedback control U we cannot expect in general to have a solution of (1.4.1) in the sense of Theorem 1.4.1. However, under certain conditions, a solution exists in a weak sense. That is, there exists some filtered probability space supporting a Brownian motion and a process X such that the desired equation holds. Note here how this notion of a solution differs from that of Theorem 1.4.1; in that theorem, the probability space, the filtration, and Brownian motion are specified in advanced; in the weak notion the probability space, the filtration, and the Brownian motion become part of the solution.

The difference between strong and weak solutions can be thought of as a difference in perspective. Consider equation (1.1.1). A strong solution represents an engineer's perspective where given W is the input to a black box whose output is X. On the other hand a weak solution represents a statistician's viewpoint in which one fits equation (1.1.1) to the data X with W being the residuals.

The notion of a weak solution of an SDE is very useful from a mathematical point of view; however, it is less intuitive than the notion of a strong solution. Thus, in order to minimize technical details, we will focus only on strong solutions. In order to do this, we must place certain conditions on the controls under consideration to guarantee a unique strong solution to (1.4.1). These requirements will become part of our conditions for admissibility, which we will introduce in the next chapter. We will see that these conditions are not

too restrictive. In particular, we will most often be interested in the class of Markov controls which, under certain conditions, will yield strong solutions.

A $\mathbb{U}$-valued U stochastic process is called a *Markov control* if $U_t = u(t, X_t)$ for some measurable function $u : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{U}$. The function u is called *Markov control function (or policy)*[4]. Clearly, any Markov control is a feedback control. If u does not depend explicitly on t — that is, if $U_t = u(X_t)$ for some measurable function $u : \mathbb{R}^n \mapsto \mathbb{U}$ — then we say that U is a *stationary Markov control* and that u is a *stationary Markov control function*. The terminology Markov control is used since the control at time t only depends on the state of the system at that time, and not on any past values of the process. Given the Markovian nature of an uncontrolled diffusion, it seems reasonable that a diffusion controlled by a Markov control should be a Markov process (and that a stationary Markov control should lead to a homogeneous Markov process). We state next some conditions under which this is true.

We wish to consider (1.4.1) where U is a Markov control given by $U_t = u(t, X_t)$ for some measurable function $u : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{U}$. As in the feedback case, the actual values of the process U will depend on X. However, now we are able to specify the Markov control function u in advance. Fix a function $u : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{U}$ and let $\varphi_u : [0, \infty) \times \mathbb{R}^n \mapsto [0, \infty) \times \mathbb{R}^n \times \mathbb{U}$ be given by $\varphi_u(t, x) = (t, x, u(t, x))'$. Define $b_u : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{R}^n$ by $b_u = b \circ \varphi_u$ where $b : [0, \infty) \times \mathbb{R}^n \times \mathbb{U} \mapsto \mathbb{R}^n$ is as given in the opening paragraphs of this section; similarly define $\sigma_u : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{R}^{n \times d}$ by $\sigma_u = \sigma \circ \varphi_u$. With the above transformations, the questions of existence and uniqueness of a solution to (1.4.1) with U a Markov control have been recast in terms of existence and uniqueness of a solution of the SDE

$$dX_t = b_u(t, X_t)dt + \sigma_u(t, X_t)dW_t \ , \quad X_0 = x, \tag{1.4.4}$$

for the fixed function u.

THEOREM 1.4.2. *Let $u : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{U}$ and suppose that b_u and σ_u as defined above satisfy* (1.1.2) *and* (1.1.3). *Then there exists a unique strong solution X of* (1.4.1), *where U is the Markov control given by $U_t = u(t, X_t)$. Suppose further that σ_u satisfies the uniform nondegeneracy condition: There exists some $\delta > 0$ such that for all (t, x) the eigenvalues of the matrix $a_u(t, x)$ are bounded below by δ. Then the controlled process X is a Markov process. In addition, if b_u and σ_u do not depend on t, then X is a homogeneous Markov process.*

The following theorem provides a converse in some sense. It says that any controlled diffusion which is Markov can be obtained through a Markov control.

[4]Often, the term Markov control is used to signify both the control process U_t and the function u for which $U_t = u(t, X_t)$. We use the terminology Markov control function for the latter term in order to avoid confusion.

THEOREM 1.4.3. *If X is a solution of* (1.4.1) *and a Markov process, then U can be taken to be a Markov control. If in addition X is time-homogeneous then U can be taken to be a stationary Markov control.*

The assumptions of Theorem 1.4.2 are fairly strong. For example, the assumption that σ_u is Lipschitz implies certain restrictive assumptions on the function u which are not made explicit in the theorem. However, attempting to relax the assumptions leads to technicalities that we wish to avoid here. For a survey of the various issues, see Section 2.2. of [**Bor05**] and references therein.

When (1.4.1) is controlled by a Markov control function u we often use the notation A^u rather than A^α, $\alpha \in \mathbb{U}$, where for suitable functions $f : [0,\infty) \times \mathbb{R}^n \mapsto \mathbb{R}$,

$$A^u f(t,x) = \frac{\partial f}{\partial t}(t,x) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t,x) b_i(t,x,u(x)) + \frac{1}{2}\sum_{i=1}^n \sum_{k=1}^n \frac{\partial^2 f}{\partial x_i \partial x_k}(t,x) a_{ik}(t,x,u(x))$$

In summary, we will be interested in three classes of controls — adapted controls, feedback controls, and Markov controls — which differ based on what information can be used to compute the value of the control U at time t. For adapted controls we may use all the information available at time t, represented by $\mathcal{F}_t$. Feedback controls can only depend on the values of the state process X_s for $s \in [0,t]$. Finally, Markov controls can only depend on the current value of the state process X_t.

Often we wish to consider (1.4.1) with a random rather than fixed initial condition. Let ξ be an $\mathcal{F}_0$-measurable random variable taking values in $\mathbb{R}^n$. We seek a solution of

$$dX_t = b(t,X_t,U_t)dt + \sigma(t,X_t,U_t)dW_t\ , \quad X_0 = \xi. \tag{1.4.5}$$

The following theorem generalizes Theorem 1.4.1 to account for the random initial condition $X_0 = \xi$, and provide bounds on the growth rates of the corresponding solution.

THEOREM 1.4.4. *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space equipped with a filtration $\{\mathcal{F}_t\}$ and a d-dimensional $\{\mathcal{F}_t\}$-Brownian motion W. Let U be an $\{\mathcal{F}_t\}$-adapted process taking values in $\mathbb{U}$. Let ξ be an $\mathcal{F}_0$-measurable random variable taking values in $\mathbb{R}^n$ which satifies $\mathbb{E}|\xi|^p < \infty$ for some $p \geq 1$. Assume that b and σ satisfy* (1.4.2) *and* (1.4.3)*. Then there exists a unique (strong) solution X to* (1.4.5) *(in the sense introduced in Theorem 1.1.1). Furthermore, for any*

$T \in (0,\infty)$ there exists a constant $C_{p,T} \in (0,\infty)$ such that

$$\mathbb{E} \sup_{0\le s\le T} |X(s)|^p \le C_{p,T}(1+\mathbb{E}|\xi|^p)\ , \tag{1.4.6}$$

and for all $s,t \in [0,T]$,

$$\mathbb{E}|X(t)-X(s)|^p \le C_{p,T}(1+\mathbb{E}|\xi|^p)|t-s|^{p/2}. \tag{1.4.7}$$

Moreover, let Let $\tilde{\xi}$ be an $\mathcal{F}_0$-measurable random variable taking values in $\mathbb{R}^n$ which satifies $\mathbb{E}|\tilde{\xi}|^p < \infty$ and let Y be the solution of (1.4.1) corresponding to initial state $Y(0) = \tilde{\xi}$ and control U. Then for any $T \in (0,\infty)$,

$$\mathbb{E} \sup_{0\le s\le T} |X(s)-Y(s)|^p \le C_{p,T}\mathbb{E}|\xi-\tilde{\xi}|^p. \tag{1.4.8}$$

CHAPTER 2

Basic Principles of Stochastic Control

In this chapter we present the basic structure of stochastic optimal control problems and key techniques for finding a solution. We will use the infinite horizon discounted cost problem as an illustration. We begin with the following motivating example from financial mathematics.

2.1. A Motivating Problem

Consider an investor can invest in two assets: a risk-free asset ("bank") with rate of return $r > 0$ and a risky asset ("stock") with mean rate of return $\mu > r$ and constant volatility $\sigma > 0$. Suppose that the price of the risk-free asset at time t, $S_0(t)$, satisfies $dS_0(t) = rS_0(t)dt$, while the price of the stock evolves according to the following SDE:

$$dS_1(t) = \mu S_1(t)dt + \sigma S_1(t)dB(t),$$

where $\{B(t)\}$ is a standard one-dimensional Brownian motion. (This is the standard Black-Scholes model.) Let $X_0(t)$ denote the investor's wealth in the bank at time $t \geq 0$, $X_1(t)$ the wealth in stock, and $X(t) = X_0(t) + X_1(t)$ the investor's total wealth, with $\Pi(t) = X_1(t)/X(t)$ the proportion of wealth invested in stock. The investor has some initial capital $X(0) = x > 0$ to invest and there are no further cash infusions. We also assume that the investor consumes wealth at nonnegative rate $C(t)$ at time $t \geq 0$.

We assume that the investor is a "small" investor whose actions do not influence the market; hence her behavior cannot influence the price of the stock. However, the investor can decide how much money to hold in stock. We call the processes C and Π *controls* – by choosing, at time t, the values $C(t)$ and $\Pi(t)$ the investor controls, or influences, the evolution of her wealth in the future. In this case, given a consumption process C and a portfolio process Π, the wealth of the investor evolves according to the SDE

$$\begin{aligned} dX_x^{\Pi,C}(t) &= [(r + \Pi(t)(\mu - r))X^{\Pi,C}(t) - C(t)]dt + \sigma\Pi(t)X^{\Pi,C}(t)dB(t), \\ X(0) &= x. \end{aligned} \tag{2.1.1}$$

The notation $X_x^{\Pi,C}$ is used to emphasize the dependence of the wealth process on the initial wealth and on the controls. Note that given an initial wealth x

and control processes C and Π, the law of the wealth process $X_x^{\Pi,C}$ is uniquely determined.

Often an investor has some goal in mind (e.g. to make as much money as possible) and would like to choose controls which will achieve this goal. Suppose that our investor derives utility from consumption and would like to maximize her total utility over the time period in which she invests. Letting g represent the investor's utility function, she may be interested in maximizing

$$\int_0^\infty e^{-\gamma t} g(C(t))dt. \tag{2.1.2}$$

Here, the constant $\gamma > 0$ is a discount factor which accounts for the time value of money (the investor would rather have \$1 today than \$1 in the future), and we have assumed the investor invests over an infinite time horizon. Unfortunately, future increments of the Brownian motion, and hence future prices of the stock, are unknown to the investor, so there is no hope of maximizing the above quantity for each underlying ω. Instead, a more realistic goal is to choose control strategies in order to maximize expected total utility:

$$\mathbb{E}\int_0^\infty e^{-\gamma t} g(C(t))dt. \tag{2.1.3}$$

Note that while the dependence on C is explicit, the above quantity also depends on x and Π; these quantities influence the investor's wealth which in turn influences future consumption and hence the reward (2.1.3). (To make this dependence more explicit, we often write $\mathbb{E}_x^{\Pi,C}$ to denote the expected value given that $X_0 = x$ and control processes Π and C are employed over the interval $[0, \infty)$.) Thus, starting from initial wealth x, the investor's goal is to find control strategies C and Π which maximize (2.1.3).

2.2. Basic Elements of a Stochastic Control Problem

The above investment-consumption problem (often called the "Merton problem") is an example of a *stochastic (optimal) control problem*. It incorporates several key elements which are common to many stochastic control problems. These include:

Time horizon. The time horizon in the investment-consumption problem is infinite, in which case we take our time index to be $[0, \infty)$. We will also consider problems with finite horizon: $[0, T]$ for $T \in (0, \infty)$; and indefinite horizon: $[0, \tau]$ for some stopping time τ (for example, the first exit time from a certain set).

(Controlled) State process. The state process is a stochastic process which describes the state of the physical system of interest. The state process is often given by the solution of a controlled stochastic differential equation of the form (1.4.1). The evolution of the state process is influenced by a control. The state

process takes values in a set $\mathbb{S}$ called the *state space*, which is typically a subset of $\mathbb{R}^n$. In the investment-consumption problem, the state process is the wealth process $X_x^{(\Pi,C)}$.

Control process. The control process is a stochastic process, chosen by the "controller" to influence the state of the system. For example, if the state is modeled by the SDE (1.4.1), the controller is able to modify the drift and diffusion coefficients. The control process takes values in the *control space* $\mathbb{U}$ which is typically a subset of $\mathbb{R}^p$. The controls in the investment-consumption problem are the processes Π and C.

Admissible controls. Typically, only controls which satisfy certain "admissibility" conditions can be considered by the controller. These conditions can be both technical, for example, integrability or smoothness requirements, and physical, for example, constraints on the values of the state process or controls. For example, in the investment-consumption problem we will only consider processes $X_x^{(\Pi,C)}$ for which a solution to (2.1.1) exists. We will also require $C(t) \geq 0$ and that the investor have nonnegative wealth at all times, which places further restrictions on the class of allowable controls.

Cost (reward) function. There is some cost associated with the system, which may depend on the system state itself and on the control used. The cost function is typically expressed as a function $J(x, U)$, representing the *expected* total cost starting from system state x if control process U is implemented. For example, a common cost function is $J(x, U) = \mathbb{E}\int_0^\infty g(t, X_t, U_t)dt$, where g is a "running cost" function. In the investment-consumption problem we have already explained the dependence of (2.1.3) on initial wealth x and control processes C and Π. Recalling that, with b and σ given, x, C, and Π are enough to determine the law of the wealth process — and since the reward (2.1.3) is defined as an expected value — these three quantities determine the value of the reward. Thus $J(x, C, \Pi) = \mathbb{E}_x^{\Pi,C}\int_0^\infty e^{-\beta t}g(C(t))dt$ denotes the reward associated with initial wealth x, consumption process C and portfolio process Π.

Value function. The value function describes the value of the minimum possible cost of the system (or maximum possible reward). It is usually denoted by V and is obtained, for initial state x, by optimizing the cost over all admissible controls. The goal of a stochastic control problem is to characterize the value function and find a control U^* whose cost attains the minimum value ($V(x) = J(x, U^*)$ for starting state x.)

Some questions of interest in a stochastic control problem include:

- Is there an optimal control?
- Is there an optimal Markov control?
- How can we find an optimal control?
- Can we compute or approximate an optimal control numerically?

- How does the value function behave?

Before attempting to answer any of these questions we make a few remarks about the cost function $J(x, U)$ associated with a control U and value function $V(x)$ and the corresponding notation. Note that since J is defined as an expected value it is a *nonrandom* function of its inputs; hence the same is true for the value function V. As we mentioned in the previous section, the Brownian motion drives the randomness in the state process of interest. Whatever our eventual definition of an admissible control, one minimal requirement will be that U is non-anticipative with respect to B. This is one why reason we define J as an expected value. If instead our cost was a random variable, like $\int_0^\infty g(t, X_t, U_t)dt$ or the quantity in (2.1.2), to be minimized (say almost surely), it is unreasonable to expect that we can find at time t a random value U_t which provides the optimal control action $U_t(\omega)$ for each future increment of the Brownian motion $B_{t+h}(\omega) - B_t(\omega)$. Thus the best we can hope to do is to minimize an expected value; we will see that this leads to a meaningful and rich theory.

The function $J(x, U)$ has two arguments: an initial state $x \in \mathbb{S}$ and a stochastic process U. In practical applications we might be interested in starting in a particular (specified) state at time 0; however, it is still useful to view the cost function J, and the value function V, as a function of a variable initial state x. We will see that considerations based on the Markovian nature of the problem (e.g. the dynamic programming principle) will play a key role in its analysis. For example, say we start from our specified starting state at time 0 and let the system run for t units of time, at which time we reach state x. If we now wish to make projections about future behavior of the system, due to the Markov property we only need to know the value x and none of the past history of the process. Therefore, we can view the system as if we are "starting" at time t in state x. For this reason, we often interpret $J(x, U)$ as the minimum "cost-to-go" if the system is in state x at some time and we implement control U from that time on.

On the other hand, fix x and consider $J(x, U)$ as a function of the stochastic process U. The notation at first might seem strange: In order to compute $J(x, U)$ at time 0 we need to know the whole future of the process U. It helps to notice that since $J(x, U)$ is defined as an expected value, e.g. $J(x, U) = \mathbb{E}\int_0^\infty g(t, X_t, U_t)dt$, we only need to know the law of (X, U) to determine the value $J(x, U)$. For example, suppose control U is defined in terms of a known Markov control function u. Then $J(x, U)$ can in fact be computed at time 0. In this case, the function u along with x, b, and σ completely specify the law of (X, U) which in turn determines any related expected values and, in particular, the value $J(x, U)$.

In general this notion — that the law of (X, U) is sufficient to determine $J(x, U)$ — is one reason why the "weak formulation" of a stochastic control problem is so convenient. While we do not detail the weak formulation here, we remark that it is analogous to the notion of a weak solution of an SDE mentioned in Section 1.4. In the weak formulation, choosing a control process U involves also choosing a filtered probability space and a Brownian motion B. Since only the law of the various processes matter in determining the cost function, the underlying probability space plays little role. Despite the usefulness of the weak formulation, in these notes we will only consider the strong formulation: the filtered probability space and Brownian motion will be fixed in advance and all definitions of admissible controls will be made with respect to these fixed objects. In many cases, the strong and weak formulations yield the same answer (i.e. optimal control); the main differerence lies in the technical details.

2.3. Dynamic Programming Principle

The primary goal of a stochastic control problem is to find an optimal control. In this section and the next we introduce one of the main tools which allows us to achieve this goal, the dynamic programming equation (DPE). We first provide a heuristic derivation of the DPE associated with a stochastic control problem with infinite horizon discounted cost. As the motivation is merely formal, we make many assumptions which will need to be verified. In the following sections, we begin to make rigorous the theory behind the DPE.

To simplify matters, we will work with the 1-dimensional controlled state process

$$dX_t = b(X_t, U_t) + \sigma(X_t, U_t)dW_t \ , \quad X_0 = x \in \mathbb{R}, \tag{2.3.1}$$

where U is the control process with control space $\mathbb{U} = \mathbb{R}$ and the state space is $\mathbb{S} = \mathbb{R}$. We will restrict our attention to Markov controls. Recall that is a control process U is a Markov control if it takes the form $U_t = u(X_t)$ for all $t \geq 0$ and some function Markov control function $u : \mathbb{R} \mapsto \mathbb{U} = \mathbb{R}$. We assume that the state process $X \equiv X_x^u$ associated with Markov control function u is a time-homogeneous Markov process with generator $A \equiv A^u$:

$$Af(x) = b(x, u(x))f'(x) + (1/2)\sigma^2(x, u(x))f''(x),$$

for f in a suitable subset of $C^2(\mathbb{R})$. From the description of A we see that the law of X is uniquely determined by the initial state x and the Markov control function u (and, of course, the given (fixed) functions b and σ).

We consider the infinite horizon discounted cost function

$$J(x, u) = \mathbb{E} \int_0^\infty e^{-\gamma s} g(X_s, u(X_s))ds.$$

(When the control process U is given by a Markov control function u we often denote the associated cost by $J(x,u)$ rather than $J(x,U)$.) With b and σ fixed, the law of X, and hence the expected value above, is completely specified by x and u. To simplify notation we write $\mathbb{E}$ rather than $\mathbb{E}_x^u$.

The value function for this stochastic control problem is

$$V(x) = \inf_u J(x,u),$$

where the infimum is taken over all Markov control functions. For this development, we will ignore questions of admissibility of controls. Furthermore, we assume that there exists an optimal Markov control function. That is, we assume that there is a Markov control function u^* which satisfies $V(x) = J(x,u^*)$ for all $x \in \mathbb{R}$. We denote by X^* the controlled process corresponding to u^*.

Using the optimal Markov control function u^*, we have for $x \in \mathbb{R}$ and $t \in (0,\infty)$,

$$V(x) = J(x,u^*) = \mathbb{E}\int_0^t e^{-\gamma s} g(X_s^*, u^*(X_s^*))ds + \mathbb{E}\int_t^\infty e^{-\gamma s} g(X_s^*, u(X_s^*))ds.$$

Considering the second term on the RHS of the above equation, we have

$$\begin{aligned}
\mathbb{E}\int_t^\infty e^{-\gamma s} g(X_s^*, u(X_s^*))ds &= e^{-\gamma t}\mathbb{E}\int_t^\infty e^{-\gamma(s-t)} g(X_s^*, u(X_s^*))ds \\
&= e^{-\gamma t}\mathbb{E}\int_0^\infty e^{-\gamma v} g(X_{v+t}^*, u(X_{v+t}^*))dv \\
&= e^{-\gamma t}\mathbb{E}\int_0^\infty e^{-\gamma v} g(\tilde{X}_v^*, u(\tilde{X}_v^*))dv.
\end{aligned} \tag{2.3.2}$$

Above we have used the change of variables $v = s - t$ and we have introduced the process $\{\tilde{X}_v^*, v \geq 0\}$, where $\tilde{X}_v^* = X_{v+t}^*$. Since X^* is a time-homogeneous Markov process, the same is true for $\tilde{X}^*$, and so the law of these processes is uniquely determined by the functions b and σ, the control function u^*, and the (starting) state at time 0. Thus, if $\tilde{X}_0^* = x$ then the processes X^* and $\tilde{X}^*$ have the same law. In particular, given $\tilde{X}_0^* = x$ we have $\mathbb{E}[\int_0^\infty e^{-\gamma v} g(\tilde{X}_v^*, u(\tilde{X}_v^*))dv | \tilde{X}_0^* = x] = \mathbb{E}\int_0^\infty e^{-\gamma s} g(X_s^*, u(X_s^*))ds = J(x,u^*)$, which is equal to $V(x)$ by optimality of u^*. Using similar reasoning applied to

the random starting state $\tilde{X}_0^* = X_t^*$ we have

$$\begin{aligned}
\mathbb{E}\int_t^\infty e^{-\gamma s} g(X_s^*, u(X_s^*))ds &= e^{-\gamma t}\mathbb{E}[\int_0^\infty e^{-\gamma v} g(\tilde{X}_v^*, u(\tilde{X}_v^*))dv] \\
&= e^{-\gamma t}\mathbb{E}[\mathbb{E}[\int_0^\infty e^{-\gamma v} g(\tilde{X}_v^*, u(\tilde{X}_v^*))dv|\tilde{X}_0^*]] \\
&= e^{-\gamma t}\mathbb{E}[J(\tilde{X}_0^*, u^*)] \\
&= e^{-\gamma t}\mathbb{E}[J(X_t^*, u^*)] \\
&= e^{-\gamma t}\mathbb{E}[V(X_t^*)].
\end{aligned}$$

Thus we have

$$V(x) = \mathbb{E}\int_0^t e^{-\gamma s} g(X_s^*, u^*(X_s^*))ds + e^{-\gamma t}\mathbb{E}[V(X_t^*)]. \tag{2.3.3}$$

Now let u be an arbitrary Markov control function. We construct a new Markov control function $\hat{u}$ which satisfies $\hat{u} = u$ from time 0 until time t and $\hat{u} = u^*$ afterwards. Letting X denote the associated controlled process, we have

$$V(x) \le J(x, \hat{u}) = \mathbb{E}\int_0^t e^{-\gamma s} g(X_s, u(X_s))ds + \mathbb{E}\int_t^\infty e^{-\gamma s} g(X_s, u^*(X_s))ds.$$

Using the same reasoning we used to derive (2.3.3), we obtain

$$V(x) \le \mathbb{E}\int_0^t e^{-\gamma s} g(X_s^u, u(X_s^u))ds + e^{-\gamma t}\mathbb{E}[V(X_t^u)], \tag{2.3.4}$$

where we have now made explicit in the notation the dependence of the RHS on the arbitrary Markov control function u. Summarizing (2.3.3) and (2.3.4), we have arrived at the following *dynamic programming principle*:

$$V(x) = \inf_u \mathbb{E}\Big[\int_0^t g(X_s^u, u(X_s^u))ds + e^{-\gamma t}V(X_t^u)\Big], \tag{2.3.5}$$

where the infinimum is over all Markov control functions.

The idea behind the dynamic programming principle is as follows. The expectation on the RHS of (2.3.5) represents the cost if we implement an arbitrary strategy from time 0 until time t and then implement the optimal strategy from time t on. Clearly, this cost will be no less than the cost associated with using the optimal strategy from the start, represented by the LHS of (2.3.5), $V(x)$. What equation (2.3.5) says is that if we determine the optimal strategy separately on each of the time intervals $[0, t]$ and $[t, \infty]$ we get the same answer as when we consider the whole time interval $[0, \infty)$ at once.

One question you may ask yourself is: How can we use the dynamic programming principle to compute an optimal control? Remember that the idea behind the DPP is that it is not necessary to optimize the control U_t over the entire

time interval $[0, \infty)$ at once; we can partition the time interval into smaller chunks, and optimize over each individually. We will see below that this idea becomes particularly powerful if we let the partition size go to zero: the calculation of the optimal control then becomes a pointwise minimization. That is, for each fixed state x we compute the optimal value of control, say $\alpha \in \mathbb{U}$, to apply whenever $X_t = x$. When considering only Markov controls, this may not seem like much. However, we will see that a DPP also applies when our value function is defined as the infinimum over larger classes of controls, e.g. adapted controls. Here, the pointwise minimization procedure provides a sharp contrast to the seemingly significantly harder task of finding an optimal control process $\{U_t, t \geq 0\}$ for all values of t simultaneously.

2.4. Dynamic programming equation

We will now let $t \to 0$ in the dynamic programming principle to obtain an ODE (in general, a PDE) which describes the local behavior of the value function. Our hope is that analysis of the resulting PDE will lead to an optimal control.

Let u be an arbitrary Markov control with corresponding state process X and generator A. Applying Ito's rule to the function $e^{-\gamma t}V(x)$ (of (t, x)) we have

$$e^{-\gamma t}V(X_t) = V(x) + \int_0^t e^{-\gamma s}[-\gamma V(X_s) + AV(X_s)]ds + \text{martingale},$$

where we have assumed that V is smooth enough and that the stochastic integral term is a martingale. Adding $\int_0^t e^{-\gamma s}g(X_s, u(X_s))ds$ and then taking expected values on both sides of the above equation we obtain

$$\begin{aligned} & \mathbb{E}\Big[\int_0^t e^{-\gamma s}g(X_s, u(X_s))ds + e^{-\gamma t}V(X_t)\Big] \\ = \quad & V(x) + \mathbb{E}\int_0^t e^{-\gamma s}[-\gamma V(X_s) + AV(X_s) + g(X_s, u(X_s))]ds. \end{aligned} \tag{2.4.1}$$

Now applying the dynamic programming principle (2.3.5) we see that the LHS of the above equation is no smaller than $V(x)$. Therefore

$$0 \leq \mathbb{E}\int_0^t e^{-\gamma s}[-\gamma V(X_s) + AV(X_s) + g(X_s, u(X_s))]ds$$

On the other hand, if u^* is an optimal Markov control function (which we assume exists), we have an equality:

$$0 = \mathbb{E}\int_0^t e^{-\gamma s}[-\gamma V(X_s^*) + A^*V(X_s^*) + g(X_s^*, u^*(X_s^*))]ds$$

where X^* denotes the corresponding controlled process with generator A^*. Now divide by t and take limits as $t \downarrow 0$. Note that $X_t \to x$, and we assume that

$u(X_t) \to u(x)$. Assuming that we can interchange the limit and the expected value and the various functions are smooth enough, we have

$$0 \leq [-\gamma V(x) + AV(x) + g(x, u(x))]$$

for any Markov control function u, with equality in the case where the control is optimal

$$0 = [-\gamma V(x) + A^* V(x) + g(x, u^*(x))]$$

We can summarize these last two displays by writing the equivalent statement:

$$0 = \inf_u [-\gamma V(x) + A^u V(x) + g(x, u(x))],$$

where, as usual, the infimum is taken over all Markov control functions. Now for fixed $x \in \mathbb{R}$, recalling the definition of A^u we see that the Markov control function u appears in the quantity to be minimized above only through the value $u(x) \in \mathbb{U}$. Thus, we can write the above equation equivalently as:

$$0 = \inf_{\alpha \in \mathbb{U}} [-\gamma V(x) + A^\alpha V(x) + g(x, \alpha)], \tag{2.4.2}$$

where

$$A^\alpha f(x) = b(x, \alpha) f'(x) + (1/2)\sigma^2(x, \alpha) f''(x).$$

Equation (2.4.2)is known as the *dynamic programming equation* (DPE). The DPE is a nonlinear ODE (more generally, a PDE) which characterizes the local behavior of the value function of the control problem. The DPE provides the key to finding an optimal control. If we can find a solution to the DPE then we should be done. That is, suppose we have found a solution to (2.4.2); such a procedure involves finding, for each $x \in \mathbb{S} = \mathbb{R}$ the minimizing value $\alpha \in \mathbb{U}$. This provides a map between $x \in \mathbb{S}$ and $\alpha \in \mathbb{U}$, which according to the development above, must coincide with an optimal Markov control u^*. Moreover, the solution V of (2.4.2) is then the value function of the control problem. Note that what we have done here is precisely the limit of the recursive procedure described at the end of the previous section when the partition size goes to zero. We have reduced the computation to a pointwise optimization for every state x separately; indeed, the minimum in (2.4.2) is merely over the set $\mathbb{U}$, not over the set of $\mathbb{U}$-valued control processes $\{U_t, t \geq 0\}$. This makes finding optimal control strategies, if not easy, at least computationally feasible.

We have obtained a dynamic programming equation which in some sense characterizes the value function of our control problem. However, in order to arrive at this equation, we made many assumptions which need to be verified if we are to develop a rigorous theory. Issues that need to be addressed include:

- **Existence of an optimal control.** We have assumed the existence of an optimal Markov control. What are some conditons under which this is true? Do the DPP and DPE still hold if we do not assume existence of an optimal control? For example, it seems that the dynamic

programming principle (2.3.5) should hold whether or not there exists an optimal control.

- **Class of admissible controls.** We have only considered Markov controls, and not the more general class of adapted controls. Is it enough to work in this smaller class of controls? Furthermore, we have not properly defined requirements for admissibility of controls. For example, we might require a control process U to satisfy $|J(x,U)| < \infty$ for all $x \in \mathbb{R}$.
- **Smoothness of the value function.** In order to apply the classical form of Itô's rule as we have, the value function must be of class C^2. However, it is often difficult to verify smoothness of the value function a priori. Moreover, there are many examples of stochastic control problems, even relatively simple ones, in which the value function is not C^2.

This last issue — lack of smoothness of the value function — is one of the main obstacles in stochastic control theory. Nevertheless, the dynamic programming equation plays a key role. There are two basic strategies: a classical approach, which involves a technique known as *verification*, and a relatively newer approach which involves the theory of *viscosity solutions*. We will present the main idea behind the classical approach now, and introduce viscosity solutions in a later chapter.

2.5. Verification

In the previous section, we started with the control problem and derived the corresponding dynamic programming equation. The verification technique, in a sense, works backwards. Rather than starting with the optimal control problem, and showing that the dynamic programming equation follows, we will start with the DPE (regarded simply as a nonlinear PDE) and suppose that we have found a solution. We will then show that this solution coincides with the value function of an optimal control problem. The solution of the DPE also provides us with a guess for the optimal control policy in the control problem associated with the PDE. We then need to verify that this control is indeed optimal by showing that its cost is equal to the value function. This verification procedure is extremely practical: It says that if we can actually find a nice solution to the DPE, then that solution gives an optimal control, which is what we care about in practice. This will allow us to solve a variety of control problems, while avoiding almost all technicalities.

The key tool used in verification is Itô's rule. When we say that we have found a solution to the DPE, we implicitly assume that this solution has all the smoothness properties required for it to make sense as a solution of the PDE (in the classical sense). It is this assumption — assumption of a smooth

solution to the PDE — which gives us the smoothness required to use Itô's rule.

As an illustration of the verification procedure, we return to the DPE associated with our control problem, equation (2.4.2), which we write as

$$0 = \inf_{\alpha \in \mathbb{U}} [-\gamma \phi(x) + A^\alpha \phi(x) + g(x, \alpha)] \,, \quad x \in \mathbb{R}. \tag{2.5.1}$$

We write the equation in this manner to emphasize the fact that, for now, we are regarding it as a nonlinear ODE for which we wish to find a solution ϕ. This solution ϕ, a priori, may not have any relationship to the value function V. It is the goal of the verification procedure to establish this relationship (i.e. it is not assumed beforehand).

Now suppose there is a C^2 function ϕ which satisfies (2.5.1). Fix $x \in \mathbb{R}$ and let α_x^* denote the value $\alpha \in \mathbb{U}$ which minimizes $-\gamma\phi(x) + A^\alpha \phi(x) + g(x, \alpha)$; that is,

$$\alpha_x^* = \arg\min_{\alpha \in \mathbb{U}} [-\gamma \phi(x) + A^\alpha \phi(x) + g(x, \alpha)]$$

We assume that for each x such an α_x^* exists and is unique. Note this minimization procedure provides a map from $\mathbb{R}$ to $\mathbb{U}$. That is, each value $x \in \mathbb{R}$ is mapped to an $\alpha_x^* \in \mathbb{U}$ via the above equation. We denote this map by $u^* : \mathbb{R} \to \mathbb{U}$; we will soon see that it defines an optimal Markov control function.

Note that these two steps — finding a solution of an ODE and finding a value at which a certain function is minimized — are essentially calculus questions. That it, no stochastic process theory is needed to solve just these questions. But they are connected intimately with our control problem, as we now see.

Returning to our control problem, we now define the class of admissible controls. A stochastic process U_t is called an *admissible control (process)* if:

(a) U is $\{\mathcal{F}_t\}$-adapted;
(b) $U_t \in \mathbb{U}$ for all $t \geq 0$;
(c) Equation (2.3.1) has a unique strong solution;
(d) The process $\int_0^t e^{-\gamma s} \phi'(X_s) \sigma(X_s, U_s) dW_s$ is a martingale; and
(e) $e^{-\gamma t} \mathbb{E}[V(X_t)] \to 0$ as $t \to \infty$.

The first three assumptions seem to be minimal requirements. For example, the value of the control at time t should be able to be determined based on current and past information; we should not need to know the future. The last two assumptions above are sufficient to make our Itô's formula calculation work. We denote by $\mathcal{A}$ the class of all admissible controls. The cost function and value function are defined similar to before, but now for the larger class of

all admissible (not necessarily Markov) controls:

$$J(x,U) = \mathbb{E}\int_0^\infty e^{-\gamma t} g(X_t, U_t)dt$$
$$V(x) = \inf_{U\in\mathcal{A}} J(x,U),$$

where now the infimum is over all admissible controls.

Now comes the verification part. Recall the map $u^* : \mathbb{R} \mapsto \mathbb{U}$ defined via the minimization procedure above. We can view this map as a Markov control function, letting U^* and X^* denote the corresponding control process and state process (with $U_t^* = u^*(X_t^*)$). For $x \in \mathbb{R}$ we assume that $U^* \in \mathcal{A}$. Applying Ito's rule to the function $e^{-\gamma t}\phi(x)$, $t \in [0,\infty)$, we have

$$e^{-\gamma t}\phi(X_t^*) = \phi(x) + \int_0^t e^{-\gamma s}[-\gamma\phi(X_s^*) + A^*\phi(X_s^*)]ds + \int_0^t e^{-\gamma s}\phi'(X_s^*)\sigma(X_s^*, U_s^*)dW_s.$$

This calculation is rigorous since ϕ is a C^2 function. Now since we have assumed $U^* \in \mathcal{A}$ the stochastic integral above is a martingale with expected value 0. Adding $\int_0^t e^{-\gamma s} g(X_s^*, U_s^*)ds$ and then taking expected values on both sides of the above equation we obtain

$$\begin{aligned} & \mathbb{E}\int_0^t e^{-\gamma s} g(X_s^*, U_s^*))ds + e^{-\gamma t}\mathbb{E}[\phi(X_t^*)] \\ = \quad & \phi(x) + \mathbb{E}\int_0^t e^{-\gamma s}[-\gamma\phi(X_s^*) + A^*\phi(X_s^*) + g(X_s^*, U_s^*)]ds. \end{aligned} \tag{2.5.2}$$

Now since ϕ solves (2.5.1) and by construction of u^* (recall $u^*(x) = \alpha_x^*$, the mimimizing value in (2.5.1)), we have $-\gamma\phi(X_s^*) + A^*\phi(X_s^*) + g(X_s^*, U_s^*) = 0$ for all $s \in [0,\infty)$ (to see this, fix $s \in [0,\infty)$ and take $x = X_s^*$ and $U_s^* = u^*(X_s^*) = u^*(x) = \alpha_x^*$ in (2.5.1)). So the second term on the RHS of the above equation vanishes. We now wish to let $t \to \infty$. Since $U^* \in \mathcal{A}$, the second term on the LHS approaches 0 as $t \to \infty$. Thus, under sufficient conditions on g (e.g. g is bounded from below), letting $t \to \infty$ we have

$$\phi(x) = \mathbb{E}\int_0^\infty e^{-\gamma s} g(X_s^*, U_s^*)ds = J(x, U^*).$$

On the other hand, suppose $U \in \mathcal{A}$ is arbitrary and repeat the above calculation. The only change is that now we have the inequality $-\gamma\phi(X_s) + A\phi(X_s) + g(X_s, U_s) \geq 0$ for all $s \in [0,\infty)$ (since ϕ solves (2.5.1), taking $x = X_s$ and $\alpha = U_s$ there). Thus proceeding as above we have

$$\phi(x) \leq \mathbb{E}\int_0^\infty e^{-\gamma s} g(X_s, U_s)ds = J(x, U).$$

Since $U \in \mathcal{A}$ is arbitrary, this last equation implies

$$\phi(x) \leq \inf_{U \in \mathcal{A}(x)} J(x, U) = V(x).$$

But we have already shown that $\phi(x) = J(x, U^*)$. Thus we have $V(x) = J(x, U^*)$ and we have verified that U^* is an optimal control.

Note that there are still assumptions in the above calculations that need to be checked. However, in the verification technique we have shifted the burden from determining properties of the value function in advance to finding a solution of the DPE and a corresponding candidate control. Once we have these objects in hand, it is a relatively straightforward task to check that the candidate control satisfies the conditions for admissibility. In particular, it is significantly easier than trying to show directly that the value function is C^2.

2.6. Infinite horizon discounted cost problem

In this section we put the arguments in the previous sections on a rigorous footing. We first summarize the main steps that were taken to obtain an optimal control.

(1) Establish a dynamic programming principle.
(2) Let the time interval approach 0 in the dynamic programming principle, to arrive at a partial derivation equation (or inequality) which "characterizes" the value function. This PDE is called the dynamic programming equation, also known as the Hamilton-Jacobi-Bellman (HJB) equation.
(3) Find a function, say ϕ, which solves the dynamic programming equation. This step also involves finding a candidate control, say U^*, whose cost is equal to ϕ.
(4) Verify that $\phi = V$, the value function for the control problem, that U^* is admissible, and that therefore U^* is an optimal control.

With some practice one can usually write down the HJB equation without taking the formal steps 1 and 2 above. Thus, we will often skip directly to step 3, stating what the HJB equation is and proceeding from there.

We now return to the infinite horizon discounted cost problem and attempt to fill in any holes leftover from the previous sections. While it may seem repetitive, we include the entire setup and arguments here for completeness.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with filtration $\{\mathcal{F}_t\}$ and a d-dimensional $\{\mathcal{F}_t\}$-Brownian motion $\{W_t\}$. The σ-field $\mathcal{F}_t$ represents all the information available to the controller at time $t \geq 0$. Our state process will be given as a solution to the controlled SDE on $\mathbb{S} = \mathbb{R}^n$:

$$dX_t = b(X_t, U_t)dt + \sigma(X_t, U_t)dW_t\ , \quad X_0 = x \in \mathbb{R}^n. \tag{2.6.1}$$

We assume that b and σ satisfy (1.4.2), (1.4.3). Let $\mathbb{U} \subset \mathbb{R}^p$. An $\{\mathcal{F}_t\}$-adapted $\mathbb{U}$-valued stochastic process is an admissible control if: (a) There exists a unique solution X of (2.6.1), called the (controlled) state process corresponding to U; and (b) $\mathbb{E}\int_0^\infty e^{-\gamma t}|g(X_s, U_s)|dt < \infty$, where g is introduced in the next paragraph. We denote the set of all admissible controls by $\mathcal{A}$.

The cost function for using control process U starting from initial state x is

$$J(x, U) = \mathbb{E}\int_0^\infty e^{-\gamma t} g(X_t, U_t)dt,$$

where g is the running cost function. We assume that there is some $m \in (-\infty, \infty)$ such that $g(x, \alpha) \geq m$ for all $x \in \mathbb{R}^n$, $\alpha \in \mathbb{U}$. (In many application $m = 0$ and so g is a true "cost".) The constant $\gamma > 0$ is a discount factor. We include a discount factor for reasons both technical (i.e. to ensure some integrability conditions on the cost and value functions) and practical (i.e. discounting to account for present values in an economic sense.) The integrability condition contained in the definition of an admissible control ensures that $J(x, U)$ is well defined and finite for $x \in \mathbb{R}^n$, $U \in \mathcal{A}$.

The value function is then

$$V(x) = \inf_{U \in \mathcal{A}} J(x, U).$$

The HJB equation associated with this control problem is

$$0 = \inf_{\alpha \in \mathbb{U}}[-\gamma\phi(x) + A^\alpha \phi(x) + g(x, \alpha)], \quad x \in \mathbb{R}^n, \tag{2.6.2}$$

where $A^\alpha \phi(x)$ for $x \in \mathbb{R}^n$, $\alpha \in \mathbb{U}$ denotes (when the various derivatives exist)

$$\begin{aligned} A^\alpha \phi(x) &= \sum_{i=1}^n \frac{\partial \phi}{\partial x_i}(x) b_i(x, \alpha) + \frac{1}{2}\sum_{i=1}^n \sum_{k=1}^n \frac{\partial^2 \phi}{\partial x_i \partial x_k}(x) a_{ik}(x, \alpha) \\ &= \nabla\phi(x) \cdot b(x, \alpha) + \frac{1}{2}\mathrm{tr}(D^2\phi(x) a(x, \alpha)). \end{aligned}$$

Since no boundary conditions are specified, there may generally be many solutions to (2.6.2). It turns out that the value function is one which does not grow too rapidly as $|x| \to \infty$.

Our main result is the following verification theorem.

THEOREM 2.6.1. *Let $\phi : \mathbb{R}^n \mapsto \mathbb{R}$ be a function of class $C^2(\mathbb{R}^n) \cap C_p(\mathbb{R}^n)$; that is, ϕ is of class C^2 and ϕ satisfies a polynomial growth condition. Suppose also that ϕ solves equation* (2.6.2). *Then for all $x \in \mathbb{R}^n$:*

(a) *$\phi(x) \leq J(x, U)$ for any $U \in \mathcal{A}$, with corresponding state process X, which satisfies*

$$\liminf_{t \to \infty} e^{-\gamma t}\mathbb{E}[\phi(X_t)] \leq 0 \tag{2.6.3}$$

(b) *Suppose there exists* $U^* \in \mathcal{A}$*, with corresponding state process* X^**, which satisfies* $\lim_{t\to\infty} e^{-\gamma t}\mathbb{E}[\phi(X_t^*)] = 0$ *and for all* $s \geq 0$*,*

$$U^*(s) \in \arg\min_{\alpha\in\mathbb{U}}[-\gamma\phi(X^*(s)) + A^\alpha\phi(X^*(s)) + g(X^*(s),\alpha)]. \tag{2.6.4}$$

Then $\phi(x) = J(x, U^*)$.

Theorem 2.6.1 is a consequence of the following lemma, which should remind you of the dynamic programming principle.

LEMMA 2.6.2. *Let* ϕ *be as in Theorem 2.6.1 and let*[1] $t \in [0,\infty)$.

(a) *If* $U \in \mathcal{A}$ *with corresponding state process* X *then*

$$\phi(x) \leq \mathbb{E}\Big[\int_0^t e^{-\gamma s}g(X_s,U_s) + e^{-\gamma t}\phi(X_t)\Big]. \tag{2.6.5}$$

(b) *If there exists* $U^* \in \mathcal{A}$ *as in part (b) of Theorem 2.6.1 with corresponding state process* X^* *then*

$$\phi(x) = \mathbb{E}\Big[\int_0^t e^{-\gamma s}g(X_s^*,U_s^*) + e^{-\gamma t}\phi(X_\tau^*)\Big].$$

PROOF. Consider $U \in \mathcal{A}$ with corresponding state process X (a semimartingale). Suppose first that τ is an $\{\mathcal{F}_t\}$-stopping time and X is bounded on $[0,\tau]$; that is, there exists a constant $r \in (0,\infty)$ such that $\sup_{0\leq s\leq\tau(\omega)}|X_s(\omega)| \leq r$ almost surely. We apply Ito's rule to the function $e^{-\gamma t}\phi(x)$, with the process X, evaluated at τ which yields

$$e^{-\gamma\tau}\phi(X_\tau) = \phi(x) + \int_0^\tau e^{-\gamma s}[-\gamma\phi(X_s) + A^{U_s}\phi(X_s)]ds + \int_0^\tau e^{-\gamma s}\nabla\phi(X_s)\cdot\sigma(X_s,U_s)dW_s.$$

Since X is bounded on $[0,\tau]$ and by assumption $\nabla\phi$ and σ are continuous with σ satisfying the linear growth condition (1.4.3), the integrand in the stochastic integral term is bounded and hence the stochastic integral above has expected value 0. Taking expected values and adding $\mathbb{E}\int_0^\tau e^{-\gamma s}g(X_s,U_s)ds$ (which is finite since $U \in \mathcal{A}$) on both sides of the above equation we obtain

$$\begin{aligned}\phi(x) = {} & \mathbb{E}\int_0^\tau e^{-\gamma s}g(X_s,U_s)ds + e^{-\gamma\tau}\mathbb{E}[\phi(X_\tau)] \\ & - \mathbb{E}\int_0^\tau e^{-\gamma s}[-\gamma\phi(X_s) + A^{U_s}\phi(X_s) + g(X_s,U_s)]ds.\end{aligned} \tag{2.6.6}$$

Now since ϕ solves (2.6.2) we have the inequality $-\gamma\phi(X_s) + A^{U_s}\phi(X_s) + g(X_s,U_s) \geq 0$ for all $s \in [0,\infty)$ (taking $x = X_s$ and $\alpha = U_s$ in (2.6.2)). Using this inequality in (2.6.6) yields (2.6.5) with t there replaced by τ, at least in the case when X is bounded, an assumption which we will now relax.

[1]Actually, we can replace t by any $\{\mathcal{F}_t\}$-stopping time τ.

Consider again $U \in \mathcal{A}$ with corresponding state process X and define for each $r = 1, 2, \ldots$ a $\{\mathcal{F}_t\}$-stopping time $\tau_r \doteq \inf\{t \geq 0 : |X_t| \geq r\}$. Then for any $t \in [0, \infty)$, X is bounded on $[0, t \wedge \tau_r]$ and so by the preceding paragraph we have

$$\phi(x) \leq \mathbb{E} \int_0^{t\wedge\tau_r} e^{-\gamma s} g(X_s, U_s) ds + \mathbb{E}[e^{-\gamma(t\wedge\tau_r)} \phi(X_{t\wedge\tau_r})] \tag{2.6.7}$$

Note that $\tau_r \to \infty$ a.s. as $r \to \infty$ and hence $t \wedge \tau_r \to t$. Since $U \in \mathcal{A}$ we have

$$\mathbb{E} \int_0^{t\wedge\tau_r} e^{-\gamma s} |g(X_s, U_s)| ds \leq \mathbb{E} \int_0^{\infty} e^{-\gamma s} |g(X_s, U_s)| ds < \infty.$$

Also, since ϕ is continuous and X has continuous sample paths we have $\phi(X_{t\wedge\tau_r}) \to \phi(X_t)$ a.s. as $r \to \infty$. In addition, from the polynomial growth of ϕ, there are positive constants $K, q < \infty$ such that

$$\mathbb{E}[\phi(X_{t\wedge\tau_r})] \leq K\mathbb{E}[1 + |X_{t\wedge\tau_r}|^q] \leq K\mathbb{E}[1 + \Big(\sup_{0\leq s\leq t} |X_s|\Big)^q]$$

This last term does not depend on r and is finite due to standard estimates on solutions of stochastic differential equations; see, for example, Appendix D of [**FS06**] for a summary. Thus part (a) of the theorem follows upon taking limits as $r \to \infty$ in (2.6.7) and applying the dominated convergence theorem. The proof of part (b) is similar; the only difference is that now we have the equality $-\gamma\phi(X_s^*) + A^{U_s^*}\phi(X_s^*) + g(X_s^*, U_s^*) = 0$ for all $s \in [0, \infty)$. ∎

Proof of Theorem 2.6.1. Let $U \in \mathcal{A}$, with corresponding state process X, satisfy (2.6.3). From part (a) of Lemma 2.6.2 we have that (2.6.5) is satisfied for any $t \in [0, \infty)$. We now wish to let $t \to \infty$ in (2.6.5). Using (2.6.3) we have

$$\liminf_{t\to\infty} \mathbb{E} \int_0^t e^{-\gamma s} g(X_s, U_s) ds \geq \phi(x) - \liminf_{t\to\infty} e^{-\gamma t} \mathbb{E}[\phi(X_t)] \geq \phi(x).$$

Note that $\int_0^t e^{-\gamma s}|g(X_s, U_s)|ds \leq \int_0^\infty e^{-\gamma s}|g(X_s, U_s)|ds$ and $\mathbb{E}\int_0^\infty e^{-\gamma s}|g(X_s, U_s)|ds$ does not depend on t and is finite since $U \in \mathcal{A}$. An application of the dominated convergence theorem then yields $\lim_{t\to\infty} \mathbb{E}\int_0^t e^{-\gamma s} g(X_s, U_s)ds = \mathbb{E}\int_0^\infty e^{-\gamma s} g(X_s, U_s)ds$. It follows that $J(x, U) \geq \phi(x)$ which proves part (a). Proof of part (b) is similar. ∎

Note that the above theorem falls a little short of showing that ϕ is equal to the value function. What it does say it that $\phi(x) = \inf_{U\in\tilde{A}} J(x, U)$, where $\tilde{\mathcal{A}} = \{U \in \mathcal{A} : \liminf_{t\to\infty} e^{-\gamma t}\mathbb{E}[\phi(X_t)] \leq 0\}$, and that U^* is optimal among this (smaller) class of controls. However, this complication can often be avoided. In particular, we have the following corollary.

COROLLARY 2.6.3. *Suppose that ϕ satisfies the assumptions of Theorem 2.6.1 and is bounded, i.e.* $\sup_{x\in\mathbb{R}^n}|\phi(x)| = M < \infty$. *Then* $\phi(x) \leq V(x)$. *If* U^* *is as in part (b) of Theorem 2.6.1 then* $\phi(x) = V(x) = J(x, U^*)$ *for any* $x \in \mathbb{R}^n$.

PROOF. For any admissible control $U \in \mathcal{A}$,

$$\liminf_{t\to\infty} e^{-\gamma t}\mathbb{E}[\phi(X_t)] \leq \liminf_{t\to\infty} e^{-\gamma t}M = 0.$$

Thus $\phi(x) \leq J(x, U)$ by part (a) of Theorem 2.6.1 and hence $\phi(x) \leq V(x)$. Part (b) of Theorem 2.6.1 implies $J(x, U^*) = \phi(x) \leq V(x)$ and the result follows. ∎

REMARK 2.6.4. Recalling equation (2.6.4) we see that at time $t \geq 0$ the control $U^*(t)$ depends on the process X only through the value $X^*(t)$, and so we expect U^* to be a Markov control. For this to be true, we need to find a function $u^* : \mathbb{R}^n \mapsto \mathbb{U}$ for which $U^*(t) = u^*(X(t))$ for all $t \geq 0$. The natural choice for u^* is

$$u^*(x) \in \arg\min_{\alpha\in\mathbb{U}}[-\gamma\phi(x) + A^\alpha\phi(x) + g(x, \alpha)] \tag{2.6.8}$$

Unfortunately, taking this u^* might not lead to an admissible control. For example, if $\mathbb{U}$ is compact and convex, then in order to have a *continuous* function u^* that satisfies (2.6.8) one generally needs that $-\gamma\phi(x) + A^\alpha\phi(x) + g(x, \alpha)$ as a function of α has a minimum on $\mathbb{U}$ at a *unique* $\alpha_x^* = u^*(x)$. When this uniqueness property fails, one must generally allow u^* to be a discontinuous function of x. However, in this case, the corresponding coefficients b_u and σ_u (see Section 1.4) may not satisfy the Ito conditions and thus equation (2.6.1) might not have a unique solution. For more on these points see Sections IV.3 – IV.5 of [**FS06**].

REMARK 2.6.5. Of course, the whole verification procedure in general depends on the existence of a solution ϕ of (2.6.2) which satisfies the assumptions of Theorem 2.6.1. Such existence results are a concern of the theory of second order PDE's. For further reading on the general theory see [**GT01**] and for some connections with stochastic control see Sections IV.4 and IV.5 of [**FS06**].

2.7. Merton Problem

We now return to the motivating example from Section 2.1. We follow the development of Section 2 of [**DN90**]. Recall that given initial wealth $x > 0$, a consumption process C, and a portfolio process Π (the controls), the wealth of the investor (the state process) evolves according to the SDE

$$dX(t) = [(r + \Pi(t)(\mu - r))X(t) - C(t)]dt + \sigma\Pi(t)X(t)dB(t). \tag{2.7.1}$$

The pair $U = (\Pi, C)$ is an admissible control for initial wealth x if it is $\{\mathcal{F}_t\}$-adapted and for all $t \geq 0$:

- $C(t) \geq 0$ and $\int_0^t C(s)ds < \infty$,
- $|\Pi(t)| \leq K_\Pi$ for some constant K_Π (which may depend on Π),
- $X(t) \geq 0$, where X is the unique strong solution of equation (2.7.1).

This last condition is a *state constraint*, which is a key feature in many stochastic control problems. In general, a solution of (2.7.1) takes values in $\mathbb{R}$; however, we are only allowed to consider controls which keep the state process in the state space $\mathbb{S} = [0, \infty)$. The control space is $\mathbb{U} = \mathbb{R} \times [0, \infty)$. Note that we do not assume that $\Pi(t) \in [0, 1]$; that is, we allow short-selling of stock and borrowing from the bank. We denote by $\mathcal{A}(x)$ controls which are admissible for initial wealth x. The dependence on x is due to the addition of the state constraints.

The investor's utility function is given by $g(c) = c^p/p$ for $c \geq 0$, where p is a constant with $0 < p < 1$. The reward associated with using control U for initial wealth x is:

$$J(x, U) = \mathbb{E} \int_0^\infty e^{-\gamma t} g(C(t))dt,$$

where $\gamma > 0$ is the constant discount factor. The value function for initial wealth x is then

$$V(x) = \sup_{U \in \mathcal{A}(x)} J(x, U).$$

It is somewhat amazing that this is one of a few examples of stochastic control problems in which a closed form solution is available for an optimal control. We will now obtain this optimal control through a verification lemma.

Let $\beta = (\mu - r)/\sigma$. We assume that[2]

$$\gamma > rp + \frac{\beta^2 p}{2(1-p)}. \tag{2.7.2}$$

The special form of the utility function leads to the following "homothetic property" of the value function which is very helpful in the analysis.

LEMMA 2.7.1. *$V(x) = V(1)x^p$ for all $x \geq 0$.*

PROOF. Consider $x > 0$ and let $a > 0$. Clearly, $g(ac) = a^p g(c)$, $c \geq 0$. Also, note that $\mathcal{A}(ax) = \{aU : U \in \mathcal{A}(x)\}$. Then

$$V(ax) = \sup_{U \in \mathcal{A}(ax)} \mathbb{E} \int_0^\infty e^{-\gamma t} g(C_t)dt = \sup_{U \in \mathcal{A}(x)} \mathbb{E} \int_0^\infty e^{-\gamma t} g(aC_t)dt = a^p V(x).$$

Now interchange the roles of a and x and set $a = 1$ to obtain the result for $x > 0$. Finally, we leave it to the reader to verify that $V(0) = 0$. ∎

[2]As stated in [**DN90**], "if this condition is violated then growth of discounted utility is possible and arbitrarily large utility may be obtained by policies of prolonged investment followed by massive consumption."

The associated HJB equation is:

$$0 = -\gamma\phi(x)+rx\phi'(x)+\sup_{\pi\in\mathbb{R}}[\pi(\mu-r)x\phi'(x)+(1/2)\pi^2\sigma^2x^2\phi''(x)]+\sup_{c\geq 0}[c^p/p-c\phi'(x)]. \tag{2.7.3}$$

Let $\phi \in C^2(\mathbb{R})$ be a solution of (2.7.3). For fixed $x > 0$, using methods from elementary calculus we see that the suprema in the terms in the HJB equation are attained at

$$\pi^*(x) = \frac{-(\mu - r)x\phi'(x)}{\sigma^2x^2\phi''(x)}\,, \quad c^*(x) = (\phi'(x))^{1/(p-1)}.$$

As usual, we would like to verify that ϕ is the value function. Since we know that V satisfies the homothetic property, it seems reasonable that the same should be true for ϕ: $\phi(x) = vx^p$, $x \geq 0$, where $v > 0$ is a constant to be determined. With this assumption on ϕ, the functions π^* and c^* simplify to:

$$\pi^*(x) = \frac{\mu - r}{\sigma^2(1-p)}\,, \quad c^*(x) = (pv)^{1/(p-1)}x.$$

Note that π^* does not depend on x. It can now be verified that $\phi(x) = vx^p$ is a solution of

$$\begin{aligned} 0 &= -\gamma\phi(x) + rx\phi'(x) + [(c^*(x))^p/p - c^*(x)\phi'(x)] \\ &\quad + [\pi^*(x)(\mu - r)x\phi'(x) + (1/2)(\pi^*(x))^2\sigma^2x^2\phi''(x)], \qquad x \geq 0. \end{aligned} \tag{2.7.4}$$

Solving for $v > 0$ we obtain the nonlinear equation:

$$\Big(\gamma - rp - \frac{\beta^2 p}{2(1-p)}\Big)v = (pv)^{1/(p-1)}(1-p)v, \tag{2.7.5}$$

which has a unique positive solution v provided that (2.7.2) is satisfied, given by

$$v = \frac{1}{p(1-p)^{p-1}}\Big(\gamma - rp - \frac{\beta^2 p}{2(1-p)}\Big)^{p-1}$$

We now use the Markov control functions π^* and c^* to define a Markov control $U^* = (\Pi^*, C^*)$: $\Pi^*(t) = \pi^*$ and $C^*(t) = c^*(X^*(t))$, $t \geq 0$, where X^* is the solution of (2.7.1) corresponding to U^*. In this case, equation (2.7.1) can be solved explicitly to obtain X^* (apply Ito's rule to $\log(x)$), a geometric Brownian motion,

$$X^*(t) = x\exp\Big\{\frac{1}{1-p}\Big(r - \gamma + \frac{\beta^2(2-p)}{2(1-p)}\Big)t - \frac{\beta^2}{2(1-p)^2}t + \frac{\beta}{1-p}B(t)\Big\}.$$

Thus $X^*(t) \geq 0$ and $C^*(t) \geq 0$ for all $t \geq 0$ and hence $U^* \in \mathcal{A}(x)$. We then verify that U^* is an optimal control by applying[3] Theorem 2.6.1.

[3]Technically, Theorem 2.6.1 as stated does not apply directly here since ϕ is not C^2 at $x = 0$. However, observing that $X^*(t) > 0$ for all $t \geq 0$ and $\phi \in C^2(0,\infty)$ we can proceed in

Summarizing the above we have:

THEOREM 2.7.2. *Assume* (2.7.2) *holds. Let*

$$v = \frac{1}{p(1-p)^{p-1}}\Big(\gamma - rp - \frac{\beta^2 p}{2(1-p)}\Big)^{p-1}$$
$$\pi^* = \frac{\mu - r}{\sigma^2(1-p)}$$
$$c^*(x) = (pv)^{1/(p-1)}x.$$

Then the value function is given by $V(x) = vx^p, x \geq 0$. *Let* $\Pi^*(t) = \pi^*$, $t \geq 0$, *and define a consumption process* C^* *via* $C^*(t) = c^*(X^*(t))$, *where* X^* *is the controlled process corresponding to* $U^* = (\Pi^*, C^*)$. *Then* $U^* \in \mathcal{A}(x)$ *and* $V(x) = J(x, U^*)$*; that is,* U^* *is an optimal control.*

The above optimal consumption-investment strategy can be described as follows. The investor trades continuously in order to keep a constant proportion of wealth π^* invested in stock. In addition, the investor consumes at a rate proportional to current wealth.

REMARK 2.7.3. The cases $g(c) = c^p/1$ for $p < 0$ and $g(c) = \log(c)$ can be handled in a manner similar to above. For the Merton problem with more general utility functions see Sections 5.8C and 5.8D of [**KS97**].

---

the usual way to verify directly that $J(x, U^*) = \phi(x) = V(x)$. See Theorem IV.5.1 of [**FS06**] for a precise statement of a verification lemma that can be applied here.

CHAPTER 3

Classical Problems in Diffusion Control

In this chapter we continue our study of "classical" problems in stochastic control. As in the previous chapter, the defining feature of such problems is that the state process is controlled through modification of its drift or diffusion coefficients. We will investigate other types of control influence in later chapters.

We begin by looking at a finite time horizon problem which is analogous to the problem in Section 2.6. We then consider problems where the state process is controlled until the time it exits some specified region.

As in the previous chapter, the main tool we will use in this chapter is the HJB equation and the verification procedure. Of course, verification involves first finding a (classical) solution of the HJB equation. We state without proof some sufficient conditions under which a solution of the HJB equation exists. Unfortunately, these assumptions are rather strong and there are many examples of stochastic control problems in which the associated HJB equation has no (classical) solution. For this reason, a relaxation of what it means to be a solution of a PDE is needed in order for dynamic programming to be a useful theory. This is accomplished through the theory of viscosity solutions, which we introduce in the next chapter,

Finally, we also consider briefly diffusion control problems with ergodic cost and risk sensitive cost criteria.

3.1. A note on infinite versus finite time horizon

In the previous chapter, we considered the infinite horizon control problem with time dependent coefficients. In that case, the value function was a function of the state variable only; it did not depend on time. Intuitively[1], starting from initial state x at any time t, the problem always looks the same; there is an infinite horizon and the future evolution of the process does not depend on the starting time.

When there are time dependent coefficients, or a finite time horizon, this "time homogeneity" is no longer present. For example, in the finite horizon case, starting from each initial time t the controller is faced with a different remaining

[1]This can be shown rigorously; assume that the value function is given by $V(t,x)$ and use a change of variables to show that $V(t,x) = V(0,x)$ for all $t \geq 0$, $x \in \mathbb{S}$.

horizon $T-t$ and thus the controller's decision may change depending on how close it is to the terminal time. For this reason, in general the cost function and the value function depend on both the initial state x and the initial time t. For example, $J(t,x,U)$ represents the expected total cost starting from system state x *at time* t if control process U is implemented; the corresponding expectation is often denoted $\mathbb{E}_{tx}^{U}$. Similarly, the value function $V(t,x)$ is the optimal value of the cost given that the initial state is x and the initial time is t. (Since the value function is often interpreted as the optimal "cost-to-go" it sometimes written as $V(T-t,x)$, a function of the remaining time, rather than $V(t,x)$, a function of the time elapsed.)

Time then becomes an additional parameter in the problem, which in some ways complicates the analysis. For example, even in the case of a one-dimensional state process, the associated HJB equation will be a PDE (recall that in the one-dimensional infinite horizon problem the HJB equation was an ODE). Of course, in the infinite horizon case we must often impose additional assumptions to ensure that the cost (defined as an integral over an infinite range of times) is well defined; but as we have seen this can be handled by incorporating a discount factor. Thus, in problems where the evolution of the state process does not depend on time (i.e. b and σ do not depend on t) an infinite horizon problem is usually easier to work with than the corresponding finite time version. While there is no explicit terminal time in an infinite horizon problems — which may some unrealistic from a practical point of view — the discount factor γ makes value fars into the future negligible, so there is at least an implied time horizon (related to $1/\gamma$).

With the above considerations in mind, we now turn our attention to the finite time horizon problem.

3.2. Finite time horizon problem

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space equipped with filtration $\{\mathcal{F}_t\}$ and a d-dimensional $\{\mathcal{F}_t\}$-Wiener process $\{W_t\}$. The time interval of interest $[0,T]$ where $T\in(0,\infty)$ is the terminal time. Starting from state $x\in\mathbb{R}^n$ at time $t\in[0,T)$ the state process X with state space $\mathbb{S}=\mathbb{R}^n$ evolves for $s\in[t,T)$ according to the controlled SDE

$$dX_s = b(s,X_s,U_s)ds + \sigma(s,X_s,U_s)dW_s\ , \quad X_t = x. \tag{3.2.1}$$

As usual, $\mathbb{U}\subseteq\mathbb{R}^p$ is the control space, the control U is a $\mathbb{U}$-valued, $\{\mathcal{F}_t\}$-adapted process, and b and σ are functions $b:[0,T)\times\mathbb{R}^n\times\mathbb{U}\mapsto\mathbb{R}^n$ and and $\sigma:[0,T)\times\mathbb{R}^n\times\mathbb{U}\mapsto\mathbb{R}^{n\times d}$. We assume that b and σ satisfy (1.4.2), (1.4.3). An $\{\mathcal{F}_t\}$-adapted $\mathbb{U}$-valued stochastic process is an admissible control if: (a) There exists a unique solution X of (3.2.1), called the (controlled) state process corresponding to U; and (b) $\mathbb{E}[\int_0^T|g(s,X_s,U_s)|ds+|h(X_t)|]<\infty$, where g and

h are introduced in the next paragraph. We denote the set of all admissible controls defined for $s \in [t, T)$ by $\mathcal{A}(t)$.

The cost (or cost-to-go) function for using control process U starting from initial state x at time t is

$$J(t, x, U) = \mathbb{E}\Big[\int_t^T g(s, X_s, U_s)ds + h(X_T)\Big], \tag{3.2.2}$$

where g is the running cost function and h is the terminal cost function. We do not make explicit assumptions on g and h; typically, these functions are assumed to be continuous with at most polynomial growth.

The value function $V : [0, \infty) \times \mathbb{R}^n \mapsto \mathbb{R}$ is given by

$$V(t, x) = \inf_{U \in \mathcal{A}(t)} J(t, x, U).$$

Using heuristic arguments similar to those in Section 2.4 we arrive at the HJB equation associated with the above finite horizon control problem:

$$0 = \inf_{\alpha \in \mathbb{U}} [A^\alpha \phi(t, x) + g(t, x, \alpha)] \ , \quad t \in [0, T), \ x \in \mathbb{R}^n,$$

where A^α is as introduced after Theorem 1.4.1. The HJB equation can also be written as

$$0 = \frac{\partial \phi}{\partial t}(t, x) + \inf_{\alpha \in \mathbb{U}} [\mathcal{L}^\alpha \phi(t, x) + g(t, x, \alpha)] \ , \quad t \in [0, T), \ x \in \mathbb{R}^n, \tag{3.2.3}$$

where $\mathcal{L}^\alpha \phi(t, x)$ for $t \geq 0$, $x \in \mathbb{R}^n$, $\alpha \in \mathbb{U}$ denotes (when the various derivatives exist)

$$\mathcal{L}^\alpha \phi(t, x) = \sum_{i=1}^n \frac{\partial \phi}{\partial x_i}(t, x) b_i(t, x, \alpha) + \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^n \frac{\partial^2 \phi}{\partial x_i \partial x_k}(t, x) a_{ik}(t, x, \alpha).$$

At the terminal time T we are charged a cost equal to $h(X_T)$; afterwards no further control is applied and no further costs accrue. Therefore, we also have the terminal condition

$$\phi(T, x) = h(x) \ , \quad x \in \mathbb{R}^n. \tag{3.2.4}$$

Comparing (3.2.3) and (2.6.2), the HJB equation for the infinite horizon problem, we see that the main difference is the presence of the time derivative term $\partial \phi / \partial t$ in (3.2.3). (Of course, there is also no $-\gamma \phi$ term in (3.2.3) because we are considering an undiscounted cost. If we included a discount factor in (3.2.2) then there would be a term $-\gamma \phi$ in (3.2.3) as there is in (2.6.2). It is certainly possible to include discounting in the finite horizon cost; it is just not as essential as it is in the infinite horizon problem.) Another significant component of the finite horizon problem is the addition of the terminal condition (3.2.4). Thus, in the verification procedure we only consider solutions of (3.2.3) which also satisfy (3.2.4).

Our main result is the following verification theorem, which is analogous to Theorem 2.6.1. Note that we no longer need to consider the "transversality conditions" that played a key role in Theorem 2.6.1.

THEOREM 3.2.1. *Let $\phi : [0,T] \times \mathbb{R}^n \mapsto \mathbb{R}$ be a function of class $C^{1,2}([0,\infty) \times \mathbb{R}^n) \cap C_p([0,\infty) \times \mathbb{R}^n)$; that is, ϕ is of class $C^{1,2}$ and ϕ satisfies a polynomial growth condition: $|\phi(t,x)| \le K_r(1+|x|^r)$ for some constants $r, K_r \in (0,\infty)$. Suppose also that ϕ solves equation (3.2.3) with terminal condition (3.2.4). Then for all $t \in [0,T]$ and $x \in \mathbb{R}^n$:*

(a) *$\phi(t,x) \le J(t,x,U)$ for any $U \in \mathcal{A}(t)$, and therefore $\phi(t,x) \le V(t,x)$.*
(b) *Suppose there exists $U^* \in \mathcal{A}(t)$, with corresponding state process X^*, which satisfies for all $s \in [t,T]$,*

$$U^*(s) \in \arg\min_{\alpha \in \mathbb{U}} [A^\alpha \phi(s, X^*(s)) + g(s, X^*(s), \alpha)]. \tag{3.2.5}$$

Then $\phi(t,x) = J(t,x,U^*) = V(t,x)$.

Theorem 3.2.1 is a consequence of the following lemma, which should remind you of a finite time horizon version of the dynamic programming principle. The proof is similar to that of Lemma 2.6.2 and is omitted.

LEMMA 3.2.2. *Let ϕ be as in Theorem 3.2.1 and let[2] $\tau \in [t,T]$.*

(a) *If $U \in \mathcal{A}(t)$ with corresponding state process X then*

$$\phi(t,x) \le \mathbb{E}\Big[\int_t^\tau g(s, X_s, U_s)ds + \phi(\tau, X_\tau)\Big]. \tag{3.2.6}$$

(b) *If there exists $U^* \in \mathcal{A}(t)$ as in part (b) of Theorem 3.2.1 with corresponding state process X^* then*

$$\phi(t,x) = \mathbb{E}\Big[\int_t^\tau g(s, X_s^*, U_s^*)ds + \phi(\tau, X_\tau^*)\Big].$$

PROOF OF THEOREM 3.2.1. Taking $\tau = T$ in part (a) of Lemma 3.2.2, we have for any $U \in \mathcal{A}(t)$,

$$\begin{aligned}\phi(t,x) &\le \mathbb{E}\Big[\int_t^T g(s, X_s, U_s)ds + \phi(T, X_T)\Big] \\ &= \mathbb{E}\Big[\int_t^T g(s, X_s, U_s)ds + h(X_T)\Big],\end{aligned}$$

where the equality in the second line follows since ϕ satisfies (3.2.4). Thus $\phi(t,x) \le J(t,x,U)$ and since $U \in \mathcal{A}(t)$ is arbitrary we have $\phi(t,x) \le V(t,x)$. This proves part (a). Part (b) of the theorem follows from part (b) of Lemma 3.2.2 in a similar manner. ∎

[2]Actually, we can replace the the non-random τ by any $\{\mathcal{F}_t\}$-stopping time τ taking values in $[t,T]$.

Merton problem with finite time horizon.

We consider a finite time horizon version of the Merton problem of Section 2.7. The time horizon is now $[0, T]$ and the cost associated with initial state x, time t, and admissible control U is

$$J(t, x, U) = \mathbb{E}\Big[\int_t^T e^{-\gamma s} g(s, X_s, U_s) ds + e^{-\gamma T} h(X_T)\Big], \tag{3.2.7}$$

where $g(c) = c^p/p$ is the running utility of consumption and $h(x) = x^p/p$ is the utility derived from terminal wealth, where $0 < p < 1$. As usual, the value function is $V(t, x) = \inf_{U \in \mathcal{A}(t)} J(t, x, U)$, where the definition of $\mathcal{A}(t, x)$, the set of admissible controls for initial condition (t, x), is an obvious modification of the definition of $\mathcal{A}(x)$ in Section 2.7.

In the infinite horizon case, we sought a solution of the form vx^p for a constant v. With a time dependent value function, we try a solution of the form $v(t)x^p$. First assume that

$$k \doteq \frac{1}{1-p}\Big(\gamma - rp + \frac{\beta^2 p}{2(1-p)}\Big) > 0, \tag{3.2.8}$$

and define

$$\kappa(t) = \frac{1 - e^{-k(T-t)}}{k} + e^{-k(T-t)}.$$

Using methods similar to those in Section 2.7, it can be shown that

$$V(t, x) = e^{-\gamma t}(\kappa(t))^{1-p} x^p/p = J(t, x, U^*),$$

where $U^* \in \mathcal{A}(t, x)$ is an optimal Markov control, with state process X^*, given by

$$\Pi^*(t) = \frac{\mu - r}{\sigma^2(1-p)}\,, \qquad C^*(t) = \frac{X^*(t)}{\kappa(t)}.$$

Note that the investment strategy is the same as in the infinite horizon case: We trade continuously in order to keep a constant fraction π^* of wealth in the stock. The consumption process is defined by the Markov control function $c^*(t, x) = (1/\kappa(t))x$. Thus, we still consume at a rate proportional to total wealth; however, now the proportion is time dependent. Note that as we approach the terminal time (as $t \to T$) $c^*(t, x) \to x$ and thus we consume at a rate equal to total wealth. Also, the terminal condition $V(T, x) = e^{-\gamma T} h(x)$ is satisfied.

We now let $T \to \infty$ in the finite horizon problem and compare the results with the infinite horizon problem. We have that $c^*(t, x) \to kx$ for any $t \in [0, T]$. Noting that $k = (pv)^{1/(p-1)}$, where v is defined in Section 2.7, we see that the optimal consumption strategy in the infinite time horizon problem is the limit of that in the finite horizon problem. Also, we see that $V(0, x) \to k^{p-1}x^p/p$, which is just the value function $V(x)$ for the infinite horizon problem.

For an analysis of the finite horizon Merton problem in a more general setting see [**KLS87**].

3.3. Indefinite time horizon problem

In the finite time horizon problem, we stop controlling the state process at the fixed time T. However, in many problems we may be forced to stop controlling the state process at an earlier, random time. For example, in applications in financial mathematics, we are often required to stop investing or consuming upon the first exit time from a certain solvency region (i.e. upon bankruptcy). In this section, we consider problems where we are required to stop controlling upon the first time of exit from a certain specified open set. We call such problems "indefinite time horizon problems." The problems in Sections 3.2 and 2.6 can be viewed as special cases of the problems considered in this section.

3.3.1. Indefinite time horizon problem with a bound on the terminal time. The setup is similar to that of Section 3.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with filtration $\{\mathcal{F}_t\}$ and a d-dimensional $\{\mathcal{F}_t\}$-Wiener process $\{W_t\}$. The time interval of interest $[0, T]$ where $T \in (0, \infty)$ is the terminal time. Starting from state $x \in O$ (where O is introduced below) at time $t \in [0, T)$ the state process X with state space evolves for $s \in [t, T)$ according to the controlled SDE

$$dX_s = b(s, X_s, U_s)ds + \sigma(s, X_s, U_s)dW_s\ , \quad X_t = x. \tag{3.3.1}$$

As usual, $\mathbb{U} \subseteq \mathbb{R}^p$ is the control space, the control U is a $\mathbb{U}$-valued, $\{\mathcal{F}_t\}$-adapted process, and b and σ are functions $b : [0, T) \times \mathbb{R}^n \times \mathbb{U} \mapsto \mathbb{R}^n$ and and $\sigma : [0, T) \times \mathbb{R}^n \times \mathbb{U} \mapsto \mathbb{R}^{n\times d}$. We assume that b and σ satisfy (1.4.2), (1.4.3). An $\{\mathcal{F}_t\}$-adapted $\mathbb{U}$-valued stochastic process is an admissible control if: (a) There exists a unique solution X of the above state equation, called the (controlled) state process corresponding to U; and (b) $\mathbb{E}[\int_0^T |g(s, X_s, U_s)|ds + |h(X_t)|] < \infty$, where g and h are introduced in the next paragraph. We denote the set of all admissible controls defined for $s \in [t, T)$ by $\mathcal{A}(t)$.

The cost function is similar to (3.2.2). However, now we only control up to the smaller of time T and the exit time of X from a given set. Let $O \subseteq \mathbb{R}^n$ be an open set and let $Q = [0, T) \times O$. Let $(t, x) \in Q$ and $U \in \mathcal{A}(t)$ with corresponding state process X and define the $\{\mathcal{F}_t\}$-stopping time

$$\tau \equiv \tau_{tx}^U = \inf\{s \geq t : (s, X(s)) \notin Q\}.$$

Like X, τ depends on the initial data and the control process U; however, we usually omit this dependence from the notation. Note that if $\tau < T$ then τ is the exit time of X from O, while $\tau = T$ if $X(s) \in O$ for all $s \in [t, T]$. Also, $(\tau, X(\tau)) \in \partial^* Q$, where

$$\partial^* Q = ([0, T] \times \partial O) \cup (\{T\} \times O).$$

The cost (or cost-to-go) function for using control process U starting from initial state x at time t is

$$J(t,x,U) = \mathbb{E}\Big[\int_t^\tau g(s,X_s,U_s)ds + h(\tau,X_\tau)\Big]. \tag{3.3.2}$$

When $\tau = T$, h is interpreted as the terminal cost, while if $\tau < T$, h is a "boundary cost" charged upon exiting the region O.

The value function $V : [0,\infty) \times \mathbb{R}^n \mapsto \mathbb{R}$ is given by

$$V(t,x) = \inf_{U\in\mathcal{A}(t)} J(t,x,U).$$

Note that some assumptions are needed on O. As in Section IV.2 of [**FS06**] we assume that either $O = \mathbb{R}^n$ or ∂O is a compact $(n-1)$-dimensional manifold of class C^3.

REMARK 3.3.1. If $O = \mathbb{R}^n$ then $\tau = T$ a.s. and thus the indefinite time horizon problem above is just the finite time horizon problem of section 3.2. In this case, $\partial^* Q = \{T\} \times \mathbb{R}^n$.

The HJB equation associated with the above control problem is:

$$0 = \frac{\partial\phi}{\partial t}(t,x) + \inf_{\alpha\in\mathbb{U}}[\mathcal{L}^\alpha\phi(t,x) + g(t,x,\alpha)]\ , \quad t\in[0,T),\ x\in O. \tag{3.3.3}$$

The only difference between (3.3.3) and (3.2.3) is the domain of the x values. On the other hand, the terminal condition is now

$$\phi(t,x) = h(t,x)\ , \quad (t,x)\in\partial^* Q. \tag{3.3.4}$$

The following verification theorem is essentially the same as Theorem (3.2.1); the only difference is that the state space is now $\overline{O}$ instead of all of $\mathbb{R}^n$.

THEOREM 3.3.2. *Let $\phi : \overline{Q} \mapsto \mathbb{R}$ be a function of class $C^{1,2}(Q) \cap C_p(\overline{Q})$. Suppose also that ϕ solves equation* (3.3.3) *with terminal condition* (3.3.4). *Then for all $t \in [0,T)$ and $x \in O$:*

(a) *$\phi(t,x) \le J(t,x,U)$ for any $U \in \mathcal{A}(t)$, and therefore $\phi(t,x) \le V(t,x)$.*

(b) *Suppose there exists $U^* \in \mathcal{A}(t)$, with corresponding state process X^* and exit time $\tau^* = \inf\{s \ge t : (s, X^*(s)) \notin Q\}$, which satisfies for all $s \in [t,\tau^*]$,*

$$U^*(s) \in \arg\min_{\alpha\in\mathbb{U}}[A^\alpha\phi(s,X^*(s)) + g(s,X^*(s),\alpha)]. \tag{3.3.5}$$

Then $\phi(t,x) = J(t,x,U^*) = V(t,x)$.

Theorem 3.3.2 is a consequence of the following lemma, which is Lemma 3.2.2 with fixed times replaced by stopping times. The idea of the proof is similar to that of Lemma 2.6.2, with appropriate modifications to account for the inclusion of the set O. For a detailed proof, we refer the reader to Lemma 3.1 of [**FS06**].

LEMMA 3.3.3. *Fix $(t,x) \in Q$. Let ϕ be as in Theorem 3.3.2 and let θ be any $\{\mathcal{F}_t\}$-stopping time taking values in $[t,T]$.*

(a) *If $U \in \mathcal{A}(t)$ with corresponding state process X then*

$$\phi(t,x) \leq \mathbb{E}\Big[\int_t^\theta g(s,X_s,U_s)ds + \phi(\theta,X_\theta)\Big].$$

(b) *If there exists $U^* \in \mathcal{A}(t)$ as in part (b) of Theorem 3.3.2 with corresponding state process X^* then*

$$\phi(t,x) = \mathbb{E}\Big[\int_t^\theta g(s,X_s^*,U_s^*)ds + \phi(\theta,X_\theta^*)\Big].$$

PROOF OF THEOREM 3.3.2. Let $U \in \mathcal{A}(t)$ with state process X and exit time τ. Taking $\theta = \tau$ in part (a) of Lemma 3.3.3, we have

$$\begin{aligned}\phi(t,x) &\leq \mathbb{E}\Big[\int_t^\tau g(s,X_s,U_s)ds + \phi(\tau,X_\tau)\Big] \\ &= \mathbb{E}\Big[\int_t^T g(s,X_s,U_s)ds + h(\tau,X_\tau)\Big],\end{aligned}$$

where the equality in the second line follows since ϕ satisfies (3.3.4). Thus $\phi(t,x) \leq J(t,x,U)$ and since $U \in \mathcal{A}(t)$ is arbitrary we have $\phi(t,x) \leq V(t,x)$. This proves part (a). Part (b) of the theorem follows from part (b) of Lemma 3.3.3 in a similar manner. ∎

3.3.2. Indefinite time horizon problem with unbounded terminal time. In the previous section control was applied until the earlier of the first exit time and some specified terminal time T. Now we remove the terminal time and control only until the first exit time, which may take the value ∞. The setup is similar to that of Section 2.6. Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space equipped with filtration $\{\mathcal{F}_t\}$ and a d-dimensional $\{\mathcal{F}_t\}$-Wiener process $\{W_t\}$. Starting from state $x \in O$ (where O is introduced below) the state process X with state space evolves according to the controlled SDE

$$dX_s = b(X_s,U_s)ds + \sigma(X_s,U_s)dW_s\ , \quad X_0 = x. \tag{3.3.6}$$

As usual, $\mathbb{U} \subseteq \mathbb{R}^p$ is the control space, the control U is a $\mathbb{U}$-valued, $\{\mathcal{F}_t\}$-adapted process, and b and σ are functions $b : \times\mathbb{R}^n \times \mathbb{U} \mapsto \mathbb{R}^n$ and and $\sigma : \times\mathbb{R}^n \times \mathbb{U} \mapsto \mathbb{R}^{n\times d}$. We assume that b and σ satisfy (1.4.2), (1.4.3). An $\{\mathcal{F}_t\}$-adapted $\mathbb{U}$-valued stochastic process is an admissible control if: (a) There exists a unique solution X of the above state equation, called the (controlled) state process corresponding to U; and (b) $\mathbb{E}\int_0^\infty e^{-\gamma s}|g(X_s,U_s)|ds < \infty$, where the cost function will be introduced below. We denote the set of all admissible controls by $\mathcal{A}$.

The cost function is similar to the infinite horizon discounted cost. However, now we only control until the exit time of X from a given set at which time we are charged an "exit cost". Let $O \subseteq \mathbb{R}^n$ be an open set. Let $x \in O$ and $U \in \mathcal{A}$ with corresponding state process X and define the $\{\mathcal{F}_t\}$-stopping time

$$\tau \equiv \tau_x^U = \inf\{s \geq 0 : X(s) \notin O\}.$$

We set $\tau = \infty$ if $X(t) \in O$ for all $t \geq 0$. The cost (or cost-to-go) function for using control process U starting from initial state x is

$$J(x, U) = \mathbb{E}\Big[\int_0^\tau e^{-\gamma s} g(X_s, U_s) ds + e^{-\gamma\tau} h(X_\tau)\, 1_{\{\tau<\infty\}}\Big]. \tag{3.3.7}$$

As usual, γ is the discount factor and g is the running cost. The function h represents a "boundary cost" charged upon exiting the region O; $1_{\{\tau<\infty\}}$ denotes the indicator function of the set $\{\tau < \infty\}$: $1_{\{\tau<\infty\}} = 1$ if $\tau(\omega) < \infty$ and it is 0 otherwise.

The value function $V : O \mapsto \mathbb{R}$ is given by

$$V(x) = \inf_{U \in \mathcal{A}} J(x, U).$$

As in the previous section, we assume that either $O = \mathbb{R}^n$ or ∂O is a compact $(n-1)$-dimensional manifold of class C^3.

REMARK 3.3.4. If $O = \mathbb{R}^n$ then $\tau = \infty$ a.s. and thus the indefinite time horizon problem above is just the infinite horizon discounted cost problem of section 2.6.

The HJB equation associated with the above control problem is:

$$0 = \inf_{\alpha \in \mathbb{U}} [-\gamma\phi(x) + A^\alpha \phi(x) + g(x, \alpha)]\,, \quad x \in O. \tag{3.3.8}$$

The only difference between (3.3.8) and (2.6.2) is the domain of the x values. In contrast to the problem of Section 2.6, we now have a boundary condition

$$\phi(x) = h(x)\,, \quad x \in \partial O. \tag{3.3.9}$$

The following verification theorem is essentially the same as Theorem (2.6.1); the only difference is that the state space is now $\overline{O}$ instead of all of $\mathbb{R}^n$.

THEOREM 3.3.5. *Let $\phi : \overline{O} \mapsto \mathbb{R}$ be a function of class $C^2(O) \cap C_p(\overline{O})$. Suppose also that ϕ solves equation* (3.3.8) *with boundary condition* (3.3.9)*. Then for all $x \in O$:*

(a) *$\phi(x) \leq J(x, U)$ for any $U \in \mathcal{A}$, with corresponding state process X and exit time τ, which satisfies*

$$\liminf_{t\to\infty} e^{-\gamma t} \mathbb{E}[\phi(X_t)\, 1_{\{\tau \geq t\}}] \leq 0. \tag{3.3.10}$$

(b) *Suppose there exists* $U^* \in \mathcal{A}$*, with corresponding state process* X^* *and exit time* τ^**, which satisfies* $\lim_{t\to\infty} e^{-\gamma t}\mathbb{E}[\phi(X_t^*)\, 1_{\{\tau^*\geq t\}}] = 0$ *and for all* $s \in [0, \tau^*)$*,*

$$U^*(s) \in \arg\min_{\alpha\in\mathbb{U}}[-\gamma\phi(X^*(s)) + A^\alpha\phi(X^*(s)) + g(X^*(s), \alpha)]. \tag{3.3.11}$$

Then $\phi(x) = J(x, U^*)$*.*

As we remarked after the proof of Theorem 2.6.1, the above theorem does not allow us to conclude that ϕ is equal to the value function V. However, this equality does hold under some additional assumptions. In particular, we have the following corollary in the case where ϕ is bounded.

COROLLARY 3.3.6. *Suppose that* ϕ *satisfies the assumptions of Theorem 3.3.5 and is bounded, i.e.* $\sup_{x\in\mathbb{R}^n} |\phi(x)| = M < \infty$*. Then* $\phi(x) \leq V(x)$*,* $x \in O$*. If* U^* *is as in part (b) of Theorem 3.3.5 then* $\phi(x) = V(x) = J(x, U^*)$ *for any* $x \in O$*.*

Comment on the Merton problem.

Recall the infinite horizon Merton Problem of Section 2.7. As we mentioned in that section, a key feature of the problem is the state constraint: in order for the control U to be admissible, the corresponding state process must satisfy $X(t) \geq 0$ for all $t \geq 0$. The presence of state constraints in a stochastic control problem usually complicates the analysis. However, in the case of the Merton problem, we can use properties of the state dynamics to reformulate the infinite horizon problem with state constraints into an equivalent problem of the exit time type described above without state constraints.

Fix $x > 0$. For a control process U, with state process X, define

$$\tau \doteq \inf\{t \geq 0 : X(t) \notin (0, \infty)\} = \inf\{t \geq 0 : X(t) = 0\}$$

(the second expression is due to the fact that X has continuous sample paths). Recall that $\mathcal{A}(x)$ denotes the set of all admissible controls for the initial state x. For $U \in \mathcal{A}(x)$, due to the state dynamics in (2.7.1), if $\tau < \infty$ we have that $C(\tau) = 0$ is the only admissible consumption value and thus $X(s) = 0$ and $C(s) = 0$ for all $s \geq \tau$. Otherwise, if $\tau = \infty$, we have $X(t) > 0$ for all $t \geq 0$. Thus, for any $U \in \mathcal{A}(x)$ we have $X(s) > 0$ for $s \in [0, \tau)$ and $X(s) = 0$, $C(s) = 0$, $s \geq \tau$. Since $g(0) = 0$ we have for any $U \in \mathcal{A}(x)$:

$$J(x, U) = \mathbb{E}\int_0^\infty e^{-\gamma t} g(C(t))dt = \mathbb{E}\int_0^\tau e^{-\gamma t} g(C(t))dt.$$

Now let $\tilde{\mathcal{A}}$ denote the set of all control processes U for which all the conditions for admissibility in Section 2.7 hold, with the possible exception of the state constraint requirement. That is, $\mathcal{A}(x)$ is the subset of controls $U \in \tilde{\mathcal{A}}$ for which $X(t) \geq 0$ for all $t \geq 0$. (In the conditions for admissibility, the state constraint

is the only condition that depends on x in a non-trivial way. For this reason we omit x from the notation $\tilde{\mathcal{A}}$.) For $U \in \tilde{\mathcal{A}}$ define the reward

$$\tilde{J}(x,U) \doteq \mathbb{E}\int_0^\tau e^{-\gamma t} g(C(t))dt.$$

Define the value function of the "unconstrained" problem with the above exit time reward as

$$\tilde{V}(x) = \sup_{U\in\tilde{\mathcal{A}}} \tilde{J}(x,U).$$

It is clear from the above considerations that the Merton problem of maximizing the infinite horizon reward over all admissible controls which satisfy the state constraint requirement is equivalent to the unconstrained problem with an exit time reward; that is,

$$V(x) = \tilde{V}(x)\ , \quad x > 0.$$

Note that for $x > 0$ and U^*, the optimal control process of Section 2.7, we have that $\tau^* = \infty$ a.s.

In the Merton problem, the special form of the state dynamics and running cost function essentially render the state constraints irrelevant. However, this is not usually the case in stochastic control problems where state constraints have a far from trivial impact on the problem.

3.4. Existence of a solution of the HJB equation

The verification theorems we have seen all take the same form: If there exists a function that is a solution of the HJB equation and has certain properties (e.g. growth or boundedness properties), then the solution must be equal to the value function. Implicitly, verification tells us that the value function is the unique solution of the HJB equation within the class of functions having the specified (growth/boundedness/etc) properties. However, the verification procedure crucially depends on the existence of a solution of the HJB equation. In this section, following the lead of Section IV.4 of [**FS06**], we state without proof results from the theory of second order PDEs which provide sufficient conditions for existence.

We first consider the indefinite time horizon problem with bounded terminal time introduced in Section 3.3.2. Recall the HJB equation (3.3.3). The main assumption we introduce below is that the HJB equation is *uniformly parabolic*: There exists a constant $c \in (0,\infty)$ such that for all $(t,x,\alpha) \in [0,T) \times \mathbb{R}^n \times \mathbb{U}$ and all $\xi \in \mathbb{R}^n$,

$$\sum_{i=1}^n \sum_{k=1}^n a_{ik}(t,x,\alpha)\xi_i\xi_k \geq c|\xi|^2. \tag{3.4.1}$$

Notice that, in particular, if $a(t,x,\alpha) = 0$ for some (t,x,α) — that is, the state process may experience *deterministic* evolution at some points in time — the uniform parabolicity condition is not satisfied.

The following theorem provides an existence result in the case when O is bounded.

THEOREM 3.4.1. *Suppose that:*

(a) *Equation* (3.3.3) *is uniformly parabolic; that is,* (3.4.1) *holds.*
(b) O *is bounded and* ∂O *is a manifold of class* C^3.
(c) $\mathbb{U}$ *is compact.*
(d) σ *and its partial derivatives* $\partial\sigma/\partial t$, $\partial\sigma/\partial x_i$, $\partial^2\sigma/\partial x_i\partial x_k$ *are continuous on* $[0,T]\times\overline{O}\times\mathbb{U}$. *We assume analogous conditions for* b *and* g.
(e) h *is of class* $C^3([0,T]\times\mathbb{R}^n)$.

Then (3.3.3) *with terminal condition* (3.3.4) *has a unique solution* $\phi\in C^{1,2}(Q)\cap C(\overline{Q})$.

Note that the assumption that $\mathbb{U}$ is compact ensures that the infimum in (3.3.3) is attained at some value $\alpha^*\in\mathbb{U}$.

We now consider unbounded O; we take $O=\mathbb{R}^n$. We require additional growth or boundedness assumptions to guarantee existence of a solution.

THEOREM 3.4.2. *Suppose that:*

(a) *Equation* (3.3.3) *is uniformly parabolic; that is,* (3.4.1) *holds.*
(b) $O=\mathbb{R}^n$.
(c) $\mathbb{U}$ *is compact.*
(d) σ *and its partial derivatives* $\partial\sigma/\partial t$, $\partial\sigma/\partial x_i$, $\partial^2\sigma/\partial x_i\partial x_k$ *are continuous and bounded on* $[0,T]\times\mathbb{R}^n\times\mathbb{U}$. *We assume analogous conditions for* b *and* g.
(e) h *is of class* $C^3(\mathbb{R}^n)$ *and* h *and its partial derivatives are bounded.*

Then (3.3.3) *with terminal condition* $\phi(T,x)=h(x)$ *has a unique solution* $\phi\in C_b^{1,2}([0,T]\times\mathbb{R}^n)$.

In the above $C_b^{1,2}$ denotes the subset of bounded $C^{1,2}$ with bounded first and second partial derivatives.

The following theorem gives sufficient conditions for existence of an optimal Markov control function. Recall that our candidate function u^* satisfies

$$u^*(t,x)\in\arg\min_{\alpha\in\mathbb{U}}[\mathcal{L}^\alpha\phi(t,x)+g(t,x,\alpha)].$$

THEOREM 3.4.3. *Under the assumptions of either Theorem* (3.4.1) *or Theorem* (3.4.2) *the Markov control function* u^* *is optimal. That is,* $V(t,x)=J(t,x,U^*)$ *for any* (t,x), *where* (X^*,U^*) *are the state process and control*

process, respectively, corresponding to initial data (t, x) *and Markov control function* u^*.

See Theorem IV.4.4 of [**FS06**] for a proof.

3.5. Ergodic control

We return briefly to the infinite horizon setting and consider an alternative cost function of the "long term time-averaged" or ergodic form. This cost function is somewhat special in that it requires a different kind of treatment compared to the cost functions we have seen already. We only provide a sketch of the problem here; for a more detailed introduction to ergodic control, we refer the reader to [**Rob83**].

As usual, for initial state $x \in \mathbb{R}^n$ and a control process U, taking values in the control space $\mathbb{U}$, the corresponding state process X is given by

$$dX(t) = b(X(t), U(t))dt + \sigma(X(t), U(t))dW(t) \ , \quad X(0) = x.$$

We are now interested in a cost function of the following form

$$J(x, U) = \limsup_{T\to\infty} \mathbb{E}\Big[\frac{1}{T}\int_0^T g(X_s, U_s)ds\Big]. \tag{3.5.1}$$

The corresponding value function is

$$V(x) = \inf_{U\in\mathcal{A}} J(x, U),$$

where $\mathcal{A}$ is the set of admissible controls (which we leave undefined).

We have written the cost function and value function as functions of the initial state x. However, as is suggested by the RHS of (3.5.1) only the time-averaged asymptotic behavior of X and U will matter in determining the cost. The influence of the initial state x will disappear in the limit and thus we should expect that $J(x, U)$ and $V(x)$ do not depend on x. We will see below that this is true, but for now we keep the variable x in the notation.

Consider the following PDE, which is the HJB equation[3] associated with the above control problem:

$$0 = \inf_{\alpha\in\mathbb{U}}[A^\alpha\phi(x) + g(x, \alpha) - \eta] \ , \quad x \in \mathbb{R}^n. \tag{3.5.2}$$

The main difference between equations (3.5.2) and (2.6.2) is the presence of the auxiliary parameter η in (3.5.2). As usual, we would like to find a solution of the HJB equation and then verify that it gives the value function and an optimal control. However, finding a solution of (3.5.2) now involves finding a pair, a function $\phi : \mathbb{R}^n \mapsto \mathbb{R}$ and a constant $\eta \in \mathbb{R}$. As mentioned above, we

[3]The HJB equation can be derived formally by considering limits as the discount factor γ approaches 0 in the infinite horizon discounted cost problem. See the example in Section 1 of [**Rob83**] for an illustration of this "vanishing discount" limit approach.

expect that the value function of the control problem does not depend on x, and thus it should be given by a constant, say $V(x) = v$ for all x. It turns out that it is the parameter η, and not the function ϕ, which gives the constant value $v = \eta$. The function ϕ still plays a role: ϕ is used to determine the optimal control action $\alpha_x^* \in \mathbb{U}$ for a given state x. While we do not expect the optimal cost to depend on the current state, it is reasonable to expect that the optimal control action is state dependent (i.e. in general, it is not true that the same α is optimal for all states x).

We now briefly sketch the verification procedure, without making explicit any of our assumptions. As usual, the main tool used in verification is Ito's rule.

Suppose that $\phi \in C^2(\mathbb{R}^n)$ and $\eta \in \mathbb{R}$ satisfy (3.5.2). Let $x \in \mathbb{R}^n$ and let U be an admissible control; let X be the corresponding state process. We apply Ito's rule to $\phi(X_T)$ and take expected values in the resulting expansion to obtain

$$\mathbb{E}\phi(X_T) = \phi(x) + \mathbb{E}\int_0^T A^{U_s}\phi(X_s)ds.$$

We have assumed that the stochastic integral in the Ito expansion is a martingale. Now add η and $\mathbb{E}\int_0^T g(X_s, U_s)ds$ and divide by T on both sides of the above equation to obtain (after some rearranging)

$$\begin{aligned} & \mathbb{E}\Big[\frac{\phi(x) - \phi(X_T)}{T}\Big] + \eta \\ = \quad & \mathbb{E}\Big[\frac{1}{T}\int_0^T g(X_s, U_s)ds\Big] - \mathbb{E}\Big[\frac{1}{T}\int_0^T [A^{U_s}\phi(X_s) + g(X_s, U_s) - \eta]ds\Big]. \end{aligned}$$

Since ϕ and η satisfy (3.5.2) we have

$$\mathbb{E}\Big[\frac{\phi(x) - \phi(X_T)}{T}\Big] + \eta \leq \mathbb{E}\Big[\frac{1}{T}\int_0^T g(X_s, U_s)ds\Big].$$

Now we take lim sup as $T \to \infty$ on both sides to obtain

$$\eta \leq \limsup_{T\to\infty} \mathbb{E}\Big[\frac{1}{T}\int_0^T g(X_s, U_s)ds\Big] = J(x, U),$$

where we have assumed that for any admissible control

$$\limsup_{T\to\infty} \mathbb{E}\Big[\frac{\phi(x) - \phi(X_T)}{T}\Big] = 0. \tag{3.5.3}$$

Since U is an arbitrary admissible control, we have $\eta \leq V(x)$ for any $x \in \mathbb{R}^n$.

Now, with ϕ and η as above, fix an initial state $x \in \mathbb{R}^n$ and suppose there exists an admissible control U^*, with corresponding state process X^*, which satisfies for all $t \geq 0$,

$$U^*(t) \in \arg\min_{\alpha\in\mathbb{U}} [A^\alpha\phi(X^*(t)) + g(X^*(t), \alpha) - \eta].$$

We can repeat the above arguments with inequalities replaced by equalities to obtain

$$\eta = J(x, U^*),$$

and hence $\eta = V(x)$ for any $x \in \mathbb{R}^n$ and the control U^* is optimal.

We have made several assumptions above that need to be verified in order to make the calculations rigorous. Namely, we have assumed that the stochastic integral terms in the Ito expansions are martingales, and that (3.5.3) holds for any admissible control. Under appropriate assumptions on the problem data and conditions for admissibility, the above arguments can be turned into a verification theorem, but we do not do so here.

We see that while the main idea of the verification technique is still applicable, the ergodic control problem requires a different kind of treatment compared to the cost functions we have seen already. With a cost function such as

$$J(x, U) = \mathbb{E} \int_0^\infty e^{-\gamma t} g(X_t, U_t) dt,$$

behavior of the processes X and U over finite time intervals plays a crucial role. However, in the ergodic cost problem behavior of the processes X and U over finite time intervals is irrelevant and only the time-averaged asymptotic behavior matters. In general, (3.5.1) implies that some kind of stability of the system is sought. Many stochastic systems do exhibit regular patterns over a sufficiently long time period, and in these cases minimizing a long term average cost may better reflect the goal of the controller. When such stability is present in the system, it may even be possible to *almost surely* minimize the long run *pathwise* average cost

$$\limsup_{T \to \infty} \frac{1}{T} \int_0^T g(X_s, U_s) ds. \tag{3.5.4}$$

3.6. Risk sensitive control

In a risk sensitive control problem, the state and control processes are as in the classical control problems defined previously. The only difference is in the cost criteria. For initial data (t, x) and control U, with state process X, define the random variable $Y \equiv Y(t, x, U)$ as

$$Y \equiv Y(t, x, U) \doteq \int_t^T g(s, X(s), U(s)) ds + h(X(T)).$$

Then the cost function in the classical stochastic control problem is

$$J(t, x, U) = \mathbb{E}[Y(t, x, U)],$$

with the value function defined as usual.

In a risk sensitive control problem, we are instead interested in the cost function

$$J(t,x,U) = \mathbb{E}[\varphi(Y(t,x,U))],$$

where φ is a specified increasing function. A typical choice is $\varphi(x) = \rho\exp(\rho x)$ where $\rho \in \mathbb{R}$. The term "risk sensitive contrl" is motivated by the following approximation.

Suppose φ is differentiable at the point $\mathbb{E}[Y]$. Then from a Taylor expansion of φ we have

$$\varphi(Y) \approx \varphi(\mathbb{E}[Y]) + \varphi'(\mathbb{E}[Y])(Y - \mathbb{E}[Y]) + \frac{1}{2}\varphi''(\mathbb{E}[Y])(Y - \mathbb{E}[Y])^2,$$

and thus taking expected values we have

$$J(t,x,U) = \mathbb{E}[\varphi(Y)] \approx \varphi(\mathbb{E}[Y]) + \frac{1}{2}\varphi''(\mathbb{E}[Y])\mathrm{Var}[Y].$$

If φ is strictly convex near $\mathbb{E}[Y]$ then $\varphi''(\mathbb{E}[Y]) > 0$. Thus since we are trying to minimize the cost, the term $\varphi''(\mathbb{E}[Y])\mathrm{Var}[Y]$ can be seen as a penalty for choosing a control which leads to a large variance. This case is called *risk averse*. On the other hand, if φ is concave, then $\varphi''(\mathbb{E}[Y]) < 0$ and we can minimize our overall cost by choosing controls with a large variance; this is the *risk seeking* case. Finally, we have the *risk neutral* case: If $\varphi''(\mathbb{E}[Y]) = 0$ then $J(t,x,U) \approx \varphi(\mathbb{E}[Y])$ and since φ is increasing, in this case, minimizing $\varphi(\mathbb{E}[Y])$ is equivalent to minizing $\mathbb{E}[Y]$, the cost function in the classical control problem.

For more on risk sensitive control see Chapter VI of [**FS06**].

CHAPTER 4

Viscosity Solutions

In any stochastic control problem, a primary goal is to find or characterize in some way an optimal control. Dynamic programming and the related HJB equations provide a tool to achieve this goal. There are two basic approaches to dynamic programming. The first is the verification procedure which we have focused on in the previous chapters. Recall that in verification we begin with the HJB equation viewed as a PDE. This equation is motivated by a formal application of the dynamic programming principle; however, in the verification procedure it is not necessary to prove the DPP beforehand. Rather, we instead begin with the HJB PDE and look for a (classical) solution of it. Once we find such a solution, we then verify that it is equal to the value function of the corresponding control problem. Finding a solution of the PDE also suggests a candidate for an optimal control.

The verification procedure is relatively straightforward; however, it has several drawbacks. First, it is an indirect method: By starting with the HJB equation viewed solely as a PDE, we are not able to take full advantage of the structure of the underlying control problem in order to find a solution. Another major obstacle to verification is that it requires us to find a sufficiently smooth solution of the HJB equation. However, relatively strong assumptions are required in order to guarantee existence of a smooth solution in general. Also, as we have mentioned several times already, the value function is in general not smooth — in fact it may not even be differentiable — and thus we can not possibly expect verification to work in these cases. We provide in the first section of this chapter an example of a relatively simple problem which exhibits this non-smooth behavior. Thus, while relatively easy to implement, the verification procedure is an ad-hoc method which is not adequate for a meaningful general theory of stochastic control.

In this chapter, we introduce the second approach to dynamic programming which is based on the theory of *viscosity solutions* of PDEs. This approach proceeds in a more direct manner. Beginning with our control problem, we first prove rigorously a dynamic programming principle. Using the DPP we then prove that the value function is a solution of the HJB equation, if our notion of solution is interpreted in a weaker sense (i.e. a viscosity solution, which we will introduce later in the chapter). Finally, we establish that the value function

is the unique solution of the HJB equation (at least among a certain class of functions). Thus we show that the HJB equation completely characterizes the value function of the control problem, and then use this characterization to find an optimal control. There are several advantages to this approach over verification. In the viscosity solution approach our starting point is the control problem itself, rather than the HJB equation, and thus we can utilize insight or structure from the control problem at any step in our analysis. More importantly, the viscosity solution approach does not require us to establish smoothness of the value function. Since the value function is often not smooth, the viscosity solution approach works in a broad class of problems while the verification procedure often breaks down.

The main disadvantage to the viscosity solution approach is that it is somwewhat non-intuitive and involves many mathematical technicalities. In this chapter we only present basic elements of theory of viscosity solutions of PDEs and their use in optimal control problems. We illustrate in detail how to apply the viscosity solution approach in the finite time horizon problem of Section 3.2.

4.1. Non-differentiability of the value function

This example is taken from Example 11.2.6 of [**Oks03**]. The state process satisfies

$$dX(s) = U(s)dW(s)\ ,\ s \geq t, \quad X(t) = x > 0.$$

Let $\tau = \inf\{s \geq t : (s, X(s)) \notin [t, T) \times (0, \infty)\}$. The reward function is $J(t, x, U) = \mathbb{E}[h(X(\tau))]$, where

$$h(x) = \begin{cases} x^2, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

The value function is $V(t, x) = \sup_U J(t, x, U)$. Intuitively, once the state process is above 1, we should turn off the control in order to stay in the current state until the terminal time, at which time we would receive the maximal reward. Thus the main control decision only occurs for state values in $(0, 1)$. Consider two strategies. The first applies no control and so the terminal state is equal to the initial state and the reward is x^2.

Now suppose instead we take U to be very large when the state process is in $(0, 1)$. Such a state process will evolve like a Brownian motion with a very large diffusion coefficient. Therefore, with U large enough the process will exit $(0, 1)$ in a small amount of time, with probability close to 1. If we hit 0 first we are forced to stop and receive a reward of 0. If we hit 1 first we set our control to 0 and then receive a terminal reward of 1. So in this case, our expected reward at the terminal time is just the probability that the state process hits 1 before 0. (Of course, there is also a probability that the terminal time is reached

before the state exits $(0,1)$; but with large enough control this probability is negligible.) Putting the above a little more rigorously, we consider Markov control functions of the following form, for $m > 0$

$$u_m(x) = \begin{cases} m, & 0 \le x < 1, \\ 0, & x \ge 1. \end{cases}$$

Let $0 < x < 1$, $X_m = x + mW(t)$ be the corresponding state process, τ_m the exit time, and define $\tilde{\tau}_m = \inf\{s \ge t : X_m(s) \notin (0,1)\}$. Then for large m, $\tilde{\tau}_m \le \tau$ with high probability. Since $X_m(\tilde{\tau}_m) \in \{0,1\}$ we have

$$J(t,x,u_m) \approx \mathbb{E}[h(X_m(\tilde{\tau}_m))] = \mathbb{P}[X_m(\tilde{\tau}_m) = 1] = \mathbb{P}(W(\tilde{\tau}_m) = (1-x)/m] = x,$$

using the "gambler's ruin" problem for Brownian motion. (Again, the RHS above is really $x\mathbb{P}(\tilde{\tau}_m < T) + X_m(T)\mathbb{P}(\tilde{\tau}_m \ge T)$ but this approaches x as $m \to \infty$.) Notice that in particular this control is superior to the first strategy of never applying control. Letting $m \to \infty$ we have

$$V(t,x) \le \begin{cases} x, & 0 \le x < 1, \\ 1, & x \ge 1. \end{cases}$$

In fact, it can be shown that equality holds in the above display. Thus, the value function is not differentiable at the point $x = 1$.

4.2. Setting and assumptions

We now illustrate the viscosity solution approach in the finite time horizon problem. We recall the setup introduced in Section 3.2. The time interval of interest is $[0,T]$ where $T \in (0,\infty)$ is the terminal time.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with filtration $\{\mathcal{F}_t\}$ and a d-dimensional $\{\mathcal{F}_t\}$-Wiener process $\{W_t\}$. Starting from state $x \in \mathbb{R}^n$ at time $t \in [0,T)$ and using control process U the state process X with state space $\mathbb{S} = \mathbb{R}^n$ evolves for $s \in [t,T)$ according to the controlled SDE

$$dX_s = b(s, X_s, U_s)ds + \sigma(s, X_s, U_s)dW_s\,, \quad X_t = x. \tag{4.2.1}$$

As usual, $\mathbb{U} \subseteq \mathbb{R}^p$ is the control space. A $\mathbb{U}$-valued, $\{\mathcal{F}_t\}$-adapted process U defined on $[t,T]$ is an admissible control, denoted $U \in \mathcal{A}(t)$, if there exists a unique solution X of (4.2.1) for (t,x,U) and

$$\mathbb{E}\Big[\int_t^T |g(s, X_s, U_s)|ds + |h(X_T)|\Big] < \infty.$$

The cost function for using control process U starting from initial state x at time t is

$$J(t,x,U) = \mathbb{E}\Big[\int_t^T g(s, X_s, U_s)ds + h(X_T)\Big], \tag{4.2.2}$$

where g is the running cost function and h is the terminal cost function. The value function $V : [0,T] \times \mathbb{R}^n \mapsto \mathbb{R}$ is given by

$$V(t,x) = \inf_{U \in \mathcal{A}(t)} J(t,x,U).$$

Throughout this chapter we assume the following.

ASSUMPTION 4.2.1. *We assume that*

(a) *The control space $\mathbb{U}$ is a compact subset of $\mathbb{R}^q$.*
(b) *The functions $b : [0,T] \times \mathbb{R}^n \times \mathbb{U} \mapsto \mathbb{R}^n$, $\sigma : [0,T] \times \mathbb{R}^n \times \mathbb{U} \mapsto \mathbb{R}^{n \times d}$, $g : [0,T] \times \mathbb{R}^n \times \mathbb{U} \mapsto \mathbb{R}$, and $h : \mathbb{R}^n \mapsto \mathbb{R}$ are continuous.*
(c) *There exists a constant $L \in (0,\infty)$ such that for all $t \in [0,T]$, $x, y \in \mathbb{R}^n$, and $\alpha \in \mathbb{U}$,*

$$|\varphi(t,x,\alpha) - \varphi(t,x,\alpha)| \leq L|x-y|, \tag{4.2.3}$$

where $\varphi(t,x,\alpha) = (b(t,x,\alpha), \sigma(t,x,\alpha), g(t,x,\alpha), h(x))$.
(d) *For all $t \in [0,T]$, $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{U}$,*

$$|\varphi(t,x,\alpha)| \leq L(1+|x|). \tag{4.2.4}$$

Note that these assumptions have been chosen for convenience and are not the most general assumptions possible. Even so, compare Assumption (4.2.1) to the much stronger assumptions of Theorems 3.4.1 or 3.4.2 which are sufficient to guarantee existence of a smooth solution of the HJB to use in the verification procedure.

4.3. Dynamic programming principle

Let $t \in [0,T)$. Equation (4.2.1) considers solutions corresponding to the fixed initial state $X(t) = x$. In dynamic programming we often need to view the initial state as a random variable. Thus we seek solutions of

$$dX_s = b(s,X_s,U_s)ds + \sigma(s,X_s,U_s)dW_s\ , \quad X_t = \xi, \tag{4.3.1}$$

where ξ is an integrable $\mathcal{F}_t$-measurable random variable. Under Assumption (4.2.1) there exists a unique solution X of (4.3.1) corresponding to initial condition $X(t) = \xi$ and control $U \in \mathcal{A}(t)$. We define the associated cost as the $\mathcal{F}_t$-measurable random variable

$$J(t,\xi,U) \doteq \mathbb{E}\Big[\int_t^T g(s,X_s,U_s)ds + h(X_T)\Big|\mathcal{F}_t\Big], \tag{4.3.2}$$

and the value function is the $\mathcal{F}_t$-measurable random variable

$$V(t,\xi) = \inf_{U \in \mathcal{A}(t)} J(t,\xi,U).$$

Note that due to the definition of V as an infimum, for any $t \in [0,t)$, $x \in \mathbb{R}^n$, and $\varepsilon > 0$ there is $U_\varepsilon \in \mathcal{A}(t)$ such that

$$V(t,x) \le J(t,x,U_\varepsilon) \le V(t,x) + \varepsilon.$$

Such a U_ε is called an "ε-optimal" control. We claim that the same is true when x is replaced by a random initial state ξ. That is, for any $t \in [0,T)$, integrable $\mathcal{F}_t$-measurable random variable ξ, and $\varepsilon > 0$, there is $U_\varepsilon \in \mathcal{A}(t)$ such that (a.s.)

$$V(t,\xi) \le J(t,\xi,U_\varepsilon) \le V(t,\xi) + \varepsilon.$$

We now prove the following dynamic programming principle. The main ideas are the same as those in Section 2.3. However, in that section we motivated the DPP only in the context of Markov controls and under the assumption of existence of an optimal control. We now prove the DPP holds in general under the current setup.

THEOREM 4.3.1. *Let $t \in [0,T)$ and $x \in \mathbb{R}^n$ and suppose Assumption 4.2.1 holds. Then for any $\tau \in [t,T)$ we have*

$$V(t,x) = \inf_{U \in \mathcal{A}(t)} \mathbb{E}\Big[\int_t^\tau g(s, X_s^U, U_s)ds + V(\tau, X_\tau^U)\Big], \tag{4.3.3}$$

where X^U denotes the state processing corresponding to (t,x,U).

SKETCH OF PROOF. We denote the RHS of (4.3.3) by $\bar{V}(t,x)$. Let $0 \le t \le \tau \le T$, $x \in \mathbb{R}^n$ and $U \in \mathcal{A}(t)$, and let X^U be the state process corresponding to (t,x,U). We have by the tower property and definition (4.3.2),

$$\begin{aligned} J(t,x,U) &= \mathbb{E}\Big[\int_t^\tau g(s, X_s^U, U_s)ds\Big] + \mathbb{E}\Big[\mathbb{E}\Big[\int_\tau^T g(s, X_s^U, U_s)ds + h(X_T)|\mathcal{F}_\tau\Big]\Big] \\ &= \mathbb{E}\Big[\int_t^\tau g(s, X_s^U, U_s)ds\Big] + \mathbb{E}[J(\tau, X_\tau^U, U)] \\ &\ge \mathbb{E}\Big[\int_t^\tau g(s, X_s^U, U_s)ds\Big] + \mathbb{E}[V(\tau, X_\tau^U)]. \end{aligned}$$

Taking infimum over $U \in \mathcal{A}(t)$ yields $V(t,x) \ge \bar{V}(t,x)$.

We now consider the reverse inequality. Let $\varepsilon > 0$ and let $U_\varepsilon \in \mathcal{A}(\tau)$ be such that

$$V(t, X^U(\tau)) \le J(t,\xi,U_\varepsilon) \le V(t, X^U(\tau)) + \varepsilon.$$

Define a process $\bar{U}$ on $[t,T]$ by

$$\bar{U}(s) = \begin{cases} U(s) & t \le s < \tau, \\ U_\varepsilon(s) & \tau \le s \le T. \end{cases}$$

We claim that $\bar{U} \in \mathcal{A}(t)$. Let $\bar{X}$ be the state process corresponding to $(t, x, \bar{U})$. Then

$$\begin{aligned} V(t,x) \leq J(t,x,\bar{U}) &= \mathbb{E}\Big[\int_t^T g(s,\bar{X}(s),\bar{U}(s))ds\Big] \\ &= \mathbb{E}\Big[\int_t^\tau g(s,X_s^U,U_s)ds\Big] + \mathbb{E}\Big[\int_\tau^T g(s,\bar{X}(s),U_\varepsilon(s))ds + h(\bar{X}(T))\Big] \\ &= \mathbb{E}\Big[\int_t^\tau g(s,X_s^U,U_s)ds\Big] + \mathbb{E}[J(\tau,X_\tau^U,U_\varepsilon)] \\ &\leq \mathbb{E}\Big[\int_t^\tau g(s,X_s^U,U_s)ds\Big] + \mathbb{E}[V(\tau,X_\tau^U)] + \varepsilon. \end{aligned}$$

The third line above follows since the second term on the RHS of second line above is the cost associated with initial state $X^U(\tau)$ when control U_ε is used on $[\tau, T]$. The fourth line is due to the ε-optimality of U_ε. Now taking infimum over $U \in \mathcal{A}(t)$ on both sides of the last line above yields $V(t,x) \leq \bar{V}(t,x) + \varepsilon$, and as ε is arbitrary it follows that $V(t,x) \leq \bar{V}(t,x)$. ∎

REMARK 4.3.2. The sketch above contains the main ideas in the proof of the DPP. However, a rigorous proof of the DPP requires the weak formulation of the stochastic control problem. We refer the reader to the middle paragraph on page 176 of [**YZ99**] for a brief explanation. For a rigorous proof of the above DPP see Theorem 3.3.3 of [**YZ99**]. A key assumption in [**YZ99**] is that the filtration is generated by the Brownian motion; see equation (3.3.5) and following. In this case, our definition (4.3.1) is actually implied by the structure of the problem; see Lemma 3.3.2 of [**YZ99**]. For an alternative proof using an approximation argument see Section IV.7 of [**FS06**].

REMARK 4.3.3. It can be shown that the dynamic programming principle (4.3.3) still holds whenever τ is a stopping time with respect to the filtration $\{\mathcal{F}_s : s \in [t,T]\}$ which takes values in $[t,T]$. See Theorem IV.7.1 of [**FS06**].

The following theorem shows that equality in the DPP holds when an arbitrary control U is replaced by an optimal control U^*. The ideas in the proof are similar to those in the proof of Theorem 4.3.1; see Theorem 4.3.4 of [**YZ99**] for details.

THEOREM 4.3.4. *Let $t \in [0,T)$ and $x \in \mathbb{R}^n$ and suppose Assumption 4.2.1 holds. Suppose that $U^* \in \mathcal{A}(t)$ is optimal for (t,x): $V(t,x) = J(t,x,U^*)$. Let X^* denote the state process corresponding to (t,x,U^*). Then for any $\tau \in [t,T]$,*

$$V(x,t) = \mathbb{E}\Big[\int_t^\tau g(s,X_s^*,U_s^*)ds + V(\tau,X_\tau^*)\Big].$$

Moreover, for any $s \in [t,T]$,

$$V(x,X^*(s)) = \mathbb{E}\Big[\int_s^T g(r,X_r^*,U_r^*)dr + h(X_\tau^*)\Big|\mathcal{F}_s\Big].$$

4.4. Continuity of the value function

We first recall Theorem 1.4.4, slightly restated to fit the current setting of a fixed finite time.

THEOREM 4.4.1. *Under the Assumption 4.2.1 the unique solution* X *of* (4.2.1) *corresponding to initial state* $x \in \mathbb{R}^n$, *initial time* $t \in [0,T)$, *and control* $U \in \mathcal{A}(t)$ *satifies the following. For any* $p \geq 1$ *there exists a constant* $C_p \in (0,\infty)$ *such that*

$$\mathbb{E}\sup_{t\leq s\leq T}|X(s)|^p \leq C_p(1+|x|^p)\ , \tag{4.4.1}$$

and for all $r,s \in [t,T]$,

$$\mathbb{E}|X(r)-X(s)|^p \leq C_p(1+|x|^p)|r-s|^{p/2}. \tag{4.4.2}$$

Moreover, let $y \in \mathbb{R}^n$ *and let* Y *be the solution of* (1.4.1) *corresponding to initial state* $Y(0)=y$, *initial time* t, *and control* U. *Then for any* $p \geq 1$,

$$\mathbb{E}\sup_{t\leq s\leq T}|X(s)-Y(s)|^p \leq C_p|x-y|^p. \tag{4.4.3}$$

The following theorem shows that the underlying structure of our control problem is in some sense "inherited" by the value function. In particular, we have that V is Lipschitz continuous in the state variable and (1/2)-Hölder continuous in the time variable.

THEOREM 4.4.2. *Under Assumption 4.2.1 the value function* V *is continuous on* $[0,T]\times\mathbb{R}^n$. *Moreover, there exists a constant* $K \in (0,\infty)$ *such that for all* $t,s, \in [0,T]$ *and* $x,y, \in \mathbb{R}^n$

$$|V(t,x)| \leq K(1+|x|)\ , \tag{4.4.4}$$

$$|V(t,x)-V(s,y)| \leq K(|x-y|+(1+|x|\vee|y|)|t-s|^{1/2})\ . \tag{4.4.5}$$

PROOF. We only prove (4.4.5); the proof of (4.4.4) is similar. Let $0 \leq t \leq s \leq T$, $U \in \mathcal{A}(t)$ and let X be the state process corresponding to (t,x,U). Then U restricted to $[s,T]$ is in $\mathcal{A}(s)$. Let Y be the state process corresponding to (s,y,U). Since X and Y differ in both initial state and initial time, we combine (4.4.2) and (4.4.3) to obtain[1]

$$\mathbb{E}\sup_{s\leq r\leq T}|X(r)-Y(r)| \leq C(|x-y|+(1+|x|\vee|y|)|s-t|^{1/2}). \tag{4.4.6}$$

[1]We have assumed $t \leq s$ in which case $|x|\vee|y|$ can be replaced by $|x|$. Similarly, if $s \leq t$ we can replace $|x|\vee|y|$ with $|y|$. In either case, $|x|\vee|y|$ provides an upper bound.

Then we have

$$\begin{aligned}
& |J(t,x,U) - J(s,y,U)| \\
\leq\ & \mathbb{E}\int_t^s |g(r,X(r),U(r))|dr + \mathbb{E}\int_s^T |g(r,X(r),U(r)) - g(r,Y(r),U(r))|dr + \mathbb{E}|h(X(T) - h(Y(T)) \\
\leq\ & L\mathbb{E}\int_t^s (1+|X(r)|)dr + L\mathbb{E}\int_s^T |X(r)-Y(r)|dr + L\mathbb{E}|X(T)-Y(T)| \\
\leq\ & L|s-t|(1+\mathbb{E}\sup_{s\leq r\leq t}|X(r)|) + LT\mathbb{E}\sup_{s\leq r\leq t}|X(r)-Y(r)| + L\mathbb{E}\sup_{s\leq r\leq t}|X(r)-Y(r)| \\
\leq\ & L(1+C)\sqrt{T}|s-t|^{1/2}(1+|X|) + (LTC+LC)(|x-y| + (1+|x|\vee|y|)|s-t|^{1/2}).
\end{aligned}$$

The second line follows from (4.2.4), (4.2.3), and the fourth line follows from (4.4.1), (4.4.6) with $p = 1$. Thus (4.4.5) follows and in particular V is continuous. ∎

4.5. Viscosity solutions defined

Recall that the HJB equation is

$$0 = \frac{\partial V}{\partial t}(s,z) + \inf_{\alpha\in\mathbb{U}}[\mathcal{L}^\alpha V(s,z) + g(s,z,\alpha)]\ , \quad s\in[0,T),\ z\in\mathbb{R}^n, \tag{4.5.1}$$

where $\mathcal{L}^\alpha V(s,z)$ for $s\geq 0$, $z\in\mathbb{R}^n$, $\alpha\in\mathbb{U}$ denotes (when the various derivatives exist)

$$\begin{aligned}
\mathcal{L}^\alpha V(s,z) &= \sum_{i=1}^n \frac{\partial\phi}{\partial x_i}(s,z)b_i(s,z,\alpha) + \frac{1}{2}\sum_{i=1}^n\sum_{k=1}^n \frac{\partial^2 V}{\partial x_i\partial x_k}(s,z)a_{ik}(s,z,\alpha). \\
&= D_x V(s,z)\cdot b(s,z,\alpha) + \frac{1}{2}\mathrm{tr}(D_x^2 V(s,z)a(s,z,\alpha)),
\end{aligned}$$

and $D_xV(s,z)$ denotes the vector of first partial derivatives in the state variable and $D_x^2V(s,z)$ denotes the matrix of second order partial derivatives in the state variable, evaluated at (s,z). Note that we have written the HJB equation in terms of V, rather than an arbitrary function ϕ, to emphasize that we now wish to show directly that V is a solution of the HJB equation. For now, we focus the HJB equation itself, without the terminal condition (we will introduce it again later).

We first introduce some notation to write the HJB equation in a more standard form. Let $\mathcal{S}_+^n$ denote the set of all symmetric, nonnegative definite $n\times n$ matrices. Define a function[2] $\mathcal{H} : [0,T)\times\mathbb{R}^n\times\mathbb{R}^n\times\mathcal{S}_+^n \mapsto \mathbb{R}$ for $s\in[0,T)$,

[2]In analogy with a term occurring in classical mechanics, $\mathcal{H}$ is often called a (generalized) Hamiltonian.

$z \in \mathbb{R}^n$, $p \in \mathbb{R}^n$, $M \in \mathcal{S}_+^n$

$$\mathcal{H}(s,z,p,M) \doteq \sup_{\alpha\in\mathbb{U}}\Big[-p\cdot b(s,z,\alpha) - \frac{1}{2}\mathrm{tr}(Ma(s,z,\alpha)) - g(s,z,\alpha)\Big]. \quad (4.5.2)$$

Since $\mathbb{U}$ is compact the above supremum is always attained. For each fixed (s,z,α), $-p\cdot b(s,z,\alpha) - \frac{1}{2}\mathrm{tr}(Ma(s,z,\alpha))$ is a linear function of (p,M). Hence for each fixed $s \in [0,T)$, $z \in \mathbb{R}^n$, $\mathcal{H}(s,z,\cdot,\cdot)$ is a convex function on $\mathbb{R}^n \times \mathcal{S}_+^n$. Also note that for $s \in [0,T)$, $z \in \mathbb{R}^n$, $p \in \mathbb{R}^n$, $M, N \in \mathcal{S}_+^n$ we have

$$\mathcal{H}(s,z,p,M) \le \mathcal{H}(s,z,p,N) \text{ whenever } M \ge N, \quad (4.5.3)$$

where by $M \ge N$ we mean that the matrix $M - N$ is nonnegative definite.

Using the above notation the HJB equation (4.5.1) can be written as

$$-\frac{\partial V}{\partial t}(s,z) + \mathcal{H}(s,z,D_xV(s,z),D_x^2V(s,z)) = 0\,, \qquad (s,z) \in Q. \quad (4.5.4)$$

Before introducing the notion of a viscosity solution, we first recall what it means to be a solution of (4.5.4) in the classical sense. Recall that $Q = [0,T)\times\mathbb{R}^n$ and $\overline{Q} = [0,T]\times\mathbb{R}^n$.

DEFINITION 4.5.1. *Suppose that $V \in C^{1,2}(Q)$.*

(a) *V is a* classical subsolution *of (4.5.4) in Q if: For all $(s,z) \in Q$*

$$-\frac{\partial V}{\partial t}(s,z) + \mathcal{H}(s,z,D_xV(s,z),D_x^2V(s,z)) \le 0.$$

(b) *V is a* classical supersolution *of (4.5.4) in Q if: For all $(s,z) \in Q$*

$$-\frac{\partial V}{\partial t}(s,z) + \mathcal{H}(s,z,D_xV(s,z),D_x^2V(s,z)) \ge 0.$$

(c) *V is a* classical solution *of (4.5.4) in Q if it is both a classical subsolution and a classical supersolution of (4.5.4) in Q.*

The following facts about classical solutions provide a motivation for the definition of viscosity solutions.

LEMMA 4.5.2. *Suppose that $V \in C^{1,2}(Q)$.*

(a) *V is a classical subsolution of (4.5.4) if and only if: For all $(s,z) \in Q$ and all $\phi \in C^{1,2}(Q)$ for which $V - \phi$ has a local maximum at (s,z) we have*

$$-\frac{\partial \phi}{\partial t}(s,z) + \mathcal{H}(s,z,D_x\phi(s,z),D_x^2\phi(s,z)) \le 0.$$

(b) *V is a classical supersolution of (4.5.4) in Q if and only if: For all $(s,z) \in Q$ and all $\phi \in C^{1,2}(Q)$ for which $V - \phi$ has a local minimum at (s,z) we have*

$$-\frac{\partial \phi}{\partial t}(s,z) + \mathcal{H}(s,z,D_x\phi(s,z),D_x^2\phi(s,z)) \ge 0.$$

(c) *The above equivalences hold with "local" replaced by "global".*

The proof is left to the reader. The key idea is that, for example, if $V-\phi$ has a local maximum at (s,z) then $D_x^2(V-\phi)\leq 0$ (i.e. it is nonpositive definite) around (s,z); then apply (4.5.3).

Considering only the second statements in the first two parts of Lemma starting with "For all $(s,z)\in Q$" we see that no regularity of V is required for these statements. In essence, we have replaced V by the smooth "test function" ϕ anywhere derivatives are involved. It is this notion that leads to our definition of viscosity solutions.

DEFINITION 4.5.3. *Suppose that $V\in C(\overline{Q})$.*

(a) *V is a* viscosity subsolution *of* (4.5.4) *in Q if: For all $(s,z)\in Q$ and all $\phi\in C^{1,2}(Q)$ for which $V-\phi$ has a local maximum at (s,z) we have*

$$-\frac{\partial\phi}{\partial t}(s,z)+\mathcal{H}(s,z,D_x\phi(s,z),D_x^2\phi(s,z))\leq 0. \tag{4.5.5}$$

(b) *V is a* viscosity supersolution *of* (4.5.4) *in Q if: For all $(s,z)\in Q$ and all $\phi\in C^{1,2}(Q)$ for which $V-\phi$ has a local minimum at (s,z) we have*

$$-\frac{\partial\phi}{\partial t}(s,z)+\mathcal{H}(s,z,D_x\phi(s,z),D_x^2\phi(s,z))\geq 0. \tag{4.5.6}$$

(c) *V is a* viscosity solution *of* (4.5.4) *in Q if it is both a viscosity subsolution and a viscosity supersolution of* (4.5.4) *in Q*

The following theorem shows that the notions of viscosity solution and classical solution coincide when the function of interest is sufficiently smooth. The proof is left to the reader.

THEOREM 4.5.4. *Any classical solution of* (4.5.4)*is a viscosity solution of* (4.5.4). *If V is a viscosity solution of* (4.5.4) *and $V\in C^{1,2}(Q)$ then V is a classical solution of* (4.5.4).

REMARK 4.5.5. In Definition 4.5.3 one may replace "local maximum (minimum)" by "global maximum (minimum)" or even by "global strict maximum (minimum)".

4.6. Viscosity solution of the HJB equation

We now show that the value function is a viscosity solution of the HJB equation. The main steps are the same as usual: Apply Ito's rule, use the DPP, and take limits. The key difference is that now we apply Ito's rule to a smooth test function ϕ, rather than to V itself (which is not known to be smooth), and so the Ito's formula calculation is rigorous.

THEOREM 4.6.1. *Under Assumption 4.2.1 the value function V is a viscosity solution of (4.5.4) in $[0,T)\times\mathbb{R}^n$.*

Continuity of V on $[0,T]\times\mathbb{R}^n$ has been established in Theorem (4.4.2). To prove the theorem, it remains to prove the subsolution and supersolution properties; we prove each property separately in the following lemmas.

LEMMA 4.6.2. *Under Assumption 4.2.1 the value function V is a viscosity subsolution of (4.5.4) in $[0,T)\times\mathbb{R}^n$.*

PROOF. Fix $(s,z)\in Q$ and let $\phi\in C^{1,2}(Q)$ be such that $V-\phi$ has a global maximum at (s,z). Assume without loss of generality that $(V-\phi)(s,z)=0$ so that $V\le\phi$. For arbitrary $\alpha\in\mathbb{U}$ consider the constant control $U(r)=\alpha, r\in[s,T]$; then $U\in\mathcal{A}(s)$, and let X be the state process corresponding to (s,z,U). Fix $\rho\in(0,\infty)$ and for each $h\in(0,T-s)$ define an $\{\mathcal{F}_s : s\in[t,T]\}$-stopping time $\tau_h\doteq(s+h)\wedge\inf\{r\ge s : |X(r)-z|\ge\rho\}$. Applying Ito's rule to $\phi(\tau_h,X(\tau_h))$ we have

$$\begin{aligned}\phi(\tau_h,X(\tau_h)) &= \phi(s,z)+\int_s^{\tau_h}[\frac{\partial\phi}{\partial t}(r,X(r))+\mathcal{L}^\alpha\phi(r,X(r))]dr\\ &\quad+\int_s^{\tau_h}D_x\phi(r,X(r))\cdot\sigma(r,X(r),U(r))dW(r).\end{aligned}$$

On $[s,\tau_h]$ the state process X is bounded and thus by continuity of $D_x\phi$ and σ we have that the integrand in the above stochastic integral is bounded. Hence, taking expected values we have

$$\mathbb{E}\phi(\tau_h,X(\tau_h))=\phi(s,z)+\mathbb{E}\int_s^{\tau_h}[\frac{\partial\phi}{\partial t}(r,X(r))+\mathcal{L}^\alpha\phi(r,X(r))]dr. \tag{4.6.1}$$

Recall that from the DPP (4.3.3) and Remark (4.3.3) we have

$$V(s,z)\le\mathbb{E}\Big[\int_s^{\tau_h}g(r,X(r),U(r))dr+V(\tau_h,X(\tau_h))\Big]. \tag{4.6.2}$$

Thus since $V(s,z)=\phi(s,z)$ and $V\le\phi$ we have

$$\phi(s,z)\le\mathbb{E}\Big[\int_s^{\tau_h}g(r,X(r),U(r))dr+\phi(\tau_h,X(\tau_h))\Big]. \tag{4.6.3}$$

Combining (4.6.1) and (4.6.3) we obtain

$$0\le\mathbb{E}\Big[\int_s^{\tau_h}(\frac{\partial\phi}{\partial t}(r,X(r))+\mathcal{L}^\alpha\phi(r,X(r))+g(r,X(r),U(r)))dr\Big]. \tag{4.6.4}$$

Note that $(1/h)\int_s^{\tau_h}(\frac{\partial\phi}{\partial t}(r,X(r))+\mathcal{L}^\alpha\phi(r,X(r))+g(r,X(r),U(r)))$ is bounded uniformly in h due to the uniform bound on X on $[s,\tau_h]$ and the continuity of the functions in the integrand. Thus dividing both sides of (4.6.4) by h and

taking limits as $h \to 0$ an application of the dominated convergence theorem yields

$$0 \leq \frac{\partial \phi}{\partial t}(s,z) + \mathcal{L}^{\alpha}\phi(s,z) + g(s,z,\alpha)).$$

Equation (4.5.5) follows since $\alpha \in \mathbb{U}$ is arbitrary. Thus V satisfies the definition of a viscosity subsolution. ∎

LEMMA 4.6.3. *Under Assumption 4.2.1 the value function V is a viscosity supersolution of* (4.5.4) *in* $[0,T) \times \mathbb{R}^n$.

PROOF. Fix $(s,z) \in Q$ and let $\phi \in C^{1,2}(Q)$ be such that $V - \phi$ has a global minimum at (s,z). Assume without loss of generality that $(V-\phi)(s,z) = 0$ so that $V \geq \phi$. Fix $\varepsilon \in (0,\infty)$, $\rho \in (0,\infty)$, and $h \in (0, T-s)$. From the DPP (4.3.3) (and Remark 4.3.3) we can find $U_{\varepsilon,h} \in \mathcal{A}(s)$ such that

$$\mathbb{E}\Big[\int_s^{\tau_{\varepsilon,h}} g(r, X_{\varepsilon,h}(r), U_{\varepsilon,h}(r))dr + V(\tau_{\varepsilon,h}, X_{\varepsilon,h}(\tau_{\varepsilon,h}))\Big] \leq V(s,z) + \varepsilon h, \quad (4.6.5)$$

where $X_{\varepsilon,h}$ is the state process corresponding to $(s,z,U_{\varepsilon,h})$ and $\tau_{\varepsilon,h} \doteq (s+h) \wedge \inf\{r \geq s : |X_{\varepsilon,h}(r) - z| \geq \rho\}$. Applying Ito's rule to $\phi(\tau_{\varepsilon,h}, X_{\varepsilon,h}(\tau_{\varepsilon,h}))$ and taking expected values (as we did in Lemma 4.6.2) and combining with (4.6.5) we have

$$\mathbb{E}\int_s^{\tau_{\varepsilon,h}} [-\frac{\partial \phi}{\partial t}(r, X_{\varepsilon,h}(r)) - \mathcal{L}^{U_{\varepsilon,h}(r)}\phi(r, X_{\varepsilon,h}(r)) - g(r, X_{\varepsilon,h}(r), U_{\varepsilon,h}(r))]dr \geq -\varepsilon h. \quad (4.6.6)$$

Therefore,

$$\mathbb{E}\int_s^{\tau_{\varepsilon,h}} [-\frac{\partial \phi}{\partial t}(r, X_{\varepsilon,h}(r)) + \sup_{\alpha \in \mathbb{U}}(-\mathcal{L}^{\alpha}\phi(r, X_{\varepsilon,h}(r)) - g(r, X_{\varepsilon,h}(r), \alpha)]dr \geq -\varepsilon h. \quad (4.6.7)$$

By Assumption 4.2.1 and the definition of $\tau_{\varepsilon,h}$, integrand on the LHS of the above inequality divided by h is uniformly bounded in h for each $\varepsilon > 0$. Thus dividing both sides of (4.6.4) by h and taking limits as $h \to 0$ an application of the dominated convergence theorem yields

$$-\frac{\partial \phi}{\partial t}(s,z) + \sup_{\alpha \in \mathbb{U}}[-\mathcal{L}^{\alpha}\phi(s,z) - g(s,z,\alpha))] \geq -\varepsilon.$$

Equation (4.5.6) follows since ε is arbitrary. Thus V satisfies the definition of a viscosity subsolution. ∎

Two fundamental issues in PDE theory are existence and uniqueness of solutions. In the verification procedure, we encountered problems related to existence of a solution to the HJB equation. Since we were looking for classical solutions, we needed to find a $C^{1,2}$ function, which in general is not possible. Of course, if we do find such a function, then the verification theorems tell us

it is unique, at least in a certain class of functions (e.g. with a specified growth or boundedness property).

When considering viscosity solutions, the notion of solution is much weaker and thus it is easier to find a solution in the viscosity sense. Thus the issue of existence is usually not a problem. In fact, we have just proved existence of a viscosity solution of the HJB equation (4.5.4) by showing that the value function is one. However, now uniqueness of the solution becomes an issue. In order for the HJB equation to be an effective characterization of the value function we need to know that the solution is unique (in a certain class of functions). Fortunately, uniqueness does hold, though proof of such a result is beyond the scope of these notes. For a proof the following theorem, see Theorem 4.6.1 of [**YZ99**].

THEOREM 4.6.4. *Under Assumption 4.2.1 the value function V is the only viscosity solution of* (4.5.4) *which satisfies:* (4.4.5), (4.4.6), *and the terminal condition*

$$V(T,x) = h(x)\ , \quad x \in \mathbb{R}^n. \tag{4.6.8}$$

Define $G : \mathbb{U} \times [0,T) \times \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{S}_+^n \mapsto \mathbb{R}$ for $\alpha \in \mathbb{U}$, $s \in [0,T)$, $z \in \mathbb{R}^n$, $p \in \mathbb{R}^n$, $M \in \mathcal{S}_+^n$ by

$$G(\alpha, s, z, p, M) \doteq -p \cdot b(s,z,\alpha) - \frac{1}{2}\mathrm{tr}(Ma(s,z,\alpha)) - g(s,z,\alpha); \tag{4.6.9}$$

then $\mathcal{H}(s,z,p,M) = \sup_{\alpha\in\mathbb{U}} G(\alpha,s,z,p,M)$. Since we know that V is a viscosity solution of (4.5.4), a verification theorem like Theorem 2.6.1 suggests a sufficient condition for a control to be optimal. For initial data (t,x) suppose that we can find a control $U^* \in \mathcal{A}(t)$, with state process X^* corresponding to (t,x,U^*), which satisfies for all $s \in [t,T]$,

$$0 = -\frac{\partial V}{\partial t}(s, X^*(s)) + G(U^*(s), s, X^*(s), D_x V(s, X^*(s)), D_x^2 V(s, X^*(s))). \tag{4.6.10}$$

Then U^* should be an optimal control. The problem with this "verification" condition is that V is not known to be smooth, so we need to replace V with some test function ϕ whenever derivatives are involved. The following

THEOREM 4.6.5. *Fix $t \in [0,T)$ and $x \in \mathbb{R}^n$. Suppose there exists a control process $U^* \in \mathcal{A}(t)$, with state process X^* corresponding to (t,x,U^*), and processes q^*, p^*, M^* taking values in $\mathbb{R}, \mathbb{R}^n, \mathcal{S}_+^n$, respectively, such that for all $s \in [t,T]$:*

$$\mathbb{E}\int_t^T |q^*(r)|^2 dr < \infty\ , \quad \mathbb{E}\int_t^T |p^*(r)|^2 dr < \infty\ , \quad \mathbb{E}\int_t^T |M^*(r)|^2 dr < \infty\ ;$$

$$(q^*(s), -p^*(s), -M^*(s)) \in D_{t+,x}^{1,2,+} V(s, X^*(s))$$

$$0 = -q(s, X^*(s)) + G(U^*(s), s, X^*(s), p^*(s), M^*(s)). \tag{4.6.11}$$

Then $V(t,x) = J(t,x,U^*)$*, i.e.* U^* *is an optimal control.*

The terms q, p^*, and M^* play the role of $-\partial/V\partial t$, D_xV and D_x^2V respectively. For fixed (s,z) the set $D^{1,2,+}_{t+,x}V(s,z)$ is called the *second-order right parabolic superdifferential of* V *at* (s,z). We have that[3] $(q,p,M) \in D^{1,2,+}_{t+,x}V(s,z)$ if and only if there exists a function $\phi \in C^{1,2}([0,T]\times\mathbb{R}^n)$ such that: $V-\phi$ attains a strict global maximum at (s,z) and

$$\begin{aligned}\phi(s,z) &= V(s,z)\\ \partial\phi\partial t(s,z) &= q\\ D_x\phi(s,z) &= p\\ D_x^2\phi(s,z) &= M.\end{aligned}$$

[3]This is not the usual definition, but it can be shown that it is equivalent. See Section 4.5 of [**YZ99**] and Lemma 4.5.4 in particular.

CHAPTER 5

Optimal Stopping Problems

In the previous chapters, we have discussed several control problems where the goal was to optimize a certain performance criterion by selecting an appropriate feedback control policy. In this chapter, we will treat a somewhat different set of control problems; rather than selecting a continuous control to be applied to an auxiliary input in the system equations, our goal will be to select an optimal stopping time to achieve a certain purpose. Such problems show up naturally in many situations where a timing decision needs to be made, e.g., when is the best time to sell a stock? When should we decide to bring an apparatus, which may or may not be faulty, in for repair (and pay the repair fee)? How long do we need to observe an unknown system to be able to select one of several hypotheses with sufficient confidence? Such problems are called optimal stopping problems, and we will develop machinery to find the optimal stopping times. These ideas can also be extended to find optimal control strategies in which feedback is applied to the system at a discrete sequence of times; we will briefly discuss such impulse control problems at the end of this chapter.

5.1. Setting

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with filtration $\{\mathcal{F}_t\}$ and a d-dimensional $\{\mathcal{F}_t\}$-Wiener process $\{W_t\}$. Starting from state $x \in \mathbb{R}^n$ at time $t \in [0, T)$ the state process X with state space $\mathbb{S} = \mathbb{R}^n$ evolves according to the (uncontrolled) SDE

$$dX(t) = b(X(t))dt + \sigma(X(t))dW(t) \ , \quad X(0) = x. \tag{5.1.1}$$

As usual, b and σ are functions $b : \mathbb{R}^n \mapsto \mathbb{R}^n$, $\sigma : \mathbb{R}^n \mapsto \mathbb{R}^{n\times d}$ which satisfy (1.1.2), (1.1.3).

In the problem we consider here, we are not able to control the values of the state process directly. Rather, we are able to choose a time at which we stop accumulating running costs (and pay some terminal cost). The main requirement on our choice of time is that we can not anticipate the future; the decision whether or not to stop should be based on current and past information alone. Mathematically speaking, we only consider *stopping times* with respect to our underlying filtration. Recall that a random variable τ taking values in

$[0,\infty)\cup\{\infty\}$ is an $\{\mathcal{F}_t\}$-stopping time if $\{\tau \le t\} \in \mathcal{F}_t$ for all $t \ge 0$. We denote by $\mathcal{S}$ the set of all $\{\mathcal{F}_t\}$-stopping times.

The cost associated with "starting" in initial state $x \in \mathbb{R}^n$ at "stopping" at time $\tau \in \mathcal{S}$ is

$$J(x,\tau) = \mathbb{E}\Big[\int_0^\tau e^{-\gamma s} g(X(s))ds + e^{-\gamma\tau} h(X(\tau))\, 1_{\{\tau<\infty\}}\Big], \tag{5.1.2}$$

where g is the running cost function and h is the terminal cost function and $\gamma > 0$ is the discount factor. (One common speical case of the above is $\gamma = 0$ and $g \equiv 0$, so that $J(x,\tau) = \mathbb{E}[h(X(\tau))]$.) The value function $V : \mathbb{R}^n \mapsto \mathbb{R}$ of the optimal stopping problem[1] is

$$V(x) = \inf_{\tau\in\mathcal{S}} J(x,\tau).$$

The goal of the optimal stopping problem is to find (or characterize) an optimal stopping time: $\tau^* \in \mathcal{S}$ such that $J(x,\tau^*) = V(x)$. Since our only requirement is that τ^* is an $\{\mathcal{F}_t\}$-stopping time, in principle, the optimal stopping time can be an arbitrarily complicated functional of the sample paths of X. However, there is a special type of stopping time which plays the same role in optimal stopping theory as Markov controls did in previous chapters. Recall that in classical control problems, our problem was formulated in terms of (the large) set of $\{\mathcal{F}_t\}$-adapted controls. However, the dynamic programming approach produced optimal feedback controls (see, for example, (2.6.4)); the value of control to apply at time t depended only on the current state $X(t)$. Thus, our control was expressed as a policy or rule: associated with each state x was an optimal control value $\alpha_x^* \in \mathbb{U}$. We encounter a similar situation in optimal stopping; the only difference is our definition of "control". Here our only control is to stop or not (i.e. continue). Thus, in an optimal stopping problem, for each state x we wish to find an optimal action — stop or continue. Our state space then is split into two regions: a continuation region D in which in the optimal action is to continue, and its complement, the stopping region. An optimal stopping strategy can be expressed, obviously, as "continue until it becomes optimal to stop". Thus starting from state x, an optimal stopping time should be $\tau_x^* = \inf\{t \ge 0 : X_x(t) \notin D\}$, where we have written X_x to emphasize the dependence on the initial state x. The continuation region D plays the same role as a Markov control function u did in the classical problem; each represents a strategy to follow. These strategies can be specified in advance: given D our optimal strategy is to stop at the first time the process exits D. Of course, the actual value of the stopping time τ will depend on the actual outcomes, just as the actual value of the control process U did in

[1]Typically, optimal stopping problems are formulated as maximization problems. We include the inf here in analogy with earlier chapters. Of course, we can formulate our problem as a maximization problem by considering $-g$ and $-h$.

previous chapters. It should be intuitive that $D = \{x \in \mathbb{R}^n : V(x) < h(x)\}$: Since at state x the cost if we stop immediately is $h(x)$, if $V(x) < h(x)$ is not optimal to stop.

5.2. Dynamic programming "equation" for optimal stopping

In this section we provide a motivation for the dynamic programming "equation", which is actually a series of equations and inequalities. In the following development we make many assumptions and neglect technical details in order to provide the main idea behind dynamic programming in optimal stopping problems.

Suppose there exists a continuation region $D \subset \mathbb{R}^n$, an open set, such that $\tau_x^* = \inf\{t \geq 0 : X_x(t) \notin D\}$ is optimal for all $x \in \mathbb{R}^n$, i.e. $V(x) = J(x, \tau_x^*)$. Fix $x \in \mathbb{R}^n$ and let $\tau \in \mathcal{S}$ be arbitrary. Define $\tilde{\tau} = \inf\{t \geq \tau : X(t) \notin D\}$. Following the stopping rule $\tilde{\tau}$, our decision is to continue until time τ, whether or not it is optimal to stop earlier; after time τ we then stop the first time the process exits D. Thus, under this stopping rule we take some arbitrary action for some length of time and then proceed optimally afterward; this should sound the beginning of a dynamic programming principle.

Proceeding as in 2.3 (see in particular since $V(x) \leq J(x, \tilde{\tau})$ we have

$$
\begin{aligned}
&V(x)\\
=&\mathbb{E}\Big[\int_0^{\tilde{\tau}} e^{-\gamma s} g(X(s))ds + e^{-\gamma\tilde{\tau}} h(X(\tilde{\tau}))\, 1_{\{\tilde{\tau}<\infty\}}\Big]\\
=&\mathbb{E}\Big[\int_0^{\tau} e^{-\gamma s} g(X(s))ds + e^{-\gamma\tau}\Big\{\int_\tau^{\tilde{\tau}} e^{-\gamma(s-\tau)} g(X(s))ds + e^{-\gamma(\tilde{\tau}-\tau)} h(X(\tilde{\tau}))\, 1_{\{\tilde{\tau}<\infty\}}\Big\}\Big]\\
=&\mathbb{E}\Big[\int_0^{\tau} e^{-\gamma s} g(X(s))ds + e^{-\gamma\tau} J(X(\tau), \tau^*_{X(\tau^*_{X(\tau)})})\Big]\\
=&\mathbb{E}\Big[\int_0^{\tau} e^{-\gamma s} g(X(s))ds + e^{-\gamma\tau} V(X(\tau)\Big], \qquad (5.2.1)
\end{aligned}
$$

where the last line follows from the optimality of the stopping rule $\tau^*(X(\tau))$: starting from state $X(\tau)$ we are stopping at time $\tilde{\tau}$ the first time (after τ) that the state process exits the continuation region. Of course, if the arbitrary τ satisfies $\tau \leq \tau_x^*$ then $\tilde{\tau} = \tau_x^*$ and so

$$V(x) = J(x, \tau_x^*) = J(x, \tilde{\tau}) = \mathbb{E}\Big[\int_0^{\tau} e^{-\gamma s} g(X(s))ds + e^{-\gamma\tau} V(X(\tau)\Big], \qquad (5.2.2)$$

where the last equality follows similarly to (5.2.1). Combining (5.2.1) and (5.2.2) we have the following dynamic programming principle for optimal stopping:

$$V(x) = \sup_{\tau \in \mathcal{S}} \mathbb{E}\Big[\int_0^{\tau} e^{-\gamma s} g(X(s))ds + e^{-\gamma\tau} V(X(\tau)\Big]. \qquad (5.2.3)$$

Now we wish to obtain a dynamic programming equation. Again, fix $x \in \mathbb{R}^n$ and let $\tau \in \mathcal{S}$. Formally applying Ito's rule to $e^{-\gamma\tau}V(X(\tau))$ we obtain

$$\mathbb{E}[e^{\gamma\tau}V(X(\tau))] = V(x) + \mathbb{E}\int_0^\tau e^{-\gamma t}[\mathcal{L}V(X(t)) - \gamma V(X(t))]dt, \qquad (5.2.4)$$

where as usual $\mathcal{L}V(x) = b(x)\cdot DV(x) + (1/2)\mathrm{tr}(a(x)D^2V(x))$ (when the various derivatives exist). First consider τ with $\tau \le \tau_x^*$. Then (5.2.4) plus (5.2.2) yield

$$0 = \mathbb{E}\int_0^\tau e^{-\gamma t}[\mathcal{L}V(X(t)) - \gamma V(X(t)) + g(X(t))]dt. \qquad (5.2.5)$$

Suppose that $x \in D$. Since D is open and X has continuous sample paths we have $\tau_x^* > 0$. Thus, dividing by τ and taking limits as $\tau \to 0$ in (5.2.5) yields:

$$0 = -\gamma V(x) + \mathcal{L}V(x) + g(x)\ , \quad x \in D. \qquad (5.2.6)$$

If $x \notin D$ then $\tau_x^* = 0$, and so $\tau = 0$ and (5.2.5) does not yield any information. However, since τ_x^* is optimal $V(x) = J(x, \tau_x^*)$, and $J(x, \tau_x^*) = h(x)$ by definition. So we have

$$0 = h(x) - V(x)\ , \quad x \notin D. \qquad (5.2.7)$$

Now with $\tau \in \mathcal{S}$ arbitrary, (5.2.4) plus (5.2.1) yield (proceeding as we did to arrive at (5.2.7))

$$0 \le -\gamma V(x) + \mathcal{L}V(x) + g(x)\ , \quad x \in \mathbb{R}^n. \qquad (5.2.8)$$

Also, since we can always stop immediately and pay the terminal cost (i.e. $V(x) \le J(x, 0) = h(x)$ for any x), so

$$0 \le h(x) - V(x)\ , \quad x \in \mathbb{R}^n. \qquad (5.2.9)$$

Thus we have two inequalities (5.2.8), (5.2.9) which are satisfied for every x. In addition, we see from (5.2.6), (5.2.7) that for every x at least one of these inequalities must be satisfied. We can summarize (5.2.6), (5.2.7), (5.2.8), (5.2.9) as

$$\min\Big\{-\gamma V(x) + \mathcal{L}V(x) + g(x)\ ,\ h(x) - V(x)\Big\}\ = 0\ \ , \quad x \in \mathbb{R}^n. \qquad (5.2.10)$$

The above HJB equation is really a *variational inequality*. Note that while we assumed existence of a continuation region D in order to motivate the HJB equation, the set D does not actually appear in (5.2). Rather, (5.2) gives us a way of distinguishing states in the continuation region from those in the stopping region. For each state we have the choice of whether to stop or continue. These two options are represented by the two terms in the minimum of (5.2). The states in the continuation region are precisely those states for which the equality in is attained by the leftmost term in (5.2); that is

$$D = \{x \in \mathbb{R}^n : h(x) - V(x) > 0\},$$

as we anticipated.

5.3. A verification theorem

In the previous section we saw that finding an optimal stopping time amounts to obtaining a characterization of the continuation region D. The HJB equation (5.2) provides the main tool for finding such a characterization: if we find a solution V of (5.2) then the continuation region is just given by $D = \{x \in \mathbb{R}^n : h(x) - V(x) > 0\}$. On the other hand, we often use some intuition from the problem to first guess the form of the continuation region and then construct a solution to (5.2). (The first approach uses the variational inequalities, while the second strategy is more closely related to a "free boundary problem".) Either way, once we have the continuation region and our candidate value function in hand, optimality of the stopping time $\tau_x^* = \inf\{t \geq 0 : X_x(t) \notin D\}$ can be established by a verification argument using Ito's rule. The main difficulty in verification is that the value function in an optimal stopping problem is almost never a C^2 function. The problem occurs on the boundary of D. Within D evolution is governed by diffusion and is therefore smooth. However, once the boundary ∂D is reached our smooth evolution comes to an abrupt stop. Thus, in general, the value function is at best C^1 across the boundary. For this reason, the classical Ito's rule can not be applied in verification arguments for optimal stopping. Luckily, there are generalizations of Ito's rule that do not require functions to be C^2 everywhere, as long as they are sufficiently smooth. One such generalized Ito's rule for one-dimensional diffusions is contained in the following lemma. For a proof see Lemma IV.45.9 of [**?**].

LEMMA 5.3.1. *Suppose that X is a solution of (5.1.1) with $n = 1$ and $X(0) = x \in \mathbb{R}$. Let $f : \mathbb{R} \mapsto \mathbb{R}$ be of class $C^1(\mathbb{R})$. Suppose that there exists a measurable function $\varphi : \mathbb{R} \mapsto \mathbb{R}$ such that: for all $a > 0$, φ is Lebesgue integrable on $[-a, a]$, and for all $y \in \mathbb{R}$,*

$$f'(y) - f'(0) = \int_0^y \varphi(z)dz.$$

Then for all $t \in [0, \infty)$, almost surely

$$f(X_t) = f(x) + \int_0^t f'(X_s)\sigma(X_s)dW_s + \int_0^t [f'(X_s)b(X_s) + (1/2)\varphi(X_s)\sigma^2(X_s)]ds.$$

Clearly, if f is C^2, then $\varphi = f''$. In general, the function φ is a second derivative in a weak or generalized sense. Note that φ is not necessarily continuous, but the assumptions of the lemma imply that f is "C^2 almost everywhere". The proof of the lemma involves approximating f with C^2 functions, applying the classical Ito's rule to the approximations, and taking limits. One key step in the proof involves showing that the diffusion process does not spend too much time at the points where φ is discontinuous.

Using the above lemma, we can prove the following verification theorem for optimal stopping problems in one-dimension. The setup is as in Section (5.1)

with $n = 1$. We make the following assumptions on g and h:

$$\mathbb{E}\int_0^\infty e^{-\gamma u}|g(X(u))|du < \infty. \tag{5.3.1}$$

$$\text{The collection } \{e^{-\gamma\tau}|h(X(\tau))|\,1_{\{\tau<\infty\}} : \tau \in \mathcal{S}\} \text{ is uniformly integrable.} \tag{5.3.2}$$

These assumptions hold, for example, if g and h satisfy a polynomial growth condition. The above assumptions guarantee that $J(x,\tau)$ is well defined and finite for any $x \in \mathbb{R}$ and $\tau \in \mathcal{S}$. Also, the uniformly integrability assumption implies that there is a constant $C \in (0,\infty)$ such that $\mathbb{E}[e^{-\gamma\tau}|h(X(\tau))|\,1_{\{\tau<\infty\}}] \le C$ for all $\tau \in \mathcal{S}$.

THEOREM 5.3.2. *Suppose there is a function $\phi : \mathbb{R} \mapsto \mathbb{R}$ which satisfies:*

(a) *ϕ of class $C^1(\mathbb{R})$.*

(b) *There exists a measurable function $\varphi : \mathbb{R} \mapsto \mathbb{R}$ such that: for all $a > 0$, φ is Lebesgue integrable on $[-a,a]$, and for all $y \in \mathbb{R}$, $\phi'(y) - \phi'(0) = \int_0^y \varphi(z)dz$.*

(c) *For all $x \in \mathbb{R}$,*

$$\min\Big\{-\gamma V(x)+\phi'(x)b(x)+(1/2)\varphi(x)\sigma^2(x)+g(x)\ ,\ h(x)-V(x)\Big\} = 0, \tag{5.3.3}$$

(d) *For all $x \in \mathbb{R}$,*

$$\lim_{t\to\infty} e^{-\gamma t}\mathbb{E}[\phi(X(t))] = 0. \tag{5.3.4}$$

Then for all $x \in \mathbb{R}$ we have $\phi(x) \le J(x,\tau)$ for any $\tau \in \mathcal{S}$ and thus $\phi(x) \le V(x)$. Furthermore, let $D = \{x \in \mathbb{R} : h(x) - V(x) > 0\}$ and define $\tau^* \equiv \tau_x^* = \inf\{t \ge 0 : X(t) \notin D\}$. Then $\phi(x) = J(x,\tau^*)$. Hence $V(x) = J(x,\tau^*)$, that is, tau* is an optimal stopping time.

PROOF. Fix $x \in \mathbb{R}$. Applying Lemma 5.3.1 to $e^{-\gamma t}\phi(X(t))$ for $t \ge 0$ yields

$$\begin{aligned}&\int_0^t e^{-\gamma u}\phi'(X_u)\sigma(X_u)dW_u\\ =\ & e^{-\gamma t}\phi(X_t) - \phi(x) - \int_0^t e^{-\gamma u}[-\gamma\phi(X_u) + \phi'(X_u)b(X_u) + (1/2)\varphi(X_u)\sigma^2(X_u)]du.\end{aligned}$$

For $r \in (0,\infty)$ define the stopping time $T_r = \inf\{t \ge 0 : |X(t) - x| \ge r\}$. Note that $T_r \to \infty$ a.s. as $r \to \infty$. Let $M(t) = \int_0^{t\wedge T_r} e^{-\gamma u}\phi'(X_u)\sigma(X_u)dW_u$. Since X is bounded on $[0,T_r]$ and ϕ', σ are continuous, the integrand in the stochastic integral is bounded; thus $\{M(t)\}$ is a martingale and in particular $\mathbb{E}[M(t)] = 0$. Let $\tau \in \mathcal{S}$ be arbitrary stopping time and let $N \in (0,\infty)$. Then $\tau \wedge N$ is a bounded stopping time and so by the optional sampling theorem

$\mathbb{E}[M(\tau \wedge N)] = 0$. Let $S = \tau \wedge N \wedge T_r$. Then taking expected values in the above Ito expansion we have

$$\begin{aligned}\phi(x) &= \mathbb{E}[e^{-\gamma S}\phi(X(S))] - \mathbb{E}\int_0^S e^{-\gamma u}[-\gamma\phi(X_u) + \phi'(X_u)b(X_u) + (1/2)\varphi(X_u)\sigma^2(X_u)]du \\ &\leq \mathbb{E}[e^{-\gamma S}\phi(X(S))] + \mathbb{E}\int_0^S e^{-\gamma u}g(X(u))du, \qquad (5.3.5)\end{aligned}$$

where the second line follows since ϕ satisfies (5.3.3). Note that as $r \to \infty$, $N \to \infty$, we have $S \to \tau$ a.s. By (5.3.1) and the dominated convergence theorem

$$\mathbb{E}\int_0^S e^{-\gamma u}g(X(u))du \to \mathbb{E}\int_0^\tau e^{-\gamma u}g(X(u))du \text{ as } r \to \infty, N \to \infty.$$

Also, since ϕ satisfies (5.3.3) we have $\phi \leq h$. Then since $S \leq N < \infty$ and by assumption (5.3.2) we have

$$\mathbb{E}[e^{-\gamma S}\phi(X(S))] \leq \mathbb{E}[e^{-\gamma S}|h(X(S))|] = \mathbb{E}[e^{-\gamma S}|h(X(S))|\,1_{\{S<\infty\}}] \leq C < \infty,$$

where C does not depend on r or N. Using this bound together with the facts that X has continuous sample paths, ϕ is continuous and satisfies (5.3.4), and $S \to \tau$ a.s., an application of the dominated convergence theorem yields

$$\begin{aligned}\mathbb{E}[e^{-\gamma S}\phi(X(S))] &= \mathbb{E}[e^{-\gamma S}\phi(X(S))\,1_{\{\tau<\infty\}}] + \mathbb{E}[e^{-\gamma S}\phi(X(S))\,1_{\{\tau=\infty\}}] \\ &\to \mathbb{E}[e^{-\gamma\tau}\phi(X(\tau))\,1_{\{\tau<\infty\}}] \text{ as } r \to \infty, N \to \infty\end{aligned}$$

Thus taking limits as $N \to \infty$, $r \to \infty$ in (5.3.5) yields

$$\begin{aligned}\phi(x) &\leq \mathbb{E}[e^{-\gamma\tau}\phi(X(\tau))\,1_{\{\tau<\infty\}}] + \mathbb{E}\int_0^\tau e^{-\gamma u}g(X(u))du \\ &\leq \mathbb{E}[e^{-\gamma\tau}h(X(\tau))\,1_{\{\tau<\infty\}}] + \mathbb{E}\int_0^\tau e^{-\gamma u}g(X(u))du \\ &= J(x,\tau),\end{aligned}$$

where the second line follows since $\phi \leq h$. Thus $\phi(x) \leq J(x,\tau)$ for any $\tau \in \mathcal{S}$.

Now consider τ^* as defined in the the theorem. Letting $S^* = \tau^* \wedge N \wedge T_r$, by a calculation similar to that above we have

$$\phi(x) = \mathbb{E}[e^{-\gamma S^*}\phi(X(S^*))] + \mathbb{E}\int_0^{S^*} e^{-\gamma u}g(X(u))du, \qquad (5.3.6)$$

where the equality follows since $X(u) \in D$ for $u < S^*$ and since ϕ solves (5.3.3). Since $S^* \to \tau^*$ as $r \to \infty$, $N \to \infty$, taking limits as in the previous paragraph

yields

$$\begin{aligned}\phi(x) &= \mathbb{E}[e^{-\gamma\tau^*}\phi(X(\tau^*))\,1_{\{\tau^*<\infty\}}] + \mathbb{E}\int_0^{\tau^*} e^{-\gamma u}g(X(u))du \\ &= \mathbb{E}[e^{-\gamma\tau^*}h(X(\tau^*))\,1_{\{\tau^*<\infty\}}] + \mathbb{E}\int_0^{\tau^*} e^{-\gamma u}g(X(u))du \\ &= J(x,\tau^*),\end{aligned}$$

where the second line follows since $X(\tau^*) \notin D$ so $\phi(X(\tau^*)) = h(X(\tau^*))$. ∎

For multi-dimensional versions of generalized Ito's rules, see Theorem 2.10.1 of [**?**]. For an application in a verification lemma for optimal stopping problems, see Theorem 10.4.1 of [**Oks03**] or Theorem 2.2. of [**ØS07**]. In any generalized Ito's rule, there is some assumption that the function is sufficiently smooth; for example, in Lemma 5.3.1 we need f to be at least a C^1 function. In general, the value function for an optimal stopping problem is not C^1, or not smooth enough for the generalized Ito's rule to be applied. When the value function is not smooth, we need to take a viscosity solution approach. We will mention viscosity solutions briefly later. Now we illustrate the verification procedure in the following example.

5.4. Example: Optimal time to sell an asset

Suppose that the price of some asset evolves according to

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t)\ , \quad X(0) = x > 0,$$

with $\mu, \sigma > 0$. Suppose that if we sell the asset we are charged a fix transaction cost $a > 0$, so that our discounted profit from sale at time t is $e^{-\gamma t}(X(t) - a)$. We wish to choose a stopping rule which maximizes our expected discounted profit, or equivalently, to minimize $\mathbb{E}[e^{-\gamma\tau}(a - X(\tau))]$. This is an optimal stopping problem of the form described in the previous sections with $g = 0$ and $h(x) = a - x$. We assume that $\mu < \gamma$; otherwise, if $\mu \geq \gamma$ it is never optimal to sell the asset.

In order to find an optimal stopping strategy, we construct a solution to (5.2) and then use a verification argument. We first guess the form of the continuation region D. It seems reasonable that when the price of the asset is small, we will not sell the asset since the transaction costs will eat up a large portion of our profits. On the other hand, if the price of the asset is large, we do not wish to wait too long to sell due to the discounting of future values. With these considerations in mind we guess that an optimal selling strategy is determined by some threshold; when the price is below the threshold we hold our asset, but once the price is above the threshold we sell it. Thus we guess that $D = \{x \in (0,\infty) : x < x^*\}$ for some $x* \in (a,\infty)$. So characterizing the continuation region amounts to finding the value x^*.

We now construct our candidate for the value function. By (5.2), in the continuation region V should satisfy

$$-\gamma V_0(x) + \mu x V_0'(x) + (1/2)\sigma^2 x^2 V_0''(x) = 0\ , \quad 0 < x < x^*,$$

where we write V_0 to emphasize that this solution will only give our candidate for the value function in the continuation region. We guess that $V_0(x) = cx^\beta$ for some constants $c < 0$ and $\beta > 0$, where

$$\beta = \sigma^{-2}[\sigma^2/2 - \mu + \sqrt{(\mu - \sigma^2/2)^2 + 2\gamma\sigma^2}].$$

(Note: we can always stop at the first time $X(t) = a$ which gives a cost of 0. Thus, $V(x) \leq 0$ which is why we take $c < 0$. We take $\beta > 0$ since $c < 0$ and V should be a decreasing function of x.) Note that $\beta > 1$ since $\mu < \gamma$.

Now for $x \geq x^*$ we have $V(x) = h(x) = a - x$. We now wish to "paste" these two solutions, $V_0(x) = cx^\beta$ for $x < x^*$ and $a - x$ for $x \geq x^*$, together in such a way that the resulting function is the value function. This idea of pasting involves finding the values of the parameters c and x^*. In general, there are many such solutions, each one corresponding to a different continuation region. However, only one such solution will be the value function. It turns out that the value function is the solution obtained by pasting in a "smooth way". In particular, guided by our generalized Ito's lemma, we see that we desire the value function to be C^1 everywhere in order to apply the verification argument. It is often the case that there is only one solution for which the pasting leads to a C^1 solution.

We now seek a solution which is C^1 at x^*. Continuity at x^* implies $c(x^*)^\beta = a - x^*$, which gives $c = (x^*)^{-\beta}(a - x^*)$. Thus V_0 becomes $V_0(x) = (a - x^*)(x/x^*)^\beta$ for $0 < x \leq x^*$. Assuming V is C^1 at x^* implies $V_0'(x^*) = -1$ which yields $x^* = (a\beta)/(\beta - 1)$; note that $x^* \in (a, \infty)$ since $\beta > 1$. Summarizing we have

$$V(x) = \begin{cases} (a - x^*)(x/x^*)^\beta, & 0 < x < x^* = \frac{a\beta}{\beta - 1} \\ a - x, & x \geq x^* = \frac{a\beta}{\beta - 1}. \end{cases}$$

Note that V is C^2 on $(0, \infty) \setminus \{x^*\}$. Thus we can proceed with a verification argument as in Theorem 5.3.2.

CHAPTER 6

Impulse Control Problems

In classical diffusion control problems, control may be applied by changing the values of the drift or diffusion coefficients continuously at all times. However, in many problems applying control at all times is not practical. For example, recall the Merton problem. We have seen that an optimal control involves continuous trading in order to keep a constant proportion of wealth in the stock. But suppose now that the investor must pay a transaction cost each time a trade is made. In this case, continuous trading becomes prohibitively expensive since the investor's wealth is consumed by the transaction costs. Thus we expect that it will be optimal to make no trades unless our wealth falls outside of some thresholds.

In some sense, classical diffusion control problems represent problems in which the controller can impose a great deal of influence on the state process, due to the ability to control continuously. On the other hand, optimal stopping problems are those in which the controller has minimal influence — the only choice is to stop or not. We introduce in this chapter the so-called impulse control problems. With the above considerations, impulse control is somewhere between optimal stopping and classical control on the spectrum of control influence. Typically, in an impulse control problem an optimal control can be described in terms of a non-intervention region, which is similar to the continuation region in optimal stopping problems. The controller takes no action when the state process is within the non-intervention region and acts only when the state exits this region. Upon exit from the non-intervention region the controller also must decide how much control to apply to the state process (in contrast to optimal stopping, in which the only decision upon exiting the continuation region is to stop). A key difference from classical control is that the controller is allowed to instantaneously move the state process by a certain amount; thus the resulting controlled process may have sample paths with jumps.

6.1. Impulse control: general formulation

In this section we consider an impulse control problem in which the underlying uncontrolled state process is a time-homogeneous diffusion and the cost is of the infinite horizon discounted form. Such a problem will introduce the main

issues related to impulse control. Similar considerations apply in the various other settings (time dependent parameters, finite or indefinite time horizons, etc).

We assume that, in the absence of control, the state process evolves according as an uncontrolled diffusion:

$$dX(t) = b(X(t))dt + \sigma(X(t))dW(t) \ , \quad X(0-) = x \in \mathbb{R}^n.$$

We write $X(0-)$ to emphasize that x is the state of the system at time 0 before any control is applied. This notation should become clearer as we proceed. An impulse control will consist of a sequence of intervention times $0 \leq \tau_1 < \tau_2 < \ldots$ and impulse values $\zeta_1, \zeta_2, \ldots$ taking values in some set $\mathcal{Z}$. Given such a sequence, the controlled state process evolves as follows. For $0 \leq t < \tau_1$, no control has been applied and thus

$$X(t) = x + \int_0^t b(X(s))ds + \int_0^t \sigma(X(s))dW(s) \ , \quad 0 \leq t < \tau_1.$$

Let $X(\tau_1-)$ denote the state of the system immediately before time τ_1; that is, $X(\tau_1-)$ represents the state of the system before control is applied. At time τ_1, we apply an impulse ζ_1 in order to instantaneously move the state of the system from $X(\tau_1-)$ to $X(\tau_1)$, where $X(\tau_1) = \Gamma(X(\tau_1-), \zeta_1)$ and $\Gamma : \mathbb{R}^n \times \mathcal{Z} \mapsto \mathbb{R}^n$ is the control function (a typical control function is $\Gamma(x, z) = x - z$). After time τ_1 no control is applied until time τ_2 and thus for $\tau_1 < t < \tau_2$ we have

$$X(t) = X(\tau_1) + \int_{\tau_1}^t b(X(s))ds + \int_{\tau_1}^t \sigma(X(s))dW(s) \ , \quad \tau_1 < t < \tau_2.$$

At time τ_2, we apply an impulse ζ_2 in order to instantaneously move the state of the system from $X(\tau_2-)$ to $X(\tau_2) = \Gamma(X(\tau_2-), \zeta_2)$. The construction of the state process for larger values of t continues in a similar manner.

Before putting an impulse control problem on a more rigorous footing, we note some differences between an impulse controlled state process and a classically controlled diffusion process, e.g. (2.6.1). Note that in classical control we effectively control the values of b and σ but we are not able to control directly the actual values of the process. The opposite is true in impulse control: b and σ are fixed but by choosing an appropriate impulse value we can directly influence the state process itself. In addition, in classical control a control process U typically imposes influence at all points in time t, while in impulse control we only control at the times $\tau_1, \tau_2, \ldots$. Finally, under suitable conditions a controlled diffusion process has continuous sample paths while an impulse controlled process may (and often does) have jumps.

DEFINITION 6.1.1. *An* impulse control $I \equiv \{(\tau_j, \zeta_j), j = 1, \ldots, M\}$, *where* M *is a random variable taking values in* $[0, 1, 2, \ldots) \cup \{\infty\}$, *consists of:*

(a) *A sequence of* $\{\mathcal{F}_t\}$*-stopping times* $\tau_1, \tau_2, \ldots$ *with* $0 \leq \tau_1 < \tau_2 < \ldots$.

(b) *A sequence of impulse values* $\zeta_1, \zeta_2, \ldots$ *such that for each* $j = 1, 2, \ldots$, ζ_j *takes values in* $\mathcal{Z} \subseteq \mathbb{R}^p$ *and* ζ_j *is measurable with respect to* $\mathcal{F}_{\tau_j}$.

Impulse control I *is* admissible *if*

(a) *The corresponding state process* $X \equiv X^I$ *exists and is unique.*
(b) *With probability one, either* $M(\omega) < \infty$ *or if* $M(\omega) = \infty$ *then* $\lim_{j\to\infty} \tau_j(\omega) = \infty$.

In this chapter we denote by $\mathcal{A}$ *the set of all admissible impulse controls.*

We allow M to be a random variable since the number of interventions may depend on the path of the process. If $M(\omega) = 0$ for some ω then the corresponding path of the state process is uncontrolled. In the problem considered here (infinite horizon cost with unbounded state space) an impulse control will typically be an infinite sequence (i.e. $M = \infty$). However, with an exit time cost criteria there may be only finitely many impulses before the exit time. (We will also consider problems where we are restricted to controlling with at most finitely many impulses in an iterative algorithm later in the chapter.) When $M = \infty$ the requirement $\lim_{j\to\infty} \tau_j = \infty$ (we say that the intervention times do not explode) ensures that we are truly only controlling at discrete points in time. For if $\lim_{j\to\infty} \tau_j < \infty$ with positive probability, for such outcomes we control infinitely many times in a finite time interval, which is akin to continuous control — a behavior that we wish to rule out in (pure) impulse control problems.

Let X denote the state process corresponding to initial state $x \in \mathbb{R}^n$ and impulse control $I \in \mathcal{A}$. (Existence and uniqueness of the state process is guaranteed under suitable conditions on b, σ, and Γ.) The corresponding cost function is defined as

$$J(x, I) = \mathbb{E}\left[\int_0^\infty e^{-\gamma s} g(X(s))ds + \sum_{j=1}^{\infty} e^{-\gamma\tau_j} \ell(X(\tau_j-), \zeta_j)\, 1_{\{\tau_j<\infty\}}\right],$$

where $g : \mathbb{R}^n \mapsto \mathbb{R}$ is the running cost and $\ell : \mathbb{R}^n \times \mathcal{Z} \mapsto \mathbb{R}$ is the intervention cost (or cost of control). The value function for the impulse control problem is

$$V(x) = \inf_{I\in\mathcal{A}} J(x, I).$$

6.2. Dynamic programming for impulse control

As in any stochastic control problem, the goal of an impulse control problem is to find an optimal control (or some characterization of one) and the corresponding value function. And as usual, the main tool we will use is dynamic programming. Now, in each state we have two decisions: whether to intervene or not, and if so what value of control impulse to apply.

A key object in impulse control is the *intervention operator* $\mathcal{K}$. The operator $\mathcal{K}$ maps a function $\phi : \mathbb{R}^n \mapsto \mathbb{R}$ to a function denoted $\mathcal{K}\phi : \mathbb{R}^n \mapsto \mathbb{R}$ given by

$$\mathcal{K}\phi(x) = \inf_{z \in \mathcal{Z}} \left[\phi(\Gamma(x,z)) + \ell(x,z)\right], \quad x \in \mathbb{R}^n. \tag{6.2.1}$$

We will be most interested in $\mathcal{K}$ applied to the value function V. In this case, $\mathcal{K}V(x)$ represents the value of the strategy that consists of taking the best immediate action in state x and behaving optimally afterward *given that you take some immediate action in state* x (i.e. under this strategy you must intervene at the present time). To see this, suppose we are in state x and we decide to take action z. Then we pay an immediate cost of $\ell(x,z)$ and instantaneously move the state process to $\Gamma(x,z)$. Starting from that state we act optimally, so the cost associated with starting from $\Gamma(x,z)$ is given by $V(\Gamma(x,z))$. Thus the total cost of this strategy is $V(\Gamma(x,z)) + \ell(x,z)$. Finally, we minimize over all possible actions z to obtain $KV(x)$. Clearly, since in any state it might not be optimal to intervene, we have

$$V(x) \leq KV(x), \quad x \in \mathbb{R}^n. \tag{6.2.2}$$

The intervention operator will play a key role in formulating a dynamic programming principle and the associated HJB equation. We proceed formally. First let T be an arbitrary $\{\mathcal{F}_t\}$-stopping time with $T < \infty$ a.s. and consider the following strategy: Before time T there is no intervention; at time T we intervene and apply the optimal impulse value at this time; we proceed optimally after time T. With the same considerations in the previous paragraph we see that starting from state $X(T-)$ the associated cost is given by $\mathcal{K}V(X(T-))$. Also, since no control is applied over $[0, T-]$ the cost of our strategy on this interval is just $\mathbb{E}\int_0^T e^{-\gamma s} g(X(s))ds$. Thus, since T is arbitrary, we have

$$V(x) \leq \mathbb{E}\left[\int_0^T e^{-\gamma s} g(X(s))ds + e^{-\gamma T}\mathcal{K}V(X(T-))\right],$$

where the RHS represents the total cost of our proposed strategy.

Now suppose we follow the same strategy above with T replaced by T^*, the first time at which it is optimal to intervene (and assume $T^* < \infty$ a.s.). Then under this strategy we are always acting optimally. Thus, letting X^* denote the corresponding controlled process, we have $\mathcal{K}V(X^*(T^*-)) = V(X^*(T^*-))$ and

$$V(x) = \mathbb{E}\left[\int_0^{T^*} e^{-\gamma s} g(X^*(s))ds + e^{-\gamma T^*} V(X^*(T^*-))\right].$$

Combining the above two displays we arrive at the following dynamic programming principle:

$$V(x) = \inf_{T \in \mathcal{S}} \mathbb{E}\left[\int_0^T e^{-\gamma s} g(X(s))ds + e^{-\gamma T}\mathcal{K}V(X(T-))\, 1_{\{T<\infty\}}\right], \tag{6.2.3}$$

where $\mathcal{S}$ is the set of all $\{\mathcal{F}_t\}$-stopping times.

We note in passing that (6.2.3) resembles the value function of an optimal stopping problem (see (5.1.2)) where the terminal cost h has been replaced by $\mathcal{K}V$. Equation (6.2.3) does not correspond exactly to an optimal stopping problem since the function $\mathcal{K}V$ is not given (it is part of the solution we are trying to obtain). Nevertheless this relationship suggests connections between optimal stopping and impulse control which we will explore later in this chapter.

We now wish to give motivation for the HJB equation to be introduced below. The development is similar to that in Section 5.2. With an arbitrary $T \in \mathcal{S}$, the DPP (6.2.3) and the usual Ito's rule and limit calculations yield

$$-\gamma V(x) + \mathcal{L}V(x) + g(x) \geq 0 \ , \quad x \in \mathbb{R}^n. \tag{6.2.4}$$

Now let T^* be as introduced above. If $T^* = 0$ then it is optimal to act immediately and so

$$V(x) = \mathcal{K}V(x) \ , \quad \text{for any } x \text{ such that } T^* = 0. \tag{6.2.5}$$

If $T^* > 0$ then, since equality holds in (6.2.3) for T^*, the same calculations used to produce (6.2.4) yield

$$-\gamma V(x) + \mathcal{L}V(x) + g(x) = 0 \ , \quad \text{for any } x \text{ such that } T^* > 0. \tag{6.2.6}$$

(Clearly, if it is not optimal to intervene immediately then the state process evolves as an uncontrolled diffusion and hence V should satisfy (6.2.6).) Hence we have two inequalities (6.2.2), (6.2.4) which must hold for all $x \in \mathbb{R}^n$, and furthermore, (6.2.5) and (6.2.6) imply that for each x one of the inequalities must be an equality. Thus summarizing (6.2.2), (6.2.4), (6.2.5), (6.2.6) we have the following HJB "equation" for our impulse control problem

$$\min\{-\gamma V(x) + \mathcal{L}V(x) + g(x), \mathcal{K}V(x) - V(x)\} = 0 \ , \quad x \in \mathbb{R}^n. \tag{6.2.7}$$

Note that this describes a "quasi-variational inequality"; the term "quasi" is used because the "terminal cost" $\mathcal{K}V$ depends on the solution V in a nontrivial way.

Given a solution V of the HJB equation (6.2.7) and its intervention operator $\mathcal{K}V$ given by (6.2.1), we are led to the following Markovian characterization of an optimal control. As remarked previously for each state x we must find two optimal decisions: whether to intervene or not, and if so the optimal impulse value. Intuitively, the states in which it is not optimal to control are those in the *non-intervention region* $D = \{x : \mathcal{K}V(x) > V(x)\}$. For $x \in D^c$ it is optimal to intervene, and the definition of $\mathcal{K}V(x)$ suggest that the optimal impulse value, denoted $z^*(x)$, is given by

$$z^*(x) \in \arg\min_{z \in \mathcal{Z}} \left[V(\Gamma(x,z)) + \ell(x,z)\right] \ , \quad x \in D^c.$$

Our candidate for an optimal impulse control $I^* = \{(\tau_j^*, \zeta_j^*)\}$, with state process X^*, can then be described as follows. First suppose that $X^*(0-) = x \in D$. Let $\tau_0^* = 0$ and define (τ_j^*, ζ_j^*) for $j = 1, 2, \ldots$ inductively by

$$\begin{aligned}\tau_j^* &= \inf\{t > \tau_{j-1}^* : X^*(t) \notin D\},\\ \zeta_j^* &\in \arg\min_{z\in\mathcal{Z}} \left[V(\Gamma(X^*(\tau_j^*-), z)) + \ell(X^*(\tau_j^*-), z)\right].\end{aligned}$$

If instead $X^*(0-) = x \in D^c$ then $\tau_1^* = 0$ and at time 0 we make an instantaneous jump, determined by the impulse ζ_1^*, to a point $X^*(0) \in D$ and proceed in constructing I^* as above. Note that if I^* is constructed in this manner, we can always assume that after applying control the state process lies within D. For suppose that starting from $x \in D^c$ we apply control and move instantaneously to another state in D^c, and from that state we move to another state in D^c, and so on. Since no time elapses between the times at which control is applied the state must eventually move to a state outside of D^c after finitely many applications of control; otherwise, we would have $\lim_{j\to\infty} \tau_j^* = 0$. Again, since no time elapses between control and we must move outside of D^c in finitely many moves, we can assume that the state moves instantaneously outside of D^c with just one application of control (the accumulation of the finitely many "smaller" moves).

6.3. Verification

Given a sufficiently smooth solution, denoted ϕ, of (6.2.7) and I^* constructed as above (with V replaced by ϕ) we can show that ϕ is the value function of the control problem and I^* is an optimal control ($\phi(x) = J(x, I^*) = V(x)$) through the usual verification procedure. We provide a sketch of the verification argument below. We do not quote a general theorem here but instead refer the reader to [**Oks03**, Theorem 6.2]. As in optimal stopping problems, the value function will almost never be C^2 everywhere; the problems occur on the boundary of the non-intervention region D. Consequently, the verification procedure relies on some generalized Ito's rule, for example Lemma 5.3.1 in the one-dimensional case. (Of course, if the candidate function ϕ is not sufficiently smooth then we must take a viscosity solution approach.)

Suppose that ϕ is a solution of (6.2.7) and I^* constructed as above, with state process X^*, satisfies $I^* \in \mathcal{A}$. Suppose further that ϕ is smooth enough to apply a generalized Ito's rule to the function $e^{-\gamma t}\phi(x)$ with the semimartingale X^*

over the time interval $[\tau_j^*, \tau_{j+1}^*-]$, which yields[1]

$$\mathbb{E}[e^{-\gamma\tau_j^*}\phi(X^*(\tau_j))]-\mathbb{E}[e^{-\gamma\tau_{j+1}^*}\phi(X^*(\tau_{j+1}^*-))] = -\mathbb{E}\int_{\tau_j^*}^{\tau_{j+1}^*} e^{-\gamma s}(-\gamma\phi(X_s^*)+\mathcal{L}\phi(X^*(s)))ds.$$

Summing over $j = 0, 1, \ldots, m$ and then letting $m \to \infty$ we get

$$\phi(x)+\sum_{j=1}^{\infty}\mathbb{E}\left[e^{-\gamma\tau_j^*}(\phi(X^*(\tau_j^*)) - \phi(X^*(\tau_j^*-))\right] = -\mathbb{E}\int_0^{\infty} e^{-\gamma s}(-\gamma\phi(X_s^*)+\mathcal{L}\phi(X^*(s)))ds,$$

and after some algebra we have

$$\begin{aligned}\phi(x) = \mathbb{E}\int_0^{\infty} e^{-\gamma s}g(X^*(s))ds + \mathbb{E}\sum_{j=1}^{\infty} e^{-\gamma\tau_j^*}\ell(X^*(\tau^*-),\zeta_j^*)\\ - \mathbb{E}\int_0^{\infty} e^{-\gamma s}(-\gamma\phi(X_s^*) + \mathcal{L}\phi(X^*(s)) + g(X^*(s)))ds\\ - \sum_{j=1}^{\infty}\mathbb{E}\left[e^{-\gamma\tau_j^*}(\phi(X^*(\tau_j^*)) - \phi(X^*(\tau_j^*-) + \ell(X^*(\tau_j^*-),\zeta_j^*)\right]. \quad (6.3.1)\end{aligned}$$

It follows from the construction of ϕ and I^* that $X^*(\tau_j^*-) \notin D$, that $\mathcal{K}\phi(X^*(\tau_j^*-)) = \phi(X^*(\tau_j^*-))$ and that $\mathcal{K}\phi(X^*(\tau_j^*-)) = \phi(\Gamma(X^*(\tau_j^*-),\zeta_j^*)) + \ell(X^*(\tau_j^*-),\zeta_j^*)$. Therefore

$$\phi(X^*(\tau_j^*)) - \phi(X^*(\tau_j^*-) + \ell(X^*(\tau_j^*-),\zeta_j^*) = 0,$$

and the fourth term on the RHS of (6.3.1) is 0. It also follows from construction of ϕ and I^* that if $X^*(s) \in D$ then $-\gamma\phi(X_s^*) + \mathcal{L}\phi(X^*(s)) + g(X^*(s)) = 0$. Furthermore, the state process X^* is always moved instantaneously back to D whenever it exits the region D. Thus X^* "spends 0 time" outside of D in the sense that $\text{Leb}\times\mathbb{P}(\{(s,\omega) : X^*(s,\omega) \notin D\}) = 0$ and it follows that

$$\mathbb{E}\int_0^{\infty} e^{-\gamma s}(-\gamma\phi(X_s^*) + \mathcal{L}\phi(X^*(s)) + g(X^*(s)))\,1_{\{X^*(s)\notin x^*\}}ds = 0.$$

Therefore, the third term on the RHS of (6.3.1) is 0. This shows that $\phi(x) = J(x, I^*)$.

[1]Technically we should replace the ϕ'' term in $\mathcal{L}\phi$ by some approximation or generalized derivative as in Theorem 5.3.2 and its proof, but we write $\mathcal{L}\phi$ for notational simplicity.

If $I = \{(\tau_j, \zeta_j)\} \in \mathcal{A}$ is arbitrary, with state process X, the same calculations which produced (6.3.1) yield

$$\begin{aligned}\phi(x) = \mathbb{E}\int_0^\infty e^{-\gamma s} g(X(s))ds + \mathbb{E}\sum_{j=1}^\infty e^{-\gamma\tau_j}\ell(X(\tau-), \zeta_j) \\ - \mathbb{E}\int_0^\infty e^{-\gamma s}(-\gamma\phi(X(s)) + \mathcal{L}\phi(X(s)) + g(X(s)))ds \\ - \sum_{j=1}^\infty \mathbb{E}\left[e^{-\gamma\tau_j}(\phi(X(\tau_j)) - \phi(X(\tau_j-) + \ell(X(\tau_j-), \zeta_j)\right]. \qquad (6.3.2)\end{aligned}$$

Since ϕ solves (6.2.7) we have $-\gamma\phi(X_s^*) + \mathcal{L}\phi(X^*(s)) + g(X^*(s)) \geq 0$ for all $s \geq 0$. Also note that if ϕ solves (6.2.7) then $\phi(x) \leq \mathcal{K}\phi(x)$ for any x. In addition, it follows from the definition of $\mathcal{K}\phi$ that $\mathcal{K}\phi(x) \leq \phi(\Gamma(x, z)) + \ell(x, z)$ for any x and $z \in \mathcal{Z}$. Then since $X(\tau_j) = \Gamma(X(\tau_j-), \zeta_j)$ we have

$$\phi(X(\tau_j-)) \leq \mathcal{K}\phi(X(\tau_j-)) \leq \phi(\Gamma(X(\tau_j-), \zeta_j)) + \ell(X(\tau_j-), \zeta_j) = \phi(X(\tau_j)) + \ell(X(\tau_j-), \zeta_j).$$

Therefore $\phi(X(\tau_j)) - \phi(X(\tau_j-) + \ell(X(\tau_j-), \zeta_j) \geq 0$. It then follows from (6.3.2) that $\phi(x) \leq J(x, I)$, and since $I \in \mathcal{A}$ is arbitrary we have $\phi(x) \leq V(x)$. Finally, since $J(x, I^*) \geq V(x)$ we have $\phi(x) = V(x) = J(x, I^*)$.

6.4. Example impulse control problem

This example is adapted from [**Oks03**, Example 6.4] and it is related to a problem of optimal dividend payments with transaction costs. Suppose that with no interventions a cash flow process is described by a Brownian motion with drift $x + \mu t + \sigma W(t)$ for $x, \mu, \sigma > 0$. At time $\tau_j \geq 0$ we can withdraw $\zeta_j \geq 0$ but we are required to pay the transaction cost $c + \lambda\zeta_j$, where $c > 0$ is the fixed transaction cost and $\lambda \geq 0$ is the proportional transaction cost. Thus if in state x we withdraw amount z the state process after the transaction (intervention) is given by $\Gamma(x, z) = x - c - (1 + \lambda)z$. Given an impulse control $\{(\tau_j, \zeta_j)\}$ the controlled state process is

$$X(t) = x + \mu t + \sigma W(t) - \sum_{j=1}^\infty (c + (1 + \lambda)\zeta_j)\, 1_{\{\tau_j \leq t\}}.$$

The goal in this problem is to maximize the expected discounted value of our withdrawals, or equivalently to minimize

$$J(x, I) = \mathbb{E}\left[\sum_{j=1}^\infty e^{-\gamma\tau_j}(-\zeta_j)\, 1_{\{\tau_j \leq T\}}\right],$$

where $\gamma > 0$ and $T = \inf\{t \geq 0 : X(t) \leq 0\}$. Impulse control I is admissible ($I \in \mathcal{A}$)[2] if it satisfies the requirements of Definition 6.1.1 and in addition $X(t) \geq 0$ all $t \geq 0$. The value function is $V(x) = \inf_{I \in \mathcal{A}} J(x, I)$. Note that $V(x) \leq -(x-c)^+/(1+\lambda)$. This can be seen by considering the strategy: if the initial state $x \leq c$ never apply control, if $x > c$ withdraw at time 0 the maximum possible amount which, taking into account transaction costs, is $(x-c)/(1+\lambda)$, in order to reduce wealth to 0.

6.4.1. Construction of candidate value function and optimal control. We now proceed to find an optimal control. We first guess that the non-intervention region D has the form $\{x : 0 < x < x^*\}$ for some $x^* > c$. Within the non-intervention region we seek a solution of

$$-\gamma\phi(x) + \mu\phi'(x) + (\sigma^2/2)\phi''(x) = 0\ , \quad 0 < x < x^*.$$

We try a solution of the form $\phi_0(x) = A(e^{r_+ x} - e^{r_- x})$ where $A < 0$ is to be determined and $r_+ > 0 > r_-$ satisfy $|r_-| > r_+$ and

$$r_+ = (-\mu + \sqrt{\mu^2 + 2\sigma^2})/(\sigma^2) \quad r_- = (-\mu - \sqrt{\mu^2 + 2\sigma^2})/(\sigma^2).$$

We have written ϕ_0 to distinguish the solution from ϕ the solution to the HJB equation we are trying construct.

For $x \notin D$ the intervention operator is key. Taking into account the state constraints, the intervention operator applied to a function ϕ yields

$$\mathcal{K}\phi(x) = \inf\left\{\phi(x - c - (1+\lambda)z) - z : 0 < z \leq \frac{x-c}{1+\lambda}\right\}.$$

For $x \geq x^*$ we wish to find $\hat{z} \equiv \hat{z}(x)$ which minimizes

$$\theta(z) = \phi_0(x - c - (1+\lambda)z) - z,$$

subject to $0 < \hat{z} \leq (x-c)/(1+\lambda)$. Note that we have ϕ_0 here since any intervention will move the state into the non-intervention region. Setting $\theta'(z) = 0$ we have $\phi_0'(x - c - (1+\lambda)z) = -1/(1+\lambda)$. Note that $\phi'(y) < 0$ for all y (since $A < 0$) and $\phi''(y) > 0$ if and only if

$$y < \frac{2(\log|r_-| - \log r_+)}{r_+ - r_-} \doteq \tilde{y}.$$

So $\phi_0'(y) = -1/(1+\lambda)$ has two solutions y_1, y_2 with $0 < y_1 < \tilde{y} < y_2$ (provided that $\phi_0'(\tilde{y}) < 1/(1+\lambda) < \phi_0'(0)$). Define $\hat{x} = y_1$. (We will see later that $x^* = y_2$.) Setting $\hat{x} = x - c - (1+\lambda)\hat{z}$ yields our candidate for $\hat{z} = (x - \hat{x} - c)/(1+\lambda)$. For $x \geq x^*$ we want $\mathcal{K}\phi(x) = \phi(x)$ which implies $\mathcal{K}\phi_0(x) = \phi(x)$ since control moves the state to the non-intervention region. Thus for $x \geq x^*$ we want

[2]Technically, given the state constraint $\mathcal{A}$ may depend on the initial state x, but we omit this dependence to simplify the notation. We will see that the constructed optimal control satisfies the admissibility conditions for any initial state $x > 0$.

$\phi(x) = \phi_0(\hat{x}) - (x - \hat{x} - c)/(1 + \lambda)$ where we have used the choice of $\hat{z}$ above. Our guess for the value function then becomes

$$\phi(x) = \begin{cases} A(e^{r_+ x} - e^{r_- x}), & 0 < x < x^*, \\ A(e^{r_+ \hat{x}} - e^{r_- \hat{x}}) - \frac{x - \hat{x} - c}{1+\lambda}, & x \geq x^*. \end{cases}$$

The constants A and x^* still need to be determined. To do so, we assume a principle of smooth fit and choose A and x^* so that ϕ is C^1. Continuity at x^* implies

$$A = \frac{x^* - \hat{x} - c}{1 + \lambda} \left[e^{r_+ \hat{x}} - e^{r_- \hat{x}} - e^{r_+ x^*} + e^{r_- x^*} \right].$$

Assuming ϕ is C^1 at x^* yields $\phi_0'(x^*) = -1/(1 + \lambda)$. We have seen that there are two candidates for x^*, which we denoted y_1 and y_2 above. However, given the expression for A above we see that $x^* = y_1 = \hat{x}$ is not a possibility. Thus we take $x^* = y_2$. Then we have $x^* > \hat{x}$, and furthermore the state constraint $\hat{z}(x^*) \geq 0$ implies $x^* - \hat{x} - c \geq 0$. Thus $A < 0$. Note that we have not expressed closed form solutions for A, x^*, $\hat{x}$, but we have three equations in three unknowns which in principle can be solved. Thus we have completely specified ϕ, our candidate for the value function of the control problem.

Our candidate optimal control $I^* = \{(\tau_j^*, \zeta_j^*)\}$ with state process X^* is defined as follows. The definition is somewhat complicated by the fact that control will only be applied until time $T^* = \inf\{t \geq 0 : X^*(t) = 0\}$. We first assume that the process is not stopped on exit from $(0, \infty)$ and construct an impulse control which behaves as described above (see also the following paragraph), with the same x^* and $\hat{x}$; we denote the resulting control by $\tilde{I}^* = \{(\tilde{\tau}_j^*, \tilde{\zeta}_j^*) : j \leq \tilde{M}\}$ and state process by $\tilde{X}^*$. Let $\tilde{\tau}_0^* = 0$; we define $\tilde{I}^*$ inductively. Suppose that we have defined $\tilde{\tau}_1^*, \ldots, \tilde{\tau}_j^*$, $\tilde{\zeta}_1^*, \ldots, \tilde{\zeta}_j^*$, and the corresponding $\tilde{X}^*$ on $[0, \tilde{\tau}_j^*]$, with the uncontrolled evolution starting again at time $\tilde{\tau}_j^*$. Define

$$\tilde{\tau}_{j+1}^* = \inf\{t > \tilde{\tau}_j^* : \tilde{X}^*(t) \geq x^*\}$$
$$\zeta_{j+1}^* = \frac{\tilde{X}^*(\tilde{\tau}_{j+1}^* -) - \hat{x} - c}{1 + \lambda}.$$

Since our uncontrolled state process is a Brownian motion with positive drift μ it follows that $\tilde{M} = \infty$ a.s. We now revise our constructed control to reflect the fact that we must stop the process when it hits 0. Let $T^* = \inf\{t \geq 0 : \tilde{X}^*(t) = 0\}$. Define the random variable $M = \sup\{j = 0, 1, 2, \ldots : \tilde{\tau}_j^* < T^*\}$. Define $\tau_j^* = \tilde{\tau}_j^*$ and $\zeta_j^* = \tilde{\zeta}_j^*$ for $j \leq M$ and let $I^* = \{(\tau_j^*, \zeta_j^*) : j \leq M\}$. Note that by construction $T^* = \inf\{t \geq 0 : X^*(t) = 0\}$, that $\tau_j^* < T^*$ for all $j \leq M$ and that $X^*(\tau_j^*) = \hat{x}$. Let X^* be the state process corresponding to I^* on $[0, T^*]$ and let $X^*(T^* + u) = 0$ for all $u \geq 0$. Thus we have constructed our candidate optimal control I^* and corresponding state process X^*.

We now wish to verify that the constructed I^* is an optimal control. We first remark that the process X^* behaves in the following way. Starting from initial state $X^*(0-) = x < x^*$ the process evolves uncontrolled, as a Brownian motion with drift (with, in particular, continuous sample paths), until the first time it either hits x^* or 0. If it reaches 0 first then the process is stopped and no further control is applied. Otherwise, upon hitting x^* we apply an impulse equal to $(x^* - \hat{x} - c)/(1+\lambda)$ and instantaneously jump the state to $\hat{x}$. Starting from $\hat{x}$ the process again evolves uncontrolled until it hits either x^* or 0, and so on. If $X^*(0-) = x \geq x^*$ the only difference is that there is an initial impulse applied at time 0 to move the state to $X^*(0) = \hat{x}$. The process then evolves as above. Note that the sample paths of X^* will be continuous on $[\tau_j^*, \tau_{j+1}^*)$ with jumps at the times τ_j^*.

We now wish to verify that $\phi(x) = J(x, I^*) = V(x)$. We first check admissibility of I^*. The main thing to show is that the intervention times do not explode, which in this case means showing that $\lim_{j\to\infty} \tau_j^* = T^*$ a.s. To sketch an argument that would also work in the unbounded state space case, we first assume that the process is not stopped on exit from $(0, \infty)$. That is, we considered the impulse control $\tilde{I}^*$ with state process $\tilde{X}^*$ as constructed above. Since the sample path of $\tilde{X}^*$ is continuous on $(\tilde{\tau}_j^*, \tilde{\tau}_{j+1}^*)$ a.s. and since $\hat{x} < x^*$ we have that $\tilde{\tau}_j^* < \tilde{\tau}_{j+1}^*$ a.s. It follows from the independent increments property of Brownian motion that $\tilde{X}^*$ "restarts" at every time $\tilde{\tau}_j^*$ from initial value $\hat{x}$. Thus $\{\tilde{\tau}_{j+1}^* - \tilde{\tau}_j^*, j = 1, 2, \ldots\}$ is an i.i.d. collection of random variables. Together, these facts imply that there is some $\varepsilon > 0$ such that $\mathbb{P}[\tilde{\tau}_{j+1}^* - \tilde{\tau}_j^* > \varepsilon] > 0$; denote this probability, which does not depend on j, by $\eta \equiv \eta(\varepsilon) > 0$. Thus we have

$$\sum_{j=1}^{\infty} \mathbb{P}[\tilde{\tau}_{j+1}^* - \tilde{\tau}_j^* > \varepsilon] = \sum_{j=1}^{\infty} \eta = \infty.$$

The second Borel-Cantelli lemma then implies that $\mathbb{P}[\tilde{\tau}_{j+1}^* - \tilde{\tau}_j^* > \varepsilon \text{ for infinitely many } j] = 1$ and it follows that $\lim_{j\to\infty} \tilde{\tau}_j^* = \infty$ a.s. Returning to the exit time problem, since $T^* < \infty$ a.s. and for the "un-stopped" case we have $\lim_{j\to\infty} \tilde{\tau}_j^* = \infty$ a.s. we must have $\lim_{j\to\infty} \tau_j^* = T^*$ a.s. It follows that $I^* \in \mathcal{A}$.

Note that ϕ is C^1 on $(0, \infty)$ and C^2 except at the point x^*. Thus we can apply the generalized Ito's rule, Lemma 5.3.1, to the function $e^{-\gamma t}\phi(x)$ with the semimartingale X^*. Verification that $\phi(x) = J(x, I^*) = V(x)$ then proceeds as in the argument sketched at the end of Section 6.3.

6.4.2. Note on fixed versus proportional transaction costs. Impulse control problems often arise in situations when there are certain transaction costs which impact the state of the system whenever control is applied. The example considered here incorporates both fixed and proportional transaction costs. We remark here that is it actually the *fixed* transaction costs which

make the above example an impulse control problem. To illustrate the point, we write $x^* \equiv x^*(\lambda, c)$ and $\hat{x} \equiv \hat{x}(\lambda, c)$, since clearly these two quantities depend on λ and c. We claim that for each fixed $\lambda > 0$ we have

$$\lim_{c \to 0} x^*(\lambda, c) = \lim_{c \to 0} \hat{x}(\lambda, c).$$

This suggests that if there are only proportional transaction costs the optimal strategy is the following: Make no transactions whenever the state is less than x^*; whenever the state hits x^* make the smallest possible transaction to push the state to be less than x^*. Thus, aside from a possible initial jump at time 0, an optimally controlled process would have continuous sample paths. Intuitively this strategy should make sense. Since there are transaction costs, we still wish to control only when the state process exits some no transaction region. However, now small transactions are preferable since the cost is proportional to amount of transaction. Problems with only proportional transaction costs can often be formulated as singular control problems, which we will introduce in the next chapter.

On the other hand, let $c > 0$ be fixed. We claim that

$$\lim_{\lambda \to 0} x^*(\lambda, c) > \lim_{\lambda \to 0} \hat{x}(\lambda, c).$$

Thus, even if there are no proportional transaction costs, in the presence of a non-zero fixed cost an optimal strategy still has the same characterization we have already seen. When the state process exits the no transaction region, there is a jump of positive size which pushes it away from the no transaction boundary. Again, this should make sense: any time we make a transaction we pay the fixed cost regardless of the size of the transaction. So we should make sure the transactions are worth the cost. That is, if we were to make a very small transaction, we pay the fixed cost but remain very close to the boundary of the no transaction region. Then we can expect to hit the no-transaction region again very soon at which time we would have to pay another fixed cost. For this reason, any time we pay the fixed cost, we want to make sure that the resulting transaction moves us "far away" from the no-transaction boundary.

6.5. Connections between impulse control and optimal stopping

Recall the setup in the impulse control problem of Section 6.1. The corresponding DPP is (see (6.2.3)):

$$V(x) = \inf_{T \in \mathcal{S}} \mathbb{E}\left[\int_0^T e^{-\gamma s} g(X(s))ds + e^{-\gamma T} \mathcal{K} V(X(T-))\, 1_{\{T<\infty\}}\right], \qquad (6.5.1)$$

The expression on the RHS to be minimized can be interpreted in terms of the following strategy. We choose a time T and decide to make no interventions until time T, thus accumulating a cost of $\mathbb{E}\int_0^T e^{-\gamma s} g(X(s))ds$ over $[0, T-]$. At

time T we then intervene by the best possible amount and proceed optimally afterward (represented by $\mathcal{K}V(X(T-))$). The DPP says that the above strategy is optimal as long as T is the first time at which it is optimal to intervene. Thus, in a sense, we have reduced the impulse control problem to an optimal stopping problem with "terminal cost" $\mathcal{K}V$. As we have previously remarked, this is not an optimal stopping problem in the true sense since $\mathcal{K}V$ is unknown (it is part of the solution). However, the above considerations suggest that connections with optimal stopping can be used in order to solve impulse control problems. In this section we present an iterative algorithm for solving impulse control problems which takes advantage of these connections.

et $X^{(0)}$ denote the state process when there are no interventions, and assume $X^{(0)}$ satisfies

$$dX^{(0)}(t) = b(X^{(0)}(t))dt + \sigma(X^{(0)}(t))dW(t)\ , \quad X^{(0)}(0-) = x.$$

Let ϕ_0 denote the cost when there is no control; that is

$$\phi_0(x) = \mathbb{E}\int_0^\infty e^{-\gamma t} g(X^{(0)}(t))dt.$$

Recalling the definition of the intervention operator $\mathcal{K}$ in (6.2.1), we define a sequence of functions $\phi_1, \phi_2, \ldots$ recursively via

$$\phi_k(x) = \inf_{T\in\mathcal{S}} \mathbb{E}\left[\int_0^T e^{-\gamma s} g(X^{(0)}(s))ds + e^{-\gamma T}\mathcal{K}\phi_{k-1}(X^{(0)}(T-))\, 1_{\{T<\infty\}}\right].$$

Defined in this way ϕ_k is obtained by solving a series of k optimal stopping problems. That is, given ϕ_0 we solve an optimal stopping problem with terminal cost ϕ_0 to obtain ϕ_1. We then use ϕ_1 as our terminal cost and solve the above optimal stopping problem to obtain ϕ_2, and so on. Thus the functions ϕ_k are computable, at least in principle. We now wish to see how these functions are related to the impulse control problem. Intuitively, solving a sequence of optimal stopping problems gives us a sequence of optimal intervention times.

Let V_k denote the value function of the impulse control problem with the restriction that no more than k interventions can be made. That is,

$$V_k(x) = \inf_{I\in\mathcal{A}_k} J(x, I),$$

where $\mathcal{A}_k = \{I \in \mathcal{A} : I = \{(\tau_j, \zeta_j) : j \le M\}$ and a.s. $M \le k\}$. We claim that under certain assumptions

$$V_k(x) = \phi_k(x), \text{ for all } x \in \mathbb{R}^n, k = 0, 1, 2, \ldots$$
$$\lim_{k\to\infty} V_k(x) = V(x), \text{ for all } x \in \mathbb{R}^n.$$

For a more precise statement of the claim and a proof see [**Oks03**, Lemma 7.1, Theorem 7.2]. ([**Oks03**, Chapter 7] provides many more details on the material in the current section.) The main idea in the proof is dynamic programming.

The main implication of the above claim is this: To get a good approximation to the solution of an impulse control problem, we may solve iteratively k optimal stopping problems for a large enough value of k. Thus we fix a large enough value of k and run the algorithm to obtain ϕ_k. The claim tells us that $\phi_k = V_k$, and that for large enough k, $V_k \approx V$. This is useful both computationally — it gives us one numerical procedure for solving impulse control problems — and analytically — the iterative algorithm can provide insight into the solution of an impulse control problem which may be difficult to solve directly.

6.5.1. Construction of an optimal control based on the iterative algorithm. We fix a large value of k and wish to use the iterative alrogithm above to find an optimal control for the restricted impulse control problem. That is, we want to construct $I = (\tau_1, \dots, \tau_k; \zeta_1, \dots, \zeta_k) \in \mathcal{A}_k$ which satisfies $J(x, I) = V_k$. The main thing to note, is that in order to do this, we must first run all k iterations of the optimal stopping-based algorithm. That is, the construction of I depends on all k iterations. (It is not true, for example, that solving for just ϕ_1 gives (τ_1, ζ_1).)

For $j = 1, \dots, k$ let

$$D_j = \{x : \phi_j(x) > \mathcal{K}\phi_{j-1}(x)\}.$$

Then D_j represents the continuation region for the optimal stopping problem with value function ϕ_j and terminal cost $\mathcal{K}\phi_{j-1}$. In the construction below we will use D_{k-j-1} to define τ_j; that is, τ_j will be the first time after τ_{j-1} that the controlled process exits D_{k-j-1}. Also, we will use ϕ_{k-j} to define ζ_j. That is, ζ_j is the minimizer z of $\phi_{k-j}(\Gamma(x, z)) + \ell(x, z)$, where x will be replaced by the value of the controlled state process at time τ_j-.

Suppose that we have run the algorithm and have obtained $\phi_1, \dots, \phi_k$ and computed $D_1, \dots, D_k$. Recall that $X^{(0)}$ is the uncontrolled state process. We begin by defining the time of the first intervention

$$\tau_1 = \inf\{t \geq 0 : X^{(0)}(t) \notin D_k\}.$$

At time τ_1 the impulse value ζ_1 is give by

$$\zeta_1 \in \arg\min_{z \in \mathcal{Z}} \phi_0(\Gamma(X^{(0)}(\tau_1-), z)) + \ell(X^{(0)}(\tau_1-), z).$$

We denote by $X^{(1)}$ the state process controlled by (τ_1, ζ_1). That is

$$X^{(1)}(t) = \begin{cases} x + \int_0^t b(X^{(1)}(s))ds + \int_0^t \sigma(X^{(1)}(s))dW(s), & t < \tau_1, \\ X^{(1)}(\tau_1) + \int_0^t b(X^{(1)}(s))ds + \int_0^t \sigma(X^{(1)}(s))dW(s), & t \geq \tau_1, \end{cases}$$

where $X^{(1)}(\tau_1) = \Gamma(X^{(1)}(\tau_1-), \zeta_1)$.

Now suppose that we have defined, for $j \leq k-1$, $\tau_1, \dots, \tau_j$, $\zeta_1, \dots, \zeta_j$, and $X^{(j)}$, the state process corresponding to control $(\tau_1, \dots, \tau_j; \zeta_1, \dots, \zeta_j)$. Then

we define the time of the $(j+1)$-st intervention

$$\tau_{j+1} = \inf\{t > \tau_j : X^{(j)}(t) \notin D_{k-j}\}.$$

The impulse value at the $(j+1)$-st intervention is given by

$$\zeta_j \in \arg\min_{z \in \mathcal{Z}} \phi_{k-(j+1)}(\Gamma(X^{(j)}(\tau_{j+1}-), z)) + \ell(X^{(j)}(\tau_{j+1}-), z).$$

Note that, under certain assumptions, $D_k \subseteq D_{k-1} \subseteq D_1$. Since D_k corresponds to τ_1 and D_1 to τ_k, intuitively, we will wait longer to intervene when we can only make one more iteration than when we still have all k iterations available to us (with similar considerations in the intermediate steps).

With I constructed in this way, we have $J(x, I) \approx V(x)$ for all $x \in \mathbb{R}^n$ as long as k is large enough.

APPENDIX A

Homework Problems

1.1. Homework 1 - due Wednesday, April 23, 2008

EXERCISE 1.1.1. *This exercise considers a one-dimensional, infinite horizon discounted cost version of the* stochastic linear quadratic regulator problem. *The controlled state process is given by*

$$dX(t) = [aX(t) + bU(t)]dt + \sigma dW(t) \ , \quad X(0) = x,$$

where $a, b, \sigma \in \mathbb{R}$ are constants with $b \neq 0$ and $\sigma > 0$. The state space and the control space is the real line. A control is admissible ($U \in \mathcal{A}$) if there exists a unique solution to the state equation for X and $J(x, U) < \infty$, where

$$J(x, U) = \mathbb{E} \int_0^\infty e^{-\gamma t}[mX^2(t) + kU^2(t)]dt,$$

for positive constants m, k, and γ. The value function is

$$V(x) = \inf_{U \in \mathcal{A}} J(x, U).$$

(a) *Show that the HJB equation becomes*

$$0 = -\gamma\phi(x) + \frac{\sigma^2}{2}\phi''(x) + ax\phi'(x) - \frac{b^2}{4k}(\phi'(x))^2 + mx^2 \ , \quad x \in \mathbb{R}.$$

(b) *Guessing a solution of the form $\phi(x) = rx^2 + q$ for constants r and q with $r > 0$, use the HJB equation to find expressions for r and q.*

(c) *From the first two parts we see that a candidate for an optimal Markov control function is given by the linear function $u^*(x) = -(br/k)x$. Use Theorem 2.6.1 to verify that u^* defines an optimal Markov control and that the value function satisfies $V(x) = rx^2 + q$.*

EXERCISE 1.1.2. *The value function often inherits nice properties from the structure of the underlying control problem. Fox example, if the control problem has a convex structure, then the value function is a convex function. We illustrate this point using the infinite horizon discounted cost problem of Section 2.6. We assume that the coefficients in the state equation are linear functions of state and control:*

$$dX(t) = [b_1X(t) + b_2U(t)]dt + [\sigma_1X(t) + \sigma_2U(t)]dW(t) \ , \quad X(0) = x \in \mathbb{R}^n,$$

where $b_1, b_2, \sigma_1, \sigma_2 \in \mathbb{R}$ are constants with $b_2 \neq 0$ and $\sigma_1, \sigma_2 > 0$. We assume that the control space $\mathbb{U} \subset \mathbb{R}^p$ is a convex set (i.e. if $\alpha, \tilde{\alpha} \in \mathbb{U}$ then $(1-\lambda)\alpha + \lambda\tilde{\alpha} \in \mathbb{U}$ for all $\lambda \in [0,1]$). A control is admissible ($U \in \mathcal{A}$) if: $U(t) \in \mathbb{U}$ for all $t \geq 0$, there exists a unique solution to the state equation, and $\mathbb{E}\int_0^\infty e^{\gamma t}|g(X_t, U_t)|dt < \infty$. The cost function is

$$J(x, U) = \mathbb{E}\int_0^\infty e^{-\gamma t} g(X_t, U_t)dt,$$

where $g : \mathbb{R}^n \times \mathbb{U} \mapsto \mathbb{R}$ is convex (i.e. $g((1-\lambda)y + \lambda\tilde{y}) \leq (1-\lambda)g(y) + \lambda g(\tilde{y})$ for all $y, \tilde{y} \in \mathbb{R}^n \times \mathbb{U}$ and $\lambda \in [0,1]$. The value function is

$$V(x) = \inf_{U \in \mathcal{A}} J(x, U).$$

Fix $x \in \mathbb{R}^n$, $U \in \mathcal{A}$ with corresponding state process X and $\tilde{x} \in \mathbb{R}^n$, $\tilde{U} \in \mathcal{A}$ with corresponding state process $\tilde{X}$. For $\lambda \in [0,1]$ let $\hat{x} = (1-\lambda)x + \lambda\tilde{x}$ and $\hat{U} = (1-\lambda)U + \lambda\tilde{U}$.

(a) *Verify that $\hat{U} \in \mathcal{A}$ and that the state process corresponding to $\hat{x}$ and $\hat{U}$, denoted $\hat{X}$, satisfies $\hat{X} = (1-\lambda)X + \lambda\tilde{X}$.*
(b) *Show that $J(\hat{x}, \hat{U}) \leq (1-\lambda)J(x, U) + \lambda J(\tilde{x}, \tilde{U})$.*
(c) *Show that V is a convex function on $\mathbb{R}^n$. (Hint: for $\varepsilon > 0$, select ε-optimal controls; that is, $U_\varepsilon, \tilde{U}_\varepsilon \in \mathcal{A}$ such that $J(x, U_\varepsilon) < V(x) + \varepsilon$ and $J(\tilde{x}, \tilde{U}_\varepsilon) < V(\tilde{x}) + \varepsilon$. (Why do such controls exist?))*

EXERCISE 1.1.3. *Here is a simple example in which no optimal control exists, in a finite horizon setting, $T \in (0, \infty)$. Suppose that the state equation is*

$$dX(t) = U(t)dt + dW(t)\ , \quad X(0) = x \in \mathbb{R}.$$

A control U is admissible ($U \in \mathcal{A}$) if: $U(t) \in \mathbb{R}$ for all $t \geq 0$, there exists a unique solution to the state equation, and $J(t, x, U) = \mathbb{E}[X_T^2] < \infty$. The value function $V(t, x) = \inf_{U \in \mathcal{A}} J(t, x, U)$ satisfies $V(t, x) = 0$ for all (t, x), but there is no admissible control U which attains this value.

(a) *Show that if $U_t = -cX_t$ for some constant $c > 0$ then $U \in \mathcal{A}$ and*

$$J(t, x, U) = \frac{1}{2c} - e^{-2c(T-t)}\Big(\frac{1 - 2cx^2}{2c}\Big).$$

(Hint: with such a U, X is an Ornstein-Uhlenbeck process.)
(b) *Conclude that $V(t, x) = 0$ for all $t \in [0, T)$, $x \in \mathbb{R}$.*
(c) *Show that there is no $U \in \mathcal{A}$ such that $J(t, x, U) = 0$. (Suppose that there is such a U and show that this leads to a contradiction. Intuitively, in order to have such a control, we must take $c = \infty$, which is clearly not admissible.)*

(d) *The verification procedure also breaks down here. The associated HJB equation is*

$$0 = \inf_{\alpha \in \mathbb{R}} [\frac{\partial \phi}{\partial t}(t,x) + \frac{1}{2}\frac{\partial^2 \phi}{\partial x^2}(t,x) + \alpha \frac{\partial \phi}{\partial x}(t,x)] \ , \quad x \in \mathbb{R}.$$

Show that there is no value $\alpha \in \mathbb{R}$ for which the infimum is attained.

(Remark: The goal in this problem is to bring the state process as close as possible to 0 at the terminal time T. However, as defined above, there is no cost of actually controlling the system; we can set U_t arbitrarily large without any negative consequences. From a modeling standpoint, there is often a trade-off between costs incurred in applying control and our overall objective. For example, a more appropriate cost function may be

$$J(t,x,U) = \mathbb{E}[\int_0^T (U_s)^2 ds + X_T^2],$$

which charges a running cost any time we apply control. It can be shown that under this cost function there is an optimal control which can be found through the usual verification procedure (in fact, this is another version of a stochastic linear quadratic regulator problem).)

EXERCISE 1.1.4. *In this exercise we use methods similar to those in Section 2.7 to solve the Merton problem with logarithmic utility. Consider the same setup as in Section 2.7, except with $g(c) = \log(c)$. Assume that $\gamma > 0$ and let*

$$\begin{aligned} \phi(x) &= \frac{1}{\gamma^2}\Big(r + \frac{\beta^2}{2} - \gamma\Big) + \frac{\log(\gamma x)}{\gamma} \\ \pi^* &= \frac{\beta}{\sigma} \\ c^*(x) &= \gamma x. \end{aligned}$$

Show that the value function is given by $V(x) = \phi(x)$. Let $\Pi^*(t) = \pi^*$, $t \geq 0$, and define a consumption process C^* via $C^*(t) = c^*(X^*(t))$, where X^* is the controlled process corresponding to $U^* = (\Pi^*, C^*)$. Show that $U^* \in \mathcal{A}(x)$ and $V(x) = J(x, U^*)$. (Hint: first consider a solution of the HJB equation of the form $\phi(x) = a\log(x) + b$.)

EXERCISE 1.1.5. *We consider Exercise 1.1.1 with an ergodic cost. Let the state equation be as in Exercise 1.1.1 with $a = 0$ for simplicity. A control is admissible ($U \in \mathcal{A}$) if there exists a unique solution to the state equation for X and if $\limsup_{t\to\infty} \mathbb{E}[X_t^2]/t = 0$. The cost associated with control U is*

$$J(U) = \limsup_{T\to\infty} \mathbb{E}\Big[\int_0^T [mX^2(t) + kU^2(t)]dt\Big],$$

for positive constants m, k. The value of the control problem is

$$V = \inf_{U \in \mathcal{A}} J(U).$$

The HJB equation is

$$0 = -\eta + \frac{\sigma^2}{2}\phi''(x) - \frac{b^2}{4k}(\phi'(x))^2 + mx^2 \ , \quad x \in \mathbb{R}.$$

By guessing a solution of the form $\phi(x) = px^2$ for a constant $p > 0$ show that $\eta = V = J(U^*) = p\sigma^2$, where $p = \sqrt{mk}/b$, and U^* is an admissible Markov control defined via the Markov control function $u^*(x) = -\sqrt{(m/k)}x$.

1.2. Homework 2 - due Friday, May 9, 2008

EXERCISE 1.2.1. *Prove that if ϕ is a viscosity solution of equation* (4.5.4) *and ϕ is of class $C^{1,2}$, then ϕ is a classical solution of* (4.5.4).

EXERCISE 1.2.2. *Consider the infinite horizon discounted cost problem of Section 2.6. The dynamic programming principle is: For any $\{\mathcal{F}_t\}$-stopping time τ taking values in $[0,\infty)$ we have*

$$V(x) = \inf_{U\in\mathcal{A}} \mathbb{E}\int_0^\tau e^{-\gamma s} g(X^U(s), U(s))ds + e^{-\gamma\tau} V(X^U(\tau))\Big]. \tag{1.2.1}$$

Using the notation of Section 4.5, with the time variable omitted, the corresponding HJB equation (2.6.2) *can be written as*

$$\gamma V(z) + \mathcal{H}(z, D_x V(z), D_x^2 V(z)) = 0\ , \qquad z\in\mathbb{R}^n. \tag{1.2.2}$$

A viscosity solution of (1.2.2) *is defined as follows. Note that we do not "replace" V itself by ϕ; we only use ϕ when derivatives of V are involved.*

DEFINITION 1.2.3. *Suppose that $V\in C(\mathbb{R}^n)$.*

(a) *V is a* viscosity subsolution *of* (1.2.2) *if: For all $z\in\mathbb{R}^n$ and all $\phi\in C^2(\mathbb{R}^n)$ for which $V-\phi$ has a local maximum at z we have*

$$\gamma V(z) + \mathcal{H}(z, D_x\phi(z), D_x^2\phi(z)) \le 0$$

(b) *V is a* viscosity supersolution *of* (4.5.4) *if: For all $z\in\mathbb{R}^n$ and all $\phi\in C^2(\mathbb{R}^n)$ for which $V-\phi$ has a local minimum at z we have*

$$\gamma V(z) + \mathcal{H}(s, z, D_x\phi(z), D_x^2\phi(z)) \ge 0$$

(c) *V is a* viscosity solution *of* (1.2.2) *if it is both a viscosity subsolution and a viscosity supersolution of* (1.2.2).

Suppose that Assumption 4.2.1 holds (with t and h omitted) and that in addition φ is bounded.

(a) *Show that the value function V is bounded and continuous on $\mathbb{R}^n$.*
(b) *Show that the value function V is a viscosity solution of* (1.2.2).

EXERCISE 1.2.4. *(Exercise 12.15 of* [**Oks03**]*.) Solve the optimal stopping problem*

$$V(t,x) = \sup_{\tau\ge 0} \mathbb{E}[e^{-\gamma(s+\tau)}(a - X(\tau))^+],$$

where

$$dX(s) = \mu X(s) + \sigma X(s) dW(s)\ , \quad X(t) = x,$$

and $\gamma > 0$, $\mu > 0$, $\sigma > 0$, $a > 0$, are constants. Proceed as in the asset selling example presented in lecture to show that

$$V(t,x) = \begin{cases} e^{-\gamma t}(a-x), & x \le x^*, \\ e^{-\gamma t}(a-x^*)\left(\frac{x}{x^*}\right)^{\beta}, & x > x^*, \end{cases}$$

where

$$\beta = \sigma^{-2}\Big[\frac{\sigma^2}{2} - \mu - \sqrt{(\sigma^2/2-\mu)^2 + 2\gamma\sigma^2}\Big] < 0,$$

and

$$x^* = \frac{a\beta}{\beta - 1} \in (0, a).$$

Moreover, show that $\tau_x^* = \inf\{t \ge 0 : X_x(t) \le x^*\}$ is an optimal stopping time. (If $\mu = \gamma$ then $V(t,x)$ gives the price of an American put option with infinite horizon (i.e. a perpetual option). Then $\beta = -2\gamma/\sigma^2$ and $x^* = (2a\gamma)/(\sigma^2+2\gamma)$.)

EXERCISE 1.2.5. *(Example 4.1 of [**ØS07**].) We are operating a plant that extracts natural gas from an underground well. The total amount of natural gas remaining in the well at time t is denoted $Q(t)$ (so the total amount of extracted natural gas is $Q(0) - Q(t)$). At time t we can choose the rate $U(t)$ at which we can extract natural gas from the well; thus Q satisfies:*

$$dQ(t) = -U(t)Q(t)dt\ , \quad Q(0) = q > 0.$$

We assume that there is some maximum extraction rate, so that U takes values in $[0, m]$. (As usual, we only consider adapted U for which the above equation has a unique solution.) After the gas has been extracted, it is sold at price $P(t)$, given by the equation

$$dP(t) = \mu P(t)dt + \sigma P(t)dW(t)\ , \quad P(0) = p > 0.$$

The cost of keeping the plant open is a constant $K > 0$ per unit time. Thus over the interval $[0, t]$ our discounted (at rate $\gamma > 0$) net profit is

$$\begin{aligned} &\int_0^t e^{-\gamma s}P(s)d(Q(0) - Q(s)) - \int_0^t e^{-\gamma s}Kds + e^{-\gamma t}\theta Q(t)P(t) \\ &= \int_0^t e^{-\gamma s}(U(s)Q(s)P(s) - K)ds + e^{-\gamma t}\theta Q(t)P(t). \end{aligned}$$

At some point in time it may no longer be profitable to keep the plant in operation and thus we are interested in finding the optimal stopping time τ at which to close the plant. Thus we are interested in choosing a control U and a stopping time τ to maximize our expected reward:

$$J(x, \tau, U) = \mathbb{E}\Big[\int_0^{\tau} e^{-\gamma s}(U(s)Q(s)P(s) - K)ds + e^{-\gamma t}\theta Q(\tau)P(\tau)\Big].$$

(A note on the notation: Since the reward depends on P and Q only through their product PQ, we anticipate that J will be a function of $x = pq$ rather than p, q individually.) The value function of this combined optimal stopping and control problem *is*

$$V(x) = \sup_{(\tau,U)} J(x, \tau, U).$$

We assume that $\mu \leq \gamma$ and $0 < \theta < m/(m+\gamma-\mu)$. Guided by the HJB equation and the "principle of smooth fit" we get a candidate solution (see Section 4.3 of [**ØS07**] *for the construction):*

$$\phi(x) = \begin{cases} \theta x & 0 \leq x \leq x^*, \\ cx^\lambda + xm/(m+\gamma-\mu) - K/\gamma & x > x^*, \end{cases}$$

$$c = \frac{[m - \theta(m+\gamma-\mu)]x^{1-\lambda}}{-\lambda(m+\gamma-\mu)} > 0,$$

$$x^* = \frac{-\lambda K(m+\gamma-\mu)}{(1-\lambda)\gamma[m-\theta(m+\gamma-\mu)]} > 0,$$

and $\lambda < 0$ is the unique negative root of

$$0 = -\gamma + (\mu - m)\lambda + \lambda(\lambda-1)\sigma^2/2.$$

Moreover, the candidate optimal control is $U^* \equiv m$ and the candidate optimal stopping time is $\tau_x^* = \inf\{t \geq 0 : P(t)Q(t) \leq x^*\}$.

The exercise: *sketch the verification argument which shows that $V(x) = \phi(x) = J(x, \tau_x^*, U^*)$ for all $x > 0$. (Hint: it will help to first write the associated HJB equation and to remember that ϕ was constructed to be a solution of it.)*

1.3. Homework 3 - due Tuesday, May 27, 2008

EXERCISE 1.3.1. *(Exercise 6.2 of* [**ØS07**]*) This exercise is similar to the "optimal dividend policy" problem presented in lecture. The main difference is that, in the absence of control, the state process is a geometric Brownian motion (rather than a Brownian motion with drift). Let* $I = \{(\tau_j, \zeta_j), j = 1, 2, \ldots\}$ *be an impulse control with state process* X*, with* $X(0) = x > 0$*, given by*

$$\begin{aligned} dX(t) &= \mu X(t)dt + \sigma X(t)dW(t)\ , \quad \tau_j < t < \tau_{j+1}, \\ X(\tau_j) &= X(\tau_j-) - c - (1+\lambda)\zeta_j, \end{aligned}$$

where $\mu > 0$*,* $\sigma > 0$*,* $c > 0$*,* $\lambda \geq 0$*. Control* I *is admissible (*$I \in \mathcal{A}$*) if the usual conditions for admissibility are satisfied and in addition* $X(t) \geq 0$ *for all* $t \geq 0$*. The cost for initial state* x *and impulse control* I *is*

$$J(x, I) = \mathbb{E} \sum_{j=1}^{\infty} e^{-\gamma\tau_j}(-\zeta_j)\ 1_{\{\tau_j \leq T\}},$$

where $\gamma > \mu$ *and we assume that the process is stopped at time* $T = \inf\{t \geq 0 : X(t) \leq 0\}$*. The value function is* $V(x) = \inf_{I \in \mathcal{A}} J(x, I)$*.*

We construct a solution ϕ *to the associated HJB equation. Begin by guessing that the non-intervention region is given by* $D = \{x : 0 < x < x^*\}$ *for some* $x^* > c$ *and that* $\phi(x) = ax^\beta$ *for* $x < x^*$*, where* $a < 0$ *is to be determined and* $\beta > 1$ *is the unique positive solution of*

$$-\gamma + \mu\beta + (\sigma^2/2)\beta(\beta - 1) = 0.$$

(a) *By guessing that for* $x \geq x^*$ *it is optimal to apply the maximum (in absolute value) possible amount of impulse control, claim that we have* $\phi(x) = (c - x)/(1 + \lambda)$*.*

(b) *Use the principle of smooth fit to find that the values of* a *and* x^* *are*

$$x^* = \frac{\beta c}{\beta - 1}\ , \quad a = \frac{c - x^*}{1 + \lambda}(x^*)^{-\beta}.$$

(c) *Check that the constructed solution* ϕ *is indeed a solution of the HJB equation.*

(d) *Let* X *denote the state process with no control (i.e. just a geometric Brownian motion). Let* $\tau_1^* = \inf\{t \geq 0 : X(t) \geq x^*\}$ *and* $\zeta_1^* = (X(\tau_1^*-)-c)/(1+\lambda)$*. Show that* $I^* = (\tau_1^*, \zeta_1^*)$ *is an admissible impulse control. Sketch the verification argument which shows that* $\phi(x) = J(x, I^*) = V(x)$*. (Note that if* X^* *is the associated control process then* $X^*(t) = 0$ *for* $t \geq \tau_1^*$*.)*

(e) *Explain why $V(x)$ is equal to the value function of the optimal stopping problem*

$$\sup_{\tau \le T} E\Big[e^{-\gamma\tau}\frac{(X(\tau)-c)^+}{1+\lambda}\Big].$$

EXERCISE 1.3.2. *(Example VIII.4.1. of* [**FS06**]*) For $x \in \mathbb{R}$ let*

$$X(t) = x + \sqrt{2}W(t) - \xi(t)$$

The control ξ is an admissible singular control ($\xi \in \mathcal{A}$) if it is adapted, non-negative, nondecreasing and RCLL. The cost for initial state x and control ξ is

$$J(x,\xi) = \mathbb{E}\Big[\int_0^\infty e^{-t}cX^2(t)dt + \int_{[0,\infty)} e^{-t}d\xi(t)\Big],$$

where $c > 0$. The value function is $V(x) = \inf_{\xi\in\mathcal{A}} J(x,\xi)$. For $a = 1 + (2c)^{-1}$ let

$$\phi(x) = \begin{cases} (1-2ca)e^{x-a} + cx^2 + 2c, & x \le a, \\ x - a + (1-2ca) + ca^2, & x \ge a \end{cases}.$$

(a) *Check that the constructed solution ϕ is indeed a solution of the HJB equation.*

(b) *Let ξ^* be given by*

$$\xi^*(t) = \sup_{0\le s\le t}[\max\{0, x + \sqrt{2}W(t) - a\}].$$

Sketch the verification argument which shows that $\phi(x) = J(x, xi^*) = V(x)$. (Hint: consider $x < a$ first. Note that $X^*(t) = x + \sqrt{2} - \xi^*(t) < a$ for all $t \ge 0$ and that

$$\int_{[0,t]} 1_{\{X^*(s)\ge a\}} d\xi^*(s) = \xi^*(t).)$$

EXERCISE 1.3.3. *Consider the controlled state process*

$$dX(t) = bU(t)dt + \sigma dW(t)\ , \quad X(0) = x = (-M, M),$$

where $b \in \mathbb{R}$ and $\sigma > 0$ and U takes values in $\mathbb{R}$. The cost function is

$$J(x,U) = \mathbb{E}\left[\int_0^\tau (pU^2(s) - q)ds\right],$$

where $p > 0$, $q > 0$. It can be shown that if $2Mb\sqrt{q} < \pi\sigma^2\sqrt{p}$ then an optimal Markov control is given by

$$u^*(x) = -\sqrt{\frac{q}{p}}\tan\left(\frac{b\sqrt{q}}{\sigma^2\sqrt{p}}x\right),$$

and the value function is

$$V(x) = \frac{2p\sigma^2}{b^2}\left[\log\left(\cos\left(\frac{b\sqrt{q}}{\sigma^2\sqrt{p}}M\right)\right) - \log\left(\cos\left(\frac{b\sqrt{q}}{\sigma^2\sqrt{p}}x\right)\right)\right].$$

In this exercise, we compute numerically the value function and optimal control using the Markov chain method (for several different choices of Markov chain) and compare our approximations to the analytic results. Set $M = 0.5$, $b = 0.7$, $\sigma = 0.5$ *and* $p = q = 1$. *Write a program to run a combined value iteration/policy iteration algorithm to compute the optimal control and value function for the corresponding discretized problem with state space* $\{-M, -M+h, -M+2h, \ldots, M\}$. *(You may use any programming language you wish.) Run the algorithm for the following 3 cases (corresponding to different Markov chain approximations.) In all cases below we take* $p(-M, -M|\alpha) = p(M, M|\alpha) = 1$ *for any* α.

	$\underline{\frac{\Delta(x,\alpha)}{h^2}}$	$\underline{p(x, x+h\|\alpha)}$	$\underline{p(x, x-h\|\alpha)}$	$\underline{p(x, x\|\alpha)}$	$\underline{\mathcal{U}}$
(1)	$\frac{1}{\sigma^2 + h\|b\alpha\|}$	$\frac{\sigma^2/2 + h(b\alpha)^+}{\sigma^2 + h\|b\alpha\|}$	$\frac{\sigma^2/2 + h(b\alpha)^-}{\sigma^2 + h\|b\alpha\|}$	0	$\mathbb{R}$
(2)	$\frac{h^2}{\sigma^2 + h\|bm\|}$	$\frac{\sigma^2/2 + h(b\alpha)^+}{\sigma^2 + h\|bm\|}$	$\frac{\sigma^2/2 + h(b\alpha)^-}{\sigma^2 + h\|bm\|}$	$\frac{h\|bm\| - h\|b\alpha\|}{\sigma^2 + h\|bm\|}$	$[-m, m]$
(3)	$\frac{h^2}{\sigma^2}$	$\frac{\sigma^2 + h(b\alpha)}{2\sigma^2}$	$\frac{\sigma^2 - h(b\alpha)}{2\sigma^2}$	0	$[-\frac{\sigma^2}{bh}, \frac{\sigma^2}{bh}]$

(a) *For each value of* $h \in \{0.20, 0.10, 0.05, 0.02\}$ *plot the computed value function on the same plot as the known value function. Do the same for the computed optimal control versus the actual optimal control.*

(b) *Compare the results under the different chains. Does one chain seem to converge faster in* h *(i.e. it gives good approximations for even the larger values of* h*)? Fixing* h *and comparing across chains, does one chain seem to converge in fewer value/policy iterations? Did you notice any difference in the computational speed?*

Be sure to hand in hardcopies of your code and your plots with your assignment. Some notes: You may wish to investigate the corresponding dynamic programming equations to see if the minimization (over α*) can be simplified before writing this part of the code. Note that you have some choice in some of the remaining parameters - e.g., the value of* m *for the bounded control space, the number of policy/value iterations to run (fixed number or certain error tolerance), initial guess for control and its cost, etc.*

APPENDIX B

Information on Paper Presentation

Each student will be required to make a 20 minute presentation on a research paper in the subject of stochastic control. Unless the instructor specifies otherwise, all presentations will take place on Tuesday, June 10, 8:30 am - 11:30 am. ALL registered students must be in attendance for ALL presentations.

Each student should choose a paper of his or her interest and notify the instructor of the selection NO LATER THAN MONDAY, MAY 26. The instructor will determine if the paper is appropriate for a presentation. If two students select the same paper, the one who notifies the instructor first will get preference.

Each presentation should include:

(a) Precise statement of the stochastic control problem
(b) Discussion of the main results of the paper
(c) Brief literature review highlighting the contributions of the paper in light of what is already known (so the paper should have been published relatively recently)
(d) Outline of key techniques or novel ideas used in proofs of main results
(e) Discussion of applications or numerical results

Specific, rigorous details do not need to be presented, but your presentation should reflect that you have carefully studied the paper. For example, you should not present a line-by-line proof of a result, but you should be able to answer questions about the issues involved in the proof. The goal of the presentations is to provide a brief overview of the various research being done in the area of stochastic control. Each presentation should emphasize aspects that make the paper interesting in terms of both theory and applications.

Be sure to practice your presentation to make sure that you can cover your material in 20 minutes. You are encouraged to prepare about 15 minutes of material to account for questions that will come up. In order to expedite time, each student is encourage to prepare a slide presentation (a maximum of about 10 slides). Please save your slides file in .pdf or Powerpoint format and email to me before the presentation. (I will bring one laptop with all the files for the presentations.)

If you choose to use the blackboard instead of slides, you should make handouts with all the details of your presentation. This way, you can just write the highlights of your talk on the board and refer to the handouts for details.

The following survey papers include excellent suggestions for papers to present. They can be found at http://www.i-journals.org/ps/

Pham, "On some recent aspects of stochastic control and their applications", *Probab. Surv.* 2 (2005) 506–549.

Borkar, "Controlled diffusion processes", *Probab. Surv.* 2 (2005), 213–244.

These papers mostly include applications in financial mathematics, but you are free to choose applications in other fields. For further suggestions, see the references on the syllabus and in the class notes (and also the bibliographies in those references). If you are interested in a particular area, you can email me for suggestions. Please be as specific as possible in describing what you are interested in.

Bibliography

[Bor89] Vivek S. Borkar, *Optimal control of diffusion processes*, Pitman Research Notes in Mathematics Series, vol. 203, Longman Scientific & Technical, Harlow, 1989. MR MR1005532 (90h:93115)

[Bor05] ———, *Controlled diffusion processes*, Probab. Surv. **2** (2005), 213–244 (electronic). MR MR2178045 (2006k:93146)

[Dem07] Amir Dembo, *Stochastic processes: Lectures notes for stat219/math136 at stanford*, 2007.

[DN90] M. H. A. Davis and A. R. Norman, *Portfolio selection with transaction costs*, Math. Oper. Res. **15** (1990), no. 4, 676–713. MR MR1080472 (92b:90036)

[EK86] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986, Characterization and convergence. MR MR838085 (88a:60130)

[FS06] Wendell H. Fleming and H. Mete Soner, *Controlled Markov processes and viscosity solutions*, second ed., Stochastic Modelling and Applied Probability, vol. 25, Springer, New York, 2006. MR MR2179357 (2006e:93002)

[GT01] David Gilbarg and Neil S. Trudinger, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001, Reprint of the 1998 edition. MR MR1814364 (2001k:35004)

[KLS87] Ioannis Karatzas, John P. Lehoczky, and Steven E. Shreve, *Optimal portfolio and consumption decisions for a "small investor" on a finite horizon*, SIAM J. Control Optim. **25** (1987), no. 6, 1557–1586. MR MR912456 (89d:90025)

[KS97] Ioannis Karatzas and Steven E. Shreve, *Brownian motion and stochastic calculus*, 2nd ed., Graduate Texts in Mathematics, vol. 113, Springer Verlag, 1997.

[Oks03] Bernt Oksendal, *Stochastic differential equations: An introduction with applications*, 6th ed., Universitext, Springer Verlag, 2003.

[ØS07] Bernt Øksendal and Agnès Sulem, *Applied stochastic control of jump diffusions*, second ed., Universitext, Springer, Berlin, 2007. MR MR2322248 (2008b:93003)

[Rob83] Maurice Robin, *Long-term average cost control problems for continuous time Markov processes: a survey*, Acta Appl. Math. **1** (1983), no. 3, 281–299. MR MR727156 (85b:49021)

[YZ99] Jiongmin Yong and Xun Yu Zhou, *Stochastic controls*, Applications of Mathematics (New York), vol. 43, Springer-Verlag, New York, 1999, Hamiltonian systems and HJB equations. MR MR1696772 (2001d:93004)